8-2019

The Effects of Active Site and Allosteric Residues on Catalysis and Inhibitor Selectivity in the Bacillus stearothennophilus Dihydrofolate Reductase

Tyler Eck
Montclair State University

Abstract

In drug discovery, building a comprehensive picture of the binding events between a drug and its enzyme target can be useful, especially in the development of new drugs or predicting the effect of a drug on a similar target from the same protein family. It is well known that a drug's selectivity is influenced by its interactions with amino acid residues within the active site, but contributions from residues situated away from (distal to) the active site are less well understood. Using a novel predictive approach PAnPredictor that employs sequence alignments, we have previously predicted residue positions in the dihydrofolate reductase (DHFR) family that play a role in drug selectivity. Interestingly, while a few of these predicted residues are located within the active site and are known drug specificity determining positions, other residues fall spatially in clusters distal to the active site. We used *Bacillus stearothermophilus* dihydrofolate reductase (*Bs* DHFR) to study the effect that introducing amino acid replacements at these predicted clusters has on the drug selectivity profile. Mutations were introduced randomly into each individual cluster and screened for functional DHFR enzymes. We determined k_{cat} and K_M values for the mutant enzymes that passed the functional screen to confirm that these mutants are functioning properly. Next, we determined K_D values of the mutant enzymes to two common DHFR competitive drugs trimethoprim (**TMP**), and pyrimethamine (**PYR**) using fluorescence equilibrium titrations to identify residues where perturbations resulted in altered drug selectivity profiles. These data will help us understand the effect that distal residues have on drug selectivity and inform us on the ability of the predictive tool, PAnPredictor, to successfully identify drug selectivity determining residue positions in the DHFR enzyme family. If we find the approach successful, we plan to employ PAnPredictor to other protein families in the future.

MONTCLAIR STATE UNIVERSITY

The Effects of Active Site and Allosteric Residues on Catalysis and Inhibitor Selectivity in The Bacillus Stearothermophilus Dihydrofolate Reductase

by

Tyler Eck

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

August 2019

College of Science and Mathematics

Department of Chemistry and Biochemistry

Thesis Committee:

[redacted]

Nina Goodey, PhD.
Thesis Sponsor

[redacted]

John Siekierka, PhD.
Committee Member

[redacted]

David Rotella, PhD.
Committee Member

[redacted]

Saliya Desilva, PhD.
Chairperson, Department of Chemistry and Biochemistry

THE EFFECTS OF ACTIVE SITE AND ALLOSTERIC RESIDUES ON CATALYSIS AND INHIBITOR SELECTIVITY IN THE BACILLUS STEAROTHERMOPHILUS DIHYDROFOLATE REDUCTASE

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

Tyler Eck

Montclair State University

Montclair, NJ

2019

Table of Contents

1. Introduction

Dihydrofolate reductase (DHFR) catalyzes the NADPH dependent reduction of dihydrofolate (DHF) into tetrahydrofolate (THF), a precursor in the synthesis of DNA bases. Because it plays an integral part in DNA synthesis, DHFR is the target of numerous inhibitors that are used to treat cancer, bacterial infections, and parasitic infections. Common DHFR inhibitors include methotrexate (MTX), trimethoprim (TMP), and pyrimethamine (PYR) which exert their effect by competing with the substrate DHF in the active site. The potency of these compounds vary across the DHFR superfamily; for example, TMP tends to bind bacterial DHFR such as *Escherichia coli*, more tightly than mammalian DHFR like *Homo sapiens*.[1] We refer to the innate inhibitor binding affinities for a single DHFR enzyme as the inhibitor specificity profile. Previously, our lab has identified three amino acid residue positions in the DHFR superfamily that we predict to play a role in determining inhibitor selectivity.[1] These residues were found in the βC – strand distal to the active site, suggesting they may act allosterically on catalysis and inhibitor binding.

Bacillus stearothermophilus (*Bs*) is a moderate thermophilic organism, whose DHFR has catalytic properties that lie between the hyperthermophilic *Thermotoga maritimita* (*Tm*) and the mesophilic *Escherichia coli* (*Ec*) DHFRs.[2] Mutagenesis based studies have been performed on the *Ec* DHFR extensively, revealing key amino acid residues involved in catalysis, conformational motions, and inhibitor binding.[3-10] In the past, a great emphasis has been put on studying active site residues and residues that are involved in dramatic conformational motions such as the Ile14 in the Met20 loop, and Gly121 in the FG loop.[4, 6, 9, 11-12] However, allosteric sites and residues such as those described here have not been considered in terms of catalysis and inhibitor binding. We sought to determine the impact of such allosteric residues in the *Bs* DHFR enzyme.

To study how active site and allosteric residues affect *Bs* DHFR catalysis and inhibitor selectivity, we determined turnover number (k_{cat}), Michaelis - constant (K_M), and the dissociation constants (K_D) of two competitive DHFR inhibitors (TMP and PYR) for nine variants of the predicted allosteric cluster. Additionally, mutations of the active site and Met20 loop were generated at Asp27 and Arg19 respectively, to gauge the magnitude that these amino acid substitutions have on catalysis in comparison to previous *Ec* DHFR studies. We determined K_D values using intrinsic tryptophan fluorescence adapted to a microplate reader format. The use of the thermophilic *Bs* DHFR is well suited for this equilibrium titration assays as the enzyme remains stable during long incubation times at room temperatures.

Typically, the effects from changing allosteric residues are thought to occur indirectly.[13-14] In other words, allosteric residues may not involve direct interactions with the substrate or inhibitors in the active site, but can impact catalysis or inhibitor selectivity greatly.[13, 15] The emerging field of allosteric regulators has suggested that allosteric effects arise by perturbing the conformational equilibrium, where the enzyme fluctuates between conformational substates with different binding characteristics.[13-16] Some states, which promote catalysis, push the enzyme along the reaction coordinate; meanwhile, other substates could have a higher affinity for one inhibitor and a reduced affinity for another.[17] Allosteric effectors are thought to control this conformational sampling, and perturbations in these areas may bias the substate equilibrium which in turn can influence inhibitor binding and selectivity. A comprehensive understanding of what residue positions can impose allosteric effects within an enzyme superfamily will be useful rational protein engineering efforts.

2. Materials and Methods

Site-directed and randomized mutagenesis of the Bs DHFR gene

Site-directed and randomized mutagenesis were performed on the *Bs* DHFR gene to obtain the required mutations for this study[2]. Mutagenic primers were created with a 3'overhang, incorporating the desired mutation at least five bases from the 5' end[18]. Oligonucleotides used for mutagenesis were obtained from Integrated DNA Technologies, Inc. and are as follows: H38Random (mutations H38Q, H38T, and H38V); 5'-GTGACAATGGGCNNKGCCATCGTGATGGGGCGCAAAACG-3' (forward) and 5'-CCCATCACGATGGCMNNGCCCATTGTCACCCGTTTAAAATACG-3' (reverse), H38N; 5'-CAATGGGCAATGCCATCGTGATGGGGCG-3'(forward) and 5'-CGCCCCATCACGATGGCATTGCCCATTG-3' (reverse), A39Random (mutations A39I and A39R); 5'-GTGACAATGGGCCATNNKATCGTGATGGGGCGCAAAACG-3' (forward) and 5'-CCCATCACGATMNNACTGCCCATTGTCACCCGTTTAAAATACG-3' (reverse), A39Y; 5'-GGGCCATTACATC GTGATGGGGCGC-3'(forward) and 5'-GCGCCCCATCACGATGTAATGGCCC-3' (reverse), I40Random (mutations I40R and I40N); 5'-GTGACAATGGGCCATGCCNNKGTGATGGGGCGCAAAACG -3' (forward) and 5'- CCCATCACMNNGGCACTGCCCATTGTCACCCGTTTAAAATACG-3' (reverse), I40A; 5'-GTGACAATGGGCCATGCCGCCGTGATGGGGCGCAAAACG-3' (forward) and 5'-CGT TTTGCGCCCCATCACGGCGGCATGGCCCATTGTCAC-3' (reverse), F102A; 5'-GGCCGAACTGGCTCGGGCGACGATGCCGATTGTC-3' (forward) and 5'-CATCGTCGCCCGAGCCAGTTCGGCCCCGCCGATG-3' (reverse), D27N; 5'-CCTTGGCATTTGCCGGCCAACTTGGCGTATTTTAAACGGGTGACAATGGGCCATGCC-3' (forward) and 5'-GGCATGGCCCATTGTCACCCGTTTAAAATACGCCAAGTTGGCCGGCAAATGCCAAGG-3'(reverse). Ambiguous base definitions are as follows: N (A, T, G, C), K (G, T), and M (C, A)

Each mutagenesis reaction consisted of 0.5 µM of forward and 0.5 µM of reverse mutagenic primers, 50 – 70 ng DNA template, 0.2 mM dNTPs, and 1X HF (Thermo Fisher Scientific, Cat # F530S) buffer in a total volume of 50 µL. Phusion DNA polymerase (Thermo Fisher Scientific, Cat # F530S) (0.5 µL) was added to the reactions, which were carried out in a thermocycler using the program in Table 2 below. Following the PCR protocol, 10 units of DpnI (Fisher Scientific, Cat # FERER1701) were added to each reaction and were incubated at 37 °C for 1 hour. XL Blue cells were transformed with 1 µL of the digested product using the protocol described below. DNA plasmids of successful transformants were isolated and sequenced at the Biology Department of Montclair State University.

Table 2.A Thermocycler program used for site-directed mutagenesis reactions

Step	**Temperature (°C)**	**Time (min.)**	**Cycles**
Initial Denaturation	98	5	1X
1. Denaturation	98	0.5	19X
2. Anneal	55	1	
3. Extension	72	3.5	
Final	72	10	1X
Storage	4		1X

Table 2.B

Step	Temperature (°C)	Time (min.)	Cycles
Initial Denaturation	98	5	1X
1. Denaturation	98	0.5	35X
2. Extension	72	6	
Final	72	10	1X
Storage	4	Hold	1X

Single primer reaction in parallel (SPRINP) for randomized mutagenesis reaction

Due to the nature of mutagenic PCR — in which complementary primers are used — there is a tendency for primer dimers to form. This can hinder the efficiency of the mutagenesis reaction, which is more important in the randomized three residue cluster reactions. To alleviate this, a reaction was performed with each primer in parallel, followed by pooling the parallel reactions into the same tube and allowing the products to reanneal.[19] In this reaction, 0.5 µM of one primer (forward or reverse), 500 ng of DNA, 0.2 mM dNTPs, Phusion DNA polymerase (Thermo Fisher Scientific, Cat # F530S) (0.5 µL), and 1X HF (Thermo Fisher Scientific, Cat # F530S) buffer in a total volume of 25 µL. Parallel reactions were run in the thermocycler using the protocol in Table 2.B. Once completed, the parallel reactions were pooled, denatured, and annealed using the protocol found in Table 3. Additionally, DpnI digestion was performed using 30 U of enzyme overnight at 37 °C before transformations were conducted.

Table 3: Denaturation and Annealing protocol for SPRINP samples

Step	Temperature (°C)	Time (min.)

