

Insight into the molecular mechanism about lowered dihydrofolate binding affinity to dihydrofolate reductase-like 1 (DHFRL1)

Jian Gao · Wei Cui · Yuguo Du · Mingjuan Ji

Received: 28 June 2013 / Accepted: 15 September 2013 / Published online: 12 October 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Human dihydrofolate reductase-like 1 (DHFRL1) has been identified as a second human dihydrofolate reductase (DHFR) enzyme. Although DHFRL1 have high sequence homology with human DHFR, dihydrofolate (DHF) exhibits a lowered binding affinity to DHFRL1 and the corresponding molecular mechanism is still unknown. To address this question, we studied the binding of DHF to DHFRL1 and DHFR by using molecular dynamics simulation. Moreover, to investigate the role the 24th residue of DHFR/DHFRL1 plays in DHF binding, R24W DHFRL1 mutant was also studied. The van der Waals interaction are more crucial for the total DHF binding energies, while the difference between the DHF binding energies of human DHFR and DHFRL1 can be attributed to the electrostatic interaction and the polar desolvation free energy. More specifically, lower DHF affinity to DHFRL1 can be mainly attributed to the reduction of net electrostatic interactions of residues Arg32 and Gln35 of DHFRL1 with DHF as being affected by Arg24. The side chain of Arg24 in DHFRL1 can extend deeply into the binding sites of DHF and NADPH, and disturb the DHF binding by steric effect, which rarely happens in human DHFR and R24W DHFRL1 mutant. Additionally, the conformation of loop I in DHFRL1 was also studied in this work. Interestingly, the loop conformation resemble to normal closed state of *Escherichia coli* DHFR other than the closed state of human DHFR. We hope this work will be useful to understand the general characteristics of DHFRL1.

Keywords Dihydrofolate · Dihydrofolate reductase-like 1 · MM/GBSA · Molecular dynamics simulation

Introduction

Dihydrofolate reductase (DHFR) is a widely recognized drug target used for antibacterial [1, 2], antiparasitic [3], and anti-cancer [4, 5]. The enzyme catalyzes the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF) through a catalytic cycle with nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor [6, 7]. DHFR on chromosome 5 was previously regarded as the only expressed and functional enzyme in the human body [8]. Nonetheless, recent evidence provided by McEntee et al. have shown that the formerly known pseudogene DHFRP4, or dihydrofolate reductase-like 1 (DHFRL1), is not only expressed but also shares the functions with normal human DHFR [9]. DHFRL1 is further considered as the second human DHFR enzyme based on its subcellular localization. Even though the sequence identity of human DHFR with DHFRL1 is up to 92 %, the enzyme activity of DHFRL1 is roughly two-thirds that of human DHFR [9, 10]. Further kinetic analysis also reveals that DHFRL1 exhibits a higher K_m values of DHF (209.3 μM) than human DHFR (20.1 μM) [9]. DHF has a remarkably lower binding affinity to DHFRL1 than human DHFR. It is very interesting to study the possible molecular mechanism and uncover the key residue(s) that accounts for the different DHF binding affinities to these two enzymes.

The molecular mechanism about the DHF binding affinity and the key residues in DHFR has been well studied by investigating the effect of a series of mutations of human DHFR. For example, the single amino acid substitutions of both Phe31 and Gln35 could moderately decrease DHF affinities by influencing the contacts the DHFR with the p-ABA and glutamate moieties of DHF [11, 12]. Similarly, the substitution of Leu22 of human

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-2018-2) contains supplementary material, which is available to authorized users.

J. Gao · W. Cui · Y. Du · M. Ji (✉)
College of Chemistry and Chemical Engineering, University of
Chinese Academy of Sciences, Beijing 100049, China
e-mail: jmj@ucas.ac.cn

DHFR by Tyr, Phe, Trp, and Arg also showed decreased DHF affinities [13]. In addition, molecular dynamics (MD) simulation becomes a powerful method to study the ligand binding [14–19] and the catalytic cycle of DHFR [20–22]. Tosso et al. [17] investigated the binding energies of DHFR inhibitors by combining MD simulation, *ab initio*, and density functional theory (DFT) calculation and obtained a significant correlation between the calculated binding energies and experimental data. A quantum mechanics/molecular mechanics MD simulation by Fan et al. [20] showed that M42W/G121V double mutation could severely restrict the conformational dynamics of the M20 loop of DHFR and reduce the entropy of activation.

However, the residues at position 22, 31, and 35 in DHFRL1 are the same as human DHFR, suggesting that the loss of DHF affinity may result from residue substitution in other positions. There are 15 residues in DHFRL1 different from human DHFR (as summarized in Fig. 1), only residue 24 locates in the binding sites of DHF and NADPH and is very close to DHF, while the other residues are far from the binding sites (Fig. 2). Generally, the residue at the position 24 is tryptophan (W), which is a strictly invariant residue in all bacterial (e.g., *Escherichia coli* DHFR) and vertebrate DHFRs [23, 24]. Moreover, previous mutagenesis studies revealed that mutation of W by phenylalanine (F) would result in a 50 % decrease in binding stability and a decrease in catalysis efficiency from 69 % to 21 % under intracellular conditions [23]. A similar site-directed mutagenesis on mouse DHFR indicated that the mutation of residue 24 from W to arginine (R) would significantly increase the K_m value of DHF and disturb the catalytic activity of enzyme [24]. As for DHFRL1, the residue at position 24 is arginine which introduces an additional -NH₃- moiety, extending deeply into the binding pocket of DHFRL1. While the impact of such a difference on the DHF binding affinity and the corresponding molecular mechanisms are still unknown. Besides, it will be interesting to investigate whether the R24W mutant would recover the binding affinity of DHF. To address these questions, MD

simulation, molecular mechanics/generalized Born surface area (MM/GBSA) free energy calculations [25–27], and MM/GBSA free energy decomposition analysis [28, 29] were employed to study the binding of DHF to human DHFR, DHFRL1, and the R24W DHFRL1 mutant. Our results may pave a new way for better understanding of the DHFRL1-DHF binding mechanism by providing atomic details and conformational dynamics that are often inaccessible in experiment.

Materials and methods

Preparation of the enzyme-substrate systems

As mentioned above, there are three systems needed to be constructed, including human DHFR-NADPH-DHF (denoted by human DHFR), DHFRL1-NADPH-DHF (denoted by DHFRL1), and R24W DHFRL1-NADPH-DHF (denoted by R24W DHFRL1). The X-ray structure of human DHFR complex (PDB ID: 2W3M), taken from the RCSB Brookhaven Protein Data Bank (PDB) [30], can be directly used for DHFR system. Considering that these two enzymes exhibit high homologous sequences (92 %) and structural similarity (rms is 0.0034) as indicated by previous homology modeling of single DHFRL1 enzyme [10], we attempted to construct the DHFRL1 system based on the human DHFR system (PDB ID: 2W3M) by simply replacing the different residues from DHFR. In addition, in order to investigate the possible role of the residue in position 24, the R24W DHFRL1 mutant was also constructed by similar modifications. All the manual modifications were accomplished by Sybyl 7.1 [31].

