Prediction of Binding Modes for Ligands in the Cytochromes P450 and Other Heme-Containing Proteins

Stewart B. Kirton, Christopher W. Murray,* Marcel L. Verdonk, and Richard D. Taylor
Astex Technology, Cambridge, United Kingdom

ABSTRACT **The cytochromes P450 (P450s) are a family of heme-containing monooxygenase enzymes involved in a variety of functions, including the metabolism of endogenous and exogenous substances in the human body. During lead optimization, and in drug development, many potential drug candidates are rejected because of the affinity they display for drug-metabolising P450s. Recently, crystal structures of human enzymes involved in drug metabolism have been determined, significantly augmenting the prospect of using structure-based design to modulate the binding and metabolizing properties of compounds against P450 proteins. An important step in the application of structure-based metabolic optimization is the accurate prediction of docking modes in heme binding proteins. In this paper we assess the performance of the docking program GOLD at predicting the binding mode of 45 heme-containing complexes. We achieved success rates of 64% and 57% for Chemscore and Goldscore respectively; these success rates are significantly lower than the value of 79% observed with both scoring functions for the full GOLD validation set. Re-parameterization of metal-acceptor interactions and lipophilicity of planar nitrogen atoms in the scoring functions resulted in a significant increase in the percentage of successful dockings against the heme binding proteins (Chemscore 73%, Goldscore 65%). The modified scoring functions will be useful in docking applications on P450 enzymes and other heme binding proteins. Proteins 2005;58:836–844.**
© 2005 Wiley-Liss, Inc.

Key words: docking; scoring functions; Cytochrome P450; heme proteins

INTRODUCTION

The cytochromes P450 (P450s) are a family of heme-containing monooxygenase enzymes involved in a variety of functions including the metabolism of exogenous and endogenous substances in the human body.[1] The P450s are of interest as they are responsible for the metabolism of more than 90% of drugs currently in clinical use.[2] Several of the major drug-metabolizing isoforms (e.g. 2D6, 2C9, 2C19) are predisposed towards genetic polymorphisms.[3–12] The polymorphisms can lead to increased and/or decreased rate of metabolism, and modify the site at which metabolism occurs in the molecule, thus limiting their clinical usage. Drug-drug interactions can also be caused by P450 inhibition and it is often desirable to avoid such interactions.[13–14] Additionally, fast turnover by P450s leads to short half-lifes for drugs, a common problem that consumes significant resource in lead optimization. Current practice in the pharmaceutical industry is to measure P450 inhibition against a number of isoforms. This data is then used to guide the optimization of drug candidates with acceptable half-lifes and P450 profiles. Despite considerable medicinal chemistry knowledge and experience, it is not always possible to optimize the P450 profile of a series of compounds and this can result in down-stream development problems or the termination of the series. In such cases it is anticipated that knowledge of the three-dimensional structure of lead molecules bound to the relevant P450s will lead to a more efficient optimization of the P450 profile. We shall refer to this process as structure-based metabolic design by analogy with structure-based drug design.

Structure-based metabolic design is now becoming a viable strategy owing to the recent structure determination of human cytochrome P450s.[15] It is now possible to determine the binding mode of relevant P450's in complex with compounds having an undesirable P450 profile. An important adjunct to experimental structure determination is the computational prediction of the binding mode of a compound. Accurate computational methods will support the redesign of compounds with diminished P450 affinity without significantly reducing the affinity of the compound for its target protein.

As well as the important application of docking in structure-based metabolic design, some heme-containing proteins are, in their own right, valid therapeutic targets (e.g., nitric oxides synthases[16] and aromatase[17]). Thus, improved methods for the prediction of binding modes to heme-binding proteins other than P450s are also of interest to the pharmaceutical industry.

There are a large number of commercially available protein–ligand docking algorithms (e.g., DOCK,[18] Autodock,[19] FlexX,[20] and GOLD[21–22]). An important characteristic of a good docking algorithm is its ability to

*Correspondence to: Christopher W. Murray, Astex Technology, 436 Cambridge Science Park, Milton Road, Cambridge, CB4 0QA, United Kingdom. E-mail: c.murray@astex-technology.com

Stewart B. Kirton's present address is De Novo Pharmaceuticals Ltd., Compass House, Vision Park, Chivers Way, Histon, Cambridge CB4 9ZR, United Kingdom.

Received 6 August 2004; Accepted 14 October 2004

Published online 13 January 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20389

© 2005 WILEY-LISS, INC.

accurately reproduce (as the top-ranked solution) the experimentally observed binding mode. Recently, in collaboration with others, we have constructed a diverse set of 224 complexes, representative of the protein families contained in the PDB.[23] A subset of 139 of these complexes contains drug-like molecules. The commercial version of the docking program GOLD[26] was shown to accurately predict 79% of these drug-like molecules to within 2 Å RMSD of the experimental binding mode with both Chemscore[24–25] and Goldscore.[21–22] Given the diverse nature of this validation set, it is to be expected that some proteins will be easier to predict than others.

We are interested in assessing the performance of GOLD in reproducing binding modes of ligands in P450's and other heme-containing proteins. Docking against heme-containing complexes might be expected to be difficult because: 1) Certain ligands (e.g., imidazoles) coordinate directly to the heme iron atom and the precise energetics of this contact for different chelating groups needs to be properly balanced with other energetic terms. 2) In the case of the P450s, the environment above the heme group is very hydrophobic compared to other enzymes and some scoring functions and docking methods perform poorly on interactions driven entirely by lipophilic contacts.[59] The performance of the different scoring functions is assessed using a set of 45 drug-like heme-containing complexes chosen from the PDB.[27] Only two of the heme containing complexes are in the original GOLD validation set. To our knowledge this is the first substantial validation of a docking method against the important class of heme-binding proteins.[60]

Two well established scoring functions—Chemscore and Goldscore—were tested for their ability to identify the correct binding mode of a docked ligand. Both functions performed disappointingly on the heme-containing complexes relative to their performance against the diverse GOLD validation set. We further demonstrate how docking performance is significantly improved by using a more sophisticated iron-chelation term and a different treatment of lipophilicity in the scoring function.

MATERIALS AND METHODS

Identification of Complexes

A search was carried out using Relibase (http://relibase.ccdc.cam.ac.uk[33]) for heme-containing complexes: Any complex was considered if it contained two ligands, the first being the heme group) and the second ligand having at least three non-hydrogen atoms within 0.5–8.0 Å of the iron atom of the heme. Eight hundred and sixty-two (862) complexes were returned from this search.

This dataset was filtered by applying the rules for identification of drug-like molecules proposed by Veber et al.[29] For a compound in the complex to be considered drug-like, the ligand has to satisfy the following criteria: Fewer than 10 rotatable bonds and a polar surface area less than or equal to 140 $Å^2$. All other protein/ligand complexes were discarded, which reduced the size of the dataset to 96 complexes.

