

PCCP

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: H. Abdizadeh, Y. T. Tamer, O. Acar, E. Toprak, A. R. Atilgan and C. Atilgan, *Phys. Chem. Chem. Phys.*, 2017, DOI: 10.1039/C7CP01458A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [author guidelines](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the ethical guidelines, outlined in our [author and reviewer resource centre](#), still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Increased Substrate Affinity in the Escherichia Coli L28R Dihydrofolate Reductase Mutant Causes Trimethoprim Resistance

Haleh Abdizadeh¹, Yusuf Talha Tamer², Omer Acar¹, Erdal Toprak², Ali Rana Atilgan¹, Canan Atilgan^{1*}

¹Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey

²Green Center for Systems Biology, UT Southwestern Medical Center, Texas, USA

Abstract

Dihydrofolate reductase (DHFR) is a ubiquitous enzyme with an essential role in cell metabolism. DHFR catalyzes the reduction of dihydrofolate to tetrahydrofolate, which is a precursor for purine and thymidylate synthesis. Several DHFR targeting antifolate drugs including trimethoprim, a competitive antibacterial inhibitor, are therefore developed and clinically used. Evolution of resistance against antifolates is a common public health problem rendering these drugs ineffective. To combat against the resistance problem, it is important to understand resistance-conferring changes in the DHFR structure and accordingly develop alternative strategies. Here, we structurally and dynamically characterize *Escherichia coli* DHFR in its wild type (WT) and trimethoprim resistant L28R mutants in the presence of the substrate and its inhibitor trimethoprim. We use molecular dynamics simulations to determine the conformational space, loop dynamics and hydrogen bond distributions at the active site of DHFR for the WT and the L28R mutant. We also report their experimental k_{cat} , K_m , and K_i values, accompanied by isothermal titration calorimetry measurements of DHFR that distinguish enthalpic and entropic contributions to trimethoprim binding. Although mutations that confer resistance to competitive inhibitors typically make enzymes more promiscuous and decrease affinity to both the substrate and the inhibitor, strikingly, we find that the L28R mutant has a unique resistance mechanism. While binding affinity differences between the WT and the mutant for the inhibitor and the substrate are small, the newly formed extra hydrogen bonds with the aminobenzoyl glutamate tail of DHF in the L28R mutant leads to increased barriers for the dissociation of the substrate and the product. Therefore, the L28R mutant indirectly gains resistance by enjoying prolonged binding times in the enzyme-substrate complex. While this also leads to slower product release and decreases the catalytic rate of the L28R mutant, the overall

effect is to maintain a sufficient product formation rate. Finally, the experimental and computational analyses together reveal the changes that occur in the energetic landscape of DHFR upon the resistance-conferring L28R mutation. We show that the negative entropy associated with the binding of trimethoprim in WT DHFR is due to water organization at the binding interface. Our study lays the framework to study structural changes in other trimethoprim resistant DHFR mutants.

Keywords: Antimicrobial resistance, selection pressure, morbidostat, dissociation constant, molecular dynamics.

Introduction

Multiple conformers of the enzyme scaffolds detected upon complex formation in different stages of a kinetic reaction cycle are triggered by the flexibility of the active site, local motions of the loops or allosteric response initiated by a distal perturbation.^{1, 2} Correlated fluctuations over a wide range of time/displacement scales that are involved in the dynamic shifts between feasible states link the collective motions to the effective chemical reactions of the enzymes in their wild type and mutated forms.^{3, 4} Conformational plasticity and allosteric regulation are essential contributors to enzyme stability, substrate specificity and catalytic efficiency.⁵⁻⁷ A comparative study of enzyme dynamics in the presence of its substrate and inhibitor is of key importance to reveal the changes in the dynamics associated with drug induced enzyme inhibition.⁸

Here we study *E. Coli* dihydrofolate reductase (DHFR) in the presence of its native substrate (DHF), and a competing inhibitor (trimethoprim). We investigate both the wild type (WT) DHFR and the L28R DHFR mutant, which is one of the most beneficial resistance-conferring mutations in the WT background.⁹ DHFR catalyzes the stereospecific reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) by utilizing nicotinamide adenine dinucleotide phosphate (NADPH) as the cofactor. The reduction involves a proton transfer from water, followed by hydride transfer where NADPH protonates DHF. The reaction is essential for maintaining cellular surplus of THF and its derivatives, vital precursors for the biosynthesis of several DNA bases and amino acids.¹⁰ Steady-state turnover of the enzyme under the cellular concentrations of substrate and the cofactor is catalyzed through five kinetic intermediates, two of which are the Michaelis complex (DHFR/DHF/NADPH) and the complex with the product (DHFR/THF/NADP⁺).¹¹ Following the hydride transfer, the release of the oxidized cofactor (NADP⁺) and rebinding of NADPH assists the release of the product (THF), which is the rate-

limiting step.¹² DHFR is always primed and its apo state is most likely short-lived and hence not observable under physiological conditions, contributing to its high catalytic efficiency.

Because of its essential functions in cell maintenance and growth, DHFR is used as the target for several important anticancer (e.g. methotrexate), antimalarial (e.g. pyrimethamine), and antimicrobial (e.g. trimethoprim) drugs. Trimethoprim (TMP) is a well-studied competitive inhibitor of bacterial DHFRs and is extensively used across the world because of its relatively lower cost. Therefore, resistance to TMP in pathogenic bacteria is very common. TMP resistance generally emerges due to spontaneous mutations that lead to amplification or overexpression of the gene encoding DHFR and multiple amino acid replacements in the DHFR enzyme.¹³⁻¹⁵

Much has been discussed on the working principles of *E. Coli* DHFR enzyme in terms of its kinetics¹⁶ and molecular dynamics.^{4, 17-20} Analysis of several complexes of DHFR with various ligands provides insight to the coupling between the short-lived dynamics and functional motions of the enzyme complexes in various states of the catalytic cycle. For example, the time scale of loop motions in complexes of *E. coli* DHFR at different stages of the kinetic cycle has been measured by nuclear magnetic resonance (NMR) experiments.^{18, 21, 22} Further NMR studies have elucidated how local loop fluctuations are coupled to kinetics and thermodynamics of DHFR.^{18, 23} Particularly, the amide backbone dynamics alterations represent different physiologically relevant time scales for the enzyme kinetics.^{21, 24} Meanwhile, molecular dynamics (MD) simulations explain the molecular level effects of different ligands on the network of coupled motions between spatially separated regions of DHFR.^{19, 25} Previously reported 10 ns long MD simulations suggest that the secondary structure arrangements of the enzyme in complex with different ligands are intact, while flexible loops move in large conformational changes.^{20, 26} Kinetic measurements show multiple mutations in and distal to the active site significantly affect

the conformational space sampled by the enzyme.²⁷ Long-range perturbations on the enzyme structure alter the kinetic parameters followed by reduced or enhanced activities. Non-additive effects of two different mutations provide another example of long range coupling between distant regions of DHFR.²⁰

While there is great amount of experimental and theoretical information available for DHFR structure and function,^{19, 25, 28, 29} despite its clinical importance, we currently do not have a dynamical/mechanical understanding of how the conformation of the enzyme is altered as a result of the binding of an inhibitor such as TMP. Previous MD simulations of *E. coli* DHFR report unbiased sampling of the unfolding pathway at temperatures up to 600 K.³⁰ Therein, DHFR in complex with NADP and folate is studied via MD simulations of one nanosecond length at increasing temperatures. MD simulations on the order of 10 ns of the wild type (WT) *E. coli* DHFR and different mutants also investigate the sampled conformational space and find the effect of mutations on the correlated motions, hydrogen bond distributions and hydride transfer reaction under the non-additive effects of mutations.^{20, 26} Combined with empirical valance bond calculations, MD simulations on the mutated DHFR structures show that the overall conformation of the enzyme is largely maintained while altering the hydride transfer reaction free energy barriers.²⁹

We have previously reported the dynamical properties of proteins in different environments using MD simulations on the order of hundreds of nanoseconds.^{7, 31-33} We have shown that local and global conformational changes delineate the molecular basis of functional pathways of various proteins (human and bacterial transferrin), calmodulin and subtilisin serine protease.^{7, 31-34} Using fluctuation patterns of residues, we have successfully predicted that distal site communications contribute to the stability and catalytic activity of these proteins.⁷ Recognizing the significance of

dynamical properties and flexibility in enzymatic reactions, here we use a similar approach to study WT DHFR and its L28R mutant bound to dihydrofolate (DHF) or trimethoprim (TMP) through MD simulations. L28R mutant is a common resistance-conferring mutation observed in experimental evolution experiments that are carried out in the morbidostat, a continuous culture device designed to evolve pathogenic bacteria under dynamically sustained drug induced selection pressure.^{14, 35} We explain the coupling between enzyme fluctuations and catalytic function by reporting the details of the molecular motions of DHFR binding to its substrate (DHF) and inhibitor (TMP), sampling the equilibrium conformations on time scales up to 200 ns. A comparative examination of the structures provides a dynamical picture for conformational changes in the presence of various ligands.

