

Copyright

by

Andrew William Ritchie

2015

**The Dissertation Committee for Andrew William Ritchie Certifies that this is the  
approved version of the following dissertation:**

**<Title of Dissertation or Treatise, Double-spaced, and centered, may be  
in ALL CAPITALS or Upper and Lower Case>**

**Committee:**

---

Lauren J. Webb, Supervisor

---

Ron Elber

---

Walter Fast

---

Graeme Henkelman

---

Pengyu Ren

---

**<Title of dissertation or treatise, double-spaced,  
may be in ALL CAPITALS or Upper and Lower Case,  
and centered>**

**by**

**Andrew William Ritchie, B.S. Chem**

**Dissertation**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**Doctorate of Philosophy**

**The University of Texas at Austin  
August 2015**

**<Title of dissertation or treatise, double-spaced,  
and centered>**

Andrew William Ritchie, PhD.

The University of Texas at Austin, 2015

Supervisor: Lauren J. Webb

<Abstract: May not exceed 350 words. It should be a continuous description, not disconnected notes or an outline.>

## Table of Contents

List of Tables .....	ix
List of Figures .....	x
List of Illustrations .....	xi
Chapter 1 Introduction .....	1
Chapter 2 Simulation Methods .....	2
2.1 Labeling and Mutating Proteins <i>in silico</i> .....	2
2.2 Enhanced Molecular Dynamics in Amber03: N-Dimensional Umbrella Sampling and Weighted Histogram Analysis Method.....	3
2.3 Electrostatic Clustering in Vibrational Chromophore Dihedral Space...	12
2.4 Probe Parameterization for AMOEBA .....	18
2.5 Small Molecule Simulations in AMOEBA .....	19
Chapter 3 Electrostatic Field Methods.....	20
3.1 Amber03 with Explicit TIP3P Water.....	20
3.1.1 Reaction Field Electrostatics.....	20

3.1.2 Hybrid Solvent Reaction Field Electrostatics and Solute Coulomb Field .....	20
3.2 Amber03 with Poisson-Boltzmann Continuum Solvent .....	21
3.2.1 Reaction Field Method .....	21
3.2.2 Grid spacing and size .....	21
3.2.3 Box Location .....	21
3.3 Amber03 with Poisson-Boltzmann Continuum Solvent and Select Explicit TIP3P Water Molecules .....	22
3.3.1 5 Å Water Sphere Around the Vibrational Chromophore .....	22
3.3.2 Single Water Molecule Nearest the Vibrational Chromophore ..	22
3.3.3 Water Molecular Hydrogen Bonding to the Vibrational Chromophore .....	22
3.4 AMOEBA .....	23
3.4.1 Poisson-Boltzmann Continuum Solvent .....	23
3.4.2 Explicit AMOEBA Water .....	23
3.4.3 Charge Penetration Field Corrections .....	23

Chapter 4 The Role of Electrostatics in Differential Binding of RalGDS to Rap	
Mutations E30D and K31E Investigated by Vibrational Spectroscopy of	
Thiocyanate Probes .....	24
4.1 Introduction .....	24
4.2 Results .....	24
4.3 Discussion .....	24
Chapter 5 Optimizing Electrostatic Field Calculations with the Adaptive Poisson-	
Boltzmann Solver to Predict Electric Fields at Protein-Protein Interfaces I:	
Sampling and Focusing .....	25
5.1 Introduction .....	25
5.2 Results .....	25
5.3 Discussion .....	25
Chapter 6 Optimizing Electrostatic Field Calculations with the Adaptive Poisson-	
Boltzmann Solver to Predict Electric Fields at Protein-Protein Interfaces II:	
Explicit Near-Probe and Hydrogen Bonding Water Molecules .....	26
6.1 Introduction .....	26
6.2 Results .....	26
6.3 Discussion .....	26

Chapter 7 Electrostatic Fields at Protein-Protein Interfaces: Increased Sampling Time and Various Electrostatic Methods: A Case for Simulating in Polarizable Force Fields .....	27
7.1 Introduction .....	27
7.2 Results .....	27
7.3 Discussion .....	27
Chapter 8 Electrostatic Fields in Small Thiocyanate Molecules with Ensembles Generated using the AMOEBA Force Field .....	28
8.1 Introduction .....	28
8.2 Results .....	28
8.3 Discussion .....	28
Appendix (or Appendices) .....	29
Glossary .....	30
References .....	31



