To the Graduate Council:

I am submitting herewith a thesis written by Robert F. Beahm entitled "Using Molecular Dynamics Simulations to Study the Dynamic and Catalytic Properties of R67 Dihydrofolate Reductase" I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science With a major in life science.

Hong Guo
Major Professor

We have read this thesis
and recommend it's acceptance:

Elizabeth E. Howell

Robert J. Hinde

Accepted for the Council

Carolin R. Hodges
Vice Provost and Dean
of the Graduate School

(Original signatures are on file with student records)

Using Molecular Dynamics Simulations to Study the Dynamic and Catalytic Properties of R67 Dihydrofolate Reductase

A Thesis Presented for the Master of Science Degree
The University of Tennessee, Knoxville

Robert Beahm
August 2007

Abstract

Dihydrofolate reductase (DHFR) catalyzes the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF) using NADPH as a cofactor. Since THF is an essential factor for nucleotide biosynthesis, inhibition of this enzyme in bacteria with folate analogs such as trimethoprim results in bacterial cell death. Plasmid encoded R67 DHFR confers resistance to trimethoprim and is both sequentially and structurally unrelated to any known chromosomal version of the enzyme. R67 DHFR is a 34,000 Da. homo-tetramer containing a rare 222 axis of symmetry in the center of its active site pore. The active site pore is contacted by residues belonging to each of the four subunits. The enzyme can nonspecifically bind 2 NADPH's, 2 DHF's, or, in the case of the productive complex, 1 NADPH and 1 DHF. R67 DHFR employs the endo transition state as opposed to the exo transition state used by *E. coli* DHFR. In this study molecular dynamics approach using the CHARMM program is employed to study the dynamics of the enzyme and energetics of the hydride transfer step catalyzed by R67 DHFR in silico. Structural and dynamic properties of four different mutants are also examined.

Table of Contents

Chapter Page

I. **Introduction**……1
Bacterial Drug Resistance……..…….1
R67 Dihydrofolate Reductase……...2
E. coli Dihydrofolate Reductase……..……..……10
Evolutionary Considerations……...13

II. **Experimental Methods**……..14
Computer Modeling……..….....14
Force Field (30 Angstrom Buffer Region)…….....16
QM/MM Molecular Dynamics Models…….....18
Energy Corrections……..….....20

III. **Results**…….....25
30 Angstrom Molecular Dynamics Results…….....25
Potential Mean Force Calculations……...….28
Comparison of Mutant Forms to Wild-type R67 DHFR…….40

IV. **Conclusions**…….53

List of References……..59

Vita…….63

List of Tables

Table Page
1. Wild-type and 4 Mutants Sequence…….……24
2. *E. coli* DHFR PMF Summary……...34
3. R67 DHFR PMF Summary…….....38
4. Dynamics Summary Data…….42

List of Figures

Figure Page

1. Hydride Transfer Step of DHFR……....2
2. R67 DHFR…….….4
3. Active Site Residues of R67 DHFR…….…....5
4. *E. coli* Chromosomal DHFR…….11
5. XN Angle……...……...18
6. QM Region of the Simulation……..….19
7. R67 and *E. coli* DHFR Model Systems……...….21
8. Linear Correction Curves for SCCDFTB Method……23
9. Movement of pABA-Glu Tail of DHF…….……....26
10. Free Energy Profile of XN Angle…….……29
11. DHF Dihedral Angle…….30
12. Pter-Nic Dihedral Angle……...……31
13. *E. coli* DHFR Free Energy Curve……...……..33
14. R67 DHFR Free Energy Curve……...……..36
15. Scatter Plots of the Hydride Transfer Distance and Angle…….……..39
16. Effects of Tail Constraints on the Energy Barrier……....41
17. Average Structure of WT R67 DHFR……..44
18. Average Structure of Y69F R67 DHFR…..……..46
19. Average Structure of I68L R67 DHFR…….……48
20. Average Structure of Q67C R67 DHFR…….……..49
21. Average Structure of M10 R67 DHFR…….…50
22. M10 vs Wild-type Free Energy Curve……...……...52

Chapter I
Introduction

Bacterial Drug Resistance

Bacterial resistance to antimicrobial agents has been an emerging problem in recent years. The use of such drugs applies selective pressure to the bacterial population causing mutations beneficial to the microorganism to occur in the protein targets of antibiotics. This situation is complicated by the fact that genes which confer resistance to antibiotics can be spread throughout a microbial population through the use of horizontal gene transfer. The three types of horizontal gene transfer are transformation, transduction and conjugation. Conjugation involves the use of plasmids or transposons to transfer genetic information from one microbe to another.

In 1989 Stokes and Hall discovered a new type of mobile genetic element known as integrons. Integrons contain the integrase gene followed by a promoter region and a recombination site (Stokes and Hall, 1989). Gene cassettes can be inserted at the recombination site by the use of the integrase enzyme. Gene cassettes contain only a promoter-less open reading frame and a recombination site. Several gene cassettes conferring resistance to several different antibiotics can be inserted in a single integron. The integrons depend on the machinery of the plasmid or chromosome in which they are located in order to replicate (Levings, *et al* 2006). By this method, novel genes can be quickly utilized by a bacterial population when it is exposed to antibiotics.

One such gene discovered in 1972 encodes a unique form of the enzyme Dihydrofolate reductase. Dihydrofolate reductase (DHFR) catalyzes the conversion of

Figure 1.

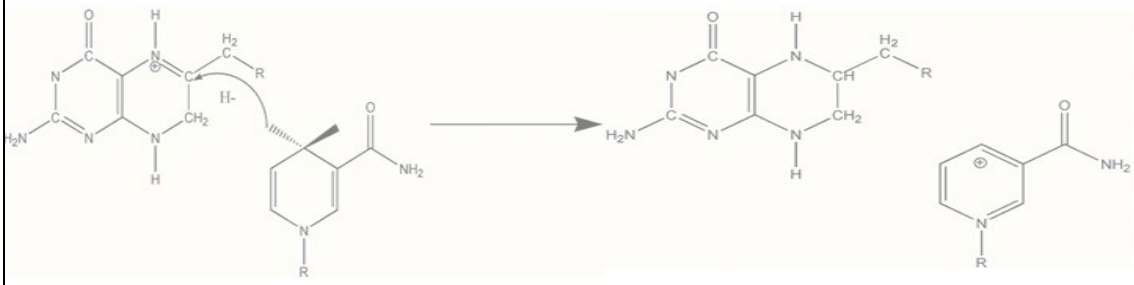


Figure 1: Hydride transfer step of DHFR

Dihydrofolate (DHF) to tetrahydrofolate (THF) using NADPH as a cofactor (Alonzo and Gready, 2006). There are two major parts to this reaction: The protonation of the N5 nitrogen on the pteridine ring and the hydride transfer from the NADPH cofactor to the pteridine ring of DHF (Figure 1). Bacterial DHFR has been the target of antibiotic drugs such as trimethoprim which block the activity of the enzyme by occupying the substrate binding site. Overuse of these drugs has led to an emergence of a type II (dhfrb) plasmid encoded form of DHFR which is resistant to trimethoprim (Howell, 2005). Recent evidence indicates that R67 DHFR, an example of a type II DHFR, is coded for in a gene cassette associated with a class 1 integron found in the R67 plasmid.

R67 Dihydrofolate Reductase

Homotetrameric R67 DHFR bears little resemblance to the monomeric

chromosomal form of the enzyme. The most striking feature is the large solvent accessible active site pore in the center of the four monomers (Figure 2). This brings about a unique 222 axis of symmetry in the center of the enzyme which allows for binding of the substrate and or cofactor in any one of four possible positions. The enzyme can bind either 2 NADPH's , 2 DHF's, or the productive ternary complex consisting of 1 NADPH and 1 DHF. The substrate and cofactor are oriented in such a way that can lead to an endo-transition state as opposed to the exo-transition state used by chromosomal DHFR (Howell, 2005).

This nonspecific binding in R67 DHFR is partially facilitated by hydrogen bonding interactions between the backbone of Ile-68 and complementary groups on the pteridine ring of DHF and the nicotinamide ring of NADPH. The four Gln-67 residues form the inner ring of the pore with two of them interacting with each other above the nicotinamide ring of NADPH and the other two interacting below the pteridine ring of DHF. This presumably acts to force the substrate and cofactor closer together and thus facilitate catalysis. Tyr-69 residues help to position Gln-67 and may also form hydrogen bonding interactions with the NADPH cofactor (Figure 3). Finally Lys-32 near the end of the pore may interact with the side chain carboxyl group at the end of DHF's pABA glu tail. Another Lys-32 on the opposite side of the pore interacts with the 2' pyrophosphate group of NADPH (Strader *et al*, 2001).

The orientation between the nicotinamide and ribose rings of NADPH also appear to be a unique feature of the R67 DHFR enzyme. Most reductases, including chromosomal DHFR, have the NADPH cofactor in the anti conformation with the

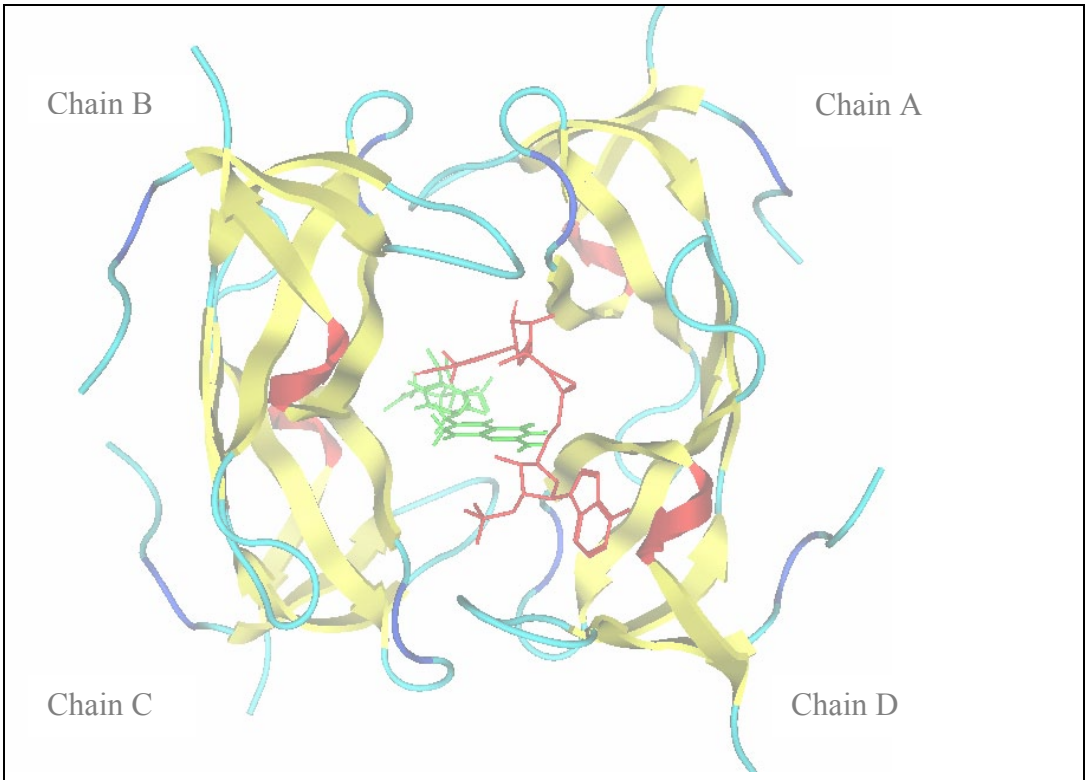
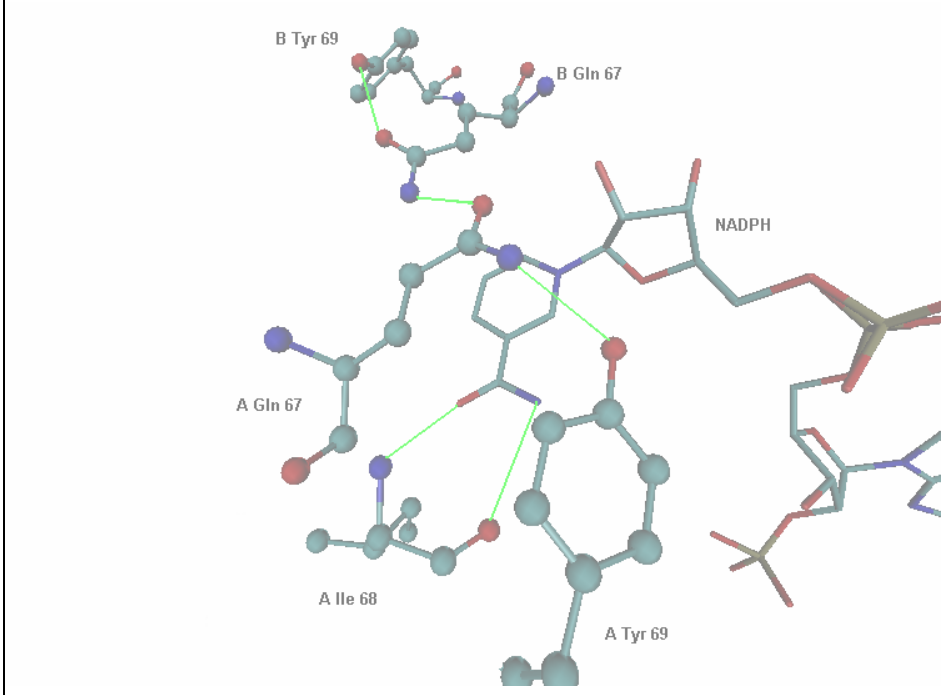


Figure 2: R67 DHFR showing bound DHF(green) and NADPH(red). For the protein structure the β-sheets are represented in yellow, the α-helices are represented in red and the loops are represented in blue. All four chains are labeled A-D. PDB file provided by Dr. Elizabeth Howell.

A.



B.

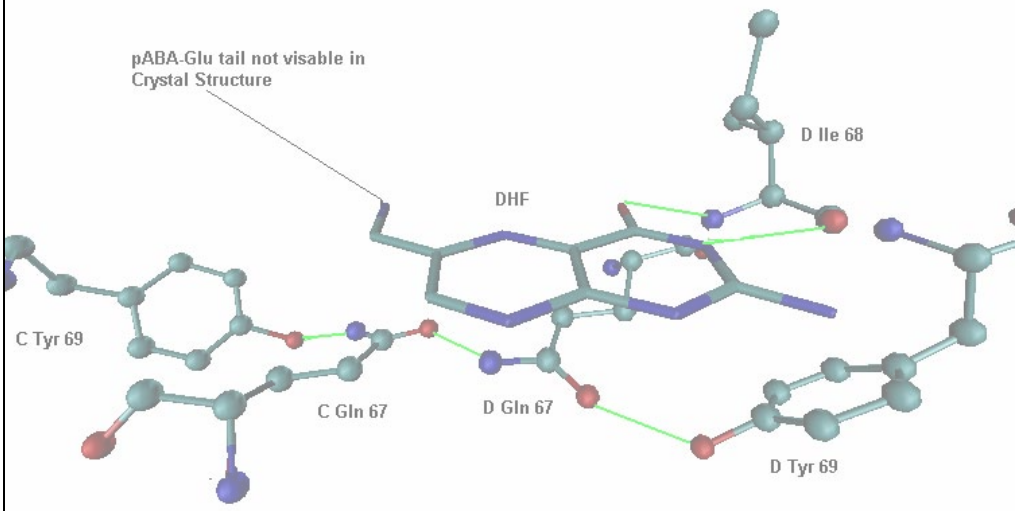


Figure 3. Active site residues of R67 DHFR showing the NADPH side of the pore (A) and the DHF side of the pore (B). Carbon atoms are represented in gray, oxygen in red, nitrogen in blue and hydrogen in white. Hydrogen bonds are represented by green lines. Note that the substrate and the cofactor are both H-bonded to the backbone of an Ile 68 residue and flanked by a Tyr-Gln-Gln-Tyr network.

binding of the substrate and or cofactor in any one of four possible positions. The enzyme can bind either 2 NADPH's , 2 DHF's, or the productive ternary complex consisting of 1 NADPH and 1 DHF. The substrate and cofactor are oriented in such a way that can lead to an endo-transition state as opposed to the exo-transition state used by chromosomal DHFR (Howell, 2005).

This nonspecific binding in R67 DHFR is partially facilitated by hydrogen bonding interactions between the backbone of Ile-68 and complementary groups on the pteridine ring of DHF and the nicotinamide ring of NADPH. The four Gln-67 residues form the inner ring of the pore with two of them interacting with each other above the nicotinamide ring of NADPH and the other two interacting below the pteridine ring of DHF. This presumably acts to force the substrate and cofactor closer together and thus facilitate catalysis. Tyr-69 residues help to position Gln-67 and may also form hydrogen bonding interactions with the NADPH cofactor (Figure 3). Finally Lys-32 near the end of the pore may interact with the side chain carboxyl group at the end of DHF's pABA glu tail. Another Lys-32 on the opposite side of the pore interacts with the 2' pyrophosphate group of NADPH (Strader *et al*, 2001).

The orientation between the nicotinamide and ribose rings of NADPH also appear to be a unique feature of the R67 DHFR enzyme. Most reductases, including chromosomal DHFR, have the NADPH cofactor in the anti conformation with the ON4—CN1'—NN1—CN2 dihedral in the range from 120 thru 180 degrees to –60 degrees (see figure 5 in methods) (Brito *et al*, 1991). R67 DHFR has this dihedral at around 55 degrees in the crystal structure which puts it in the syn conformation.

The apo enzyme possesses 222 symmetry, but fluorescence and ITC experiments demonstrate that binding of a single ligand (or cofactor) produces asymmetry in the total structure which can lead to various cooperative effects on subsequent binding. Binding of DHF induces positive cooperativity in the binding of a second DHF molecule. NADPH causes negative cooperativity in the binding of a second NADPH molecule but positive cooperativity in the binding of a DHF molecule to the other half of the pore. The differences are most likely to be caused by interactions between the two ligands in the center of the pore. NADPH, which binds first, contains a bulky head group that can occupy the center of the pore and potentially block a second NADPH molecule from binding. A DHF molecule fits better in the second half of the pore leading to the productive ternary complex (Bradrick *et al*, 1996).