Force field

The minimization and molecular dynamics simulation of the three systems were conducted by AMBER 9.0 [32] molecular simulation package with the standard AMBER03 force field

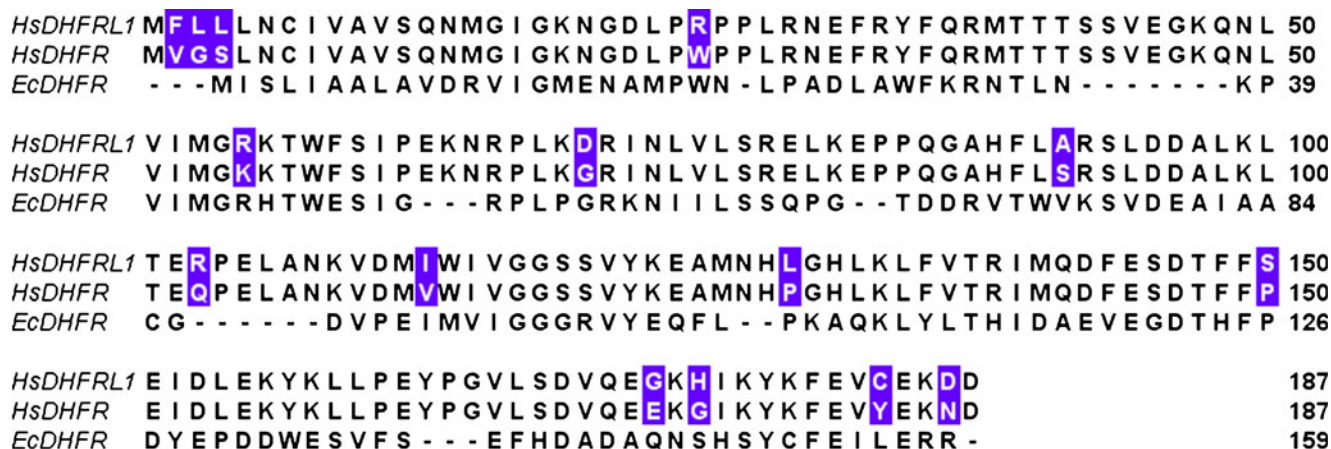


Fig. 1 Sequence alignment of DHFRL1, human DHFR, and ecDHFR. The different residues between DHFRL1 and human DHFR are shown in purple

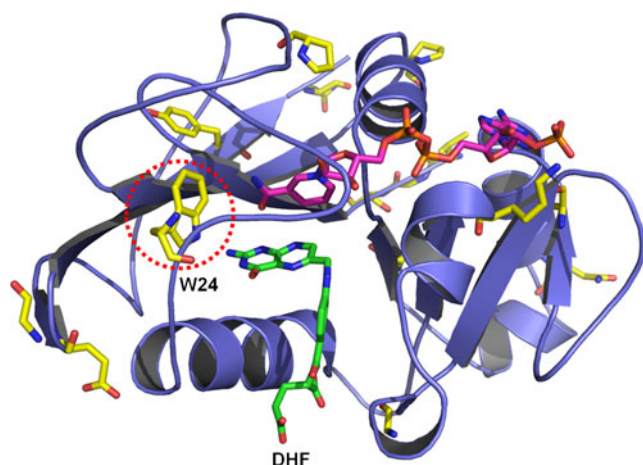


Fig. 2 Illustration of the different residues between DHFRL1 and human DHFR (shown in yellow stick). Only residue 24 is located within the binding site of DHF and is labeled by red dotted line. Figure generated using Pymol

(*ff03*) [33]. The general AMBER force field (*gaff*) [34] was used for the parameter of substrate. The substrate DHF was protonated to form two strong interactions with residue Glu30. Due to lack of electrostatic parameters of substrate in the *gaff* force field, the substrate DHF was firstly minimized using HF/6-31G* optimization with the Gaussian03 program [35], and then the atom partial charges were obtained by fitting the electrostatic potentials derived by the restrained electrostatic potential (RESP) [36] fitting technique in AMBER 9.0. Then, the partial atomic charges and *gaff* force field parameters for DHF were accomplished using the *antechamber* module in AMBER 9.0 [37]. Force field parameter for the cofactor NADPH was adopted from previous simulations [38]. To neutralize the charge of each system, counter ions of Na^+ were placed in grids with the largest negative Coulombic potential around the protein. Last, the whole system was immersed in a rectangular box of TIP3P [39] water molecules, which extended 10 Å away from any solute atoms. The buffer zone of 10 Å is reasonable and acceptable in this study, which will be discussed in the section of results and discussion.

Molecular dynamics simulation

Since the initial structure of each system was taken from crystal structure or simple modification, the substrate DHF and the enzyme along with cofactor NADPH may have partial shift from the initial structure during the MD simulation, relatively long MD simulation of 20 ns was carried out on each system. Before the MD simulation, energy optimization was firstly conducted using the *sander* program in AMBER 9.0 to relax each system via the following three steps: first, the water molecules/ions were optimized by restraining the protein (2500 cycles of steepest descent minimization and 2500 cycles of conjugated gradient minimization); then, the side chains of the protein were relaxed by restraining all

backbone atoms (2500 steps of steepest descent minimization and 2500 steps of conjugated gradient minimization); finally, the whole system was relaxed without any restrain (5000 steps of steepest descent minimization and 5000 steps of conjugated gradient minimization). After that, each system was gradually heated in the NVT ensemble from 0 to 310 K within 60 ps. Last, 20 ns MD simulation with a 2.0 fs time step was performed under a constant temperature of 310 K. The temperature of MD simulation is set to 310 K in view of the enzymes studied are from human. During MD simulation, the SHAKE [40] procedure and particle mesh Ewald (PME) [41] were employed to constrain all bonds involving at least one hydrogen atom and deal with the long-range electrostatic interactions, respectively. In the sampling process, the coordinates were saved every 1 ps for the following binding free energy calculation and binding affinity analysis.

MM/GBSA calculation and MM/GBSA free energy decomposition analysis

In this work, the binding free energy of each system was calculated using the MM/GBSA method via the following equation (1) [42, 43]:

$$\begin{aligned}\Delta G_{\text{bind}} &= G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}} \\ &= \Delta E_{\text{MM}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S \\ &= \Delta E_{\text{vdw}} + \Delta E_{\text{ele}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S\end{aligned}\quad (1)$$

where ΔE_{MM} is the gas-phase interaction energy between protein and ligand, which includes the van der Waals (ΔE_{vdw}) and the electrostatic (ΔE_{ele}) contributions; ΔG_{GB} and ΔG_{SA} are the polar and non-polar components of the desolvation free energy, respectively; $-T\Delta S$ is the entropy contribution at temperature T .

In detail, the polar desolvation free energy was calculated by the generalized born (GB) models developed by Onufriev et al. [44], with dielectric constants of the solvent and the solute were set to 80 and 4, respectively. The non-polar contribution was estimated from the solvate-accessible surface area (SASA) by LCPO method: $\Delta G_{\text{SA}} = 0.0072 \times \Delta \text{SASA}$ [45]. The binding free energy of each system was calculated based on 400 snapshots evenly extracted from 10 to 20 ns MD trajectories. It is necessary to point out that the MM/GBSA calculation is based on the single trajectory approach [46, 47]. Finally, the normal-mode analysis was applied to calculate the changes of conformational entropy upon the ligand binding using the *nmode* module in AMBER9.0 [32]. Due to the extremely time-consuming and high computational demand, only 40 snapshots evenly extracted from 10 to 20 ns were used to estimate the entropy contribution.

To calculate the interaction between ligand and each residue of receptor, MM/GBSA free energy decomposition analysis was carried out using the *mm_pbsa* program in

AMBER9.0 [48]. The binding interaction of each ligand–residue pair consists of three energy terms: van der Waals contribution (ΔE_{vdw}), electrostatic contribution (ΔE_{ele}), and the desolvation term ($\Delta E_{\text{desolvation}}$), which includes the polar (ΔG_{GB}) and the non-polar (ΔG_{SA}) terms. The ΔE_{vdw} and ΔE_{ele} contributions were calculated using the *sander* program in AMBER9.0 [32]. The ΔG_{GB} contribution was calculated by the generalized Born (GB) model, while the non-polar contribution (ΔG_{SA}) was computed according to SASA determined with ICOSA model [48]. All 400 snapshots generated for the binding free energy calculations were also used for the energy decomposition analysis.

