J. Enzyme Inhibition, 1999, Vol. 14, pp. 175–192
Reprints available directly from the publisher
Photocopying permitted by license only

MOLECULAR MODELLING OF LANOSTEROL 14α-DEMETHYLASE (CYP51) FROM *SACCHAROMYCES CEREVISIAE* VIA HOMOLOGY WITH CYP102, A UNIQUE BACTERIAL CYTOCHROME P450 ISOFORM: QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIPS (QSARs) WITHIN TWO RELATED SERIES OF ANTIFUNGAL AZOLE DERIVATIVES

DAVID F.V. LEWIS[a,*], ALAN WISEMAN[a] and MIKE H. TARBIT[b]

[a]*School of Biological Sciences, University of Surrey, Guildford, Surrey, GU2 5XH, UK;* [b]*Glaxo Wellcome Research and Development Limited, Park Road, Ware, Hertfordshire, SG12 0DP, UK*

(Received 20 August 1998)

The construction of a three-dimensional molecular model of the fungal form of cytochrome P450 (CYP51) from *Saccharomyces cerevisiae*, based on homology with the haemoprotein domain of CYP102 from *Bacillus megaterium* (a unique bacterial P450 of known crystal structure) is described. It is found that the endogenous substrate, lanosterol, can readily occupy the putative active site of the CYP51 model such that the known mono-oxygenation reaction, leading to C_{14}-demethylation of lanosterol, is the preferred route of metabolism for this particular substrate. Key amino acid contacts within the CYP51 active site appear to orientate lanosterol for oxidative attack at the C_{14}-methyl group, and the position of the substrate relative to the haem moiety is consistent with the phenyl–iron complexation studies reported by Tuck *et al.* [*J. Biol. Chem.*, **267**, 13175–13179 (1992)]. Typical azole inhibitors, such as ketoconazole, are able to fit the putative active site of CYP51 by a combination of haem ligation, hydrogen bonding, π–π stacking and hydrophobic interactions within the enzyme's haem environment. The mode of action of azole antifungals, as described by the modelling studies, is supported by quantitative structure–activity relationship (QSAR) analyses on two groups of structurally related fungal inhibitors. Moreover, the results of molecular electrostatic isopotential (EIP)

* Corresponding author. Tel.: 01483 300800. Fax: 01483 300803.
E-mail: d.lewis@surrey.ac.uk.

energy calculations are compatible with the proposed mode of binding between azole antifungal agents and the putative active site of CYP51, although membrane interactions may also have a role in the antifungal activity of azole derivatives.

Keywords: CYP51; Lanosterol; 14α-demethylase; Inhibitors; QSAR; CYP102; Antifungals; Azole derivatives

INTRODUCTION

The cytochromes P450 (CYP) constitute a superfamily of haem-thiolate enzymes which catalyse a large number and variety of mixed function oxidations (and some reductions) for a vast array of structurally diverse chemicals, both exogenous and endogenous.[1–3] The yeast P450 isoform (CYP51), which mediates the formation of the principal fungal steroid ergosterol by a pathway including C_{14}-demethylation of lanosterol, has been the target for antifungal drug design over many years,[4–14] and it is thought that ergosterol may act in the regulation of yeast cell growth under anaerobic conditions.[9] In addition to the fungal forms, CYP51 isozymes are also present in mammalia and other animal species, where lanosterol 14α-demethylase is involved in *de novo* cholesterol biosynthesis.[9] According to the most recent update on P450 nomenclature,[1] CYP51 proteins have been sequenced from human, rat, wheat and a total of nine different species of fungi, including several yeast species such as: *Saccaromyces cerevisiae*, *Candida albicans* and *C. tropicalis*. Figure 1 presents an alignment between the amino acid sequences of these three yeast forms of CYP51 produced using the GCG package (Genetics Computer Group, Madison, Wisconsin) where there is clear evidence of significant homology (~89%) between the three isoforms in agreement with earlier work.[15] In fact, *C. tropicalis* has been shown to possess at least seven additional P450s which have been assigned to a separate multigene family, CYP52.[16] Unlike CYP51, however, the CYP52 isoforms are associated with the ω-terminal oxidations of alkanes,[10] although the two families are likely to be evolutionarily related despite possessing different substrate specificities.

