Grand Valley State University
ScholarWorks@GVSU

Student Summer Scholars Manuscripts | Student Summer Scholars

9-3-2021

Computational Analysis of BshC, an Enzyme Responsible for Antibiotic Resistance in Firmicute Bacteria

Matthew R. Martin
Grand Valley State University, martima6@mail.gvsu.edu

Follow this and additional works at: https://scholarworks.gvsu.edu/sss

 Part of the Enzymes and Coenzymes Commons

ScholarWorks Citation

Martin, Matthew R., "Computational Analysis of BshC, an Enzyme Responsible for Antibiotic Resistance in Firmicute Bacteria" (2021). *Student Summer Scholars Manuscripts*. 218.
https://scholarworks.gvsu.edu/sss/218

This Open Access is brought to you for free and open access by the Student Summer Scholars at ScholarWorks@GVSU. It has been accepted for inclusion in Student Summer Scholars Manuscripts by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

Computational Analysis of BshC,

an Enzyme Responsible for Antibiotic Resistance in Firmicute Bacteria

Matthew Martin, Dr. Mary Karpen- Grand Valley State University

Abstract:

Bacillithiol is a compound synthesized by certain gram-positive bacteria called firmicutes, such as *B. Subtilis.* This compound is used to protect the bacterium from oxidative stress and establish antibiotic resistance. There are three enzymes in the synthesis pathway of bacillithiol: BshA, BshB, and BshC[1].BshC, unlike the other two, has a potentially novel mechanism that is not yet well understood. This enzyme is a putative cysteine ligase, however the ligand that donates the cysteine is unknown. The structure of BshC has been previously solved and has a unique dumbbell shape suggesting a hinging motion. Previous Small Angle X-ray Scattering (SAXS) results indicate that the structure in solution may adopt a slightly different conformation compared to the crystal structure. Using computational biochemistry methods, including molecular dynamics and normal mode analysis, we found a conformer that better fits the SAXS data and identified structural hinges. By performing structural homology searches, we were able to find potential ligands for the HUP domain Rossmann fold of the BshC active site. The results of this project will be used to formulate hypotheses about ligand interactions and possible enzyme mechanisms.

Introduction:

Antibiotic resistance is a major concern across the globe. Many bacteria have evolved to defend themselves against antibiotic drugs. Certain gram-positive bacteria, or firmicutes, can effectively combat oxidative stress brought on by antibiotics such as fosfomycin using molecules containing a thiol group (-SH)[1]. Bacillithiol is one of these thiol containing molecules. Much is still unknown about bacillithiol and its biochemical pathway. Previous studies have concluded that there are three enzymes used to create bacillithiol: BshA, BshB, and BshC. The functional mechanisms of BshA and BshB are well understood; previously our lab used computational methods to support these studies[2]. However, BshC's function has yet to be fully understood; better understanding of its mechanism is important in combating the antibiotic resistance it establishes.

BshC has been reported to be a putative cysteine ligase, using the product of BshB (malyl-glucosamine) as its substrate and attaching a cysteine from an unknown donor to the amine group of the substrate to create Bacillithiol[1], Figure 1. Structurally, BshC appears in solution as a "dumbbell-shaped" homodimer, Figure 2. This unique shape is brought on by the dimerization of the protein along the axis of its coiled-coil domain. Additionally, BshC has a well conserved Rossmannoid fold that houses the

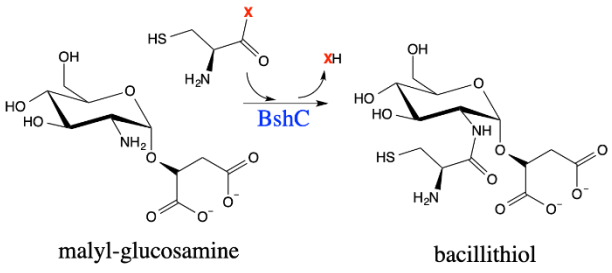

Figure 1: The reaction scheme of BshC is shown with cysteine donated to the substrate, malyl-glucosamine, to form the product, bacilliothiol. The red "X" represents the unknown cysteine donor.