## List of Tables

Table 2-1: Sample Parameters for Ryckaert Bellemans dihedral potential function used for validating 2D WHAM code .....	9
---	---

## List of Figures

Figure 2-1: 2D WHAM Validation .....	11
Figure 2-2: PB Solvent Reaction Field vs. Solute Analytic Coulomb Field .....	15
Figure 2-3: Field Values using Clustering Vs. Field Values using All Frames.....	16
Figure 2-4: Correlations and Slopes at Various Cutoff Values .....	17

## **List of Illustrations**

**Word did not find any entries for your table of contents.**

In your document, select the words to include in the table of contents, and then on the Home tab, under Styles, click a heading style. Repeat for each heading that you want to include, and then insert the table of contents in your document. To manually create a table of contents, on the Document Elements tab, under Table of Contents, point to a style and then click the down arrow button. Click one of the styles under Manual Table of Contents, and then type the entries manually.

## **Chapter 1 Introduction**

<Body text to begin here.>

## **Chapter 2 Simulation Methods**

### **2.1 LABELING AND MUTATING PROTEINS *IN SILICO***

<Body text to begin here.>

## 2.2 ENHANCED MOLECULAR DYNAMICS IN AMBER03: N-DIMENSIONAL UMBRELLA SAMPLING AND WEIGHTED HISTOGRAM ANALYSIS METHOD

An umbrella sampling strategy was used to obtain a Boltzmann-weighted statistical ensemble of thiocyanate probe orientations for all MD sampling. Through examining simulated protein structures, it became increasingly apparent that a second degree of freedom,  $\chi_1$ , was relevant to our probe conformational distributions. Therefore, two different umbrella sampling strategies were tested: one-dimensional sampling about the thiocyanate  $\chi_2$  dihedral angle and two-dimensional umbrella sampling about the thiocyanate  $\chi_1$  and  $\chi_2$  dihedral angles, shown in Figure 2-1. All molecular dynamics were completed using the GROMACS<sup>1</sup> software package at 300 K with the AMBER03<sup>2</sup> force field and periodic boundary conditions.

Six probe locations on RalGDS were examined: N27C<sub>SCN</sub>, G28C<sub>SCN</sub>, N29C<sub>SCN</sub>, Y31C<sub>SCN</sub>, K32C<sub>SCN</sub>, and N54C<sub>SCN</sub>, in the monomeric state and docked to each GTPase system examined. We have therefore examined all probe locations and mutated constructs for which experimental data are available. Six structures for each system modeled were generated by fixing the thiocyanate  $\chi_2$  dihedral angle from 0° to 300° in 60° increments. Each structure was sampled with a dihedral potential that was flat within  $\pm 30^\circ$  of the fixed-dihedral position and quadratic with a force constant of 1000 kJ mol<sup>-1</sup> rad<sup>-2</sup> otherwise. These restraining potentials were carried through for the duration of the system set-up and simulation. Each structure was energy minimized with cut-off electrostatics, solvated with tip3p water<sup>3</sup> in a dodecahedron box, charge balanced by

randomly replacing the appropriate number of water molecules with sodium or chloride ions using the genion GROMACS utility, and solvent relaxed by sampling for 20 ps with position restraints on all non-solvent heavy atoms with a force constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup> using PME<sup>4, 5</sup> electrostatics with a real-space cut-off of 0.9 nm, spacing of 0.12 nm, and interpolation order 4. Each rotamer of each system was then sampled using the GROMACS stochastic dynamics integrator, constraints on hydrogen-bonds using the LINCS algorithm<sup>6</sup>, and PME electrostatics for 3 ns, recording snapshots every 5 ps, for a total of 18 ns of simulation and 3606 frames for each system.

Each frame was assigned to one of 72 5° bins from -180° to 175° based on the  $\chi_2$  dihedral angle. The weighted histogram analysis method (WHAM)<sup>7, 8</sup> was then used to calculate a torsional potential of mean force (PMF) for each of  $N$  bins  $i$ , which is related to the torsional probability distribution for each bin  $i$  ( $P_i$ ) described by equation (2-1):

$$P_i = \frac{e^{-\beta \cdot PMF_i}}{\sum_{j=0}^N e^{-\beta \cdot PMF_j}}, \beta = \frac{1}{k_b T} \quad (2-1)$$

which is the typical Boltzmann distribution function for a state  $i$  divided by the partition function, where  $T$  is the temperature in Kelvin,  $k_b$  is the Boltzmann constant, and  $PMF_j$  is the PMF for some state  $j$ .