CYP51 from *S. cerevisiae* exhibits typical P450 spectroscopic properties[10] and, upon binding lanosterol, produces a Type I spectral change which can be used to determine the enzyme–substrate dissociation constant (K_D) of 6.0 μM; and this is in good agreement with the apparent Michaelis constant (K_m) of 6.25 μM. The binding affinity for lanosterol to CYP51 can, therefore, be estimated as lying within the range −7.096 to −7.120 kcal·mole^{-1} at 25°C,

```
cyp102    .......... .......... .......... .......... .........T
cyp51cal  ....MAIVET VID....GIN YFLSLSVTQQ ISILLGVPFV YNLVWQYLYS
cyp51ctr  ....MAIVDT AID....GIN YFLSLSLTQQ ITILVVFPFI YNIAWQLLYS
cyp51sce  MSATKSIVGE ALEYVNIGLS HFLALPLAQR ISLIIIIPFI YNIVWQLLYS
                    10        20        30        40
cyp102    I.KE.MPQP. KTFGELKNLP LLNTDKPVQALMKIADELGEI FKFEAPGRVT
cyp51cal  LRKDRAPLVF YWIPWFGSAA SY.GQQPYEFFESCRQKYGDV FSFMLLGKIM
cyp51ctr  LRKDRVPMVF YWIPWFGSAA SY.GMQPYEFFEKCRLKYGDV FSFMLLGKVM
cyp51sce  LRKDRPPLVF YWIPWVGSAV VY.GMKPYEFFEECQKKYGDI FSFVLLGRVM
          50        60        70        80                  90
cyp102    RYLSSQRLIKEACDESRF.DK NL.SQALKFV RDFAGDGLFT S..WTHEKNW
cyp51cal  TVY.LGPKGHEFVFNAKLSDV SAEDAYKHLT TPVFGKGVIY DCPNSRLMEQ
cyp51ctr  TVY.LGPKGHEFIYNAKLSDV SAEEAYTHLT TPVFGKGVIY DCPNSRLMEQ
cyp51sce  TVY.LGPKGHEFVFNAKLADV SAEAAYAHLT TPVFGKGVIY DCPNSRLMEQ
             100       110       120       130       140
cyp102    KKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADEHI EV.PEDMTRL
cyp51cal  KKFAKFAL.T.TD.SFKRYVPKIREEILNYFVTDESFKLKEKT HGVANVMKTQ
cyp51ctr  KKFAKFAL.T.TD.SFKTYVPKIREEVLNYFVNDVSFKTKERD HGVASVMKTQ
cyp51sce  KKFVKGAL.T.KE.AFKSYVPLIAEEVYKYFRDSKNFRLNERT TGTIDVMVTQ
           150       160       170       180       190       200
cyp102    TLDTIGLCGFNYRFNSFYRDQPHPFITSMVRALDEAMNKLQRA.NPDDP.AYDE
cyp51cal  PEITIFTASRSL.F....GDEMRRIFDRSFAQLYSDLDKGFTPINFVFPNLPLP
cyp51ctr  PEITIFTASRCL.F....GDEMRKSFDRSFAQLYADLDKGFTPINFVFPNLPLP
cyp51sce  PEMTIFTASRSL.L....GKEMRAKLDTDFAYLYSDLDKGFTPINFVFPNLPLE
                  210       220       230       240
cyp102    NKRQFQEDIKVMNDLVDKIIA DRKASGE.Q. SDDLLTHMLN GKDPETGEPL
cyp51cal  HYWRRDAAQKKISATYMKEIK LRRERGDIDP NRDLIDSLLI HSTYKDGVKM
cyp51ctr  HYWRRDAAQRKISAHYMKEIK RRRESGDIDP KRDLIDSLLV NSTYKDGVKM
cyp51sce  HYRKRDHAQKAISGTYMSLIK ERRKNNDIQ. DRDLIDSLMK NSTYKDGVKM
          250       260       270       280       290
cyp102    DDENIRYQII TFLIAGHETT SGLLSFALYF LVKNPHVLQK AAEEAARVLV
cyp51cal  TDQEIANLLI GILMGGQHTS ASTSAWFLLH LGEKPHLQDV IYQEVVELLK
cyp51ctr  TDQEIANLLI GVLMGGQHTS ASTSAWFLLH LAEQPQLQDD LYEELTNLLK
cyp51sce  TDQEIANLLI GVLMGGQHTS AATSAWILLH LAERPDVQQE LYEEQMRVL.
          300        310       320       330       340
cyp102    D.P....VPS YKQVKQLKYV GMVLNEALRL W.PTAPAFSL YAKEDTVLGG
cyp51cal  EKGGDLNDLT YEDLQKLPSV NNTIKETLRM HMPLHSIFRK VTNPLRIPET
cyp51ctr  EKGGDLNDLT YEDLQKLPLV NNTIKETLRM HMPLHSIFRK VMNPLRVPNT
cyp51sce  D.GG.KKELT YDLLQEMPLL NQTIKETLRM HHPLHSLFRK VMKDMHVPNT
        350       360       370       380
cyp102    FYPLEKGDEL MVLIPQLHRD KTIWGDDVEEFRPERFE.... NPSAI.....
cyp51cal  NYIVPKGHYV LVSPGYAHTS ERYF.DNPEDFDPTRWDTAAA KANSVSFNSS
cyp51ctr  KYVIPKGHYV LVSAGYAHTS DRWF.EHPEHFNPRRWESDDT KASAVSFNSE
cyp51sce  SYVIPAGYHV LVSPGYTHLR DEYF.PNAHQFNIHRWN.... KDSASSYSVG
                  390       400       410       420
cyp102    .......... ...PQHAFKP FGNGQRACIG QQFALHEATL VLGMMLKHFD
cyp51cal  DEVDYGFGKV SKGVSSPYLP FGGGRHRCIG EQFAYVQLGT ILTTFVYNLR
cyp51ctr  DTVDYGFGKI SKGVSSPYLP FGGGRHRCIG EQFAYVQLGT ILTTYIYNFK
cyp51sce  EEVDYGFGAI SKGVSSPYLP FGGGRHRCIG EHFAYCQLGV LMSIFIRTLK
                  430       440       450
cyp102    FE..DHTNY. ELDIKETLTLKPEGFVVKAKS KKIPLGGI
cyp51cal  WTI.DGYKVP DPDY.SSMVVLPTEPAEIIWE KRETCMF.
cyp51ctr  WRL.NGDKVP DVDY.QSMVTLPLEPAEIVWE KRDTCMV.
cyp51sce  WHYPEGKTVP PPDF.TSMVTLPTGPAKIIWE KRNPEQKI
```

