

## Curriculum Vitae

Meagan C. Small  
20 Penn St., Rm 601, Baltimore MD 21201  
Email: meagan.small@gmail.com

## Education

*University of Maryland, Baltimore, Baltimore MD*

### **Ph. D. in Pharmaceutical Sciences, July 2015**

Dissertation: "Molecular Modelling of Macrolide Antibiotic Conformational Sampling and Interactions in the 50S Ribosomal Subunit for the Development of Novel Antibiotics"

Area of Specialization: Computer-aided drug design, molecular dynamics, force field development, docking

*University of South Florida, Tampa FL*

### **M.S. in Chemistry, August 2010**

Thesis: "Time-resolved Thermodynamics of Heme Copper Oxidase Models"

Areas of Specialization: Spectroscopy, thermodynamics, and biophysics

*University of South Florida, Tampa FL*

### **B.S. in Chemistry, May 2007**

Areas of Specialization: Physical chemistry

Honors: Graduated Cum Laude

*University of South Florida, Tampa FL*

### **B.S. in Biology, May 2007**

Honors Thesis: "Microbial Diversity Using Phylogenetic Analysis of 16S rRNA from Jewfish Sink"

Areas of Specialization: Molecular biology and microbiology

Honors: Graduated Cum Laude

## Publications

### *Peer-Reviewed Journal Articles*

1. "Desmethyl Macrolides: Synthesis and Evaluation of 4-Desmethyl Telithromycin"  
Ian Glassford, Miseon Lee, Bharat Wagh, Venkata Velvadapu, Tapas Paul, Gary Sandelin, Charles DeBrosse, Dorota Klepacki, Meagan C. Small, Alexander D. MacKerell, Jr. and Rodrigo B. Andrade, **ACS Med Chem Let (2014)** 5(29): 1021-1026

2. “Desmethyl Macrolides: Synthesis and Evaluation of 4,8,10-Tridesmethyl Cethromycin” Bharat Wagh, Tapas Paul, Charles DeBrosse, Dorota Klepacki, Meagan C. Small, Alexander D. MacKerell, Jr. and Rodrigo B. Andrade, **ACS Med Chem Let (2013)**, 4(11): 1114-1118
3. “Impact of Ribosomal Modification on the Binding of Antibiotic Telithromycin Using a Combined Grand Canonical Monte Carlo/Molecular Dynamics Simulation Approach” Meagan C. Small, Pedro Lopes, Rodrigo B. Andrade, Alexander D. MacKerell, Jr., **PLOS Computational Biology (2013)**, 9(6):e1003113
4. “Desmethyl Macrolides: Synthesis and Evaluation of 4,10-Didesmethyl Telithromycin” Venkata Velvadapu, Ian Glassford, Miseon Lee, Tapas Paul, Charles DeBrosse, Dorota Klepacki, Meagan C. Small, Alexander D. MacKerell, Jr. and Rodrigo B. Andrade, **ACS Med Chem Let (2012)**, 3(3):211-215
5. “Desmethyl Macrolides: Synthesis and Evaluation of 4,8-Didesmethyl Telithromycin” Bharat Wagh, Tapas Paul, Ian Glassford, Charles DeBrosse, Dorota Klepacki, Meagan C. Small, Alexander D. MacKerell, Jr. and Rodrigo B. Andrade, **ACS Med Chem Let (2012)**, 3(12):1013-1018
6. “Scaled-Up Syntheses of Cuprilig for Removal of Nuisance Anions” Dean F. Martin, Nicole Rubin, Meagan Small, **Florida Scientist (2010)**, 73:39-46.
7. “Use of Model Compounds to Study Removal of Pharmaceuticals Using Octolig” Wen-Shan Chang, Dean F. Martin, Meagan Small, **Technology and Innovation (2010)**, 12(1):71-77
8. “Removal of Selected Pharmaceuticals Using Octolig: A Supported Chelating Agent” Wen-Shan Chang, Dean F. Martin, Meagan Small, **Technology and Innovation (2010)**, 12(2):143-152

### *Book Chapters*

1. “Force field representation of biomolecular systems” Meagan C. Small and Alexander D. MacKerell, Jr., *In Silico Drug Discovery and Design: Theory, Methods, Challenges and Applications* (2015, in press), CRC Press: Taylor & Francis Group, ISBN (978-1-4822-1783-4)

### **Presentations**

*59<sup>th</sup> Annual Biophysical Society Meeting, Baltimore MD*

Poster Presentation, February 2015

“Molecular Understanding of the Binding of Macrolide Antibiotics to the Ribosome Using Site-Identification via Ligand Competitive Saturation”

*Computer-Aided Drug Design Symposium, Baltimore MD*

Poster Presentation, June 2014

“Conformational Sampling of Desmethyl Telithromycin Using Hamiltonian Replica Exchange Molecular Dynamics”

*Second Nanomedicine Day, Baltimore MD*

Poster Presentation, July 2013

“Impact of Ribosomal Modification on the Binding of Antibiotic Telithromycin using Empirical Force Field Simulations: Implications for Drug Design”

*1<sup>st</sup> Annual Regional AAPS, West Virginia University, Morgantown WV*

Poster Presentation, May 2013

“Impact of Ribosomal Modification on the Binding of Antibiotic Telithromycin using Computer Simulations”

*35<sup>th</sup> Annual Graduate Research Conference, Baltimore MD*

Oral Presentation, April 2013

“Impact of Ribosomal Modification on the Binding of Antibiotic Telithromycin using Empirical Force Field Simulations: Implications for Drug Design”

*244<sup>th</sup> American Chemical Society National Meeting and Exposition, Philadelphia PA*

Poster Presentation, August 2012

“Investigations of the Impact of Ribosomal Modification on the Binding of the Antibiotic Telithromycin Using Molecular Dynamics Simulations”

*Mid-Atlantic Regional Meeting of ACS, Baltimore MD*

Poster Presentation, May 2012

“Investigations of the Impact of Ribosomal Modification on the Binding of the Antibiotic Telithromycin Using Molecular Dynamics Simulations”

*34<sup>th</sup> Annual Graduate Research Conference, Baltimore MD*

Poster Presentation, April 2012

“Investigations of the Impact of Ribosomal Modification on the Binding of the Antibiotic Telithromycin Using Molecular Dynamics Simulations”

*4th Annual Frontiers at the Chemistry-Biology Interface Symposium, Newark DE*

Poster Presentation, April 2011

“Conformational Sampling of Telithromycin/Cethromycin and Desmethyl Analogs”

*33rd Annual Graduate Research Conference, Baltimore MD*

Poster Presentation, April 2011

“Conformational Sampling of Telithromycin/Cethromycin and Desmethyl Analogs”

*8<sup>th</sup> Annual Castle Conference, Tampa FL*

Poster Presentation, April 2010

“Synthesis and Characterization of a Novel Calix[4]resorcinarene”

*7<sup>th</sup> Annual Castle Conference, Tampa FL*

Session Chair, April 2009

Presided over talks, mediated questions

*53<sup>rd</sup> Annual Biophysical Society Meeting, Boston MA*

Poster Presentation, February 2009

“Time-resolved Thermodynamics of CO Photolysis from an Engineered CuB Center in Sperm Whale Myoglobin”

*60<sup>th</sup> Southeast Regional Meeting of ACS, New Orleans LA*

Oral Presentation: Physical Chemistry Session, November 2008

“Time-resolved thermodynamics of CO Photolysis from an Engineered Copper Center in Sperm Whale Myoglobin”

### **Awards and Honors**

H.A.B. Dunning Research Merit Award, University of Maryland, Baltimore, **2014-2015**

Dr. Frank J. Slama Graduate Fellowship Award, University of Maryland, Baltimore,  
**2013-2014**

Fred L & Helen M Tharp Endowed Scholarship, University of South Florida, **2009**

## **Abstract**

Dissertation Title: Molecular Modeling of Macrolide Antibiotic Conformational Sampling and Interactions in the 50S Ribosomal Subunit for the Development of Novel Antibiotics

Meagan C. Small, Doctor of Philosophy, 2015

Dissertation Directed By: Alexander D. MacKerell, Jr., Professor, Department of Pharmaceutical Sciences

Overcoming microbial resistance is a major challenge in the development of antibiotics. Bacteria limit the effectiveness of antibiotics using three major mechanisms: extrusion of the drug via efflux pumps, metabolism to an inactive metabolite, or inhibition of binding by modification of the drug target. The macrolides are an important class of antibiotics that target the ribosome and recent generation macrolides have largely addressed resistance stemming from the first two mechanisms. However, they remain susceptible to resistance due to modification of the ribosome, mainly modification of base A2058 (*E. coli* numbering throughout) that resides within the heart of the macrolide binding pocket. While crystal structures are available for bacterial 70S ribosomes with macrolides bound, there are none available for A2058-modified ribosomes. Thus, the molecular details underlying A2058 modification-based resistance are unclear.

The motivation underlying the present work is to address the need for novel antibiotics, including those addressing A2058 modification-based resistance. To accomplish this, a three-pronged approach has been employed that incorporates both

ligand- and structure-based drug design. First, utilizing a ligand-based strategy, the effects of macrolide desmethylation are investigated using molecular dynamics and a pharmacophore-based method known as Conformationally Sampled Pharmacophore (CSP). This will be the subject of Chapter 2. In Chapter 3, the focus shifts to the structure. Molecular dynamics simulations of the 50S subunit are used to understand the impact of A2058 modification on the binding of third generation macrolide antibiotic telithromycin. And, to complete the three-pronged approach, a fragment-based computer-aided drug design method known as Site-Identification by Ligand Competitive Saturation (SILCS) is applied to the ribosome leading to macrolide antibiotics with novel functionality and the potential for enhanced activity against A2058-modified ribosomes. This is the subject of Chapter 4.

The methodology underlying all of this work is the use of empirical force field-based simulations, which will be the focus of Chapter 1. As an extension of force fields, Chapter 5 will deal with the optimization of small molecule aldehydes and ketones as well as acyclic sugars toward the development of a comprehensive CHARMM polarizable biomolecular force field based on the classical Drude oscillator.

Molecular Modelling of Macrolide Antibiotic Conformational Sampling and Interactions  
in the 50S Ribosomal Subunit for the Development of Novel Antibiotics

by  
Meagan C. Small

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, Baltimore in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2015

## **Acknowledgements**

*The author would like to acknowledge the Computer-Aided Drug Design Center of the University of Maryland, Baltimore for computational resources and the Department of Pharmaceutical Sciences at the University of Maryland, Baltimore, School of Pharmacy for financial support. In addition, sincerest gratitude is extended to all the members (past and present) of the MacKerell laboratory.*

## Table of Contents

Chapter 1 Introduction: Biomolecular Force Fields in Computer-Aided Drug Design .....	1
1.1 Potential Energy Function.....	2
1.1.1 Functional Form of the Potential Energy Function.....	2
1.1.2 Extensions to the Classical Potential Energy Function .....	7
1.1.2.1 Bonded Terms.....	7
1.1.2.2 Non-Bonded Terms.....	8
1.2 Force Field Parametrization.....	11
1.2.1 Parameter Optimization for Model Compounds.....	11
1.2.2 Additional Optimization for Biomolecular Force Fields.....	16
1.3 Classes of Biomolecular Force Fields.....	17
1.3.1 Protein Force Fields .....	18
1.3.2 Nucleic Acid Force Fields .....	19
1.3.3 Carbohydrate Force Fields.....	21
1.3.4 Lipid Force Fields .....	23
1.3.5 Small Molecule Force Fields .....	25
1.3.6 Polarizable Force Fields.....	27
1.4 Applications of Force Fields to Drug Design .....	28
1.4.1 Conformationally Sampled Pharmacophore (CSP).....	29
1.4.2 Site-Identification by Ligand Competitive Saturation (SILCS) .....	30

1.5 References.....	33
Chapter 2 The Effects of Macrolide Desmethylation on Bioactivity using the	
Conformational Sampled Pharmacophore (CSP) Approach.....	
2.1 Introduction.....	58
2.2 Research Design and Methods.....	63
2.2.1 Conformationally Sampled Pharmacophore .....	63
2.2.2 Ligand-bound Molecular Dynamics .....	65
2.3 Results.....	67
2.3.1 Wild-type ribosomes.....	68
2.3.2 A2058G mutant.....	72
2.3.3 N6-methylated A2058.....	74
2.4 Discussion.....	74
2.5 References.....	79
Chapter 3 Impact of Ribosomal Modification on the Binding of the Antibiotic	
Telithromycin using a Combined Grand Canonical Monte Carlo/Molecular Dynamics	
Simulation Approach .....	
3.1 Introduction.....	91
3.2 Experimental Section .....	95
3.2.1 Parametrization details for methylated A2058 .....	99
3.2.2 System Set-up and Equilibration .....	106

3.2.3 Testing MD simulation lengths during each GCMC/MD iteration. ....	109
3.2.4 Root mean square fluctuations to determine conformational flexibility .....	113
3.3 Results and Discussion .....	114
3.4 Conclusion .....	130
3.5 References.....	135
Chapter 4 Investigation of <i>in situ</i> Click Chemistry Derived Macrolide Antibiotic Analogs in the Ribosome Using Site Identification by Ligand Competitive Saturation (SILCS) 147	
4.1 Introduction.....	148
4.2 Computational Methods.....	151
4.2.1 SILCS Simulations.....	151
4.2.2 SILCS-MC Sampling.....	154
4.3 Results and Discussion .....	156
4.4 References.....	166
Chapter 5 Polarizable Force Field for Small Ketones and Aldehydes Based on the Classical Drude Oscillator .....	
5.1 Introduction.....	174
5.2 Computational Methods.....	177
5.3 Results and Discussion .....	180
5.3.1 Nonbonded parameters .....	181
5.3.2 Bonded parameters.....	188

5.4 Conclusion .....	195
5.5 References .....	195
Chapter 6 Conclusion.....	207
Chapter 7 Bibliography.....	212

## List of Tables

<b>Table 2.1</b> Minimum inhibitory concentration (MIC) values in $\mu\text{g/mL}$ for 4,8,10-tridesmethyl TEL; 4,10-didesmethyl TEL; 4,8-didesmethyl TEL; 4-desmethyl TEL; and telithromycin (TEL). All MIC data was obtained by members of the Andrade laboratory.	69
<b>Table 2.2</b> Interaction energies (kcal/mol) between telithromycin and the <i>E. coli</i> crystal structure (PDB 3OAT). Calculations were performed for telithromycin and C4-desmethyl telithromycin with the WT, A2058G, MAD1, MAD2, and DMAD ribosomes. Interaction energies were calculated for selected ligand atoms with all the bases/residues within the truncated ribosome (see Methods in Section 2.2) as well as with base 2058 alone. Selected ligand atoms include the neutral group surrounding the C4-methyl [C3(=O)-C4(H <sub>2</sub> )-C5]. <sup>a</sup>	71
<b>Table 3.1</b> Interaction Energies (kcal/mol) and Distances ( $\text{\AA}$ ) for N6-monomethyl Adenine and Water Complexes.	101
<b>Table 3.2</b> Interaction Energies (kcal/mol) and Distances ( $\text{\AA}$ ) for N6,N6'-dimethyl Adenine and Water Complexes.	101
<b>Table 3.3</b> N6-monomethyl Adenine QM and MM Vibrational Spectra for the modes involving the methylated amino group.	103
<b>Table 3.4</b> N6,N6'-dimethyl Adenine QM and MM Vibrational Spectra for the modes involving the methylated amino group.	104
<b>Table 3.5</b> QM versus MM relative energies for MAD1 versus MAD2.	106
<b>Table 3.6</b> The number of GCMC waters within 10 $\text{\AA}$ of telithromycin averaged over the last 25 ns of GCMC/MD for varying lengths of MD. Water molecules with oxygen atoms within 10 $\text{\AA}$ of the center are considered.	111
<b>Table 4.1</b> Rank-ordering of triazoles <b>1</b> , <b>19-32</b> and azide fragment <b>2</b> by dissociation constants ( $K_d$ ) for 70S <i>E. coli</i> ribosomes in nM by fluorescence polarization along with experimental $\Delta G$ values (kcal/mol), calculated normalized Ligand Grid Free Energies (LGFEs, kcal/mol) and ligand efficiencies from SILCS. LGFEs and ligand efficiencies are calculated for the total molecule, ARM, and macrolactone+desosamine (Macro+Des) parts, where the ARM is defined as the four carbon alkyl linker and heterocycle extending from N11. The predictive indices <sup>48</sup> are calculated for each type of LGFE and ligand efficiency.	158
<b>Table 4.2</b> Sum of mass count % increases, natural logarithm of the sum %, logP, molecular weight (MW) and predicted natural logarithm of the sum % for triazoles <b>1</b> , <b>19-32</b> and azide fragment <b>2</b> (with rank ordering preserved from Table 4.1). The predictive indices <sup>48</sup> are calculated for the predicted mass count % increase versus the experimental mass count % increase.	163
<b>Table 5.1</b> Dipole moments and molecular polarizability for acetone (ACO), acetaldehyde (AALD), butanone (BTON) gauche and trans conformations, and propionaldehyde (PALD) gauche and trans conformations. QM values were obtained at the MP2/6-31+G(d)//MP2/cc-PVQZ levels of theory. C36 values were as reported previously for the CHARMM36 additive force field. <sup>52</sup>	182
<b>Table 5.2</b> Water interactions for acetone (ACO) and acetaldehyde (AALD). QM values were obtained at the MP2/6-31G(d)//MP2/cc-PVQZ levels of theory with counterpoise	

correction for basis set superposition error (BSSE).<sup>63, 64</sup> C36 values were obtained from<sup>52</sup>. Water interaction pairs are shown in **Figure 5.2** and **5.3**. Root mean square differences (RMSD) are shown. Hydrogen water interactions are not included for the methyl groups. .... 184

**Table 5.3** Water interactions for butanone (BTON) and propionaldehyde (PALD) in trans conformations. QM values were obtained at the MP2/6-31G(d)//MP2/cc-PVQZ levels of theory with counterpoise correction for basis set superposition error (BSSE).<sup>63, 64</sup> Water interaction pairs are shown in **Figure 5.2** and **5.3**. Root mean square differences (RMSD) are shown. Hydrogen water interactions are not included for the methyl groups. .... 185

**Table 5.4** Calculated (Drude) heats of vaporization ( $\Delta H_{\text{vaporization}}$ ) in kcal/mol, molar volumes  $V_m$  in  $\text{\AA}^3$ ,  $\Delta G_{\text{solvation}}$  and dielectric constants ( $\epsilon$ ) for acetone (ACO), butanone (BTON), acetaldehyde (AALD), and propionaldehyde (PALD) compared to experimental values (Exp). Properties were obtained at the same temperature as experimental conditions: ACO 298 K, BTON 314 K, AALD 283 K, and PALD 298 K for heats of vaporization and molar volumes; ACO and BTON 293.2 K, AALD 291.2 K, PALD 290.2 K for dielectric constants. Standard deviations were calculated based on the average of 4 simulations. .... 187

**Table 5.5** Bond lengths in  $\text{\AA}$  and angle lengths in degrees for acetone (ACO) and acetaldehyde (AALD) from QM optimizations at the MP2/6-31G(d) level of theory, crystallographic surveys as reported in<sup>52</sup> and Drude calculations. Crystal survey error is shown in parenthesis. .... 189

**Table 5.6** Bond lengths in  $\text{\AA}$  and angle lengths in degrees for butanone (BTON) and propionaldehyde (PALD) gauche and trans conformations from QM optimizations at the MP2/6-31G(d) level of theory and Drude calculations. .... 189

**Table 5.7** Vibrational analysis of acetone (ACO) using the internal coordinate system of Pulay<sup>77</sup>. QM frequencies were obtained at the MP2/6-31G(d) level of theory. .... 191

**Table 5.8** Vibrational analysis of acetaldehyde (AALD) using the internal coordinate system of Pulay<sup>77</sup>. QM frequencies were obtained at the MP2/6-31G(d) level of theory. .... 192

**Table 5.9** The averages and RMSF for the dipole moments of acetone (ACO), acetaldehyde(AALD), butanone (BTON), and propionaldehyde(PALD) from gas and condensed phase simulations (pure liquid and aqueous solution). .... 194

## List of Figures

<b>Figure 1.1</b> Cartoon representation of a generic molecule, where $b$ represents the bond length between two atoms, $\theta$ represents the angle between three atoms, $\chi$ is the dihedral angle between four atoms, and $r_{ij}$ is the intermolecular distance between two atoms. ....	4
<b>Figure 1.2</b> Flow chart representing the steps in the optimization of a force field. ....	12
<b>Figure 1.3</b> A) Model compounds for the parametrization of a histidine dipeptide. The dipeptide can be broken down into the backbone represented by <i>N</i> -methyl acetamide (NMA) and the side chain represented by 4-methylimidazole (MIMI). The $C\alpha$ hydrogen is omitted for clarity. After parametrization of the two model compounds, they are connected and the covalent linkage between $C\alpha$ and $C\beta$ as well as the $\phi$ , $\psi$ dihedral angles are optimized. B) An example CHARMM topology for the model compound MIMI. It contains the atom names and types, partial atomic charges, connectivity information, and specific improper terms acting on the hydrogens. C) Interactions between water and MIMI are used to obtain the supramolecular target data for optimization of the partial atomic charges. Note that while four waters are shown in the figure, the water-MIMI interactions are each evaluated individually. ....	13
<b>Figure 2.1</b> 2D representation of first and second generation macrolides: erythromycin, azithromycin, and clarithromycin. ....	60
<b>Figure 2.2</b> 2D representation of third generation macrolides: telithromycin, solithromycin, and cethromycin. ....	61
<b>Figure 2.3</b> (A) Telithromycin and A2058 interactions in <i>H. marismortui</i> with select distances in Angstroms (Steitz et. al. PDB=1YIJ). (B) Steric consequences of A2058G mutation. Image produced with VMD <sup>80</sup> . ....	62
<b>Figure 2.4</b> CSP probability distributions for telithromycin (black); C4,8,10-tridesmethyl TEL (blue); C4,10-didesmethyl TEL (purple); 4,8-didesmethyl TEL (red); and 4-desmethyl TEL (green). The vertical line corresponds to the crystallographic distances from PDB. ....	69
<b>Figure 2.5</b> CSP probability distributions for cethromycin (black) and C4,8,10-tridesmethyl CETH (red). The vertical line corresponds to the crystallographic distances from PDB 1NWX. Atom pairs represented in A through D are shown in the inset figure. ....	71
<b>Figure 2.6</b> CSP probability distributions for selected atom pairs in telithromycin (solid black) and C4-desmethyl telithromycin (solid green) in the WT <i>E. coli</i> ribosome and in A2058G mutant (telithromycin, dashed black; C4-desmethyl telithromycin, dashed green). The vertical line corresponds to the crystallographic distances from PDB 3OAT. Atom pairs represented in A through D are shown in the inset of <b>Figure 2.4</b> . ....	73
<b>Figure 3.1</b> 2D representation of telithromycin and bases within the macrolide binding pocket showing the three biologically relevant telithromycin-ribosome interactions that are the subject of this study: 2'-OH to 2058/2059 hydrogen bonding, stacking between telithromycin's 3'-protonated dimethylamine and G2505. Dotted lines represent hydrogen bonding. Hydroxyl groups of the ribose sugars are omitted for clarity. ....	94
<b>Figure 3.2</b> 2D representation of N6-monomethyl adenine (left) and N6,N6'-dimethyladenine (right) showing the numbering scheme pertaining to the parametrization. ....	100

**Figure 3.3** Potential energy scan for (left) N6-monomethyl adenine for the dihedral C5-C6-N6-C7A and (right) N6,N6'-dimethyl adenine dihedral C5-C6-N6-C7B, where C7A/B represent the exocyclic methyl carbons as shown in **Figure 3.2**. ..... 105

**Figure 3.4** The number of GCMC waters as a function of dynamics simulation time for each system studied. Number of GCMC waters were those present at the start of MD during each cycle. WT=wild-type (black), A2058G=A2058 mutated to G (red), MAD1/MAD2= the N6- monomethyl adenine with the methyl cis (green) and trans (blue) to the N1 atom, DMAD= N6,N6'-dimethyl adenine (magenta). ..... 108

**Figure 3.5** The RMSD (in Ang) for RNA (top panel), protein (middle), and telithromycin (bottom) for the WT and mutant/modifications over the 30 ns GCMC/MD simulation. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. .... 108

**Figure 3.6** The potential energy as a function of cycle for the WT and mutant/modifications over the 30 ns GCMC/MD simulation. 1500 cycles is equivalent to 30 ns dynamics simulation time. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. .... 109

**Figure 3.7** The number of GCMC waters as a function of time for varying lengths of MD performed to check for convergence. Original refers to the GCMC/MD simulation in which all 1500 cycles were run continuously using 10,000 MC steps and 20 ps MD. The 20, 50, 100, 500 ps, and 1 ns lines refer to the GCMC/MD simulations used to test for convergence and were performed by first running 100 cycles of GCMC/MD (10,000 MC steps/20 ps MD), then using the final snapshot to start 5 separate GCMC/MD simulations with varying MD steps as described in the text. For this reason, the number of GCMC waters in the latter GCMC/MD simulations are shown starting at 2 ns with the number of GCMC waters during the first 2 ns shown in cyan. The Original starts at the origin. ... 111

**Figure 3.8** The probability distributions of the telithromycin 2'OH to N1, N6 (A2058) and N6 (A2059) for the various lengths of molecular dynamics simulations performed to determine convergence. The 2'-OH – N1 (A2058) is shown in black, 2'-OH to N/O6 (A/G2058) in red, and 2'-OH to N6 (A2059) in green. Corresponding crystal structure distances are illustrated as vertical lines using the same color scheme. The two GCMC/MD simulations obtained with 20 ps MD are shown. Original refers to the GCMC/MD simulation in which all 1500 cycles were run continuously using 10,000 MC steps and 20 ps MD. The second simulation in which the coordinates were obtained from an initial 100 cycle GCMC/MD simulation is designated as such and is described in the text. .... 112

**Figure 3.9** RMS Fluctuations (Å) for RNA (top panel) and protein (bottom panel) within 10 Å of telithromycin in the WT and mutant/modifications. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, DMAD in magenta, and crystal structure values in orange. RMSF values reported are for the entire nucleotide or residue (including backbone). .... 113

**Figure 3.10** RMS Fluctuations (Å) for telithromycin in the WT and mutant/modifications. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, DMAD in magenta, and crystal structure values in orange. RMSF values reported are for non-hydrogen atoms only. .... 114

**Figure 3.11** The probability distributions of the telithromycin 2'OH to A2058 N1, A2058 N6 and A2059 N6 distances for WT, A2058G mutant and methyl modifications. The 2'-OH – A2058 N1 distance is shown in black, 2'-OH to A2058 N6/O6 in red, and 2'-OH to A2059 N6 in green. Distances reported are measured between heavy atoms. Corresponding crystal structure distances are shown as vertical lines using the same color scheme. .... 116

**Figure 3.12** The probability distributions for WT, A2058G mutant and methyl modifications of the (A) WC A752 N1 to U2609 N3 distances, (B) WC A752 N1 to U2609 N3 distances, (C) distance between the centers of masses of A752-U2609 and telithromycin's imidazole-pyridine (ARM) moiety and (D) angle between the planes defined A752-U2609 and the imidazole-pyridine moieties. The crystal structure values are shown as a vertical black line, while the distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta). .... 120

**Figure 3.13** The probability distributions of distances between telithromycin's 3'-protonated dimethylamine and G2505 P for WT, A2058G mutant and methyl modifications. The crystal structure distance is shown as a vertical black line, while the distributions are from the GCMC/MD simulations: WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta). .... 120

**Figure 3.14** Proposed communication pathways from A2058 to the A752-U2609 and ARM region: the telithromycin pathway (A), G2057 pathway (B), or A2059 pathway (C). .... 122

**Figure 3.15** The probability distributions for WT, A2058G mutant and methyl modifications of telithromycin intramolecular distances (A) C3-O to C16-O, (B) C16-O to C6-O, (C) C3-O to C6-O, and (D) C3'-N and C26 as shown in the inset figure. The crystal structure values are shown as vertical black lines, while distributions are from the GCMC/MD simulations: WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta). The probability distributions from telithromycin gas phase simulations without the ribosome are shown in orange. .... 123

**Figure 3.16** The number of conformations for WT, A2058G mutant and methyl modifications of (A) COM distances between the base atoms of 2057 and 2058 and (B) angles between planes comprised by 2057 and 2058 base atoms. The probability distributions for (C) WC G2057 N1 to C2611 N3 distances. The crystal structure values are shown as vertical black lines, while distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta). .... 124

**Figure 3.17** The probability distributions for WT, A2058G mutant and methyl modifications of the telithromycin C2-methyl to C2610 C4' (black), C4-methyl to C2611 C5 (red), and C4-methyl to A2058 C2 (green). Corresponding crystal structure values are shown as vertical lines using the same color scheme. .... 126

**Figure 3.18** The number of conformations for WT, A2058G mutant and methyl modifications of (A) COM distances between base atoms of 2058 and 2059 and (B) angles between planes comprised by 2058 and 2059 base atoms. The probability distributions for all systems studied of the (C) telithromycin C26 to A2062 C2 distances and (D) telithromycin C6-methyl to 2059 C2 (black), C6-methyl to 2062 C2 (red), C8-

methyl to 2058 C2' (green), and C10-methyl to 2058 C2' (blue). The crystal structure values are shown as vertical black lines in (A – C) and in (D) using the same color scheme, while distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta) in (A – C)..... 128

**Figure 3.19** The probability distributions of distances for A2059 C2 to A2062 N6. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. The crystal structure distance is shown as a black vertical line..... 129

**Figure 3.20** The probability distributions for the distances between telithromycin O16 to A2059 P (black), O16 to Lys90 NZ (red), and O18 to A2062 N6 (green). Corresponding crystal structure distances are illustrated as vertical lines using the same color scheme. The inset figure illustrates the distances and orientations of the atoms from the crystal structure. O16 refers to C16-ketone, O18 refers to C9-ketone. .... 129

**Figure 3.21** The probability distribution of distances between 2058 N6/O6 and a nearby  $Mg^{2+}$  ion that chelates the phosphate groups of G2056 and G2057 and two crystallographic waters. Note that  $Mg^{2+}$  and crystallographic waters were included in the simulations. Waters are excluded in the image for clarity. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. The crystal structure distance is shown as a vertical black line. .... 132

**Figure 4.1** 2D structures of fragment precursors for the *in situ* click chemistry. (A) The full ligand after the cycloaddition reaction including the regioisomers anti (1,4)- and syn (1,5)- analogs and (B) alkyne-derived fragments. .... 151

**Figure 4.2** The individual contributions to the total LGFE attributed to the macrolactone+desosamine and ARM moieties. .... 159

**Figure 4.3**  $\Delta G$ s (kcal/mol) versus the normalized macrolactone+desosamine LGFEs (kcal/mol). The red line indicates the linear regression of the data, which represents an  $R^2$  of 0.07. **Equation for  $\Delta G$  versus Normalized Macrolactone+Desosamine LGFE:**  
 $y = 0.37 x - 30.75$ . .... 160

**Figure 4.4** The correlation of the  $\Delta G$  values for the full ligand to the experimental click enhancements for the alkyne-derived fragments..... 162

**Figure 4.5** The experimental mass count % increases versus the predicted mass count % increases. The red line indicates the linear regression of the data, which represents an  $R^2$  of 0.64. **Equation for predicted click enhancement versus experimental click enhancement from multiple regression analysis:**  $\ln \text{sum \%} = 1.12 \text{ LGFE} - 0.12 (\log P)^2 + 0.06 \text{ MW} + 6.30$ ..... 165

**Figure 5.1** Acetone (ACO), acetaldehyde (AALD), butanone (BTON), and propionaldehyde (PALD) with corresponding atom types from the CHARMM Drude FF. Methyl hydrogens are omitted for clarity. .... 177

**Figure 5.2** Water interaction pairs for acetone (ACO) corresponding to pairs in **Table 5.2**..... 184

**Figure 5.3** Water interaction pairs for acetaldehyde (AALD) corresponding to pairs in **Table 5.2**..... 184

**Figure 5.4** Dipole moments as a function of time for acetone (ACO), acetaldehyde (AALD), butanone (BTON), and propionaldehyde (PALD) in gas phase (left column), pure liquid (middle column), and aqueous solution (right column). .... 194

## **Chapter 1 Introduction: Biomolecular Force Fields in Computer-Aided Drug**

### **Design**

Force field-based simulations have become an essential component of *in silico* drug design and hence will be a major focus of this introduction. Simulations are used to generate an ensemble of protein and/or drug conformations that represent the accessible conformations of the molecules, from which various properties are calculated, as in determining ligand binding affinities or conformational properties that may be used to relate structure and physical features to activity. Such information allows for improved interpretation of experimental results, thereby facilitating qualitative ligand design, as well as making more quantitative predictions about changes in binding affinities and efficacies. Accordingly, *in silico* methods save money, time, and resources.

Ultimately, the accuracy of a simulation is based on the quality of the force field utilized with respect to the biomolecules and ligands of interest. Force fields vary in the form of the potential energy functions and the assumptions and methodologies used in the optimization of the parameters in those energy functions. In addition, they are often designed for a specific purpose or subset of molecules. Though a detailed description of all available force fields is beyond the scope of this chapter, its aim is to present the foundations of force fields (FF) commonly used in computer-aided drug design (CADD), their representations of biomolecules, and their use in selected applications in drug design. The chapter begins with an overview of the potential energy functions and summarizes extensions to potential energy functions that have been developed to increase

their accuracy. Then, the different classes of biomolecular force fields will be discussed including a description of how force fields are optimized, with emphasis on the CHARMM additive force fields that have been used in all of the studies presented herein. Finally, the chapter will conclude with a discussion of a number of applications of biomolecular force fields to drug design that have been developed in our laboratory and applied here to the ribosome and several macrolide antibiotics: Site Identification by Ligand Competitive Saturation (SILCS) and Conformationally Sampled Pharmacophore (CSP).

## **1.1 Potential Energy Function**

### ***1.1.1 Functional Form of the Potential Energy Function***

A force field is the potential energy function and the collection of parameters that have been optimized for use in that potential energy function. Together, the function and parameters are used to evaluate the potential energies and forces of molecules in a simulation. In general, a potential energy function is comprised of bonded terms that describe the bond, valence angle, dihedral or torsion angle, and, in some cases, out-of-plane torsion (or improper dihedral angle) contributions to the potential energy and forces, and non-bonded terms that describe the electrostatics and van der Waal's (VDW) contributions. Force fields vary based on the form of the potential energy function that is employed, but the most widely used in CADD is shown in **Equation 1.1**. This general form is used in AMBER<sup>1</sup>, CHARMM<sup>2</sup>, OPLS<sup>3</sup>, and GROMOS<sup>4</sup>, and is known as a Class

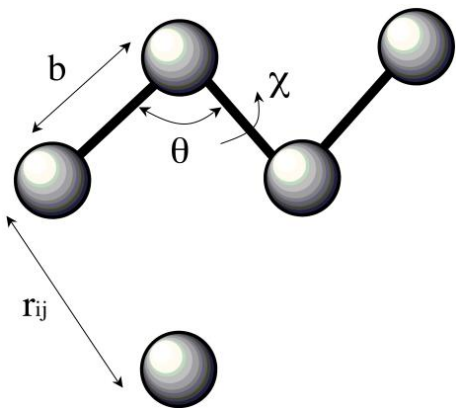
I force field.

$$U(\vec{R}) = \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi (1 + \cos(\eta\chi - \delta)) + \sum_{impropers} K_{imp} (\varphi - \varphi_0)^2 + \sum_{nonbond} \left( \varepsilon_{ij} \left[ \left( \frac{Rmin_{ij}}{r_{ij}} \right)^{12} - \left( \frac{Rmin_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\varepsilon r_{ij}} \right) \quad \text{Equation 1.1}$$

**Equation 1.1** may be broken down into terms describing the parameters in the model and the 3D structure,  $R$ , of the molecules. For a generic molecule as schematically shown in **Figure 1.1**, the 3D structure may be considered in terms of the bond lengths,  $b$ ; valence angles,  $\theta$ ; dihedral angles,  $\chi$ ; out-of-plane torsion angles,  $\varphi$ ; and distances between atoms  $i$  and  $j$ ,  $r_{ij}$ . These terms may come from a crystal or NMR structure or from a molecular model. The remaining terms in **Equation 1.1** are force field parameters, which may be separated into the bonded contributions (internal/intramolecular terms) or non-bonded (external/intermolecular/interaction terms). The energy function and parameters combined with the 3D structure of the molecule allow determination of the potential energy and forces.

Parameters are unique for different types of atoms so that the simple, computationally efficient function shown in **Equation 1.1** can be used for a wide range of chemical entities. For example, parameters for the C-C bond in ethane differ from the C=C bond in ethene. Since the two types of carbons have their own parameters (namely equilibrium bond length and bond force constant), the energy and forces for the two molecules can be calculated using **Equation 1.1**. The simplicity of the classical potential

energy function not only makes large-scale simulations computationally accessible, but exhibits a high degree of accuracy for biomolecules given an appropriately optimized collection of parameters.<sup>5</sup>



**Figure 1.1** Cartoon representation of a generic molecule, where  $b$  represents the bond length between two atoms,  $\theta$  represents the angle between three atoms,  $\chi$  is the dihedral angle between four atoms, and  $r_{ij}$  is the intermolecular distance between two atoms.

As seen in **Equation 1.1**, the bond and valence angle contributions to the potential energy are calculated using a harmonic functional form, which assumes values oscillating relatively close to equilibrium. For most simulations of biomolecules, harmonic treatment is appropriate because the simulations are performed at room temperature and do not include breaking or formation of bonds, thus their conformations remain close to their equilibrium values. The associated parameters include the bond force constant,  $K_b$ , and equilibrium bond length,  $b_0$ , and the valence angle force constant,  $K_\theta$ , and equilibrium valence angle,  $\theta_0$ . The potential energy due to rotation about bonds (or torsions), **Figure 1.1**, is computed through the dihedral angle term. Because rotation is through  $360^\circ$  and oscillates between a range of values, the mathematical form is a

cosine function, in which the dihedral force constant,  $K_{\chi}$ , is the amplitude indicating the barrier to rotation; the multiplicity,  $n$ , is the periodicity dictating the number of cycles of rotation per  $360^{\circ}$ ; and the phase angle,  $\delta$ , is the constant controlling the location of the maxima of the energy surface. In most force fields the phase angle is kept at 0 or 180 degrees to maintain symmetry as required to apply the parameters to diastereomers. However, the option of using other values for the phase is available in most programs.<sup>6</sup> For rotation about the C-C bond in ethane from the aforementioned example, the multiplicity  $n$  equals 3 and the rotations generate two low energy staggered conformations and a high energy eclipsed conformation. The dihedral energy about an individual bond may be described by more than one dihedral term by summing the cosine functions into a Fourier series, **Equation 1.1**, that is useful in reproducing, for example, a target quantum mechanical (QM) dihedral potential energy surface. The potential energy associated with out-of-plane torsions (or improper angles) is also calculated harmonically and its parameters include an improper angle force constant,  $K_{imp}$ , and equilibrium improper angle,  $\varphi_o$ . The existence of the improper angle term is primarily to maintain planarity and prevent inversion of the molecule's chirality, especially in force fields that do not explicitly contain all hydrogen atoms (e.g. the CHARMM Param19 FF)<sup>7</sup>.

The non-bonded terms in **Equation 1.1** represent the most important contributions to the potential energy and include the van der Waals (VDW) and electrostatic interaction energy. For both the VDW and electrostatic interaction energies, the interactions between atoms covalently bonded to each other or separated by two covalent bonds (i.e. 1,2 and 1,3 interactions, respectively) are ignored. In most force fields, the VDW contribution to the potential energy is calculated using a Lennard-Jones

6-12 function (LJ 6-12, **Equation 1.1**), though alternative terms have been used (see below). In the LJ 6-12 function the repulsive term is a “hard” wall potential of  $1/r^{12}$  enforcing the Pauli exclusion principle that electron clouds cannot overlap and the attractive term is  $1/r^6$ , which approximates the London dispersion energy. The interaction parameters  $\epsilon_{ij}$  and  $R_{min,ij}$  describe the Lennard-Jones well depth and the minimum interaction distance between two atoms  $i$  and  $j$ , respectively. The well depth is a measure of how favorable the London dispersion force is between the atoms, while the minimum interaction distance is dependent on the VDW radius of the atoms. Generally  $\epsilon_i$  and  $R_{min,i}$  are obtained for individual atom types (such as for the carbon and hydrogen atom types of ethane), then combined according to a set of combining rules to yield the pairwise  $\epsilon_{ij}$  and  $R_{min,ij}$ . The combining rules vary among force fields as the arithmetic mean or the geometric mean. These differences, along with other considerations (see below), make it inappropriate to mix parameters between force fields. For pairwise interactions that are not described well by use of the standard combination rules, atom-pair specific (NBFIX) LJ parameters can be applied.<sup>8</sup>

The electrostatic contribution to the potential energy is evaluated using Coulomb's law. In additive, pairwise force fields the electrostatic parameters include the partial atomic charges of the atoms,  $q$ . In force fields that explicitly treat electronic polarizability, additional terms to treat the atomic polarizability are included, as discussed below. When applying Coulomb's law with an explicit solvent representation the dielectric constant  $\epsilon$  is set equal to 1. However, there are a wide range of approximations, referred to as implicit solvent models, to avoid the use of explicit solvent thereby offering significant savings in computational costs. A number of recent reviews have addressed

these models.<sup>9, 10</sup>

### ***1.1.2 Extensions to the Classical Potential Energy Function***

When selecting a force field to use, the system under study and the desired outcomes must be taken into consideration. In certain cases a more complex form of the potential energy function may be desired. These are known as Class II force fields. Historically, limited computer resources as well as increasing difficulty in optimization of the parameters in these force fields made expanded treatment of the potential energy function difficult. Extensions of the potential energy function can be made to either the bonded or non-bonded terms, which are described below. A recent review describes advances in potential energy models in more detail.<sup>11</sup>

#### ***1.1.2.1 Bonded Terms***

The harmonic form of the bond potential energy is less accurate for molecular geometries far from equilibrium, which can be accounted for using the Morse potential (**Equation 1.2**) that includes bond breaking and bond dissociation energy. Bond stretching asymmetry can be treated using cubic and quartic extension terms. Linear molecules such as hydrogen cyanide and carbon dioxide are better treated with a cosine functional form for the angular potential energy rather than the harmonic valence angle terms.<sup>12, 13</sup> Cross terms that relate the interdependence of the bonds, angles, and dihedrals to each other can also be used. For instance, the conformational properties associated

with correlation of two adjacent dihedrals can be treated using a grid-based dihedral energy correction map (CMAP), used mostly to enhance the conformational properties of protein backbone dihedrals  $\phi$  and  $\psi$ .<sup>14-17</sup> Recently, CMAP in conjunction with Hamiltonian Replica Exchange has been used to improve conformational sampling of peptides and carbohydrates.<sup>18-20</sup> While these enhancements more accurately treat the change in energy with intramolecular geometry, for the most part the Class I additive potential energy function is satisfactory for CADD.

$$U(\text{bond}) = D_e \left( 1 - e^{-\sqrt{\frac{K_b}{2D_e}}(b - b_0)} \right)^2 \quad \textbf{Equation 1.2}$$

#### 1.1.2.2 Non-Bonded Terms

Improvements to the accuracy of the potential energy function can also be achieved via higher order non-bonded functions. Non-bonded terms represent the most important contribution to the potential energy because electrostatics and van der Waal's terms dominate interactions dictating the structure and dynamics of biopolymers as well as ligand-protein interactions and related phenomena. They also are the most computationally expensive to calculate and non-bonded terms of a higher level of accuracy can quickly become cost-prohibitive due to the  $N \times N/2$  calculations required, where  $N$  is the number of particles in the system. Concerning the VDW interactions, in most cases, the LJ 6-12 term is sufficient to treat biomolecules, but given interatomic

distances  $r_{ij}$  smaller than  $R_{min,ij}$  the energy increase is over estimated as  $r$  decreases.<sup>21</sup> Modified potentials have sought to “soften” the repulsive part of the VDW contribution using a 9th order repulsive term (LJ 6-9) or modifying the functional form as in the Buckingham potential (or Exp-6 potential),<sup>22</sup> which treats the repulsion as an exponential that gives a more gradual increase in energy as  $r$  decreases.<sup>23</sup> Additionally, the buffered 14-7 potential<sup>21</sup> uses buffering constants that modulate the repulsive and attractive terms of the LJ potential and is the functional form of the VDW potential energy employed in the Merck Molecular Force Field (MMFF),<sup>24</sup> with a variation of that function used in the polarizable AMOEBA FF.<sup>17, 25, 26</sup>

The inclusion of higher order terms for the electrostatic interactions warrants a more extensive discussion than the previous potential energy terms. A majority of force fields utilize Coulomb's law as in **Equation 1.1**. Partial atomic charges on each atom are static and the contribution of the electrostatic interactions to the potential is the sum of all of the pairwise Coulombic interactions (excluding 1,2 and 1,3 interactions). Hence, this is known as the additive model. Additive force fields do not capture the contributions of polarizability explicitly. This is a major drawback because the electron distribution of a molecule is influenced by the surrounding environment. Thus, models that include the explicit treatment of polarization have been the subject of ongoing work for over 30 years.<sup>27</sup> To overcome the fixed charged approximation in additive force fields polarization is treated in a mean-field way by adjusting the partial atomic charges to overestimate the gas phase molecular dipoles and hence the electrostatic interactions of the molecule in aqueous systems are approximated. While including polarizability implicitly has fortuitously shown good agreement between condensed phase and

experimental molar volumes and heats of vaporization<sup>28-32</sup> as well as solvation free energies<sup>33-36</sup>, there is room for improvement for biomolecules.<sup>37</sup> Indeed, recent advances in polarizable force fields are now yielding improvements in a range of systems.<sup>25, 38-45</sup>

Explicit treatment of polarizability introduces a new term into the potential energy function,  $U_{polar}$ .<sup>46, 47</sup> The polarizability needed to determine  $U_{polar}$  may be calculated using induced dipoles,<sup>48-52</sup> fluctuating charges,<sup>53-59</sup> and the classical Drude oscillator.<sup>46, 60, 61</sup> In the induced dipole model, the functional form of  $U_{polar}$  is based on introducing an induced dipole onto each atom in addition to the partial atomic charge. This is performed in the AMOEBA polarizable FF, in addition to treatment of the static contribution to the electrostatics with a multipole expansion out to quadrupoles.<sup>17, 25, 26</sup> Fluctuating charge models are based on allowing the partial atomic charges to fluctuate in response to the electric field. The polarization energy is related to the absolute (Mulliken) electronegativity<sup>62, 63</sup> and the hardness of the atom<sup>64</sup>, which themselves are dependent on the electron affinity and ionization potential of an atom. Though their names imply a change in the number of electrons, in the context of chemical polarization these are measures of how likely the electronic distribution is polarized and are easier to understand in terms of two atoms.

Models based on the classical Drude oscillator<sup>65-67</sup> (also known as Shell or charge-on-spring (COS) models) account for polarization by introducing a massless, partially charged Drude particle that is harmonically attached to the nucleus of the parent atom. The atomic polarizability is simply the Drude charge squared divided by the force constant on the harmonic term between the Drude particle and the atomic core. In practice, the charge on the atom and the Drude particle are corrected to account for the

partial atomic charge on the atom. Thus, for a fixed set of atomic positions, the Drude particles can relax in the surrounding electric field yielding the polarization response. This relaxation can be performed via energy minimization which is equivalent to a self-consistent field calculation, although computationally efficient methods are used to treat the Drude particles in MD simulations, as discussed in the next section.

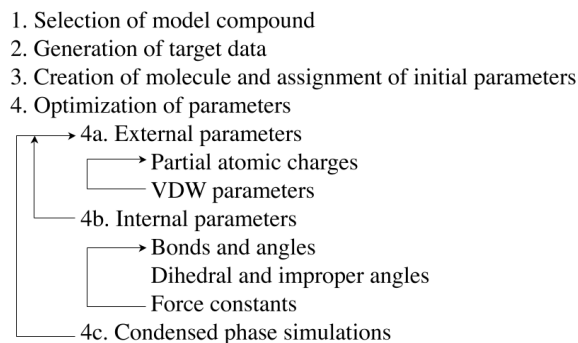
## **1.2 Force Field Parametrization**

The usefulness of a force field (FF) in target-based drug design is based on the availability of parameters for the target molecule, typically a protein, and the ligands under study. These aspects of the force field must be compatible with each other, which is based on their parameters being optimized using a consistent approach. Such an approach has been taken with different components of the CHARMM additive force field. In the remainder of this section we give an overview of the parameter optimization approach for model compounds representative of biomolecules and of the range of functional groups in drug-like molecules. This will be followed by an overview of the additional optimization steps required for macromolecules focusing on the recent update of the CHARMM36 protein FF.<sup>14-16</sup>

### ***1.2.1 Parameter Optimization for Model Compounds***

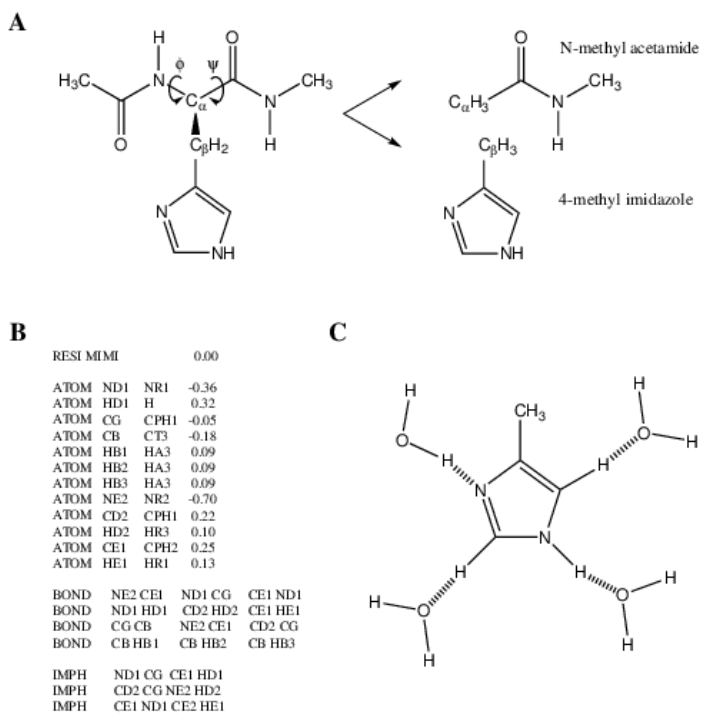
Optimization of a force field follows the general steps in **Figure 1.2**; a more detailed overview of the parameter optimization process in the context of the CHARMM

additive model may be obtained elsewhere.<sup>68</sup> To illustrate this procedure, we use histidine as an example. In Step 1, histidine, which is treated in its dipeptide form (i.e. N-acetyl, C-methylamide), is broken down into two model compounds: N-methyl acetamide (NMA) for the backbone and 4-methylimidazole (MIMI) for the side chain, **Figure 1.3A**.



**Figure 1.2** Flow chart representing the steps in the optimization of a force field.

To generate target data (Step 2), we utilize experimental data such as crystal structures and condensed phase properties, which we supplement with QM data. QM calculations include minimized NMA and MIMI structures and vibrational spectra, as well as supramolecular data for water or dimer interactions. In Step 3, the topology is created and initial parameters are assigned as illustrated in the example CHARMM topology in **Figure 1.3B**. The MIMI topology entry contains the atom names and types, partial atomic charges, connectivity information, and, in special cases, specific improper terms.



**Figure 1.3** A) Model compounds for the parametrization of a histidine dipeptide. The dipeptide can be broken down into the backbone represented by *N*-methyl acetamide (NMA) and the side chain represented by 4-methylimidazole (MIMI). The  $C_{\alpha}$  hydrogen is omitted for clarity. After parametrization of the two models compounds, they are connected and the covalent linkage between  $C_{\alpha}$  and  $C_{\beta}$  as well as the  $\phi$ ,  $\psi$  dihedral angles are optimized. B) An example CHARMM topology for the model compound MIMI. It contains the atom names and types, partial atomic charges, connectivity information, and specific improper terms acting on the hydrogens. C) Interactions between water and MIMI are used to obtain the supramolecular target data for optimization of the partial atomic charges. Note that while four waters are shown in the figure, the water-MIMI interactions are each evaluated individually.

Once the topology information is in place the appropriate parameters must be identified. There are a number of available algorithms that can assign initial guess parameters by analogy and are compatible with existing force fields.<sup>69-78</sup> These include the ParamChem/CHARMM General Force Field engine<sup>74, 75</sup> and MATCH<sup>77</sup>, both of which may be used with the CHARMM additive force fields. When performing this task manually, the program CHARMM<sup>79</sup> automatically identifies missing parameters during

the structure generation process. With the topology built and initial parameters assigned, the partial atomic charges and VDW parameters are then optimized as in Step 4a. Water interactions for NMA and MIMI (**Figure 1.3C**) provide interaction energies and geometries that are used to adjust the partial atomic charges until good agreement with the QM supramolecular distances and energies is observed. VDW parameters are optimized using gas and condensed phase simulations of NMA and MIMI to target molar volumes and heats of vaporization.<sup>36</sup> Alternatively, crystal or hydration free energies as well as QM data may be used for VDW parameter optimization, which may be considered the most difficult aspect of the parameter optimization process. However, in the majority of cases VDW parameters may be directly transferred from known molecules without additional optimization.

Once the external parameters have been optimized, the bonded parameters are adjusted (Step 4b). Bond and angle equilibrium values are adjusted until good agreement with geometries from QM calculations or surveys of the Cambridge crystallographic database<sup>80</sup> is achieved. The bond and angle force constants are parametrized by comparing to experimental and QM vibrational spectra. Vibrational spectra in CHARMM may be analyzed using the MOLVIB module<sup>81</sup>, which creates a potential energy distribution (PED) matrix<sup>82</sup> that breaks down each frequency into its contributing normal modes and allows straightforward analysis of the contribution of internal degrees of freedom to the spectra. Dihedral parameters are optimized against vibrational spectra and, for the “softer” degrees of freedom (e.g. rotation about single bonds), using potential energy surfaces (PES) for rotations around specific bonds. PES are typically determined for torsions involving only non-hydrogen atoms, with the dihedral parameters adjusted

until the molecular mechanical (MM) PES agrees with the QM surface. In addition, dihedrals associated with terminal hydroxyl or sulfhydryl groups are also often optimized. Typically the phase for each dihedral angle is limited to either 0 or 180, with the focus of the optimization on the force constant and multiplicity, with additional dihedral terms added to the Fourier series as necessary. For example, the dihedral Fourier series for the peptide bond typically includes the expected 2-fold term ( $n = 2$ ) to treat the double-bond character of the amide bond, with a 1-fold term included to fine tune the relative energies of the cis versus trans states of the peptide bond.

