Three-Dimensional Quantitative Structure−Activity Relationship Analysis of a Set of *Plasmodium falciparum* Dihydrofolate Reductase Inhibitors Using a Pharmacophore Generation Approach

Marco Daniele Parenti, Sara Pacchioni, Anna Maria Ferrari, and Giulio Rastelli*

Dipartimento di Scienze Farmaceutiche, Università di Modena e Reggio Emilia, Via Campi 183, 41100 Modena, Italy

Received January 12, 2004

A 3D pharmacophore model able to quantitatively predict inhibition constants was derived for a series of inhibitors of *Plasmodium falciparum* dihydrofolate reductase (PfDHFR), a validated target for antimalarial therapy. The data set included 52 inhibitors, with 23 of these comprising the training set and 29 an external test set. The activity range, expressed as K_i, of the training set molecules was from 0.3 to 11 300 nM. The 3D pharmacophore, generated with the HypoGen module of Catalyst 4.7, consisted of two hydrogen bond donors, one positive ionizable feature, one hydrophobic aliphatic feature, and one hydrophobic aromatic feature and provided a 3D-QSAR model with a correlation coefficient of 0.954. Importantly, the type and spatial location of the chemical features encoded in the pharmacophore were in full agreement with the key binding interactions of PfDHFR inhibitors as previously established by molecular modeling and crystallography of enzyme−inhibitor complexes. The model was validated using several techniques, namely, Fisher's randomization test using CatScramble, leave-one-out test to ensure that the QSAR model is not strictly dependent on one particular compound of the training set, and activity prediction in an external test set of compounds. In addition, the pharmacophore was able to correctly classify as active and inactive the dihydrofolate reductase and aldose reductase inhibitors extracted from the MDDR database, respectively. This test was performed in order to challenge the predictive ability of the pharmacophore with two classes of inhibitors that target very different binding sites. Molecular diversity of the data sets was finally estimated by means of the Tanimoto approach. The results obtained provide confidence for the utility of the pharmacophore in the virtual screening of libraries and databases of compounds to discover novel PfDHFR inhibitors.

Introduction

Dihydrofolate reductase (DHFR) catalyzes a sequential reaction in the biosynthesis of thymidylate, which is known to be essential for DNA synthesis and for maintaining intracellular pools of tetrahydrofolate cofactors. Inhibition of DHFR blocks the reduction of dihydrofolate to tetrahydrofolate and prevents DNA synthesis, resulting in cell death. The dihydrofolate reductase domain of the bifunctional enzyme dihydrofolate reductase−thymidylate synthase of *Plasmodium falciparum* (PfDHFR-TS) is a validated target of antimalarial antifolates such as pyrimethamine (Pyr), cycloguanil (Cyc), and other antifolates used for prophylaxis and treatment of *Plasmodium falciparum* infections.[1−4] Unfortunately, Pyr and Cyc are facing the emergence of resistant *Plasmodium falciparum* strains, limiting their utility in the treatment of malaria. The compromised clinical efficacy of these drugs and the lack of a suitable vaccine have an enormous economic and social impact, particularly in Africa where malaria causes about 2 million deaths per year and a very high level of morbidity. Therefore, there is urgent need to search for new targets and/or new effective antifolate antimalarials to combat the resistant parasites.[5]

During the past years, the structural basis of antifolate resistance has been investigated; homology models of PfDHFR[6−8] have been reported and have been exploited to investigate the molecular interactions between the active site residues of the enzyme and antifolates such as Pyr, Cyc, and WR99210 (6,6-dimethyl-1-[3-(2,4,5-trichlorophenoxy)propoxy]-1,6-dihydro[1,3,5]-triazine-2,4-diamine).[1,6−8] We previously proposed a "steric constraint hypothesis" to rationalize antifolate resistance in PfDHFR. In agreement with the hypothesis, structure-based design of inhibitors devoid of substituents at a position where steric clash can be formed resulted in a significant reduction of resistance.[7,9] More recently, the crystal structures of PfDHFR-TS from the wild type and a quadruple drug-resistant mutant strain N51I−C59R−S108N−I164L in complex with WR99210, and the resistant double mutant C59R−S108N in complex with Pyr, have been reported.[10] The analysis of these crystal structures confirmed and strengthened the previously reported models of interaction of antifolates with PfDHFR. It is now established that resistant mutant PfDHFRs could be efficiently inhibited by compounds that are much more flexible than Pyr and Cyc (WR99210 is a good example of this successful approach) and that are devoid of bulky groups in proximity of the S108 residue.[7,10]

Very recently, we have developed and applied a combined pharmacophore screening−molecular docking strategy to discover new inhibitors of PfDHFR.[11] The strategy involved prefiltering the Available Chemicals

* To whom correspondence should be addressed. Phone: 0039-059-2055145. Fax: 0039-059-2055131. E-mail: rastelli.giulio@unimo.it.

10.1021/jm040769c CCC: $27.50
Published on Web 07/16/2004

Directory (ACD) database to search for molecules that match specific 3D-pharmacophore requirements derived by the analysis of the complexes of PfDHFR with the three classical antifolate inhibitors Pyr, Cyc, and WR99210. This process served to exclude the molecules that lack the chemical features crucial for interacting with key active site residues, mainly D54, and to reduce the number of structures to be screened in the docking process. The study led to the identification of 12 new compounds whose chemical structures are completely unrelated to classical antifolates but that inhibit the wild type and resistant PfDHFRs harboring A16V, S108T, A16V + S108T, C59R + S108N + I164L, and N51I + C59R + S108N + I164L mutations.[11]

An interesting application of pharmacophore-based approaches is that the experimentally determined activity of a set of known inhibitors can be used to drive the generation of pharmacophores. In this context, pharmacophores can be built under the condition that they explain the observed trend of activity within a training set of compounds. The advantage of this approach is that the pharmacophores, once validated, can be used to *quantitatively* predict the activity of new compounds. Therefore, they constitute a powerful and fast tool to estimate the biological activity of new potential ligands in 3D databases of compounds. This methodology has proven to be successful in a number of virtual screenings of databases and libraries.[12−18] Other approaches were successfully applied to screen a large number of compounds for DHFR activity. Wyss et al.[19] explored a large set of substituted 2,4-diaminopyrimidine derivatives using structure-based (docking) and diversity-based criteria and concluded that structure-based screening provided a greater number of hits and significantly more active inhibitors of *S. aureus* and *S. pneumonia* DHFRs compared to the diversity-based screening. Jurs et al.[20] used computational neural networks to generate quantitative structure−activity (QSAR) models for inhibitors of *rat liver*, *Pneumocystis carinii*, and *Toxoplasma gondii* DHFRs, employing a large number of topological, geometric, and electronic descriptors.

The aim of the present work is to generate a predictive pharmacophore model for PfDHFR inhibitors. To develop the model, we have used the program Catalyst 4.7[21] and the Hypogen algorithm for automated pharmacophores generation. Starting from a collection of conformational models of PfDHFR inhibitors spanning 4−5 orders of magnitude in inhibition constants (training set), a pharmacophore model (hypothesis) able to quantitatively correlate the estimated activities with the measured activities was generated. The model was then validated for its predictive ability of the activities of a set of different PfDHFR inhibitors (test set) and other inhibitors taken from the MDL Drug Data Report (MDDR)[22] database. We found that the pharmacophore was able to correctly predict as active the 174 DHFR inhibitors present in the MDDR database and to coherently predict as inactive or moderately active the 882 inhibitors of aldose reductase, an enzyme in which the structural requirements for inhibition clearly differ from those of DHFR. Taken together, these results suggest that the present pharmacophore can be useful to screen databases and libraries of compounds to retrieve structures that can be used as new potentially active candidates.