For those cases where there were duplications of a ligand bound to the same protein (e.g. 1akd, 1dz4, 1dz9, and 5cp4 are all examples of camphor bound to P450cam) the complex with the highest resolution was chosen to be in the validation set. This resulted in an initial validation set containing 52 complexes.

The validation set was then examined in greater detail. Some complexes showed multiple binding modes of ligands in the active site (e.g. 1qmq) and hence were discarded. 1jio, which is a complex with a macrocyclic ligand (6-deoxyerythronolide), and 1k2o, which is a complex with a ruthenium-containing ligand were also omitted from the set, because of the complexity of the ligands.

Some complexes showed ligands bound in strained conformations. 1jip was removed from the test set because of an unusual C–C bond angle in the ligand—a tetrahedral carbon atom was represented as having bond angles close to 90°. 1phb, a complex containing a dichlorinated ligand was also rejected because the crystal structure conformation placed one of the chlorine atoms in close proximity to a carbon atom (interatomic distance of 2.08 Å).

Finally, the recently published 2C5 structures complexed with sulfaphenazole and diclofenac (PDB accession codes 1nb6[30] and 1nr6[31]) and the human 2C9 structure complexed with s-warfarin (1og5[15]) were specifically added to the test set giving a "clean" set of 45 complexes (Table I).

Preparation of Complexes for Docking Studies

The active site of each protein–ligand complex was carefully examined. The pH of the active site was assumed to be approximately 7. As a result, any basic residue in the active site was assumed to be protonated (Arg/Lys → Arg^+/Lys^+), and any acidic residues were considered to be deprotonated (Asp/Glu → Asp^-/Glu^-). Charging of these residues was carried out using InsightII (Accelrys Inc., 2002). Residues forming direct hydrogen bonds with ligands were manually protonated/deprotonated. Protonation states of histidine residues in contact with ligands were also assessed on an individual basis, based on the hydrogen bonds formed by these groups.

Having assigned protonation states of active site residues in each of the protein–ligand complexes the ligands were extracted from the active sites. Each ligand was examined with regards to its own likely protonation state at pH 7. Any formal charges resulting from protonation/deprotonation were assigned manually using Insight II.

Dockings

The ligands were docked back into the native active site of the protein structure, using the docking program GOLD. The active site was defined as all protein atoms within 6 Å of any ligand atom in the experimental protein–ligand complex (this generally encompasses the entire active site). The scoring functions used to rank the dockings were Goldscore[21] and our implementation of Chemscore.[26] Dockings were considered to have been successful if the heavy atom RMSD of the highest ranked solution was $\leq$ 2.0Å from the experimentally determined structure. A maximum of twenty docking solutions were generated for each

TABLE I. Heme-Containing Complexes Validation Set

PDB code	Protein family	Authors/reference
1CK6	Peroxidase	Tsukamoto et al.[36]
1D0C	Bovine E-Nos	Raman et al.[37]
1D1V	Bovine E-Nos	Raman et al.[38]
1D1X	Bovine E-Nos	Raman et al.[38]
1D1Y	Bovine E-Nos	Raman et al.[38]
1DD7	Murine E-Nos	McMillan et al.[39]
1DM6	Bovine E-Nos	Raman et al.[37]
1DM7	Bovine E-Nos	Raman et al.[40]
1DM8	Bovine E-Nos	Raman et al.[40]
1DZ4	P450 Cam	Schlichting et al.[1]
1EA1	P450 51	Podust et al.[41]
1ED4	Bovine E-Nos	Li et al.[42]
1ED5	Bovine E-Nos	Raman et al.[37]
1ED6	Bovine E-Nos	Li et al.[43]
1EGY	P450 Ery-F	Cupp-Vickery et al.[44]
1EUP	P450 Ery-F	Cupp-Vickery et al.[44]
1EWA	Peroxidase	Lacount et al.[45]
1FOI	Bovine E-Nos	Raman et al.[37]
1GDK	Hemoglobin	Harutyunyan et al.[46]
1HSR	Peroxidase	Itakura et al.[47]
1I83	Bovine E-Nos	Raman et al.[38]
1IKE	Nitrophorin	Roberts et al.[48]
1J51	P450 cam	Chen et al.[49]
1LH6	Hemoglobin	Arutyunyan et al.[50]
1LH7	Hemoglobin	Arutyunyan et al.[50]
1M8D	Murine I-Nos	Rosenfeld et al.[51]
1M9K	Human E-Nos	Rosenfeld et al.[51]
1M9M	Human E-Nos	Rosenfeld et al.[51]
1M9Q	Human E-Nos	Rosenfeld et al.[51]
1MPW	P450 cam	Bell et al.[52]
1N6B	P450 2C5	Wester et al.[30]
1NOO	P450 cam	Li et al.[53]
1NR6	P450 2C5	Wester et al.[31]
1OG5	P450 2C9	Williams et al.[15]
1PHF	P450 cam	Poulos and Howard[34]
1PHG	P450 cam	Poulos and Howard[34]
3NOD	Murine I-Nos	Crane et al.[54]
4CPP	P450 cam	Raag and Poulos[55]
4NSE	Bovine E-Nos	Raman et al.[56]
5CPP	P450 cam	Raag and Poulos[57]
5NSE	Bovine E-Nos	Raman et al.[58]
6NSE	Bovine E-Nos	Raman et al.[58]
7CPP	P450 cam	Raag and Poulos[57]
7NSE	Bovine E-Nos	Raman et al.[58]
8CPP	P450 cam	Raag and Poulos[55]

structure, with early termination of the process if the respective RMSDs of the three highest ranked docking solutions were within 1.5 Å RMSD of one another (GOLD default 1 setting: 100,000 Genetic Algorithm (GA) Operations, 5 islands). Flipping of ring free corners, amide bonds and planar nitrogen atoms were allowed.

GOLD attempts to classify each metal in a binding site as either tetrahedral or octahedral, based on the coordinating protein atoms. The coordination sites not occupied by protein atoms are available to ligand acceptors. GOLD classifies the Fe in a heme group as octahedral, which means GOLD considers the coordination site above the plane of the heme to be available to ligand acceptors.

TABLE II. μ_{MA}/C Values for the Different Acceptor Types Derived From CSD-Based Searches

Acceptor type	μ_{MA}/C[a]
N (sp2)	1.4
N (sp3)	1.2
CO (any carbonyl)	0.4[b]
COO (carboxylate)	1.4
OH (any hydroxyl)	0.5
COC (any ether)	0.0
CS-	1.4

[a]C is a constant; we used $C = -6.4$ for Goldscore and $C = 0.64$ for Chemscore.
[b]This value was also used for C = S acceptors

Local Optimization

A local optimization strategy[26] was employed for some test runs in this work. Here a normal GOLD (GA) run is performed initially but all atoms are kept fixed, except terminal protein and ligand OH and NH_3^+ groups. These groups are allowed to spin around to optimize hydrogen bonds. This initial run is followed by a simplex optimization in which all ligand (and protein OH/NH_3^+) torsions and the position and orientation of the ligand are refined to the nearest local optimum.