To explore the energetics of DHF and trimethoprim binding to DHFR, we analyze the free energy changes in response to ligand dissociation from the enzyme. The results are supported by TMP and substrate binding (k_{cat} , K_m , and K_i) measurements for WT and the mutant. We present a structural view of how trimethoprim binding inhibits DHFR function, which may provide guidelines for future drug design efforts. We also report CharmM22 parameters for neutral and protonated trimethoprim for the first time in the literature.

Materials and methods

Protein structures. Numerous structures have been reported for complexes of the enzyme from various organisms, in the presence of a wide range of ligands, including the substrate, product, and cofactor, as well as inhibitors.^{11, 36, 37} *E. coli* DHFR is a small (~18 kD) protein formed from a mixture of eight stranded β -sheets and four contiguous α -helices (see Figure 1a).¹¹ The protein is divided by the active site cleft into two structural subdomains: the adenosine binding

subdomain and the major subdomain. The adenosine binding subdomain (residues 38–88) provides the binding site for the adenosine moiety of the cofactor (NADPH). The major subdomain consists of ~100 residues from the N- and C-termini and is dominated by a set of three loops on the ligand binding face that surrounds the active site. These loops construct approximately half of the major subdomain; hence the name “loop subdomain.” These loops are termed as M20 (residues 9–24), FG (residues 116–132), and GH (residues 142–150) loops. In addition, the CD loop (residues 64–71) on the adenosine binding subdomain is also deemed important for DHFR function. The M20 loop is located directly over the active site, protecting it from the solvent, and is primarily responsible for regulating the active site.¹¹ Three well-defined conformations of the M20 loop in the crystal are the open, occluded, and closed states.^{11, 38} In our work, the initial structure is that having the PDB code 1rx2 with the closed M20 loop conformation.¹¹

Molecular dynamics simulations. The NAMD package is used to model the dynamics of the protein-water systems.³⁹ Solvation is achieved via the VMD 1.9.1 program solvate plug-in version 1.2.⁴⁰ The protein is soaked in a cubic solvent box such that there is at least a 10 Å layer of solvent in each direction from any atom of the protein to the edge of the box. Ionic strength in all the simulations is 150 mM, maintained by adding the suitable number of K⁺ and Cl[−] ions for the box size of the pertaining simulation. The CharmM22 all-atom force field is used to model the protein and the TIP3P potential for water.⁴¹ We have made use of the 5-protonated 7,8-dihydrofolate force field parameters reported in the literature,⁴² while TMP parameters are developed in this study by using the Force Field ToolKit (fftk) plugin⁴³ implemented in VMD as described in the next subsection.

During the simulations, periodic boundary conditions are imposed on the simulation boxes that have $60 \times 67 \times 58 \text{ \AA}^3$ dimensions. Long range electrostatic interactions are calculated by the particle mesh Ewald method,⁴⁴ with a cutoff distance of 12 \AA and a switching function at 10 \AA . The RATTLE algorithm⁴⁵ is applied and a time step size of 2 fs in the Verlet algorithm is used. Temperature control is carried out by Langevin dynamics with a dampening coefficient of $5/\text{ps}$. Pressure control is attained by a Langevin piston. All systems are first subjected to 10000 steps of energy minimization with the conjugate gradients algorithm. The resulting structures are then run in the NPT ensemble at 1 atm and 310 K until volumetric fluctuations are stabilized and the desired average pressure is maintained.

MD simulation of the ternary complex of the DHF bound WT system (WT^{DHF}) is constructed based on the crystallographic structure with PDB code 1rx2 (Figure 1a).¹¹ DHFR is complexed with folate and oxidized NADP (NADP^+) in this native form. In a separate set of MD simulations, we study the effect of trimethoprim binding in its unprotonated (TMP) and ground state (TMP^+) on the WT DHFR conformation (WT^{TMP} and WT^{TMP^+}). We protonate NADP and folate so that the former is in reduced form (NADPH) and the latter is 5-protonated 7,8-dihydrofolate to model the stable state prior to the hydride transfer step.

Since there are no crystal structures of *E. coli* DHFR with TMP, we have docked the inhibitor based on the coordination of equivalent residues of the trimethoprim binding region of *Staphylococcus Aureus* DHFR (PDB code: 2w9g) that has 36.6% sequence identity.³⁶ We have used AutoDock 4.2 for this purpose.^{46, 47} The center of mass of the docked molecule was selected from the coordinates of the C_α atoms of the TMP contacting residues in *Staphylococcus Aureus* DHFR and is consistent with DHF binding site of *E. Coli* dihydrofolate binding complex. Accordingly, we choose residues 5, 6, 7, 27, 28, 30, 31, 50, 54, 94, 100, 111, and 113. TMP was

enclosed in a box with number of grid points $40 \times 40 \times 40$ with a grid spacing of 0.375 Å. Lamarckian genetic algorithms, as implemented in AutoDock were employed to perform the calculations. All other parameters were set to their default values. For each of the docking cases, the docked conformation having the lowest Autodock scoring function value was selected as the binding mode.

For MD simulations of L28R mutant of DHF, TMP and TMP^+ bound forms of DHFR (L28R^{DHF} , L28R^{TMP} , and $\text{L28R}^{\text{TMP}^+}$ respectively), we mutated the WT structures in silico via BIOVIA Discovery Studio 4.0 package.⁴⁸ The minimization and equilibration was performed as described for the WT systems. All MD simulations are 210 ns long, with the first 10 ns discarded for equilibration. Simulations for the WT cases were repeated to confirm reproducibility of the results.

Determination of trimethoprim force field parameters. We parameterized trimethoprim molecules in two protonation states (TMP and TMP^+) using Force Field ToolKit (fftk) plugin implemented in VMD.⁴³ For necessary quantum mechanical calculations, Gaussian package was utilized.⁴⁹ We followed the protocol developed by Mayne et al.⁴³ to calculate the force field parameters. As the initial three-dimensional structure of the inhibitor, we adopted the coordinates for TMP bound to *Staphylococcus aureus* DHFR structure. Missing parameter determination, geometrical optimization and bond, angle and dihedral potential optimizations were done by quantum mechanical calculations. The resulting parameters are listed in supplementary text S1 and S2.

Quantifying ligand motions by contributions from principal axes. The mobility of the equilibrium structures of trimethoprim and DHF is quantified by calculating the region spanned

by the principal axes of pyrimidine tail of trimethoprim and p-aminobenzoyl glutamate tail of DHF. For TMP, a vector is defined between the NA2 and C7 atoms (NA2-C7 vector) (see Figure S1 for atom labeling scheme), while for DHF, a vector is defined between N10 and CD atoms (N10-CD vector) (see Figure S2 for atom labeling scheme).

The $A_{3 \times N}$ matrix of the NA2-C7 and N10-CD vectors are formed where N denotes number of snapshots collected during the 210 ns of MD simulation. The three principal axes of the molecules are calculated and the projections of the molecule along these axes are considered as follows:

$$\mathbf{A}_{3 \times N} = \mathbf{U}_{3 \times 3} \mathbf{W}_{3 \times 3} \mathbf{V}_{3 \times N}^T \quad (1)$$

Contribution fraction of the shape along the three principal axes is evaluated from the size of the eigenvalues belonging to each principal axis, per equation 1. In this context, equal contributions along the three axes represent spanning a spherical region, a large contribution to one axis and insignificant contributions to the remaining two represent a rod like shape, and significant contributions to two axes, and a small contribution to the third represent a planar shape. Examples are provided in Figure S3, where the regions spanned by the NA2-C7 unit vector are plotted for the WT and the L28R mutants with the contribution fractions of the major axis being 0.42 and 0.69, respectively.

Steered molecular dynamics (SMD) simulations. The constant velocity SMD algorithm implemented in NAMD package³⁹ was used by fixing the C_α atoms of DHF/TMP/TMP⁺ binding region in WT and L28R mutant forms and external force was applied to the ligand (DHF or trimethoprim). For SMD simulations, several snapshots from the equilibrated enzyme described in the previous section is soaked in a cubic solvent box such that there is at least a 40 Å layer of

solvent in each direction from any atom of the protein towards the edge of the box, leading to dimensions of $124 \times 125 \times 127 \text{ \AA}^3$. The larger water box is used to dissociate ligands from the protein completely and reach a plateau in the cumulative work profile. The ionic strength is set to 150 mM by adding K^+ and Cl^- ions. DHFR was fixed in 3D space using ligand binding site as the reference to prevent it from translations and rotations during the pulling experiment. The ligand was pulled with a harmonic constraint of $20 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$ and a constant velocity of $0.05 \text{ \AA} \cdot \text{ps}^{-1}$. The large value of the spring constant satisfies the stiff spring approximation⁵⁰ and the pulling speed is slow enough to reduce the non-equilibrium effects. With this constraint, no conformational change is detected in the ligands in the SMD simulation. The pulling direction was defined based on the vector from the center of mass of the fixed atoms in DHFR toward the center of mass of the ligand. The pulling process was applied to the equilibrated DHFR positions at 10, 50, 90, 130, 170 and 210 ns frames as the starting structures. Each SMD simulation lasted 0.8 ns in the NPT ensemble with periodic boundary conditions, the Langevin thermostat and isotropic position scaling, at 310 K and 1 bar. For each initial structure, ten independent SMD simulations were performed. Therefore, the total number of trajectories to define the work done by the external force to dissociate the ligands from their respective active sites in each system is 60. The force exerted on the ligand is defined by:

$$\mathbf{F}(t) = k\{vt - [\mathbf{r}(t) - \mathbf{r}_0] \cdot \mathbf{n}\} \quad (2)$$

where k is the spring constant, v is the constant velocity of pulling, \mathbf{r}_0 and $\mathbf{r}(t)$ are the ligand center of mass position at initial and current time t respectively, \mathbf{n} is unit vector for the pulling force.