Results and discussion

Overall structure and dynamics

As described above, the initial structures of all three systems are based on the crystal structure of human DHFR complex with NADPH and DHF (PDB ID: 2W3M). To explore the conformational dynamics of these three systems, the root mean square deviation (RMSD) of the protein backbone atoms were calculated with the reference of the starting structure (after energy optimization and temperature rise period) (Fig. 3). The RMSD plot indicates that the RMSD values of all three systems achieve equilibrium in a short time and fluctuate around ~ 1.3 Å. Whole conformational changes of all three complexes are little in view of the smaller RMSD (~ 1.3 Å). It is noteworthy that the RMSD of human DHFR is slightly smaller than that of DHFRL1 and its mutant, which may be due to the fact that the initial structures of both DHFRL1 and its mutant are constructed based on the human DHFR, and the mutated residues may cause partial conformational adjustments.

Moreover, the RMSD of the binding site was also calculated to further support the thermodynamic stabilities and the results were shown in the Supplementary data. It can be seen that the residues of binding site remain stable during the MD simulation for the three systems (Fig. S1). As mentioned above, the buffer zone was set to 10 Å in this study, rather

than 12 Å in general. The total energy of each system was calculated to check the thermodynamic stabilities and further validate the buffer zone value. As shown in Fig. S2, the total energy of each system remains constant during MD simulation, which implies the buffer zone of 10 Å is reasonable. Nevertheless, all three systems are well equilibrated after 10 ns, so it is reasonable to do the binding free energy calculation and free energy decomposition based on the trajectories from 10 to 20 ns.

Binding free energies estimated by MM/GBSA

MM/GBSA free energy calculation was employed to estimate the binding affinities of ligand with enzymes, and the results were based on the trajectories extracted from 10 to 20 ns. Table 1 shows the binding free energies and the energy components of three systems. Previous studies have shown that the K_m value of DHF for human DHFR and DHFRL1 are 20.1 and 209.3 μM , respectively [9]. The corresponding experimental binding free energies are -6.41 and -5.02 kcal mol^{-1} , respectively. These data imply that DHF has stronger binding affinity to human DHFR than DHFRL1. The calculated binding free energy of DHF to human DHFR is -42.40 kcal mol^{-1} , stronger than the value of DHFRL1 (-38.92 kcal mol^{-1}). The calculated binding free energy of R24W DHFRL1 mutant decreases to -44.15 kcal mol^{-1} , similar to the value of human DHFR (-42.40 kcal mol^{-1}) rather than DHFRL1 (-38.92 kcal mol^{-1}). This means that the mutation of R24W could considerably recover the binding affinity of DHF, which is in accord with previous hypothesis [9].

It is obvious that the predicted binding free energies can correctly rank the binding affinities of ligands, but the predicted values are much stronger than the experimental ones. MM/GBSA have made reasonable predictions for some systems, while the absolute binding free energy calculated by this method usually have large deviation from the experimental value [14, 46, 49]. Our overestimated binding affinities may be due to the fact that the snapshots from single trajectory of MD simulation were used to calculate the binding affinities, which tends to exaggerate the binding free energy values [46, 47]. Moreover,

Fig. 3 Time evolution of the backbone RMSD of the complexes with respect to the first snapshots

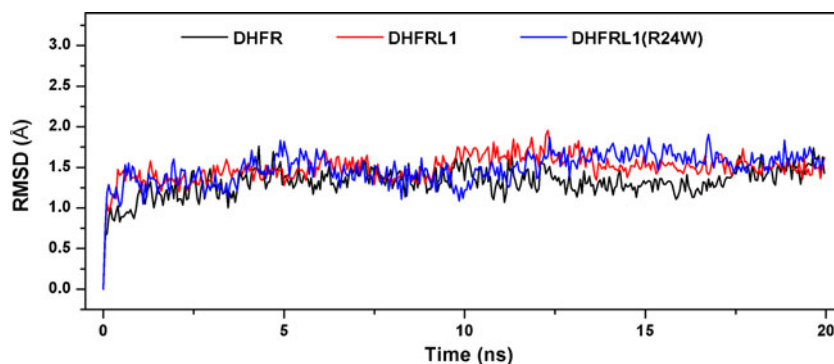


Table 1 Binding free energy of components of the three systems calculated by MM/GBSA (all values are given in kcal mol⁻¹)

System	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{SA}	$\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$	$T\Delta S$	ΔG_{bind}
Human DHFR	-48.63±5.26	-40.74±5.44	25.54±3.93	-7.17±0.19	-15.20±2.63	-28.59±2.58	-42.40±4.33
DHFRL1	-49.19±4.04	-38.43±4.57	27.81±3.75	-6.93±0.15	-10.62±1.87	-27.82±2.75	-38.92±4.01
R24W DHFRL1	-49.14±4.13	-33.07±4.31	18.84±3.55	-7.11±0.27	-14.23±2.18	-26.34±2.65	-44.15±4.26

the ligand DHF is protonated in this study so that the pterin ring of DHF can form two stable interactions with residue Glu30 (one is hydrogen bond, and another is ion-ion interaction, rather than two hydrogen bonds in the crystal structure). In this way, the interactions of pterin ring and the residue Glu30 may enlarge, and ultimately the total binding affinities of ligand may be overestimated. The aim of our study is to rank the binding affinities of ligands and investigate the different binding modes in three systems, rather than give the accurate predictions of the absolute binding free energies. Previous studies also showed that MM/GBSA is usually used to rank the binding affinity for some systems [46]. Thus the predicted binding free energies in this study may be suitable to rank the binding affinity of ligand and investigate the following binding modes analysis.

To further investigate the contribution of each energy term to the binding affinity, the four individual energy components (ΔE_{vdw} , ΔE_{ele} , ΔG_{GB} , and ΔG_{SA}) were carefully compared. The van der Waals and non-polar contributions of DHF binding to human DHFR are -48.63 and -7.17 kcal mol⁻¹, respectively, similar to DHFRL1 (-49.19 and -6.93 kcal mol⁻¹), but the net electrostatic ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) contributions of DHF binding to human DHFR and DHFRL1 are different with the corresponding values of -15.20 and -10.62 kcal mol⁻¹. It is obvious that the van der Waals interaction are more crucial for the total DHF binding energies, while the difference between the DHF binding energies of human DHFR and DHFRL1 can be attributed to the net electrostatic ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) interactions. In addition, the contributions of conformational entropy ($T\Delta S$) of human DHFR and DHFRL1 are -28.59 and -27.82 kcal mol⁻¹, respectively, the contributions of conformational entropy is also minor. As for the R24W DHFRL1 mutant, the van der Waals and non-polar contributions are -49.14 and -7.11 kcal mol⁻¹, respectively, similar to the values of wild-type DHFRL1 (-49.19 and -6.93 kcal mol⁻¹).

More interestingly, the increased binding affinity of DHF to the R24W DHFRL1 mutant can also be attributed to the net electrostatic ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) contributions with the value changing from -10.62 to -14.23 kcal mol⁻¹. Meanwhile, the values of conformational entropy ($T\Delta S$) for the three systems are very close. As suggested by above results, the loss of DHF binding to DHFRL1 may be due to the difference of residues at position 24 and the reduced binding energies arise from the decrease of net electrostatic ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) contributions. The following free energy decomposition and binding mode analysis will elucidate the corresponding structural features of

the binding regions around the 24th residue in DHFRL1 enzyme.

