informa
healthcare

ORIGINAL ARTICLE

A common feature-based 3D-pharmacophore model generation and virtual screening: identification of potential *Pf*DHFR inhibitors

Legesse Adane, Prasad V. Bharatam, and Vikas Sharma

National Institute of Pharmacuetical Education and Research (NIPER), S.A.S. Nagar, Mohali, India

Abstract
A four-feature 3D-pharmacophore model was built from a set of 24 compounds whose activities were reported against the V1/S strain of the *Plasmodium falciparum* dihydrofolate reductase (*Pf*DHFR) enzyme. This is an enzyme harboring Asn51Ile + Cys59Arg + Ser108Asn + Ile164Leu mutations. The HipHop module of the Catalyst program was used to generate the model. Selection of the best model among the 10 hypotheses generated by HipHop was carried out based on rank and best-fit values or alignments of the training set compounds onto a particular hypothesis. The best model (hypo1) consisted of two H-bond donors, one hydrophobic aromatic, and one hydrophobic aliphatic features. Hypo1 was used as a query to virtually screen Maybridge2004 and NCI2000 databases. The hits obtained from the search were subsequently subjected to FlexX and Glide docking studies. Based on the binding scores and interactions in the active site of quadruple-mutant *Pf*DHFR, a set of nine hits were identified as potential inhibitors.

Keywords: *Pharmacophore model; Catalyst/HipHop; virtual screening; molecular docking; P. falciparum DHFR*

Introduction

Plasmodium falciparum dihydrofolate reductase (*Pf*DHFR) enzyme is one of the several targets in the treatment of malaria[1–5]. This enzyme catalyzes the nicotinamide adenine nucleotide phosphate (NADPH) dependent reduction of dihydrofolate (DHF) to tetrahydrofolate (THF), which is essential for DNA synthesis in the parasitic cell. It has been recognized as a specific target for common antifolate-based antimalarial drugs: pyrimethamine (**1**) and cycloguanil (**2**) (Figure 1). These drugs competitively inhibit DNA synthesis, and ultimately lead to parasitic death. However, due to the accumulation of genetic mutations[6–11] at one or more amino acid residues 16, 51, 59, 108, and 164, the parasite has become highly resistant to these antifolate drugs. Thus, the clinical uses of these drugs have been compromised in many parts of the world.

Steric clashes between a particular inhibitor and mutated amino acid residues result in the displacement of a drug from its optimal orientation in its interaction with residue Asp54, which is crucial for inhibitor binding, with consequent loss of inhibitory activity. For instance, in the parasite harboring Ala16Val + Ser108Thr mutation, one of the methyl groups at the 6-position of **2** experiences a severe steric interaction with Val16. This results in the development of resistance of the parasite to **2**. On the other hand, in the case of a multiple-mutant parasite, steric clashes of the *para*-Cl groups of both **1** and **2** with Asn108 (mutated from Ser108) result in the loss of inhibitory activities of these drugs or increase the chance of the parasite to develop cross-resistance[7]. These observations led to the conception of steric constraint hypothesis. According to this hypothesis, (i) changing the position of the *para*-Cl group to the *meta* position and (ii) the use of

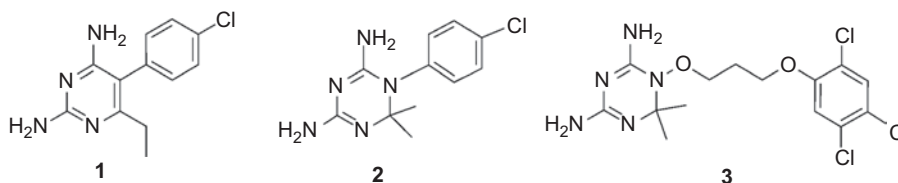http://www.informahealthcare.com/enz

RIGHTS LINK

Figure 1. Chemical structures of pyrimethamine (**1**), cycloguanil (**2**), and WR99210 (**3**).

compounds which are more flexible than **1** and **2** could avoid steric clashes, and subsequently result in agents that effectively inhibit mutant *Pf*DHFRs[7,12–18]. WR99210 (**3**) is one of the typical examples of compounds designed to avoid steric clash with Asn108 of the mutant *Pf*DHFR enzyme (Figure 1). Inhibitory activity tests indicated that this compound has similar activities against wild-type and multiple-mutant parasites carrying the Asn51Ile + Cys59Arg + Ser108Asn + Ile164Leu mutated DHFR enzyme. However, because of toxicity problems, this compound could not advance into clinical stages[19,20]. However, it is still used as a model compound for designing competitive *Pf*DHFR inhibitors of flexible nature.

Virtual screening is a computer-based strategy used to identify new potential lead compounds from large databases of compounds that are subsequently synthesized and tested for their biological activities[21–23]. It has also been employed in the discovery of *Pf*DHFR enzyme inhibitors[7,24]. Rastelli *et al.*[25] employed virtual screening, which led to the identification of 12 new leads whose chemical structures are different from those of classical antifolates. Enzyme assay and inhibition studies indicated that the compounds are generally more active toward mutant enzyme than wild-type *Pf*DHFR enzyme. Recently, Dasgupta *et al.*[26] carried out high-throughput *in silico* screening of a database with consequent *in vitro* enzymatic assay and cellular culture studies. They identified three novel compounds (RJF001302, RJF00670, and RJF00719) of biguanide analogs. The compounds were found to be active against both wild-type and quadruple-mutant *Pf*DHFRs. Inspite of the accumulation of a wealth of information and several efforts, no successful drug has come on to the market to combat malaria. Thus, there is always an urgent and unmet need for new drugs to cure malaria.

In this article, we present some hits obtained from a database search. Catalyst/HipHop software was used to generate the 3D-pharmacophore model from a set of compounds whose antimalarial activities are reported in literature[27]. HipHop, also known as a common-feature hypothesis, is used as a method for 3D-pharmacophore model generation in a training set of compounds having high structural homology and a narrow activity range. In this method, no activity data are required. This is because only chemical features common to all training set molecules are taken into account for model generation[28,29]. A set of nine hits were identified from our screening. Selection of the hits was based on their (i) best-fit values, (ii) binding scores, (iii) binding modes, and (iv) interactions with important amino acid residues in the active site such as Asp54, Leu164, Asn108, and Ile14, which are known to be essential for inhibitory activities of antifolate-based antimalarial drugs. Two docking programs, namely, FlexX[30,31] and Glide[32,33], were used to carry out calculations on the hits obtained from the database search.