Free energy difference estimator in terms of the potential of mean force, defined as the free energy profile along a reaction coordinate was initially calculated by Jarzynski's equality to relate the probability distribution of non-equilibrium work and equilibrium free energy differences obtained during the constant velocity SMD processes:⁵¹

$$\exp\left(\frac{-\Delta G}{k_B T}\right) = \langle \exp\left(\frac{-W}{k_B T}\right) \rangle \quad (3)$$

where W is the work performed, k_B is Boltzmann's constant, T is the bulk temperature, and ΔG is the free energy difference. Angle brackets denote the ensemble average over the independent constant velocity SMD trajectories for each starting position. The bias corrected Jarzynski free energy difference estimator (the final ΔG) was then obtained considering cumulative work of the near-equilibrium trajectories whose probability distribution is checked to fit Gaussian. Using the dissipated work, the difference between the average work and the free energy difference estimator, large statistical errors and biases evolving due to the sensitivity of the Jarzynski's approach to rare events was calculated as described by Gore et al.⁵² (see ref⁵³ for details of our implementation). The error bars with these choices are in the range of 0.6 – 1.0 kcal/mol. For an exact calculation of the standard free energy changes, hence K_d values, corrections for relaxing the constraints imposed on the binding site residues and the pulled molecule must be considered. However, this would further increase the error bars. We therefore only report the $\Delta\Delta G$ values where these corrections will cancel, and we discuss the relative stabilities between the WT and the mutant enzyme in the text.

Purification of DHFR. *E. coli* *FolA* gene was cloned into the pet24a-*folA*-kanR plasmid including T7 promoter and six His-tags on the C terminus is constructed (Full map of the plasmid is provided in Figure S4). *E. coli* BL21-DE3 strain is chemically transformed with pet24a-*folA*-

kanR plasmid and grown overnight in LB + Kanamycin (50 $\mu\text{g/mL}$) media. 100 μL of o/n culture is incubated in 400 mL Terrific Broth media (catalog #: Fisher Scientific BP2468) until OD600 reaches 0.5-0.8. Then 250 μM IPTG is added to induce the production of DHFR o/n. Induced cells are sonicated and the DHFR is purified using Ni-NTA affinity column. After purification proteins are dialyzed in the storage buffer (50 mM Tris HCl pH 8.0, 300 mM NaCl, 1% glycerol). Concentrations of dialyzed proteins are measured by absorption at 280 nm using the molar extinction coefficient of 33585 $\text{M}^{-1}\text{cm}^{-1}$.

Isothermal Titration Calorimetry. The binding energetics of trimethoprim to WT and L28R DHFR were measured using an MCS titration calorimeter (MicroCal iTC200 Malvern Inc.) Proteins were dialyzed exhaustively before titration; ~ 500 μL of protein sample in 2 L of storage buffer including 50mM Tris HCl pH 8.0, 300 mM NaCl and 1% Glycerol in one dialysis step overnight (~ 12 -16h). TMP solution is prepared by dissolving it in dialysis buffer to minimize the heats of dilution when injection happens with the concentration of 0.2mg/mL (688.94 μM). In each experiment, 25 μM of protein + 250 μM NADPH complex are titrated with 250 μM of TMP. Protein concentrations were determined by absorbance at 280 nm. Each measurement was carried out in triplicate.

Results and discussion

We analyze the structural changes in the WT (WT^{DHF} , WT^{TMP^+} and WT^{TMP}) and L28R mutants (L28R^{DHF} , $\text{L28R}^{\text{TMP}^+}$ and L28R^{TMP}) of *E. coli* DHFR that are induced by DHF and TMP binding via 210 ns long MD simulations for each system. The simulations for the WT systems are

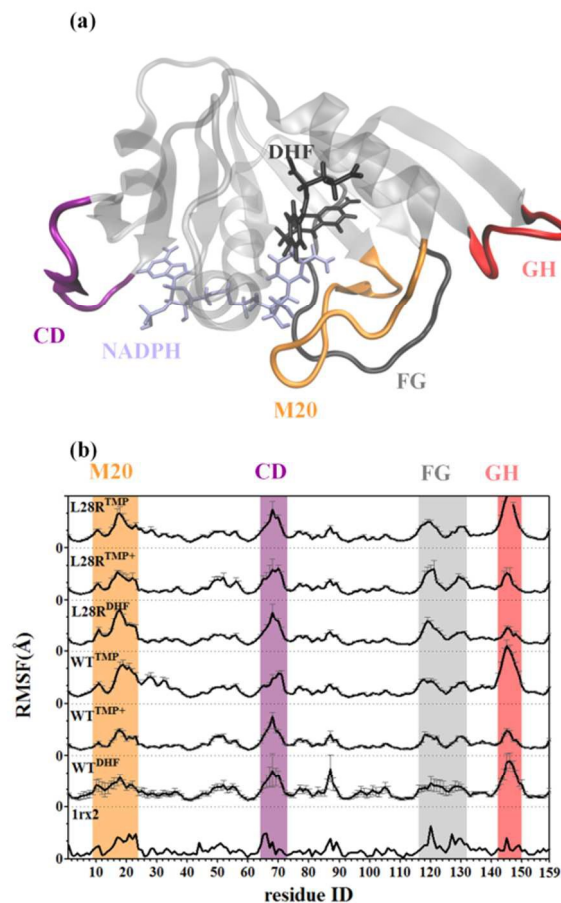


Figure 1. (a) Cartoon representation of *E.coli* DHFR for the PDB structure 1rx2 is displayed in transparent gray. The characteristic loops of the enzyme are shaded in orange, purple, dark gray and red for the M20, CD, FG and GH loops, respectively. DHFR is complexed with DHF (black) and NADPH (lavender) in its native form. (b) RMSF profiles of WT^{DHF}, WT^{TMP+}, WT^{TMP}, L28R^{DHF}, L28R^{TMP+} and L28R^{TMP} averaged over all equally spaced chunks of 40 ns are displayed; error bars for each residue are shown in gray. The average fluctuations obtained from crystallographic B-factors are displayed in bottom panel. The regions of the characteristic loops of the enzyme are shaded in orange, purple, gray and red indicating M20, CD, FG and GH loops, respectively.

duplicated to assess the reproducibility of our observations. The root mean square deviation (RMSD) profiles of the six systems studied with respect to the 1rx2 structure are displayed in Figure S5. The RMSD from the initial, closed x-ray structure (protein data bank, PDB, code 1rx2¹¹), for WTDHF, WT^{TMP} and WT^{TMP+} are 3.8 ± 0.6 , 3.5 ± 0.6 and 2.2 ± 0.3 Å, respectively, averaged

over the whole trajectory for the duplicate MD simulations. For the L28R mutants, these are 2.2 ± 0.5 , 3.3 ± 0.1 and 2.7 ± 0.4 Å, respectively, for L28R^{DHF}, L28R^{TMP} and L28R^{TMP+}. Besides, we closely monitored the M20 loop, which controls the catalytic activity of DHFR in all runs by computing the average RMSD from the closed, open and occluded structures (PDB codes 1rx2, 1ra1 and 1rx7¹¹, respectively). We verified that it remained in the closed conformation in all runs.

While studying the enzyme's conformational behavior during the MD simulations by analyzing the auto and cross correlations (see below), we discard the first 10 ns of the simulations accounting for the time required for equilibration (Figure S5). We then divide the rest of the constant pressure MD simulation trajectory (10-210 ns) of each system into equally spaced and equilibrated pieces of 40 ns to evaluate the characteristic motions averaged over all chunks. As such, the results will reproduce the motions on the time scale up to tens of nanoseconds.