FIGURE 1 An alignment between CYP102 haemoprotein domain sequence and those of three CYP51 proteins from *S. cerevisiae*, *C. albicans* and *C. tropicalis*, respectively using information from Ref. [1].

using the expression:

$$\Delta G = RT \ln K$$

where R is the gas constant, T is the absolute temperature, and K may be either K_D or K_{app} in this instance.

Based on sequence comparisons, the CYP51 proteins from *S. cerevisiae* and *C. tropicalis* exhibit a 66.5% identity and an overall homology of 89.6% when conservatively replaced amino acids are taken into account.[17] It is possible to identify four highly homologous regions in the CYP51 sequences which appear to correspond with the putative active site of the enzyme, including the haem-binding domain. In fact there is a close correspondence between conserved regions in the CYP51 alignment (Figure 1) and the substrate recognition sites (SRSs) identified by Gotoh[18] in a multiple sequence alignment of several CYP2 family proteins. Site-directed mutagenesis of glycine-310 to aspartate (i.e., the change G310D) in CYP51 from *S. cerevisiae* leads to a catalytically inactive form of the enzyme.[19] Sequence alignment with $P450_{cam}$ (CYP101), which was the only P450 available at the time to have been characterized crystallographically, indicated that position 310 in CYP51 lies within the putative I helix which is orientated distal to the haem moiety.[19] Moreover, molecular modelling of part of this distal helix suggested that a rotation of this segment of secondary structure could occur in the mutant form of the enzyme, such that a downstream histidine residue (His-317) would readily ligate the haem iron,[19] thus providing an explanation for the inactivity of the G310D mutant.

It is possible that histidine-317 may play a key role in the action of azole antifungal agents, as it appears that the more potent CYP51 inhibitors contain either an imidazole or triazole 5-membered ring[4,11–13] which is analogous to the side-chain of a histidine residue. Crystallographic studies on $P450_{cam}$ containing a bound imidazole antifungal compound[20] indicate that conformational changes in the enzyme take place upon inhibition, although these appear to affect active site amino acid side-chains rather than the protein backbone itself. However, it is clear from this and other reported studies on imidazole inhibitors (reviewed in Ref. [20]) that the imidazole ring sp^2 nitrogen ligates the haem of P450 at a distance of about 2 Å from the iron atom. Consequently, His-317 in CYP51 could be involved in the access of azole inhibitors to the active site, where its normal endogenous function is likely to be associated with the oxygenation mechanism, which requires that the substrate remains enzyme-bound during the entire reaction sequence, involving three successive mono-oxygenations without the release of intermediates.[9,10]

Morris and Richards[21] reported the results of active site modelling for CYP51 based on homology with CYP101, although sequence alignments using the CYP101 template indicate several regions of poor homology and large gaps due to stretches of unmatched peptide between the two protein sequence[22] which hamper the production of a satisfactory model. However, since the publication of the three-dimensional structure of the CYP102 haemoprotein domain[23] it has been apparent that this represents an improved template for the construction of microsomal P450s[2] since the above problems are largely eliminated. Consequently, homology modelling of CYP51 from the unique bacterial P450 CYP102 crystal structure is likely to advance our understanding of the yeast lanosterol C-14 demethylase system.