MM/GBSA free energy decomposition analysis

To better characterize the detailed binding mode between enzyme and ligand, MM/GBSA free energy decomposition analysis was employed to decompose the total binding free energies into ligand-residue pairs, which would provide more quantitative information about the contribution of each residue (shown in Fig. 4). The residues Ile7, Val8, Ala9, Glu30, Phe31, Arg32, Phe34, Gln35, Ile60, Asn64, Leu67, and Val115 of human DHFR together with cofactor NADPH, have strong interactions with the ligand DHF (Fig. 4a). The key residues captured from the MM/GBSA free energy decomposition are in agreement with the contacted residues observed in the crystal structure (PDB ID: 2W3M). In the case of DHFRL1, the residues Ile7, Val8, Ala9, Arg28, Glu30, Phe31, Phe34, Ile60, Asn64, Leu67, Lys68, Arg70, Val115, and cofactor NADPH, are key residues for DHF binding (Fig. 4b). It can be seen that the contributions of Arg24, Phe31, Arg32, Gln35, Leu67, Lys68, and Arg70 in DHFRL1 are distinct from human DHFR. To further obtain the detailed information, the van der Waals and net electrostatic interactions between ligand and residue are separately calculated and the results are shown in Fig. 4d and e. There are four residues in DHFRL1 which have considerably different van der Waals interactions with ligand: Arg24, Phe31, Gln35 and Leu67 (Fig. 4d). Specifically, van der Waals interactions between DHF and residues Phe31, Gln35 in human DHFR (-3.21 and -1.21 kcal mol⁻¹) are stronger than those in DHFRL1 (-2.34 and -0.36 kcal mol⁻¹), while the interactions between DHF and residues Arg24 and Leu67 in human DHFR (-0.15 and -1.64 kcal mol⁻¹) are weaker than those in DHFRL1 (-0.99 and -2.18 kcal mol⁻¹). On the whole, the total van der Waals interactions of DHF with human DHFR are close to DHFRL1.

On the other hand, there are three residues in DHFRL1 that were identified to have different net electrostatic interactions with DHF: Arg28, Arg32, and Gln35 (Fig. 4e). The net electrostatic interactions of Arg32 and Gln35 in human DHFR are -1.01 and -0.7 kcal mol⁻¹, respectively, which are stronger than the corresponding values in DHFRL1 (-0.16 and -0.25 kcal mol⁻¹). However, the residue Arg28 in human DHFR (-0.79 kcal mol⁻¹) is weaker than in DHFRL1 (-1.35 kcal mol⁻¹). These data indicate that the loss of DHF affinity observed in DHFRL1 is mainly attributed to

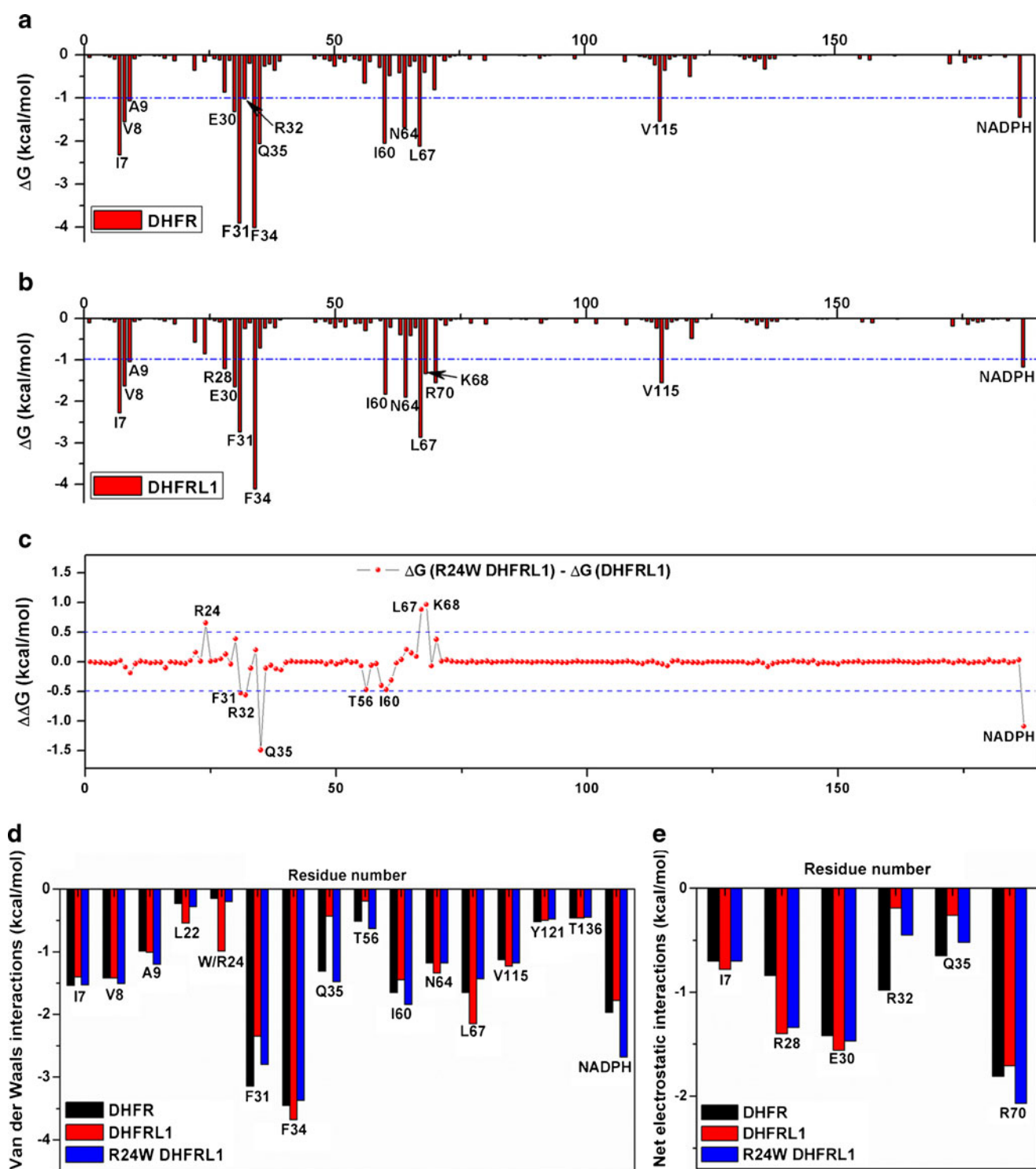


Fig. 4 Decomposition of the total binding free energies per residue for human DHFR (a) and DHFRL1 (b). The difference of free energy decomposition between R24W DHFRL1 and DHFRL1 (c). The van der Waals (d) and net electrostatic (e) interaction spectrums of three systems

the reduction of the net electrostatic interaction, especially for the residues Arg32 and Gln35.

Moreover, the MM/GBSA free energy decomposition analysis of the R24W DHFRL1 mutant was also studied. To better

understand the impact of amino acid mutations on the binding mode of DHF, the differences of the contribution of each residues between R24W DHFRL1 mutant and DHFRL1 were calculated and the results are depicted in Fig. 4c. The R24W

mutation largely increases the interactions of DHF with the residues Phe31, Arg32, Gln35, Thr56, Ile60, and NADPH, while moderately decreases the interactions with residues Arg24, Leu67, and Lys68. More specifically, the R24W mutation enhances van der Waals contributions of the residues Phe31, Gln35, Thr56, and Ile60 but decreases those of the residues Arg24 and Leu67 (Fig. 4d). Meanwhile, the mutation also increases net electrostatic contributions of the residues Arg32 and Gln35 (Fig. 4e). As mentioned above, the differences of binding modes between human DHFR and DHFRL1 can also be attributed to these residues (Fig. 4d and e). The binding of DHF to R24W DHFRL1 mutant resembles to human DHFR. In other words, the R24W mutation may induce significant change of the structure of binding site in DHFRL1 so as to recover the binding mode of DHF in human DHFR.