Ligand binding studies conducted on alternative substrates reveal the importance of several functional groups. Methotrexate, which has an amine group instead of the O4 oxygen in DHF, binds significantly more weakly. Removal of the pABA-Glu tail as in dihydrobiopterin eliminates the ITC signal altogether indicating the importance of this group in binding. Alterations in buffer conditions which change the pKa of the N3 nitrogen such that it is deprotonated increase the strength of binding as well. This is consistent with crystallography and docking studies that show a hydrogen bonding interaction between N3 and O4 of the pteridine ring and the backbone C=O and NH groups of Ile-68 (Jackson *et al*, 2005).

In order to determine what factors are involved in R67 DHFR catalysis, Strader and his associates conducted ITC experiments on several mutants. These experiments reveal that there is a linear correlation between the enthalpy change of folate binding and

log (k_{cat}/K_m): The greater the enthalpy change, the better the catalysis. This indicates that formation of the ternary complex is at least in part enthalpy driven and that binding interactions in the active site play an important role in facilitating the hydride transfer (Strader *et al*, 2004). The overlap of the pteridine and nicotinamide rings, which has been calculated to be more stable in the endo versus the exo-transition state, may also be a factor in addition to interactions between the substrate and enzyme (Andres *et al*, 1996).

The role of water in substrate binding to R67 DHFR has also been studied by examining the effects of various osmolyte concentrations on binding of the substrate and cofactor. The effects were then used to quantify the number of water molecules forced in or out of the active site upon binding. Calculations indicate that 48 ± 12 water molecules are taken up upon DHF binding which could serve to both increase the entropic factor and decrease the desolvation penalty (Chopra, personal communication, 2007).

Mutation analysis supports the assertion that residues 66-69 have an important role in both binding and catalysis. The results of mutations in the R67 DHFR active site are generally dramatic as a single mutation affects all four subunits. The Q67C mutation results in a functional enzyme that has a lower k_{cat} and a much higher K_m then the wild-type for both the substrate and cofactor. Mutation of Ile-68 to leucine also decreases k_{cat} and increases K_m although not as much for the DHF substrate. The Y69F mutation actually increases the k_{cat} although the K_m for both the substrate and cofactor goes up considerably (Strader *et al*, 2001).

Mutations at Lys-32 (on all four chains) disrupt the formation of the homo-tetramer. The role of Lys-32 has been studied by the generation of asymmetric mutations on a "Quad" protein which has all four subunits linked together in a single chain. While a

K32M mutation at a single site does not significantly affect catalysis, mutations at two sites do. Mutations on opposite sides of the active site pore (K32M:1+2 or K32M:1+4) decrease k_{cat}. When the two K32M mutations (K32M:1+3) are placed on the same side of the pore, k_{cat} is increased while binding is weaker. Comparison of the k_{cat}/K_m values of the double mutants suggest that an ionic interaction formed during binding must be broken in order to form the transition state. Since NADPH binds first, it will most likely bind to the non-mutated side of the K32M:1+3 pore with DHF binding to the other side. Because of this, it is likely that the broken ionic interaction involves the pABA-Glu tail of DHF (Hicks *et al* , 2004).

Although each of these point mutations leads to a decrease in activity, combinatorial mutagenesis studies conducted by Schmitzer and her associates showed that functional mutants can be created by making multiple mutations in residues 66-69 in the active site pore. Three of the mutants she created actually have a higher activity (as measured by k_{cat}/K_m) than the wild-type. Surprisingly there seemed to be no consistency in the characteristics of the residues at each position in the functional mutants. Neither size nor hydrophobicity seems to be a factor at position 66 (WT = Val). Position 67 (WT = Gln) can have a positive, negative or no charge. The characteristics of position 68 (WT = Ile) seem to be more important in that long aliphatic chains seem to be favored. This is most likely due to the need to properly position the backbone of this residue for hydrogen bonding with the substrate and cofactor. With a few exceptions, aromatic side chains seem to be favored at position 69 (WT = Tyr) (Schmitzer *et al* , 2004).

The biggest mystery in the catalytic mechanism is when and how the N5 nitrogen in the pteridine ring of DHF becomes protonated. For *E. coli* chromosomal DHFR there

are two lines of thought. Some simulation data indicates that the proton comes directly from water (Thorpe and Brooks, 2003). In other simulations the proton is thought to come from an Asp residue and is shuttled across several water molecules to the N5 atom (Ferrer *et al*, 2003). R67 DHFR, however, has no obvious proton donor in the active site. The pKa of N5 in solution is 2.59 and studies show that the pKa is not altered much by the R67 active site environment unlike the chromosomal form of the enzyme (Deng *et al*, 2001). Also the activity of the R67 enzyme increases with a decrease in pH indicating that N5 protonated DHF is the substrate for the reaction (Park *et al*, 1997).

Crystallography studies have been challenging due to the rapid motion of DHF's pABA-Glu tail (Narayana *et al*, 1995). Docking studies show many possible positions for the pABA-Glu portion most of which include ionic interactions between the negatively charged groups of the DHF pABA-Glu tail and Lys-32 residues on two adjacent chains (Alonzo *et al*, 2005; Howell *et al*, 2001). What effect, if any, the motion of the tail has on the hydride transfer reaction is not yet known. However, salt effects studies indicate the loss of a salt bridge during the formation of the transition state. This could occur either at the pABA-Glu tail of DHF or at one of the phosphate groups of NADPH which also forms interactions with the Lys-32 residues on that side of the pore (Hicks *et al*, 2003). A Lys-32 forms an ionic interaction with the 2' phosphate group of NADPH and the adjacent Lys-32 forms an ionic interaction with the pyrophosphate group of NADPH.

E. coli Dihydrofolate Reductase

E. coli DHFR has a single subunit and has more specific interactions between the enzyme and the substrate and cofactor (Figure 4). Numerous specific interactions exist

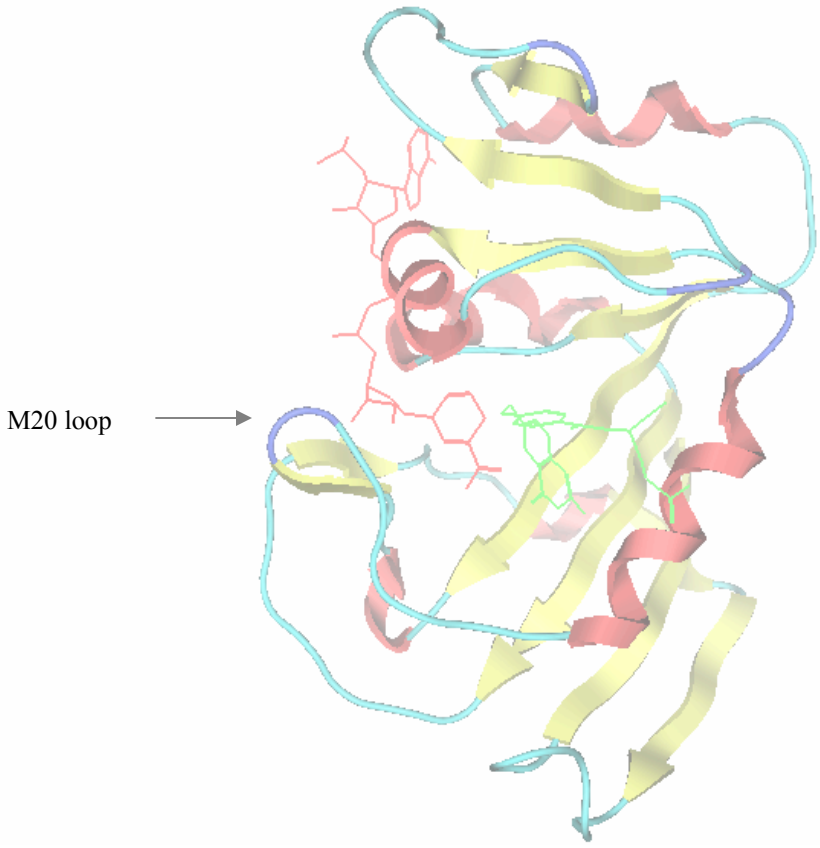


Figure 4: *E. coli* DHFR showing bound DHF(green) and NADPH(red). For the protein structure the β-sheets are represented in yellow, the α-helices are represented in red and the loops are represented in blue. The M20 loop region is indicated by the arrow. PDB Id: 1RX2.

between the substrate, cofactor, and enzyme. Arg-52, Arg-57 and Lys-32 interact with the pABA-Glu tail of DHF. Asp-27 forms important interactions with the N3 and N2 nitrogens of the pteridine ring of DHF. His-44 and Thr-45 interact with the pyrophosphate groups of NADPH. The chromosomal form of the enzyme is generally more capable of accommodating mutations. If mutations are made in the DHF binding pocket they affect the binding of DHF to a significant degree but affect NADPH binding to a much lesser degree. The reverse is also true, suggesting that active site interactions are local. Mutations at Asp-27 can profoundly affect the rate of the hydride transfer step. This residue has also been implicated in a possible mechanism for the protonation of the N5 nitrogen.

A significant number of computer simulations have been conducted on *E. coli* DHFR (Agarwal *et al,* 2002, 2005; Garcia-Vilocia *et al,* 2003, 2004; Cummins *et al* , 1998). These investigations, along with evidence from crystallographic and NMR studies suggest that several conformational changes of the enzyme are important for binding and catalysis. Residues 9-24 form what is called the M-20 loop, shown in figure 4, which can exist in opened, closed and occluded conformations. The closed state is stabilized by hydrogen bonding interactions between Glu-17 and Asp-122. This conformation is thought to bring the substrate and cofactor into the proper position for the reaction to occur. When NADPH is not bound, the enzyme exists in the occluded conformation during which the active site is partially blocked. An intermediate conformation consisting of an unstable opened loop structure may facilitate product release of THF. Evidence for this includes crystallographic data which shows the oscillation frequency of the M-20 helix to be the same as the rate of product release (Sawaya and Kraut, 1997).

Molecular dynamics studies on the hydride transfer step of *E. coli* dihydrofolate reductase indicate that brief puckering of the nicotinamide ring is correlated with achievement of the transition state (Agarwal *et al*, 2002). Alteration of the environment in such a way that changes the puckering event also changes the magnitude of the energy barrier. The pteridine ring also puckers to a greater degree such that the C6 carbon of DHF becomes closer to the CN4 carbon of the nicotinamide ring of NADPH at the transition state. Primary isotope effect experiments indicate that hydrogen tunneling takes place during the hydrogen transfer process (Wang *et al*, 2006). Previous computational studies also predict this effect (Agarwal *et al*, 2002; Garcia-Vilocia *et al,* 2003).

Evolutionary Considerations

E. coli chromosomal DHFR is a monomeric enzyme which employs the exo-type transition state as opposed to the endo-type transition state employed by R67. As a result there is less overlap between the nicotinamide ring of NADPH and the pteridine ring of DHF and the orientation is slightly less stable (Castillo *et al*, 1999). This could be due to the evolutionary optimization of the N5 protonation step in addition to the hydride transfer step in the case of E. coli DHFR. When the suspected *E. coli* DHFR proton donor Asp-27 is mutated to Serine, k_{cat} decreases 70 fold and the K_m increases 50 fold. (Cannon *et al* , 1997). Various attempts to add a proton donor in the R67 DHFR active site have failed to increase the catalytic efficiency (Park *et al*, 1997).

Chapter II
Experimental Methods

Computer Modeling

Molecular Mechanics is a computational tool used to study the motion of proteins and other large molecules over time. The potential energy of the system is determined by the sum of bonded and non-bonded forces. Bonded forces consist of functions that describe the change in potential energy with respect to changes in bond length, bond angle, dihedral angle and improper dihedral. For example the function describing bond length is $E_{bond\text{-}stretch} = \Sigma K_b (b - b_0)^2$ where E is the energy, K_b is the force constant, b_0 is the ideal bond length and b is the current bond length. The equation describing the change in angle is $E_{angle} = \Sigma K_\theta (\theta - \theta_0)^2$ where K_θ is the force constant, θ is the angle and θ_0 is the ideal angle. For the dihedral angle the equation is $E_{dihedral} = \Sigma K_\chi (1 + \cos(n\chi - \delta)$ where n is the multiplicity, χ is the angle and δ is the phase shift. For the improper dihedral the equation is $E_{improper} = \Sigma K_\Phi (\Phi - \Phi_0)^2$ where K_Φ is the force constant, Φ is the improper dihedral and Φ_0 is the equilibrium improper dihedral (Brooks et al., 1983).

The non-bonded forces consist of electrostatic forces and Van der Waals forces. The electrostatic forces are determined by the equation $E_{electrostatic} = \Sigma q_1 q_2 / D r_{12}$ where q_1 and q_2 are the charges of atoms 1 and 2, D is the dielectric constant, and r_{12} is the distance between atoms 1 and 2. The Van der Waals forces are modeled after the Lennard-Jones potential and are determined by the equation $E_{van\text{-}der\text{-}Waals} = \Sigma ((A_{ik} / r_{ik}^{12}) - (C_{ik} / r_{ik}^{6}))$ where r_{ik} is the radius between atoms i and k and A and C are both atom type dependent constants. A third type non-bonded factor called the Urey-Bradley

term describes the distance between atoms separated by two bonds with the equation $E_{uray\text{-}bradley} = \Sigma K_\mu (\mu - \mu_0)^2$ where K_μ is the force constant, μ_0 is the ideal distance and μ is the actual distance. The force constants for the bonded forces and the atom dependent forces are both derived from empirical data and higher level calculations (Brooks et al., 1983). They are sufficient for representing most types of molecular motion but cannot be used when bonds are made or broken. This requires the use of molecular orbital theory based on quantum mechanics.

A less time consuming form of the ab initio calculation is the semiempirical method. Semiempirical methods use empirical data to substitute for the more time consuming terms of the molecular orbital theory such as the two electron integrals. Because of this the semiempirical methods are generally less accurate than ab initio calculations but they can be as much as a thousand times faster. In a simulation where the potential energy must be determined up to millions of times, the semiempirical approach is the more ideal type of method to use. The semiempirical method used in this study is known as the self-consistent-charge density-functional tight-binding or SCCDFTB method which is based on density functional theory (Cui *et al*., 2001).

In order to set up a typical QM/MM simulation of an enzyme, the molecule must be divided up into quantum mechanics and molecular mechanics regions. The QM region generally includes the substrate(s) and any residues of the protein involved in the catalytic mechanism. The potential energies of these atoms are determined by semi-empirical calculations. The rest of the protein is described by molecular mechanics. If a QM atom is bonded to a MM atom, a generalized hybrid orbital (GHO) or link atom must be used to facilitate proper energy determination on both sides of the bond.

Due to the limitations of semi-empirical methods in energy evaluation, corrections can be added based on fixed point ab initio calculations of model systems. These corrections can either be added to the resulting free energy profile or implemented as additional force potentials during the simulation itself. The validity of the results must be based on correlation with all available experimental data regarding the kinetics of the reaction. This presents an interesting challenge in optimizing each model system to accurately describe a specific reaction.

In order to examine both the dynamic and catalytic properties of R67 DHFR, two different models were built. The first model has a 30 Å buffer region and contains only MM atoms. This model is used to examine dynamic properties of the complex that may be further from the center of the pore (such as the XN angle and the pABA-Glu tail). A second model was built with a QM region and an MM region (QM//MM) which had a smaller buffer region of 20 Å. This model was used to examine the catalytic properties of the enzyme as well as the dynamic properties at the active site.

Force Field (30 angstrom buffer region)

The dynamics simulation was set up using the initial crystal structure of R67 DHFR complexed with folate and NADP+ (provided by E. E. Howell, 2005). The pABA-glu tail was added to the substrate and the folate was converted to dihydrofolate and the cofactor was converted to NADPH using the Molecular Operating Environment (MOE) program (MOE Chemical Computing Group Inc.) . Because the position of the pABA-glu tail is unknown, minimization of the tail region was performed in the MOE environment using the CHARMM22 force field prior to dynamics (MacKerell *et al*, 1998). From this

point forward, all manipulations and calculations were made using the CHARMM program and based on the CHARMM33 force field (Brooks *et al,* 1983). Molecular Mechanics parameters for the DHF molecule were taken from Garcia-Vilocia's 2003 study of *E. coli* DHFR. The CHARMM force field potential described above was used for the rest of the system. A 30 Å solvation shell was added to the model using the TIP3 model in CHARMM.The stochastic boundary method was used for all simulations (Brooks *et al*, 1985). A buffer region of 28 Å from the C6 atom of DHF was constructed for the standard molecular mechanics region. The area between 28 and 30 angstroms was treated with Langevin dynamics. The system was minimized using an adapted basis Newton-Raphson (ABNR) method until convergence and then heated to 310 degrees Kelvin over 70 ps followed by an additional 40 ps equilibration. The dynamics simulation was run for 2 nanoseconds using femptosecond timesteps. The trajectory file was updated every 0.05 picoseconds.

The approximate conformation of the pABA -Glu tail of DHF was monitored using the distances between Lysine 32 of chain B or C and the carboxyl group associated with the CD or CT carbons in the DHF glutamate region. The average structures from several of the 40 ps windows were selected as a means of comparison. The XN angle, previously defined by Almarsson and Bruice (1993), was also monitored (Figure 5). This dihedral refers to the orientation between the nicotinamide and ribose rings of the NADPH cofactor (Almarasson and Bruice, 1993). Frequency data from the trajectory files were used to construct a free energy plot of this angle using the equation: $\Delta G = -RTLn(n/N)$ where n is the frequency of a specific conformation and N is the total sample size. A Potential of Mean Force (PMF) calculation on the XN angle was also performed

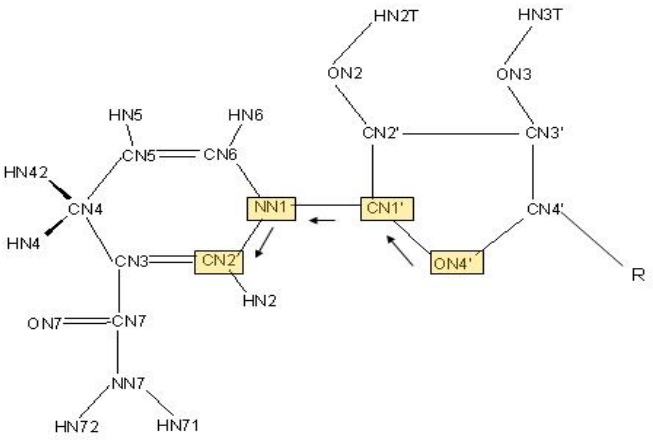


Figure 5: XN angle. The angle formed by the ON4',CN1', NN1, and CN2 atoms of NADPH. R represents the remainder of the NADPH molecule.

in order to confirm the results, however the cofactor moved out of the active site as the dihedral angle was forced further from the energy minimum by the external constraints that were applied.