Materials and methods

Hypothesis generation and validation

A data set of 24 compounds for which *in vitro* inhibitory activities against the multiple-drug esistant (V1/S strain) *Pf*DHFR enzyme reported in the literature were used[27]. The chemical structures of the molecules were constructed using the 2D/3D-editor sketcher of Catalyst, and standard 3D-structures were generated and minimized to the nearest local minimum using the molecular mechanics CHARMm force field implemented in the Catalyst program[34]. In order to represent flexibility, a collection of conformational models were generated for each compound. These models were generated automatically, starting from the local minimized structures, using a "polling" algorithm[35,36]. The "best conformer generation" option of the Catalyst software was employed. The default value of 255 was used for the maximum number of conformer generations per molecule. All the conformers within a range of 20 kcal/mol, with respect to the global minimum, were employed to build a set of pharmacophore hypotheses. Pharmacophore model generation was carried out using the HipHop module of Accelrys Catalyst software (version 4.10)[34]. The training set molecules, with their associated conformational models, were rearranged into a spreadsheet and subsequently submitted to Catalyst hypothesis generation using the HipHop module. HipHop accommodates hypotheses with up to 10 different features which are necessary for biological activities of compounds, and are distributed within a three-dimensional (3D) space. Chemical features that are surface-accessible and common in hypothesis generation include hydrogen bond donor, hydrogen bond acceptor, ring aromatic, hydrogen bond acceptor lipid, hydrophobic, hydrophobic aromatic, hydrophobic aliphatic, negative ionizable, positive ionizable, negative charge, and positive charge[37]. Different hypotheses were generated by varying the default values of the control parameters and altering the feature selection in order to build the best 3D-pharmacophore model. The hypothesis used in this study was obtained by setting three parameters: misses, feature

misses, and complete misses to 3, 2, and 2, respectively. The Principal value of 2 and MaxOmitFeat value of 0 were assigned to the most active compounds (**3** and **17**). This was to allow the two compounds to map all the features in the hypotheses. For the rest of the compounds, the Principal value of 1 and MaxOmitFeat value of 1 were assigned. Details of these control parameters are given in the Catalyst users' manual[37]. The selected features were hydrogen bond donor, positive ionizable, hydrophobic aromatic, ring aromatic, and aliphatic aromatic.

The validity of the model was evaluated based on (i) the presence of important chemical features required for interaction with key amino acid residues (Asp54, Ile14, Leu164) in the active site of the Pf DHFR enzyme and (ii) the best-fit values of the training set molecules and/or matching of their chemical features with those of the model.

Database search
The 3D-pharmacophore model (hypo1) was used as a query to search hits from the Maybridge2004 database (59,652 compounds) and NCI2000 database (238,819 compounds). The "best and flexible database search" option of the Catalyst program[34] was employed to perform the virtual screening. Primary filters such as Lipinski's rule-of-five[38] and restricting the number of rotatable bonds[39,40] to ≤7 were applied to reduce the data set. Further screening of the hits was carried out using two docking algorithms, FlexX[30,31] and Glide[32,33].

Molecular docking studies
The X-ray crystal structure of the quadruple-mutant Pf DHFR enzyme (1J3K:pdb)[11] was obtained from the RCSB Protein Data Bank and was used in order to model the protein structure in this study. It contains **3** bound into the active site in the presence of NADPH. FlexX- and Glide-based docking programs were used to carry out screening of the hits obtained from the database search. FlexX is a fast, flexible docking method that uses an incremental construction algorithm to place ligands into an active site[30,31]. During preparation of the receptor description file (rdf) for FlexX docking, a radius of 6.5Å around the bound ligand was used to define the active site[41]. The formal charges were assigned to the molecules before submission for docking calculation. Default parameters were used, and 30 solutions were generated for each compound. For the purpose of prioritization of hits during selection, G-score[42], PMF-score[43], D-score[44], and ChemScore[45] were also estimated.

Another method used for the molecular docking study was Glide. Glide (Grid-based Ligand Docking with Energetics) uses a series of hierarchical filters to search for possible locations in the active site region of the receptor[32,33]. The properties of a receptor/active site region are represented by a grid that has different sets of fields that provide progressively more accurate scoring of the ligand pose. It uses a Glide score (Gscore) for predicting binding affinity and rank ordering of ligands in the database screen. Another scoring function, which is much better than G-score in selecting the correct pose, and is used

along with Gscore, is Emodel. For details of the methodology, refer to the original literature[32,33]. Glide v4.5 was used for the calculations[46]. During protein preparation for Glide docking in this study, only chain **A** and cofactor NDP610 of Pf DHFR-TS (thymidylate synthase) were retained. All water molecules and the rest of the chains (B, C, and D) were removed from the complex, and the protein was minimized using the "protein preparation wizard" and default parameters. Partial atomic charges were assigned according to the OPLS-AA force field. A radius of 15 Å was selected for the active site cavity during receptor grid generation[46]. Molecules were drawn in SYBYL7.1, and were imported to the Glide window and subsequently converted into Maestro format using the LigPrep module of Glide. The output was set to give 10 docking poses/ligand, whereas default values were used for other parameters. The flexible docking option and XP mode were used in all calculations.

Results and discussion

Pharmacophore generation
The pharmacophore model was built using a set of 24 compounds (with diverse chemical structures) and the Catalyst program. The data set of compounds consisted of derivatives of 2,4-diaminoquinazoline, 2,4-diamino-5,6,7,8-teterahydroquinazoline, 2,4-diamonopteridine, and also **3** (Figure 2). The *in vitro* inhibitory activities of these compounds against the multiple-drug resistant (V1/S strain) Pf DHFR enzyme were carried out using the same method[27]. The compounds with their associated conformational models were submitted to the Catalyst hypothesis generation (HipHop) run.

In order to generate the best model, several HipHop runs were carried out by varying the control parameters (data not shown). Among those several runs, the selected HipHop run was obtained by setting three parameters: misses, feature misses, and complete misses to 3, 2, and 2, respectively. The Principal value of 2 and MaxOmitFeat value of 0 were assigned to the most active compounds (**3** and **17**). For the rest of the compounds, the Principal value of 1 and MaxOmitFeat value of 1 were assigned[37]. The features selected for this run were hydrogen bond donor (D), positive ionizable (PI), hydrophobic aromatic (Z), ring aromatic (R), and hydrophobic aliphatic (H). This HipHop run resulted in 10 hypotheses (hypo1–10). Their ranking scores ranged from 241.11 to 255.74 kcal/mol (Table 1). All of them were four-feature hypotheses. Hypo1–4, 6, 9, and 10 consisted of Z, H, and two Ds whereas hypo5, 7, and 8 each consisted of R, H, and two Ds (Table 1).