METHODS

Figure 1 shows an alignment between three CYP51 sequences from *C. tropicalis*, *C. albicans* and *S. cerevisiae* with that of the CYP102 haemoprotein domain. This alignment was produced using the GCG package (Genetics Computer Group, Maidson, Wisconsin) from the primary structures of the relevant proteins obtained from the most recent update of P450 accession numbers[1] including that of *S. cerevisiae*.[24] Apart from an N-terminal membrane-binding peptide of about 40 residues, there is a close match between CYP102 and the three CYP51 sequences with only a small number of short gaps, sometimes involving only one or two amino acid residues, with the exception of an 18-residue stretch upstream of the P450 signature motif containing the invariant cysteine which is the proximal haem ligand. As this octadecapeptide does not feature in the CYP52 sequences[16] it would appear that it is a characteristic of CYP51, and may relate to the mode of membrane binding and/or topology of the enzyme's orientation within the phospholipid bilayer, which could have a bearing on catalytic activity. Nevertheless, there is clear homology between CYP102 and CYP51 which supports the use of the former as a structural template and, moreover, we have shown[25] previously that the CYP102 crystal structure is satisfactory for homology modelling of other steroidogenic P450s including aromatase (CYP19) which exhibits significant sequence similarity[1] with CYP51, together with an analogous mechanism of substrate oxygenation.[9,10] Furthermore, CYP52A1, which catalyses the ω-hydroxylation of lauric acid, appears to cluster with CYP102 in a dendrogram of eukaryotic P450s[1] where the substrate specificity of CYP102 for long chain fatty acids may be related to the similarities in primary sequences of the two proteins.

A three-dimensional model of CYP51 was constructed from the CYP102 structure, based on the alignment shown in Figure 1, using the Biopolymer module of the Sybyl package version 6.4 (Tripos Associates, St. Louis, Missouri) for residue replacement, deletion and insertion as required by the sequence alignment. Molecular modelling via Sybyl version 6.4 was carried out on a Silicon Graphics Indigo^2 IMPACT 10000 workstation. All residue insertions were incorporated via loop-searching of the Brookhaven Protein Databank irrespective of the length of peptide required, and the raw structure of CYP51 was then energy minimized using the Tripos force field to produce a low energy minimum geometry of $-1508.9\,\text{kcal mole}^{-1}$ in the absence of substrate. The final CYP51 model was then probed using the endogenous substrate, lanosterol, and azole inhibitors, such as ketoconazole and itraconazole, using either crystal structures of the compounds themselves[26–28] or pre-minimized structural fragments obtained from the Sybyl Fragment Library. Crystallographic data for itraconazole was kindly supplied by Henri Moereels (Janssen Pharmaceutica, Beerse, Belgium) from the original crystal coordinates produced by Professor C.J. De Ranter. In both ketoconazole and itraconazole, the active enantiomer was used, namely, the *RS* form. Interactive docking of substrate and inhibitors within the putative active site of CYP51 was carried out using the DOCK routine in Sybyl 6.4, and involved a 3 Å distance constraint between the C-14 methyl and haem iron in the case of lanosterol whereas a 2 Å distance constraint between ligating nitrogen and iron was employed for the azole inhibitors, in keeping with crystallographic data on P450 complexes. Quantitative structure–activity relationships (QSARs) for two series of azole antifungals (see for Figure 2 structures) were generated using the reported biological data on CYP51 inhibition[11,29] and from CNDO/2-calculated electronic structures[30] via the COSMIC modelling framework[31] as implemented on a MicroVAX II mini computer operating under VMS. The COSMIC software package was a generous gift from Dr J.G. Vinter.

RESULTS AND DISCUSSION

1. CYP51 Homology Model

Figure 3 shows a view of the energy minimized structure of CYP51 where the lanosterol molecule can be visualized as occupying an essentially hydrophobic pocket in the vicinity of the putative active site bounded by the haem moiety, part of the distal I helix, the F and B′ helices, and two β-sheet regions corresponding to alignment positions 328 and 437

