Journal of Computer-Aided Molecular Design, 11 (1997) 395–408.

J-CAMD 399

Three-dimensional modelling of human cytochrome P450 1A2 and its interaction with caffeine and MeIQ

J.J. Lozano, E. López-de-Briñas, N.B. Centeno, R. Guigó and F. Sanz*

Research Group on Medical Informatics, Institut Municipal d'Investigació Mèdica, Universitat Autònoma de Barcelona, C/ Dr. Aiguader 80, E-08003 Barcelona, Spain

Received 4 November 1996
Accepted 19 March 1997

Keywords: Sequence alignment strategies; Assessment of secondary structure prediction methods; Solvation; AUTODOCK; Molecular electrostatic potential

Summary

The three-dimensional modelling of proteins is a useful tool to fill the gap between the number of sequenced proteins and the number of experimentally known 3D structures. However, when the degree of homology between the protein and the available 3D templates is low, model building becomes a difficult task and the reliability of the results depends critically on the correctness of the sequence alignment. For this reason, we have undertaken the modelling of human cytochrome P450 1A2 starting by a careful analysis of several sequence alignment strategies (multiple sequence alignments and the TOPITS threading technique). The best results were obtained using TOPITS followed by a manual refinement to avoid unlikely gaps. Because TOPITS uses secondary structure predictions, several methods that are available for this purpose (Levin, Gibrat, DPM, NnPredict, PHD, SOPM and NNSP) have also been evaluated on cytochromes P450 with known 3D structures. More reliable predictions on α-helices have been obtained with PHD, which is the method implemented in TOPITS. Thus, a 3D model for human cytochrome P450 1A2 has been built using the known crystal coordinates of P450 BM3 as the template. The model was refined using molecular mechanics computations. The model obtained shows a consistent location of the substrate recognition segments previously postulated for the CYP2 family members. The interaction of caffeine and a carcinogenic aromatic amine (MeIQ), which are characteristic P450 1A2 substrates, has been investigated. The substrates were solvated taking into account their molecular electrostatic potential distributions. The docking of the solvated substrates in the active site of the model was explored with the AUTODOCK programme, followed by molecular mechanics optimisation of the most interesting complexes. Stable complexes were obtained that could explain the oxidation of the considered substrates by cytochrome P450 1A2 and could offer an insight into the role played by water molecules.

Introduction

Cytochromes P450 are a superfamily of isoenzymes that catalyse the metabolism of a large number of compounds of both exogenous and endogenous origin. Cytochromes P450 play a key role in the metabolism of steroids, vitamins, prostaglandins and leukotrienes, in the metabolic excretion of many drugs and toxins, and in the activation of some compounds to toxins or carcinogens [1]. Cytochrome P450 1A2 is a member of the CYP1 family that is responsible for the metabolism of planar conjugated compounds, such as caffeine [2]. The metabolic activity of P450 1A2 is inhibited by several 8-methyl-xanthines [3] as well as by some quinolones [4]. Other substrates of P450 1A2 are a series of mutagenic heterocyclic amines that are present in cooked meat and fish. Their mutagenic activity is much higher than that of benzopyrenes [5]. These amines are metabolically activated by cytochrome P450 1A2 by conversion of the amino radical into a hydroxyamino group [6]. The result-

*To whom correspondence should be addressed.

Abbreviations: 6-DEB, 6-deoxyerythronolide B; MEP, molecular electrostatic potential; MeIQ, 2-amino-3,4-dimethyl-imidazol[4,5-*f*]quinolone; PFZ, 1-(*N*-imidazol)-2-hydroxy-2-(2,3-dichlorophenyl)octane; rms, root mean square; SCR, structurally conserved region; SRS, substrate recognition site.

ing hydroxyamino derivatives are further activated by forming esters that are the ultimate form that produces DNA adducts [7].

In order to improve insight on the interaction mechanisms of P450 1A2 with substrates and inhibitors, it would be useful to have topological information about this enzyme, including data on its tertiary structure. Experimental methods, such as X-ray crystallography or NMR, are being used to obtain 3D structures of proteins and ligand–protein complexes. However, these techniques are difficult to apply to large membrane proteins like P450 1A2. The only X-ray coordinates of cytochromes P450 available are not membrane bound and belong to prokaryotic microorganisms: P450cam from *Pseudomonas putida* [8], cytochrome P450 BM3 from *Bacillus megaterium* [9], cytochrome P450terp from *Pseudomonas sp* [10] and cytochrome P450eryF from *Saccaropolyspora erythrea* [11]. The 3D structures of these cytochromes show a similar topology, with the greatest structural variability being located in the substrate-binding region [12]. A 3D superposition of the above-mentioned P450s permits the definition of structurally conserved regions (SCRs) [12].

Initial studies on the 3D modelling of microsomal P450s [13–21] used P450cam as the template. Currently, P450 BM3 is the preferred available template, because P450cam is a class I cytochrome that receives its NADH-derived electrons from a two-protein redox chain (FAD reductase → iron–sulphur protein → P450), whereas P450 BM3 is a class II cytochrome that receives electrons directly from an FAD/FMN-containing reductase [22] in a similar manner as microsomal cytochromes such as human P450 1A2 do.

When the degree of homology between the studied protein and the available 3D templates is low, as occurs in members of the cytochrome P450 superfamily, model building becomes a difficult task and the reliability of the results depends critically on the correctness of the sequence alignment. In this case, the accuracy of the alignments may be improved by taking into account further topological information, such as secondary structural elements, mainly α-helices, which can be experimentally known or computationally predicted. The high similarity among experimental 3D structures of prokaryote cytochromes P450, despite their low sequence homology [12], suggests a common fold for the entire family. Furthermore, it has been postulated that many of the secondary structure elements have been preserved across the entire P450 family, including the prokaryotic P450s [23]. These findings provide the rationale for a reliable 3D modelling of mammalian P450s despite their low degree of homology with prokaryotic P450s. On the other hand, site-directed mutagenesis, antibody recognition information and photoaffinity labelling are useful experiments to validate and improve the proposed models.

The first aim of this study was to develop a reliable 3D model of human cytochrome P450 1A2 with special attention to the quality of sequence alignment. As already pointed out by others [24,25], accurate sequence alignment is a crucial step in the homology modelling of proteins. When the sequence alignment is incorrect, the 3D model is guaranteed to be wrong, despite the quality and sophistication of the subsequent methods used in the modelling process. Although the use of sophisticated techniques, such as molecular dynamics computations, could refine the detailed structure of a model, all efforts would have no practical significance if the starting point is inaccurate.

The second objective was to provide a preliminary insight on the interaction of human cytochrome P450 1A2 with their substrates. For this purpose, the docking of caffeine and one of the most active carcinogenic amines (MeIQ) into the binding site of the model was explored. The possible role of water in the complexes was considered by solvating the substrates according to their molecular electrostatic potential distributions that were previously analysed to find a common pharmacophore of the P450 1A2 substrates and inhibitors [26,27].

Methods

The procedure used to build a 3D model of P450 1A2 consists of several steps, which are described in detail in the next sections. The first step is a comparison of results obtained with different alignment strategies in order to select the most reliable choice. Because one of the considered strategies (TOPITS) uses predictions of secondary structure, a study to assess the quality of the available secondary structure prediction methods applied to cytochromes P450 has also been performed. Using the best alignment obtained, a 3D model was built with cytochrome P450 BM3 as the template. Substrate docking models were then generated and analysed.

Sequence alignment

Two different strategies to obtain an alignment of the sequences of human P450 1A2 and a template with known crystal structure have been considered and compared: multiple sequence alignment and the TOPITS threading technique.

Multiple sequence alignment This strategy includes three steps:

(1) A multiple sequence alignment of the four cytochromes with experimentally known 3D structures, obtained by a manual merging of the previously described alignment of three of them (P450cam, P450terp and P450 BM3) [28] with that of P450cam and P450eryF [11].