QM/MM Molecular Dynamics Models

The QM/MM models were constructed using the same initial setup described above. However, the buffer region was cut to 22 Å with the region between 20 and 22 Å described by Langevin dynamics. The QM region consisted of 65 atoms including two QM/MM link atoms which form the boundary between the QM and MM regions (Field *et al*, 1990) (Figure 6). The link atoms were placed between CN4' and CN5' of NADPH and between C11 and C of DHF. Assignment of the link atoms in these locations were based on several requirements. The first requirement is that link atoms must be placed next to carbon atoms in the QM region. The second requirement is that the QM region is big enough to effectively describe the quantum physics of the reaction but small enough

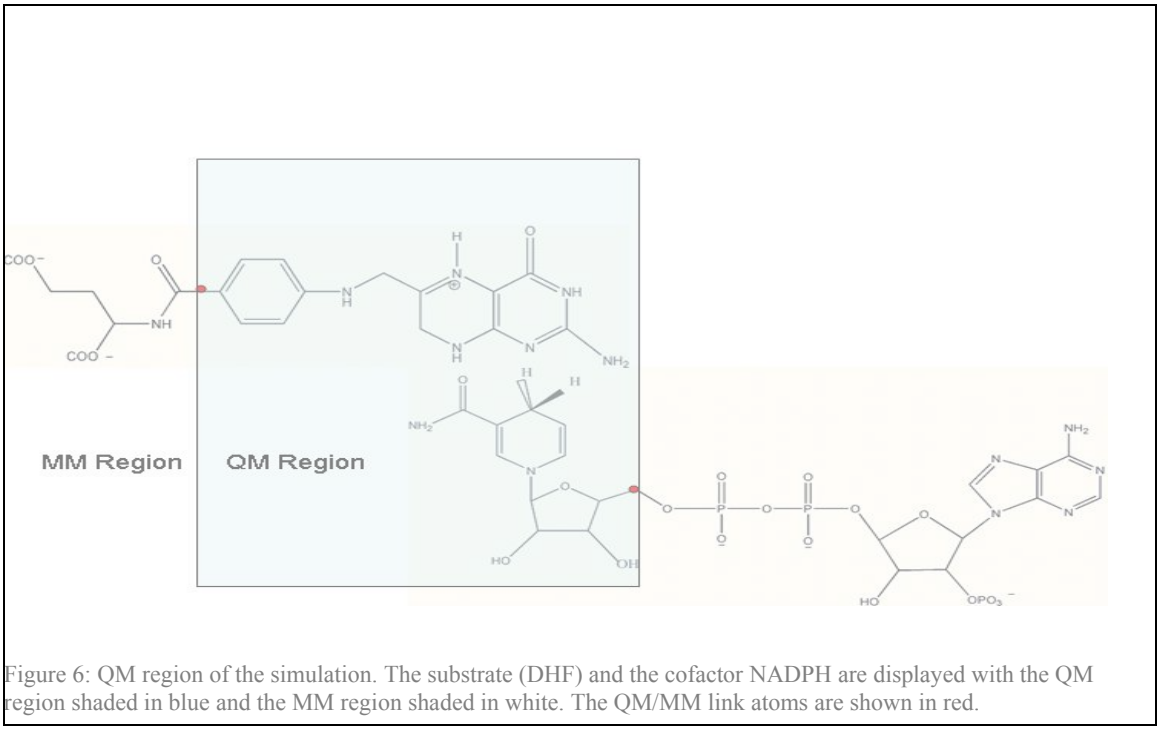


Figure 6: QM region of the simulation. The substrate (DHF) and the cofactor NADPH are displayed with the QM region shaded in blue and the MM region shaded in white. The QM/MM link atoms are shown in red.

to not be computationally over expensive. This arrangement puts the pteridine and aromatic rings of DHF and the nicotinamide and ribose rings of NADPH in the QM region.

The systems were minimized using the ABNR method and then heated to 300 degrees K over 70 ps. The hydrogen atoms were constrained using the SHAKE method for 500 picoseconds of equilibration (Rychaert *et al*, 1977). Following this, the SHAKE constraints were taken off the HN4 and HN42 atoms attached to the CN4 atom of NADPH. For wild-type DHFR and the mutants, the system was run for an additional 1 nanosecond using one femptosecond time steps during which time trajectory information was stored every 0.05 picoseconds for later data analysis.

A Potential of mean force (PMF) calculation was performed using CHARMM in order to determine the free energy profile for the hydride transfer step in both *E. coli*

DHFR and R67 DHFR. This was done by using the umbrella sampling method (Torrie and Valleau, 1974). The frequency data are analyzed by the Weighted Histogram Analysis Method (WHAM), (Kumar *et al,* 1992). The reaction coordinate was defined as: RC = d(C6---HN4) – d(CN4---HN4) where d(C6---HN4) is the distance between DHF: C6 and NADPH: HN4 and d(CN4---HN4) is the distance between NADPH:CN4 and NADPH:HN4.

The simulation was conducted both forward and backwards using 22 windows in each direction for *E. coli* DHFR and 23 windows in each direction for R67 DHFR. Each window was 100 ps in length or 100,000 time steps. The data from the first 50,000 time steps in window were discarded for equilibration and the remaining 50,000 were used for data analysis. The average PMF free energy of the forward and backward plots was calculated.

Energy Corrections

Initial studies of the self-consistent-charge density-functional tight-binding (SCC-DFTB) (Cui *et al*., 2001) method revealed major problems in its description of the energy profile of the hydride transfer step. When using this method uncorrected the free energy actually drops when approaching the transition state (data not shown). In order to correct for this, model systems were created. The model system consists of the pteridine and aromatic rings of DHF and the nicotinamide ring of NADPH (Figure 7). Because R67 DHFR employs the endo-transition state as opposed to the exo-transition state used by *E. coli* DHFR, two separate model systems were created. Constrained minimization was performed using the SCCDFTB method at 40 points along the reaction coordinate

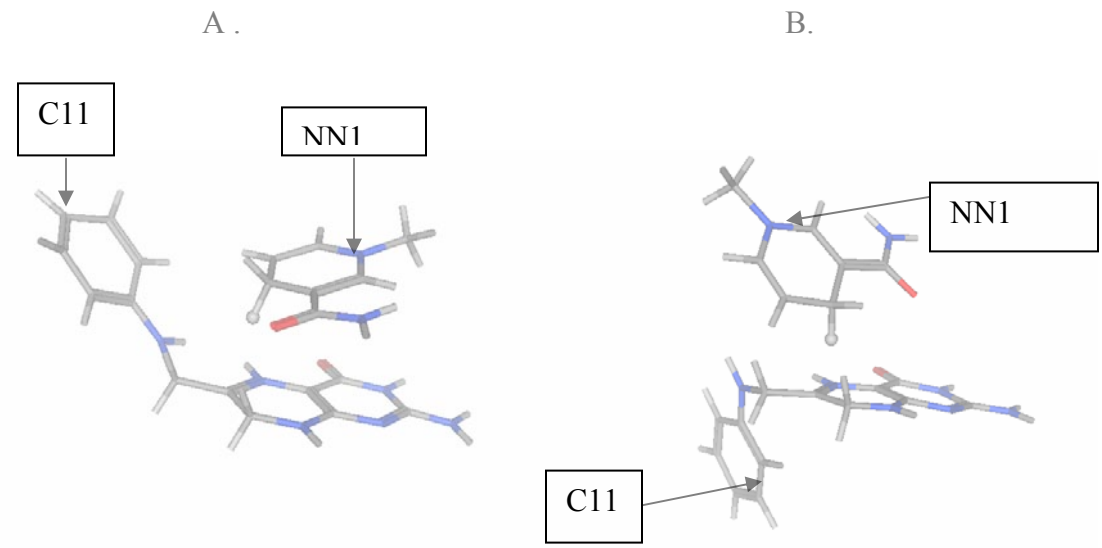


Figure 7: R67 DHFR (A) and *E. coli* DHFR (B) model systems. The R67 DHFR (A) model represents the endo transition state while the *E. coli* DHFR (B) model represents the exo transition state. The carbon atoms are shown in grey, the hydrogens are shown in white, the oxygens are red and the nitrogens blue. The hydrogen that is transferred during the reaction is shown as a ball.

from 2.0 to -1.9 Å. The C11 and NN1 atoms were fixed in order to prevent the substrate and cofactor from floating apart. The minimization was done in both the forward and reverse directions to insure that the same structures and SCCDFTB energies could be obtained each time. The energy for the resulting fixed point structures was then calculated at the B3LYP/6-31G* level using the GAMESS-US program.

The initial correction was performed by starting at the calculated ground state and adding the difference between the model system SCCDFTB potential energy and the B3LYP/6-31G* energy to the SCCDFTB determined free energy such that:

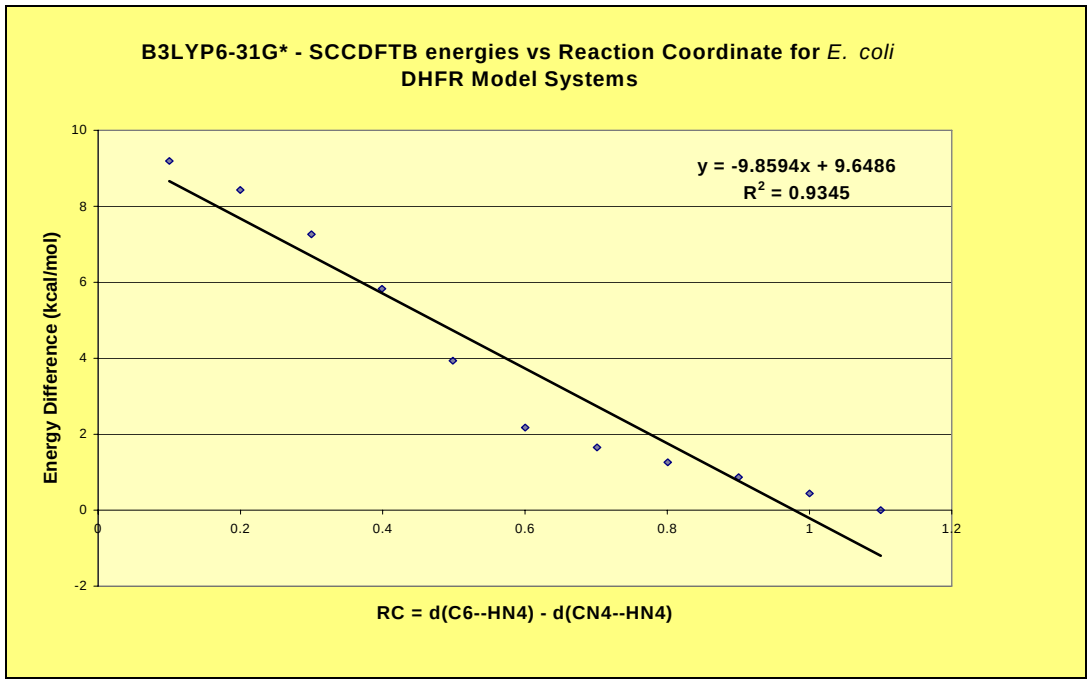
$PMF_CORR_n = PMF_SC_n + (\Delta E_B3_n - \Delta E_SC_n)$ where PMF_SC is the SCCDFTB determined free energy, E_SC is the SCCDFTB determined potential energy of the model system and E_B3 is the B3LYP/6-31G* determined potential energy of the same model. PMF_CORR is the final corrected value for the free energy and n is the point along the

reaction coordinate.

The relationship between the SCCDFTB and B3LYP/6-31G* energies was not linear. However, a linear approximation could be made between the ground state and the transition state. The value of ($\Delta E_B3_n - \Delta E_SC_n$) was plotted against the reaction coordinate between these two states and fit to a linear function (Figure 8). This linear function was then used as a biasing potential for a second PMF calculation. The results of this calculation cannot be used to determine free energy outside this specific range however, they can be used to estimate the activation energy.

Three single mutants and one functional triple mutant in addition to the wild-type were chosen for further dynamics studies. The Y69F, I68L, and Q67C mutants were chosen because Tyr 69, Ile 68, and Gln 67 are all important active site residues that retain some activity when these mutations are made. The M10 mutant is a functional triple mutant discovered by Schmitzer and her associates during combinatorial mutagenesis studies. The k_{cat}/K_m for M10 is actually higher then the activity of the wild-type. M10 is a composite of V66I, Q67N, and I68R mutations (Table 1). It should be noted that for the purpose of this study the mutations will be made on all four chains.

A.



B.

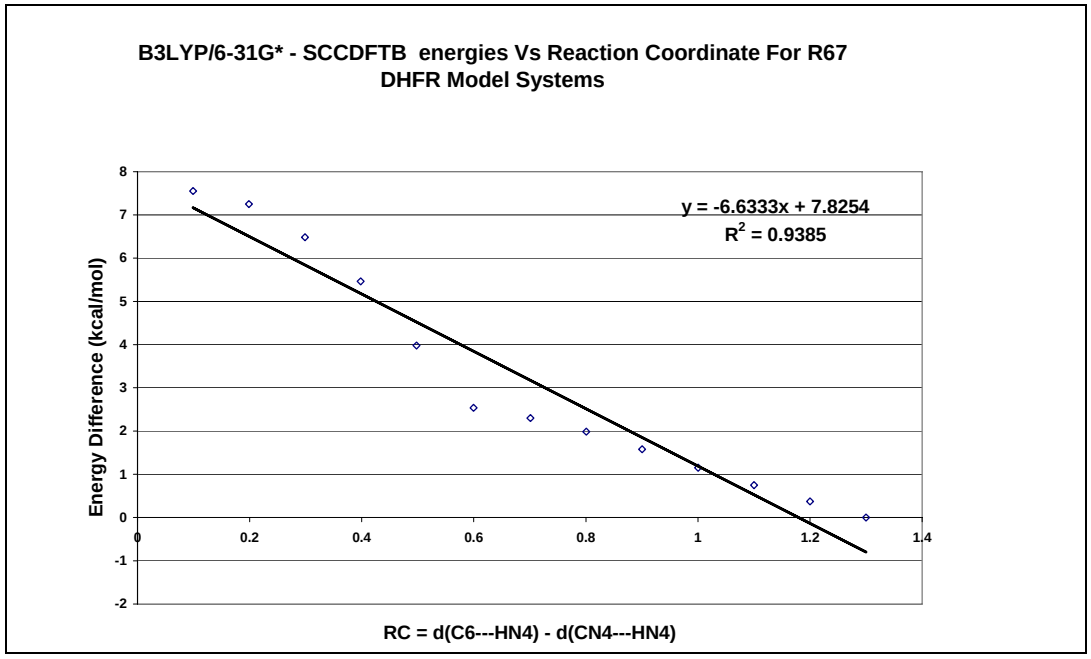


Figure 8: Linear Correction Curves for *E. coli* DHFR (A) and R67 DHFR (B). For each model system the difference in B3LYP6-31G* and SCCDFTB energy is plotted against the reaction coordinate from the ground state on the right to the transition state on the left. The plot is fitted to a linear curve shown on the chart which is used to bias the simulations during PMF data collection. The R^2 value is also shown on the chart.

	Sequence 66-69	k_{cat} s^{-1}	K_m (DHF) *u*M	K_m (NADPH) *u*M
WT [a]	Val - Gln - Ile - Tyr	1.3	5.8	3
Y69F [a]	Val - Gln - Ile - Phe	2.5	44	66
I68L [a]	Val - Gln - Leu - Tyr	0.32	24	26
Q67C [a]	Val - Cys - Ile - Tyr	0.1	55	26
M10 [b]	Ile - Asn - Arg - Tyr	1.2	3.5	5.3

Table 1: Wild-type and 4 mutants sequence. Sequence for residues 66-69 for Wild-type R67 DHFR and 4 mutants. Mutated residues are highlighted followed by the experimental values for k_{cat} , K_m (DHF), and K_m(DHF).

[a] Values from Strader *et al*. 2001.

[b] Values from Schmitzer *et al*. 2004.

Chapter III
Results

30 Angstrom Molecular Dynamics Results

The dynamics simulation was conducted for 2 nanoseconds during which time the trajectory was monitored. Initial analysis revealed the pABA-Glu tail of DHF to be quite mobile (see Figure 9). The average structure for the 240 to 280 ps interval shows the carboxyl terminus (O1) oxygen in an ionic interaction with B Lys-32. The Glu side chain (OE1) is bound via a water bridge to C Lys-32. This water bridge is broken by the 400 to 440 ps interval where the side chain moves further into the solvent but the ionic interaction between the carboxyl terminus end and B Lys-32 remains. At the 600 – 640 ps interval this interaction is weakened as a water molecule bridges the O1 oxygen to DHF: O. In the 840 – 880 ps structure, the ionic interaction is reformed. By the 960 to 1000 ps interval, the carboxyl-terminus end is bound to B Lys-32 and the side chain is pointing to the outer end of the pore. During the 1200-1240 ps interval, the carboxyl terminus end is no longer bound and the side chain is bound to Lys-32 by a water bridge. This bridge is short lasting and the carboxyl terminus end is rebound to Lys-32 by the 1240-1280 ps interval average. From this point there are no direct ionic interactions until the 1960 – 2000 ps interval where the side chain is once again bound to B Lys-32 (Figure 9).