The two-dimensional umbrella sampling was done in much the same way as the one-dimensional sampling, with few minor changes. The  $\chi_1$  angle was fixed from 0° to 300° in 30° increments, resulting in 12 structures. Each of these structures then had the  $\chi_2$  angle fixed from 0° to 300° in 30°, resulting in 144 total structures. To avoid steric

clashes in the starting structures, for each  $\chi_2$  rotation, the distance between the center of mass coordinate of each rotated atom and every non-rotated atom was calculated. If a distance was found to be under 1.5 Å, the  $\chi_2$  angle was rotated  $\pm 1.5^\circ$  from the dihedral center and the distances were recalculated. This was done until all non-bonded atoms were at least 1.5 Å from each rotated atom. Next, two harmonic dihedral restraining potentials were generated for each structure, one for the  $\chi_1$  dihedral angle and one for the  $\chi_2$  dihedral angle. Following the same set-up strategy used in the one-dimensional sampling, the system was then energy minimized with cut-off electrostatics and dihedral force constants of 1000 kJ mol<sup>-1</sup> rad<sup>-2</sup>, solvated in tip3p water in a dodecahedron box, and charge balanced as described above. The system underwent solvent relaxation using PME electrostatics for 20 ps with a force constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup> on heavy backbone atoms, dihedral force constants of 150 kJ mol<sup>-1</sup> rad<sup>-2</sup>, and unrestrained side-chain atoms. Each of the 144  $\chi_1$ - $\chi_2$  rotamers were then sampled for 400 ps using the GROMACS stochastic dynamics integrator with PME electrostatics (again, with a real-space cut-off of 0.9 nm, spacing of 0.12 nm, and interpolation order 4) and dihedral restraining potentials of 70 kJ mol<sup>-1</sup> rad<sup>-2</sup>, for a total of 57.6 ns of simulation and 11664 frames for each system. It is worth reiterating that each step used a progressively smaller dihedral restraining potential. This was done to ensure that the dihedral angles of the final structure before sampling were as close as possible to the umbrella-sampling window while still allowing nearby residues to relax to orientations that accommodate the inclusion of our probe. Starting with a large restraining potential fixes the probe to a



specific location orientation and forcibly moves nearby residues to accommodate the probe to minimize interaction energies. Subsequent weakening of the restraining potential allows the probe to respond to its surroundings in a manner more typical of MD, allowing both the probe as well as the residues near the probe to relax to energy minimized orientations. Without this subsequent weakening, many simulations resulted in dihedral forces becoming larger than the integrator can or is expected to handle, which were usually caused by steric clashes between the probe and side-chain atoms.

A Boltzmann-weighted statistical ensemble of structures obtained from two-dimensional sampling was then assembled using WHAM. Each frame was assigned to one of 5184 5° by 5° bins. The bins were assigned based on equation (2-2), where  $b_1$  is the one-dimensional  $\chi_1$  bin number,  $b_2$  is the one-dimensional  $\chi_2$  bin number,  $B_i$  is the total number of bins in degree of freedom  $i$  (72 for all two-dimensional sampling done), and  $i_{index}$  refers to the subscript on  $b$ . Conditional probability was assumed.

$$\text{Bin}(b_1, b_2) = \sum_{i=b_1, b_2} B_i^{2-i_{index}} \cdot i \quad (2-2)$$

To validate our 2D WHAM code, we examined a Ryckaert-Bellemans dihedral potential, shown in equation (2-3), where  $\phi$  is some angle and  $n$  and  $C_n$  are some example parameters obtained from the GROMACS manual, shown in Table 2-1. This is a very simple potential function with no contributions from any other source, unlike the potential energy calculation in a protein, which will be influenced by various force field parameters such as bond force constants. However in WHAM, the sources of the

potentials are not distinguished, and we are able to use the simple Ryckaert-Bellemans model to validate the code. We constructed the PMF from the sum of the potentials for a given pair of coordinates and the unbiased probability distribution,  $p^\circ(\chi_1, \chi_2)$ , using the ratio of the Boltzmann distribution function to the partition function, shown in equation (2-4), over an array of 2D dihedral angles ranging from  $-180^\circ$  to  $180^\circ$ .

