Assessing Protein-Drug Resistance Due to Mutations via a Rigidity Analysis in silico Approach

Dylan Carpenter, Tess Thackray, Cecilia Kalthoff, and Filip Jagodzinski

Western Washington University, Bellingham, WA
filip.jagodzinski@wwu.edu

Abstract

A mutation to the amino acid sequence of a protein can cause a biomolecule to be resistant to the intended effects of a drug. Assessing the changes of a drug's efficacy in response to mutations via mutagenesis wet-lab experiments is prohibitively time consuming for even a single point mutation, let alone for all possible mutations. Existing approaches for inferring mutation-induced drug resistance are available, but all of them reason about mutations of residues at or very near the protein-drug interface. However, there are examples of mutations far away from the region where the ligand binds, but which nonetheless render a protein resistant to the effects of the drug. We present a proof-of-concept computational pipeline that generates *in silico* the set of all possible single point mutations in a protein-ligand complex. We assess drug resistance using a graph theoretic rigidity analysis approach. Unlike existing methods, we are able to assess the impact of mutations far away from the protein-drug interface. We introduce several visualizations for exploring how amino acid substitutions both near and far away from where the ligand interacts with a protein target have a stabilizing or destabilizing effect on the protein-drug complex. We discuss our analytical approach in the context of experimental data from the literature about clinically known protein-drug interactions.

Keywords: drug resistance, protein-drug complex, rigidity analysis

1 Introduction & Related Works

The sequence of amino acids in a protein determines a biomolecule's structure. Any change to the sequence – often even a single amino acid substitution – can alter a protein's stability and thereby its function. A mutation can be the cause of a debilitating disease, such as Alexander disease, which often results in fatal neurological complications [15]. Conversely, a mutation can also increase a protein's resistances to the effects of a drug. Therefore, identifying and understanding mutation-induced drug resistance can guide research efforts for developing long-lasting effective medicines.

1.1 Rational Drug Design and Drug Discovery

Rational drug design is the process by which a ligand – a drug – is designed and synthesized in order to have an activating or inhibitory effect on a biomolecule – usually a protein – that plays

a role in a disease [18]. The goal of rational drug design is to create a medicine that interacts with a target biomolecule which yields benefits to the patient in a medically significant way. The drug design process has improved vastly in the last few decades, with improvements in both *in vitro* approaches, and in computational methods [2, 6]. A variety of techniques are employed during the drug design process, such as incorporating one or more well-studied molecular substructures with known biological behaviors – pharmacophore [10] – into the drug's scaffold. Unfortunately wet-lab approaches require enormous amounts of time, resources, and money. Modern computational imaging technologies have allowed for the emergence of many protein structure determination methodologies such as X-ray Crystallography and Nuclear Magnetic Resonance spectroscopy. These techniques, paired with atomic-level detail maps of molecules and a deeper understanding of genetic-level factors that impact diseases, have enabled the development of a wide range of computational analysis tools to aid in the drug design process.

1.2   Limitations of Existing Computational Methodologies

Analytical tools for elucidating protein-drug interactions are available. Many of these methods rely on molecular dynamics simulations [22], and includes PIRSpred, which explores the docking and dynamics of a ligand with a target biomolecule [12, 14]. Such approaches explore the interaction of a drug with a protein target, and do not seek to investigate how a protein mutation induces drug resistance. A handful of tools for assessing drug resistance are available. They include screens for emergent drug resistance [13], and scanning approaches that generate *in silico* mutations and evaluate their resistances to known drugs [23]. Unfortunately, all of these computational tools reason about the effects of mutations to residues that are close to the ligand binding site, despite the fact that mutations that are far removed from where the ligand interfaces with the protein can give rise to drug resistance [21].

1.3   Motivation

We aim to explore how amino acid substitutions – both close and far away from the protein-drug interface – alter the stability of the protein-drug complex, and thus may induce the protein to be resistant to the effects of a drug. A secondary goal is to enable a quick screen of all possible mutants generated *in silico*, and to perform an analysis of the variant protein-drug complexes to help reason about drug resistance due to mutations.