In Figure 1a, the DHFR-DHF-NADPH ternary (1rx2) structure is displayed with loops color coded. In Figure 1b, we display the root mean square fluctuations (RMSF) of all the complexes averaged over the five equally spaced chunks of 40 ns along with those from experimental B-factors for the 1rx2 PDB structure. The most significant changes in RMSF are observed in the M20, CD, FG and GH loops (Figure 1b) which are highlighted in orange, purple, gray and red, respectively. The overall pattern reproduced by the RMSFs for all six systems are also in accord with the experimental B-factors despite shifts in relative magnitudes. The importance of these loops in *E. coli* DHFR has been addressed in the literature. The closed to occluded transition of the M20 loop that occurs on the micro to milliseconds time scale involves breaking M20 hydrogen bonds with the FG loop and establishing new ones with the GH loop.¹¹ A triad of interactions is known to exist between these loops. In addition, residues 67-69 on the CD loop

have been shown to exhibit a consistent increase in rigidity upon trimethoprim binding on the ps-ns time scale motions of the backbone dynamics.¹⁰ They are coupled in stability and function to G121 on the FG loop, despite separation exceeding 25 Å.⁵⁴

Protonated state of trimethoprim in WT enzyme and L28R mutant. Previous studies utilizing ¹³C NMR spectra suggest that in WT DHFR structure, when bound to DHFR, ¹³C labeled trimethoprim is protonated (TMP⁺) on position N1 (Figure S2).⁵⁵ Therein, it is reported that trimethoprim molecules bound to D27N mutant lose a proton at this position at all pH values and is in deprotonated form (TMP). Interestingly, the inhibitor bound to D27S mutant displays pH-dependent protonation state. On the other hand, force field parameters for trimethoprim⁵⁶ molecule in the literature utilize united atom models for CH₂ and CH₃ groups and thus are not adequate for the level of detail required in this study. Therefore, we have parameterized trimethoprim in its neutral and protonated states (TMP and TMP⁺). We have described the parameterization protocol used for trimethoprim force field under the Materials and Methods section; the resulting parameters are listed in supplementary texts S1 and S2.

We verified that the WT enzyme in its inhibitor bound state reproduces structures that are compatible with a protonated TMP⁺ which has been shown by NMR to be the state found when bound to *E. coli* DHFR.⁵⁵ After docking trimethoprim into DHFR's substrate binding site, residues I5, D27, I94, and Y100 are in close proximity of the ligand and form hydrogen bonds with TMP/TMP⁺ (Figure 2a-b). I5 and I94 make the hydrogen bonds through their backbone oxygen atoms, while D27 and Y100 form hydrogen bonds with the ligand via their side chains.

We monitor the average hydrogen bond distances throughout the WT^{TMP⁺} trajectory (Figure 2c). All the initial hydrogen bonds between TMP⁺ and DHFR fluctuate around the initial distances

reducing the motions of the inhibitor in the binding site. The initial distances between hydrogen bond donor and acceptor for I5, D27 (two hydrogen bonds), I94, and Y100 are 1.8, 1.6, 1.6, 2.0 and 3.5 Å and average values of 2.8 ± 0.6 , 2.7 ± 0.9 , 1.9 ± 0.3 , 3.2 ± 0.4 and 3.1 ± 0.4 Å (mean \pm standard deviation), respectively, are recorded to validate that all the hydrogen bonds are preserved. The C4-C5-C7-C1P dihedral (see Figure S1 for the atom labels) characterizing the trimethoxyphenyl methyl tail motions sweep the $180 \pm 30^\circ$ range. The C5-C7-C1P-C2P dihedral fluctuates around $-91.0 \pm 20.0^\circ$ throughout the 210 ns. Lack of librational motions in WT^{TMP+} is quantified by the contribution fraction of the rotation along the three principal axes of the vector connecting the NA2 and C7 atoms as described and visually demonstrated under Models and Methods (see also equation 1) where the major axis of motion has a fraction of 0.80, with the two minor axes contributing only 0.11 and 0.09. A large contribution to one axis and insignificant contributions to the remaining two represent a stiff molecule.⁵⁷

Although TMP is not in its neutral form when bound to WT DHFR, WT^{TMP} structure is studied as a necessary step for evaluating the changes due to the deprotonation. In fact, while TMP⁺ is prevented from going through free movements in the binding cavity due to strong hydrogen bonding of position N1 and the carboxylate group of D27, the opposite is observed for TMP due to the loss of the aforementioned hydrogen bonds (Figure 2d). We observe that all the initial hydrogen bonds between TMP and DHFR are broken in WT^{TMP} complex due to the tumbling motions of TMP inside the active site without exiting on the time scale of the simulations. In alignment with this observation, we find that the contribution fraction of the three principal axes of the vector connecting the NA2 and C7 atoms to TMP movements in this case is 0.42, 0.31 and 0.27. The close values along the three axes of TMP bound to WT^{TMP} represent a free rotation of the molecule inside the binding cavity.⁵⁷ Upon such free rotational movements of TMP inside the

binding site, the initial distances between hydrogen bond donor and acceptor for I5, D27 (two hydrogen bonds), I94, and Y100 increase to the average values of 10.8 ± 2.5 , 9.6 ± 2.7 , 9.7 ± 2.7 , 10.7 ± 2.7 and 10.3 ± 2.3 Å, respectively (Figure 2d). Furthermore, the internal movements of TMP that involve the dihedrals connecting the two rings present in the antibiotic structure change in time. The C4-C5-C7-C1P dihedral characterizing the trimethoxyphenyl methyl also moves in a wider range.

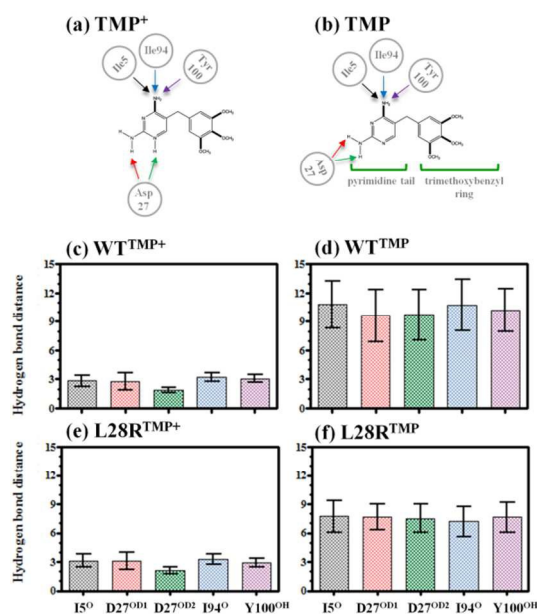


Figure 2. (a), (b) Schematic representation of hydrogen bonds in TMP⁺ and TMP binding site in *E. coli* DHFR. Hydrogen bonds to the indicated amino acid on DHFR (gray circles) are shown by the colored arrows. (c-f) Average hydrogen bond distances formed between trimethoprim and binding site for WT-TMP⁺, WT-TMP, L28R-TMP⁺, L28R-TMP, respectively. Standard deviations are shown by vertical lines. Color coding follows the colors used in the arrows in (a) and (b).

Consequently, the pyrimidine ring of TMP changes its configuration, occasionally returning to the vicinity of the initial binding pose. While adopting multiple configurations, the very mobile TMP can establish a plethora of short-lived hydrogen bonds with other residues in the pocket,

namely A6, I14, M16, E17, N18, M20, W22, N23, L24, S49 and W113. These favorable interactions prevent the drug from leaving the binding site despite the broken initial contacts, and the protein can accommodate the highly mobile TMP inside the active site, occasionally flipping into and out of its original orientation. Thus, the extra positive charge ensures tight binding of trimethoprim (TMP⁺).

It should be noted that increased TMP mobility in the binding pocket has experimentally been observed in the complex of *Lactobacillus casei* (*L. casei*) DHFR with trimethoprim and NADP⁺.⁵⁸ Ring-flipping is achieved upon substantial conformational changes in the protein around the trimethoxybenzyl ring.⁵⁹ NMR studies of binary and ternary complexes of *L. casei* DHFR reveals that trimethoxybenzyl ring conformational rearrangements occur on the millisecond range and they require transient breaking of the hydrogen bonds to disrupt the interactions of trimethoprim which requires conformational changes in the protein to satisfy the predicted barrier heights.⁶⁰ Meanwhile, NMR measurements on *E. coli* DHFR complex that point to two conformations for trimethoprim detects librational motions in trimethoprim which are restricted upon coenzyme binding.¹⁷ The librations are consistent with the binding modes of TMP⁺ observed on the tens of ns time scales observed in the current MD simulations.

We also study the L28R mutant in the presence of both TMP (L28R^{TMP}) and TMP⁺ (L28R^{TMP+}) to determine the protonation state of the inhibitor in this mutant under the prevailing conditions. As in the WT protein, the hydrogen bond donor–acceptor pair distances in L28R^{TMP+} stay near their initial values (Figure 2e). The hydrogen bond distances between I5, D27 (two hydrogen bonds), I94, and Y100 that are originally at 1.8, 1.6, 1.6, 2.0 and 3.6 Å have the average values of 3.1±0.7, 3.1±0.9, 2.1±0.3, 3.3±0.5 and 2.9±0.4 Å, respectively, over the course of the 210 ns trajectory. Thus, while the hydrogen bonds are maintained as in the WT, the distances are

elongated by an average of 0.2 Å in the mutant. C4-C5-C7-C1P dihedral rotates from its initial value of 156° and equilibrates around -100±9°. Meanwhile, the C5-C7-C1P-C2P dihedral remains near its initial value of -108° with the average value of -93±12°. Overall, the motions of TMP⁺ in the pocket are as constrained as in WT, with the contribution fractions of the major axes being 0.83, 0.09, and 0.07. Because of the lengthened average hydrogen bond distances and the distorted C4-C5-C7-C1P dihedral angle, we conclude that TMP⁺ binding is weakened by the L28R mutation, consistent with the increased K_i value (24.63±1.34 nM) with respect to the WT (2.39±1.06 nM).