At this point, the first round of parameter optimization is completed. To check whether the changes to the bonded terms have affected our external parameters, it is necessary to loop over Step 4 again and determine whether the deviations are within defined convergence criteria. For instance, in the CHARMM General Force Field (CGenFF)<sup>68</sup>, deviations for the bond and valence angles of 0.03 Å and 3° from the QM target data, respectively, are acceptable while differences of 5% or less between the QM and MM vibrational frequencies is ideal. The final step in the parametrization of histidine is to connect the model compounds and optimize the parameters that are involved in the linkages, for instance the C $\alpha$  – C $\beta$  bond and  $\phi$ ,  $\psi$  dihedrals as noted in **Figure 1.3A**. For the latter, QM target data for dipeptides serves as a useful model for optimizing the backbone dihedrals.<sup>15, 16</sup>

As a final note on parametrization, it is important that parameters from different force fields not be combined due to the treatment of the VDW and electrostatic terms. For example, OPLS and CHARMM partial atomic charges are targeted to water interactions with the compound obtained at the HF/6-31G\* level of theory while charges in AMBER

are fit to restrained electrostatic potentials (RESPs)<sup>83</sup> of the same QM level. This leads to variations in the charge distributions between the force fields, which may impact the nature of the atomistic interactions of the molecule with its surroundings.<sup>33</sup> Additionally, for the LJ combining rules CHARMM and AMBER use the geometric mean for epsilon and arithmetic mean for  $R_{min}$ , while OPLS uses the geometric mean for both.<sup>3</sup> OPLS also defines the radius, sigma, where the LJ energy between two atoms is zero, whereas the  $R_{min}$  in CHARMM and AMBER is where the LJ energy is at its minimum. Force fields also use different scaling factors for 1,4 non-bonded interactions.<sup>3, 84</sup>

### ***1.2.2 Additional Optimization for Biomolecular Force Fields***

Once the model compounds have been parametrized, they are compiled into a force field. Following our example of histidine, the dipeptides of all the amino and imino acids are combined to yield a protein force field. At this stage it is crucial to validate the quality of the protein force field using condensed phase simulations of peptide and proteins for which experimental data is available. When necessary, further optimization is performed to balance local and global macromolecular properties. For instance, additional optimization<sup>14, 16</sup> of the CHARMM22 protein force field<sup>2</sup> was undertaken to reduce the relative population of  $\pi$  helix that was observed in longer simulations.<sup>85</sup> First, the alanine, glycine, and proline dipeptides were used as model systems to refine the parameters for the  $\phi$ ,  $\psi$  backbone dihedrals based on QM data. The  $\phi$ ,  $\psi$  backbone distributions were improved by introducing a new  $\phi$ ,  $\psi$  dihedral cross-term to the potential energy function that is a grid-based energy correction map (CMAP)<sup>16</sup> that

allows for near ideal reproduction of dipeptide QM energy surfaces. Then, additional adjustments were performed to fine tune the conformational sampling, yielding the CHARMM22-CMAP protein force field in 2004.<sup>15, 16</sup> Subsequently, additional optimization was undertaken to correct for the overstabilization of helices leading to CHARMM36.<sup>14-16</sup> CHARMM36 refines the CMAP by further adjusting the Ala dipeptide CMAP surface targeting the conformational sampling of the (Ala)<sub>5</sub>, GB1, and CH<sub>3</sub>-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptides in aqueous solution. In addition, higher level QM calculations on the glycine and proline dipeptides were used to define their CMAP terms. Additional optimization of CHARMM36 involved the dihedral parameters associated with the  $\chi_1$ ,  $\chi_2$  side chain torsions. This optimization targeted both QM data, NMR data and  $\chi_1$ ,  $\chi_2$  distributions from a survey of the protein databank.<sup>86</sup> The parameters were then validated using simulations of test proteins in crystal, aqueous, and denaturing environments. This effort illustrated that longer simulations are required to guide refinements of force field parameters and that additional optimization leads to improved representation of biomolecules by a force field.

### **1.3 Classes of Biomolecular Force Fields**

Force fields (FF) vary in the numerical approximations and target data used in their development as discussed above. They are often designed targeting a subset of molecules. In this section, the major force fields available for biomolecular simulations will be briefly discussed with particular emphasis given to the CHARMM, AMBER, OPLS, and to a lesser extent the GROMOS force field. Additional force fields will be

presented as they relate to the biomolecules. This summary is not intended to be a thorough review on all of the available biomolecular force fields. For that, the reader is referred to additional articles.<sup>5, 87-90</sup>

### ***1.3.1 Protein Force Fields***

The first molecular dynamics (MD) simulation of a biological macromolecule was a protein<sup>91</sup> and proteins continue to be the most studied biological molecules to date for *in silico* drug design. The AMBER<sup>1, 92-95</sup>, CHARMM<sup>2, 15, 16</sup>, and OPLS<sup>31, 96</sup> protein force fields are all-atom, while GROMOS<sup>97-101</sup> incorporates the hydrogens into the parent heavy atoms (united atom), with the exception of polar hydrogens. These are discussed in great detail in a comprehensive protein force field review by Ponder and Case<sup>90</sup>.

In general, all the aforementioned force fields tend to reproduce experimental 3D structures of proteins.<sup>102</sup> CHARMM and AMBER were optimized to reproduce QM data and experimental data based on surveys of crystallographic databases on model compounds representative of proteins. OPLS uses the internal parameters from AMBER (ff94),<sup>1</sup> from which select torsions are optimized using QM data.<sup>31, 96</sup> For the LJ parameters, all of the force fields use condensed phase simulations. While CHARMM and AMBER use the TIP3P water model,<sup>103</sup> OPLS was parametrized to work primarily with TIP3P,<sup>103</sup> though it has been used with TIP4P<sup>103</sup> and related models.<sup>30, 103-106</sup> GROMOS is typically used with the SPC model.<sup>105</sup>

One of the considerations in force field development is that the data used during the parametrization process can bias the results. Therefore, the user must critically

analyze the outcomes of force field simulations to determine whether particular interactions may be artifacts of force field bias. In a comparison of OPLS, AMBER, GROMOS, and CHARMM,<sup>85</sup> it was found that the relative energy barrier between the  $\alpha$  to  $\pi$  helix state was underestimated compared to high level QM calculations by 1 to 2 kcal/mol. Thus, in many force field simulations of proteins, the relative population of the  $\pi$  helix was higher than expected. This observation led to the creation of the CMAP terms in CHARMM<sup>16</sup> as discussed above. Similarly, force fields have been found to be either too helical or not helical enough.<sup>107</sup> The CHARMM36 protein FF has addressed this targeting the aqueous phase conformational sampling of multiple peptides, as discussed above. With the AMBER FF a number of variants have been developed (e.g. ff96<sup>94</sup>, ff99<sup>95</sup>, AmberGS<sup>108</sup>, Amber03<sup>92</sup>, Amber03\*<sup>109</sup>, ff99SB<sup>93</sup>, ff99SB\*<sup>109</sup>, ff99SB-ildn<sup>110</sup>, ff99SBnmr1<sup>111</sup>) with the more recent force fields being better behaved with respect to polypeptides in solution, as with CHARMM36. Limitations in all the protein force fields have become evident because of the growing computational resources and longer simulation times for biomolecules, observations that ultimately lead to improvements in the models over time.

### ***1.3.2 Nucleic Acid Force Fields***

Early on, the charged nature of the nucleic acid backbone made optimization of nucleic acid force fields a challenge.<sup>5, 112</sup> Simulations using the early nucleic acid force fields such as AMBER<sup>113</sup>, CHARMM<sup>114</sup>, and GROMOS<sup>4</sup> achieved mediocre success due to the stability of oligonucleotides. The use of appropriate truncation methods<sup>115</sup> and

particle mesh Ewald to treat long range electrostatics<sup>116</sup> later allowed for longer, stable MD simulations to be performed.<sup>117, 118</sup> Updates to the AMBER and CHARMM nucleic acid force fields yielded AMBER (ff94)<sup>1</sup> and CHARMM22.<sup>119</sup> These were subsequently updated to AMBER (ff99)<sup>120</sup> and CHARMM27<sup>121, 122</sup> that were developed to improve the sugar pucker and equilibrium between the A and B canonical forms, respectively. Notably, CHARMM27 was a full reoptimization of the CHARMM22 nucleic acid force field to balance the energetics of the model compounds based on QM data with the conformational properties of duplex DNA in solution, while AMBER (ff99) updated the torsion parameters pertaining to the sugar pucker. Most recently, updates have been made to CHARMM27 to better treat the 2'OH group on RNA and the equilibrium between the BI and BII DNA states, leading to CHARMM36.<sup>123, 124</sup> At the present time, the most recommended AMBER nucleic acid force field is ff99bsc0<sup>125</sup> for canonical DNA and RNA, which is a refinement of ff99 intended to correct overpopulation of the  $\gamma$ =trans state. For noncanonical RNA, the ff99bsc0 $\chi_{OL}$  which involves adjustments of the glycosidic linkage parameters is recommended.<sup>126</sup>

Another nucleic acid force field of note is the Bristol-Myers-Squibb (BMS) nucleic acid force field.<sup>127</sup> BMS uses the internal parameters from CHARMM27 for the nucleobases and CHARMM/Quanta (a commercial CHARMM force field by Accelrys, Inc.) for the sugar and phosphate backbone internal parameters. To date, the BMS force field has not seen wide use, in part due to it not being compatible with a wider range of biomolecules.

One caveat of force fields, as mentioned in the previous section, is that the target data can bias the results from a simulation. For instance, the extent to which the above

force fields utilize QM target data for the larger model compounds and experimental structures for DNA/RNA duplexes varies. The BMS nucleic acid force field uses mostly the experimental nucleic acid structures as obtained from the Nucleic Acid Database (NDB).<sup>128, 129</sup> For this reason, simulations done with the BMS force field tend to have a predominantly B form of DNA. CHARMM27/36 uses both the A and B form structures from the NDB and the QM conformational energies of larger model compounds. Hence, the model yields an improved equilibrium between A and B form DNA. The most recent CHARMM36, in fact, is optimized to yield an equilibrium between the BI and BII states of DNA as well as properly treating the equilibrium between A and B form DNA. AMBER (ff94) did not use experimental duplex DNA/RNA structures during parametrization, though AMBER (ff99) revision did incorporate this into the optimization of the final dihedrals. Still, it was found that AMBER (ff99) led ultimately to degradation of the B form, leading to the most recent variant of the AMBER force field (ff99bsc0) that addresses B form stabilization using high level QM calculations of larger model compounds that are more representative of nucleotides. For noncanonical structures ff99bsc0 has been noted to lead to incorrect loop geometries and other structural inconsistencies,<sup>130</sup> which have led to  $\chi$  modifications introduced in the  $\chi$ OL variant<sup>126</sup> and ff99 $\chi$ <sup>131</sup>.

### ***1.3.3 Carbohydrate Force Fields***

The parametrization of sugar FFs has been challenging for two reasons. First, there are many types of monosaccharides and, secondly, the chemical nature of sugars

makes it challenging to balance the inter- and intramolecular hydrogen bonding that occurs via the sugar hydroxyls in water. For polysaccharides, this effect is amplified by the different types of glycosidic linkages as well as the  $\alpha$  and  $\beta$  anomers. The two most used force fields for sugars are the AMBER GLYCAM<sup>132-135</sup> and CHARMM carbohydrate<sup>136-140</sup> force fields. The original AMBER GLYCAM\_93<sup>135</sup> was optimized for model compounds derived from tetrahydrofuran. Partial atomic charges were fit to RESP maps. In the most recent GLYCAM06<sup>132, 134</sup> problems with solvation and diffusion rates were overcome. The force field was optimized to generate a transferable set of parameters that were not specific to one type of monosaccharide using a large training and test set of molecules from a diverse group of chemical families. A recent update to the GLYCAM06 force field has been developed for lipopolysaccharide membranes.<sup>132</sup>

Another complication in force field-based simulations of carbohydrates is that a carbohydrate force field must be compatible with existing protein and lipid force fields because carbohydrates involved in recognition are generally covalently attached to lipids or proteins. Therefore, simulations are done using combinations of carbohydrate and lipid/protein parameters. While GLYCAM06 is general enough to study a diverse set of biomolecules such as proteins, nucleic acids, lipids, and small molecules, there are some inconsistencies with the AMBER protein FF because the 1,4 scaling factor of 1/1.2 is excluded in GLYCAM06 potentially leading to altered conformational energies. The CHARMM carbohydrate force fields are designed specifically to be compatible with the remainder of the additive CHARMM biomolecular force field. The advantage is that when heterogeneous systems are to be studied, glycoprotein or glycolipid simulations can be performed using parameters that have been specifically optimized for each

biomolecule, including 1-4 interactions between the different biomolecules, while maximizing the accuracy of the interactions between each class of biomolecule. Presently, the CHARMM carbohydrate force fields have been optimized for hexopyranose monosaccharides<sup>136</sup>, furanoses<sup>139</sup>, acyclic polyalcohols<sup>138</sup>, as well as glycosidic linkages between hexopyranoses<sup>137</sup> and furanoses<sup>140</sup>. Other carbohydrate force fields include OPLS<sup>141</sup> and GROMOS<sup>142, 143</sup>, though these are not discussed in detail here and the reader is referred to a comprehensive review.<sup>87</sup>

### ***1.3.4 Lipid Force Fields***

Lipid FFs have been difficult to optimize because accurate treatment of the gel-liquid phase properties of the bilayer is challenging. The most widely used force fields for membrane simulations are the all-atom CHARMM force field<sup>36, 144-147</sup> and to a lesser extent the united-atom OPLS-UA<sup>148</sup> and GROMOS96<sup>101</sup> force fields, and the coarse-grained MARTINI force field<sup>149</sup>. Among these, the CHARMM lipid force field is the most popular mainly because it was the only all-atom force field highly optimized specifically for lipid molecules. The General AMBER force field (GAFF)<sup>150</sup> has also been used for lipid simulations. Though not specifically optimized for lipids, simulations on various membranes using GAFF have shown that it yields reasonable agreement with certain experimental values, namely membrane thickness, area per lipid, and deuterium order parameters.<sup>151-153</sup> However, the level agreement was achieved only after application of a surface tension to the interface (e.g. to obtain the correct surface area/head group) and that agreement tended to be limited to a few properties. A more

recent AMBER lipid force field named LIPID11 was developed using GAFF parameters that were refined for the phospholipid tails and headgroups.<sup>154</sup>

United-atom force fields are popular for lipid simulations mainly due to their computational efficiency. The most commonly used united-atom force field is the GROMOS96 force field<sup>101</sup>, which in this context refers to the original 43A1 force field and its derivatives. The VDW parameters in 43A1 were optimized targeting condensed phase properties of alkanes<sup>155</sup>, then the hydrocarbon chains were updated<sup>156</sup> by adjusting the LJ parameters to target pentadecane properties. Subsequent updates<sup>157, 158</sup> have focused on improving bilayers, though most of these are derived from the 43A1<sup>99</sup> parameter set.

The CHARMM27 lipid force field<sup>36, 144-146</sup> was a reoptimization of the original CHARMM22 lipid force field.<sup>159</sup> The torsional parameters for the aliphatic groups were optimized to high level QM target data for butane and hexane and the phosphate headgroups were optimized in the context of the lipids and nucleic acids since the phosphate is common to both. The most recent CHARMM36<sup>147</sup> force field is the recommended CHARMM lipid force field. It addresses two major drawbacks of CHARMM27, as well as other lipid force fields, which are the need for an applied surface tension to prevent bilayer shrinkage and the limited reproduction of deuterium order parameters. These issues have been addressed by optimizing the partial atomic charges of the headgroup and select torsional parameters for the headgroup and glycerol linker. Results show that simulations of select lipid bilayers using the CHARMM36 lipid force field result in stable liquid crystalline bilayers, as judged by the correct surface area/head group,<sup>147</sup> which is of great importance in studying membrane proteins using *in*

*silico* drug design methods that often include lipids, proteins and drug-like molecules along with water and ions.

### ***1.3.5 Small Molecule Force Fields***

Development of a small molecular FF that can cover the vastness of the chemical universe is challenging. A few of the most widely used force fields specifically parametrized for small drug-like molecule in condensed phase include: CHARMM's CGenFF<sup>68</sup>, Merck's MMFF<sup>160</sup>, and AMBER's GAFF<sup>150</sup>. All three force fields utilize the same basic optimization scheme, which was discussed in Section 2.1. MMFF was developed with condensed phase protein-ligand simulations in mind, however, the ability of the FF to accurately reproduce condensed phase properties, including structural features of proteins is limited. This is due to the lack of optimization of parameters for proteins as compared to specific biomolecular force fields. Furthermore, because it uses a different functional form for the potential energy function than that shown in **Equation 1.1**, combining it with a more specific biomolecular force field such as CHARMM will most likely result in inaccurate interaction energies due to differences in the way non-bonded interactions are treated. GAFF overcomes this obstacle since it was parametrized to be used with the existing AMBER force field. Similarly, CGenFF was optimized to be compatible with the existing CHARMM force fields. Ultimately this is advantageous due to wide variety of biomolecules represented by the additive CHARMM force field, allowing the biomolecular portion of the system to be treated using parameters specifically optimized for the biomolecule of interest, while the small molecule can be

treated using CGenFF. In fact, because all of the CHARMM force fields are optimized in a similar fashion, calculation of the non-bonded terms is expected to be reasonably accurate and balanced, which is crucial in computer-aided drug design. It is especially important in the work presented here because the ribosome is a heterogeneous system of protein and RNA and small molecules in the case of the macrolide-bound structure.

There are a number of automated parameter assignment engines available for selected small molecule FFs.<sup>69-78</sup> AnteChamber<sup>76</sup> was designed for use with GAFF and generates AMBER topologies. It assigns bonded parameters by analogy from GAFF and partial atomic charges based on Mulliken charges first calculated at the AM1 level and then corrected to RESP-like charges using bond charge correction parameters derived from a training set of molecules<sup>161, 162</sup>. Alternatively, AnteChamber can also utilize user-supplied Gaussian output files to generate actual RESP charges with QM calculations done at the HF/6-31G\* level of theory to maintain compatibility with GAFF parameters. ParamChem<sup>74, 75</sup> is an engine designed for use with CGenFF and generates CHARMM topologies. Bonded parameters are assigned by analogy to CGenFF parameters, while charges are assigned using a bond charge increment scheme optimized targeting the partial atomic charges of a training set of model compounds. This scheme is loosely based off of MMFF's charge assignment implementation and is a more modular approach to the assignment of partial atomic charges.<sup>160, 163-165</sup> MATCH<sup>77</sup> is also available to generate CHARMM topologies and parameters by analogy. The key feature of MATCH is that it represents each molecule as a graphical tree that it uses to make comparisons between molecules in order to obtain a set of parameters for a novel compound. Other parameter assignment engines include: ATB<sup>70</sup> and PRODRG<sup>73</sup> for GROMOS,

SwissParam<sup>78</sup> for CHARMM and MMFF, YASARA<sup>69</sup> for GAFF, and GENRTF<sup>71</sup> for CHARMM.

### ***1.3.6 Polarizable Force Fields***

The parametrization of force fields that treat polarizability is an active area in force field development. Reviews of polarizable force fields have been published elsewhere.<sup>46, 47</sup> The inclusion of polarizability into the potential energy function was discussed in Section 1.2.2. Adding to that, implementation of polarizable force fields will be briefly discussed for the CHARMM Drude<sup>42, 43</sup> and AMOEBA<sup>17, 25, 26</sup> polarizable force fields. The CHARMM Drude polarizable FF is a product of work in this laboratory in collaboration with Roux and coworkers, and parametrization of it has required significant effort and time, dating back over 10 years.<sup>66</sup> Since then, major progress has been made and at this time lipid<sup>38</sup>, protein<sup>42</sup> and nucleic acid<sup>43</sup> parameters are available, with parameters also available for a subset of the carbohydrates including polyalcohols<sup>39</sup> and aldopento/hexapyranoses<sup>20, 166</sup>. The implementation of the Drude FF in CHARMM overcomes the demand of treating the Drude particles via self-consistent field calculations in MD simulations by instead treating them as classical dynamic variables in the context of an extended Lagrangian formalism.<sup>167</sup> In practice, this is achieved by assigning a small mass (0.4 AMU) to the Drude particle from the parent atom and applying specific thermostats to the real versus the virtual Drude particles. The AMOEBA<sup>17, 26</sup> force field is based on atomic multipoles and induced dipoles and is available for proteins<sup>26</sup> and general organic molecules.<sup>168</sup> Each atomic center is

comprised of a partial charge, dipole vector, and quadrupole tensor. MD simulations with AMOEBA are computationally demanding, although the recent development of a water model utilizing a direct polarization approximation (iAMOEBA)<sup>169</sup> partially overcomes this issue. However, such treatment represents a significant approximation whereby the induced dipoles do not relax with respect to each other as compared to the more rigorous use of an extended Lagrangian which mimics the SCF regimen.<sup>51, 170, 171</sup> While water models that reproduce a range of experimental properties may be obtained, recently using the ForceBalance method,<sup>172</sup> it remains to be seen whether neglecting mutual polarization will result in the same accuracy for the presence of solutes, including ions and biomolecules.

In Chapter 5, the parametrization of acyclic sugars D-allose and D-psicose is presented, along with optimization of small model aldehydes and ketones (acetone, butanone, acetaldehyde, and propionaldehyde) as well as myo-inositol. The availability of these parameters continues the work in developing a comprehensive polarizable biomolecular force field.

#### **1.4 Applications of Force Fields to Drug Design**

Force field-based simulations are an important part of *in silico* drug design. Simulations are used to generate an ensemble of protein and/or drug conformations that represent the dynamics of the molecule in solution, from which thermodynamic or conformational properties are measured. A prime example being the calculation of absolute and relative free energies of binding using free energy perturbation

methods,<sup>173,174</sup> among a variety of other methods.<sup>37, 175-178</sup> In the remainder of this section emphasis will be placed on two methodologies developed in this laboratory and used in the work that will be presented in the following chapters: CSP (Conformationally Sampled Pharmacophore) and SILCS (Site-Identification by Ligand Competitive Saturation).

#### ***1.4.1 Conformationally Sampled Pharmacophore (CSP)***

CSP is a pharmacophore-based method in which ensembles of ligand conformations are generated using MD simulations from which probability distributions are calculated for select distances and angles in the molecule (i.e. pharmacophore features).<sup>179</sup> Analyses are then performed to correlate biological activity of the ligand with the CSP pharmacophoric features. The method may be used qualitatively to identify biologically important geometric features as well as quantitatively to predict the activity of other ligands. Notably, physical properties of ligands may be readily incorporated into CSP models. In an approach such as CSP, the utility of the method is based on the accurate treatment of conformational sampling by the ligands, a property directly related to the FF used in the MD simulations.

CSP has been applied previously in this laboratory to two important classes of ligands: opioids and bile acids. For the opioid receptor  $\delta$ <sup>179-181</sup> CSP was applied to develop a model able to distinguish agonists from antagonists, while the CSP model applied to  $\mu$  receptor ligands<sup>182</sup> provided a comprehensive structure-activity relationship explaining how ligand modification altered activity. Notable was the use of CSP to

facilitate the design of an opioid that acts as both a mu agonist and a delta antagonist, thereby having a decreased tolerance profile. The application of CSP-SAR to bile acid transporters yielded a model that indicated ligands with higher intramolecular hydrogen bonding were more active against human apical sodium-dependent bile acid transporter (hASBT) due to increased hydrophobicity.<sup>183</sup> Interestingly, the model also predicted that dianionic bile acid conjugates can achieve high binding affinities using a molecular switch controlled by the location of a carboxyl group that is involved in intramolecular hydrogen bonding.

In the work that will be presented in Chapter 2, CSP was applied to analogs of the macrolide antibiotics telithromycin and cethromycin to show that removal of methyl groups from the antibiotic ring reduced activity compared to telithromycin/cethromycin by increasing the conformational flexibility of the analog.<sup>184-188</sup>

#### ***1.4.2 Site-Identification by Ligand Competitive Saturation (SILCS)***

The final application that will be discussed in this chapter is the SILCS method.<sup>189-191</sup> SILCS harnesses the power of force field-based simulations to map the functional group affinities of a protein or any other macromolecule, an approach that has been used in other laboratories for individual ligands,<sup>192-199</sup> thereby incorporating macromolecule flexibility and solvation into the maps. This is achieved by first performing a number of parallel MD simulations of the protein in a box of explicit water with ~ 0.25 to 1 M fragments. The fragments are a diverse set of small molecules selected such that they capture the chemical features common to drugs, namely hydrophobic and

hydrogen bonding groups, including both charged and neutral species. During the simulations, the fragments interact with all regions on the protein, thus allowing the generation of 3D probability maps (FragMaps), which are normalized with respect to the fragments in solution alone and converted to free energies (called Grid Free Energies, GFEs). The 3D distribution of GFE FragMaps identify regions of the protein, including occluded pockets,<sup>200</sup> that bind fragments favorably, as well as regions where specific fragment types are not favored. Notably, the normalization procedure leads to the FragMaps including the energetic penalty for ligand desolvation and since the fragments must compete with water on the protein, it also accounts for the energetic penalty required to compete with water solvating the protein. Thus, the approach inherently includes the information in water-mapping methods, such as WaterMap,<sup>201, 202</sup> while also identifying the types of functional groups that can successfully interact with different regions of the protein surface. In addition, the binding affinity of ligands may be estimated based on the ligand-based GFE (LGFE).

Initial studies from this laboratory using SILCS involved the use of two fragments (Tier I SILCS): benzene to represent aromatic groups and propane to represent aliphatic groups, with water used to represent hydrogen bond donor and acceptor.<sup>189, 190</sup> More recently, the number of fragments was expanded to include benzene, propane, formamide, acetaldehyde, acetate, methylammonium, and methanol, which is referred to a Tier II SILCS.<sup>191</sup> Benzene and propane again represent aromatic and aliphatic groups, respectively, while formamide, acetaldehyde, and methanol represent neutral hydrogen bond donors (via their polar hydrogens) and acceptors (via their oxygens). Methylammonium and acetate are included to represent charged hydrogen bond donors

and acceptors, respectively. In addition, imidazole can be included to represent a heterocycle as well as both an additional neutral donor and acceptor as is done in the present work. Though Tier I SILCS showed that the crystallographic binding modes of ligands for a diverse set of proteins could be reproduced even with a simple set of drug-like fragments, it also indicated that the simple set of fragments used limited the ability to correctly rank LGFE scores for protein ligands.<sup>189, 190</sup> This, along with the need to map the functional requirements for specific types of hydrogen bonding groups, stimulated the use of more specific fragment types. Importantly, the presence of chemically distinct hydrogen bonding fragments in SILCS is advantageous because the fragments compete with each other and with water, such that the resulting FragMaps include the energetic cost of the displacement of water, as mentioned above. It was shown that the Tier II SILCS fragment maps are also able to identify the classes of protein-ligand interactions that are observed in crystal structures.<sup>191</sup> With respect to quantitative ranking of ligands, it was shown that the use of an ensemble of ligand conformations leads to significant improvements. Notable in that study was the use of the FragMap LGFE scores in conjunction with Monte Carlo sampling to generate ensembles of ligand conformations. Such an approach essentially represents the use of the SILCS method for ligand docking. Towards this end a SILCS pharmacophore-based database screening protocol has been presented and shown to yield improved enrichment rates in three test proteins.<sup>203</sup> Future extensions of the SILCS method may involve the use of different types a functional groups, a process that will be facilitated by the availability of parameters for the molecules of interest, as may be obtained from CGenFF.

In Chapter 4, the application of SILCS to the macrolide binding pocket of the *E.*

*coli* 50S subunit will be explored.

## 1.5 References

1. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* **1995**, 117, 5179-5197.
2. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Josephy-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **1998**, 102, 3586-3616.
3. Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Function for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J Amer Chem Soc* **1988**, 110, 1657-1666.
4. van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*. University of Groningen: Groningen, The Netherlands, 1987; p 1-221.
5. MacKerell, A. D., Jr. Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* **2004**, 25, 1584-604.
6. Blondel, A.; Karplus, M. New formulation for derivatives of torsion angles and

improper torsion angles in molecular mechanics: elimination of singularities. *J Comput Chem* **1996**, 17, 1132-1141.

7. Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J Chem Phys* **1996**, 105, 1902-1919.
8. Baker, C. M.; Lopes, P. E.; Zhu, X.; Roux, B.; MacKerell, A. D., Jr. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J Chem Theory Comput* **2010**, 6, 1181-1198.
9. Feig, M.; Brooks, C. L., III. Recent advances in the development and application of implicit solvent models in biomolecular simulations. *Curr Opin Struct Biol* **2004**, 14, 217-224.
10. Roux, B.; Simonson, T. Implicit Solvent Models. *Biophys Chem* **1999**, 78, 1-20.
11. Demerdash, O.; Yap, E.-H.; Head-Gordon, T. Advanced Potential Energy Surfaces for Condensed Phase Simulations. *Annu Rev Phys Chem* **2014**.
12. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. DREIDING: a generic force field for molecular simulations. *J Phys Chem* **1990**, 94, 8897-8909.
13. Rappe, A. K.; Casewit, C. J.; Colwel, K. S.; Goddard, W. A., III; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J Am Chem Soc* **1992**, 114, 10024-10035.
14. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* **2012**, 8, 3257-3273.

15. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **2004**, 25, 1400-15.
16. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* **2004**, 126, 698-699.
17. Ren, P.; Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J Comput Chem* **2002**, 23, 1497-506.
18. Mallajosyula, S. S.; Adams, K. M.; Barchi, J. J.; MacKerell, A. D., Jr. Conformational determinants of the activity of antiproliferative factor glycopeptide. *J Chem Inf Model* **2013**, 53, 1127-1137.
19. Mallajosyula, S. S.; MacKerell, A. D., Jr. Influence of solvent and intramolecular hydrogen bonding on the conformational properties of o-linked glycopeptides. *J Phys Chem B* **2011**, 115, 11215-29.
20. Patel, D. S.; He, X.; MacKerell, A. D., Jr. Polarizable empirical force field for hexopyranose monosaccharides based on the classical drude oscillator. *J Phys Chem B* **2015**, 119, 637-652.
21. Halgren, T. A. Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters. *J Am Chem Soc* **1992**, 114, 7827-7843.
22. Buckingham, R. A. The classical equation of state of gaseous helium, neon and argon. *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* **1938**, 168, 264-283.

23. White, D. N. J. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *J Computer-Aided Mol Des* **1997**, 11, 517-521.
24. Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J Comput Chem* **1999**, 20, 730-748.
25. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J Phys Chem B* **2010**, 114, 2549-64.
26. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. The Polarizable Atomic Multipole-based AMOEBA Force Field for Proteins. *J Chem Theory Comput* **2013**, 9, 4046-4063.
27. Warshel, A.; Levitt, M. Theoretical studies of enzymatic reactions: dielectric, electrostatic, and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* **1976**, 103, 227-249.
28. Fox, T.; Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J Phys Chem B* **1998**, 102, 8070-8079.
29. Gough, C. A., DeBolt, S. E. and Kollman, P. A. Derivation of Fluorine and Hydrogen Atom Parameters Using Liquid Simulations. *J Comput Chem* **1992**, 13, 963-970.
30. Jorgensen, W. L. Optimized Intermolecular Potential Functions for Liquid Alcohols. *J Phys Chem* **1986**, 90, 1276-1284.

31. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* **1996**, 118, 11225-11236.
32. MacKerell, A. D., Jr.; Karplus, M. Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids. *J Phys Chem* **1991**, 95, 10559-10560.
33. Chen, I. J.; Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of Lennard-Jones parameters for polar-neutral compounds. *J Comput Chem* **2002**, 23, 199-213.
34. Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model. *J Phys Chem* **1994**, 98, 13077-13082.
35. Rizzo, R. C.; Jorgensen, W. L. OPLS all-atom model for amines: resolution of the amine hydration problem. *J Am Chem Soc* **1999**, 121, 4827-4836.
36. Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of lennard-jones parameters. *J Comput Chem* **1998**, 19, 334-338.
37. Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys* **2003**, 119, 5740-5761.
38. Chowdary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D., Jr.; Roux, B. A polarizable force field of dipalmitoylphosphatidylcholine based on the classical drude model for molecular dynamics simulations of lipids. *J Phys Chem B* **2013**, 117, 9142-9160.

39. He, X.; Lopes, P. E.; MacKerell, A. D., Jr. Polarizable empirical force field for acyclic polyalcohols based on the classical drude oscillator. *Biopolymers* **2013**, 99, 724-38.
40. Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc Natl Acad Sci U S A* **2008**, 105, 6290-5.
41. Jiao, D.; Zhang, J.; Duke, R. E.; Li, G.; Schnieders, M. J.; Ren, P. Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential. *J Comput Chem* **2009**, 30, 1701-11.
42. Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr. Force Field for Peptides and Proteins based on the Classical Drude Oscillator. *J Chem Theory Comput* **2013**, 9, 5430-5449.
43. Savelyev, A.; MacKerell, A. D., Jr. All-atom polarizable force field for DNA based on the classical drude oscillator model. *J Comput Chem* **2014**, 10, 1652-1664.
44. Shi, Y.; Zhu, C. Z.; Martin, S. F.; Ren, P. Probing the effect of conformational constraint on phosphorylated ligand binding to an SH2 domain using polarizable force field simulations. *J Phys Chem B* **2012**, 116, 1716-27.
45. Zhang, J.; Yang, W.; Piquemal, J. P.; Ren, P. Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J Chem Theory Comput* **2012**, 8, 1314-1324.
46. Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. Theory and applications. *Theor Chem Acc* **2009**, 124, 11-28.

47. Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev Comp Chem* **2002**, 18, 89-146.
48. Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. An Anisotropic Polarizable Water Model: Incorporation of All-Atom Polarizabilities into Molecular Mechanics Force Fields. *J Phys Chem* **1994**, 98, 4180-4187.
49. Caldwell, J.; Dang, L. X.; Kollman, P. A. Implementation of Nonadditive Intermolecular Potentials by Use of Molecular Dynamics: Development of a Water-Water Potential and Water-Ion Cluster Interactions. *J Am Chem Soc* **1990**, 112, 9144-9147.
50. Dang, L. X. Importance of Polarization Effects in Modeling Hydrogen Bond in Water Using Classical Molecular Dynamics Techniques. *J Phys Chem B* **1998**, 102, 620-624.
51. Sprik, M.; Klein, M. L. A polarizable model for water using distributed charge sites. *J Chem Phys* **1988**, 89, 7556-7560.
52. Wallqvist, A.; Berne, B. J. Effective potentials for liquid water using polarizable and nonpolarizable models. *J Phys Chem* **1993**, 97, 13841-13851.
53. Asensio, J. L.; Canada, F. J.; Chen, X. H.; Khan, N.; Mootoo, D. R.; Jimenez-Barbero, J. Conformational differences between O- and C-glycosides: The alpha-O-Man-(1 -> 1)-beta-Gal/alpha-C-Man-(1 -> 1)-beta-Gal case - A decisive demonstration of the importance of the exo-anomeric effect on the conformation of glycosides. *Chem European Journal* **2000**, 6, 1035-1041.
54. Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. Cooperative effects in the structure of fluoride water clusters: Ab initio hybrid quantum

mechanical/molecular mechanical model incorporating polarizable fluctuating charge solvent. *J Chem Phys* **1998**, 109, 3077-3085.

55. Llanta, E.; Ando, K.; Rey, R. Fluctuating charge study of polarization effects in chlorinated organic liquids. *J Phys Chem B* **2001**, 105, 7783-7791.

56. Patel, S.; Brooks, C. L., 3rd. CHARMM fluctuating charge force field for proteins: I Parameterization and application to bulk organic liquid simulations. *J Comput Chem* **2003**, 25, 1-15.

57. Rick, S. W.; Berne, B. J. Dynamical fluctuating charge force fields: the aqueous solvation of amides. *J Am Chem Soc* **1996**, 118, 672-679.

58. Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J. Fluctuating Charge Force Fields for Aqueous Solutions. *Studies in Physical and Theoretical Chemistry* **1995**, 83, 31-40.

59. Yoshii, N.; Miyauchi, R.; Miura, S.; Okazaki, S. A molecular-dynamics study of the equation of state of water using a fluctuating-charge model. *Chem Phys Let* **2000**, 317, 414-420.

60. Kunz, A.-P. E.; van Gunsteren, W. F. Development of a nonlinear classical polarization model for liquid water and aqueous solution: COS/D. *J Phys Chem A* **2009**, 113, 11570-11579.

61. van Maaren, P. J.; van der Spoel, D. Molecular dynamics simulations of water within novel shell-model potentials. *J Phys Chem B* **2001**, 105, 2618-2626.

62. Iczkowski, R. P.; Margrave, J. L. Electronegativity. *J Am Chem Soc* **1961**, 83, 3547-3551.

63. Mulliken, R. S. A new electroaffinity scale; together with data on valence states

and on valence ionization potentials and electron affinities. *J Chem Phys* **1934**, 2, 782-793.

64. Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *J Am Chem Soc* **1983**, 105, 7512-7516.

65. Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J Chem Theory Comput* **2005**, 1, 153-168.

66. Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J Chem Phys* **2003**, 119, 5185-5197.

67. Yu, H.; Hansson, T.; Van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J Chem Phys* **2003**, 118, 221-234.

68. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **2010**, 31, 671-90.

69. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA-a self-parameterizing force field. *Proteins: Structure, Function, and Genetics* **2002**, 47, 393-402.

70. Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J Chem Theory Comput* **2011**, 7, 4026-4037.

71. Miller, B. T.; Singh, R. P.; Klauda, J. B.; Hodoseck, M.; Brooks, B. R.; Woodcock, H. L., III. CHARMMing: a new, flexible web portal for CHARMM. *J Chem*

*Inf Model* **2008**, 48, 1920-1929.

72. Ribiero, A. A. S. T.; Horta, B. A. C.; de Alencastro, R. B. MKTOP: a program for automatic construction of molecular topologies. *J Braz Chem Soc* **2008**, 19, 1433-1435.

73. Schuttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **2004**, 6, 1355-1363.

74. Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* **2012**, 52, 3144-54.

75. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* **2012**, 52, 3155-68.

76. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **2006**, 25, 247-60.

77. Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L., 3rd. MATCH: an atom-typing toolset for molecular mechanics force fields. *J Comput Chem* **2012**, 33, 189-202.

78. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem* **2011**, 32, 2359-68.

79. Brooks, B. R.; Brooks III, C. L.; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.;

- Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J Comput Chem* **2009**, 30, 1545-1614.
80. Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-Based Search, Retrieval, Analysis and Display of Information. *Acta Cryst* **1979**, B35, 2331-2339.
81. Kuczera, K.; Wiorkiewicz, J. K.; Karplus, M. *MOLVIB: Program for the Analysis of Molecular Vibrations*, CHARMM, Harvard University: 1993.
82. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives. *J Am Chem Soc* **1979**, 101, 2550-2560.
83. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J Phys Chem* **1993**, 97, 10269-10280.
84. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J Am Chem Soc* **1984**, 106, 765-784.
85. Feig, M.; MacKerell, A. D., Jr.; Brooks, C. L., III. Force field influence on the observation of p-helical protein structures in molecular dynamics simulations. *J Phys Chem B* **2003**, 107, 2831-2836.

86. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* **2000**, *28*, 235-242.
87. Fadda, E.; Woods, R. J. Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov Today* **2010**, *15*, 596-609.
88. French, A. D.; Johnson, G. P. Computerized molecular modeling of carbohydrates. *Methods Mol Biol* **2011**, *715*, 21-42.
89. Guvench, O.; MacKerell, A. D., Jr. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol Biol* **2008**, *443*, 63-88.
90. Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Advances in Protein Chemistry* **2003**, *66*, 27-85.
91. McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267*, 585-590.
92. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J Comput Chem* **2003**, *24*, 1999-2012.
93. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712-25.
94. Kollman, P. A. Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc Chem*

*Res* **1996**, 29, 461-469.

95. Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* **2000**, 21, 1049-1074.
96. Kaminski, G.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J Phys Chem B* **2001**, 105, 6474-6487.
97. Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* **2004**, 25, 1656-76.
98. Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* **2011**, 40, 843-56.
99. Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J Comput Chem* **2001**, 22, 1205-1218.
100. Soares, T. A.; Hunenberger, P. H.; Kastholz, M. A.; Krautler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F. An improved nucleic acid parameter set for the GROMOS force field. *J Comput Chem* **2005**, 26, 725-37.
101. van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. BIOMOS b.v.: Zürich, 1996.

102. Price, D. J.; Brooks, C. L., 3rd. Modern protein force fields behave comparably in molecular dynamics simulations. *J Comput Chem* **2002**, *23*, 1045-57.
103. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, *79*, 926-936.
104. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. The Missing Term in Effective Pair Potentials. *J Phys Chem* **1987**, *91*, 6269-6271.
105. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*, Pullman, B., Ed. Reidel Publishing Co.: Dordrecht, Holland, 1981; pp 331-342.
106. Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-EW. *J Chem Phys* **2004**, *120*, 9665.
107. Best, R. B.; Buchete, N. V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys J* **2008**, *95*, L07-9.
108. Garcia, A. E.; Sanbonmatsu, K. Y. alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci U S A* **2002**, *99*, 2782-2787.
109. Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to helix-coil transition of polypeptides. *J Phys Chem B* **2009**, *113*, 9004-9015.
110. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950-8.
111. Li, D.-W.; Bruschweiler, R. NMR-based protein potentials. *Angew Chem Int Ed*

*Engl* **2010**, 49, 6778-6780.

112. MacKerell, A. D., Jr. Atomistic Models and Force Fields. In *Computational Biochemistry and Biophysics*, Becker, O. M.; MacKerell, A. D., Jr.; Roux, B.; Watanabe, M., Eds. Marcel Dekker, Inc.: New York, 2001; pp 7-38.
113. Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J Comput Chem* **1981**, 2, 287-303.
114. Nilsson, L.; Karplus, M. Empirical energy functions for energy minimization and dynamics of nucleic acids. *J Comput Chem* **1986**, 7, 591-616.
115. Norberg, J.; Nilsson, L. On the truncation of long-range electrostatics in DNA. *Biophys J* **2000**, 79, 1537-1553.
116. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J Chem Phys* **1993**, 98, 10089-10092.
117. MacKerell, A. D., Jr. Influence of magnesium ions on duplex DNA structural, dynamic, and solvation properties. *J Phys Chem B* **1997**, 101, 646-650.
118. Norberg, J.; Nilsson, L. Constant pressure molecular dynamics simulations of hte dodecamers: d(GCGCGCGCGCGC)<sub>2</sub> and r(GCGCGCGCGCGC)<sub>2</sub>. *J Chem Phys* **1996**, 104, 6052-6057.
119. MacKerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. An all-atom empirical energy function for the simulation of nucleic acids. *J Am Chem Soc* **1995**, 117, 11946-11975.
120. Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol*

*Struct Dyn* **1999**, 16, 845-861.

121. Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **2000**, 21, 86-104.
122. MacKerell, A. D., Jr.; Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* **2000**, 21, 105-120.
123. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **2011**, 32, 1929-43.
124. Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; Mackerell, A. D., Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* **2012**, 8, 348-362.
125. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J* **2007**, 92, 3817-29.
126. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., 3rd; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* **2011**, 7, 2886-2902.
127. Langley, D. R. Molecular dynamics simulations of environment and sequence dependent DNA conformation: The development of the BMS nucleic acid force field and comparison with experimental results. *J Biomol Struct Dyn* **1998**, 16, 487-509.

128. Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. The nucleic acid database: a comprehensive relational database of the three-dimensional structures of nucleic acids. *Biophys J* **1992**, *63*, 751-759.
129. Coimbatore Narayanan, B.; Westbrook, J.; Ghosh, S.; Petrov, A. I.; Sweeney, B.; Zirbel, C. L.; Leontis, N. B.; Berman, H. M. The Nucleic Acid Database: new features and capabilities. *Nucleic Acids Res* **2014**, *42*, D114-22.
130. Cheatham, T. E., 3rd; Case, D. A. Twenty-five years of nucleic acid simulations. *Biopolymers* **2013**, *99*, 969-77.
131. Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. Reparameterization of RNA chi torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *J Chem Theory Comput* **2010**, *6*, 1520-1531.
132. Kirschner, K. N.; Lins, R. D.; Maass, A.; Soares, T. A. A Glycam-Based Force Field for Simulations of Lipopolysaccharide Membranes: Parametrization and Validation. *J Chem Theory Comput* **2012**, *8*, 4719-4731.
133. Kirschner, K. N.; Woods, R. J. Solvent interactions determine carbohydrate conformation. *Proc Natl Acad Sci U S A* **2001**, *98*, 10541-5.
134. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem* **2008**, *29*, 622-55.
135. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1.

- GLYCAM\_93 parameter development. *J Phys Chem* **1995**, 99, 3832-3846.
136. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* **2008**, 29, 2543-64.
137. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* **2009**, 5, 2353-2370.
138. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, 5, 1315-1327.
139. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* **2009**, 113, 12466-76.
140. Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* **2010**, 114, 12981-12994.
141. Kony, D.; Damm, W.; Stoll, S.; Van Gunsteren, W. F. An improved OPLS-AA force field for carbohydrates. *J Comput Chem* **2002**, 23, 1416-29.
142. Hansen, H. S.; Hunenberger, P. H. A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers. *J Comput Chem* **2011**, 32, 998-1032.
143. Lins, R. D.; Hunenberger, P. H. A new GROMOS force field for hexopyranose-

- based carbohydrates. *J Comput Chem* **2005**, 26, 1400-12.
144. Feller, S. E.; MacKerell, A. D., Jr. An improved empirical potential energy function for molecular simulations of phospholipids. *J Phys Chem B* **2000**, 104, 7510-7515.
145. Klauda, J. B.; Brooks, B. R.; MacKerell, A. D., Jr.; Venable, R. M.; Pastor, R. W. An Ab Initio Study on the Torsional Surface of Alkanes and its Effect on Molecular Simulations of Alkanes and a DPPC Bilayer. *J Phys Chem B* **2005**, 109, 5300-5311.
146. Klauda, J. B.; Pastor, R. W.; Brooks, B. R. Adjacent gauche stabilization in linear alkanes: implications for polymer models and conformational analysis. *J Phys Chem B Letters* **2005**, 109, 15684-15686.
147. Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* **2010**, 114, 7830-43.
148. Ulmschneider, J. P.; Ulmschneider, M. B. United atom lipid parameters for combination with the optimized potentials for liquid simulations all-atom force field. *J Chem Theory Comput* **2009**, 5, 1803-1813.
149. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* **2007**, 111, 7812-7824.
150. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J Comput Chem* **2004**, 25, 1157-1174.
151. Jojart, B.; Martinek, T. A. Performance of the general amber force field in

- modeling aqueous POPC membrane bilayers. *J Comput Chem* **2007**, 28, 2051-8.
152. Rosso, L.; Gould, I. R. Structure and dynamics of phospholipid bilayers using recently developed general all-atom force fields. *J Comput Chem* **2008**, 29, 24-37.
153. Siu, S. W. I.; Vacha, R.; Jungwirth, P.; Bockman, R. A. Biomolecular simulations of membranes: physical properties from different force fields. *J Chem Phys* **2008**, 128, 125103.
154. Skjevik, A. A.; Madej, B. D.; Walker, R. C.; Teigen, K. LIPID11: a modular framework for lipid simulations using amber. *J Phys Chem B* **2012**, 116, 11124-36.
155. Daura, X.; Mark, A. E.; van Gunsteren, W. F. Parametrization of Aliphatic CH<sub>n</sub> United Atoms of GROMOS96 Force Field. *J Comp Chem* **1998**, 19, 535-547.
156. Berger, O.; Edholm, O.; Jahnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* **1997**, 72, 2002-2013.
157. Chiu, S.-W.; Pandit, S. A.; Scott, H. L.; Jakobsson, E. An improved united atom force field for simulation of mixed lipid bilayers. *J Phys Chem B* **2009**, 113, 2748-2763.
158. Poger, D.; van Gunsteren, W. F.; Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J Comput Chem* **2010**, 31, 1117-25.
159. Schlenkrich, M.; Brinkman, J.; MacKerell, A. D., Jr.; Karplus, M. An Empirical Potential Energy Function for Phospholipids: Criteria for Parameter Optimization and Applications. In *Membrane Structure and Dynamics*, Merz, K. M.; Roux, B., Eds. Birkhauser: Boston, 1996; pp 31-81.
160. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J Comput Chem* **1996**, 17, 490-519.

161. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: 1. Method. *J Comput Chem* **2000**, *21*, 132-146.
162. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* **2002**, *23*, 1623-41.
163. Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J Comput Chem* **1996**, *17*, 520-552.
164. Halgren, T. A. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules. *J Comput Chem* **1996**, *17*, 616-641.
165. Halgren, T. A. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94. *J Comput Chem* **1996**, *17*, 553-586.
166. Jana, M.; Alexander D. MacKerell, J. CHARMM Drude polarizable force field for aldehydofuranoses and methyl-aldehydofuranosides. *J Phys Chem B* **2015**, just accepted.
167. Lamoureux, G.; Roux, B. Modeling Induced Polarization with Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J Chem Phys* **2003**, *119*, 3025-3039.
168. Shi, Y.; Wu, C.; Ponder, J. W.; Ren, P. Investigation of multipole electrostatics in hydration free energy calculations. *J Comput Chem* **2011**, *32*, 967-977.
169. Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman,

- P. K.; Martinez, T. J.; Pande, V. S. Systematic improvement of a classical molecular model of water. *J Phys Chem B* **2013**, 117, 9956-9972.
170. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Molecular Physics* **1996**, 87, 1117-1157.
171. Tuckerman, M. E.; Martyna, G. J. Understanding Modern Molecular Dynamics: Techniques and Applications. *J Phys Chem B* **2000**, 104, 159-178.
172. Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J Chem Theory Comput* **2013**, 9, 452-460.
173. Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem Rev* **1993**, 93, 2395-2417.
174. Straatsma, T. P.; McCammon, J. A. Computational Alchemy. *Annu Rev Phys Chem* **1992**, 43, 407-435.
175. Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, 303, 1813-1818.
176. Shim, J.; MacKerell, J. A. D. Computational ligand-based rational design: role of conformational sampling and force fields in model development. *Med Chem Comm* **2011**, 2, 356-370.
177. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Jr. Computational methods in drug discovery. *Pharmacological Reviews* **2014**, 66, 334-395.
178. Zhong, S.; MacKerell, J., A.D. . Binding Response: A Descriptor for Selecting Ligand Binding Site on Protein Surfaces. *J Chem Infor Modeling* **2007**, 47, 2303-2315.
179. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. 2D Conformationally Sampled Pharmacophore: A Ligand-Based Pharmacophore To Differentiate delta Opioid Agonists

- from Antagonists. *J Am Chem Soc* **2003**, 125, 3101-3107.
180. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. Conformationally sampled pharmacophore for peptidic delta opioid ligands. *J Med Chem* **2005**, 48, 7773-80.
181. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. Quantitative Conformationally Sampled Pharmacophore for  $\delta$  Opioid Ligands: Reevaluation of Hydrophobic Moieties Essential for Biological Activity. *J Med Chem* **2007**, 50, 1799-1809.
182. Shim, J.; Coop, A.; MacKerell, A. D., Jr. Consensus 3D model of mu-opioid receptor ligand efficacy based on a quantitative Conformationally Sampled Pharmacophore. *J Phys Chem B* **2012**, 115, 7487-96.
183. Rais, R.; Acharya, C.; Tririyaa, G.; MacKerell, A. D., Jr.; Polli, J. E. Molecular switch controlling the binding of anionic bile acid conjugates to human apical sodium-dependent bile acid transporter. *J Med Chem* **2010**, 53.
184. Glassford, I.; Lee, M.; Wagh, B.; Velvadapu, V.; Paul, T.; Sandelin, G.; DeBrosse, C.; Klepacki, D.; Small, M. C.; A.D. MacKerell, J.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4-desmethyl telithromycin. *ACS Med Chem Lett* **2014**, 5, 1021-1026.
185. Velvadapu, V.; Glassford, I.; Lee, M.; Paul, T.; DeBrosse, C.; Klepacki, D.; Small, M. C.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,10-didesmethyl telithromycin. *ACS Med Chem Lett* **2012**, 3, 211-215.
186. Velvadapu, V.; Paul, T.; Wagh, B.; Klepacki, D.; Guvench, O.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolide analogues to address antibiotic resistance: total synthesis and biological evaluation of 4,8,10-tridesmethyl telithromycin. *ACS Med Chem Lett* **2011**, 2, 68-72.

187. Wagh, B.; Paul, T.; DeBrosse, C.; Klepacki, D.; Small, M. C.; A.D. MacKerell, J.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,8,10-tridesmethyl cethromycin. *ACS Med Chem Lett* **2013**, 4.
188. Wagh, B.; Paul, T.; Glassford, I.; DeBrosse, C.; Klepacki, D.; Small, M. C.; MacKerell, A. D., Jr.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,8-didesmethyl telithromycin. *ACS Med Chem Lett* **2012**, 3, 1013-1018.
189. Guvench, O.; MacKerell, A. D., Jr. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput Biol* **2009**, 5, e1000435.
190. Raman, E. P.; Yu, W.; Guvench, O.; Mackerell, A. D. Reproducing crystal binding modes of ligand functional groups using Site-Identification by Ligand Competitive Saturation (SILCS) simulations. *J Chem Inf Model* **2011**, 51, 877-96.
191. Raman, E. P.; Yu, W.; Lakkaraju, S. K.; Mackerell, A. D., Jr. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J Chem Inf Model* **2013**, 53, 3384-98.
192. Bakan, A.; Nevins, N.; Lakdawala, A. S.; Bahar, I. Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules. *J Chem Theory Comput* **2012**, 8, 2435-2447.
193. Ben-Shimon, A.; Eisenstein, M. Computational mapping of anchoring spots on protein surfaces. *J Mol Biol* **2010**, 402, 259-277.
194. Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, 25, 621-627.
195. Halgren, T. A. New method for fast and accurate binding site identification

- analysis. *Chem Biol Drug Des* **2007**, 69, 146-148.
196. Lexa, K. W.; Carlson, H. A. Full protein flexibility is essential for proper hot-spot mapping. *J Am Chem Soc* **2011**, 133, 200-202.
197. Miranker, A.; Karplus, M. Functionality Maps of Binding Sites: A Multiple Copy Simultaneous Search Method. *Proteins* **1991**, 11, 29-34.
198. Tan, Y. S.; Sledz, P.; Lang, S.; Stubbs, C. J.; Spring, D. R.; Abell, C.; Best, R. B. Using ligand-mapping simulations to design a ligand selectively targeting a cryptic surface pocket of polo-like kinase I. *Angew Chem Int Ed Engl* **2012**, 51, 10078-10081.
199. Wang, S.; Yang, C.-Y. Hydrophobic binding hot spots of Bcl-xL protein-protein interfaces by cosolvent molecular dynamics simulations. *ACS Med Chem Lett* **2011**, 2, 280-284.
200. Lakkaraju, S. K.; Raman, E. P.; Yu, W.; MacKerell, A. D., Jr. Free energies of organic solutes in aqueous and heterogeneous environments using grand canonical monte carlo-molecular dynamics simulations. *J Chem Theory Comput* **2014**, 10, 2281-2290.
201. Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of Factor Xa ligand binding. *J Am Chem Soc* **2008**, 130, 2817-2831.
202. Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci U S A* **2007**, 104, 808-813.
203. Yu, W.; Lakkaraju, S. K.; Raman, E. P.; MacKerell, A. D., Jr. Site-identification by ligand competitive saturation (SILCS) assisted pharmacophore modeling. *J Computer-Aided Mol Des* **2014**.