2   Our Past Work

We previously developed our ProMuteHT [1] software that receives as input an RCSB protein data bank (PDB) [3] file, and which performs an *in silico* mutation involving one or more amino acid substitutions. ProMuteHT relies on homology modeling in combination with brief runs of molecular dynamics to generate energetically feasible mutants. It has been used in a variety of settings, including predicting the effects that point mutations have on protein stability [8].

In previous work, we have relied on rigidity analysis to gain a deeper understanding of the structural and biological properties of various biomolecules. Rigidity analysis is a fast, graph-based method for analyzing bond networks in biomolecules and identifying flexible and rigid regions [9]. Rigidity analysis reasons about the degrees of freedom of an associated graph that represents the mechanical model of a protein. Pebble game algorithms [11] determine the maximal rigid components and over-constrained regions of the associated graph. The output of rigidity analysis identifies atoms that are part of rigid clusters, which vary in size from just a few

atoms, to the largest rigid cluster (LRC), which may encompass the majority of the atoms in a protein. In prior work we used rigidity analysis in combination with machine learning models to predict with almost 80% accuracy the effects of mutations on protein stability [5]. Most recently we have integrated rigidity analysis into an approach that reasons about mutations to a ligand, to infer which atoms of a drug are most responsible for the effect that the drug has on a protein target [20]. In this work we leverage rigidity analysis to examine drug resistances due to point mutations in the protein.

2.1 Rigidity Distance

In order to quantitatively compare a mutated protein to its wild type, we utilize the rigidity distance, RD, metric, which we developed previously to assess how *in silico* amino acid substitution alter a protein's stability [17, 20]. The RD metric evaluates the effect of mutating a residue to one of 19 other possible amino acids. It measures the difference in counts of the rigid clusters of the wild type and a mutant, and has the following form:

$$RD_{WT \rightarrow Mut} : \sum_{i=1}^{i=LRC} i \times [WT_i - Mut_i]$$

where WT refers to the wild type, Mut refers to the mutant, and LRC refers to the largest rigid cluster (in atoms). Each summation term calculates the difference in counts between the wild type and the mutant for a single cluster size, and multiplies that difference by the cluster size in order to weigh the term. The $RD_{WT \rightarrow Mut}$ metric is calculated for each mutant, yielding $19 \times n$ rigidity distance scores, where n is the count of protein residues in the protein-ligand complex. For this work, the ligand is left unaltered. By leveraging rigidity distance we are able to assess how amino acid substitutions affect the rigidity of a protein-ligand complex.

3 Methods

Our computational pipeline (Fig. 1) integrates protein mutation [1] and rigidity analysis [9] software to generate and analyze all possible single amino acid substitution mutants of a protein-drug complex. The analysis includes a variety of metrics, as well as plots and figures, to help reason about drug resistances due to mutations.
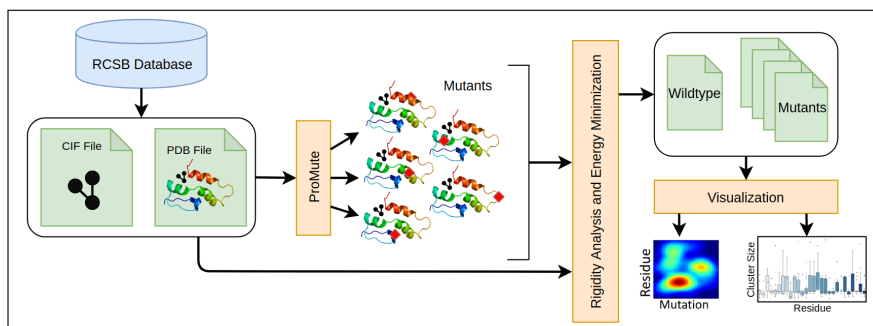

Figure 1: Computational pipeline for generating all possible mutants for a protein-ligand complex, and analyzing them using rigidity analysis to help reason about drug resistance.

PDB ID	Num residues	Ligand ID	Run Time (hrs)	Output Size (GB)	Num Files
3TQ8	173	TOP	15.5	16.6	133,409
6QNB	129	NAG	11.5	9.16	102140
4BVV	80	CPF	5	3.7	61288
3FRE	157	TOP	12.1	13.8	120806
6OIL	168	NAG	11.8	12.1	116465
6NNX	137	RET	9.8	10.5	107846

Table 1: Benchmark run times generating mutants, energy minimization, and rigidity analysis.