1	95	5
2	90	1
3	80	1
4	70	0.5
5	60	0.5
6	50	0.5
7	40	0.5
8	37	Hold

Chemical transformation of XL1 Blue and BL21 strains

Using chemically competent cells, 1 - 5 µL of wild type or mutated pET21a+ DHFR plasmid was added to 50 µL of the competent cells. The cells were incubated on ice for 30 minutes then placed into a 42 °C water bath for 45 seconds. The cells were placed on ice for 2 minutes before adding 0.9 mL of sterile LB broth preheated to 37 °C. The cells were allowed to recover for 1 hour at 37 °C with shaking of 225 RPM. The mixture (100 µL) was spread on sterile LB agar plates with 100 µg/mL of ampicillin and incubated overnight at 37 °C.

Transformations involving the randomized mutant products had their mixtures suspended in 25 mL of LB, containing 100 µg/mL of ampicillin, and grown overnight at 37 °C and 225 RPM[20]. The plasmid DNA was purified using a QIAGEN MiniPrep kit to obtain a pooled sample of the mutated plasmids. To confirm a randomized pool, this sample was sequenced at the Biology Department of Montclair State University by Adam Parker.

Electrocompetent folA- cell preparation

The folA$^-$ strain (*E. coli* MH829 *folA::*kan *recA::*cm) was a kind gift from Alex Horswill.[21] A culture was grown in LB with 25 µg/mL kanamycin and 50 µg/mL thymidine overnight at 37 °C with shaking at 225 RPM. The overnight culture (8 mL) was used to inoculate an 800 mL volume of LB with 25 µg/mL kanamycin and 50 µg/mL thymidine and grown to an $OD_{600\ nm}$ of 0.7 – 1.0. The culture was quickly placed on ice for 1 hour then centrifuged at 4,500 RPM, 4 °C for 10 minutes using a Beckman Avanti J-26 XP High Speed Centrifuge. The cell pellet was gently resuspended and washed with 25 mL of a 10 % glycerol solution then centrifuged at 4,500 RPM, 4 °C for 10 minutes and the supernatant was poured off. Wash steps were repeated two times. After the final wash, the entire pellet was resuspended in 6 mL of 10 % glycerol solution, aliquoted as 200 µL, and stored at -80 °C.

Electroporation of folA- cells

Transformations involving the folA- strain were conducted in 1 mm electroporation cuvettes (BioRad, Cat # 1652083) using the 1.8 kV electrical stimulus protocol of a Gene Pulser Xcell Electroporation System (BioRad, Cat # 165-2660). For each reaction, 1 µL of purified plasmid DNA was mixed with 50 - 100 µL of electrocompetent cells. The mixture was added to a 1 mm electroporation cuvette and a 1.8 V electrical pulse was applied to the cuvette. The reaction was recovered with 0.90 - 0.95 mL of LB + 50 µg/mL thymidine and incubated at 37 °C for 1 hour shaking at 225 RPM. The mixture (100 µL) was plated onto selective LB plates containing 25 µg/mL kanamycin, 100 µg/mL ampicillin, and 50 µg/mL thymidine. Overnight cultures were made of the mixtures for cells that would be later used in selection experiments. Unless otherwise stated, all steps were performed at room temperature to increase the transformation efficiency[22].

Developing a functional selection method using the folA- strain

To test our functional selection design, the untransformed folA-, folA- with wild type *Bs* DHFR, and folA- with *Bs* DHFR mutants were grown overnight in LB broth with 100 µg/mL of ampicillin 25 µg/mL Kanamycin and 50 µg/mL thymidine. The untransformed folA- lacks ampicillin resistance so ampicillin was not added to this culture. The culture was pelleted and washed three times in 1 X minimal media A (1 g/L $(NH_4)_2SO_4$, 4.5 g/L KH_2PO_4, 10.5 g/L $K_2H_2PO_4$, and 0.5 g/L sodium citrate with additions of 0.2% glucose, 0.6 mM arginine, and 1 mM $MgSO_4$). The samples were plated onto two environments: MMA with 100 µM of isopropyl β-D-thiogalactose (IPTG), 25 µg/mL kanamycin and 50 µg/mL thymidine and MMA with 25 µg/mL kanamycin and 50 µg/mL thymidine. Parallel plates were incubated at 37 °C or room temperature for five days.

Functional Selection of Randomized Libraries

Overnight cultures of folA+ transformation recoveries were washed with 10 mL 1 X MMA solution three times and suspended in 10 mL of 1 X MMA solution. The cell count was approximated using the conversion of an OD_{600} of 1 to indicate 10^8 cells/mL.[20] As described previously, 5 x 10^2 – 10^3 cells were plated onto selective MMA plates (25 µg/mL kanamycin and 50 µg/mL thymidine) with and without 100 µg/mL ampicillin present.[20] The plates were left to incubate at room temperature or 37 °C for five days. Only colonies that emerged at day two of incubation were assumed to express functional DHFR, whereas colonies that emerged after day four were not.

Expression and purification of Bs DHFR proteins

Cloned wild-type and mutant *Bs* DHFR genes were expressed in BL21 cells under the inducible T7 promotor[2]. Starter cultures for each protein to be expressed included 5 mL LB with 100 μg/mL Ampicillin and were grown to saturation overnight at 37 °C and 225 RPM shaking. The following day, fresh LB (125 – 500 mL) was inoculated with a 1 % saturated overnight culture volume and grown to an OD600 of 0.6 – 0.8 at 37 °C and 225 RPM. Expression was induced with 1.0 mM IPTG, followed by an overnight incubation at 30°C and 225 RPM. Cells were pelleted and resuspended in B-Per Protein Extraction Solution (Thermo Fisher Scientific, Cat # 78243), at a ratio of 4 mL reagent for every 1 g of wet cell pellet, and incubated at room temperature for 15 minutes. The cell lysate was centrifuged at 15,000 x g for 20 minutes. Protein purification was done in one step with a SP Sepharose Fast Flow cation exchange column (Fisher Scientific, Cat # 45-002-934), using 40 mM Hepes, 1 mM DTT at pH 6.8 to wash and 40 mM Hepes, 1 mM DTT, and 0.2 M NaCl to elute the protein from the column.[2] Enzyme concentrations were spectroscopically determined at 280 nm using the molar extinction coefficient 25,565 $M^{-1}cm^{-1}$.[23] Purity was determined by SDS-PAGE analysis.

Michaelis Menten Kinetics

DHFR reaction rates were obtained by tracking the disappearance of DHF and NADPH cofactor through a decrease in absorbance at 340nm.[24] Saturating DHF and NADPH (100 μM) were used in HEPES buffer (40 mM HEPES, 1 mM DTT, pH 6.8) and the reaction was initiated with 0.05 – 2.05 μM of the enzyme. The slope of the graphs (absorbance at 340 nm/ s) were converted to velocities (mM/s) using the molar extinction coefficient, 13.2 (mM^{-1} cm^{-1}) and a 0.46 cm.[24] The pathlength of 0.46 cm was determined using the Synergy H1 pathlength determination software, which measures the absorbance of water molecules at 900 nm and 977

nm in the 200 μL reaction mixture. The wild type enzyme was used as a positive control at 50 nM for every mutant enzyme experiment.

K_M values were determined under similar conditions as above; however, enzyme concentration ranged between 5 - 15 nM and DHF concentrations ranged from 234 nM to 120 μM. In this instance, the reaction was monitored using the tetrazolium dye 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium (MTS) (Fisher Scientific, Cat # NC1508746).[25] Reduction of the MTS reagent by THF produces a formazan product that has a greater molar extinction coefficient at 450 nm than that of combined NADPH and DHF extinction coefficient, making the assay more sensitive in the low DHF concentration range.[25-27] For each experiment, a negative control containing enzyme, NADPH, and MTS was monitored and the signal subtracted from all other reactions containing DHF. The K_M was determined by plotting the velocities against DHF concentration, and fitting to the Michaelis-Menten equation in KaleidahGraph.

Determining K_D by Fluorescent Equilibrium Titration

Bs DHFR enzymes (600 nM) were mixed with 100 μM NADPH and trimethoprim or pyrimethamine at concentrations ranging from 22 nM to 2,000 nM in 40 mM HEPES buffer with 1 mM DTT at pH 6.8. The enzyme cofactor solution was incubated with the inhibitor at room temperature for 30 minutes to allow adequate time for an equilibrium to be reached.[28] Data was collected using a Synergy H1 microplate reader in fluorescence mode using a Xenon flash light source from the top face of the microtiter plate. The temperature was set to 25 °C before taking measurements by exciting at 290 nm and collecting 340 nm emission, then exciting at 290 nm and collecting at 450 nm. The read height was 7 mm, and gain was 140 for

340 nm emission and 110 for the 450 nm emission. Fluorescence emission intensity at 340 nm or 450 nm was graphed against the corresponding inhibitor concentration using Microsoft Excel. Signal averaging of the fluorescence emission intensity was performed with three trials that yielded an averaged graph which constituted one trial. An inner filter correction was applied to account for any effect the titrating inhibitor has on fluorescence signal.[29] K_D values were calculated from the data by using a modified Morrison equation (Equation 1) to fit an increase in fluorescent intensity with titrating inhibitor.[30]The K_D value is a result of at least three replicate measurements.

$$F = (F_{max} - F_{min}) * \left(\frac{([E]+[L]+K_D)-\sqrt{([E]+[L]+K_D)^2-4[E][L]}}{2[E]}\right) + F_{min} \qquad \textbf{Equation 1}$$

3.1 Role of Asp27 and Arg19 in *Bs* DHFR catalysis

Asp27 in *E. coli* (*Ec*) DHFR has been found to play a critical role in protonation of DHF in catalysis; the Asp27Asn mutation resulted in a 300-fold decrease in the k_{cat}.[4, 31] Because Asp27 in *E. coli* DHFR aligns with Asp27 in *Bs* DHFR (Figure 1), we replaced Asp27 of *Bs* DHFR by Asn to investigate whether Asp27 also plays a role in catalysis in *Bs* DHFR. The D27N mutation resulted in a 980-fold decrease in *Bs* DHFR k_{cat}, which was found to be 0.0083 ± 0.0005 SE, suggesting that Asp27 performs a similar role in *Bs and Ec* DHFRs (Table 4). This finding is not surprising because many DHFRs show conservation of a negatively charged residue at this alignment position and the active site aspartic acid is the only ionizable residue within the active site of both the *Bs* and *Ec* DHFRs (Figure 1).[32] X-ray crystallography structures of *Ec* DHFR (PDB code: 7DFR) have shown that Asp27 can form hydrogen bonds

with the pterin ring of folate and water molecules and these observations to the suggestion that Asp27 serves to donate a proton during catalysis.[4, 6, 31-32] Since Asp27 is located more than 5Å away from the N5 atom of DHF that is protonated during catalysis, others have proposed that Asp27 helps to orient the substrate while the solvent acts as a proton donor.[8, 31-32] Wan et. al suggested that the negatively charged Asp27 facilitates solvent-mediated protonation of DHF by increasing the pK_a of the active site environment and the N5 atom of DHF from 2.4 to 6.5, allowing the N5 atom to become protonated more easily.[33] Replacing Asp27 by Asn removed the negative charge, likely resulting in a lower active site pK_a and thus contributing to the observed reduction in catalytic activity.