Results and Discussion

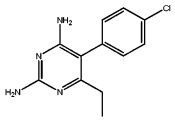
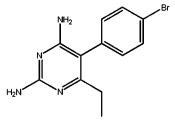
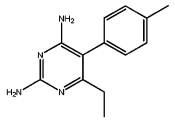
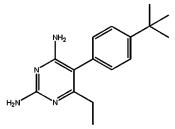
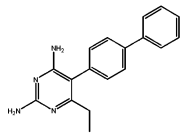
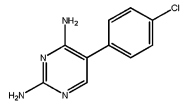
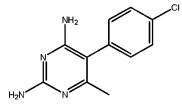
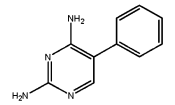
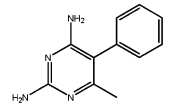
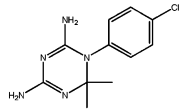
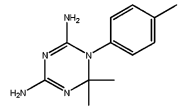
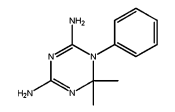
In the present work, a 3D-pharmacophore model with quantitative predictive ability in terms of inhibition of PfDHFR was derived.

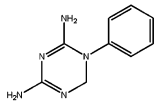
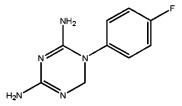
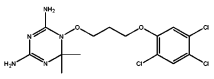
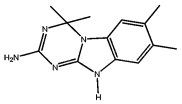
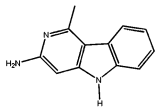
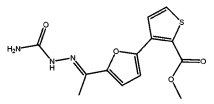
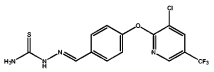
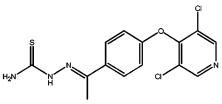
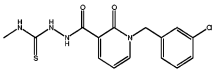
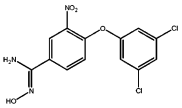
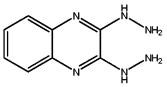
A 3D pharmacophore is a collection of chemical features in space that are required for a desired biological activity. These include hydrophobic groups, charged/ionizable groups, hydrogen bond donors/acceptors, and others, properly assembled in 3D to reflect structural requirements for interaction with the target. Even when a protein structure-based approach is made possible by knowledge of the structure of the target from crystallography or modeling studies, a ligand-based approach like that for 3D pharmacophores may provide an alternative and complementary tool for drug design. We recently proposed a combined structure-based and ligand-based approach in which 3D pharmacophores dictated by the structure of the target were used to filter the ACD database for molecules possessing specific requirements, after which docking of the resulting focused database into the active site of PfDHFR was performed.[11] In this work, we extend our investigation to the design of PfDHFR inhibitors by developing a more sophisticated pharmacophore model with the additional ability to quantitatively predict the inhibition constants. One interesting application of this approach is that 3D pharmacophores, once validated, can be used as 3D queries in searching 3D databases to retrieve and rank new potential ligands.

Pharmacophore Generation. Pharmacophores have been generated with the Hypogen module of Catalyst 4.7 using a training set of 23 PfDHFR inhibitors, with inhibition constants (K_i) ranging from 0.3 to 11 300 nM. The training set included 15 derivatives of Pyr, Cyc, and WR99210 found in the literature[23,24] (compounds **1**−**1h**, **2**−**2d**, **3** in Table 1), 2 inhibitors retrieved by docking studies on a PfDHFR model reported in the literature[25] (compounds **4a**,**b** in Table 1), and 6 new inhibitors not related to classical antifolates disclosed in our previous molecular docking study[11] (compounds **5a**−**f** in Table 1).

Hypogen allows a maximum of five features in pharmacophore generation. Therefore, all the features included in the Catalyst's features dictionary were first considered and subsequently reduced to a maximum of 5. First, we excluded all the chemical features that do not map the molecules of the training set, like "positive charge" (all inhibitors were considered in their neutral form), "negative charge", and "negative ionizable". Also, preliminary hypothesis generation including the "hydrogen bond acceptor" and "hydrogen bond acceptor lipid" features revealed that these features are never used in the pharmacophore models, though present and properly mapped on many molecules of the training set; therefore, the "hydrogen bond acceptor" and "hydrogen bond acceptor lipid" features were removed from the list. Moreover, considering that most of the molecules in the training set possess both hydrophobic aromatic and hydrophobic aliphatic groups, the specific "hydrophobic aliphatic" and "hydrophobic aromatic" features were

Table 1. Chemical Structures, Experimental and Estimated Activities (K_i, nM), and Relative Errors of the PfDHFR Inhibitors in the Training Set

Cpd.	Structure	Exp. Act. (Ki,nM)	Est. Act. (Ki,nM)	Error (Est/Exp)
1 (Pyr)		0.6	0.87	+1.4
1a		0.3	0.84	+2.8
1b		0.4	1.1	+2.7
1c		0.6	0.77	+1.3
1d		188	90	-2.1
1e		2.4	0.74	-3.2
1f		2	0.84	-2.4
1g		41.9	100	+2.4
1h		10.9	110	+9.6
2 (Cyc)		1.5	1.1	-1.4
2a		1.8	1.6	-1.1
2b		20	140	+6.9

Cpd.	Structure	Exp. Act. (Ki,nM)	Est. Act. (Ki,nM)	Error (Est/Exp)
2c		329	85	-3.9
2d		270	88	-3.1
3 (WR99210)		0.5	1.1	+2.2
4a		540	180	-3
4b		8700	2300	-3.8
5a		9700	20000	+2.1
5b		5200	9000	+1.7
5c		900	1400	+1.5
5d		11300	900	-13
5e		2400	1700	-1.5
5f		3600	8600	+2.4

preferred over the more generic "hydrophobic" feature. On these bases, the following five chemical features, hydrogen bond donor (HBD), hydrophobic aromatic (HYAr), hydrophobic aliphatic (HYAl), positive ionizable

Table 2. Composition, Ranking Score, Statistical Parameters, and Cost Analysis (Expressed in Bits) of the Top 10 Generated Hypotheses[a]

compd	total cost	cost difference (null cost − total cost)	error cost	rmsd	correlation coefficient	features
1	115.6	93.0	89.1	1.009	0.954	HBD, HBD, PI, HYAl, HYAr
2	115.8	92.8	89.7	1.035	0.951	HBD, PI, RA, HYAl
3	115.8	92.8	88.2	0.972	0.958	HBD, PI, RA, HYAl
4	115.8	92.8	89.7	1.036	0.951	HBD, HBD, PI, HYAl
5	116.2	92.4	89.8	1.041	0.951	HBD, HBD, PI, HYAl, HYAr
6	116.3	92.3	90.1	1.052	0.950	HBD, HBD, PI, HYAl
7	116.4	92.2	90.3	1.060	0.950	HBD, HBD, PI, HYAl
8	116.4	92.1	89.8	1.037	0.951	HBD, HBD, PI, HYAl
9	119.1	89.4	92.7	1.155	0.939	HBD, HBD, PI, HYAl, HYAr
10	119.5	89.1	93.1	1.169	0.938	HBD, HBD, HYAl, HYAr

[a] Null cost of top 10 hypotheses is 208.6 bits; fixed cost value is 102.5 bits.

