Computational Design of Novel Inhibitors of Dihydrofolate Reductase in Three Bacterial Species

Allison Adams[1], Derek J. Cashman[1]

[1] Tennessee Technological University Department of Chemistry, 55 University Drive, Cookeville, TN 38501-0001

Introduction

Dihydrofolate Reductase (DHFR) is an enzyme present in all living organisms that is essential for cell growth. The enzyme catalyzes the conversion of dihydrofolate to tetrahydrofolate using NADP as an electron donor (Figure 1). The aims of this project are the design of high affinity small molecule inhibitors of bacterial DHFR for the purpose of obtaining broad-spectrum antibiotics against multiple bacterial diseases, including *Bacillus anthracis* (anthrax), *Staphylococcus aureus*, *and Mycobacterium tuberculosis*.

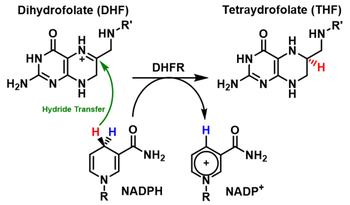


Figure 1. Oxidation-Reduction Reactions catalyzed by DHFR (Source: Wikipedia).

In order to provide a new antibacterial drug novel DHFR inhibitors were systematically designed and analyzed to find a wide range drug which can target multiple species of bacteria. This analysis was conducted using the Protein Frustratometer (http://frustratometer.qb.fcen.uba.ar/, EMBNet & University of Buenos Aires, Buenos Aires, Argentina) and Evolutionary Trace (http://lichtargelab.org/software/ETserver, Baylor College of Medicine, Baylor University, Houston, Texas USA). Evolutionary trace and frustration are both important for defining the active site, as they are useful in determining binding specificity, and areas of the molecule in high energetic states, respectively. 189 small organic molecules were designed based off of a (database review) and designed to interact with these amino acid functional groups based on complementary, non-covalent functional group interactions. They were then docked into the 3D structure of DHFR using the Small Molecular Docking module of MOE2020. The compound were also analyzed according to the ligand interactions in order to determine if it was interacting with the active site and Lipinski's Rules of 5 which helps to determine if an antibiotic would be effective in humans.

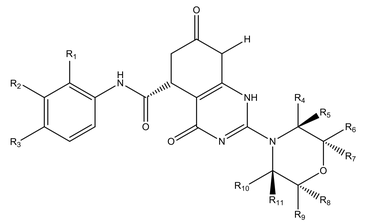
Methods

The three dimensional structures of dihydrofolate reductase were downloaded from the Protein Data Bank (RCSB PDB, Rutgers University, New Brunswick, NJ USA; https://www.rcsb.org/) for three bacterial species: *B. anthraces* (PDB accession code: 3sa2) (3), *M. tuberculosis* (PDB accession code: 1dg8) (4) and *S. aureus* (PDB accession code: 3frd) (5). Using the molecular modeling software MOE 2020 (Chemical Computing, Ltd., Montreal, Quebec CANADA), each molecule was protonated to physiological conditions (temperature = 300 K, pH = 7, salt concentration = 0.1 M), and water molecules present were removed. The AMBER14:EHT force field was used for all calculations. To analyze the active site for potential ligand interactions, the Protein Frustratometer and Evolutionary Trace software was employed. The modified PDB structure of each bacterial species was used as input for each. For this project, a residue was considered frustrated if the frustration density was 20% or greater and the evolutionary trace was noted if the importance from the evolutionary trace was within the top 25%. 189 potential inhibitors were computationally designed from two lead compounds obtained from a previous study (source).

Each potential inhibitor was modeled in MOE 2020 and subjected to energy minimization using the AMBER14:EHT force field until a gradient of 0.01 kcal/mol was reached, indicating convergence. Each molecule was docked into each of the three bacterial species using the Docking module of MOE 2020. Before the molecules were docked in the DHFR active site, the active sites were manually selected. The docking placement method was done using the triangle matcher and London dG used for scoring, and the final docking and refinement was accomplished using the induced fit method with the GBVI/WSA dG scoring method to calculate the binding free energy for each compound (kcal/mol).

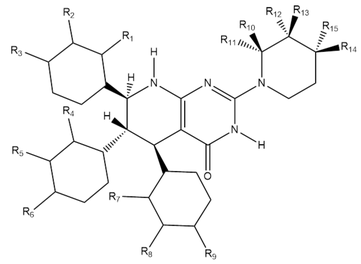
Inhibitors

A.