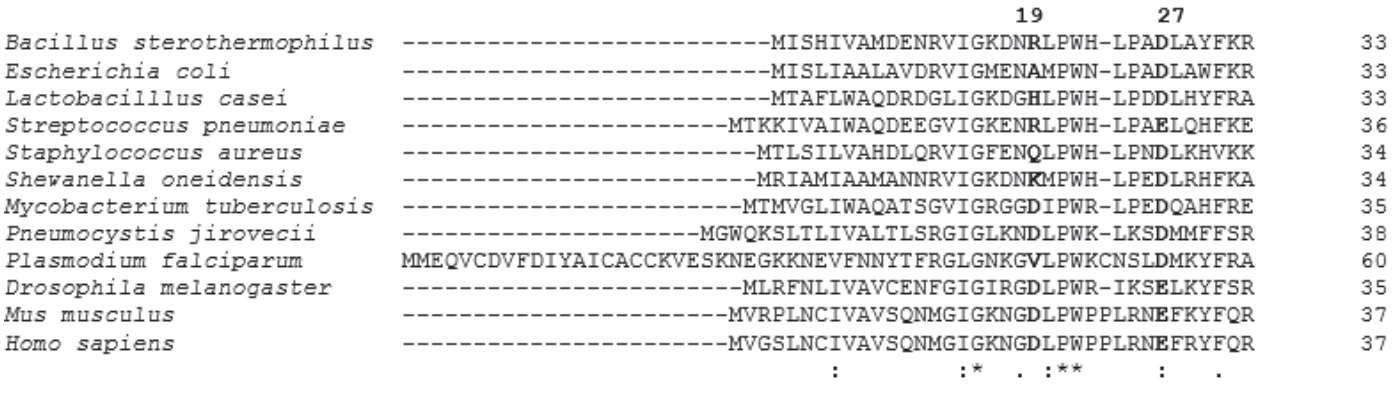


Figure 1. CLUSTAL Omega alignment between *Bacillus stearothermophilus* (*Bs*), *Escherichia coli* (*Ec*), *Lactobacilllus casei* (*Lc*), *Streptococcus pneumoniae* (*Sp*), *Staphylococcus aureus* (*Sa*), *Shewanella oneidensis* (*So*), *Mycobacterium tuberculosis* (*Mt*), *Pneumocystis jirovecii* (*Pj*), *Plasmodium falciparum* (*Pf*), *Drosophila melanogaster* (*Dm*), *Mus musculus* (*Mm*), and *Homo sapiens* (*Hs*) DHFRs. Residues 19 and 27 (*Bs* numbering) are indicated in bold font.

Table 4. Turnover number (k_{cat}) for the wild type, R19D, and D27N Bs DHFR proteins. Slopes of absorbance at 340 nm vs time (s) were converted to velocities using the extinction coefficient, 13.2 $mM^{-1}s^{-1}$ and a 0.46 cm pathlength.[24] Values represent averages of three replicate measurements and are presented with standard error.

	k_{cat} (s^{-1})	$k_{cat\ (wild\ type)}/k_{cat\ (mutant)}$
Wild Type	8.1 ± 0.5	1
R19D	0.072 ± 0.003	110
D27N	0.0083 ± 0.0005	980

The Met20 loop is a highly mobile region of bacterial DHFR enzymes that has been reported to be involved in ligand binding and catalysis.[34] The *Ec* DHFR Met20 loop (residues 9 – 24) has been extensively studied and critical residues that stabilize the major conformers in catalysis (the open, closed, and occluded conformations) have been identified in this loop.[32, 35] In *E. coli* DHFR, the Met20 loop adopts different conformations in response or preparation for NADPH and DHF binding (open), orienting the two molecules for catalysis (closed), and for THF release (occluded).[32] It has been suggested that movement of the Met20 loop is a limiting step in catalysis for bacterial but not vertebrate DHFR, owing to a lack of a poly-proline region found immediately following the Met20 loop residues in vertebrates.[32, 36-37] To investigate the impact of the Met20 loop on catalysis in *Bs* DHFR, we replaced Arg19 in *Bs* DHFR by Asp. This amino acid replacement results in charge reversal and Asp is found at the corresponding aligned position in the *Hs*, *Mm*, *Mt*, and *Dm* DHFRs (Figure 1). Table 4 shows that the k_{cat} for

R19D is 0.072 ± 0.003 s^{-1}, indicating a 110-fold decrease in catalytic activity compared to the k_{cat} of the wild type enzyme (8.1 ± 0.5 s^{-1}).

Hydrogen-deuterium exchange experiments have ascribed motions of the *Bs* DHFR Met20 loop to a preorganization term in calculations of the hydride transfer mechanism.[38] Preorganization in DHFR is thought to relate to a sampling of conformational substates, of which few states are acceptable for the hydride transfer; and enzyme flexibility, which promotes conformational fluctuations.[34, 38] It stands to reason that perturbing the equilibrium of these substates could impact hydride transfer and ultimately, catalysis by limiting the preponderance of catalytic substates. In conformational state sampling, many atoms can move and cause the breaking and forming of new noncovalent interactions that would contribute to an enthalpic barrier between two states.[38] The *Bs* DHFR (PDB#: 1ZDR) Arg19 residue spatially overlaps with the Asp21 residue in *Hs* DHFR (PDB#: 2W3A) and the Ala19 residue of *Ec* DHFR (PDB#: 1RH3) in the closed conformation (Figure 2). In the *Bs*, *Ec*, and *Hs* DHFR structures, the residue at this alignment position points outward toward the solvent. The *Ec* DHFR and Hs DHFR structures with $NADP^+$ bound do not show any interactions between the side chain of this residue and $NADP^+$ (Figure 2). Without obvious noncovalent interactions involving residues aligned to residue 19 in *Bs* DHFR, cofactor binding is unlikely to be the cause for reduced activity in the R19D mutant. Rather, this mutation may have impacted conformational sampling by reducing Met20 loop flexibility, which distorts the population of conformational substates and ultimately reduces activity.

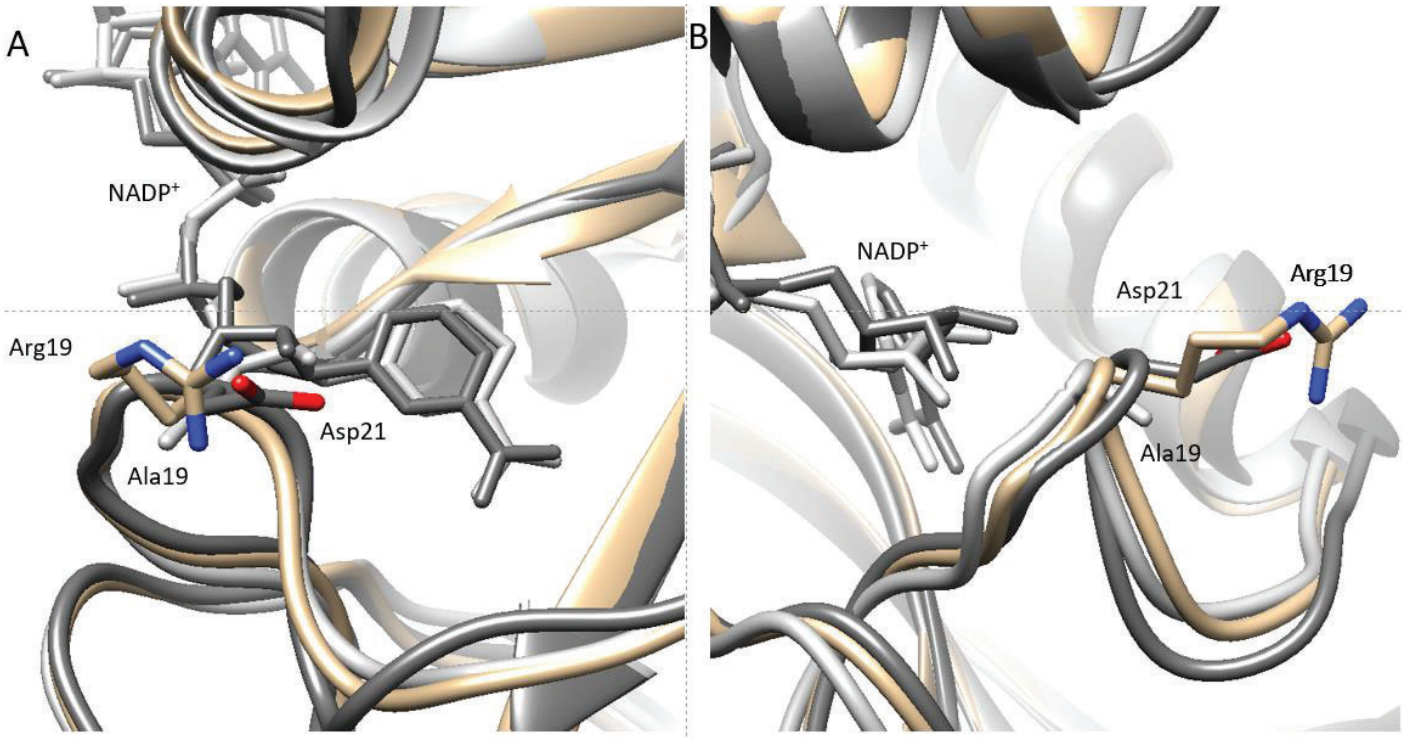


Figure 2. Cartoon representation of the *Bs* DHFR (tan, pdb: 1ZDR), *Hs* DHFR (dark grey, pdb: 2W3A) and the *Ec* DHFR with the Met20 loop the closed (light grey, pdb: 1RH3) conformation. (A) Labels accompany aligned residues with their native numbering. $NADP^+$ is shown and labeled from the both the *Hs* (dark grey) and *Ec* (light grey) structures. (B) Another view that shows all residues pointing towards the solvent in crystal structures. Both figures were made in Chimera (version 1.11.2) using the Match Maker tool.[39]

3.2 DHFR mediated functional selection

When mutations are randomly introduced into an enzyme, some amino acid replacements can result in inactive enzymes, due to the presence of premature stop codons or amino acid substitutions that cause improper protein folding. While inactive mutants can provide useful information on residues that play a role in folding or catalysis, they are not

useful when examining their effects on allosteric regulation or inhibitor selectivity. This is mostly because it is not possible to determine inhibition constants for an inactive enzyme and compare these to wild type values. Saraf et. al. developed a functional selection protocol that utilizes an *E. coli* strain (MH829) lacking the folA (DHFR) gene that cannot survive on minimal media.[40] However, the growth of this strain on minimum media can be rescued by the coexpression of a plasmid encoding a functional DHFR.[21, 40-41] We used the *E. coli* strain (MH829), a kind gift from Alex Horswill, and plasmid-based coexpression of either wild type or mutated *Bs* DHFRs to screen for functional activity.

As reported in the literature, we have found that found that an exogenously introduced gene encoding DHFR can cause an increased growth rate of the MH829 strain in minimal media that is supplemented with thymidine (supplemental figure 1).[42] This finding suggests that MH829 cells when expressing functional DHFR would grow faster than non-functional DHFR expressing variants when plated onto a minimal media source supplemented with thymidine. Figure 3 depicts the growth patterns of the MH829 strain, with and without wild type and mutated *Bs* DHFR expression, over the course of a four-day incubation at 37 °C on minimal media A plates supplemented with 100 μM of isopropyl β-D-thiogalactose (IPTG), 25 μg/mL kanamycin and 50 μg/mL thymidine. We observed vigorous cell growth for all *Bs* DHFR expressing cells within 2 – 3 days following plating. The MH829 cells that completely lacked DHFR activity and thus have a slower growth rate, still showed growth after four days of incubation. Interestingly, the previously discussed mutations R19D and D27N exhibit the same growth pattern as the wild type and other variants, despite the two (110 – fold) and nearly three orders (980 – fold) of magnitude reduction in DHFR activity. This finding is in line with a report by Reynolds et. al. who have observed the same phenomenon when using a D27N variant of *Ec* DHFR in the same plate-based rescue assay.[42] These results suggest that only

minimal DHFR activity is needed to cause an appreciable increase in growth rate when compared to the MH829 strain. The appearance of the cells containing *Bs* DHFR variants (H38N, H38Q, H38T, H38V, A39I, A39R, A39Y, I40N, I40R, I40S, F102A, and K145V) on minimal media A in the same time frame as cells expressing wildtype Bs DHFR indicate that these mutations did not render the enzyme non-functional and can thus be useful in our study of ligand selectivity determining residues (Figure 3).