3.1 Compute Pipeline, Datasets, and Runtimes

The compute pipeline receives from the user a PDB file ID, as well as the ligand ID(s) for which analysis is desired. The software retrieves the PDB file, as well as CIF files, which contain details about a ligand's atoms. A user has the option to specify if energy minimization on the *in silico* generated mutant should be performed. Multiple energy minimization runs can be performed, in the case that a user wants to confirm that the random velocities of atoms at that start of the simulation do not bias the molecule into a local energy minimum or maximum [19]. The pipeline invokes the ProMuteHT [1] software to generate all possible mutants of the protein-ligand complex with single amino acid substitutions. For example, a 150 residue protein-ligand complex, would yield $150 \times 19 = 2850$ protein-ligand mutants. The Kinari [9] open-source software is used to perform rigidity analysis. The rigidity properties of all mutants are tallied, and visualizations are generated to aid in inferring how mutations in the protein alter the ligand's effect on a protein's stability.

We benchmarked our pipeline on a system with an 8-core *Intel Core i7* processor at 3.40GHZ with 4GHz Turbo Boost and 31GB of RAM. The proteins in Table 1 were selected because they have no missing residues, are relatively small (fewer than 200 residues), and because they include one or more ligands for which wet-lab experiments have been performed and are reported in the literature. These runs did not leverage multiple cores and thus reflect single-core performance.

4 Results and Discussion

To assess the utility of using our computational pipeline, we performed 3 case studies, each highlighting potential analytical uses of our tool for assessing drug resistance due to mutations. We include several visualizations of the results of rigidity analysis, along with a brief discussion of the results. A last subsection discusses our approach in the context of wet-lab or clinical data from the literature about the protein-ligand complexes that we studied.

4.1 Case Study 1: Trimethoprim & Dihydrofolate Reductase

We analyzed the drug trimethoprim in complex with *Mycobacterium avium* Dihydrofolate Reductase (PDB 2W3V) and Dihydrofolate Reductase of *Mycobacterium tuberculosis* (PDB 1DG5). Trimethoprim (ligand ID TOP) is a common antibiotic used to treat a variety of bacterial infections; it was initially designed as an inhibitor of Dihydrofolate Reductase (DHFR) [7].

To identify which mutations to residues in Dihydrofolate Reductase greatly alter how trimethoprim affects the biomolecule's stability, we generated all mutants of PDB files 2W3V and 1DG5. We calculated the *RD* scores for all mutants, and generated figures of box-plots (Fig. 2). One of the distinguishing capabilities of our approach is that it is able to reason
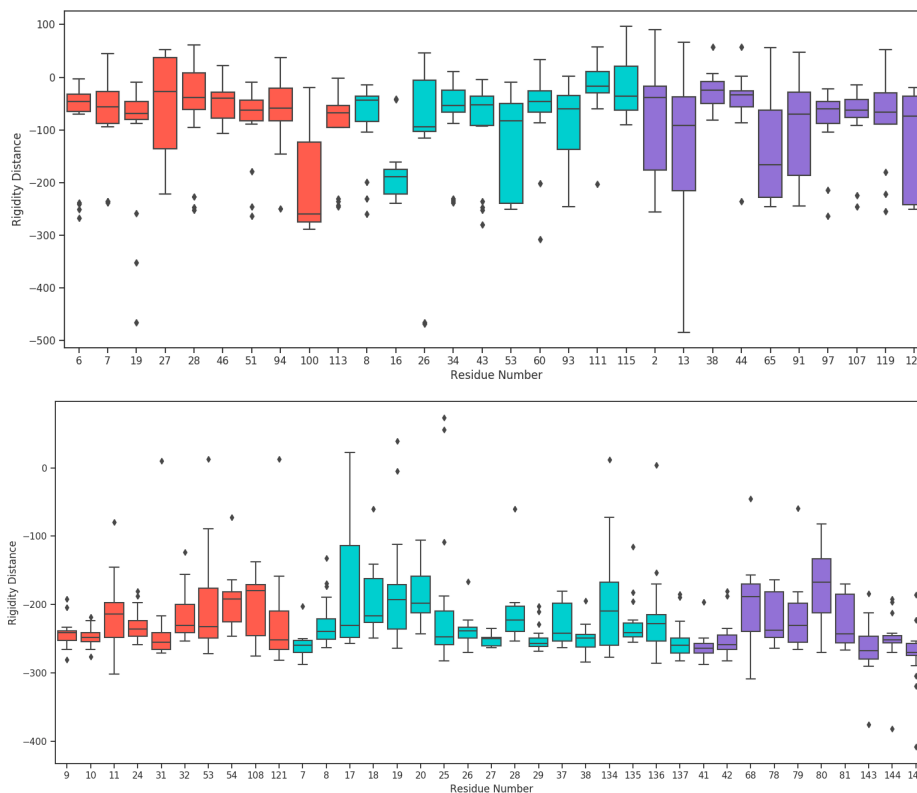
Figure 2: Box plots of distributions of RD scores (y-axis) for residues in PDB file 1DG5 (top) and 2W3V (bottom) whose alpha carbon atoms are 0-4Å (red), 4-8Å (blue), and 8-12Å (purple) from the C3 atom of trimethoprim. Not all residues are shown.