This simulation shows the pABA-Glu tail of DHF to be quite mobile which is consistent with its lack of density in the X-Ray crystallography experiments. The formation of transient water bridges between the substrate and enzyme indicates the dynamic role that water plays in this interaction. Osmolality studies shows that about 30

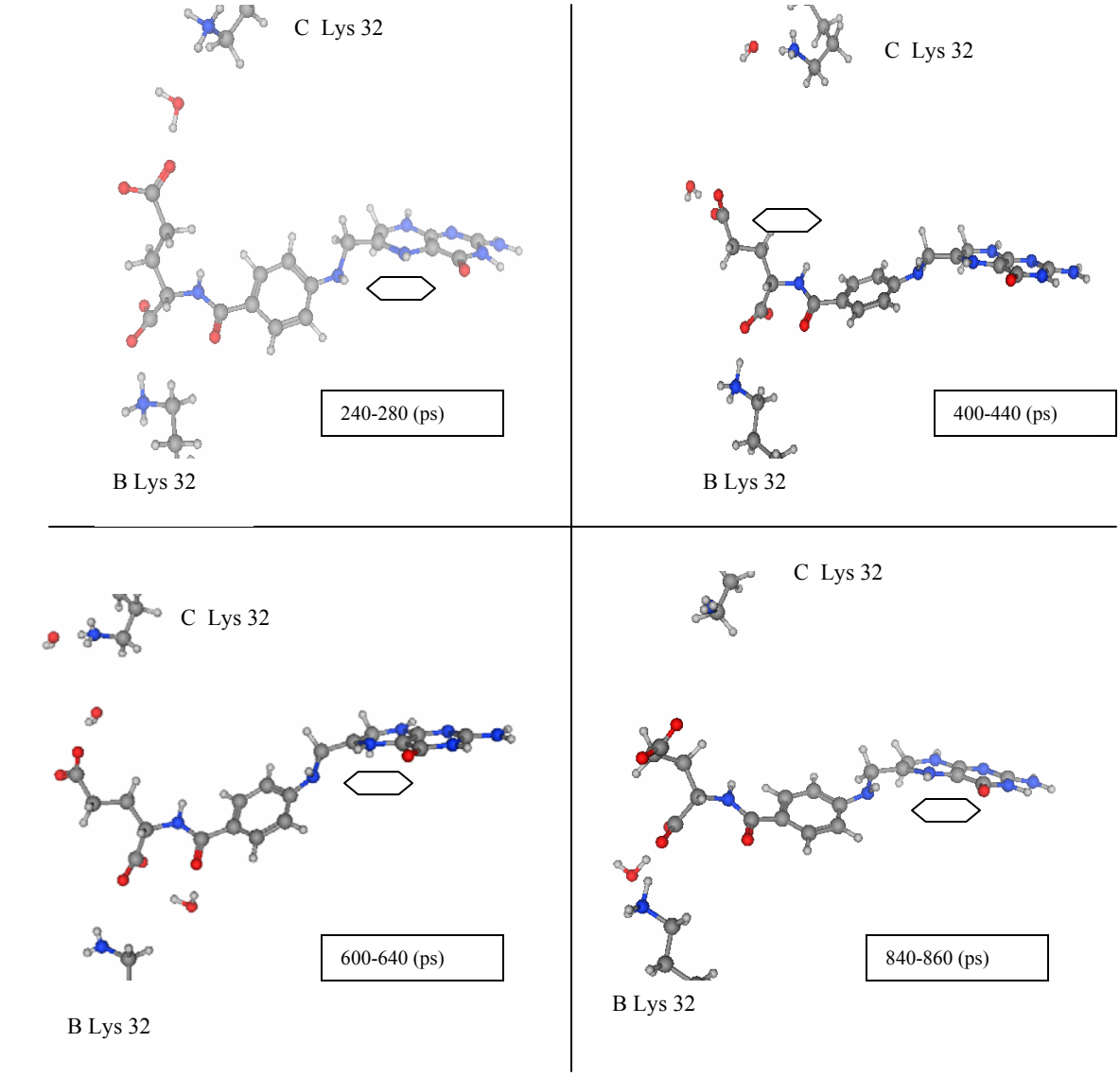


Figure 9 Movement of the pABA-Glu tail of DHF. DHF is shown in various conformations over the 2 nanosecond dynamics simulation time with the adjacent B Lys 32 and C Lys 32 residues. The black hexagon represents the

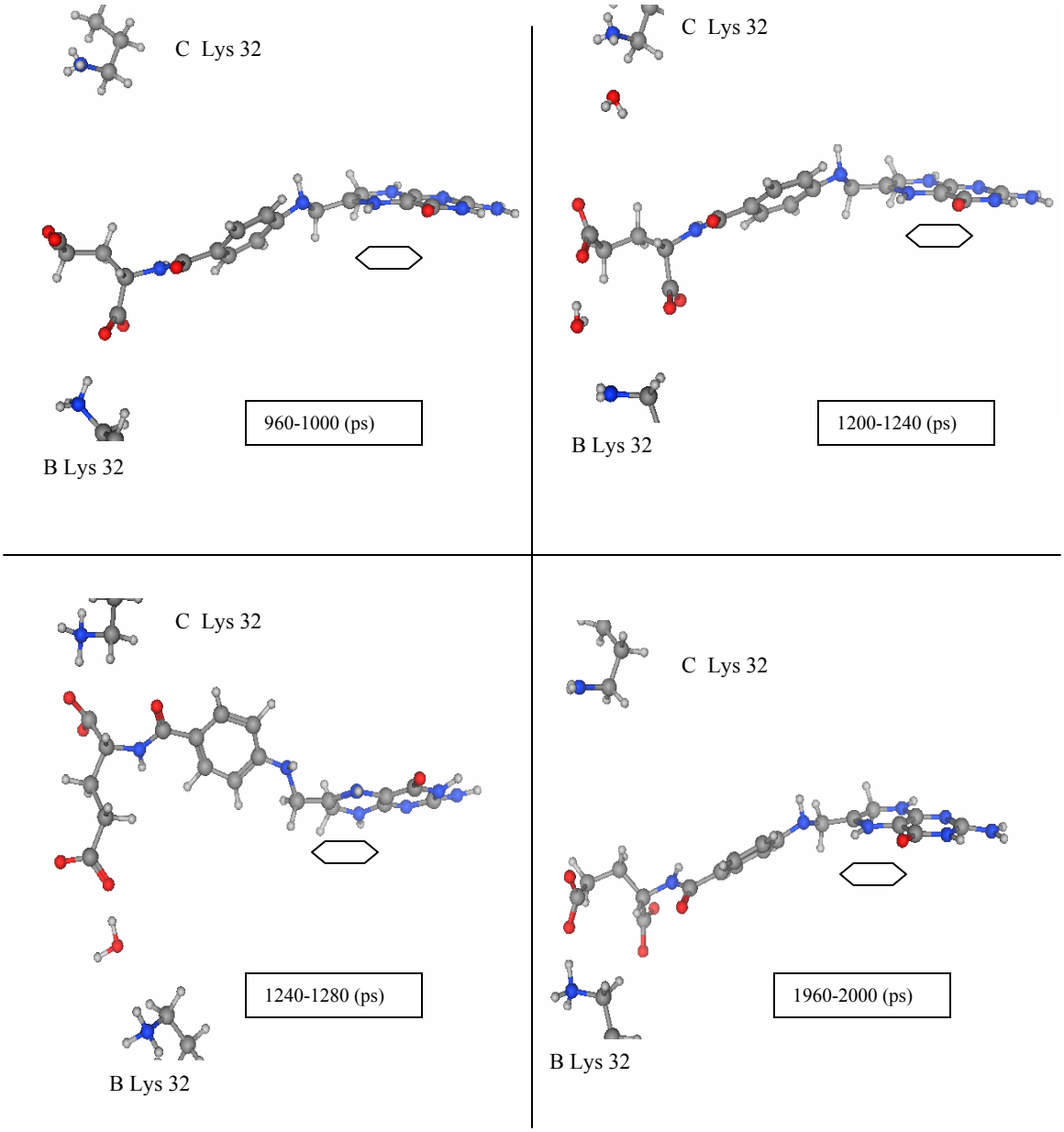


Figure 9 continued …

–50 water molecules are taken up upon DHF binding (Chopra *et al*, 2006). This most likely mitigates the entropic binding penalty by increasing the entropy of the system.

The position of the XN angle appears to be partially determined by interactions between B Val 66 O and HN atoms and the HN2T and ON3'atoms of the ribose ring. Free energy analysis of the trajectory data using 5 degree bins shows a minimum free energy at 60 degrees. PMF studies using umbrella sampling show a similar profile with a minimum at an XN angle of 55 degrees (Figure 10). An attempt to fully rotate the XN angle in the active site resulted in cofactor dissociation. This indicates that although the enzyme may be able to bind the enzyme in both the syn and anti conformations, it is unlikely that the cofactor can switch conformations while still bound to the active site.

In addition to the XN angle, 3 additional dihedral angles were monitored in the PMF and dynamics simulations in order to observe conformational changes in the substrate and cofactor. Puckering of the Pteridine ring was monitored by the N8—C7—C6—N5 dihedral which will be referred to as the DHF Dihedral (Figure 11 a). Puckering of the Nicotinamide ring is monitored by the CN2—CN3—CN4—CN5 dihedral which is referred to as the NADP dihedral (Figure 11 b). Finally the orientation between the pteridine ring of DHF and the Nicotinamide ring of NADP was monitored by another angle referred to as the PTR_NIC dihedral. This angle is the dihedral angle made from the C8A and C6 atoms of the pteridine ring and the CN4 and NN1 atoms of the nicotinamide ring (Figure 12).

Potential of Mean Force Calculations

The initial corrected PMF curve for *E. coli* DHFR shows activation energy of

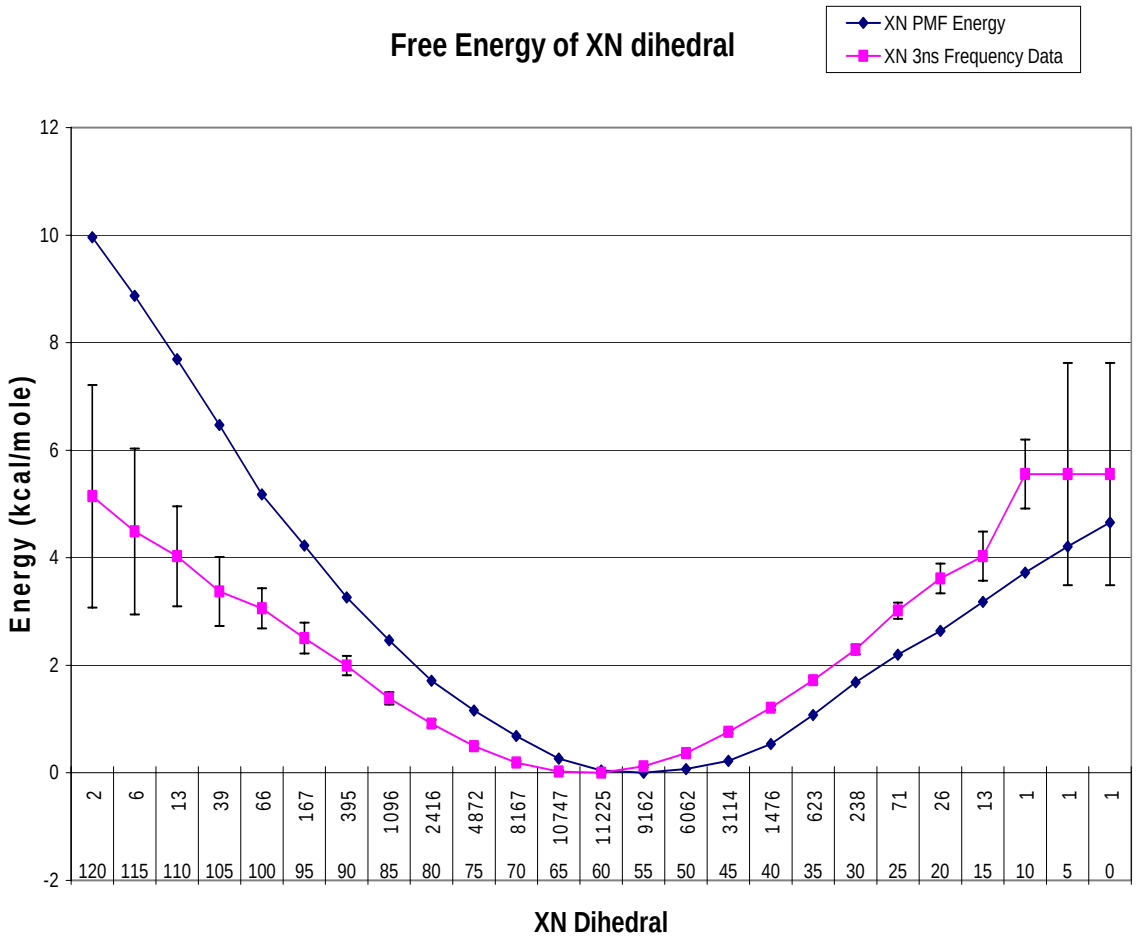


Figure 10: Free energy profile of XN angle using PMF (Blue) and Dynamics Frequency Data (Pink). For the frequency data, the curve is divided into 24 bins from 0 to 120 degrees. The frequency for each bin is displayed directly below the X axis. The free energy is determined from the equation $\Delta G = -RTLn(n/N)$ (see text for details). Error bars for the frequency data are calculated using the method described by Kobrak, 2003.

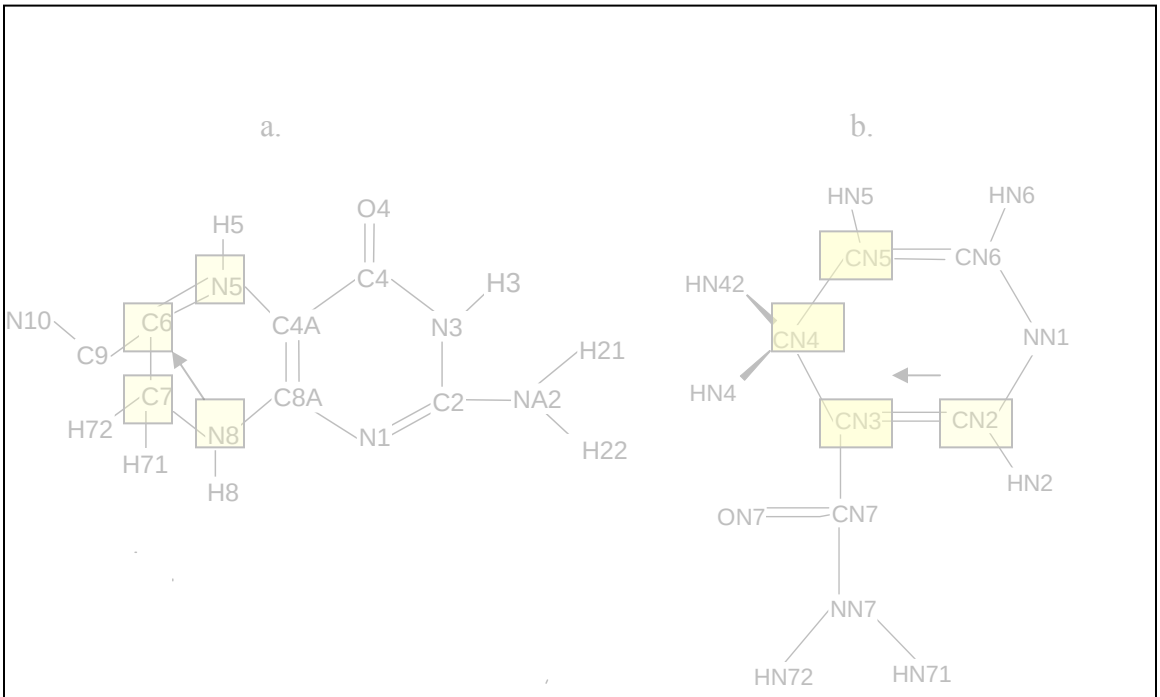


Figure 11: DHF Dihedral angle (a) and NADPH dihedral angle (b). The DHF dihedral is the N8—C7—C6—N5 angle (highlighted) and the NADPH dihedral is the CN2—CN3—CN4—CN5 angle. These angles are used to measure the puckering of the pteridine and nicotinamide rings respectively.

