

6-2012

Prediction of Metal Charge-Dependent Amide Hydrogen Exchange Rates for the Proteins Rubredoxin and Azurin

Stephen Scoglio

Union College - Schenectady, NY

Follow this and additional works at: <https://digitalworks.union.edu/theses>



Part of the [Chemistry Commons](#)

Recommended Citation

Scoglio, Stephen, "Prediction of Metal Charge-Dependent Amide Hydrogen Exchange Rates for the Proteins Rubredoxin and Azurin" (2012). *Honors Theses*. 897.

<https://digitalworks.union.edu/theses/897>

This Open Access is brought to you for free and open access by the Student Work at Union | Digital Works. It has been accepted for inclusion in Honors Theses by an authorized administrator of Union | Digital Works. For more information, please contact digitalworks@union.edu.

Prediction of Metal Charge-Dependent Amide Hydrogen Exchange Rates for the Proteins Rubredoxin and Azurin

By

Stephen T. Scoglio

Submitted in partial fulfillment of the requirements for Honors in
the Department of Chemistry

Union College

June, 2012

Scoglio, Stephen '12

New York, NY

Abstract

Prediction of Metal Charge-Dependent Amide Hydrogen Exchange Rates for the Proteins Rubredoxin and Azurin

Stephen T. Scoglio, Union College, Department of Chemistry, June 2012

Although it has been shown that the hydrogen exchange rates between the most slowly exchanging amide protons of a protein and the solvent can be used to predict the global thermodynamic stability of a protein, the great majority of amide hydrogens along the backbone of a protein exchange from either partially or fully folded conformations, and their rate of exchange strongly depends upon the environment in that folded state. We are examining the electrostatic and conformational dependence of the hydrogen exchange kinetics of *Pseudomonas aeruginosa* azurin, which can bind either the diamagnetic Cu(I) or Zn(II) in its active site and *Pyrococcus furiosus* (Pf) rubredoxin which can be substituted with Zn(II) or Ga(III). Molecular dynamics and electrostatic calculations were used to determine the theoretical hydrogen exchange rates for these metal-substituted proteins in order to assess the relationship between metal ion charge and exchange rate.

Table of Contents

<u>TITLE</u>	<u>PAGE</u>
Abstract.....	ii
Table of Figures.....	iv
INTRODUCTION.....	1-7
EXPERIMENTAL.....	8-11
RESULTS.....	12-20
DISCUSSION.....	21-23
REFERENCES.....	24
ACKNOWLEDGEMENTS.....	25
APPENDIX.....	26-43
<i>RMSD Plots after Molecular Dynamics.....</i>	<i>26-28</i>
<i>Metal Binding Site Angles and Distances after MD.....</i>	<i>29-32</i>
<i>Setting up NAMD for Molecular Dynamic Simulations.....</i>	<i>33-39</i>
<i>Setting up Delphi Calculations.....</i>	<i>40-43</i>

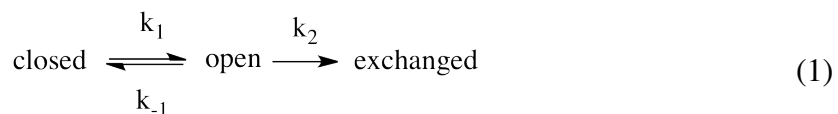
Table of Figures

Figure 1: Chimera 3D model of <i>Pf</i> rubredoxin.....	2
Figure 2: Chimera 3D model of <i>Pf</i> rubredoxin metal binding sit.....	3
Figure 3: Backbone structure of <i>Pf</i> rubredoxin.....	4
Figure 4: Secondary structure of azurin.....	5
Figure 5: Chimera 3D model of azurin metal binding site.....	5
Figure 6: Diagram of N-methylacetamide reference molecule.....	7
Figure 7: Overview of the computational methods used.....	8
Figure 8: RMSD for Zn(II) rubredoxin MD run for production 1.....	12
Figure 9: Metal binding site angles for Zn(II) azurin MD run, production 1.....	13
Figure 10: Metal binding site distances for Zn(II) azurin MD run, production 1.....	13
Figure 11: Predicted vs. experimental exchange rates for Zn(II) rubredoxin.....	14
Figure 12: Predicted vs. experimental exchange rates for Ga(III) rubredoxin.....	15
Figure 13: Chimera 3D image of azurin secondary structure, α -helix 2.....	16
Figure 14: Chimera 3D image of Zn(II) and Cu(I) azurin α -helix 2 structure with rates..	17
Figure 15: Chimera 3D image of azurin secondary structure, β -strands 6, 3, and 1.....	18
Figure 16: Chimera 3D image of Cu(I) azurin β -strands 6, 3, and 1with rates.....	19
Figure 17: Chimera 3D image of Zn(II) azurin β -strands 6, 3, and 1with rates.....	20

Introduction

It has been shown that the hydrogen exchange rates between the most slowly exchanging amide protons of a protein and the solvent can be used to predict information about protein structure and dynamics. Specifically, it can provide input about an amino acid's conformation and interactions [1]. Hydrogen exchange is a proton transfer reaction where freely exposed amide hydrogen exchanges with solvent water catalyzed by a hydroxide anion [1]. Because the exchange rate depends on the electrostatic environment around the amide group, one can calculate the rate constant for that exposed amide by calculating the electrostatic energy of the protonated and deprotonated form.

Experimental hydrogen exchange rates depend on pH, temperature, isotope effects, nearest-neighbor inductive effects and blocking effects [1]. Represented in Equation 1 is the reaction mechanism of protein amide hydrogen exchange rates.



In the first step, the native structure of the protein opens to allow access to the solvent, and hydrogen exchange occurs in the second step. Berger and Linderstrom-Lang [2] proposed that if the open state returns to the closed state faster than to the exchanged state, ($k_{-1} \gg k_2$), a pre-equilibrium exists and the overall exchange rate constant equals $(k_1/k_{-1})k_2$ [3,4].

Hydrogen exchange rates can be measured experimentally using NMR by placing a protein sample into D₂O [1]. In this work, computational methods will be used to theoretically predict hydrogen exchange rates for the proteins *Pyrococcus furiosus* (Pf)

rubredoxin and *Pseudomonas aeruginosa* azurin using molecular dynamics (MD) simulations and Delphi calculations.

Pf rubredoxin is a thermostable small electron transfer 54-residue protein [4]. Shown in Figure 1 is the X-ray structure of *Pf* rubredoxin (pdb code 1bq8).

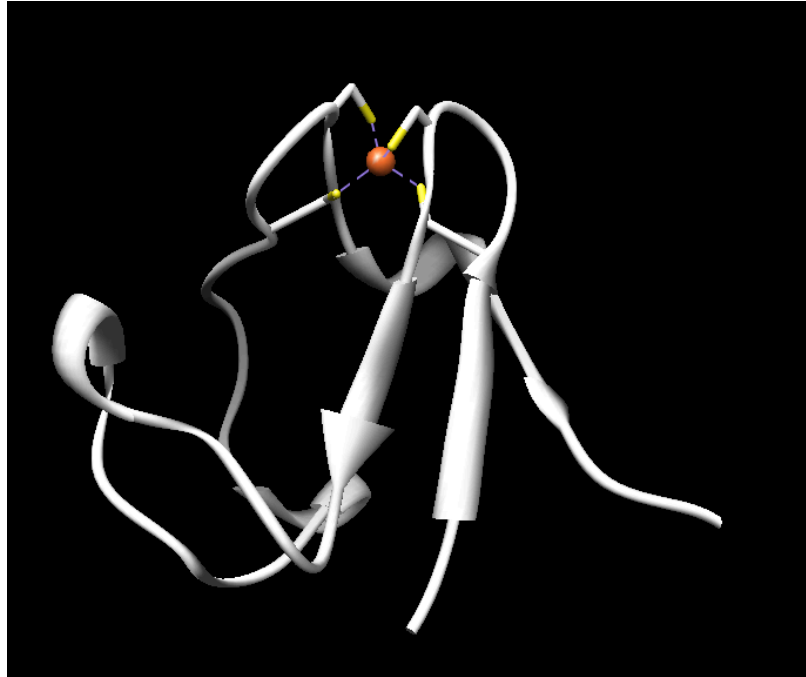


Figure 1: Chimera 3D model of *Pf* rubredoxin. Shown within the metal binding site are four cysteine residues surrounding a metal ion (orange) [5].

Represented in Figure 2 is the X-ray structure of *Pf* rubredoxin's metal binding site with the metal ion, Fe^{3+} [5].

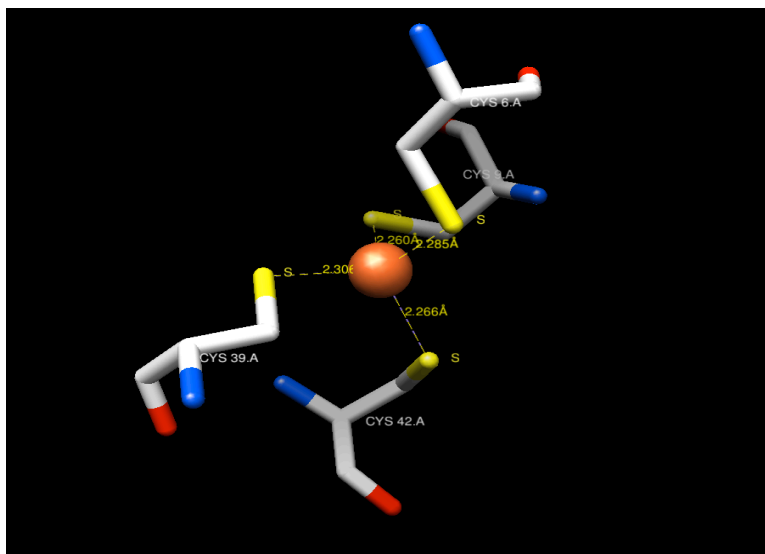


Figure 2: Chimera 3D model of (*Pf*) rubredoxin's tetrahedral tetracysteine-coordinated metal binding site [6].

Zn^{2+} or Ga^{3+} can be substituted for Fe^{3+} , forming a stable protein and similar X-ray crystal structures [5]. Thus, the charge on the metal ion does not alter the spatial orientation of the amide groups in each protein and the metal charge-dependent differences in the rate of hydrogen exchange of the protein should primarily depend on the electrostatic stabilization of the amide anion that is formed during the exchange reaction. Anderson *et. al.* [7] predicted the kinetics of hydrogen exchange using Poisson-Boltzmann continuum dielectric calculations using the Delphi program for the amide groups of *Pf* rubredoxin that are exposed to solvent in the high resolution X-ray structure. These rates were determined to depend strongly on the electrostatic potential around the amide group, and the rates spanned a billion-fold range.

LeMaster *et. al.* [4] showed that the metal binding site of *Pf* rubredoxin, shown in Figure 2, greatly influenced the hydrogen exchange rates due to electrostatic interactions. In a previous experiment, LeMaster *et al.* [8] created the first reported “biomacromolecular complex” with *Pf* rubredoxin having the metal ion, Ge^{4+} . Interestingly, when the metal ion charge increased from Zn^{+2} to Ge^{+4} , the hydrogen

exchange rates for amide protons located near rubredoxin's metal binding site increased up to a million-fold [7]. On the other hand, amides that were greater than 12Å away from the binding site were not affected by the change in ionic charge. Shown in Figure 3 is the backbone of *Pf* rubredoxin with its metal binding site and the effect of the metal ion impacting the hydrogen exchange rates [7].



Figure 3: Backbone of *Pf* rubredoxin with the metal binding site shown in pink [4]. Amides whose exchange rates changed the most for the different metal ions were those closest to the metal binding site, shown in red. Intermediate exchange rate changes are shown in yellow, which were further away from the metal binding site, and amides whose exchange rates were not affected by altering the metal ion are shown in black (Farther than 12 Å away from the metal ion). The residues marked in gray that lack exchange rate measurement are primarily prolines that have no amide hydrogen.

This is significant because the electrostatic environment caused by *Pf* rubredoxin's metal binding site greatly influenced the rate of exchange for exposed amide hydrogens. Thus, electrostatic environments are important influential factors for hydrogen exchange rates on a protein.

Azurin is a small 128-residue protein with a Cu^{1+} metal binding site located between two β sheets, as shown in Figure 3 [9].



Figure 4: Secondary structure of azurin with its metal binding site, which is shown in a bead-like structure. In violet is the Cu^{+} metal ion [9].

When Zn^{2+} is substituted for Cu^{+} , azurin's natural structure is still mostly retained [9].

Five residues located in the metal binding site form a distorted trigonal bipyramidal structure around the metal ion, as shown in Figure 4.

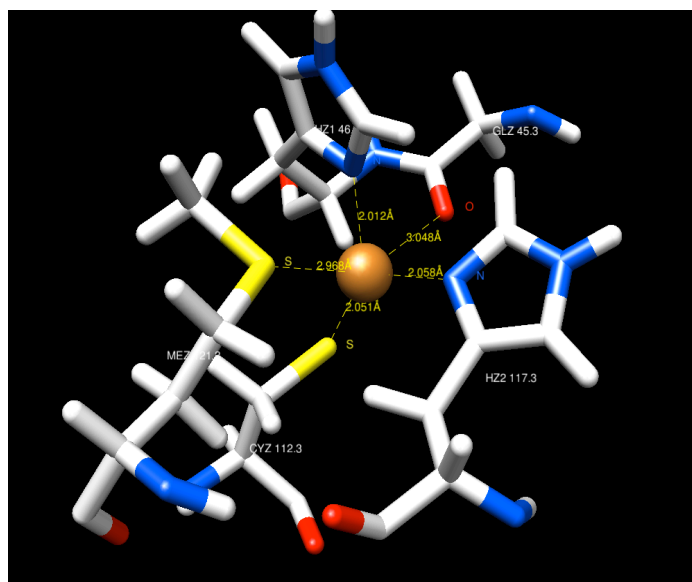


Figure 5: Chimera 3D model of azurin showing the distorted trigonal bipyramidal metal binding site [6].