Name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}
2020-01	H	Cl	H	H	H	H	H	H	H	H	C_3H_5
2020-02	H	H	Br	H	H	H	H	H	H	H	C_3H_5
2020-03	I	H	H	H	H	H	H	H	H	H	C_3H_5
2020-04	H	I	H	H	H	H	H	H	H	H	C_3H_5
2020-05	H	H	CO_2H	H	H	H	H	H	H	H	C_3H_5
2020-06	H	H	H	H	H	H	H	H	C_3H_5	H	H
2020-07	H	H	H	H	C_3H_5	H	H	H	H	H	H
2020-08	H	H	H	H	H	C_3H_5	H	H	H	H	H
2020-09	H	H	H	H	H	H	H	H	F	H	C_3H_5
2020-10	H	H	H	H	H	H	H	H	Cl	H	C_3H_5
2020-11	H	H	H	H	H	I	H	H	H	H	C_3H_5
2020-12	H	H	H	CO_2H	H	H	H	H	H	H	C_3H_5
2020-13	H	H	H	H	H	H	H	H	CO_2H	H	C_3H_5
2020-14	H	H	H	H	H	H	H	CO_2H	H	H	C_3H_5
2020-15	H	H	H	H	H	H	COH	H	H	H	C_3H_5
2020-16	H	H	H	H	H	H	H	H	COH	H	C_3H_5

B.



Name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}
2020-17	H	CO_2H	H	H	H	H	H	H	H	H	H	H	H	H	H
2020-18	H	COH	H	H	H	H	H	H	H	H	H	H	H	H	H
2020-19	H	H	H	H	H	F	H	H	H	H	H	H	H	H	H
2020-20	H	H	H	H	I	H	H	H	H	H	H	H	H	H	H
2020-21	H	H	H	H	CO_2H	H	H	H	H	H	H	H	H	H	H
2020-22	H	H	H	H	H	H	H	OH	H	H	H	H	H	H	H
2020-23	H	H	H	H	H	H	H	H	H	H	H	H	H	H	OH
2020-24	H	H	H	H	H	H	H	H	H	CO_2H	H	H	H	H	H
2020-25	H	H	H	H	H	H	H	H	H	H	H	CO_2H	H	H	H
2020-26	H	H	H	H	H	H	H	H	H	H	H	H	CO_2H	H	H
2020-27	H	H	H	H	H	H	H	H	H	H	H	H	H	CO_2H	H
2020-28	H	H	H	H	H	H	H	H	H	H	H	H	H	H	CO_2H

Figure 2. The two base compounds used and the modifications associated with each compound. The listed compounds were among the top 10 compounds docked in at least one of the bacteria. COH represents an aldehyde. CO_2H represents a carboxylic acid. C_3H_5 represents a cyclopropane.

Binding Free Energy

	Binding Free Energy (kcal/mol)		
Name	B. anthraces	M. tuberculosis	S. aureus
2020-01	-7.9582	-7.7995	-8.0317
2020-02	-7.9672	-7.8252	-8.3235
2020-03	-7.9988	-7.6975	-8.2180
2020-04	-7.9439	-7.7351	-8.5786
2020-05	-8.4230	-8.8191	-7.7935
2020-06	-7.2620	-7.8535	-8.4688
2020-07	-8.5291	-8.1826	-7.6993
2020-08	-8.5652	-7.7882	-7.8985
2020-09	-7.5335	-8.1934	-8.2271
2020-10	-8.0863	-8.0435	-8.0391
2020-11	-8.8365	-7.8890	-7.6541
2020-12	-8.5611	-8.2192	-7.1090
2020-13	-8.6389	-8.3833	-7.3790
2020-14	-8.6273	-7.9950	-7.0040
2020-15	-8.3831	-8.1175	-8.1098
2020-16	-7.6622	-8.7250	-8.0325
2020-17	-8.5410	-9.3239	-7.3994
2020-18	-7.3216	-9.2053	-7.6162
2020-19	-8.4452	-8.7769	-7.1218
2020-20	-7.9159	-9.2379	-6.9815
2020-21	-7.9565	-9.4470	-7.4236
2020-22	-7.9136	-8.6332	-11.1083
2020-23	-7.7814	-9.2738	-6.9919
2020-24	-7.7635	-9.5140	-7.2780
2020-25	-7.3701	-9.4608	-6.5988
2020-26	-8.0172	-9.4454	-7.8276
2020-27	-7.7370	-9.4193	-7.7456
2020-28	-8.6374	-9.3087	-7.9629

Table 1. The binding free energy of each compound in all three bacteria. A darker green indicates a more favorable (negative) values and red represents a less favorable value.