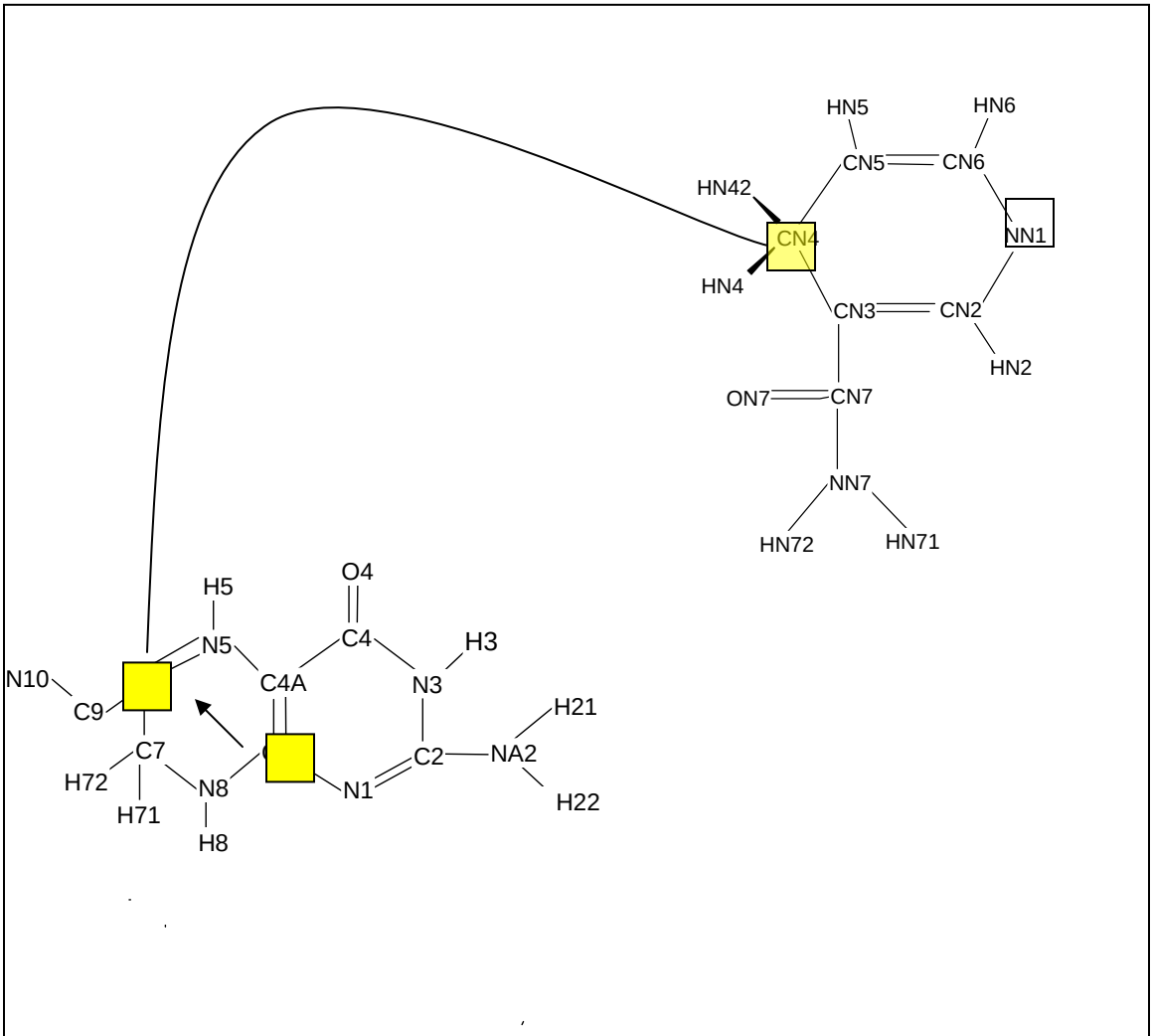
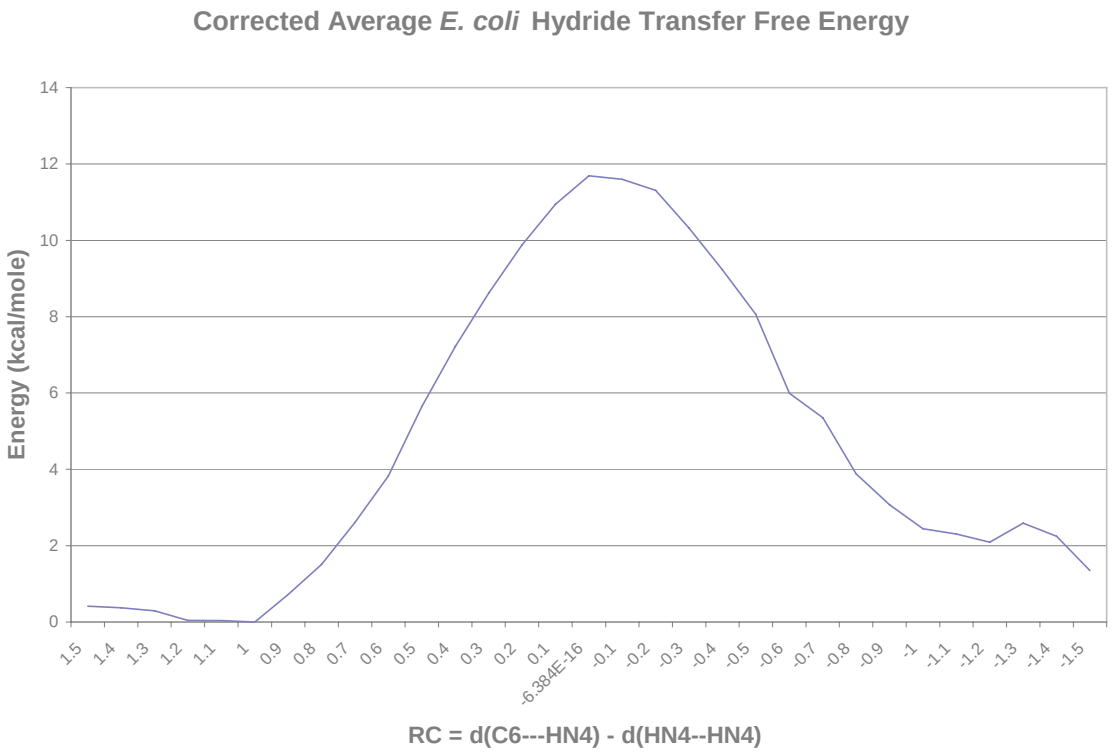


Figure 12: Pter-Nic Dihedral angle between Pteridine and Nicotinamide Rings. The Pter-Nic angle is the C8A—C6—CN4—NN1 angle. This angle is used to determine the orientation between the pteridine and nicotinamide rings.

11.7 kcal/mole. This is relatively close to the experimental value for the activation energy of 13.4 kcal/mole. The calculated free energy change for the reaction, however, is 2.09 kcal/mole which is considerably different from the experimental value of -4.4 kcal/mole (Figure 13 a). This could be due in part to some observed substrate-cofactor movement after the hydride transfer that is not accounted for by the simple reaction coordinate used in this experiment. The transition state is located at RC = 0.0 which is different from the value of 0.1 obtained from both the model system (not shown) and previous computational experiments by Garcia-Vilocia (2003) who obtained a value of 0.145. This is most likely a result of the significant error of the uncorrected SCCDFTB method in energetically describing the reaction around the transition state. The free energy curve with the linear correction function from the ground state to the product state gives a activation energy of 13.0 kcal/mole which is much closer to the experimental value of 13.4 (Figure 13 b).

Average values for certain parameters were obtained from 100 ps simulations at the ground state (GS), transition state (TS), and product state (PS) of the reaction and are given in table 2. The C6---CN4 distance was 3.26 Å for the ground state, it decreased to 2.67 Å for the transition state and then went back up to 3.09 Å for the product state. The C6—HN4—CN4 angle average changed some between the GS and TS but was still within one standard deviation. But between the TS and the PS the angle drops from 157.2 to 141.6 degrees. Neither the XN dihedral nor the PTR_NIC dihedral appears to change much during the reaction. Some slight puckering of the Nicotinamide ring was observed. However the high standard deviation indicates that the NADPH dihedral is quite dynamic. Table 2 shows a significant change in the DHF dihedral over the course of the

a.



b.

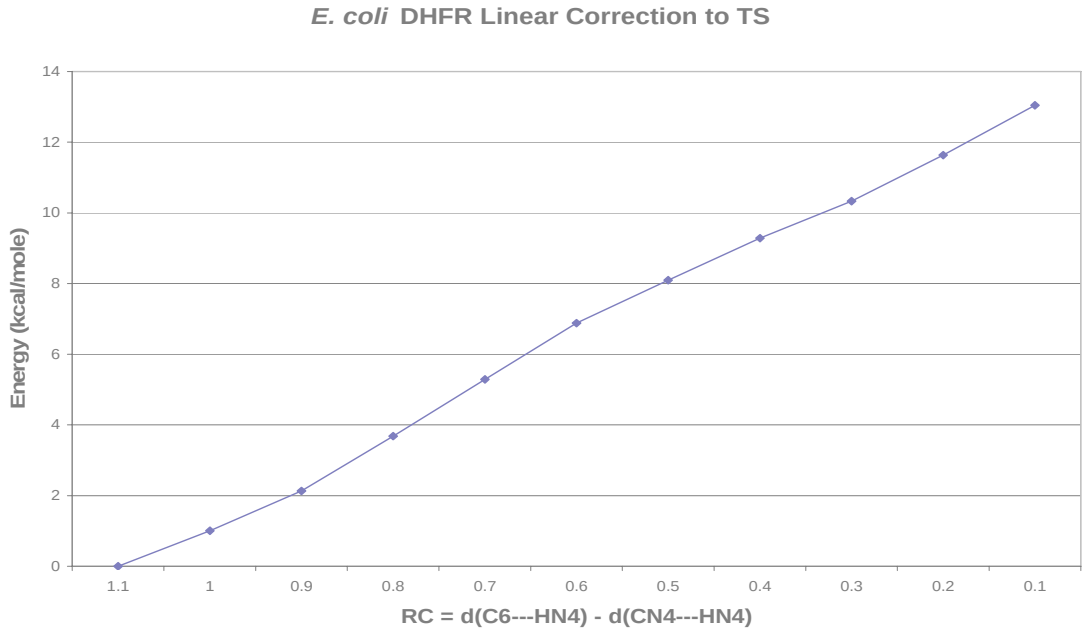


Figure 13: Free energy curve for the hydride transfer step of *E. coli* DHFR using a point correction (a) and a linear correction to the transition state (b). The reaction proceeds from left to right.

				E. coli DHFR PMF Summary			
		C6---NC4	C6--NH4--NC4	XN Dihedral	PTR-NIC Dihedral	NADP Dihedral	DHF Dihedral
GS	AVERAGE	3.26	153.6	-121.9	152.7	-7.2	24.0
	STDEV	0.10	11.2	11.0	7.1	14.0	9.1
TS	AVERAGE	2.67	157.2	-125.0	150.4	-13.2	46.6
	STDEV	0.06	6.6	9.6	5.2	7.2	6.1
PS	AVERAGE	3.09	141.6	-126.5	150.0	-3.2	53.8
	STDEV	0.08	6.7	9.8	5.4	6.9	5.9

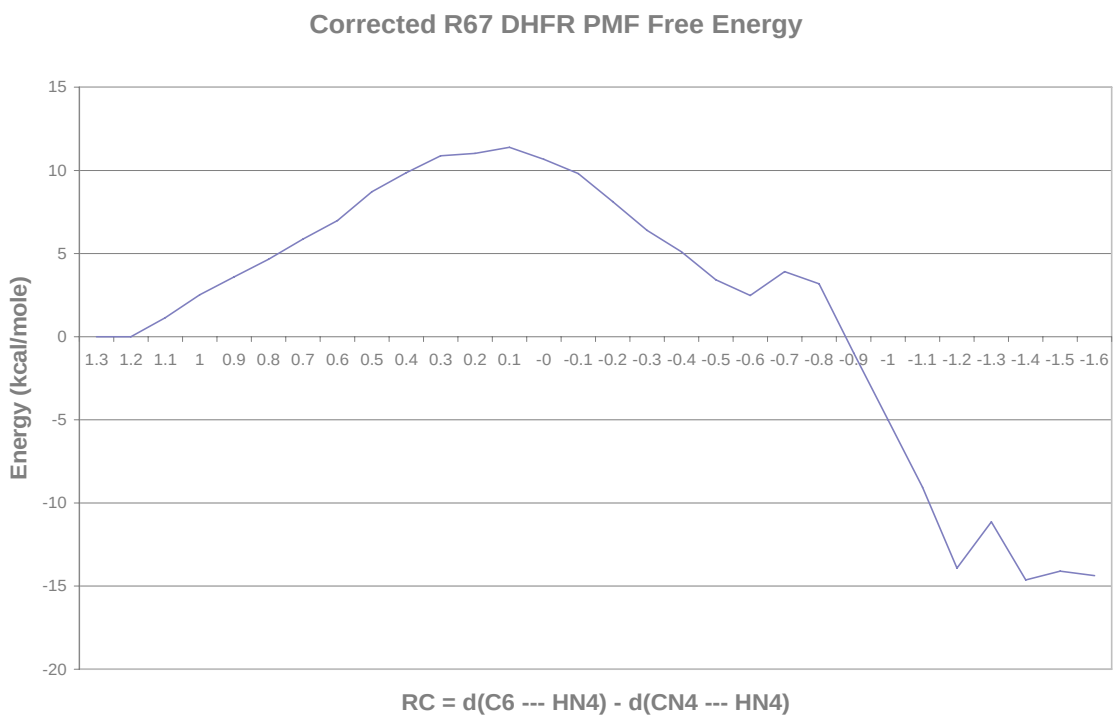
Table 2: *E. coli* DHFR PMF Summary. Average values for the *E. coli* DHFR hydride transfer step at the Ground State (GS), the Transition State (TS), and the Product State (PS). The standard deviation (STDEV) is shown for each value in order to indicate the variation during dynamics.

reaction. The DHF dihedral starts at 24.0 degrees at the GS, goes to 46.6 at the TS and ends up at 53.8 degrees at the PS.

Other experimenters who did free energy calculations on *E. coli* DHFR either obtained an unrealistic energy barrier due to errors in the semiempirical method (Pang *et al*., 2006), or used Empirical Valence Bond theory in their calculations (Agarwal *et al*., 2002). Because of this it is not possible to validate this model by comparing the energy barrier to that of other models. A Better method of validation is to compare structures at the transition state. Garcia-Vilocia (2003) had a C6 -- CN4 distance of 2.68 Å , and a C6 – HN4 –CN4 angle of 157 degrees at the transition state which matches extremely well with the data from this study of 2.67 Å for the C6 -- CN4 distance and 157.2 for the C6 – HN4 –CN4 angle (see table 2). Agarwal *et al*. (2002) obtained values of 2.73 Å and 164 degrees respectively at the transition state. The differences are most likely due to the methods as Agarwal used a GROMOS force field while Garcia-Vilocia used a CHARMM force field like that which was used in this experiment.

For wild-type R67 DHFR, the activation energy for the initial corrected PMF curve is 11.4 kcal/mole. The TS is calculated to be at RC = 0.1 Å which is consistent with the model system. The free energy change for the reaction is calculated to be -12.9 kcal/mol (Figure 14 a). Again there could be some motion of the substrate and cofactor after the hydride transfer reaction that is not accounted for by the reaction coordinate. Experimental analysis shows that the THF product binds significantly more weakly to the enzyme than the DHF substrate (Jackson, 2005). This could account for the significantly low free energy change of the reaction. The free energy curve with the linear correction from the GS to the PS gives a value of 10.0 kcal/mole, which is similar to the value

a.



b.

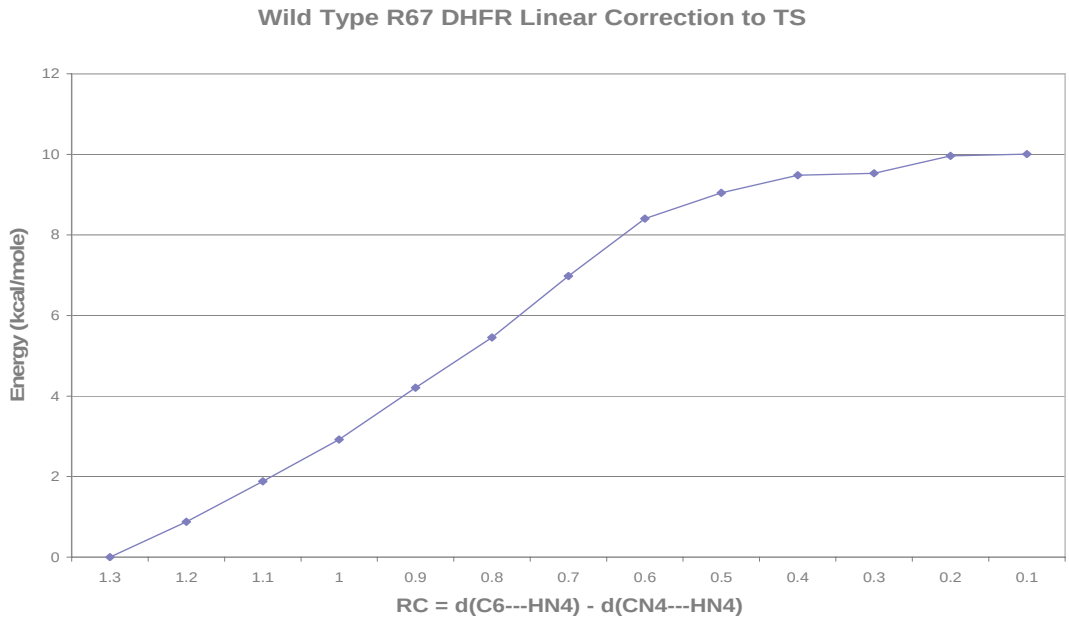


Figure 14: Free energy curve for the hydride transfer step of R67 DHFR using a point correction (a) and a linear correction to the transition state (b).

obtained in the original corrected PMF curve (Figure 14 b).

The C6---CN4 distance starts out at 3.32 Å for the GS, goes to 2.67 Å at the TS and then ends up at 2.72 Å at the PS. A much more significant change in the C6—HN4—CN4 angle occurs in R67 DHFR then in the *E. coli* form. The angle starts out at 135.8 º at the GS, goes up to 169.0 º at the TS then drops to 98.7 ºat the PS. This indicates significant change in the orientation of these atoms after the hydride transfer occurs (Table 3). Figure 15 shows scatter plots representing how the C6—CN4 distance (Figure 15 A) and the C6—HN4—CN4 angle (Figure 15 B) change during the reaction. The XN and PTR_NIC dihedrals change some over the course of the reaction but the changes are not significant with regard to the standard deviation values. The NADP dihedral actually appears to become more positive at the TS and even more so at the PS. The standard deviation values indicate that this dihedral is dynamic, however this could indicate that puckering of the nicotinamide ring is not as much of a factor in R67 DHFR as it is in *E. coli* DHFR. Also the overlap of the nicotinamide and pteridine rings in R67 DHFR may favor nicotinamide puckering away from the pteridine at the transition state in order to avoid steric interference. The DHF dihedral at the GS of 6.6 º indicates that the pteridine ring is almost in a plane at this stage. This dihedral goes up to 36.6 º at the TS and drops slightly to 29.4º at the PS. The high standard deviation of the DHF angle at the PS of 26.0º indicates that the pteridine ring has a wide range of motion at this stage (Table 3). This is consistent with binding studies which show that the tetrahydrofolate product binds much more weakly to the complex then the DHF substrate (Jackson et al., 2005).