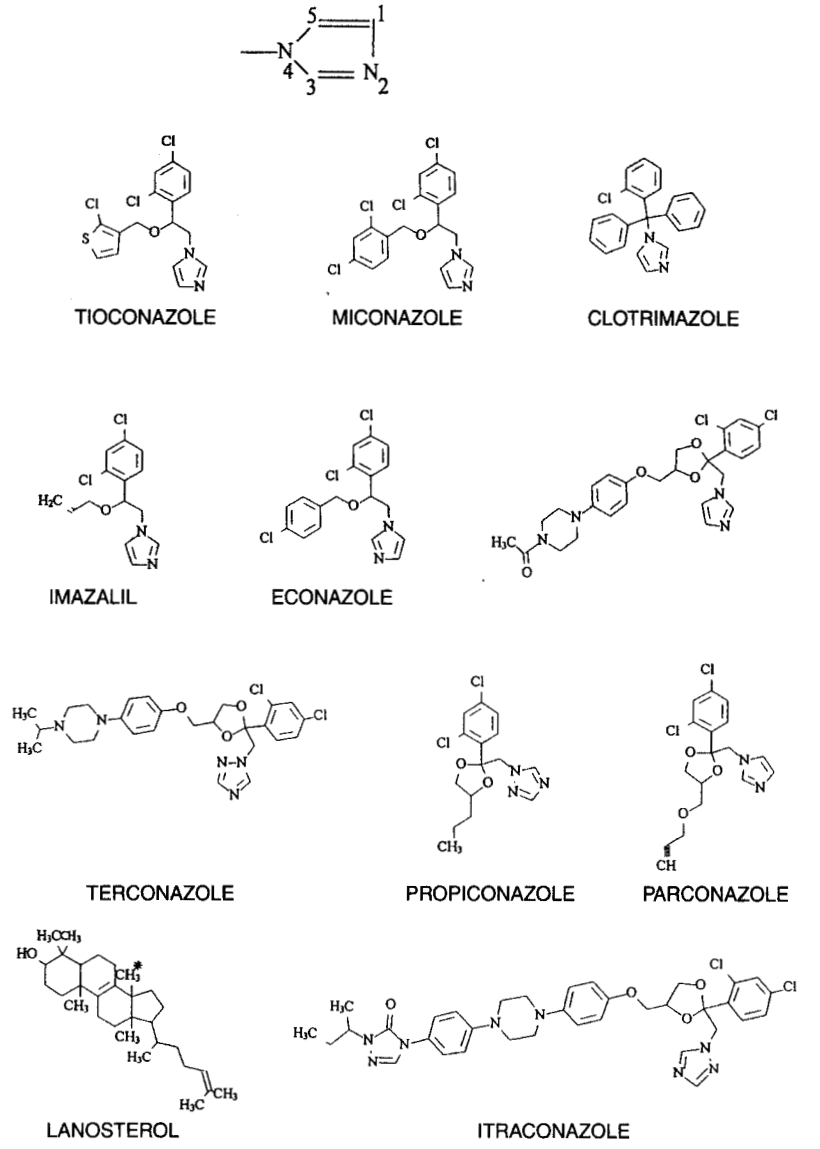

FIGURE 2 Structures of all compounds studied in this work together with the numbering system of the azole ring.

(CYP102 nomenclature). Tuck *et al.*[32] have shown that lanosterol is likely to become orientated relative to the haem of CYP5l such that its 3-hydroxyl group can form a hydrogen bond with an active site residue and, in the current model, it appears that serine-382 (at alignment position 329)

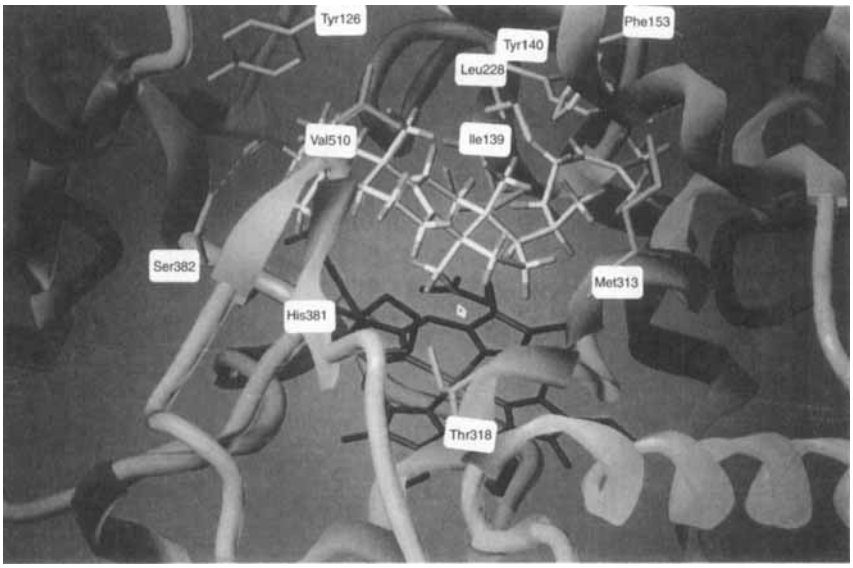

FIGURE 3 A view of the putative active site of CYP51 showing the disposition of the endogenous substrate, lanosterol, orientated for 14α-demethylation. See Color Plate I.