It is not expected that the locally optimized solution will differ significantly from the crystal structure. Any such deviations could be an indication of a poor structure, or a problem with the searching algorithm. Additionally, for complexes where docking fails, comparison of the scores obtained from the local search with those obtained by the full GA search will indicate if there is a scoring problem or a searching problem.

Adaptations of Scoring Functions

Metal–ligand interactions

Searches were carried out in the Cambridge Structural Database (CSD[28]) and in the Protein Data Bank (PDB[27]) for contacts between Fe atoms and various acceptor atoms. CSD searches were carried out using ConQuest,[32] storing all contacts between iron atoms and acceptor atoms within 6 Å; the types of acceptors considered are listed in Table II. PDB searches were done using the public domain version of Relibase (http://relibase.ccdc.cam.ac.uk[33]). The public domain version of Relibase does not allow searching for contacts between two ligand atoms (the Fe is a ligand atom), so all contacts between Fe atoms and protein acceptor atoms within 6 Å were stored. For each acceptor type, histograms were generated for the Fe···A distances. The frequencies in the histograms were converted into relative propensities, $p_{MA}(r)$, as follows:

$$p_{MA}(r) = \frac{\rho_{MA}(r)}{\langle \rho_{MA}(r') \rangle_{r'=5}^{r'=6}} \text{ with } \rho_{MA}(r) = \frac{n_{MA}(r)}{4\pi r^2 \Delta r}$$

where $\rho_{MA}(r)$ is the contact density at an Fe···A distance r; $\langle \rho_{MA}(r') \rangle_{r'=5}^{r'=6}$ is the average contact density between 5 and 6 Å; $n_{MA}(r)$ is the number of contacts between $r-\Delta r/2$ and $r+\Delta r/2$; Δr is the bin size (we used 0.2 Å). Each propensity

TABLE III. μ_{MA}/C[a] Values for the Different Acceptor Types Derived From PDB-Based Searches

Acceptor type	μ_{MA}/C[a]
N (any N acceptor)	1.2
CO (any carbonyl)	−1.0[b]
COO (carboxylate)	1.0
OH (any hydroxyl)	−0.2
CS-	1.2

[a]C is a constant; we used $C = -6.4$ for Goldscore and $C = 0.64$ for Chemscore.
[b]This value was also used for C = S acceptors.

TABLE IV. Results of Initial Dockings

Scoring function	Successful dockings (%)	Comparison to GOLD "Drug-like List"
Goldscore	57 (4)	79 (1)
Chemscore	64 (2)	79 (1)

Percentage of successful dockings averaged over 10-experiments. Standard deviations are shown in parentheses and take into account the nondeterministic nature of the search algorithm but do not include sampling errors related to the size of the test sets. Comparison to drug-like list refers to results achieved for the GOLD test set.[26]

distribution was converted to a relative interaction strength, μ_{MA}, via:

$$\mu_{MA} = C \cdot \ln[p_{MA}(r_o)]$$

where r_0 is the distance at which $p_{MA}(r)$, is maximal, and C is a constant that brings the interaction strengths at the correct scale for each of the two scoring functions; we used $C = -6.4$ for Goldscore and $C = 0.64$ for Chemscore; C is negative for Goldscore because in this scoring function favorable contributions have a positive sign; conversely, C is positive for the Chemscore function because in that scoring function favorable contributions are negative.

The GOLD source code was modified such that it can read in a parameter file that specifies the μ_{MA} values for each combination of metal and acceptor. The acceptor atoms can be defined quite precisely, e.g. the ether oxygen in an ester linkage, which makes the parameterization very flexible. The Goldscore function was adapted such that the specified μ_{MA} values simply replace the original strengths of these interactions. In the original Chemscore function, the contribution of a metal-acceptor contact is independent of the type of metal and acceptor involved. Hence, we adapted the Chemscore metal term, such that the specified μ_{MA} values can be used:

$$S_{metal} = \sum_{MA} \mu_{MA} \cdot f(r_{MA}, r_{m,1}, r_{m,2})$$

$$f(r_{MA}, r_{m,1}, r_{m,2}) = \begin{cases} 1 & r_{MA} \le r_{m,1} \\ (r_{m,2} - r_{MA})/(r_{m,2} - r_{m,1}) & r_{m,1} < r_{MA} \le r_{m,2} \\ 0 & r_{MA} > r_{m,2} \end{cases}$$

where $r_{m,1}$ and $r_{m,2}$ are constants controlling the range of metal-acceptor interactions; we used $r_{m,1} = 2.6$ Å and $r_{m,2} = 3.0$ Å; the summation is over all protein–ligand metal-acceptor pairs.

Table II shows all μ_{MA} values (divided by the constant C) used in this study for the different acceptor types for CSD-based searches; Table III shows the μ_{MA} values (divided by the constant C) derived from PDB-based searches. Positive μ_{MA}/C values reflect favorable metal-acceptor interactions, whereas negative values indicate unfavorable contacts. A low μ_{MA}/C value does not indicate a low *absolute* number of observed short contacts, but rather a low number of observed short contacts, *relative* to the number of "contacts" at longer distance. Differences in μ_{MA}/C values derived from the CSD and PDB (e.g. for carbonyl oxygens) could reflect the distribution of different types of Fe ions (e.g. in terms of coordination geometry or oxidation state) in these databases.

Lipophilic terms

The lipophilic term in Chemscore has the same functional form as the metal term but is much longer range.

$$S_{lipo} = \sum_{LL} f(r_{LL}, r_{l,1}, r_{l,2})$$

The summation here is over all pairs of lipophilic atoms in protein and ligand. r_{LL} is the distance between protein and ligand for a given pair of lipophilic atoms; r_{l1} and r_{l2} are constants controlling the range of lipophilic atoms.

In the newly modified versions of Chemscore, described in this text, the lipophilic contribution term was altered so that any planar nitrogen atoms, except planar nitrogen atoms in amide groups, were classed as lipophilic.

The adaptations to the scoring functions described above are quite straightforward. They do, however, all involve modifications to the GOLD source code, and therefore the described functionality is currently unavailable in the standard version of GOLD.

Scoring functions

Eight scoring function variants were produced as a result of the re-parameterization of metal and lipophilic terms. Each of these was investigated. In brief the variants are, original Chemscore (CO), original Goldscore (GO), Chemscore + PDB derived metal parameters (CP), Goldscore + PDB metal derived parameters (GP), Chemscore + CSD derived metal parameters (CC), Goldscore + CSD derived metal parameters (GC), Chemscore + PDB derived metal parameters + lipophilic planar nitrogen term (CPN) and Chemscore + CSD derived metal parameters + lipophilic planar nitrogen term (CCN). Each variant was tested with the docking protocols described previously.