As given in Table 1, the first four hypotheses (hypo1 to hypo4) had the same molecular features and close ranking scores. Analysis of best-fit values of the training set compounds against these hypotheses was carried out to choose the best model. The calculated best-fit values indicated hypo1 to be the best model (Table 2). Each of the compounds in the training set showed a good best-fit value with respect to this model. Thus, the model was chosen as a query for the
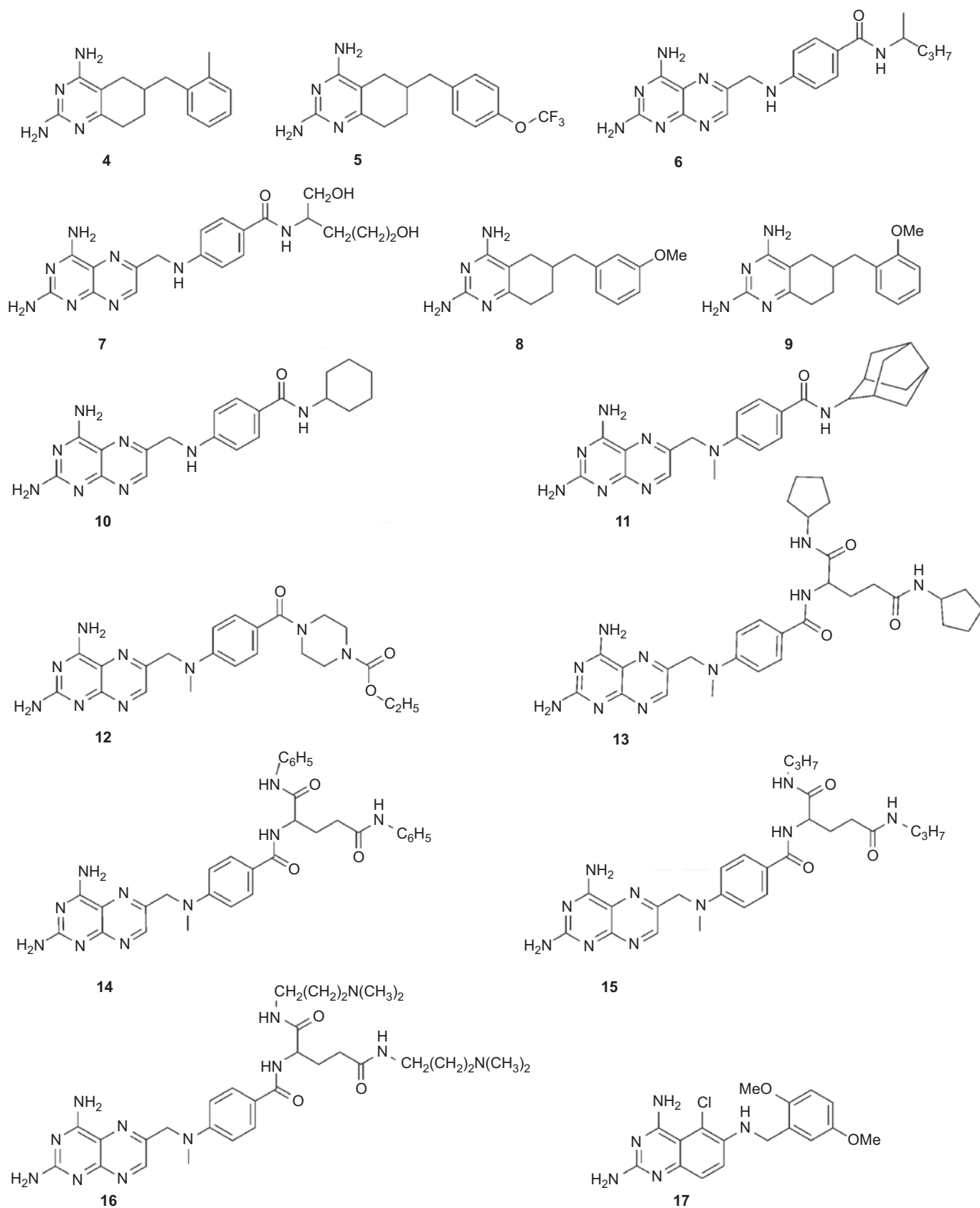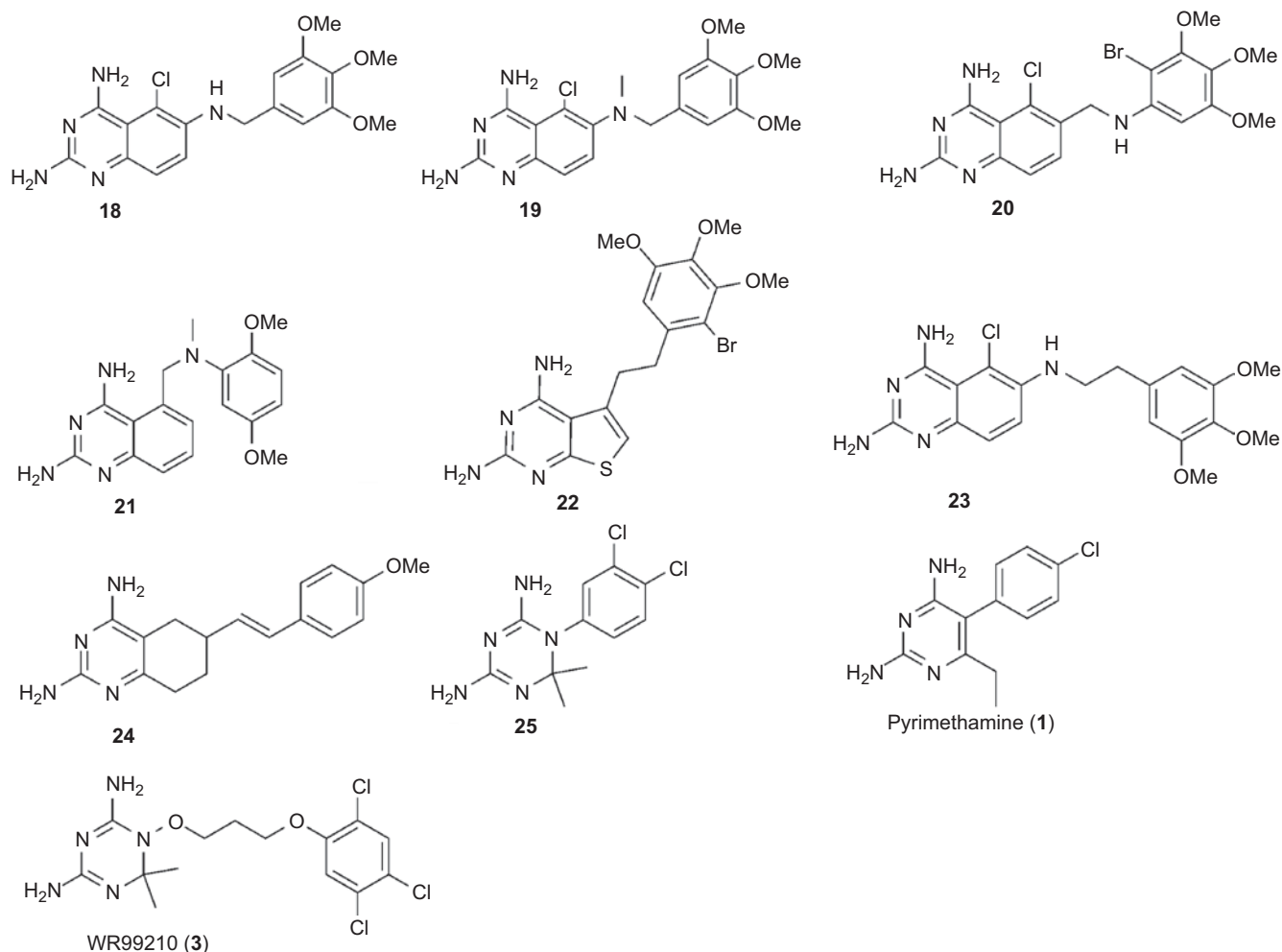
Figure 2. Chemical structures of the training set compounds used for pharmacophore model generation.

Figure 2. Continued.

Table 1. Summary of the hypotheses generated by a Catalyst/HipHop run.

Hypothesis	Molecular features[a]	Ranking score[b]	Direct hit (DH)[c]	Partial hit (PH)[d]
1	ZHDD	255.74	0111111111101111111110111	1000000000010000000001000
2	ZHDD	253.50	0111111111101111111110111	1000000000010000000001000
3	ZHDD	252.88	0111111111101111111110111	1000000000010000000001000
4	ZHDD	252.48	0111111111101111111110111	1000000000010000000001000
5	RHDD	249.25	1111111111111111111110110	0000000000000000000001001
6	ZHDD	249.14	0111111111101111111110111	1000000000010000000001000
7	RHDD	246.00	1111111111111111111110110	0000000000000000000001001
8	RHDD	243.91	1111111111111111111110110	0000000000000000000001001
9	ZHDD	243.10	0111111111101111111110111	1000000000010000000001000
10	ZHDD	241.11	1111111111111111111110110	0000000000000000000001001

[a]Z, hydrophobic aromatic; H, hydrophobic aliphatic; D, hydrogen bond donor; R, ring aromatic.