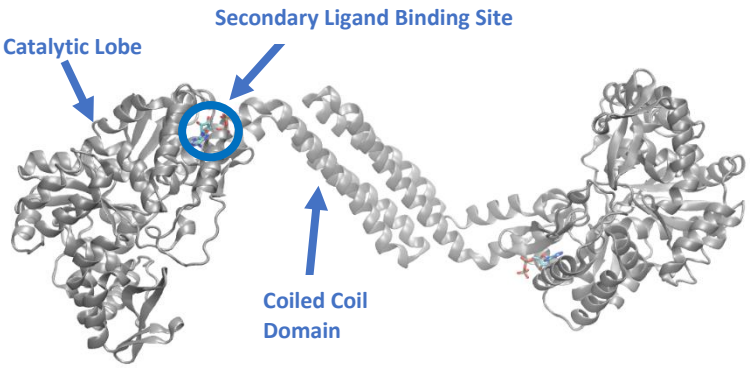

Figure 2: PDB entry for BshC (4WBD), ADP bound in secondary ligand binding site. Catalytic lobe, coiled coil domain, and secondary ligand binding site labeled for one monomer of dimer.

active site as well as a secondary ligand binding site theorized to bind ADP[1]. These structures may hold key information necessary to understand the mechanism of action for BshC. So far, only one crystal structure has been published in the PDB for BshC (4WBD). In this crystal structure, there is a citrate and a glycerol molecule bound in the active site. This suggests good placements for the malyl moiety of the substrate (citrate) as well as the cysteine from the unknown donor (glycerol).

Our approach to furthering the knowledge of how BshC functions is to use computational biochemistry to explore structure, dynamics, and potential ligand binding partners for the protein. Computational methods that will be used include homology searches, molecular dynamics simulations, and multiple analysis techniques. The results of this study will be used to help experimentalists formulate further hypothesis that can be used to inspire further wet-lab experiments on BshC.

Materials & Methods:

Structural homology searches

Structural homology searches for BshC were done using various sources. The RCSB Protein Data Bank (PDB) was used to explore any enzymes that had a coiled-coil motif and their functional relevance. The Dali server uses atomic coordinates of the crystal structure, as well as the PDB,

to explore possible structurally similarities not detectable using amino acid sequences. This server also provides the user with a list of possible similar proteins as well as calculating the conservation of each residue[3]. Using this server, we established a list of structurally similar proteins that can be used to hypothesize potential ligands, as well as explore conserved residues that BshC had compared to these other proteins. The ECOD server gives the user the option to search for evolutionarily conserved structures in a protein as well as providing links to superfamilies that these structures belong to[4]. Using this, we were able to further explore related proteins and establish biologically relevant information about the structural domains of BshC.

Molecular Dynamics Simulations

The CHARMM-GUI input generator was used to create the initial files for molecular dynamics simulations of BshC. This generator uses a pdb file of a protein and converts it to CHARMM readable format[5]. The output of this generator was then used to write a crd and psf file that would be used to create the initial CHARMM input file for BshC[6]. Initially, three molecular dynamics simulations were ran using these coordinates in a generalized Borne implicit solvent system. The first simulation had ADP in the secondary ligand binding site, the second simulation had ATP in the secondary ligand binding site, and the third simulation had no ligand in the secondary binding site. No simulations had the substrate docked into the active site at this time. All three simulations were initially minimized using 1,000 steps steepest descent (SD) and 1,500 steps adopted basis Newton-Raphson (ABNR). After minimization all simulations were equilibrated with 20,000 steps of molecular dynamics, each timestep being 2 fs. The implicit solvent simulation of BshC with ADP in the secondary ligand binding site was ran for a total of

3,140,000 timesteps, or 3.14 ns. The implicit solvent simulation of BshC with ATP in the secondary ligand binding site was ran for a total of 940,000 timesteps, or 0.94 ns. The implicit solvent simulation of BshC with no ligand in the secondary ligand binding site was ran for a total of 2,040,000 timesteps, or 2.04 ns. The output of each simulation provided a pdb file, cor file, psf file, and a dcd file that would be used for visualization and analysis.