RESULTS

The initial dockings, for both Chemscore and Goldscore showed a lower rate of success than was obtained for the standard GOLD validation set (Table IV). When comparing success rates from a small test set (the 45 heme-containing complexes) with those from a large test set (the

TABLE V. Comparison Between RMSDs Achieved Via Unrestrained Docking and Simplex Minimization Using Chemscore†

	RMSD (Å)		ΔG (kJ/mol)	
PDB code	Local optimization	Docking	Local optimization	Docking
1CK6	0.20	3.86	−17.01	−20.47
1D0C	0.21	2.80	−29.90	−31.22
1D1Y	0.36	2.23	−31.92	−38.50
1DD7	0.37	9.09	−32.32	−36.82
1DM6	0.24	5.50	−29.41	−32.32
1DM8	0.49	8.06	−13.78	−13.92
1EA1	0.35	5.36	−15.34	−16.61
1ED5	0.29	6.58	−22.07	−24.95
1EGY	1.06	2.92	−31.80	−35.12
1EUP	0.51	6.42	−33.63	−35.55
1FOI	0.34	4.98	−25.78	−30.84
1M9K	0.79	2.72	−27.30	−28.10
3NOD	0.38	5.81	−16.73	−22.17
8CPP	0.74	3.28	−25.42	−31.70

†Comparison between RMSDs achieved via unrestrained docking and simplex minimization for complexes consistently predicted incorrectly over a range of chemscore functions. The simplex RMSDs suggest that there is no problem with respect to the GA searching method. This is supported by the fact that the Chemscore (ΔG) values are more favorable for the failed dockings than the simplex minimized structures—highlighting a problem with the scoring function.

139 drug-like complexes), care must be taken to assess whether any differences in the success rates are of statistical significance and not artifacts of the reduction in set size.[23] For example, if one has a very large set of complexes for which the success rate is 80% and one chooses five complexes at random then on average one would expect four (80%) of the complexes to be predicted correctly. However, a significant number of random choices could give three (60%) or five (100%) correctly predicted complexes, which, if examined in isolation, could be misleading. If we consider the heme-containing protein test set, the expected success rate for 45 complexes sampled randomly from a very large test set with a success rate of 79% is 79 ± 6%. We can therefore state that the differences in success rates reported in Table IV are likely to be significant.

The possibility that the difference reported in Table IV was due to a problem with the searching algorithm was excluded by performing a local optimization from the experimental binding mode. The Chemscore value returned from the local optimization was compared to the Chemscore value of the initial docking experiments for a selection of compounds that had failed to dock accurately (Table V). It is apparent from these results that the searching algorithm was very unlikely to be responsible for the failure in the dockings. Each of the docked compounds returned a more favorable Chemscore value than the corresponding locally optimized Chemscore value. This indicates that the incorrect dockings were due to the scoring function, and were not due to inefficient sampling of conformational space, in the protein active site, by the search algorithm.

TABLE VI. Comparison of Different Scoring Functions†

Scoring function	Percentage successful dockings (%)
CO	64 (2)
GO	57 (4)
CP	67 (2)
GP	65 (3)
CC	68 (2)
GC	65 (2)
CPN	74 (2)
CCN	73 (1)

†Percentage of successful dockings averaged over 10 experiments. Standard deviations are shown in parentheses. In brief the variants of the scoring functions are: original Chemscore (CO), original Goldscore (GO), Chemscore + PDB-derived metal parameters (CP), Goldscore + PDB metal derived parameters (GP), Chemscore + CSD derived metal parameters (CC), Goldscore + CSD derived metal parameters (GC), Chemscore + PDB derived metal parameters + lipophilic planar nitrogen term (CPN) and Chemscore + CSD derived metal parameters + lipophilic planar nitrogen term (CCN).

An obvious difference between proteins in the current test set and the majority of the proteins in the standard GOLD validation set is the presence of the heme. Both Goldscore and Chemscore lack stringent and explicit parameterization for metal–ligand and heme–ligand contacts, as they are under-represented in the current validation sets. This was reflected in the poor results from the initial runs.

Goldscore dockings

One of the prevailing features of the failed Goldscore dockings was an interaction between oxygen atoms of carbonyl groups on ligands and the iron atom at the centre of the heme [e.g. Fig. 1(A)]. The hypothesis was formed that such interactions did not exist in the structural data. To support this, a search was carried out using ConQuest in an effort to find any structures containing examples of carbonyl oxygen atoms coordinating directly to an iron atom. No structures were returned (for those complexes with carbon monoxide bound the coordination to the metal was shown to be via the carbon atom). It therefore became apparent that the metal-acceptor interaction term in the Goldscore function needed to be re-parameterized in order to eliminate the erroneous metal-ligand contacts. This re-parameterization was based upon empirical data derived from the Protein DataBank and the Cambridge Structural Databases (see Methods section for details).

The results from incorporating the newly parameterized CSD-based and PDB-based metal terms show an increase in the percentage of successful dockings for both Chemscore and Goldscore (Table VI). Some complexes that were previously incorrectly predicted are now consistently predicted in their correct binding mode as a result of these alterations [e.g. Fig. 1(B)]. The percentage improvement for Chemscore is less than for Goldscore. However, this can be rationalized by the fact that there are fewer erroneous metal–ligand contacts observed for the original Chemscore dockings when compared to the Goldscore dockings.

In addition to an overall increase in the percentage of correctly predicted binding modes, the number of com-