[b]The higher the ranking score, the less likely it is that the molecules in the set fit the hypothesis by a chance correlation. The best hypothesis shows the highest value.

[c]Direct hit indicates whether (1) or not (0) a molecule in the training set mapped every feature in the hypothesis.

[d]Partial hit indicates whether (1) or not (0) a particular molecule in the training set mapped all but one feature in the hypothesis. Numeration of molecules is from right to left in both DH and PH[34].

virtual screening in order to identify structures that matched its functional and spatial constraints. Figure 3 shows the 3D mappings of **4**, **8**, **9**, **17**, **18**, **19**, and WR99210 (**3**) onto hypo1.

Validation of hypo1 and virtual screening

Before performing the actual virtual screening, the model was validated based on (i) the presence of chemical features required for interactions with key amino

Table 2. pIC_{50}, number of conformers, best-fit values, and ΔE values of the training set compounds.

Compound[a]	pIC_{50}[b]	Conf.[c]	Best-fit value[d]				ΔE (kcal/mol) of conformation[e]			
			Hypo1	Hypo2	Hypo3	Hypo4	Hypo1	Hypo2	Hypo3	Hypo4
4	7.30	32	2.960	1.985	1.440	3.326	17.420	1.397	13.346	10.971
5	6.85	147	3.590	2.897	2.342	2.886	9.260	16.361	3.189	16.631
6	5.77	113	3.332	3.744	2.899	3.107	3.064	7.553	8.836	1.979
7	5.55	206	2.691	2.954	2.589	2.886	16.589	13.317	13.737	4.274
8	7.54	105	3.959	2.968	2.986	3.886	10.116	9.651	8.354	11.252
9	7.44	69	2.951	2.897	1.806	3.759	18.626	19.471	13.049	15.347
10	5.89	82	3.031	3.056	2.609	2.821	5.664	15.421	11.947	0.327
11	5.75	66	3.422	3.483	2.579	2.943	12.770	1.847	1.506	10.735
12	6.00	103	3.039	3.007	2.627	2.918	6.265	15.884	5.685	7.028
13	6.20	240	2.451	3.655	3.152	3.347	11.229	11.059	5.091	5.377
14	5.92	226	2.937	2.960	2.602	2.820	18.223	9.043	14.27	18.223
15	5.85	254	3.516	3.906	3.137	3.385	9.227	18.264	8.995	19.165
16	5.00	249	2.972	2.840	2.893	2.923	2.483	14.678	10.491	13.438
17	8.05	103	4.000	2.955	3.752	4.000	11.430	15.610	9.664	5.496
18	7.49	117	3.988	2.919	3.469	3.970	10.965	6.858	13.295	7.682
19	7.51	127	3.992	2.838	3.443	3.949	8.337	16.648	10.719	3.698
20	6.65	87	3.979	2.825	3.468	3.961	12.051	11.733	11.421	5.835
21	5.77	116	3.907	2.550	3.262	3.653	5.702	7.052	10.486	19.096
22	5.00	86	3.938	2.659	3.065	3.505	19.126	11.977	5.028	15.964
23	5.85	175	3.984	2.931	3.827	3.976	18.843	19.645	12.593	5.389
24	6.77	25	2.129	3.845	2.789	2.213	0.978	19.790	15.113	0.303
25	7.15	3	2.646	0.286	2.542	1.997	0.003	0.003	0.000	0.000
Pyrimethamine (**1**)	5.38	6	1.988	0.565	0.798	1.438	0.086	0.059	0.000	0.059
WR99210 (**3**)	8.57	158	3.868	3.494	3.899	3.898	9.027	6.845	4.231	9.449

[a]IUPAC names of the compounds are given in ref. 27.

[b]IC_{50} (in nM ranges) values are given in ref. 27.

[c]Number of conformational models of training set compounds within the range of 20 kcal/mol from the global minimum.

[d]Best-fit value of the conformer used for mapping.

[e]Energy difference between the conformer used for mapping and the global minimum calculated by Catalyst software[34].
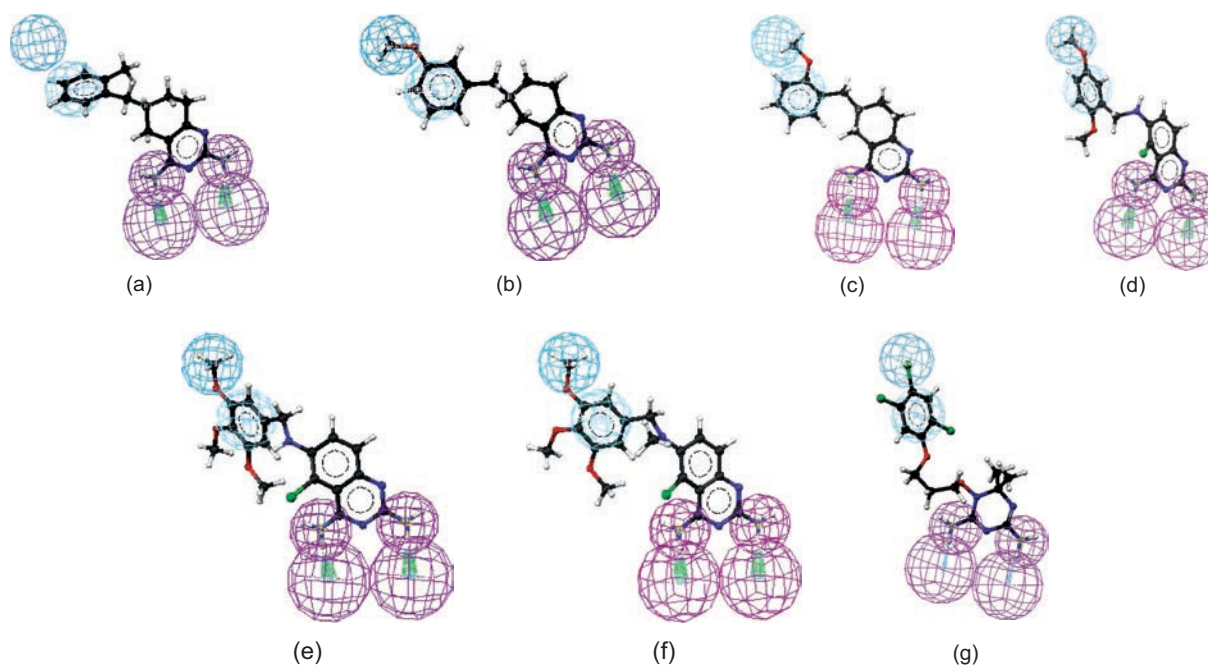
Figure 3. Mapping of **4** (a), **8** (b), **9** (c), **17** (d), **18** (e), **19** (f), and WR99210 (**3**) (g) onto hypo1. Pharmacophoric features are color-coded (violet, hydrogen bond donor; blue, hydrophobic aliphatic; light blue, aromatic hydrophobic).