## **Chapter 2 The Effects of Macrolide Desmethylation on Bioactivity using the Conformational Sampled Pharmacophore (CSP) Approach**

Overcoming microbial resistance presents a major challenge in the development of new antibiotics. Approximately 50% of all antibiotics target the peptidyl transfer center (PTC) of the bacterial ribosome, inhibiting protein synthesis and thereby disrupting cellular growth. A recent class of 3<sup>rd</sup> generation macrolides, the ketolides, has been demonstrated to be effective against bacterial strains with macrolide- lincosamide-streptogramin B (MLS<sub>B</sub>) resistance due to the addition of an imidazole-pyridine bearing alkyl chain that increases interactions within the binding pocket. However, telithromycin is still susceptible to resistance via modification of base A2058 (*E.coli* numbering). Crystal structures of the *E. coli* ribosome suggest that A2058G mutation leads to a steric clash with a C4-methyl group of telithromycin. To probe the role of the C-4 methyl group, 50 ns Hamiltonian replica exchange molecular dynamics simulations were performed on telithromycin and its desmethyl analogs. Analysis of selected distances distributions revealed that removal of methyl groups from the telithromycin macrocycle leads to increased conformational flexibility, which is used to explain the decrease in minimum inhibitory concentration values for the desmethyl analogs.

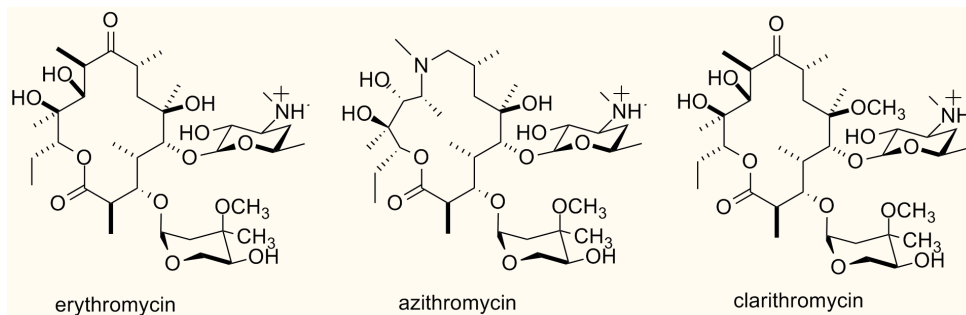
### **2.1 Introduction**

The bacterial ribosome is an important target for a number of existing

antibiotics.<sup>1,2</sup> These bind at vital sites within both the 30S and 50S ribosomal subunits and inhibit processes that are essential for cell survival.<sup>3-10</sup> Macrolide antibiotics (or macrolides hereafter) bind at the entrance to the nascent peptide exit tunnel in the 50S subunit of the bacterial ribosome, thereby blocking elongation of the growing polypeptide chain.<sup>7, 10-12</sup> Macrolides belong to a large class of natural products known as polyketides. These compounds are synthesized by living organisms in order to confer a competitive advantage over other organisms and thus many polyketides are highly pharmaceutically active antimicrobials. Macrolides represent a class of molecules containing a 14- to 16-membered macrocyclic lactone ring that is functionalized by one or more sugar molecules. These molecules are chemically complex, containing upwards of 10 stereospecific centers that, in addition to the sugar moieties, make macrolide total synthesis a major challenge. Erythromycin is the most notable among the macrolide antibiotics and it has been widely used since its isolation from the soil bacterium *Saccharopolyspora erythraea* in the 1950s. Erythromycin (**Figure 2.1**) is a 14-membered macrolactone ring functionalized with two sugar moieties, L-cladinose and D-desosamine. Because erythromycin readily undergoes hydrolysis at low pH and hence is easily inactivated by gastric acid,<sup>13</sup> early analogs of erythromycin sought to increase its acid stability using conjugated salts.

Further investigation into the acid stability of erythromycin and the improvement of its pharmacokinetic profile led to the second-generation macrolide antibiotics clarithromycin and azithromycin (**Figure 2.1**). The acid instability had been attributed to an intramolecular dehydration reaction to form erythromycin-6,9-hemiketal and then anhydroerythromycin<sup>14</sup> and thus second-generation macrolides were designed to avoid

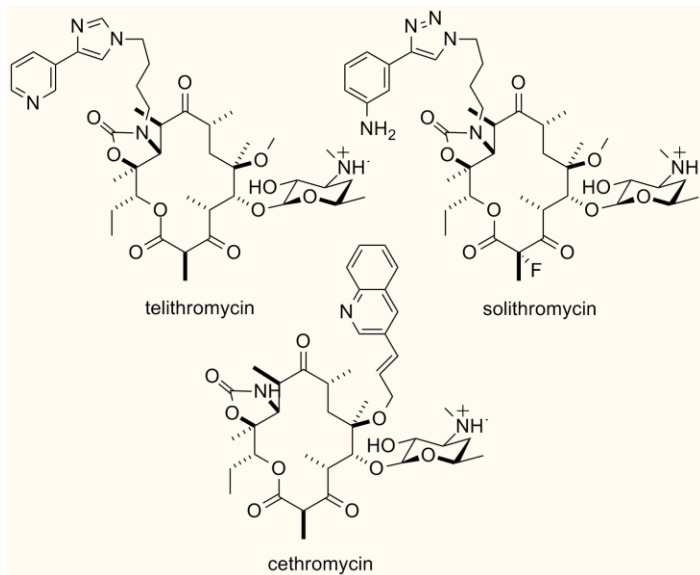
this dehydration. These molecules are very similar to erythromycin with the exception of the replacement of the C6-hydroxyl in erythromycin for a C6-methyl ether in clarithromycin and in azithromycin, the C9-ketone is replaced with a methylene and the 14-membered macrolactone ring is extended by a nitrogen to yield a 15-membered ring.



**Figure 2.1** 2D representation of first and second generation macrolides: erythromycin, azithromycin, and clarithromycin.

The most recent third-generation macrolides arose due to growing antibacterial resistance of the previous generations. Bacteria achieve resistance to antibiotics in three major ways: extrusion of the drug from the cell using efflux pumps, metabolism of the drug to an inactive metabolite, or modification of the drug target. Early generation macrolide antibiotics were found to induce ribosomal modification as well as be susceptible to extrusion via macrolide efflux pumps (such as those encoded by *mef*<sup>15</sup> genes). Removal of the L-cladinose and its subsequent replacement with a keto-group led to a group of semisynthetic macrolides called ketolides, named for the C3-keto group, that partly overcame these problems due to a higher affinity for the ribosome.<sup>8,16</sup> Ketolides include telithromycin (FDA approved in 2001), cethromycin, and solithromycin, as shown in **Figure 2.2**. While ketolides have mitigated macrolide

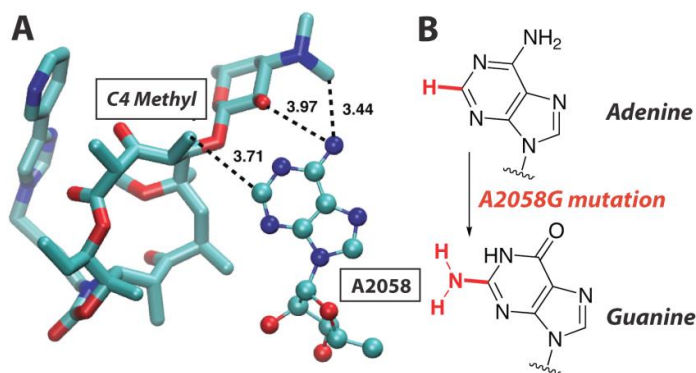
resistance due to the A2058G modification to a minor extent, they have not addressed resistance due to ribosomal methylation.<sup>17,18</sup> This is predominantly ribosomal modification of 23S rRNA at base A2058 (*Escherichia coli* numbering throughout), which confers cross-resistance to antibiotics in the macrolide, lincosamide and streptogramin B classes (MLS resistance).<sup>2,19-21</sup> A2058-based resistance includes mutation of A2058 to G as well as methylation of the exocyclic N6 of A2058 via the expression of erythromycin resistant methyltransferases encoded by the *erm* gene, which add one or two methyl groups and is the most effective mechanism of macrolide resistance.<sup>22-24</sup>



**Figure 2.2** 2D representation of third generation macrolides: telithromycin, solithromycin, and cethromycin.

Crystal structures of macrolides bound to *E. coli*<sup>25</sup> (as well as *Thermus thermophilus*,<sup>26</sup> *Haloarcula marismortui*,<sup>2</sup> and *Deinococcus radiodurans*<sup>10</sup>) have

revealed several interactions that are important for ligand binding. Macrolides possess a D-desosamine sugar whose 2'-hydroxyl group forms a hydrogen bond with N1 of A2058. The 2'-OH has also been proposed to hydrogen bond with N6 of A2058 as well as N6 of neighboring A2059.<sup>10, 27</sup> Additionally, the protonated dimethylamine of the desosamine sugar is suggested to form ionic interactions with the backbone phosphate oxygens of G2505.<sup>27</sup> (Though there is some debate as to whether the protonated form is the clinically relevant species.<sup>28</sup>) And, despite the lack of crystallographic data for A2058 mutant ribosomes, these crystal structures have also provided structural insight into the binding interactions that are impacted by A2058 modification. For example, in the A2058G mutants, the C4 methyl group on erythromycin is believed to clash with the N2 of guanine 2058.<sup>2,29,30</sup> (**Figure 2.3**) Binding affinities increase by  $\sim 10^4$  upon mutation of G2099A in *H. marismortui*, which has a naturally occurring guanine in the 2099 position (corresponding to 2058 in *E. coli*).<sup>2</sup>



**Figure 2.3** (A) Telithromycin and A2058 interactions in *H. marismortui* with select distances in Angstroms (Steitz et al. PDB=1YIJ). (B) Steric consequences of A2058G mutation. Image produced with VMD<sup>80</sup>.

Given these findings, Andrade et al. has embarked on a macrolide desmethylation

strategy, whereby the C4 methyl group is removed from the macrolactone ring in order to regain activity against the A2058G mutation. The third-generation macrolide antibiotics telithromycin and cethromycin have been selected for desmethylation. The methyl groups are also removed from the C8 and/or C10 positions to facilitate total synthesis. Because the removal of methyl groups alters the VDW interactions of macrolide antibiotics with the bases delineating the binding pocket, it is important to understand the impact of desmethylation. In the present work, molecular dynamics simulations of telithromycin and cethromycin and a series of their desmethyl analogs are used to relate desmethylation to antimicrobial activity. The approach utilizes a method developed in this laboratory known as conformationally sampled pharmacophore (CSP), which was detailed in Chapter 1.

## **2.2 Research Design and Methods**

### ***2.2.1 Conformationally Sampled Pharmacophore***

Calculations were performed with the program CHARMM, version C36a2.<sup>32</sup> Force field parameters were obtained using a combination of the CHARMM carbohydrate<sup>33-37</sup> and CGenFF<sup>38-40</sup> force fields and the TIP3P water model<sup>41</sup>. All structures were initially minimized by 500 steps of steepest descent (SD) in the gas phase using infinite nonbond lists. The minimized structures were then immersed in a cubic waterbox with a side length of approximately 48 Å. Any waters with an oxygen within 2.5 Å of solute non-hydrogen atoms were deleted. The box length was chosen on the

basis that it extends 14 Å beyond the maximum distance between the solute non-hydrogen atoms. The entire system was then subjected to 2000 steps of SD minimization, with a harmonic restraint of 50 kcal/mol/Å on the solute non-hydrogen atoms, followed by 1000 steps of conjugate gradient (CG) minimization with a harmonic restraint of 0.5 kcal/mol/Å on the solute non-hydrogen atoms. In preparation for the production simulation, the systems were equilibrated for 100 ps in the NVT ensemble (T = 298.0 K) followed by 400 ps in the NPT ensemble (T = 298.0 K, P = 1 atm with a piston mass of 1000.0 amu and gamma value of 25.0 per picosecond)<sup>42-46</sup> allowing equilibration of the water molecules around the solutes. All dynamics were performed using SHAKE for the covalent bonds involving hydrogens<sup>47</sup> and a 2 fs integration timestep with a force switching function applied from 10 to 12 Å for the Lennard-Jones interactions and a non-bonded cutoff list at 14 Å. The electrostatics were treated using Particle Mesh Ewald with a kappa equal to 0.29, a sixth order spline and ~ 1 Å grid spacing, with a real space cutoff of 12 Å and a switching function applied to the forces from 10 to 12 Å.<sup>48</sup> Non-bond lists were updated heuristically during dynamics. Conformational sampling was achieved using Hamiltonian Replica Exchange Molecular Dynamics (HREX MD)<sup>49-51</sup> in the NPT ensemble. In this method, a set of simulations is run in parallel, in which to each simulation (or replica) a biasing potential is applied to dihedral angles along the compound's backbone. Comprehensive sampling is achieved by gradually increasing the biasing potential in the replicas, allowing the compound to surmount dihedral transition energy barriers and escape local minima. For these simulations the dihedral angles were arbitrarily selected to be representative of the whole backbone and correspond to the following: C3-C4-C5-C6, C2-C1-O-C13, O-C13-C12-

C11, and C8-C9-C10-C11. The biasing potential was applied to the force constants of the dihedral angles using 5 replicas so that the first replica was unperturbed and the final replica contained dihedral angle force constants of opposite sign from their original values. Force constants for the intermediate replicas were obtained by linear interpolation. Exchanges were attempted every 1 ps, with coordinates saved every 2 ps for a total simulation time of 10 ns/replica. Only coordinates stored in the unperturbed replica were used for analysis. Conformationally Sampled Pharmacophore (CSP) analysis included the probability distributions for distances between relevant atoms in the compound. Distributions reported correspond to bin sizes of 0.2 Å for **Figures 2.4A-C** and 0.7 Å for **Figure 2.4D**.

### ***2.2.2 Ligand-bound Molecular Dynamics***

Calculations were performed with the program CHARMM, version C35b6<sup>32</sup> and the CHARMM additive force field including the protein<sup>52-55</sup>, nucleic acid<sup>56-59</sup>, carbohydrate, and CGenFF<sup>38-40</sup> parameters and the TIP3P water model<sup>41</sup>. Coordinates were obtained from the protein crystal database (PDB ID 3OAT), with hydrogens added using the HBUILD facility in CHARMM. All molecular dynamics (MD) simulations were performed using a stochastic boundary-based approach that has been presented previously.<sup>60</sup> Briefly, the system was truncated to the region of interest around telithromycin by deleting residues outside of 40 Å of telithromycin's center of mass. Residues were considered within 40 Å if one atom was within the distance criterion. This truncation scheme reduced the number of atoms, making the MD simulations less

computationally expensive. Then, three regions within the sphere were defined. Bases and residues containing one or more atoms within 28 Å comprised the dynamic region, those not in the dynamic region containing one or more atoms within 34 Å comprised the buffer region, and the remainder comprised the outer reservoir region. Atoms within the reservoir region were fixed for all calculations, while varying harmonic restraints were used on atoms within the buffer and dynamic regions as described below. Water was maintained within the sphere using a spherical, quartic restraining potential as implemented in the MMFP module of CHARMM<sup>61</sup> using a 1 kcal/mol/Å force constant and offset parameter (P1) of 2.5 that was applied to the water oxygen atoms.

Prior to dynamics, the entire system was first subjected to 250 steps of steepest descent (SD) minimization with a harmonic restraint of 5 kcal/mol/Å on non-hydrogen atoms within the dynamic region and a mass-weighted harmonic restraint of 10 kcal/mol/Å on non-hydrogen atoms within the buffer region, followed by 250 steps of Adopted-Basis Newton Rhapsod (ABNR)<sup>62</sup> using the same restraints. Equilibration consisted of 400 ps (20 cycles) of Grand Canonical Monte Carlo/Molecular Dynamics (GCMC/MD) using the aforementioned restraints. GCMC/MD is implemented within the MC module in CHARMM and has been described previously<sup>63</sup> as well as will be presented in Chapter 3.

Following equilibration, the C4-desmethyl telithromycin and mutant/modified A2058 ribosomes were generated. Inactive water molecules from the GCMC/MD equilibration were deleted and patches were applied to telithromycin and A2058 in order to generate C4-desmethyl telithromycin with WT, A2058G, N6-monomethyl (MAD), and N6, N6'-dimethyl A2058 (DMAD). Atoms modified during the patch were subjected to

minimization for 200 steps SD and 200 steps CG, and the entire system was allowed to relax for 50 steps SD and 50 steps CG. Parameters for the N6-mono and N6,N6'-dimethyl A2058 have been developed in our lab previously.<sup>60</sup> Two monomethyl systems were studied due to the high energy barrier for the C6-N6 torsion, in which the methyl group in MAD1 is oriented toward telithromycin's desosamine sugar and away from it in MAD2.

All systems were then subjected to 5 ns of Langevin dynamics at 298 K with a friction coefficient of 5/ps and a 2 fs integration timestep using the “leapfrog” Verlet integrator.<sup>64</sup> All dynamics were performed using SHAKE for the covalent bonds involving hydrogens.<sup>47</sup> Nonbond lists were updated heuristically during dynamics with a cutoff of 16 Å, the forces truncated at 12 Å and a switching function applied to the forces from 10 to 12 Å for both electrostatic and van der Waals energy terms. Interaction energies and associated RMS fluctuations reported were calculated using the last 4 ns of the simulation, with the same non-bonded cutoffs as used during dynamics. Snapshots were written every 10 ps. The neutral group surrounding the C4 methyl [C3(=O)-C4(H<sub>2</sub>)-C5] was used so as not to calculate the interaction between species with non-integer charge.

### 2.3 Results

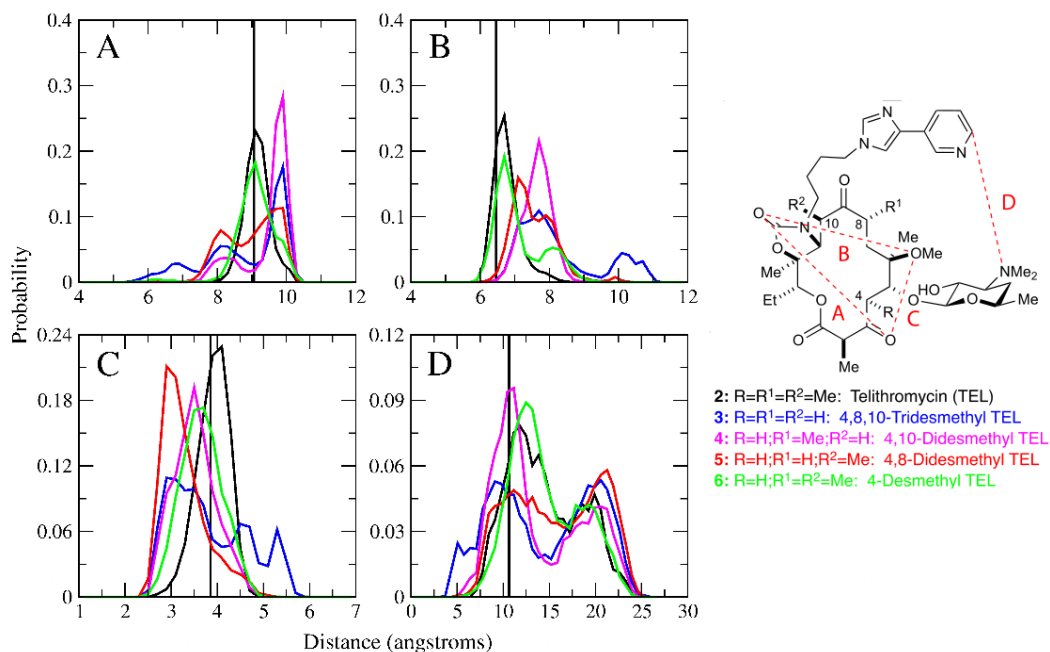
Hamiltonian replica exchange MD simulations of telithromycin and cethromycin and their desmethyl analogs in explicit solvent were undertaken to investigate the impact of removing the C4, C8, and C10 methyl groups on conformational sampling. Employing the CSP approach, probability distributions of select intramolecular distances were

measured to compare bioactivity with conformational flexibility. In the present work, four sets of distances are selected : 1) O16 to O3, 2) O16 to O6, 3) O3 to O6, and 4) N3' to C26, where the atom numbering is as used in<sup>65</sup> and are illustrated in **Figure 2.4** and **Figure 2.5**. These distances were selected because they span the macrolactone ring as well as probe the flexibility of the alkyl-aryl chain extending from N11 and thus are representative of the conformational sampling of the entire ligand. As a measure of bioactivity, minimum inhibitory concentrations (MIC,  $\mu\text{g/mL}$ ) are obtained for telithromycin, cethromycin, and the desmethyl analogs against wild-type (WT), A2058G, and methylated A2058 ribosomes (i.e. bacteria expressing *ermA*). All experimental work was performed by members in the laboratory of Rodrigo B. Andrade at Temple University: Venkata Velvadapu, Bharat Wagh, and Ian M. Glassford.

### ***2.3.1 Wild-type ribosomes***

Telithromycin is active against WT ribosomes with an MIC value of 0.5  $\mu\text{g/mL}$ , **Table 2.1**. In the CSP analysis, **Figure 2.4**, the distribution of distances sampled by telithromycin are in excellent agreement with the crystal structure distances (PDB 3OAT). As methyl groups are removed from the macrolactone ring, the probability distributions become more broad, which indicates an increase in conformational flexibility compared to telithromycin. This broadening appears to be inversely related to bioactivity. The C-4,8,10 tridesmethyl conformer has the least bioactivity with an MIC value that is 64-fold higher than telithromycin in the WT ribosome, while the C-4 desmethyl conformer (C4-des) is just as active as telithromycin and demonstrates the

highest degree of overlap in the ensemble of conformations sampled.



**Figure 2.4** CSP probability distributions for telithromycin (black); C4,8,10-tridesmethyl TEL (blue); C4,10-didesmethyl TEL (purple); 4,8-didesmethyl TEL (red); and 4-desmethyl TEL (green). The vertical line corresponds to the crystallographic distances from PDB

**Table 2.1** Minimum inhibitory concentration (MIC) values in  $\mu\text{g/mL}$  for 4,8,10-tridesmethyl TEL; 4,10-didesmethyl TEL; 4,8-didesmethyl TEL; 4-desmethyl TEL; and telithromycin (TEL). All MIC data was obtained by members of the Andrade laboratory.

Strain	Bacteria	WT/ mutant	4,8,10- trides	4,10- dides	4,8- dides	4-des	TEL
DK/pKK3535	<i>E. coli</i>	WT	32	8	4	0.5	0.5
SQ171/2058G	<i>E. coli</i>	A2058G	>512	>256	>256	>256	>256
DK/2058G	<i>E. coli</i>	A2058G	64	16	32	4	1
ATC33591	<i>S. aureus</i>	ermA	>128	>128	>64	>128	>128

Two didesmethyl telithromycin analogs are studied in the present work: C-4,8 and C-4,10 didesmethyl telithromycin (C4,8- and C4,10-dides, respectively). The probability distributions from these two analogs show intermediate overlap compared to the tridesmethyl and monodesmethyl analogs. Similarly, their antimicrobial activity is in between that of telithromycin and the C-4,C8,C10 tridesmethyl conformer (C4,8,10-trides).

Interaction energies were calculated for the ligands and the 50S ribosomal subunit, **Table 2.2**. The interaction energy approximates the VDW and electrostatic contributions to the enthalpy. Here, the system is a truncated version of the 50S subunit that will be explained in greater detail in Chapter 3, in which all bases and residues outside of 40 Angstroms from telithromycin are removed. As shown in **Table 2.2**, the interactions between telithromycin/C4-des and base 2058 are similar in the WT ribosome. However, when all of the bases and residues of the truncated system are included, the interaction energy is lower for telithromycin than C4-desmethyl telithromycin, which suggests telithromycin binds more favorably.

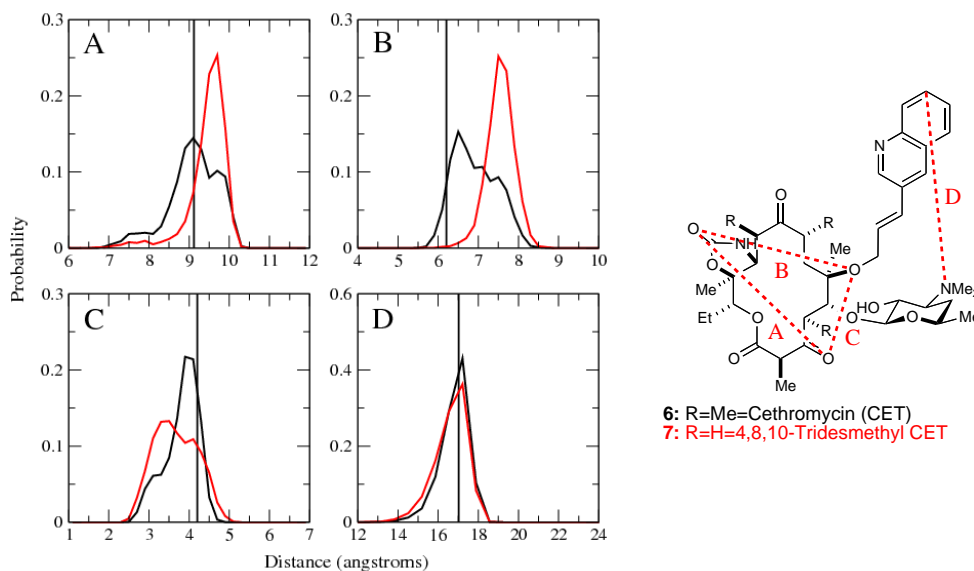
The CSP approach was also applied to another ketolide cethromycin and its tridesmethyl analog, C-4,8,10 tridesmethyl cethromycin (C4,8,10-trides ceth), **Figure 2.5**. The probability distributions for cethromycin and the tridesmethyl conformer show a high degree of overlap suggesting that the C-4,8,10-trides ceth will bind to the ribosome. However, the shift to longer lengths for the distance distributions measured in panels A and B indicate that the macrolactone ring is in a more extended conformation in the tridesmethyl analog. While the MIC values were only measured for C4,8,10-trides ceth in these studies, the analog is shown to be more potent in WT ribosomes (with an MIC

value of 2  $\mu\text{g/mL}$ ) than any of the desmethyl telithromycin analogs.

**Table 2.2** Interaction energies (kcal/mol) between telithromycin and the *E. coli* crystal structure (PDB 3OAT). Calculations were performed for telithromycin and C4-desmethyl telithromycin with the WT, A2058G, MAD1, MAD2, and DMAD ribosomes. Interaction energies were calculated for selected ligand atoms with all the bases/residues within the truncated ribosome (see Methods in Section 2.2) as well as with base 2058 alone. Selected ligand atoms include the neutral group surrounding the C4-methyl [C3(=O)-C4(H<sub>2</sub>)-C5].<sup>a</sup>

Ribosome	Telithromycin		C4-desmethyl telithromycin	
	All Bases/Residues	Base 2058	All Bases/Residues	Base 2058
WT	-9.2 (1.0)	-1.8 (0.4)	-7.7 (1.0)	-1.4 (0.6)
A2058G	-8.5 (1.1)	1.2 (0.5)	-8.6 (0.9)	1.1 (0.3)
MAD1	-9.7 (1.1)	-2.1 (0.5)	-7.3 (0.8)	-2.2 (0.3)
MAD2	-11.7 (1.5)	-3.1 (0.5)	-7.5 (1.2)	-2.1 (0.6)
DMAD	-6.9 (1.4)	-0.9 (0.4)	-7.5 (0.9)	-0.9 (0.3)

<sup>a</sup> Units in kcal/mol. RMS fluctuations shown in parenthesis.



**Figure 2.5** CSP probability distributions for cethromycin (black) and C4,8,10-tridesmethyl CET (red). The vertical line corresponds to the crystallographic distances from PDB 1NWX. Atom pairs represented in A through D are shown in the inset figure.

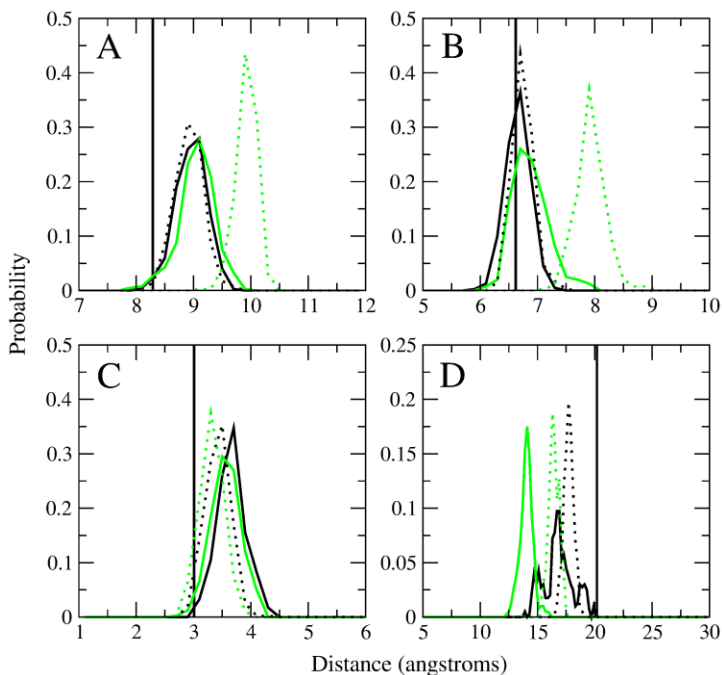
### 2.3.2 A2058G mutant

MIC values were obtained for two types of *E. coli* strains containing A2058G ribosomes, SQ171/2058G and DK/2058G. These strains differ in the relative population of WT and A2058G ribosomes. Namely, SQ171/2058G contains a pure population of mutant ribosomes, while DK/2058G contains a mixed population of WT and A2058G ribosomes. As shown in **Table 2.1**, all of the molecules tested are inactive against the SQ171/2058G strain, with MIC values  $>256 \mu\text{g/mL}$ . For the mixed population strain DK/2058G, telithromycin demonstrates the most potent bioactivity with an MIC of  $1 \mu\text{g/mL}$  and the desmethyl analogs are increasingly less active. C4-des is four-fold less bioactive than telithromycin in the DK/2058G strain, which will be discussed below.

To determine whether the conformational space sampled by telithromycin and the C-4 desmethyl conformer are similar in the WT and A2058G mutant, the CSP approach was applied to macrolide-bound ribosome simulations (**Figure 2.5**) as above. As shown in **Figure 2.6**, there is little overlap in the conformations sampled by telithromycin and C4-des in the A2058G mutant. The probability distributions are shifted to longer distances in C4-des, suggesting that the macrolactone ring adopts a more extended conformation in the A2058G mutant compared to the WT ribosome. Notably, there is a high degree of overlap between the conformations sampled by telithromycin in the WT and A2058G ribosomes suggesting that the A2058G mutation affects the C-4 desmethyl conformer but not telithromycin.

The interaction energy between G2058 and both telithromycin and C4-des is 2.5 to 3 kcal/mol less favorable than the WT A2058 interaction, which is inconsistent with

the hypothesis that C-4 desmethylation enhances binding in G2058-containing bacterial strains by alleviating a steric clash between G2058 C2-amine and telithromycin C4-methyl. Furthermore, the interaction energies between the ribosome and telithromycin/C4-des are similar suggesting that desmethylation does not impact the overall VDW and electrostatic interactions of the ligand within the ribosome. For the tridesmethyl cethromycin analog, its bioactivity as measured by MIC is as potent against the mixed population of A2058G and WT ribosomes in strain DK/2058G as telithromycin, **Table 2.1**. Notably, C4,8,10-trides ceth is more potent than any of the desmethyl telithromycin analogs studied.



**Figure 2.6** CSP probability distributions for selected atom pairs in telithromycin (solid black) and C4-desmethyl telithromycin (solid green) in the WT *E. coli* ribosome and in A2058G mutant (telithromycin, dashed black; C4-desmethyl telithromycin, dashed green). The vertical line corresponds to the crystallographic distances from PDB 3OAT. Atom pairs represented in A through D are shown in the inset of **Figure 2.4**.

### 2.3.3 N6-methylated A2058

Telithromycin and the desmethyl analogs were tested against *Staphylococcus aureus* (*S. aureus*) strain ATCC33591 to determine the MIC values in N6-methylated ribosomes. All of the molecules tested were inactive against ATCC33591 with MIC values >128  $\mu\text{g/mL}$  (or >64  $\mu\text{g/mL}$  for the C4,8-dides conformer), including the C-4,8,10 tridesmethyl cethromycin analog. The CSP approach was not applied to ligand-bound N6-methylated ribosomes since the present study was focused on addressing the A2058G mutation, however interaction energies as shown in **Table 2.1** indicate that the interaction of telithromycin with the dimethylated N6-A2058 (DMAD) ribosome is nearly 2.5 kcal/mol less favorable than with the WT ribosome.

## 2.4 Discussion

The motivation behind the present work was to remove the C4-methyl group in order to recapture telithromycin and cethromycin antimicrobial activity against bacteria containing A2058G 50S ribosomes. In addition, the removal of the C8 and/or C10 methyl groups was attempted to facilitate the total synthesis. Molecular dynamics simulations were performed for each of the analogs to understand the impact of desmethylation on the antimicrobial activity. As shown in **Table 2.1**, the C4-desmethyl analogs display lowered antimicrobial activity compared to the parent macrolides, while the C4,C8/C10-dides and C4,C8,C10-tridesmethyl analogs show even less activity. Following the CSP approach, the lowered activity is explained by the increase in conformational flexibility that arises

from the removal of methyl groups. The presence of an additional methyl in the mono- vs. di- vs. tridesmethyl analogs reduces conformational flexibility, indicating that the former is sampling more of the bioactive conformations, thereby contributing to the improved MIC values. Altogether, these results indicate a model where removal of the methyl groups leads to increased conformational flexibility and therefore lower biological activity.

Binding affinity is determined by the enthalpic and entropic contributions to the Gibbs free energy, **Equations 2.1** and **2.2** (where  $G$  is Gibbs free energy,  $H$  is enthalpy,  $T$  is temperature,  $S$  is entropy,  $R$  is the gas constant, and  $K_D$  is the dissociation constant). The CSP approach is inherently an approximation of configurational entropy because it measures distributions of distances obtained from MD simulations. In the case of the desmethyl analogs shown in **Figure 2.4**, because they are more conformationally flexible than telithromycin in solution they experience a larger decrease in entropy when they bind to the ribosome, which leads to a higher  $\Delta G$  and hence lower binding affinity. Telithromycin and cethromycin are less conformationally flexible in solution, meaning they have a lower entropy and the subsequent decrease in entropy upon binding to the ribosome does not have as large of an unfavorable impact on the Gibbs free energy. In fact, as shown in **Figures 2.4** and **2.6**, the probability distributions of telithromycin and C4-desmethyl telithromycin in solution and bound to the 50S *E. coli* ribosome are very similar, which suggests that the entropic contribution to the binding affinity is small for these two ligands.

$$\Delta G = \Delta H - T\Delta S \quad \text{Equation 2.1}$$

$$\Delta G = RT \ln K_D \quad \text{Equation 2.2}$$

The second contribution to the binding affinity is from the enthalpy, which is approximated by the interaction energies presented in **Table 2.2**. While the interaction energies are a measure of the intermolecular VDW and electrostatic contributions to the enthalpy and do not account for intramolecular terms (though these should be similar in a congeneric series), they are used as a comparison here for telithromycin and C4-desmethyl telithromycin in the various A2058-modified ribosomes. The interaction between C4-desmethyl telithromycin and the bases/residues of the truncated WT ribosome is 1.5 kcal/mol less favorable than telithromycin indicating a lower contribution to the enthalpy. Since the entropic contribution to  $\Delta G$  as measured by CSP probability distributions are similar for C4-des and telithromycin in the WT ribosome, these results indicate that C4-des has a higher  $\Delta G$  and hence binds less favorably. The less favorable enthalpic contribution in the C4-desmethyl analog likely arises from the removal of a methyl group that can form favorable interactions within the binding pocket. While the MIC values indicate that C4-des and telithromycin are equally potent against the WT strain, MIC values are not true dissociation constants, which would provide a more direct comparison with the interaction energy and CSP-based entropic arguments presented here. Furthermore, though the interaction energies for the other desmethyl analogs was not measured here, it is likely that the energies would be even less favorable owing to the removal of methyl groups that form favorable VDW interactions within the binding pocket. Combined with their less favorable entropy contributions upon binding, the di- and tridesmethyl analog would be expected to have even less favorable binding affinities,

as supported by their reduced potency.

Given the original hypothesis that removal of the C4 methyl group would lead to an increase in antimicrobial activity against the A2058G mutation, the results from the current work were difficult to interpret. Early conclusions on the mechanism underlying A2058G-based resistance to macrolides were based on crystal structures of the 50S ribosome in extremophile bacterial species (i.e. *Thermus thermophilus*,<sup>26</sup> *Haloarcula marismortui*,<sup>2</sup> and *Deinococcus radiodurans*<sup>10</sup> from the domain Archaea) because these were the only structures available at the time. Like humans, extremophiles naturally contain guanine at position 2058, rather than adenine as most pathogenic bacteria do. Thus, it was suggested that G2058 ribosomes were unable to bind macrolide antibiotics, presumably due to steric clashes between guanine's C2-amino group and the ligand.<sup>2,10</sup> Furthermore, these conclusions were drawn based on crystal structures of erythromycin, which contains an additional sugar moiety compared to third-generation macrolide antibiotics such as telithromycin. Telithromycin has been found to maintain effectiveness against the A2058G mutation that confers resistance to its parent erythromycin.<sup>2, 66-68</sup> Similarly, the MIC values, as shown in **Table 2.1**, confirm that telithromycin activity is only slightly attenuated in the DK/2058G strain. In the SQ171/2058G strain that contains a pure population of A2058G ribosomes, telithromycin activity is completely abolished. The reason for this is unclear but may be due to other species-specific factors.

Telithromycin and other ketolides have increased affinity for 50S ribosomes owing to alkyl-aryl moieties (ARMs) that extend from their macrolactone rings and form pi-pi stacking interactions with a Watson-Crick base pair A752-U2609. Even so, telithromycin has been relatively unsuccessful in combating microbial resistance

resulting from methylation of the exocyclic N6 of A2058, especially dimethylated ribosomes.<sup>17,69-71</sup> We also observe this in which the MIC values for telithromycin against the methylated ribosomes (*ermA* strain) are over 250-fold greater than against WT ribosomes. Further investigations into the mechanism underlying 2058 methylation-based resistance will be presented in Chapter 3.

It is important to note that comparing the effectiveness of antibiotics, especially against modified ribosomes, is challenging because there are species-specific differences that are often poorly understood. Rarely is the level of 2058 methylation reported when antimicrobial activity of a compound is measured and there is large variation in these levels across bacterial species. In a comprehensive investigation into telithromycin activity as a function of the levels of mono- and di-methylated A2058 ribosomes, it was found that in general telithromycin maintained activity against monomethylated A2058 ribosomes.<sup>17, 72</sup> In fact, high levels of dimethylated ribosomes are required to completely abrogate telithromycin activity. However, notable exceptions were found that presumably arise due to variation in methylation mechanisms among species.

The expression of methyltransferases encoded by *erm* genes can be constitutive or inducible depending on the type of *erm* gene (see <sup>21</sup> and references therein), which ranges among species.<sup>73</sup> Macrolides are known<sup>20,21</sup> to induce the expression of methyltransferase enzymes. However, the level of induction depends on the structure of the macrolide within the binding pocket. For example, erythromycin bound to WT ribosomes has been found to induce *erm* gene expression, while telithromycin does not.<sup>74</sup> The mechanism responsible for erythromycin-induction of *ermC* (*Staphylococcus aureus*) has been described by a translation attenuation model (described thoroughly in <sup>20,75</sup>). The model

involves the following: a sequence-specific leader peptide translated from an *ermC* gene transcript interacts with the erythromycin-bound ribosome causing it to stall during translation. As the ribosome is stalled, *ermC* mRNA isomerizes to a conformation that contains the mRNA sequence encoding a methyltransferase enzyme. The methyltransferase-containing sequence is translated and the resulting methyltransferase enzymes are able to modify free ribosomes, resulting in erythromycin resistance. Other mechanisms exist depending on the *erm* gene type, including transcriptional attenuation that involves the stalling of the ribosome due to a sequence that leads to antitermination of transcription<sup>76, 77</sup> as well as translational attenuation due to conformational changes of the binding pocket in the presence of macrolides.<sup>78</sup>

Variations in sequences of the macrolide binding pocket can also affect the level of methyltransferase enzymes. In a study of the conformation of rRNA surrounding 2058,<sup>79</sup> it was shown that interactions of the Watson-Crick base pair G2057-C2611 modulated the conformation of the macrolide binding pocket such that the levels of methylated ribosomes were affected due to how easily accessible 2058 was to methyltransferase enzymes. Since not all bacterial ribosomes contain the G2057-C2611 base pair, it is reasonable that there are subtle conformational changes that impact A2058 methylation levels.<sup>1</sup> The impact of 2058 modification on these conformational changes will be discussed more in Chapter 3.

## 2.5 References

1. Poehlsgaard, J.; Douthwaite, S. The bacterial ribosome as a target for antibiotics.

*Nat Rev Microbiol* **2005**, 3, 870-81.

2. Tu, D.; Blaha, G.; Moore, P. B.; Steitz, T. A. Structures of MLSBK antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance. *Cell* **2005**, 121, 257-70.
3. Brandi, L.; Marzi, S.; Fabbretti, A.; Fleishcer, C.; Hill, W.; Gualerzi, C. O.; Lodmell, J. S. The translation initiation functions of IF2: targets for thiostrepton inhibition. *J Mol Biol* **2004**, 335, 881-894.
4. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000**, 407, 340-8.
5. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **1996**, 274, 1367-71.
6. Ganoza, M. C.; Kiel, M. C. A ribosomal ATPase is a target for hygromycin B inhibition on *Escherichia coli* ribosomes. *Antimicrob Agents Chemother* **2001**, 45, 2813-9.
7. Hansen, J. L.; Ippolito, J. A.; Ban, N.; Nissen, P.; Moore, P. B.; Steitz, T. A. The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol Cell* **2002**, 10, 117-28.
8. Katz, L.; Ashley, G. W. Translation and protein synthesis: macrolides. *Chem Rev* **2005**, 105, 499-528.
9. Rodnina, M. V.; Savelsbergh, A.; Matassova, N. B.; Katunin, V. I.; Semenov, Y. P.; Wintermeyer, W. Thiostrepton inhibits the turnover but not the GTPase of elongation

factor G on the ribosome. *Proc Natl Acad Sci U S A* **1999**, 96, 9586-90.

10. Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001**, 413, 814-21.

11. Kirillov, S.; Porse, B. T.; Vester, B.; Woolley, P.; Garrett, R. A. Movement of the 3'-end of tRNA through the peptidyl transferase centre and its inhibition by antibiotics. *FEBS Lett* **1997**, 406, 223-33.

12. Tenson, T.; Lovmar, M.; Ehrenberg, M. The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *J Mol Biol* **2003**, 330, 1005-14.

13. Boggiano, B.; Gleeson, M. Gastric acid inactivation of erythromycin stearate in solid dosage forms. *J Pharm Sci* **1976**, 65, 497-502.

14. Fiese, E.; Steffen, S. Comparison of the acid stability of azithromycin and erythromycin A. *J Antimicrob Chemother* **1990**, 25, 39-47.

15. Clancy, J.; Petitpas, J.; Dib-Hajj, F.; Yuan, W.; Cronan, M.; Kamath, A.; Bergeron, J.; Retsema, J. Molecular cloning and functional analysis of a novel macrolide-resistance determinant, *mefA*, from *Streptococcus pyogenes*. *Mol Microbiol* **1996**, 22, 867-879.

16. Douthwaite, S. Structure-activity relationships of ketolides vs. macrolides. *Clin Microbiol Infect* **2001**, 7 Suppl 3, 11-7.

17. Liu, M.; Douthwaite, S. Activity of the ketolide telithromycin is refractory to Erm monomethylation of bacterial rRNA. *Antimicrob Agents Chemother* **2002**, 46, 1629-33.

18. Pfister, P.; Jenni, S.; Poehlsgaard, J.; Thomas, A.; Douthwaite, S.; Ban, N.;

Bottger, E. C. The structural basis of macrolide-ribosome binding assessed using mutagenesis of 23S rRNA positions 2058 and 2059. *J Mol Biol* **2004**, 342, 1569-81.

19. Vester, B.; Douthwaite, S. Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrob Agents Chemother* **2001**, 45, 1-12.

20. Weisblum, B. Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrob Agents Chemother* **1995**, 39, 797-805.

21. Weisblum, B. Erythromycin resistance by ribosome modification. *Antimicrob Agents Chemother* **1995**, 39, 577-85.

22. Denoya, C.; Dubnau, D. Mono- and dimethylating activities and kinetic studies of the ermC 23 S rRNA methyltransferase. *J Biol Chem* **1989**, 264, 2615-24.

23. Nielsen, A. K.; Douthwaite, S.; Vester, B. Negative in vitro selection identifies the rRNA recognition motif for ErmE methyltransferase. *Rna* **1999**, 5, 1034-41.

24. Zhong, P.; Pratt, S. D.; Edalji, R. P.; Walter, K. A.; Holzman, T. F.; Shivakumar, A. G.; Katz, L. Substrate requirements for ErmC' methyltransferase activity. *J Bacteriol* **1995**, 177, 4327-32.

25. Dunkle, J. A.; Xiong, L.; Mankin, A. S.; Cate, J. H. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc Natl Acad Sci U S A* **2010**, 107, 17152-7.

26. Bulkley, D.; Innis, C. A.; Blaha, G.; Steitz, T. A. Revisiting the structures of several antibiotics bound to the bacterial ribosome. *Proc Natl Acad Sci U S A* **2010**, 107, 17158-63.

27. Schlunzen, F.; Harms, J. M.; Franceschi, F.; Hansen, H. A.; Bartels, H.; Zarivach, R.; Yonath, A. Structural basis for the antibiotic activity of ketolides and azalides.

*Structure* **2003**, 11, 329-38.

28. Mao, J. C.-H.; Putterman, M. The intermolecular complex of erythromycin and ribosome. *J Mol Biol* **1969**, 44, 347-361.

29. Velvadapu, V.; Glassford, I.; Lee, M.; Paul, T.; DeBrosse, C.; Klepacki, D.; Small, M. C.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,10-didesmethyl telithromycin. *ACS Med Chem Lett* **2012**, 3, 211-215.

30. Velvadapu, V.; Paul, T.; Wagh, B.; Klepacki, D.; Guvench, O.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolide analogues to address antibiotic resistance: total synthesis and biological evaluation of 4,8,10-tridesmethyl telithromycin. *ACS Med Chem Lett* **2011**, 2, 68-72.

31. Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics *J Molec Graphics* **1996**, 14, 33-38.

32. Brooks, B. R.; Brooks III, C. L.; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J Comput Chem* **2009**, 30, 1545-1614.

33. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* **2008**, 29, 2543-64.

34. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D.

CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* **2009**, *5*, 2353-2370.

35. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, *5*, 1315-1327.

36. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* **2009**, *113*, 12466-76.

37. Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* **2010**, *114*, 12981-12994.

38. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **2010**, *31*, 671-90.

39. Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* **2012**, *52*, 3144-54.

40. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* **2012**, *52*, 3155-68.

41. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**,

79, 926-936.

42. Andersen, H. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* **1980**, 72, 2384.

43. Berendsen, H.; Postma, J.; Gundersen, W. v.; DiNola, A.; Haak, J. Molecular dynamics with coupling to an external bath. *J Chem Phys* **1984**, 81, 3684-3690.

44. Feller, S.; Zhang, Y.; Pasto, R.; Brooks, B. Constant pressure molecular dynamics simulation: The langevin piston method. *J Chem Phys* **1995**, 103, 4613-4621.

45. Hoover, W. G.; Berendsen, H.; Postma, J. M.; Gunsteren, W.; DiNola, A.; Haak, J. R. Canonical dynamics: Equilibrium phase-space distributions  
Molecular dynamics with coupling to an external bath. *Phys Rev A* **1985**, 31, 1695-1697.

46. Nose, S.; Klein, M. Constant pressure molecular dynamics for molecular systems. *Molec Phys* **1983**, 50, 1055-1076.

47. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **1977**, 23.

48. Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh Ewald method. *J Chem Phys* **1995**, 103, 8577-8593.

49. Jiang, W.; Hodoscek, M.; Roux, B. Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics (FEP/REMD). *J Chem Theory Comput* **2009**, 5, 2583-2588.

50. Kannan, S.; Zacharias, M. Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing potential. *Proteins* **2007**, 66, 697-706.

51. Zuckerman, D. M.; Lyman, E. A second look at canonical sampling of biomolecules using replica exchange simulation. *J Chem Theory Comput* **2006**, *2*, 1200-1202.
52. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **2011**, *7*, 3162-3180.
53. MacKerell Jr., A. D.; Bashford, D. B. M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J. G., H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I., W.E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamic studies of protein. *J Phys Chem B* **1998**, *102*, 3586-3616.
54. Mackerell Jr., A. D.; Feig, M.; Brooks III, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **2004**, *25*, 1400-15.
55. MacKerell Jr., A. D.; Feig, M.; Brooks III, C. L. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* **2004**, *126*, 698-9.
56. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **2011**, *32*, 1929-43.

57. Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; Mackerell, A. D., Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* **2012**, *8*, 348-362.
58. MacKerell, A. D., Jr.; Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* **2000**, *21*, 105-120.
59. Mackerell Jr., A. D.; Nilsson, L. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr Opin Struct Biol* **2008**, *18*, 194-9.
60. Small, M. C.; Lopes, P.; Andrade, R. B.; Mackerell, A. D., Jr. Impact of ribosomal modification on the binding of the antibiotic telithromycin using a combined grand canonical monte carlo/molecular dynamics simulation approach. *PLoS Comput Biol* **2013**, *9*, e1003113.
61. Deng, Y.; Roux, B. Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J Chem Phys* **2008**, *128*, 1-8.
62. van Gunsteren, W. F.; Karplus, M. A Method for Constrained Energy Minimization of Macromolecules. *Journal of Computational Chemistry* **1980**, *1*, 266-274.
63. Hu, J.; Ma, A.; Dinner, A. R. Monte Carlo simulations of biomolecules: The MC module in CHARMM. *J Comput Chem* **2006**, *27*, 203-16.
64. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*. Clarendon Press: Oxford, 1987; p 385.
65. Evrard-Todeschi, N.; Gharbi-Benarous, J.; Gaillet, C.; Verdier, L.; Bertho, G.; Lang, C.; Parent, A.; Girault, J.-P. Conformations in solution and bound to bacterial

ribosomes of ketolides, HMR 347 (telithromycin) and RU 72366: a new class of highly potent antibacterials. *Bioorg Med Chem* **2000**, 8, 1579-1597.

66. Berisio, R.; Harms, J.; Schluenzen, F.; Zarivach, R.; Hansen, H. A.; Fucini, P.; Yonath, A. Structural insight into the antibiotic action of telithromycin against resistant mutants. *J Bacteriol* **2003**, 185, 4276-9.

67. Douthwaite, S.; Hansen, L. H.; Mauvais, P. Macrolide-ketolide inhibition of MLS-resistant ribosomes is improved by alternative drug interaction with domain II of 23S rRNA. *Mol Microbiol* **2000**, 36, 183-93.

68. Hansen, L. H.; Mauvais, P.; Douthwaite, S. The macrolide-ketolide antibiotic binding site is formed by structures in domains II and V of 23S ribosomal RNA. *Mol Microbiol* **1999**, 31, 623-31.

69. Kresken, M.; Henrichfreise, B.; Bagel, S.; Brauers, J.; Wiedemann, B. High prevalence of the ermB gene among erythromycin-resistant streptococcus pneumoniae isolates in Germany during the winter of 2000-2001 and in vitro activity of telithromycin. *Antimicrob Agents Chemother* **2004**, 48, 3193-5.

70. Wolter, N.; Smith, A. M.; Farrell, D. J.; Northwood, J. B.; Douthwaite, S.; Klugman, K. P. Telithromycin resistance in Streptococcus pneumoniae is conferred by a deletion in the leader sequence of erm(B) that increases rRNA methylation. *Antimicrob Agents Chemother* **2008**, 52, 435-40.

71. Wolter, N.; Smith, A. M.; Low, D. E.; Klugman, K. P. High-level telithromycin resistance in a clinical isolate of Streptococcus pneumoniae. *Antimicrob Agents Chemother* **2007**, 51, 1092-5.

72. Douthwaite, S.; Jalava, J.; Jakobsen, L. Ketolide resistance in Streptococcus

pyogenes correlates with the degree of rRNA dimethylation by Erm. *Mol Microbiol* **2005**, 58, 613-22.

73. Park, A. K.; Kim, H.; Jin, H. J. Phylogenetic analysis of rRNA methyltransferases, ERm and KsgA, as related to antibiotic resistance. *FEMS Microbiol Lett* **2010**, 309, 151-162.

74. Bonnefoy, A.; Girard, A. M.; Agouridas, C.; Chantot, J. F. Ketolides lack inducibility properties of MLS(B) resistance phenotype. *J Antimicrob Chemother* **1997**, 40, 85-90.

75. Dubnau, D. Translational attenuation: the regulation of bacterial resistance to the macrolide-lincosamide-streptogramin B antibiotics. *CRC Crit Rev Biochem* **1984**, 16, 103-32.

76. Kolter, R.; Yanofsky, C. Attenuation in amino acid biosynthesis operons. *Annu Rev Genet* **1982**, 16, 113-134.

77. Kwak, J.-H.; Choi, E.-C.; Weisblum, B. Transcriptional attenuation control of ermK, a macrolide-lincosamide-streptogramin B resistance determinant from *Bacillus licheniformis*. *J Bacteriol* **1991**, 173, 4725-4735.

78. Sothiselvam, S.; Liu, B.; Han, W.; Ramu, H.; Klepacki, D.; Atkinson, G. C.; Brauer, A.; Remm, M.; Tenson, T.; Schulten, K.; Vasquez-Laslop, N.; Mankin, A. S. Macolide antibiotics allosterically predispose the ribosome for translation arrest. *PNAS* **2014**, 111, 9804-9809.

79. Vester, B.; Hansen, L. H.; Douthwaite, S. The conformation of 23S rRNA nucleotide A2058 determines its recognition by the ErmE methyltransferase. *Rna* **1995**, 1, 501-9.

### **Chapter 3 Impact of Ribosomal Modification on the Binding of the Antibiotic Telithromycin using a Combined Grand Canonical Monte Carlo/Molecular Dynamics Simulation Approach**

Resistance to macrolide antibiotics is conferred by mutation of A2058 to G or methylation by *erm*-encoded methyltransferases of the exocyclic N6 of A2058 (*E. coli* numbering) that forms the macrolide binding site in the 50S subunit of the ribosome. Ketolides such as telithromycin mitigate A2058G resistance yet remain susceptible to *erm*-based resistance. Molecular details associated with macrolide resistance due to the A2058G mutation and methylation at N6 of A2058 were investigated using empirical force field-based simulations. To address the buried nature of the macrolide binding site, the number of waters within the pocket was allowed to fluctuate via the use of a Grand Canonical Monte Carlo (GCMC) methodology. The GCMC water insertion/deletion steps were alternated with Molecular Dynamics (MD) simulations to allow for relaxation of the entire system. From this GCMC/MD approach information on the interactions between telithromycin and the 50S ribosome was obtained. In the wild-type (WT) ribosome, the 2'-OH to A2058 N1 hydrogen bond samples short distances with a higher probability, while the effectiveness of telithromycin against the A2058G mutation is explained by a rearrangement of the hydrogen bonding pattern of the 2'-OH to 2058 that maintains the overall antibiotic-ribosome interactions. In both the WT and A2058G mutation there is significant flexibility in telithromycin's imidazole-pyridine side chain (ARM), indicating that entropic effects contribute to the binding affinity. Methylated ribosomes show lower sampling of short 2'-OH to 2058 distances and also demonstrate enhanced G2057-A2058

stacking leading to disrupted A752-U2609 Watson-Crick (WC) interactions as well as hydrogen bonding between telithromycin's ARM and U2609. This information will be of utility in the rational design of novel macrolide analogs with improved activity against methylated A2058 ribosomes.