$$V_{rb} = \sum_{n=0}^5 C_n \left( \cos(\phi - \pi) \right)^n \quad (2-3)$$

Next, we constructed the PMF and probability distribution landscapes, shown in Figure 2-1a. We then applied biasing potential windows to each dimension in a manner that mimics the methods used in GROMACS. The biased probability,  $p(\chi_1, \chi_2)$ , is given by equation (2-5). We then performed a Monte Carlo simulation centered on each of 144 biased windows, each with dimensions of  $30^\circ \times 30^\circ$ , with a probability  $p(\chi_1, \chi_2)$  of sampling a given pair of dihedral angles, which was then analyzed using our 2D WHAM code to return the unbiased PMF and  $p^\circ(\chi_1, \chi_2)$ .

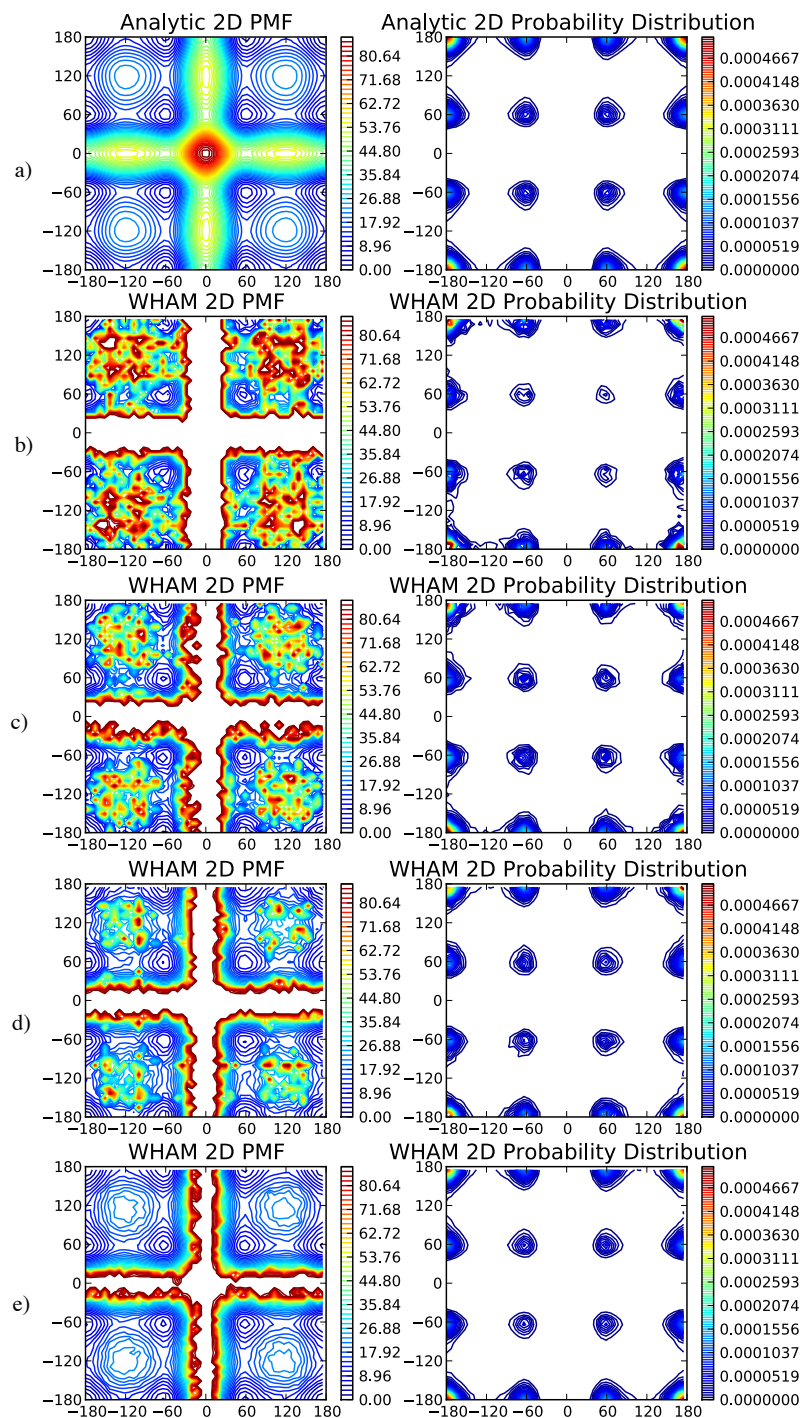
$$p^\circ(\chi_1, \chi_2) = \frac{e^{-\beta(V_{rb}(\chi_1) + V_{rb}(\chi_2))}}{\iint e^{-\beta(V_{rb}(\chi_1) + V_{rb}(\chi_2))} d\chi_1 d\chi_2} \quad (2-4)$$

