Testing Predicted Specificity Determinants in the Dihydrofolate Reductase Enzyme Family

Seema J. Patel

Follow this and additional works at: https://digitalcommons.montclair.edu/etd

Part of the Biochemistry Commons, and the Chemistry Commons
Abstract

Dihydrofolate Reductase (DHFR) is an essential enzyme for most organisms, ranging from bacteria to humans. DHFR has essential functions in DNA biosynthesis and cell replication; as a result, cell growth can be inhibited by the inhibition of DHFR. Evaluating the use of well-known DHFR inhibitors is becoming essential in treating infections in the developing world as DHFR is a known target of antibacterial and antiparasitic drugs. Understanding determinants of DHFR inhibitor specificity in terms of amino acid sequence and structure will allow repurposing or designing of new compounds that selectively target DHFR from the pathogenic organism of interest over the Hs DHFR. Previously, a computational analysis was developed to predict allosteric residues involved in ligand discrimination using DHFR as a model system. The approach was based on inhibitor specificity and amino acid composition for sets of protein homolog pairs, predicting eighteen alignment positions. The residues were clustered as follows: three of the residues are found in the active site; four of the residues are proximal to the active site, four of the residues are clustered together in the adenosine binding domain and five of the residues are on the ββG loop. Many of the predicted residues are located in allosteric region away from the active site. My role in this project was to experimentally validate these predictions using site specific mutations in B. Stearothermophilus DHFR gene as a model system. To ensure all mutants were correctly folded and active, turnover numbers \((k_{cat})\) and Michaelis constants \((K_M)\) were measured for wildtype and mutants. \(K_I\) values of the 12 single mutants against four DHFR inhibitors, methotrexate, trimethoprim, pyrimethamine, and raltitrexed were determined. Interestingly, comparisons of the ligand binding profiles of the mutants to...
those of the wildtype enzyme revealed significant changes in ligand specificity, supporting the predictions. In addition, the effects of mutations on $K_1$ values are ligand specific.
MONTCLAIR STATE UNIVERSITY

Testing Predicted Specificity Determinants in the Dihydrofolate Reductase Enzyme Family

by

Seema J. Patel

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 2014

College of Science and Mathematics

Department of Chemistry and Biochemistry

Certified by

Robert Prezant, PhD.
Dean of College of Science and Mathematics

Thesis Committee:

Nina M. Goody, PhD.
Thesis Sponsor

John Siekierka, PhD.
Committee Member

David P. Rotella, PhD.
Committee Member

Johannes P. Schelvis, PhD.
Chairperson, Department of Chemistry and Biochemistry

*August 2014
Testing Predicted Specificity Determinants in the Dihydrofolate Reductase Enzyme Family

A THESIS

Submitted in partial fulfillment of the requirements
For the degree of Master of Science

by
Seema J. Patel
Montclair State University
Montclair, NJ
2014
Acknowledgements

I would like to dedicate this thesis to my family (Jayanti Patel (father), Kapila Patel (mother), Jigisha Patel (sister), and Chirag Patel (brother) for their endless love, support, and encouragement. I would like to thank my mentor; Dr. Nina Goodey for her advice, support, and for giving me the opportunity to work with her. I would also like to thank Dr. Goodey’s previous student: Sarah Hall and Michael Little for providing me with the mutant genes and Michael Little, Dina Edani and Maureen Malivert for helping me by providing purified mutants. I would like to thank Graduate Summer Research Fellowship, the Margaret E. Herman Sokol Institute and Women in Science (WIS) for funding. I am grateful to have performed molecular graphics and analyses with the freely available UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311). I would like to thank my readers, Dr. D. Rotella and Dr. J. Siekierka for critical thinking and contribution to my thesis.
Table of Contents

1. Introduction 8
   1.1. Dihydrofolate Reductase - An important drug target 8
   1.2. Predicted specificity determining determinants 9
   1.3. *Bacillus stearothermophilus* - model system 12

2. Materials and Methods 13
   2.1. Expression and Purification of *Bs* DHFR 13
   2.2. Synthesis of dihydrofolate 14
   2.3. Kinetic analysis 15
   2.4. Inhibition assay 15
   2.5. Sequence and structural alignment 16

3. Results and Discussion 16
   3.1. Predicted residues in model system *Bs* DHFR 16
   3.2. Structural alignments 18
   3.3. Steady-state kinetic analysis 22
   3.4. Determination of $K_1$ values 27
   3.5. Effects of active site and second shell mutations 32
   3.6. Effects of distal mutations on $K_1$ 34
   3.7. Effects of mutations are drug specific 36
   3.8. Ratios of $K_1$ values of different inhibitors 37

4. Summary 40

References 41
1. Introduction

1.1. Dihydrofolate Reductase - An important drug target

Dihydrofolate Reductase (DHFR) is an essential enzyme that plays important roles in DNA biosynthesis and cell replication and is therefore an important therapeutic drug target for many diseases, including cancer and microbial infections (1-3). DHFR has gained interest as a drug target against parasitic infections caused by protozoa *Leishmania major* (*Lm*), *Trypanosoma cruzi* (*Tc*) and *brucei* (*Tb*), and *Plasmodium falciparum* (*Pf*) and opportunistic infections caused by *Pneumocystis carinii* (*Pc*) and *Toxoplasma gondii* (*Tg*) which are often catastrophic for patients with immune deficiencies (4-7). Inhibitor potency varies for different DHFR homologs. It is critical to understand inhibitor specificity of DHFR homologs from different organisms and to design compounds that selectively inhibit DHFR from infectious organisms while leaving *Homo sapiens* (*Hs*) DHFR alone.