### 3.1 Introduction

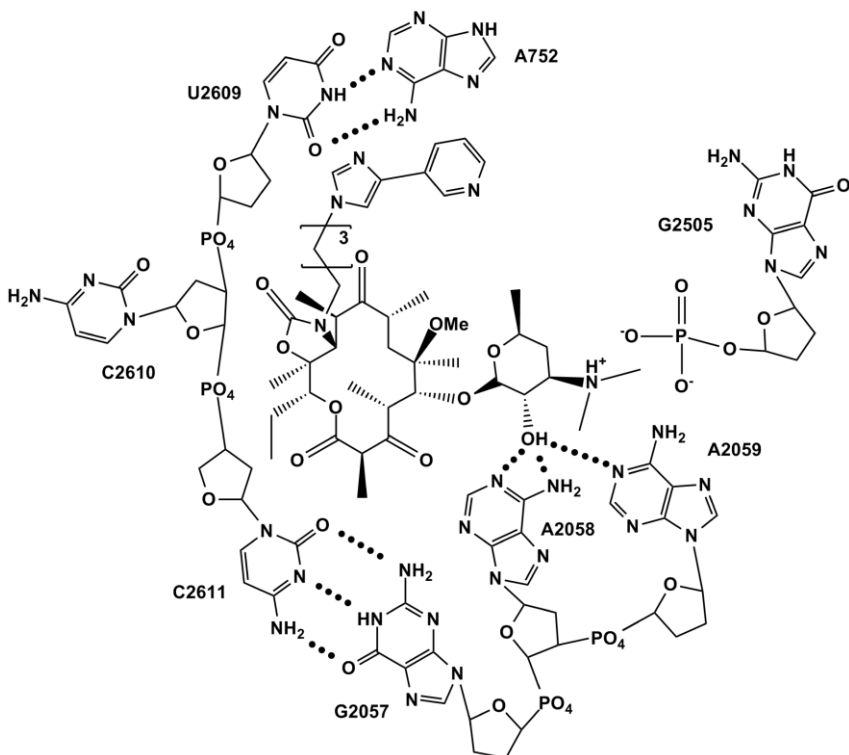
Microbial resistance presents a major challenge in the development of novel antibiotics because bacteria are continually developing new resistance mechanisms.<sup>1</sup> The bacterial ribosome is a target for over 60% of antibiotics,<sup>2, 3</sup> which bind at vital sites within both the 30S and 50S ribosomal subunits and inhibit processes that are essential for cell survival.<sup>4-11</sup> One important class of antibiotics is the macrolides (discussed in detail in Chapter 2), which bind at the beginning of the exit tunnel in the 50S subunit of the bacterial ribosome and block elongation of the nascent polypeptide chain.<sup>11-13</sup> For the macrolide class of antibiotics, bacteria achieve resistance by modifying or mutating bases within the binding pocket as well as by other mechanisms such as drug metabolism and overexpression of efflux pumps. Telithromycin is the first of a new class of macrolides, called the ketolides, named for the substitution of the C3 L-cladinose sugar with a ketone. Ketolides have largely addressed resistance due to drug efflux and metabolism yet still remain susceptible to ribosomal modification,<sup>9, 14</sup> the most clinically relevant of which is modification of A2058 in 23S rRNA (*E. coli* numbering throughout) that confers cross-resistance to antibiotics in the macrolide, lincosamide and streptogramin B classes (MLS).<sup>15-18</sup> A2058-based resistance includes mutation of A2058 to G as well as methylation of the exocyclic N6 of A2058 via the expression of *erm* genes encoding

methyltransferases (Erm methyltransferases) that add one or two methyl groups and represent the most effective mechanism of macrolide resistance.<sup>19-21</sup> Telithromycin and other ketolides have been found to mitigate macrolide resistance due to the A2058G mutation,<sup>22-24</sup> though ketolides that are widely effective against Erm methylation-based modifications are not known. Still, telithromycin has shown improved activity against monomethylated ribosomes compared to previous generation macrolides,<sup>23,25</sup> which has been attributed to secondary contacts made between telithromycin's alkyl-aryl group and bases A752 and U2609 within the binding pocket.<sup>15, 25-27</sup>

Telithromycin's improved activity compared to erythromycin can also partly be explained by its absence of *erm* gene induction.<sup>28</sup> Erythromycin has been shown to increase levels of Erm methyltransferases in species carrying inducible *erm* genes<sup>17,18,25,29-33</sup> consequently increasing the levels of dimethylated ribosomes, whereas telithromycin and other ketolides have been found to mostly bypass the *erm* gene induction pathway.<sup>28,33,34</sup> In a comprehensive study looking at antibiotic activity as a function of mono-/dimethylated 2058 ribosome levels, it was found that in general increased monomethyl A2058 levels alone were not sufficient to mitigate telithromycin activity and that upon increasing the percentage of dimethylated ribosomes via induction with erythromycin was significant telithromycin susceptibility observed.<sup>25</sup> However, species-dependent exceptions were noted and therefore it appears that a combination of factors is involved in enhancing telithromycin's effectiveness in monomethylated ribosomes. Even so, dimethylated ribosomes exhibit the most potent resistance to telithromycin<sup>25</sup> as well as other antibiotics in the MLS class<sup>17</sup> and higher levels of dimethylated ribosomes confer the greatest resistance. Given the widespread transfer of

resistant genes between species<sup>35, 36</sup> it is important to develop analogs with improved affinity toward Erm-mediated resistant phenotypes.

Crystal structures of telithromycin bound to *E. coli*<sup>37</sup> (as well as *T. thermophilus*,<sup>38</sup> *H. marismortui*,<sup>15</sup> and *D. radiodurans*<sup>11</sup>) have revealed several interactions that are important for telithromycin binding, as shown schematically in **Figure 3.1**. Multiple van der Waals (VDW) contacts between hydrophobic alkyl groups on the macrolactone and residues forming the walls of the exit tunnel are present that contribute to binding. Telithromycin's desosamine 2'-hydroxyl group forms hydrogen bonds with N1 of A2058, N6 of A2058, and N6 of neighboring A2059.<sup>11</sup> In addition, desosamine's 3'-dimethylamino moiety forms a salt bridge with the phosphate of G2505.<sup>11</sup> The imidazole-pyridine side-chain (ARM) extending from the C11-C12 cyclic carbamate engages in  $\pi$  stacking with A752 and U2609, increasing telithromycin's binding affinity compared to earlier generation macrolides such as clarithromycin, roxithromycin, and azithromycin.<sup>15,26,27,39</sup>



**Figure 3.12D** representation of telithromycin and bases within the macrolide binding pocket showing the three biologically relevant telithromycin-ribosome interactions that are the subject of this study: 2'-OH to 2058/2059 hydrogen bonding, stacking between telithromycin's 3'-protonated dimethylamine and G2505. Dotted lines represent hydrogen bonding. Hydroxyl groups of the ribose sugars are omitted for clarity.

The present study addresses the molecular details associated with macrolide resistance due to the A2058G mutation and methylation at N6 of A2058 by *erm*-encoded methyltransferases. To compare these interactions in the wild type and telithromycin-resistant strains, empirical force field-based simulations were performed on truncated versions of the wild type and mutant/modified *E. coli* 50S ribosomal subunits. To address the buried nature of the macrolide binding site, the number of waters within the pocket were allowed to fluctuate via the use of a combined Grand Canonical Monte Carlo (GCMC)/Molecular Dynamics (MD) methodology. GCMC/MD has been successfully

applied to systems with deeply buried binding pockets<sup>40, 41</sup> and has been applied to the ribosome to determine the binding free energy of sparsomycin.<sup>42, 43</sup> It is utilized here as a means to assure proper solvation of the telithromycin binding site. The GCMC water insertion/deletion steps are alternated with MD simulations to allow for relaxation of the entire system. From this GCMC/MD approach information on the nature of the interactions between telithromycin and the 50S ribosome was obtained, yielding a detailed understanding of these therapeutically relevant antibiotic resistance mechanisms at the molecular level.

### 3.2 Experimental Section

Calculations were performed with the program CHARMM, version C36a6<sup>51</sup> and the CHARMM additive force field including the protein with the CMAP correction,<sup>52-54</sup> nucleic acid,<sup>55-58</sup> carbohydrate,<sup>59-64</sup> and CGenFF<sup>65</sup> parameters and the TIP3P water model.<sup>66</sup> Coordinates were obtained from the protein database (PDB ID 3OAT<sup>37</sup>), with hydrogens added using the HBUILD facility in CHARMM. Since only the region around the telithromycin binding site was of interest, residues without one or more atoms within 40 Å of the center of system, as defined by the center of mass of telithromycin, were deleted. The system was then overlaid with a 28 Å water sphere and any waters bearing an oxygen within 2.8 Å of a solute non-hydrogen atom were deleted. Mutants were generated using this truncated system, with patches applied to A2058 in order to generate the A2058G, N6-monomethyl, and N6,N'6-dimethyl A2058 variants. Initial guess parameters for the base methylations were obtained from ParamChem.<sup>67, 68</sup> Bond lengths,

angles, and dihedrals specific to the mutations were optimized via comparison to quantum mechanical (QM) geometries, water interactions, molecular vibrations and dihedral potential energy scans. Details of the parametrization are included in the Supporting Information (SI).

MD simulations were performed using a stochastic boundary-based approach<sup>69-72</sup> in which three regions within the sphere were defined. Nucleotides and residues containing one or more atoms within 28 Å comprise the dynamic region, those not in the dynamic region containing one or more atoms within 34 Å comprise the buffer region, and the remaining atoms based on the 40 Å cutoff represent the constrained outer reservoir region. Nucleotides 732, 696, 2458 and 2459 are found along the border of the buffer and reservoir regions, and were manually assigned to the reservoir region; nucleotides 766 and 1324 border the dynamic and buffer regions and were assigned to the buffer region. Atoms within the reservoir region were fixed for all calculations, while varying harmonic restraints were used on atoms within the buffer and dynamic regions as described below. Water was maintained within the sphere using a spherical, quartic restraining potential as implemented in the MMFP module<sup>73</sup> of CHARMM using a 1 kcal/mol/Å force constant and offset parameter (P1) of 2.5 that was applied to the water oxygen atoms.

The entire system was first subjected to 250 steps of steepest descent (SD)<sup>74</sup> minimization with a harmonic restraint of 5 kcal/mol/Å on non-hydrogen atoms within the dynamic region and a mass-weighted harmonic restraint of 10 kcal/mol/Å on non-hydrogen atoms within the buffer region, followed by 250 steps of Adopted-Basis Newton Rhapsion (ABNR)<sup>74</sup> using the same restraints. Equilibration consisted of two

phases of Grand Canonical Monte Carlo/Molecular Dynamics (GCMC/MD), which is implemented within the MC module in CHARMM.<sup>75</sup> The MC module has been described in detail elsewhere,<sup>41,75-78</sup> thus the details of the general methodology will only briefly be addressed as they pertain to the simulations presented here.

A total of 8997 water molecules were used for the bath of water molecules accessible to the GCMC move set. This number was determined by calculating the number of water molecules that would result in a spherical volume of 35 Å radius with density 0.0334 molecules per Å<sup>3</sup>, then multiplied by 1.5 to guarantee that an excess of water molecules were available. The move set was comprised of rigid body translations, rigid body rotations, as well as insertion/deletions that were performed using an excess chemical potential of -5.8 kcal/mol. All moves were equally weighted. The GCMC water pool includes the waters overlaid on the system as described above and the additional water molecules that were placed in a single coordinate set slightly offset from the heterocyclic ARM of telithromycin. All crystallographic waters were set as active throughout the entire simulation. The GCMC waters were initially set as inactive, thereby removing them from all calculations. Each GCMC/MD cycle is defined as 10,000 MC steps and 10,000 MD steps. During the GCMC steps, waters can undergo moves as defined in the move set, in which particle insertions were set to be blocked by all active atoms including hydrogens. The initial hydration phase consisted of 20 GCMC/MD cycles using 5 kcal/mol/Å harmonic restraints on non-hydrogen atoms within the dynamic region and a 10 kcal/mol/Å mass-weighted harmonic restraint on non-hydrogen atoms within the buffer region. For equilibration, the list of active GCMC waters was reset and the harmonic restraints were removed for atoms within the dynamic region and

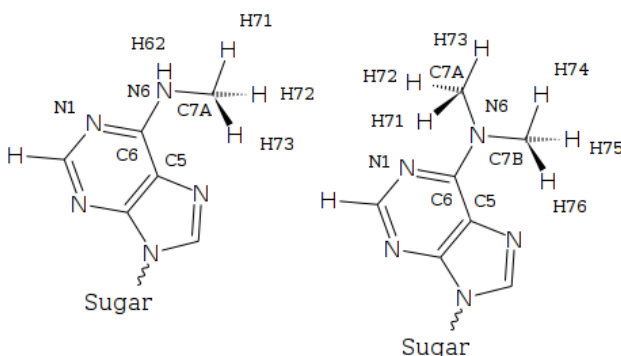
reduced to 2 kcal/mol/Å on non-hydrogen atoms within the buffer region. Iterative GCMC/MD cycles were performed until the number of waters reached adequate convergence (approximately 5 ns cumulative dynamics simulation time). Once the number of waters reached convergence, production consisted of GCMC/MD for a total dynamics simulation time of 25 ns using the same restraints/constraints. Only coordinates from production GCMC/MD were used for analysis. Additional GCMC/MD simulations were performed in which longer MD sampling was performed per cycle (i.e. 100, 250, 500 and 1000 ps) to test convergence of the method with respect to water insertions (see below and SI). As described in the SI, these different simulations of 25 ns sampling each were ultimately combined for the final analysis yielding a total of 150 ns of cumulative MD sampling for each of the studied systems.

MD simulations were performed using Langevin dynamics<sup>79, 80</sup> at 298 K with a friction coefficient of 5/ps, SHAKE<sup>81</sup> was applied to covalent bonds involving hydrogens, and a 2 fs integration timestep was used with the “leapfrog” Verlet integrator.<sup>82</sup> Nonbonded lists were updated heuristically during dynamics with a cutoff of 16 Å, the forces truncated at 12 Å, and a switching function applied to the forces from 10 to 12 Å for both electrostatic and van der Waals energy terms.<sup>83</sup> Coordinates were saved every 10 ps for analysis. Hydrogen bonds were considered present if the hydrogen donor and acceptor atoms came within 2.4 Å<sup>84</sup> and have an occupancy greater than 10%.

### 3.2.1 Parametrization details for methylated A2058

Parametrization of the N6-mono- and N6,N6'-dimethyl adenine moieties followed the CGenFF optimization protocol.<sup>65</sup> For efficiency, the ribose and phosphate groups were removed and N9 was capped with a hydrogen atom. Initial parameters were obtained from the CGenFF program, version 0.9.1, via the ParamChem interface (available online).<sup>67,68</sup> Mol2 files were generated using MOE (Molecular Operating Environment<sup>85</sup>). Charges obtained from ParamChem were replaced by standard charges from the CHARMM36 nucleic acid force field<sup>57, 86-88</sup> for all atoms with the exception of the exocyclic N6 and N9, the latter of which whose ParamChem “guess” charge was maintained because of the capping hydrogen atom. Structures were initially generated in CHARMM<sup>51</sup> using initial guess parameters, then the geometry optimized at the quantum mechanical (QM) MP2 level of theory with the 6-31G\* basis set and tight optimization with the program Gaussian03.<sup>89</sup> Charges were first optimized by comparison of QM and (CHAR)MM interaction energies with water. QM interaction energies were calculated using Gaussian03, with the Hartree-Fock level of theory and 6-31G\* basis set as previously described.<sup>65,88</sup> Because methylation occurs at the exocyclic N6 and presumably would have little effect on the charges of other atoms, the charges for all other adenine atoms were maintained and only the methyl carbon and N6 charges were optimized. Hydrogens of the exocyclic methyl group were maintained at 0.09. **Tables 3.1 and 3.2** show the interaction energies and distances for N6-mono and N6,N6'-dimethyl adenine and water in various geometries, respectively. Only the most favorable geometries were selected as target QM data.

Bond lengths and angle parameters for the exocyclic methyl groups were optimized using the QM target geometry. The N6-monomethyl adenine molecule was optimized first and the parameters transferred to the dimethyl analog. In general, the CGenFF parametrization protocol permits deviations of  $3^\circ$  for the angles and  $0.03 \text{ \AA}$  for the bonds. However, the difference in the angle N1-C6-N6 was found to be  $3.7^\circ$ . The associated parameter belongs to the existing parameters within the nucleic acid force field and not subject to optimization, therefore the slight increase in deviation was deemed acceptable. For N6,N6'-dimethyl adenine, parameters for the bond and angle lengths were initially directly transferred for the monomethyl analog. However, because the agreement with QM data for the bonds and angle was poor, the second methyl group was defined as a unique atom type from the first methyl carbon, allowing the dimethyl analog to be specifically parametrized separately. Again, deviations between the QM and MM bond and angle lengths for the atoms of interest were found to be less than  $3^\circ$  for the angles and  $0.03 \text{ \AA}$  for the bonds, with the exception of N1-C6-N6 which was found to be  $4.7^\circ$ , see **Figure 3.2** for atom numbering.



**Figure 3.2** 2D representation of N6-monomethyl adenine (left) and N6,N6'-dimethyladenine (right) showing the numbering scheme pertaining to the parametrization.

**Table 3.1** Interaction Energies (kcal/mol) and Distances (Å) for N6-monomethyl Adenine and Water Complexes.

Interaction geometry	$\Delta E$ (HF)*	$\Delta E$ (36NA)	$\Delta\Delta E$ (36NA-HF)	r (HF)	r (36NA)	$\Delta r$ (36NA- HF)
H71...OHH	-1.71	-1.54	0.17	2.65	2.65	0.00
H72...OHH	0.06	-0.32	-0.38	2.77	2.71	-0.06
H73...OHH	-0.78	-0.32	0.46	2.75	2.71	-0.04
H62...OHH	-1.74	-1.51	-0.23	2.17	2.02	-0.15
N6.....HOH	-2.88	-2.95	-0.07	2.35	2.31	-0.04
<b>AD</b>	0.08	<b>RMSD</b>	0.29	<b>AAD</b>	0.26	

\* HF energies are scaled by a factor of 1.16. Water interactions are shown only for a subset of atoms pertaining to the added methyl group and surrounding atoms. Included are average deviation (AD), root mean square deviation (RMSD), as well as absolute average deviation (AAD). Atom numbering is as follows (**Figure 3.2**): H71, H72, and H73 correspond to the exocyclic methyl hydrogens, H62 is the exocyclic amine hydrogen, and N6 is the exocyclic amine nitrogen.

**Table 3.2** Interaction Energies (kcal/mol) and Distances (Å) for N6,N6'-dimethyl Adenine and Water Complexes.

Interaction geometry	$\Delta E$ (HF)*	$\Delta E$ (36NA)	$\Delta\Delta E$ (36NA- HF)	r (HF)	r (36NA)	$\Delta r$ (36NA- HF)
H71...OHH	-1.59	-1.13	0.46	2.67	2.70	0.03
H72...OHH	0.64	0.31	-0.33	3.28	2.85	-0.43
H73...OHH	-0.74	-0.13	0.61	2.79	2.76	-0.03
H74...OHH	-1.37	-1.29	0.08	2.69	2.69	0.00
H75...OHH	-0.94	-0.16	0.78	2.78	2.74	-0.04
H76...OHH	-1.51	0.64	2.15	5.68	2.82	-2.86
N6.....HOH	-2.28	-2.80	-0.52	2.33	2.19	-0.14
<b>AD</b>	0.46	<b>RMSD</b>	0.82	<b>AAD</b>	0.71	

\* HF energies are scaled by a factor of 1.16. Water interactions are shown only for a subset of atoms pertaining to the added methyl group and surrounding atoms. Included are average deviation (AD), root mean square deviation (RMSD), as well as absolute average deviation (AAD). Atom numbering as in **Figure 3.2**.

Bond and angle force constants were optimized by comparison of vibrational spectra for the QM and MM geometries, in which QM frequency values were scaled by 0.943.<sup>90</sup> Vibrational spectra were calculated via the MOLVIB module in CHARMM. Frequencies and contributions of the internal degrees of freedom were compared for both the QM and MM geometries and bond/angle force constants were manually optimized until the vibrational spectra were in good agreement, **Tables 3.3** and **3.4** for the mono- and dimethyl adenine molecules.

The force constants for the dihedral angles C5-C6-N6-C7A and C5-C6-N6-C7B, where C7A/B represent the methyl carbons, were optimized using potential energy surfaces (PES). QM scans were performed with Gaussian03<sup>89</sup> from 0° to 180° in 12 steps of 15° for each targeted dihedral. CHARMM was utilized to generate the MM PES. The relative energies at each scan point were compared between QM and MM and the force constants and phases were manually changed until good agreement was achieved (**Figure 3.3**). Phases were constrained to 0° or 180° as per the CGenFF convention. During the optimization emphasis was placed on the low energy regions of the PES as these are sampled during the MD simulations.

**Table 3.3** *N6-monomethyl Adenine QM and MM Vibrational Spectra for the modes involving the methylated amino group.*

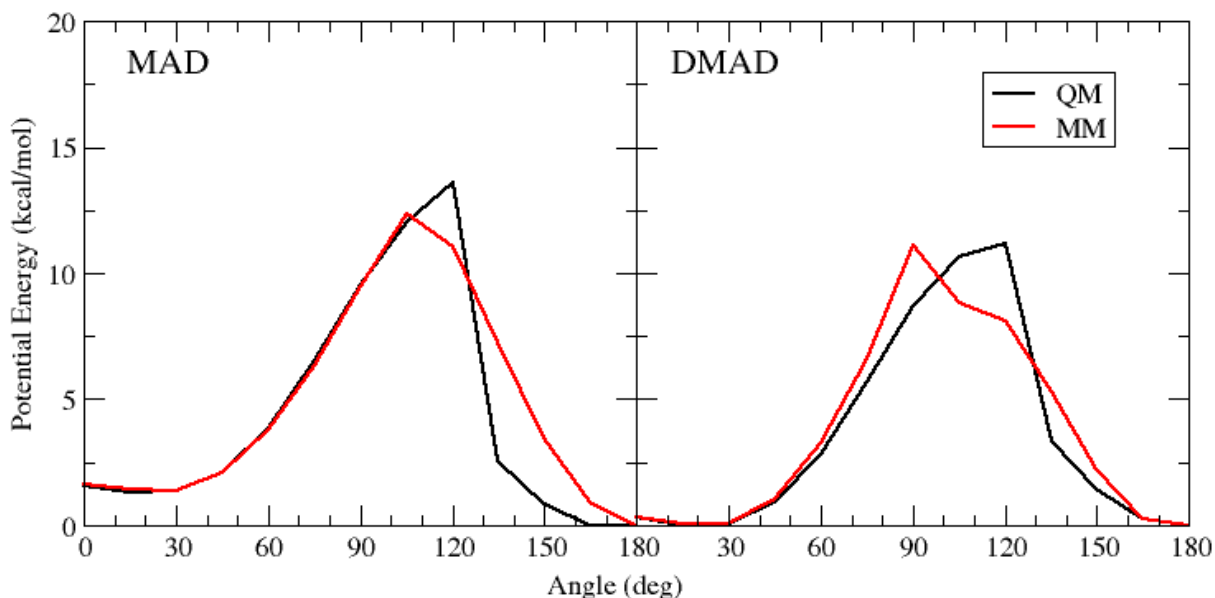
Assignment	MP2/6-31G*		CGenFF	
	Frequency	%	Frequency	%
sN6-H	3437.7	100	3503.6	100
sN6-C7A	994.5	47	955.2	16
	1127.6	19	1087.9	37
	1174.3	16		
sC7A-H	2925.8	100	2978.5	99
	3011.2	100	3022.2	100
	3050.4	99	3022.9	100
dC6N6C7A	170.9	20	184.5	32
	326.7	34	334.0	37
dN6H	1447.8	22	1385.3	18
			1507.2	23
wN6H62	82.2	44	417.3	52
	428.6	23	471.8	19
dsC7AH	1411.6	67	1449.9	98
daC7AH	1479.7	57	1406.3	84
	1502.6	25		
daC7AH'	1454.1	86	1413.6	84
rC7AH	1072.3	25	1043.6	32
	1174.3	33	1098.4	27
rC7AH'	1116.6	78	1065.5	81
			1413.6	15
torN6-C6	82.2	25	129.2	26
	428.6	53	280.4	39
			332.5	27
torN6-C7A	96.2	73	96.0	99

\* MP2/6-31G\* values are scaled by 0.943<sup>90</sup>

**Table 3.4** *N6,N6'*-dimethyl Adenine QM and MM Vibrational Spectra for the modes involving the methylated amino group.

Assignment	MP2/6-31G*		CGenFF	
	Frequency	%	Frequency	%
sN6-C7A(B)	830.3	15	913.6	34
	914.0	38	1289.8	53
	1049.7	23	1335.0	19
	1245.2	46		
sC7B-H	2899.3	98	2977.7	79
	2985.5	100	2978.4	20
	3071.2	99	3020.0	48
			3020.8	93
			3023.4	54
dN6C7A	368.7	50	331.0	34
			349.3	18
			488.1	20
dN6C6	356.6	50	229.4	17
			334.4	19
			359.0	34
wN6C6	91.5	28	208.0	21
	203.0	24	378.9	22
	224.3	20		
dsC7BH	1423.7	79	1462.0	44
			1483.9	28
daC7BH	1474.4	56	1411.5	62
	1477.3	20		
daC7BH'	1462.6	17	1418.5	56
	1477.3	51	1425.2	18
rC7BH	1090.8	20	1041.9	47
	1133.2	26		
rC7BH'	1090.8	21	1022.1	47
	1133.2	23	1094.6	30
torN6-C6	67.4	80	74.9	78
torN6-C7A(B)	124.5	77	122.6	75
	148.1	60	148.9	94
	231.2	18		

\* MP2/6-31G\* values are scaled by 0.943.<sup>90</sup>



**Figure 3.3** Potential energy scan for (left) N6-monomethyl adenine for the dihedral C5-C6-N6-C7A and (right) N6,N6'-dimethyl adenine dihedral C5-C6-N6-C7B, where C7A/B represent the exocyclic methyl carbons as shown in **Figure 3.2**.

Because optimization of the dihedral yielded high potential energy barriers for this dihedral, it was necessary to consider two separate conformations for the N6-monomethyl mutant, one in which the methyl group was oriented cis to the N1 atom and the other in which the methyl is trans. QM and MM energies were calculated for both orientations, in which the methyl group oriented cis to N1 was designated MAD1 and trans to N1 as MAD2. **Table 3.5** shows the QM and MM energies for both orientations. QM energies were obtained using Gaussian03<sup>89</sup> using the MP2 level of theory and 6-31G\* basis set and tight optimization. MM energies were obtained in CHARMM using the optimized parameters aforementioned and performing an energy minimization in vacuo with conjugate gradient for 200 steps and Newton-Raphson minimization for 50 steps with a tolerance of 0.00001 using infinite non-bond cutoffs.

**Table 3.5** QM versus MM relative energies for MAD1 versus MAD2.

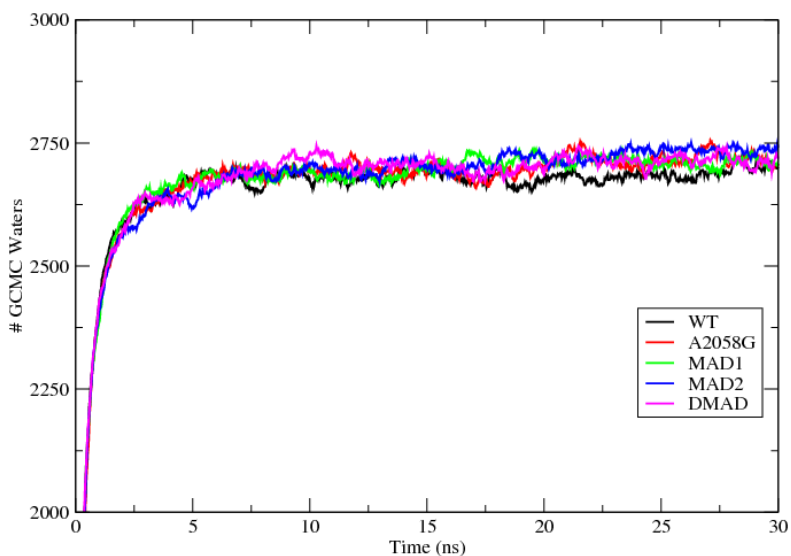
Energy Type (kcal/mol)*	MAD1	MAD2
QM	0.00	1.25
MM	0.00	1.30

\* Energy values shown are relative to MAD1.

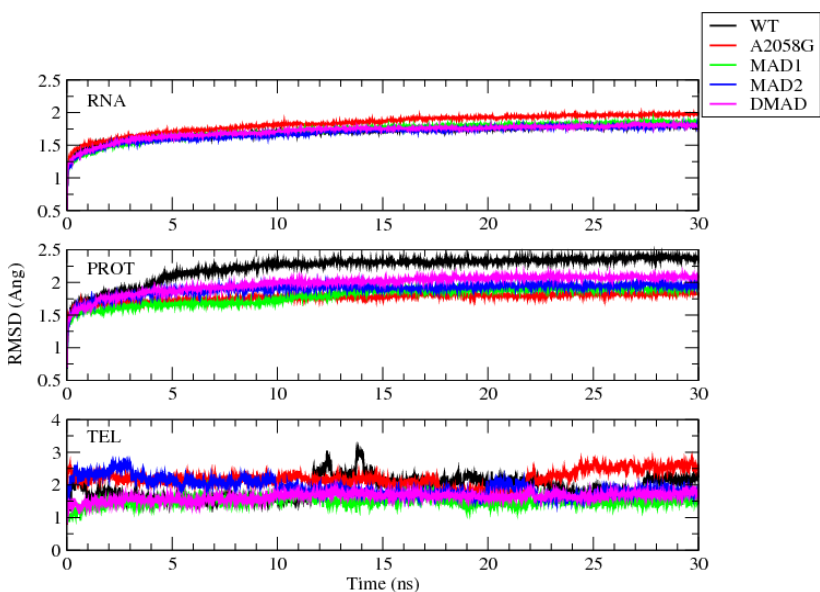
### 3.2.2 System Set-up and Equilibration

Simulations of the entire ribosomal complex are computationally expensive- e.g. solvating the complex within a box of water would result in a system of over 1 million atoms.<sup>91,92</sup> Since only atoms pertaining to the telithromycin binding site are of interest for the present study, a stochastic boundary approach<sup>69-72,77</sup> is utilized whereby atoms of interest encompassing the telithromycin binding site “see” an external field that is representative of atoms outside of the region of interest. Furthermore, because of the buried nature of the binding site and the fact that it is inaccessible to the bulk region, a Grand Canonical Monte Carlo-Molecular Dynamics (GCMC/MD) approach is employed.<sup>41,75</sup> GCMC/MD allows for the number of waters to fluctuate during the course of the simulation resembling an open system at a constant chemical potential. Hence it is utilized here as a means to robustly determine the degree of hydration within the telithromycin binding site during the simulation. GCMC/MD is divided into two stages. In the first (GCMC) waters are inserted, deleted, rotated, or translated according to a Metropolis acceptance criterion<sup>93</sup> that is dependent on the excess chemical potential of water and the change in energy between the putative position and existing position. The

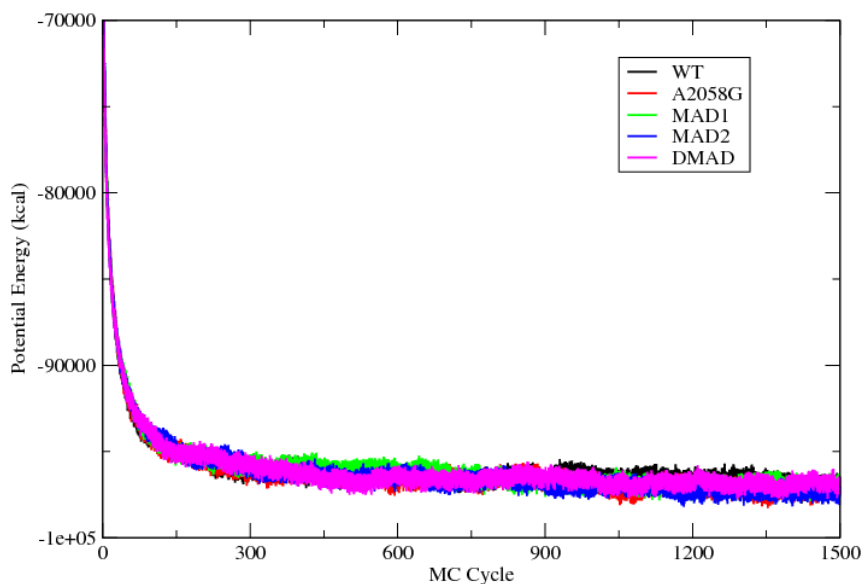
second stage (MD) allows for subsequent relaxation of the system's dynamic regions, including the newly placed waters. These stages are repeated iteratively until the number of waters within the system reaches convergence, with this stage of the simulation considered equilibration. All the systems (i.e. WT, mutant, and modifications) were treated in this manner and found to reach convergence similarly. Shown in **Figure 3.4** are the number of waters as a function of simulation time for all the systems studied. The number of active GCMC waters approaches convergence by 5 ns (a total of 250 GCMC/MD cycles) and this portion of the trajectory is considered equilibration. At this point, the running average of GCMC waters within the 35 Å sphere was determined to be 2115 with a standard deviation of 2 water molecules for WT. Equilibration of the system was confirmed by root mean square deviation analysis (**Figure 3.5**), which shows that the ribosome (RNA and protein) is stable and has relaxed around the newly placed waters by 5 ns. Though the dynamics is not continuous as in standard molecular dynamic simulations, because the positions of the system (namely, RNA, protein, and ligand) are unchanged during the GCMC portion, it is treated as continuous for the purposes of analysis. From the RMSD, the deviation is less than 2.5 Å in all cases and was assumed to be adequately converged by 5 ns such that the remaining 25 ns was considered production and was used for analysis. As a check, the potential energy as a function of cycle was calculated and was also found to stabilize by 250 cycles (or 5 ns dynamics time, **Figure 3.6**). Following equilibration, production GCMC/MD is performed which allows the systems to sample a range of conformations and degrees of hydration.



**Figure 3.4** The number of GCMC waters as a function of dynamics simulation time for each system studied. Number of GCMC waters were those present at the start of MD during each cycle. WT=wild-type (black), A2058G=A2058 mutated to G (red), MAD1/MAD2= the N6- monomethyl adenine with the methyl cis (green) and trans (blue) to the N1 atom, DMAD= N6,N6'-dimethyl adenine (magenta).



**Figure 3.5** The RMSD (in Ang) for RNA (top panel), protein (middle), and telithromycin (bottom) for the WT and mutant/modifications over the 30 ns GCMC/MD simulation. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta.



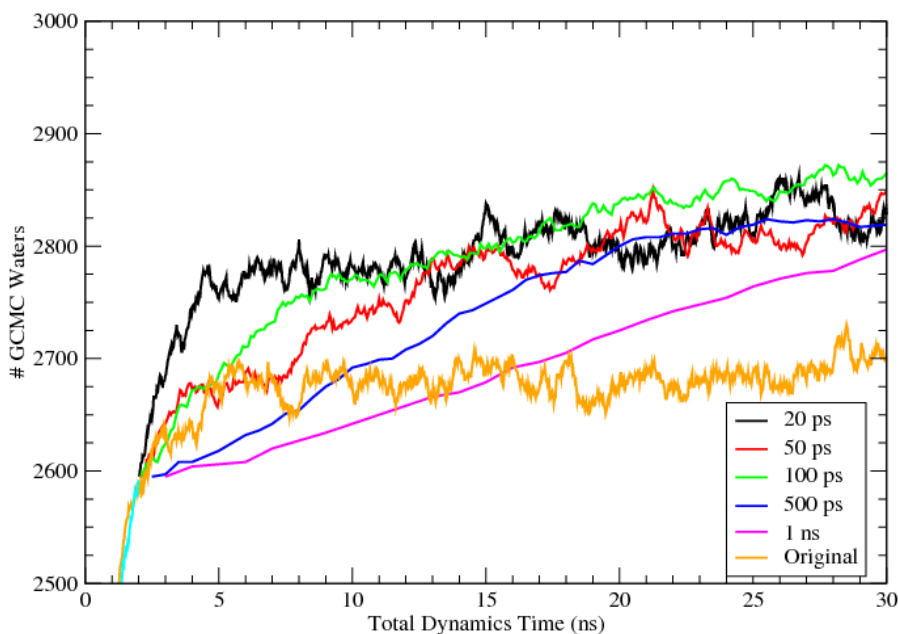
**Figure 3.6** The potential energy as a function of cycle for the WT and mutant/modifications over the 30 ns GCMC/MD simulation. 1500 cycles is equivalent to 30 ns dynamics simulation time. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta.

### 3.2.3 Testing MD simulation lengths during each GCMC/MD iteration.

The convergence of the GCMC/MD simulations was confirmed by running multiple simulations using longer MD during the GCMC/MD cycles. The number of MC steps remained unchanged at 10,000. Starting from the coordinates obtained after the initial hydration phase (which involves 20 cycles of GCMC/MD with restraints/constraints as defined in the methods), these additional simulations consisted of an initial 100 cycles of GCMC/MD using 10,000 steps-20 ps MD per cycle for a total of 2 ns. At this point the majority of GCMC waters were inserted. The coordinates were then used to start 4 separate simulations in which the number of MD steps was increased to 25,000 steps, 50,000 steps, 250,000 steps, or 500,000 steps for total dynamics simulation times of 50 ps, 100 ps, 500 ps, and 1 ns per cycle, respectively. A fifth

GCMC/MD simulation was also performed using the same number of MD steps per cycle (10,000 steps/20 ps) so as to test the original simulation protocol. GCMC/MD simulations were continued for 28 ns bringing the total dynamics time to 30 ns with the last 25 ns used for analysis. This procedure was applied to all the WT and A2058 mutant/modified systems.

Analysis showed the number of waters inserted during GCMC to be correlated to the length of the MD, **Figure 3.7**. Thus, the longer MD simulations appear to allow for additional relaxation of the structure allowing for a higher number of waters to be inserted overall. To determine whether this effect would alter the dynamics within the macrolide binding pocket, the number of waters within 10 Å of telithromycin was counted over the last 25 ns and averaged for each simulation (**Table 3.6**). Waters were counted if the oxygen atom resided within 10 Å of the center and the final snapshot of each trajectory was used to determine the water coordinates. The number of waters was found to be similar irrespective of the length of MD and therefore not likely to affect the dynamics within the pocket. RMSD analysis was also performed for the varying MD simulations and found to be within 2.5 Å (not shown).

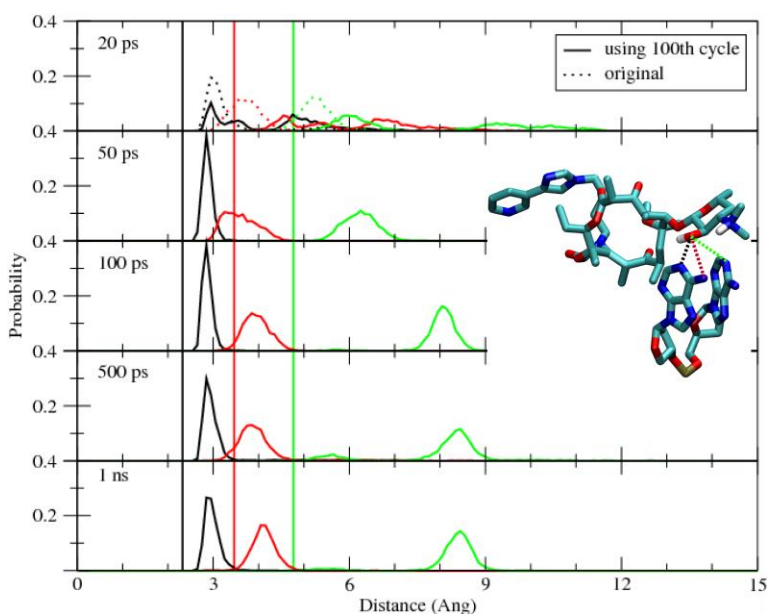


**Figure 3.7** The number of GCMC waters as a function of time for varying lengths of MD performed to check for convergence. Original refers to the GCMC/MD simulation in which all 1500 cycles were run continuously using 10,000 MC steps and 20 ps MD. The 20, 50, 100, 500 ps, and 1 ns lines refer to the GCMC/MD simulations used to test for convergence and were performed by first running 100 cycles of GCMC/MD (10,000 MC steps/20 ps MD), then using the final snapshot to start 5 separate GCMC/MD simulations with varying MD steps as described in the text. For this reason, the number of GCMC waters in the latter GCMC/MD simulations are shown starting at 2 ns with the number of GCMC waters during the first 2 ns shown in cyan. The Original starts at the origin.

**Table 3.6** The number of GCMC waters within 10 Å of telithromycin averaged over the last 25 ns of GCMC/MD for varying lengths of MD. Water molecules with oxygen atoms within 10 Å of the center are considered.

System	# waters
Original (MD=20 ps)	52
New MD=20 ps	42
MD=50 ps	42
MD=100 ps	43
MD=500 ps	42
MD=1 ns	42

The impact of the length of the MD simulations, as well as convergence of the sampling, was also assessed by analyzing probability distributions of the telithromycin 2'-OH to N1/N6 A2058/A2059 distances for the wild-type and mutant simulations (**Figure 3.8**). Overall, good agreement was observed between the simulations using different duration MD simulations indicating that the ensemble of conformations sampled during the different simulations were sampling similar, though not identical regions of conformational space. For this reason, the conformations from all of the simulations were combined and used for analysis, which yielded a total of 150 ns cumulative sampling time for each of the studied systems.

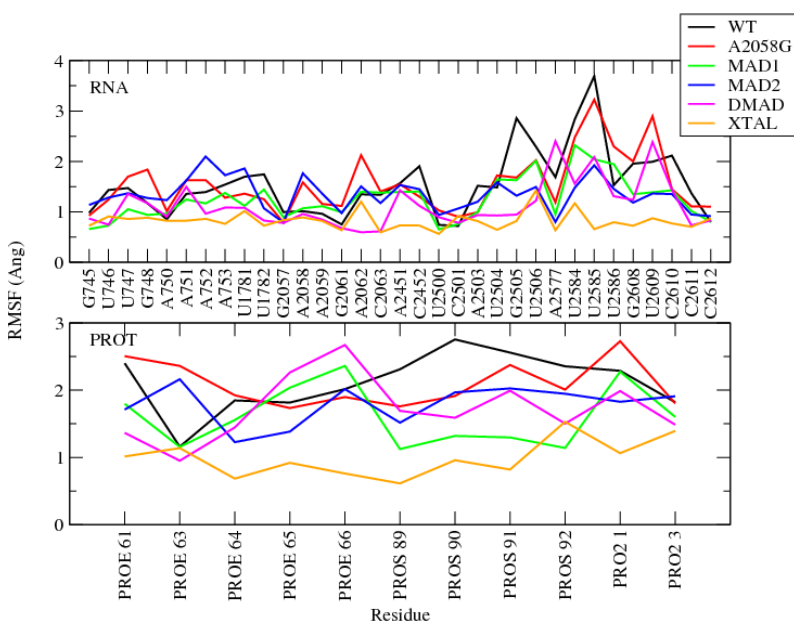


**Figure 3.8** The probability distributions of the telithromycin 2'-OH to N1, N6 (A2058) and N6 (A2059) for the various lengths of molecular dynamics simulations performed to determine convergence. The 2'-OH – N1 (A2058) is shown in black, 2'-OH to N/O6 (A/G2058) in red, and 2'-OH to N6 (A2059) in green. Corresponding crystal structure distances are illustrated as vertical lines using the same color scheme. The two GCMC/MD simulations obtained with 20 ps MD are shown. Original refers to the GCMC/MD simulation in which all 1500 cycles were run continuously using 10,000 MC steps and 20 ps MD. The second simulation in which the coordinates were obtained from an initial 100 cycle GCMC/MD simulation is designated as such and is described in the text.

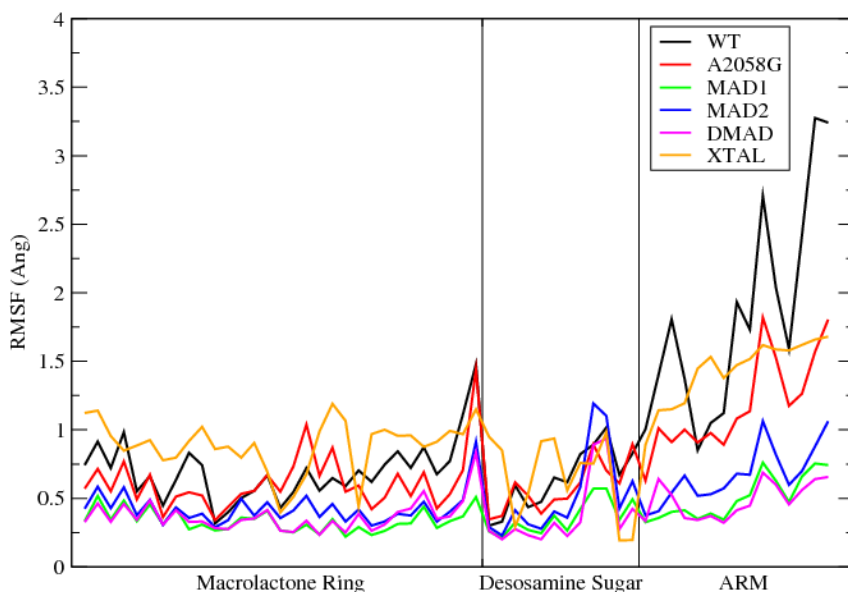
### 3.2.4 Root mean square fluctuations to determine conformational flexibility

Root mean square fluctuations (RMSF) were calculated for RNA and protein within 10 Å of telithromycin as well as telithromycin to determine whether A2058 modification impacts the conformational flexibility of the macrolide binding pocket or of the macrolide. RMSF values were calculated for non-hydrogen atoms of the nucleotide (including base, sugar, and phosphate group). B-factors are also shown for completeness and are calculated by averaging over the atomic B-factors for each nucleotide.

**Figure 3.9** shows that RNA and protein are in general more conformationally flexible in the WT ribosome. Moreover, **Figure 3.10** indicates that the arm of telithromycin is more flexible in the WT and A2058G mutant and that this increased flexibility coincides with the B-factors from the crystal structure as discussed in the text.



**Figure 3.9** RMS Fluctuations (Å) for RNA (top panel) and protein (bottom panel) within 10 Å of telithromycin in the WT and mutant/modifications. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, DMAD in magenta, and crystal structure values in orange. RMSF values reported are for the entire nucleotide or residue (including backbone).



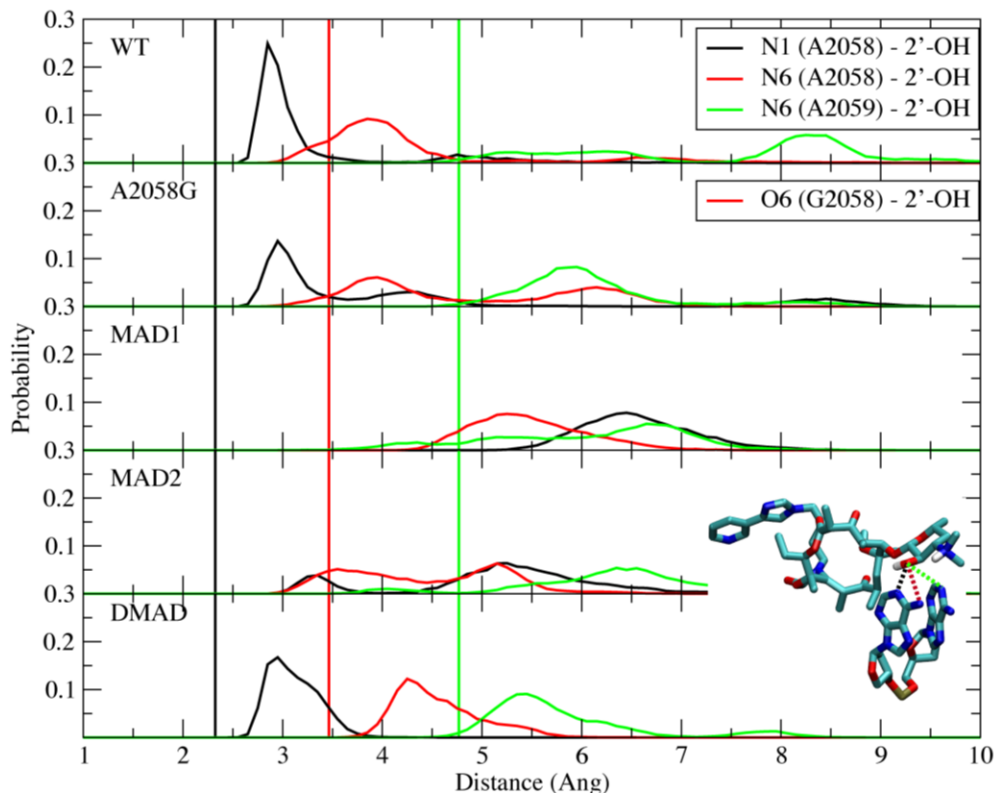
**Figure 3.10** RMS Fluctuations ( $\text{\AA}$ ) for telithromycin in the WT and mutant/modifications. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, DMAD in magenta, and crystal structure values in orange. RMSF values reported are for non-hydrogen atoms only.

### 3.3 Results and Discussion

The present work sets out to understand details of the impact of ribosome modifications on the ketolide antibiotic telithromycin using MD simulations. For computational expediency, as required to obtain adequate conformational sampling, the ribosomal system was truncated to a 40  $\text{\AA}$  radius spherical system centered on telithromycin and a GCMC/MD approach was used to assure proper solvation of the system. Validation of the methodology along with tests of convergence of the simulations are detailed in the Supporting Information (SI). Applying this methodology allowed for 150 ns of MD sampling of each of the studied species, with analysis focusing on the interactions between telithromycin and the ribosome shown in **Figure 3.1** and how those

interactions impact the local structure of the ribosome, thereby leading to altered telithromycin-ribosome binding modes.

Initial analysis involved the hydrogen bonds between telithromycin's desosamine hydroxyl (2'-OH) and A2058 N1/N6 or A2059 N6 as these have been indicated to play a vital role in telithromycin binding and are presumably disrupted in A2058-modified ribosomes.<sup>15,37,38</sup> Probability distributions of the 2'-OH to N distances are shown in **Figure 3.11** for the WT and A2058 mutated/modified bacterial ribosomes. Distances for the WT simulation are in good agreement with the crystal structure distances. The maximum in the 2'-OH to A2058 N1 probability distribution occurs at 2.95 Å versus the crystal structure distance of 2.32 Å. The crystal distance, which is short for a hydrogen bond (i.e. WC N1-N3 interaction distances are approximately 2.9 Å<sup>44</sup>) is probably explained by the resolution of the crystal structure (3.25 Å) obtained at a temperature of 100 K. In addition, the average B factors for telithromycin and A2058 (base atoms only) are 28.5 and 31.4, respectively, which correspond to root mean square fluctuations of nearly 1 Å each at the temperature of 100 K. The 2'-OH to A2058 N6 WT distance is also in good agreement with the crystal data. Notably the 2'-OH to N6 interaction is not direct, as indicated by the interaction distances being > 3 Å. There is also an interaction between the 2'-OH and A2059 N6, which samples a range of distances in the simulation compared to a crystal structure interaction distance of ~4.7 Å indicating that it is a relatively weak interaction. The 2'-OH to N6 A2059 hydrogen bond was proposed based on crystal structures of other macrolides (e.g., erythromycin, clarithromycin, and roxithromycin) bound to the *Deinococcus radiodurans* 50S subunit,<sup>11</sup> but does not play a significant role here.



**Figure 3.11** The probability distributions of the telithromycin 2'-OH to A2058 N1, A2058 N6 and A2059 N6 distances for WT, A2058G mutant and methyl modifications. The 2'-OH – A2058 N1 distance is shown in black, 2'-OH to A2058 N6/O6 in red, and 2'-OH to A2059 N6 in green. Distances reported are measured between heavy atoms. Corresponding crystal structure distances are shown as vertical lines using the same color scheme.

Telithromycin activity is slightly decreased in the A2058G mutant.<sup>23, 24</sup> Consistent with this is the distance probability distributions in **Figure 3.11** showing the hydrogen bonding to be largely maintained. This is somewhat surprising given that upon going from A to G the N1 atom becomes a donor and the exocyclic N6 amine of A is replaced by a keto group in G. The maintenance of the 2'-OH to N1 interaction in A2058G is due to the 2'-hydroxyl now acting as an acceptor and the corresponding

reorientation of the hydroxyl proton allowing it to form a favorable hydrogen bond with the O6 of G. Thus, the 2'-hydroxyl in telithromycin can accommodate the switch in the location of the donor and acceptor groups in the A2058G mutation thereby maintaining favorable hydrogen bonding and explaining how telithromycin remains effective against A2058G mutants.

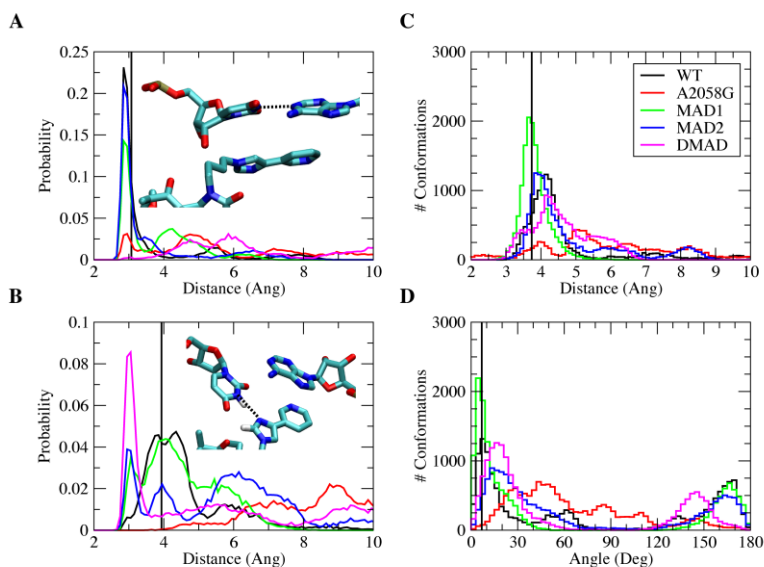
Erm-dependent methylation of A2058 leads to telithromycin resistance, with activity diminished in monomethylated ribosomes and almost completely abolished in dimethylated ribosomes.<sup>23</sup> For monomethylated ribosomes, the N6-methyl group of A2058 can be oriented toward the desosamine sugar or away from it and QM calculations (**Table 3.5**) indicate both orientations to be energetically accessible, though the barrier between the orientations is high. Accordingly, both orientations of the N6 monomethyl modifications were explicitly studied, where the methyl group is directed either toward (MAD1) or away from (MAD2) the desosamine sugar (i.e., cis or trans to A2058 N1, respectively). The hydrogen bonding distributions for the methylated species are included in **Figure 3.11**. With the N6 monomethylations (i.e., MAD1 and MAD2), the hydrogen bonding is significantly perturbed, with that perturbation larger in MAD1, consistent with the methyl being directed towards telithromycin. However, in the N,N'-dimethylated species the hydrogen bonding pattern is well maintained. The N1 to 2'-OH distribution is only slightly perturbed from that in the WT and the N6 to 2'-OH interaction is also well maintained, with a shift to longer distances as compared to the WT due to the presence of the methyls. Thus, while the perturbation of the A2058-telithromycin interactions in the monomethyl modifications is consistent with their lower activity, the maintenance of

interactions with the N,N'-dimethyl modification indicate additional factors are impacting the efficacy of telithromycin against the methylated species.