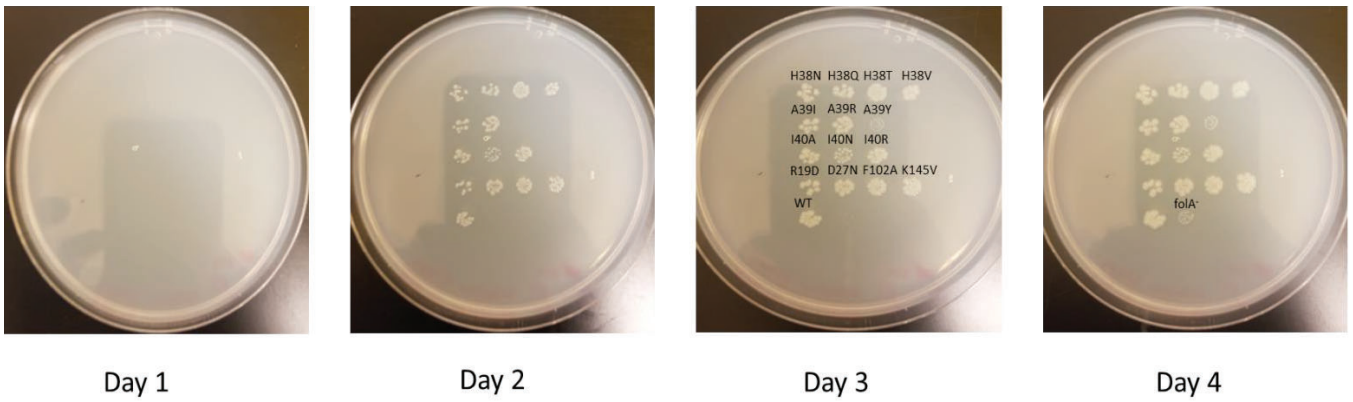
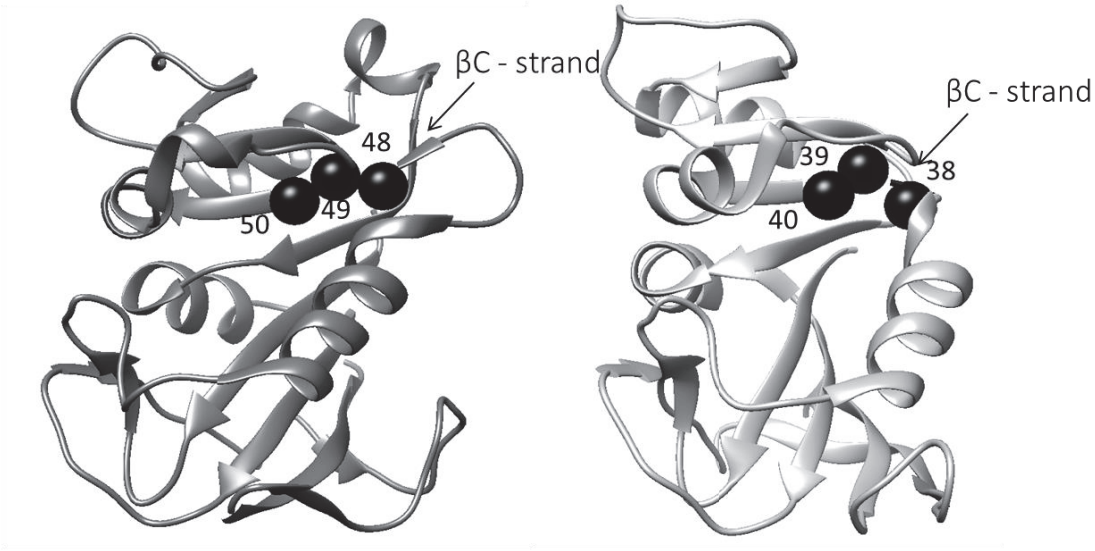


Figure 3. Daily observations of functional selection performed on MMA plates supplemented with 100 µM of isopropyl β-D-thiogalactose (IPTG), 25 µg/mL kanamycin and 50 µg/mL thymidine, and incubated at 37 °C. Overnight cultures of each variant were washed with 1 X MMA before 10^2 – 10^3 cells were plated, assuming an $OD_{600\,nm}$ of 1.0 is equal to 10^8 cells.[20] Colonies with exogenous DHFR would appear within 2 – 3 days of incubation, while the $folA^-$ cells would require four days to appear.

3.3 Role of three allosteric residues on protein-ligand selectivity

Allosteric residues, located away from an enzyme's active site, can play a role in ligand binding and catalysis.[13-14] Here we focus on the three amino acid positions (38, 39, and 40), which we call allosteric cluster 1, in *Bs* DHFR that cluster in the βC – strand found above the active site (Figure 4.A). Oyeyemi and colleagues found that the βC – strand in *Bs* DHFR was more flexible than the *Ec* DHFR counterpart via hydrogen-deuterium exchange, and suggested that this region plays a role in the preorganizational sampling of conformations for *Bs* DHFR.[43] Conformational substate sampling has been implicated in catalysis and inhibitor binding for the DHFR enzyme, previously, which would suggest that these residues in the βC – strand might play a similar role.[16, 44] These three amino acid positions were also predicted a play role in protein-ligand selectivity in the DHFR family.[1] The prediction was family-wide and uses *Hs* DHFR numbering as do many other papers in the DHFR literature. For this reason, Figure 5 shows the *Bs* DHFR residues in allosteric cluster 1 superimposed with the residues in the same alignment positions as the human DHFR sequence. By protein-ligand selectivity, we here mean a protein or variant's ligand affinity profile; for example, *Hs* DHFR binds to methotrexate with greater affinity than to TMP and PYR. Mutations at allosteric cluster 1 were generated using a pool of mutagenic primers that were randomized at the corresponding codon positions. The resulting variants were screened for activity using the functional screen described in section 3.2. We chose mutants to use for further analysis from among the variants we obtained from the randomized PCR and the differences in size and side chain characteristics.

A



B

```
                                                                38 39 40
Bacillus sterothermophilus      VTMG-----------------------------------HAIVMGRKTFEAIGR--P-LP      55
Escherichia coli                NTLN-----------------------------------KPVIMGRHTWESIGR--P-LP      55
Lactobacilllus casei            QTVG-----------------------------------KIMVVGRRTYESFPK--RPLP      56
Streptococcus pneumoniae        TTLN-----------------------------------HAILMGRVTFDGMGR--RLLP      59
Staphylococcus aureus           LSTG-----------------------------------HTLVMGRKTFESIGK--P-LP      56
Shewanella oneidensis           MTLG-----------------------------------KPVVMGRKTFESIGR--P-LP      56
Mycobacterium tuberculosis      ITMG-----------------------------------HTIVMGRRTWDSLPAKVRPLP      60
Pneumocystis jirovecii          VTSGLLV--------------------------TRSTGQMNVVLMGRKTWESLPAHSRPLK      73
Plasmodium falciparum           VTTYVNESKYEKLKYKRCKYLNKETVDNVNDMPNSKKLQNVVVMGRTNWESIPKKFKPLS     120
Drosophila melanogaster         TTKR-----------------------------TSDPTKQNAVVMGRKTYFGVPESKRPLP      67
Mus musculus                    MTTT---------------------------SSVEGKQNLVIMGRKTWFSIPEKNRPLK      69
Homo sapiens                    MTTT---------------------------SSVEGKQNLVIMGKKTWFSIPEKNRPLK      69
                                 :                                 : :::*: .: ..      *
```

Figure 4. Structural alignment of the *Hs* DHFR (PDB #: 2W3A) in dark grey lines and *Bs* DHFR (PDB #: 1ZDR) in light grey lines.[2] The alpha carbons of the allosteric cluster 1 residues are represented in the *Hs* DHFR by black spheres, with equivalent residues in the *Bs* DHFR also shown as black spheres and numbered accordingly. The allosteric cluster 1 is labeled in both structures. This figure was made using the Match Maker tool of Chimera Software (version 1.11.2).[39] (B) Clustal Omega multiple sequence alignment with the allosteric cluster 1 residues bolded and underlined. Numbering corresponds to the *Bs* DHFR sequence.

Figure 5. Structural alignment of *Hs* DHFR (PDB #: 2W3A) and *Bs* DHFR (PDB #: 1ZDR) outlined in grey.[2] Aligned residue sidechains of allosteric cluster 1 are shown for *Hs* DHFR in dark grey and the *Bs* DHFR in black. Panel (A) shows a histidine at position 38 of *Bs* DHFR aligned with the asparagine at position 48 in the *Hs* DHFR. Panel (B) shows an alanine at position 39 of *Bs* DHFR aligned with the asparagine at position 49 in the *Hs* DHFR. Panel (C) shows an isoleucine at position 40 of *Bs* DHFR aligned with the asparagine at position 50 in the *Hs* DHFR. Residues are labeled according to their respective *Hs* or *Bs* numbering. This figure was made using the Match Maker tool of Chimera Software (version 1.11.2).[39]

3.4 Steady-state kinetics

The turnover number (k_{cat}) and Michaelis constant K_M were determined for each mutant and compared to the wild type value (Table 5). All *Bs* DHFR variants had measurable k_{cat} values, the largest change in catalytic activity was observed for the I40R mutation, which had a 5-fold reduction in activity (wild type DHFR k_{cat} = 8.1± 0.8, I40R DHFR k_{cat} = 1.5 ± 0.4). The other allosteric cluster variants (H38Q, H38V, H38T, A39I, A39R, A39Y, I40A, and I40N) were found to have a less than two-fold change in catalytic activity. The K_M values were determined using the redox sensitive reagent, 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium (MTS) to track the production of THF by reduction of the

tetrazolium ring into formazan product, which causes an increase in absorbance at 450 nm.[25, 27] By this method, we determined the K_M of the wild type *Bs* DHFR to be 1.6 ±0.4 SD μM, which is similar to the value determined by Guo et. al of 2.4 ±0.4 μM.[23] Nearly all of the K_M values for the allosteric mutants were similar to the wild type *Bs* DHFR (1.6 ±0.4 SD μM), except for the I40R mutation (32. ±13. SD μM), which caused a significant ($t(2) = 3.93$, $p = 0.062$) increase in the K_M (20-fold). The k_{cat} of the I40R variant (1.5 ± 0.4 s^{-1}) may not be the maximum value because the high K_M (32. ±13. SD μM) could mean that our "saturating" concentration of DHF at 100 μM may not be saturating for this variant. We are limited to this DHF concentration (100 μM) because increasing the DHF concentration would exceed the linear range of the spectrophotometer.