$$p(\chi_1, \chi_2) = \frac{e^{-\beta(V_{rb}(\chi_1) + V_1 + V_{rb}(\chi_2) + V_2)}}{\iint e^{-\beta(V_{rb}(\chi_1) + V_1 + V_{rb}(\chi_2) + V_2)} d\chi_1 d\chi_2} \quad (2-5)$$

Figure 2-1 shows the analytical PMF and probability distributions (a), WHAM PMF and probability distributions on 144 windows each containing 40 frames (b), WHAM PMF and probability distributions on 144 windows each containing 80 frames (c), WHAM PMF and probability distributions on 144 windows each containing 160 frames (d), and WHAM PMF and probability distributions on 144 windows each containing 1000 frames (e). The major features of the probability distribution become clear after only 40 frames; after 80 frames the probability distributions look very similar to the analytical distribution, and after 160 frames very little improvement is seen. We can also see that the high probability regions, representing the staggered orientations, have PMF landscapes that look like the analytical PMF, although the gauche regions appear to be ill characterized still. After 1000 frames the probability distribution is nearly identical to the analytical distribution and the moderate ranges of the PMF (light blue, > 17.92 kJ mol<sup>-1</sup>) now quantitatively approach the predictions of the analytical expression. Exact analytical PMF matching of low probability regions is very slow, requiring  $\geq 10000$  frames. From these results we conclude that the PMF converges to the analytical expression slowly for regions of low probability and quickly for regions of high probability.

Table 2-1: Sample Parameters for Ryckaert Bellemans dihedral potential function used for validating 2D WHAM code

	Constraint (kJ mol <sup>-1</sup> )		Constraint (kJ mol <sup>-1</sup> )
<b>C<sub>0</sub></b>	9.28	<b>C<sub>3</sub></b>	-3.06
<b>C<sub>1</sub></b>	12.16	<b>C<sub>4</sub></b>	-26.24
<b>C<sub>2</sub></b>	-13.12	<b>C<sub>5</sub></b>	-31.5



### Figure 2-1: 2D WHAM Validation

Comparison between the PMF and probability distributions of a) an analytic Ryckaert-Bellemans dihedral potential and Monte Carlo 2D umbrella sampling for b) 40 frames/biasing window; c) 80 frames/biasing window; d) 160 frames/biasing window; e) 1000 frames/biasing window. Units on the PMF are  $\text{kJ mol}^{-1}$ .

### 2.3 ELECTROSTATIC CLUSTERING IN VIBRATIONAL CHROMOPHORE DIHEDRAL SPACE

The largest bottleneck for these sorts of calculations we do are the electrostatics. A single node on Stampede can generate >10 ns of simulation per day. That number can be increased (logarithmically) by using additional nodes. However, the continuum solvent electrostatics calculations take anywhere from 45-60 seconds (APBS) to  $\approx$  20 minutes (AMOEBA) per frame. If we keep every 4 ps and collect 250 frames per nanosecond, then the electrostatics require 5-8 ns/day for APBS calculations and approximately 0.3 ns/day for AMOEBA calculations. This can be decreased further by running the serial calculations in parallel. Regardless, it would be convenient to find some method of pruning the total number of frames for continuum electrostatics calculations while ensuring that the average field does not differ significantly from the average using every frame.

For convenience, the vacuum electrostatic field at the nitrile due to solute only was chosen as an indicator of total electrostatic field. In the absence of solvent, this is trivial to calculate for both point charge force fields (Amber03) as well as multipole force fields (AMOEBA). This was chosen because it was 1) intuitive and 2) there is consistently a good correlation between the Coulomb field and the solvent reaction field, as seen in Figure 2-2, suggesting that frames which well represent the Coulomb field also well represent the reaction field.