could fulfil the role of a hydrogen bond-forming side-chain which is able to donate a hydrogen bond to lanosterol in CYP51 (Figure 2). Moreover, a second hydrogen bond donor/acceptor (as required by the currently accepted mechanism which is described in Ref. [9]) may be afforded by histidine-381 (at position 328 in the alignment), although histidine-317 (corresponding to alignment position 267) could also fulfil this requirement. The likelihood of a hydrogen bond-forming amino acid side-chain at about this region of the CYP51 active is suggested by the finding that the enzyme exhibits higher affinity for both C-14 hydroxymethyl lansterol and the corresponding aldehyde than for lanosterol itself.[32]

In addition to these hydrogen bond contacts, it is clear from Figure 3 that there are several hydrophobic interaction points between substrate and enzyme, including some evidence of $\pi-\pi$ stacking interactions. In particular, the planar dimethyl alkene function terminating the D-ring side-chain grouping of lanosterol is able to π stack with the aromatic ring of phenylalanine-153 (position 99 in the alignment). Also, a number of aliphatic amino acid side-chains are implicated in the binding of lanosterol via complementary contacts with substrate methyl groups, in addition to hydrocarbon segments of the steroid nucleus; these hydrophobic amino acid residues include:

Val-138, Ile-139, Leu-228, Leu-257, Met-313, Met-509 and Val-510. Furthermore, the aromatic ring of Tyr-140 could also form a $\pi-\pi$ stacking interaction with a planar unsaturated region of lanosterol (see Figure 3). Consequently, the putative active site of CYP51 represents a closely complementary surface for the endogenous substrate, lanosterol, to bind with high affinity and in such a manner that the molecule is orientated by the CYP51 protein for C_{14}-demethylation and with the required stereospecificity.

When one considers a typical CYP51 inhibitor, such as ketoconazole, for example, the situation is fairly analogous to that evidenced with lanosterol binding to CYP51 described above. Figure 4 shows a view of ketoconazole interacting within the putative active site of CYP51 such that inhibition of the enzyme would be brought about via imidazole ligation of the haem iron. To some extent, the same type of amino acid residues are involved in the binding of ketoconazole to the active site of CYP51 as described previously for lanosterol. In particular, the side-chains of Phe-134 and Tyr-126 are able to form π-stacking interactions with the two aromatic rings on the inhibitor, as shown in Figure 4. Furthermore, histidine-381 donates a hydrogen bond to the phenoxy oxygen atom on the ketoconazole side-chain. These, together

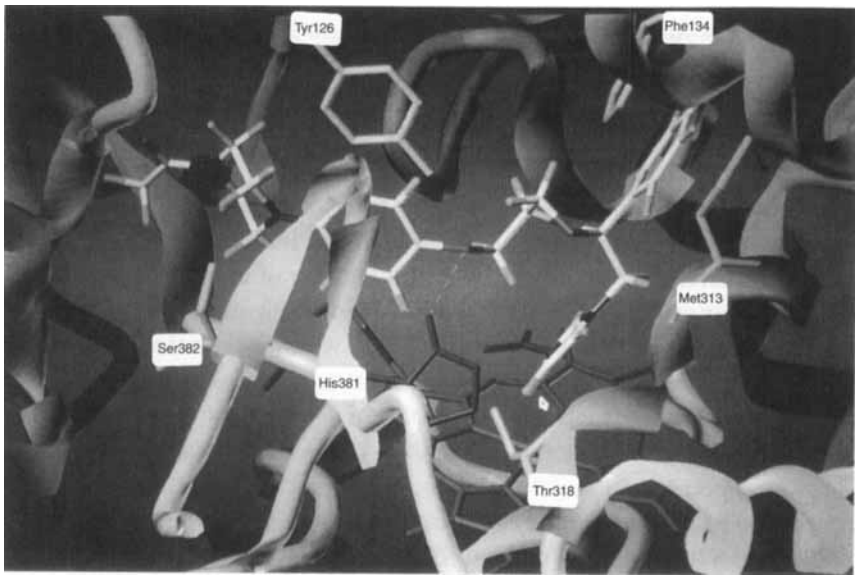

FIGURE 4 The inhibitor, ketoconazole, is shown docked within the putative active site of CYP51 where complementary interactions orientate the molecule for haem ligation. See Color Plate II.

with a number of complementary hydrophobic interactions, orientate the ketoconazole molecule relative to the haem moiety such that the imidazole group lies directly above the central iron atom. In such a position, as shown in Figure 4, the sp^2 nitrogen on imidazole is relatively close to the haem iron at a distance of 3.458 Å. Since many azole antifungal agents possess a 2,4-dichlorophenyl group at an analogous location relative to their imidazole (or triazole) ring as that of ketoconazole, it is likely that these compounds will also occupy the CYP51 active site in a similar manner. Furthermore, the variations in their antifungal activity can be rationalized by the extent of favourable contacts made by the remainder of the molecule, which also extends into the binding site channel like ketoconazole.