Binding modes of DHF in DHFR, DHFRL1, and DHFRL1 mutant

To further investigate the binding modes of three systems, the last snapshot of each system is shown in Fig. 5. As indicated from Fig. 5a and b, although DHF is well accommodated in the binding pocket of DHFRL1, its corresponding binding

mode is apparently different from that of human DHFR. The obvious difference is the conformation of pterin ring of DHF. In detail, pterin ring of DHF in human DHFR has strong hydrogen bond interactions with Ile7, Val8, and Glu30 (Fig. 5a), while pterin ring of DHF in DHFRL1 has hydrogen bond interaction with Val8, Arg24, Glu30, and Thr136 (Fig. 5b). Moreover, two new hydrogen bonds are created between Arg24 and NADPH (Fig. 5b) in the case of DHFRL1. This may be interpreted by the fact that the side chain of Arg24 extends deeply into the binding pocket of DHFRL1, and disturbs the normal hydrogen bond networks. In addition, the hydrogen bond interactions of DHF in R24W DHFRL1 mutant are very similar to human DHFR (Fig. 5c).

The differences of hydrogen bond interactions are probably caused by the structural differences of the side chain of residue 24. The side chain of Arg24 is not only longer than that of Trp24, but also more flexible than Trp24. To elucidate the possible effect of residue 24 on hydrogen bond network, the distances between the side chain of residue 24 and DHF, NADPH were separately calculated (Fig. 5d and e). Figure 5d depicts the case of DHF (Trp24@NE1-DHF@O1 for human DHFR and R24W DHFRL1; Arg24@CZ-DHF@O1 for DHFRL1), and Fig. 5e shows the case of

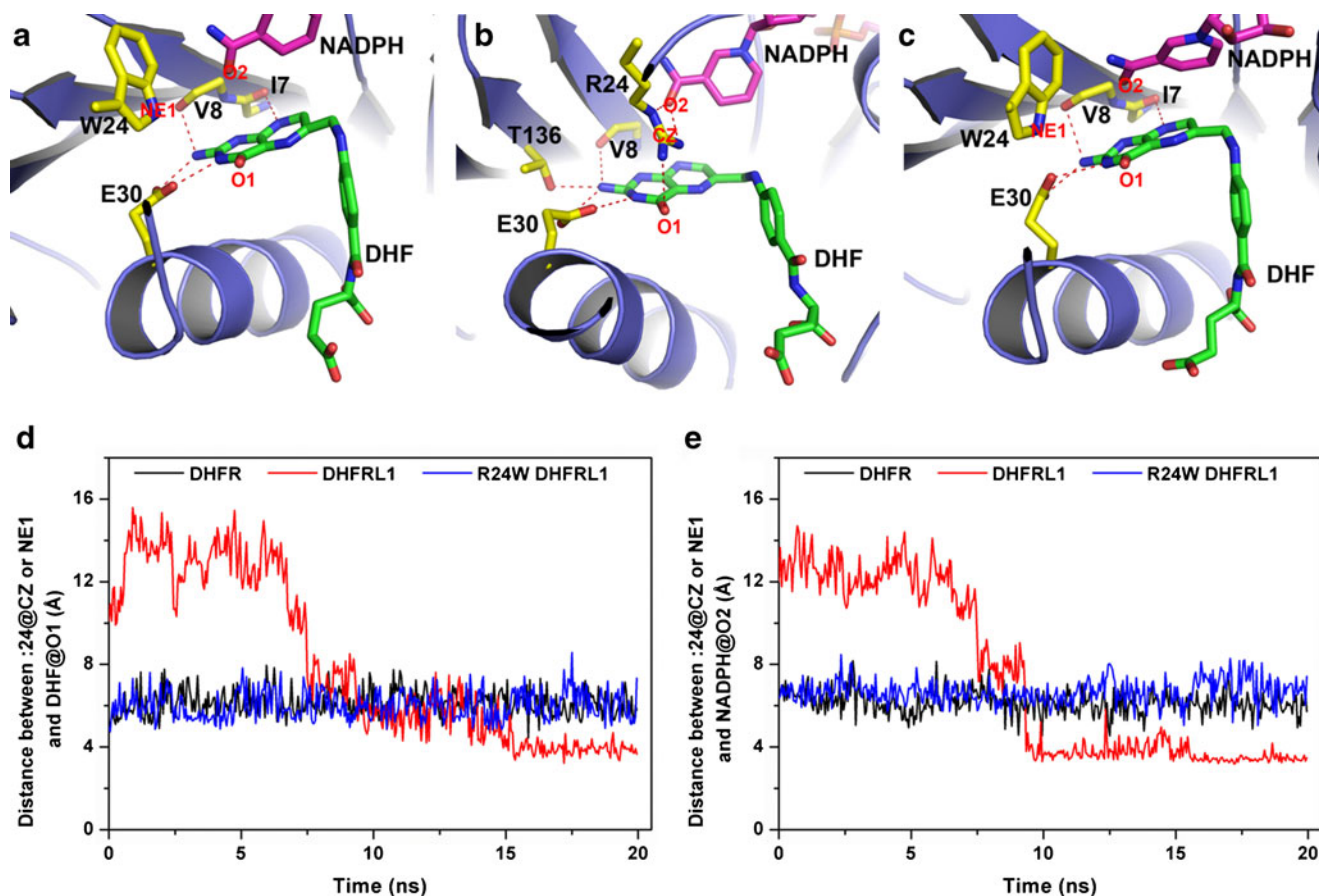


Fig. 5 Illustration of the interactions of DHF with DHFR (a), DHFRL1 (b), and R24W DHFRL1 (c). The distances between the side chain of residue 24 and DHF (d) and NADPH (e). Figures generated using Pymol

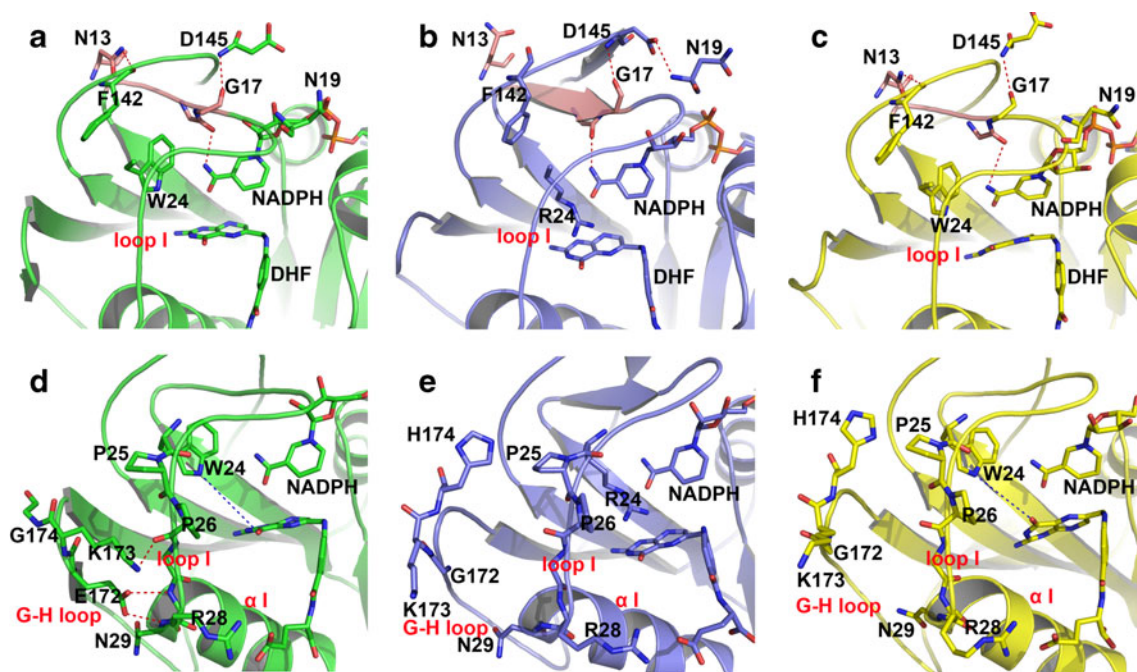


Fig. 6 Illustration of the hydrogen bond interactions between the N-terminal of loop I and its neighboring loop in human DHFR (**a**), DHFRL1 (**b**), and R24W DHFRL1 (**c**), and the corresponding ones of