acid residues (Ile14, Asp54, Leu164) in the active site of Pf DHFR. The model had four features (one H, one Z, and two Ds). This observation is consistent with that of known antifolate Pf DHFR inhibitors (e.g. **3**)[6,7,11,25]. In this inhibitor, it has been observed that two hydrogen bond donors interact with Asp54 and backbone amino acids (Ile14 or/and Leu164). On the other hand, hydrophobic features interact with amino acid residues such as Met55, Ser111, and Pro113 near the opening of the active site via hydrophobic interactions. For instance, as shown in Figure 2, hydrogen bond donor features of hypo1 mapped to 2-amino and 4-amino groups. The hydrophobic aromatic feature mapped on the phenyl side chain, whereas the hydrophobic aliphatic feature of the model mapped on the Cl group. Moreover, the conformation of **3**, which was extracted from the X-ray crystallographic structure of the quadruple-mutant Pf DHFR enzyme, was also found to perfectly overlap onto hypo1 (data not given). Similarly, the features of hypo1 mapped onto the appropriate chemical features of other active compounds (**4**, **8**, **9**, **17**, and **18**) as illustrated in Figure 3, and (ii) the best-fit values of the training set molecules obtained by mapping onto hypo1 were generally higher than the corresponding values for other hypotheses (Table 2). Thus, hypo1 was selected and subsequently used as a query for database searching. The Maybridge2004 and NCI2000 databases were searched using the "best and flexible database search" option of the Catalyst program. The virtual screening resulted in 267 and 255 hits from Maybridge2004 and NCI2000 databases, respectively. To reduce the number of molecules for further analysis, several filtering criteria were used. First, Lipinski's rule-of-five[31] was used. According to this rule, any drug-like molecule should be characterized by molecular weight <500, calculated octanol–water partition coefficient (CLogP) <5, number of hydrogen bond donors (HBDs) (OH + NH groups) <5, and number of H-bond acceptors (HBAs) (O + N atoms) <10. Thus, hits with molecular weight >500, number of HBDs >5, and number of HBAs >10 were eliminated from further analysis. Second,

to avoid too flexible compounds, those hits having more than seven rotatable bonds[39,40] were also eliminated. The remaining 191 hits were subjected to docking into the active site of quadruple-mutant Pf DHFR using FlexX and Glide docking programs in order to investigate their binding modes and interactions. To verify the reproducibility of the docking calculations, the bound ligand (**3**) was extracted from the complex and submitted for one-ligand run calculation. This reproduced top scoring conformations of **3** falling within RMS (root mean square) values of 0.96 Å and 0.489 Å from the bound X-ray conformation for FlexX and Glide docking, respectively. The binding modes of the hits were ranked according to FlexX score and Glide score values. In the case of FlexX docking, other scoring functions such as G-score, PMF-score, D-score, and C-score were also used for further refining of the hits. The hits with better or comparative scores were retained for further analysis. The Catalyst "best-fit" values were calculated for each retained hit using hypo1.

Those hits with "best-fit" value <1.5 were rejected. Based on visual inspection of the binding modes of the hits from both FlexX and Glide docking calculations, together with binding scores and "best-fit" values, a set of nine hits were identified, which belonged to different classes (Table 3). The chemical structures of the identified hits are given in Figure 4, whereas the mappings of representative hits to hypo1 are given in Figure 5.

Binding modes (interactions) of the identified hits

As described in previous sections, the identified hits show all the necessary interactions in the active site which are required by antifolate Pf DHFR inhibitors for effective inhibition of this enzyme[6,7,11,18,25]. Their hydrogen bond donor features interact with Asp54 and Ile14 or/and Leu164, whereas the hydrophobic portions interact with amino acid residues Met55, Ser111, Pro113, and other amino acids via hydrophobic interactions. Some of the hits also show H-bond interaction with Asn108. Their binding modes also indicated that there is no unfavorable steric clash with Asn108 (mutated from Ser108), which is

Table 3. Docking scores of hits identified from the virtual screening study.

S. no.	Hit	Glide Gscore (Emodel)	FlexX FlexX score	G-score	PMF-score	D-score	ChemScore	C-score
1	**NCI00043568**	−11.34 (−64.5)[b]	−24.78[a]	−211.26	−73.26	−109.68	−40.28	4
2	**NCI0029604**	−10.21 (−54.6)[a]	−23.82[a]	−134.47	−72.56	−103.24	−26.30	5
3	**GK03630**	−10.05 (−65.3)[a]	−19.43[a]	−192.00	−55.54	−97.01	−26.11	5
4	**GK03628**	−9.96 (−59.4)[b]	−17.14[a]	−187.77	−15.10	−102.22	−24.27	3
5	**NCI00129588**	−9.92 (−63.7)[a]	−19.96[a]	−184.61	−68.31	−100.70	−24.03	5
6	**CD00706**	−9.53 (−62.5)[c]	−24.64[a]	−31.02	−11.97	−114.57	−33.07	4
7	**NCI0027612**	−9.26 (−50.2)[a]	−23.11[a]	−123.23	−72.94	−95.66	−24.72	5
8	**NCI0037722**	−8.69 (−61.6)[a]	−23.35[a]	−171.53	−47.26	−102.20	−29.30	5
9	**NCI0014710**	−6.33 (−51.6)[d]	−22.40[d]	−38.00	−64.64	−97.19	−26.18	2
	WR99210 (ref.)	−8.44 (−80.20)[a]	−17.86[a]	−209.23	−67.80	−134.79	−28.91	5