When TMP is neutralized in the presence of the L28R mutation, the distances between the donor - acceptor pair distances of I5, D27 (two hydrogen bonds), I94, and Y100 in L28R^{TMP} have the average values of 7.7±1.6, 7.6±1.3, 7.5±1.4, 7.2±1.5 and 7.6±1.6 Å, respectively (Figure 2f). The initial hydrogen bonds between TMP and DHFR are no longer permanent. Nevertheless, compared to WT^{TMP}, the donor-acceptor pair distances shorten. While the oscillations around the dihedrals C4-C5-C7-C1P and C5-C7-C1P-C2P are maintained in L28R^{TMP}. The most significant difference from its WT counterpart is that the rigid body rotational motions of TMP inside the binding cavity are much more restricted despite the lack of hydrogen bonds. The latter is quantified by the contribution fractions of the major axes of 0.69, 0.23, and 0.08 (Figure S3). We note that this trajectory may not be fully converged (figure S5); however, lengthening it will not guarantee convergence and will also not include any additional information to that discussed above.

We have also carried out 50 ns long MD simulations on TMP⁺ and TMP bound D27N and D27S mutants to validate the approach we use here to determine the stable form of the drug by monitoring the five hydrogen bonds. The former mutant is known to bind trimethoprim in its

neutral state, while the latter carries significant populations of both forms of the drug, as shown by ^{13}C NMR resonance spectra.⁵⁵ Thus, in our simulations, the D27N^{TMP} should maintain significant amounts of the characteristic hydrogen bonds compared to $\text{D27N}^{\text{TMP}^+}$. Furthermore, D27S^{TMP} to $\text{D27S}^{\text{TMP}^+}$ should both maintain the hydrogen bonds. The relevant hydrogen bond distances are displayed in Figure S6 and are in excellent agreement with these expectations. In summary, the hydrogen bond distribution patterns summarized in this section delineate that the TMP^+ form is the more stable binding partner for the WT and the L28R enzyme. We will therefore confine our discussion to the TMP^+ bound forms of both the WT and mutant enzyme in the rest of the manuscript.

Conformational preferences of DHF-bound DHFR. In Table 1, local structural changes of the systems are reported in terms of individual RMSD of major loops (color coded in Figure 1a). The average structures WT^{DHF} and L28R^{DHF} are displayed in Figure 3a. The RMSD between them is 2.7 Å and we trace the difference mainly to the mobility of the GH loop. While M20 and FG loop conformations of WT^{DHF} are close to the crystal structure, the GH loop displays a major conformational change from its initial position (Figure 3a, grey). Conversely, in the L28R^{DHF} , the GH loop maintains its initial shape (figure 3a, red).

We attribute the lower RMSD of L28R^{DHF} as compared to WT^{DHF} mainly to its sluggish GH loop with an RMSD of 1.5 ± 0.2 Å in the mutant as opposed to 4.2 ± 0.6 Å in the WT. The enhanced mobility of the GH loop in WT^{DHF} is quantified by the RMSFs (Figure 1b). While the fluctuations in this region are relatively higher than the B-factors recorded in the 1rx2 coded x-ray structure, we note that the state simulated in this work is not in any of the x-ray structures. Instead, the cofactor is in reduced form as opposed to the oxidized form in the 1rx2 structure, and DHF is represented as 5-protonated 7,8-dihydrofolate for the reactant in the hydride transfer step

as opposed to folate in the experimental structure.¹¹ Meanwhile, the total ion current abundance % per residue plots in recent ultraviolet photo dissociation mass spectroscopy measurements confirm enhanced motions in the ternary complex in solution, corroborating our observation of the mobile GH loop under some conditions.⁶¹

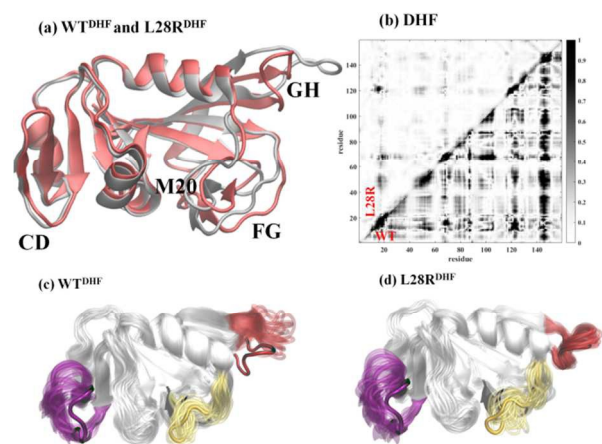


Figure 3. (a) Average structure of WT^{DHF} (gray) is superposed on that of L28R^{DHF} (red). The main difference is observed in M20, FG and GH loops. (b) The residue-residue cross correlation maps between WT^{DHF} (lower diagonal) and L28R^{DHF} (upper diagonal), averaged over chunks of 40 ns. Correlations are much stronger, extending over the whole protein in the WT compared to the L28R mutant. (c,d) Superposed structures of WT^{DHF} and L28R^{DHF}, clearly display the mobility of the loops. 20 equally spaced snapshots are shown.

Table 1. Loop and overall RMSD (Å) in the systems with respect to initial structure, averaged over the 210 ns trajectories.

	M20	CD	FG	GH	overall
WT ^{DHF}	2.7±0.7	2.9±0.8	2.3±0.5	4.2±0.6	3.8±0.6
L28R ^{DHF}	2.6±0.7	2.8±0.7	2.8±0.5	1.5±0.2	2.2±0.5
WT ^{TMP+}	2.4±0.2	1.8±0.4	2.9±0.5	1.5±0.2	2.2±0.3
L28R ^{TMP+}	2.6±0.4	3.2±0.6	2.7±0.3	1.5±0.2	2.7±0.4

The overall fluctuation patterns of DHF bound complexes of DHFR represent the role of loop dynamics in global motions of the enzyme. However, to gain insight on how the loop motions are

coupled to each other and the rest of the residues, we construct the covariance matrix between the fluctuations of the residue networks.⁶² The residue-residue cross correlation maps averaged over all the chunks are displayed in Figure 3b providing information if the processes are coupled over the ten nanoseconds time scale; the lower triangular region is for WT^{DHF}, while the upper part is for L28R^{DHF}. Equally spaced, superposed snapshots from this trajectory are displayed in Figure 3c.

In WT^{DHF}, there are strong correlations between almost all parts of the protein (Figure 3b, lower triangle). In particular, the M20 loop is correlated to residues 47-59, CD loop, residues 85-89, 95-107, FG and GH loops. Meanwhile, residues 119-126 of FG loop and residues 142-149 located on GH loop are highly correlated with the entire structure of the system. A strong coupling is also detected between residues 95-107 and the CD loop.

In Figure 3b, upper triangle, we display the effect of the L28R mutation on the residue cross correlations. The mutation suppresses the overall cross correlations between the different regions of the enzyme in the presence of DHF compared to the WT systems. In particular, the concerted movements of the GH loop with the rest of the protein disappear. However, we observe a remaining correlation of the M20 loop (Figure 3d, yellow loop) with some CD and FG loop residues as well as residues 42-55. The enhanced motions of the FG loop and the overall more ordered structure of the L28R^{DHF} system are apparent in Figure 3d.

Comparing the local and global dynamics of the WT and L28R mutant enzyme, we next monitor the local interactions of the native ligand through the hydrogen bonds established with the enzyme. We assess the changes in hydrogen bond distributions of the active site in all systems (Figure 4). The residues in the DHF binding site of DHFR are I5, A6, A7, M20, D27, L28, F31,

R52, R57, I94 and T113. Among them, backbone of residue I5 and side chains of residues D27, R52 and R57 can form hydrogen bonds with DHF atoms of NA2, N3 and N8 as the hydrogen bond donors and O, O1 and O2 as the hydrogen bond acceptors (Figure 4a; see Figure S2 for atom labeling scheme). Average acceptor – donor distances for each hydrogen bond formed between DHF and the binding site in WTDHF and L28RDHF are displayed in Figure 4b-c. In both enzymes, $I5^O$ -DHF^{H8} (2.5 ± 0.6 Å), $D27^{OD1}$ -DHF^{NA2} (2.1 ± 0.4 Å), and $D27^{OD2}$ -DHF^{N3} (1.9 ± 0.3 Å) hydrogen bonds are maintained throughout the MD simulations, keeping I5 and D27 in constant contact with DHF. I5 and D27 are located on a rigid and stable α -helix and β -strand, respectively. In the WT system, the remaining monitored hydrogen bonds on R52 and R57 begin to break by 10 ns of MD simulations with occasional reformation and breaking events throughout the trajectory due to the mobility of the C6-C9-N10-C14 dihedral angle connecting the pterin ring and p-aminobenzoyl glutamate tail of DHF. The dihedral angle changes in the range $[-154^\circ, 113^\circ]$ while the hydrogen bond occurs only for the dihedral angle range of $[-109^\circ, 113^\circ]$. More negative angles result in twisting the p-aminobenzoyl glutamate tail, turning the hydrogen bond acceptor site away from R57, losing the $R57^{HH11}$ -DHF^{O1} hydrogen bond. R52 and R57 reside on a highly mobile and flexible loop (residues 51-58) which displays large fluctuations and its RMSD increases up to 2.7 Å. Consequently, p-aminobenzoyl glutamate tail of DHF is provided with enough freedom to move in large swinging motions while the pterin ring is kept in position by I5 and D27. Because of a freely moving tail, the distance of the $R52^{HH21}$ -DHF^O hydrogen bond increases up to 16.4 Å by 10 ns of MD simulations, with an average value of 6.2 ± 2.1 Å. The $R57^{HH11}$ -DHF^{O1} and $R57^{HH21}$ -DHF^{O2} hydrogen bonds are kept intact until 30 ns, after which the distances increase up to 8.3 ± 3.1 and 9.0 ± 3.5 Å, respectively. We find that the movements of the pterin tail increases the $R57^{HH11}$ -DHF^{O1} and $R57^{HH21}$ -DHF^{O2} distances. We quantify the mobility of DHF by calculating the principle axis of p-aminobenzoyl glutamate tail in the WT