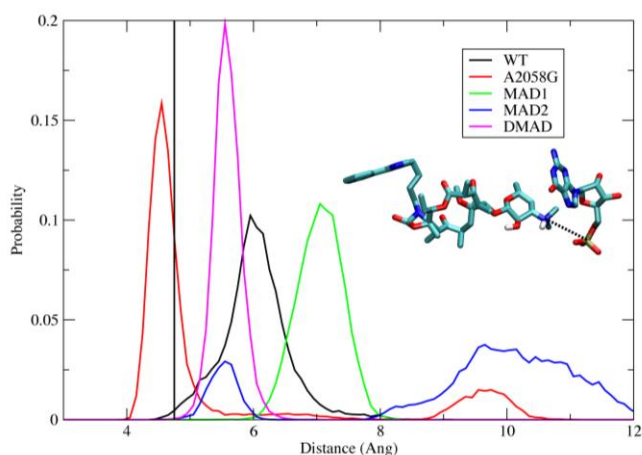
With this in mind, the interactions of telithromycin's imidazole-pyridine moiety (ARM) with A752-U2609 were investigated. Telithromycin's increased affinity for the 50S ribosome macrolide binding pocket compared to earlier generation macrolides has been linked to these secondary interactions.<sup>15,37,38</sup> Crystal structures show that the heterocyclic ARM stacks with U2609 and A752,<sup>37,38</sup> (**Figure 3.1**), which themselves form a Watson-Crick (WC) pair that bridge domains II and IV of 23S rRNA.<sup>37</sup> Accordingly, the properties of A752-U2609 and the ARM were compared between the WT and mutant/modified ribosomes. The A752 N1 to U2609 N3 WC distance (**Figure 3.12A**) is well maintained in both the WT and MAD2 systems, though some sampling of longer distances is evident. A2058G and the other methylated species undergo a significant loss of WC base pairing, with MAD1 followed by A2058G still sampling some of the WC state. Concerning A752-U2609 to ARM interactions, in the simulations stacking was observed to varying degrees in the different systems. Probability distributions of the COM distance and angles between least-square planes through the A752-U2609 and the ARM (**Figure 3.12C,D**) show a significant amount of stacking in all the systems except A2058G. However, in all the systems a wide range of spatial arrangements of the bases and the ARM beyond the stacked state are being sampled. This is associated with high fluctuations of the ARM, as evidenced by the RMS fluctuations of telithromycin in the different systems (**Figure 3.10**). The highest fluctuations occur with the WT and A2058G systems, with the magnitude of the fluctuations consistent with those obtained from the crystallographic study based on the B factors. Notably, the

fluctuations of the ARM, which were obtained at 100 K in the experimental study, are larger than the remainder of the ketolide, indicating that the ARM is likely sampling a range of conformations, including the stacked state, as observed in the simulations. Interestingly, the methylations have the lowest fluctuations of the ARM. Further analysis of A752-U2609 to ARM interactions revealed the presence of hydrogen bonds between U2609 and the imidazole ring of the ARM (**Figure 3.12B**). Thus, the methylations lead to perturbations of the region of the ribosome around the ARM that lead to ARM-U2609 hydrogen bonding, thereby contributing to the lowered fluctuations of the ARM, while the WT and A2058G systems have a highly flexible ARM. Given the activity of telithromycin against these species, these results indicate that the flexibility is important for activity. However, the presence of the hydrogen bonds between the ARM and U2609 in the methylated species has interesting implications in the context of ligand design, as discussed below.

The third interaction considered is the electrostatic interaction between telithromycin's 3'-protonated dimethylamine and the phosphate of G2505 (**Figure 3.1**).<sup>11</sup> Distance distributions for the 3'-nitrogen to the phosphate are shown in **Figure 3.13**. WT and MAD1 sample the shortest distances ( $\sim 3.5$  Å) with A2058G and DMAD also forming stable interactions. While WT and MAD1 do sample shorter distances than the other species, it is not evident that this interaction discriminates between the WT, A2058G, MAD1 and DMAD species. The favorable interactions of MAD1 may contribute to telithromycin maintaining weak activity against the monomethylated species,<sup>25</sup> while the systematically longer distances with MAD2 may contribute to a loss of binding of telithromycin to that species.



**Figure 3.12** The probability distributions for WT, A2058G mutant and methyl modifications of the (A) WC A752 N1 to U2609 N3 distances, (B) WC A752 N1 to U2609 N3 distances, (C) distance between the centers of masses of A752-U2609 and telithromycin's imidazole-pyridine (ARM) moiety and (D) angle between the planes defined A752-U2609 and the imidazole-pyridine moieties. The crystal structure values are shown as a vertical black line, while the distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta).

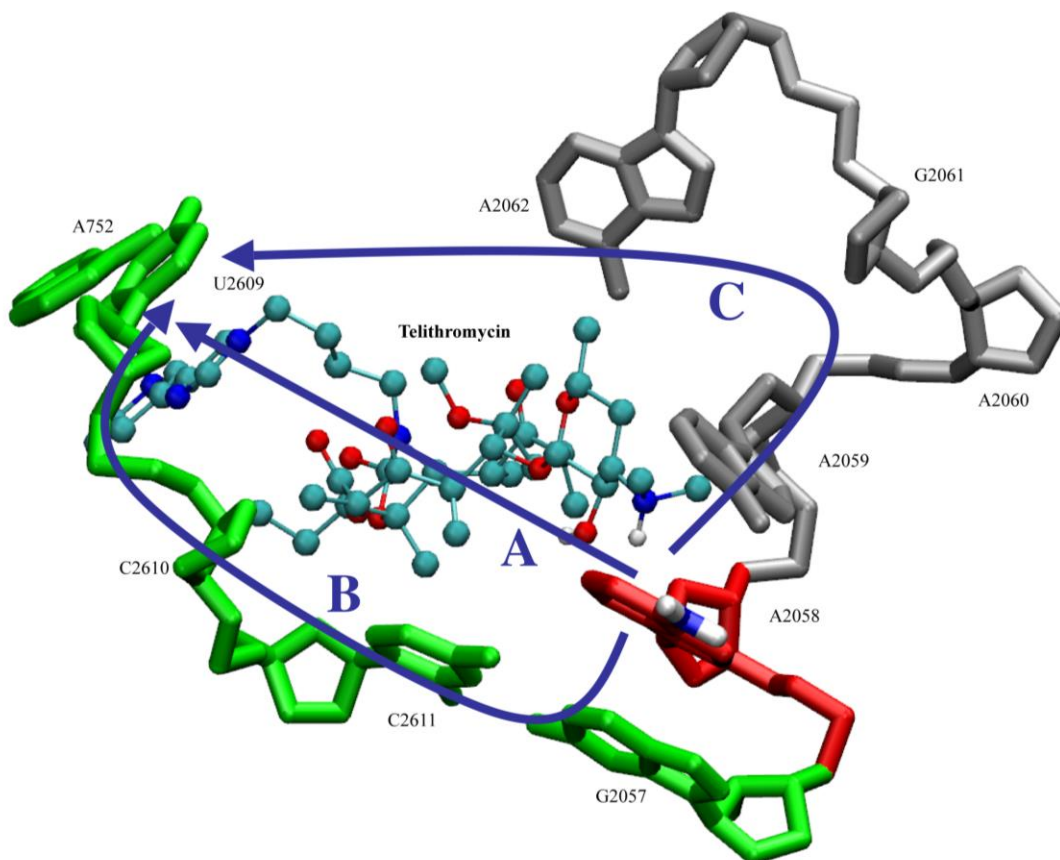


**Figure 3.13** The probability distributions of distances between telithromycin's 3'-protonated dimethylamine and G2505 P for WT, A2058G mutant and methyl modifications. The crystal structure distance is shown as a vertical black line, while the distributions are from the GCMC/MD simulations: WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta).

While the hydrogen bonding of 2058 with telithromycin and the ionic interaction involving G2505 are direct interactions between the ribosome and the antibiotic that may be directly perturbed by mutation or methylation, the significant changes in the A752-U2609 region were somewhat unexpected. Accordingly, further analysis was performed to understand how alteration of A2058 leads to long-range changes in telithromycin-ribosome interactions. Inspection of the telithromycin-ribosome structure suggested three potential “communication” pathways, as shown in **Figure 3.14**. The first, the telithromycin pathway, involves communication through telithromycin. The second pathway, or G2057 pathway, involves A2058 stacking with G2057 that may be communicated to A752-U2609 via the covalent connectivity from C2611 through U2609. And the third pathway, or A2059 pathway, proceeds via the covalent connectivity from A2058 through A2062 to telithromycin's ARM via VDW interactions. To determine possible contributions of these three possibilities, each pathway was subjected to further analysis.

Analysis of the telithromycin pathway involved inspection of the conformational sampling of the antibiotic in the simulations. Conformational sampling of the macrolactone ring was similar for the studied species, being consistent with that observed in the crystal structure (**Figure 3.15**), although a second conformation is observed in DMAD. Significantly, a much greater range of distances was sampled by the ARM, with those distances varying significantly from that in the crystal structure, consistent with the stacking and RMS fluctuation analysis discussed above. Given the similarities of the conformation of telithromycin in the different systems, the telithromycin pathway does

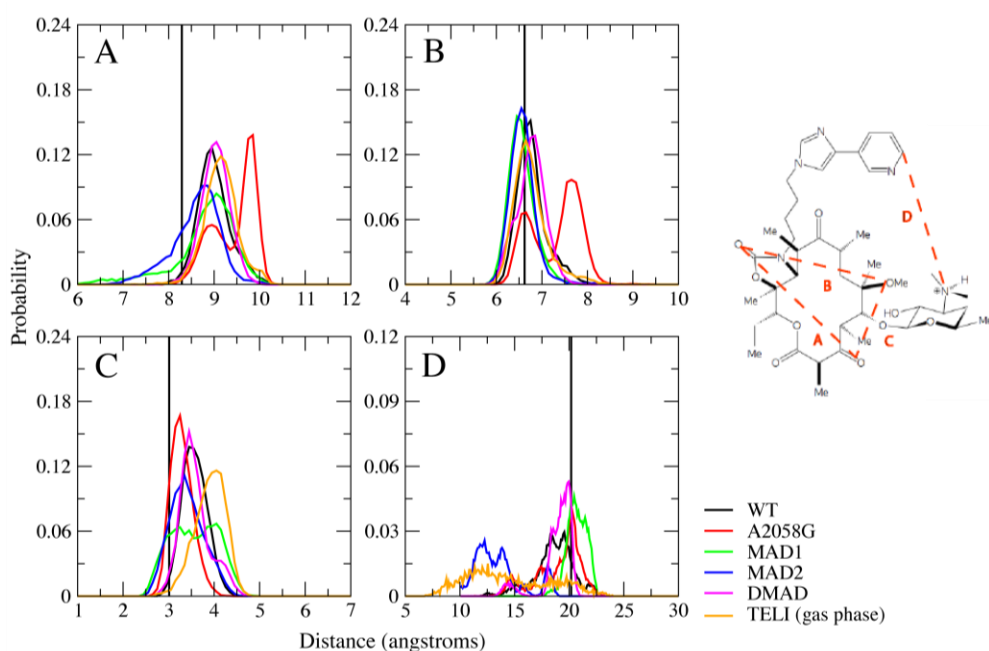
not appear to be the primary communication pathway between the A2058 and the ARM regions, with a possible exception of the DMAD species.



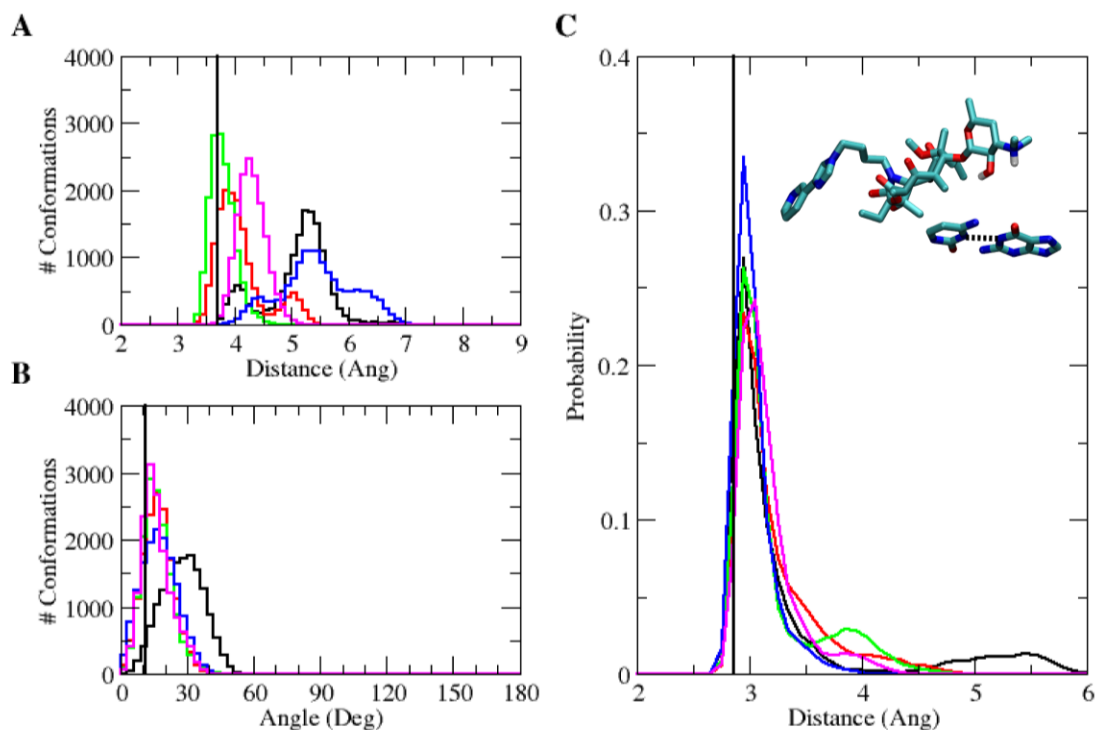
**Figure 3.14** Proposed communication pathways from A2058 to the A752-U2609 and ARM region: the telithromycin pathway (A), G2057 pathway (B), or A2059 pathway (C).

Central to the G2057 (i.e., second pathway) is the stacking of A2058 with G2057 and the WC interaction between G2057 and C2611 (**Figure 3.1**). This offers a logical, through-bond path by which perturbations of A2058 may impact RNA-telithromycin ARM interactions. Analysis of the stacking interactions involved investigation of the center of mass (COM) distance and the angle between least square planes through the

respective bases (**Figure 3.16A,B**). The probability distributions show that the WT samples a bimodal distance distribution with one mode corresponding to the stacked state observed in the crystal. In contrast, A2058G, MAD1 and DMAD sample single distributions in the vicinity of the crystal value, indicating improved stacking. A single distribution also occurs in MAD2 though a wide range of COM distances are sampled in that system. These results indicate increased stacking between A2058 and G2057 in A2058G, MAD1 and DMAD. Less favorable stacking with MAD2 is due to the N6 methyl directed towards the G2057 base, thereby perturbing stacking, while the second methyl in DMAD overcomes this leading to enhanced stacking. Thus, in all modified systems alteration of the interactions between the bases of A2058 and G2057 are present.



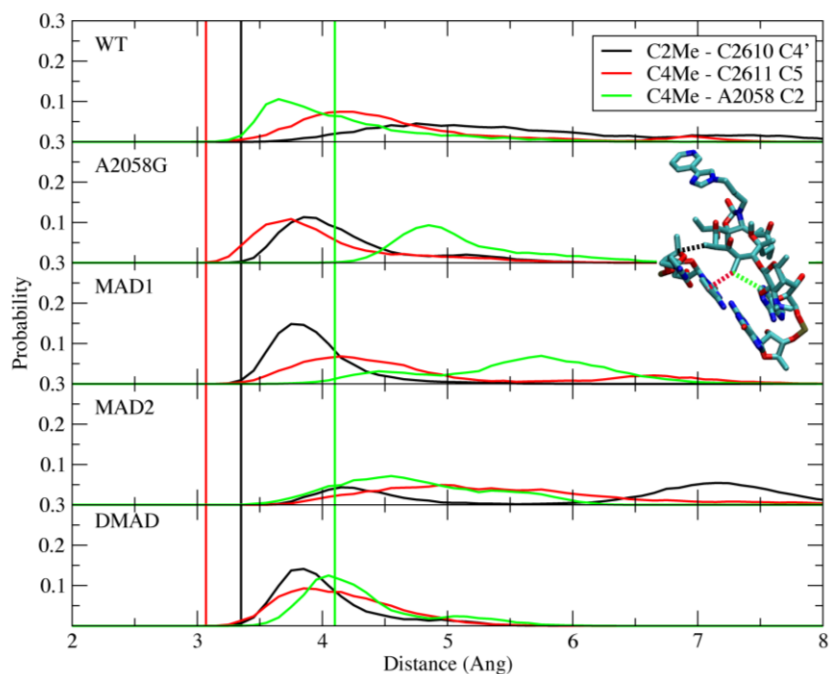
**Figure 3.15** The probability distributions for WT, A2058G mutant and methyl modifications of telithromycin intramolecular distances (A) C3-O to C16-O, (B) C16-O to C6-O, (C) C3-O to C6-O, and (D) C3'-N and C26 as shown in the inset figure. The crystal structure values are shown as vertical black lines, while distributions are from the GCMC/MD simulations: WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta). The probability distributions from telithromycin gas phase simulations without the ribosome are shown in orange.



**Figure 3.16** The number of conformations for WT, A2058G mutant and methyl modifications of (A) COM distances between the base atoms of 2057 and 2058 and (B) angles between planes comprised by 2057 and 2058 base atoms. The probability distributions for (C) WC G2057 N1 to C2611 N3 distances. The crystal structure values are shown as vertical black lines, while distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta).

Perturbation of A2058-G2057 interactions would be communicated along the G2057 pathway via WC interactions of that base with C2611. WC interactions between GC are high in all the systems indicating maintenance of base pairing (**Figure 3.16C**). This is consistent with experimental studies indicating that the orientation of G2057 and its base pairing with C2611 are important for maintaining the orientation of 2058, with the overall conformation of RNA being perturbed upon mutation of these bases.<sup>45,46</sup> This allows the mutation/modifications to impact the orientation of C2611, C2610 and U2609.

C2610 and C2611 make VDW contacts with two of telithromycin's methyl groups, C2 and C4 (**Figure 3.1**), which are 3.3 and 3.0 Å from the nearest non-hydrogen atoms on C2610 and C2611, respectively. In addition, C4 is also close to A2058, being 4.2 Å away in the crystal structure. Notably, in earlier generation macrolides the disruption of this interaction was thought to contribute to the loss of activity in A2058G mutants.<sup>16</sup> To investigate the impact of the A2058 mutation and methylations on these interactions, their probability distributions were calculated (**Figure 3.17**). In all the systems there is an outward shift in the two shortest interactions by ~1 Å suggesting that the 100 K crystal structure distances are slightly too short, while in the WT the C4 to A2058 C2 interaction became shorter. In the A2058G mutant, the two short interactions are well maintained, consistent with the activity of telithromycin against the mutant. The shift of the C4 to A2058G C2 distance is expected given the presence of an amine group at the C2 position in G. In the methyl modifications, the C2 is not substituted and the shift to larger distances is indicative that the methylations weaken the interactions and that the conformation of 2611 relative to telithromycin is altered in these systems. The less pronounced broadening and shift of the distance distribution in DMAD is presumably a result of more restrained conformational sampling as suggested by the lower RMSF for both telithromycin and C2611 (**Figures 3.9, 3.10**). Overall, the results indicate that perturbation of the G2057 communication pathway impacts interactions with the methyl groups to the largest extent with the monomethylations.



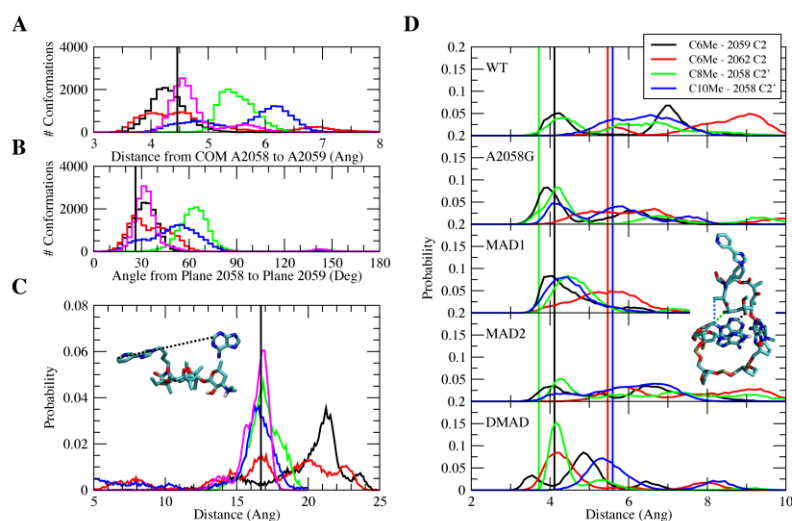
**Figure 3.17** The probability distributions for WT, A2058G mutant and methyl modifications of the telithromycin C2-methyl to C2610 C4' (black), C4-methyl to C2611 C5 (red), and C4-methyl to A2058 C2 (green). Corresponding crystal structure values are shown as vertical lines using the same color scheme.

The third and final communication network considered was the A2059 pathway (**Figure 3.14**). This involves the covalent connectivity of A2058 to A2062, which is facilitated by stacking of A2058 and A2059 as well as a short interaction between A2059 C2 and A2062 N6 of 4 Å in the crystal structure. As performed above, stacking of 2058 to 2059 was analyzed based on the COM distance and angle distributions (**Figure 3.18A,B**). The stacking is well maintained in the WT and A2058G mutant, consistent with the activity of telithromycin in those species. The stacking is also well maintained in the double methylation while significantly perturbed in both the monomethylations. Concerning the A2059 C2 to A2062 N6 interaction, in the DMAD system this interaction is well maintained (**Figure 3.19**). Sampling of short distances also occurs in the other

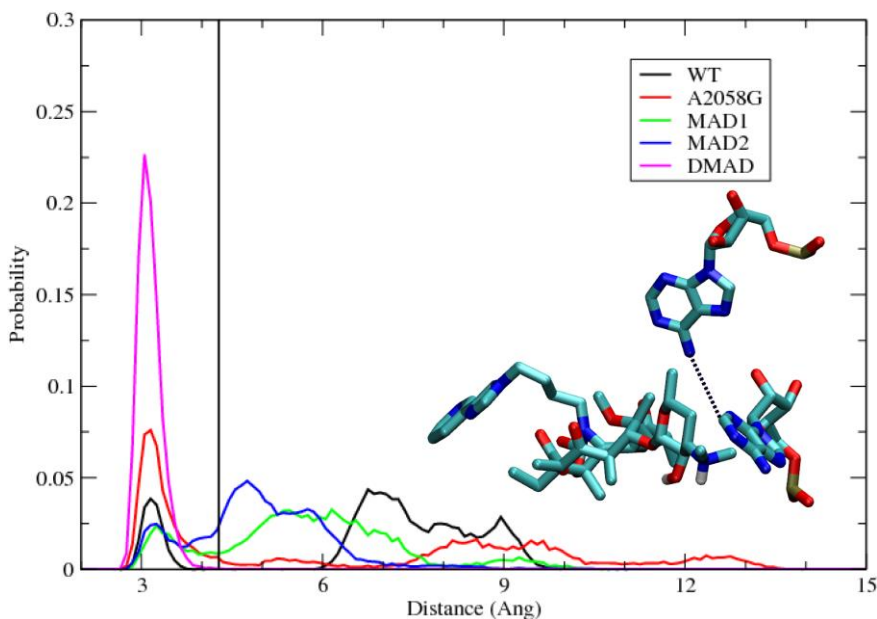
systems, but sampling of longer distances is present. Sampling of the longest distances occurs with WT and the A2058G mutant, which may contribute to the larger fluctuations of the ARM in those systems. This differential behavior could impact interaction of the nucleotides in the A2059 pathway with telithromycin. Nucleotides 2058, 2059, and 2062 come into VDW contact with the macrolactone of telithromycin. Shown in **Figure 3.18D** are probability distributions of selected distances between these bases and methyls extending from the macrocycle. All systems tend to sample short distances, with DMAD having a tendency to sample shorter distances to a larger extent than the other systems. Similar trends were observed for interactions between telithromycin's C9 and C16 keto groups with A2059 and A2062 as well as with a ribosomal Lys residue whose Nz atom is within 4 Å of one of the keto groups (**Figure 3.20**).

The final step in the A2059 communication pathway from 2062 to the A752-U2609 region involves indirect contacts. The region bordered by 2062, telithromycin, and 752/2609 is empty in the crystal structure and, thus, may be assumed to be filled with water, which likely contributes to the high flexibility of telithromycin's ARM. Such flexibility may allow for VDW contacts between 2062 and the ARM. Shown in **Figure 3.18C** are the distance distributions between C26 of telithromycin's pyridine and A2062 C2. The WT samples the longest distances indicating that the imidazole-pyridine moiety does not come close to A2062 and thus communication via this interaction does not occur. The A2058G mutant samples the widest range of distances, including some short distances (~5 Å) that may perturb the heterocyclic ARM, potentially leading to the WC base pairing of A752-U2609 being altered. The methylated species mostly sample C26 and A2062 C2 distances similar to those observed in the crystal structure. This limited

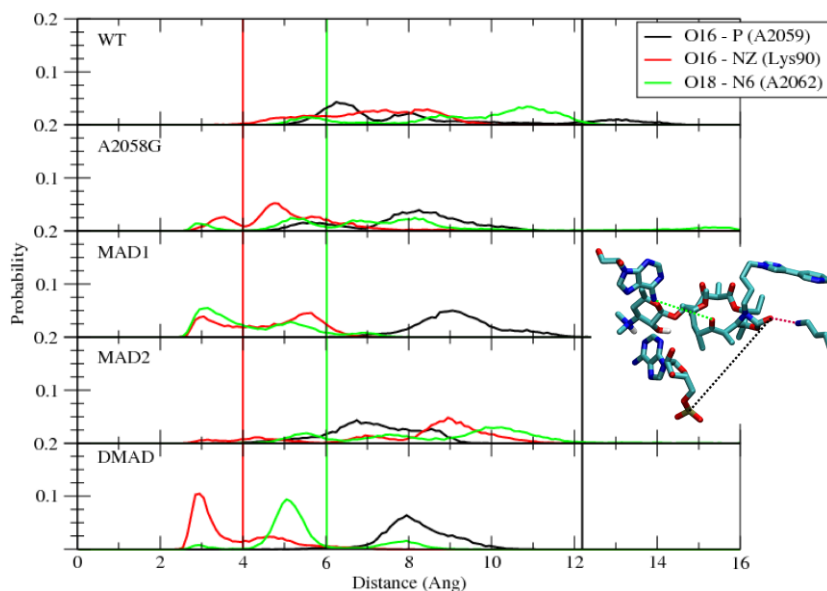
sampling is consistent with the RMS fluctuations of the ARM in these species. The exception is MAD2 that also samples shorter distances like the A2058G mutant and suggests that the MAD2 2058 modification may be propagated to A752-U2609 by a slightly different mechanism than the other methylated modifications. However, the lack of contact between A2062 and telithromycin's ARM heterocycle in the modifications indicates that the A2059 pathway is not as predominant as the G2057 pathway in the communication of 2058 mutation/modifications to A752-U2609.



**Figure 3.18** The number of conformations for WT, A2058G mutant and methyl modifications of (A) COM distances between base atoms of 2058 and 2059 and (B) angles between planes comprised by 2058 and 2059 base atoms. The probability distributions for all systems studied of the (C) telithromycin C26 to A2062 C2 distances and (D) telithromycin C6-methyl to 2059 C2 (black), C6-methyl to 2062 C2 (red), C8-methyl to 2058 C2' (green), and C10-methyl to 2058 C2' (blue). The crystal structure values are shown as vertical black lines in (A – C) and in (D) using the same color scheme, while distributions from the GCMC/MD simulations are shown as WT (black), A2058G (red), MAD1 (green), MAD2 (blue), and DMAD (magenta) in (A – C).



**Figure 3.19** The probability distributions of distances for A2059 C2 to A2062 N6. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. The crystal structure distance is shown as a black vertical line.



**Figure 3.20** The probability distributions for the distances between telithromycin O16 to A2059 P (black), O16 to Lys90 NZ (red), and O18 to A2062 N6 (green). Corresponding crystal structure distances are illustrated as vertical lines using the same color scheme. The inset figure illustrates the distances and orientations of the atoms from the crystal structure. O16 refers to C16-ketone, O18 refers to C9-ketone.

### 3.4 Conclusion

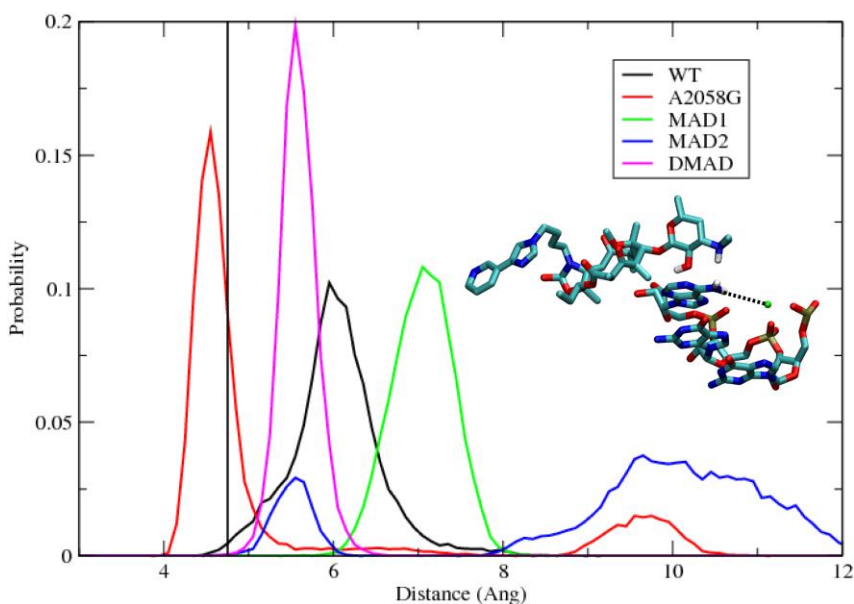
Altogether, our findings indicate that hydrogen bonding between 2'-OH and A2058 is important for telithromycin activity. Telithromycin maintains activity in the A2058G mutant via hydrogen bonds between the 2'OH and the WC hydrogen bonding groups of guanine, while decreased sampling of short 2'-OH to A2058 distances in the monomethylated species contributes to their lowered activity. In addition, the mutation/methylations are predicted to alter base stacking interactions with 2057 and to a lesser extent with 2059. These perturbed stacking interactions are communicated to more remote regions of the ribosome that comprise the telithromycin binding pocket thereby contributing to changes in telithromycin's activity in the methylated species. These changes occur through three potential pathways identified in this study: the telithromycin pathway, the G2057 pathway and the A2059 pathway. Analysis indicates the G2057 pathway leads to the largest conformational changes, including alterations of interactions of nucleotides 2058, 2610 and 2611 with methyl groups on telithromycin and, importantly, perturbation of A752-U2609 base pairing and interactions of those bases with the ARM of telithromycin.

In the context of telithromycin activity in wild type and A2058-methylated ribosomes, studies suggest that high levels of A2058-dimethylated ribosomes are required to confer telithromycin resistance.<sup>25</sup> Minimum inhibitory concentration (MIC) values for telithromycin are only marginally increased (~ 4-fold) in monomethylated ribosomes as opposed to 256-fold for erythromycin.<sup>23</sup> The ability of telithromycin to maintain activity against monomethylated ribosomes has been proposed to result from

imidazole-pyridine to A752-U2609 stacking interactions that mitigate the effects of disrupted 2'-OH – A2058 hydrogen bonds<sup>15,26,27,39</sup> and our findings coincide with this assessment. While hydrogen bonding does occur between telithromycin's imidazole and 2609 in MAD1, the ARM and 752-2609 maintain distances and plane angles that are indicative of stacking in a large number of conformations and 752-2609 WC distances are sampled with a higher probability than non base-paired or non-stacked distances.

Recently, Melman and Mankin<sup>47</sup> suggested that disruption of the 2'-OH – A2058 N1 hydrogen bond was not the major reason that telithromycin activity is reduced in A2058-dimethylated ribosomes. They found that removal of the 2'-OH from telithromycin did not increase MIC values to the extent that Erm(A) expression did. In other words, disruption of the hydrogen bond was not the predominant explanation for loss of telithromycin activity. They propose that the major explanation for reduced telithromycin activity is more likely that the overall structure within the binding site is perturbed upon the introduction of methyl groups onto A2058 as a result of nonbonded interactions between the methyl groups and nearby crystallographic waters. The exocyclic A2058 N6 is within 4 Å of water molecules that are coordinated to Mg<sup>+2</sup> that chelates the phosphate groups on G2056 and G2057. Accordingly, the authors suggest that interactions between the N6-methyl groups and water molecules lead to structural changes in the dimethylated ribosomes that reduce telithromycin activity. The results presented here further suggest that the conformation of RNA in the macrolide binding pocket is perturbed in both the 2058 mutant and modified ribosomes. However, the present results indicate that this is due to altered base stacking interactions with 2057 and 2059 that are propagated to other bases in the G2057 and A2059 pathways (**Figure 3.14**),

respectively. To test the hypothesis of Melman and Mankin,<sup>48</sup> the distance between the exocyclic N6 of A2058 and the  $Mg^{+2}$  ion were compared in the WT and mutant/modified systems (O6 in A2058G mutant). In all the systems except the mutant, the interaction is shifted by  $>1 \text{ \AA}$  indicating that the crystal structure distance is slightly too short (**Figure 3.21**). The shift to shorter distances for the A2058G mutant is expected given that the exocyclic N6 amine of A is replaced by a keto group in G. The distribution of distances show a high degree of overlap between WT and DMAD indicating that the relative position of the magnesium ion does not change upon introduction of two methyl groups onto A2058. Moreover, the distribution for DMAD is sharper suggesting that dimethylation restricts the conformational sampling of the 2058 N6 to  $Mg^{2+}$  distance.



**Figure 3.21** The probability distribution of distances between 2058 N6/O6 and a nearby  $Mg^{2+}$  ion that chelates the phosphate groups of G2056 and G2057 and two crystallographic waters. Note that  $Mg^{2+}$  and crystallographic waters were included in the simulations. Waters are excluded in the image for clarity. WT is shown in black, A2058G in red, MAD1 in green, MAD2 in blue, and DMAD in magenta. The crystal structure distance is shown as a vertical black line.

A particularly interesting result is the flexibility of the ARM in WT and the A2058G mutant versus that occurring in the methylated species. The fluctuations of the ARM are significantly higher than the rest of telithromycin in the WT crystal structure, corresponding to an average RMS fluctuation of 1.55 Å at 100 K over the non-hydrogen atoms in the heterocycles in the ARM. Thus, while stacking of the ARM with A752-U2609 is occurring to some extent, this is clearly a dynamic region of the system. The flexibility of the ARM is consistent with kinetic studies of telithromycin binding to the E. coli ribosome, which show that telithromycin's shift from a low to a high-affinity state results from reorganization of the ARM.<sup>49</sup> Given that this flexibility is large in both the WT and A2058G mutant and significantly lower in the methylated species, it suggests a scenario where entropic contributions associated with the flexibility of the ARM, and possibly the surrounding environment, makes a favorable contribution to binding. Upon methylation the mobility of the ARM is decreased, thereby contributing to a decrease in binding affinity. However, this is due to increased hydrogen bonding between the ARM and the A752 and U2609 nucleotides, interactions that could be exploited to improve the binding affinity. Essentially, a favorable entropic contribution to binding is being switched to a potentially favorable enthalpic contribution. Indeed, this appears to be the case with the recently published analog, solithromycin, in which the ARM imidazole-pyridine moiety was replaced with a triazolyl-aminophenyl group allowing for additional hydrogen bonding with 752 as well as nearby 748.<sup>48</sup> Namely, the hydrogen bonds lead to a favorable enthalpic contribution while decreasing the favorable entropic contribution which is consistent with the experimentally observed decrease in the RMS fluctuations from 1.5 to 1.0 Å for the ARM non-hydrogen atoms upon going from telithromycin to

solithromycin. While solithromycin shows improved activity compared to telithromycin against Erm-based modifications, which may be attributed to this hydrogen bonding, the presence of a C2-fluoro group not present on telithromycin may also contribute to improved binding. The C2-fluoro group appears to form a favorable (2.7 Å), hydrophobic interaction with C2611 thereby complicating interpretation of the contribution of ARM hydrogen bonding to affinity.

The results presented herein may be used to suggest modifications to telithromycin, or future ketolides, that could improve its binding to Erm-methylated ribosomes. These include modifying telithromycin's ARM to 1) engage in hydrogen bonding interactions with 752, 2609 and adjacent nucleotides, leading to an enthalpic contribution to binding or, on the other hand, 2) decrease the potential for such hydrogen bonding and/or increase the conformational flexibility in the ARM, leading to an entropy gain. Another region for potential improvement of binding against the methylated species involve modifying the macrolactone of telithromycin. The present calculations indicate the telithromycin methyl-base interactions to be longer in the modified ribosome. Accordingly, modifications to telithromycin that add steric bulk to the methyl groups nearby 2058, 2057, and 2611 may enhance VDW interactions with the ribosome that may also gain from an increased hydrophobic contribution to binding. Such modifications will be guided by previous work showing that ketolides bearing C2 groups larger than F are devoid of activity<sup>50</sup> as well as the results with solithromycin<sup>48</sup> showing that C2-fluorination may increase ketolide activity against Erm-based modifications.

### 3.5 References

1. Wright, G. D. Molecular mechanisms of antibiotic resistance. *Chem Commun (Camb)* **2011**, 47, 4055-61.
2. Poehlsgaard, J.; Douthwaite, S. The bacterial ribosome as a target for antibiotics. *Nat Rev Microbiol* **2005**, 3, 870-81.
3. Tenson, T.; Mankin, A. Antibiotics and the ribosome. *Mol Microbiol* **2006**, 59, 1664-77.
4. Brandi, L.; Marzi, S.; Fabbretti, A.; Fleishcer, C.; Hill, W.; Gualerzi, C. O.; Lodmell, J. S. The translation initiation functions of IF2: targets for thiostrepton inhibition. *J Mol Biol* **2004**, 335, 881-894.
5. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000**, 407, 340-8.
6. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **1996**, 274, 1367-71.
7. Ganoza, M. C.; Kiel, M. C. A ribosomal ATPase is a target for hygromycin B inhibition on *Escherichia coli* ribosomes. *Antimicrob Agents Chemother* **2001**, 45, 2813-9.
8. Hansen, J. L.; Ippolito, J. A.; Ban, N.; Nissen, P.; Moore, P. B.; Steitz, T. A. The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol Cell* **2002**, 10, 117-28.

9. Katz, L.; Ashley, G. W. Translation and protein synthesis: macrolides. *Chem Rev* **2005**, 105, 499-528.
10. Rodnina, M. V.; Stark, H.; Savelsbergh, A.; Wieden, H. J.; Mohr, D.; Matassova, N. B.; Peske, F.; Daviter, T.; Gualerzi, C. O.; Wintermeyer, W. GTPases mechanisms and functions of translation factors on the ribosome. *Biol Chem* **2000**, 381, 377-87.
11. Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001**, 413, 814-21.
12. Kirillov, S.; Porse, B. T.; Vester, B.; Woolley, P.; Garrett, R. A. Movement of the 3'-end of tRNA through the peptidyl transferase centre and its inhibition by antibiotics. *FEBS Lett* **1997**, 406, 223-33.
13. Tenson, T.; Lovmar, M.; Ehrenberg, M. The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *J Mol Biol* **2003**, 330, 1005-14.
14. Douthwaite, S. Structure-activity relationships of ketolides vs. macrolides. *Clin Microbiol Infect* **2001**, 7 Suppl 3, 11-7.
15. Tu, D.; Blaha, G.; Moore, P. B.; Steitz, T. A. Structures of MLSBK antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance. *Cell* **2005**, 121, 257-70.
16. Vester, B.; Douthwaite, S. Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrob Agents Chemother* **2001**, 45, 1-12.
17. Weisblum, B. Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrob Agents Chemother* **1995**, 39, 797-805.

18. Weisblum, B. Erythromycin resistance by ribosome modification. *Antimicrob Agents Chemother* **1995**, 39, 577-85.
19. Denoya, C.; Dubnau, D. Mono- and dimethylating activities and kinetic studies of the ermC 23 S rRNA methyltransferase. *J Biol Chem* **1989**, 264, 2615-24.
20. Nielsen, A. K.; Douthwaite, S.; Vester, B. Negative in vitro selection identifies the rRNA recognition motif for ErmE methyltransferase. *Rna* **1999**, 5, 1034-41.
21. Zhong, P.; Pratt, S. D.; Edalji, R. P.; Walter, K. A.; Holzman, T. F.; Shivakumar, A. G.; Katz, L. Substrate requirements for ErmC' methyltransferase activity. *J Bacteriol* **1995**, 177, 4327-32.
22. Champney, W. S.; Tober, C. L. Structure-activity relationships for six ketolide antibiotics. *Curr Microbiol* **2001**, 42, 203-10.
23. Liu, M.; Douthwaite, S. Activity of the ketolide telithromycin is refractory to Erm monomethylation of bacterial rRNA. *Antimicrob Agents Chemother* **2002**, 46, 1629-33.
24. Pfister, P.; Jenni, S.; Poehlsgaard, J.; Thomas, A.; Douthwaite, S.; Ban, N.; Bottger, E. C. The structural basis of macrolide-ribosome binding assessed using mutagenesis of 23S rRNA positions 2058 and 2059. *J Mol Biol* **2004**, 342, 1569-81.
25. Douthwaite, S.; Jalava, J.; Jakobsen, L. Ketolide resistance in *Streptococcus pyogenes* correlates with the degree of rRNA dimethylation by Erm. *Mol Microbiol* **2005**, 58, 613-22.
26. Berisio, R.; Harms, J.; Schluenzen, F.; Zarivach, R.; Hansen, H. A.; Fucini, P.; Yonath, A. Structural insight into the antibiotic action of telithromycin against resistant mutants. *J Bacteriol* **2003**, 185, 4276-9.

27. Hansen, L. H.; Mauvais, P.; Douthwaite, S. The macrolide-ketolide antibiotic binding site is formed by structures in domains II and V of 23S ribosomal RNA. *Mol Microbiol* **1999**, 31, 623-31.
28. Bonnefoy, A.; Girard, A. M.; Agouridas, C.; Chantot, J. F. Ketolides lack inducibility properties of MLS(B) resistance phenotype. *J Antimicrob Chemother* **1997**, 40, 85-90.
29. Champney, W. S.; Chittum, H. S.; Tober, C. L. A 50S ribosomal subunit precursor particle is a substrate for the ErmC methyltransferase in *Staphylococcus aureus* cells. *Curr Microbiol* **2003**, 46, 453-60.
30. Dubnau, D. Translational attenuation: the regulation of bacterial resistance to the macrolide-lincosamide-streptogramin B antibiotics. *CRC Crit Rev Biochem* **1984**, 16, 103-32.
31. Lai, C. J.; Weisblum, B. Altered methylation of ribosomal RNA in an erythromycin-resistant strain of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **1971**, 68, 856-60.
32. Usary, J.; Champney, W. S. Erythromycin inhibition of 50S ribosomal subunit formation in *Escherichia coli* cells. *Mol Microbiol* **2001**, 40, 951-62.
33. Wolter, N.; Smith, A. M.; Farrell, D. J.; Northwood, J. B.; Douthwaite, S.; Klugman, K. P. Telithromycin resistance in *Streptococcus pneumoniae* is conferred by a deletion in the leader sequence of erm(B) that increases rRNA methylation. *Antimicrob Agents Chemother* **2008**, 52, 435-40.
34. Bryskier, A. Ketolides-telithromycin, an example of a new class of antibacterial agents. *Clin Microbiol Infect* **2000**, 6, 661-9.

35. Salyers, A. A.; Amabile-Cuevas, C. F. Why are antibiotic resistance genes so resistant to elimination? *Antimicrob Agents Chemother* **1997**, *41*, 2321-5.
36. Shoemaker, N. B.; Vlamakis, H.; Hayes, K.; Salyers, A. A. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl Environ Microbiol* **2001**, *67*, 561-8.
37. Dunkle, J. A.; Xiong, L.; Mankin, A. S.; Cate, J. H. Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc Natl Acad Sci U S A* **2010**, *107*, 17152-7.
38. Bulkley, D.; Innis, C. A.; Blaha, G.; Steitz, T. A. Revisiting the structures of several antibiotics bound to the bacterial ribosome. *Proc Natl Acad Sci U S A* **2010**, *107*, 17158-63.
39. Douthwaite, S.; Hansen, L. H.; Mauvais, P. Macrolide-ketolide inhibition of MLS-resistant ribosomes is improved by alternative drug interaction with domain II of 23S rRNA. *Mol Microbiol* **2000**, *36*, 183-93.
40. Deng, Y.; Roux, B. Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J Chem Phys* **2008**, *128*, 1-8.
41. Woo, H. J.; Dinner, A. R.; Roux, B. Grand canonical Monte Carlo simulations of water in protein environments. *J Chem Phys* **2004**, *121*, 6392-400.
42. Ge, X.; Roux, B. Calculation of the standard binding free energy of sparsomycin to the ribosomal peptidyl-transferase P-site using molecular dynamics simulations with restraining potentials. *J Mol Recognit* **2010**, *23*, 128-41.
43. Ge, X.; Roux, B. Absolute binding free energy calculations of sparsomycin analogs to the bacterial ribosome. *J Phys Chem B* **2010**, *114*, 9525-39.

44. Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr. Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *The journal of physical chemistry. B* **2011**, *115*, 580-96.
45. Vester, B.; Hansen, L. H.; Douthwaite, S. The conformation of 23S rRNA nucleotide A2058 determines its recognition by the ErmE methyltransferase. *Rna* **1995**, *1*, 501-9.
46. Villsen, I. D.; Vester, B.; Douthwaite, S. ErmE methyltransferase recognizes features of the primary and secondary structure in a motif within domain V of 23 S rRNA. *J Mol Biol* **1999**, *286*, 365-74.
47. Melman, A.; Mankin, A.
48. Llano-Sotelo, B.; Dunkle, J.; Klepacki, D.; Zhang, W.; Fernandes, P.; Cate, J. H.; Mankin, A. S. Binding and action of CEM-101, a new fluoroketolide antibiotic that inhibits protein synthesis. *Antimicrob Agents Chemother* **2010**, *54*, 4961-70.
49. Kostopoulou, O. N.; Petropoulos, A. D.; Dinos, G. P.; Choli-Papadopoulou, T.; Kalpaxis, D. L. Investigating the entire course of telithromycin binding to Escherichia coli ribosomes. *Nucleic Acids Res* **2012**, *40*, 5078-87.
50. Denis, A.; Bretin, F.; Fromentin, C.; Bonnet, A.; Piltan, G.; Bonnefoy, A.; Agouridas, C. Beta-keto-ester chemistry and ketolides. Synthesis and antibacterial activity of 2-halogeno, 2-methyl and 2,3 enol-ether ketolides. *Bioorg Med Chem Lett* **2000**, *10*, 2019-22.
51. Brooks, B. R.; Brooks III, C. L.; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.;

- Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J Comput Chem* **2009**, *30*, 1545-1614.
52. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Josephy-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **1998**, *102*, 3586-3616.
53. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **2004**, *25*, 1400-15.
54. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* **2004**, *126*, 698-699.
55. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **2011**, *32*, 1929-43.
56. Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **2000**, *21*, 86-104.

57. Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; Mackerell, A. D., Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* **2012**, *8*, 348-362.
58. MacKerell, A. D., Jr.; Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* **2000**, *21*, 105-120.
59. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* **2008**, *29*, 2543-64.
60. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* **2009**, *5*, 2353-2370.
61. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **2011**, *7*, 3162-3180.
62. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, *5*, 1315-1327.
63. Hatcher, E.; Guvench, O.; Mackerell Jr., A. D. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* **2009**, *113*, 12466-76.

64. Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* **2010**, 114, 12981-12994.
65. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **2010**, 31, 671-90.
66. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, 79, 926-936.
67. Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* **2012**, 52, 3144-54.
68. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* **2012**, 52, 3155-68.
69. Brooks III, C. L.; Brunger, A.; Karplus, M. Active site dynamics in protein molecules: a stochastic boundary molecular-dynamics approach. *Biopolymers* **1985**, 24, 843-65.
70. Brooks III, C. L.; Karplus, M. Deformably stochastic boundaries in molecular dynamics. *J Chem Phys* **1983**, 79, 6312-6325.

71. Brooks III, C. L.; Karplus, M. Solvent effects on protein motion and protein effects on solvent motion. Dynamics of the active site region of lysozyme. *J Mol Biol* **1989**, 208, 159-81.
72. Brunger, A.; Brooks III, C. L.; Karplus, M. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett* **1984**, 105, 495-500.
73. Deng, Y.; Roux, B. Hydration of amino acid side chains: nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J Phys Chem B* **2004**, 108, 16567-16576.
74. van Gunsteren, W. F.; Karplus, M. A Method for Constrained Energy Minimization of Macromolecules. *Journal of Computational Chemistry* **1980**, 1, 266-274.
75. Hu, J.; Ma, A.; Dinner, A. R. Monte Carlo simulations of biomolecules: The MC module in CHARMM. *J Comput Chem* **2006**, 27, 203-16.
76. Beglov, D.; Roux, B. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J Chem Phys* **1994**, 100, 9050-9063.
77. Im, W.; Berneche, S.; Roux, B. Generalized solvent boundary potential for computer simulations. *J Chem Phys* **2001**, 114, 2924-2937.
78. Lynch, G. C.; Pettitt, B. M. Semi-grand canonical molecular dynamics simulation of bovine pancreatic trypsin inhibitor. *Chem Phys* **2000**, 258, 405-413.
79. Langevin, P. Sur la theorie du mouvement brownien. *C R Acad Sci (Paris)* **1908**, 146, 530-533.

80. Lemons, D. S.; Gythiel, A. Paul Langevin's 1908 paper "On the theory of brownian motion" ["Sur la theorie du mouvement brownien," C.R. Acad. Sci. (Paris) 146, 530-533 (1908)]. *Am J Phys* **1997**, 65, 1079-1081.
81. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **1977**, 23.
82. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*. Clarendon Press: Oxford, 1987; p 385.
83. Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods of Long-Range Forces in Macromolecular Simulations. *J. Comp. Chem.* **1994**, 15, 667-683.
84. De Loof, H.; Nilsson, L.; Rigler, R. Molecular dynamics simulation of galanin in aqueous and nonaqueous solution. *J Am Chem Soc* **1992**, 114, 4028-4035.
85. MOE, M. O. E. Chemical Computing Group. 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A, 2R7, 2012 **2012**.
86. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **2012**, 32, 1929-43.
87. Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* **2000**, 21, 86-104.
88. MacKerell Jr., A. D.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2000**, 56, 257-65.

89. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03*, Gaussian, Inc.: Wallingford, CT, 2004.
90. Scott, A. P.; Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* **1996**, 100, 16502-16513.
91. Sanbonmatsu, K. Y.; Joseph, S.; Tung, C. S. Simulating movement of tRNA into the ribosome during decoding. *Proc Natl Acad Sci U S A* **2005**, 102, 15854-9.
92. Sanbonmatsu, K. Y.; Tung, C. S. High performance computing in biology: multimillion atom simulations of nanoscale systems. *J Struct Biol* **2007**, 157, 470-80.
93. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; H., T. A. Equation of state calculations by fast computing machines. *J Chem Phys* **1953**, 21, 1087-1092.

## **Chapter 4 Investigation of *in situ* Click Chemistry Derived Macrolide Antibiotic**

### **Analogs in the Ribosome Using Site Identification by Ligand Competitive**

#### **Saturation (SILCS)**

Investigations were undertaken to understand the binding and *in situ* click chemistry efficiency of macrolide antibiotic analogs in the 50S *E. coli* ribosome using the Site-Identification by Ligand Competitive Saturation (SILCS) approach. Using a Monte Carlo sampling algorithm, the ligands were docked into the SILCS FragMaps to obtain ligand grid free energy scores (LGFEs). In the present study separate models were developed for the 1) overall binding affinity of macrolides and 2) the ability of alkyne-derived moieties to undergo the click reaction. In the first model, LGFEs were correlated to experimental dissociation constants for the assembled macrolide analogs and the predictive power of the LGFEs were shown to be dependent on the chemical structure of the alkyl-R group (ARM) moieties, which impacts the overall interaction of the macrolactone and desosamine moieties of macrolides with the ribosome. In the second model, click enhancement for the alkyne-derived fragments was determined to be dependent on normalized LGFE values, molecular weight, and hydrophobicity as estimated by logP values. These two models will allow for the prediction of ARM modifications that maximize the interaction of the macrolactone and desosamine moieties with the ribosome and whose alkyne analogs have a high probability of accessing the peptidyl-transfer pocket of the ribosome and undergoing the click reaction. In particular, this approach should be used with SILCS maps generated based on modified ribosomes

that are known to be resistant to the macrolide class of antibiotics, including A2058-modified ribosomes.

#### 4.1 Introduction

There is a critical need for antibiotics that address antibiotic resistance. Developing novel antibiotics is severely challenged by the short lag time between the availability of a new antibiotic and its associated clinical resistance. Bacteria reproduce rapidly and readily share genetic information thereby allowing for the widespread transfer of resistance mechanisms between bacterial species. Hence, an expedited process for discovering highly potent antibiotics is required.