The five residues are: His46, Cys112, His117, Gly45 and Met121. Three of the amino acids are in the equatorial position, His46, Cys112 and His117, and two are axial, Gly45 and Met121 [9]. As shown in Figure 4, N_{His46} , N_{His117} , S_{Cys112} and S_{Met121} , and O_{Gly45} interact with the metal ion.

This research will examine the theoretical electrostatic and conformational dependence of hydrogen exchange kinetics for Cu and Zn^{2+} *Pseudomonas aeruginosa* azurin and Zn^{2+} and Ga^{3+} *Pf* rubredoxin using molecular dynamic simulated (MD) conformations [10] and Delphi Poisson-Boltzmann calculations of electrostatic energy [11,12]. A high-resolution X-ray structure is only a single molecular conformation, but MD simulations generate thousands of low energy conformations, and as a result, having more conformations will increase the number of amide hydrogens exposed to solvent due to small conformational changes in the protein structure during its MD simulation. High-resolution X-ray protein structures, (pdb codes 3azu and 1bq8) will be used as the starting structures for the MD simulations for the proteins azurin [13] and *Pf* rubredoxin [5], respectively.

We will generate two 1000 (50ns) or 2000 (100ns) frame MD trajectories for each protein using Union College's IBM cluster. For each production run, constraints will be used to maintain the metal binding site distances and angles. The program SurfV will be used to analyze the surface area of each amide atoms throughout each of the protein's production runs [14]. The amide hydrogens exposed to solvent in each production frame will be used in Delphi calculations to determine the electrostatic energy of both the protonated and deprotonated state of the amide. Using the results of these Delphi calculations, theoretical hydrogen exchange rate constants can be determined [7]. In

order to keep the overall charge of the system constant, a reference molecule, N-methylacetamide (or N-methylacetamide anion), was added to the system at least 16Å away from any formal charges, as shown in Figure 5.

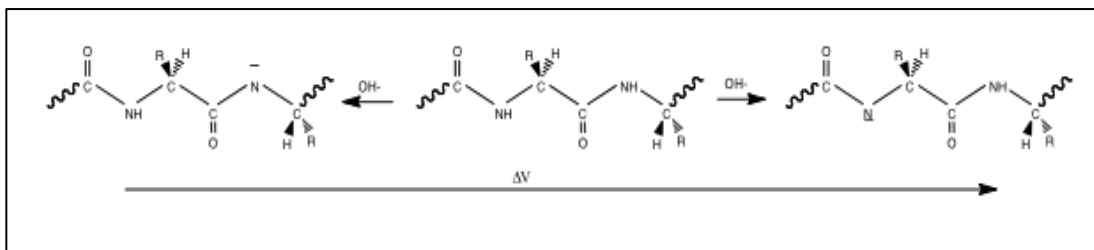


Figure 6: A diagram of a protein's backbone where the solvent exposed amide hydrogens are exchanging with a hydroxide anion. Each protonated and deprotonated amide states create its own distinct electrostatic potential. Differences between the protonated and deprotonated protein state electrostatic potentials are calculated, ΔV (Reference).

The energy differences, ΔV , for the protonated and deprotonated amides are then converted to the theoretical hydrogen exchange rates.

After calculating the theoretical hydrogen exchange rates for the exposed amide hydrogens, experimental data is used to see if there is a positive correlation between the experimental vs. theoretical results. Experimental hydrogen exchange rate data, for the protein *Pf* rubredoxin, will be provided from researchers Dr. David LeMaster and Griselda Hernández. For each purified protein sample, NMR experiments measured the experimental hydrogen exchange rates using D₂O [3]. Overall, we hope to successfully theoretically predict the exposed amide hydrogen hydrogen exchange rates from the proteins rubredoxin and azurin using the computational methods of MD and Delphi calculations and to test the correlation between the experimental and computational results.

Experimental

Represented in Figure 7 is the overview of the computational methods used to calculate the theoretical hydrogen exchange rates for each exposed amides in each of the four MD production runs for either Cu^+ and Zn^{2+} *Pseudomonas aeruginosa* azurin and Zn^{2+} and Ga^{3+} *Pf* rubredoxin.

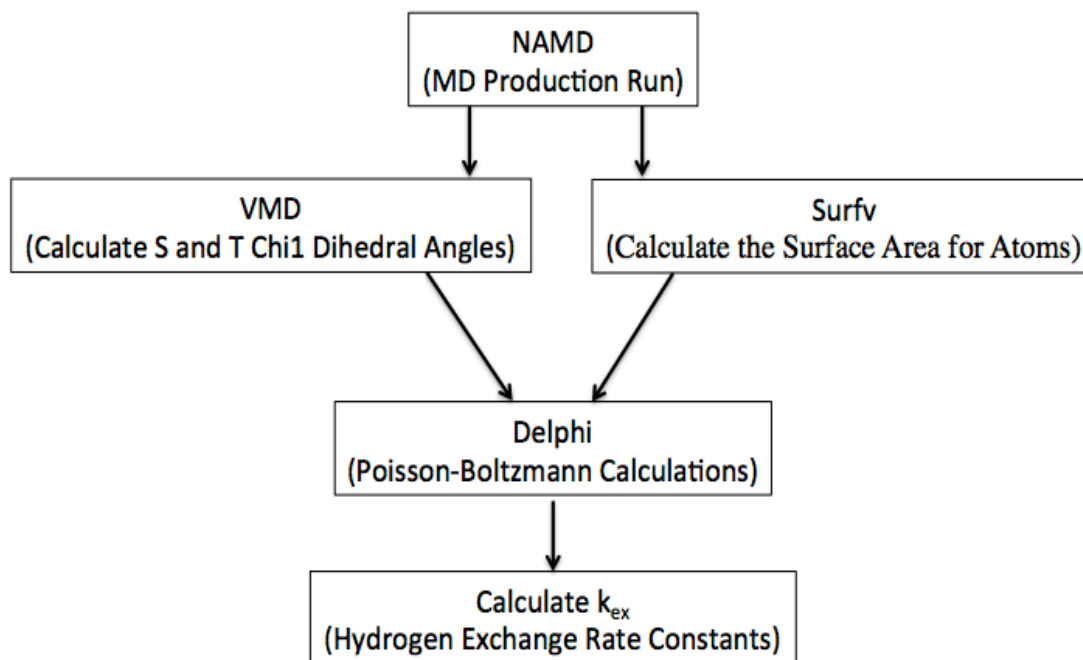


Figure 7: Overview of the programs used to calculate the theoretical hydrogen exchange rates for each exposed amides in each of the four MD production runs for either azurin or rubredoxin.

The Appendix labeled “Setting up NAMD for molecular dynamic simulations” gives details about the individual steps used for each of the computational methods shown in Figure 7. The programs and files used for azurin were derived from the files and programs used for rubredoxin. Individual Fortran programs and scripts are located in an electronic folder labeled “Programs_scripts.”

For the production runs, the starting conformations were the X-ray structures from pdb files 3azu and 1bq8, respectively. To build the protein in NAMD, patches, ZNC4, GAC4, CUZ5 and ZNZ4 were used for Zn^{2+} rubredoxin, Ga^{3+} rubredoxin, Cu^{+1} azurin, and Zn^{2+} azurin respectively. These patches (Shown in the Appendix) are found in the files: top_all27_prot_lipid_az_v2.inp and top_all27_prot_lipid_zn2_ga3-inp.txt. Each topology file has atomic charges designated for each atom and metal ion located near or within the metal binding site. The parameter file, par_all27_prot_lipid_az_rub.prm, has the bond lengths and atomic radius used for each production run. Separate pdb files, converted into CHARMM22 format, as shown in the Appendix, for metal, protein and crystal waters, were needed to create each designated protein structure for MD productions. Na^+ and Cl^- ions were added using autoionize plugin version 1.30m to bring the ionic strength to 0.150m and a decreasing constraint step run was done, where the protein was constrained in a water cube and the environment was heated from 0 to 298K. The decreasing constraints run for both proteins was done on MacPro computers and MD productions on Union College's IBM cluster. Directions on how to perform MD production runs and the decreasing constraint runs are found in the Appendix labeled: *"Setting up NAMD and Delphi to calculate Hydrogen Exchange Rates."*

As shown in the Appendix, the program surfv2 was used to calculate the surface area for each atom within a protein's MD production run. The programs and files needed to set up for surfv2 calculations were: center_999_pdb_fix.f, separate_pdbs_delphi_2.f, make_delhi_input_7.f and LeeRich.siz. To extract the surface-accessible area data from surfv2, the program ensembleSASA.f was used.

The program Delphi using Poisson-Boltzmann calculations determined the electrostatic potential using a N-methylacetamide reference molecule. The programs and files needed for setting up Delphi were: make_dir_labels_110211.f, 3azu_sequence.txt, surfv2 output text file, make_dir_names_ST_110211.f, make_directories_accF and make_delphi_input_10.f. To run Delphi, the script, start_delphi_all and run_delphi scripts were used to run Delphi using the IBM cluster, as shown in the Appendix.

Chi1 dihedral angles for the S and T residues were calculated using VMD software. Since S and T residues contain an alcohol group, it can readily undergo hydrogen exchange with the solvent depending on the chi1 dihedral angle. If the chi1 angles for the S and/or T residues are gauche + or gauche -, 60° or -60°, or trans, 180°, then the OH group is replaced with a hydrogen to simulate this exchange. These altered residues are labeled Z or X to replace the S and T residues, respectively. The software VMD and the program calc_STchi1_script_test.tcl extracted each MD production's chi1 S and T residues and created an output text file needed for Delphi. Additional information about calc_STchi1_script_test.tcl program's output file is located in the Appendix.

After calculating the electrostatic potentials using Delphi, the programs, run_grep and read_grep_allcalc_azurin_v4.f, were used to extract the energies and calculate the hydrogen exchange rates for the exposed amide hydrogens. For rubredoxin, the scale factor yielding the smallest rmsd between experimental and calculated log k values was used for calculating the theoretical rates. Excel was also used to analyze the exchange rate data. Experimental data for rubredoxin, provided by Griselda Hernández and David LeMaster, was used to see if the theoretical computational data positively correlates to

the experimental NMR data. Because no experimental exchange rates were analyzed for azurin, the scale factor used for azurin is 7.6.

Results

Following each molecular dynamic production run, root mean square deviations (RMSD) from the initial conformation for the heavy atoms were calculated. Represented in Figure 8 is the RMSD values for production 1 Zn(II) rubredoxin throughout the molecular dynamic production run.

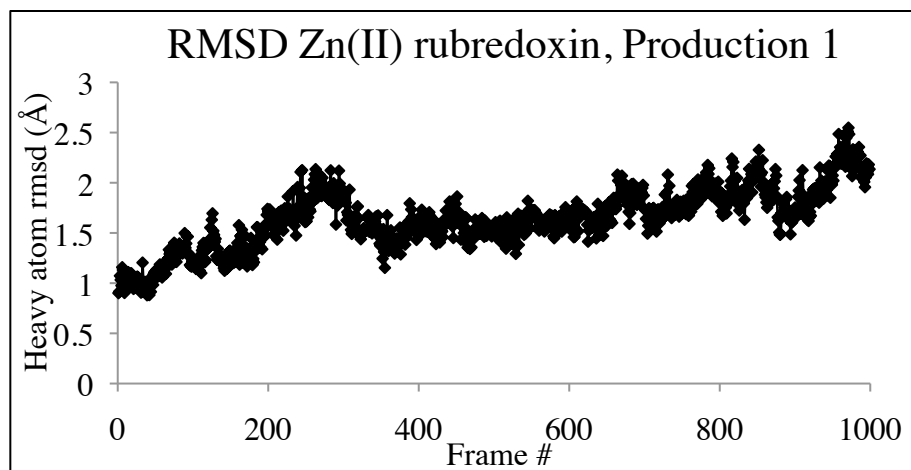


Figure 8: Root mean square deviation for production one Zn(II) rubredoxin's heavy atoms throughout its 50ns trajectory. The average RMSD value is 1.64Å. VMD was used to calculate the RMSD values.

Additional RMSD figures for the other production runs are shown in the Appendix.

The metal binding site angles and distances of Cu(I) and Zn(II) azurin were calculated throughout each of their MD production runs. Shown in Figure 9 are the results for the Zn(II) azurin, production run #1.