Table 5. Steady-state kinetic values of allosteric cluster mutants. Michaelis constant (K_M) is reported in μM while the turnover number (k_{cat}) is reported in s^{-1}. All measurements were performed in triplicate and standard deviations are shown. *Represents k_{cat} for I40R using the highest DHF concentration (100 μM) we could use in this assay, which may not be fully saturating given the high K_M for I40R.

	k_{cat} (s-1)	K_M (μM)	k_{cat} (wild type) / k_{cat} (mutant)	K_M (wild type) / K_M (mutant)
Wild Type	8.1 ±0.8	1.6 ± 0.4	1	1
H38Q	10 ± 2	1.6 ± 0.5	0.8	1.0
H38V	5.5 ± 0.4	1.3 ± 0.2	1.5	1.2
H38T	5.4 ± 1.1	0.98 ± 0.07	1.5	1.6
A39Y	6.7 ± 0.4	2.7 ± 1.0	1.2	0.6
A39R	4.6 ± 1.0	1.1 ± 0.3	1.8	1.4
A39I	4.0 ± 0.7	1.5 ± 0.3	2.0	1.1
I40A	6.6 ± 1.9	1.8 ± 1.1	1.2	0.9
I40N	5.0 ± 0.5	1.4 ± 0.5	1.6	1.1
I40R	1.5 ± 0.4*	32. ± 13	5.4	0.05

To explain the observed increase in K_M for the I40R *Bs* DHFR variant, we replaced the aligned residue in the *Ec* DHFR structure (PDB #: 7DFR) — a valine at position 40 — using the Rotamers function in Chimera (version 1.11.2) (Figure 6).[39, 45] We used the ternary complex *Ec* DHFR (DHFR:$NADP^+$:folate) in this analysis to include possible steric clashes between V40R and folate, and a structure with *Bs* DHFR bound with folate is not available. The two highest probability rotamers for the V40R mutations, determined by the "Dunbrack backbone-dependent rotamer library" within the Rotamers function of Chimera (version 1.11.2), were chosen in this analysis.[45] Using the Find Clashes/Contacts tool in Chimera (version 1.11.2), the highest probability rotamer for V40R is predicted to sterically clash with multiple nearby residues, including the active site residue Phe31 and the folate molecule.[39, 45] These possible steric clashes with the substrate in the active site can potentially explain the observed increase in KM for the *Bs* DHFR I40R variant. The second highest probability rotamer does not predict the same steric clashes with the folate molecule, however, clashes are still predicted between the Phe31 and the V40R sidechain. Mutational studies of *Hs* DHFR at the Phe34 residue (Phe31 in *Bs* DHFR) have implicated the residue as a ligand-selectivity determinant, as it forms important contacts with methotrexate and DHF.[46] Therefore, a mutation that can potentially alter the position of Phe31 (*Bs* DHFR numbering) within the active site may result in the observed decrease in substrate affinity via a domino effect. Aside from the I40R mutation, the relatively small changes in k_{cat} and K_M values imply that these mutations did not cause major structural changes that deviate from the typical DHFR structure.

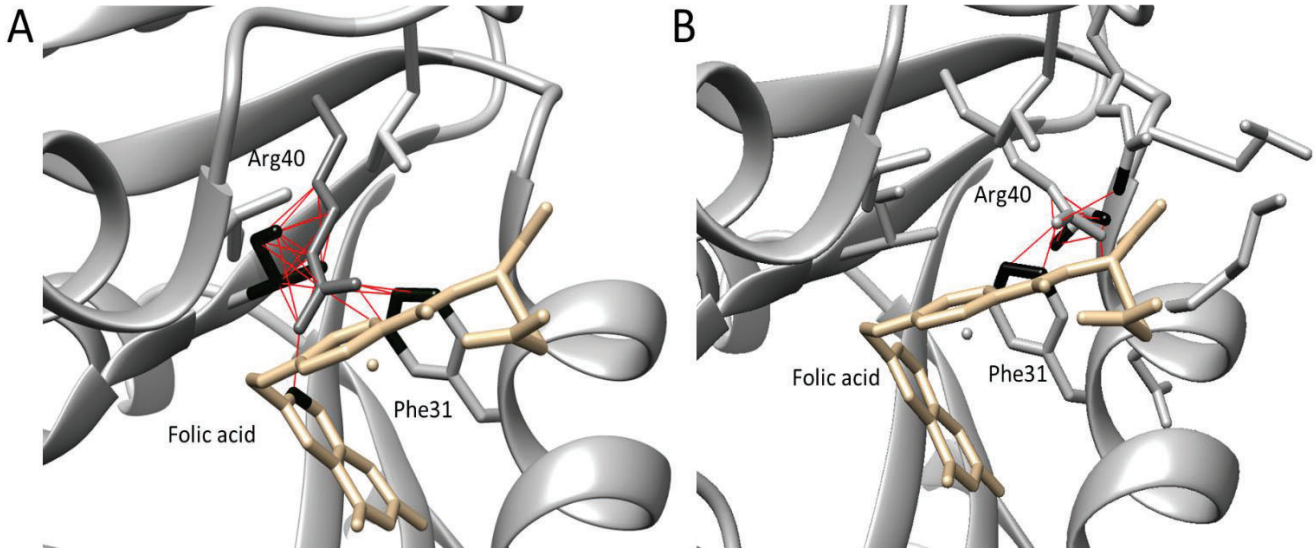


Figure 6. (A) The V40R mutation was modeled into the *Ec* DHFR:$NADP^+$:folate structure (PDB #: 7DFR) using the Rotamer tool in Chimera (version 1.11.2) and choosing the highest probability rotomer.[45] Predicted steric clashes are indicated by red lines determined using the Find Clashes/Contacts tool. Atoms affected by clashes with Arg40 are colored black. (B) The V40R mutation modeled into the *Ec* DHFR:$NADP^+$:folate structure (PDB #: 7DFR) using the Rotamer tool in Chimera (version 1.11.2) and choosing the second highest probability rotomer.[45] Predicted steric clashes are indicated by red lines determined using the Find Clashes/Contacts tool. Atoms affected by clashes with Arg40 are colored black.

3.5 Dissociation constants (K_D) of allosteric cluster 1 variants and two antifolates

To investigate the effect of amino acid replacements in allosteric cluster 1 on inhibitor binding and selectivity, we measured dissociation constants (K_D) of the *Bs* DHFR variants for two common antifolates, trimethoprim (TMP) and pyrimethamine (PYR), by intrinsic tryptophan fluorescence.[24, 46] The physiological cofactor NADPH was included in the assays at 100 μM because it is present in the cellular environment. Including it in the assays ensured that the K_D

values measured were relevant to physiological conditions. Excess NADPH in the assay solution interfered with our ability to measure tryptophan fluorescence at 340 nm because NADPH absorbs at 340 nm. We therefore excited the tryptophan residues at 290 nm, which emitted radiation at 340 nm, exciting the NADPH molecules, which emitted radiation at 450 nm as was done previously.[6] We measured emission intensity of NADPH at 450 nm and corrected the signal for inner filter effects as described previously.[24, 29] We adapted the intrinsic tryptophan fluorescence experiment to a Synergy H1 microplate reader to improve throughput improvement compared to the original assay performed in a cuvette. An additional advantage of this approach is that the emission intensity values for each drug concentration are measured simultaneously rather than subsequently, allowing for a longer incubation time (30 minutes or longer) and for equilibrium between inhibitor and enzyme to establish before measurement. It is important for equilibrium to be reached when capturing the fluorescence signal so that the K_D is not under - or overestimated.[28]

3.5.1. Impact of amino acid replacements in allosteric cluster 1 on TMP and PYR binding

K_D values of the wild type *Bs* DHFR were determined to be 8.0 ± 1.3 nM for trimethoprim and 12 ± 4 nM for pyrimethamine (Table 6). We observed significant changes in the K_D for TMP for two *Bs* DHFRs variants, I40N and I40R: When compared to the wild type enzyme, TMP K_Ds for I40N (120 ± 30 nM) and I40R (150 ± 20 nM) variants increased 15- and 19– fold, respectively ($t(2) = 4.33$, $p < 0.051$ and $t(2) = 6.26$, $p = 0.023$). Such increases in K_Ds were not observed for PYR for these variants. Decreased affinity for TMP and a relatively similar affinity for PYR resulted in significant preferences for PYR over TMP binding for these variants (See *3.5.2 Impacts of amino acid replacements of allosteric cluster 1 on inhibitor*

selectivity). Additional changes, while not statistically significant, in both TMP and PYR K_D values were observed for the other seven *Bs* DHFRs variants (H38Q, H38T, H38V, A39I, A39R, A39Y, and I40A). These effects ranged from a 1.5 – fold reduction (H38Q) to a 3.5 – fold increase (I40A) in TMP K_D values and from a 4.3 – fold reduction (H38V) to 2 – fold increase (H38Q and I40R) in PYR K_D values. Our results show TMP binding to be more affected by perturbations at allosteric cluster 1, especially in the case of position 40, whose variants (I40A, I40N, and I40R) had the largest increases in TMP K_D overall. Interestingly, allosteric cluster 1 appears to have greater influence on TMP binding, with minimal effects to PYR binding. One possible explanation for this would suggest changes in conformational equilibrium, which may stabilize conformations that bind to TMP less tightly.

Table 6. Determined K_D values for trimethoprim and pyrimethamine for wild type *Bs* DHFR and allosteric cluster mutants. Values are the average of three measurements unless otherwise noted in parenthesis. K_D values as concentration in nM and standard error are reported.

	$K_{D\ (TMP)}$ (nM)	$K_{D\ (TMP)}$ (wild type) / $K_{D\ (TMP)}$ (mutant)	$K_{D\ (PYR)}$ (nM)	$K_{D\ (PYR)}$ (wild type) / $K_{D\ (PYR)}$ (mutant)	$K_{D\ (TMP)}$ / $K_{D\ (PYR)}$
Wild Type	8.0 ± 1.3 (4)	1	12 ± 4 (4)	1	0.7
H38Q	7.8 ± 1.1	1.0	23 ± 4	0.5	0.3
H38T	18. ± 2.	0.4	2.8 ± 1.2	4.3	6.5
H38V	5.3 ± 1.9	1.5	7. ± 2.	1.7	0.8
A39I	22 ± 7	0.4	4.8 ± 0.6	2.5	4.5
A39R	17.8 ± 1.5	0.4	6 ± 4	2.0	3.0
A39Y	14.2 ± 0.9	0.6	5.5 ± 1.7	2.2	2.6
I40A	28 ± 14	0.3	8.4 ± 1.8	1.4	3.3
I40N	120 ± 30	0.07	18 ± 4	0.7	6.7
I40R	150 ± 20	0.05	23 ± 11	0.5	6.9