Active site residues are extensively studied in protein-ligand interactions as they can form direct intermolecular interactions with a ligand. Generally, changes in kinetic parameters and binding interactions are examined that result from mutations of active site residues that are replaced with residues possessing different chemical properties to explore the determinants of ligand binding and specificity. In the DHFR family, numerous conserved active site residues have been identified through the sequencing of drug resistant cell lines that are known to be involved in ligand binding (8). These studies are essential in understanding the role of specific residues in protein-ligand interactions and ligand selectivity.
Enzymologists have recently begun to view protein structure as a combination of conformers rather than one single conformer obtained from crystal structures. The structure of the conformation can have an effect on the active site structure and environment. The equilibrium of structural conformers can be influenced by primary protein structure, which as a result can impact the population of active sites. The replacement of amino acids in the allosteric region can alter enzyme properties, which has been observed in E. coli DHFR and others (9, 10). Thus, it is essential to consider the role of the entire protein rather than just the binding or active site to understand enzyme structure-function relationships in protein-ligand binding.

1.2. Predicted specificity determining determinants

A homology-based, computational method was developed to predict inhibitor (drug) specificity determinants in a protein family (11). This method was applied to the DHFR enzyme family and 18 amino acid positions were predicted to influence drug specificity. The prediction was done using 13 organisms which were Hs, E. coli (Ec), M. tuberculosis (Mt), T. brucei (Tb), T. gondii (Tg), L. major (Lm), T. cruzi (Tc), P. falciparum (Pf), P. vivax (Pv), R. norvegicus (Rn), P. carinii (Pc), L. casei (Lc) and M. avium (Ma). The predicted ligand specificity determining residues are located in the active site and allosteric region (labeled using the Hs numbering system). Residues located in the allosteric region, outside the ligand binding domain, have no direct intermolecular interactions with a ligand. Allosteric residues 48, 49, 50, and 55 are located on the βC strand that is located above the ligand binding site which could influence ligand binding through second shell effects as shown in figure 1A. Likewise,
the binding site environment could be indirectly affected by residue 15 which is located close to the active site on the M20 loop. In addition, ten other allosteric residues (88, 93, 99, 100, 140, 144, 146, 147, 148, and 180) were predicted. As these are distal residues, they cannot influence ligand specificity directly through specific interactions. In the adenosine binding domain, residues 93, 99, and 100 are found on the Ca-helix and residue 88 is in close proximity to the Ca-helix. Residues 140, 144, 146, 147, and 148 are located on the βFβG loop which is predicted to undergo important motions during binding and catalysis for the E. coli DHFR (12-14). The effects of distal mutations on enzyme properties have been observed through site specific mutagenesis (15-18). Enzyme properties could be influenced through mutations of residues located on the M20 and βFβG loops as these loops are involved in the hydride transfer step required for catalysis (18). In NMR relaxation experiments, it has been observed that upon substrate and the coenzyme binding, the M20 and βFβG loops undergo conformational changes (14, 19, 20).
Allostery is the regulation of protein properties influenced by the binding of a small molecule at a distal region that is located away from the active site (10). One way to look at the allosteric effect is to introduce a mutation in the allosteric region and observe the effects (9, 10). In most cases, protein mutants have comparable properties as
their wildtype structures. Active site residue mutations have direct influence on enzyme properties as these amino acid sidechains have direct intermolecular interactions with the ligand. However, this might not be the case when a distal residue is mutated.

The aim of this study was to experimentally validate the previously developed method of predicting determinants of inhibitor specificity. \textit{Bs} is used as a model system to experimentally test the predicted residues. Protein structure and function can be correlated using mutagenesis (9). Enzyme properties could be influenced due to mutation to alter the local interactions. The local interactions in turn could affect the overall structure resulting in changes in kinetic parameters. The following mutations were introduced into the \textit{Bs} DHFR gene and purified: V6P, V13Y, H38N, A39Y, I40A, G51I, L74R, W85L, I86A, T125A, F126L, and Y127A as outlined previously. Inhibition constants (K_\text{I}) were measured for methotrexate (MTX), trimethoprim (TMP), pyrimethamine (PYR), and raltitrexed (RLX) for wildtype and mutants. The Inhibition constant (IC_{50}) is the inhibitor concentration at which maximal rate is decreased by half. As mentioned, to ensure all mutants were correctly folded and active, turnover numbers (k_{cat}) and Michaelis constants (K_M) were measured for wildtype and mutants. The turnover number, also known as, k_{cat} is the number of substrate molecules converted to product per molecule of enzyme per unit of time. Michaelis constant (K_M) is the substrate concentration at half of the maximum rate.

1.3. \textit{Bacillus stearothermophilus} - model system

\textit{Bacillus stearothermophilus (Bs)} DHFR is a 18.7 kD monomeric thermophilic protein (21, 23). \textit{Bs} DHFR was used as the model system to experimentally test whether
these residues are indeed ligand specificity determinants in the DHFR enzyme family. *Bs* DHFR was chosen as the model system because it was not one of the 13 DHFRs used for the computational predictions, thus, serving as a useful and independent model (11). The $\alpha$-carbons of predicted specificity determining residues in *Hs* DHFR are shown as gray spheres in figure 1A and the corresponding residues in *Bs* DHFR are shown in figure 1B. This was done to ensure the *Hs* DHFR residues positions correspond to *Bs* DHFR residue positions in space. For example, residue position 8 in *Hs* DHFR corresponds to residue position 6 in *Bs* DHFR. The two structures show that residue positions in *Hs* DHFR correspond to residue positions in *Bs* DHFR in space.

2. Materials and Methods

2.1. Expression and Purification of *Bs* DHFR

*Bs* DHFR (wildtype and mutants) were expressed in *E. coli* BL21 cells and grown in LB broth which contained 100 mg/mL ampicillin at 37 °C and 225 RPM. Expression was induced using 1 mM IPTG when OD$_{600}$ reached 0.6 and the cultures were grown for 12-16 h at 30 °C and 225 RPM. Cells were pelleted by centrifuging the cultures using Beckman Avanti J-26 XP High Speed Centrifuge with a rotor of AJ 10 at 5,000 RPM and 4 °C for 20 minutes. The pelleted cells were washed with 0.9 % NaCl and centrifuged. The pellet was dissolved in HEPES buffer, lysed by sonication, and centrifuged as mentioned above. The supernatant was loaded on to Sephadex C$_{25}$ column and eluted with HEPES buffer containing 0.2 M NaCl. NaCl was removed by several cycles of concentrating followed by diluting the protein in HEPES buffer containing 10 % glycerol and 1 mM DTT. Protein purity was confirmed to be >90% by SDS-PAGE.
Pure protein fractions were concentrated and stored at -80 °C in HEPES buffer with 10% glycerol and 1 mM DTT.