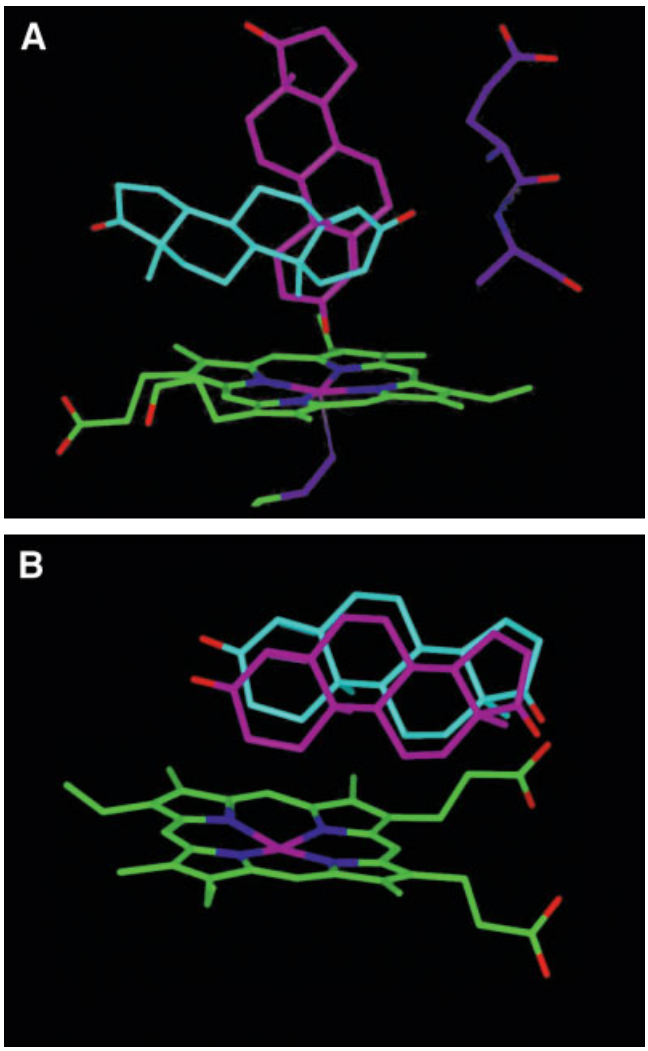

Fig. 1. Example of Goldscore dockings in 1eup. The conformation of 4-androstene-3,17-dione in the 1eup crystal structure is shown in blue; oxygen atoms are shown in red. The conformation of the highest ranked docking is shown in pink. **A:** The docking fails due to an incorrectly parameterized interaction between the oxygen atom of a carbonyl group on the ligand and the iron atom at the center of the heme (RMSD = 5.12 Å). **B:** Incorporation of the new iron chelating parameters results in a successful docking (RMSD = 0.79 Å).

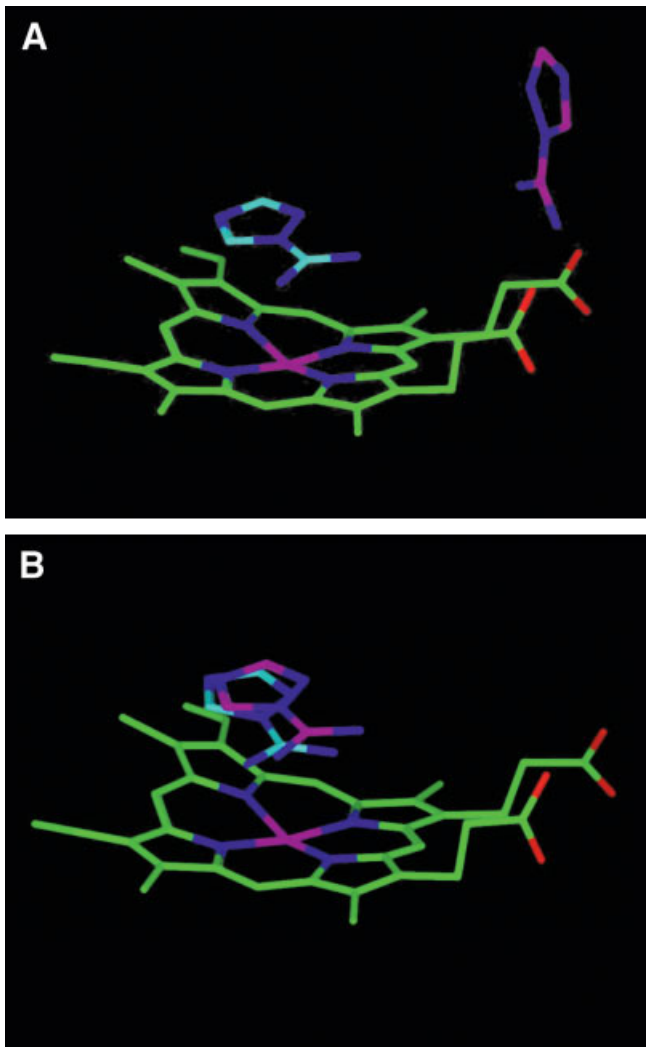

Fig. 2. Example of Chemscore dockings of 1,2,4,-triazole carboxamidine in bovine eNos (1dm8). **A:** The crystal structure (blue) shows the triazole ring parallel and interacting with the pyrole rings of the heme (a predominantly lipophilic interaction). The docked structure (pink) is incorrect (RMSD = 7.2 Å). Insufficient parameterization with regards to the lipophilic nature of nitrogen atoms is responsible for this failure. **B:** Incorporating a term to make planar nitrogen atoms lipophilic results in a successful docking (RMSD = 0.42 Å).

pounds observed to be making erroneous contacts with iron in the Goldscore dockings is significantly reduced. Prior to re-parameterization of the metal terms, 21% of the dockings displayed an incorrect ligand–metal contact. After re-parameterization this figure was reduced to 9% with both the CSD and PDB derived parameters.

Chemscore Dockings

Although there was an improvement for the Chemscore dockings with the incorporation of the re-parameterized metal terms, this improvement was small. It was likely that the Chemscore dockings were failing for reasons other than poor parameterization of metal–ligand interactions. One possible explanation for the failed Chemscore dockings was that the crystal structures involved often had planar lipophilic nitrogen atoms binding in the lipophilic environment above the heme [Fig. 2(A)]. Chemscore does not treat such nitrogen atoms in the ligand or the heme as hydrophobic and so the lipophilic contribution in these cases is significantly under-estimated. Searches in Relibase supported the hypothesis that such atoms were implicated in lipophilic interactions in and around the heme, and this observation was incorporated into Chemscore. Goldscore does not contain a lipophilic term, and hence could not be altered in this manner.

The re-parameterization had a marked effect on successful prediction of binding modes using Chemscore (Table VI), recording a greater than 70% success rate for both CSD and PDB derived parameters. As with the Goldscore results, some complexes that had consistently failed to