the C-terminal of loop I with G-H loop in human DHFR (**d**), DHFRL1 (**e**), and R24W DHFRL1 (**f**). Figures generated using Pymol

NADPH (Trp24@NE1-NADPH@O2 for human DHFR and R24W DHFRL1; Arg24@CZ-NADPH@O2 for DHFRL1). In the cases of human DHFR and R24W DHFRL1 mutant, the distances between the NE1 atom of Trp24 and two oxygen atoms O1 (DHF), O2 (NADPH) remain nearly constant around 6 Å during the entire 20 ns MD simulation, indicating that the side chain of Trp24 do not form specific contacts with DHF and NADPH.

As for DHFRL1, the side chain of Arg24 could extend deeply into the binding sites of DHF and NADPH as shown in

Fig. 5d and e. The distances between CZ atom of Arg24 and the two oxygen atoms (O1 and O2) largely fluctuate around 12 Å at the first 7 ns, then quickly decrease to 8 Å within 2 ns, and finally achieve equilibrium at 10 ns with an average value of 3.9 Å. These data well support the hypothesis that Arg24 forms strong hydrogen bond interactions with DHF and NADPH. In addition, Arg24 becomes stable in the binding pocket of DHFRL1 after 10 ns, which implies that the above binding free energy calculation based on the snapshots extracted from 10 to 20 ns is reasonable and acceptable.

Table 2 Hydrogen bond analyses of loop I and its neighboring loop calculated from last 10 ns trajectories

Donor	Acceptor	Human DHFR		DHFRL1		DHFRL1(R24W)	
		Occupied (%)	Distance (Å)	Occupied (%)	Distance (Å)	Occupied (%)	Distance (Å)
:I16@O	:NADPH@H72:-NADPH@N7N	67.44	2.97±0.12	96.04	2.88±0.12	69.90	2.95±0.13
:F142@O	:N13@HD21:-N13@ND2	78.42	2.94±0.13	0.02	2.82±0.01	81.06	2.95±0.12
:N13@OD1	:F142@H:-F142@N	26.44	3.00±0.12	—	—	17.02	3.04±0.11
:G17@O	:E145@H:-E145@N	88.58	2.91±0.12	98.20	2.87±0.11	94.38	2.92±0.12
:E145@OD1	:N19@H:-N19@N	—	—	16.10	2.93±0.12	—	—
:E145@OD2	:N19@H:-N19@N	—	—	7.76	2.97±0.13	—	—
:P26@O	:K173@HZ1:-K173@NZ	16.30	2.88±0.13	—	—	0.60	2.89±0.14
:P26@O	:K173@HZ2:-K173@NZ	15.74	2.87±0.13	—	—	1.14	2.89±0.14
:P26@O	:K173@HZ3:-K173@NZ	12.54	2.88±0.13	—	—	0.42	2.89±0.16
:E172@OE1	:N29@H:-N29@N	69.94	2.93±0.12	—	—	—	—
:E172@OE2	:R28@H:-R28@N	64.62	2.93±0.13	—	—	—	—

Conformations of loop I in human DHFR, DHFRL1, and DHFRL1 mutant

It has been well accepted that loop I, residues 11–27 (equivalent to the M20 loop of *Escherichia coli* DHFR) [50–52] plays an important role in enzyme's catalysis and substrate binding. The residue at position 24 is also located in loop I. It will be very interesting to study the possible conformations of loop I in DHFRL1 and its mutant. Compared to *Escherichia coli* DHFR, human DHFR appears to be more rigid, especially for loop I. The conformation of loop I in human DHFR is invariably closed [50]. To better characterize the conformation of loop I, hydrogen bond analysis was applied on the three systems based on the last 10 ns trajectories, and the results are shown in Fig. 6 and Table 2. For each complex, the N-terminal (residues 11–19) and C-terminal (residues 20–27) of loop I are depicted, respectively. For human DHFR two hydrogen bonds are formed between Asn13 and Phe142, and another one hydrogen bond is formed between Asp145 and Gly17 (Fig. 6a and Table 2). The three hydrogen bonds maintain the conformation of the N-terminal of loop I. It can be seen that similar hydrogen bond patterns are found in the R24W DHFRL1 mutant (Fig. 6c and Table 2). As for DHFRL1, besides the above three hydrogen bonds, a new one is formed between Asp145 and Asn19, with hydrogen bond occupancies of 16.10 and 7.76 %, respectively (Fig. 6b and Table 2). It is likely that the conformation of N-terminal of loop I in DHFRL1 is similar to the normal closed state of *Escherichia coli* DHFR [50] due to the additional hydrogen bond formed by Asp145 and Asn19 (corresponding residues in *Escherichia coli* DHFR is Asp122 and Glu17). Moreover, the new hydrogen bond might facilitate the formation of a short β -sheet (residues 13–17) in DHFRL1 (Fig. 6b). In addition, Gly17 forms a strong hydrogen bond with the nicotinamide group of NADPH, and the corresponding hydrogen bond occupancies in human DHFR, DHFRL1, and R24W DHFRL1 are 67.44, 96.04, and 69.90 %, respectively.

The conformations of C-terminal of loop I were also studied. The C-terminal of loop I of the human DHFR has weak hydrogen bond interaction with G-H loop (residues 159–174) via Pro26 and Lys173 (Fig. 6d and Table 2). It is in good agreement with previous studies [50, 51]. As for DHFRL1 and R24W DHFRL1 mutant (Fig. 6 e and f), the C-terminal of loop I is disengaged from G-H loop. It may be attributed to the fact that the residue 172 in human DHFR is glutamic acid, which has strong hydrogen bond interaction with two residues (Arg28 and Asn29) in α -helical segment α I (residues 28–39), while the corresponding residue at the position of 172 in DHFRL1 and its R24W mutant is glycine. The interactions of G-H loop with α I in human DHFR help narrow the distance between G-H loop and the C-terminal of loop I. Importantly, the C-terminal of loop I in all three systems consists of a 'XPXX' fragment (X represents Trp24 in human DHFR, or

Arg24 in DHFRL1), which can increase the rigidity of loop I. Meanwhile, in the cases of human DHFR and R24W DHFRL1 mutant, the distance between the side chain of Trp24 and the oxygen atom (O1) of DHF is almost the same (Fig. 5d), which also implies the C-terminal of loop I is rigid enough and may not be affected by the G-H loop. Taking the two terminals of loop I together, the conformation of loop I of DHFRL1 is very similar to normal closed state of *Escherichia coli* DHFR other than the closed state of human DHFR.