2.2. Synthesis of dihydrofolate

The synthesis was carried out in containers covered with foil as dihydrofolate (DHF) is light sensitive. Ascorbate solution (0.57 M) at pH 7.0 was prepared using cold (4 °C) distilled water, the pH was adjusted using 1 M or 6 M NaOH. Folate solution (0.0566 M) was prepared using 0.01 M NaOH. Ascorbate solution (20 mL) and folate solution (3.2 mL) were combined to which was added 941 mg of sodium hydrosulfite, which was then stirred for 5 minutes. The pH was raised to 6.8 to 7.1 using 1 M NaOH; the pH of the solution was checked until it was constant (pH 6.8 – 7.1). The temperature of the solution was lowered to 5 °C using an ice bath. The solution was titrated (4 drops/min) using 1 M HCl over a period of approximately two hours until pH reached 2.8. The solution was then aliquoted to 1 mL eppendorf tubes and centrifuged for 5 minutes (4 °C and 14,000 RPM). The pellet was resuspended in fresh ascorbate solution (2 g of sodium ascorbate/20 mL of distilled water), stirred for 5 minutes, and pH was measured using pH paper. The pH was adjusted to 6.6 (clear color) and recrystallization was done with 1 M HCl (pH ~5). The solution was again aliquoted to 1 mL eppendorf tubes and centrifuged for 5 minutes (4 °C and 14,000 RPM). The pellet was washed four times with ~200 mL of 5 mM HCl (each time centrifuged at 4 °C and 14,000 RPM). The pellet was resuspended in 5 mL of 5 mM HCl. DHF concentration was measured at 283 nm and 340 nm using 0.01 M NaOH as a reference. The extinction coefficient used to measure DHF concentration was 28 000 M⁻¹ cm⁻¹ at 283 nm. After concentration determination, 7.5 µL
of 14 M 2-mercaptoethanol (Sigma M- 6250) was added to the DHF solution and was stored in -80 °C freezer in small aliquots [Blakely (1960) Nature 188,231-232].

2.3. Kinetic analysis

DHFR (wildtype and mutant) activity ($k_{cat}$) was measured at 100 µM NADPH, 100 µM dihydrofolate (DHF), and 25 °C in HEPES buffer (40 mM HEPES, 1 mM DTT at pH 6.8) using a Thermo Scientific Evolution 201/220 UV-Visible Spectrophotometer. The reaction was initiated by the addition of DHF and NADPH and DHF depletion was recorded at 340 nm. Molar differential extinction coefficient of 13.2 mM$^{-1}$ cm$^{-1}$ at 340 nm, which takes into consideration the reduction of DHF and oxidation of NADPH, was used to calculate the catalytic activity ($k_{cat}$). Michaelis constant ($K_M$) for DHF was determined for DHFR (wildtype and mutants) by measuring DHFR activity at 25 °C, 15-30 nM DHFR, 100 µM NADPH and various DHF concentrations, ranging from 0 to 120 µM in HEPES buffer (21). Rate ($\mu$M*s$^{-1}$) was plotted as a function of DHF concentration (µM) using the KaleidaGraph software (Synergy software). Michaelis-Menten equation was fitted to the data using KaleidaGraph software to obtain the $K_M$ and $V_{max}$ values. * NADPH was commercially available; the buffer and NADPH solution was prepared in the lab.

2.4. Inhibition assay

$K_I$ values of $Bs$ DHFR (wildtype and mutants) for MTX, TMP, PYR, and RLX were determined using inhibitor concentrations that resulted in 10 – 40 % reduction in activity. The range of inhibitor concentrations were as follows: MTX (0.005 – 0.5 nM),
TMP (0.05 – 1.0 nM), PYR (0.0025 – 0.4 nM), and RLX (0.5 – 2.0 nM). DHFR activity was measured at 25 °C and 15-30 nM DHFR, 100 µM NADPH and various DHF concentrations (5 - 50 µM) in HEPES buffer. The absorbance was recorded at 340 nm for 1-3 minutes. $K_I$ values were determined by plotting the inverse rate ($s/\mu$M) as a function of inhibitor concentration (nM) using a Dixon plot (24, 25).

2.5. Sequence and structural alignment

The sequence alignment of DHFR homologs were done using Uniport (default settings) and the DHFR homolog sequences were obtained from Uniprot. The Chimera software (version 1.8.1) was used to make the structural alignments (22). Structural alignments were created using the following commands in Chimera: Tools → Structure Comparison → MatchMaker command. The structural alignments were created using the $Hs$ DHFR structure complexed with NADPH and TMP in the active site (PDB Code: 2W3A) and the $Bs$ DHFR (PDB Code: 1ZDR).