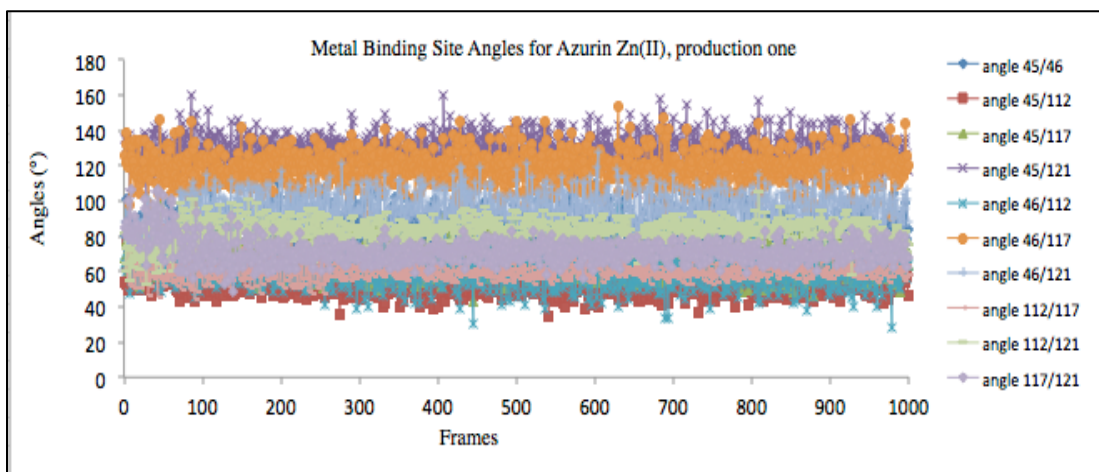


Figure 9: Metal binding site angles for the protein Zn(II) azurin, production one. The program used was *calc_metal_angles_dist_azurin.f*. Residue numbers are shown.

The angles calculated were between the residues inside the metal binding, as shown in Figure 5. Additional metal binding site angles for the proteins rubredoxin and azurin can be found in the Appendix. The metal binding site distances were also calculated throughout each of MD production runs. Represented in Figure 10 are the metal binding site distances for Zn(II) azurin, production run #1.

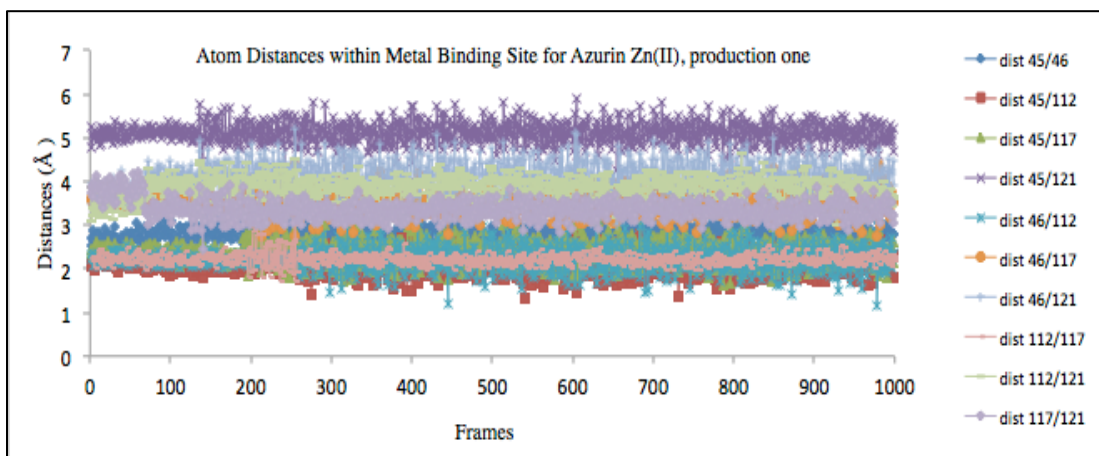


Figure 10: Metal binding site distances for the protein Zn(II) azurin, production one. The program used was *calc_metal_angles_dist_azurin.f*. Residue numbers are shown.

Additional metal binding site distances for the other respective MD production runs for the protein azurin are shown in the Appendix.

After each production run, Surfv, and Delphi calculations were used to calculate hydrogen exchange rates. Shown in Figures 11 and 12 are the predicted hydrogen exchange rates and the experimental hydrogen exchange rates for the proteins Zn(II) rubredoxin and Ga(III) rubredoxin, respectively.

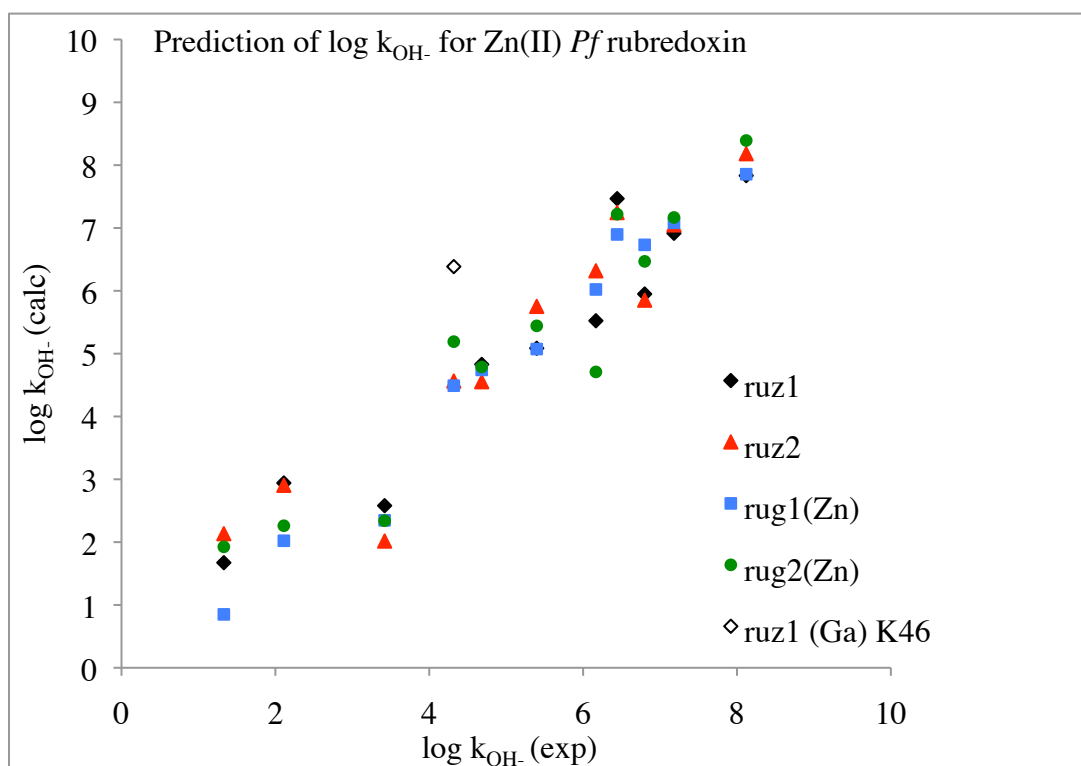


Figure 11: The predicted amide hydrogen exchange rates (*calc*) vs. the experimental hydrogen exchange rates (*exp*) for the protein Zn(II) rubredoxin. Zn(II) rubredoxin production one and two are indicated as *ruz1* and *ruz2* respectively.

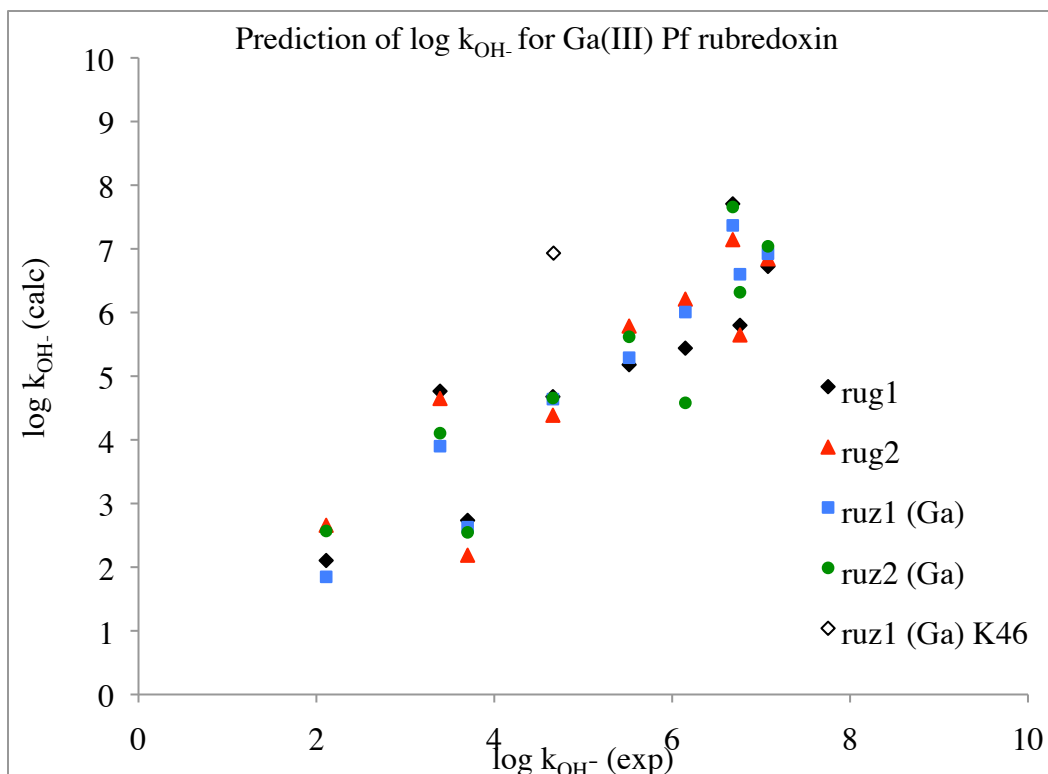


Figure 12: The predicted amide hydrogen exchange rates (*calc*) versus the experimental hydrogen exchange rates (*exp*) for the protein Ga(III) rubredoxin. Ga(III) rubredoxin production one and two are indicated as *rug1* and *rug2* respectively.

As shown in Figures 11 and 12, log k values were calculated for each of the hydrogen exchange rates. For the theoretical hydrogen exchange rates for Zn(II) and Ga(III) rubredoxin, the rates were averaged over both production runs. In addition, Figure 11 shows the comparison of electrostatic calculations from two MD simulations of Zn(II) rubredoxin and for two simulations of Ga(III) rubredoxin with Zn(II) coordination charges substituted for Ga(III) charges, shown as blue squares and green circles respectively. Figure 12 shows the corresponding comparison for Ga(III) rubredoxin.

Because no experimental data was available for Cu(I) and Zn(II) azurin, the theoretical hydrogen exchange rates, averaged over both runs, were analyzed for the α -helix and β -sheet region. Shown in Figure 13, helix 2, is the largest α -helix in azurin.

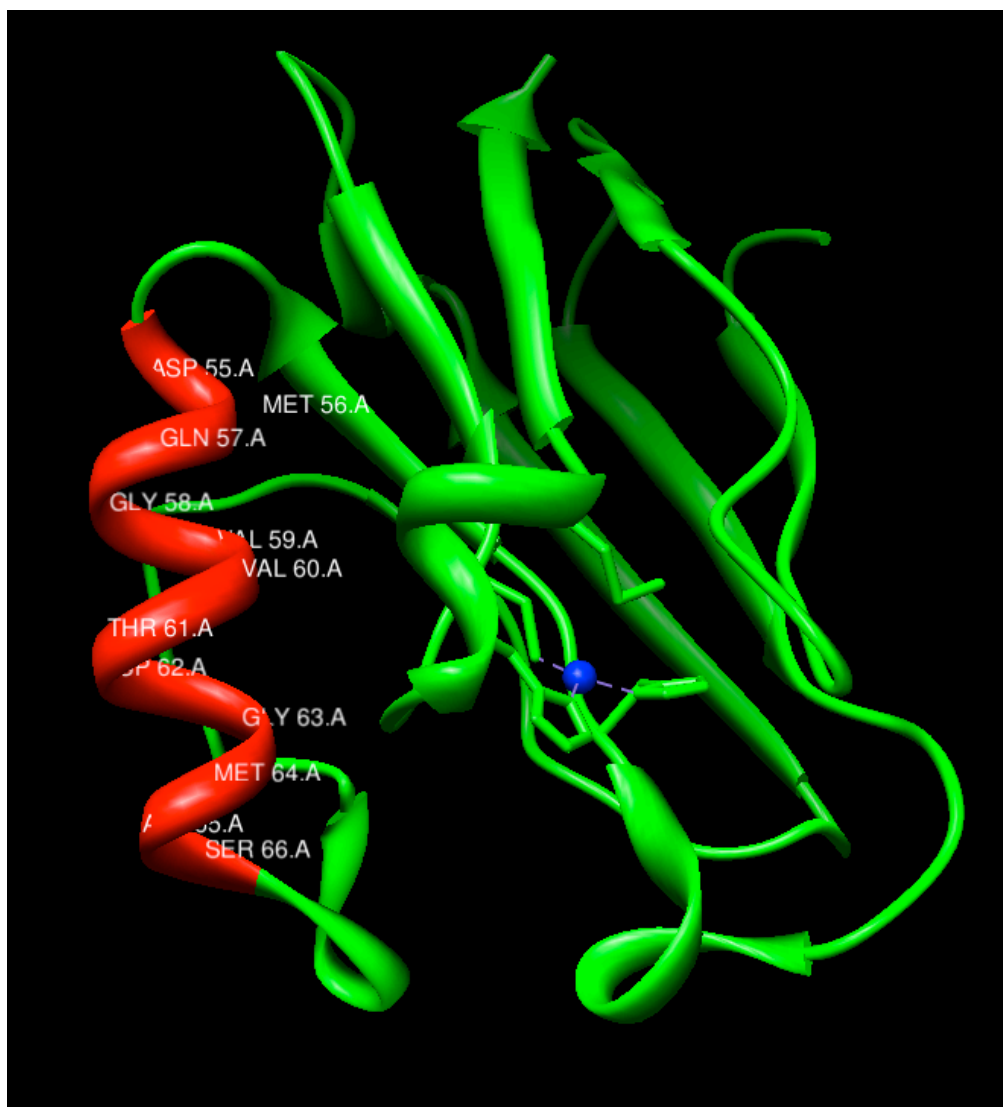


Figure 13: Chimera 3D image of azurin secondary structure. Helix 2 is shown in red. Blue is the metal ion.