2. Quantitative Structure–Activity Relationships (QSARs) in Antifungals

Several comparative studies have been reported for azole antifungal agents[11,29,33] including some investigations in mammalian systems.[34–38] It is generally accepted that the evidence from Type II binding spectra indicates azole ligation of the haem moiety in P450 as the mechanism of inhibition[39,40] in contrast to the Type I binding exhibited by the substrate, lanosterol. For the generation of QSARs between antifungal activity and structural descriptors on the compounds concerned, we have used the biological data produced by Pye and Marriott[29] and Vanden Bossche *et al.*[11] as these comprise inhibitory information on a sufficient number of structurally related azoles to permit QSAR analysis.

Table I presents the relevant data set for a series of azoles exhibiting antifungal activity towards *C. albicans*[29] where a number of structural descriptors from a total of 40 independent variables gave rise to statistically significant correlations with both Minimum Inhibitory Concentration (MIC) for arresting fungal growth and inhibition of ergosterol production (IC_{50}). As far as the former activity is concerned, the only single descriptor correlation ($R = 0.77$) was that involving Q_3, which is the net atomic charge on atom 3 of the azole nucleus and adjacent to the haem-ligating nitrogen (see Figure 2 for nomenclature). The value of Q_3 does not vary dramatically within the series, apart from in the triazole compound, butaconazole, which is the least potent antifungal in the series. However, elimination of the data for tioconazole from the analysis results in a substantially improved correlation ($R = 0.97$) between Q_3 and log MIC, as shown in Table I (Eq. (1)). Plotting out the relationship reveals, however, that it consists of a cluster of five points such that the strength of the regression is primarily determined by the remaining point, corresponding to the triazole butaconazole. Consequently,

TABLE I Electronic structural and physicochemical data for 7 azole antifungals v. *C. albicans*

	Compound	Q_3	Q_2H	$\alpha \cdot \Delta E$	$\log D_{7.4}$	log MIC	pIC_{50}
1	Ketoconazole	0.127	0.0001	678.97	3.30	1.6628	1.6900
2	Miconazole	0.121	0.0610	480.30	6.26	1.4771	2.3010
3	Clotrimazole	0.127	0.0416	474.51	5.68	1.5185	2.0792
4	Econazole	0.122	0.0525	471.17	5.48	1.5185	1.9542
5	Parconazole	0.126	0.0602	474.44	3.19	1.5315	1.7782
6	Butaconazole	0.184	0.0920	475.50	3.34	2.0792	1.7782
7	Tioconazole	0.120	0.0299	437.67	5.26	0.9243	1.6990
	Regression equation	*n*	*s*	*R*	*F*		
(1)	$\log MIC = 9.12\,Q_3 + 0.40$ (± 1.06)	6	0.058	0.97	73.7		
(2)	$\log MIC = 13.98\,Q_2H + 0.045\alpha\Delta E - 1.40$ (± 2.51) (± 0.001)	7	0.134	0.95	17.3		
(3)	$pIC_{50} = 0.15 \log D_{7.4} + 1.24$ (± 0.03)	6	0.090	0.94	28.2		
(4)	$pIC_{50} = 3.14\,Q_1H - 5648.95\,Q_2L + 1.76$ (± 0.57) (± 1034.74)	7	0.088	0.95	17.8		

Q_3 = net atomic charge on azole ring carbon atom 3;
Q_1H, Q_2H = HOMO electron density on azole ring atoms 1 and 2, respectively (ΔE) where HOMO is the highest occupied MO;
$\alpha \cdot \Delta E$ = product of molar polarizability (α) and excitation energy (ΔE) where $\Delta E = E_{LUMO} - E_{HOMO}$;
$\log D_{7.4}$ = common logarithm of the distribution coefficient at pH 7.4;
Q_2L = LUMO electron population on azole ring nitrogen atom 3 where LUMO is the lowest unoccupied MO;
log MIC = log minimum inhibitory concentration (μM) against *C. albicans* C66 growth;[29]
pIC_{50} = −log concentration (nM) required for 50% inhibition of ergosterol synthesis in *C. albicans*.[29]
Correlation Eqs. (1) and (3) resulted after the removal of the data for tioconazole (compound **7**) from the regression analysis. This compound was excluded on the basis of its thiophene ring system, which is not present in the other inhibitors, and may mask the activity due solely to the azole grouping.

this apparently important relationship has to be regarded with considerable suspicion.