(2) A multiple sequence alignment of a sample of eukaryote cytochromes (CYP1, CYP2, CYP3, CYP17, CYP21 families; n = 120), obtained using two different

programmes: the multiple sequence alignment option of the CLUSTALW package [29] and the PILEUP programme of the GCC package [30].

(3) Finally, each of the two multiple sequence alignments obtained in the previous step was merged with that of the first step by carrying out a profile alignment using the corresponding option of the CLUSTALW package [29].

TOPITS threading technique This strategy, used to obtain an alignment of the sequences of human P450 1A2 and a template with known crystal structure, was based on the TOPITS threading technique [31]. This technique performs an automatic search for the best remote homologue in the Brookhaven Protein Database (PDB), by taking into account not only the sequence but a PHD [32] prediction of secondary structure and solvent accessibilities of the studied protein. TOPITS supplies a sequence alignment and a confidence score for every proposed solution.

Comparison of sequence alignment strategies The three alignments obtained from these strategies were critically analysed and compared. For this purpose, we considered the percentage of identical and conserved residues along the alignment and the number and location of the insertions and deletions (hereafter called 'gaps'). Ideally, these gaps should not be located in secondary structure motives (i.e. α-helices or β-sheets) and they should not be concentrated in a short segment of the sequence. The best alignment automatically obtained was manually refined in order to further improve the abovementioned criteria.

Evaluation of the accuracy of secondary structure prediction methods

The TOPITS procedure is based on secondary structure predictions carried out with the PHD method. However, many other methods are available for this purpose, and an assessment of the accuracy of predictions resulting from these methods has been considered appropriate. In order to have a P450-customised evaluation, the assessment was carried out using a test set constituted by the four experimentally known cytochromes P450. Seven methods were considered:

(1) The Levin method, based on a search for homologies of local sequences [33].

(2) The Gibrat method, in which the influence of the adjacent sequence upon the conformation of a given residue is calculated using information theory [34].

(3) The double prediction method (DPM), which improves the success rate in secondary structure prediction by taking into account a protein class prediction [35].

(4) The NnPredict method, based on a two-layer, feed-forward neural network [36].

(5) The PHDsec method, a profile network method that uses evolutionary information in the form of multiple sequence alignments [32].

(6) The self-optimised prediction method (SOPM), which optimises its parameters by analysing the predictions of a selected subset of proteins with known structures [37].

(7) The NNSP method, a neural-network method that combines nearest-neighbour algorithms and multiple sequence alignments [38].

The first three methods were used as implemented in the ANTHEPRO package [39]. The remaining methods were used via WWW pages: SOPM at the Protein Sequence Analysis Server of IBCP (Lyon, France) [40]; NnPredict at the Server of the Department of Cellular and Molecular Pharmacology of UCSF (San Francisco, CA, U.S.A.) [41]; PHDsec at the Predict Protein Server of EMBL (Heidelberg, Germany) [42]; and NNSP at the Protein Secondary Structure Prediction Server of the BCM (Houston, TX, U.S.A.) [43].

Secondary structures of the four cytochromes P450 with known X-ray structures (P450cam, P450terp, P450eryF and P450 BM3) that have been used as test sample were obtained from the Database of Secondary Structures of Proteins (DSSP) [44].

The accuracy of the predictions was evaluated separately on α-helices (h) and β-sheets (b). The following accuracy measures were computed:

(1) S1, a measure of sensitivity defined as the proportion of residues actually being in a given structural element, α-helix (S1h) or β-sheet (S1b), that are predicted as members of such a structural element:

$$\mathrm{S1} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

(2) S2, a measure of specificity defined as the proportion of residues predicted as members of a given structural element, α-helix (S2h) or β-sheet (S2b), that actually are in such a structural element:

$$\mathrm{S2} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

(3) Cramer's correlation coefficient, which is a standard measure of attribute correlation. Somehow, CC summarises both S1 and S2, although it cannot be computed from S1 and S2. CC is explicitly defined as

$$\mathrm{CC} = \frac{(\mathrm{TP}\,\mathrm{TN}) - (\mathrm{FP}\,\mathrm{FN})}{\sqrt{(\mathrm{TN}+\mathrm{FN})\,(\mathrm{TN}+\mathrm{FP})\,(\mathrm{TP}+\mathrm{FN})\,(\mathrm{TP}+\mathrm{FP})}}$$

where TP is the number of residues in the sequence correctly predicted as members of a given structural element (α-helix or β-sheet), TN is the number of residues correctly predicted as not being in a given structural element, FP is the number of residues wrongly predicted as members of a given structural element, and FN is the number of residues wrongly predicted as not being in a given structural element.

3D model building

Based on mechanistic reasons explained in the Introduction section and the TOPITS automatic selection of the best homologue, the crystal structure of P450 BM3 (PDB entry 2hpd) was used as the template for the 3D model building of human P450 1A2. No attempts were made to model the first 37 residues which constitute the transmembrane segment and are not present in P450 BM3. The following steps were carried out:

(1) Waters of crystallisation of P450 BM3 were deleted. The oxygen atom complexed with the iron of the heme group was preserved as a molecule of water.

(2) Residues of the 3D structure of P450 BM3 were substituted by those of P450 1A2 using the default options of the homology modelling module of the WHAT IF package [45]. This programme tries to maintain the coordinates of the backbone and carries out a preliminary correction of the side-chain collisions.

(3) Using the homology module of Biosym/MSI [46], gaps between P450 1A2 and P450 BM3 were solved by searching fragments constituted by the concerned residues in the PDB select database (a representative subset of the Brookhaven Protein Databank) [47,48]. The programme provides the 10 best options according to the minimum rms deviation of the C^{α} trace between the residues adjacent to the fragment considered (five adjacent residues per side). We have selected the solution with the maximum number of conserved residues, a low rms of the C^{α} trace and a shape that fits well with the rest of the protein. Once the fragment had been added, the structure obtained was subjected to splice-repair by energy minimisation of the fragment and the adjacent regions.

(4) The protein structure obtained was preliminarily refined using the coordinate regularisation module of the WHAT IF package [45].

(5) A molecular mechanics energy minimisation of the resulting model was carried out with the DISCOVER module of the Biosym/MSI package [49] using a CVFF force field [50] modified to take into account the iron atom of the heme group (M. Paulsen, personal communication). A four-step procedure was used. First, hydrogen atoms were added using the Builder module of INSIGHT II [51]. In order to allow the reorientation of hydrogen atoms according to their local environment, a refinement of their positions was performed using 1000 steps of steepest descent minimisation, while the rest of the protein was kept fixed. Secondly, a constrained optimisation where only the gap zones were free to move was performed using 1000 steps of steepest descent. Thus, 2000 steps of steepest descent were performed keeping only the backbone atoms of the SCRs fixed, as well as the side chains of residues of the SCRs that are conserved among the CYP1A family and are identical in the alignment with BM3. Finally, a steepest descent optimisation of the full protein was performed until the rms gradient energy was lower than 1 kcal/mol Å. During the whole process, the heme group was kept fixed, a distance-dependent dielectric constant of 4r was used and a nonbonded cutoff of 25 Å was applied.

The accuracy assessment of the model along the different steps of the process was monitored with the PROCHECK programme [52] and the structure validation module of the WHAT IF package [53]. In order to have a convenient reference, the same tests were carried out with P450 BM3.

Substrate docking models

Two typical substrates of P450 1A2 were docked into the binding site of the 3D model obtained: caffeine and MeIQ. Both substrates were selected taking into account their high activity in front of the studied enzyme, measured as their capability of inhibiting the phenacetin metabolism by cytochrome P450 1A2 (R. de la Torre and J. Segura, unpublished data).

The following procedure was carried out to obtain putative relative positions of the substrates into the binding site:

(1) 3D structures of caffeine and MeIQ were built using the Builder module of INSIGHT II [51] and were geometrically optimised using the GAUSSIAN 94 [54] programme with the STO-3G basis set.