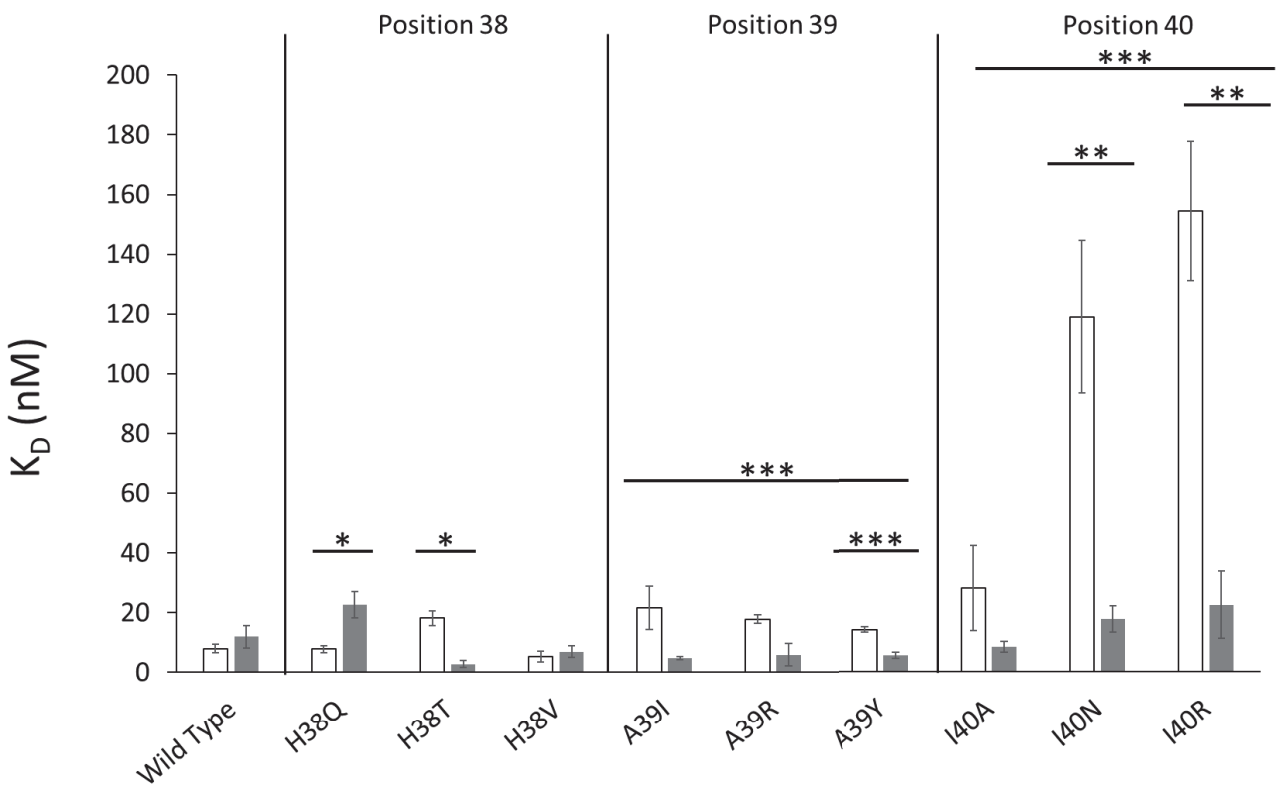


Figure 7. Bar chart representation of K_D values for trimethoprim (white bars) and pyrimethamine (grey bars) of *Bs* DHFR wild type and allosteric cluster 1 mutants. Error bars accompanying K_D values represent the standard error of the measurement. Bars over *Bs* DHFR variant indicate a significant difference between TMP and PYR K_D values for that mutant. Bars that span over multiple variants of the same residue position indicate significant difference between TMP and PYR K_D values for the entire position. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

3.5.2 Impacts of amino acid replacements of allosteric cluster 1 on inhibitor selectivity

We found no statistically significant difference between *Bs* DHFR K_D (TMP) (8.0 ± 1.3 nM) and *Bs* DHFR K_D (PYR) (12 ± 4 nM) values, indicating that the wild type enzyme does not selectively bind one of the two inhibitors (Table 6). The data in Table 6 suggest that the ratio of K_D (TMP) to K_D (PYR), which we here define as a measure of ligand specificity, is influenced by amino acid replacements. Although most variants had not produced significantly different K_D values from the wild type DHFR, the concomitant decrease of one inhibitor affinity and increase in the other resulted in a significant inhibitor selectivity among these variants. Out of nine allosteric *Bs* DHFR variants studied here, we found five variants (H38Q, H38T, A39Y, I40N, and I40R) that show a significant preference for one ligand over the other ($p < 0.1$). The H38T, I40N, and I40R mutations resulted in the largest selectivity for PYR over TMP, showing 6.5-, 6.7-, and 6.9-fold selectivity for PYR over TMP (H38T: $t(2) = 4.19$, $p = 0.052$; I40N: $t(2) = 4.34$, $p = 0.049$; I40R: $t(2) = 5.58$, $p = 0.031$). The A39Y variant also showed highly significant selectivity for PYR over TMP (2.6-fold) (A39Y: $t(2) = 14.151$, $p = 0.005$). In total, four allosteric variants (H38T, A39Y, I40N, and I40R) were found to be selective for PYR, however, one variant (H38Q) bound TMP more tightly than PYR (H38Q: $t(2) = 3.25$, $p = 0.083$). The wild type residue His38 points outward towards the solvent and is within hydrogen bonding distance of the Asp91 sidechain. It is possible that an interaction, or lack thereof, between Asp91 and the position 38 sidechain that influences inhibitor selectivity.

3.5.3 Dependencies of allosteric perturbations on inhibitor selectivity

At some residue positions, we noticed that the change in selectivity upon amino acid replacement did not appear to depend on the identity of the side chain that was introduced. For example, the three mutations we made at position 39 (A39I, A39R, and A39Y) all resulted in

preferences for PYR binding over TMP. A similar trend in inhibitor selectivity was also observed for position 40. *Bs* DHFR variants I40A, I40N, and I40R all exhibit selectivity for PYR over TMP, with their $K_{D\ (TMP)}$ / $K_{D\ (PYR)}$ ratios being much greater than the wild type ratio of 0.7. Preferences for PYR observed in response to perturbations at specific residue positions (39 and 40) suggest that these sites are ligand-specificity determinants as was previously predicted.[1]

At position 40, a larger decrease in TMP affinity was observed for the I40N and I40R variants compared to the I40A mutation. This is reasonable if we consider that this position's side chain points towards the binding site and could clash with residues of the active site that directly interact with the inhibitor, such as Phe31.[46-47] Where the larger I40 sidechain variants sterically clash with the wild type positioning of Phe31, and cause Phe31 to position in a way that would impact TMP and PYR binding. In the *Mycobacterium tuberculosis* (*Mt*) ternary complexes with TMP (PDB #: 1DG5) and PYR (PDB #: 6NNI) bound, Phe31 (the same residue as Phe31 in *Bs* DHFR) aligns well in both structures (Figure 8). This would suggest that the I40N and I40R variants would have similar increases in PYR as observed in TMP binding. However, PYR also interacts with NADPH via van der Waals forces, while TMP does not.[48] We can hypothesize that the asymmetric effect on TMP and PYR binding occurs by a domino effect, as described above, but that this might be outweighed by the interaction PYR has with NADPH.[48] The similar decrease in TMP affinity for both the charged I40R variant and neutral I40N variant also suggests that charge does not influence inhibitor selectivity at this position. While, the alanine mutation introduced in this position is smaller than the native isoleucine residue, the moderate increase in TMP affinity could be due to disrupting potential van der Waals interactions with neighboring Met42, Leu54, and Ile96 in the *Bs* structure.

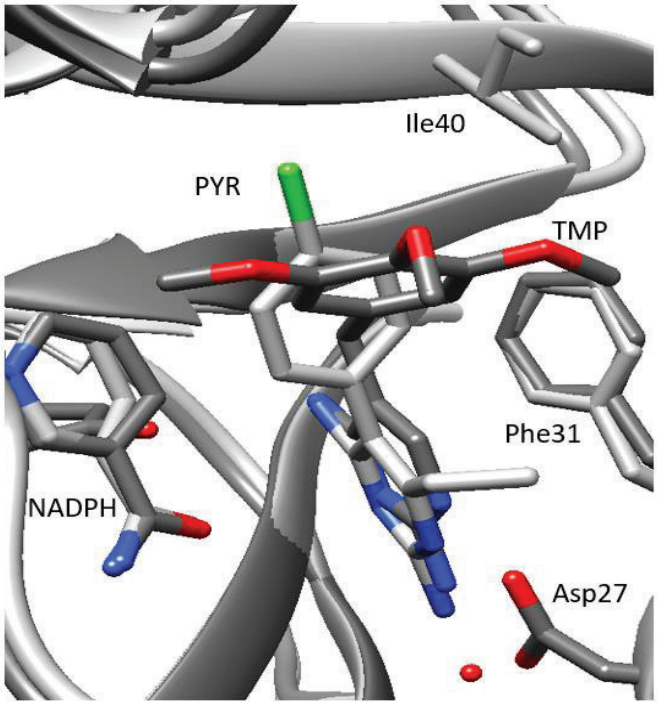


Figure 8. Structures of the *Mycobacterium tuberculosis* (*Mt*) DHFR with NADPH and either TMP (PDB #: 1DG5) or PYR (PDB #: 6NNI) bound. The two structures were aligned using the "MatchMaker" tool in Chimera (version 1.11.2)

The effects of perturbations at position 38 appear to depend on the amino acid substitution. This is demonstrated by the H38T variant preferring PYR binding ($K_{D\ (TMP)}$ / $K_{D\ (PYR)}$: 6.5) while the H38Q variant prefers TMP ($K_{D\ (TMP)}$ / $K_{D\ (PYR)}$: 0.3). The amino acid side chain dependence is further exemplified by the H38V variant; whose preference for TMP was insignificant ($K_{D\ (TMP)}$ / $K_{D\ (PYR)}$: 0.8), but reductions in both TMP and PYR K_Ds were also observed. Admittedly, this analysis is limited to the sample size of mutations generated, and it is difficult to extrapolate these observations to those that describe the full nature of these allosteric positions. This is especially the case for position 38, which does not show a particular trend in inhibitor selectivity in the variants we have generated. Further characterization of this

position and allosteric cluster would require a larger sample size that includes a complete array of sidechain chemistries (e.g. sizes, charge, polarity). Currently, our data suggests that perturbations within allosteric cluster 1 have an effect on ligand-selectivity ($t(26) = 3.27$, $p = 0.003$). These effects appear to be nuanced for each position, where a position may be amino acid sidechain dependent (position 38), residue position dependent (position 39), or a combination of the two dependencies (position 40).

Possible mechanisms by which allosteric cluster 1 mutations can impact ligand selectivity of Bs DHFR.

Overall, mutations made in allosteric cluster 1 caused a shift in inhibitor selectivity towards PYR binding ($t(26) = 3.27$, $p = 0.003$). The effects of perturbations in allosteric cluster 1 appear to be drug specific, that is, TMP affinity was impacted to a greater extent than PYR affinity. With the emergence of TMP resistant pathogenic organisms, numerous studies have investigated the TMP resistance mechanism as it relates to DHFR mutations.[7, 10, 25, 49-50] Notable findings include, the Phe98Tyr strain of *Staphylococcus aureus* where a single amino acid substitution which imparts TMP resistance.[49] The directed evolution of *Escherichia coli* DHFR has also revealed positions 10, 20, 26, 30, 45, 95, and 109 to influence TMP resistance.[7] Finally, clinical strains of trimethoprim-sulfamethoxazole treatment resistant *Pneumocystis jirovecii* have been speculated to due to mutations found within the DHFR gene of these strains.[50-51] The aforementioned studies include mostly residues involved in TMP binding, but fail to implicate the region we describe here as allosteric cluster 1 in any resistance mechanism. Likewise, studies of emerging PYR resistance in parasitic organisms such as *Plasmodium falciparum* (*Pf*) and *Toxoplasma gondii* (*Tg*) found that mutations within the active site that

reduce PYR affinity by either widening or blocking the site.[52] Because residues in the allosteric cluster 1 do not have interactions within the active site, we speculated that this site acts as an allosteric determinant in inhibitor selectivity. Therefore, some principles behind allosteric regulation could potentially explain the observed TMP resistance pattern.