At this point the question remains as to what effect the tail motion has on the reaction . The pABA-GLU tail has been shown to be mobile (see Figure 9) but the effect

				R67 DHFR PMF Summary			
		C6---NC4	C6--NH4--NC4	XN Dihedral	PTR-NIC Dihedral	NADP Dihedral	DHF Dihedral
GS	AVERAGE	3.32	135.8	62.1	22.1	-3.2	6.6
	STDEV	0.12	10.4	8.6	5.4	12.6	12.4
TS	AVERAGE	2.67	169.0	58.5	27.4	10.4	36.6
	STDEV	0.06	5.9	9.3	6.1	7.8	9.1
PS	AVERAGE	2.72	98.7	49.5	20.1	12.9	29.4
	STDEV	0.06	3.7	9.2	9.2	6.4	26.0

Table 3: R67 DHFR PMF Summary. Average values for the R67 DHFR hydride transfer step at the Ground State (GS), the Transition State (TS), and the Product State (PS). The standard deviation (STDEV) is shown for each value to indicate the variation

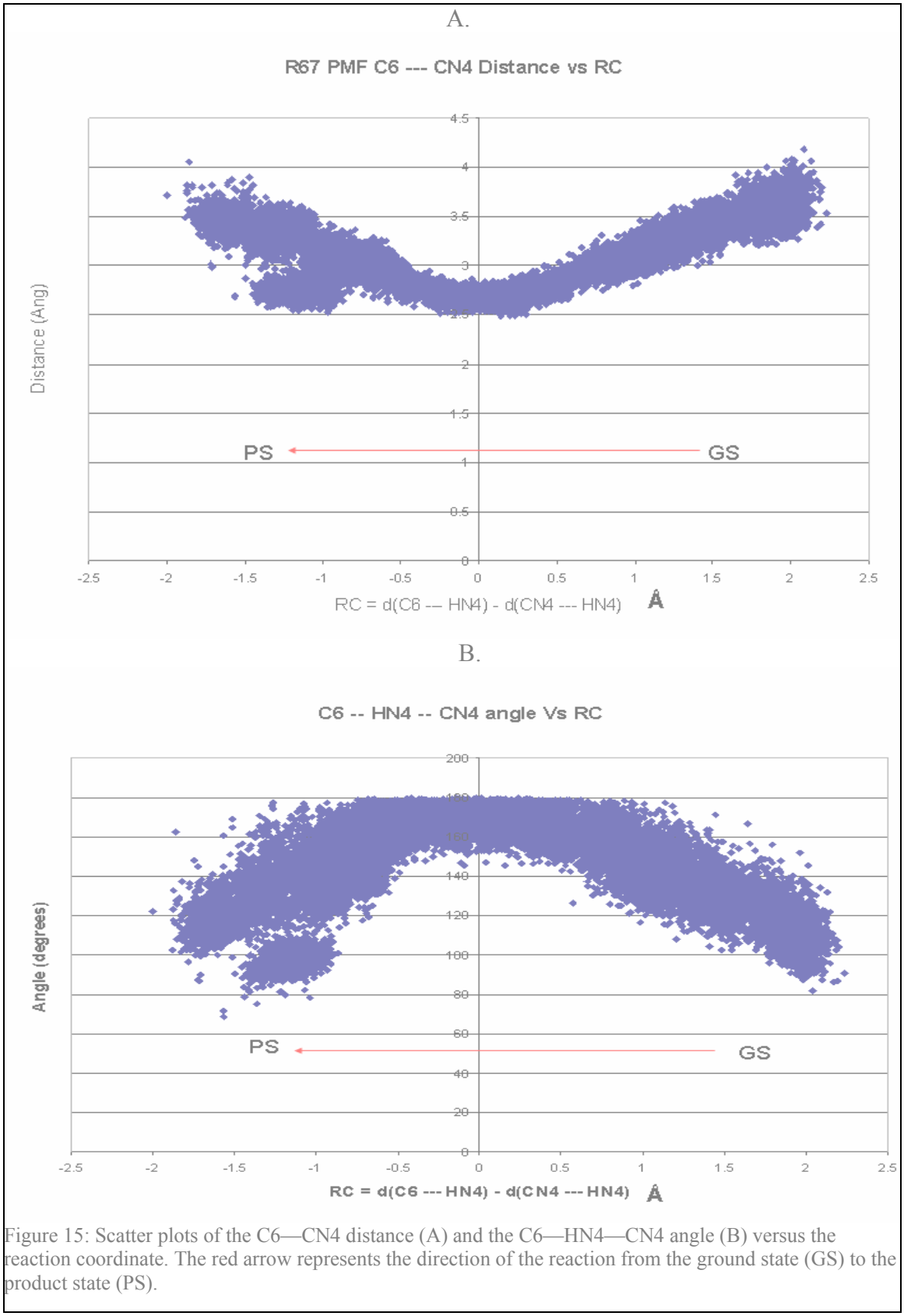


Figure 15: Scatter plots of the C6—CN4 distance (A) and the C6—HN4—CN4 angle (B) versus the reaction coordinate. The red arrow represents the direction of the reaction from the ground state (GS) to the product state (PS).

of this on the reaction remains unclear. The relatively short time scale of the PMF calculation prevents the tail from sampling all possible configurations. To examine this further, two additional PMF calculations with linear correction functions were performed with added constraints on the tail position. In the first simulation the CD carboxyl carbon (at the end of the Glu sidechain of DHF) was restricted within 4 angstroms of the NZ nitrogen of B Lys 32. The second simulation is performed with the CT tail (at the c-terminus eng of the Glu sidechain) constrained within 4 angstroms of C Lys 32:NZ. The CD constrained curve shows a energy barrier at 0.2 on the reaction coordinate as opposed to 0.1 in the unconstrained wild-type. It is also slightly higher (by about 0.4 kcal/mole) than the unconstrained model. In the CT constrained curve, the energy barrier is higher by about 1.5 kcal/mole (Figure 16). This indicates that the tail position may have only small effects on the catalysis of the reaction. Previous studies show that Lys 32 mutations at a single site do not significantly reduce the catalytic rate (Hicks *et al* , 2004). This could mean the pABA-glu tail acts more like a guide to bring the substrate to the active site rather than a contributor to catalysis once the substrate is in place.

Comparison of Mutant forms to Wild-type R67 DHFR

At this point the origin of mutational effects on R67 DHFR are studied by constructing models of 3 single point less functional mutants (Y69F, I68L and Q67C) and one fully functional triple point mutant (M10). The results of these dynamics studies will be used to examine the determinants of binding and catalysis. Average values for specific interactions, angles and dihedrals are shown in Table 4 followed by a further in depth structural analysis.

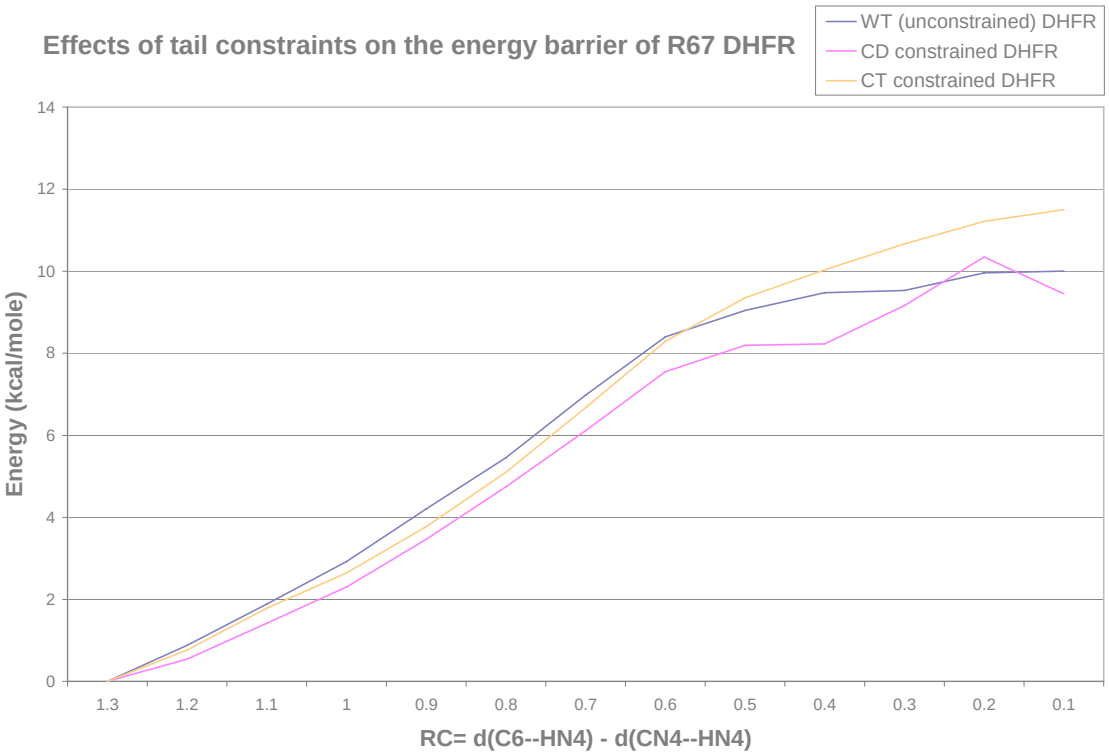


Figure 16: Effects of tail constraints on the energy barrier of R67 DHFR unconstrained (blue), CD constrained (pink), and CT constrained (orange).

				Dynamics Summary Data			
WT DHFR		C6---NH4	NC4---NH4	C6---NC4	C6---NH4---NC4	PTR_NIC Dihedral	XN Dihedral
	AVERAGE	2.60	1.13	3.43	131.6	25.5	55.9
	STDEV	0.44	0.06	0.29	16.1	6.1	8.7
Y69F							
	AVERAGE	2.51	1.12	3.38	135.2	43.0	51.8
	STDEV	0.27	0.03	0.21	13.3	6.5	9.7
I68L							
	AVERAGE	2.63	1.12	3.50	135.9	42.6	-4.3
	STDEV	0.32	0.03	0.26	13.2	10.2	11.8
Q67C							
	AVERAGE	2.68	1.12	3.57	138.2	21.8	39.0
	STDEV	0.31	0.03	0.03	13.7	5.8	11.0
M10							
	AVERAGE	2.85	1.12	3.78	141.6	29.1	40.2
	STDEV	0.34	0.03	0.29	12.4	5.1	9.8

Table 4: Dynamics Summary Data. Dynamics summary data for WT R67 and 4 mutants showing values and standard devation (STDEV)

Average structural parameters from the 1 nanosecond dynamics simulations of the WT and 4 mutants are shown on Table 4. Molecular dynamics of WT DHFR reveal interesting changes to the structure of the active site. As expected, the backbone O and NH groups of Ile 68 (on the A and D) chains form hydrogen bonding interactions with N3 and O4 atoms of the pteridine ring and ON7 and N7 atoms of the nicotinamide ring respectively. On the NADPH side of the pore, the ribose ring of NADPH is stabilized in the syn conformation by hydrogen bonds between ON3' and NH backbone group of A Val 66. The ON2'-HN2T group rapidly switches back and forth between being an H-bond donor to A Gln-67:OE1 and B Val 45:O. A Gln 67:OE1 also H-bonds to B Gln-67: HE21 forming a H-bonding triangle above the nicotinamide ring. A Tyr-69:OH forms transient H-bonds with AGln 67:HE22. Finally B Gln-67: OE1 hydrogen bonds to the OH group of B Tyr-69 (Figure 17 a). This bond is broken, however, when the pABA-Glu tail moves to the NADPH side of the pore and DHF:OE1 becomes the H-bond acceptor.

On the DHF side of the pore the hydrogen bonding network appears to be reversed. The oxygen of the OH group on C Tyr 69 H-bonds with CGln 67: HE22. CGln 67: OE1 H-bonds with DGln 67: HE22. The OE1 atom of DGln67 H-bonds with DTyr 69: HH to complete the H-bonding network. There do not appear to be any direct ionic interactions between the pABA-Glu tail of DHF and either Lys 32 residue. ALys 32 forms an ionic interaction with ON2 of the pyrophosphate group of NADPH. DLys 32 forms an ionic interaction with O1P of the phosphate group off of the adenosine ring (Figure 17 b).

The C6---NC4 distance averages 3.43 angstroms with a stºandard deviation of approximately 0.30 Å . The C6—NH4—NC4 angle averages 131.58 º with a standard

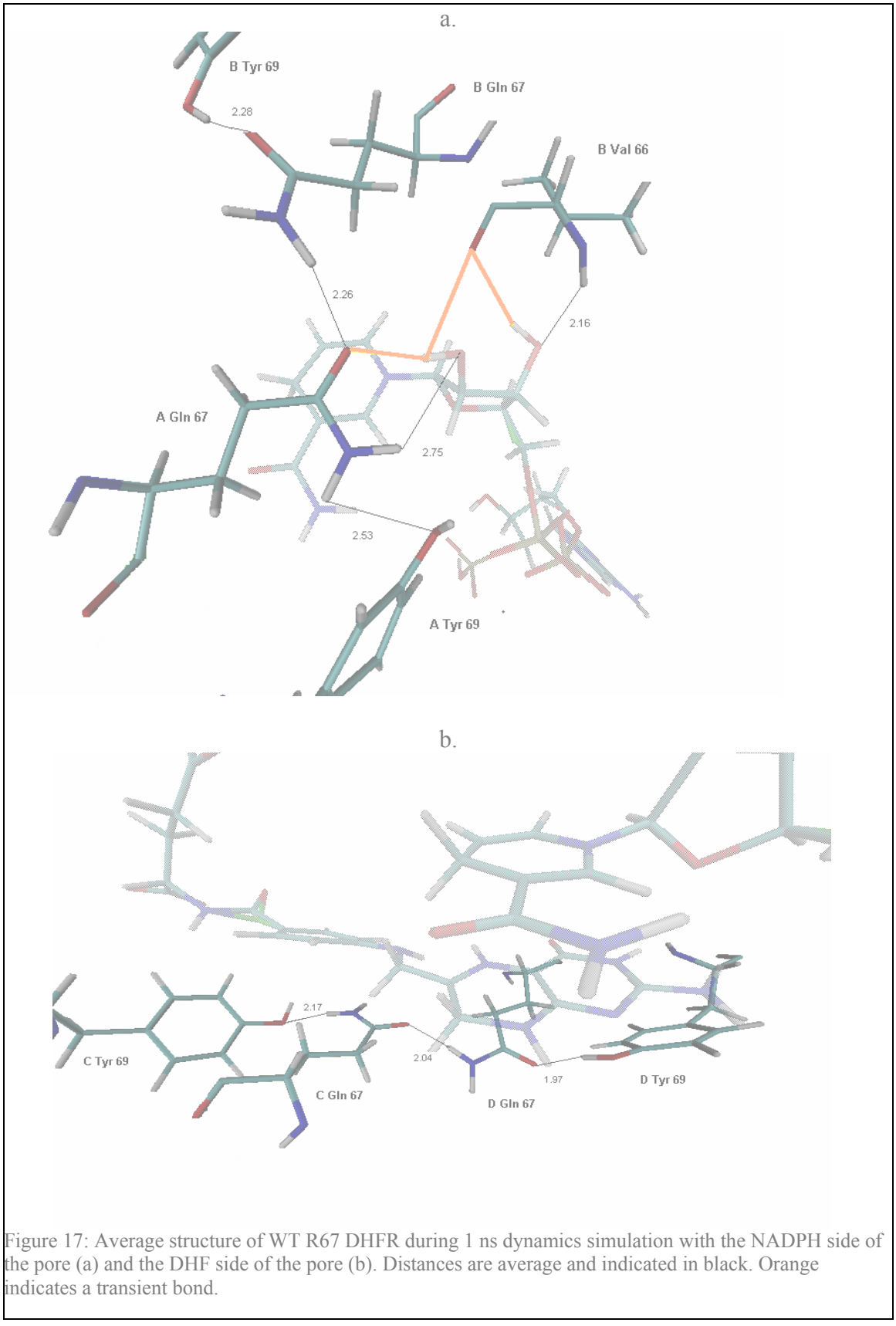


Figure 17: Average structure of WT R67 DHFR during 1 ns dynamics simulation with the NADPH side of the pore (a) and the DHF side of the pore (b). Distances are average and indicated in black. Orange indicates a transient bond.

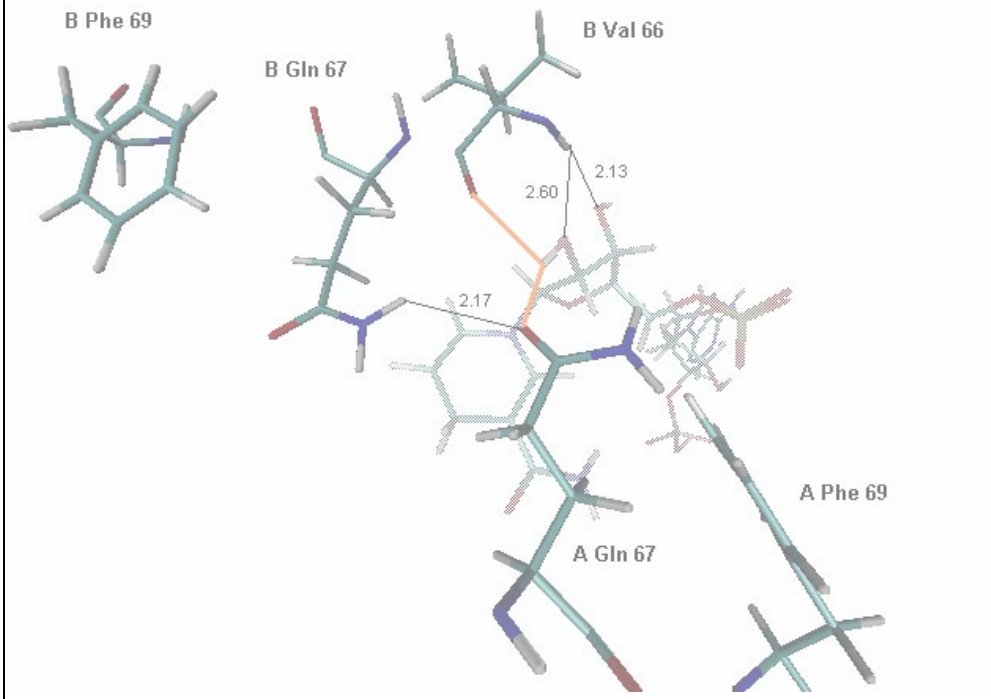
deviation of 16.13 º. The dihedral angle between the nicotinamide and ribose rings averages 25.47 º and the XN dihedral averages 55.90 º. This is consistent with the energy minimum of 55 º obtained from the previous XN PMF curve (Table 4).