(2) Water molecules were added to the substrates to take into account the importance of structural waters in the recognition process. The position and number of water molecules were chosen according to the location of the MEP minima of the substrates, taking into account the results of previous studies made by our group [26,27], in which the MEPMIN module of the MEPSIM package was used [55,56]. The MEP computations were performed employing the exact MEP definition and electron density distributions obtained from 3-21G ab initio wave functions computed with the GAUSSIAN 94 programme [54].

(3) The relative position of water molecules with respect to the substrates as well as the significant rotatable angles of the substrates (i.e. methyl groups in caffeine and the amino group in MeIQ) were optimised using the GAUSSIAN 94 programme [54] with the 3-21G basis set.

(4) Docking of the solvated substrates, as rigid systems, into the binding site of the enzyme was explored using the AUTODOCK 2.4 programme [57]. Before using this programme for P450 1A2, its ability to reproduce the X-ray positions of several P450 complexes was tested. Besides the P450cam–camphor structure (PDB entry 2cpp) chosen by Olson and co-workers [57] as a test system for AUTODOCK, we also considered the P450cam–metyrapone (PDB entry 1phg) and P450eryF–6-DEB (PDB entry 1oxa) complexes. Initially, the substrates were placed in an arbitrary position in the binding site. The exploration of docking positions included 100 runs of simulated annealing calculation using an initial temperature of 310 K,

TABLE 1
CHARACTERISTICS OF THE SEQUENCE ALIGNMENTS

	CLUSTALW	PILEUP	TOPITS	Refined TOPITS
Percentage of identical residues	15	16	20	18
Percentage of conserved residues	14	14	16	18
Percentage of (identical + conserved) residues	29	30	36	36
Gaps in loops	14	13	12	13
Gaps in α-helix	6	6	6	1
Gaps in β-sheet	4	5	4	2
Combinations of very close gaps	3	3	1	0

which was linearly decreased along 50 cycles. Each cycle was performed until 2000 structures were either accepted or rejected, and the minimum energy state obtained in one cycle was used as the starting point for the next cycle. During the process, an effective dielectric constant of 4r was used. The 200 resulting positions were clustered according to an rms criterion of 1 Å and the most energetically favourable position of every cluster was visually analysed.

(5) Docking complexes having the chemical moiety that is to be oxidised (i.e. one of the three methyl groups in caffeine or the amine group in MeIQ) properly oriented towards the heme group were selected.

(6) The docking positions selected were subjected to 2000 steps of steepest descent minimisation optimising both internal and relative geometries of the substrate, the binding site residues and the water molecules.

Results and Discussion

Sequence alignment

Three different sequence alignments between human cytochrome P450 1A2 and P450 BM3 were obtained using the procedures described in the Methods section. We will refer to them hereafter as CLUSTALW, PILEUP and TOPITS alignments. Table 1 summarises the most significant features that characterise their reliability: (i) the percentage of identical and conserved residues; and (ii) the number and distribution of the gaps along the alignment. TOPITS generates the best alignment because (i) it maximises the percentages of identical and conserved residues; (ii) it minimises the total number of gaps; and (iii) it produces only one combination of very close gaps. A combination of very close gaps has been defined as two or more gaps concentrated in a short sequence segment. This kind of combinations of gaps should be particularly avoided if they are located in structural motifs. The combination generated by TOPITS involves only very close gaps in β-sheets, whereas those produced by CLUSTALW and PILEUP include this type of gaps not only in β-sheets but also in α-helices, which constitute the most reliable predictions of structural motifs (see the next section).

Although multiple sequence alignments take advantage of an implicit consideration of the relative importance of the sequence regions along the evolution of the P450 superfamily, the avoiding of secondary structure information and the search of a consensus between a large number of sequences generate a distribution of gaps that is often structurally and genetically improbable (gaps including a large number of residues, gaps located in secondary structure motives or combinations of very close gaps). By contrast, TOPITS analysis benefits from being based on pairwise alignments where not only sequence homology but also secondary structure and even solvent accessibility similarities are taken into account.

Not surprisingly, TOPITS suggests P450 BM3 as the best remote homologue for human P450 1A2. The 3.95 value of the Zali score that TOPITS assigns to this alignment implies a correctness confidence greater than 60%. There are further arguments supporting the choice of BM3 as the template instead of any other cytochrome with known crystal structure. As mentioned in the Introduction section, P450 BM3 is a class II cytochrome that receives electrons directly from an FAD/FMN-containing NADPH reductase [14] in a similar manner as microsomal cytochromes such as human P450 1A2 do. On the other hand, P450 BM3 is phylogenetically closer to the CYP1 family than the remaining crystallised P450s [58]. Accordingly, the percentage of sequence identity between human P450 1A2 and the crystallised cytochromes is 21% for P450 BM3, 16% for P450terp and P450cam, and 13% for P450eryF.

A careful inspection of TOPITS alignment, keeping in mind the degree of conservation of the residues along the P450 superfamily and their structural and functional features (see the legend of Fig. 1), has allowed us to perform a manual refinement of the alignment (Fig. 1). Some significant features of such refinement are (i) elimination of a combination of gaps in the β 1-4 sheet; (ii) elimination of five out of six gaps located in α-helices; and (iii) conservation of the sum of identical and conserved residues (see Table 1). The resulting alignment presents 12 insertions and five deletions in P450 1A2 with respect to P450 BM3. Table 2 shows each insertion or deletion by giving their position, length and the secondary structure element involved. This manually refined TOPITS alignment has been used in subsequent steps of this study.

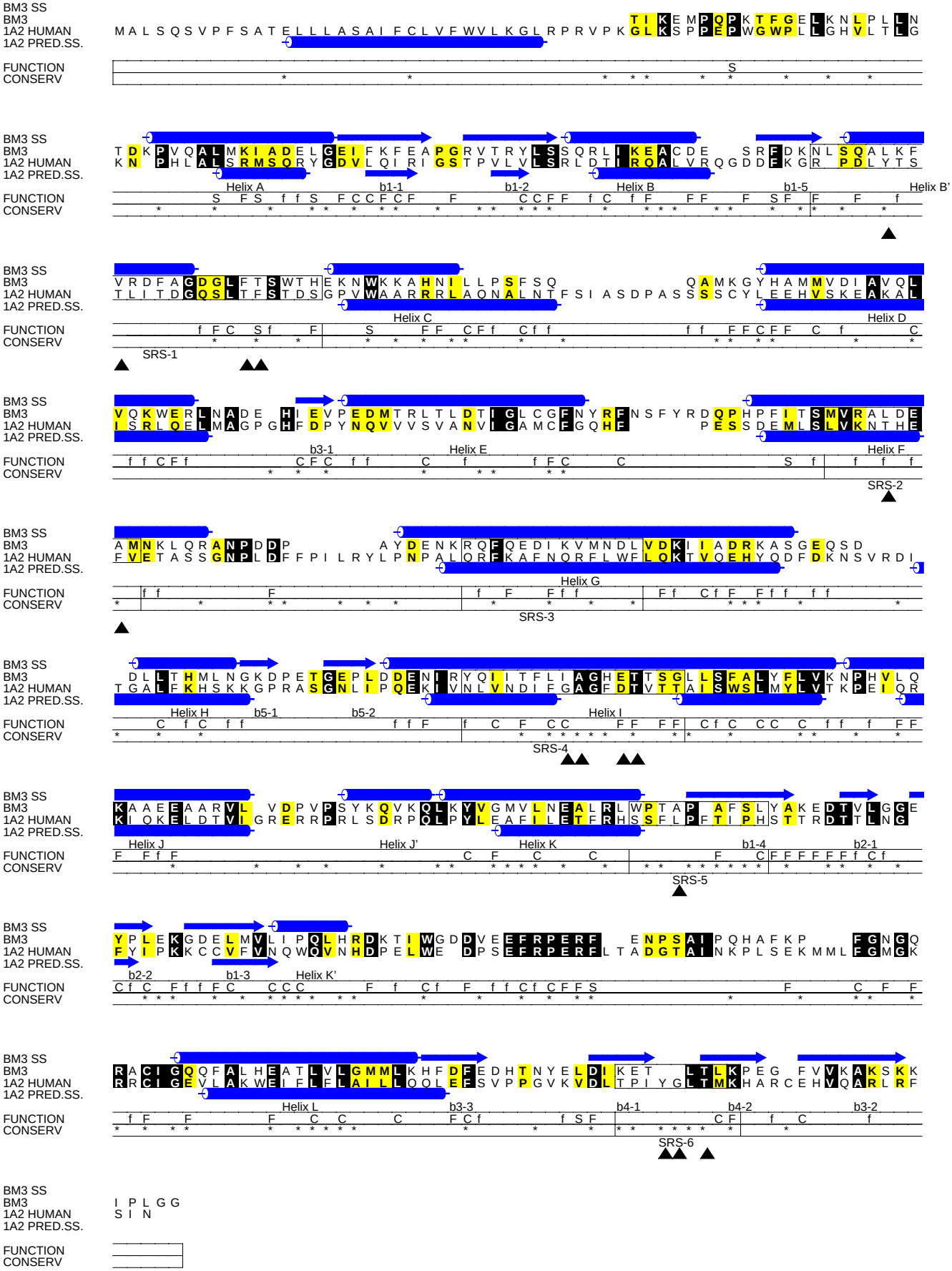
BM3 SS
BM3
1A2 HUMAN
1A2 PRED.SS.
FUNCTION
CONSERV
Helix A
b1-1
b1-2
Helix B
b1-5
Helix B'
Helix C
Helix D
SRS-1
b3-1
Helix E
Helix F
SRS-2
Helix G
SRS-3
Helix H
b5-1
b5-2
Helix I
SRS-4
Helix J
Helix J'
Helix K
b1-4
b2-1
SRS-5
b2-2
b1-3
Helix K'
Helix L
b3-3
b4-1
b4-2
b3-2
SRS-6