system. The principal axis calculations were carried out for a vector connecting the N10 position to CD position as described in the Models and Methods section (also see equation 1). The contribution fraction of the three principal axes to DHF p-aminobenzoyl glutamate tail movements in WTDHF is 0.45, 0.33 and 0.22 (Figure S3a). Comparable contributions along the three axes of DHF bound to WT system represent free rotation of the tail.⁵⁷

We note that the strengthening the hydrogen bonds in the binding pocket only explains the enthalpic contribution to the Gibbs free energy difference. We have used pulling experiments to predict the relative binding affinity of DHF to DHFR in the WT and the mutant. To facilitate ligand dissociation in each system, we have used snapshots at the 10, 50, 90, 130, 170 and 210 ns time points of the trajectories as described in the methods section. The total free energy differences, ΔG , between ligand free and ligand bound states are 17.0 ± 1.0 , and 15.8 ± 1.0 kcal/mol, for WT^{DHF} and $L28R^{DHF}$, respectively, which imply that the systems have indistinguishable stability, within error bounds ($\Delta\Delta G = 1.2 \pm 1.4$ kcal/mol), assuming the L28R mutation has similar free energetic cost in the DHF-bound and DHF-free enzyme. Thus, despite the hydrogen bond stabilization of the ligand in the L28R mutant, the entropy loss due to the lack of fluctuations in the p-aminobenzoyl glutamate tail and the long side chain of residue 28, along with the suppressed mobility of the GH loop (Figures 1b and 3, Table 1), compensate for the energetic gain and the two systems have comparable overall stability. On the other hand, the extra hydrogen bonds established between the ligand and the enzyme in the mutant (figure 4d-g) are expected to lead to a larger energetic cost to reach the transition state ($\Delta H_{L28R}^\ddagger > \Delta H_{WT}^\ddagger$). Therefore, the barrier heights for ligand dissociation from the L28R mutant is expected to be higher. In fact, using the relation $K_m = (k_f + k_{cat})/k_f$ where k_f and k_r are the forward and reverse reaction rates for substrate binding, k_{cat} is the reaction rate for product formation, and substituting

the experimentally measured values for K_m and k_f listed in Table 2 as well as an $\Delta\Delta G$ of 1.2 kcal/mol disfavoring DHF binding in the mutant, we find that the ratio of reaction rate for DHF dissociation, $k_f^{\text{WT}}/k_f^{\text{L28R}}$, is expected to increase in the range 2-20 depending on the actual value of k_f^{WT} . Furthermore, since THF is only a single hydrogen different than the 5-protonated 7,8-dihydrofolate in the current simulations, one would expect similar hydrogen bonding patterns for THF, and a similar increase in barrier heights.

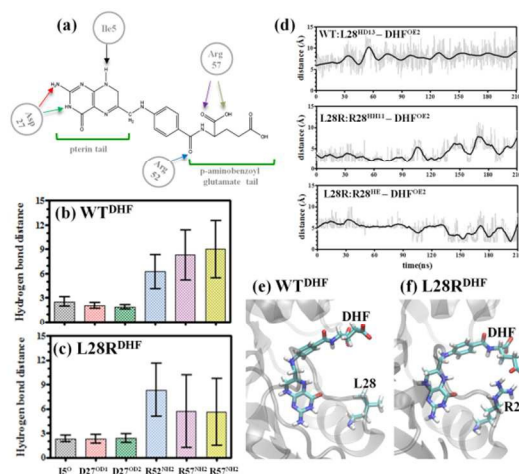


Figure 4. (a) Schematic representation of hydrogen bonds in DHF binding site in *E. coli* DHFR. Hydrogen bonds to the indicated amino acid on DHFR (gray circles) are shown by the colored arrows. (b), (c) Average hydrogen bond distances formed between DHF and binding site for WT^{DHF} and L28R^{DHF}, respectively. Standard deviations are shown by vertical lines. Color coding follows those used in the arrows in (a). (d) Sample donor-acceptor distance between the methyl group of residue 28 and DHF. For the WT, L28 makes no hydrogen bonds with the p-aminobenzoyl glutamate tail of the substrate. Sample donor-acceptor distances between methyl group of R28 and DHF indicates there are sustained hydrogen bonds between the α -amino group of R28 and p-aminobenzoyl glutamate tail of DHF in the L28R mutant. (f), (g) sample pose of DHF relative to residue 28 in WT^{DHF} and L28R^{DHF}, respectively. R28 interacts with DHF p-aminobenzoyl glutamate tail via its side chain and further stabilizes the ligand.

This is in conformity with k_{cat} values we measure which are 8.16 ± 3.27 and 1.30 ± 0.01 s⁻¹, respectively for the WT and the L28R mutant enzyme (Table 2). We note that THF release is the rate limiting step in the catalytic cycle. Nevertheless, the thermodynamics of the inhibitor bound states must also be assessed to evaluate the impact of the mutation in the competitive inhibition of DHFR which we address next.

Table 2. Experimental results for the competitive inhibition of DHFR.

ITC measurements					Competitive inhibition measurements		
	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal/mol/K)	K_d (nM)	k_{cat} (s ⁻¹)	K_m (μM)	K_i (nM)
WT	-11.2±0.5	-13.1±1.4	-6.5±3.0	4.5±0.9	8.16±3.27	1.49±0.14	2.39±1.06
L28R	-10.8±0.2	-6.8±0.5	13.4±2.6	13.1±1.5	1.30±0.01	0.62±0.01	24.63±1.34
Difference	+0.4±0.5 (+0.6±1.1)*	+6.3±1.5	+19.9±4.0				

*free energy difference calculated via pulling experiments

L28R mutation disrupts water organization at the trimethoprim – enzyme interface. The overall structure of trimethoprim-bound DHFR is nearly unaltered upon the inclusion of the L28R mutation, as evident from the average structures displayed in Figure 5a. The RMSD between WT^{TMP+} and L28R^{TMP+} average structures is only 1.1 Å. The largest change with respect to the x-ray structure occurs in the CD loop (Table 1). Introducing TMP⁺ into the DHFR structure leads to a system where the motions of the loops are weakly dependent on each other in both the WT and the L28R mutant (Figure 5b). In the WT, M20 loop is correlated with residues 45-55, CD and FG loop. CD loop also displays small, but detectable correlations with residues 76-105. These results are in conformity with NMR spectroscopy findings on *E. coli* DHFR in complex with trimethoprim where the closed conformation was found to predominate the ensemble of structures upon drug binding.⁸ It was proposed therein that, while filling the active site sends a

signal for conformational change on the μs – ms time scale, the drugged-complex remains locked in the closed conformation. Moreover, NMR measurements on backbone and side chain methyl groups showed that the motional landscape was also simplified on the ps – ns time scale that is also sampled in the current work. In particular, side-chain methyl order parameters measuring the ps – ns dynamics in the closed, trimethoprim bound protein indicates a general rigidification. This is in accordance with the lack of strong correlations in protein fluctuations of WT^{TMP^+} system (Figure 5b) as well as the lower RMSF values (Figure 1b).

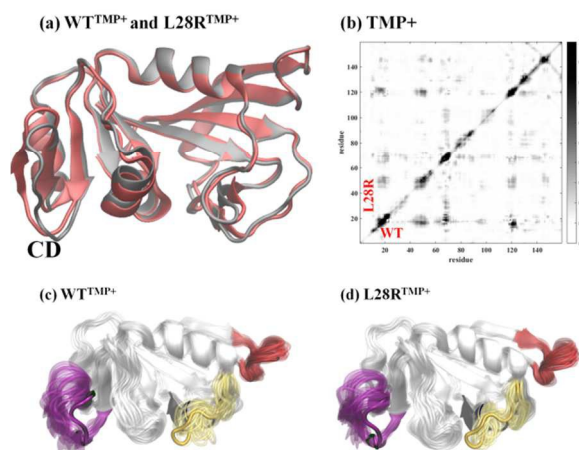


Figure 5. (a) Average structure of WT^{TMP^+} (gray) is superposed on that of $\text{L28R}^{\text{TMP}^+}$ (red). The main difference is observed in the CD loop. (b) The residue-residue cross correlation maps between WT^{TMP^+} (lower diagonal) and $\text{L28R}^{\text{TMP}^+}$ (upper diagonal), averaged over chunks of 40 ns. Correlations are subsided in TMP^+ bound forms, slightly more so in the L28R mutant than the WT. (c,d) Superposed structures of WT^{TMP^+} and $\text{L28R}^{\text{TMP}^+}$ clearly display loop motions. 20 equally spaced snapshots are shown.