3. Results and Discussion

3.1. Predicted residues in model system $Bs$ DHFR

Previously, eighteen residues were predicted to modulate ligand specificity in the DHFR enzyme family (11). The predicted residues are located in the active site and the allosteric region which is away from the active site. Twelve residues were experimentally tested, which are displayed on the $Hs$ DHFR structure (PDB code: 2W3A) using $Hs$ numbering (Figure 1A). $K_I$ values were measured for four drugs (MTX, TMP, PYR, and RLX) for wildtype and mutants. The inhibition analysis was performed to investigate
whether the mutations alter ligand specificity. A sequence alignment was done between
the *Hs* and *Bs* DHFR sequences. In figure 2, the sequence alignment shows which
residues in *Bs* DHFR are aligned with the predicted specificity determinants in the *Hs*
sequence. For example, residue position V8 in *Hs* DHFR corresponds to residue
position V6 in *Bs* DHFR. Similarly, residue position L99 in *Hs* DHFR corresponds to
residue position W85 in *Bs* DHFR. Previous students made mutations at the
corresponding residue positions in *Bs* DHFR; these residues are presented in figure 1B
using *Bs* numbering and residue names. The figure shows residue V6 in the active site,
V13 located on the proximal M20 loop, and G51 as a second shell residue. Using the
Chimera software (version 1.8.1), a hydrogen bond (2.87 Å) is found between the
sidechain amine group of N64 (G51 in *Bs* DHFR) and the carbonyl oxygen of ρ-
aminobenzoate ring of folate (PDB code: 1DRF) in the *Hs* DHFR. If a similar hydrogen
bond is present in *Bs* DHFR, then G51 could be considered an active site residue as it
would have direct intermolecular interaction with the ligand. Residues H38, A39, I40,
L74, W85, I86, T125, F126, and Y127 lie in allosteric regions away from the active site.
The twelve studied predicted specificity determinants are located in five clusters: The
active site residue (V6), second shell residues (V13 and G51), allosteric cluster 1 (H38,
A39, and I40), allosteric cluster 2 (L74, W85, and I86) and allosteric cluster 3 (T125,
F126, and Y127) which is located on the βFβG loop.
Figure 2. Amino acid sequence alignment between *Hs* and *Bs* DHFRs, created using Uniprot (default settings) (27). The alignment was done using 14 organisms (*Bs*, *Hs*, *Ec*, *Mt*, *Tb*, *Tg*, *Lm*, *Tc*, *Pf*, *Pv*, *Rn*, *Pc*, *Lc*, and *Ma*); only the *Hs* DHFR and *Bs* DHFR sequences are shown for clarity. *Bs* (top) and *Hs* DHFR (bottom) alignment positions included in this study are indicated by black boxes with residue labels using the *Bs* (top) and *Hs* (bottom) numbering. In *Hs* DHFR sequence, the first position is counted as M0, rather than M1; this numbering system has been used historically for *Hs* DHFR and is being used in this paper for continuity (28). In the *Bs* DHFR sequence, gaps were moved so that residue H38 is positioned adjacent to A39 for a better alignment of the residues. In the structural alignment, H38 (*Bs* DHFR) was a better match for N48 (*Hs* DHFR).

3.2. Structural alignments

The results of the amino acid sequence alignment were confirmed by performing structure alignment for the predicted residues to ensure the residues in *Bs* DHFR most closely corresponded to predicted residue positions in *Hs* DHFR. *Hs* DHFR (PDB code: 2W3A) and *Bs* DHFR (PDB code: 1ZDR) were used for the structural alignment, which
was performed using Chimera software (version 1.8.1) (22). Figure 3A shows active site and second shell residues and allosteric residues are displayed in figure 3B. Eight residues were found to fully overlap in space: V8 (Hs) and V6 (Bs), G15 (Hs) and V13 (Bs), L49 (Hs) and A39 (Bs), V50 (Hs) and I40 (Bs), F88 (Hs) and L74 (Bs), T146 (Hs) and T125 (Bs), Fl47 (Hs) and F126 (Bs), and F148 (Hs) and Y127 (Bs). Residues N48 (Hs) and H38 (Bs), N64 (Hs) and G51 (Bs), L99 (Hs) and W85 (Bs), T100 (Hs) and I86 (Bs), do not completely overlap in space because the loops the residues are located on adopt different local conformations in the Hs and Bs DHFR structures. These Bs DHFR residues are structurally the closest matches to corresponding Hs DHFR positions.
Figure 3A. The cartoon representation shows the structural alignment of active site and second shell residues of Bs DHFR and Hs DHFR. The backbones of Hs DHFR and Bs DHFR are shown in light gray, Hs DHFR residue in dim gray, and Bs DHFR in black. The residues are numbered using the Hs numbering system. The Hs DHFR is in complex with NADPH and trimethoprim. V6 is found in the DHF and NADPH binding site. The figure was created using the Chimera software (version 1.8.1) using the Tools → Structure Comparison → MatchMaker command. PDB Code: Hs DHFR (2W3A) and Bs DHFR (1ZDR).
Figure 3B. The cartoon representation shows the structural alignment of active site and second shell residues of Bs DHFR and Hs DHFR. The backbones of Hs DHFR and Bs DHFR are shown in light gray, Hs DHFR residue in dim gray, and Bs DHFR in black. The residues are numbered using the Hs numbering system. The Hs DHFR is in complex with NADPH and trimethoprim. The figure was created using the Chimera software (version 1.8.1) using the Tools → Structure Comparison → MatchMaker command. PDB Code: Hs DHFR (2W3A) and Bs DHFR (1ZDR).
3.3. Steady-state kinetic analysis

The turnover values ($k_{\text{cat}}$) (wildtype and mutants) were measured to ensure the residues selected for replacements did not result in unfolding of the protein and/or cause the protein to be inactive. The $k_{\text{cat}}$ analysis was done by measuring reaction rates at saturated DHF and NADPH concentration. The reaction mixture contained HEPES buffer (40 mM HEPES, 1 mM DTT at pH 6.8), 100 µM NADPH, 15-30 nM DHFR. The reaction was initiated by the addition of 100 µM DHF. The NADPH and DHF depletion was recorded at 340 nm, a representative decrease is shown in figure 4. The data was fitted to a liner equation to obtain slope (mM/s$^{-1}$) which was then used to calculate the catalytic activity ($k_{\text{cat}}$) using molar differential extinction coefficient of 13.2 mM$^{-1}$ cm$^{-1}$ at 340 nm. The molar differential extinction coefficient accounts for the reduction of DHF and oxidation of NADPH.

The $k_{\text{cat}}$ value of the mutant was compared to the value for wildtype $B_s$ DHFR (Table 1). From the data, it could be concluded that all 12 variants retained $> 10\%$ of wildtype activity. In addition, it could be safely assumed that all mutants are catalytically active and therefore likely share the typical DHFR structural fold, an eight-stranded $\beta$ sheet and four $\alpha$ helices connected by flexible loops (29, 30). Not surprisingly, some of the largest decreases in catalytic activity were observed for active site mutant V6P and M20 loop mutant V13Y which were approximately 4 and 6-fold compared to wildtype DHFR, respectively as seen in figure 5. The decrease in catalytic activity could be due to the absence or weaker direct interaction of these residues with the ligand. For example, for active site mutant V6P, the interaction with the ligand could change as the amino acid was changed from a valine to a proline. Proline is unique compared to other amino acids.
in the sense that it has cyclic sidechain which introduces rigidity in the local environment of the protein structure. The interactions of the neighboring residues with the ligand could be disturbed due to the rigidly of proline.