TABLE 2
INSERTIONS AND DELETIONS OF HUMAN P450 1A2 VERSUS P450 BM3

Gap no.	Insertion[a]	Deletion[b]	Localisation
1	M1 → K37		N-terminal transmembrane fragment
2		K24	Helix A
3	Q101 → G102		Loop (helix B–β 1-5)
4		L71	Loop (β 1-5–helix B')
5	F147 → S159		Loop (helix C–helix D)
6	G184		Loop (helix D → β 3-1)
7		N163 → R167	Loop (helix E–helix F)
8	F239 → Y244		Loop (helix F–helix G)
9	V280 → T284		Loop (helix G–helix H)
10	G352		Loop (helix J–helix J')
11	F384		β 1-4
12		E344	β 2-2
13		D369	Loop (helix K'–helix L)
14	L433 → T435		Loop (helix K'–helix L)
15	M448 → L450		Loop (helix K'–helix L)
16	Y495 → G496		Loop (β 4-1–β 4-2)
17	C506		Loop (β 4-2–β 3-2)

[a] Sequence numbering of P450 1A2.
[b] Sequence numbering of P450 BM3.

Evaluation of secondary structure prediction methods

Once we have observed the importance of taking into account secondary structure predictions in protein sequence alignments, a question that arises is how good is the secondary structure prediction method implemented in TOPITS and whether there is a better alternative. Table 3 shows the results of the evaluation of a series of secondary structure prediction methods as has been described in the Methods section. It is noteworthy that this evaluation was carried out in a P450-customised manner, so that no general conclusions can be drawn for other protein families. The results in Table 3 are expressed as means and standard deviations (in parentheses) of the results obtained for the four experimentally known structures of cytochromes P450. PHD has yielded the best predictions of α-helices in P450 proteins. On the other hand, poor results in β-sheet predictions have been obtained with all methods. These results suggest that PHD is a good option to carry out predictions of α-helices on cytochromes P450 with unknown secondary structures. In the case of P450 1A2 (see Fig. 1), PHD is able to predict a pattern of α-helices which can be matched with those found in BM3. The only exceptions are helices B' and K', which are not predicted, and helix I which is split into two helices with a central segment predicted as random coil. Interestingly, helix B' is the most variably positioned secondary structure element among the cytochromes with known structure, presenting different lengths and orientations as well as low sequence homology [11,12]. Similarly, P450 BM3 is the only cytochrome with known structure which presents the α-helix K' [9,11,12]. On the other hand, we are almost sure that the prediction as random coil of the central part of helix I is an error of the PHD method because the involved residues are highly conserved.

3D model building

Figure 2 displays the model obtained for human cytochrome P450 1A2. The most evident differences as compared with the template are obviously located in the loops that present gaps in the sequence alignment. Although P450 1A2 presents 12 insertions and five deletions with respect to P450 BM3 (see Table 2), only three gaps of one residue are located in the protein core, causing a minor disruption of the protein fold. The remaining gaps are in loops, most of them being located on the protein surface, where there is enough space to accommodate the insertions or to reconnect the segments when deletions occur.

The geometrical optimisation procedure used has been designed to attain a good balance between the necessity of obtaining an energetically feasible model and the aspiration of maintaining the information coming from the template, particularly on residues conserved in BM3 and the CYP1A family. In such a way, the steepest descent minimisation method has been chosen because it performs minor shifts in the coordinates, and the use of restraints that are gradually removed ensures that the modelled protein will preserve the principal features of the template, especially the whole fold and the orientation of the side chains of conserved residues.

In order to have a more realistic model, it is desirable to take into account the effect of the solvent. In this study, we have implicitly considered solvation through the use of a distance-dependent dielectric constant of 4r plus the explicit addition of a water molecule as the sixth ligand of the iron atom. There are many reasons that

←

Fig. 1. Sequence alignment of human P450 1A2 versus P450 BM3, where structural elements are also shown. β-Sheets are represented as arrows and α-helices as cylinders. BM3 SS indicates the experimental secondary structure elements of P450 BM3, whereas IA2 PRED.SS. indicates the predicted secondary structure of P450 1A2 obtained with TOPITS [31]. Residues in black boxes are identical and those in grey boxes correspond to conservative mutations. Taking into account the conservation of the residues along the CYP1A family and the BM3 structure, and following a procedure similar to that of Valencia et al. [59], residues have been classified into four categories displayed in the FUNCTION row: C = structural core residues; S = structure fine-tuning residues; F = residues of the functional interface regions; f = residues responsible for interspecie differences. On the other hand, residues conserved among all CYP1A members are marked with an asterisk in the CONSERV row. Substrate recognition segments postulated by Gotoh [60] are indicated by SRSn labels (n = 1,2,...,6). In addition, residues involved in our substrate docking models are shown with black up-arrows. Only sequence numbers of P450 1A2 are shown.

TABLE 3
EVALUATION (MEAN + SD) OF SECONDARY STRUCTURE PREDICTION METHODS ON EXPERIMENTALLY KNOWN STRUCTURES OF CYTOCHROMES P450

Method	S1h	S1b	S2h	S2b	CCh	CCb
Levin	0.555	0.270	0.697	0.239	0.360	0.151
	(0.137)	(0.061)	(0.087)	(0.017)	(0.148)	(0.031)
Gibrat	0.684	0.261	0.625	0.207	0.319	0.119
	(0.112)	(0.123)	(0.084)	(0.059)	(0.130)	(0.087)
DPM	0.552	0.112	0.686	0.159	0.335	0.045
	(0.052)	(0.081)	(0.087)	(0.089)	(0.106)	(0.072)
NnPredict	0.473	0.117	0.766	0.173	0.383	0.053
	(0.108)	(0.053)	(0.049)	(0.072)	(0.088)	(0.056)
PHD	0.680	0.239	0.823	0.377	0.560	0.227
	(0.035)	(0.190)	(0.105)	(0.110)	(0.127)	(0.143)
SOPM	0.621	0.437	0.790	0.339	0.493	0.293
	(0.092)	(0.137)	(0.086)	(0.083)	(0.142)	(0.117)
NNSP	0.659	0.249	0.758	0.272	0.485	0.169
	(0.123)	(0.078)	(0.089)	(0.072)	(0.171)	(0.087)

justify this approach. The explicit solvation by soaking the protein with bulk water is a computationally very expensive method that makes it unwieldy for medium-sized proteins [61]. Another approach is to consider only structural waters. However, when the modelled protein exhibits low homology with respect to the template, it is difficult to ascertain the positions of these water molecules. Indeed, the analysis of the P450 substrate-free X-ray structures available indicates that the number and location of structural waters are different in each cytochrome, the only exception being the water molecule near the iron atom [8–10], which is the only explicit molecule of water that has been included in our model. On the other hand, the use of a distance-dependent dielectric constant has become customary in molecular mechanics protein computations to implicitly account for the effect of the solvent [19,61–63]. We have chosen a distance-dependent dielectric constant equal to 4r since it is especially suitable to mimic the solvent environment in the inner part of the protein. Consequently, some uncertainties are expected at the protein surface, but they are not crucial because we are mainly interested in the binding site that is located in the protein core.