Equally spaced snapshots from the WT^{TMP^+} and $\text{L28R}^{\text{TMP}^+}$ simulations are superimposed in Figure 5c,d for comparison. The covariance matrix of $\text{L28R}^{\text{TMP}^+}$ in the upper triangle of Figure 5b, shows that the drug binding leads to very similar correlations to those of WT^{TMP^+} system in corresponding residue positions. Together with the fact that the hydrogen bonds in the binding

pocket are nearly unaltered in the WT and mutant enzyme (Figure 2 c,e), we conclude that the L28R mutation has a minute effect on the trimethoprim binding propensity of the enzyme.

To measure the relative stability of TMP^+ in the binding pocket, we have again resorted to free energy of binding calculations. Using snapshots at the 10, 50, 90, 130, 170 and 210 ns time points for both systems as described in the Materials and Methods section, ΔG between ligand free and ligand bound states are calculated as 11.6 ± 0.9 and 11.0 ± 0.7 kcal/mol for WT^{TMP} and L28R^{TMP} systems, respectively. We note that since TMP is found in neutral form in the water environment, these values cannot be used to calculate the exact K_d . However, the difference between them may be assessed to determine the relative stability of TMP in the binding pocket, which is implied to be very similar, within error bounds ($\Delta\Delta G = 0.6 \pm 1.1$ kcal/mol). In fact, we have carried out ITC experiments for the binding of trimethoprim to the WT and L28R mutant. The experimental data are presented in Figure S7, and the calculated ΔG , ΔH and ΔS values are listed in Table 2.⁶³ The experimental free energy difference between the mutant and the WT (0.4 ± 0.5 kcal/mol) are in conformity with the calculated value. ITC experiments find K_d of 4.5 and 13.1 nM respectively for WT and L28R systems in the presence of trimethoprim. The nM range calculated dissociation constants are in agreement with experimental K_d values reported for trimethoprim bound WT system (9.1 nM).¹⁶

While the free energy differences for trimethoprim dissociation are similar for the WT and the mutant, we find that there is energy/entropy compensation. The enthalpy of binding of the drug is less favorable in the mutant, which is not directly apparent from a mere consideration of the average 0.2 Å increase in the hydrogen bond distances within the binding pocket (Figure 2c,e).

This energetic loss is compensated by the increased entropy change (Table 2) from -6.5 ± 3.0 cal/mol/K to 13.4 ± 2.6 cal/mol/K.

The entropy change due to binding is associated with the following contributions: The largest is due to the solvent entropy increase arising from the release of solvent molecules upon binding. Conformational entropy change accounts for the variations in the degrees of freedom of the protein (e.g. restricted side chain mobility) and the ligand; it may make a positive or a negative contribution. Furthermore, rotational and translational degrees of freedom loss in both the protein and the ligand upon binding will have a negative effect. Finally, ordering of water at the complex interface may contribute unfavorably to entropy change and favorably to enthalpy change.⁶⁴

We find that the WT protein in fact harbors an environment for ordered water when bound to trimethoprim, corroborating the measured entropy and enthalpy changes listed in Table 2. The average solvent accessible surface area of the trimethoprim in the WT protein is 112 \AA^2 in the WT while it is reduced to 99 \AA^2 (standard deviation is 20 \AA^2 in both cases), leading to a tightening of $\sim 10 \%$. When we monitor the trajectories, we find that a structural water molecule is located next to the NA2 atom of the TMP^+ molecule that is stable for the last 50 ns of the first trajectory and the last 100 ns of the second trajectory. No such water molecules are found in the mutant. The environment of the coordinating water molecule in the WT is displayed and compared with that of the mutant in Figure S7. Residues 5, 111 and 113 from the β sheet beneath the TMP binding site accommodate the interfacial water molecule that makes an additional hydrogen bond with trimethoprim while not disrupting the one with D27. In the L28R mutant, the binding site is slightly tightened so that a water molecule may no longer squeeze between the β sheet, D27 and trimethoprim.

Conclusions

In this study, thermodynamic and kinetic measurements with all atom classical and steered MD simulations on substrate and inhibitor bound DHFR are combined. The role of ligand dynamics, and its local and allosteric effects on the enzyme provide a detailed mechanism for altering inhibitory activity of trimethoprim on DHFR. The main findings of the study are outlined in Figure 6.

Measurements and calculations in both the E-I and the E-S ternary complexes depict that the positions of the minima are not greatly altered. In the E-I state, TMP^+ rather than TMP is the stable form in both WT and L28R systems. The loops maintain their initial shape and establish much-rigidified structures on the ps–ns time scales, in conformity with NMR spectroscopy findings where the closed conformation was found to predominate the ensemble of structures upon drug binding.⁸ Thus, the lost correlations across the enzyme contributes to the inhibition process in WT^{TMP^+} . The inability to retrieve long-range correlations in $\text{L28R}^{\text{TMP}^+}$ shows that the L28R substitution may not be fully effective in developing a resistance mechanism in the presence of TMP^+ . While the binding free energy difference between the two systems is not large enough to confer resistance to the L28R mutant by destabilizing the drug, there are local rearrangements (Figure 6, left side). The loss in the enthalpic contributions in the mutant is compensated by entropic gain in the mutant as measured by ITC experiments (Table 2). The cause of this observed change is deciphered by monitoring the MD trajectories and found to be due to the release of an interfacial water molecule from the slightly tightened binding site of the mutant.

In the presence of DHF (Figure 6, E-S state), while the M20 and GH loop motions are highly correlated in the WT, the extent of long-range coupling between different regions significantly decreases in L28R^{DHF}, mainly due to the diminished mobility of the GH loop. The emergence of new hydrogen bonds in the binding site of L28R^{DHF} demonstrates substrate mobility is impaired by the mutation. Thus, in the E-S state, barriers to reach both E+I+S and to E+P states are increased in the L28R mutant due to the additional hydrogen bonds established between the R28 side chain and the substrate. The latter is directly measured by the decrease in k_{cat} , while the former is expected by combining the ~ 1 kcal/mol increase in the position of the by the measured K_m values. The increased barrier heights for the DHFR/NADPH/DHF ternary complex prolong the lifetime of this complex, leading to a slightly increased product formation rate for the mutant over the WT at low drug concentrations, providing the necessary condition for viability.

Since trimethoprim is a competitive inhibitor for bacterial DHFR, one would expect that the resistance-conferring mutations around the active site of DHFR would decrease both substrate and drug specificity, making DHFR more promiscuous as was previously observed in several enzymes such as beta-lactamases.⁶⁵⁻⁶⁷ Interestingly, the mechanism for trimethoprim resistance in the L28R mutant increases substrate specificity while slightly weakening drug binding. Therefore, in the presence of both the substrate and the drug, substrate binding is preferred over the drug binding and strong trimethoprim resistance emerges. However, this leads to an evolutionary trade off, as catalytic rate of the L28R mutant decreases as a result of slower product release. Diminished DHFR activity in the L28R mutant may explain why this mutation is not observed in trimethoprim free environments.

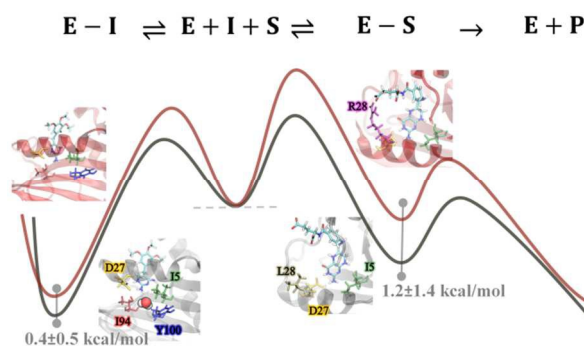


Figure 6. Landscape shift in the L28R mutant (red) with respect to the WT (gray) in competitive inhibition of the substrate (S) binding to the enzyme (E) by the inhibitor (I) to alter product (P) formation (see text for details). Both E-I and E-S states are ternary complexes, also accommodating the cofactor (not shown). The energy levels are relative to the E+I+S state.

Acknowledgements

This work was supported by the Scientific and Technological Research Council of Turkey Projects (grant number 114Z527) and supported in part by the University of Texas Southwestern Medical Center Endowed Scholars Program (Toprak). We thank Chad Brautigam, PhD and Shih-Chia (Scott) Tso, PhD from University of Texas Southwestern Macromolecular Biophysics Resource Core Laboratory for their generous help in ITC.