Lipinski's Rules of 5 and Ligand Interactions

B. anthraces

Name	Molecular Weight (g/mol)	LogP	H-Bond Acceptors	H-Bond Donors	ΔG Bind (kcal/mol)	Ligand Interactions
2020-11	534.35	2.13	8	2	-8.8365	L25 L21 L29 F96 I51 A50 N20 N47 A8 W23 E28 R24 M6 P26 N19
2020-13	451.46	0.35	10	2	-8.6389	K33 R53 E28 L25 L21 L29 L55 V32 M6 F96 V7 Y102 A8 W23 P26 I51 R24
2020-28	505.60	4.07	7	2	-8.6374	K33 R58 L21 L29 V32 R53 E28 L25 N47 N20 I51 A50 W23 M6 V7 A8 F96 G57 P56
2020-14	451.46	0.35	10	2	-8.6273	R53 L25 W23 L21 L29 I51 F96 L55 M6 V7 Y102 A8 K33 E28 R24 V32
2020-08	408.46	1.06	8	2	-8.5652	L25 E28 N47 L21 L29 F96 R24 I51 A50 N20 V32 Y102 W23 A8 V7 M6
2020-12	451.46	0.44	10	2	-8.5611	L21 L29 N20 A50 N19 I51 F96 M6 N47 V7 Y102 A8 W23 E28 L25 R24
2020-17	505.60	4.21	7	2	-8.5410	R58 K33 R53 L29 L21 R24 A8 V7 F96 L55 A50 P56 V32 M6 I51
2020-07	408.46	1.09	8	2	-8.5291	L21 E28 L29 F96 I51 L25 N47 W23 R24 P26 V32 A8 V7 M6 A50 N20
2020-19	480.59	5.02	6	2	-8.4452	I51 L21 L29 F96 L55 K33 R53 V32 N47 A50 N20 A8 M6 L25 W23 V7 R56
2020-05	451.46	0.27	10	2	-8.4230	R58 K33 R53 L29 L21 L25 A8 V7 N47 M6 F96 I51 P26 R24 E28 L55 P56 V32

M. tuberculosis

Name	Molecular Weight (g/mol)	LogP	H-Bond Acceptors	H-Bond Donors	ΔG Bind (kcal/mol)	Ligand Interactions
2020-24	505.60	4.23	7	2	-9.5140	R60 Q28 I20 F31 V54 P51 S49 L50 T49 W22 I5 L24 A7 W6 P58 R32 L57 A7
2020-25	505.60	4.10	7	2	-9.4608	R32 R60 I20 Q28 F31 V54 S49 P51 L50 T46 I5 W22 A8 D27 L24 P58 L58 M36 W6
2020-21	505.60	4.21	7	2	-9.4470	Q28 I20 F31 W6 Y100 L50 A7 W22 R23 L24 L50 P25 I94 D27 S49 V54 P52 T46 R32 R60 L57 P58
2020-26	505.60	4.10	7	2	-9.4454	R60 R32 I20 F31 Q28 V55 L24 P51 S49 L50 T46 I5 A9 W22 W6 D27 L59
2020-27	505.60	4.07	7	2	-9.4193	R32 R60 Q28 F31 I20 D27 V54 L24 I5 A7 W6 W22 S49 T46 L50 L57 P51 M36
2020-17	505.60	4.21	7	2	-9.3239	Q28 I20 F31 W22 L24 L50 E111 Y100 H30 W6 D27 R60 R32 A7 T113 L57 V54 S49 P51 T46 L50
2020-28	505.60	4.07	7	2	-9.3087	Q28 I20 R32 R60 F31 P58 I5 W6 W22 A8 P51 S49 T46 L50 L24 V54 D27 L57 M36
2020-23	477.59	4.04	6	2	-9.2738	Q28 I20 R32 R60 F31 M36 V54 T46 L50 S49 P51 W22 W6 L24 A7 I5 D27 P58 L57
2020-20	588.49	6.07	5	2	-9.2379	Q28 I20 F31 W22 I5 W6 Y100 A7 T46 S49 L24 L57 L50 D27 V54 R32 P51 R23 P25 P58 R60
2020-18	490.61	4.94	6	2	-9.2053	Q28 I20 F31 V54 L50 D27 L57 P51 R32 L24 R60 W6 E111 T113 T46 A7 W22 I5 H30 Y100