Table 4 presents a brief summary of the quality tests of the model obtained for P450 1A2 in comparison with P450 BM3. The ϕ,φ distributions of the Ramachandran plots are summarised as percentages of the number of non-glycine, non-proline residues found in the four different categories following the PROCHECK definition [52]. As expected, the Ramachandran plot of P450 1A2 is

Fig. 2. Model of human cytochrome P450 1A2.

worse than that of P450 BM3, but the values are still suited. The peptide bond planarity is given as the rms deviation from 180° for all the ω angles. The good deviation found in our model is due to the use of a force constant of 50 kcal/mol for the ω angles along the minimisation procedure in order to disallow large deviations that are characteristic in other protein models [24]. χ^1–χ^2 plots were used to assess the quality of the side-chain orientations. Although some side chains were found to adopt an unfavourable conformation, the percentage (2.8%) is really low. Most of them are on the protein surface, where modelling of side chains is expected to be poor. When they are located inside the protein, the unfavourable conformations are the result of avoiding bumps between neighbour residues. No close contacts are found in our model at distances shorter than 3.1 Å. The packing environment of the residues was analysed using the WHAT IF/Quality Control module [64]. A similar percentage of poor packing residues was found in our model as in BM3. Furthermore, a quality packing score was calculated. A score equal to or above −1.0 indicates a reliable structure. Accordingly, our value equal to −1.24 is slightly inaccurate, but it is the same that has been found in a recent related work [25].

In order to check the biological consistency of the present model, data from site-directed mutagenesis [65–68], antibody recognition [69,70], chemical modification [71] and photoaffinity labelling [72] on the CYP1A family have been considered. Based on studies with a chemically modified cytochrome P450 1A1, the involvement of different lysine residues (97, 271, 279 and 407) in the interaction with cytochrome P450 reductase has been postulated [71]. On the other hand, studies with monoclonal antibodies [69,70] have led to the conclusion that two segments of cytochrome P450 1A1 (294–301 and 380–385) should be on the protein surface because of their interaction with cytochrome P450 reductase. In addition, His^{163}Glu [68] and a series of Lys^{453} mutants [69] of rat cytochrome P450 1A2 have confirmed that these residues play a critical role in cytochrome–reductase interaction and should be located at the protein surface. Accordingly, the corresponding residues of P450 1A2 (Arg^{95}, Lys^{266}, Lys^{277}, Phe^{376}–Phe^{381}, Arg^{392}–Gly^{398}, Lys^{403} and Lys^{455}) are located on the surface of our model. Furthermore, as predicted by Edwards et al. [70], segment 392–398 is one of the most exposed loops. In accordance with the proposals of Mayuzumi et al. [68], the superposition of our model with the crystallographic structure of P450cam shows that His^{164} of human P450 1A2 (analogous to His^{163} of rat P450 1A2) and Arg^{130} of P450cam are located at the same site and may play the same biological role.

Both the putative binding site formed by residues located near the heme group and the hydrophobic channel, which allows the access of ligands, are larger in our model than in P450 BM3. Significant differences in the sequence and geometric characteristics of these regions are not surprising because of substrate specificity of the different members of the P450 superfamily. Even then, in the vicinity of the heme group there is approximately the same chemical environment in both cytochromes (1A2 and BM3). Hydrophobic and aromatic residues are predominant, although some charged and polar (Thr) residues that have been postulated to play a catalytic role [65–67] are also found. Among the residues conserved in the two cytochromes, there are the Cys-pocket residues (Phe^{451}, Gly^{452}, Arg^{456}, Cys^{458}, Ile^{459} and Gly^{460}), the residues surrounding the heme propionate groups (Lys^{106}, Trp^{133} and Arg^{456}) and several hydrophobic residues mainly located in the substrate recognition segments (SRSs) postulated by Gotoh [60]: Leu^{123} at SRS-1, Ala^{317}, Gly^{318} and Thr^{321} in SRS-4, Pro^{383} in SRS-5 and Ala^{464} which is located in helix L, near the Cys-pocket. The remaining residues located in the SRSs exhibit a high degree of sequence variability. They could account for differences

TABLE 4
QUALITY TESTS OF HUMAN P450 1A2 VERSUS P450 BM3

	P450 BM3	Human P450 1A2
Ramachandran plot results		
% of residues in most favoured regions	90.1	82.3
% of residues in additional allowed zones	8.9	15.3
% of residues in generously allowed regions	0.5	1.9
% of residues in disallowed regions	0.5	0.5
Standard deviation of ω angles	6.0	5.5
Number of cis residues	0	0
Number of D-aa	0	0
% of residues with poor packing	4.0	5.2
% of residues with favourable conformation (χ^1-χ^2 plots)	100	97.2
Number of bad contacts		
<2.5 Å	0	0
2.5 to 3.1 Å	50	0
Packing quality Z-score	−0.80	−1.24

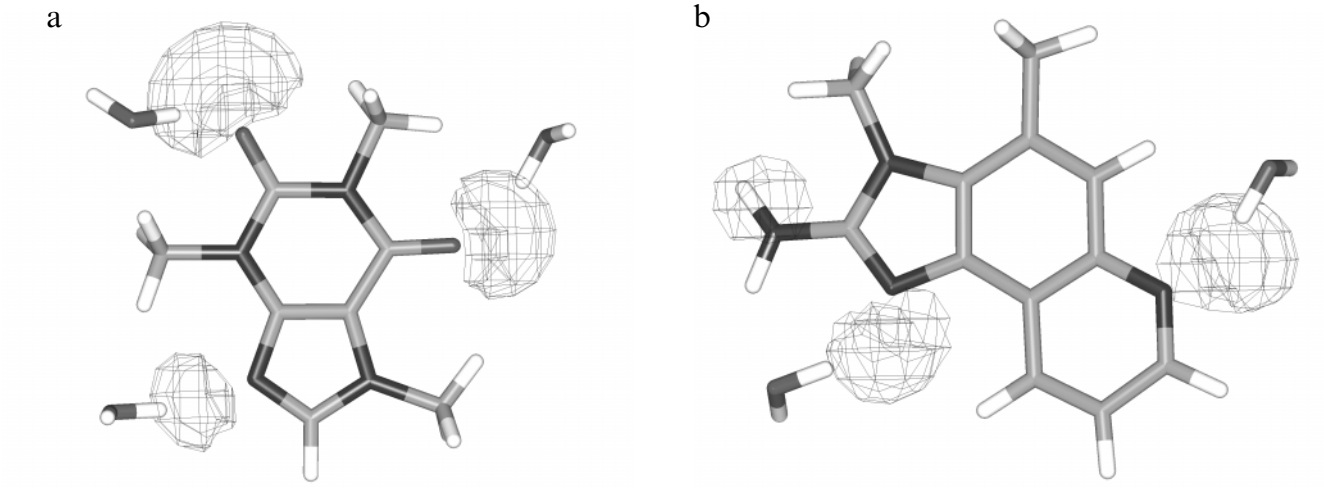

Fig. 3. Solvated substrates as used in the docking analysis: (a) caffeine; (b) MeIQ. MEP surfaces of −30 kcal/mol are shown.

in substrate affinities between the two cytochromes. On the other hand, although the hydrophobic nature of the channel is preserved, there is only one residue conserved (Leu[51]), whereas a series of substitutions changes the topology of the channel as a consequence of the different side-chain sizes (i.e. Phe[87] in BM3 corresponds to Thr[124] in our model). Once again, these differences are probably related to the different nature of substrates that could come into the active site of each cytochrome.