Effect of residue 24 on the DHF binding

It is clear that the binding modes of DHF in three systems are closely related with residue 24. To identify structural

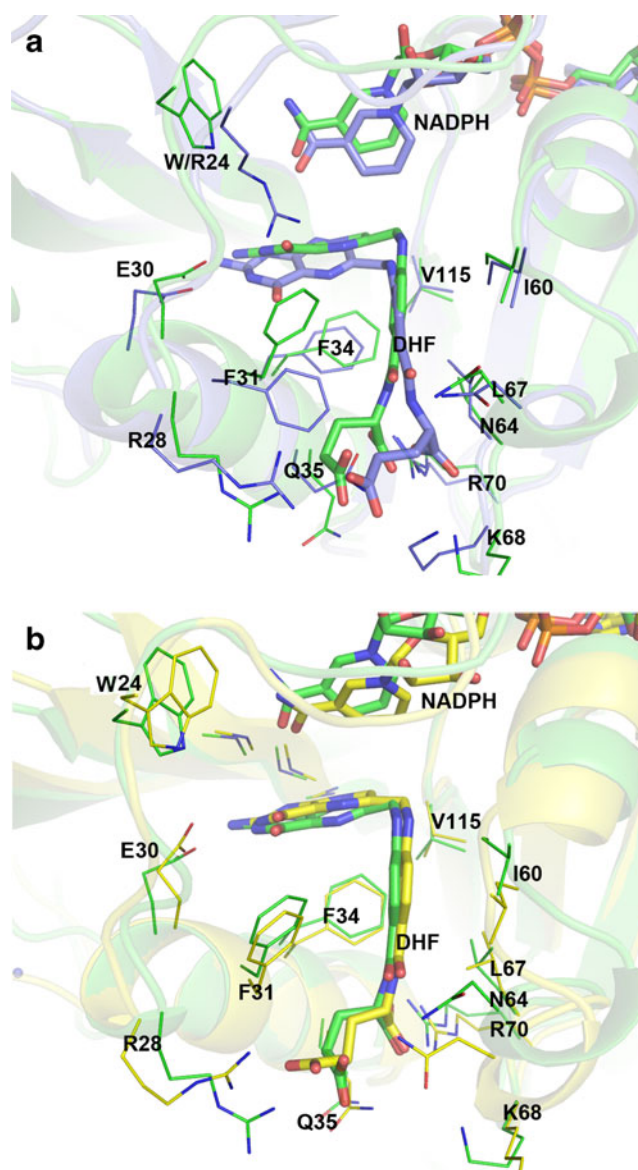


Fig. 7 The superposition of the binding modes of DHF in DHFR and DHFRL1 (a), and in DHFR and R24W DHFRL1 (b). Figure generated using Pymol

differences and elucidate possible effect of residue 24 on DHF binding, the last snapshots of DHF in three systems were superimposed (Fig. 7). It can be seen that the binding mode of DHF in each complex agrees well with the results from binding free energy decomposition analysis (Fig. 4a, b, and c). The Arg24 in DHFRL1 inserts into the binding sites of DHF and NADPH and causes a significant steric hindrance on DHF binding. More specifically, the hydrogen bonds formed by Arg24 and NADPH destabilize the pterin ring of DHF so that the latter is rotated and close to Phe34 (Fig. 7a). Meanwhile, the glutamate moiety of DHF is switched near to the residues Asn64, Leu67, Lys68, and Arg70, and far away from the residues Gln35 and Phe31 (Fig. 7a). The differences of the DHF binding in two enzymes also agree with the results from the MM/GBSA free energy decomposition analysis. For example, Phe31 in human DHFR is close to DHF with a strong binding free energy contribution ($-3.9 \text{ kcal mol}^{-1}$), while Phe31 in DHFRL1 is far from DHF with a decreased binding free energy contribution ($-2.74 \text{ kcal mol}^{-1}$) (Fig. 4a and b). The contribution of Gln35 in human DHFR ($-2.12 \text{ kcal mol}^{-1}$) is stronger than that of DHFRL1 ($-0.81 \text{ kcal mol}^{-1}$).

Superimposition of R24W DHFRL1 mutant and human DHFR (Fig. 7b) shows that the binding modes of DHF in these two complexes are almost the same excepting a little rotation of the terminal of glutamate moiety of DHF. In other words, although there are other different residues existed between human DHFR and R24W DHFRL1, the mutation at residue 24 may recover the binding affinity of DHF to DHFRL1.

Conclusions

In this work, combination of different molecular modeling techniques were applied to study the molecular mechanism of DHFRL1 with lowered DHF binding affinity. The initial structure of DHFRL1 complex with DHF was constructed based on the crystal structure of human DHFR complex with DHF. In addition, to investigate the possible impact of residue at position 24 on DHF binding, the R24W DHFRL1 mutant was designed. Twenty ns molecular dynamics simulation was performed on the systems of human DHFR, DHFRL1, and R24W DHFRL1, and the results show that the van der Waals interaction is more crucial for the total binding energies, but the net electrostatic contributions ($\Delta E_{ele} + \Delta G_{GB}$) are responsible for the different binding affinities of DHF between human DHFR and DHFRL1. In the case of R24W DHFRL1 mutant, the binding affinity of DHF largely recovered. And the calculated binding free energy of DHF to mutant ($-44.15 \text{ kcal mol}^{-1}$) is similar to the value of human DHFR ($-42.40 \text{ kcal mol}^{-1}$) rather than DHFRL1 ($-38.92 \text{ kcal mol}^{-1}$). Interestingly, the increased binding affinity of DHF with R24W DHFRL1 can be attributed to the enhanced net

electrostatic interactions. According to the free energy decomposition and structure analysis, the loss of DHF affinity in DHFRL1 is mainly characterized by the reduction of net electrostatic interactions of residues Arg32 and Gln35. It is interesting to find that the R24W DHFRL1 mutant would increase the net electrostatic interactions, especially for residues Arg32 and Gln35. Moreover, hydrogen bond analysis reveals that the conformation of loop I in DHFRL1 may be similar to normal closed state of *Escherichia coli* DHFR rather than the closed state of human DHFR.

Structure analysis also reveals that the distinction of binding modes between human DHFR and DHFRL1 may attribute to the different behaviors of residue 24. The side chain of residue 24 in DHFRL1 may extend deeply into the binding sites of DHF and NADPH, then disturb the normal hydrogen bond networks, and finally affect the DHF binding by steric hindrance. As for the R24W DHFRL1 mutant, the steric effect would disappear, and the binding affinity of DHF in DHFRL1 would recover. We hope this work gives some valuable information on understanding the molecular mechanism of DHF binding to DHFRL1.

Acknowledgments We thank Prof. Jingyuan Li for helpful discussions. This work was supported by the Natural Science Foundation of China (No. 21173264) and the Foundation of Knowledge Innovative Engineering of Chinese Academy of Sciences (No. ZNWH-2011-011).