S. aureus

Name	Molecular Weight (g/mol)	LogP	H-Bond Acceptors	H-Bond Donors	ΔG Bind (kcal/mol)	Ligand Interactions
2020-22	477.59	4.53	6	2	-11.1083	K32 L54 R57 L28 K29 K52 P25 P55 S35 T36 V31
2020-04	534.35	2.18	8	2	-8.5786	L20 L29 F92 I50 S49 T46 V6 V31 A7 L5 D27 W22 L24 H23
2020-06	408.46	1.06	8	2	-8.4688	L28 L20 L54 F92 P25 I50 S49 T46 K52 K32 K29 R57 V31
2020-02	487.35	1.88	8	2	-8.3235	L20 F92 L28 H23 I51 S50 T46 D27 V31 V6 A7 L5 W22
2020-09	426.45	1.10	9	2	-8.2271	L20 L28 S49 F92 I50 Q19 H23 T46 D27 W22 T111 V31 L54 A7 V6
2020-03	534.35	2.19	8	2	-8.2180	L20 L28 I50 H23 F92 S51 T46 L5 A7 V6 W22 V31 L54 D27
2020-15	436.47	0.99	10	2	-8.1098	I50 L20 L29 F92 L54 K53 S49 T46 N18 Q19 A7 V6 V31 L5 D27
2020-10	442.90	1.68	8	2	-8.0391	I50 L20 L28 L54 S49 K52 F92 N18 Q19 T46 L5 A7 V6 V31 D27
2020-16	436.47	1.00	10	2	-8.0325	R57 L28 L54 L20 I50 F92 L54 V6 V31 A7 K32 D27 P55
2020-01	442.90	1.79	8	2	-8.0317	L20 L28 F92 H23 I50 S49 T46 L5 V6 A7 V31 D27

Table 2. The top 10 compounds for each bacteria analyzed according to Lipinski's Rules of 5, as well as according to the ligand interactions. For the ligand interactions, red indicates the amino acid is among the top 25% evolutionary trace in the bacteria, green indicates the frustration density is over 20%, blue indicates an overlap of the two, white indicates it falls into neither category.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (March 2001). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings". Adv. Drug Deliv. Rev. 46 (1–3): 3–26. doi:10.1016/S0169-409X(00)00129-0. PMID 11259830.

Top Compounds

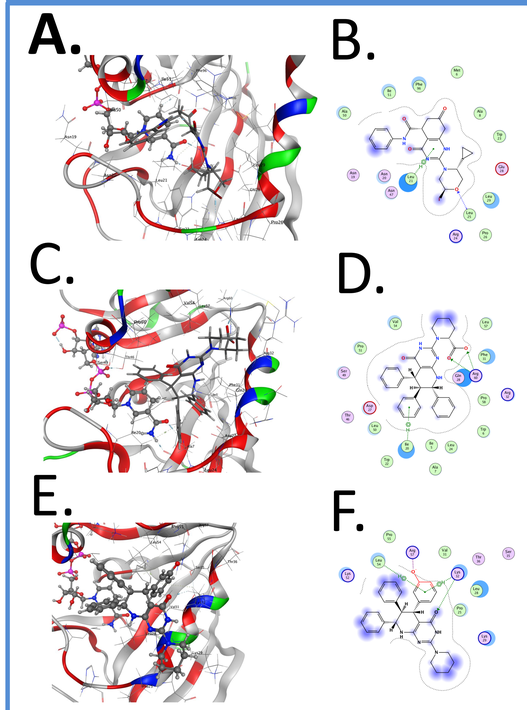


Figure 3. The top compound for each bacteria docked in the bacteria and the ligand interactions for each compound. A and B display 2020-11 in *B. anthraces*. C and D display 2020-24 in *M. tuberculosis*. E and F display 2020-22 in *S. aureus*. Green ribbons indicate amino acid residues with 20% or greater energetic frustration, red ribbons indicate residues in the top 20% of evolutionary importance scores, and blue ribbons indicate residues with both.

Conclusion and Future Research

- The top compounds all tended to bond in areas of high frustration and evolutionary trace. This indicates that these metrics can be used to predict binding sites.
- The compounds generally follow Lipinski's Rules of Five (small molecules should have a molecular weight of less than 500 g/mol, a LogP < 5, less than five hydrogen bond donors, and less than 10 hydrogen bond acceptors). There are some that are outside of some of these guidelines, but they are generally close.
- Of the top 28 compounds, two were among the top 10 for two bacteria, 2020-17 and 2020-28, which were both among the top 10 for *B. anthraces* and *M. tuberculosis*.
- While there was not a compound that was among the top 10 for all three, there were several that had high binding free energy for all three, including 2020-15, 2020-10, 2020-22, and 2020-28.
- Future research would include conducting molecular dynamic simulations on the top compounds in order to assess the protein dynamic and stability of each conformation.

Acknowledgements

This research is supported by the Tennessee Tech University Department of Chemistry and Tennessee Tech University Honors Department.

MOE 2020 software was provided by a grant from Chemical Computing, Ltd. (Montreal, Quebec, Canada). The authors gratefully acknowledge Dr. Scott Northrup (Professor Emeritus) for providing computer hardware in the lab.