An 8-fold decrease in catalytic activity was observed compared to wildtype for second shell mutant G51I. This could be explained by the probable loss of flexibility as glycine is replaced by isoleucine. Furthermore, mutant G51I is proximal to the active site which could influence the binding of the cofactor, NADPH and/or the substrate, DHF through second shell effects. In the Hs DHFR, N64 (G51 in Bs DHFR) interacts with the “tails” of the bound folate molecules as it is present on the edge lining of the active site (31). In addition, a hydrogen bond (2.87 Å) is present between the carbonyl oxygen of p-aminobenzoate ring of folate and the sidechain amine group of N64 (PDB code: 1DRF).

Distal mutants W85L (7-fold) and I86A (5-fold) had decreases in activity compared to wildtype. Allosteric mutants I40A and T125A had approximately 4-fold decreases in activity compared to wildtype as well. On the other hand, mutants H38N, A39Y, L74R, F126L, and Y127A have comparable activities to wildtype which are distal to the active site.
<table>
<thead>
<tr>
<th>Mutant</th>
<th>$k_{cat}$ (s⁻¹)</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat(mutant)} / k_{cat(wildtype)}$</th>
<th>$K_M(mutant)/K_M(wildtype)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>11 ± 0.81</td>
<td>2.3 ± 0.27</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>V6P</td>
<td>2.4 ± 0.11</td>
<td>11 ± 0.30</td>
<td>0.2</td>
<td>4.8</td>
</tr>
<tr>
<td>V13Y</td>
<td>1.9 ± 0.07</td>
<td>2.8 ± 0.60</td>
<td>0.2</td>
<td>1.2</td>
</tr>
<tr>
<td>H38N</td>
<td>9.8 ± 0.66</td>
<td>3.7 ± 0.59</td>
<td>0.9</td>
<td>1.6</td>
</tr>
<tr>
<td>A39Y</td>
<td>11 ± 1.01</td>
<td>1.2 ± 0.12</td>
<td>1.0</td>
<td>0.51</td>
</tr>
<tr>
<td>I40A</td>
<td>2.6 ± 0.40</td>
<td>8.5 ± 0.53</td>
<td>0.2</td>
<td>3.7</td>
</tr>
<tr>
<td>G51I</td>
<td>1.5 ± 0.19</td>
<td>5.6 ± 1.1</td>
<td>0.1</td>
<td>2.4</td>
</tr>
<tr>
<td>L74R</td>
<td>5.4 ± 1.0</td>
<td>10 ± 2.1</td>
<td>0.5</td>
<td>4.4</td>
</tr>
<tr>
<td>W85L</td>
<td>1.6 ± 0.26</td>
<td>2.2 ± 0.23</td>
<td>0.1</td>
<td>0.97</td>
</tr>
<tr>
<td>I86A</td>
<td>2.4 ± 0.08</td>
<td>1.8 ± 0.17</td>
<td>0.2</td>
<td>0.76</td>
</tr>
<tr>
<td>T125A</td>
<td>2.7 ± 0.16</td>
<td>2.2 ± 0.51</td>
<td>0.2</td>
<td>0.96</td>
</tr>
<tr>
<td>F126L</td>
<td>6.1 ± 0.42</td>
<td>20 ± 0.51</td>
<td>0.5</td>
<td>8.6</td>
</tr>
<tr>
<td>Y127A</td>
<td>10 ± 0.45</td>
<td>3.9 ± 0.27</td>
<td>0.9</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table 1. The turnover ($k_{cat}$) and Michaelis constant ($K_M$) values for $B_{sr}$ DHFR wildtype and mutants are shown. The $k_{cat}$ (s⁻¹) and $K_M$ (µM) values are averages of at least 3 independent measurements.

**Figure 4.** Absorbance vs. time data. The y-axis shows absorbance at 340 nm and the x-axis shows time in seconds. The blue line represents the decrease in NADPH and DHF over time, which is recorded as absorbance at 340 nm. The data was fitted to a linear equation (red line) to obtain slope. The slope (mM/s⁻¹) was converted to rate (mM/s⁻¹) using the extinction coefficient of 13.2 mM⁻¹ * cm⁻¹. The rate (mM/s⁻¹) was divided by enzyme concentration (mM) to obtain a $k_{cat}$ (s⁻¹).
Figure 5. Catalytic activity ($s^{-1}$) for wildtype and mutants along with standard deviation error bars. The $k_{cat}$ values are averages of at least 3 independent measurements.

The Michaelis constant ($K_M$) was measured for the substrate, DHF. The data was fitted to Michaelis-Menten equation ($V_{max}*[S]/(K_M+[S])$) using Kaleidagraph, a representative example is shown in figure 6. The $K_M$ is the substrate concentration at half of the maximal velocity. $K_M$ gives direct measurement of affinity of the ligand for enzyme and indicates a change in substrate binding between wildtype and mutants. The rates were obtained in the same manner as $k_{cat}$ and were plotted as a function of DHF concentrations using Kaleidagraph software. As expected, 5-fold increase in $K_M$ compared to wildtype was observed for mutant V6P which is located in the active site. Minimal changes in $K_M$ were observed for M20 loop mutant V13Y and many of the
allosteric mutations. Surprisingly, the $K_M$ values for distal mutants I40A and L74R also increased by 5-fold compared to wildtype. On the contrary, little change in $K_M$ values were observed for allosteric mutants V13Y, H38N, A39Y, G51I, W85L, I86A, T125A and Y127A. Interestingly, allosteric cluster 3 mutant F126L showed an approximately 9-fold increase compared to wildtype despite of being away from the active site as seen in figure 7. In previous site specific mutagenesis studies, it has been observed that allosteric mutations (away from the DHFR active site) can influence substrate binding (15-18, 32).