Analysis Techniques

All visualizations of molecular dynamics were done using the Visualizing Molecular Dynamics (VMD) software[7]. Loading in the psf, cor, and dcd files provided by the simulations allowed the visualization of the dynamics of BshC.

Normal modes were created using the Prody[8] Normal Mode Wizard tutorial which calculated an anisotropic network model (ANM) that could be loaded into VMD using the Normal Mode Wizard plug-in. The first 50 normal modes were visualized in VMD, with a theoretically infinite amount possible. Normal modes were then individually visualized by creating an animation for each. Additional Prody tutorials that were used for analysis of the protein were: Elastic Network Models, Ensemble Analysis, Structure Analysis, Trajectory Analysis, and ESSA. All of these tutorials were done using the Jupyter Lab Notebook with Python script.

Each molecular dynamics simulation was clustered individually based on the root-mean-squared deviation for the main chain and C_β atoms of all conformers to find a better fit conformation of BshC in solution compared to the previously obtained SAXS data. The simulations were clustered multiple times with cluster centers ranging from 0.4 Å to 0.9 Å. A

representative frame of each cluster was then extracted as a pdb file and compared to the experimental SAXS data using the online FoXS server[9].

Potential ligands extracted from the Dali server were superimposed onto BshC using the RMSD Visualizer Tool in VMD. The substrate, malyl-glucosamine, was parameterized using the Automatic PSF Builder and Forcefield Toolkit plug-ins in VMD. The parameterized substrate was then docked to BshCs active site by appending the psf and cor file created during parameterization and then minimizing using CHARMM with 1,000 steps SD and 1,000 steps ABNR.

Results:

Structural Homology

Structural homology searches using the PDB to look for other coiled-coil exhibiting proteins did not produce many fruitful results. This suggests that the overall function of the coiled-coil region in BshC may be novel to this enzyme. However, using the ECOD[4] online server did answer some questions about BshC from an evolutionary view. This server determined that there were two domains in BshC that were evolutionarily conserved, the Rubredoxin domain (residues 274-329) and the HUP domain (residues 86-377, Rossmanoid fold region). Further study of the HUP domain superfamily showed that almost all proteins in this family bind a nucleotidyl-phosphate in the active site near a conserved glycine; no known BshC sequence has[10] this glycine, most have a tyrosine in this position. Tyrosine would clash with the nucleotide ligand's base (Figure 3). This suggests that BshC is novel in mechanism even compared to its evolutionary superfamily. In addition, most HUP-Domain proteins have a

conserved motif of K-M-S-K-S in the second β-sheet of the Rossmanoid fold. BshC has only one conserved residue in this region, which is phenylalanine. The conservation of these residues were checked using the Dali server, which showed conserved residues of BshC as well as the other structurally similar proteins. Most structurally similar proteins shared the conserved (K-M-S-K-S) motif while BshC did not.