The Y69F mutation results in a change in the active site environment due to the loss of electrostatic binding potential at that residue. This results in a slight change in the orientation of the pteridine and nicotinamide rings such that the Pter-Nic angle is 43.0 º. The hydrogens of both OH groups of the ribose ring form transient interactions with B Val 66:O. NADPH: HN2T also forms a transient bond with A Gln67: OE1 which forms a more stable interaction with B Gln67:OE1.This results in the positioning of the A GLN 67: OE1 atom generally close to the plane of the nicotinamide ring (Figure 18 a). On the DHF side of the pore C Gln 67 :OE1 is bound to D Gln 67:HE22 (Figure 18 b).

The striking feature of the I68L dynamics simulation is that the XN dihedral has an average value of -4.27 º as opposed to the wild-type for which it is 55 º (see Figure 18). The reason for this appears to be steric hindrance between the D Leu-68 and B Val-66 side chains. This effect changes the position of the backbone of B Val-66 which interacts with the ribose ring of NADPH. B Val-66:O H-bonds with NADPH HN2T and B Val-66:O H-bonds with ON2' of the same OH group on the NADPH ribose ring. A Gln-67:OE1 is bound to B Gln-67: HE22, however the B Gln-67 residue is oriented sideways with respect to A Gln-67. A Gln-67:OE1 also participates in an H-bond with NADPH:HN2T while A Tyr-69: forms a H-bond with NADPH: ON5'. As in the wild-type, the H-bonds involving HN2T are transient. B Tyr-69:HH also transiently interacts with the OE1 oxygen of the DHF tail (Figure 18 a).

On the NADPH side of the pore, A Gln-67:OE1 is bound to B Gln-67: HE22.

a



b.

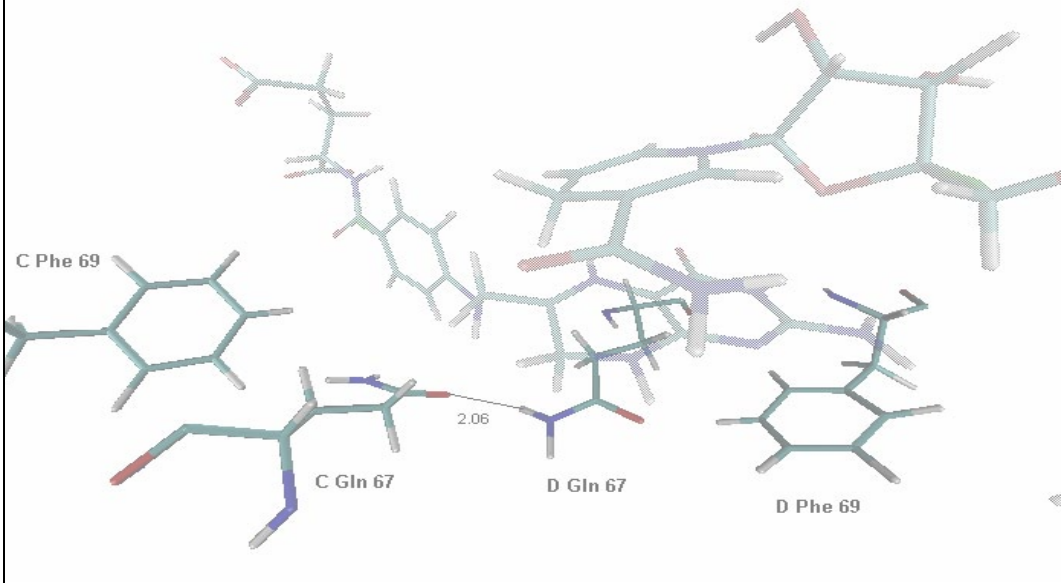


Figure 18: Average structure of Y69F R67 DHFR during 1 ns dynamics simulation with the NADPH side of the pore (a) and the DHF side of the pore (b). Distances are average and orange indicates a transient bond.

Unlike the wild-type, The B Gln-67 side chain is oriented sideways with respect to A Gln-67. This results in the electronegative OE1 atom being positioned further from the plane of the nicotinamide ring (Figure 19 a). On the DHF side of the pore, the hydrogen bonding network is more consistent with the wild-type. The exception to this is that C Tyr-67 is hydrogen bound to DHF:O. This provides further indication of the importance of Tyr 69 in binding of both the substrate and cofactor. C Tyr-69 is bound to C Gln-67:HE22 and C Gln-67:OE1 is bound to D Gln-67:HE22. The H-bonding network is completed with D Gln 67:OE1 binding to D Tyr-69:HH (Figure 19 b).

The hydrogen bonding network of the Q67C mutant is much more transient and dynamic than the wild-type and the other mutants. On the NADP side of the pore A Cys-67: HG1 forms a transient interaction with A Tyr-69:OH. A Tyr-69:HH forms another transient interaction with NADP:ON2'. NADP:HN2T forms a transient interaction with B Val-66:O. The only stable hydrogen bonding interaction that is formed is between B Val-66:HN and NADP:ON3'. Another transient interaction is formed between B Cys-67:S and B Tyr-69:HH (Figure 20 a). On the DHF side of the pore there are no hydrogen bonding interactions formed within the enzyme (Figure 20 b).

The M10 triple mutant provides an example of how a significant change in the active site environment can lead to an alternative functional structure. The XN angle is approximately 40 degrees, but the surrounding atoms are different. B Ile-66:HN is bound to NADPH:ON2 of the ribose ring. NADPH:HN2T forms transient interactions with both A Tyr-69:OH and B Ile-67:O. The side chains of A Asn-67 and B Asn-67 form a tight interaction. A Asn-67: OD1 is bound to B Asn-67: HD21 and B Asn-67:OD1 is bound to A Asn-67:HD21. B Asn-67:HD22 (Figure 21 a).

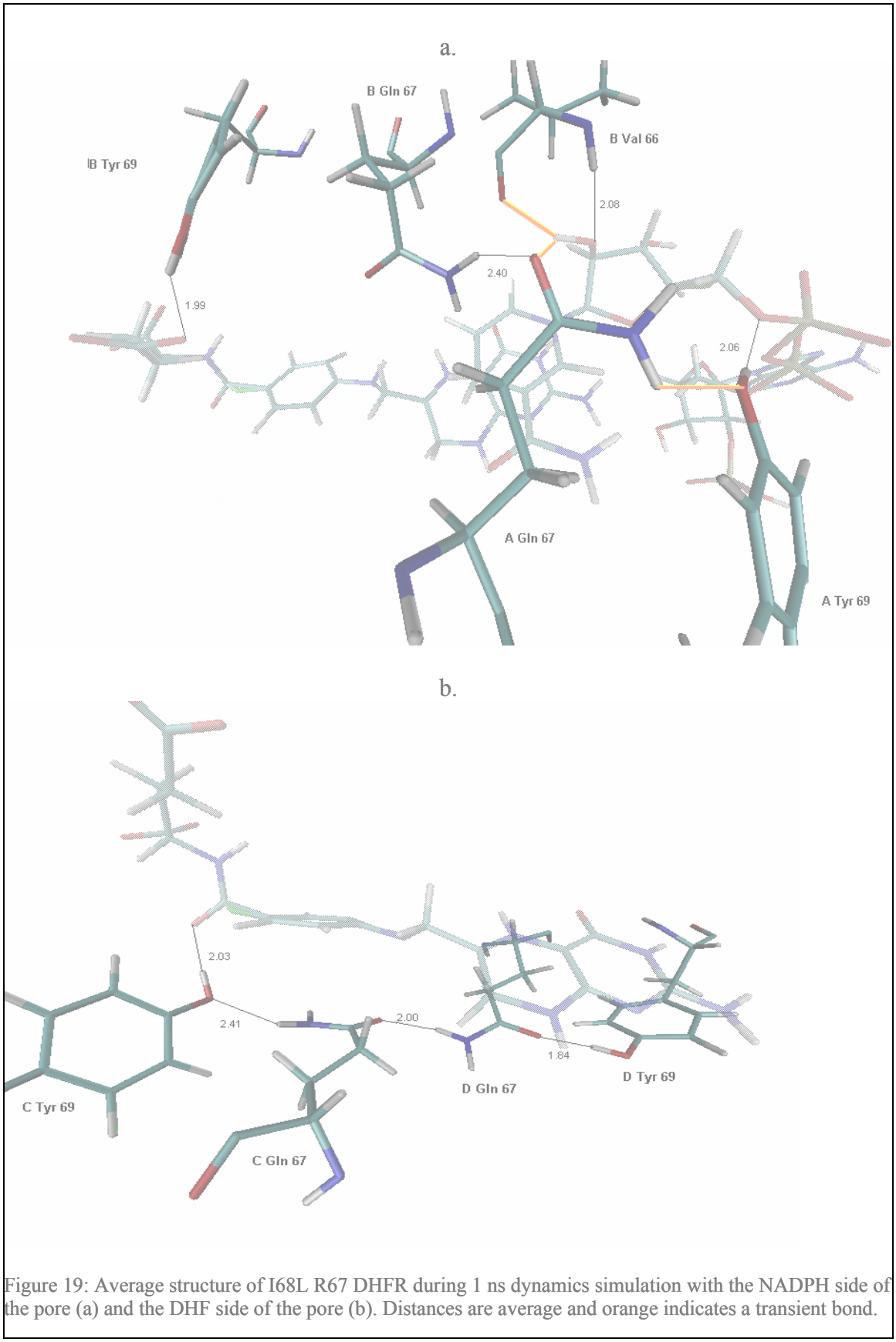


Figure 19: Average structure of I68L R67 DHFR during 1 ns dynamics simulation with the NADPH side of the pore (a) and the DHF side of the pore (b). Distances are average and orange indicates a transient bond.

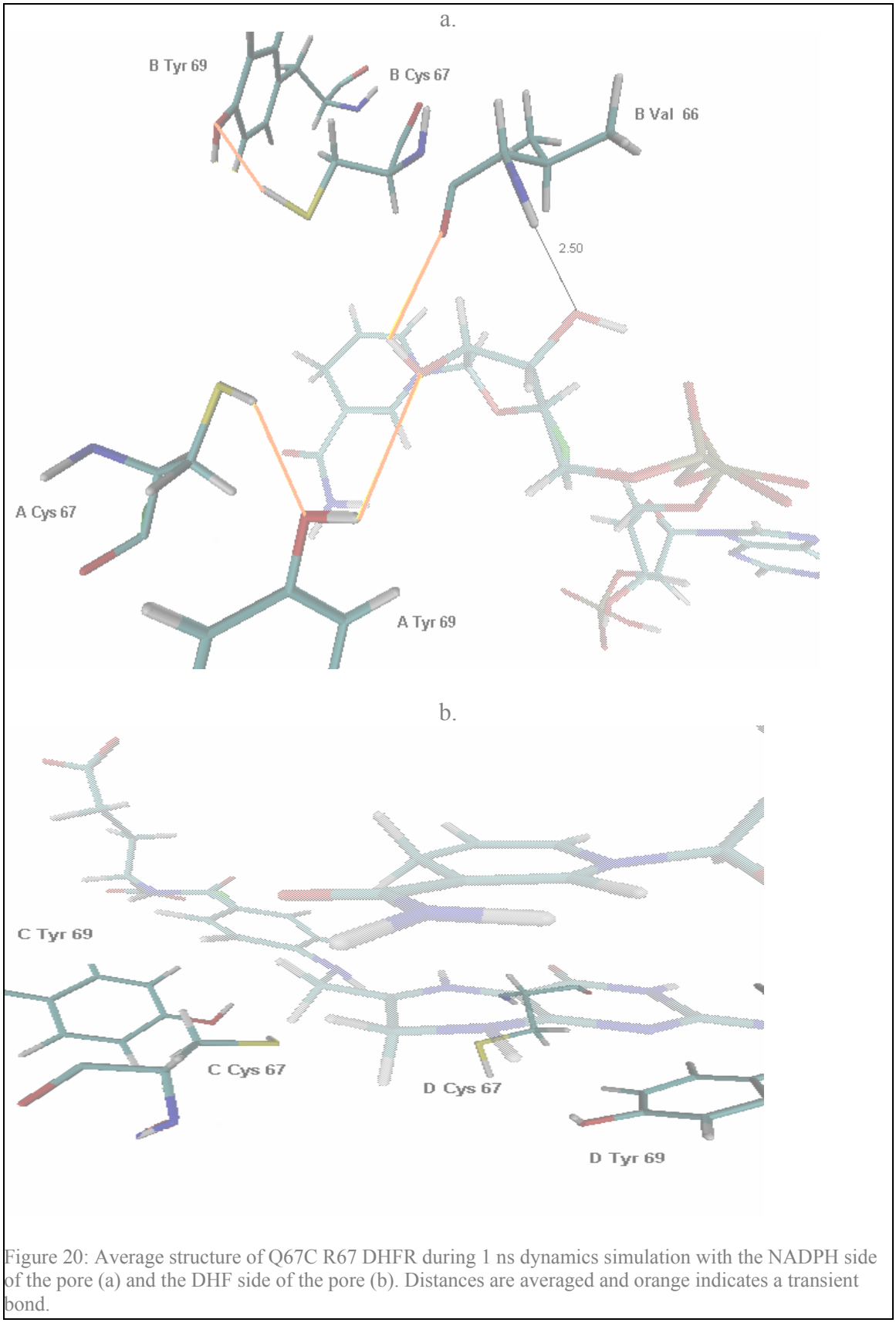


Figure 20: Average structure of Q67C R67 DHFR during 1 ns dynamics simulation with the NADPH side of the pore (a) and the DHF side of the pore (b). Distances are averaged and orange indicates a transient bond.

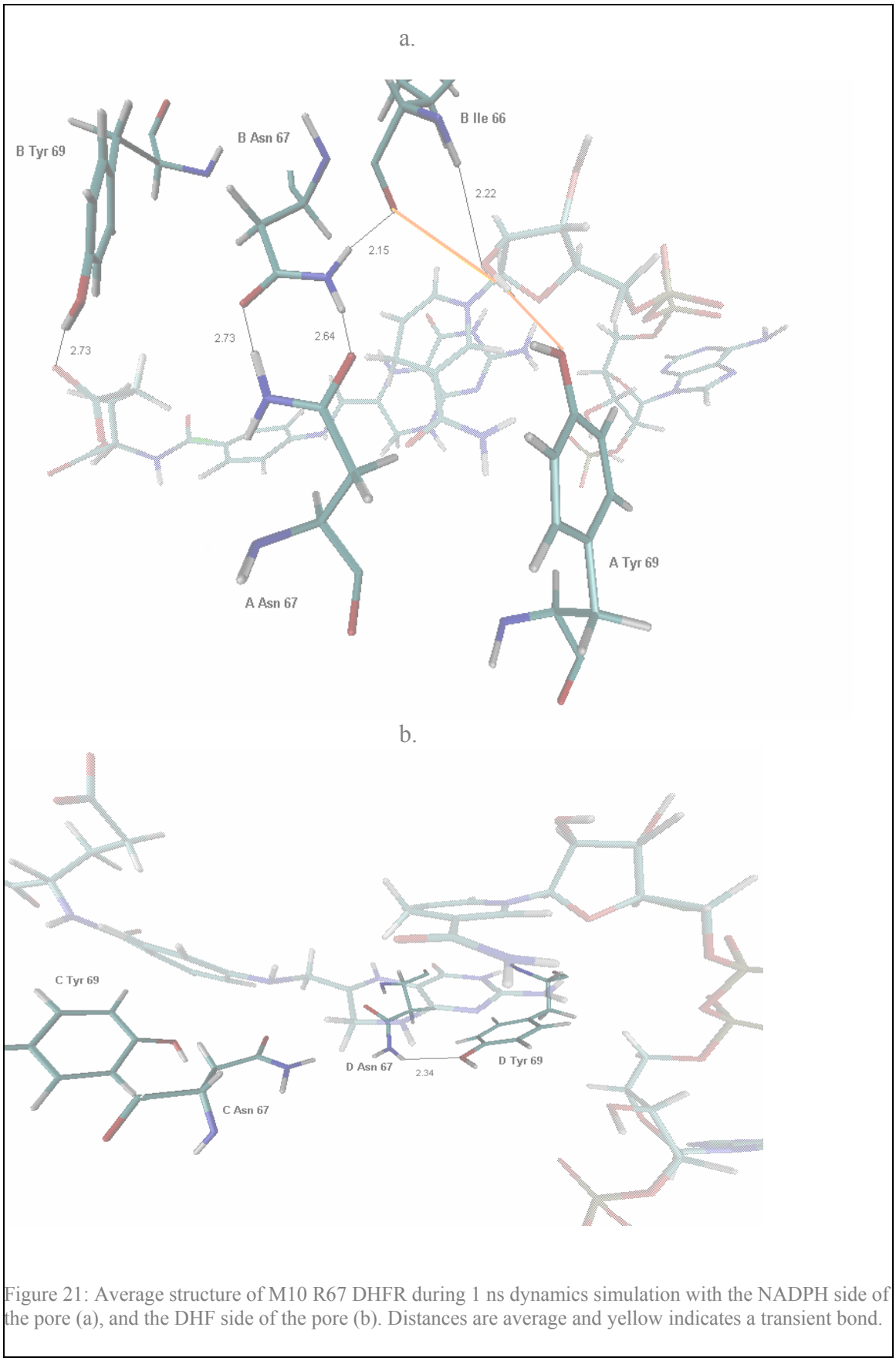


Figure 21: Average structure of M10 R67 DHFR during 1 ns dynamics simulation with the NADPH side of the pore (a), and the DHF side of the pore (b). Distances are average and yellow indicates a transient bond.