(PI), and ring aromatic (RA), were selected for hypothesis generation.

A total of 10 pharmacophore models were thus generated. The difference between the null and the fixed costs, which should be higher than 70 to indicate a robust correlation, is 106 in our case. Pharmacophore features, ranking scores, and statistical parameters associated with the generated hypotheses are listed in Table 2. The total hypothesis cost, expressed in bits, of the 10 best hypotheses varies from 115.6 to 119.5. Such a small range, covering only 4 bits, suggests that the set of the generated hypothesis is homogeneous and that the selected training set was adequate for pharmacophore design. Moreover, the finding that the total costs were much closer to the fixed cost (102.5) than to the null cost (208.6) indicates that significant models were obtained. Notably, the correlation coefficients between the experimental and the calculated activities range between 0.954 and 0.938. Within the 10 models generated, three models consisted of five features while seven models consisted of four features. All pharmacophores have at least one hydrogen bond donor (HBD) function (8 out of 10 have two HBD functions), one hydrophobic aliphatic group (HYAl), and one positive ionizable (PI) function (except for model 10). Six models include aromatic features, four models with a hydrophobic aromatic (HYAr) function and two models with a ring aromatic (RA) function. As indicated by the small difference in total cost, the generated hypotheses were very similar in terms of composition and 3D orientation of chemical features. All hypotheses were visually inspected by fitting one of the most active compounds in the training set (Pyr) on each generated model to investigate recurrent features. Pyr mapped the PI function with the N_1 nitrogen and the HYAl function with the *p*-chloro substituent (see below). Depending on the model, the phenyl ring was mapped by RA or HYAr functions. The major difference among the 10 models was in the location of HBD features, which mapped both the 2-amino and 4-amino groups of Pyr. Notably, both amino groups form hydrogen bonds with PfDHFR in the structure of the complex.[7,10] Considering that all the generated pharmacophores map the molecules of the training set in a similar way, the first model (Hypo1), characterized by the highest cost difference and the best correlation coefficient, was chosen for subsequent analyses. Hypo1 contains two hydrogen bond donors, one positive ionizable feature, one hydrophobic aliphatic feature, and one hydrophobic aromatic feature. The activities estimated using Hypo1 are reported in Table 1 along with the experimental activities and the errors (expressed as the ratio between the estimated and the experimental activities). Comparison between the estimated activities of the compounds in the training set and their experimentally measured values shows that 20 molecules out of the 23 molecules in the training set have errors less than 4, while the remaining three have errors not higher than 13. These findings indicate a reliable ability to estimate the affinity of the molecules in the training set. Moreover, the cost difference between Hypo1 and the null hypothesis is 93 bits, corresponding to more than a 90% chance of true correlation in the data.[26,27]

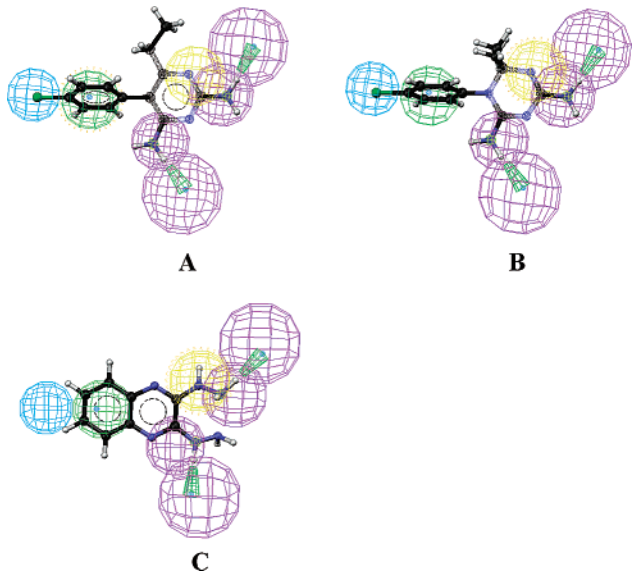


Figure 1. Pharmacophore mapping of some molecules in the training set. Pharmacophore model Hypo1 is aligned to **1** (Pyr) (A), **2** (Cyc) (B), and **5f** (C). Pharmacophore features are purple for hydrogen bond donors (HBD), yellow for positive ionizable (PI), green for hydrophobic aromatic (HYAr), and blue for hydrophobic aliphatic (HYAl).

Figure 1 shows Hypo1 mapped onto Pyr (Figure 1A), Cyc (Figure 1B), and a dihydrazine inhibitor reported by us[11] (Figure 1C). Pyr (Figure 1A) maps onto the selected pharmacophore as follows. The two hydrogen bond donor functions are mapped onto the 2-amino and 4-amino substituents, the positive ionizable group is mapped onto the N_1 nitrogen, and the hydrophobic aromatic function and hydrophobic aliphatic function are mapped onto the *p*-chlorophenyl ring. Pyr derivatives (compounds **1a**−**h**) match the pharmacophore similarly to **1**, and their activities are correctly estimated. A similar mapping was observed for Cyc (Figure

1B), and its derivatives (compounds **2a**−**d**). The estimated activity for cycloguanil is 1.1 nM versus an experimental activity of 1.5 nM. Figure 1C is an example of pharmacophore mapping of a compound less active than Pyr and Cyc; coherently, this compound does not map all the chemical features encoded in Hypo1. Compound **5f** (Figure 1C) maps the two HBD functions by means of the two hydrazine substituents, and the HYAr function with the benzene ring, but it does not map the PI and the HYAl functions. According to its partial mapping, this compound is predicted to be less active.

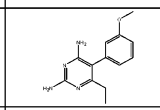
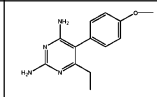
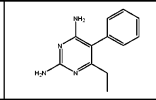
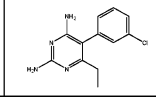
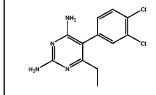
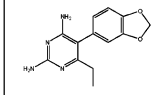
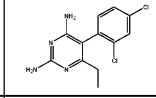
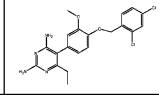
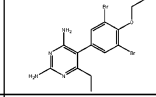
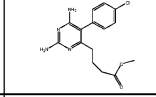
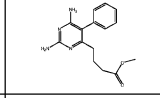
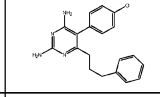
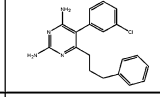
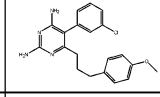
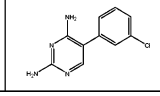
Pharmacophore Validation. Validation is a crucial aspect of pharmacophore design, particularly when the model is built for the purpose of predicting activities of compounds in external test series. Within the compounds of the training set, two statistical methods were used to check the robustness of the correlation. In the first test, the pharmacophore model Hypo1 was evaluated for statistical significance using a randomization trial procedure derived from the Fischer method. The experimental activities of the molecules in the training set were scrambled 19 times (using the catScramble utility) to obtain spreadsheets with randomized activity data. Nineteen hypothesis generations were then performed using the scrambled training sets. Among the 190 resulting hypothesis (10 for each computational run), none of these were found with a cost lower than that of Hypo1. From a statistical point of view, this results in at least a 95% probability that Hypo1 represents true correlation in the data. As a second statistical test, the "leave-one-out" method, which consists of recomputing the hypothesis by excluding from the training set one molecule at a time, was performed. This validation served to check if the correlation was tightly dependent on one particular molecule in the training set and to verify if the activity of each excluded molecule is predicted to be in agreement with experiment. In each case, we have achieved no significant differences between Hypo1 and the hypotheses resulting from the exclusion of one molecule at a time, in terms of correlation coefficient, composition of the pharmacophore, and predicted activity of the excluded molecule.