Shown in Figure 14, helix 2 is labeled with theoretical hydrogen exchange rates, if the amide was exposed during the MD simulation: if the amide was not exposed, no exchange rates could be calculated.

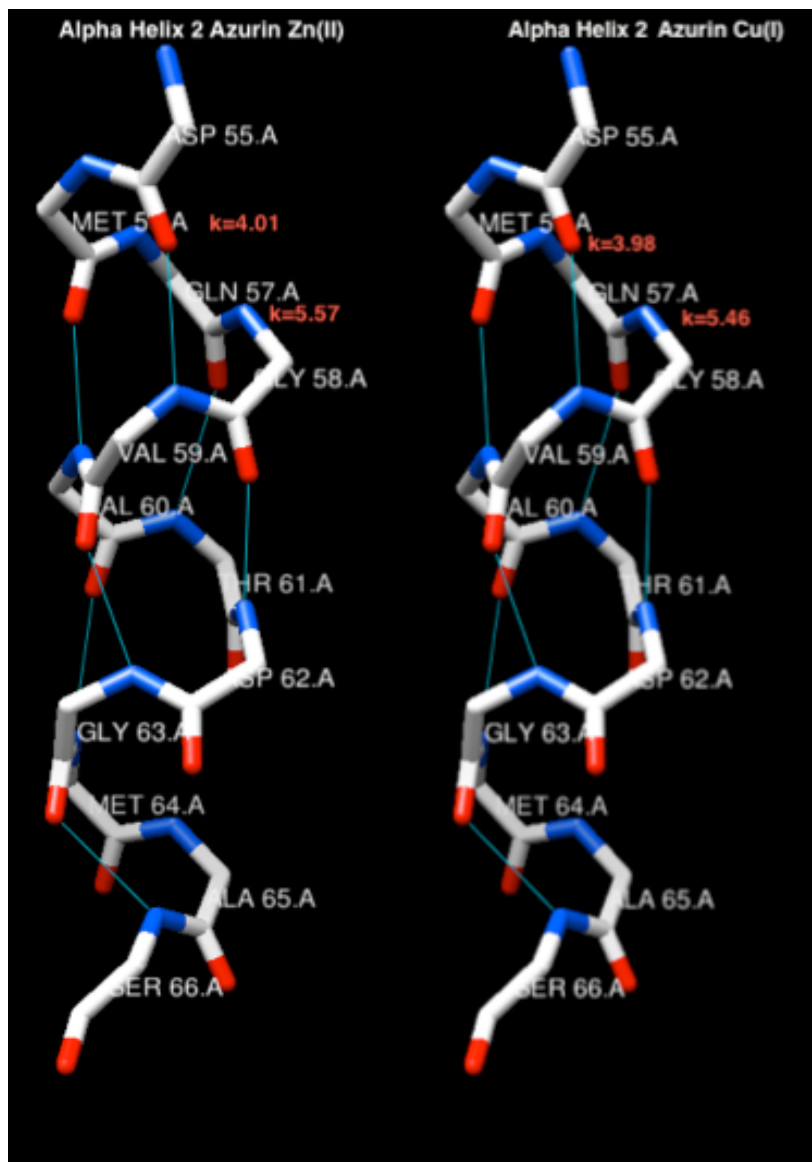


Figure 14: 3D chimera image of the backbone structure of helix 2 of azurin. Hydrogen exchange rates for the exposed amide protons are shown in red and hydrogen bonds are represented in blue lines. Amide nitrogens are represented in blue and carbonyl oxygens are in red.

The largest β -sheet in azurin was also analyzed. Represented in Figure 15 are β -strand 6, 3, and 1 forming azurin's largest β -sheet.



Figure 15: 3D chimera image of the backbone structure of azurin. In blue is the largest β -sheet composed of β -strands 6, 3, and 1. The metal ion is shown in red.

As shown in Figure 15, β -strands 6, 3, and 1 form a mixed anti-parallel and parallel β -sheet where β -strand number 6 and 3 are anti-parallel with each other and β -strand 3 and 1 are parallel. The hydrogen exchange rates were calculated for each of the amide hydrogens located on β -strands 6, 3, and 1 for both Cu(I) and Zn(II) azurin. Represented in Figure 16 are the backbone structure and theoretical hydrogen exchange rates for Cu(I) azurin's largest β -sheet.

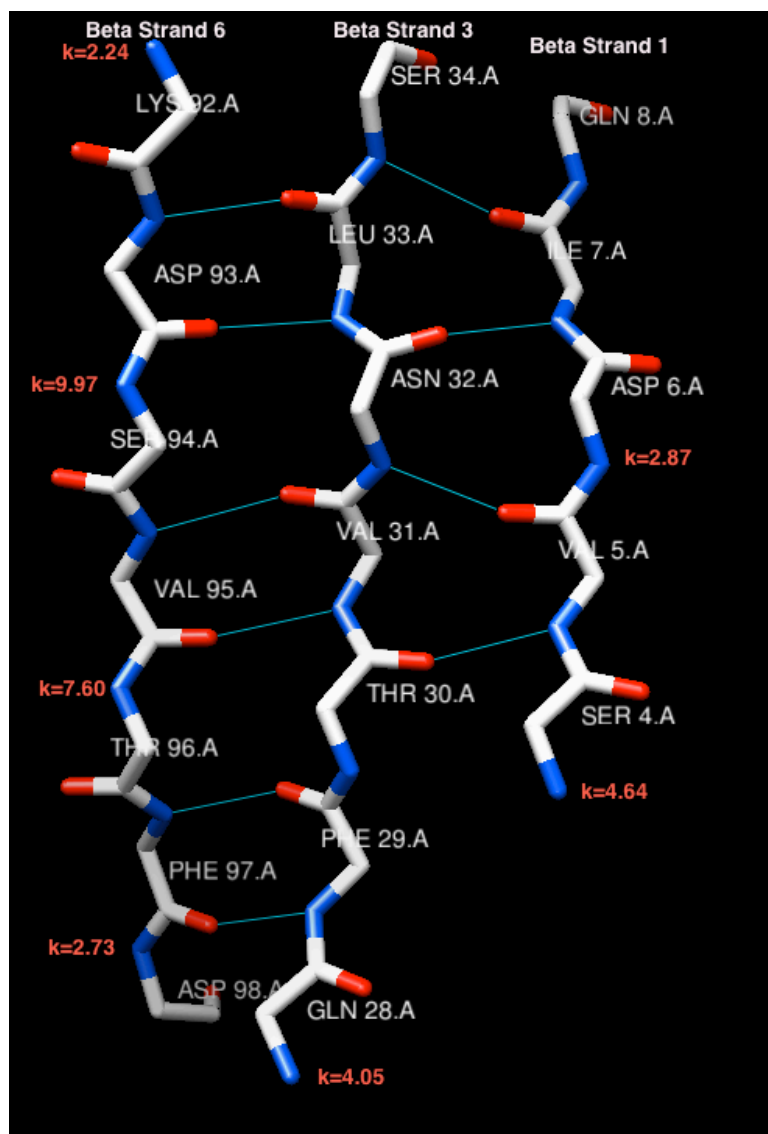


Figure 16: 3D chimera image of the backbone structure for Cu(I) azurin's largest β -sheet. In red are the exposed amide hydrogen's hydrogen exchange rates and hydrogen bonds are shown in blue lines.

Represented in Figure 17 is β -strand 6, 3, and 1 for Zn(II) azurin with its exposed amide hydrogen's theoretical hydrogen exchange rates.

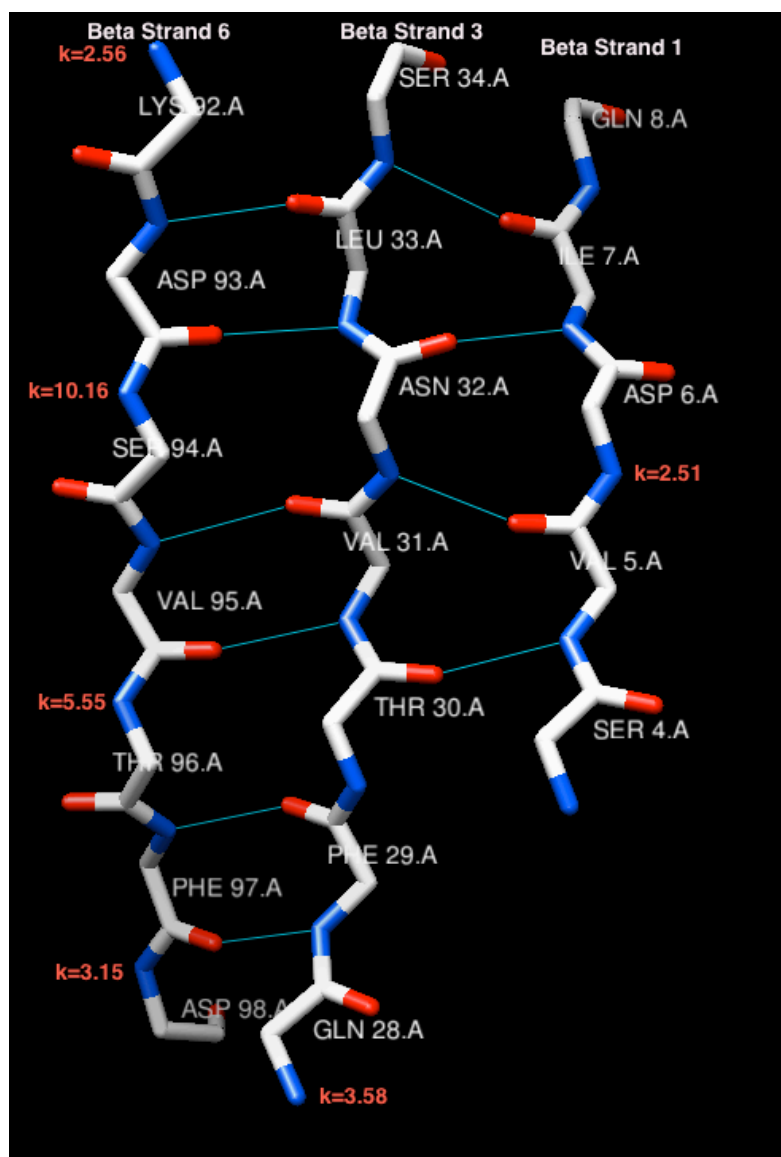


Figure 17: 3D chimera image of the backbone structure for Zn(II) azurin's largest β -sheet. In red are the exposed amide hydrogen's hydrogen exchange rates and hydrogen bonds are shown in blue lines.

Discussion

Each of the protein structures remained stable during their molecular dynamic 50 to 100ns trajectories. For instance, each RMSD plot was relatively flat after an initial increase in RMSD from the first structure. In addition, their average RMSD values were less than 2.0 Å. These results indicate that each molecule was stable throughout the trajectory.

Another indication of stability was the measurement of each protein's metal binding site's geometry over the trajectory. Shown in Figures 9 and 10 are distances and angles for production one Zn(II) azurin metal binding site. The residues within the metal binding site for Zn(II) azurin maintained a stable geometry throughout the trajectory. This is significant because it indicates that Zn(II) azurin's metal binding site was stable throughout its production run and that the parameter's force field, CHARMM22, worked correctly to maintain the binding site stability. Similar plots for the other production runs, shown in the Appendix, were stable.

After undergoing molecular dynamic runs, surfv and Delphi calculations for each protein, their predicted hydrogen exchange rates were determined. Figure 11 and 12 shows theoretical hydrogen exchange rates for rubredoxin compared to the experimental hydrogen exchange rates. These data have a linear correlation indicating that the hydrogen exchange rates determined by the computational methods successfully predicted the rates for the exposed amide hydrogens located on a moderately small thermo stable protein [4]. For both Zn(II) and Ga(III) rubredoxin, lysine 46 was an outlier. Lysine contains a large side chain with a protonated amine group located at the

very end of the chain. As a result, the positive charge on the side chain affected the electrostatic environment around the K46 amide group.

A similar occurrence was seen in previous research with the theoretical amide hydrogen exchange rates of serine and threonine residues. These residues disagreed with experimental results because they have hydroxyl groups, which can readily undergo hydrogen exchange with the solvent. To correct for this, residues with gauche \pm chi1 angles had oxygen atoms replaced with a hydrogen.

Because the same X-ray structure was used for both the Zn(II) and Ga(III) models of rubredoxin, it should be possible to replace the Zn(II) and ligand charges of the two Zn(II) production runs with the Ga(III) ion and ligand charges, and then calculate the exchange rates. When this was done, the rate constants calculated from the Zn(II) trajectories with Ga(III) charges agreed with the experimental Ga(III) rate constants as well as the original calculated values. The other case, switching Ga(III) charges to Zn(II) charges in the Ga(III) rubredoxin trajectories also gave reasonable results which confirmed our assumption that the structure of rubredoxin did not depend on the metal atom charge.