Docking models

Before using the AUTODOCK programme to analyse the interaction of caffeine and MeIQ with cytochrome P450 1A2, we carried out tests of this software on three experimentally known complexes: P450cam with camphor and metyrapone, and P450eryF with 6-DEB. Our results have shown that AUTODOCK is able to reproduce, without ambiguity, the X-ray structures of the complexes of camphor and metyrapone with P450cam. That is, for the two complexes, the lowest energy docking position, which is also the most populated in the AUTODOCK cluster analysis, is very similar to the X-ray position. A more complicated situation arose when the docking of 6-DEB into P450eryF was simulated. When crystallographic water molecules present in the complex were not considered, AUTODOCK provided three docking positions. The lowest energy position, which is also the most populated (92%), does not match with the X-ray structure, whereas this experimental position appears as the third possibility, which is 5.3 kcal/mol less stable and is less populated (4%). Conversely, when crystallographic water molecules were taken into account as a part of the protein, two positions were obtained and, now, the lowest energy position, which is also the most populated (84%), corresponds to the X-ray structure. Consequently, in the absence of experimental data, the results obtained when water molecules are not taken into account could lead to an erroneous description of the interaction.

Despite the importance of considering water molecules in order to get a proper description of the enzyme–ligand interaction, the prediction of the number and position of water molecules of the active site is an extremely difficult task. From the available experimental data on P450 cytochrome complexes, a variety of situations have been observed. In the case of cytochrome P450cam, which has six molecules of water in the binding site in the substrate-free

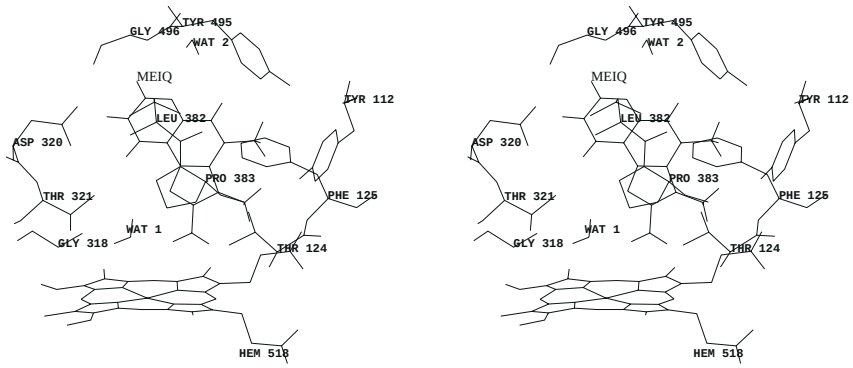

Fig. 4. Docking model of MeIQ in the binding site of human P450 1A2, showing the residues involved in the interaction.

structure [8], there are three complexes (camphor, 4-phenylimidazol and metyrapone) that do not show any water molecule in the vicinity of the ligand [73,74]. Nevertheless, the complex of this cytochrome with 2-phenylimidazol shows a water molecule that enables hydrogen bonds between the ligand and two residues of the binding site [74]. In the same way, the complex with PFZ exhibits two water molecules in the vicinity of the ligand, although only one of them mediates the interaction between the ligand and the protein [75]. In the case of cytochrome P450eryF, only the bound form with DEB-6 is available, and it shows three water molecules stabilising the complex between the substrate and the enzyme [10]. From these experimental data, we can assume that the number, location and role of the water molecules depend on the nature of the substrate. It seems that the higher the hydrophilicity of the substrate, the greater the number of water molecules involved. These findings suggest the strategy of docking solvated substrates into a water-free binding site model, instead of determining the number and position of the water molecules in the active site upon binding the ligand. Despite their differences, the P450eryF–6-DEB complex is the nearest in terms of hydrophilicity to the complexes studied in the present work. Consequently, it is feasible to expect that a certain number of water molecules could partake in the substrate–P450 1A2 interactions.

As a preliminary approach to consider the solvation of the substrates, water molecules were placed in the vicinity of the MEP minima and a 3-21G geometrical optimisation was performed. A test of this approach with the P450eryF–6-DEB complex showed that every experimental water molecule is placed in the vicinity of a deep MEP minimum of 6-DEB, although not all the MEP minima indicate water molecules in the complex. Following this strategy, caffeine was solvated with three molecules of water (Fig. 3a) and MeIQ with two (Fig. 3b). Thus, the exploration of the docking position of the solvated substrates, as a rigid system, was performed with AUTODOCK, which also considers the protein as a rigid body. Obviously, this is quite a rough approach, and, for this reason, after selecting the solutions having biological meaning, they were optimised within the molecular mechanics framework.

The docking positions obtained for MeIQ were clustered into 33 groups. Only one of them corresponds to a position that is biologically outstanding. This position is the lowest energy and the most populated (22%) one. Figure 4 shows this position after molecular mechanics optimisation. The substrate appears centred in the binding site, with the amino group to be oxidised properly oriented towards the heme group. This docking position is mainly stabilised by van der Waals interactions of MeIQ with hydrophobic (Thr^{124}, Gly^{316}, Ala^{317} and Leu^{382}) and aromatic (Tyr^{112}, Phe^{125} and Tyr^{495}) residues. Interestingly, Tyr^{112} is the only conserved residue of SRS-1 along the CYP1A family and Tyr^{495} is located in a similar 3D arrangement as Tyr^{96} in P450cam, which is known to play

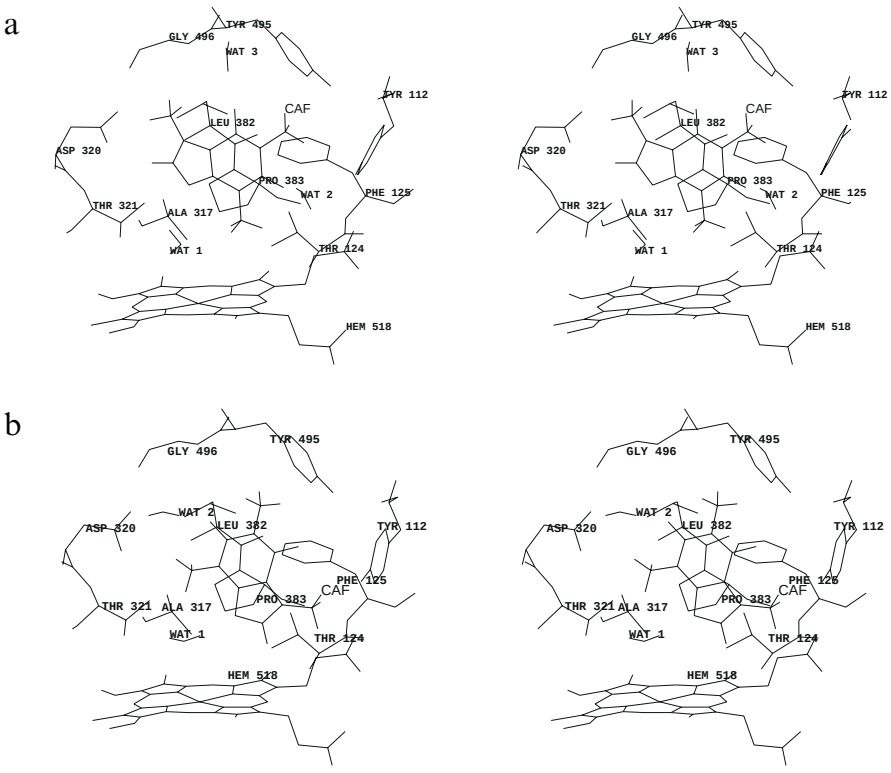

Fig. 5. Docking models of caffeine in the binding site of human P450 1A2, showing the residues involved in the interaction.