[a]Hydrogen bond interactions of hits with Ile14, Leu164, and Asp54.
[b]Hydrogen bond interactions of hits with Asp54, Asn108, and Ile14 or Leu164.
[c]Hydrogen bond interactions of hits with Asp54 and Ile14 or Leu164.
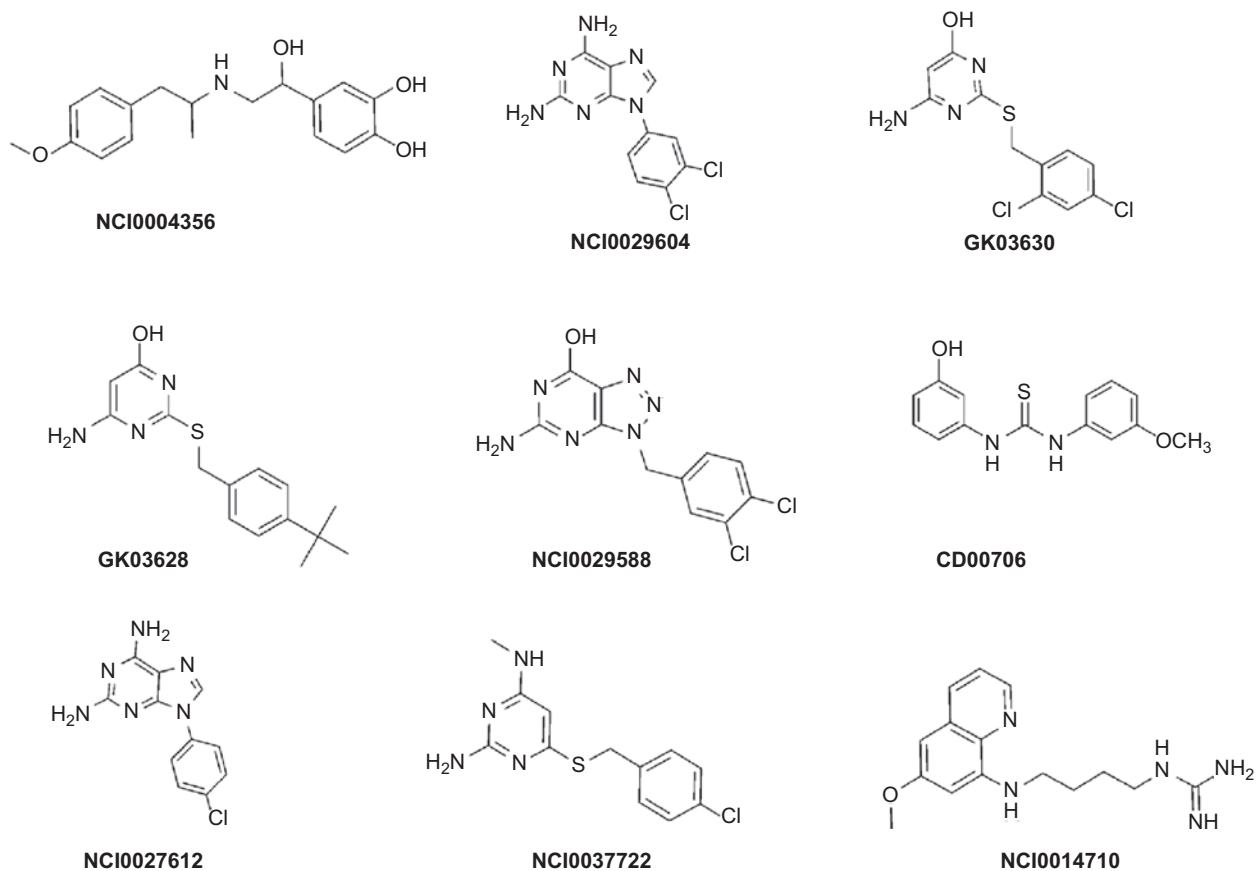[d]Hydrogen bond interactions of hits with Asp54 and Asn108.

Figure 4. Chemical structures of the identified hits.

known to be responsible for resistance. The detailed binding interactions of some of the hits (**NCI0004356, GK-03628**, **NCI0029588, CD00706, NCI0037722**, and **NCI0014710**) are given in the following sections. Unless and otherwise mentioned specifically, discussion is based on the Glide docking result.

Binding mode of NCI0004356
Figure 6a shows the binding mode of **NCI0004356** within the active site of the quadruple-mutant *Pf*DHFR enzyme. As illustrated in this binding model, one of the OH substituents of the aromatic ring forms an H-bond with Asp54 with a bond length value of 2.092 Å. The OH group of the aliphatic side chain interacts with Asn108 and Leu164. This OH group donates an H-bond to the carbonyl group of Leu164, and it accepts an H-bond from the amino group of Asn108. The corresponding H-bond lengths are 2.896 and 1.876 Å, respectively. The long side chain of this hit interacts hydrophobically with Met55, Ser111, and Pro113.

Binding mode of GK-03628
The binding mode of **GK-03628** within the active site of the quadruple-mutant *Pf*DHFR enzyme is given in Figure 6b. It interacts with Asp54, Ile14, and Asn108. The OH group interacts with Asp54 (1.949 Å), whereas its amino group interacts with Ile14 (2.897 Å). The S atom of

the linker unit accepts an H-bond from the amino group of Asn108 (2.313 Å) (Figure 6b). Visual inspection also showed that its side chain interacts with amino acids (Met55, Ser111, and Pro113) in the hydrophobic region of the active site.

Binding mode of NCI0029588
The binding mode of **NCI0029588** within the active site of the quadruple-mutant *Pf*DHFR enzyme is given in Figure 6c. Similar to that of the bound ligand (**3**), **NCI0029588** interacts with Asp54 as well as backbone amino acid residues (Ile14 and Leu164). The H-bond length between the OH group of **NCI0029588** and oxygen atom of the carboxylic group of Asp54 is 2.128 Å. The amino group forms H-bonds with Ile14 and Leu164. The corresponding H-bond length values are 1.709 and 1.929 Å, respectively (Figure 6c).

Binding mode of CD00706
The binding mode of **CD00706** within the active site of the quadruple-mutant *Pf*DHFR is given in Figure 6d. The OH group interacts with Asp54 via H-bonding (1.714 Å) and one of the NH groups of the thiourea linker unit forms an H-bond with Leu164 (2.260 Å). The second aromatic ring and its OMe substituent interact hydrophobically with Met55, Ser111, and Pro113.
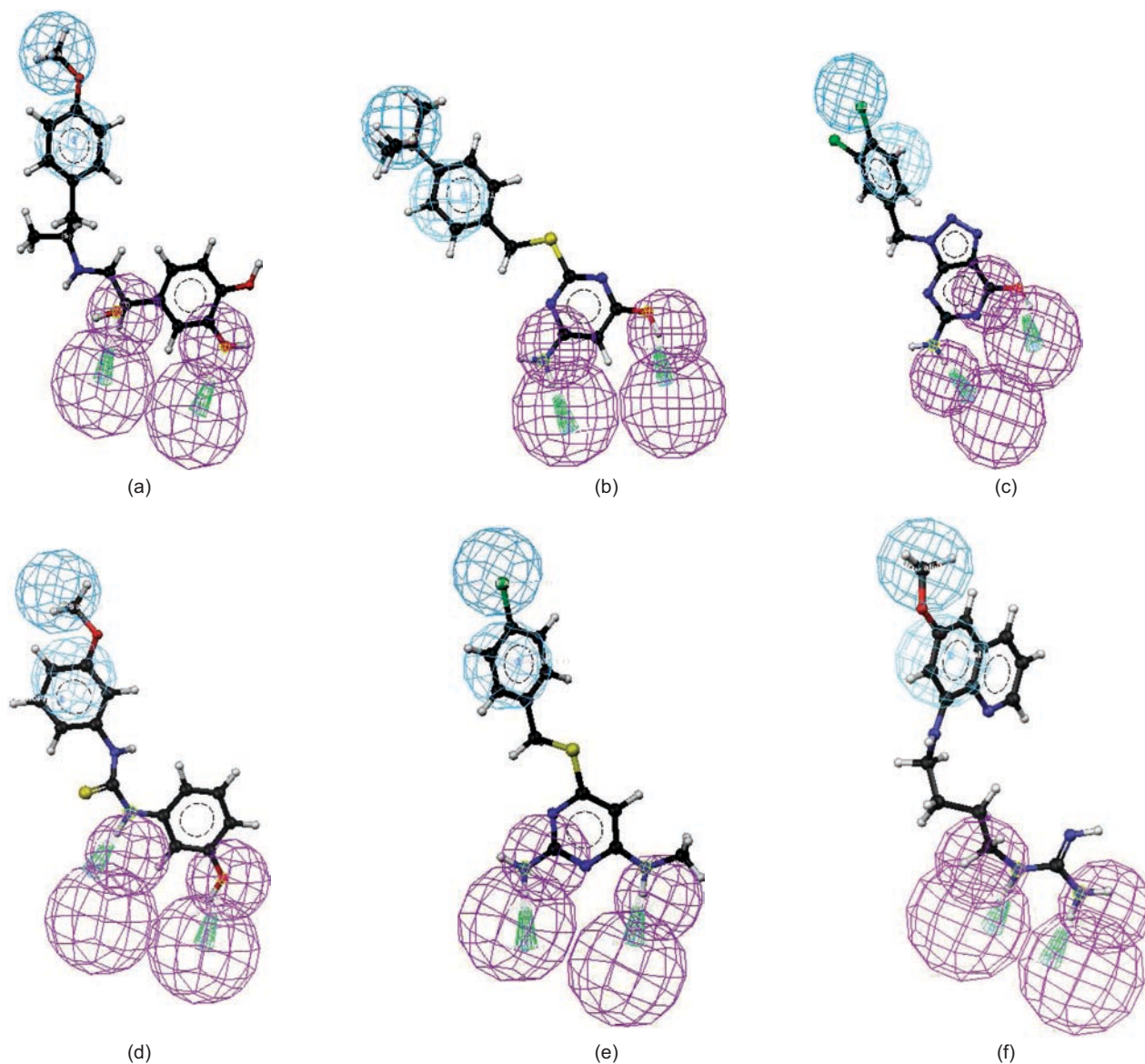
Figure 5. Mapping of **NCI00043568** (a), **GK-03628** (b), **NCI0029588** (c), **CD00706** (d), **NCI0037722** (e), and **NCI0014710** (f) onto hypo1. Pharmacophoric features are color-coded (violet, hydrogen bond donor; blue, hydrophobic aliphatic; light blue, aromatic hydrophobic).