We then took advantage of the weighted averaging over binned data. The Boltzmann weighted average is calculated as in equation (2-6),

$$\langle x \rangle = \sum_{i=1}^{\text{nbins}} \rho_i \sum_{j=1}^{c_i} \frac{x_{ij}}{c_i} \quad (2-6)$$

where the probability of being in each bin  $i$  is  $\rho_i$ , the number of times bin  $i$  is visited is  $c_i$ , and each value in bin  $i$  is  $x_{ij}$  for  $j = 1$  to  $j = c_i$ . There exists some subarray of values in bin

$i$  that has  $k_i$  values, where  $k_i \geq c_i$  entries and  $\left| \sum_{j=1}^{k_i} \frac{x_{ij}}{k_i} - \sum_{j=1}^{c_i} \frac{x_{ij}}{c_i} \right| \leq \chi$ , where  $\chi$  is some

threshold. As  $\chi$  approaches 0,  $k_i$  approaches  $c_i$ , and the subarray is the full array and the averages are identical. Using the a set of test data where umbrella windows were centered every 120 degrees (at 60°, 180°, and 300°, the expected alkane maximum probability torsions) with a flat biasing potential  $\pm 60^\circ$  of the window center and a force constant of **\*\*XXX**, the clustered average field for each Ral probe in the monomeric state, docked to each of the four Rap1a mutants and each of the four Ras mutants is plotted against their average field for various values of  $\chi$ , indicated in the upper-left corner of each subplot, from the full data sets in Figure 2-3. From this, it's clear that the clustered averages are linearly correlated to the full averages. The correlation coefficients and best-fit slopes have also been plotted as a function of the cutoff,  $\chi$ , in Figure 2-4. Even for a relatively large  $\chi = 1$ , the clustered correlation coefficient and slope is *very* close to 1.0. In general, as  $\chi$  approaches zero, the correlation coefficient and slope also approach zero. Moreover, at a  $\chi = 0.01$ , only approximately 20% of all frames are used, which is a significant decrease in computation requirement. This method also has the advantage of



guaranteeing that the property the cluster is based on always has a clustered average nearly identical to the full average, which is a useful sanity check.

**\*\*REMAKE THIS FIGURE USING FEWER DOTS (LOWER RESOLUTION)**

Figure 2-2: PB Solvent Reaction Field vs. Solute Analytic Coulomb Field

Comparison between the analytic Coulomb field at the nitrile bond midpoint due to solute (less the contributions due to the probe itself) (x-axis) and the solvent reaction field at the nitrile bond midpoint (y-axis) for each frame, grouped by the probe identity, both using Amber03 point charges. Correlation coefficients ( $r$ ), slopes ( $m$ ), and y-intercepts ( $int$ ), are indicated in the upper-right corner of each figure.