about the effects of mutations that are far away from the protein-ligand interface. As such, we grouped residues into one of three bins: those whose alpha carbon is within 4Å of the C3 atom of TOP, 4-8Å from the C3 atom of TOP, and 8-12Å from the C3 atom of TOP.

Both subplots of Fig. 2 reveal that there are residues far from the ligand, which when mutated have an impact on the stability of the protein as inferred using the RD scores. Because the subplots in Fig. 2 are for two different proteins (in complex with the same drug), we are not able to compare specific residues of one protein with that of another. We can, however, note that the RD values of the mutants of the protein-complex in 2W3V (bottom plot in Fig. 2) range from -300 to -100, while the RD values for the complex in 1DG5 (top plot in Fig. 2) are mostly between -200 and 100 (with some outliers such as residues 100 and 13). Our preliminary interpretation of this data is that the mutations to the *Mycobacterium avium* Dihydrofolate Reductase in PDB file 2W3V impose a greater degree of resistance to trimethoprim than when mutations are performed on *Mycobacterium tuberculosis* Dihydrofolate Reductase in PDB file 1DG5.

4.2   Case Study 2: Assessing Roles of Different Ligands

To explore the extent that a protein is resistant to different drugs, we compared the effect of trimethoprim (Ligand ID TOP, same as Case Study 1) versus Dihydronicotinamide-adenine dinucleotide phosphate (NADPH, Ligand ID NDP) on *Mycobacterium avium* Dihydrofolate Reductase (PDB 2W3V). We relied on the capability of our compute pipeline to permit a user to designate which ligand(s) to test resistance for. In one experiment, NADPH only was retained in complex with Dihydrofolate Reductase, and in another TOP only was retained.
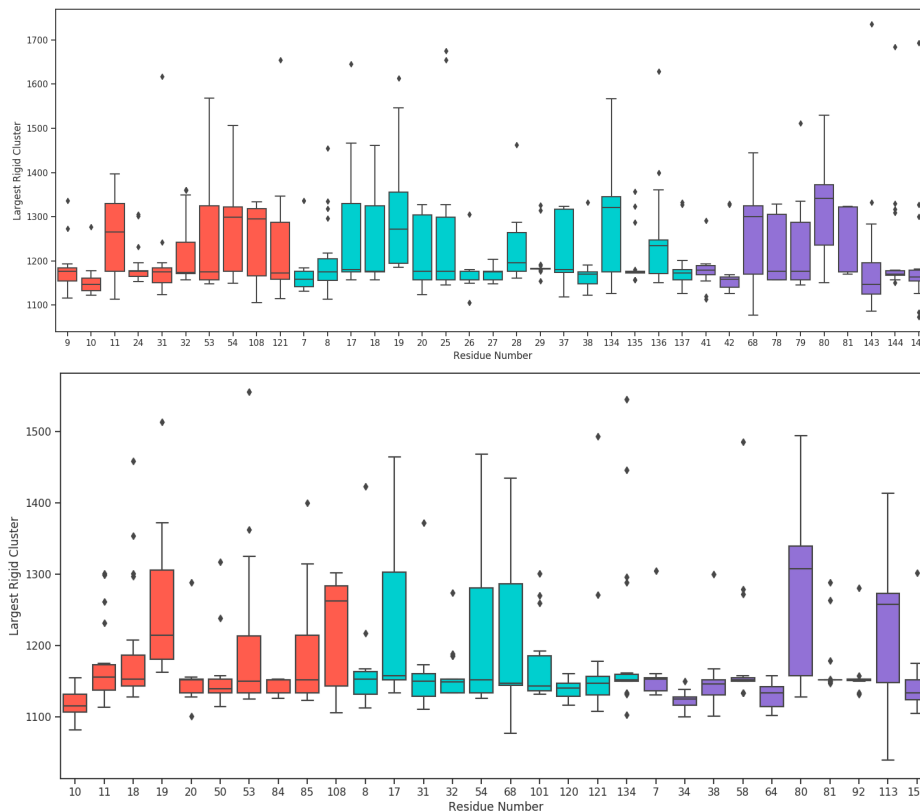