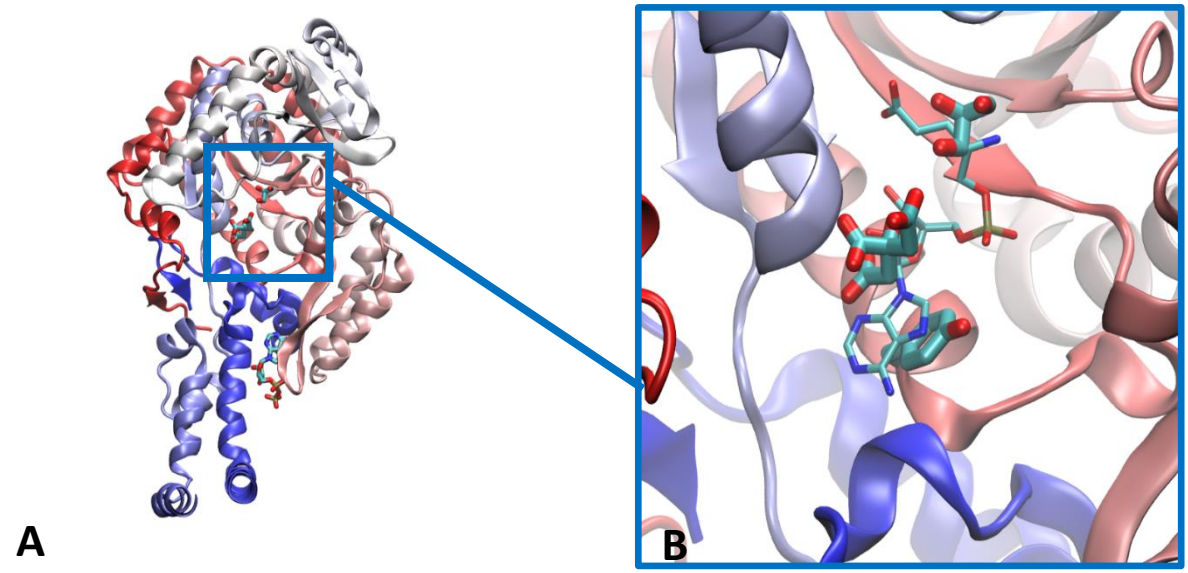

Figure 3: Ribbon diagram of a monomer of BshC from PDB 4WBD colored by conserved residues; red most conserved and blue least conserved.
A: BshC backbone with citrate and glycerol in active site.
B: Glu-AMP (thinner bonds) from structurally similar HUP protein (1n78) with Tyr-101, citrate, and glycerol shown in thicker bonds. Note the clash between the adenine base of the AMP ligand and BshC's Tyr-101.

Molecular Dynamics/ SAXS comparison

Three implicit solvent CHARMM[5] molecular dynamics simulations were ran, each varying by molecule bound in secondary ligand binding site. The first simulation, with ADP bound, was run for 3.14 ns. The second simulation, with ATP bound, was run for 0.94 ns. The third simulation, with no ligand bound, was run for 2.04 ns. Visualization of the three implicit solvent molecular dynamics simulations in VMD did not result in an overall motion of the protein easily analyzed by the eye. Longer simulations may result in more significant motion that can be visually

analyzed. Some noticeable events did happen, however. In the implicit solvent simulation with ATP in the secondary ligand binding site, the ATP molecule did not stay in this binding site and was expelled from the protein. In the simulation with ADP in the secondary ligand binding site ADP remained bound over the course of 3.14 ns. In the simulation with no ligand in the secondary binding site, there was less overall motion of the protein compared to the simulations with ADP and ATP. Calculating the radius of gyration using Prody[8] supported this observation with the range of the simulation with no ligand bound being 53.15- 56.41 Å, compared to the ones with ADP (54.9- 61.4 Å) and ATP (48.7- 58.3 Å) bound (Figure 4).