![Graph](image.png)

**Figure 6.** G51I Michaelis-Menten curve. The rate ($\mu$M/s) was plotted as a function of dihydrofolate concentrations ($\mu$M). Rates were obtained at various DHF concentrations as described in figure 4. The data was fitted to Michaelis-Menten equation ($V_{max}*[S]/(K_M+[S])$) and resulted in a $K_M$ of 6.02 $\mu$M and $V_{max}$ of 0.05 (s$^{-1}$).
Figure 7. Graphical representation of the $K_M$ values (µM) for wildtype and mutants along with standard deviation error bars. The $K_M$ values are averages of at least 3 independent measurements.

3.4. Determination of $K_I$ values

$K_I$ values were obtained using a Dixon plot, an example is shown in figure 8. The $K_I$ data was normalized for visual comparison of the ligand specificity profiles of all variants which is displayed in figure 9. As observed from kinetic studies, all active site and allosteric mutants retained high DHFR activities and were therefore considered appropriate models for wildtype $B_5$ DHFR. To validate the previously made predictions, the $K_I$ values were measured for all of the single-mutants for four competitive DHFR
inhibitors (MTX (chemotherapeutic agent), TMP (antibacterial agent), PYR (antiprotozoal agent), and RLX (chemotherapeutic agent)) and then compared to the ligand specificity profile of the wildtype enzyme (34). The ligand specificity profile shows the preference an enzyme has for one ligand over another ligand. A change in ligand preference by a variant would support the previously made predictions (8). In other words, if a mutation at the predicted residue position alters the ligand specificity profile, it would mean that the position mutated is important for ligand specificity. The $K_I$ analysis shows that the predicted residue position mutants remarkably alter inhibitor potency. Significant changes (> 11-fold) in $K_I$ are observed for the 12 mutants studied for at least one of the four inhibitors. Compared to wildtype $K_I$ value, the changes range from a 176-fold increase (I40A-MTX) to a 5-fold decrease (T125A-PYR). Minimal changes are observed for RLX, ranging from a 2-fold increase to a 3-fold decrease for V13Y and I86A, respectively as shown in $K_I$ data table 2 and figure 10.
Figure 8. A representative Dixon plot (V6P). The inverse rate (s/µM) plotted as a function of trimethoprim concentrations (0 nM, 0.05 nM, and 0.1 nM) at 50 µM (blue line), and 10 µM (red line) dihydrofolate. Rates were measured at 50 and 10 µM of DHF with three inhibitor concentrations. The inverse rate was then plotted as the function of inhibitor concentration. A linear regression line was fitted to each DHF concentration data set. The equation of the linear regression line was set equal to each other and solved for X which would give the $K_I$ value. The interaction point of the two lines is the $K_I$ value which calculated to be 0.133 nM (24, 25).
Table 2. $K_I$ values of $B_s$ DHFR wildtype and mutants for MTX, TMP, PYR, and RLX.

$K_I$ values were measured at two different DHF concentration and three different inhibitor concentrations. The ratio of $K_I$(mutant)/$K_I$(wildtype) is reported in parenthesis. The $K_I$ values are averages of at least 3 independent measurements.
Figure 9. Normalized $K_I$ values for MTX, TMP, PYR, and RLX for $B_s$ DHFR wildtype and mutants. To allow for a visual comparison of the ligand specificity profiles of all variants, the data was normalized. Each $K_I$ value was divided by the highest $K_I$ value measured for that inhibitor; the values used to normalize the data were MTX: 1.8 nM, TMP: 4.2 nM, PYR: 0.77 nM, and RLX: 7.2 nM. The normalized $K_I$ values are presented along with the standard deviation error bars to indicate the standard deviation. The $K_I$ values are averages of at least 3 independent measurements.
3.5. Effects of active site and second shell mutations

It is not surprising to observe active site mutants to alter inhibitor potencies as these mutants may have altered intermolecular interactions with inhibitors. For instance, large increase in $K_I$ value was observed for MTX (29-fold) and PYR (27-fold) for the active site mutant V6P. Interestingly, TMP inhibition was barely affected and RLX inhibition was unaffected by V6P. In addition, the M20 loop mutant V13Y resulted in an
increase in $K_I$ value for all inhibitors except TMP compared to wildtype. Loss of inhibition by MTX was observed for mutants V6P and V13Y by 20 to 30-fold compared to wildtype. As previously mentioned, a hydrogen bond is present between one of the amine groups of folate and the carbonyl oxygen of the backbone of V8 folate in Hs DHFR (1DRF). As a result, it is likely for this interaction to be altered by a proline replacement at V6 in Bs DHFR as the mutation may have altered the local structure of the backbone. Moreover, V13 could also influence the binding site environment as it is located on the M20 loop close to the active site (21). A number of DHFR ligands form intermolecular interactions with V8 in Hs DHFR as observed in crystal structures; which corresponds to residue V6 in Bs DHFR. For instance, a hydrogen bond (3.27 Å) is present between the 2-amino group of the folate and the carbonyl oxygen of the V8 backbone in Hs DHFR (1DRF). It is possible that this backbone to ligand interaction is influenced due to changes in the local structure by the V6P replacement in Bs DHFR. V6P replacement could change the local structure by possibly introducing a bend in the backbone.

Furthermore, mutant V6P could possibly affect interactions of its neighboring residues. Two hydrogen bonds are present in Hs DHFR (PDB code: 4M6J): one of them being between the amide H of nicotinamide ring and the carbonyl oxygen of the backbone of A9 (2.84 Å) and the other being between the carbonyl oxygen of nicotinamide ring and the amine hydrogen of the backbone of A9 (2.98 Å). Similarly, two hydrogen bonds are present in E. coli DHFR (PDB code: 1RX2): one of them being between the amide H of nicotinamide ring and the carbonyl oxygen of the backbone of A7 (3.03 Å) and the other being between the carbonyl oxygen of nicotinamide ring and the amine hydrogen of the backbone of A7 (2.80 Å). It is possible that active site mutant
V6P alters the neighboring interaction of A9 thorough an introduction of a bend in the backbone based on the interactions found in the two DHFR homologs. Additionally, the binding site environment could be influenced by the M20 loop residue V13 as the M20 loop is observed to play a vital role in catalysis (14, 21).