Figure 3: Box plots of distributions of the Largest Rigid Clusters (LRC, y-axis) for trimethoprim (top) and Dihydronicotinamide-adenine dinucleotide phosphate (bottom) in complex with *Mycobacterium avium* Dihydrofolate Reductase. Residues (x-axis) are binned based on carbon alpha distance to the retained drug: 0-4Å(red), 4-8Å(blue), and 8-12Å(purple).

The box plots of the Largest Rigid Clusters among all mutants (Fig. 3) reveal that *Mycobacterium avium* Dihydrofolate Reductase in complex with trimethoprim produces more large rigid clusters (range from 1100 to 1600) than when in complex with NADPH (range from 1100 to 1500). This suggests that 2W3V has a greater resistance to TOP than it does to NADPH because rigidity analysis identifies larger rigid clusters (inflexible regions) in the mutated 2W3V-TOP complexes than in the mutated 2W3V-NADPH complexes. Also mutations to residues that are far removed from the ligand binding site (e.g. residues 80 and 113 in Fig. 3 bottom, and residues 68, 78, 79, 80, 81, and 143 in Fig. 3 top), have an impact on how the ligand affects the structural stability of the protein.

74

4.3   Case Study 3: Stabilizing Interactions and Drug Resistance

We investigated if mutating a residue that hydrogen bonds and/or forms a hydrophobic interaction, versus a residue that engages in neither interaction, has any bearing on the extent that a protein is resistant to the effects of a drug. We again selected PDB file 2W3V and retained NADPH (ligand ID NDP). For this case study we developed a box plot that distinguishes between residues that in the wild type engage either in a hydrogen bond or a hydrophilic interaction, and which do not engage in either interaction. This plot was chosen to explore correlations between the distance of residue from the ligand, stabilizing interaction status, and rigidity distance, RD, distribution. Fig. 4 reveals that, not surprisingly, residues that are close to the ligand active site (left-most residues 10, 11, 18, 19, 20, etc.) do not need to engage in stabilizing interactions in order to have resistance to the effect of the drug when mutated. Residues close to the protein-ligand interface, when mutated, may alter the geometry of the ligand binding site, and thus alter how the drug affects the protein. It is surprising, however, that residues 8 and 32 (which are between 4-8Å from the ligand), and especially residue 7 (8-12Å from the ligand), do not engage in stabilizing interactions, yet when they are mutated, they nonetheless alter the extent of the effect of the drug NADPH on the protein target. This means that our approach might provide insights about drug resistance which cannot be attained via simply identifying which residues engage in stabilizing interactions.
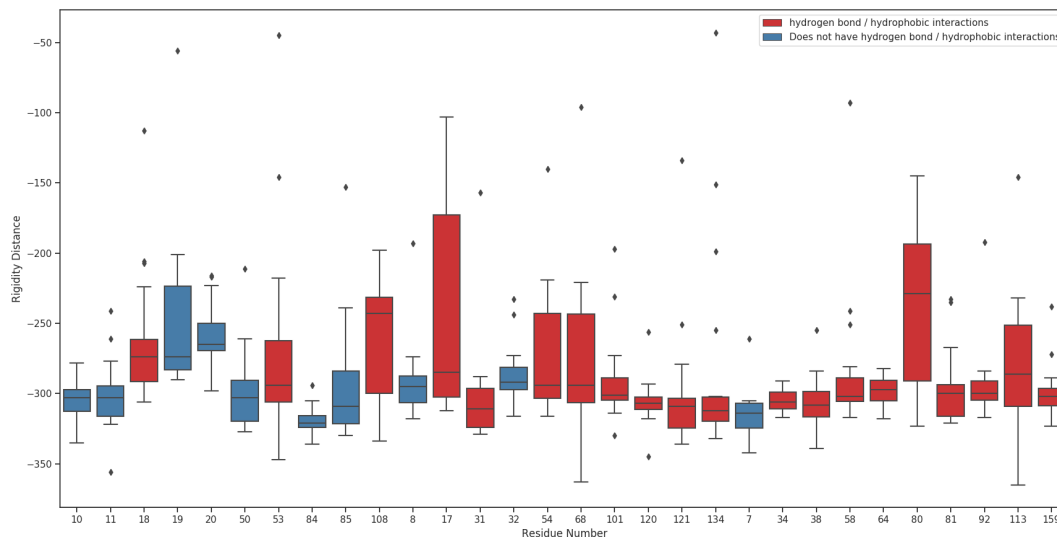