TABLE VII. Representative Docking Runs for Each of the Eight Scoring Functions†

Complex	CO	GO	CP	GP	CC	GC	CPN	CCN
1CK6	***3.8***	***4.7***	***2.0***	***4.7***	***2.0***	***3.2***	***3.8***	***3.8***
1D0C	***2.8***	0.4	***2.8***	0.5	***2.8***	0.4	0.1	0.2
1D1V	0.6	0.7	0.4	0.7	0.4	0.6	0.5	0.5
1D1X	0.8	1.0	0.6	0.7	0.7	0.8	0.8	0.7
1D1Y	***2.2***	1.0	***2.5***	***3.9***	1.0	0.9	***2.3***	***2.3***
1DD7	***9.0***	***6.3***	***7.0***	***5.8***	***7.0***	6.4	***6.8***	0.9
1DM6	***5.5***	0.9	***5.4***	1.2	***5.5***	1.2	***5.4***	***5.5***
1DM7	0.4	0.8	0.6	0.8	0.7	0.5	0.4	0.7
1DM8	***8.0***	0.8	***7.5***	0.8	***8.0***	0.7	0.5	1.4
1DZ4	0.9	***2.9***	0.9	***2.9***	0.9	***3.4***	1.0	0.9
1EA1	***5.3***	***2.3***	***5.4***	1.9	***5.4***	***2.2***	***5.4***	***5.4***
1ED4	0.4	0.5	0.5	0.4	0.5	0.5	0.4	0.5
1ED5	***6.5***	0.5	***6.7***	0.6	***6.7***	0.6	***6.6***	***6.9***
1ED6	0.2	0.5	0.4	0.5	0.3	0.4	0.2	0.3
1EGY	***2.9***	***3.1***	0.9	***5.2***	0.6	***5.2***	0.9	0.6
1EUP	***6.4***	***6.3***	***6.4***	0.9	***6.4***	0.7	***6.4***	***6.4***
1EWA	0.7	1.3	0.5	1.3	0.4	1.3	0.7	0.5
1FOI	***4.9***	0.6	***4.9***	0.8	***4.9***	0.5	***4.9***	***5.0***
1GDK	0.8	***3.8***	***3.8***	***3.8***	***3.8***	***3.8***	0.9	0.6
1HSR	0.5	***4.3***	0.6	1.9	0.4	***2.3***	0.4	0.4
1I83	1.6	1.5	1.9	1.4	1.5	1.5	1.7	1.4
1IKE	0.8	***2.1***	0.7	0.8	0.6	0.7	0.7	0.5
1J51	0.9	1.1	0.6	1.1	0.8	1.1	0.5	0.6
1LH6	0.4	0.3	0.6	0.1	0.6	0.2	0.5	0.6
1LH7	0.7	***3.1***	0.5	0.8	0.6	***3.3***	0.6	0.3
1M8D	0.3	0.6	0.2	0.5	0.3	0.6	0.2	0.2
1M9K	***2.7***	0.7	***2.7***	0.6	***2.6***	0.7	***2.6***	***2.7***
1M9M	1.0	***4.8***	1.1	***4.8***	1.0	1.2	1.1	1.0
1M9Q	***4.3***	***4.3***	1.2	***4.3***	1.2	***4.8***	1.1	1.1
1MPW	1.7	***2.6***	1.7	***2.6***	1.8	***2.6***	1.8	1.8
1N6B	***2.5***	***3.5***	***2.7***	1.4	***2.3***	1.4	***2.5***	***2.5***
1NOO	0.6	***3.2***	1.6	***3.2***	1.6	***3.5***	1.6	***3.2***
1NR6	1.1	***3.8***	1.2	***3.7***	1.0	0.8	1.3	0.8
1OG5	0.8	0.4	0.8	0.4	0.9	0.4	0.9	0.9
1PHF	1.2	0.2	0.8	1.5	0.8	1.5	1.2	1.3
1PHG	***4.7***	0.4	***4.8***	0.2	0.6	0.2	***4.8***	***4.8***
3NOD	***5.8***	1.7	1.6	0.7	1.5	0.5	0.7	6.2
4CPP	1.6	1.5	1.6	1.4	1.7	1.3	1.6	1.7
4NSE	0.2	0.3	0.2	0.4	0.3	0.3	0.2	0.2
5CPP	1.4	***2.6***	1.4	***2.5***	1.4	***3.0***	1.3	1.4
5NSE	0.3	0.4	0.3	0.5	0.5	0.4	0.3	0.4
6NSE	0.4	0.5	0.6	0.6	0.3	***2.9***	0.3	0.4
7CPP	1.8	***3.4***	1.7	***3.4***	1.7	***3.8***	1.8	1.8
7NSE	0.2	0.5	0.3	0.5	0.2	0.5	0.3	0.3
8CPP	***3.2***	***3.1***	***3.2***	***3.2***	***3.2***	***3.1***	***3.2***	***3.2***
% Success	64	55	67	67	68	67	73	71

†Representative docking runs for each of the eight scoring functions. Numbers indicate RMSD (Å) from crystal structure. Failures to dock accurately (RMSD > 2.0 Å) are in bold italics.

dock prior to re-parameterization of the lipophilic terms were consistently predicted correctly [Fig. 2(B)].

Overall

A summary of representative results for the dockings across all of the scoring functions is shown in Table VII. Consideration so far has centered on how well the scoring functions and docking protocols reproduce the correct binding mode (as the top ranked solution) with the best scoring functions achieving success rates of over 70%. It is however noticeable that the correct solution is generated by at least one of the scoring functions for all but two of the complexes (1ck6 and 8cpp). GOLD will therefore be good at generating useful putative binding modes in over 95% of cases even if one protocol cannot reliably predict the observed binding mode in all these examples. It is interesting to ask whether there is anything special about 1ck6 and 8cpp that makes them particularly difficult test cases for GOLD.

Examination of the crystal structures showed that for both 1ck6 and 8cpp there was a water molecule located at the sixth coordination site of the heme that was not displaced when the ligand bound in the active site.[36,55] In addition it is feasible that, if oriented appropriately, these water molecules are involved in hydrogen bonding with the ligands, thus mediating the interaction of the ligand with the protein. All water molecules were stripped from the active sites of proteins prior to docking. To have arbitrarily allowed single water molecules to remain in the active site, while discarding others would have arguably biased the investigation. As such, it is expected that accurate binding mode predictions for 1ck6 and 8cpp in the absence of this water molecule are unlikely.

Three complexes, 1m9k, 1n6b, and 8cpp show static disorder of the ligand, i.e., there are two proposed experimental binding modes. It was hypothesized that any failures for these compounds may be attributed to the scoring functions having identified the alternate binding mode. RMSD calculations based on the alternative binding mode of the ligand for these three complexes showed that this was not the case.

Our recent publications[26] advocate the re-scoring of docked solutions using a different scoring function as a way of increasing success rates. Re-scoring of Goldscore dockings with Chemscore for the best performing Goldscore variant increased the percentage success rate from 65 to 70%, suggesting that although Goldscore is capable of identifying the correct binding mode of a ligand it is less successful in ranking the mode accurately than Chemscore. For Chemscore dockings re-scored using Goldscore there is no significant change in success rates. This is consistent with previous findings for default 1 settings with drug-like compounds.[26]

DISCUSSION AND CONCLUSIONS

Recently the first human cytochrome P450 structure was determined.[15] This has opened up the possibility of performing structure-based metabolic design, in which the crystal structures of a lead molecule in relevant P450 isoforms are used to design out P450 liabilities in the lead series. As with structure-based drug design, more accurate computational docking methods are important to optimal exploitation of the crystallographic information. However, current docking procedures have not been substantially validated against heme-binding complexes and so the expected accuracy of the methods is unknown. Additionally it is to be expected that the specialized environment

around the heme will present some specific problems for docking.

We have conducted the first substantial validation of a docking program on a set of heme-binding complexes. We have demonstrated with the two scoring functions in GOLD (Goldscore and Chemscore) that success rates are about 20% lower than would be expected from success rates against general validation sets. We ascribe this difference to interactions with the heme group that are not well treated with current scoring functions. Additionally, this specialized environment is not well represented in previous validation sets; for example, only two out of the 45 heme-binding complexes studied in this paper are contained in the "clean set".[23]

We have improved the scoring functions by using a more rigorous knowledge-based potential for contacts between the heme iron and different acceptor groups. In the case of Chemscore, we have also implemented a better description of lipophilicity that is likely to be more appropriate for the highly lipophilic environment above the heme. These improvements have increased the success rates by around 10%. It is anticipated that such improvements will enable greater confidence in using docking tools in future applications against heme-binding proteins.