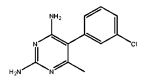
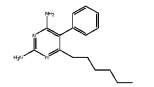
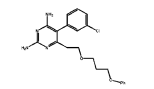
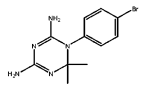
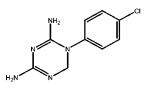
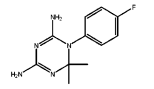
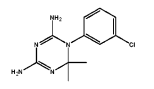
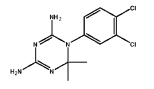
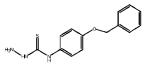
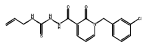
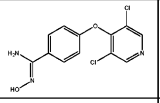
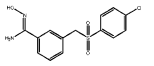
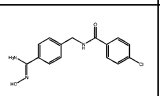
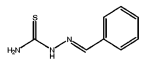
Since the purpose of the pharmacophore hypothesis generation is not only to predict the activity of the training set compounds but also to predict the activities of external molecules, a test set of PfDHFRs inhibitors with known K_i values was prepared. Our target was to verify if the pharmacophore was able to predict the activity of test set molecules in agreement with the experimentally determined value. For external predictions, it is useful to classify the compounds using an activity scale. Therefore, all the selected compounds were classified as follows: <10 nM (+++++), 10−100 nM (++++), 100−1000 nM (+++), 1000−10000 nM (++), 10000−100000 nM (+), >100000 nM (−). They were then submitted to activity prediction. The activities were estimated using Hypo1 as previously described. The chemical structures, the experimental and estimated activities, and activity scales are reported in Table 3. Out of the 29 compounds of the test set, 9 were predicted to be in the correct activity range with a mean error of 4 ± 5, 17 were classified in an adjoining activity range with a mean error of 33 ± 22, and only 3 were estimated to be in separate activity ranges with a mean error of 64 ± 55. By exclusion of compound **5j**, which had the worst prediction, the mean error in predicting activity was 24 ± 22; therefore, given the inherent simplicity of the pharmacophore approach and considering that experimental activities are also affected by errors, we can conclude that the performance of the pharmacophore in predicting activity is more than satisfactory.

One of the main goals of 3D pharmacophores is the generation of a predictive model useful for the identification of active and structurally diverse compounds in large 3D databases of molecules.With the aim of testing further the predictive ability of the generated pharmacophore in database screening, we have performed an activity estimate of two focused databases of compounds extracted from the MDL Drug Data Report.[22] To make the test, the MDDR database was filtered for dihydrofolate reductase and aldose reductase inhibitors, and the two sets were placed in separate focused databases. While the pharmacophore should be able to classify as active most of the compounds in the DHFR database, most of the AR inhibitors should result in being inactive or poorly active. The rationale for this is that AR is an enzyme in which the structural requirements for inhibition differ markedly from those of DHFR.[28−30] The AR active site, in fact, is mainly composed of hydrogen bond donors (Tyr48, His110, and Trp111) that bind compounds such as negatively charged carboxylates, as opposed to DHFR which is mainly composed of hydrogen bond acceptors (Asp54 and the two backbone carbonyls of Ile14 and Ile164) and preferentially binds positively charged compounds such as protonated Pyr and Cyc. The activities of the molecules in the two reduced databases (174 for DHFR and 882 for AR) were predicted using Hypo1 (Table 4). For the DHFR database, 31% of the molecules were predicted to be in the highest activity range (<10 nM), 40% in the activity range 10 nM to 1 μM, and only 29% in the range 1−100 μM; none of the compounds were predicted with activity worse than 100 μM. In contrast, none of the molecules in the AR database were predicted with activity better than 100 nM, and only 10% were predicted to be in the range 100 nM to 1 μM. The majority of the compounds were classified in the lower activity scale.

It is worth noting the ability of the present pharmacophore to pick up compounds selective for PfDHFR versus DHFRs of different species. In fact, the DHFR database from MDDR includes inhibitors of DHFRs of different organisms, and the question is if and how the pharmacophore is able to discriminate among them. Unfortunately, selectivity data for the DHFR inhibitors in MDDR are not available, but interesting observations can be made. First, we found that all compounds based on a 2,4-diaminepyrimidine scaffold substituted with aromatic rings are predicted to be in the highest activity range (+++++), in agreement with their structural similarity with Pyr. Second, compounds with a 2,4-diaminepteridine ring are predicted to be in lower activity ranges (from ++ to +++). Remarkably, these compounds structurally resemble methotrexate, a selective inhibitor of vertebrate DHFRs. Third, previous 3D-pharmacophore modeling made with Catalyst on a set

Table 3. Chemical Structures, Experimental and Estimated Activities (K_i, nM), Relative Errors, and Experimental and Estimated Activity Scales (See Text) of the PfDHFR Inhibitors in the Test Set

Cpd	Structure	Exp. Act. (Ki,nM)	Est. Act. (Ki,nM)	Error	Act. scale	Est. Act. scale
1i		1.07	36	+33	+++++	++++
1j		0.9	1.3	+1.4	+++++	+++++
1k		2.3	100	+43	+++++	++++
1l		0.8	72	+90	+++++	++++
1m		1	0.79	-1.3	+++++	+++++
1n		1.1	72	+65	+++++	++++
1o		12.6	0.92	-13.7	++++	+++++
1p		1	9.6	+9.6	+++++	+++++
1q		7	37	+5.3	+++++	++++
1r		0.3	0.14	-2.1	+++++	+++++
1s		0.6	10	+17	+++++	+++++
1t		0.7	0.74	+1.06	+++++	+++++
1u		1.2	42	+35	+++++	++++
1v		2.2	11	+5	+++++	++++
1w		6.3	75	+11.9	+++++	++++