References

1. Z. D. Nagel and J. P. Klinman, *Nat. Chem. Biol.*, 2009, 5, 543-550.
2. S. J. Benkovic and S. Hammes-Schiffer, *Science*, 2003, 301, 1196-1202.
3. K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus and D. Kern, *Nature*, 2007, 450, 913-916.
4. G. Bhabha, J. Lee, D. C. Ekiert, J. Gam, I. A. Wilson, H. J. Dyson, S. J. Benkovic and P. E. Wright, *Science*, 2011, 332, 234-238.
5. C. Baysal and A. R. Atilgan, *Proteins: Struct. Funct. Bioinform.*, 2001, 45, 62-70.
6. C. Baysal and A. R. Atilgan, *Biophys. J.*, 2002, 83, 699-705.
7. H. Abdizadeh, G. Guven, A. R. Atilgan and C. Atilgan, *J. Enzyme Inhib. Med. Chem.*, 2015, 30, 867-873.
8. R. V. Mauldin, M. J. Carroll and A. L. Lee, *Structure (London, England : 1993)*, 2009, 17, 386-394.
9. A. C. Palmer, E. Toprak, M. Baym, S. Kim, A. Veres, S. Bershtein and R. Kishony, *Nature Communications*, 2015, 6, 7385.
10. J. R. Schnell, H. J. Dyson and P. E. Wright, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, 33, 119-140.
11. M. R. Sawaya and J. Kraut, *Biochemistry*, 1997, 36, 586-603.
12. C. A. Fierke, K. A. Johnson and S. J. Benkovic, *Biochemistry*, 1987, 26, 4085-4092.
13. P. Huovinen, L. Sundström, G. Swedberg and O. Sköld, *Antimicrob. Agents Chemother.*, 1995, 39, 279.
14. E. Toprak, A. Veres, J.-B. Michel, R. Chait, D. L. Hartl and R. Kishony, *Nat. Genet.*, 2012, 44, 101-105.
15. M. N. Alekshun and S. B. Levy, *Cell*, 2007, 128, 1037-1050.
16. M. Watson, J.-W. Liu and D. Ollis, *FEBS J.*, 2007, 274, 2661-2671.
17. Q. X. Yang, F. Y. Huang, T. H. Lin, L. Gelbaum, E. E. Howell and T. H. Huang, *Solid State Nucl. Magn. Reson.*, 1996, 7, 193-201.
18. D. D. Boehr, D. McElheny, H. J. Dyson and P. E. Wright, *Science*, 2006, 313, 1638-1642.
19. P. K. Agarwal, S. R. Billeter, P. T. R. Rajagopalan, S. J. Benkovic and S. Hammes-Schiffer, *Proceedings of the National Academy of Sciences*, 2002, 99, 2794-2799.

20. T. H. Rod, J. L. Radkiewicz and C. L. Brooks, *Proceedings of the National Academy of Sciences*, 2003, 100, 6980-6985.
21. M. J. Osborne, J. Schnell, S. J. Benkovic, H. J. Dyson and P. E. Wright, *Biochemistry*, 2001, 40, 9846-9859.
22. M. J. Osborne, R. P. Venkitakrishnan, H. J. Dyson and P. E. Wright, *Protein Science : A Publication of the Protein Society*, 2003, 12, 2230-2238.
23. D. McElheny, J. R. Schnell, J. C. Lansing, H. J. Dyson and P. E. Wright, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, 102, 5032-5037.
24. C. J. Falzone, P. E. Wright and S. J. Benkovic, *Biochemistry*, 1994, 33, 439-442.
25. K. F. Wong, T. Selzer, S. J. Benkovic and S. Hammes-Schiffer, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, 102, 6807-6812.
26. J. L. Radkiewicz and C. L. Brooks, *Journal of the American Chemical Society*, 2000, 122, 225-231.
27. I. F. Thorpe and C. L. Brooks, *Proteins: Structure, Function, and Bioinformatics*, 2004, 57, 444-457.
28. C. T. Liu, K. Francis, J. P. Layfield, X. Huang, S. Hammes-Schiffer, A. Kohen and S. J. Benkovic, *Proceedings of the National Academy of Sciences*, 2014, 111, 18231-18236.
29. C. T. Liu, P. Hanoian, J. B. French, T. H. Pringle, S. Hammes-Schiffer and S. J. Benkovic, *Proceedings of the National Academy of Sciences*, 2013, 110, 10159-10164.
30. Y. Y. Sham, B. Ma, C. J. Tsai and R. Nussinov, *Proteins*, 2002, 46, 308-320.
31. H. Abdizadeh, A. Atilgan and C. Atilgan, *JBIC Journal of Biological Inorganic Chemistry*, 2015, 20, 705-718.
32. G. Guven, A. R. Atilgan and C. Atilgan, *The Journal of Physical Chemistry B* 2014, 118, 11677-11687.
33. H. Abdizadeh and C. Atilgan, *Phys. Chem. Chem. Phys.*, 2016, 18, 7916-7926.
34. A. O. Aykut, A. R. Atilgan and C. Atilgan, *PLoS Comput. Biol.*, 2013, 9, e1003366.
35. J. V. Rodrigues, S. Bershtein, A. Li, E. R. Lozovsky, D. L. Hartl and E. I. Shakhnovich, *Proceedings of the National Academy of Sciences*, 2016, 113, E1470-E1478.
36. H. Heaslet, M. Harris, K. Fahnoe, R. Sarver, H. Putz, J. Chang, C. Subramanyam, G. Barreiro and J. R. Miller, *Proteins: Structure, Function, and Bioinformatics*, 2009, 76, 706-717.
37. T. Dams, G. Auerbach, G. Bader, U. Jacob, T. Ploom, R. Huber and R. Jaenicke, *J. Mol. Biol.*, 2000, 297, 659-672.
38. G. P. Miller, D. C. Wahnnon and S. J. Benkovic, *Biochemistry*, 2001, 40, 867-875.

39. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, 26, 1781-1802.
40. W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics* 1996, 14, 33-38.
41. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and e. al, *J. Comput. Chem.*, 1983, 4, 187-217.
42. M. Garcia-Viloca, D. G. Truhlar and J. Gao, *Biochemistry*, 2003, 42, 13558-13575.
43. C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid and J. C. Gumbart, *J. Comput. Chem.*, 2013, 34, 2757-2770.
44. T. Darden, L. Perera, L. P. Li and L. Pedersen, *Struc Fold Des*, 1999, 7, R55-R60.
45. H. C. Andersen, *J. Comput. Phys.*, 1983, 52, 24-34.
46. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, 30, 2785-2791.
47. W. R. Harris, *Biochim. Biophys. Acta*, 2012, 1820, 348-361.
48. Dassault Systèmes BIOVIA, 2015.
49. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford, CT, USA, 2009.
50. S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *The Journal of chemical physics*, 2003, 119, 3559-3566.
51. C. Jarzynski, *Phys. Rev. Lett.*, 1997, 78, 2690.
52. J. Gore, F. Ritort and C. Bustamante, *Proceedings of the National Academy of Sciences*, 2003, 100, 12564-12569.
53. O. Sensoy, A. R. Atilgan and C. Atilgan, *Phys. Chem. Chem. Phys.*, under review.
54. E. Ohmae, K. Iriyama, S. Ichihara and K. Gekko, *J. Biochem.*, 1998, 123, 33-41.
55. R. E. London, E. E. Howell, M. S. Warren, J. Kraut and R. L. Blakley, *Biochemistry*, 1986, 25, 7229-7235.

56. C. L. Brooks and S. H. Fleischman, *J. Am. Chem. Soc.*, 1990, 112, 3307-3312.
57. C. Baysal, B. Erman and B. Chu, *The Journal of Chemical Physics*, 2001, 114, 5444-5449.
58. J. Feeney, *Angew. Chem. Int. Ed.*, 2000, 39, 290-312.
59. V. I. Polshakov, B. Birdsall and J. Feeney, *Biochemistry*, 1999, 38, 15962-15969.
60. M. S. Searle, M. J. Forster, B. Birdsall, G. C. Roberts, J. Feeney, H. T. Cheung, I. Kompis and A. J. Geddes, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, 85, 3787-3791.
61. M. B. Cammarata, R. Thyer, J. Rosenberg, A. Ellington and J. S. Brodbelt, *J. Am. Chem. Soc.*, 2015, 137, 9128-9135.
62. C. Atilgan, O. B. Okan and A. R. Atilgan, *Annual Review of Biophysics*, 2012, 41, 205-225.
63. I. Jelesarov and H. R. Bosshard, *J. Mol. Recognit.*, 1999, 12, 3-18.
64. G. A. Holdgate, A. Tunnicliffe, W. H. J. Ward, S. A. Weston, G. Rosenbrock, P. T. Barth, I. W. F. Taylor, R. A. Pauptit and D. Timms, *Biochemistry*, 1997, 36, 9663-9673.
65. S. Bershtein, M. Segal, R. Bekerman, N. Tokuriki and D. S. Tawfik, *Nature*, 2006, 444, 929-932.
66. E. Dellus-Gur, M. Elias, E. Caselli, F. Prati, M. L. Salverda, J. A. de Visser, J. S. Fraser and D. S. Tawfik, *J. Mol. Biol.*, 2015, 427, 2396-2409.
67. Michael A. Stiffler, Doeke R. Hekstra and R. Ranganathan, *Cell*, 2015, 160, 882-892.