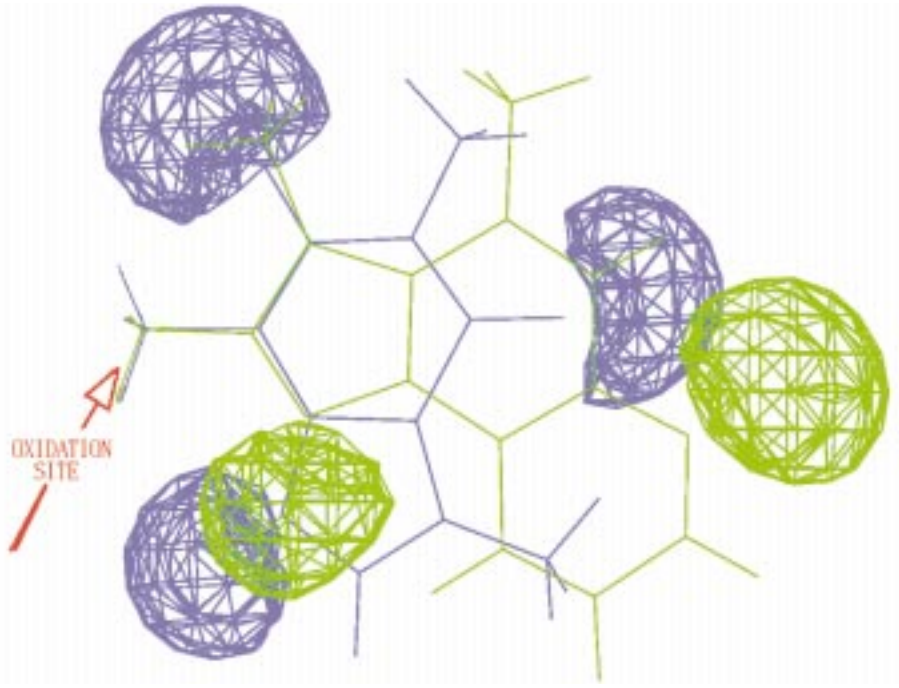

Fig. 6. Fitting of caffeine and MeIQ based on their MEP distributions. MEP surfaces of −30 kcal/mol are shown.

an important role in binding some substrates [73–75]. The water molecule initially placed near the amino group remains in this position and makes a hydrogen-bonded bridge between the heme group, Thr^{321} and MeIQ. Thr^{321} is a highly conserved residue that has been postulated to play a catalytic role in P450 cytochromes [66,76]. The second water molecule goes far away from MeIQ after minimisation and does not seem to be important for the interaction. Indeed, the same docking position was found as the result of an AUTODOCK exploration when this second molecule of water was removed from the substrate.

The AUTODOCK positions of caffeine were clustered into 28 groups. In the case of this substrate, more complex results could be expected, since it has several metabolic pathways catalysed by cytochrome P450 1A2 (N1, N3 and N7 demethylations, and C8 hydroxylation) [2,77]. Accordingly, different docking positions of caffeine in the binding site were obtained. Each position has one of the possible groups to be oxidised oriented towards the heme group. Since N3 demethylation is the main metabolic pathway, we have focused our analysis on two positions having this moiety properly oriented towards the heme group.

The first above-mentioned docking position is shown in Fig. 5a after molecular mechanics optimisation. In this complex, all the water molecules form hydrogen bonds between caffeine and the enzyme. Besides the heme group, the residues involved in the hydrogen bonds are Ala^{317} and Tyr^{495}. The complex is further stabilised by van der Waals interactions between caffeine and hydrophobic (Thr^{124}, Ala^{317}, Leu^{382}, Pro^{383} and Gly^{496}) and aromatic (Tyr^{112}, Phe^{125} and Tyr^{495}) residues. These hydrophobic interactions are almost the same as those observed in MeIQ. Actually, the docking position found for MeIQ and this docking position of caffeine are in agreement with a fitting of the two substrates previously postulated on the basis of their MEP distributions [27], which is shown in Fig. 6.

The second docking position constitutes a completely different arrangement of caffeine into the binding site. Thus, a changed pattern of interactions is observed, although the residues implicated are almost the same. Among the three water molecules, there is one that leaves the substrate after minimisation and, consequently, it does not seem directly involved in the interaction. After removing this water molecule, another AUTODOCK exploration was carried out and an improved docking position was found (Fig. 5b). In this position, a water molecule is located in between the heme group and the N3-methyl, and the second water molecule makes a hydrogen-bonded bridge between caffeine and Asp^{320} and Gly^{496}. The stabilising interaction between caffeine and Asp^{320} agrees with experimental evidence of the importance of a negatively charged residue (Asp or Glu) at this position in P450 cytochromes [66,76]. This position is further stabilised by the same above-mentioned van der Waals interactions.

Conclusions

A homology 3D model of human cytochrome P450 1A2 was built using P450 BM3 as the template. Since the

degree of homology between them is low, a careful analysis of the methodologies considered was made. Special efforts were devoted to obtain a reliable sequence alignment, since it is the most critical step in the homology modelling process. The reliability of the model was assessed by a systematic comparison of its stereochemical and packing properties versus those of the template, and by analysing the consistency of the model with experimental data on the CYP1A family. Thus, the purpose of obtaining a model P450 1A2 to improve our understanding of this important biological system seems to be accomplished. However, there is still room for further methodological studies, since there are some aspects (i.e. geometric optimisation procedures, docking strategies, solvation) that could be undertaken following other approaches. Although one cannot have the same degree of confidence in the model as in an experimentally determined structure, the careful building process and the results obtained from the tests indicate that our model is reliable enough to be used for exploring substrate docking. Preliminary work has been done on this issue and promising results have been obtained, although further investigation is needed in order to gain insights into the molecular mechanism of the interaction between human cytochrome P450 1A2 and its substrates and inhibitors. The next step in this line of research would be the docking analysis of a series of substrates and inhibitors in order to derive quantitative structure–activity relationships. Finally, it should be pointed out that the robust strategy developed could serve as a guide to model other mammalian cytochromes P450 and related systems.

Acknowledgements

We thank Marta Pulido, M.D., for editing the manuscript and editorial assistance. We thank Dr. A. Valencia for helpful discussions. This work was partially funded by the CICYT, grant SAF93-0722-C02-02.