Cpd	Structure	Exp. Act. (Ki,nM)	Est. Act. (Ki,nM)	Error	Act. scale	Est. Act. scale
1x		1.9	83	+43.7	+++++	++++
1y		0.3	10	+33.3	+++++	++++
1z		0.4	13	+32.5	+++++	++++
2e		1.1	1.2	+1.1	+++++	+++++
2f		24.4	0.84	-29	++++	+++++
2g		4.6	170	+37	+++++	++++
2h		3.7	90	+24.3	+++++	++++
2i		1.1	1.1	1	+++++	+++++
5g		13700	2000	-6.9	+	++
5h		15200	410	-37	+	+++
5i		25200	930	-27	+	+++
5j		102000	800	-127.5	-	+++
5k		77200	1500	-51.5	+	++
5l		23600	31000	+1.3	+	+

of *Mycobacterium avium* and human DHFRs[16] resulted in pharmacophores composed of chemical features that significantly differ from the ones proposed here for *Plasmodium falciparum* DHFR: two hydrogen bond acceptors, one hydrophobic and one ring-aromatic feature for *Mycobacterium avium* DHFR, and three hydrogen bond donors and one hydrophobic feature for human DHFR. To gain additional confidence on the ability of our pharmacophore to discriminate between inhibitors of DHFR of different species, two of the most potent inhibitors from Debnath's work[16] were mapped on our pharmacophore (Figure 2). Compounds **65** and **62** in ref 16, which are nanomolar inhibitors of *Mycobacterium avium* and human DHFR, respectively, mapped only three of the five functions of the pharmacophore (the two hydrogen bond donors and the hydrophobic aromatic function), with estimated PfDHFR inhibition constants of 1,4 and 1 μM, respectively. Therefore, there is a good chance that the pharmacophore generated on the PfDHFR training set is capable of retrieving com-

Table 4. Activity Predictions of the DHFR and AR Inhibitors Selected from MDDR Database Using Hypo1[a]

predicted activity range (nM)	no. of MDDR DHFR inhibitors	no. of MDDR AR inhibitors	% of DHFR inhibitors	% of AR inhibitors
<10 (+++++)	54	0	31	0
10−100 (++++)	13	0	8	0
100−1000 (+++)	56	92	32	10
1000−10000 (++)	39	102	22	12
10000−100000 (+)	12	560	7	63
>100000 (−)	0	128	0	15

[a] The table shows the number, and the related percentages, of molecules in each predicted activity range. The DHFR and AR databases contain 174 and 882 molecules, respectively.

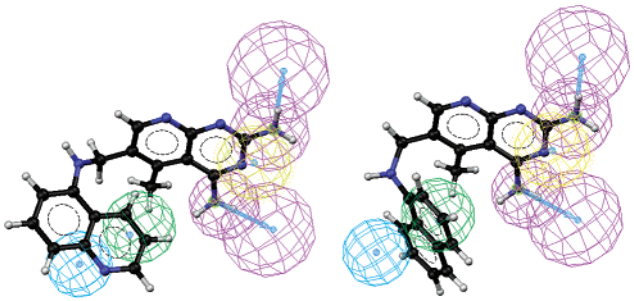


Figure 2. Selectivity of the pharmacophore. The figure shows how the PfDHFR-derived pharmacophore maps two potent inhibitors of different species, one of *Mycobacterium avium* DHFR (A) (compound **65** in ref 16) and one of human DHFR (B) (compound **62** in ref 16). In both cases, only three of the five functions of the pharmacophore could be mapped, and the activities of these molecules were estimated to be much lower than those of Pyr and Cyc.

pounds selective for this particular target, though the DHFR active site is generally conserved. Further work will be needed to include selectivity criteria in pharmacophore design.

To check if the success in predicting activity is somewhat dependent on molecular diversity, the pharmacophore validation was finally completed by calculation of self-similarity using Tanimoto coefficients.[31−33] We believe that this is a critical point in pharmacophore screening of large, and usually redundant, databases of compounds. The estimates included the training and the test sets, the two MDDR subsets of inhibitors (DHFR and AR), and the commercial database Available Chemicals Directory (ACD). Tanimoto coefficients are defined as $N_{AB}/(N_A + N_B - N_{AB})$, where N_A and N_B are the number of bit sets on (i.e., 1) in bit strings (binary representation of molecular structure) of molecules A and B, respectively, and N_{AB} is the number of bits that are in common. The value of the Tanimoto coefficient varies between 0 and 1. The lower the coefficient, the lesser is the similarity between the molecules being compared. The clustering of the molecules in the two subsets of DHFR and AR inhibitors and in the ACD database was evaluated for Tanimoto coefficients ranging from 0 to 1 with 0.1 increments. Comparison of the diversity of the five different databases was achieved by dividing the number of clusters by the number of entries in each database, and the results are graphically reported in Figure 3. First, the number of clusters/entry of the training and the test sets is always much higher than that of the DHFR, AR, and ACD databases,

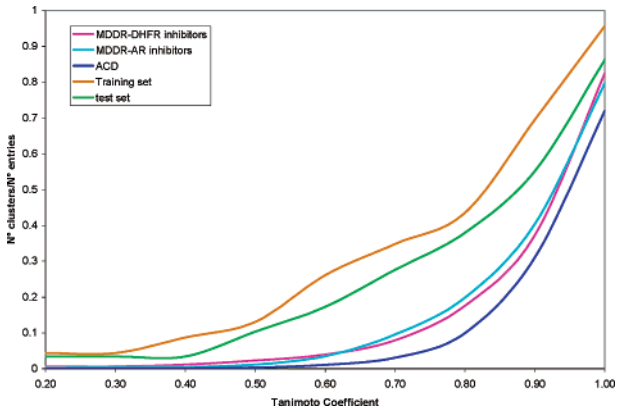


Figure 3. Molecular diversity analysis. Molecular diversity of the training and test sets, of the two databases of DHFR and AR inhibitors and of the ACD database according to Tanimoto coefficients analysis. The number of clusters divided by the number of databases entries is plotted against the Tanimoto coefficient values.

suggesting a better molecular diversity; this confirms that the training and the test sets that were used in the present work are nonredundant, thus meeting one of the conditions for pharmacophore generation. Second, even though the DHFR and AR databases show very similar self-similarity (Figure 3), the pharmacophore proved to be able to correctly discriminate between the two classes of inhibitors. Third, even though the ACD database contains more than 230 000 compounds, its molecular diversity is rather close to that of the DHFR and AR databases for which we obtained good activity predictions, suggesting that our pharmacophore can be applied to screen such a large databases. All these encouraging results, taken together, highlight that the generated pharmacophore can, indeed, be useful to identify novel DHFR inhibitors from databases of compounds.

A final comment on the structural significance of the present pharmacophore is worth noting. We have compared the pharmacophore mapping of Pyr with the corresponding enzyme−inhibitor complex obtained from homology modeling and crystallography.[7,10] For comparison, Figure 4 shows Hypo1 superposed on Pyr (Figure 4A), a diagram of the most important molecular interactions established by Pyr with active site residues (Figure 4B), and a 3D view of the active site of PfDHFR in complex with Pyr (Figure 4C). It is interesting to note that all key interactions between Pyr and PfDHFR (i.e., bidentate hydrogen bonds with the carboxylate side chain of D54 via the N1-protonated nitrogen and the N2 amino groups, hydrogen bonds with the backbone carbonyls of I14 and I164 via the N4 amino group, stacking interaction of the phenyl ring with nicotinamide ring of NADPH, and hydrophobic interactions close to residue S108)[7,10] are all encoded in the pharmacophore. As can be inferred from the comparison of parts A−C of Figure 4, the positive ionizable function located at N1 represents a protonated amine (Pyr, Cyc, and WR99210 bind in protonated form) able to form a charge-reinforced hydrogen bond with residue D54. The adjacent hydrogen bond donor function (mapped on the N2 amino) corresponds to the second hydrogen bond with D54 and completes the bidentate hydrogen bonding with this residue as observed in both Pyr and WR99210 complexes.[10] The pharmacophore also includes a hydro-