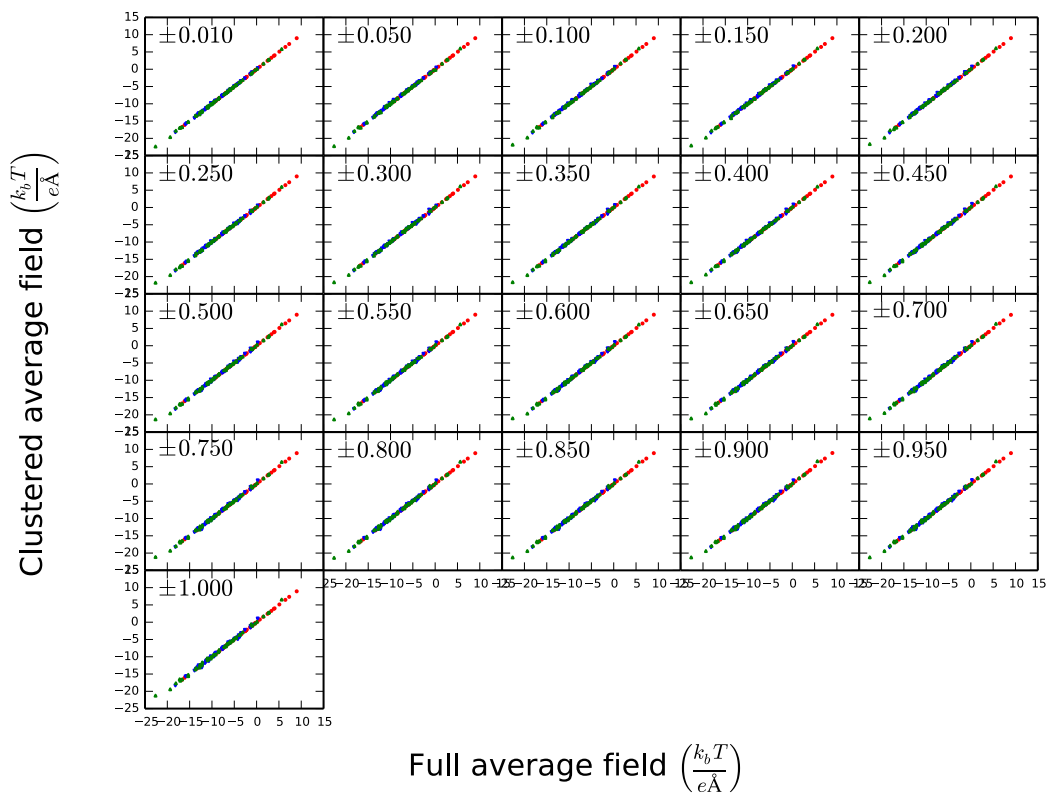


Figure 2-3: Field Values using Clustering Vs. Field Values using All Frames

Average Coulomb electrostatic field (red), solvent reaction field (blue), and the electrostatic field calculated using the AMOEBA force field (green) from clustered frames versus the respective full averages for various cutoff values,  $\chi$ .

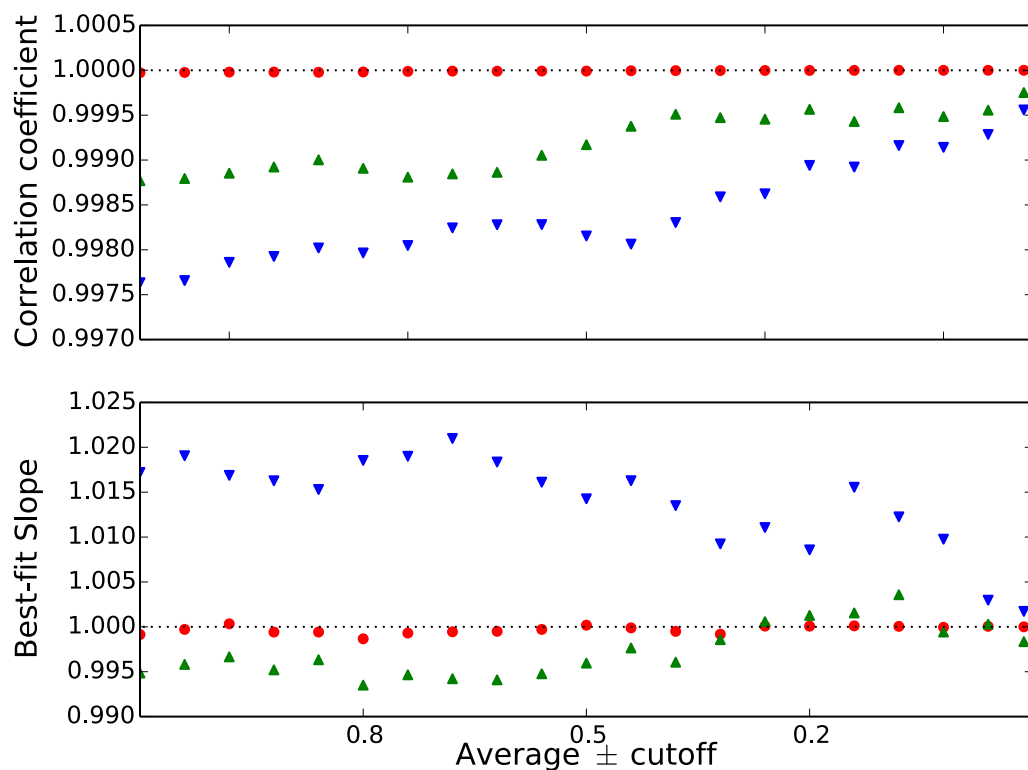


Figure 2-4: Correlations and Slopes at Various Cutoff Values

Correlation coefficients of the Coulomb electrostatic field (red), solvent reaction field (blue), and the electrostatic field calculated using the AMOEBA force field (green) as a function of the cutoff,  $\chi$ . (Bottom) Best-fit slopes of the Coulomb electrostatic field (red), solvent reaction field (blue), and the electrostatic field calculated using the AMOEBA force field (green) as a function of the cutoff,  $\chi$ .