Figure 4: Box plots of RD metrics for a selection of residues (x-axis) in 2W3V in complex with NADPH. X-axis residues (not all shown) are grouped by relative distance from the ligand (see case studies 1 and 2). Coloring signifies if in the wild type a residue was involved (blue) in a hydrogen bond or hydrophobic interaction, or if it was involved in neither (red).

4.4   Correlations Between Our Results and Experimental Data

To explore the biological relevance of the results of our computational experiments, we assessed our results in the context of wet lab studies that are reported in the literature about mutations performed in physical protein-ligand complexes. The genus *mycobacterium* contains pathogens

associated with tuberculosis and leprosy, as well as other serious mammalian diseases [16]. *Mycobacterium avium* Dihydrofolate Reductase has been found to play a role in a variety of disorders, including urinary tract and kidney infections, and progression of the late stages of AIDS because of ease of infection due to a weak immune system. *Mycobacterium avium* has a strong cell envelope which allows its wild type form to be moderately drug resistant [24].

The drug trimethoprim (ligand TOP in our three case studies) is a common antibiotic used to treat a variety of infections caused by *mycobacterium avium* [16]. Trimethoprim was the first designed inhibitor of Dihydrofolate Reductase (DHFR) [7]. However, studies have shown a resistance of DHFR to the drug [7]. Examining how trimethoprim interacts with and alters that stability of *mycobacterium avium* provides valuable information about the protein-ligand complex. By focusing on mutants that show a drastic change in rigid clusters compared to the wild type of the protein, we can begin to better understand which residues of the protein are the cause of resistance to trimethoprim.

The rigidity analysis of *Mycobacterium avium* Dihydrofolate Reductase in PDB file 2W3V reveals that certain residues in the protein, when mutated to some target amino acids, produce RD scores that differ significantly from when other residues are mutated. For this discussion, we focus on those specific residues – residues 31 and 32 – and create scatter plots of rigidity distance, RD, values for each residue (Figs. 5a, 5b). Mutating residue 31 to a Valine results in an RD score whose value (approximately 50) is far different than when residue 31 is mutated to all other amino acids (whose RD values are -200 to -300). This mutation to Valine and its corresponding RD score is revealed as the outlier in the box plot in Fig. 3. Residue 32, when mutated to all other amino acids, results in RD scores in the range of -120 to -250 for the protein-ligand complex. Comparing the RD scores for residues 31 and 32 (Fig. 3), shows that residue 31 produces a greater variance in the change of rigid clusters than residue 32. We infer this to mean that residue 31, when mutated, is more disruptive, and hence may be the cause of drug resistance, than when residue 32 is mutated.
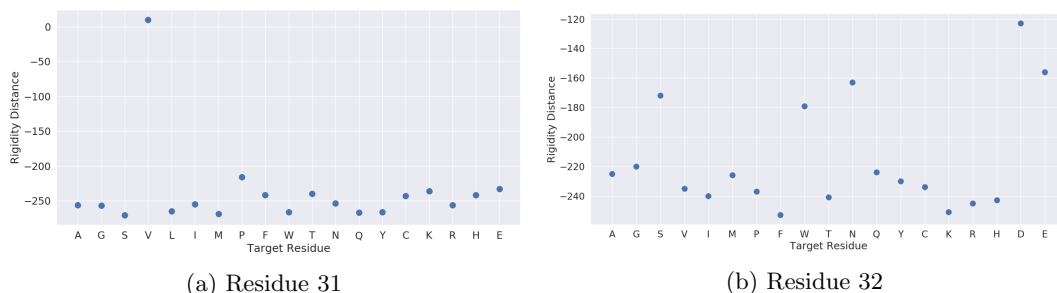