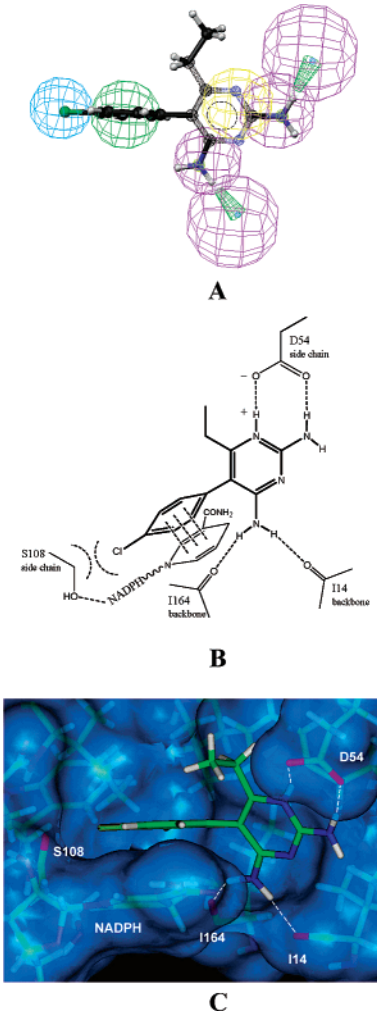


Figure 4. Comparison between generated pharmacophore mapping and active site interaction of Pyr. (A) Pharmacophore model Hypo1 aligned with Pyr, one of the most active compounds in the training set. Pharmacophore features are color-coded: purple for hydrogen bond donors (HBD), yellow for positive ionizable (PI), green for hydrophobic aromatic (HYAr), and blue for hydrophobic aliphatic (HYAl). (B) Diagram showing the most important interactions between Pyr and PfDHFR active site residues as inferred from the enzyme structures. (C) 3D view of active site of PfDHFR in complex with Pyr. The dashed lines show the hydrogen bonds with key residues D54, I14, and I164.

gen bond donor at the N4 amino, which serves as the hydrogen bond donor for the backbone carbonyl of I14 in the structures. Because Catalyst cannot map more than one hydrogen bond donor function per amino group, a function directed toward I164 could not be located. The hydrophobic aromatic function (mapped onto the phenyl ring) encodes for aromatic−aromatic interactions with the nicotinamide ring of the cofactor NADPH, which is close to the phenyl ring in the crystal structures.[10] Finally, the hydrophobic aliphatic function (mapped on the para chlorine atom) accounts for hydrophobic contacts with the methylene group of the S108 side chain (the S108 hydroxyl, in turn, is hydrogen-bonded to the cofactor). In conclusion, the chemical features that a molecule should possess in order to be a good enzyme inhibitor are all represented in our pharmacophore, although it has been generated using a ligand-based approach and without considering structural information of the target.

Conclusions

The work presented in this study shows that, starting from a set of PfDHFR inhibitors with a broad range of activity, a 3D pharmacophore with quantitative predictive ability in terms of activity can be obtained.

The best-generated pharmacophore (Hypo1) consists of a five-feature model (two hydrogen bond donors, one positive ionizable feature, one hydrophobic aliphatic feature, and one hydrophobic aromatic feature). Hypo1 correctly estimates the activity of all the molecules of the training set and shows the best correlation between estimated and experimental activities and the best statistical significance among all of the generated models. Two validation methods, the Fisher test and the "leave-one-out" test, confirmed the statistical significance of the 3D pharmacophore and excluded the chance of a casual correlation between predicted and experimental activity.

The model was predictive not only within the training set but also for a test set of 29 different PfDHFR inhibitors, and with acceptable error. Moreover, it was able to discriminate, and correctly estimate as active or inactive, a number of various DHFR and AR inhibitors extracted from the MDDR database. The validation made on MDDR inhibitors revealed that our model is indeed predictive for compounds that are different from those used to generate and validate the pharmacophore and that it is able to retrieve selective inhibitors of *Plasmodium falciparum* DHFR. The evaluation of molecular diversity of the databases by means of a Tanimoto approach provided confidence that the success in predicting activity does not depend on the diversity or similarity of the molecules in the databases. This finding has important implications for the screening of large and usually redundant databases of compounds.

Finally, even though the pharmacophore was generated using a ligand-based approach, the type and spatial location of the chemical features in our pharmacophore agrees perfectly with the pattern of enzyme−inhibitor interactions identified from homology modeling and crystallography.[7,10] Therefore, the pharmacophore completely fulfills the requirements for an effective interaction of the inhibitors with the active site of PfDHFR.

Together, these results suggest that our model can be used as a 3D query in large 3D databases of compounds to identify and quantitatively estimate new potential DHFR inhibitors. Compared with other drug discovery tools, the pharmacophore approach has the significant advantage that it is fast and able to quantitatively estimate the activity of large number of compounds in a short amount of time. Our model is expected to be useful not only in the identification of new inhibitors from 3D databases but also in the development and optimization of a known series of inhibitors; in particular, the model can be useful to estimate the potential activity of virtual libraries of newly designed PfDHFR inhibitors prior to synthesis or biological testing.

Materials and Methods

Training and Test Sets. In automated pharmacophore generation with Catalyst, the selection of the molecules in the training set must follow certain rules; the minimum number of molecules necessary to ensure statistical predictive power should be 16, the activity data should span over 4−5 orders of

magnitude, each order should be represented by at least three compounds, and finally, each compound in the training set should provide new information concerning binding modalities relevant for predicting biological activity.[34] On these bases, a training set of 23 inhibitors of PfDHFR, with inhibition constants (K_i) between 0.3 and 11 300 nM, was defined. These include 15 derivatives of Pyr, Cyc, and WR99210 found in the literature,[23,24] 2 inhibitors reported in a previous study,[25] and 6 of the 12 new inhibitors disclosed in our previous docking study.[11] The test set used for the validation of the pharmacophore model consisted of 23 Pyr and Cyc derivatives found in the literature[23,24] and the remaining 6 inhibitors reported in our previous work.[11]

All the compounds in the training set and test set were built using the 2D and 3D sketcher of Catalyst; a conformational set was generated for each molecule using the poling algorithm and the "best-quality conformational analysis" method, based on CHARMm force field.[35] All conformations within 20 kcal/mol in energy from the global minimum were saved. An uncertainty factor of 3, representing the ratio range of uncertainty in biological activity value, was set for each compound.

Pharmacophore Generation. The chemical features selection is an important step in pharmacophore generation. Each feature in Catalyst is defined in a "feature dictionary" by a chemical function, a "weight" factor, a location and orientation in 3D space, and a tolerance in location. Features in Catalyst include hydrogen bond acceptor (HBA), hydrogen bond acceptor lipid (HBAl), hydrogen bond donor (HBD), hydrophobic (HY), hydrophobic aliphatic (HYAl), hydrophobic aromatic (HYAr), negative charge (NC), negative ionizable (NI), positive charge (PC), positive ionizable (PI), and ring aromatic (RA). In our case, the generic functions HYAr and PI were modified to map specific functional group in some molecules of the training set. Catalyst allows selection of up to a maximum of five features and definition of the minimum and maximum number of each function to be used. As explained in detail in Results and Discussion, the following five chemical features were selected: hydrogen bond donor (HBD), hydrophobic aromatic (HYAr), hydrophobic aliphatic (HYAl), positive ionizable (PI), and ring aromatic (RA).