A limitation of the docking method and the way we have performed the validation study is the treatment of receptor flexibility. GOLD allows polar hydrogens on the protein (e.g. serine hydroxyls) to move but keeps all other atoms fixed.[21] By their nature, the cytochromes P450 engaged in xenometabolism are promiscuous and are therefore likely to be flexible in their substrate recognition. This mobility has recently been confirmed in studies on 2C9.[15] In the validation study, ligands are removed from the enzyme co-complex, the ligand conformation and orientation is randomized and then the ligand is flexibly docked back in to the native enzyme conformation. In essence, any protein movement occurring on ligand binding has already been taken into account. Despite this bias, our current work is still useful in assessing the performance of docking program against this difficult enzyme class and the improvements we have implemented are real and useful. In applications, we would recommend docking against multiple enzyme conformations where available. In future work we plan to investigate the effect of receptor flexibility and provide mechanisms for including partial receptor flexibility in the docking. We also hope to further improve the scoring functions as more structural information on human P450s becomes available.

In summary we have performed the first substantial validation of docking methods on heme-binding complexes, and improved the success rates by providing a better treatment of the environment around the heme. For example, the modified Chemscore function now reproduces the experimental binding mode for 74% of the heme-binding complexes. The new scoring functions should be useful in docking applications on P450 enzymes and other heme-binding proteins.

ACKNOWLEDGMENTS

The authors thank Pamela Williams for helpful comments with regard to this manuscript.

REFERENCES

1. Schlichting I, Berendzen J, Chu K, Stock AM, Maves SA, Benson DE, Sweet RM, Ringe D, Petsko GA, Sligar SG. The catalytic pathway of cytochrome p450cam at atomic resolution. Science 2000;287:1615–1622.
2. Spatzenegger M, Jaeger W. Clinical importance of hepatic cytochrome P450 in drug metabolism. Drug Metab Rev 1995;27:397–417.
3. Mahgoub A, Idle JR, Dring LG, Lancaster R, Smith RL. Polymorphic hydroxylation of Debrisoquine in man. Lancet 1977;2:584–586.
4. Eichelbaum M, Spannbrucker N, Steincke B, Dengler HJ. Defective N-oxidation of sparteine in man: a new pharmacogenetic defect. Eur J Clin Pharm 1979;16:183–187.
5. Kroemer HK, Eichelbaum M. "It's the genes, stupid." Molecular bases and clinical consequences of genetic cytochrome P450 2D6 polymorphism. Life Sci 1995;56:2285–2298.
6. Sullivan-Klose TH, Ghanayem BI, Bell DA, Zhang ZY, Kaminsky LS, Shenfield GM, Miners JO, Birkett DJ, Goldstein JA. The role of the CYP2C9-Leu359 allelic variant in the tolbutamide polymorphism. Pharmacogenetics 1996;6:341–349.
7. Aithal GP, Day CP, Kesteven PJ, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. Lancet 1999;353:717–719.
8. Kidd RS, Straughn AB, Meyer MC, Blaisdell J, Goldstein JA, Dalton JT. Pharmacokinetics of chlorpheniramine, phenytoin, glipizide and nifedipine in an individual homozygous for the CYP2C9*3 allele. Pharmacogenetics 1999;9:71–80.
9. Kidd RS, Curry TB, Gallagher S, Edeki T, Blaisdell J, Goldstein JA. Identification of a null allele of CYP2C9 in an African-American exhibiting toxicity to phenytoin. Pharmacogenetics 2001;11:803–808.
10. Takahashi H, Echizen H. Pharmacogenetics of CYP2C9 and interindividual variability in anticoagulant response to warfarin. Pharmacogenomics J 2003;3:202–214.
11. Kawamura M, Ohara S, Koike T, Iijima, K, Suzuki J, Kayaba S, Noguchi K, Hamada S, Noguchi M, Shimosegawa T. The effects of lansoprazole on erosive reflux esophagitis are influenced by CYP2C19 polymorphism. Aliment Pharmacol Ther 2003;17:965–973.
12. Sohn D-R. Clinical implications of CYP2C19 polymorphism for tailor-made pharmacotherapy. International Congress Series 2002;1244 (Pharmacogenetics: Tailor-Made Pharmacotherapy). p 41–49.
13. Black DJ, Kunze KL, Wienkers LC, Gidal BE, Seaton TL, McDonnell ND, Evans JS, Bauwens JE, Trager WF. Warfarin-fluconazole. II. A metabolically based drug interaction: in vivo studies.
14. Guengerich FP. Role of cytochrome P450 enzymes in drug-drug interactions. Adv Pharmacol 1997;43:7–35.
15. Williams PA, Cosme J, Ward A, Angove C, Vinkovic DM, Jhoti H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature 2003:424;464–468.
16. Hobbs AJ, Higgs A, Moncada S. Inhibition of nitric oxide synthase as a potential therapeutic target. Annu Rev Pharmacol Toxicol 1999;39:191–220.
17. Brodie AM, Njar VC. Aromatase inhibitors and their application in breast cancer treatment. Steroids 2000;65:171–179.
18. Makino S, Kuntz ID. Automated flexible ligand docking method and its application for database search. J Comput Chem 1997;18:1812–1825.
19. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 1998;19:1639–1662.
20. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–489.
21. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267:727–748.

22. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 1995;245:43–53.
23. Nissink JWM, Murray CW, Hartshorn MJ, Verdonk ML, Cole JC, Taylor R. A new test set for validating predictions of protein-ligand interaction. Proteins 2002;49:457–471.
24. 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 1997:11:425–445.
25. Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD. Flexible docking using Tabu search and an empirical estimate of binding affinity. Proteins 1998;33:367–82.
26. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Proteins 2003;52: 609–623.
27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
28. Allen FH, Motherwell WDS. Applications of the Cambridge Structural Database in organic chemistry and crystal chemistry. Acta Crystallogr B: Structural Science 2002;3:407–422.
29. Veber, DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615–2623.
30. Wester MR, Johnson EF, Marques-Soares C, Dansette PM, Mansuy D, Stout CD. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å. Resolution: evidence for multiple substrate binding modes. Biochemistry 2003;42:6370–6379.
31. Wester MR, Johnson EF, Marques-Soares C, Dijols S, Dansette PM, Mansuy D, Stout CD. Structure of Mammalian Cytochrome P450 2C5 Complexed with Diclofenac at 2.1 Å. Resolution: evidence for an induced fit model of substrate binding. Biochemistry 2003;42:9335–9345.
32. Bruno IJ, Cole JC, Edgington PR, Kessler M, Macrae CF, McCabe P, Pearson J, Taylor R. New software for searching the Cambridge Structural Database and visualizing crystal structures. Acta Crystallogr B: Structural Science 2002;3:389–397.
33. Hendlich M, Bergner A, Gunther J, Klebe G. Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. J Mol Biol 2003;326:607–620.
34. Poulos TL, Howard AJ. Crystal structures of metyrapone- and phenylimidazole-inhibited complexes of cytochrome P-450cam. Biochemistry 1987;26:8165–8174.
35. Raag, R, Poulos TL. The structural basis for substrate-induced changes in redox potential and spin equilibrium in cytochrome P-450CAM. Biochemistry 1989;28:917–922.
36. Tsukamoto K, Itakura H, Sato K, Fukuyama K, Miura S, Takahashi S, Ikezawa H, Hosoya T. Binding of salicylhydroxamic acid and several aromatic donor molecules to Arthromyces ramosus peroxidase, investigated by x-ray crystallography, optical difference spectroscopy, NMR relaxation, molecular dynamics, and kinetics. Biochemistry 1999;38:12558–12568
37. Raman CS, Li H, Martasek P, Southan G, Masters BSS, Poulos TL. Crystal structure of nitric oxide synthase bound to nitro indazole reveals a novel inactivation mechanism. Biochemistry 2001;40:13448–13455.
38. Raman CS, Li H, Martasek P, Babu BR, Griffith OW, Masters BSS, Poulos TL. Implications for isoform-selective inhibitor design derived from the binding mode of bulky isothioureas to the heme domain of endothelial nitric-oxide synthase. J Biol Chem 2001;276:26486–26491.
39. McMillan K, Adler M, Auld DS, Baldwin JJ, Blasko E, Browne LJ, Chelsky D, Davey D, Dolle RE, Eagen KA, and others. Allosteric inhibitors of inducible nitric oxide synthase dimerization discovered via combinatorial chemistry. Proc Natl Acad Sci USA 2000;97: 1506–1511.
40. Raman CS, Li H, Martasek P, Southan G, Masters BSS, Poulos TL. Crystal structure of bovine endothelial nitric oxide synthase complexed with various inhibitors. Unpublished.
41. Podust LM, Poulos TL, Waterman MR. Crystal structure of cytochrome P450 14α-sterol demethylase (CYP51) from Mycobacterium tuberculosis in complex with azole inhibitors. Proc Natl Acad Sci USA 2001;98:3068–3073.
42. Li H, Raman CS, Martasek P, Kral V, Masters BSS, Poulos TL. Mapping the active site polarity in structures of endothelial nitric oxide synthase heme domain complexed with isothioureas. J Inorg Biochem 2000;81:133–139.
43. Li H, Raman CS, Martasek P, Masters BSS, Poulos TL. Crystallographic studies on endothelial nitric oxide synthase complexed with nitric oxide and mechanism-based inhibitors. Biochemistry 2001;40:5399–5406.
44. Cupp-Vickery J, Anderson R, Hatziris Z. Crystal structures of ligand complexes of P450eryF exhibiting homotropic cooperativity. Proc Natl Acad Sci USA 2000;97:3050–3055.
45. Lacount MW, Zhang E, Chen Y-P, Han K, Whitton MM, Lincoln DE, Woodin SA, Lebioda L. The crystal structure and amino acid sequence of dehaloperoxidase from Amphitrite Ornate indicate common ancestry with globins. Unpublished.
46. Safonova TN, Teplyakov AV, Obmolova GV, Popov AN, Kuranova IP, Arutyunyan EG. Crystal structure of ferric complex of the yellow lupin leghemoglobin with isoquinoline at 1.8 Å. resolution. Bioorganicheskaya Khimiya 1991;17:1605–1612.
47. Itakura H, Oda Y, Fukuyama K. Binding mode of benzhydroxamic acid to Arthromyces ramosus peroxidase shown by X-ray crystallographic analysis of the complex at 1.6 Å. resolution. FEBS Lett 1997;412:107–110.
48. Roberts SA, Weichsel A, Qiu Y, Shelnutt JA, Walker FA, Montfort WR. Ligand-induced heme ruffling and bent NO geometry in ultra-high-resolution structures of nitrophorin 4. Biochemistry 2001;40:11327–11337.
49. Chen X, Christopher A, Jones JP, Bell SG, Guo Q, Xu F, Rao Z, Wong L-L. Crystal structure of the F87W/Y96F/V247L Mutant of Cytochrome P-450cam with 1,3,5-trichlorobenzene bound and further protein engineering for the oxidation of pentachlorobenzene and hexachlorobenzene. J Biol Chem 2002;277:37519–37526.
50. Arutyunyan EG, Kuranova IP, Vainshtein BK, Steigemann W. X-ray structural investigation of leghemoglobin. VI. Structure of acetate- ferrileghemoglobin at a resolution of 2.0 7Aring;. Kristallografiya (Russian) 1980;25:80.
51. Rosenfeld RJ, Garcin ED, Panda K, Andersson G, Aberg A, Wallace AV, Morris GM, Olson AJ, Stuehr DJ, Tainer JA, Getzoff ED. Conformational changes in nitric oxide synthases induced by chlorzoxazone and nitroindazoles: crystallographic and computational analyses of inhibitor potency. Biochemistry 2002;41:13915–13925.
52. Bell SG, Chen X, Sowden RJ, Xu F, Willams JN, Wong L-L, Rao Z. Molecular recognition in (+)-A-pinene oxidation by cytochrome P450Cam. Unpublished.
53. Li H, Narasimhulu S, Havran LM, Winkler JD, Poulos TL. Crystal structure of cytochrome p450cam complexed with its catalytic product, 5-exo-hydroxycamphor. J Am Chem Soc 1995;117:6297–6299.
54. Crane BR, Arvai AS, Ghosh DK, Wu C, Getzoff ED, Stuehr DJ, Tainer JA. Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. Science 1998;279:2121–2126.
55. Raag R, Poulos TL. Crystal structures of cytochrome P-450CAM complexed with camphane, thiocamphor, and adamantane: factors controlling P-450 substrate hydroxylation. Biochemistry 1991; 30:2674–2684.
56. Raman CS, Li H, Martasek P, Kral V, Masters BSS, Poulos TL. Crystal structure of constitutive endothelial nitric oxide synthase: a paradigm for pterin function involving a novel metal center. Cell 1998;95:939–950.
57. Raag R, Poulos TL. The structural basis for substrate-induced changes in redox potential and spin equilibrium in cytochrome P-450CAM. Biochemistry 1989;28:917–922.
58. Raman CS, Li H, Martasek P, Masters BSS, Poulos TL. Crystal Structures of the heme domain of bovine endothelial nitric oxide synthase complexed with arginine analogues. Unpublished.
59. Tame JRH. Scoring functions—a view from the bench. J Comput Aided Mol Des 1993;13:99–108.
60. Rosenfeld RJ, Goodsell DS, Musah RA, Morris GM, Goodin DB, Olson AJ. Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. J Comput Aid Mol Des 2003;17:525–536.