Binding mode of NCI0037722

Figure 6e shows the binding mode of **NCI0037722** within the active site of the quadruple-mutant *Pf*DHFR. As illustrated in the model, it forms all the key H-bond interactions in the active site. The Me substituent bearing amino group interacts with Asp54 via H-bonding (1.922 Å), whereas the 4-amino group forms an H-bond with the backbone amino acids Ile14 (2.229 Å) and Leu164 (2.225 Å). As in the case of the hits discussed above, the side chain of **NCI0037722** is oriented toward the hydrophobic region of the active site. Thus, it is expected to interact with amino acid residues in this region via hydrophobic interaction (Figure 6e).

Binding mode of NCI0014710

Figure 6f shows the binding mode of **NCI0014710** within the active site of the quadruple-mutant *Pf*DHFR. Its two amino

groups form a bidentate H-bond interaction with Asp54 (1.616 and 1.867 Å). The N atom of the ring system accepts an H-bond from the NH_2 group of Asn108 (2.291 Å). The ring system and its OMe substituents are also expected to interact hydrophobically with amino acid residues such as Ser111, Pro113, and Met55 (Figure 6f).

Conclusions

Pharmacophore modeling is a powerful method in drug design and discovery since it enables medicinal chemists to rapidly identify new potential drugs. In this study we employed Catalyst/HipHop software to generate a hypothesis having four features: two hydrogen bond donors, one aromatic hydrophobic, and one hydrophobic aliphatic. Use of this hypothesis as a query for searching Maybridge2004
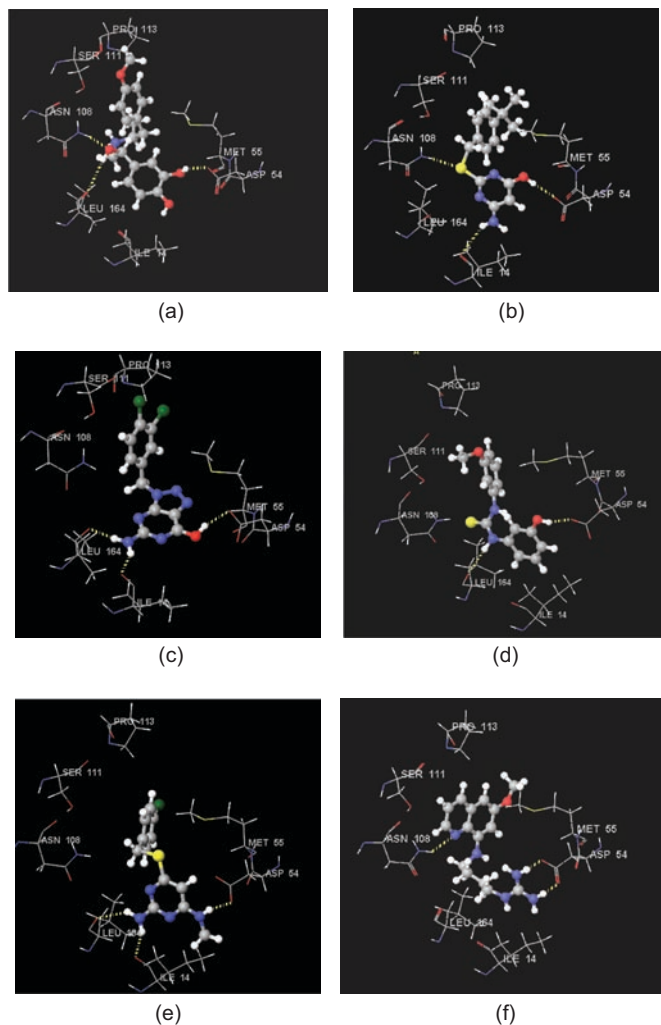
Figure 6. Stereoview of XPGlide predicted binding poses of (a) **NCI004356**, (b) **GK03628**, (c) **NCI0029588**, (d) **CD00706**, (e) **NCI0037722**, and (f) **NCI0014710** in the active site of the quadruple-mutant *Pf*DHFR enzyme. For the sake of clarity, only important amino acid residues are given.

and NCI2000 databases helped us to identify a set of nine potential inhibitors of the quadruple-mutant *Pf*DHFR enzyme. The hits were selected based on their docking scores, binding orientations, and interactions with key amino acid residues in the active site (Asp54, Ile14, and Leu164) as well as their interactions in the hydrophobic region of the active site. Because of the flexible nature of their side chains, there are no potential steric clashes with Asn108, which is known to be responsible for resistance of the mutant parasite to antimalarial drugs **1** and **2**[6,7,10–13,18,25]. The hits also fulfill all the properties required by drug-like molecules[47,48]. To the best of our knowledge, there are no reports on antimalarial activities of these hits. Thus, we encourage those research teams working in the area of antimalarial drug discovery to carry out activity tests of these hits so that compounds with high activity may be identified to overcome resistance.