The hydrogen exchange rates of Cu(I) and Zn(II) azurin were examined in relation to their secondary structures. Shown in Figure 14 are the calculated hydrogen exchange rates for the largest α -helix of azurin. The majority of the amide hydrogens did not undergo hydrogen exchange with the solvent because they were either hydrogen-bonded with the carbonyl oxygen to maintain the stability of the α -helix or they were buried inside the protein (such as alanine 65), where the amide hydrogen was not exposed to the solvent in the 1000-2000 frame molecular dynamic trajectories.

A similar trend can be seen with azurin's largest β -sheet, represented in Figures 16 and 17. Amide hydrogen exchange did not occur with the amide protons that maintained the structure of the β -sheet. These amide protons were hydrogen-bonded with the carbonyl oxygens that were directly across from each other. On the other hand, amide protons on the exterior parts of the β -sheet, such as on β -strands 6 and 1 did undergo hydrogen exchange with the solvent. As a result, amide protons that are not buried within the protein and are not hydrogen bonded have a higher hydrogen exchange rate.

Although we expected the hydrogen exchange rates of azurin to depend on the metal ion charge as with rubredoxin, the hydrogen exchange rates of Zn(II) azurin (shown in Figures 16 and 17) were not greatly different from those of Cu(I) azurin. This indicates that the metal ion did not influence azurin's amide hydrogen exchange rates located on α -helices, β -strands and other areas on azurin's backbone. Azurin is a much larger structure in comparison to rubredoxin and the majority of the amide hydrogens undergoing hydrogen exchange rates are located farther away from the metal binding site. As a result, these amide hydrogens do not experience the metal ion charge as much as amide hydrogens from rubredoxin as shown in LeMaster *et al.* [4], which indicated that amide protons that are closer to the metal binding site experience a greater shift in hydrogen exchange rates when the metal ion shifted charges.

In the future, we hope to have experimental data for azurin hydrogen exchange rates to compare to our calculated hydrogen exchange rates and determine whether our models of Cu(I) and Zn(II) azurin are correct.

References

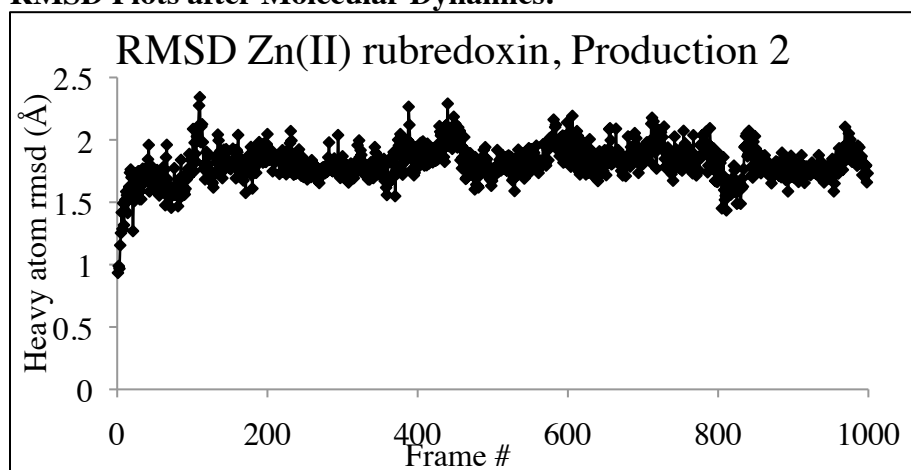
1. Maity, Haripada, Lim K., Woon, Rubley N., Jon and S. Walter Englander. (2002). "Protein hydrogen exchange mechanism: Local fluctuations." *Protein Science*. 12, 153-160.
2. Hvidt, A., and Linderstrom-Lang, K. (1954). "Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions." *Biochem. Biophys. Acta*. 14, 574-575.
3. LeMaster, D.M., Anderson, J.S., and Hernández, G. (2009) "Peptide Conformer Acidity Analysis of Protein Flexibility Monitored by Hydrogen Exchange." *Biochemistry*. 48, 9256-9265.
4. LeMaster, D.M., Anderson, J.S., and Hernández, G. (2006) "The Role of Native-State Structure in Rubredoxin Native-State Hydrogen Exchange", *Biochemistry*. 45, 9956-9963.
5. Bau, R., Rees, D.C., Kurtz, D.M., Scott, R.A., Huang, H., Adams, M.W.W. and M.K. Eidsness. (1998) "Crystal Structure of Rubredoxin from *Pyrococcus Furiosus* at 0.95 Angstroms Resolution, and the structures of N-terminal methionine and formylmethionine variants of Pf Rd. Contributions of N-terminal interactions to thermostability." *J. Biol.Inorg.Chem*. 484-493.
6. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. J Computational Chem. (2004). 25(13), 1605-1612.
7. Anderson, Janet S.; Hernández, Griselda; LeMaster, David M. (2008). "A Billion-fold Range in Acidity for the Solvent-Exposed Amides of *Pyrococcus furiosus* Rubredoxin", *Biochemistry*. 47, 6178-6188.
8. LeMaster, D.M., Minnich, M., Parsons, P.J., Anderson, J.S., and Hernández, G. (2006) "Tetrathiolate Coordination of Germanium (IV) in a Protein Active Site", *J. Inorg. Biochem*. 100, 1410-1412.
9. Rizzuti, Bruno, Swart, Marcel, Sportelli, Luigi and Rita Guzzi. (2003). "Active Site Modeling in Copper Azurin Molecular Dynamics Simulations." *J mol model*. 10(1), 25-31.
10. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) "Scalable molecular dynamics with NAMD." *J. Comp. Chem*. 26, 1781-1802.
11. Rocchia, W., Alexov, E., Honig, B. (2001) "Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions" *J. Phys. Chem. B*. 105, 6507-6514.
12. Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., and Honig, B. (2002) "Rapid Grid-based Construction of the Molecular Surface for both Molecules and Geometric Objects: Applications to the Finite Difference Poisson-Boltzmann Method" *J. Comp. Chem*. 23, 128-137.
13. Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M. and G.W. Canters. (1991) "X-ray crystal structure of the two site-specific mutants His35Gln and His35Leu of azurin from *Pseudomonas aeruginosa*." *J.Mol.Biol*. 218, 427-447.
14. AutoCAD. (2009). 19(10), 1203-43

Acknowledgments

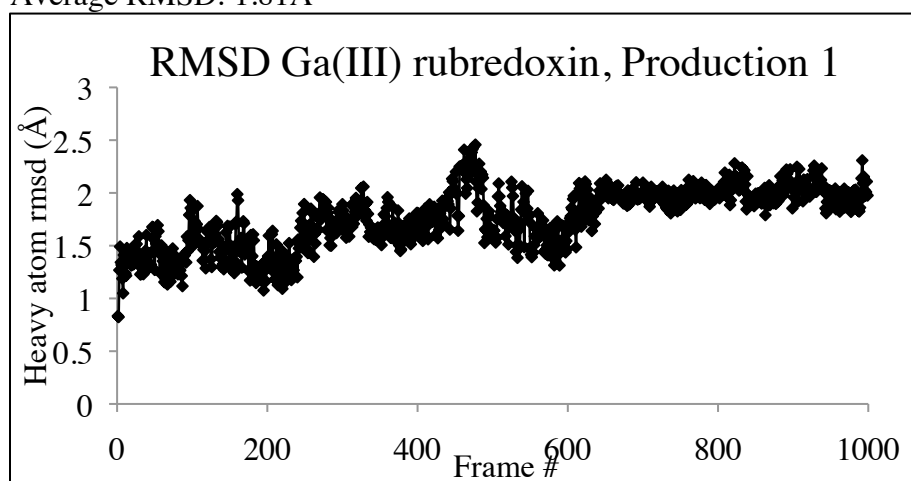
I would like to give special thanks to Professor Janet Anderson for her guidance and support throughout this research project. Her assistance with debugging the countless number of Fortran programs and imparting the methodology of this research is more than greatly appreciated. I would also like to thank Dr. David LeMaster and Griselda Hernández for their amide hydrogen experimental data conducted by the Wadsworth Center of the New York State Department of Health in Albany, New York, and also for their helpful discussions. Finally, Union College's IMB intelligent cluster has been of immeasurable help to my project. This research would not be where it is today without the high computing power for molecular dynamic simulations and Delphi calculations.

Appendix

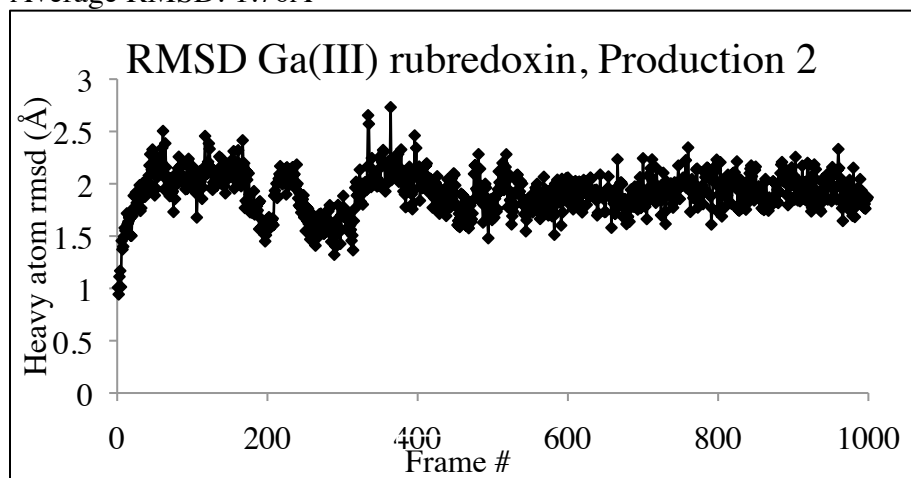
RMSD Plots after Molecular Dynamics:



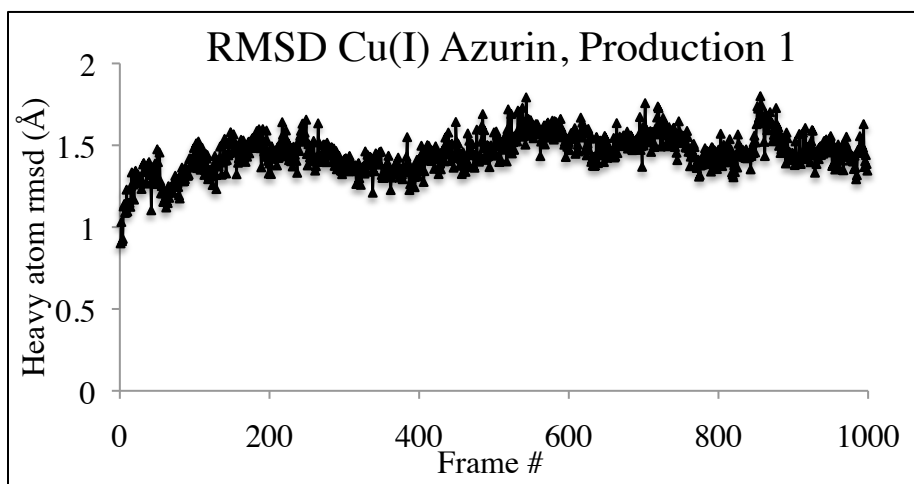
Average RMSD: 1.81Å



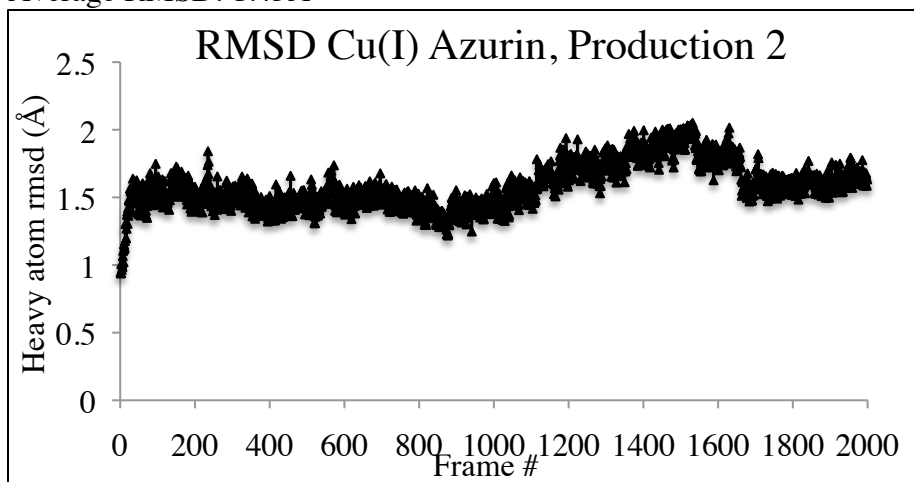
Average RMSD: 1.76Å



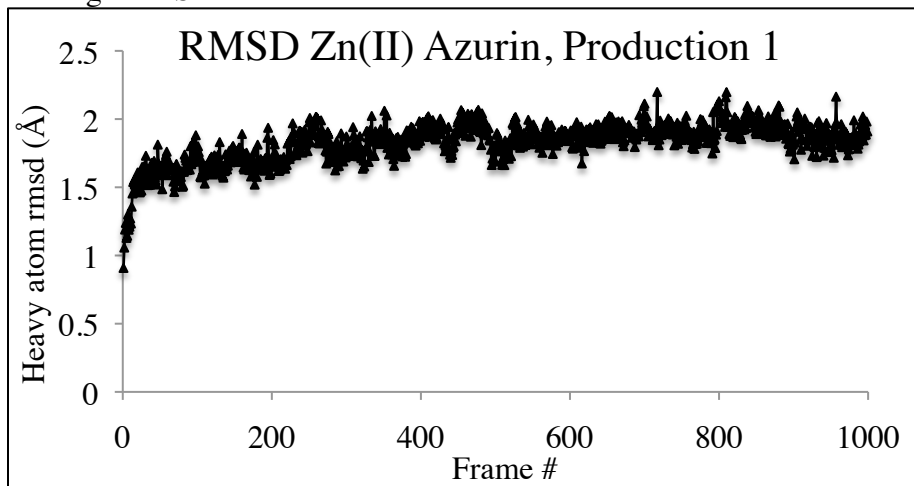
Average RMSD: 1.91Å



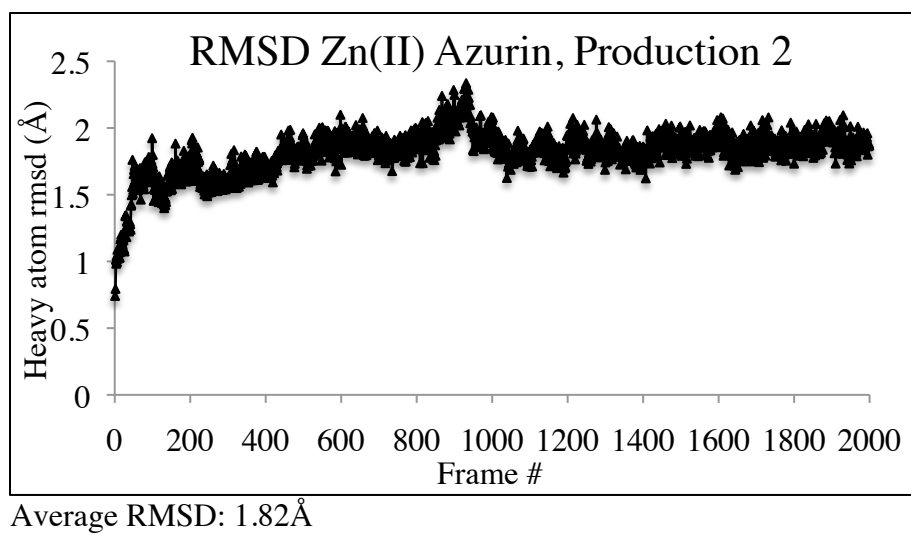
Average RMSD: 1.45Å



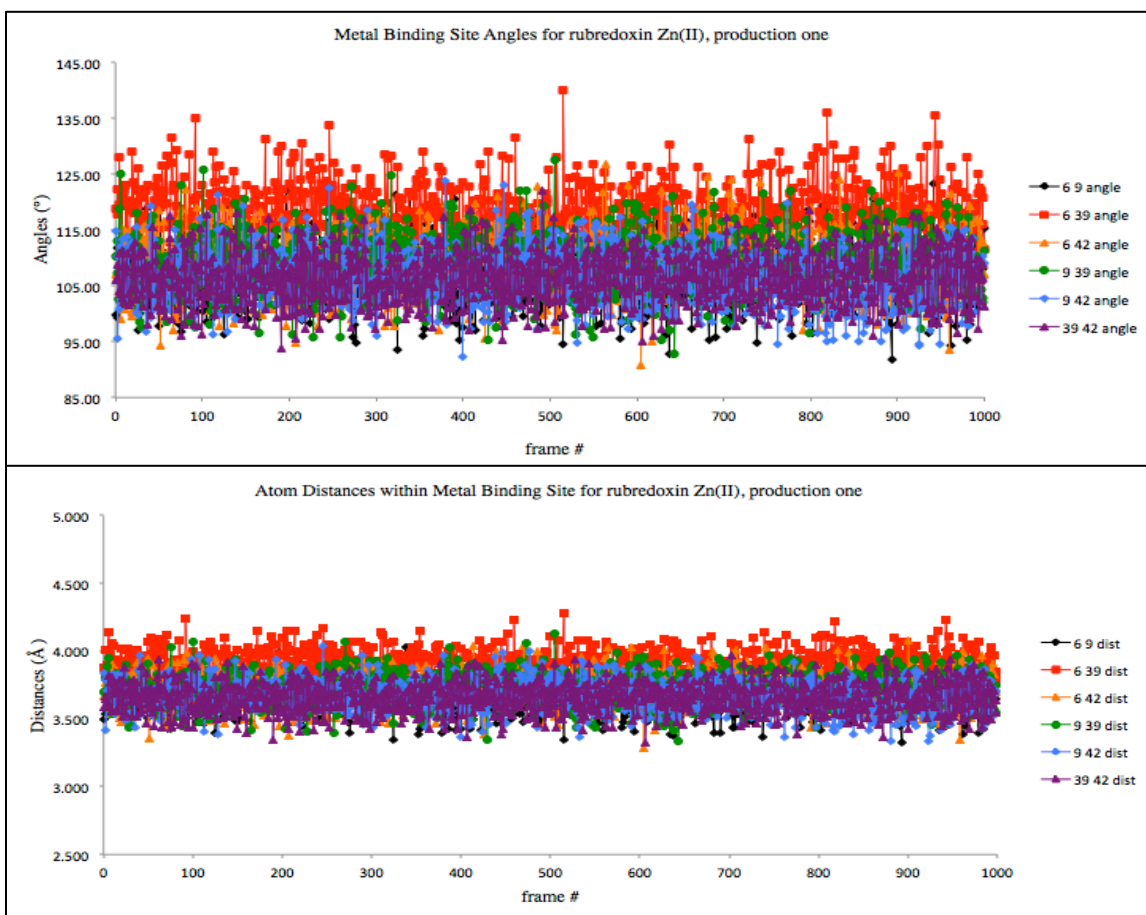
Average RMSD: 1.41Å



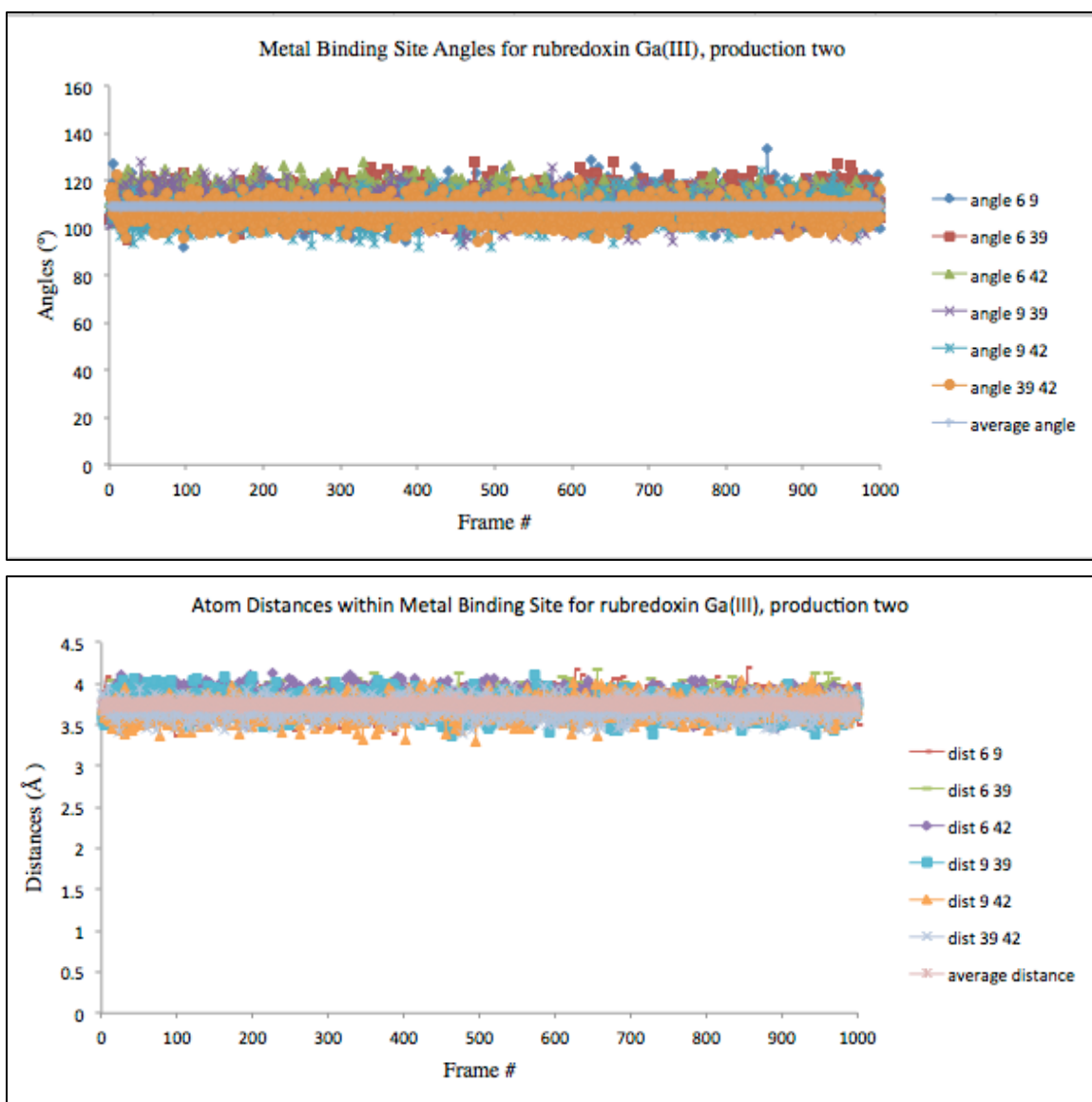
Average RMSD: 1.83Å



Zn(II) and Ga(III) rubredoxin Metal Binding Site's Angles and Distances after MD:

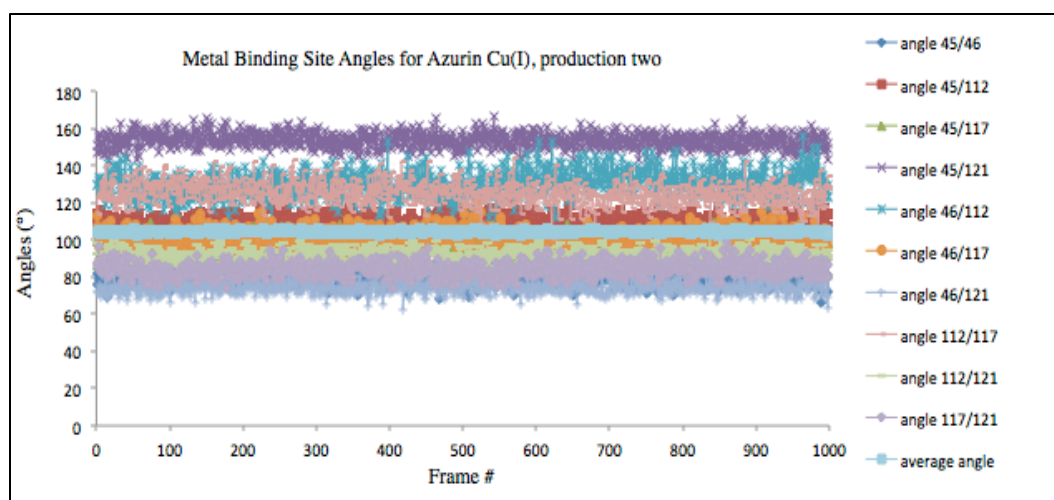
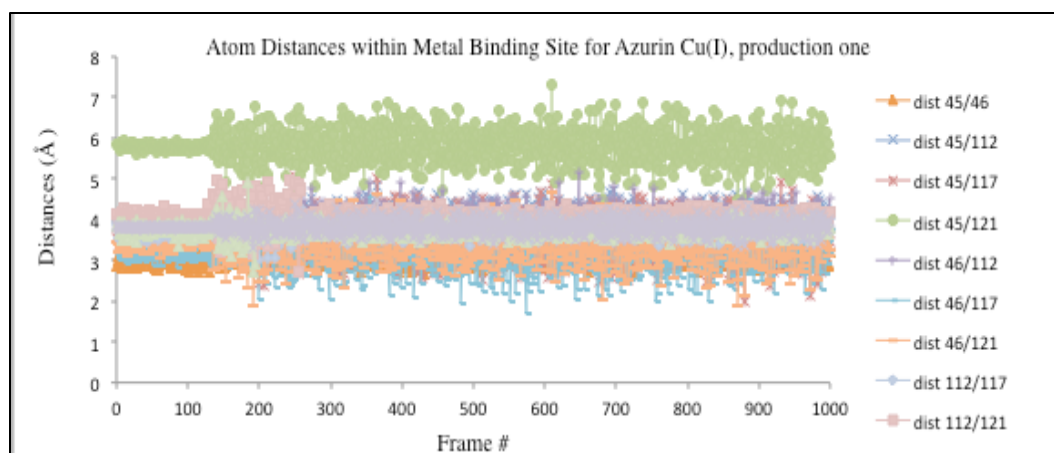
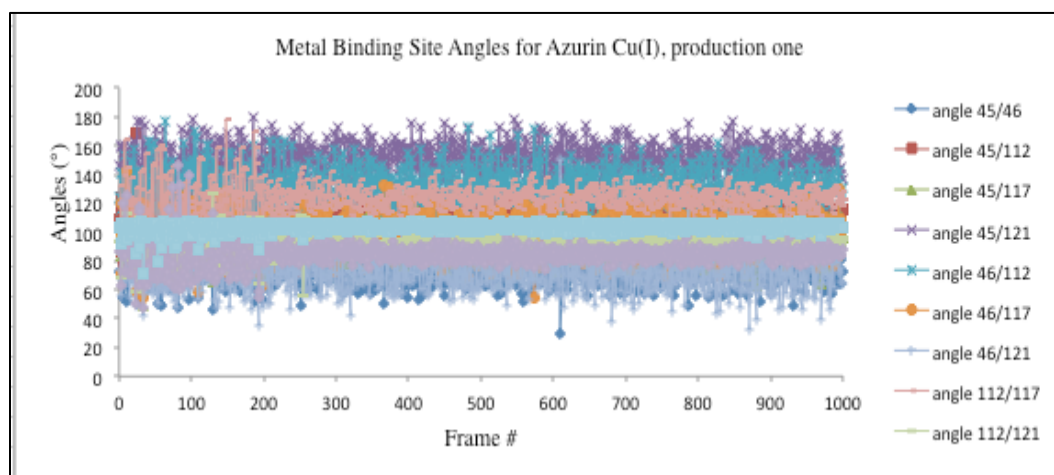