(a) Residue 31   (b) Residue 32

Figure 5: Rigidity distance, RD, values for exhaustive *in silico* mutations for residues 31 (a) and 32 (b) *mycobacterium avium* Dihydrofolate Reductase in complex with trimethoprim.

Residue 31 is highly conserved and is directly involved in the binding of a substrate to DHFR. That residue has been found to likely be involved in the catalysis of DHFR activity [4]. Wet lab experiments involving mutagenesis of residue 31 to target mutations A, E, Q, N, and L (with wildtype D) have shown a decrease in DHFR activity. Conversely, mutations to residue 32 show little to no change in enzyme activity [4]. Rigidity distance can give us indications of specific residues in a protein, that with specific target mutations, alter a drug's effect. In the case of 2W3V, one might analyze the extent that a mutation generates an outlier RD score relative to the mean RD score as a marker of the resistance of the protein to the drug. In 2W3V, the mutation D31V produces an RD score that is more than 3 standard deviations from the

average, while the mutations L32S, L32W, L32N, L32D, and L32E result in RD scores that are less than 2 standard deviations from the average RD score.

5   Conclusion and Future Work

Drug design and synthesis is difficult and both time and cost intensive. Most existing computational approaches that aid the drug design process aim to better understand how a drug binds to a target protein. Most approaches do not reason how mutations in the protein might alter the drug's effect. Of those few computational methods for screening drug resistance caused by mutations, all assess the effects of amino acid substitutions near the protein-drug interface only. This is despite known mutations to residues far away from where the drugs acts, which cause resistance to known drugs. We have developed a computational pipeline that exhaustively *in silico* mutates all residues in a protein-ligand complex. We analyze the effects of those mutations using a graph theoretical approach. This allows us to assess how mutating a residue – regardless how far from the ligand binding site – affects the stability of a protein-drug complex.

In ongoing work we are using clinical trial and wet-lab data of mutations to develop a benchmark dataset of protein-drug complexes. Using this dataset we are developing machine learning models such as Random Forest, and employing a more robust statistical analysis of the rigidity of mutants, to infer which amino acids, when mutated, alter the effect that a drug has on the protein-drug complex. Work towards these goals involves refining the current software to leverage High Performance Computing (HPC) best principles to reduce run-times and size of the data generated. The software is available upon request, and in the near future will be distributed via a Docker environment.

References

[1] Erik Andersson and Filip Jagodzinski. ProMuteHT: A high throughput compute pipeline for generating protein mutants in silico. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 655–660. ACM, 2017.

[2] Robert E Babine and Steven L Bender. Molecular recognition of protein- ligand complexes: Applications to drug design. *Chemical reviews*, 97(5):1359–1472, 1997.

[3] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank, 1999–. *International tables for Crystallography*, 2000.

[4] Ronnie A Böck, Jose L Soulages, and William W Barrow. Substrate and inhibitor specificity of mycobacterium avium dihydrofolate reductase. *The FEBS journal*, 274(13):3286–3298, 2007.

[5] Ramin Dehghanpoor, Evan Ricks, Katie Hursh, Sarah Gunderson, Roshanak Farhoodi, Nurit Haspel, Brian Hutchinson, and Filip Jagodzinski. Predicting the effect of single and multiple mutations on protein structural stability. *Molecules*, 23(2):251, 2018.

[6] Sean Ekins, Jordi Mestres, and Bernard Testa. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British journal of pharmacology*, 152(1):9–20, 2007.

[7] Alexavier Estrada, Dennis L Wright, and Amy C Anderson. Antibacterial antifolates: from development through resistance to the next generation. *Cold Spring Harbor perspectives in medicine*, 6(8):a028324, 2016.