During the pharmacophore generation, the "weight variation" option was not allowed. The feature "weights", which represent the orders of magnitude in terms of activity to which the presence of a hypothesis feature contributes, was set as default.

To improve the correlation between estimated and experimental activity, the location tolerance was optimized during generation, using the "variable tolerance" option set to 1. This option compensates for deficiencies in the conformational coverage of flexible compounds and tightens the tolerances when large tolerance values are not needed.

Given that the molecules in our training set are generally small, the minimum spacing between two neighboring functions was set to 1 Å instead of the standard value of 3 Å.

During a hypothesis generation, Catalyst considers and discards many thousands of models and selects the best hypotheses from many possibilities by applying a cost analysis. For each model three cost values, expressed in bits, are assessed: the "null", the "fixed", and the "total" costs. The null and the fixed costs are calculated prior to model generation based on the training set data, the features selected, and the run options. The fixed cost corresponds to the simplest model that fits all data perfectly, while the null cost corresponds to the model with no features and estimates activity as the average of the activity data of the training set. To be significant, a pharmacophore should have a difference between the null and the fixed costs of at least 70 bits. The total cost is calculated as the sum of the weight, error, and configuration costs for each generated model. The weight cost increases in a Gaussian form as the weight of the features in the model deviates from an ideal value (2.0), thus favoring the hypothesis where the feature weight is close to 2. The error cost increases as the rms difference between estimated and measured activities for the training set increases and favors the model with the better correlation coefficient. The configuration cost corresponds to the entropy of hypothesis space and depends on the complexity of the hypothesis space being optimized. The total cost lies between the fixed and the null costs. To be statistically significant, one hypothesis should have a total cost close to the fixed cost and far from the null cost. When the difference between the total cost and the null cost is 60 bits or higher, the predictive correlation probability is 90% or higher. Cost differences of 40−60 bits lead to a predictive correlation probability of 75−90%, while for values below 40 the probability decreases to less than 50%.[26,27]

Pharmacophore Validation. Several validation techniques were used to test the robustness, the accuracy, and the statistical significance of the pharmacophore hypothesis generated from the training set molecules. The first validation technique was based on Fisher's randomization test and was done using the CatScramble utility of Catalyst/HypoGen module. The purpose of this test was to validate the correlation between the structures of the training set molecules and their biological activities. With this aim, the activity values were reassigned after randomization using the CatScramble utility and new spreadsheets were created. Different levels of statistical significance can be achieved by varying the number of the scrambled spreadsheets obtained. A 95% confidence level was chosen for our validation, corresponding to 19 new training sets having randomized activity data. Nineteen hypothesis generation runs were thus performed on the scrambled training sets using the same features and parameters used to generate the original pharmacophore. The second statistical validation was the "leave-one-out" cross-validation test. This method checks if the correlation between experimental and estimated activities is tightly dependent on one particular molecule of the training set by recomputing the pharmacophoric model with the exclusion of one molecule at a time; 23 new training sets were built, each composed of 22 molecules, and 23 HypoGen computations were performed under the usual conditions. For each run, the hypothesis having the best total cost was used to estimate the activity of the excluded molecule and to calculate the new correlation coefficient. The third validation consisted of predicting the activities of compounds included in an external test set. The test set was prepared as described in the training and test set section and subjected to activity prediction using the best-generated pharmacophore.

To get additional confidence on the usefulness of the pharmacophore model for database screening, we validated our model for its ability to predict the activities of dihydrofolate reductase and aldose reductase inhibitors derived from the MDL Drug Data Report.[22] MDDR is an annotated database covering the patent literature, journals, meetings, and congresses, which contains over 132 000 biologically relevant compounds and well-defined derivatives such as drugs launched or under development phase. The MDDR database was filtered for DHFR and AR inhibitors, and the two sets were placed in two separate databases in MDL format. The two subsets of inhibitors were imported in Catalyst and converted into Catalyst databases using the utility CatDB; for each molecule, a conformational set within 20 kcal from the global minimum was generated. Pharmacophore mappings and activity prediction of the DHFR and AR inhibitors were done using the Catalyst "compare fit" option.

Molecular Diversity Analysis. The structural diversity of the training and test sets, of the two DHFR and AR databases as described above, and of the Available Chemicals Directory database was determined and compared. The diversities of the five databases were estimated by means of a Tanimoto distance metric analysis.[31] First, counterions were removed and each database was converted into smiles format, then fingerprints were calculated using the 458 bit descriptors (E_SCREEN) implemented in the CACTVS system.[32] The Tanimoto analysis of the fingerprints was performed using SUBSET 1.0,[33] which estimates the molecular diversity of a database using a clustering approach. The number of clusters

was calculated for Tanimoto coefficients ranging from 0 to 1, with 0.1 increments. Comparison of the diversity of the five databases was achieved by dividing the number of clusters by the number of entries in each database.

All molecular modeling work was performed on Silicon Graphics O_2 workstations.

Acknowledgment. We thank B. Shoichet and Binqing Wei for giving us access, at UCSF, to molecule conformations from the MDDR database. The use of the MDDR in the Shoichet lab is generously supported by MDL Inc., San Leandro, CA. This work was supported by a grant from the European Union (INCO-DEV Grant No. ICA4-CT-2001-10077).