References

1 Black, S.D. and Coon, M.J., Adv. Enzymol., 60 (1987) 35.
2 Grant, D.M., Campbell, M.E., Tang, B.K. and Kallow, W., Biochem. Pharmacol., 36 (1987) 1251.
3 Segura, J., Roberts, D.J. and Tarrús, E., J. Pharm. Pharmacol., 41 (1988) 129.
4 Fuhr, V., Strobl, G., Manaut, F., Anders, E.-M., Sörgel, F., López-de-Briñas, E., Chu, D.T.W., Pernet, A.G., Mahr, G., Sanz, F. and Staib, H., Mol. Pharmacol., 43 (1993) 191.
5 Wakabayashi, K., Nagao, M., Esumi, H. and Sugimura, T., Cancer Res., 52 (Suppl.) (1992) 2092.
6 Shimada, T., Iwasaki, M., Martin, M.V. and Guengerich, F.P., Cancer Res., 49 (1989) 3218.
7 Yamashita, K., Umemoto, A., Grivas, S., Kato, S., Sato, S. and Sugimura, T., Nucleic Acids Res., 19 (1988) 111.
8 Poulos, T.L., Finzel, B.C. and Howard, A.J., Biochemistry, 25 (1986) 5314.
9 Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. and Deisenhofer, J., Science, 261 (1993) 731.
10 Hasemann, C.A., Ravichandran, K.G., Peterson, J.A. and Deisenhofer, J., J. Mol. Biol., 236 (1994) 1169.
11 Cupp-Vickery, J.R. and Poulos, T.L., Nat. Struct. Biol., 2 (1995) 144.
12 Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J., Structure, 3 (1993) 41.
13 Laughton, C.A., Neidle, S., Zvelebil, M.J.J.M. and Sternberg, M.J.E., Biochem. Biophys. Res. Commun., 171 (1990) 1160.
14 Morris, G.M. and Richards, W.G., Biochem. Soc. Trans., 19 (1991) 793.
15 Zvelebil, M.J.J.M., Wolf, C.R. and Sternberg, M.J.E., Protein Eng., 4 (1991) 271.
16 Lewis, D.F.V. and Moereels, H., J. Comput.-Aided Mol. Design, 6 (1992) 235.
17 Vijayakumar, S. and Salerno, J.C., Biochim. Biophys. Acta, 1160 (1992) 281.
18 Laughton, C.A., Zvelebil, M.J.J.M. and Neidle, S., J. Steroid Biochem. Mol. Biol., 44 (1993) 399.
19 Koymans, L.M.H., Vermeulen, N.P.E., Baarslag, A. and Donné-Op den Kelder, G.M., J. Comput.-Aided Mol. Design, 7 (1993) 281.
20 Boscott, P.E. and Grant, G.H., J. Mol. Graph., 12 (1994) 185.
21 Szklarz, G.D., Ornstein, R.L. and Halpert, J.R., J. Biomol. Struct. Dyn., 12 (1994) 61.
22 Narhi, L.O. and Fulco, A.J., J. Biol. Chem., 262 (1987) 6683.
23 Nelson, D.R. and Strobel, H.W., Biochemistry, 28 (1989) 656.
24 Mosimann, S., Meleshko, R. and James, M.N.G., Proteins Struct. Funct. Genet., 23 (1995) 301.
25 Chang, Y.-T. and Loew, G.H., Protein Eng., 9 (1996) 755.
26 Sanz, F., López-de-Briñas, E., Rodríguez, J. and Manaut, F., Quant. Struct.–Act. Relatsh., 13 (1994) 281.
27 Lozano, J.J., López-de-Briñas, E., Manaut, F. and Sanz, F., In Sanz, F., Giraldo, J. and Manaut, F. (Eds.) QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications, Prous Science, Barcelona, Spain, 1995, pp. 325–328.
28 Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J., Structure, 2 (1995) 41.
29 Thompson, J.D., Higgins, J.D. and Gibson, T.J., Nucleic Acids Res., 22 (1994) 4673.
30 PILEUP, Genetics Computer Group Inc., University Research Park, Madison, WI, U.S.A., 1995.
31 Rost, B., In Rawlings, C., Clark, D., Altman, R., Hunter, L., Lengauer, T. and Wodak, S. (Eds.) Proceedings of the Third International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park, CA, U.S.A., 1995, pp. 314–321.
32 Rost, B. and Sander, C., J. Mol. Biol., 232 (1993) 584.
33 Levin, J.M., Robson, B. and Garnier, J., FEBS Lett., 205 (1986) 303.
34 Gibrat, J.F., Garnier, J. and Robson, B., J. Mol. Biol., 198 (1987) 425.
35 Deleage, G. and Roux, B., Protein Eng., 1 (1987) 289.
36 Kneller, G., Cohe, F.E. and Langridge, R., J. Mol. Biol., 214 (1990) 171.
37 Geourjon, C. and Deleage, G., Protein Eng., 7 (1994) 157.
38 Salamov, A.A. and Solovyev, V.V., J. Mol. Biol., 247 (1995) 11.
39 Deleage, G., Clerc, F.F., Roux, B. and Gautheron, D.C., Comput. Appl. Biosci., 4 (1988) 351.
40 SOPM, http://www.ibcp.fr.
41 NnPredict, http://www.cmpharm.ucsf.edu.
42 PHDsec, http://www.embl-heidelberg.de.
43 NNSP, http://www.bcm.tcm.edu.

44 Kabsch, W. and Sander, C., Biopolymers, 22 (1983) 2577.
45 Vriend, G., J. Mol. Graph., 8 (1990) 52.
46 HOMOLOGY (v. 95.0) from Biosym/MSI, San Diego, CA, U.S.A., 1995.
47 Hobohm, U., Scharf, M., Schneider, R. and Sander, C., Protein Sci., 1 (1992) 409.
48 Hobohm, U. and Sander, C., Protein Sci., 3 (1994) 522.
49 DISCOVER (v. 2.9.5) from Biosym/MSI, San Diego, CA, U.S.A., 1995.
50 Dauber-Osguthorpe, P., Roberst, V.A. and Wolff, J., Proteins Struct. Funct. Genet., 4 (1988) 31.
51 INSIGHT II (v. 95.0) from Biosym/MSI, San Diego, CA, U.S.A., 1995.
52 Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 283.
53 Hooft, R.W.W., Vriend, G., Sander, C. and Abola, E.E., Nature, 381 (1996) 272.
54 GAUSSIAN 94, Gaussian Inc., Pittsburgh, PA, U.S.A.
55 Sanz, F., Manaut, F., José, J., Segura, J., Carbó, M. and De la Torre, R., J. Mol. Struct. (THEOCHEM), 170 (1988) 171.
56 Sanz, F., Manaut, F., Rodríguez, J., Lozoya, E. and López-de-Briñas, E., J. Comput.-Aided Mol. Design, 7 (1993) 337.
57 Morris, G.M., Goodsell, D., Huey, R. and Olson, A.J., J. Comput.-Aided Mol. Design, 10 (1996) 293.
58 Degtyarenko, K.N. and Archakov, A.I., FEBS Lett., 322 (1993) 1.
59 Valencia, A., Kjeldgaard, M., Pai, E.F. and Sander, C., Proc. Natl. Acad. Sci. USA, 88 (1991) 5443.
60 Gotoh, O., J. Biol. Chem., 267 (1992) 83.
61 Arnold, G.E. and Ornstein, R.L., Proteins Struct. Funct. Genet., 18 (1994) 19.
62 Guenot, J.M. and Kollman, P.A., Protein Sci., 1 (1992) 1185.
63 Guenot, J.M. and Kollman, P.A., J. Comput. Chem., 14 (1993) 295.
64 Vriend, G. and Sander, C., J. Appl. Crystallogr., 26 (1993) 47.
65 Furuya, H., Shimizu, T., Hirano, K., Hatano, M. and Fujii-Kuriyama, Y., Biochemistry, 28 (1989) 6848.
66 Shimizu, T., Sadeque, A.J.M., Sadeque, G.N., Hatano, M. and Fujii-Kuriyama, Y., Biochemistry, 30 (1991) 1490.
67 Tuck, F.S., Hiroya, K., Shimizu, T., Hatano, M. and Ortiz de Montellano, P.R., Biochemistry, 32 (1993) 2548.
68 Mayuzumi, H., Shimizu, T., Sambongi, C., Hiroya, K. and Hatano, M., Arch. Biochem. Biophys., 310 (1994) 367.
69 Edwards, R.J., Singleton, A.M., Murray, B.P., Murray, S., Boobis, A.R. and Davies, D.S., Biochem. J., 278 (1991) 749.
70 Edwards, R.J., Sesardic, D., Murray, B.P., Singleton, A.M., Davies, D.S. and Boobis, A.R., Biochem. Pharmacol., 43 (1992) 1737.
71 Shen, S. and Strobel, H.W., Arch. Biochem. Biophys., 294 (1992) 83.
72 Cvrk, T., Hodek, P. and Strobel, H.W., Arch. Biochem. Biophys., 330 (1996) 142.
73 Poulos, T.L., Finzel, B.C. and Howard, A.J., J. Mol. Biol., 195 (1987) 687.
74 Poulos, T.L. and Howard, A.J., Biochemistry, 26 (1987) 8165.
75 Raag, R., Li, H., Jones, B.C. and Poulos, T.L., Biochemistry, 32 (1993) 4571.
76 Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C. and Kraut, J., J. Biol. Chem., 260 (1985) 16122.
77 Fuhr, U., Dohmer, J., Battula, N., Wölfel, C., Kudla, C., Keita, Y. and Staib, H., Biochem. Pharmacol., 43 (1992) 225.