[8] Roshanak Farhoodi, Max Shelbourne, Rebecca Hsieh, Nurit Haspel, Brian Hutchinson, and Filip Jagodzinski. Predicting the effect of point mutations on protein structural stability. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health*

Informatics, ACM-BCB â17, page 247â252, New York, NY, USA, 2017. Association for Computing Machinery.

[9] Naomi Fox, Filip Jagodzinski, and Ileana Streinu. Kinari-lib: A C++ library for mechanical modeling and pebble game rigidity analysis. *Minisymposium on Publicly Available Geometric/-Topological Software*, pages 29–32, 2012.

[10] Osman F Güner. *Pharmacophore perception, development, and use in drug design*, volume 2. Internat'l University Line, 2000.

[11] Donald J Jacobs, Andrew J Rader, Leslie A Kuhn, and Michael F Thorpe. Protein flexibility predictions using graph theory. *Proteins: Structure, Function, and Bioinformatics*, 44(2):150–165, 2001.

[12] Ekachai Jenwitheesuk, Kai Wang, John E Mittler, and Ram Samudrala. Pirspred: a web server for reliable hiv-1 protein-inhibitor resistance/susceptibility prediction. *Trends in microbiology*, 13(4):150–151, 2005.

[13] Hervé Le Nagard, Corine Vincent, France Mentré, and Jacques Le Bras. Online analysis of in vitro resistance to antimalarial drugs through nonlinear regression. *Computer methods and programs in biomedicine*, 104(1):10–18, 2011.

[14] Aeri Lee, Seungpyo Hong, and Dongsup Kim. Krds: a web server for evaluating drug resistance mutations in kinases by molecular docking. *Journal of cheminformatics*, 10(1):20, 2018.

[15] Rong Li, Anne B Johnson, Gajja Salomons, James E Goldman, Sakkubai Naidu, Roy Quinlan, Bruce Cree, Stephanie Z Ruyle, Brenda Banwell, Marc D'Hooghe, et al. Glial fibrillary acidic protein mutations in infantile, juvenile, and adult forms of alexander disease. *Annals of neurology*, 57(3):310–326, 2005.

[16] Rongbao Li, Rachada Sirawaraporn, Penchit Chitnumsub, Worachart Sirawaraporn, Jason Wooden, Francis Athappilly, Stewart Turley, and Wim GJ Hol. Three-dimensional structure of m. tuberculosis dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs. *Journal of molecular biology*, 295(2):307–323, 2000.

[17] Nicholas Majeske and Filip Jagodzinski. Elucidating which pairwise mutations affect protein stability: An exhaustive big data approach. In *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, volume 1, pages 508–515. IEEE, 2018.

[18] Soma Mandal, Sanat K Mandal, et al. Rational drug design. *European journal of pharmacology*, 625(1-3):90–100, 2009.

[19] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with namd. *Journal of computational chemistry*, 26(16):1781–1802, 2005.

[20] Hunter Read, Dylan Carpenter, Sam Herr, and Filip Jagodzinski. Petra: Protein-ligand complex engineering through rigidity analysis. In *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, pages 568–573. ACM, 2019.

[21] Soo-Yon Rhee, Jonathan Taylor, W Jeffrey Fessel, David Kaufman, William Towner, Paolo Troia, Peter Ruane, James Hellinger, Vivian Shirvani, Andrew Zolopa, et al. Hiv-1 protease mutations and protease inhibitor cross-resistance. *Antimicrobial agents and chemotherapy*, 54(10):4253–4261, 2010.

[22] Wei Wang and Peter A Kollman. Computational study of protein specificity: the molecular basis of hiv-1 protease drug resistance. *Proceedings of the National Academy of Sciences*, 98(26):14937–14942, 2001.

[23] Feng-Xu Wu, Fan Wang, Jing-Fang Yang, Wen Jiang, Meng-Yao Wang, Chen-Yang Jia, Ge-Fei Hao, and Guang-Fu Yang. Aimms suite: a web server dedicated for prediction of drug resistance on protein mutation. *Briefings in bioinformatics*, 2018.

[24] Sabrina Zywno-van Ginkel, Thomas P Dooley, William J Suling, and William W Barrow. Identification and cloning of the mycobacterium avium fol a gene, required for dihydrofolate reductase activity. *FEMS microbiology letters*, 156(1):69–78, 1997.