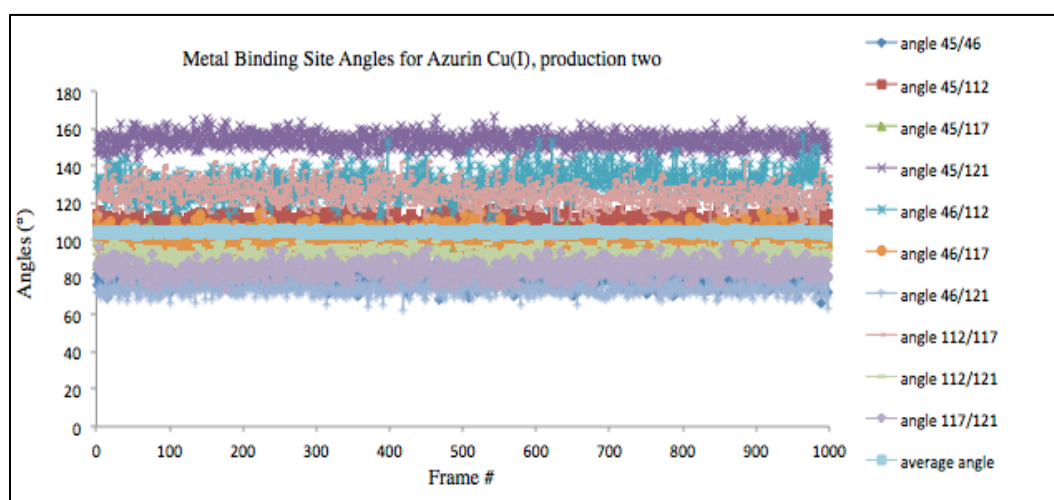
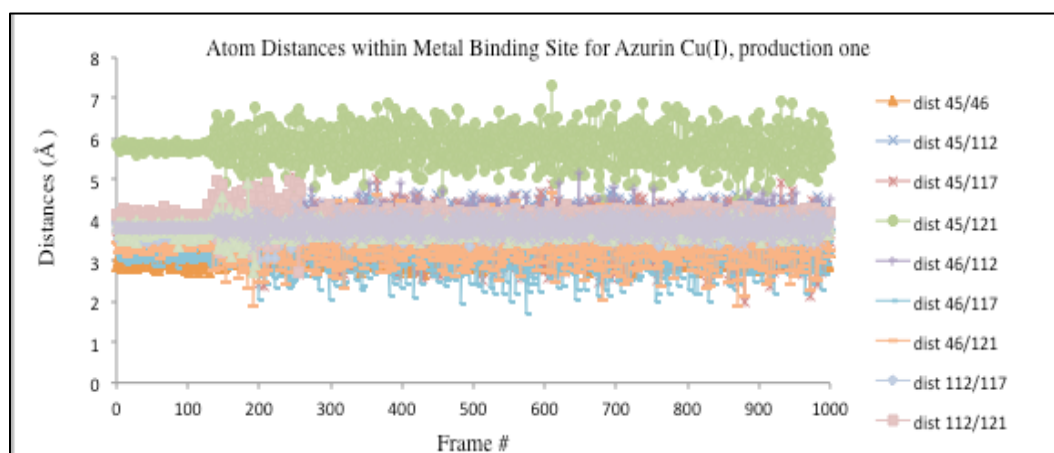
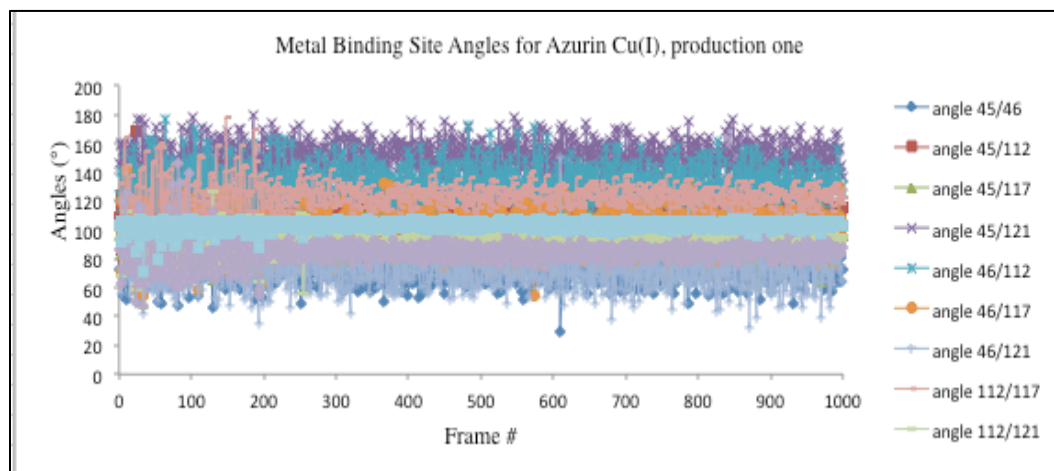
Note: Rubredoxin Zn(II) production two metal binding site's distances and angles (Not shown) had a similar trend as shown in rubredoxin Ga(III) production one metal binding distances and angles.



Note: Rubredoxin Ga(III) production one metal binding site's distances and angles (Not Shown) had a similar trend as shown in rubredoxin Ga(III) production two metal binding distances and angles.

Cu(I) and Zn (II) azurin Metal Binding Site's Angles and Distances after MD:





Setting up NAMD for Molecular Dynamic Simulations:

1. Download 3azu.pdb and 1bq8.pdb from PDB for azurin and rubredoxin respectively. Use chain A for 3azu.pdb.
2. Make patches ZNC4, GAC4, CUZ5 and ZNZ4 for Zn^{2+} rubredoxin, Ga^{3+} rubredoxin, Cu^{+1} azurin and Zn^{2+} azurin respectively. Patches came from the topology files top_all27_prot_lipid_az_v2.inp and top_all27_prot_lipid_zn2_ga3.inp.

Patch ZNC4:

```

PRES ZNC4      -2.00 ! patch for [Zn(Cys)4]2-. Patch must be 1-ZN and 2-CYS 3-CYS 4-CYS 5-CYS
                ! use CY2 CY2 CY2 CY2
                !
                ! use in a patch statement
                !
                !      -4CB--4SG  5SG--5CB-
                !          \ /
                !          1ZN
                !          / \
                !      -2CB--2SG  3SG--3CB-
                !
GROUP          ! zinc ion (read in separate segment)
ATOM 1ZN  ZN   1.24
!GROUP        ! cysteine ligand2
ATOM 2CB  CT2  -0.38
ATOM 2SG  SSS  -0.61
!GROUP        ! cysteine ligand3
ATOM 3CB  CT2  -0.38
ATOM 3SG  SSS  -0.61
!GROUP        ! cysteine ligand4
ATOM 4CB  CT2  -0.38
ATOM 4SG  SSS  -0.61
!GROUP        ! cysteine ligand5
ATOM 5CB  CT2  -0.38
ATOM 5SG  SSS  -0.61
BOND 1ZN 2SG
BOND 1ZN 3SG
BOND 1ZN 4SG
BOND 1ZN 5SG

```

Patch GAC4:

PRES GAC4 -1.00 ! patch for [Ga(Cys)4]-. Patch must be 1-GA and 2-CYS 3-CYS 4-CYS 5-CYS

! use CY3 CY3 CY3 CY3

!

! use in a patch statement

!

! 4HB1,4HB2-4CB--4SG 5SG--5CB-5HB1,5HB2

! \ /

! 1GA

! / \

! 2HB1,2HB2-2CB--2SG 3SG--3CB-3HB1,3HB2

!

GROUP ! gallium ion (read in separate segment)

ATOM 1GA GA 1.24

!GROUP ! cysteine ligand2

ATOM 2CB CT2 -0.34

ATOM 2SG SSG -0.42

ATOM 2HB1 HA 0.10

ATOM 2HB2 HA 0.10

!GROUP ! cysteine ligand3

ATOM 3CB CT2 -0.34

ATOM 3SG SSG -0.42

ATOM 3HB1 HA 0.10

ATOM 3HB2 HA 0.10

!GROUP ! cysteine ligand4

ATOM 4CB CT2 -0.34

ATOM 4SG SSG -0.42

ATOM 4HB1 HA 0.10

ATOM 4HB2 HA 0.10

!GROUP ! cysteine ligand5

ATOM 5CB CT2 -0.34

ATOM 5SG SSG -0.42

ATOM 5HB1 HA 0.10

ATOM 5HB2 HA 0.10

BOND 1GA 2SG

BOND 1GA 3SG

BOND 1GA 4SG

BOND 1GA 5SG

Patch CUZ5:

```
PRES CUZ5      +0.00 ! patch for [Zn(Gly 45 O)(His 46 ND1)(Cys 112 SG)
                ! (His 117 ND1)(Met 121 SD)]+1. Patch must be 1-ZN and
                ! 2-GLY 3-HIS46 4-CYS 5-HIS117 6-MET
                ! use GLZ for Gly 45
                ! use HZ1 for His 46 (has HE2 but not HD1)
                ! use CYZ for Cys 112 (no HG, -1 charge)
                ! use HZ2 for His 117 (has HE2 but not HD1)
                ! use MEZ for Met 121

                !
                ! use in a patch statement
                !
                !      -4CB--4SG  5ND1--5CG-
                !          \ /
                !          1Cu -- 6SD--6CG
                !          / \
                !      -2C--2O  3ND1--3CG-
                !

GROUP          ! zinc ion (read in separate segment)
ATOM 1CU  CU   0.414
!GROUP        ! glycine 45 ligand2
ATOM 2C  C     0.51
ATOM 2O  OZ   -0.457
!GROUP        ! histidine 46 ligand3
ATOM 3CG  CPH1  0.22
ATOM 3ND1 NR2 -0.555
!GROUP        ! cysteine 112 ligand4
ATOM 4CB  CT2  -0.11
ATOM 4SG  SSZ  -0.88
!GROUP        ! histidine 117 ligand5
ATOM 5CG  CPH1  0.22
ATOM 5ND1 NR2 -0.555
!GROUP        ! methionine 121 ligand6
ATOM 6CG  CT2  -0.14
ATOM 6SD  SMZ  -0.037
BOND 1CU 2O
BOND 1CU 3ND1
BOND 1CU 4SG
BOND 1CU 5ND1
BOND 1CU 6SD
```


Patch ZNZ4:

```
PRES ZNZ4      +1.00 ! patch for [Zn(Gly 45 O)(His 46 ND1)(Cys 112 SG)
                ! (His 117 ND1)(Met 121 SD)]+1. Patch must be 1-ZN and
                ! 2-GLY 3-HIS46 4-CYS 5-HIS117 6-MET
                ! use GLZ for Gly 45
                ! use HZ1 for His 46 (has HE2 but not HD1)
                ! use CYZ for Cys 112 (no HG, -1 charge)
                ! use HZ2 for His 117 (has HE2 but not HD1)
                ! use MEZ for Met 121

                !
                ! use in a patch statement
                !
                !      -4CB--4SG  5ND1--5CG-
                !          \  /
                !          1ZN -- 6SD--6CG
                !          /  \
                !      -2C--2O  3ND1--3CG-
                !

GROUP          ! zinc ion (read in separate segment)
ATOM 1ZN  ZN   1.414
!GROUP        ! glycine 45 ligand2
ATOM 2C  C     0.51
ATOM 2O  OZ    -0.457
!GROUP        ! histidine 46 ligand3
ATOM 3CG  CPH1  0.22
ATOM 3ND1 NR2  -0.555
!GROUP        ! cysteine 112 ligand4
ATOM 4CB  CT2   -0.11
ATOM 4SG  SSZ   -0.88
!GROUP        ! histidine 117 ligand5
ATOM 5CG  CPH1  0.22
ATOM 5ND1 NR2  -0.555
!GROUP        ! methionine 121 ligand6
ATOM 6CG  CT2   -0.14
ATOM 6SD  SMZ   -0.037
BOND 1ZN 2O
BOND 1ZN 3ND1
BOND 1ZN 4SG
BOND 1ZN 5ND1
BOND 1ZN 6SD
```

3. Generate the protein structure file (psf file) for chain A of 3azu.pdb (H35Q azurin with Cu^{2+}), but use Zn^{2+} instead of Cu^{2+} .

a. Put the protein, metal, and crystal waters in separate pdb files. Remove all lines except ATOM lines for protein and HETATM lines for metal and crystallographic water. If there are multiple conformations, keep only one and convert to CHARMM format.