Enzymes are known to exist in a dynamic equilibrium of conformations, whose conformers have innate catalytic characteristics and ligand-binding affinities.[13, 16, 53] Studies of DHFR and methotrexate (MTX) binding have shown that MTX can bind to at least two conformations of *Ec* DHFR and *Bs* DHFR.[6, 17] For the *Bs* DHFR in particular, a 300 – fold decrease in k_{on} was observed for a secondary conformer binding to MTX.[17] Recent work has shown that many enzymes exist in a variety of conformers in statistically distinct populations at equilibrium.[13] Mutations introduced at allosteric positions can affect ligand selectivity by disrupting this dynamic equilibrium via enthalpic or entropic contributions to the free energy of certain populations of conformation, asymmetrically.[13-15] For example, while not affecting a direct interaction with the ligand of interest, a mutation may stabilize an enzyme conformation that binds a ligand weakly or alternatively destabilize a conformation that has high affinity for the ligand, resulting in a shift or redistribution of native protein conformations.[14] If the populations of conformers that have lower affinities to a ligand increase relative to other conformers, the average conformer across all conformational states will have a lower binding affinity, which can be observed as an increase in K_D. Hydrogen-deuterium exchange experiments have shown that movements of the βC – strand are related to the preorganization process, where the enzyme samples conformational substates before hydride transfer and, we speculate, to inhibitor binding.[43] Therefore, a possible explanation for the reported data could be that cluster 1 mutations perturb the preorganization process by altering amino acid networks. In other

words, allosteric cluster 1 mutations could cause a shift in the proportion of conformers that have different PYR and TMP binding affinities.

4. Summary

In summary, we generated eleven *Bs* DHFR variants that interrogate amino acid residues within and distal to the active site to study catalysis and allostery in *Bs* DHFR. Our results indicate that active site residue, Asp27, is critical for *Bs* DHFR activity, which conforms to previously found results in the *Ec* DHFR. The Met20 loop is also catalytically relevant in the *Bs* DHFR, noted by the 113-fold reduction in activity for the R19D mutation studied in this work. This residue has not previously been studied, presumably as it does not make contacts with the NADPH cofactor; however, Arg19 may play a role in Met20 loop dynamics for the *Bs* DHFR. The K_D analysis of allosteric cluster 1 indicates that variants with small changes in k_{cat} and K_M values could have significant effects on ligand affinity and ligand selectivity, particularly influencing TMP binding. We found that these effects can be side chain independent at certain residue positions (39), but also be side chain dependent at other positions (38 and 40). These observations are interesting because these residues do not directly interact with ligands in the binding site, yet they are involved in inhibitor-selectivity. It is possible that this region affects ligand-selectivity by perturbing the preorganization process, where the enzyme samples multiple conformations in preparation to ligand binding and catalysis. Further understanding of allosteric regions that dictate these qualities has implications for the discovery of novel therapeutic allosteric compounds and for protein engineering.

5.1 Supplemental methods

Constitutive CAT promotor construct

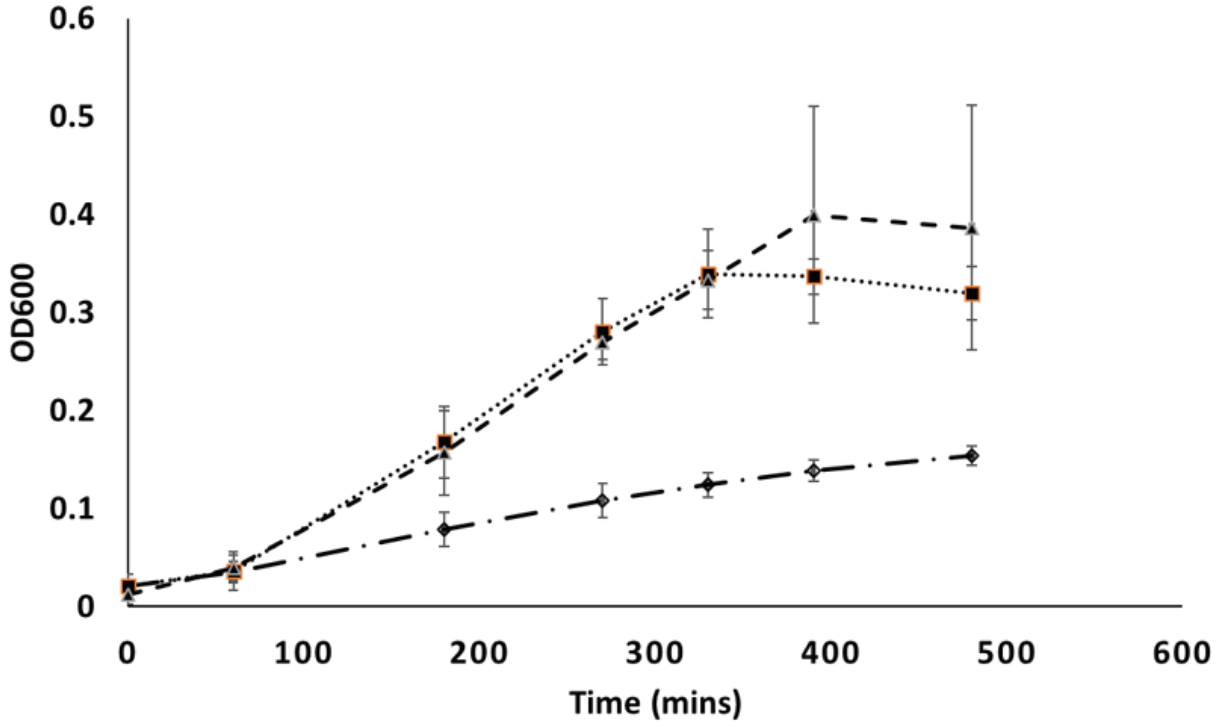
The pET21a+ plasmid containing *Bs* DHFR was cut using BglII/XbaI[2]. The CAT promoter sequence was taken from the Registry of Standard Biological Parts and manufactured by Integrated DNA Technologies, Inc. and includes the appropriate restriction enzyme sites (CAT Promoter Forward: 5'-gatcggcacgtaagaggttccaactttcaccataatgaaaca-3'), (CAT Promoter Reverse: 5'- ctagtgtttcattatggtgaaagttggaacctcttacgtgcc – 3')[54]. The CAT promotor was inserted into the BglII/XbaI cut pET21a+ DHFR vector using a 3:1 insert to vector molar ratio and 1 µL of T4 ligase in a 20 µL reaction volume. The ligase reaction was incubated overnight at 16 °C before being used for further applications.

Liquid culture growth curves

Two isolated folA- transformants (folA- with wild type *Bs* DHFR) under the control of the constitutive CAT promotor and the folA- strain were grown to saturation in 5 mL LB with 25 µg/mL kanamycin, 50 µg/mL thymidine, and 100 µg/mL ampicillin for the transformants. Following growth overnight, the cultures were washed three times with 1 X minimal media A (1 g/L $(NH_4)_2SO_4$, 4.5 g/L KH_2PO_4, 10.5 g/L $K_2H_2PO_4$, and 0.5 g/L sodium citrate with additions of 0.2% glucose, 0.6 mM arginine, and 1 mM $MgSO_4$) and resuspended to the same density as determined by an OD600 reading. Three individual flasks containing 100 mL of 1 X MMA supplemented with 25 µg/mL kanamycin and 50 µg/mL thymidine were inoculated with 1 mL of the washed overnight cultures. The flasks were incubated at 37 °C with shaking at 225

RPM for 8 hours. OD600 readings were taken throughout the incubation period and plotted to create the growth curve. Measurements were performed in triplicate.

5.2 Supplemental Figures



Supplemental Figure 1. Growth differential experiment. MH829 cell line (◊) and 2 individually isolated cell lines that have been transformed with wild-type *Bs DHFR* (■, ▲). Cells were grown in 1 X MMA solution supplemented with 50 µg/mL thymidine and OD600 readings were taken at set time points. Error bars represent the standard deviation of three replicate experiments.

6. References

1. Goodey, N. M.; Herbert, K. G.; Hall, S. M.; Bagley, K. C., Prediction of residues involved in inhibitor specificity in the dihydrofolate reductase family. *Biochimica et biophysica acta* **2011,** *1814* (12), 1870-9.
2. Hui Sun, K.; Damo, S. M.; Seok-Yong, L.; Wemmer, D.; Klinman, J. P., Structure and Hydride Transfer Mechanism of a Moderate Thermophilic Dihydrofolate Reductase from Bacillus stearothermophilus and Comparison to Its Mesophilic and Hyperthermophilic Homologues. *Biochemistry* **2005,** *44* (34), 11428-14439.
3. Singer, S.; Ferone, R.; Walton, L.; Elwell, L., Isolation of a dihydrofolate reductase-deficient mutant of Escherichia coli. *Journal of Bacteriology* **1985,** *164* (1), 470-472.
4. ELIZABETH E. HOWELL, J. E. V., MARK S. WARREN, STUART J. OATLEY, JOSEPH KRAum, Functional Role of Aspartic Acid-27 in Dihydrofolate Reductase Revealed by Mutagenesis. *Science* **1985,** *231*, 1123-1128.
5. Perry, K. M.; Onuffer, J. J.; Touchette, N. A.; Herndon, C. S.; Gittelman, M. S.; Matthews, R. C.; Chen, J. T.; Mayer, R. J.; Taira, K., Effect of single amino acid replacements on the folding and stability of dihydrofolate reductase from Escherichia coli. *Biochemistry* **1987,** *26* (10), 2674-2682.
6. Appleman, J. R.; Howell, E. E.; Kraut, J.; Kühl, M.; Blakley, R. L., Role of aspartate 27 in the binding of methotrexate to dihydrofolate reductase from Escherichia coli. *Journal of Biological Chemistry* **1988,** *263* (19), 9187-9198.
7. Watson, M.; Liu Jw Fau - Ollis, D.; Ollis, D., Directed evolution of trimethoprim resistance in Escherichia coli. **2007,** (1742-464X (Print)).
8. Liu, C. T.; Kevin, F.; Joshua, P. L.; Xinyi, H.; Sharon, H.-S.; Amnon, K.; Stephen, J. B., Escherichia coli dihydrofolate reductase catalyzed proton and hydride transfers: Temporal order and the roles of Asp27 and Tyr100. *Proceedings of the National Academy of Sciences of the United States of America* **2014,** *111* (51), 18231.
9. Mhashal, A. R.; Vardi-Kilshtain, A.; Kohen, A.; Major, D. T., The role of the Met20 loop in the hydride transfer in Escherichia coli dihydrofolate reductase. *Journal of Biological Chemistry* **2017,** *292* (34), 14229-14239.
10. Abdizadeh, H.; Tamer, Y. T.; Acar, O.; Toprak, E.; Atilgan, A. R.; Atilgan, C., Increased substrate affinity in the Escherichia coli L28R dihydrofolate reductase mutant causes trimethoprim resistance. *Physical Chemistry Chemical Physics* **2017,** *19* (18), 11416-11428.
11. Swanwick, R. S.; Shrimpton, P. J.; Allemann, R. K., Pivotal Role of Gly 121 in Dihydrofolate Reductase from Escherichia coli: The Altered Structure of a Mutant Enzyme May Form the Basis of Its Diminished Catalytic Performance. *Biochemistry* **2004,** *43* (14), 4119-4127.
12. Stojković, V.; Perissinotti, L. L.; Lee, J.; Benkovic, S. J.; Kohen, A., The effect of active-site isoleucine to alanine mutation on the DHFR catalyzed hydride-transfer. *Chemical Communications (Cambridge, England)* **2010,** *46* (47), 8974-8976.
13. Goodey, N. M.; Benkovic, S. J., Allosteric regulation and catalysis emerge via a common route. *Nature chemical biology* **2008,** *4* (8), 474-82.
14. Naganathan, A. N., Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. *Current Opinion in Structural Biology* **2019,** *54*, 1-9.
15. Swain, J. F.; Gierasch, L. M., The changing landscape of protein allostery. *Current Opinion in Structural Biology* **2006,** *16* (1), 102-108.
16. Seo, M.-H.; Park, J.; Kim, E.; Hohng, S.; Kim, H.-S., Protein conformational dynamics dictate the binding affinity for a ligand. *Nature Communications* **2014,** *5*, 3724.
17. Goodey, N. M.; Alapa, M. T.; Hagmann, D. F.; Korunow, S. G.; Mauro, A. K.; Kwon, K. S.; Hall, S. M., Development of a fluorescently labeled thermostable DHFR for studying conformational changes associated with inhibitor binding. *Biochemical and biophysical research communications* **2011,** *413* (3), 442-7.
18. Lei Zheng, U. B. a. J.-L. R., An efficient one-step-directed and site-saturation mutagenesis protocol. *Nucleic Acids Research* **2004,** *Vol. 32* (14), 5.
19. Edelheit, O.; Hanukoglu, A.; Hanukoglu, I., Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology* **2009,** *9*, 61-61.
20. Cindy Schulenburg1, Y. S., Matthias Ku¨nzle, and Donald Hilvert2, Comparative Laboratory Evolution of Ordered and Disordered Enzymes. *The Journal of Biological Chemistry* **2015**.
21. Herrington, M. B.; Chirwa, N. T., Growth properties of a folA null mutant of Escherichia coli K12. *Canadian Journal of Microbiology* **1999,** *45* (3), 191-200.