3.6. Effects of distal mutations on $K_I$

Interestingly, significant changes in $K_I$ were observed for distal mutants. It is interesting because it is unlikely for these distal residues to form any intermolecular interactions with the ligands. The largest changes in $K_I$ were observed for MTX, a 176-fold increase for I40A and 162-fold increase for L74R. In other words, the I40A and L74R mutation caused MTX to be less effective by a 176-fold and 162-fold, respectively, compared to wildtype. Allosteric mutants H38N and A39Y had comparably similar values as wildtype, whereas, mutants G51I (24-fold), W85L (99-fold), I86A (15-fold), T125A (67-fold), F126L (27-fold), and Y127A (17-fold) had significantly larger $K_I$ values for MIX. For distal mutant W85, a tryptophan residue is replaced by isoleucine. Two hydrogen bonds are present in $B_s$ DHFR (IZDR): one of them being between the carbonyl oxygen of the backbone of E81 and the nitrogen of the backbone of W85 (3.03 Å) and the other being between the carbonyl oxygen of the backbone of W85 and the nitrogen of the backbone of S88 (3.31 Å). The backbone interactions can be influenced by W85L mutation as there is a change in the size of the sidechain. To put it differently, the inhibitor potency is decreased by mutations at these allosteric positions (excluding cluster 1), implying their indirect importance in MTX inhibitor potency. Interestingly, in the $B_s$ DHFR structure, both W85 and H38 are located in allosteric regions. Both residues
have comparable $k_{cat}$ values to wildtype: a 4-fold decrease for W85L and no change was observed for H38N. Strangely enough, the DHFR reaction is catalyzed with similar efficiencies as wildtype by these allosteric variants, but significant changes are observed in their interactions with inhibitors, as was predicted previously (11).

The loss of inhibition by TMP is significant for all allosteric mutants except T125A. Unlike the other mutants, allosteric cluster 1 has a big effect on TMP inhibition. On the other hand, inhibition by PYR is increased for allosteric cluster 1 (except I40A) and T125A. Nonetheless, PYR inhibition decreased for mutants G51I (10-fold), L74R (4-fold), W85L (16-fold), I86A (13-fold), F126L (4-fold), and Y127A (5-fold). For PYR, largest changes in $K_I$ were observed for M20 loop mutant V13Y (43-fold) and allosteric cluster 2 mutant W85L (16-fold). The large change in $K_I$ for mutant V13Y is not surprising as it is located on the M20 loop that covers the DHFR active site. However, residue 186 is distal to the active site and is found in the same cluster as W85, discussed above. These allosteric residue observations are interesting because it is highly unlikely that the changes are simply due to altered short-range intermolecular interaction.

For distal mutants A39Y, G51I, W85L, I86A, and Y127A, minimal (1 to 3-fold) increase in RLX inhibition is observed compared to wildtype. Surprisingly, mutants A39Y, W85L, I86A, and Y127A are better inhibited by RLX compared to wildtype. To put it another way, mutations at these specific positions in the allosteric region increase the inhibitor potency. Allosteric mutants H38N, I40A, L74R, T125A, and F126L had similar RLX inhibitor potency as wildtype. For allosteric cluster 2 (W85L and I86A) the $K_I$ value decreased for RLX, whereas, the $K_I$ value increased for MTX, TMP and PYR. For allosteric cluster 3 mutant T125A, significant changes were observed for inhibitors.
MTX and PYR, but not for TMP and RLX compared to wildtype. Mutants F126L and Y127A, only minimally influenced inhibition for all inhibitors except MTX.

The magnitudes of the influence of mutations on $K_I$ are significant for MTX, TMP, and PYR compared to RLX. Compared to wildtype, allosteric mutants W85L, I86A, and Y127A are better inhibited by RLX. RLX is the weakest inhibitor, in other words, has the lowest potency compared to other drugs and has the least influence on $K_I$. From the results, it could be concluded that the influence of mutations are drug specific. For instance, the effect of mutants V6P and V13Y is greater on TMP compared to other inhibitors.

3.7. Effects of mutations are drug specific

A conclusion drawn from the $K_I$ analysis was that the mutation effects are drug specific. To put it another way, mutations only affected binding of certain drugs and not others. For example, active site mutant (V6P) and the M20 loop mutant (V13Y) have influence on TMP inhibition, yet, negligible changes in $K_I$ are observed for other drugs. Also, mutations at allosteric cluster 1 (H38N and A39Y) alter TMP but not MTX, PYR, and RLX inhibition. Interestingly, allosteric cluster 1 mutant (I40A) affects MTX inhibition and TMP inhibition. On the contrary, mutations at allosteric cluster 2 (L74R, W85L and I86A) result in changes in $K_I$ values for all of the four drugs when compared to wildtype. For mutant G51I, loss of inhibition was observed for MTX (24-fold) and PYR (10-fold) but inhibition was barely affected for TMP and RLX. It is noteworthy that the mutants alter inhibition potency differently for different drugs despite all drugs binding in the same region.
3.8. Ratios of $K_I$ values of different inhibitors