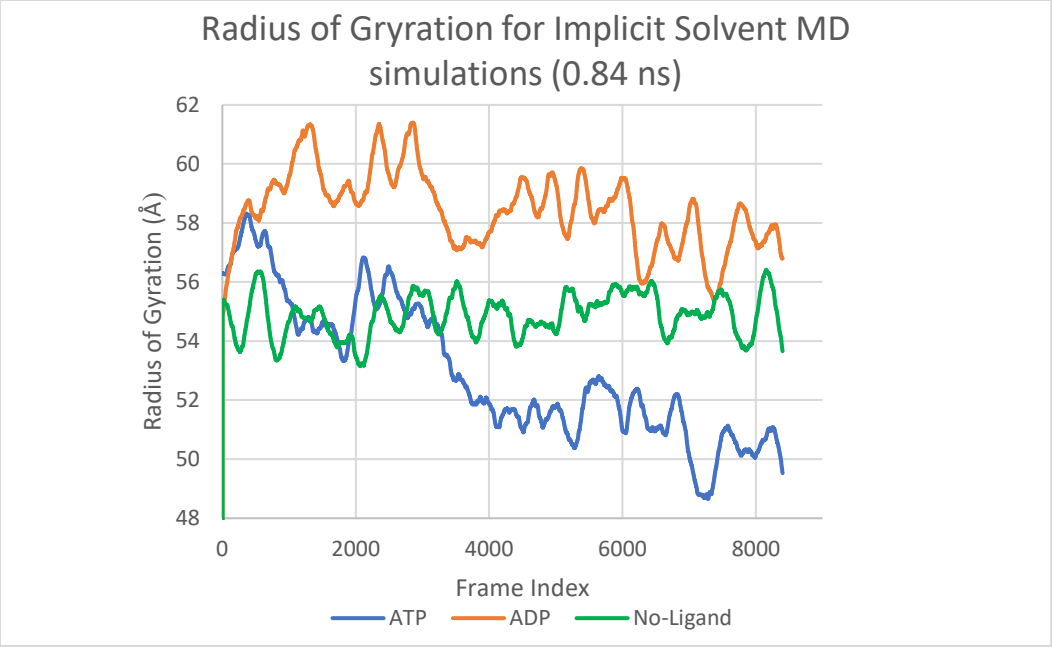

Figure 4: Plot of radius of gyration for the first 8,400 timesteps (0.84 ns) for each of the three molecular dynamics simulations (ADP, ATP, or No ligand in secondary ligand binding site). ATP (blue) ranges from 48.7- to 58.3 Å, ADP (orange) ranges from 54.9- to 61.4 Å, and no ligand (green) ranges from 53.15 to 56.41 Å. Large range of the ATP simulation could be result of ATP being expelled from the protein early on in simulation.

During the simulations, BshC flexing of the catalytic lobes toward and away from each other was observed, as well as a twisting motion of an apparent hinge between the coiled-coil domain and the catalytic lobe.

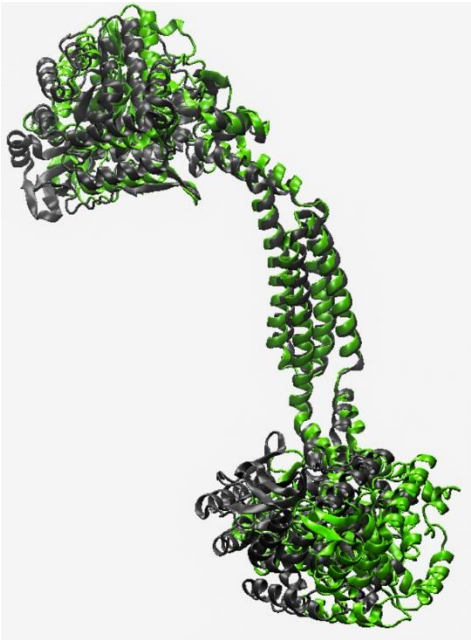

Figure 5: The above photo shows the crystal structure (gray) and the best fit molecular dynamics conformation (lime) superimposed at the coiled coil domain (residues 400 to 480).

When comparing the previously obtained experimental SAXS data with the published crystal structure of BshC, the FoXS[9] server calculated a χ^2 value of 2.17. This suggests that the crystal structure has a different conformation than the structure of BshC in solution. Using the data that was clustered based on RMSD of C_α positions of BshC from the molecular dynamics simulations, we were able to find multiple better-fitting conformations for BshC compared to the crystal structure. The best fit structure had a X^2 value of 1.31 compared to the SAXS data. The difference between the crystal structure and this best-fit structure appears to be that the best-fit structure exists in a more "open" state than the crystal structure (Figure 5). Calculating the radius of gyration for the best-fit structure and published crystal structure, 60.8 Å (best fit) and 54.5 Å (crystal) suggests the enzyme exists in solution in a more "open" state.