Supporting Information Available: Tables of the DHFR and AR structures extracted from the MDDR database along with activity predictions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Warhurst, D. C. Antimalarial drug discovery development of inhibitors of dihydrofolate reductase active in drug desistance. *Drug Discovery Today* **1998**, *3*, 538−546.
(2) Warhurst, D. C. Resistance to antifolates in *Plasmodium falciparum*, the causative agent of tropical malaria. *Sci. Prog.* **2002**, *85*, 89−111.
(3) Sirawaraporn, W. Dihydrofolate reductase and antifolate resistance in malaria. *Drug Resist. Updates* **1998**, *1*, 397−406.
(4) Yuthavong, Y. Basis for antifolates action and resistance in malaria. *Microbes Infect.* **2002**, *4*, 175−182.
(5) Wirth, D. Malaria: a 21st century solution for an ancient disease. *Nat. Med.* **1998**, *4*, 1360−1362.
(6) Lemcke, T.; Christensen, I. T.; Jorgensen, F. S. Towards an understanding of drug resistance in malaria: three-dimensional structure of *Plasmodium falciparum* dihydrofolate reductase by homology building. *Bioorg. Med. Chem.* **1999**, *7*, 1003−1011.
(7) Rastelli, G.; Sirawaraporn, W.; Sompornpisut, P.; Vilaivan, T.; Kamchonwongpaisan, S.; Quarrell, R.; Lowe, G.; Thebtaranonth, Y.; Yuthavong, Y. Interaction of pyrimethamine, cycloguanil, WR99210 and their analogues with *Plasmodium falciparum* dihydrofolate reductase: structural basis of antifolate resistance. *Bioorg. Med. Chem.* **2000**, *8*, 1117−1128.
(8) Delfino, R. T.; Santos-Filho, O. A.; Figueroa-Villar, J. D. Molecular modeling of wild-type and antifolate resistant mutant *Plasmodium falciparum* DHFR. *Biophys. Chem.* **2002**, *98*, 287−300.
(9) Yuthavong, Y.; Vilaivan, T.; Chareonsethakul, N.; Kamchonwongpaisan, S.; Sirawaraporn, W.; Quarrell, R.; Lowe, G. Development of a lead inhibitor for the A16V+S108T mutant of dihydrofolate reductase from the cycloguanil-resistant strain (T9/94) of *Plasmodium falciparum*. *J. Med. Chem.* **2000**, *43*, 2738−2744.
(10) Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtananukul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M. D.; Yuthavong, Y. Insight into antifolate resistance from malarial DHFR-TS structures. *Nat. Sruct. Biol.* **2003**, *10*, 357−365.
(11) Rastelli, G.; Pacchioni, S.; Sirawaraporn, W.; Sirawaraporn, R.; Parenti, M. D.; Ferrari, A. M. Docking and database screening reveal new classes of *Plasmodium falciparum* dihydrofolate reductase inhibitors. *J. Med. Chem.* **2003**, *46*, 2834−2845.
(12) Palomer, A.; Cabre, F.; Pascual, J.; Campos, J.; Trujillo, M. A.; Entrena, A.; Gallo, M. A.; Garcia, L.; Mauleon, D.; Espinosa, A. Identification of Novel Cyclooxygenase-2 Selective Inhibitors Using Pharmacophore Models. *J. Med. Chem.* **2002**, *45*, 1402−1411.
(13) Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of Three-Dimensional Quantitative Structure−Activity Relationships of P-Glycoprotein Inhibitors and Substrates. *Mol. Pharmacol.* **2002**, *61*, 974−981.
(14) Barbaro, R.; Betti, L.; Botta, M.; Corelli, F.; Giannacchini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. Synthesis, Biological Evaluation, and Pharmacophore Generation of New Pyridazinone Derivates with Affinity toward α_1- and α_2-Adrenoceptores. *J. Med. Chem.* **2001**, *44*, 2118−2132.
(15) Bureau, R.; Daveu, C.; Lancelot, J.; Rault, S. Molecular Design Based on 3D-Pharmacophore. Application to 5-HT Subtypes Receptors. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 429−436.
(16) Debnath, A. K. Pharmacophore mapping of a series of 2,2-diamino-5-deazapteridine inhibitors of *Mycobacterium avium* complex dihydrofolate reductase. *J. Med. Chem.* **2002**, *45*, 41−53.
(17) Debnath, A. K. Generation of predictive pharmacophore models for CCR5 antagonists: study with pteridine- and piperazine-based compounds as a new class of HIV-1 entry inhibitors. *J. Med. Chem.* **2003**, *46*, 4501−4515.
(18) Krovat, E. M.; Langer, T. Non-peptide angiotensin II receptor antagonists: chemical feature based pharmacophore identification. *J. Med. Chem.* **2003**, *46*, 716−726.
(19) Wyss, P. C.; Gerber, P.; Hartman, P. G.; Hubschwerlen, C.; Locher, H.; Marty, H.-P.; Stahl, M. Novel dihydrofolate reductase inhibitors. Structure-based versus diversity-based library design and high-throughput synthesis and screening. *J. Med. Chem.* **2003**, *46*, 2304−2312.
(20) Mattioni, B. E.; Jurs, P. C. Prediction of dihydrofolate reductase inhibition and selectivity using computational neural networks and linear discriminant analysis. *J. Mol. Graphics Modell.* **2003**, *21*, 391−419.
(21) *Catalyst*, version 4.7; Accelrys Inc.: San Diego, CA, 2002.
(22) MDL Information Systems Inc., www.mdli.com.
(23) Yuthavong, Y.; Vilaivan, T.; Chareonsethakul, N.; Kamchonwongpaisan, S.; Sirawaraporn, W.; Quarrell, R.; Lowe, G. Development of a Lead Inhibitor for the A16V+S108T Mutant of Dihydrofolate Reductase from the Cycloguanil-Resistant Strain (T9/94) of *Plasmodium falciparum*. *J. Med. Chem.* **2000**, *43*, 2738−2744.
(24) Tarnchompoo, B.; Sirichaiwat, C.; Phupong, W.; Intaraudom, C.; Sirawaraporn, W.; Kamchonwongpaisan, S.; Vanichtananukul, J.; Thebtaranonth, Y.; Yuthavong, Y. Development of 2,4-Diaminopyrimidines as Antimalarials Based on Inhibition of the S108N and C59R+S108N Mutants of Dihydrofolate Reductase from Pyrimethamine-Resistant *Plasmodium falciparum*. *J. Med. Chem.* **2002**, *45*, 1244−1252.
(25) Toyoda, T.; Brobey, R. K. B.; Sano, G.; Horii, T.; Tomioka, N.; Itai, A. Lead discovery of inhibitors of dihydrofolate reductase domain of *plasmodium falciparum* dihydrofolate reductase-thymidylate sintase. *Biochem. Biophys. Res. Commun.* **1997**, *235*, 515−519.
(26) Sutter, J.; Guner, O. F.; Hoffman, R.; Li, H.; Wadman, M. Effect of variable weight and tolerances on predictive model generation. In *Pharmacophore Perception, Development, and Use in Drug Design*; Guner, O. F., Ed.; International University Line: La Jolla, CA, 1999; pp 501−511.
(27) For cost analysis, see Catalyst 4.7 documentation at www.accelrys.com/doc/life/catalyst47/help/HypoGenAlgRef.doc.html.
(28) Wilson, D. K.; Tarle, I.; Petrash, J. M.; Quiocho, F. A. Refined 1.8 Å structure of human aldose reductase complexed with the potent inhibitor zopolrestat. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9847−9851.
(29) Costantino, L.; Rastelli, G.; Gamberini, M. C.; Vinson, J. A.; Bose, P.; Iannone, A.; Staffieri, M.; Antolini, L.; Del Corso, A.; Mura, U.; Albasini, A. 1-Benzopyran-4-one antioxidants as aldose reductase inhibitors. *J. Med. Chem.* **1999**, *42*, 1881−1893.
(30) Costantino, L.; Rastelli, G.; Vianello, P.; Cignarella, G.; Barlocco, D. Diabetes complications and their potential prevention: aldose reductase inhibition and other approaches. *Med. Res. Rev.* **1999**, *19*, 3−23.
(31) Dean, P. M., Ed. *Molecular Similarity in Drug Design*; Blackie Academic & Professional: London, 1995; pp 111−137.
(32) Ihlenfeldt, W. D.; Takahashi, Y.; Abe, S.; Sasaki, S. Computation and management of chemical properties in CACTVS: an extensible network approach toward modularity and compatibility. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 109−116.
(33) SUBSET is available at http://cactus.nci.nih.gov/SUBSET/.
(34) Catalyst 4.7 documentation at www.accelrys.com/support/life/catalyst/hypogen.html.
(35) Smellie, A.; Teig, S. L.; Towbin, P. Poling: promoting conformational variation. *J. Comput. Chem.* **1995**, *16*, 171−187.

JM040769C