On the DHF side of the pore the hydrogen bonding network is less complete. C Asn-67 and D Asn-67 do not interact. D Asn-67:HD22 bonds to D Tyr-69:OH. Of particular interest is the ionic interaction between C Arg-68:HH11 and DHF:OE2 (not shown). This in addition to the interactions with Tyr-69 and Lys-32 could be part of the reason that DHF binding to the M10 mutant is better than the wild-type (K_m = 3.5 uM vs 8.2 for wild-type). A QM/MM potential mean force calculation was performed on the M10 mutant in order to determine if the energy barrier is similar to the wild-type because the k_{cat} values are similar (1.2 s-1 vs 1.3 s-1 for the wild-type) (Schmitzer, 2004). The resulting curve shows more differences than expected from the wild-type. The energy barrier appears to be higher by 1.5 kcal/mole and at 0.2 on the reaction coordinate (Figure 22) . The differences may be an indication that the hydride transfer step is not rate limiting and thus cannot be correlated with $k_{cat.}$

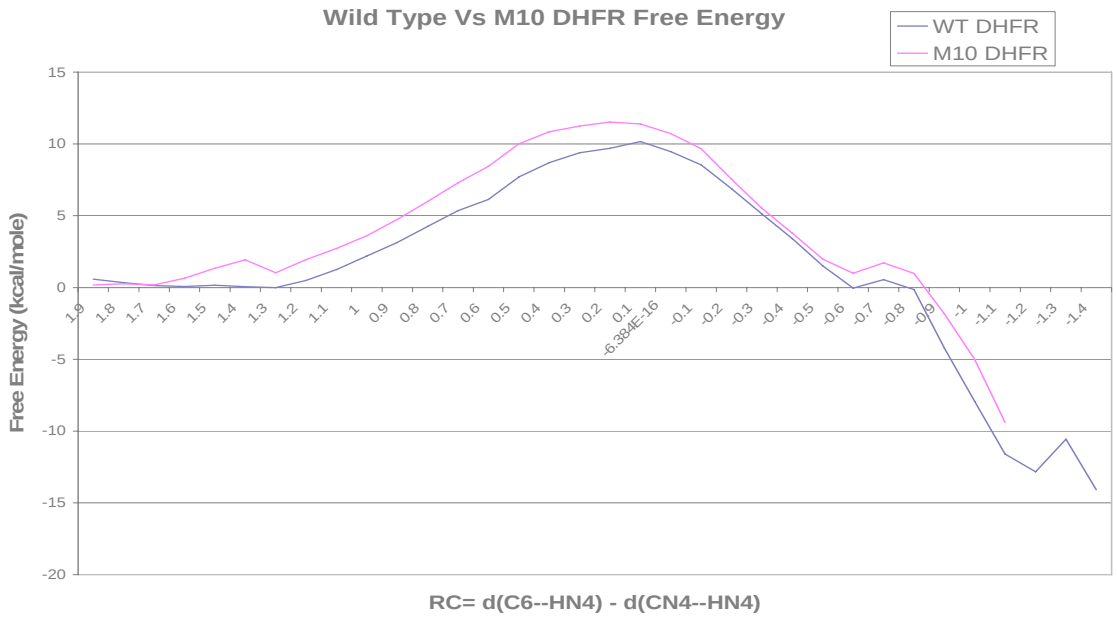


Figure 22: PMF Curve of M10 Vs WT R67 DHFR. The wild-type Potential Mean Force curve is shown in blue and the M10 mutant curve is shown in pink.

Chapter IV
Conclusions

The initial dynamics simulations demonstrate that the tail of DHF, when bound to R67 DHFR is mobile on the nanosecond time scale. This explains the failure of x-ray crystallography in determining its exact position. Without the tail, the ligand does not function as a substrate, (Chopra *et al*, 2006), so the question is what role it plays. More specifically, is there a motion of the tail which correlates with catalysis, or is the function of the tail relegated to binding the substrate to the active site? Mutation studies suggest that an ionic interaction needs to be broken in order to form the transition state (Hicks *et al* , 2004). The PMF studies show that a constrained tail can have some effects on the energy barrier of the hydride transfer step but the differences are small and within the computational error.

The XN angle between the nicotinamide and ribose rings of NADPH remains relatively stable between 55 and 65 degrees and defines a syn orientation. This is due primarily to interactions between the OH groups of the ribose ring and the backbone NH and C=O groups of B Val 66. Another simulation was performed to determine if the angle would be stable at -55 º (anti-conformation). The corresponding model was very unstable on the picosecond timescale (data not shown) and quickly returned closer to its original conformation. The bound syn orientation appears to be a unique feature of the R67 DHFR enzyme (Almarsson and Bruice, 1993). In the case of the I68L mutant, changes in the side chain at position 68 cause movement of the Val 66 residue and thus a change in the orientation of the ribose ring.

The PMF studies indicate that the energy barrier for the hydride transfer step starting with N5 protonated DHF is between 0.3 and 3.4 kcal/mol lower for R67 DHFR than for *E. coli* DHFR depending on which correction method is used. These simulations were complicated by the fact that there was motion of the substrate and cofactor close to the product state that was not accounted for by the reaction coordinate. Because of this, the given orientation of the complex at the product state should be considered less reliable. However, the ground state and transition state confirmations should be considered more valid.

Remarkably the distance between the C6 and CN4 atoms at the transition state was the same for both the *E. coli* DHFR and R67 DHF hydride transfer reactions at 2.67 Å. The C6—HN4—CN4 angle at this point was greater at ~169 degrees for the R67 form of the enzyme than ~157 degrees for the *E. coli* form. This is most likely due to the exo transition state used by *E. coli* DHFR versus the endo transition state used by R67 DHFR. For *E. coli* DHFR, the NADP dihedral angle became more negative going from -7.2 degrees at the ground state to -13.2 degrees at the transition state indicating puckering of the CN4 portion of the nicotinamide ring toward the C6 acceptor on the pteridine ring. This is consistent with the results of other *E. coli* DHFR hydride transfer simulations (Agarwal *et al,* 2002, 2005; Garcia-Vilocia *et al,* 2003, 2004). Curiously, the CN4 portion of the nicotinamide puckers away from the C6 acceptor atom during the transition state of R67 DHFR. Again the endo versus exo conformations dictate what adjustments the substrate and cofactor have to make in order to reach the transition state. In both cases the DHF dihedral becomes more positive at the transition state indicating that the C6

portion of the pteridine ring is puckering upward toward the NC4 donor (see tables 2 and 3).

From an evolutionary standpoint, *E. coli* DHFR is a more advanced enzyme which is better at carrying out the overall reaction including the N5 protonation. R67 DHFR does not have an obvious mechanism for transferring the proton to the N5 nitrogen (Park *et al*, 1997) so the enzyme has to get the proton from solution which is energetically costly. As a result it can only serve to optimize the hydride transfer portion in order to make the reaction easier. R67 DHFR appears to accomplish this by forming a Tyr-Gln-Gln-Tyr hydrogen bonding network which forces the pteridine ring of DHF and the nicotinamide ring of NADPH closer together.

It is necessary to closely examine the wild-type and each of the mutants in order to determine what factors are involved in binding and catalysis. It is initially obvious that the electrostatic interactions which involve Tyr-69 are important for binding as mutation to Phe at this residue causes a sharp increase in K_m. In the I68L and M10 mutants, B Tyr 69 forms an hydrogen bonding interaction with the Glu sidechain of DHF's pABA-Glu tail (see figures 19a and 21a). This indicates that both Lys 32 and Tyr 69 can form interactions with the pABA-Glu tail of DHF. In the I68L mutant C Tyr 69 forms a hydrogen bond with the DHF: O oxygen (see figure 19b). In the I68L, Q67C and M10 mutants D Tyr 69 is shown to hydrogen bond, at least transiently, to portions of the NADPH cofactor (see figures 19a, 20a, 21a). This is consistent with the mutation studies conducted by Stinnett and her associates showing the importance of Tyr-69 in the binding of both DHF and NADPH (Stinnett *et al*, 2004).

These mutant dynamics simulations show that Tyr-69 can form such interactions with both the substrate and cofactor. However the sharp decrease in K_m in the Q67H mutant (not studied here) indicates that stacking interactions between the pteridine and nicotinamide rings along with enzyme also play a role in binding. The Q67H mutation results in binding of both DHF and NADPH that is over one hundred times tighter then that of the wild-type (Smiley *et al*., 2002). Formation of a contractile ring in the center of the pore by the Tyr-Gln-Gln-Tyr hydrogen bonding network in the wild-type could be important for maximizing favorable contact between substrate and cofactor. In the wild-type and the M10 mutant, which have similar K_m values (see table 1), the residues at position 67 of the A and B chains form a stable interaction above the nicotinamide ring. In the Y69F mutant a similar interaction is formed though not as part of a tight hydrogen bonding network because there is not a Tyr residue there to complete the network. The binding for both DHF and NADPH to the Y69F mutant is much looser then would be expected. However the k_{cat} is significantly higher then the wild-type at 2.5 s^{-1} (see table 1).

Determining the factors involved in catalysis is much more challenging due to the difficulty in differentiating the hydride transfer step from the N5 protonation step. It is not agreed upon weather the two steps happen sequentially or concurrently. In *E. coli* DHFR experiments involving NADPD show product release to be the rate limiting step at pH 7.0 but at pH 8.1 the hydride transfer step becomes rate limiting and demonstrates an isotope effect (Karginov *et al*., 1997). For R67 the pH profile which displays higher activity at low pH suggesting that there the enzyme does not possess a proton donor (Park et al., 1997). In the current study N5 protonated DHF was used as the substrate for both

E. coli and R67 DHFR simulations. The results show that, for this isolated portion of the reaction, the energy barrier is slightly lower for the R67 form. This indicates that the ability of *E. coli* DHFR to effectively protonate the N5 nitrogen of DHF is a big part of the reason it has a much higher activity then the R67 form.

Therefore it would be improper to draw direct relationships between k_{cat} values and the energy barrier for the hydride transfer step of the reaction. The PMF free energy curve for the M10 mutant versus the wild-type shows that while both have similar k_{cat} values (1.2 s^{-1} vs 1.3 respectively) the M10 mutant has a higher energy barrier by 1.5 kcal/mol. Generally speaking, the amide groups interactions at position 67 (Gln in the wild-type, Asn in the M10 mutant) appear to play an important role in catalysis. The isoforms studied here which have highest catalytic rates (the wild-type, Y69F and M10 mutants) all have amide groups that bind in a regular orientation above the nicotinamide ring (see figures 17a, 18a, and 21a). The Q67C mutant does not have an amide group and its k_{cat} value is 0.1 s^{-1} (see figure 20a). In the I68L mutant, the two Gln residues above the nicotinamide ring are at an angle with respect to each other (see figure 19a). The k_{cat} value for the I68L mutant is 0.32 s^{-1}. It should be noted that several of the functional mutants discovered in the Schmitzer *et al's* 2004 mutagenesis studies do not have an amide group at the 67 position. The structural characteristics of these mutants were not determined nor were they part of this study.

DHFR performs an essential role in most living things. Human intervention targeting the *E coli* form of the enzyme led to the emergence of the much simpler, yet evolutionarily distinct, R67 form of the enzyme. The symmetry of the enzyme and the fact that it is a homotetramer allow it to introduce dramatic effects with a single mutation.

However, some multiple mutations can increase the activity of the enzyme to above the activity of the wild-type and still have resistance to trimethoprim (Schmitzer *et al*., 2004) This implies that the R67 form DHFR is a simple solution to a complex problem.

List of References

List of References

Agarwal, P. K., Billeter, S. R., Rajagopalan, P. T., Benkovic, S. J., Hammes-Schiffer, S. (2002) Proceedings of the National Academy of Sciences USA 99, 2794-2799.

Agarwal, P. K. (2005) Journal of the American Chemical Society 127, 15248-15256.

Almarsson, O., Bruice, T. C. (1993) Journal of The American Chemical Society 115, 2125-2138.

Alonzo, H., Gillies, M. B., Cummins, P. L., Bliznyuk, A. A., Gready, J. E. (2005) Computer-Aided Molecular Design 19, 165-187.

Alonzo, H., Gready, J. E. (2006) Trends in Microbiology 14 (5), 236-242.

Andres, J., Moliner, V., Safront, V. S., Domingo, L. R., Picher, M. T., Krechl, J. (1996) Bioorganic Chemistry 24, 10-18.

Bradrick, T. D., Beechem, J. M., Howell, E. E. (1996) Biochemistry 35, 11414-11424.

Brito, R. M. M., Rudolph, F. B., Rosevear, P. R. (1991) Biochemistry 30, 1461-1469.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. (1983) Journal of Computational Chemistry 4, 187-217.

Brooks, C. L., III, Brunger, A. Karplus, M. (1985) Biopolymers 24, 843-865.

Cannon, W. R., Garrison, B. J., Benkovic, S. J. (1997) Journal of the American Chemical Society 119, 2386-2395.

Castillo, R., Andres, J., Moliner, V. (1999) Journal of The American Chemical Society 121, 12140-12147.

Chopra, S., Lynch, R., Kim, S., Jackson, M., Howell, E. E. (2006) Biochemistry 45, 6596-6605.

Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., Karplus, M. (2001) Journal of Physical Chemistry B 105, 569-585.

Cummins, P. L., Gready. J. E (1998) Journal of Computational Chemistry 19 (8), 977-988.

Deng, H., Callender, R., Howell, E. (2001) The Journal of Biological Chemistry 276 (52), 48956-48960.

Ferrer, S., Silla, E., Tunon, I., Marti, S., Moliner, V. (2003) Journal of Physical Chemistry B 107, 14036-14041.

Field, M. J., Bash, P. A. and Karplus, M. (1990) A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations, J Comput. Chem. 11, 700-733.

Fierke, C. A., Johnson, K. A., Benkovic, S. J. (1987) Biochemistry 26, 4085-4092.

Garcia-Vilocia, Gao, J., Karplus, M. (2004) Science 303, 186-195.

Garcia-Vilocia, M., Truhlar, D. G., Gao, J. (2003) Biochemistry 42, 13558-13575.

Hicks, S. N., Smiley R. D., Hamilton, B., Howell, E. E. (2003) Biochemistry 42, 10569-10578.

Hicks, S. N., Smiley R. D., Stinnett, L. G., Minor K. H., Howell, E. E. (2004) The Journal of Biological Chemistry 279 (45), 46995-47002.

Howell, E. E. (2005) ChemBioChem 6, 590-600.

Howell, E. E., Shukla, U., Hicks, S. N., Smiley, R. D., Kuhn, L. A., Zavodszky, M. I. (2001) Journal of Computer-Aided Molecular Design 15, 1035-1052.

Jackson, M., Chopra, S., Smiley, R. D., Maynord, P. O., Rosowsky, A., London, R. E., Levy, L., Kalman, T. I., Howell, E. E. (2005) Biochemistry 44, 12420-12433.

Karginov, V. A., Mamaev, S. V., An, H., Van Cleve, M. D., Hecht, S. M., Komatsoulis, G. A., Abelson, J. N. (1997) Journal of the American Chemical Society 119, 8166-8176.

Kobrak, M. N. (2003) Journal of Computational Chemistry 24, 1437-1446.

Kumar, S., Bouzida, D., Swendsen, R. H., et al. (1992) Journal of Computational Chemistry 13 (8), 1011-1121.

Levings, R. S., Lightfoot, D., Elbourne, L. D. H., Djordjevic, S. P., Hall, R. M. (2006) Antimicrobial Agents and Chemotherapy, 2863-2865.

MacKerell, A. D., Jr., Bashford, D., Bellott, M., Bunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen D. T., Prodhom, B., Reiher, W. E., III, Roux B., Schlenkrich, M., Smith, J. C., Stote, R., Staub J., Watanabe, M., Wiorkiewicz- Kuczera, J., Yin, D., Karplus, M. (1998) Journal of Physical Chemistry B 102, 3586-3616.

Narayana, N., Matthews, D. A., Howell, E. E., Xuong, N. H. (1995) Nature Structural Biology 2, 1018-1025.

Park, H., Zhuang, P., Nichols, R., Howell, E. E. (1997) Journal of Biological Chemistry 272 (4), 2252-225.

Rychaert, J. P., Ciccotti, G., Berendsen, H. J. C. (1977) Journal of Computational Physics 23, 327-341.

Sawaya, M. R., Kraut, J. (1997) Biochemistry 36, 586-603.

Schmitzer, A. R., Lepine F., Pelletier, J. N. (2004) Protein Engeneering Design and Selection 17 (11), 809-819.

Smiley, R. D., Stinnett, L. G., Saxton, A. M., Howell, E. E., (2002) Biochemistry 41, 15664-15675.

Stinnett, L. G., Smiley, R. D., Hicks, S. N., Howell, E. E. (2004) The Journal of Biological Chemistry 279 (45), 47003-47009.

Stokes, H. W., Hall, R. M. (1989) Molecular Microbiology 3 (12), 1669-1683.

Strader, M. B., Chopra, S., Jackson, M., Smiley, R. D., Stinnett, L., Wu, J., Howell, E. E. (2004) Biochemistry 43, 7403-7412.

Strader, M. B., Smiley, R. D., Stinnett, L. G., VerBerkmoes, N. C., Howell, E. E. (2001) Biochemistry 40, 11344-11352.

Thorpe, I. F., Brooks, C. L. (2003) Journal of Physical Chemistry B 107, 14042-14051.

Torrie, G. M., Valleau, J. P. (1974) Chemical Physics Letters 28, 578-581.

Wang, L., Tharp, S., Selzer, T., Benkovic, S. J., Kohen, A. (2006) Biochemistry 45, 1383-1392.

Vita

Robert Beahm was born in 1978 in Oak Ridge Tennessee. He graduated from Oak Ridge High School in 1997 and went on to the University of Tennessee, Knoxville where he received a B.A. in Biochemistry in 2001. In 2003 he received an A.S. in computer programming from Pellissippi State Technical Community College also in Knoxville. He is currently working toward a masters degree at the University of Tennessee.