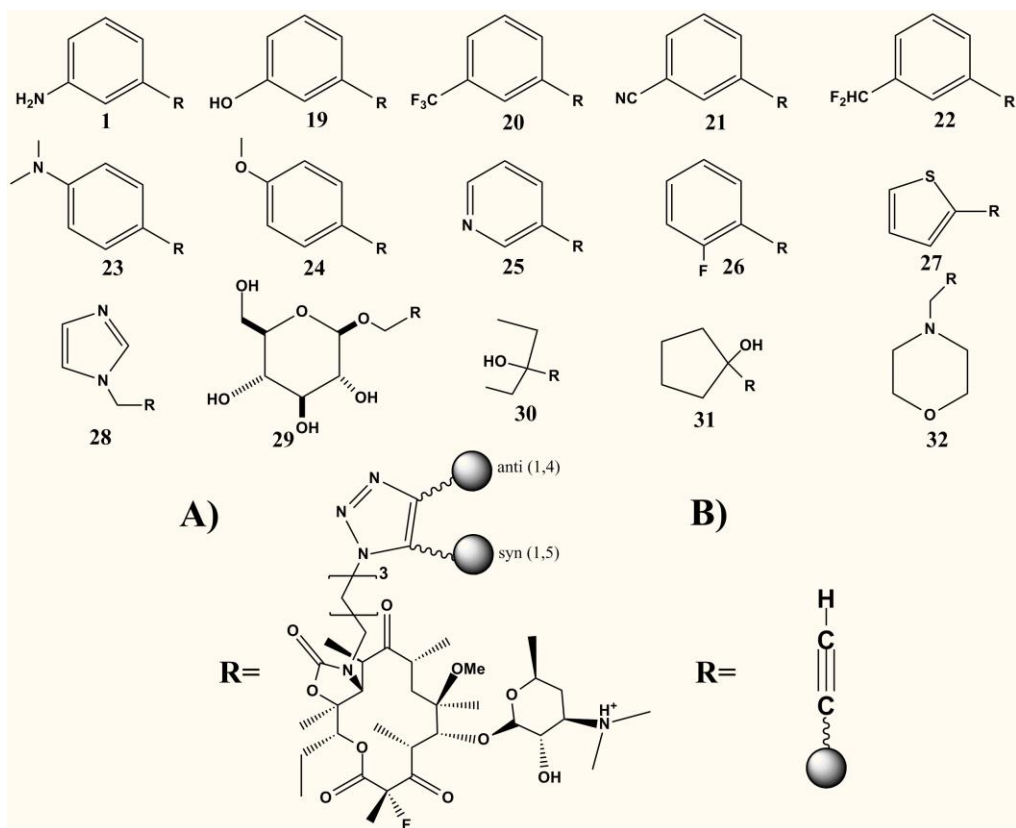
Fragment-based drug design (FBDD) is based on the idea that weakly-bound fragments that complex with the drug target may be linked to create a new drug. The use of fragments in drug design has been shown to yield a higher percentage of hits in high throughput screens (HTS) compared to traditional compound screens,<sup>1-3</sup> meaning fragments are better able to predict activity than standard chemical libraries. *In-situ* click chemistry is an extension of fragment-based drug discovery that uses the drug target to template the synthesis of pharmaceuticals using smaller fragments that bind to the target. It has been demonstrated in a range of systems including HIV protease,<sup>4</sup> protein-protein interactions,<sup>5</sup> and transcriptional regulation<sup>6,7</sup> as well as others.<sup>8-14</sup> One of the most popular click chemistry reactions is the azide-alkyne Huisgen cycloaddition reaction.<sup>15</sup> When performed under *in situ* conditions, it yields both the 1,4 and 1,5 triazole thus allowing for the drug target to template synthesis of the regioisomer with the highest

binding affinity.

While the power of *in situ* click chemistry as a fragment-based drug discovery tool is without doubt, it suffers from the limitation that only a few fragments can be screened at a given time. Therefore, it is not amenable to screening of hundreds of fragments, let alone the millions that comprise the fragment universe.<sup>1,16,17</sup> Computational FBDD methods, however, are not limited in this way and thus represent an ideal method for reducing the fragment universe to a tractable size for *in situ* click reactions. The Site-Identification by Ligand Competitive Saturation (SILCS) method<sup>18-20</sup> is a computational FBDD-based method that relies on molecular dynamics simulations of a drug target immersed in an aqueous solution of drug-like fragments and was presented in Chapter 1. It is applied here to the macrolide binding pocket of the *E. coli* 50S ribosomal subunit as a means to identify the types of chemical groups that enhance ligand binding affinities. Notably, this study employs two aspects of SILCS that will be briefly discussed in more detail. First, the application of SILCS to systems with occluded binding pockets<sup>21</sup> is made possible by the use of an in-house Grand Canonical Monte Carlo-like (GCMC) algorithm that drives the insertion of the Tier II fragments<sup>20</sup>. Specifically, the excess chemical potential ( $\mu_{\text{ex}}$ ) of the SILCS fragments and water are oscillated over many iterations of MC simulations, which allows for increased sampling of all fragments within the binding pocket leading to greater convergence. The excess chemical potential values are calculated from oscillating  $\mu_{\text{ex}}$  GCMC-MD simulations of the SILCS fragments in bulk water (no protein) as the  $\mu_{\text{ex}}$  needed to maintain a target concentration of 0.25M for each fragment and have been shown to be in good agreement with the experimental solvation free energies for each fragment.<sup>21</sup> Furthermore, the use of an oscillating  $\mu_{\text{ex}}$  facilitates

convergence for fragment insertions due to higher acceptance probabilities than traditional Monte Carlo (MC) simulations.<sup>21-23</sup> The second notable aspect of SILCS employed in this study is the use of FragMap Ligand Grid Free Energy (LGFE) scores in conjunction with MC sampling<sup>20</sup> to generate ensembles of ligand conformations as a means of ligand docking. The ligands sample translational, rotational, and torsional degrees of freedom within the “field” of the FragMaps using a Metropolis-based MC sampling algorithm, which provides an ensemble of conformations whose overlap with the FragMaps can be evaluated to give the LGFE score (discussed in Chapter 1).

In the present work, the ligands are derived from the Huisgen cycloaddition reaction that has been applied to the 50S ribosome by Andrade et al.<sup>24</sup> using a macrolactone-based azide and 15 alkyne-derived fragments, **Figure 4.1**. The assembled macrolide analogs and the alkyne-derived moieties are subjected to MC sampling in order to develop a model by which macrolide analogs and alkyne-derived fragments can be evaluated *a priori* to narrow down the number of fragments for the *in situ* reaction.



**Figure 4.1** 2D structures of fragment precursors for the *in situ* click chemistry. (A) The full ligand after the cycloaddition reaction including the regioisomers *anti* (1,4)- and *syn* (1,5)- analogs and (B) alkyne-derived fragments.

## 4.2 Computational Methods

### 4.2.1 SILCS Simulations

Calculations were performed using GROMACS,<sup>25</sup> version 5.0 (developmental) and the CHARMM additive force field including the CHARMM22 protein with the CMAP correction,<sup>26-28</sup> CHARMM36 nucleic acid,<sup>29-32</sup> CHARMM36 carbohydrate,<sup>33-38</sup> and CGenFF<sup>39</sup> (2b8) parameters and the TIP3P water model.<sup>40</sup> The initial coordinates were obtained from a snapshot at 5 ns from a previous GCMC-MD simulation.<sup>41</sup> The set-

up was modified slightly to be performed in GROMACS: non-hydrogen atoms in the outer region (from 34 to 40 Å) were restrained using a harmonic positional restraint of 10000 kJ/mol/nm<sup>2</sup> (24 kcal/mol/Å<sup>2</sup>) and non-hydrogen atoms in the buffer region were restrained using a harmonic positional restraint of 2500 kJ/mol/nm<sup>2</sup> (6 kcal/mol/Å<sup>2</sup>). Atoms in the inner region of 28 Å were not restrained. In all cases, ribosome positional restraints used the initial coordinates (above) as reference.

For SILCS simulations, telithromycin and all waters were removed. Water and the remaining fragments were maintained within a sphere using a spherical flat-bottomed potential of 4184 kJ/mol/nm<sup>2</sup> (10 kcal/mol/Å<sup>2</sup>), which is implemented in GROMACS version 5.0. Fragments were subjected to a flat-bottom potential extending 20 Å from the center of the system (set at the origin) with the restraint assigned to the following atoms of each fragment: propane C2, benzene CG, methylammonium CE, acetate C1, methanol CB, formamide C, imidazole CG, and acetaldehyde C. Waters were subjected to a flat-bottom potential extending 45 Å from the center with the positional restraint applied to the oxygen.

The GCMC-SILCS procedure has been described in detail elsewhere (Chapter 2),<sup>21</sup> such that the details of the general methodology will only briefly be addressed as they pertain to the simulations presented here. The macrolide binding pocket is primarily RNA and hence the charged nucleic acid backbone tends to favor the insertions of water, thereby limiting the insertions of hydrophobic fragments. Hence, an initial phase of 150,000 MC steps was performed in which all fragments except water were inserted, then an additional “flooding” phase of 4,000,000 MC steps was performed for water only using an insertion radius of 45 Å. This radius was selected because it extends beyond the

ribosome sphere and guarantees that all  $\text{Mg}^{2+}$  ions are solvated. The two phases were followed by a steepest decent (SD) minimization of 2000 steps using a force tolerance of 10 and maximum step size of 0.01 nm, in which positional restraints on the ribosome were removed, and 75 ps of molecular dynamics (MD) with a 1 fs time step to relax the system around the newly placed fragments.

GCMC-SILCS simulations were performed using 10 runs of 500 cycles each (a total of 500 ns cumulative MD simulation time) with each cycle consisting of MC, energy minimization, equilibration, and production. MC was performed for 150,000 steps using a radius of 20 Å. All fragments were given equal sampling time with a random assignment of the fragment insertion order. A cutoff of 15 Å beyond the sphere was used for calculation of the energies for each MC step. After each cycle of MC, relaxation of the system around the newly placed fragments and waters was accomplished with 500 steps of SD using a force tolerance of 10 and maximum step size of 0.01 nm. MD simulations were performed at 298 K using the LINCS<sup>42</sup> algorithm for covalent bonds involving hydrogens and the leapfrog integrator (GROMACS integrator “md”). Nonbonded interactions were treated using grid neighbor searching employing a group cutoff-scheme with a cutoff of 10 Å. Neighbor lists were updated every 10 steps. A switching function was applied to the van der Waals energy terms from 5 to 8 Å, while the electrostatics were treated using Reaction-Field-Zero with a cutoff of 8 Å. Equilibration was accomplished using 50 ps of MD employing a 1 fs time step. Generation of the fragment maps was based on 100 ps production MD using a 2 fs time step, with coordinates saved every 2 ps. Coordinates from the last snapshot of the production phase were used to initiate a new cycle of GCMC. FragMaps were generated

for the atom types: benzene carbons, propane carbons, acetaldehyde oxygen, formamide oxygen, formamide nitrogen, imidazole nonpolar nitrogen, imidazole polar nitrogen, methanol oxygen, methylammonium nitrogen, and acetate oxygens following the protocol outlined in previous SILCS papers<sup>18-21,43-45</sup>. Notably, the generation of hydrogen bond donor maps (i.e. imidazole, methylammonium, formamide, methanol) is now based on the non-hydrogen atom directly bonded to the hydrogen, which prevents the overestimation of the LGFE for ligands bearing multiple hydroxyl groups (e.g. ligand 29).

#### ***4.2.2 SILCS-MC Sampling***

Each ligand was generated using MOE<sup>46</sup>. The macrolactone ring coordinates were obtained from PDB ID 3OAT<sup>47</sup> with the aniline ring replaced by the R group of interest (**Figure 4.1**). The alkyne-derived fragment coordinates were generated by removing the macrolactone and desosamine rings and replacing the triazole ring with an alkyne. For each ligand, an ensemble of conformations was generated using an in-house Monte Carlo sampling algorithm described previously.<sup>20</sup> Sampling of translational, rotational, and torsional degrees of freedom was allowed using a flat-bottom potential of 10 Å. For the assembled macrolide analogs, the initial coordinates of the macrolactone and desosamine were based on the crystal structure coordinates. Sampling was accomplished using 100 runs each consisting of two phases. In the first phase 100,000 MC steps are performed with translations allowed to vary by 1 Å, rotations by 180°, and torsions by 180° in each MC step. The second phase involves a simulated-annealing (SLOWCOOL) phase of

50,000 MC steps with translations allowed to vary by 0.2 Å, rotations by 9°, and torsions by 9° in each MC step with conformations saved every 1,000 steps. Only conformations from the second phase are collected for analysis giving a total of 5000 conformations for each ligand. For the alkyne-derived moieties, 1000 runs were performed in which the alkyne was randomly oriented within 15 Å of the center of the binding pocket until an orientation is found with LGFE < 0. Sampling for each run was accomplished using 40,000 MC steps in which the move set was randomly generated in each step. The lowest energy conformation from each run was then subjected to a simulated annealing-based MC minimization of 10,000 steps. The final coordinates of each run were collected for analysis.

The calculation of LGFE values has been described previously<sup>20</sup> and is presented here to highlight details specific to this study. Unnormalized LGFE values were calculated by summing the GFE values for each atom in the congeners over each snapshot to calculate the Boltzmann average LGFE. Normalized LGFE values were calculated by first dividing the Boltzmann average LGFE by the total number of classified atoms as determined using a set of rules that classifies each atom type into a FragMap type. As opposed to previous SILCS studies, the rules used for classification of the ligand atom types into FragMap classes in this study did not include hydrogens. Then the values were multiplied by the total number of non-hydrogen atoms to yield the normalized LGFE. For the macrolide analogs, normalized LGFE values were also calculated for subsets of ligand atoms: macrolactone+desosamine and ARM. The ARM is defined as the alkyl-R group moiety extending from N11 (**Figure 4.1A**), while the macrolactone+desosamine atoms correspond to the macrolactone ring and D-desosamine

ring extending from C5. Predictive indices values were calculated according to Pearlman and Charifson<sup>48</sup> and rank from -1 (negative correlation) to +1 (positive correlation).

Multiple regression analysis for the alkyne-derived moieties was performed in Microsoft Excel using the normalized LGFE, molecular weight, and logP values. The logP values were calculated using MOE<sup>46</sup>. Analysis was performed to fit the natural logarithm of the sum % of click enhancement, as determined by the Andrade laboratory<sup>24</sup>.

### **4.3 Results and Discussion**

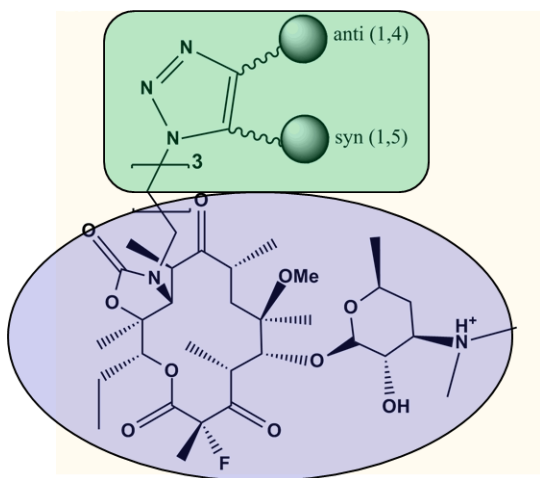
Dissociation constants were compared to ligand grid free energies (LGFES) obtained from the Site-Identification by Ligand Competitive Saturation (SILCS) approach.<sup>18-20</sup> SILCS harnesses the power of molecular dynamics simulations to map the functional group affinities of the 50S ribosome. This is achieved by performing a number of MD simulations of the 50S ribosome in the presence of explicit water and ~ 0.25 to 1 M fragments. The fragments include benzene, propane, formamide, acetaldehyde, acetate, methylammonium, and methanol, which is referred to as Tier II SILCS<sup>20</sup>. They represent a diverse set of small molecules representing hydrophobic and hydrogen bonding groups, including both charged and neutral species. In addition, imidazole is included to represent a heterocycle as well as both an additional neutral donor and acceptor. During the simulations, the fragments interact with all regions within the macrolide binding pocket of the 50S ribosome using a Grand Canonical-like Monte Carlo approach<sup>21</sup> optimized for occluded pockets like that in the 50S ribosome. 3D probability

maps (FragMaps) are generated that are normalized with respect to the fragments in solution alone and converted to free energies (called Grid Free Energies, GFEs). Ligand-based GFEs (LGFEs) are used to estimate the binding affinity for each product from the *in situ* click chemistry and are computed using an ensemble of ligand conformations, which has been shown to lead to significant improvements in the quantitative ranking of ligands.<sup>20</sup> Ensembles are generated for each congener using a Metropolis-based Monte Carlo sampling algorithm<sup>20</sup> that samples translational, rotational, and torsional degrees of freedom within the “field” of the FragMaps. LGFEs are calculated based on the overlap of the ensemble with the FragMaps and compared to experimental dissociation constants.

Dissociation constants for each of the *in situ* click products were converted to  $\Delta G$ s (kcal/mol) for direct comparison to the calculated LGFE values via regression analysis. In addition, predictive indices (PI)<sup>48</sup> were calculated to measure how well the LGFE values quantitatively rank the experimental data. PI values indicate how well a metric orders the relative binding of a congeneric series. The normalized LGFE values for each compound are shown in **Table 4.1** for the full ligands, macrolactone+desosamine, and ARM moieties. Their predictive index is calculated to be -0.22 indicating that the LGFE values are not an accurate predictor of binding affinity. In the present congeners, the macrolactone and desosamine sugar remain unchanged while the ARM moieties extending from N11 differ. These are derived from the alkyne precursors (**Figure 4.1**) in the *in situ* click reaction. Given the relatively small differences in the dissociation constants (i.e.  $K_d = 0.6$  to  $3.5$  nanomolar) despite the structural differences in these moieties, further analysis was undertaken to determine the individual contributions to the binding affinity, **Figure 4.2**.

**Table 4.1** Rank-ordering of triazoles **1**, **19-32** and azide fragment **2** by dissociation constants ( $K_d$ ) for 70S E. coli ribosomes in nM by fluorescence polarization along with experimental  $\Delta G$  values (kcal/mol), calculated normalized Ligand Grid Free Energies (LGFEs, kcal/mol) and ligand efficiencies from SILCS. LGFEs and ligand efficiencies are calculated for the total molecule, ARM, and macrolactone+desosamine (Macro+Des) parts, where the ARM is defined as the four carbon alkyl linker and heterocycle extending from N11. The predictive indices<sup>48</sup> are calculated for each type of LGFE and ligand efficiency.

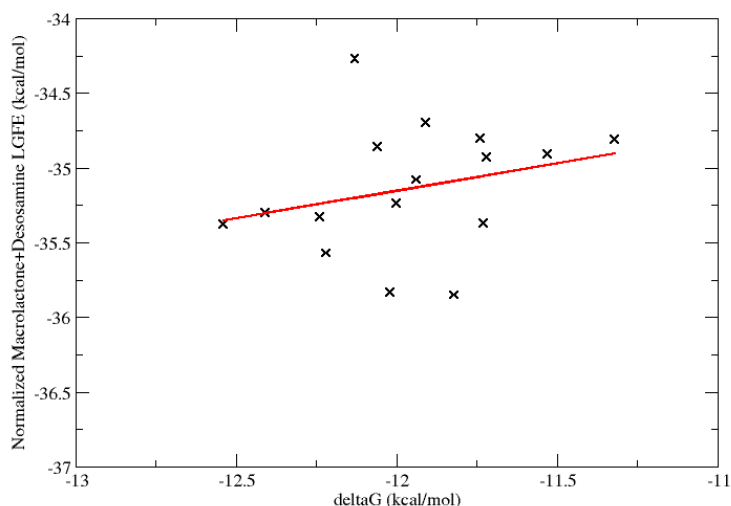
ID	$K_d$ (nM)	$\Delta G$ (kcal/mol)	Total LGFE (kcal/mol)	ARM LGFE (kcal/mol)	Macro+Des LGFE (kcal/mol)
<b>SOL (1)</b>	0.63 ± 0.1	-12.54	-49.67	-16.20	-35.38
<b>32</b>	0.79 ± 0.12	-12.41	-51.83	-17.09	-35.30
<b>19</b>	1.05 ± 0.14	-12.24	-48.72	-14.01	-35.33
<b>25</b>	1.09 ± 0.12	-12.22	-48.70	-14.87	-35.57
<b>23</b>	1.26 ± 0.15	-12.13	-52.16	-17.19	-34.27
<b>24</b>	1.41 ± 0.15	-12.06	-50.83	-15.50	-34.86
<b>28</b>	1.53 ± 0.14	-12.02	-47.80	-14.27	-35.83
<b>21</b>	1.57 ± 0.15	-12.00	-50.82	-16.23	-35.24
<b>22</b>	1.74 ± 0.21	-11.94	-52.61	-17.81	-35.08
<b>27</b>	1.84 ± 0.16	-11.91	-47.93	-13.64	-34.70
<b>Azide 2</b>	2.12 ± 0.39	-11.82	-38.85	-5.04	-35.85
<b>31</b>	2.43 ± 0.23	-11.74	-49.49	-15.15	-34.80
<b>30</b>	2.48 ± 0.25	-11.73	-48.96	-13.91	-35.37
<b>26</b>	2.52 ± 0.23	-11.72	-49.18	-15.39	-34.93
<b>20</b>	3.50 ± 0.51	-11.53	-52.86	-18.55	-34.91
<b>29</b>	4.96 ± 0.39	-11.32	-52.90	-18.10	-34.81
<b>PI</b>			<b>-0.22</b>	<b>-0.14</b>	<b>0.37</b>



**Figure 4.2** The individual contributions to the total LGFE attributed to the macrolactone+desosamine and ARM moieties.

Determining the individual contributions of the macrolactone+desosamine and ARM moieties exploits the atom-based scoring method using in computing the LGFE values. The LGFE values are a sum of the individual atomic grid free energy contributions of classified atoms in each ligand, as previously described.<sup>20</sup> Based on the atom-based assumption, the LGFEs were decomposed into two constituent parts: 1) the unperturbed macrocycle + desosamine sugar and 2) the ARM, which is defined as the four carbon alkyl linker extending from N11 + the alkyne-derived moiety, including the triazole (**Figure 4.1**). The normalized LGFE values calculated based on these parts are shown in **Table 4.1**. The macrolactone+desosamine LGFE values are somewhat predictive of the  $\Delta G$  values. Their predictive index is calculated to be 0.37 and a correlation plot of macrolactone+desosamine LGFE versus  $\Delta G$  values, **Figure 4.3**, shows an  $R^2$  value of 0.07. This indicates that the LGFE values calculated based on the

macrolactone+desosamine alone have some predictive capabilities of the experimental trends in dissociation constants as evidenced by positive predictive indices. Yet, the correlation of the  $\Delta G$  to the macrolactone+desosamine LGFEs is low, which most likely is due to the small differences between the  $\Delta G$  values. The magnitude of the LGFE values compared to the total LGFE values indicates that the macrolactone+desosamine have the largest contribution to binding. Presumably, these favorable interactions arise from hydrophobic interactions of the macrolactone ring with the walls of the exit tunnel, as well as hydrogen bonding and electrostatic interactions (viz. salt bridges) between the desosamine and bases 2058 and 2505, respectively.



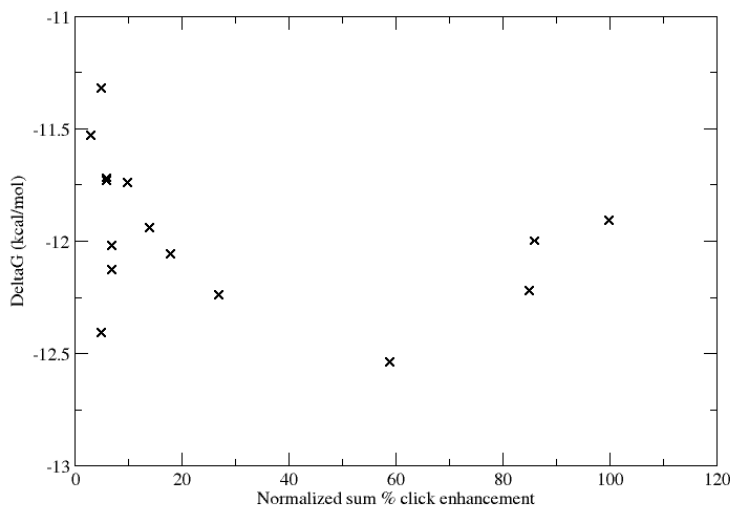
**Figure 4.3**  $\Delta G$ s (kcal/mol) versus the normalized macrolactone+desosamine LGFEs (kcal/mol). The red line indicates the linear regression of the data, which represents an  $R^2$  of 0.07. Equation for  $\Delta G$  versus Normalized Macrolactone+Desosamine LGFE:  $y = 0.37x - 30.75$ .

The normalized LGFE values for the ARM show a predictive index of -0.14, indicating that the LGFE values are negatively correlated with the  $\Delta G$  values. As mentioned, the small differences in the  $\Delta G$  values despite structural changes in the

alkyne-derived moieties points to the fact that the ARM is not the driving force for binding. Instead, the binding affinity of the congeners comes from the macrocycle+desosamine. This is further evidenced when the congeners are compared to other macrolides, such as clarithromycin and azithromycin that lack ARMs yet bind with affinities similar or better to those presented here ( $\Delta G_{\text{azithromycin}} = -14.17$  kcal/mol and  $\Delta G_{\text{clarithromycin}} = -11.95$  kcal/mol). Since the major contribution to binding is derived from the macrolactone+desosamine, whose normalized LGFE values show a positive predictive index, it is expected that the normalized total LGFEs would also show some positive predictive power. However, the opposite is true. This suggests that while the ARM is not directly impacting the binding affinity, it is somehow affecting the affinity of the molecule as a whole. Previous molecular dynamics studies of telithromycin bound within the 50S *E. coli* subunit<sup>41</sup> have demonstrated that the ARM region is dynamic. Similarly, it is likely that the sampling of the ARM in these studies is affecting the conformation of the macrocycle+desosamine and thus is indirectly impacting the overall binding affinity of the congener. These results lead us to suggest the following scenario. The binding of solithromycin and related macrolides are dominated by the macrolactone+desosamine moieties with the impact of the ARM moiety not being dominated by its own binding affinity, but by its impact on the affinity of the macrolactone+desosamine moieties. This is important as it allows for implementation of a rational drug design strategy in which variations in the ARM are selected to maximize interactions of the macrolactone+desosamine moieties with the ribosome. Of course, in such a scenario, any improvement in the affinity of the ARM for the ribosome would contribute to the overall binding affinity of the macrolide, and therefore should also be

pursued simultaneously.

The click chemistry methodology is based on the binding of the individual species to the target system, which leads to their alignment and subsequent reaction to generate a product. This templating process will thus yield compounds that are specific for the target system. It may also be anticipated that the tighter the binding of the individual species to the target system would lead to higher click enhancement as well as high binding affinities of the products. To test this assumption we compared the binding free energies (**Table 4.1**) with the overall enhancement rates, which were calculated as the sum of the normalized anti, syn and mixed enhancements as reported to us from the Andrade laboratory<sup>24</sup> and shown in **Table 4.2**. As may be seen from **Figure 4.4**, there is no evident correlation between the affinities of the products and the enhancement rates, which is not unexpected given the observation that the ARM moieties are not directly contributing to binding. However, it motivated analysis of the relationship of the LGFEs and other physical properties of the individual alkynes in the enhancement rates.



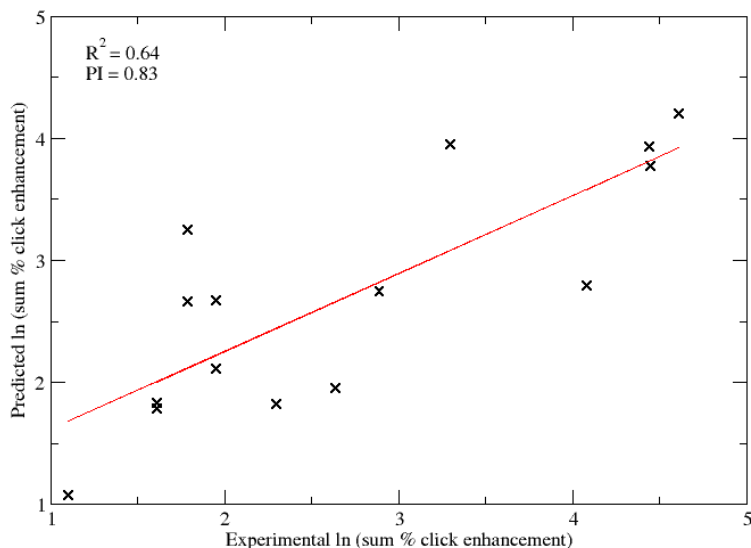
**Figure 4.4** The correlation of the  $\Delta G$  values for the full ligand to the experimental click enhancements for the alkyne-derived fragments.

**Table 4.2** Sum of mass count % increases, natural logarithm of the sum %, logP, molecular weight (MW) and predicted natural logarithm of the sum % for triazoles **1**, **19-32** and azide fragment **2** (with rank ordering preserved from Table 4.1). The predictive indices<sup>48</sup> are calculated for the predicted mass count % increase versus the experimental mass count % increase.

<b>ID</b>	<b>K<sub>d</sub> (nM)</b>	<b>Sum mass count % increase</b>	<b>ln (sum mass count % increase)</b>	<b>logP</b>	<b>Normalized LGFE (kcal/mol)</b>	<b>MW</b>	<b>Predicted ln (sum mass count % increase)</b>
<b>SOL (1)</b>	0.63 ± 0.1	59	4.08	1.75	-8.88	117.15	2.79
<b>32</b>	0.79 ± 0.12	5	1.61	0.04	-10.52	125.17	1.78
<b>19</b>	1.05 ± 0.14	27	3.30	2.11	-7.75	118.13	3.95
<b>25</b>	1.09 ± 0.12	85	4.44	1.15	-7.32	103.12	3.93
<b>23</b>	1.26 ± 0.15	7	1.95	2.30	-10.71	145.21	2.11
<b>24</b>	1.41 ± 0.15	18	2.89	2.34	-9.45	132.16	2.66
<b>28</b>	1.53 ± 0.14	7	1.95	0.34	-8.73	106.13	2.67
<b>21</b>	1.57 ± 0.15	86	4.45	2.08	-8.39	127.15	3.77
<b>22</b>	1.74 ± 0.21	14	2.64	3.01	-10.82	152.14	1.95
<b>27</b>	1.84 ± 0.16	100	4.61	1.91	-7.10	108.16	4.20
<b>31</b>	2.43 ± 0.23	10	2.30	1.51	-9.47	110.16	1.82
<b>30</b>	2.48 ± 0.25	6	1.79	2.00	-8.64	112.17	2.74
<b>26</b>	2.52 ± 0.23	6	1.79	2.54	-8.27	120.13	3.25
<b>20</b>	3.50 ± 0.51	3	1.10	3.36	-12.30	170.13	1.07
<b>29</b>	4.96 ± 0.39	5	1.61	-2.01	-14.88	218.21	1.83
<b>PI</b>							<b>0.83</b>

For a congener to have a high click enhancement (as indicated by a high percent increase in mass counts), the azide and alkyne precursors must bind with high affinity to the binding pocket on similar time scales and also have orientations that favor the Huisgen cycloaddition reaction. The azide is constant in the *in situ* click reaction and since it does bind with high affinity the mass count % increase is based on the alkyne precursor. **Table 4.2** shows the mass count % increase for each alkyne fragment from mass spectrometry, logP(octanol/water) squared, normalized LGFE for the alkyne fragment, and molecular weight (MW). Multiple regression analysis was undertaken against the natural log for each mass count % increase, which is analogous to a Boltzmann transformation. The predicted click enhancement obtained from regression against normalized LGFE, logP squared, and MW are plotted against the experimental % increase in **Figure 4.5**, yielding an  $R^2$  value of 0.64 and predictive index of 0.83, which indicates that a combination of these properties is contributing to the click enhancement rate.

As mentioned, the mass count % increase of a congener depends on whether the alkyne fragment favorably binds in the vicinity of the azide reactant. Notably, the normalized LGFE values for the alkyne fragments are a measure of the binding affinity for each moiety. However, if the binding is too favorable in one region, the alkyne-derived moiety will not effectively sample the entire binding pocket and therefore may not find an orientation suitable for the cycloaddition reaction. This is exemplified in the slope for the normalized LGFE contribution to the predicted mass count % increase being positive (+1.12) meaning that the more favorable the LGFE (more negative), the smaller the mass count % increase.



**Figure 4.5** The experimental mass count % increases versus the predicted mass count % increases. The red line indicates the linear regression of the data, which represents an  $R^2$  of 0.64. **Equation for predicted click enhancement versus experimental click enhancement from multiple regression analysis:**  $\ln(\text{sum } \%) = 1.12 \text{ LGFE} - 0.12 (\log P)^2 + 0.06 \text{ MW} + 6.30$ .

The ability of the alkyne-derived moiety to get into the binding pocket also affects the mass count % increase because the cycloaddition reaction will not proceed in the absence of the alkyne precursor. Bulkier fragments have a harder time entering the binding pocket, which is evidenced by the slope of the MW contribution to the predicted mass count % increase being positive (+0.06). Though this value is small, suggesting that the size of the alkyne-derived fragment is not the dominant factor in determining the predicted % increases. As a measure of hydrophobicity, logP is also indicative of whether an alkyne fragment will enter the pocket or not. The macrolide binding pocket is predominantly RNA, thus hydrophilic moieties have higher binding affinities because they bear hydrogen bonding functional groups. The slope for the  $(\log P)^2$  contribution to

the mass count % increase is -0.12, indicating that alkyne-derived fragments that are more hydrophilic (more negative logP) will increase the mass count % increase. However, as indicated by the parabolic term in the logP values, alkyne-derived fragments that are too hydrophobic or hydrophilic will lead to lower click enhancement, presumably because these fragments are unable to partition themselves from water to the hydrophobic macrolide binding pocket of the ribosome.

The motivation underlying the *in situ* click method is to identify novel high affinity macrolides that are active against resistance bacteria. In the present study separate models were developed for the 1) overall binding affinity of macrolides and 2) the ability of alkyne-derived moieties to undergo the click reaction. These two models will allow for the prediction of ARM modifications that maximize the interaction of the macrolactone and desosamine moieties with the ribosome and whose alkyne analogs have a high probability of accessing the peptidyl-transfer pocket of the ribosome and undergoing the click reaction. In particular, this approach should be used with SILCS maps generated based on modified ribosomes that are known to be resistant to the macrolide class of antibiotics, including A2058G mutants and methylated A2058 ribosomes.

#### 4.4 References

1. Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **2007**, 6, 211-9.
2. Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on

- the probability of finding leads for drug discovery. *J Chem Inf Model* **2001**, 41, 856-864.
- Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E. Library design for fragment based screening. *Curr Top Med Chem* **2005**, 5, 751-762.
  - Whiting, M.; Muldoon, J.; Lin, Y. C.; Silverman, S. M.; Lindstrom, W.; Olson, A. J.; Kolb, H. C.; Finn, M. G.; Sharpless, K. B.; Elder, J. H.; Fokin, V. V. Inhibitors of HIV-1 protease by using in situ click chemistry. *Angew Chem Int Ed* **2006**, 45, 1435-1439.
  - Namelikonda, N. K.; Manetsch, R. Sulfo-click reaction via in situ generated thioacids and its application in kinetic target-guided synthesis. *Chem Commun* **2012**, 48, 1526-1528.
  - Sharpless, K. B.; Manetsch, R. In situ click chemistry: a powerful means for lead discovery. *Expert Opinion on Drug Discovery* **2006**, 1, 525-538.
  - Willand, N.; Desroses, M.; Toto, P.; Dirié, B.; Lens, Z.; Villeret, V.; Rucktooa, P.; Locht, C.; Baulard, A.; Deprez, B. Exploring Drug Target Flexibility Using in Situ Click Chemistry: Application to a Mycobacterial Transcriptional Regulator. *ACS Chemical Biology* **2010**, 5, 1007-1013.
  - Grimster, N. P.; Stump, B.; Fotsing, J. R.; Weide, T.; Talley, T. T.; Yamauchi, J. G.; Nemezc, A.; Kim, C.; Ho, K. Y.; Sharpless, K. B.; Taylor, P.; Fokin, V. V. Generation of candidate ligands for nicotinic acetylcholine receptors via in situ click chemistry with a soluble acetylcholine binding protein template. *J Am Chem Soc* **2012**, 134, 6732-6740.
  - Hirose, T.; Sunazuka, T.; Sugawara, A.; Endo, A.; Iguchi, K.; Yamamoto, T.; Ui,

H.; Shiomi, K.; Watanabe, T.; Sharpless, K. B.; Omura, S. Chitinase inhibitors: extraction of the active framework from natural argifin and use of in situ click chemistry. *J Antibiotic* **2009**, 62, 277-282.

10. Krasinski, A.; Radić, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In Situ Selection of Lead Compounds by Click Chemistry: Target-Guided Optimization of Acetylcholinesterase Inhibitors. *J Am Chem Soc* **2005**, 127, 6686-6692.

11. Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radic, Z.; Carlier, P. R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. Click chemistry in situ: Acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. *Angew Chem Int Ed* **2002**, 41, 1053.

12. Manetsch, R.; Krasinski, A.; Radic, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In situ click chemistry: enzyme inhibitors made to their own specifications. *J Am Chem Soc* **2004**, 126, 12809-12818.

13. Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. H.; Kolb, H. C. In situ click chemistry: enzyme-generated inhibitors of carbonic anhydrase II. *Angew Chem Int Ed* **2004**, 44, 116-120.

14. Poulin-Kerstien, A. T.; Dervan, P. B. DNA-templated dimerization of hairpin polyamides. *J Am Chem Soc* **2003**, 125, 15811-15821.

15. Kolb, H.; Sharpless, K. The growing impact of click chemistry on drug discovery. *Drug Discov Today* **2003**, 8, 1128-1137.

16. Fink, T.; Bruggesser, H.; Reymond, J. L. Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew Chem Int Ed Engl* **2005**, 44, 1504-8.

17. Martin, Y. C. Challenges and prospects for computational aids to molecular diversity. *Perspectives in Drug Discovery and Design* **1997**, 7-8, 159-172.
18. Guvench, O.; MacKerell, A. D., Jr. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput Biol* **2009**, 5, e1000435.
19. Raman, E. P.; Yu, W.; Guvench, O.; Mackerell, A. D. Reproducing crystal binding modes of ligand functional groups using Site-Identification by Ligand Competitive Saturation (SILCS) simulations. *J Chem Inf Model* **2011**, 51, 877-96.
20. Raman, E. P.; Yu, W.; Lakkaraju, S. K.; Mackerell, A. D., Jr. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J Chem Inf Model* **2013**, 53, 3384-98.
21. Lakkaraju, S. K.; Raman, E. P.; Yu, W.; MacKerell, A. D., Jr. Free energies of organic solutes in aqueous and heterogeneous environments using grand canonical monte carlo-molecular dynamics simulations. *J Chem Theory Comput* **2014**, 10, 2281-2290.
22. Jayaram, B.; Beveridge, D. Grand canonicla Monte Carlo simulations on aqueous solutions of sodium chloride and sodium DNA: excess chemical potentials and sources of nonideality in electrolyte and polyelectrolyte solutions. *J Phys Chem* **1991**, 95, 2506-2516.
23. Mezei, M. A cavity-biased (T,V,m) Monte Carlo method for the computer simulation of fluids. *Mol Phys* **1980**, 40, 901-906.
24. Glassford, I.; Teijaro, C. N.; Daher, S. S.; Weil, A.; Redhu, S. K.; Small, M. C.; Buttaro, B.; Nicholson, A. W.; Alexander D. MacKerell, J.; Cooperman, B. S.; Andrade, R. B. Ribosome-templated azide–alkyne cycloadditions: synthesis of potent macrolide antibiotics by in situ Click chemistry. **in preparation.**

25. van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*. University of Groningen: Groningen, The Netherlands, 1987; p 1-221.
26. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **1998**, *102*, 3586-3616.
27. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **2004**, *25*, 1400-15.
28. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* **2004**, *126*, 698-699.
29. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **2012**, *32*, 1929-43.
30. Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **2000**, *21*, 86-104.
31. Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; Mackerell, A. D.,

- Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* **2012**, 8, 348-362.
32. MacKerell Jr., A. D.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2000**, 56, 257-65.
33. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* **2008**, 29, 2543-64.
34. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* **2009**, 5, 2353-2370.
35. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **2011**, 7, 3162-3180.
36. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, 5, 1315-1327.
37. Hatcher, E.; Guvench, O.; Mackerell Jr., A. D. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* **2009**, 113, 12466-76.
38. Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* **2010**, 114, 12981-12994.

39. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **2010**, *31*, 671-90.
40. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, *79*, 926-936.
41. Small, M. C.; Lopes, P.; Andrade, R. B.; Mackerell, A. D., Jr. Impact of ribosomal modification on the binding of the antibiotic telithromycin using a combined grand canonical monte carlo/molecular dynamics simulation approach. *PLoS Comput Biol* **2013**, *9*, e1003113.
42. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J Comp Chem* **1997**, *18*, 1463-1472.
43. Raman, E. P.; Vanommeslaeghe, K.; Alexander D. MacKerell, J. Site-Specific Fragment Identification Guided by Single-Step Free Energy Perturbation Calculations. *J Chem Theory Comput* **2012**, *8*, 3513-3525.
44. Yu, W.; Lakkaraju, S.; EP, E. R.; L, L. F.; AD MacKerell, J. Pharmacophore modeling using site-identification by ligand competitive saturation (SILCS) with multiple probe molecules. *J Chem Inf Model* **2015**, *55*, 407-420.
45. Yu, W.; Lakkaraju, S.; Raman, E.; AD MacKerell, J. Site-Identification by Ligand Competitive Saturation (SILCS) assisted pharmacophore modeling. *J Comput Aided Mol Des* **2014**, *28*, 491-507.
46. MOE, M. O. E. Chemical Computing Group. *1010 Sherbooke St. West, Suite*

#910, Montreal, QC, Canada, H3A, 2R7, 2012 **2012**.

47. Dunkle, J. A.; Xiong, L.; Mankin, A. S.; Cate, J. H. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc Natl Acad Sci U S A* **2010**, 107, 17152-7.

48. Pearlman, D.; Charifson, P. Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *J Med Chem* **2001**, 44, 3417-3423.

49. Hansch, C. A quantitative approach to biochemical structure-activity relationships. *Acc Chem Res* **1969**, 2, 232-239.

## Chapter 5 Polarizable Force Field for Small Ketones and Aldehydes Based on the Classical Drude Oscillator

### 5.1 Introduction

Molecular mechanics force fields (FFs) have been used extensively in computer simulations for exploring complex biological phenomena. The majority of these simulations employ Class I additive force fields that is based on the potential energy function shown in **Equation 5.1**. While the use of Class I additive FFs have yielded excellent agreement with experimental properties for a variety of systems, they exclude explicit polarizability. This is a major drawback because the electron distribution of a molecule is influenced by the surrounding environment. Thus, models that include the explicit treatment of polarization have been the subject of ongoing work for over 30 years.<sup>1</sup>

To account for the fixed charged approximation in additive force fields polarization is treated in a mean-field way by adjusting the partial atomic charges to overestimate the gas phase molecular dipoles and hence the electrostatic interactions of the molecule in aqueous systems are more accurately treated. While including polarizability implicitly has fortuitously shown good agreement between condensed phase and experimental molar volumes and heats of vaporization<sup>2-6</sup> as well as solvation free energies<sup>7-10</sup>, there is room for improvement in these properties as well as for use in biomolecules.<sup>11</sup> Indeed, recent advances in polarizable force fields are now yielding

improvements in a range of systems.<sup>12-20</sup>

$$U(\vec{R}) = \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi (1 + \cos(\eta\chi - \delta)) + \sum_{impropers} K_{imp} (\varphi - \varphi_0)^2 + \sum_{nonbond} \left( \varepsilon_{ij} \left[ \left( \frac{R_{min_{ij}}}{r_{ij}} \right)^{12} - \left( \frac{R_{min_{ij}}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\varepsilon r_{ij}} \right) \quad \text{Equation 5.1}$$

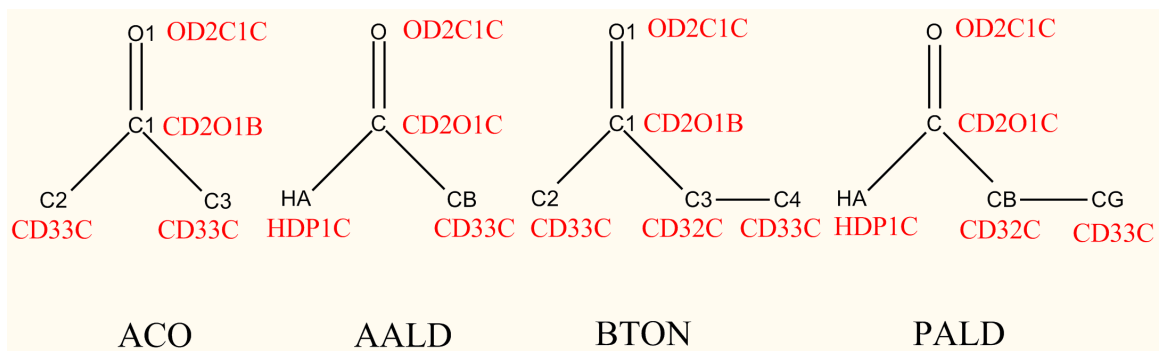
Parametrization of polarizable FFs is an active area in FF development. Explicit treatment of polarizability introduces a new term into the potential energy function,  $U_{polar}$ .<sup>21,22</sup> Current polarizable FFs differ in the method used to treat  $U_{polar}$  including the use of induced dipoles,<sup>23-27</sup> fluctuating charges,<sup>28-34</sup> or the classical Drude oscillator.<sup>21,35,36</sup> Briefly, in the induced dipole model, the functional form of  $U_{polar}$  is based on introducing an induced dipole onto each atom in addition to the partial atomic charge. This is performed in the AMOEBA polarizable FF, in addition to treatment of the static contribution to the electrostatics with a multipole expansion out to quadrupoles.<sup>17,37,38</sup> The fluctuating charge model, also known as the charge equilibration model, is based on allowing the partial atomic charges to fluctuate in response to the electric field. The polarization energy is related to the absolute (Mulliken) electronegativity<sup>39,40</sup> and the hardness of the atom,<sup>41</sup> which themselves are dependent on the electron affinity and ionization potential of an atom and are measures of the ability of the electronic distribution to distort in response to the electric field. The CHARMM CHEQ polarizable force field is based on the fluctuating charge model and work by Berne, Friesner and coworkers using both fluctuating charge and induced dipole treatments of electronic

polarization. Other efforts include the POSSUM model that is based on induced dipoles.

Models based on the classical Drude oscillator<sup>42-44</sup> (also known as Shell or charge-on-spring (COS) models) account for polarization by introducing a massless, charged particle, referred to as the Drude oscillatory, that is harmonically attached to the nucleus of the parent atom. The atomic polarizability in the Drude FF is simply the Drude charge squared divided by the force constant on the harmonic term between the Drude particle and the atomic core. In practice, the charge on the nucleus is adjusted to account for the partial atomic charge on the atom. Thus, for a fixed set of atomic positions, the Drude particles can relax in the surrounding electric field yielding the polarization response. This relaxation can be performed via energy minimization which is equivalent to a self-consistent field (SCF) calculation. However, due to the computational demand of treating the Drude particles via a SCF calculation, in practice Drude particles are instead treated as classical dynamic variables in the context of an extended Lagrangian formalism.<sup>45</sup> This is achieved by assigning a small mass (0.4 AMU) to the Drude particle from the parent atom and applying specific thermostats to the real versus the virtual Drude particles. Additionally, for computational expediency, Drude particles are not attached to hydrogen atoms in the Drude force field developed in this study. More detailed descriptions of the polarizable force fields have been published elsewhere.<sup>21,22</sup>

Development and optimization of the CHARMM Drude polarizable FF has been the subject of ongoing work in our laboratory in collaboration with Roux and coworkers since 2001.<sup>43</sup> To date, lipid<sup>12</sup>, protein<sup>16</sup>, DNA<sup>18</sup> and limited carbohydrate<sup>13,46,47</sup> parameters are available. The work presented here details parametrization of small

molecule aldehydes and ketones (viz. acetaldehyde, propionaldehyde, acetone, and butanone) as shown in **Figure 5.1** as well as their associated acyclic sugars (D-allose and D-psicose) as part of the development of a comprehensive polarizable force field for all biomolecule types. Given that Drude FF parameters are optimized with transferability in mind, the acyclic sugars can be considered as an extension of the polyols previously parametrized in this laboratory<sup>13</sup> with the addition of a ketone or aldehyde functionality. Currently, the CHARMM Drude polarizable FF does not support small ketones and aldehydes, thus the present work describes their optimization. Subsequently, these parameters are used in combination with the existing polyol parameters in the optimization of D-allose and D-psicose.



**Figure 5.1** Acetone (ACO), acetaldehyde (AALD), butanone (BTON), and propionaldehyde (PALD) with corresponding atom types from the CHARMM Drude FF. Methyl hydrogens are omitted for clarity.

## 5.2 Computational Methods

Calculations were performed with the program CHARMM, version c38a1<sup>48</sup>, and the Drude force field<sup>12,13,16,18,46,47</sup>. All MD simulations employed a 1 fs time step and

were performed at 298 K unless otherwise noted. Periodic boundary conditions and the velocity Verlet integrator were used in all condensed phase MD simulations.<sup>45</sup> Quantum mechanical (QM) calculations were performed using Gaussian03<sup>49</sup> and QCHEM<sup>50</sup>.

Internal parameters (bonds and valence angles, equilibrium values and force constants, and the dihedral force constant, phase and multiplicities) were optimized targeting QM geometries and vibrational data performed at the MP2/6-31G(d) level of theory. Bond and angle parameters were verified using crystal survey data<sup>51</sup> as performed previously for the additive CHARMM model for acetone and acetaldehyde<sup>52</sup>. Force constants for the bonds and angles were adjusted to fit QM vibrational target data. MM vibrational frequencies were obtained using the MOLVIB<sup>53</sup> utility in CHARMM which was also used to determine the spectra potential energy distributions. QM frequencies were scaled by 0.9434.<sup>54</sup> Dihedral parameters were fit using a least squares fitting algorithm that has recently been described in detail.<sup>55</sup>

For transferability, optimization of the electrostatic parameters (i.e. charges, polarizabilities (alpha) and Thole scale factors<sup>56</sup>) was first performed for acetone and acetaldehyde, with the final parameters transferred to butanone and propionaldehyde, respectively, by adding a methyl group whose electrostatic parameters were derived from ethane. Initial nonbonded parameters for acetone and acetaldehyde were obtained using GAAMP, a web utility that utilizes CHARMM's FITCHARGE module. Default values for the 1-4 interaction scaling factors, weights for water interactions, and weights for Thole and polarizability constraints were employed, while the MP2/6-31G(d) level of theory was used to generate potential energy scans used as target data to fit the torsion parameters. For acetone, the methyl carbons and methyl hydrogens were considered

equivalent, while only geminal hydrogens were considered equivalent for acetaldehyde. The partial atomic charge of oxygen was restrained at 0 using a charge restraint of 10.

Optimization of the partial atomic charges, alpha and Thole values were performed targeting dipole moments and molecular polarizabilities obtained at the MP2/6-31+G(d)//MP2/cc-PVQZ model chemistry. The position of the lone pairs (LP) of the oxygen atom was directly taken from the Drude model of N-methyl acetamide (NMA)<sup>57</sup>. Anisotropic polarizabilities<sup>58</sup> on the oxygen atoms were used as obtained from GAAMP. Partial atomic charges were fit to reproduce QM dipole moments using an automated Monte Carlo simulated annealing procedure<sup>59</sup> and refined targeting molecular polarizabilities such that the MM molecular polarizability was roughly 70-85% of the QM molecular polarizability.<sup>21,43,60-62</sup> Partial charges were verified for acetone and acetaldehyde using water interactions obtained at the MP2/6-31G(d)//MP2/cc-PVQZ levels of theory with counterpoise correction for basis set superposition error (BSSE).<sup>63,64</sup> Acetone and acetaldehyde electrostatic parameters were then transferred to butanone and propionaldehyde, respectively, as described above, with manual adjustments made to the partial atomic charge, alpha and Thole values to better reproduce the dipole moments of multiple conformations of these molecules.

Lennard-Jones (LJ) parameters were optimized targeting experimental data for all small model compounds (acetone, butanone, acetaldehyde, and propionaldehyde), including neat liquid heats of vaporization ( $\Delta H_{\text{vaporization}}$ ) and molar volumes ( $V_m$ ) using condensed phase molecular dynamics (MD) simulations as in previously published parametrization studies.<sup>13,65-71</sup> Gas-phase simulations were used to calculate the intramolecular contributions to  $\Delta H_{\text{vaporization}}$ . Condensed phase MD simulations used an

isotropic long range correction (LRC).<sup>72</sup> Neat liquid simulations were performed using a box of 125 pre-equilibrated solute molecules. Solvated liquid simulations were performed using a solute molecule immersed in a box of 250 SWM4-NDP water molecules.<sup>73</sup> All properties were calculated based on the average of 4 independent simulations that differed based on the random seed used to assign the velocities. Heats of vaporization and molecular volumes were calculated from the last 600 ps of 1 ns MD simulations performed at the following temperatures (acetone=298 K, butanone=314 K, acetaldehyde=283 K, propionaldehyde=298 K). The free energies of solvation and dielectric constants were used to verify the external parameters and, when necessary, tweak the LJ parameters. Dielectric constants were calculated using 2 ns simulations at temperatures which experimental dielectric constants were obtained (acetone=293.2 K, butanone=293.2 K, acetaldehyde=291.2 K, propionaldehyde=290.2 K). Averages were obtained from the last 1.9 ns. Free energies of solvation were obtained from free energy perturbation (FEP)<sup>74</sup> simulations as previously described,<sup>65,75</sup> by performing 50 ps equilibration and 100 ps production simulations for each lambda window. Averages were obtained from the production phase.

### **5.3 Results and Discussion**

The results presented here represent those based on the final set of optimized parameters after self-consistent optimization of both the nonbonded and bonded terms.

### 5.3.1 Nonbonded parameters

**Table 5.1** shows the QM and MM dipole moments for acetone (ACO), butanone (BTON), acetaldehyde (AALD), and propionaldehyde (PALD). For AALD, the individual contributions to the MM dipole are in excellent agreement with QM, while for ACO the contribution to the dipole arises from the C=O group due to the symmetry of the molecule and this is slightly underestimated in the Drude model. Compared to the C36 additive FF<sup>52</sup>, the Drude model is better able to reproduce the QM dipole moments demonstrating the importance of incorporating polarizability as well as the C36 additive FF charges being developed primarily targeting HF/6-31G(d) interactions with water. As mentioned, for transferability the charges, alphas and Thole factors from ACO and AALD were transferred to BTON and PALD, respectively. The additional methyl group's charge and alpha parameters were obtained from ethane and the Thole factors were manually adjusted to fit the QM dipole moments of the gauche and trans conformations of BTON and PALD. While the difference between the MM and QM dipole moments is small for each conformation, the relative differences between the gauche and trans conformations is not maintained in the MM dipoles. In both BTON and PALD, the gauche MM dipole moment is lower than the trans when it should be greater based on the QM dipole moments. However, given that the difference in the QM dipoles between the two states was less than 0.3 D, this level of disagreement was considered acceptable.