The ratio of $K_I$ values of different inhibitors was determined to observe changes in ligand specificity; the data is displayed in figure 11, using a log scale. Ligand specificity is the preference of one inhibitor over another by a DHFR homolog. Changes in ligand specificity were observed for mutants (Table 3). In other words, the preference of one drug over another changed for a mutant compared to wildtype. Between TMP and MTX, wildtype is better inhibited by MTX (positive ratio), whereas, allosteric mutant T125A is better inhibited by TMP (negative ratio) as observed in the $K_I(TMP)/K_I(MTX)$ ratio (Figure 11A). Furthermore, wildtype has selectivity for MTX over PYR as observed in the $K_I(MTX)/K_I(PYR)$ ratio (figure 11D). On the other hand, allosteric cluster 1 mutants H38N, A39Y, and I40A have selectivity for PYR over MTX. Similarly, allosteric 3 mutants T125A, F126L, and Y127A have selectivity for PYR over MTX. The changes in ligand specificity profile for these allosteric mutants support the previous predictions.
<table>
<thead>
<tr>
<th></th>
<th>$K_{1($TMP$)/K_{1(MTX)}$</th>
<th>$K_{1($TMP$)/K_{1(PYR)}$</th>
<th>$K_{1($RLX$)/K_{1($TMP$)$</th>
<th>$K_{1($MTX$)/K_{1(PYR)}$</th>
<th>$K_{1($RLX$)/K_{1($MTX$)$</th>
<th>$K_{1($RLX$)/K_{1(PYR)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>18 (1.0)</td>
<td>10 (1.0)</td>
<td>16 (1.0)</td>
<td>0.58 (1.0)</td>
<td>296 (1.0)</td>
<td>171 (1.0)</td>
</tr>
<tr>
<td>V6P</td>
<td>0.49 (0.027)</td>
<td>0.31 (0.030)</td>
<td>21 (1.3)</td>
<td>0.63 (1.1)</td>
<td>10 (0.034)</td>
<td>6.4 (0.037)</td>
</tr>
<tr>
<td>V13Y</td>
<td>0.87 (0.048)</td>
<td>0.23 (0.022)</td>
<td>41 (2.5)</td>
<td>0.26 (0.46)</td>
<td>36 (0.12)</td>
<td>9.4 (0.055)</td>
</tr>
<tr>
<td>H38N</td>
<td>328 (18)</td>
<td>367 (35)</td>
<td>0.98 (0.060)</td>
<td>1.1 (1.9)</td>
<td>321 (1.1)</td>
<td>359 (2.1)</td>
</tr>
<tr>
<td>A39Y</td>
<td>91 (5.0)</td>
<td>380 (37)</td>
<td>0.86 (0.052)</td>
<td>4.2 (7.3)</td>
<td>77 (0.26)</td>
<td>325 (1.9)</td>
</tr>
<tr>
<td>I40A</td>
<td>1.1 (0.062)</td>
<td>32 (3.1)</td>
<td>3.3 (0.20)</td>
<td>29 (50)</td>
<td>3.7 (0.013)</td>
<td>107 (0.63)</td>
</tr>
<tr>
<td>G51I</td>
<td>1.4 (0.075)</td>
<td>1.9 (0.19)</td>
<td>7.7 (0.47)</td>
<td>1.4 (2.5)</td>
<td>11 (0.035)</td>
<td>15 (0.087)</td>
</tr>
<tr>
<td>L74R</td>
<td>1.1 (0.061)</td>
<td>28 (2.7)</td>
<td>1.7 (0.11)</td>
<td>25 (44)</td>
<td>1.9 (0.0064)</td>
<td>48 (0.28)</td>
</tr>
<tr>
<td>W85L</td>
<td>0.62 (0.034)</td>
<td>2.2 (0.21)</td>
<td>1.7 (0.11)</td>
<td>3.6 (6.2)</td>
<td>1.1 (0.0036)</td>
<td>3.9 (0.023)</td>
</tr>
<tr>
<td>I86A</td>
<td>4.3 (0.24)</td>
<td>2.9 (0.28)</td>
<td>1.5 (0.091)</td>
<td>0.68 (1.2)</td>
<td>6.4 (0.022)</td>
<td>44 (0.026)</td>
</tr>
<tr>
<td>T125A</td>
<td>0.17 (0.0092)</td>
<td>30 (2.8)</td>
<td>53 (3.2)</td>
<td>177 (308)</td>
<td>8.9 (0.030)</td>
<td>1570 (9.2)</td>
</tr>
<tr>
<td>F126L</td>
<td>5.5 (0.31)</td>
<td>22 (1.9)</td>
<td>4.5 (0.28)</td>
<td>3.6 (6.3)</td>
<td>25 (0.085)</td>
<td>92 (0.54)</td>
</tr>
<tr>
<td>Y127A</td>
<td>5.8 (0.32)</td>
<td>12 (1.1)</td>
<td>1.1 (0.067)</td>
<td>2.0 (3.5)</td>
<td>6.4 (0.021)</td>
<td>13 (0.076)</td>
</tr>
</tbody>
</table>

Table 3. The ratio of $K_1$ values of different inhibitors is reported and the fold change is reported in parenthesis.
Figure 11. The ratios of $K_i$ values of different inhibitors are shown. Panel A shows the ratio of $K_i(\text{TMP})$ over $K_i(\text{MTX})$ values. Panel B shows the ratio of $K_i(\text{TMP})$ over $K_i(\text{PYR})$ values. Panel C shows the ratio of $K_i(\text{RLX})$ over $K_i(\text{MTX})$ values. Panel D shows the ratio of $K_i(\text{MTX})$ over $K_i(\text{PYR})$ values. Panel E shows the ratio of $K_i(\text{RLX})$ over $K_i(\text{MTX})$ values. Panel F shows the ratio of $K_i(\text{RLX})$ over $K_i(\text{PYR})$ values.
4. Summary

In summary, I experimentally tested 12 Bs DHFR variants predicted to serve as specificity determinants in the DHFR family by performing inhibition analysis on wildtype and mutants against four drugs, which were methotrexate (MTX), trimethoprim (TMP), pyrimethamine (PYR), and raltitrexed (RLX). I measured turnover numbers ($k_{cat}$) and found that all of the variants were active and thus were appropriate models for testing the previously made predictions. In addition, Michaelis constants were measured and it was found that of the 12 mutants studied, the $K_M$ increased at the most by 5-fold and thus, active site binding properties are relatively similar to wildtype. From the $K_I$ analysis, it was concluded that significant effects on inhibitor specificity profiles by distal mutations were observed, yet, they had relatively small effects on catalytic activity. These findings support the previous predictions and the approach for the DHFR enzyme family for determining ligand specificity determinants. From the inhibition analysis, it was found that the effects of mutations are ligand specific even though the ligands are all competitive and bind at the active site.
References