## **2.4 PROBE PARAMETERIZATION FOR AMOEBA**

<Body text to begin here.>

## **2.5 SMALL MOLECULE SIMULATIONS IN AMOEBA**

<Body text to begin here.>

## **Chapter 3 Electrostatic Field Methods**

### **3.1 AMBER03 WITH EXPLICIT TIP3P WATER**

#### **3.1.1 Reaction Field Electrostatics**

#### **3.1.2 Hybrid Solvent Reaction Field Electrostatics and Solute Coulomb Field**

<Body text to begin here.>

## **3.2 AMBER03 WITH POISSON-BOLTZMANN CONTINUUM SOLVENT**

### **3.2.1 Reaction Field Method**

<Body text to begin here.>

### **3.2.2 Grid spacing and size**

<Body text to begin here.>

### **3.2.3 Box Location**

<Body text to begin here.>



### **3.3 AMBER03 WITH POISSON-BOLTZMANN CONTINUUM SOLVENT AND SELECT EXPLICIT TIP3P WATER MOLECULES**

#### **3.3.1 5 Å Water Sphere Around the Vibrational Chromophore**

<Body text to begin here.>

#### **3.3.2 Single Water Molecule Nearest the Vibrational Chromophore**

<Body text to begin here.>

#### **3.3.3 Water Molecular Hydrogen Bonding to the Vibrational Chromophore**

<Body text to begin here.>

## **3.4 AMOEBA**

### **3.4.1 Poisson-Boltzmann Continuum Solvent**

<Body text to begin here.>

### **3.4.2 Explicit AMOEBA Water**

<Body text to begin here.>

### **3.4.3 Charge Penetration Field Corrections**

<Body text to begin here.>

# **Chapter 4 The Role of Electrostatics in Differential Binding of RalGDS to Rap Mutations E30D and K31E Investigated by Vibrational Spectroscopy of Thiocyanate Probes**

## **4.1 INTRODUCTION**

<Body text to begin here.>

## **4.2 RESULTS**

<Body text to begin here.>

## **4.3 DISCUSSION**

<Body text to begin here.>

# **Chapter 5 Optimizing Electrostatic Field Calculations with the Adaptive Poisson-Boltzmann Solver to Predict Electric Fields at Protein-Protein Interfaces I: Sampling and Focusing**

## **5.1 INTRODUCTION**

<Body text to begin here.>

## **5.2 RESULTS**

<Body text to begin here.>

## **5.3 DISCUSSION**

<Body text to begin here.>

# **Chapter 6 Optimizing Electrostatic Field Calculations with the Adaptive Poisson-Boltzmann Solver to Predict Electric Fields at Protein-Protein Interfaces II: Explicit Near-Probe and Hydrogen Bonding Water Molecules**

## **6.1 INTRODUCTION**

<Body text to begin here.>

## **6.2 RESULTS**

<Body text to begin here.>

## **6.3 DISCUSSION**

<Body text to begin here.>

# **Chapter 7 Electrostatic Fields at Protein-Protein Interfaces: Increased Sampling Time and Various Electrostatic Methods: A Case for Simulating in Polarizable Force Fields**

## **7.1 INTRODUCTION**

<Body text to begin here.>

## **7.2 RESULTS**

<Body text to begin here.>

## **7.3 DISCUSSION**

<Body text to begin here.>

## **Chapter 8 Electrostatic Fields in Small Thiocyanate Molecules with Ensembles Generated using the AMOEBA Force Field**

### **8.1 INTRODUCTION**

<Body text to begin here.>

### **8.2 RESULTS**

<Body text to begin here.>

### **8.3 DISCUSSION**

<Body text to begin here.>

## **Appendix (or Appendices)**

This page is optional—must be placed in this order if it is included in the dissertation. If you don't want to include an appendix, then delete the entire page and the following page break.



## **Glossary**

This page is optional—must be placed in this order if it is included in the dissertation. If you don't want to include a glossary, then delete the entire page and the following page break.

## References

1. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701-1718.
2. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T., et al., 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.
3. 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* (2), 926-935.
4. Darden, T.; York, D.; Pedersen, L. G., Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
5. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.
6. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., - LINCS: A linear constraint solver for molecular simulations. **1997**, - *18* (- 12), - 1472.
7. Roux, B., The calculation of the potential of mean force using computer simulations. *Comp. Phys. Commun.* **1995**, *91*, 275-282.

8. Gallicchio, E.; Andrec, M.; Felts, A. K.; Levy, R. M., Temperature Weighted Histogram Analysis Method, Replica Exchange, and Transition Paths. *J. Phys. Chem. B* **2005**, *109*, 6722-6731.