References

- Roth B, Aig E, Rauckman BS, Strelitz JZ, Phillips AP, Ferone R, Bushby SR, Sigel CW (1981) 2,4-Diamino-5-benzylpyrimidines and analogues as antibacterial agents. 5. 3',5'-Dimethoxy-4'-substituted-benzyl analogues of trimethoprim. *J Med Chem* 24:933–941
- Bertino JR, Sobrero A, Mini E, Moroson BA, Cashmore A (1987) Design and rationale for novel antifolates. *NCI Monogr* 5:87–91
- Plowe CV, Kublin PG, Dzinjalimala FK, Kamwendo DS, Mukadam RAG, Chimpeni P, Molyneux ME, Taylor TE (2004) Sustained clinical efficacy of sulfadoxine-pyrimethamine for uncomplicated falciparum malaria in Malawi after 10 years as first line treatment: five year prospective study. *Brit Med J* 328:545–548
- Fleming GF, Schilsky RL (1992) Antifolates: the next generation. *Semin Oncol* 19:707–719
- Bertino JR (2009) Cancer research: from folate antagonism to molecular targets. *Best Pract Res Clin Haematol* 22:577–582
- Anderson AC (2005) Targeting DHFR in parasitic protozoa. *Drug Discov Today* 10:121–128
- Benkovic SJ (1980) On the mechanism of action of folate- and bipterin-requiring enzymes. *Annu Rev Biochem* 49:227–251
- Maurer BJ, Barker PE, Masters JN, Ruddle FH, Attardi G (1984) Human dihydrofolate reductase gene is located in chromosome 5 and is unlinked to the related pseudogenes. *Proc Natl Acad Sci U S A* 81: 1484–1488
- McEntee G, Minguzzi S, O'Brien K, Ben Larbi N, Loscher C, O'Fagain C, Parle-McDermott A (2011) The former annotated human pseudogene dihydrofolate reductase-like 1 (DHFRL1) is expressed and functional. *Proc Natl Acad Sci U S A* 108:15157–15162

10. Anderson DD, Quintero CM, Stover PJ (2011) Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proc Natl Acad Sci U S A* 108:15163–15168
11. Chunduru SK, Cody V, Luft JR, Pangborn W, Appleman JR, Blakley RL (1994) Methotrexate-resistant variants of human dihydrofolate reductase. Effects of Phe31 substitutions. *J Biol Chem* 269:9547–9555
12. Volpato JP, Yachnin BJ, Blanchet J, Guerrero V, Poulin L, Fossati E, Berghuis AM, Pelletier JN (2009) Multiple conformers in active site of human dihydrofolate reductase F31R/Q35E double mutant suggest structural basis for methotrexate resistance. *J Biol Chem* 284:20079–20089
13. Lewis WS, Cody V, Galitsky N, Luft JR, Pangborn W, Chunduru SK, Spencer HT, Appleman JR, Blakley RL (1995) Methotrexate-resistant variants of human dihydrofolate reductase with substitutions of leucine 22. Kinetics, crystallography, and potential as selectable markers. *J Biol Chem* 270:5057–5064
14. Rastelli G, Rio AD, Degliesposti G, Sgobba M (2010) Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J Comput Chem* 31:797–810
15. Cummins PL, Gready JE (2000) QM/MM and SCRF studies of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase: existence of a low-barrier hydrogen bond. *J Mol Graph Model* 18:42–49
16. Tosso RD, Andujar SA, Gutierrez L, Angelina E, Rodriguez R, Noguera M, Baldoni H, Suvire FD, Cobo J, Enriz RD (2013) Molecular modeling study of dihydrofolate reductase inhibitors. molecular dynamics simulations, quantum mechanical calculations, and experimental corroboration. *J Chem Inf Model*. doi:10.1021/ci400178h
17. Oliveira AA, Rennó MN, de Matos CAS, Bertuzzi MD, Ramalho TC, Fraga CAM, França TCC (2011) Molecular modeling studies of *Yersinia pestis* dihydrofolate reductase. *J Biomol Struct Dyn* 29:351–367
18. Gorse AD, Gready JE (1997) Molecular dynamics simulations of the docking of substituted N5-deazapterins to dihydrofolate reductase. *Protein Eng* 10:23–30
19. Choowongkamon K, Theppabutr S, Songtawee N, Day N, White N, Woodrow C, Imwong M (2010) Computational analysis of binding between malarial dihydrofolate reductases and anti-folates. *Malar J* 9:65
20. Fan Y, Cembran A, Ma S, Gao J (2013) Connecting protein conformational dynamics with catalytic function as illustrated in dihydrofolate reductase. *Biochemistry* 52:2036–2049
21. Schnell JR, Dyson HJ, Wright PE (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct* 33:119–140
22. Swanwick RS, Shrimpton PJ, Allemann RK (2004) Pivotal role of Gly 121 in dihydrofolate reductase from *Escherichia coli*: the altered structure of a mutant enzyme may form the basis of its diminished catalytic performance. *Biochemistry* 43:4119–4127
23. Beard WA, Appleman JR, Huang S, Delcamp TJ, Freisheim JH, Blakley RL (1991) Role of the conserved active site residue tryptophan-24 of human dihydrofolate reductase as revealed by mutagenesis. *Biochemistry* 30:1432–1440
24. Thillet J, Absil J, Stone SR, Pictet R (1988) Site-directed mutagenesis of mouse dihydrofolate reductase. Mutants with increased resistance to methotrexate and trimethoprim. *J Biol Chem* 263:12500–12508
25. Wang W, Kollman PA (2000) Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *J Mol Biol* 303:567–582
26. Kuhn B, Kollman PA (2000) Binding of a diverse set of ligands to Avidin and Streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J Med Chem* 43:3786–3791
27. Hou T, Zhu L, Chen L, Xu X (2002) Mapping the binding site of a large set of quinazoline type EGF-R inhibitors using molecular field analyses and molecular docking studies. *J Chem Inf Comput Sci* 43:273–287
28. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras–Raf. *J Comput Chem* 25:238–250
29. Hou T, Zhang W, Case DA, Wang W (2008) Characterization of domain-peptide interaction interface: a case study on the amphiphysin-1 SH3 domain. *J Mol Biol* 376:1201–1214
30. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242
31. SYBYL molecular simulation package. 2004. <http://www.sybyl.com>.
32. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
33. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24:1999–2012
34. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
35. Frisch MJ TG, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz J, Komaromi I, Martin RL, Fox DJ, Keith T, Laham A, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) Gaussian 03, revision C.02. Wallingford
36. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 97:10269–10280
37. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25:247–260
38. Cummins PL, Ramnarayan K, Singh UC, Gready JE (1991) Molecular dynamics/free energy perturbation study on the relative affinities of the binding of reduced and oxidized NADP to dihydrofolate reductase. *J Am Chem Soc* 113:8247–8256
39. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
40. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341
41. Essmann U, Perera L, Berkowitz M, Darden T, Lee H, Pedersen L (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593
42. Wang J, Hou T, Xu X (2006) Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Curr Comput Aided Drug Des* 2:287–306
43. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 33:889–897
44. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 55:383–394

45. Weiser J, Shenkin PS, Still WC (1999) Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J Comput Chem* 20:217–230
46. Hou T, Wang J, Li Y, Wang W (2010) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 51:69–82
47. Lee J, Kim J-S, Seok C (2010) Cooperativity and specificity of Cys2His2 zinc finger protein-DNA interactions: a molecular dynamics simulation study. *J Phys Chem B* 114:7662–7671
48. Gohlke H, Kiel C, Case DA (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J Mol Biol* 330:891–913
49. Bag S, Tawari NR, Degani MS, Queener SF (2010) Design, synthesis, biological evaluation and computational investigation of novel inhibitors of dihydrofolate reductase of opportunistic pathogens. *Bioorg Med Chem* 18:3187–3197
50. Sawaya MR, Kraut J (1997) Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. *Biochemistry* 36:586–603
51. Davies JF, Delcamp TJ, Prendergast NJ, Ashford VA, Freisheim JH, Kraut J (1990) Crystal structures of recombinant human dihydrofolate reductase complexed with folate and 5-deazafolate. *Biochemistry* 29: 9467–9479
52. Oefner C, D'Arcy A, Winkler FK (1988) Crystal structure of human dihydrofolate reductase complexed with folate. *Eur J Biochem* 174: 377–385