Normal Mode Analysis

We also analyzed the dynamics of BshC by predicting protein motion using normal mode analysis. The first 50 anisotropic normal modes, calculated with Prody[8] and visualized in VMD[7], provided a useful representation of the different motions BshC could possibly exhibit. These motions included flexing of the catalytic domains toward and away from each other with an

apparent hinging region located within the coiled-coil domain as well as a twisting motion of the catalytic lobe, both vertically and horizontally, with an apparent hinge between the coiled-

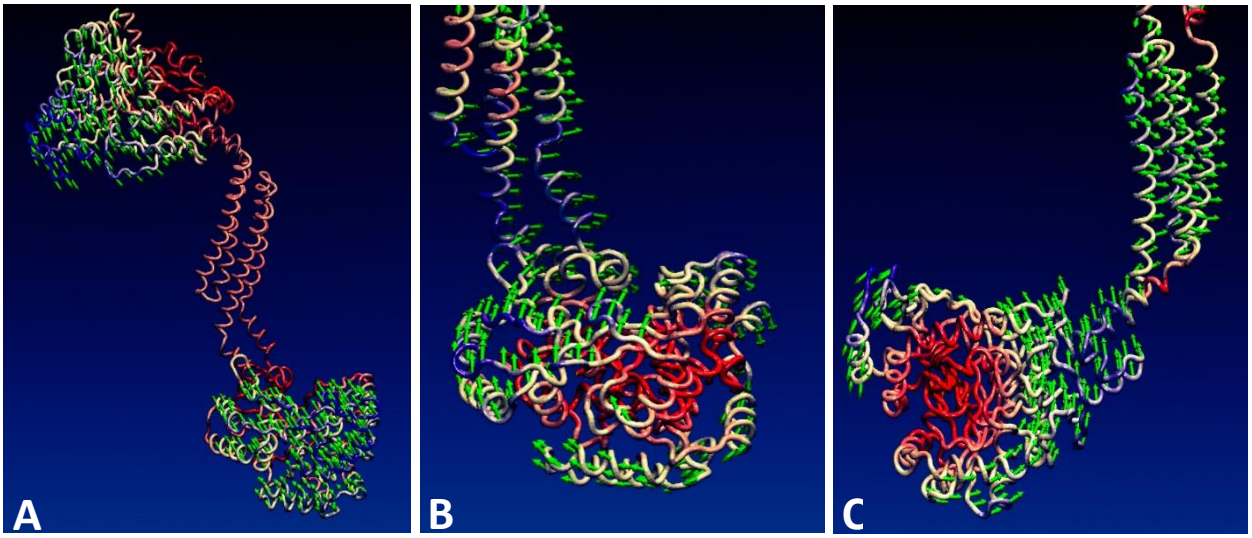

Figure 6: The backbone is colored by mobility; blue most mobile, white intermediate mobility, red least mobile. Green arrows indicate direction of motion.
A: Lowest frequency motion (normal mode 1). Coiled-coil region flexes moving the two catalytic domains relative to each other.
B: Normal mode 4, showing side-to-side twisting motion of the catalytic domain.
C: Normal mode 7, shows the up-and-down twisting motion of the catalytic domain.

coil and catalytic lobe (near residues 381 and 500). These motions could best be compared to the motion of a hand about a wrist (Figure 6). Higher frequency normal modes (modes 20-50) resulted in more precise and independent motions of specific areas throughout the protein. One interesting observation is that in all calculated normal modes (1-50) there is no apparent movement of the active site or the conserved HUP-Domain Rossmanoid fold.