Declaration of interest

References

1. Kompis IM, Islam K, Then RL. DNA and RNA synthesis: antifolates. Chem Rev 2005;105:593–620.
2. Da Cuhan EF, Ramlho TC, Maia ER, de Alenncastro RB. The search for new DHFR inhibitors: a review of patents, January 2001–February 2005. Expert Opin Ther Patents 2005;15:1–20.
3. Nzila A. Inhibitors of de novo folate enzymes in P. *falciparum*. Drug Discov Today 2006;11:939–44.
4. Anderson AC, Wright DL. Targeting DHFR in parasitic protozoa. Drug Discov Today 2005;10:121–8.
5. Schenell JR, Dyson HJ, Wright PE. Structure, dynamics, and catalytic function of dihydrofolate reductase. Annu Rev Biophys Biomol Struct 2004;33:119–40.
6. Lemeke T, Christensen IT, Jorgensen FS. Towards understanding of drug resistance in malaria: three-dimensional structure of P. *falciparum* DHFR by homology building. Bioorg Med Chem 1999;7:1003–11.
7. Rastelli G, Sirawaraporn W, Sompornpisut P, Vilaivan T, Kamchonwongpaisan S, Quarrell R, et al. Interaction of pyrimethamine, cycloguanil, WR99210 and their analogues with *P. falciparum* dihydrofolate reductase: structural basis of antifolate. Bioorg Med Chem 2000;8:1117–28.
8. Delfino RT, Santos-Filho OA, Figueroa-Villar JD. Molecular modeling of wild-type and antifolate resistant mutant P. *falciparum* DHFR. Biophys Chem 2002;98:287–300.
9. Santos-Filho OA, de Alencastrob RD, Figueroa-Villar JD. Homology modeling of wild-type and pyrimethamine/cycloguanil-cross-resistant mutant type P. *falciparum* DHFR: a model for antimalarial chemotherapy resistance. Biophys Chem 2001;91:305–17.
10. Sirawaraporn W, Sathitkul T, Sirawaraporn R, Yuthavong Y, Santi DV. Antifolate-resistant mutants of P. *falciparum* dihydrofolate reductase. Proc Natl Acad Sci USA 1997;94:1124–9.
11. Yuvaniyama J, Chitnumsub P, Kamchonwongpaisan S, Vanichatanankul J, Sirawaraporn W, Taylor P, et al. Insights into antifolate resistance from malarial DHFR-TS structures. Nat Struct Biol 2003;10:357–65.
12. Warhurst DV. Antimalarial drug discovery: development of inhibitors of dihydrofolate reductase active in drug resistance. Drug Discov Today 1998;3:538–46.
13. Warhurst DV. Resistance to antifolate in P. *falciparum* DHFR, the causative agent of tropical malaria. Sci Prog 2002;85:89–111.
14. McKie JH, Douglas KT, Chan C, Roser SA, Yates R, Read M, et al. Rational drug design approach for overcoming drug resistance: application to pyrimethamine resistance in malaria. J Med Chem 1998;41:1367–70.
15. Sardarian A, Douglas KT, Read M, Sims PFG., Hyde JE, Chitnumsub P, et al. Pyrimethamine analogs as strong inhibitors of double and quadruple mutants of dihydrofolate reductase in human malaria. J Biol Chem 2003;1:960–4.
16. Kamchonwongpaisan S, Quarrell R, Charoensetakul N, Ponsinet R, Vilaivan T, Vanichtanankul J, et al. Inhibitors of multiple mutants of P. *falciparum* dihydrofolate reductase and their antimalarial activities. J Med Chem 2004;47:673–80.
17. Yuthavong Y, Vilainvan T, Chareonsethakul N, Kamchonwongpaisan S, Sirawaraporn W, Quarrell R, et al. Development of a lead inhibitor for the A16V+S108T mutant of dihydrofolate reductase from the cycloguanil-resistant strain (T9/94) of P. *falciparum*. J Med Chem 2000;43:2738–44.
18. Sichaiwat C, Intaraudom C, Kamchonwongpaisan S, Vanichtanankul J, Thebtaranonth Y, Yuthavong Y. Target guided synthesis of 5-benzyl-2,4-diamonopyrimidines: their antimalarial activities and binding affinities to wild type and mutant dihydrofolate reductases from *P. falciparum*. J Med Chem 2004;47:345–54.
19. Hunt SY, Detering C, Varani G, Jacobus DP, Schiehser GA, Shieh HM, et al. Identification of the optimal third generation antifolate against P. *falciparum* and P. *vivax*. Mol Biochem Parasitol 2005;144:198–205.
20. Japrung D, Leartsakulpanich U, Chusacultanachai S, Yuthavong Y. Conflicting requirements of P. *falciparum* dihydrofolate reductase mutations conferring resistance to pyrimethamine-WR99210 combination. Antimicrob Agents Chemother 2007;51:4356–60.

21. Klebe G. Foundation review: virtual ligand screening: strategies, perspectives and limitations. Drug Discov Today 2006;11:580–94.

22. Muegge I, Oloff S. Advances in virtual screening. Drug Discov Today Tech 2006;3:405–11.

23. Alvarez J, Shoichet B. Virtual Screening in Drug Design. Boca Raton, FL: Taylor & Francis, 2005.

24. Toyoda T, Brobey RK, Sano G, Horii T, Tomioka N, Itai A. Lead discovery of inhibitors of the dihydrofolate reductase domain of *P. falciparum* dihydrofolate reductase-thymidylate synthase. Biochem Biophys Res Commun 1997;235:515–19.

25. Rastelli G, Pacchioni S, Sirawaraporn W, Sirawaraporn R, Parenti MD, Ferrari AM. Docking and database screening reveal new classes of *P. falciparum* dihydrofolate reductase inhibitors. J Med Chem 2003;46:2834–45.

26. Dasgupta T, Chitnumsub P, Kamchonwongpaisan S, Maneeruttanarungroj C, Nichols SE, Lyons TM, et al. Exploiting structural analysis, *in silico* screening, and serendipity to identify novel inhibitors of drug-resistant *Falciparum malaria*. ACS Chem Biol 2009;4:29–40.

27. Ommeh S, Nduati E, Mberu E, Kokwaro G, Marsh K, Rosowsky A, et al. *In vitro* activities of 2,4-diaminoquinazoline and 2,4-diaminopteridine derivatives against *P. falciparum*. Antimicrob Agents Chemother 2004;48:3711–14.

28. Guner OF. Pharmacophore: Perception, Development and Use in Drug Design. La Jolla, CA: International University Line, 2000.

29. Kurogi Y, Guner OF. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. Curr Med Chem 2001; 8:1035–55.

30. Rarey M, Kramer B, Lengauer T. Multiple automatic base selection: protein-ligand docking based on incremental construction without manual intervention. J Comput Aided Mol Des 1997;11:369–84.

31. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–89.

32. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004;47:1739–49.

33. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 2004;47:1750–9.

34. Catalyst, Version 4.10. San Diego, CA: Accelrys Inc., 2005.

35. Smellie A, Teig SL, Towbin P. Polling: promoting conformational variation. J Comput Chem 1995;16:171–87.

36. Smellie A, Kahn SD, Teig SL. An analysis of conformational coverage 2. Applications of conformational models. J Chem Inf Comput Sci 1995;35:295–304.

37. Catalyst manual, available at http://www.accelrys.com.

38. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–25.

39. Muegge I. Selection criteria for drug-like compounds. Med Res Rev 2003;23:302–21.

40. Oprea TI. Property distribution of drug-related chemical databases. J Comput Aided Mol Des 2000;14:251–64.

41. SYBYL 7.1. St. Louis, MO: Tripose Inc., 2005.

42. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267:727–48.

43. Muegge I., Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. J Med Chem 1999;42:791–804.

44. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. J Mol Biol 1982;161:269–88.

45. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput Aided Mol Des 1997;11:425–45.

46. Schrödinger Suite 2007. New York: Schrödinger, LLC, 2007.

47. Proudfoot JR. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. Bioorg Med Chem Lett 2002;12:1647–50.

48. Oprea TI, Allu TK, Fara DC, Rad RF, Ostopovici L, Bologa CG. Lead-like, drug-like or "pub-like": how different are they? J Comput Aided Mol Des 2007;21:113–19.