22. Tu, Q.; Yin, J.; Fu, J.; Herrmann, J.; Li, Y.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y., Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Scientific Reports* **2016,** *6*, 24648.
23. Guo, J.; Luk, L. Y. P.; Loveridge, E. J.; Allemann, R. K., Thermal Adaptation of Dihydrofolate Reductase from the Moderate Thermophile Geobacillus stearothermophilus. *Biochemistry* **2014,** *53* (17), 2855-2863.
24. Perez-Abraham, R.; Sanchez, K. G.; Alfonso, M.; Gubler, U.; Siekierka, J. J.; Goodey, N. M., Expression, purification and enzymatic characterization of Brugia malayi dihydrofolate reductase. *Protein expression and purification* **2016,** *128*, 81-5.
25. Bourne, C. R.; Barrow, E. W.; Bunce, R. A.; Bourne, P. C.; Berlin, K. D.; Barrow, W. W., Inhibition of Antibiotic-Resistant Staphylococcus aureus by the Broad-Spectrum Dihydrofolate Reductase Inhibitor RAB1. *Antimicrobial Agents and Chemotherapy* **2010,** *54* (9), 3825.
26. Bourne, C. R.; Wakeham, N.; Webb, N.; Nammalwar, B.; Bunce, R. A.; Berlin, K. D.; Barrow, W. W., The Structure and Competitive Substrate Inhibition of Dihydrofolate Reductase from Enterococcus faecalis Reveal Restrictions to Cofactor Docking. *Biochemistry* **2014,** *53* (7), 1228-1238.
27. Dunigan, D.; Waters, S. B.; Owen, T. C., *Aqueous soluble tetrazolium/formazan MTS as an indicator of NADH- and NADPH-dependent dehydrogenase activity*. 1995; Vol. 19, p 640-9.
28. Hulme, E. C.; Trevethick, M. A., Ligand binding assays at equilibrium: validation and interpretation. *Br J Pharmacol* **2010,** *161* (6), 1219-1237.
29. BERRY BIRDSALL, R. W. K., * MIRIAM R. WHEELER,? CHARLES A. LEWIS, JR.,? SCOTT R. GOODE,? R. BRUCE DUNLAP,? AND GORDON C. K. ROBERTS*, Correction for light absorption in fluorescence studies of Protein-Ligand Interactions. **1983**.
30. Williams, J. W.; Morrison, J. F., [17] The kinetics of reversible tight-binding inhibition. In *Methods Enzymol*, Academic Press: 1979; Vol. 63, pp 437-467.
31. Shrimpton, P.; Allemann, R. K., Role of water in the catalytic cycle of E. coli dihydrofolate reductase. *Protein Sci* **2002,** *11* (6), 1442-1451.
32. Schnell, J. R.; Dyson, H. J.; Wright, P. E., STRUCTURE, DYNAMICS, AND CATALYTIC FUNCTION OF DIHYDROFOLATE REDUCTASE. *Annual Review of Biophysics* **2004,** *33* (1), 119.
33. Wan, Q.; Bennett, B. C.; Wilson, M. A.; Kovalevsky, A.; Langan, P.; Howell, E. E.; Dealwis, C., Toward resolving the catalytic mechanism of dihydrofolate reductase using neutron and ultrahigh-resolution X-ray crystallography. PNAS: 2014; Vol. 111, pp 18225-18230.
34. Luk, L. Y. P.; Ruiz-Pernía, J. J.; Dawson, W. M.; Loveridge, E. J.; Tuñón, I.; Moliner, V.; Allemann, R. K., Protein Isotope Effects in Dihydrofolate Reductase From Geobacillus stearothermophilus Show Entropic–Enthalpic Compensatory Effects on the Rate Constant. *Journal of the American Chemical Society* **2014,** *136* (49), 17317-17323.
35. McElheny, D.; Schnell, J. R.; Lansing, J. C.; Dyson, H. J.; Wright, P. E., Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (14), 5032.
36. Liu, C. T.; Hanoian, P.; French, J. B.; Pringle, T. H.; Hammes-Schiffer, S.; Benkovic, S. J., Functional significance of evolving protein sequence in dihydrofolate reductase from bacteria to humans. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (25), 10159-10164.
37. Bhabha, G.; Ekiert, D. C.; Jennewein, M.; Zmasek, C. M.; Tuttle, L. M.; Kroon, G.; Dyson, H. J.; Godzik, A.; Wilson, I. A.; Wright, P. E., Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nature Structural & Molecular Biology* **2013,** *20* (11), 1243-1249.
38. Oyeyemi, O. A.; Sours, K. M.; Lee, T.; Resing, K. A.; Ahn, N. G.; Klinman, J. P., Temperature dependence of protein motions in a thermophilic dihydrofolate reductase and its relationship to catalytic efficiency. *Proceedings of the National Academy of Sciences* **2010,** *107* (22), 10074.
39. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004,** *25* (13), 1605-1612.
40. Saraf, M. C.; Moore, G. L.; Goodey, N. M.; Cao, V. Y.; Benkovic, S. J.; Maranas, C. D., IPRO: an iterative computational protein library redesign and optimization procedure. *Biophys J* **2006,** *90* (11), 4167-80.
41. Elizabeth Ehrhardt Howell, t. P. G. F., and Lisa M. Foster, Construction of a Dihydrofolate Reductase-Deficient Mutant of Escherichia coli by Gene Replacement. *Journal of Bacteriology* **1988,** *170*.
42. Reynolds, K. A.; McLaughlin, R. N.; Ranganathan, R., Hot spots for allosteric regulation on protein surfaces. *Cell* **2011,** *147* (7), 1564-1575.
43. Oyeyemi, O. A.; Sours, K. M.; Lee, T.; Kohen, A.; Resing, K. A.; Ahn, N. G.; Klinman, J. P., Comparative Hydrogen-Deuterium Exchange for a Mesophilic vs Thermophilic Dihydrofolate Reductase at 25

°C: Identification of a Single Active Site Region with Enhanced Flexibility in the Mesophilic Protein. *Biochemistry* **2011,** *50* (38), 8251-8260.
44. Agarwal, P. K.; Doucet, N.; Chennubhotla, C.; Ramanathan, A.; Narayanan, C., Conformational Sub-states and Populations in Enzyme Catalysis. *Methods Enzymol* **2016,** *578*, 273-297.
45. Dunbrack, R. L., Jr., Rotamer libraries in the 21st century. (0959-440X (Print)).
46. Volpato, J. P.; Fossati, E.; Pelletier, J. N., Increasing Methotrexate Resistance by Combination of Active-site Mutations in Human Dihydrofolate Reductase. *Journal of Molecular Biology* **2007,** *373* (3), 599-611.
47. Lamb, K. M.; G-Dayanandan, N.; Wright, D. L.; Anderson, A. C., Elucidating Features That Drive the Design of Selective Antifolates Using Crystal Structures of Human Dihydrofolate Reductase. *Biochemistry* **2013,** *52* (41), 7318-7326.
48. Kongsaeree, P.; Khongsuk, P.; Leartsakulpanich, U.; Chitnumsub, P.; Tarnchompoo, B.; Walkinshaw, M. D.; Yuthavong, Y., Crystal structure of dihydrofolate reductase from Plasmodium vivax: Pyrimethamine displacement linked with mutation-induced resistance. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (37), 13046.
49. Dale, G. E.; Broger, C.; D' Arcy, A.; Hartman, P. G.; DeHoogt, R.; Jolidon, S.; Kompis, I.; Labhardt, A. M.; Langen, H.; Locher, H.; Page, M. G. P.; Stüber, D.; Then, R. L.; Wipf, B.; Oefner, C., A single amino acid substitution in Staphylococcus aureus dihydrofolate reductase determines trimethoprim resistance 1 1 Edited by T.Richmond. *Journal of Molecular Biology* **1997,** *266* (1), 23-30.
50. Queener, S. F.; Cody, V.; Pace, J.; Torkelson, P.; Gangjee, A., Trimethoprim resistance of dihydrofolate reductase variants from clinical isolates of Pneumocystis jirovecii. *Antimicrobial agents and chemotherapy* **2013,** *57* (10), 4990-4998.
51. Eliopoulos, G. M.; Huovinen, P., Resistance to Trimethoprim-Sulfamethoxazole. *Clinical Infectious Diseases* **2001,** *32* (11), 1608-1614.
52. Anderson, A. C., Targeting DHFR in parasitic protozoa. *Drug Discovery Today* **2005,** *10* (2), 121-128.
53. Tzeng, S.-R.; Kalodimos, C. G., Protein dynamics and allostery: an NMR view. *Current Opinion in Structural Biology* **2011,** *21* (1), 62-67.
54. Vijayan, V. H., Allen; Fomundam, Lawrence Part:BBa_I14033. http://parts.igem.org/Part:BBa_I14033 (accessed March).