Ligand Binding:

By using the ESSA tutorial for Prody[8], we were able to predict multiple possible ligand binding pockets throughout BshC. Most were in the active site, except for one near the secondary ligand binding site (potential ADP binding region). This allowed the visualization of possible areas of function within the catalytic lobe of BshC. Importantly, there were predicted pockets of

appropriate size for malyl-glucosamine and cysteine, with plenty of space for other potential ligands or binding partners. Studies for potential ligand binding partners of BshC will continue in the future.

Conclusion:

BshC is an enzyme responsible for production of bacillithiol in firmicutes. It is evolutionarily classified as a HUP-Domain superfamily protein but appears to function in a way that is novel compared to other proteins in this same family, since it appears to not bind an adenosine phosphate or other nucleotide in its active site. The substrate, malyl-glucosamine, appears to fit into the active site well with plenty of room for a cysteine to enter and react. However, the cysteine donor is still unknown. Longer dynamics simulations and more conformations to compare to the experimental SAXS data may result in a better understanding of how the protein exists in solution, opening the door to more in-depth ligand docking experiments. Further studies of the HUP-domain, the coiled-coil domain, and the Rubredoxin domain may result in hypotheses of potential ligand binding partners and cysteine donors. Finding the cysteine donor will be crucial in understanding the mechanism of BshC and allowing experimentalists and drug designers to find a way of inhibiting the production of bacillithiol.

References:

1. VanDuinen, Andrew J., Kelsey R. Winchell, Mary E. Keithly, and Paul D. Cook. X-Ray Crystallographic Structure of BshC, a Unique Enzyme Involved in Bacillithiol Biosynthesis. *Biochemistry* **2015**, *54* (2), 100–103. https://doi.org/10.1021/bi501394q.

2. VanDuinen, Andrew J., Kelsey R. Winchell, Mary E. Keithly, and Paul D. Cook. A Structural, Functional, and Computational Analysis of BshA, the First Enzyme in the Bacillithiol Biosynthesis Pathway. *Biochemistry* **2016**, *55* (33), 4654–4665. https://doi.org/10.1021/acs.biochem.6b00472.

3. Holm, Kääriäinen, Wilton et al. Using Dali for Structural Comparison of Proteins. *Curr Protoc Bioinformatics* **2006**, *Chapter 5*, Unit 5.5. https://doi.org/10.1002/0471250953.bi0505s14.

4. Cheng, Schaeffer, Liao et al. ECOD: An Evolutionary Classification of Protein Domains. *PLOS Computational Biology* **2014**, *10* (12), e1003926. https://doi.org/10.1371/journal.pcbi.1003926.

5. S. Jo, T. Kim, V.G. Iyer, and W. Im (2008) CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem. 29:1859-1865*.

6. Brooks, Brooks, MacKerell et al. CHARMM: The Biomolecular Simulation Program. *J Comput Chem* **2009**, *30* (10), 1545–1614. https://doi.org/10.1002/jcc.21

7. Humphrey, W., A. Dalke, and K. Schulten. VMD: Visual Molecular Dynamics. *J Mol Graph* **1996**, *14* (1), 33–38, 27–28. https://doi.org/10.1016/0263-7855(96)00018-5.

8. Bakan, Ahmet, Lidio M. Meireles, and Ivet Bahar. ProDy: Protein Dynamics Inferred from Theory and Experiments | Bioinformatics | Oxford Academic https://academic.oup.com/bioinformatics/article/27/11/1575/217006 (accessed Jul 24, 2021).

9. Schneidman-Duhovny, Hammel, Tainer et al. FoXS, FoXSDock and MultiFoXS: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAXS Profiles. *Nucleic Acids Res* **2016**, *44* (W1), W424–W429. https://doi.org/10.1093/nar/gkw389.

10. Orengo, C. A., & Bateman, A. (Eds.). (2014). Protein families: Relating protein sequence, structure, and function. ProQuest