Change to CHARMM Format for Azurin:

- Change CD1 of ILE to CD.
- Change C-terminal O and OXT to OT1 and OT2.
- Change GLY 45 to GLZ 45 for azurin patch.
- Change HIS 46 to HZ1 46 for azurin patch.
- Change CYS 112 to CYZ for azurin patch.
- Change HIS 117 to HZ2 117 for azurin patch.
- Change MET 121 to MEZ 121 for azurin patch.
- Change HIS 83 to HSE 83.
- Change water residue to TIP3 and water O atoms to OH2; change HETATM to ATOM.
- Add TER at the end of each pdb file.
- Edit the metal pdb file.

EXAMPLE:

```
ATOM 3897 ZN ZN A 129 29.493 19.064 -10.206 1.00 21.89 Zn
TER 3898
```

- Put topology files in directory and patch in the respective patches.
- Open Terminal and cd to directory where your files are. Make a text file 3azu.pgn with the following lines:

```
package require psfgen
topology top_all27_prot_lipid_az_v2.inp
pdbalias atom ILE CD1 CD
pdbalias residue ZN ZN2
segment 3azu {pdb 3azu_A_proteinonly.pdb}
segment metl {pdb 3azu_A_zn.pdb}
patch ZN4 metl:129 3azu:45 3azu:46 3azu:112 3azu:117 3azu:121
patch DISU 3azu:3 3azu:26
coordpdb 3azu_A_proteinonly.pdb 3azu
coordpdb 3azu_A_zn.pdb metl
segment xwat {
  auto none
  pdb 3azu_A_xtalwater.pdb
}
coordpdb 3azu_A_xtalwater.pdb xwat
guesscoord
writepdb 3azu_1.pdb
writepsf 3azu_1.psf
```

- Save the text file. To run the script, in the Terminal window:
vmd -dispdev text -e 3azu.pgn
- Edit pdb file (protein and xtal water segments) to replace with occupancy of 1.00. Ex: (1.00 0.00)
- Exit the vmd > prompt.

4. Solvate with water box and add ions:

- Open up graphical VMD and read in combined protein, Zn, and crystal water psf and pdb files.
- In VMD, select Extensions → Tk Console → cd to the appropriate directory.
- Enter in Tk Console:

```

package require solvate (loads version 1.4)
solvate 3azu_1.psf 3azu_1.pdb -t 10 -o 3azu_wb
set everyone [atomselect top all]
measure minmax $everyone
measure center $everyone
    • Add ions with autoionize plugin, version 1.3 by entering in Tk-console:
package require autoionize
autoionize -psf 3azu_wb.psf -pdb 3azu_wb.pdb -sc 0.15 -o 3azu_wb_ion -between 8
    • Copy new pdb and psf files into new directory "common."

```

5. Do short minimization with constrained protein and crystal water.

- Copy new pdb file to constraint xwat.pdb file and put 5.0 in the B-factor column of all heavy atoms of protein and crystal waters.
- Sorting files by atom name:

Example:

```

sort -k3 (ion pdb) > sorted1.pdb
    • edit sorted1.pdb: For heavy atoms (C, N, O, S) in protein and xray water, replace 1.00 0.00 with 1.00 5.00
sort -k2 sorted1.pdb > sorted2.pdb
    • edit sorted2.pdb to put END in last place and CRYST1 in first place.
mv sorted2.pdb 3azu_constraint_5prot_xwat.pdb
    • Put parameter file par_all27_prot_lipid_az.prm in common directory.
    • Create metal binding site geometry constraint file using 3azu or 1bq8 for distances between ligands.

```

Example:

```

(contents of file azurin_metal_site_zn.txt)
bond 662 669 20 2.831
bond 662 1672 20 3.428
bond 662 1737 20 3.180
bond 662 1794 20 5.365
bond 669 1672 20 3.884
bond 669 1737 20 3.237
bond 669 1794 20 3.237
bond 1672 1737 20 3.782
bond 1672 1794 20 4.450
bond 1737 1794 20 3.622
#
# Add ligand to ligand bond length restraints to maintain geometry
extrabonds on
extrabondsfile ../common/azurin_metal_site_zn.txt
    • Edit 3azu_min.conf for file names and for size of box:
# Periodic Boundary Conditions (insert numbers from VMD)
cellBasisVector1 59.04 0.0 0.0
cellBasisVector2 0.0 59.04 0.0
cellBasisVector3 0.0 0.0 54.59
# PME (for full-system periodic electrostatics)
PME yes
PMEGridSizeX 64
PMEGridSizeY 64
PMEGridSizeZ 64
nohup /usr/local/NAMD_2.8_MacOSX-x86-multicore/namd2 +p4 3azu_min.conf > 3azu_min.log &

```

6. Heat from 0 to 298 K in steps of 25 K with constrained protein and crystal water.

Example: Enter-> `namd2 +p4 3azu_heat2.conf > 3azu_heat2.log &`

7. Do run where constraints are decreased over time using Union College's IBM cluster. Use 1000 steps x 11 to reduce constraints to zero; then 1×10^6 steps of equilibrium without constraints (1 ns).

- Create a new folder with dec_constraints output files, ion.psf file and parameters file for each designated protein.
- Transfer the newly made folder into Union College's IBM donated cluster.
- Run the MD production run for 1000-2000 frames using the program runnamd2 (Enter: `squb runnamd2`).

Setting up Delphi Calculations:

1. Convert NAMD dcd file to concatenated pdb file (999 pdb files) of protein atoms (no ions or solvent water).

- Open VMD, load psf and dcd files.
- In VMD Main window, highlight the name of the .psf file.
- In VMD Main window, File/Save Coordinates. In Save Trajectory window, Selected atoms: protein or rename ZN2 (Ga: protein or rename GA3)
- In Frames, First: 1
Last: 999 (these numbers will depend on the length of the MD run)
Stride 1
- Save (eg. azz1_30_999_mdresults.pdb)
- To check to see how many frames were saved, open the new pdb file with TextWrangler. Search/Find for END, Find All. Should be 999 copies of the END line.

2. While in VMD, use the Extensions/Analysis/RMSD Trajectory Tool to calculate the rmsd over the trajectory (“protein or rename ZN2” or “protein or rename GA3”).

3. While in VMD, calculate the S and T chi 1 dihedral angles.

- Use tk-console and cd to directory with psf dcd files and tcl program.
- Calculate the angles. Ex: source calc_STchi1_script.tcl
- Start Excel, then open each .txt file with Excel.
- Use Delimited/Space, then Copy, Paste /Special/Transpose to put the chi 1 angles for a given residue in a column. Be sure that they are in the same order as the residue numbers, ex. Column 1 S004, column 2 T017, column 3 T021, etc. Make each number have two decimal places and each column width 10.
- In Excel, Save As/Format: Space Delimited Text (filename.prn)
- Open this filename.prn file in TextWrangler, then Save As: filename.txt, Line breaks Unix (LF).
- Save chi1 anlge TextWrangler (LF) file.

4. Calculate the metal binding site distances and angles using the program calc_metal_angles_dist_azurin.f for azurin. (Note: no program for rubredoxin)

Input to the program:

- the combined pdb file
- the number of conformations in the file (usually 999)
- the number of atoms in each conformer (check the file—usually 1929 for 3azu azurin)
- Start Excel, then open the angle_dist text file with Excel. Use Text to Columns. Plot each angle and each distance vs. conformation #. Save as an Excel file.

5. Use program “center_999_pdb_fix.f” to re-center protein.

- To compile: in a Terminal window, type the line below and hit Return.
f95 center_999_pdb_fix.f -o center_999_pdb_fix
- To run the program, type the line below and hit Return.
./center_999_pdb_fix

Input:

- name of combined pdb file
- # protein atoms + metal ions + END lines
- name for centered pdb file
- name for center of mass info

6. Use “separate_pdbs_delphi_2.f” to write individual pdb files for the surfv program.

Input:

- # of conformations

- protein descriptor
- identifying suffix
- name of centered pdb file

Output:

- List of pdbfilenames "pdbfilename."
- List of atmnamefiles "atmnamefile."

7. In the same directory as the pdb files, put the parameter file "LeeRich.siz", run the "run_surfv2."

./run_surfv2

[script run_surfv2:]

```
#!/bin/sh
# This script (run_surfv2) uses files
# to run the surfv program.

exec 3< 'pdbnamefile'

while read <&3 file1name
do
echo $file1name
/usr/local/surfv_dir/SURFV/surfv -s 3 1.4 LeeRich_az.siz "$file1name"
0      1 1 > "$file1name".out"
done
```

8. Run the program "ensembleSASA_azurin.f" to extract the surface-accessible area information.

Input:

- # of pdb files
- # of atoms in each pdb file
- surfv atm filename list (ex. **Atmnamefile**)
- name of outputfile

Output:

- The first line of **surfv_azz1.txt** lists the exposed amides to be used in Delphi calculations.
- Each residue has two columns:
 1. First column has 1 if solvent-accessible (area > 0.5 Å²).
 2. Second column has 1 if accessible at all (area > 0.0 Å²).

NOTE:

- Count the **number of exposed amides** in the first line of surfv file.
- Clean up the working directory by making a directory for the files generated by surfv and move those files into it.

Example:

```
mkdir surfv_output_files
mv *_298* surfv_output_files
```

9. Make a text file with the protein sequence listed by three-letter code by selecting the lines in the original pdb file found in: <http://www.rcsb.org/pdb/home/home.do>, search for 3azu and 1bq8.

Example:

(contents of 3azu_sequence.txt)

```
SEQRES 1 A 128 ALA GLN CYS SER VAL ASP ILE GLN GLY ASN ASP GLN MET
SEQRES 2 A 128 GLN PHE ASN THR ASN ALA ILE THR VAL ASP LYS SER CYS
SEQRES 3 A 128 LYS GLN PHE THR VAL ASN LEU SER GLN PRO GLY ASN LEU
SEQRES 4 A 128 PRO LYS ASN VAL MET GLY HZ1 ASN TRP VAL LEU SER THR
```

```

SEQRES   5  A   128   ALA ALA ASP MET GLN GLY VAL VAL THR ASP GLY MET ALA
SEQRES   6  A   128   SER GLY LEU ASP LYS ASP TYR LEU LYS PRO ASP ASP SER
SEQRES   7  A   128   ARG VAL ILE ALA HSE THR LYS LEU ILE GLY SER GLY GLU
SEQRES   8  A   128   LYS ASP SER VAL THR PHE ASP VAL SER LYS LEU LYS GLU
SEQRES   9  A   128   GLY GLU GLN TYR MET PHE PHE CYS THR PHE PRO GLY HZ2
SEQRES  10  A   128   SER ALA LEU MET LYS GLY THR LEU THR LEU LYS

```

10. Run program make_dir_labels_azurin.f.

Input:

- name of surfv output file
- # of exposed amides
- protein sequence file name

Output:

- new_dir_labels (# of anion directories)
- all_ST_res (res # of all S and T in protein. Ex: 20 for azurin and 4 for rubredoxin)

11. Transfer programs and files to Union College's IBM donated cluster.

- Create a folder and place the programs and files: charm22_11.crg, charm22_11.siz, all_ST_res, centered concatenated pdb file, chi1 angle text file, new_dir_labels, surfv output file, make_directories_accF, directory_names and make_delphi_input_10.f.
- Tar the newly created folder and transfer it to Union College's IBM donated cluster.

12. Run the script "make_directories_accF" with the **directory_names** list to create the directories in Union College's IBM donated cluster.

Example of make_directories_accF:

```

#  !/bin/sh
#  This script (make_directories_dl1) creates a directory for every
#  name in the directory_names list, and puts the parameter files in
#  each directory.
exec 3< 'directory_names'
while read <&3 file1name
do
#  echo $file1name
mkdir "$file1name"
done

```

13. Run the program make_delphi_input_10.f, in Union College's IBM donated cluster, to generate pdb files for amide anions of solvent-exposed amides for the protein. (Only those conformations with exposed amides will have prm and complete pdb files created. The rest will have prm files and single-line pdb files.)

Input:

- Name of centered pdb file
- 1 for metal; 0 for no metal
- Name of metal
- Name of output file generated by "ensembleSASA_azurin.f"
- Protein identifier
- Number of conformations

- Anion information file
- S and T chi1 values
- Enter the number of S and T residues
- Two files are read in without prompts: (**all_ST_res** and **ST_mapping**)

Output:

- cluster_delphi_names, start_delphi_all and run_grep.

14. Start Delphi in Union College's IBM donated cluster:

- Run the program, start_delphi_all with the script run_delphi. Ex: nohup ./start_delphi_all

15. Take the Delphi information and transfer it to home computer.

- Run the program run_grep to extract data from the newly make .out files from delphi.
- Create a new file with all of the grep output information.
- Tar the newly created folder and transfer it to home computer.

16. Run the program, read_grep_allcalc_azurin_v4.f, to extract the grep file information into exchange rates.

Input:

- Surfv output information, chi1 angle file, ST_mapping, directory_names and grep directory files.

Output:

- Combined_data: Exchange rate information from the exposed amide hydrogens.

17. Use Excel to analyze the exchange rate information from combined_data file.