**Table 5.1** Dipole moments and molecular polarizability for acetone (ACO), acetaldehyde (AALD), butanone (BTON) gauche and trans conformations, and propionaldehyde (PALD) gauche and trans conformations. QM values were obtained at the MP2/6-31+G(d)//MP2/cc-PVQZ levels of theory. C36 values were as reported previously for the CHARMM36 additive force field.<sup>52</sup>

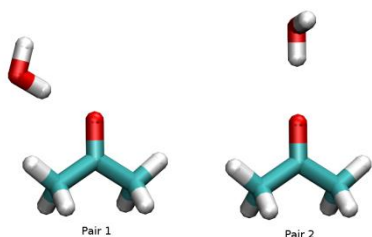
		Dipole Moments				Polarizability				
		x	y	z	Total	MM- QM	xx	yy	zz	Total
ACO										
	Drude	0.000	-2.728	0.000	2.728	-0.226	6.754	4.740	3.040	14.53
	C36	0.000	-3.570	0.000	3.570	1.191				
	QM	0.000	-2.954	0.000	2.954		6.489	6.661	4.830	17.98
AALD										
	Drude	2.565	1.111	0.000	2.795	0.007	4.167	2.493	2.188	8.85
	C36	3.280	0.640	0.000	3.350	1.021				
	QM	2.539	1.153	0.000	2.788		5.073	4.362	3.321	12.76
BTON										
<i>gauche</i>	Drude	-1.443	1.723	0.781	2.379	-0.633	8.985	6.622	4.577	20.18
	QM	-2.039	2.083	0.757	3.011		8.547	7.998	6.558	23.10
<i>trans</i>	Drude	0.649	-2.559	0.003	2.640	-0.171	10.925	5.680	4.021	20.63
	QM	0.206	-2.804	0.000	2.812		8.819	8.058	6.299	23.18
PALD										
<i>gauche</i>	Drude	2.196	-0.222	0.742	2.329	-0.569	6.795	3.876	3.576	14.25
	QM	2.661	0.076	1.146	2.898		7.044	5.439	5.525	18.01
<i>trans</i>	Drude	1.039	2.465	0.000	2.675	-0.002	5.778	4.776	3.313	13.87
	QM	-1.549	2.183	0.000	2.677		6.544	6.522	4.837	17.90

The molecular polarizability values are also shown in **Table 5.1**. In a previous study by Lamoreaux and coworkers<sup>45</sup>, it was found that the molecular polarizability needed to be scaled down to prevent overestimation of the dielectric constant. While the reason for the scaling is still point of discussion, (see <sup>21</sup>and discussion therein) for consistency and transferability of the parameters it was maintained for the carbonyls. Accordingly, for all of the compounds in the present work, the MM molecular polarizability is roughly 70 to 90% that of the QM values.

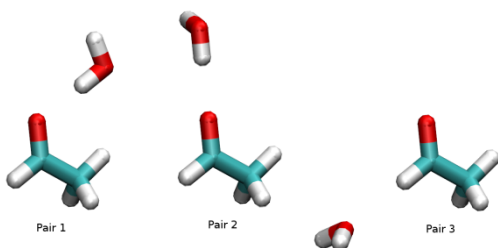
The water interactions for ACO and AALD are presented in **Table 5.2** for Drude, C36, and QM. The relative differences between the QM and MM interaction energies are all less than 0.7 kcal/mol, while the differences in the interaction distances are 0.1 Å or less. The root mean square deviations (RMSD) for the interaction energies and distances between ACO and water are 0.71 kcal/mol and 0.04 Å, respectively. Those for AALD and water are 0.728 kcal/mol and 0.12 Å, respectively. Notable is that the Drude interaction energies and distances are in better agreement with the QM than the C36 results, which in part arises from the introduction of lone pairs on the carbonyl oxygen in the Drude FF. Also, the interaction of water with the aldehydic hydrogen in Drude is in satisfactory agreement with the QM data, while the additive C36 FF interaction is too favorable. Finally, it should be noted that optimization of the additive C36 charges targeted QM HF/6-31G(d) data, which would lead to difference versus the QM data being targeted for the polarizable force field.

**Table 5.2** Water interactions for acetone (ACO) and acetaldehyde (AALD). QM values were obtained at the MP2/6-31G(d)//MP2/cc-PVQZ levels of theory with counterpoise correction for basis set superposition error (BSSE).<sup>63, 64</sup> C36 values were obtained from<sup>52</sup>. Water interaction pairs are shown in **Figure 5.2** and **5.3**. Root mean square differences (RMSD) are shown. Hydrogen water interactions are not included for the methyl groups.

Pair	$E_{QM}$	$E_{Drude}$	$E_{C36}$	$E_{Drude} - E_{QM}$	$R_{QM}$	$R_{Drude}$	$R_{C36}$	$R_{Drude} - R_{QM}$
ACO								
Pair 1	-5.52	-4.81	-7.19	0.71	2.06	2.03	1.71	-0.03
Pair 2	-4.03	-4.08	-6.49	-0.05	2.10	2.07	1.69	-0.03
RMSD				0.71				0.04
AALD								
Pair 1	-5.05	-5.07	-6.07	-0.02	2.08	2.01	1.81	-0.07
Pair 2	-3.58	-3.69	-4.96	-0.11	2.15	2.09	1.83	-0.06
Pair 3	-1.41	-1.11	-2.48	0.30	2.56	2.63	2.18	0.07
RMSD				0.78				0.12



**Figure 5.2** Water interaction pairs for acetone (ACO) corresponding to pairs in **Table 5.2**.



**Figure 5.3** Water interaction pairs for acetaldehyde (AALD) corresponding to pairs in **Table 5.2**.

Water interactions for BTON and PALD are shown in **Table 5.3**. The trans conformations were selected for water interaction calculations because these have lower relative energies compared to the gauche conformations. The RMSD values for the interaction energies in BTON and PALD are 0.65 and 1.19 kcal/mol, respectively, and the interaction distances have an RMSD of 0.07 and 0.13 Å, respectively. Though the interaction energy of Pair 1 could be improved in the ketones, modification of the charges would impact the dipole moments and polarizabilities and further refinement of the charges was not undertaken. However, since the water interactions impact the hydration free energies of each compound, an NBFIX term was required as discussed below.

**Table 5.3** Water interactions for butanone (BTON) and propionaldehyde (PALD) in trans conformations. *QM* values were obtained at the MP2/6-31G(d)//MP2/cc-PVQZ levels of theory with counterpoise correction for basis set superposition error (BSSE).<sup>63, 64</sup> Water interaction pairs are shown in **Figure 5.2** and **5.3**. Root mean square differences (RMSD) are shown. Hydrogen water interactions are not included for the methyl groups.

Pair	$E_{QM}$	$E_{Drude}$	$E_{Drude} - E_{QM}$	$R_{QM}$	$R_{Drude}$	$R_{Drude} - R_{QM}$
BTON						
Pair 1	-5.55	-4.90	0.65	2.06	2.02	-0.04
Pair 2	-4.12	-4.16	0.04	2.12	2.06	-0.06
RMSD			0.65			0.07
PALD						
Pair 1	-2.63	-1.91	-0.72	2.83	2.94	0.11
Pair 2	-4.71	-3.76	0.95	2.14	2.08	-0.06
Pair 3	-1.36	-1.39	-0.03	2.56	2.60	0.04
RMSD			1.19			0.13

**Table 5.4** shows the heats of vaporization (kcal/mol) and molecular volumes ( $\text{\AA}^3$ ) for ACO, BTON, AALD, and PALD. The reported heats of vaporization and molar volumes represent a compromise between ACO and BTON and between AALD and PALD. The MM heats of vaporization are overestimated in ACO and AALD, but underestimated in the larger BTON and PALD. Similarly, the molar volumes are too small in ACO and AALD, yet too large in BTON and PALD. The LJ parameters were specifically optimized to achieve this balance. Notably, the oxygen atom type of the ketones and aldehydes is the same, however, it was found to be necessary to vary the carbonyl carbon atom type to obtain reasonable agreement with the experimental data. The heats of vaporization and molar volumes for acetone in the C36 additive FF<sup>76</sup> are 7.37 kcal/mol and 124.47  $\text{\AA}^3$ , respectively. For acetaldehyde, they are 6.21 kcal/mol and 94.84  $\text{\AA}^3$ , respectively. While these values represent a slight improvement over the current Drude FF, the larger deviation in the Drude FF is the result of balancing the LJ parameters for BTON and PALD as mentioned above, which were not considered when optimizing the additive FF.

The optimized LJ and electrostatic parameters were further validated against experimental free energies of solvation and dielectric constants as shown in **Table 5.4**. To achieve reasonable agreement between the experimental and Drude free energies of solvation, the introduction of an NBFIX term was required between the carbonyl oxygen (OD2C1C) and water oxygen (ODW) in which the well depth (epsilon) was increased from -0.2054 to -0.3554 and the minimum distance ( $R_{\min}$ ) was decreased from 3.5869 to 3.5269. In addition, the  $R_{\min}$  value of the ketone carbon (CD2O1B) and water oxygen (ODW) was decreased from 3.5569 to 3.5069, while the value for the aldehyde carbon

(CD2O1C) and water oxygen was increased from 3.3169 to 3.3669. The well depths for the ketone/aldehyde carbon to water oxygen LJ were unchanged. This allowed a balance between the MM and QM solvation free energies of the present compounds to be maintained, though the MM  $\Delta G_{\text{solvation}}$  for AALD is somewhat too favorable. The MM dielectric constants are all in good agreement with the QM values and no further optimization was deemed necessary.

**Table 5.4** Calculated (Drude) heats of vaporization ( $\Delta H_{\text{vaporization}}$ ) in kcal/mol, molar volumes  $V_m$  in  $\text{\AA}^3$ ,  $\Delta G_{\text{solvation}}$  and dielectric constants ( $\epsilon$ ) for acetone (ACO), butanone (BTON), acetaldehyde (AALD), and propionaldehyde (PALD) compared to experimental values (Exp). Properties were obtained at the same temperature as experimental conditions: ACO 298 K, BTON 314 K, AALD 283 K, and PALD 298 K for heats of vaporization and molar volumes; ACO and BTON 293.2 K, AALD 291.2 K, PALD 290.2 K for dielectric constants. Standard deviations were calculated based on the average of 4 simulations.

	$\Delta H_{\text{vaporization}}$	% (Drude - Exp)	$V_m$	% (Drude - Exp)	$\Delta G_{\text{solvation}}$	% (Drude - Exp)	$\epsilon$
ACO							
Drude	7.55 ± 0.02	0.9	119.82 ± 0.08	-2.3	-4.01±0.07	4.2	16.5 ± 0.4
Exp	7.48		122.67		-3.85		21.01
BTON							
Drude	7.89 ± 0.03	-2.4	154.7 ± 0.2	1.1	-3.5±0.1	-3.8	14.2 ± 0.4
Exp	8.08		153.07		-3.64		18.6
AALD							
Drude	6.69 ± 0.02	1.4	88.4 ± 0.3	-4.1	-3.9±0.2	11.4	22.1 ± 0.2
Exp	6.60		92.14		-3.50		21.00
PALD							
Drude	6.95 ± 0.02	-2.7	122.9 ± 0.3	1.5	-3.3±0.2	-4.1	15.8 ± 0.5
Exp	7.14		121.14		-3.44		18.5

### 5.3.2 Bonded parameters

**Tables 5.5** and **5.6** shows the bond and angle lengths for ACO, BTON, AALD, PALD as measured in the Drude FF, QM optimizations, and from a survey of crystallographic data (CSD)<sup>51</sup> for ACO and AALD as previously reported in reference <sup>52</sup>. For the ketones, only bonds and angles involving non-hydrogen atoms are shown, while those containing the aldehydic hydrogen are included for the aldehydes. All bonds and angles are in excellent agreement with the QM values, as indicated by small relative differences. Bonds are within 0.02 Å, while angles are within 3°. For ACO and AALD, the MM bond length and valance angles are also in good agreement with the CSD values with the largest deviation coming from the HA-C=O in acetaldehyde at 3.8°, though the standard deviation of this angle from the crystallographic survey is over 3° and in general hydrogens are poorly resolved in X-ray crystallographic studies.

The MM and QM molecular vibrations for ACO and AALD are shown in **Tables 5.7** and **5.8** using the internal coordinate system of Pulay<sup>77</sup>. MM frequencies were obtained by optimizing the MM force constants to reproduce the scaled QM frequencies. Most notable are the low frequency vibrations that describe larger conformational changes in the molecule as these dominate during MD simulations. While the MM frequency for the O=C-CB-HB1 torsion in AALD is larger than ideal, the remaining MM frequencies are in excellent agreement with the QM scaled frequencies. Moreover, the MM frequencies for ACO are all within 5% of their QM target values.

**Table 5.5** Bond lengths in Å and angle lengths in degrees for acetone (ACO) and acetaldehyde (AALD) from QM optimizations at the MP2/6-31G(d) level of theory, crystallographic surveys as reported in <sup>52</sup> and Drude calculations. Crystal survey error is shown in parenthesis.

Atom pair	QM	Crystal	MM	MM-QM	MM-Crystal
ACO					
C1-O1	1.228	1.207 (0.044)	1.217	-0.011	0.010
C1-C2/C1-C3	1.514	1.472 (0.048)	1.500	-0.014	0.028
C2/C3-C1-O1	121.75	121.04 (3.96)	121.73	-0.02	0.69
C2-C1-C3	116.50	117.46 (3.67)	116.53	0.03	-0.93
AALD					
C-O	1.224	1.217 (0.054)	1.218	-0.006	0.001
C-CB	1.503	1.476 (0.048)	1.498	-0.005	0.022
C-HA	1.109	0.973 (0.059)	1.104	-0.005	0.131
CB-C-O	124.30	124.93 (3.71)	122.45	-1.85	-2.48
HA-C-O	120.33	117.15 (3.76)	120.96	0.63	3.81
HA-C-CB	115.37	117.53 (3.22)	116.59	1.22	-0.94

**Table 5.6** Bond lengths in Å and angle lengths in degrees for butanone (BTON) and propionaldehyde (PALD) gauche and trans conformations from QM optimizations at the MP2/6-31G(d) level of theory and Drude calculations.

	Atom pair	QM	MM	MM-QM
	C1-O1	1.229	1.218	-0.011
	C2-C1	1.515	1.501	-0.013
	C1-C3	1.516	1.505	-0.012
BTON	C3-C4	1.532	1.537	0.005
<i>gauche</i>	C2-C1-O	121.21	121.11	-0.10
	O-C1-C3	121.47	121.07	-0.40
	C2-C1-C3	117.32	117.80	0.49
	C1-C3-C4	113.13	113.89	0.76

*Table 5.6 continued*

	C1-O1	1.228	1.218	-0.010
	C2-C1	1.514	1.500	-0.014
	C1-C3	1.518	1.504	-0.014
BTON	C3-C4	1.522	1.538	0.017
<i>trans</i>	C2-C1-O	121.64	121.33	-0.31
	O-C1-C3	121.69	122.13	0.44
	C2-C1-C3	116.67	116.54	-0.13
	C1-C3-C4	113.08	114.50	1.43
<hr/>				
	C-O	1.224	1.218	-0.006
	C-CB	1.506	1.496	-0.009
	CB-Cg	1.532	1.533	0.001
PALD	C-HA	1.111	1.104	-0.007
<i>gauche</i>	HA-C-O	120.35	121.12	0.77
	HA-C-CB	114.94	116.55	1.61
	CB-C-O	124.70	122.29	-2.4
	C-CB-CG	111.06	111.80	0.74
<hr/>				
	C-O	1.224	1.218	-0.006
	C-CB	1.507	1.500	-0.007
	CB-Cg	1.521	1.536	0.015
PALD	C-HA	1.110	1.104	-0.006
<i>trans</i>	HA-C-O	120.28	120.84	0.56
	HA-C-CB	115.51	115.86	0.36
	CB-C-O	124.22	123.30	-0.91
	C-CB-CG	113.10	114.18	1.08
<hr/>				

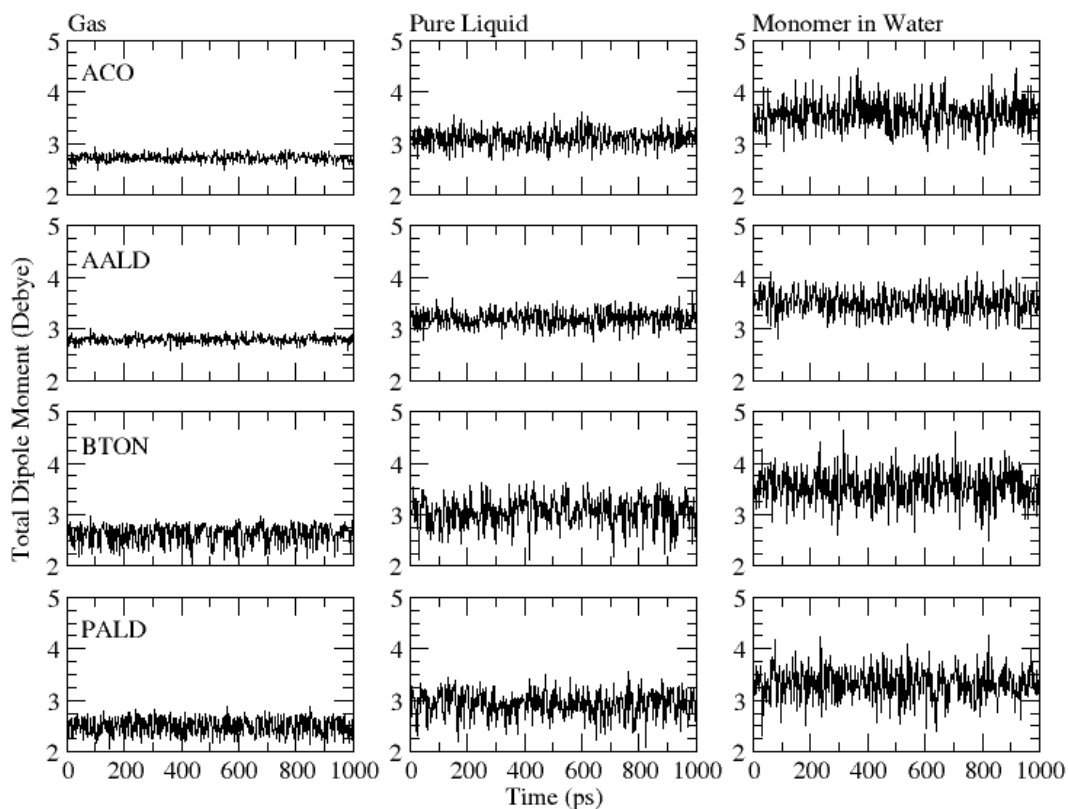
**Table 5.7** Vibrational analysis of acetone (ACO) using the internal coordinate system of Pulay<sup>77</sup>. QM frequencies were obtained at the MP2/6-31G(d) level of theory.

Mode	QM	MM
CHtors	84.5	85.8
dCCC	463.7	471.6
wCO	659.1	627.8
rCO	665.6	661.4
sCC	805.3	813.2
rCH3	951.6	990.0
rCH3'	977.9	940.7
sCCas	1036.5	1044.9
dCH3	1360.4	1342.6
dCH3a	1384.5	1422.0
dCH3a'	1386.0	1426.7
sCO	1564.6	1588.5
sCH3	2849.2	2934.1
sCH3a	2913.4	3048.3
sCH3a'	2914.2	3006.2

**Table 5.8** *Vibrational analysis of acetaldehyde (AALD) using the internal coordinate system of Pulay<sup>77</sup>. QM frequencies were obtained at the MP2/6-31G(d) level of theory.*

Mode	QM	MM
tCH3	81.1	195.3
dCCC	590.9	632.1
sCC	927.9	974.6
rCH3'	929.0	916.2
rCH3	982.9	974.8
wCHO	1004.5	955.4
rCHa	1300.5	1372.6
dCH3	1348.5	1339.2
dCH3a	1384.9	1391.5
dCH3a'	1399.2	1407.0
sCO	1562.0	1622.0
sCHa	2818.3	2823.1
sCH3	2848.4	2933.5
sCH3a'	2912.0	3005.6
sCH3a	2914.5	3048.8

**Figure 5.4** shows the dipole moments as a function of time for ACO, AALD, BTON, and PALD in gas phase, pure liquid, and aqueous solution. The strength of polarizable force fields rests in the ability of the dipole to fluctuate in response to the environment. This is clearly illustrated by the increase in the magnitude and the variations of the dipole moment in moving from gas phase to condensed phase. The averages and root mean square fluctuations (RMSF) of the dipole moments are shown in **Table 5.9**. In ACO and AALD the averages are 2.71 and 2.79, respectively, with an RMSF of 0.07 D. In the condensed phase simulations, the average increases to 3.1 and 3.2 and the RMSF increases to 0.2-0.3 D, which are indicative of dipole moments of the individual monomers in the condensed phase responding to the molecules in the local environment, whether those molecules are other monomers (pure liquid) or water molecules (aqueous solution). For BTON and PALD, the RMSF is 0.2 to 0.3 D for gas and condensed phase simulations. Unlike ACO and AALD, these two molecules can assume a trans or gauche conformation during MD and therefore two populations of dipole moments are represented in the gas phase plots, which increases the RMSF. In going from gas to condensed phase the average dipole moment increases as expected due to the presence of surrounding molecules that induce the dipole moment. Within the condensed phase simulations, the monomer in aqueous solution has a larger dipole moment because water is a polar solvent as opposed to the pure liquids and thus the polarization response to water is larger. Notably, this is a phenomenon that the additive FF is intrinsically unable to reproduce, which highlights the importance of including polarizability.



**Figure 5.4** Dipole moments as a function of time for acetone (ACO), acetaldehyde (AALD), butanone (BTON), and propionaldehyde (PALD) in gas phase (left column), pure liquid (middle column), and aqueous solution (right column).

**Table 5.9** The averages and RMSF for the dipole moments of acetone (ACO), acetaldehyde (AALD), butanone (BTON), and propionaldehyde (PALD) from gas and condensed phase simulations (pure liquid and aqueous solution).

Molecule	Gas	Pure Liquid	Water
ACO	$2.71 \pm 0.07$	$3.1 \pm 0.2$	$3.6 \pm 0.3$
AALD	$2.79 \pm 0.07$	$3.2 \pm 0.2$	$3.5 \pm 0.2$
BTON	$2.6 \pm 0.2$	$3.0 \pm 0.3$	$3.5 \pm 0.3$
PALD	$2.5 \pm 0.2$	$2.9 \pm 0.2$	$3.3 \pm 0.3$

## 5.4 Conclusion

The work presented here details parametrization of small molecule aldehydes and ketones (viz. acetaldehyde, propionaldehyde, acetone, and butanone) as part of the development of a comprehensive Drude polarizable FF for biomolecules, including acyclic sugars such as D-allose and D-psicose. The presented models for ketones and aldehydes are shown to be in good agreement with QM and experimental target data and will be used in combination with the recently parametrized polyols<sup>13</sup> for the optimization of D-allose and D-psicose. Future parametrization will focus on optimization of the covalent linkages between the polyols and ketones/aldehydes.

## 5.5 References

1. Warshel, A.; Levitt, M. Theoretical studies of enzymatic reactions: dielectric, electrostatic, and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* **1976**, 103, 227-249.
2. Fox, T.; Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J Phys Chem B* **1998**, 102, 8070-8079.
3. Gough, C. A., DeBolt, S. E. and Kollman, P. A. Derivation of Fluorine and Hydrogen Atom Parameters Using Liquid Simulations. *J Comput Chem* **1992**, 13, 963-970.
4. Jorgensen, W. L. Optimized Intermolecular Potential Functions for Liquid Alcohols. *J Phys Chem* **1986**, 90, 1276-1284.

5. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* **1996**, 118, 11225-11236.
6. MacKerell, A. D., Jr.; Karplus, M. Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids. *J Phys Chem* **1991**, 95, 10559-10560.
7. Chen, I. J.; Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of Lennard-Jones parameters for polar-neutral compounds. *J Comput Chem* **2002**, 23, 199-213.
8. Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model. *J Phys Chem* **1994**, 98, 13077-13082.
9. Rizzo, R. C.; Jorgensen, W. L. OPLS all-atom model for amines: resolution of the amine hydration problem. *J Am Chem Soc* **1999**, 121, 4827-4836.
10. Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of lennard-jones parameters. *J Comput Chem* **1998**, 19, 334-338.
11. Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys* **2003**, 119, 5740-5761.
12. Chowdary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D., Jr.; Roux, B. A polarizable force field of dipalmitoylphosphatidylcholine based on the classical drude model for molecular dynamics simulations of lipids. *J Phys Chem B* **2013**, 117, 9142-9160.

13. He, X.; Lopes, P. E.; MacKerell, A. D., Jr. Polarizable empirical force field for acyclic polyalcohols based on the classical drude oscillator. *Biopolymers* **2013**, *99*, 724-38.
14. Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc Natl Acad Sci U S A* **2008**, *105*, 6290-5.
15. Jiao, D.; Zhang, J.; Duke, R. E.; Li, G.; Schnieders, M. J.; Ren, P. Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential. *J Comput Chem* **2009**, *30*, 1701-11.
16. Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr. Force Field for Peptides and Proteins based on the Classical Drude Oscillator. *J Chem Theory Comput* **2013**, *9*, 5430-5449.
17. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J Phys Chem B* **2010**, *114*, 2549-64.
18. Savelyev, A.; MacKerell, A. D., Jr. All-atom polarizable force field for DNA based on the classical drude oscillator model. *J Comput Chem* **2014**, *10*, 1652-1664.
19. Shi, Y.; Zhu, C. Z.; Martin, S. F.; Ren, P. Probing the effect of conformational constraint on phosphorylated ligand binding to an SH2 domain using polarizable force field simulations. *J Phys Chem B* **2012**, *116*, 1716-27.

20. Zhang, J.; Yang, W.; Piquemal, J. P.; Ren, P. Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J Chem Theory Comput* **2012**, 8, 1314-1324.
21. Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. Theory and applications. *Theor Chem Acc* **2009**, 124, 11-28.
22. Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev Comp Chem* **2002**, 18, 89-146.
23. Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. An Anisotropic Polarizable Water Model: Incorporation of All-Atom Polarizabilities into Molecular Mechanics Force Fields. *J Phys Chem* **1994**, 98, 4180-4187.
24. Caldwell, J.; Dang, L. X.; Kollman, P. A. Implementation of Nonadditive Intermolecular Potentials by Use of Molecular Dynamics: Development of a Water-Water Potential and Water-Ion Cluster Interactions. *J Am Chem Soc* **1990**, 112, 9144-9147.
25. Dang, L. X. Importance of Polarization Effects in Modeling Hydrogen Bond in Water Using Classical Molecular Dynamics Techniques. *J Phys Chem B* **1998**, 102, 620-624.
26. Sprik, M.; Klein, M. L. A polarizable model for water using distributed charge sites. *J Chem Phys* **1988**, 89, 7556-7560.
27. Wallqvist, A.; Berne, B. J. Effective potentials for liquid water using polarizable and nonpolarizable models. *J Phys Chem* **1993**, 97, 13841-13851.
28. Asensio, J. L.; Canada, F. J.; Chen, X. H.; Khan, N.; Mootoo, D. R.; Jimenez-Barbero, J. Conformational differences between O- and C-glycosides: The alpha-O-Man-

(1 → 1)-beta-Gal/alpha-C-Man-(1 → 1)-beta-Gal case - A decisive demonstration of the importance of the exo-anomeric effect on the conformation of glycosides. *Chem European Journal* **2000**, 6, 1035-1041.

29. Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A.

Cooperative effects in the structure of fluoride water clusters: Ab initio hybrid quantum mechanical/molecular mechanical model incorporating polarizable fluctuating charge solvent. *J Chem Phys* **1998**, 109, 3077-3085.

30. Llanta, E.; Ando, K.; Rey, R. Fluctuating charge study of polarization effects in chlorinated organic liquids. *J Phys Chem B* **2001**, 105, 7783-7791.

31. Patel, S.; Brooks, C. L., 3rd. CHARMM fluctuating charge force field for proteins: I Parameterization and application to bulk organic liquid simulations. *J Comput Chem* **2003**, 25, 1-15.

32. Rick, S. W.; Berne, B. J. Dynamical fluctuating charge force fields: the aqueous solvation of amides. *J Am Chem Soc* **1996**, 118, 672-679.

33. Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J. Fluctuating Charge Force Fields for Aqueous Solutions. *Studies in Physical and Theoretical Chemistry* **1995**, 83, 31-40.

34. Yoshii, N.; Miyauchi, R.; Miura, S.; Okazaki, S. A molecular-dynamics study of the equation of state of water using a fluctuating-charge model. *Chem Phys Let* **2000**, 317, 414-420.

35. Kunz, A.-P. E.; van Gunsteren, W. F. Development of a nonlinear classical polarization model for liquid water and aqueous solution: COS/D. *J Phys Chem A* **2009**, 113, 11570-11579.

36. van Maaren, P. J.; van der Spoel, D. Molecular dynamics simulations of water within novel shell-model potentials. *J Phys Chem B* **2001**, 105, 2618-2626.
37. Ren, P.; Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J Comput Chem* **2002**, 23, 1497-506.
38. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. The Polarizable Atomic Multipole-based AMOEBA Force Field for Proteins. *J Chem Theory Comput* **2013**, 9, 4046-4063.
39. Iczkowski, R. P.; Margrave, J. L. Electronegativity. *J Am Chem Soc* **1961**, 83, 3547-3551.
40. Mulliken, R. S. A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. *J Chem Phys* **1934**, 2, 782-793.
41. Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *J Am Chem Soc* **1983**, 105, 7512-7516.
42. Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J Chem Theory Comput* **2005**, 1, 153-168.
43. Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J Chem Phys* **2003**, 119, 5185-5197.
44. Yu, H.; Hansson, T.; Van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J Chem Phys* **2003**, 118, 221-234.

45. Lamoureux, G.; Roux, B. Modeling Induced Polarization with Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J Chem Phys* **2003**, 119, 3025-3039.
46. Jana, M.; Alexander D. MacKerell, J. CHARMM Drude polarizable force field for aldepentofuranoses and methyl-aldopentofuranosides. *J Phys Chem B* **2015**, just accepted.
47. Patel, D. S.; He, X.; MacKerell, A. D., Jr. Polarizable empirical force field for hexopyranose monosaccharides based on the classical drude oscillator. *J Phys Chem B* **2015**, 119, 637-652.
48. Brooks, B. R.; Brooks III, C. L.; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J Comput Chem* **2009**, 30, 1545-1614.
49. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven Jr., T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, K.; Kitao, O.; Nakai, H.; Klene, M.; Li, T. W.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.;

Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.04; Gaussian, Inc.: Pittsburgh, PA, 2003.

50. Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; Jr., R. A. D.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Voorhis, T. V.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; III, H. L. W.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; III, H. F. S.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M.; Gan, Z.; Zhao, Y.; Schultz, N. E.; Truhlar, D.; Epifanovsky, E.; Oana, M.; Q-Chem. *Q-Chem 3.1*; Q-Chem, Inc.: Pittsburgh, PA, 2007.

51. Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-

Based Search, Retrieval, Analysis and Display of Information. *Acta Cryst* **1979**, B35, 2331-2339.

52. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, 5, 1315-1327.

53. Kuczera, K.; Wiorkiewicz, J. K.; Karplus, M. *MOLVIB: Program for the Analysis of Molecular Vibrations*, CHARMM, Harvard University: 1993.

54. Scott, A. P.; Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* **1996**, 100, 16502-16513.

55. Vanommeslaeghe, K.; AD MacKerell, J. Robustness in the fitting of molecular mechanics parameters. *J Comp Chem* **2015**, in press.

56. Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem Phys* **1981**, 59, 341-350.

57. Lin, B.; Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Kirkwood-Buff analysis of aqueous N-methylacetamide and acetamide solutions modeled by the CHARMM additive and Drude polarizable force fields. *Journal of Chemical Physics* **2013**, 139, 084509.

58. Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; Jr., A. D. M.; Roux, B. Atomic Level Anisotropy in the Electrostatic Modeling of Lone Pairs for a Polarizable Force Field Based on the Classical Drude Oscillator. *J Chem Theory Comput* **2006**, 2, 1587-1597.

59. Guvench, O.; MacKerell, A. D., Jr. Automated conformational energy fitting for force-field development. *J Mol Model* **2008**, 14, 667-79.

60. Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; AD MacKerell, J. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem Phys Lett* **2006**, 418, 245-249.
61. Morita, A.; Kato, S. An ab initio analysis of medium perturbation on molecular polarizabilities. *Journal of Chemical Physics* **1999**, 110, 11987-11998.
62. Schropp, B.; Tavan, P. The polarizability of point-polarizable water models: density functional theory/molecular mechanics results. *J Phys Chem B* **2008**, 112, 6233-40.
63. Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interaction by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, 19, 553-566.
64. Ransil, B. Studies in molecular structure. IV. Potential curve for the interaction of two helium atoms in single-configuration LCAO MO SCF approximation. *J Chem Phys* **1961**, 34, 2109.
65. Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; Mackerell, A. D., Jr. Polarizable empirical force field for the primary and secondary alcohol series based on the classical Drude model. *J Chem Theory Comput* **2007**, 3, 1927-1946.
66. Baker, C. M.; Lopes, P. E.; Zhu, X.; Roux, B.; MacKerell, A. D., Jr. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J Chem Theory Comput* **2010**, 6, 1181-1198.

67. Harder, E.; Anisimov, V. M.; Whitfield, T.; MacKerell, A. D., Jr.; Roux, B. Understanding the dielectric properties of liquid amides from a polarizable force field. *J Phys Chem B* **2008**, 112, 3509-21.
68. Lopes, P. E.; Lamoureux, G.; Mackerell, A. D., Jr. Polarizable empirical force field for nitrogen-containing heteroaromatic compounds based on the classical Drude oscillator. *J Comput Chem* **2009**, 30, 1821-38.
69. Lopes, P. E.; Lamoureux, G.; Roux, B.; Mackerell, A. D., Jr. Polarizable empirical force field for aromatic compounds based on the classical drude oscillator. *J Phys Chem B* **2007**, 111, 2873-85.
70. Vorobyov, I.; Anisimov, V. M.; Greene, S.; Venable, R. M.; Moser, A.; Pastor, R. W.; MacKerell, A. D., Jr. Additive and Classical Drude Polarizable Force Fields for Linear and Cyclic Ethers. *J. Chem. Theory Comp.* **2007**, 3, 1120-1133.
71. Zhu, X.; MacKerell, A. D., Jr. Polarizable empirical force field for sulfur-containing compounds based on the classical Drude oscillator model. *J Comput Chem* **2010**, 31, 2330-41.
72. Lagüe, P.; Pastor, R. W.; Brooks, B. R. Pressure-based long-range correction for lennard-jones interactions in molecular dynamics simulations: applications to alkanes and interfaces. *J Phys Chem B* **2004**, 108, 363-368.
73. Yu, W.; Lopes, P. E.; Roux, B.; A.D. MacKerell, J. Six-site polarizable model of water based on the classical drude oscillator. *J Chem Phys* **2013**, 138, 034508-034513.
74. Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem Rev* **1993**, 93, 2395-2417.

75. Deng, Y.; Roux, B. Hydration of amino acid side chains: nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J Phys Chem B* **2004**, 108, 16567-16576.
76. Hatcher, E.; Guvench, O.; Mackerell Jr., A. D. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* **2009**, 5, 1315-1327.
77. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives. *J Am Chem Soc* **1979**, 101, 2550-2560.

## Chapter 6 Conclusion

Overcoming microbial resistance is a major challenge in the development of antibiotics. The macrolides are an important class of antibiotics that target the ribosome and recent generation macrolides have largely addressed resistance stemming from drug efflux pumps and metabolic enzymes. However, they remain susceptible to resistance due to modification of the ribosome. In the work that has been presented here, the focus was on resistance due to modification of base A2058 (*E. coli* numbering). As there are no crystal structures available yet for A2058-modified ribosomes, computational models were developed to study the ribosome as well as macrolide antibiotic analogs aimed at addressing the need for novel antibiotics.

Chapter 1 introduced force fields and force field-based simulations, which is the methodology underlying all of the computational models in this body of work. Two notable applications of force fields were discussed including the conformationally sampled pharmacophore (CSP) and site-identification by ligand competitive saturation (SILCS) approaches. CSP is a pharmacophore-based method in which ensembles of ligand conformations are generated using MD simulations from which probability distributions are calculated for select distances and angles in the molecule (i.e. pharmacophore features). Analyses are then performed to correlate biological activity of the ligand with the CSP pharmacophoric features. SILCS harnesses the power of force field-based simulations to map the functional group affinities of a protein. The fragments are a diverse set of small molecules selected such that they capture the chemical features

common to drugs, namely hydrophobic and hydrogen bonding groups, including both charged and neutral species. During the simulations, the fragments interact with all regions on the protein, thus allowing the generation of 3D probability maps (FragMaps), which are normalized with respect to the fragments in solution alone and converted to free energies (called Grid Free Energies, GFEs). The 3D distribution of GFE FragMaps identify regions of the protein that bind fragments favorably, thus identifying the types of functional groups that can successfully interact with the protein. In addition, a Monte Carlo based sampling algorithm allows for the docking of ligands in the field of the FragMaps, in which binding affinity is estimated based on the ligand-based GFE (LGFE).

In Chapter 2, the CSP approach was employed to investigate the effects of macrolide desmethylation on pharmacological activity. A model based on select intramolecular distances spanning the macrolactone ring as well as the alkyl-aryl (ARM) moiety was developed. The CSP-based model indicated that reduced bioactivity as measured in minimum inhibitory concentrations (MIC) resulted from an increase in conformational flexibility when methyl groups were removed from the macrocycle. Contrary to the hypothesis that removal of the C4 methyl group would increase activity against A2058G mutant ribosomes by alleviating a steric clash between the C4 methyl and G2058 N2, no improvement in bioactivity was observed. However, the studies did confirm that VDW interactions between methyl groups extending from the macrocycle are important for binding, which is an important consideration in the design of novel macrolide analogs in the future.

Chapter 3 shifted the focus from the ligand to the ribosome. Molecular dynamics (MD) simulations of truncated version of the *E. coli* 50S subunit were presented with the

aim of understanding the impact of A2058 modification on the binding of telithromycin. It was shown that modification of 2058 perturbs the conformation of the entire macrolide binding pocket through a pathway of bases extending from 2058 to a A752-U2609 Watson-Crick base pair. Furthermore, these conformational changes affect the VDW interactions of the macrocycle with the bases delineating the pocket, which is an important driving force for macrolide binding. In fact, when compared to telithromycin activity against 2058-modified ribosomes, the presented studies indicate that these conformational perturbations impact binding affinity to a greater extent than disruption of the telithromycin 2'-OH to A2058 N1 hydrogen bond, which was maintained to a reasonable extent even in the dimethylated 2058 systems. Future MD simulations of the ribosome may focus on other macrolide antibiotics bound within the 50S subunit, including the most recently developed solithromycin that has greater antimicrobial activity and lesser toxicity than existing telithromycin. Even so, simulations would greatly benefit from the availability of 2058-modified ribosome crystal structures.

Chapter 4 highlighted the application of the Site-Identification by Ligand Competitive Saturation (SILCS) approach to the ribosome, including a Monte Carlo sampling algorithm that docks the ligands into the SILCS FragMaps to obtain ligand grid free energy scores (LGFEs). Two separate models were developed for a congeneric series of the compounds synthesized by ribosome-templated *in situ* click chemistry using 1) the overall binding affinity of macrolides and 2) the ability of alkyne-derived moieties to undergo the Click reaction. In the first model, LGFEs correlated to experimental dissociation constants for the assembled macrolide analogs indicated that the predictive power of the LGFEs were dependent on the chemical structure of the alkyl-R group

(ARM) moieties, which impacts the overall interaction of the macrolactone and desosamine moieties of macrolides with the ribosome. In the second model, click enhancement for the alkyne-derived fragments was determined to be dependent on normalized LGFE values, molecular weight, and hydrophobicity as estimated by logP squared values. Going forward, these two models will allow for the prediction of ARM modifications that maximize the interaction of the macrolactone and desosamine moieties with the ribosome and whose alkyne analogs have a high probability of accessing the peptidyl-transfer pocket of the ribosome to undergo the Click reaction. Moreover, as the novel macrolide antibiotics are intended to be active against bacteria with 2058 modified ribosomes, in the future this approach should be used with SILCS maps generated based on modified ribosomes that are known to be resistant to the macrolide class of antibiotics, including A2058-modified ribosomes.

Potential future directions for the development of novel macrolide antibiotics include pursuing congeners based on the desosamine sugar, which engages in hydrogen bonding and electrostatic interactions with neighboring RNA bases, or building upon interactions of the ARM with the ribosome by adding a second ARM. The advantage of the SILCS approach in these cases is that the FragMaps are already generated, saving significant time and effort. Preliminary tests of macrolide analogs bearing multiple ARM moieties indicate more favorable LGFEs compared to ligands bearing a single ARM. Such analogs may exploit the ribosome-templated *in situ* click chemistry approach discussed in Chapter 4 using multiple types of precursors, thereby expanding the number of potential congeners possible in a single *in situ* reaction. Future work will have to take into consideration the impact on pharmacokinetic properties due to the presence of an

additional ARM.

In Chapter 5, the optimization of small molecule aldehydes and ketones for the CHARMM Drude polarizable force field was presented, which will serve as the basis of optimization of the acyclic sugars D-allose and D-psicose in the future. The addition of acyclic sugars as well as inositol will continue the development of a comprehensive CHARMM Drude polarizable for all biomolecule types. In the context of the computer-aided drug design theme presented in these bodies of work, the use of a polarizable force field on the ribosome has yet to be explored. The macrolide binding pocket is dominated by RNA, which is highly charged due to the backbone and may benefit from the inclusion of explicit polarization. Furthermore, macrolide analogs that exploit modifications of the positively charged desosamine sugar will be enhanced by the consideration of polarization effects as the sugar engages in electrostatic interactions with neighboring RNA. These future directions will rely on the continued development of the CHARMM Drude polarizable force field as well as the extension of the SILCS methodology to include Drude simulations.

In conclusion, the body of work presented here provides a comprehensive model by which future macrolide antibiotic analogs can be evaluated prior to synthesis. These models exploit the power of computational simulations to focus down the fragment universe into a more tractable size for chemical and biological evaluation. With the support of approaches like SILCS and CSP, a potent antibiotic with activity against macrolide-resistant bacteria is certainly achievable.

## Chapter 7 Bibliography

1. Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of Factor Xa ligand binding. *J Am Chem Soc* 2008, 130, 2817-2831.
2. Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-Based Search, Retrieval, Analysis and Display of Information. *Acta Cryst* 1979, B35, 2331-2339.
3. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*. Clarendon Press: Oxford, 1987; p 385.
4. Andersen, H. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 1980, 72, 2384.
5. Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J Chem Theory Comput* 2005, 1, 153-168.
6. Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D., Jr. Polarizable empirical force field for the primary and secondary alcohol series based on the classical Drude model. *J Chem Theory Comput* 2007, 3, 1927-1946.
7. Åqvist, J.; Hansson, T. On the validity of electrostatic linear response in polar solvents. *J Phys Chem* 1996, 100, 9512-9521.
8. Åqvist, J.; Medina, C.; Samuelsson, J.-E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng* 1994, 7, 385-391.
9. Åqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng* 1994, 7, 385-391.
10. Asensio, J. L.; Canada, F. J.; Chen, X. H.; Khan, N.; Mootoo, D. R.; Jimenez-Barbero, J. Conformational differences between O- and C-glycosides: The alpha-O-Man-(1 -> 1)-beta-Gal/alpha-C-Man-(1 -> 1)-beta-Gal case - A decisive demonstration of the importance of the exo-anomeric effect on the conformation of glycosides. *Chem European Journal* 2000, 6, 1035-1041.
11. Bakan, A.; Nevins, N.; Lakdawala, A. S.; Bahar, I. Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules. *J Chem Theory Comput* 2012, 8, 2435-2447.
12. Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr. Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *The journal of physical chemistry. B* 2011, 115, 580-96.
13. Baker, C. M.; Lopes, P. E.; Zhu, X.; Roux, B.; MacKerell, A. D., Jr. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J Chem Theory Comput* 2010, 6, 1181-1198.
14. Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E., 3rd; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J Chem Theory Comput* 2010, 6,

3836-3849.

15. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J Phys Chem* 1993, 97, 10269-10280.
16. Becker, O. M.; MacKerell, A. D., Jr.; Roux, B.; Watanabe, M. *Computational Biochemistry and Biophysics*. Marcel-Dekker, Inc.: New York, 2001; p 512.
17. Beglov, D.; Roux, B. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J Chem Phys* 1994, 100, 9050-9063.
18. Ben-Shimon, A.; Eisenstein, M. Computational mapping of anchoring spots on protein surfaces. *J Mol Biol* 2010, 402, 259-277.
19. Berendsen, H. Molecular dynamics with couple to an external bath. *J Chem Phys* 1984, 81, 3684-3690.
20. Berendsen, H.; Postma, J.; Gundsteren, W. v.; DiNola, A.; Haak, J. Molecular dynamics with coupling to an external bath. *J Chem Phys* 1984, 81, 3684-3690.
21. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. The Missing Term in Effective Pair Potentials. *J Phys Chem* 1987, 91, 6269-6271.
22. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*, Pullman, B., Ed. Reidel Publishing Co.: Dordrecht, Holland, 1981; pp 331-342.
23. Berger, O.; Edholm, O.; Jahnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 1997, 72, 2002-2013.
24. Berisio, R.; Harms, J.; Schluenzen, F.; Zarivach, R.; Hansen, H. A.; Fucini, P.; Yonath, A. Structural insight into the antibiotic action of telithromycin against resistant mutants. *J Bacteriol* 2003, 185, 4276-9.
25. Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. The nucleic acid database: a comprehensive relational database of the three-dimensional structures of nucleic acids. *Biophys J* 1992, 63, 751-759.
26. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* 2000, 28, 235-242.
27. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. 2D Conformationally Sampled Pharmacophore: A Ligand-Based Pharmacophore To Differentiate delta Opioid Agonists from Antagonists. *J Am Chem Soc* 2003, 125, 3101-3107.
28. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. Conformationally sampled pharmacophore for peptidic delta opioid ligands. *J Med Chem* 2005, 48, 7773-80.
29. Bernard, D.; Coop, A.; MacKerell, A. D., Jr. Quantitative Conformationally Sampled Pharmacophore for  $\delta$  Opioid Ligands: Reevaluation of Hydrophobic Moieties Essential for Biological Activity. *J Med Chem* 2007, 50, 1799-1809.
30. Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. An Anisotropic Polarizable Water Model: Incorporation of All-Atom Polarizabilities into Molecular Mechanics Force Fields. *J Phys Chem* 1994, 98, 4180-4187.
31. Best, R. B.; Buchete, N. V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys J* 2008, 95, L07-9.

32. Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to helix-coil transition of polypeptides. *J Phys Chem B* 2009, 113, 9004-9015.
33. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* 2012, 8, 3257-3273.
34. Blondel, A.; Karplus, M. New formulation for derivatives of torsion angles and improper torsion angles in molecular mechanics: elimination of singularities. *J Comput Chem* 1996, 17, 1132-1141.
35. Boggiano, B.; Gleeson, M. Gastric acid inactivation of erythromycin stearate in solid dosage forms. *J Pharm Sci* 1976, 65, 497-502.
36. Bohm, H.-J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Computer-Aided Mol Des* 1992, 6, 61-78.
37. Bonnefoy, A.; Girard, A. M.; Agouridas, C.; Chantot, J. F. Ketolides lack inducibility properties of MLS(B) resistance phenotype. *J Antimicrob Chemother* 1997, 40, 85-90.
38. Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interaction by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* 1970, 19, 553-566.
39. Brandi, L.; Marzi, S.; Fabbretti, A.; Fleishcer, C.; Hill, W.; Gualerzi, C. O.; Lodmell, J. S. The translation initiation functions of IF2: targets for thiostrepton inhibition. *J Mol Biol* 2004, 335, 881-894.
40. Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* 2009, 25, 621-627.
41. Brooks, B. R.; Brooks III, C. L.; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J Comput Chem* 2009, 30, 1545-1614.
42. Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J Comput Chem* 1983, 4, 187-217.
43. Brooks III, C. L.; Brunger, A.; Karplus, M. Active site dynamics in protein molecules: a stochastic boundary molecular-dynamics approach. *Biopolymers* 1985, 24, 843-65.
44. Brooks III, C. L.; Karplus, M. Deformably stochastic boundaries in molecular dynamics. *J Chem Phys* 1983, 79, 6312-6325.
45. Brooks III, C. L.; Karplus, M. Solvent effects on protein motion and protein effects on solvent motion. Dynamics of the active site region of lysozyme. *J Mol Biol* 1989, 208, 159-81.
46. Brunger, A.; Brooks III, C. L.; Karplus, M. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett* 1984, 105, 495-500.

47. Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crysta* 1998, D54, 905-921.
48. Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. Cooperative effects in the structure of fluoride water clusters: Ab initio hybrid quantum mechanical/molecular mechanical model incorporating polarizable fluctuating charge solvent. *J Chem Phys* 1998, 109, 3077-3085.
49. Bryskier, A. Ketolides-telithromycin, an example of a new class of antibacterial agents. *Clin Microbiol Infect* 2000, 6, 661-9.
50. Buckingham, R. A. The classical equation of state of gaseous helium, neon and argon. *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* 1938, 168, 264-283.
51. Bulkley, D.; Innis, C. A.; Blaha, G.; Steitz, T. A. Revisiting the structures of several antibiotics bound to the bacterial ribosome. *Proc Natl Acad Sci U S A* 2010, 107, 17158-63.
52. Caflisch, A.; Miranker, A.; Karplus, M. Multiple copy simultaneous search and construction of ligands in binding sites: application to inhibitors of HIV-1 aspartic proteinase. *J Med Chem* 1993, 36, 2142-2167.
53. Caldwell, J.; Dang, L. X.; Kollman, P. A. Implementation of Nonadditive Intermolecular Potentials by Use of Molecular Dynamics: Development of a Water-Water Potential and Water-Ion Cluster Interactions. *J Am Chem Soc* 1990, 112, 9144-9147.
54. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 2000, 407, 340-8.
55. Champney, W. S.; Chittum, H. S.; Tober, C. L. A 50S ribosomal subunit precursor particle is a substrate for the ErmC methyltransferase in *Staphylococcus aureus* cells. *Curr Microbiol* 2003, 46, 453-60.
56. Champney, W. S.; Tober, C. L. Structure-activity relationships for six ketolide antibiotics. *Curr Microbiol* 2001, 42, 203-10.
57. Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J Med Chem* 1999, 42, 5100-5109.
58. Cheatham, T. E., 3rd; Case, D. A. Twenty-five years of nucleic acid simulations. *Biopolymers* 2013, 99, 969-77.
59. Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* 1999, 16, 845-861.
60. Chen, I. J.; Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of Lennard-Jones parameters for polar-neutral compounds. *J Comput Chem* 2002, 23, 199-213.
61. Chen, W.; Gilson, M. K.; Webb, S. P.; Potter, M. J. Modeling protein-ligand binding by mining minima. *J Chem Theory Comput* 2010, 6, 3540-3557.
62. Chiu, S.-W.; Pandit, S. A.; Scott, H. L.; Jakobsson, E. An improved united atom

- force field for simulation of mixed lipid bilayers. *J Phys Chem B* 2009, 113, 2748-2763.
63. Chowdary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D., Jr.; Roux, B. A polarizable force field of dipalmitoylphosphatidylcholine based on the classical drude model for molecular dynamics simulations of lipids. *J Phys Chem B* 2013, 117, 9142-9160.
64. Clancy, J.; Petitpas, J.; Dib-Hajj, F.; Yuan, W.; Cronan, M.; Kamath, A.; Bergeron, J.; Retsema, J. Molecular cloning and functional analysis of a novel macrolide-resistance determinant, *mefA*, from *Streptococcus pyogenes*. *Mol Microbiol* 1996, 22, 867-879.
65. Coimbatore Narayanan, B.; Westbrook, J.; Ghosh, S.; Petrov, A. I.; Sweeney, B.; Zirbel, C. L.; Leontis, N. B.; Berman, H. M. The Nucleic Acid Database: new features and capabilities. *Nucleic Acids Res* 2014, 42, D114-22.
66. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 1995, 117, 5179-5197.
67. Dang, L. X. Importance of Polarization Effects in Modeling Hydrogen Bond in Water Using Classical Molecular Dynamics Techniques. *J Phys Chem B* 1998, 102, 620-624.
68. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 1993, 98, 10089-10092.
69. Daura, X.; Mark, A. E.; van Gunsteren, W. F. Parametrization of Aliphatic CHN United Atoms of GROMOS96 Force Field. *J Comp Chem* 1998, 19, 535-547.
70. De Loof, H.; Nilsson, L.; Rigler, R. Molecular dynamics simulation of galanin in aqueous and nonaqueous solution. *J Am Chem Soc* 1992, 114, 4028-4035.
71. Demerdash, O.; Yap, E.-H.; Head-Gordon, T. Advanced Potential Energy Surfaces for Condensed Phase Simulations. *Annu Rev Phys Chem* 2014.
72. Deng, Y.; Roux, B. Hydration of amino acid side chains: nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J Phys Chem B* 2004, 108, 16567-16576.
73. Deng, Y.; Roux, B. Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J Chem Phys* 2008, 128, 1-8.
74. Denis, A.; Bretin, F.; Fromentin, C.; Bonnet, A.; Piltan, G.; Bonnefoy, A.; Agouridas, C. Beta-keto-ester chemistry and ketolides. Synthesis and antibacterial activity of 2-halogeno, 2-methyl and 2,3 enol-ether ketolides. *Bioorg Med Chem Lett* 2000, 10, 2019-22.
75. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; MacKerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* 2011, 32, 1929-43.
76. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; MacKerell, A. D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* 2012, 32, 1929-43.
77. Denoya, C.; Dubnau, D. Mono- and dimethylating activities and kinetic studies of the *ermC* 23 S rRNA methyltransferase. *J Biol Chem* 1989, 264, 2615-24.
78. DeWitte, R. S.; Shakhnovich, E. I. SMOG: de novo design method based on

- simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. *J Am Chem Soc* 1996, 118, 11733-11744.
79. Dey, F.; Caflisch, A. Fragment-based de novo ligand design by multiobjective evolutionary optimization. *J Chem Inf Model* 2008, 48, 679-690.
80. Di Nola, A.; Roccatano, D.; Berendsen, H. J. C. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins* 1994, 19, 174-182.
81. Douthwaite, S. Structure-activity relationships of ketolides vs. macrolides. *Clin Microbiol Infect* 2001, 7 Suppl 3, 11-7.
82. Douthwaite, S.; Hansen, L. H.; Mauvais, P. Macrolide-ketolide inhibition of MLS-resistant ribosomes is improved by alternative drug interaction with domain II of 23S rRNA. *Mol Microbiol* 2000, 36, 183-93.
83. Douthwaite, S.; Jalava, J.; Jakobsen, L. Ketolide resistance in *Streptococcus pyogenes* correlates with the degree of rRNA dimethylation by Erm. *Mol Microbiol* 2005, 58, 613-22.
84. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J Comput Chem* 2003, 24, 1999-2012.
85. Dubnau, D. Translational attenuation: the regulation of bacterial resistance to the macrolide-lincosamide-streptogramin B antibiotics. *CRC Crit Rev Biochem* 1984, 16, 103-32.
86. Dunkle, J. A.; Xiong, L.; Mankin, A. S.; Cate, J. H. Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc Natl Acad Sci U S A* 2010, 107, 17152-7.
87. ECDC, E. C. f. D. P. a. C.; EMeA, E. M. A. Technical report. The bacterial challenge: time to react. 2009, 1-54.
88. Ertle, P.; Jelfs, S.; Muhlbacher, J.; Schuffenhauer, A.; Selzer, P. Quest for the rings. In silico exploration of ring universe to identify novel bioactive heteroaromatic scaffolds. *J Med Chem* 2006, 49, 4568-4573.
89. Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh Ewald method. *J Chem Phys* 1995, 103, 8577-8593.
90. Evrard-Todeschi, N.; Gharbi-Benarous, J.; Gaillet, C.; Verdier, L.; Bertho, G.; Lang, C.; Parent, A.; Girault, J.-P. Conformations in solution and bound to bacterial ribosomes of ketolides, HMR 347 (telithromycin) and RU 72366: a new class of highly potent antibacterials. *Bioorg Med Chem* 2000, 8, 1579-1597.
91. Fadda, E.; Woods, R. J. Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov Today* 2010, 15, 596-609.
92. Feig, M.; Brooks, C. L., III. Recent advances in the development and application of implicit solvent models in biomolecular simulations. *Curr Opin Struct Biol* 2004, 14, 217-224.
93. Feig, M.; MacKerell, A. D., Jr.; Brooks, C. L., III. Force field influence on the observation of p-helical protein structures in molecular dynamics simulations. *J Phys Chem B* 2003, 107, 2831-2836.
94. Feller, S.; Zhang, Y.; Pasto, R.; Brooks, B. Constant pressure molecular dynamics

- simulation: The langevin piston method. *J Chem Phys* 1995, 103, 4613-4621.
95. Feller, S. E.; MacKerell, A. D., Jr. An improved empirical potential energy function for molecular simulations of phospholipids. *J Phys Chem B* 2000, 104, 7510-7515.
96. Fiese, E.; Steffen, S. Comparison of the acid stability of azithromycin and erythromycin A. *J Antimicrob Chemother* 1990, 25, 39-47.
97. Fink, T.; Bruggesser, H.; Reymond, J. L. Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew Chem Int Ed Engl* 2005, 44, 1504-8.
98. Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 2000, 21, 86-104.
99. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* 1996, 274, 1367-71.
100. Fox, T.; Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J Phys Chem B* 1998, 102, 8070-8079.
101. French, A. D.; Johnson, G. P. Computerized molecular modeling of carbohydrates. *Methods Mol Biol* 2011, 715, 21-42.
102. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004, 47, 1739-1749.
103. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven Jr., T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, K.; Kitao, O.; Nakai, H.; Klene, M.; Li, T. W.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.04*; Gaussian, Inc.: Pittsburgh, PA, 2003.
104. Ganoza, M. C.; Kiel, M. C. A ribosomal ATPase is a target for hygromycin B inhibition on *Escherichia coli* ribosomes. *Antimicrob Agents Chemother* 2001, 45, 2813-9.
105. Garcia, A. E.; Sanbonmatsu, K. Y. alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci U S A* 2002, 99, 2782-2787.
106. Ge, X.; Roux, B. Absolute binding free energy calculations of sparsomycin analogs to the bacterial ribosome. *J Phys Chem B* 2010, 114, 9525-39.
107. Ge, X.; Roux, B. Calculation of the standard binding free energy of sparsomycin

- to the ribosomal peptidyl-transferase P-site using molecular dynamics simulations with restraining potentials. *J Mol Recognit* 2010, 23, 128-41.
108. Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys J* 1997, 1997, 1047-1069.
109. Glassford, I.; Lee, M.; Wagh, B.; Velvadapu, V.; Paul, T.; Sandelin, G.; DeBrosse, C.; Klepacki, D.; Small, M. C.; A.D. MacKerell, J.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4-desmethyl telithromycin. *ACS Med Chem Lett* 2014, 5, 1021-1026.
110. Glassford, I.; Tejjaro, C. N.; Daher, S. S.; Weil, A.; Redhu, S. K.; Small, M. C.; Buttaro, B.; Nicholson, A. W.; Alexander D. MacKerell, J.; Cooperman, B. S.; Andrade, R. B. Ribosome-templated azide-alkyne cycloadditions: synthesis of potent macrolide antibiotics by in situ Click chemistry. in preparation.
111. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* 1985, 28, 849-857.
112. Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognition* 1996, 9, 1-5.
113. Gough, C. A., DeBolt, S. E. and Kollman, P. A. Derivation of Fluorine and Hydrogen Atom Parameters Using Liquid Simulations. *J Comput Chem* 1992, 13, 963-970.
114. Grimster, N. P.; Stump, B.; Fotsing, J. R.; Weide, T.; Talley, T. T.; Yamauchi, J. G.; Nemezc, A.; Kim, C.; Ho, K. Y.; Sharpless, K. B.; Taylor, P.; Fokin, V. V. Generation of candidate ligands for nicotinic acetylcholine receptors via in situ click chemistry with a soluble acetylcholine binding protein template. *J Am Chem Soc* 2012, 134, 6732-6740.
115. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* 2008, 29, 2543-64.
116. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell Jr., A. D. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* 2008, 29, 2543-64.
117. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* 2009, 5, 2353-2370.
118. Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; MacKerell Jr., A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput* 2009, 5, 2353-2370.
119. Guvench, O.; MacKerell, A. D., Jr. Automated conformational energy fitting for force-field development. *J Mol Model* 2008, 14, 667-79.
120. Guvench, O.; MacKerell, A. D., Jr. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol Biol* 2008, 443, 63-88.
121. Guvench, O.; MacKerell, A. D., Jr. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput Biol* 2009, 5, e1000435.
122. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe,

- K.; Foster, T. J.; Jamison, F. W., 2nd; MacKerell, A. D., Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* 2011, 7, 3162-3180.
123. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison II, F. W.; MacKerell Jr., A. D. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* 2011, 7, 3162-3180.
124. Hagler, A. T.; Stern, P. S.; Sharon, R.; Becker, J. M.; Naider, F. Computer Simulation of the Conformational Properties of Oligopeptides. Comparison of Theoretical Methods and Analysis of Experimental Results. *J Am Chem Soc* 1979, 101, 6842-6852.
125. Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* 2007, 6, 211-9.
126. Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. NMR-based screening in drug discovery. *Quarterly Reviews in Biophysics* 1999, 32, 211-240.
127. Halgren, T. A. Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters. *J Am Chem Soc* 1992, 114, 7827-7843.
128. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J Comput Chem* 1996, 17, 490-519.
129. Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J Comput Chem* 1996, 17, 520-552.
130. Halgren, T. A. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94. *J Comput Chem* 1996, 17, 553-586.
131. Halgren, T. A. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules. *J Comput Chem* 1996, 17, 616-641.
132. Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J Comput Chem* 1999, 20, 730-748.
133. Halgren, T. A. New method for fast and accurate binding site identification analysis. *Chem Biol Drug Des* 2007, 69, 146-148.
134. Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J Chem Inf Model* 2001, 41, 856-864.
135. Hansch, C. A quantitative approach to biochemical structure-activity relationships. *Acc Chem Res* 1969, 2, 232-239.
136. Hansen, H. S.; Hunenberger, P. H. A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers. *J Comput Chem* 2011, 32, 998-1032.
137. Hansen, J. L.; Ippolito, J. A.; Ban, N.; Nissen, P.; Moore, P. B.; Steitz, T. A. The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol Cell* 2002, 10, 117-28.
138. Hansen, L. H.; Mauvais, P.; Douthwaite, S. The macrolide-ketolide antibiotic

- binding site is formed by structures in domains II and V of 23S ribosomal RNA. *Mol Microbiol* 1999, 31, 623-31.
139. Hansmann, U. H. E. Parallel tempering algorithm for conformational studies of biological molecules. *Chem Phys Lett* 1997, 281, 140-150.
140. Hansson, T.; Marelius, J.; Aqvist, J. Ligand binding affinity prediction by linear interaction energy methods. *J Computer-Aided Mol Des* 1998, 12, 27-35.
141. Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; Jr., A. D. M.; Roux, B. Atomic Level Anisotropy in the Electrostatic Modeling of Lone Pairs for a Polarizable Force Field Based on the Classical Drude Oscillator. *J Chem Theory Comput* 2006, 2, 1587-1597.
142. Harder, E.; Anisimov, V. M.; Whitfield, T.; MacKerell, A. D., Jr.; Roux, B. Understanding the dielectric properties of liquid amides from a polarizable force field. *J Phys Chem B* 2008, 112, 3509-21.
143. Hardy, L. W.; Malikayil, A. The impact of structure-guided drug design on clinical agents. *Cur Drug Disc* 2003.
144. Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D., Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J Chem Theory Comput* 2012, 8, 348-362.
145. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* 2009, 5, 1315-1327.
146. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* 2009, 113, 12466-76.
147. Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. Supporting information for CHARMM additive all-atom force field for acyclic polyalcohols, acyclic carbohydrates and inositol. *J Chem Theory Comput* 2009, 5, available online.
148. Hatcher, E.; Guvench, O.; MacKerell Jr., A. D. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates and Inositol. *J Chem Theory Comput* 2009, 5, 1315-1327.
149. Hatcher, E.; Guvench, O.; MacKerell Jr., A. D. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* 2009, 113, 12466-76.
150. He, X.; Lopes, P. E.; MacKerell, A. D., Jr. Polarizable empirical force field for acyclic polyalcohols based on the classical drude oscillator. *Biopolymers* 2013, 99, 724-38.
151. Head, M. S.; Given, J. A.; Gilson, M. K. "Mining Minima": Direct Computation of Conformational Free Energy. *J Phys Chem A* 1997, 101, 1609-1618.
152. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J Comp Chem* 1997, 18, 1463-1472.
153. Hirose, T.; Sunazuka, T.; Sugawara, A.; Endo, A.; Iguchi, K.; Yamamoto, T.; Ui, H.; Shiomi, K.; Watanabe, T.; Sharpless, K. B.; Omura, S. Chitinase inhibitors: extraction of the active framework from natural argifin and use of in situ click chemistry. *J Antibiotic* 2009, 62, 277-282.
154. Homeyer, N.; Gohlke, H. Free Energy Calculations by the Molecular Mechanics

- Poisson–Boltzmann Surface Area Method. *Molecular Informatics* 2012, 31, 114-122.
155. Hoover, W. G.; Berendsen, H.; Postma, J. M.; Gunsteren, W.; DiNola, A.; Haak, J. R. Canonical dynamics: Equilibrium phase-space distributions  
Molecular dynamics with coupling to an external bath. *Phys Rev A* 1985, 31, 1695-1697.
156. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-EW. *J Chem Phys* 2004, 120, 9665.
157. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 2006, 65, 712-25.
158. Hu, J.; Ma, A.; Dinner, A. R. Monte Carlo simulations of biomolecules: The MC module in CHARMM. *J Comput Chem* 2006, 27, 203-16.
159. Huang, Y. M.; Chen, W.; Potter, M. J.; Chang, C. E. Insights from free-energy calculations: protein conformational equilibrium, driving forces, and ligand-binding modes. *Biophys J* 2012, 103, 342-51.
160. Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics *J Molec Graphics* 1996, 14, 33-38.
161. Hwang, M. W.; Stockfisch, T. P.; Hagler, A. T. Derivation of class II force fields. 2. Derivation and characterization of a class II force field ,CFF93, for the alkyl functional group and alkane molecules. *J Am Chem Soc* 1994, 116, 2515-2525.
162. Iczkowski, R. P.; Margrave, J. L. Electronegativity. *J Am Chem Soc* 1961, 83, 3547-3551.
163. Im, W.; Berneche, S.; Roux, B. Generalized solvent boundary potential for computer simulations. *J Chem Phys* 2001, 114, 2924-2937.
164. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: 1. Method. *J Comput Chem* 2000, 21, 132-146.
165. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 2002, 23, 1623-41.
166. Jana, M.; Alexander D. MacKerell, J. CHARMM Drude polarizable force field for aldehydofuranoses and methyl-aldehydofuranosides. *J Phys Chem B* 2015, just accepted.
167. Jayaram, B.; Beveridge, D. Grand canonical Monte Carlo simulations on aqueous solutions of sodium chloride and sodium DNA: excess chemical potentials and sources of nonideality in electrolyte and polyelectrolyte solutions. *J Phys Chem* 1991, 95, 2506-2516.
168. Jiang, W.; Hodoscsek, M.; Roux, B. Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics (FEP/REMD). *J Chem Theory Comput* 2009, 5, 2583-2588.
169. Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc Natl Acad Sci U S A* 2008, 105, 6290-5.
170. Jiao, D.; Zhang, J.; Duke, R. E.; Li, G.; Schnieders, M. J.; Ren, P. Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable

- potential. *J Comput Chem* 2009, 30, 1701-11.
171. Jojart, B.; Martinek, T. A. Performance of the general amber force field in modeling aqueous POPC membrane bilayers. *J Comput Chem* 2007, 28, 2051-8.
172. Jones, G.; Willett, P.; R.C., G.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 1997, 267, 727-748.
173. Jones-Hertzog, D. K.; Jorgensen, W. L. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with linear response method. *J Med Chem* 1997, 40, 1539-1549.
174. Jorgensen, W. L. Optimized Intermolecular Potential Functions for Liquid Alcohols. *J Phys Chem* 1986, 90, 1276-1284.
175. Jorgensen, W. L. The many roles of computation in drug discovery. *Science* 2004, 303, 1813-1818.
176. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983, 79, 926-936.
177. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 1996, 118, 11225-11236.
178. Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Function for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J Amer Chem Soc* 1988, 110, 1657-1666.
179. Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model. *J Phys Chem* 1994, 98, 13077-13082.
180. Kaminski, G.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J Phys Chem B* 2001, 105, 6474-6487.
181. Kannan, K.; Vazquez-Laslop, N.; Mankin, A. S. Selective protein synthesis by ribosomes with a drug-obstructed exit tunnel. *Cell* 2012, 151, 508-520.
182. Kannan, S.; Zacharias, M. Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing potential. *Proteins* 2007, 66, 697-706.
183. Katz, L.; Ashley, G. W. Translation and protein synthesis: macrolides. *Chem Rev* 2005, 105, 499-528.
184. Kirillov, S.; Porse, B. T.; Vester, B.; Woolley, P.; Garrett, R. A. Movement of the 3'-end of tRNA through the peptidyl transferase centre and its inhibition by antibiotics. *FEBS Lett* 1997, 406, 223-33.
185. Kirschner, K. N.; Lins, R. D.; Maass, A.; Soares, T. A. A Glycam-Based Force Field for Simulations of Lipopolysaccharide Membranes: Parametrization and Validation. *J Chem Theory Comput* 2012, 8, 4719-4731.
186. Kirschner, K. N.; Woods, R. J. Solvent interactions determine carbohydrate conformation. *Proc Natl Acad Sci U S A* 2001, 98, 10541-5.
187. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: a generalizable biomolecular

- force field. Carbohydrates. *J Comput Chem* 2008, 29, 622-55.
188. Klauda, J. B.; Brooks, B. R.; MacKerell, A. D., Jr.; Venable, R. M.; Pastor, R. W. An Ab Initio Study on the Torsional Surface of Alkanes and its Effect on Molecular Simulations of Alkanes and a DPPC Bilayer. *J Phys Chem B* 2005, 109, 5300-5311.
189. Klauda, J. B.; Pastor, R. W.; Brooks, B. R. Adjacent gauche stabilization in linear alkanes: implications for polymer models and conformational analysis. *J Phys Chem B Letters* 2005, 109, 15684-15686.
190. Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 2010, 114, 7830-43.
191. Kolb, H.; Sharpless, K. The growing impact of click chemistry on drug discovery. *Drug Discov Today* 2003, 8, 1128-1137.
192. Kolb, H. C.; Sharpless, K. B. The growing impact of click chemistry on drug discovery. *Drug Discov Today* 2003, 8, 1128-1137.
193. Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem Rev* 1993, 93, 2395-2417.
194. Kollman, P. A. Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc Chem Res* 1996, 29, 461-469.
195. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc Chem Res* 2000, 33, 889-897.
196. Kolter, R.; Yanofsky, C. Attenuation in amino acid biosynthesis operons. *Annu Rev Genet* 1982, 16, 113-134.
197. Kony, D.; Damm, W.; Stoll, S.; Van Gunsteren, W. F. An improved OPLS-AA force field for carbohydrates. *J Comput Chem* 2002, 23, 1416-29.
198. Kool, E. T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu Rev Biophys Biomol Struct* 2001, 30, 1-22.
199. Kostopoulou, O. N.; Petropoulos, A. D.; Dinos, G. P.; Choli-Papadopoulou, T.; Kalpaxis, D. L. Investigating the entire course of telithromycin binding to *Escherichia coli* ribosomes. *Nucleic Acids Res* 2012, 40, 5078-87.
200. Krasiński, A.; Radić, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In Situ Selection of Lead Compounds by Click Chemistry: Target-Guided Optimization of Acetylcholinesterase Inhibitors. *J Am Chem Soc* 2005, 127, 6686-6692.
201. Kresken, M.; Henrichfreise, B.; Bagel, S.; Brauers, J.; Wiedemann, B. High prevalence of the ermB gene among erythromycin-resistant streptococcus pneumoniae isolates in Germany during the winter of 2000-2001 and in vitro activity of telithromycin. *Antimicrob Agents Chemother* 2004, 48, 3193-5.
202. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA-a self-parameterizing force field. *Proteins: Structure, Function, and Genetics* 2002, 47, 393-402.
203. Kuczera, K.; Wiorkiewicz, J. K.; Karplus, M. MOLVIB: Program for the Analysis

- of Molecular Vibrations, CHARMM, Harvard University: 1993.
204. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J Mol Biol* 1982, 161, 269-288.
205. Kunz, A.-P. E.; van Gunsteren, W. F. Development of a nonlinear classical polarization model for liquid water and aqueous solution: COS/D. *J Phys Chem A* 2009, 113, 11570-11579.
206. Kwak, J.-H.; Choi, E.-C.; Weisblum, B. Transcriptional attenuation control of ermK, a macrolide-lincosamide-streptogramin B resistance determinant from *Bacillus licheniformis*. *J Bacteriol* 1991, 173, 4725-4735.
207. Lagüe, P.; Pastor, R. W.; Brooks, B. R. Pressure-based long-range correction for lennard-jones interactions in molecular dynamics simulations: applications to alkanes and interfaces. *J Phys Chem B* 2004, 108, 363-368.
208. Lai, C. J.; Weisblum, B. Altered methylation of ribosomal RNA in an erythromycin-resistant strain of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 1971, 68, 856-60.
209. Lakkaraju, S. K.; Raman, E. P.; Yu, W.; MacKerell, A. D., Jr. Free energies of organic solutes in aqueous and heterogeneous environments using grand canonical monte carlo-molecular dynamics simulations. *J Chem Theory Comput* 2014, 10, 2281-2290.
210. Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; AD MacKerell, J. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem Phys Lett* 2006, 418, 245-249.
211. Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J Chem Phys* 2003, 119, 5185-5197.
212. Lamoureux, G.; Roux, B. Modeling Induced Polarization with Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J Chem Phys* 2003, 119, 3025-3039.
213. Langevin, P. Sur la theorie du mouvement brownien. *C R Acad Sci (Paris)* 1908, 146, 530-533.
214. Langley, D. R. Molecular dynamics simulations of environment and sequence dependent DNA conformation: The development of the BMS nucleic acid force field and comparison with experimental results. *J Biomol Struct Dyn* 1998, 16, 487-509.
215. Lapatto, P.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. X-Ray Analysis of HIV-1 Proteinase at 2.7 Å Resolution Confirms Structural Homology among Retroviral Enzymes. *Nature* 1989, 342, 299-302.
216. Lemons, D. S.; Gythiel, A. Paul Langevin's 1908 paper "On the theory of brownian moetion" ["Sur la theorie du mouvement brownien," *C.R. Acad. Sci. (Paris)* 146, 530-533 (1908)]. *Am J Phys* 1997, 65, 1079-1081.
217. LeTourneau, N.; Vimal, P.; Klepacki, D.; Mankin, A.; Melman, A. Synthesis and antibacterial activity of desosamine-modified macrolide derivatives. *Bioorg Med Chem Lett* 2012, 22, 4575-8.
218. Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radic, Z.; Carlier, P. R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. Click chemistry in situ: Acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. *Angew Chem Int Ed* 2002, 41, 1053.

219. Lexa, K. W.; Carlson, H. A. Full protein flexibility is essential for proper hot-spot mapping. *J Am Chem Soc* 2011, 133, 200-202.
220. Li, D.-W.; Bruschiweiler, R. NMR-based protein potentials. *Angew Chem Int Ed Engl* 2010, 49, 6778-6780.
221. Lin, B.; Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Kirkwood-Buff analysis of aqueous N-methylacetamide and acetamide solutions modeled by the CHARMM additive and Drude polarizable force fields. *Journal of Chemical Physics* 2013, 139, 084509.
222. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 2010, 78, 1950-8.
223. Lins, R. D.; Hunenberger, P. H. A new GROMOS force field for hexopyranose-based carbohydrates. *J Comput Chem* 2005, 26, 1400-12.
224. Liu, M.; Douthwaite, S. Activity of the ketolide telithromycin is refractory to Erm monomethylation of bacterial rRNA. *Antimicrob Agents Chemother* 2002, 46, 1629-33.
225. Llano-Sotelo, B.; Dunkle, J.; Klepacki, D.; Zhang, W.; Fernandes, P.; Cate, J. H.; Mankin, A. S. Binding and action of CEM-101, a new fluoroketolide antibiotic that inhibits protein synthesis. *Antimicrob Agents Chemother* 2010, 54, 4961-70.
226. Llanta, E.; Ando, K.; Rey, R. Fluctuating charge study of polarization effects in chlorinated organic liquids. *J Phys Chem B* 2001, 105, 7783-7791.
227. Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr. Force Field for Peptides and Proteins based on the Classical Drude Oscillator. *J Chem Theory Comput* 2013, 9, 5430-5449.
228. Lopes, P. E.; Lamoureux, G.; MacKerell, A. D., Jr. Polarizable empirical force field for nitrogen-containing heteroaromatic compounds based on the classical Drude oscillator. *J Comput Chem* 2009, 30, 1821-38.
229. Lopes, P. E.; Lamoureux, G.; Roux, B.; MacKerell, A. D., Jr. Polarizable empirical force field for aromatic compounds based on the classical drude oscillator. *J Phys Chem B* 2007, 111, 2873-85.
230. Lopes, P. E.; Roux, B.; MacKerell, A. D., Jr. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. *Theory and applications. Theor Chem Acc* 2009, 124, 11-28.
231. Lopes, P. E. M.; Harder, E.; Roux, B.; MacKerell, J., A.D. Formalisms for the explicit inclusion of electronic polarizability in molecular modeling and dynamics studies. In *Multi-scale Quantum Models for Biocatalysis: Modern Techniques and Applications*, York, D.; Lee, T.-S., Eds. 2009; Vol. 7.
232. Lynch, G. C.; Pettitt, B. M. Semi-grand canonical molecular dynamics simulation of bovine pancreatic trypsin inhibitor. *Chem Phys* 2000, 258, 405-413.
233. MacKerell, A. D., Jr. Influence of magnesium ions on duplex DNA structural, dynamic, and solvation properties. *J Phys Chem B* 1997, 101, 646-650.
234. MacKerell, A. D., Jr. Protein Force Fields. In *Encyclopedia of Computational Chemistry*, Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; H.F. Schaefer, I.; Schreiner, P. R., Eds. John Wiley & Sons: Chichester, 1998; Vol. 3, pp 2191-2200.
235. MacKerell, A. D., Jr. Atomistic Models and Force Fields. In *Computational Biochemistry and Biophysics*, Becker, O. M.; MacKerell, A. D., Jr.; Roux, B.; Watanabe,

- M., Eds. Marcel Dekker, Inc.: New York, 2001; pp 7-38.
236. MacKerell, A. D., Jr. Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* 2004, 25, 1584-604.
237. MacKerell, A. D., Jr.; Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* 2000, 21, 105-120.
238. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Josephy-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 1998, 102, 3586-3616.
239. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 2004, 25, 1400-15.
240. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 2004, 126, 698-699.
241. MacKerell, A. D., Jr.; Karplus, M. Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids. *J Phys Chem* 1991, 95, 10559-10560.
242. MacKerell, A. D., Jr.; Wiorkiewicz-Kuczera, J.; Karplus, M. An all-atom empirical energy function for the simulation of nucleic acids. *J Am Chem Soc* 1995, 117, 11946-11975.
243. MacKerell Jr., A. D.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 2000, 56, 257-65.
244. MacKerell Jr., A. D.; Bashford, D. B. M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J. G., H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I., W.E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamic studies of protein. *J Phys Chem B* 1998, 102, 3586-3616.
245. MacKerell Jr., A. D.; Feig, M.; Brooks III, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 2004, 25, 1400-15.
246. MacKerell Jr., A. D.; Feig, M.; Brooks III, C. L. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 2004, 126, 698-9.
247. MacKerell Jr., A. D.; Nilsson, L. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr Opin Struct Biol* 2008, 18, 194-9.
248. Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 2000, 112, 8910-8922.

249. Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J Chem Theory Comput* 2011, 7, 4026-4037.
250. Mallajosyula, S. S.; Adams, K. M.; Barchi, J. J.; MacKerell, A. D., Jr. Conformational determinants of the activity of antiproliferative factor glycopeptide. *J Chem Inf Model* 2013, 53, 1127-1137.
251. Mallajosyula, S. S.; Guvench, O.; Hatcher, E.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Phosphate and Sulfate Linked to Carbohydrates. *J Chem Theory Comput* 2012, 8, 759-776.
252. Mallajosyula, S. S.; MacKerell, A. D., Jr. Influence of solvent and intramolecular hydrogen bonding on the conformational properties of o-linked glycopeptides. *J Phys Chem B* 2011, 115, 11215-29.
253. Manetsch, R.; Krasinski, A.; Radic, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In situ click chemistry: enzyme inhibitors made to their own specifications. *J Am Chem Soc* 2004, 126, 12809-12818.
254. Mangoni, M.; Roccatano, D.; Di Nola, A. Docking flexible ligands to flexible receptors in solution by molecular dynamics simulation. *Proteins: Structure, Function, and Genetics* 1999, 35, 153-162.
255. Mao, J. C.-H.; Putterman, M. The intermolecular complex of erythromycin and ribosome. *J Mol Biol* 1969, 44, 347-361.
256. Maple, J. R.; Dinur, U.; Hagler, A. T. Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc Natl Acad Sci U S A* 1988, 85, 5350-5354.
257. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 2007, 111, 7812-7824.
258. Martin, Y. C. Challenges and prospects for computational aids to molecular diversity. *Perspectives in Drug Discovery and Design* 1997, 7-8, 159-172.
259. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Molecular Physics* 1996, 87, 1117-1157.
260. Massova, I.; Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in Drug Discovery and Design* 2000, 18, 113-135.
261. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. DREIDING: a generic force field for molecular simulations. *J Phys Chem* 1990, 94, 8897-8909.
262. McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* 1977, 267, 585-590.
263. Melman, A.; Mankin, A.
264. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; H., T. A. Equation of state calculations by fast computing machines. *J Chem Phys* 1953, 21, 1087-1092.
265. Mezei, M. A cavity-biased (T,V,m) Monte Carlo method for the computer simulation of fluids. *Mol Phys* 1980, 40, 901-906.
266. Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P. Influence of stacking on hydrogen bonding: quantum chemical study on pyridine-benzene model complexes. *J Phys Chem A* 2004, 108, 6038-6044.

267. Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P. Influence of the pi-pi interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. *Nucleic Acids Res* 2005, 33, 1779-89.
268. Miller, B. T.; Singh, R. P.; Klauda, J. B.; Hodoscek, M.; Brooks, B. R.; Woodcock, H. L., III. CHARMMing: a new, flexible web portal for CHARMM. *J Chem Inf Model* 2008, 48, 1920-1929.
269. Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. Structure of Complex of Synthetic HIV-1 Protease with a Substrate-Based Inhibitor at 2.3 Å Resolution. *Science* 1989, 246, 1149-1152.
270. Miranker, A.; Karplus, M. Functionality Maps of Binding Sites: A Multiple Copy Simultaneous Search Method. *Proteins* 1991, 11, 29-34.
271. Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. H.; Kolb, H. C. In situ click chemistry: enzyme-generated inhibitors of carbonic anhydrase II. *Angew Chem Int Ed* 2004, 44, 116-120.
272. MOE, M. O. E. Chemical Computing Group. 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A, 2R7, 2012 2012.
273. Morita, A.; Kato, S. An ab initio analysis of medium perturbation on molecular polarizabilities. *Journal of Chemical Physics* 1999, 110, 11987-11998.
274. Morris, G. M.; Goodsell, D. S.; Pique, M. E.; Lindstrom, W.; Huey, R.; Forli, S.; Hart, W. E.; Halliday, S.; Belew, R.; Olson, A. J. AutoDock Version 4.2 Updated for version 4.2.5. 2012.
275. Mulliken, R. S. A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. *J Chem Phys* 1934, 2, 782-793.
276. Murphy, R. B.; Philipp, D. M.; Friesner, R. A. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale model of chemistry in protein environments. *J Comput Chem* 2000, 21, 1442-1457.
277. Namelikonda, N. K.; Manetsch, R. Sulfo-click reaction via in situ generated thioacids and its application in kinetic target-guided synthesis. *Chem Commun* 2012, 48, 1526-1528.
278. Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J Chem Phys* 1996, 105, 1902-1919.
279. Nielsen, A. K.; Douthwaite, S.; Vester, B. Negative in vitro selection identifies the rRNA recognition motif for ErmE methyltransferase. *Rna* 1999, 5, 1034-41.
280. Nilsson, L. Efficient table lookup without inverse square roots for calculation of pair wise atomic interactions in classical simulations. *J Comput Chem* 2009, 30, 1490-8.
281. Nilsson, L.; Karplus, M. Empirical energy functions for energy minimization and dynamics of nucleic acids. *J Comput Chem* 1986, 7, 591-616.
282. Norberg, J.; Nilsson, L. Constant pressure molecular dynamics simulations of the dodecamers: d(GCGCGCGCGCGC)<sub>2</sub> and r(GCGCGCGCGCGC)<sub>2</sub>. *J Chem Phys* 1996, 104, 6052-6057.
283. Norberg, J.; Nilsson, L. Solvent influence on base stacking. *Biophys J* 1998, 74, 394-402.
284. Norberg, J.; Nilsson, L. On the truncation of long-range electrostatics in DNA.

- Biophys J 2000, 79, 1537-1553.
285. Nose, S.; Klein, M. Constant pressure molecular dynamics for molecular systems. *Molec Phys* 1983, 50, 1055-1076.
286. Novotny, G. W.; Jakobsen, L.; Andersen, N. M.; Poehlsgaard, J.; Douthwaite, S. Ketolide antimicrobial activity persists after disruption of interactions with domain II of 23S rRNA. *Antimicrob Agents Chemother* 2004, 48, 3677-83.
287. Oashi, T.; Ringer, A. L.; Raman, E. P.; MacKerell, A. D. Automated Selection of Compounds with Physicochemical Properties To Maximize Bioavailability and Druglikeness. *J Chem Inf Model* 2011, 51, 148-158.
288. Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 2004, 25, 1656-76.
289. Park, A. K.; Kim, H.; Jin, H. J. Phylogenetic analysis of rRNA methyltransferases, Erm and KsgA, as related to antibiotic resistance. *FEMS Microbiol Lett* 2010, 309, 151-162.
290. Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *J Am Chem Soc* 1983, 105, 7512-7516.
291. Parthasarathy, P.; Murthy, M. R. N. Analysis of temperature factor distribution in high-resolution protein structures. *Protein Science* 1997, 6, 2561-2567.
292. Patel, D. S.; He, X.; MacKerell, A. D., Jr. Polarizable empirical force field for hexopyranose monosaccharides based on the classical drude oscillator. *J Phys Chem B* 2015, 119, 637-652.
293. Patel, S.; Brooks, C. L., 3rd. CHARMM fluctuating charge force field for proteins: I Parameterization and application to bulk organic liquid simulations. *J Comput Chem* 2003, 25, 1-15.
294. Pearlman, D.; Charifson, P. Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *J Med Chem* 2001, 44, 3417-3423.
295. Pearson, R. G. Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc Natl Acad Sci U S A* 1986, 83, 8440-8441.
296. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Loughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J* 2007, 92, 3817-29.
297. Perez, C.; Ortiz, A. R. Evaluation of docking functions for protein-ligand docking. *J Med Chem* 2001, 44, 3768-3785.
298. Pfister, P.; Jenni, S.; Poehlsgaard, J.; Thomas, A.; Douthwaite, S.; Ban, N.; Bottger, E. C. The structural basis of macrolide-ribosome binding assessed using mutagenesis of 23S rRNA positions 2058 and 2059. *J Mol Biol* 2004, 342, 1569-81.
299. Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys J* 2011, 100, L47-9.
300. Poehlsgaard, J.; Douthwaite, S. The bacterial ribosome as a target for antibiotics. *Nat Rev Microbiol* 2005, 3, 870-81.
301. Poger, D.; van Gunsteren, W. F.; Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J Comput Chem* 2010, 31, 1117-25.
302. Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Advances in*

- Protein Chemistry 2003, 66, 27-85.
303. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N.; Johnson, M. E.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J Phys Chem B* 2010, 114, 2549-64.
304. Poulin-Kerstien, A. T.; Dervan, P. B. DNA-templated dimerization of hairpin polyamides. *J Am Chem Soc* 2003, 125, 15811-15821.
305. Price, D. J.; Brooks, C. L., 3rd. Modern protein force fields behave comparably in molecular dynamics simulations. *J Comput Chem* 2002, 23, 1045-57.
306. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives. *J Am Chem Soc* 1979, 101, 2550-2560.
307. Rais, R.; Acharya, C.; Tririya, G.; MacKerell, A. D., Jr.; Polli, J. E. Molecular switch controlling the binding of anionic bile acid conjugates to human apical sodium-dependent bile acid transporter. *J Med Chem* 2010, 53.
308. Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* 2010, 114, 12981-12994.
309. Raman, E. P.; Guvench, O.; MacKerell Jr., A. D. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* 2010, 114, 12981-94.
310. Raman, E. P.; Vanommeslaeghe, K.; Alexander D. MacKerell, J. Site-Specific Fragment Identification Guided by Single-Step Free Energy Perturbation Calculations. *J Chem Theory Comput* 2012, 8, 3513-3525.
311. Raman, E. P.; Yu, W.; Guvench, O.; MacKerell, A. D. Reproducing crystal binding modes of ligand functional groups using Site-Identification by Ligand Competitive Saturation (SILCS) simulations. *J Chem Inf Model* 2011, 51, 877-96.
312. Raman, E. P.; Yu, W.; Lakkaraju, S. K.; MacKerell, A. D., Jr. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J Chem Inf Model* 2013, 53, 3384-98.
313. Ransil, B. Studies in molecular structure. IV. Potential curve for the interaction of two helium atoms in single-configuration LCAO MO SCF approximation. *J Chem Phys* 1961, 34, 2109.
314. Rappe, A. K.; Casewit, C. J.; Colwel, K. S.; Goddard, W. A., III; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J Am Chem Soc* 1992, 114, 10024-10035.
315. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 1996, 261, 470-489.
316. Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J Comput Chem* 2010, 31, 797-810.
317. Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J Chem Theory Comput* 2012, 8, 3705-3723.
318. Ren, P.; Ponder, J. W. Consistent treatment of inter- and intramolecular

- polarization in molecular mechanics calculations. *J Comput Chem* 2002, 23, 1497-506.
319. Ribiero, A. A. S. T.; Horta, B. A. C.; de Alencastro, R. B. MKTOP: a program for automatic construction of molecular topologies. *J Braz Chem Soc* 2008, 19, 1433-1435.
320. Rick, S. W.; Berne, B. J. Dynamical fluctuating charge force fields: the aqueous solvation of amides. *J Am Chem Soc* 1996, 118, 672-679.
321. Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev Comp Chem* 2002, 18, 89-146.
322. Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J. Fluctuating Charge Force Fields for Aqueous Solutions. *Studies in Physical and Theoretical Chemistry* 1995, 83, 31-40.
323. Rizzo, R. C.; Jorgensen, W. L. OPLS all-atom model for amines: resolution of the amine hydration problem. *J Am Chem Soc* 1999, 121, 4827-4836.
324. Rodnina, M. V.; Savelsbergh, A.; Matassova, N. B.; Katunin, V. I.; Semenov, Y. P.; Wintermeyer, W. Thiostrepton inhibits the turnover but not the GTPase of elongation factor G on the ribosome. *Proc Natl Acad Sci U S A* 1999, 96, 9586-90.
325. Rodnina, M. V.; Stark, H.; Savelsbergh, A.; Wieden, H. J.; Mohr, D.; Matassova, N. B.; Peske, F.; Daviter, T.; Gualerzi, C. O.; Wintermeyer, W. GTPases mechanisms and functions of translation factors on the ribosome. *Biol Chem* 2000, 381, 377-87.
326. Rosso, L.; Gould, I. R. Structure and dynamics of phospholipid bilayers using recently developed general all-atom force fields. *J Comput Chem* 2008, 29, 24-37.
327. Roux, B.; Simonson, T. Implicit Solvent Models. *Biophys Chem* 1999, 78, 1-20.
328. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 1977, 23.
329. Salyers, A. A.; Amabile-Cuevas, C. F. Why are antibiotic resistance genes so resistant to elimination? *Antimicrob Agents Chemother* 1997, 41, 2321-5.
330. Sanbonmatsu, K. Y.; Joseph, S.; Tung, C. S. Simulating movement of tRNA into the ribosome during decoding. *Proc Natl Acad Sci U S A* 2005, 102, 15854-9.
331. Sanbonmatsu, K. Y.; Tung, C. S. High performance computing in biology: multimillion atom simulations of nanoscale systems. *J Struct Biol* 2007, 157, 470-80.
332. Savelyev, A.; MacKerell, A. D., Jr. All-atom polarizable force field for DNA based on the classical drude oscillator model. *J Comput Chem* 2014, 10, 1652-1664.
333. Schlenkrich, M.; Brinkman, J.; MacKerell, A. D., Jr.; Karplus, M. An Empirical Potential Energy Function for Phospholipids: Criteria for Parameter Optimization and Applications. In *Membrane Structure and Dynamics*, Merz, K. M.; Roux, B., Eds. Birkhauser: Boston, 1996; pp 31-81.
334. Schlunzen, F.; Harms, J. M.; Franceschi, F.; Hansen, H. A.; Bartels, H.; Zarivach, R.; Yonath, A. Structural basis for the antibiotic activity of ketolides and azalides. *Structure* 2003, 11, 329-38.
335. Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* 2001, 413, 814-21.
336. Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* 2011, 40, 843-56.

337. Schropp, B.; Tavan, P. The polarizability of point-polarizable water models: density functional theory/molecular mechanics results. *J Phys Chem B* 2008, 112, 6233-40.
338. Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E. Library design for fragment based screening. *Curr Top Med Chem* 2005, 5, 751-762.
339. Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J Comput Chem* 2001, 22, 1205-1218.
340. Schuttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 2004, 6, 1355-1363.
341. Scott, A. P.; Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* 1996, 100, 16502-16513.
342. Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; C. Ochsenfeld; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; Jr., R. A. D.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Voorhis, T. V.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; III, H. L. W.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; III, H. F. S.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M.; Gan, Z.; Zhao, Y.; Schultz, N. E.; Truhlar, D.; Epifanovsky, E.; Oana, M.; Q-Chem. Q-Chem 3.1; Q-Chem, Inc.: Pittsburgh, PA, 2007.
343. Sharpless, K. B.; Manetsch, R. In situ click chemistry: a powerful means for lead discovery. *Expert Opinion on Drug Discovery* 2006, 1, 525-538.
344. Shi, Y.; Wu, C.; Ponder, J. W.; Ren, P. Investigation of multipole electrostatics in hydration free energy calculations. *J Comput Chem* 2011, 32, 967-977.
345. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. The Polarizable Atomic Multipole-based AMOEBA Force Field for Proteins. *J Chem Theory Comput* 2013, 9, 4046-4063.
346. Shi, Y.; Zhu, C. Z.; Martin, S. F.; Ren, P. Probing the effect of conformational constraint on phosphorylated ligand binding to an SH2 domain using polarizable force field simulations. *J Phys Chem B* 2012, 116, 1716-27.
347. Shim, J.; Coop, A.; MacKerell, A. D., Jr. Consensus 3D model of mu-opioid receptor ligand efficacy based on a quantitative Conformationally Sampled Pharmacophore. *J Phys Chem B* 2012, 115, 7487-96.
348. Shim, J.; MacKerell, J. A. D. Computational ligand-based rational design: role of conformational sampling and force fields in model development. *Med Chem Comm* 2011, 2, 356-370.
349. Shirts, M. R.; Mobley, D. L. An introduction to best practices in free energy

- calculations. *Methods Mol Biol* 2013, 924, 271-311.
350. Shirts, M. R.; Pitner, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys* 2003, 119, 5740-5761.
351. Shoemaker, N. B.; Vlamakis, H.; Hayes, K.; Salyers, A. A. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl Environ Microbiol* 2001, 67, 561-8.
352. Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. Lead discovery using molecular docking. *Curr Opin Chemical Biol* 2002, 6, 439-446.
353. Siu, S. W. I.; Vacha, R.; Jungwirth, P.; Bockman, R. A. Biomolecular simulations of membranes: physical properties from different force fields. *J Chem Phys* 2008, 128, 125103.
354. Skjevik, A. A.; Madej, B. D.; Walker, R. C.; Teigen, K. LIPID11: a modular framework for lipid simulations using amber. *J Phys Chem B* 2012, 116, 11124-36.
355. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Jr. Computational methods in drug discovery. *Pharmacological Reviews* 2014, 66, 334-395.
356. Small, M. C.; Lopes, P.; Andrade, R. B.; Mackerell, A. D., Jr. Impact of ribosomal modification on the binding of the antibiotic telithromycin using a combined grand canonical monte carlo/molecular dynamics simulation approach. *PLoS Comput Biol* 2013, 9, e1003113.
357. Soares, T. A.; Hunenberger, P. H.; Kastenholtz, M. A.; Krautler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F. An improved nucleic acid parameter set for the GROMOS force field. *J Comput Chem* 2005, 26, 725-37.
358. Song, Q.; Qiu, Z.; Wang, H.; Xia, Y.; Shen, J.; Zhang, Y. The effect of methylation on the hydrogen-bonding and stacking interaction of nucleic acid bases. *Structural Chemistry* 2012, 24, 55-65.
359. Sothiselvam, S.; Liu, B.; Han, W.; Ramu, H.; Klepacki, D.; Atkinson, G. C.; Brauer, A.; Remm, M.; Tenson, T.; Schulten, K.; Vasquez-Laslop, N.; Mankin, A. S. Macrolide antibiotics allosterically predispose the ribosome for translation arrest. *PNAS* 2014, 111, 9804-9809.
360. Spöner, J.; Leszczynski, J.; Hobza, P. Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. *Nucleic Acids Sci* 2002, 61, 3-31.
361. Sprik, M.; Klein, M. L. A polarizable model for water using distributed charge sites. *J Chem Phys* 1988, 89, 7556-7560.
362. Srinivasan, J.; Cheatham, T. E., 3rd; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *J Am Chem Soc* 1998, 120, 9401-9409.
363. Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods of Long-Range Forces in Macromolecular Simulations. *J. Comp. Chem.* 1994, 15, 667-683.
364. Straatsma, T. P.; McCammon, J. A. Computational Alchemy. *Annu Rev Phys Chem* 1992, 43, 407-435.
365. Su, Y.; Gallicchio, E.; Das, K.; Arnold, E.; Levy, R. M. Linear interaction energy (LIE) models for ligand binding in implicit solvent: Theory and application to the binding of NNRTIs to HIV-1 reverse transcriptase. *J Chem Theory Comput* 2007, 3, 256-277.
366. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics methods for

- protein folding. *Chem Phys Lett* 1999, 314, 141-151.
367. Tan, Y. S.; Sledz, P.; Lang, S.; Stubbs, C. J.; Spring, D. R.; Abell, C.; Best, R. B. Using ligand-mapping simulations to design a ligand selectively targeting a cryptic surface pocket of polo-like kinase I. *Angew Chem Int Ed Engl* 2012, 51, 10078-10081.
368. TATFAR, T. t. o. a. r. Recommendations for future collaborations between the U.S. and E.U. 2011, 1-47.
369. Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J Computer-Aided Mol Des* 2002, 16, 151-166.
370. Tenson, T.; Lovmar, M.; Ehrenberg, M. The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *J Mol Biol* 2003, 330, 1005-14.
371. Tenson, T.; Mankin, A. Antibiotics and the ribosome. *Mol Microbiol* 2006, 59, 1664-77.
372. Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem Phys* 1981, 59, 341-350.
373. Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J Comput Phys* 1977, 23, 187-199.
374. Tu, D.; Blaha, G.; Moore, P. B.; Steitz, T. A. Structures of MLSBK antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance. *Cell* 2005, 121, 257-70.
375. Tuckerman, M. E.; Martyna, G. J. Understanding Modern Molecular Dynamics: Techniques and Applications. *J Phys Chem B* 2000, 104, 159-178.
376. Ulmschneider, J. P.; Ulmschneider, M. B. United atom lipid parameters for combination with the optimized potentials for liquid simulations all-atom force field. *J Chem Theory Comput* 2009, 5, 1803-1813.
377. Usary, J.; Champney, W. S. Erythromycin inhibition of 50S ribosomal subunit formation in *Escherichia coli* cells. *Mol Microbiol* 2001, 40, 951-62.
378. van Gunsteren, W. F.; Berendsen, H. J. C. Groningen Molecular Simulation (GROMOS) Library Manual. University of Groningen: Groningen, The Netherlands, 1987; p 1-221.
379. van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. Biomolecular Simulation: The GROMOS96 Manual and User Guide. BIOMOS b.v.: Zürich, 1996.
380. van Gunsteren, W. F.; Karplus, M. A Method for Constrained Energy Minimization of Macromolecules. *Journal of Computational Chemistry* 1980, 1, 266-274.
381. van Maaren, P. J.; van der Spoel, D. Molecular dynamics simulations of water within novel shell-model potentials. *J Phys Chem B* 2001, 105, 2618-2626.
382. Vanommeslaeghe, K.; AD MacKerell, J. Robustness in the fitting of molecular mechanics parameters. *J Comp Chem* 2015, in press.
383. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 2010, 31, 671-90.
384. Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model*

2012, 52, 3144-54.

385. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* 2012, 52, 3155-68.
386. Varekova, R. S.; Koca, J.; Zhan, C.-G. Complexity and convergence of electrostatic and van der Waals energies within PME and cutoff methods. *Int J Mol Sci* 2004, 5, 154-173.
387. Velvadapu, V.; Glassford, I.; Lee, M.; Paul, T.; DeBrosse, C.; Klepacki, D.; Small, M. C.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,10-didesmethyl telithromycin. *ACS Med Chem Lett* 2012, 3, 211-215.
388. Velvadapu, V.; Paul, T.; Wagh, B.; Klepacki, D.; Guvench, O.; MacKerell Jr., A. D.; Andrade, R. B. Desmethyl macrolide analogues to address antibiotic resistance: total synthesis and biological evaluation of 4,8,10-tridesmethyl telithromycin. *ACS Med Chem Lett* 2011, 2, 68-72.
389. Vester, B.; Douthwaite, S. Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrob Agents Chemother* 2001, 45, 1-12.
390. Vester, B.; Hansen, L. H.; Douthwaite, S. The conformation of 23S rRNA nucleotide A2058 determines its recognition by the ErmE methyltransferase. *Rna* 1995, 1, 501-9.
391. Villsen, I. D.; Vester, B.; Douthwaite, S. ErmE methyltransferase recognizes features of the primary and secondary structure in a motif within domain V of 23 S rRNA. *J Mol Biol* 1999, 286, 365-74.
392. Vorobyov, I.; Anisimov, V. M.; Greene, S.; Venable, R. M.; Moser, A.; Pastor, R. W.; MacKerell, A. D., Jr. Additive and Classical Drude Polarizable Force Fields for Linear and Cyclic Ethers. *J. Chem. Theory Comp.* 2007, 3, 1120-1133.
393. Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D., Jr. Polarizable empirical force field for alkanes based on the classical Drude oscillator. *J Phys Chem B* 2005, 109, 18988-18999.
394. Wagh, B.; Paul, T.; DeBrosse, C.; Klepacki, D.; Small, M. C.; A.D. MacKerell, J.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,8,10-tridesmethyl cethromycin. *ACS Med Chem Lett* 2013, 4.
395. Wagh, B.; Paul, T.; Glassford, I.; DeBrosse, C.; Klepacki, D.; Small, M. C.; MacKerell, A. D., Jr.; Andrade, R. B. Desmethyl macrolides: synthesis and evaluation of 4,8-didesmethyl telithromycin. *ACS Med Chem Lett* 2012, 3, 1013-1018.
396. Wallqvist, A.; Berne, B. J. Effective potentials for liquid water using polarizable and nonpolarizable models. *J Phys Chem* 1993, 97, 13841-13851.
397. Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000, 21, 1049-1074.
398. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 2006, 25, 247-60.
399. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J Comput Chem* 2004, 25, 1157-1174.
400. Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable

- Force Fields from Quantum Chemistry Data. *J Chem Theory Comput* 2013, 9, 452-460.
401. Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic improvement of a classical molecular model of water. *J Phys Chem B* 2013, 117, 9956-9972.
402. Wang, S.; Yang, C.-Y. Hydrophobic binding hot spots of Bcl-xL protein-protein interfaces by cosolvent molecular dynamics simulations. *ACS Med Chem Lett* 2011, 2, 280-284.
403. Warshel, A.; Levitt, M. Theoretical studies of enzymatic reactions: dielectric, electrostatic, and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 1976, 103, 227-249.
404. Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J Comput Chem* 1981, 2, 287-303.
405. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J Am Chem Soc* 1984, 106, 765-784.
406. Weisblum, B. Erythromycin resistance by ribosome modification. *Antimicrob Agents Chemother* 1995, 39, 577-85.
407. Weisblum, B. Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrob Agents Chemother* 1995, 39, 797-805.
408. Weisblum, B.; Siddhikol, C.; Lai, C. J.; Demohn, V. Erythromycin-inducible resistance in *Staphylococcus aureus*: requirements for induction. *J Bacteriol* 1971, 106, 835-47.
409. White, D. N. J. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *J Computer-Aided Mol Des* 1997, 11, 517-521.
410. Whiting, M.; Muldoon, J.; Lin, Y. C.; Silverman, S. M.; Lindstrom, W.; Olson, A. J.; Kolb, H. C.; Finn, M. G.; Sharpless, K. B.; Elder, J. H.; Fokin, V. V. Inhibitors of HIV-1 protease by using in situ click chemistry. *Angew Chem Int Ed* 2006, 45, 1435-1439.
411. Willand, N.; Desroses, M.; Toto, P.; Dirié, B.; Lens, Z.; Villeret, V.; Rucktooa, P.; Loch, C.; Baulard, A.; Deprez, B. Exploring Drug Target Flexibility Using in Situ Click Chemistry: Application to a Mycobacterial Transcriptional Regulator. *ACS Chemical Biology* 2010, 5, 1007-1013.
412. Wolter, N.; Smith, A. M.; Farrell, D. J.; Northwood, J. B.; Douthwaite, S.; Klugman, K. P. Telithromycin resistance in *Streptococcus pneumoniae* is conferred by a deletion in the leader sequence of erm(B) that increases rRNA methylation. *Antimicrob Agents Chemother* 2008, 52, 435-40.
413. Wolter, N.; Smith, A. M.; Low, D. E.; Klugman, K. P. High-level telithromycin resistance in a clinical isolate of *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 2007, 51, 1092-5.
414. Woo, H. J.; Dinner, A. R.; Roux, B. Grand canonical Monte Carlo simulations of water in protein environments. *J Chem Phys* 2004, 121, 6392-400.
415. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1.

- GLYCAM\_93 parameter development. *J Phys Chem* 1995, 99, 3832-3846.
416. Wright, G. D. Molecular mechanisms of antibiotic resistance. *Chem Commun (Camb)* 2011, 47, 4055-61.
417. Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L., 3rd. MATCH: an atom-typing toolset for molecular mechanics force fields. *J Comput Chem* 2012, 33, 189-202.
418. Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. Reparameterization of RNA chi torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *J Chem Theory Comput* 2010, 6, 1520-1531.
419. Yin, D.; MacKerell, A. D., Jr. Combined ab initio/empirical approach for optimization of lennard-jones parameters. *J Comput Chem* 1998, 19, 334-338.
420. Yoshii, N.; Miyauchi, R.; Miura, S.; Okazaki, S. A molecular-dynamics study of the equation of state of water using a fluctuating-charge model. *Chem Phys Lett* 2000, 317, 414-420.
421. Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci U S A* 2007, 104, 808-813.
422. Yu, H.; Hansson, T.; Van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J Chem Phys* 2003, 118, 221-234.
423. Yu, W.; Lakkaraju, S.; EP, E. R.; L, L. F.; AD MacKerell, J. Pharmacophore modeling using site-identification by ligand competitive saturation (SILCS) with multiple probe molecules. *J Chem Inf Model* 2015, 55, 407-420.
424. Yu, W.; Lakkaraju, S.; Raman, E.; AD MacKerell, J. Site-Identification by Ligand Competitive Saturation (SILCS) assisted pharmacophore modeling. *J Comput Aided Mol Des* 2014, 28, 491-507.
425. Yu, W.; Lakkaraju, S. K.; Raman, E. P.; MacKerell, A. D., Jr. Site-identification by ligand competitive saturation (SILCS) assisted pharmacophore modeling. *J Computer-Aided Mol Des* 2014.
426. Yu, W.; Lopes, P. E.; Roux, B.; A.D. MacKerell, J. Six-site polarizable model of water based on the classical drude oscillator. *J Chem Phys* 2013, 138, 034508-034513.
427. Zell, B. L.; Goldmann, D. A. Healthcare-associated infection and antimicrobial resistance: moving beyond description to prevention. *Infection Control and Hospital Epidemiology* 2007, 28, 261-264.
428. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., 3rd; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* 2011, 7, 2886-2902.
429. Zhang, J.; Yang, W.; Piquemal, J. P.; Ren, P. Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J Chem Theory Comput* 2012, 8, 1314-1324.
430. Zhong, P.; Pratt, S. D.; Edalji, R. P.; Walter, K. A.; Holzman, T. F.; Shivakumar, A. G.; Katz, L. Substrate requirements for ErmC' methyltransferase activity. *J Bacteriol* 1995, 177, 4327-32.
431. Zhong, S.; Macias, A. T.; MacKerell, A. D., Jr. Computational identification of

- inhibitors of protein-protein interactions. *Curr Top Med Chem* 2007, 7, 63-82.
432. Zhong, S.; MacKerell, J., A.D. . Binding Response: A Descriptor for Selecting Ligand Binding Site on Protein Surfaces. *J Chem Infor Modeling* 2007, 47, 2303-2315.
433. Zhu, X.; MacKerell, A. D., Jr. Polarizable empirical force field for sulfur-containing compounds based on the classical Drude oscillator model. *J Comput Chem* 2010, 31, 2330-41.
434. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem* 2011, 32, 2359-68.
435. Zuckerman, D. M.; Lyman, E. A second look at canonical sampling of biomolecules using replica exchange simulation. *J Chem Theory Comput* 2006, 2, 1200-1202.