Fortunately, the regression analysis also generated another statistically significant correlation which involved a combination of two independent variables, Q_2H and $\alpha \cdot \Delta E$ that are, respectively, the HOMO electron population density on Q_2 and the product of polarizability (α) and ΔE, which is the difference between the energies of the lowest unoccupied and highest occupied molecular orbitals, respectively. As the Q_2H parameter represents a measure of the electron-donating capacity of the azole nitrogen, which is probably involved in haem ligation of the yeast P450, the appearance of this structural descriptor in a QSAR governing inhibition of fungal growth is to be expected. However, the rationale for the combined parameter $\alpha \cdot \Delta E$ requires a little more explanation. Essentially, the product of molar polarizability (α) and the excitation energy (ΔE) represents an approximation for the dispersion forces (sometimes referred to as London forces or van der Waals forces) of attraction experienced by one molecule interacting with another (reviewed in Ref. [41]). This contribution to the overall binding energy is highly distance-dependent, being inversely proportional to the sixth power of the distance between interacting species, and relates to the number and distribution of electrons over the whole molecule.

This so-called polarization term also represents a component of the solvation energy and, consequently, the partitioning energy between two solvents experienced by a solute molecule,[42] such that it can be regarded as one of the three main factors contributing to the partition coefficient of a compound between, for example, octanol and water. Consequently, the lipophilicity parameter ($\log P_{oct}$), for a particular chemical, contains a polarizability component which can be the major factor of its overall lipophilicity.[43,44] One might anticipate, therefore that polarizability will correlate with $\log P_{oct}$ but it is found that this only occurs for neutral non-polar molecules.[43] Returning to the QSAR expression Table I, Eq. (2), it would appear that the antifungal potency of the seven azoles is governed by the relative ability of the azole nitrogen to ligate the haem group of the yeast P450 isoform partially involved in fungal growth, together with the extent of binding between the inhibitor molecule and the relevant P450 protein (or membrane phospholipid) as determined by the polarization component of compound lipophilicity.

The possibility of compound lipophilicity being relevant to the antifungal activity of the azoles is reinforced by the finding that inhibition of CYP51 for six compounds is proportional to $\log D_{7.4}$, the ionization-corrected $\log P$ value, as shown in Table I, Eq. (3). However, as in the case of the QSAR

relating to MIC Table I, Eq. (1) the data for the thiophene compound, tioconazole, have been excluded from the analysis although the correlation is still acceptable ($R = 0.76$) for all seven antifungals. The log distribution coefficient, $\log D_{7.4}$, represents the compound lipophilicity (i.e. $\log P_{oct}$) corrected for ionization at pH 7.4[45] and, consequently, contains a component relating to the chemical's basicity due to the presence of the azole ring, together with being a measure of the overall hydrophobic character of the molecule. The good correlation ($R = 0.94$) with this parameter which, unlike Q_3, exhibits a satisfactory spread of values within the series of 7 compounds, provides a degree of confidence in the significance of this result which, moreover, can be readily rationalized in terms of the likely mode of binding between the azole antifungals and the CYP51 binding site where hydrophobic interactions appear to feature prominently, in addition to the haem binding ability of the azole ring.

However, the only significant two-parameter correlation obtained for all 7 compounds in the data set involves a combination between two frontier orbital population densities on the azole ring, as shown by Eq. (4), Table I. This expression bears some resemblance to Eq. (2), in that a HOMO electron density is involved and the azole nitrogen likely to be involved in haem iron ligation also features. However, in this case, it is the LUMO density on the nitrogen at position 2 (Q_2L) which is paired with the HOMO electron population on the adjacent carbon (Q_1H) to produce a satisfactory correlation ($R = 0.95$) with CYP51 inhibition. Both of these expressions, which are shown as Eqs. (3) and (4) in Table I, exhibit close agreement between calculated and experimental values for pIC_{50}. Here econazole represents the only slight outlier in each case in that, with Eq. (3) it gives an overestimate whereas Eq. (4) produces an underestimate of econazole's activity. Nevertheless, it is important to find that structural descriptors for this series provide fairly satisfactory correlations with antifungal potency, as the latter depends also on other interactions, including those involving yeast cell membranes both pericytoplasmic and organelle-located.

For the second series of seven azoles, however, some similarities and differences are apparent in the QSARs produced against *S. cerevisiae* inhibition, as presented in Table II, which also lists the relevant parameters involved. Although the biological data stem from an independent study to the one described previously, two compounds (ketoconazole and miconazole) are common to both investigations, although these antifungals appear to be somewhat less potent than, for example, itraconazole.[4] Of the two correlations listed in Table II, Eq. (1) appears to be slightly more significant (Figure 5 compares experimental and calculated values) when one compares

TABLE II Electronic and molecular structural data for 7 azole antifungals v. *S. cerevisiae*

	Compound	Q_4	E_{min}	a/d^2	E_{LUMO}	pIC_{50}	$pIC_{50}^{calc.}$ (1)	$pIC_{50}^{calc.}$ (2)
1	Itraconazole	−0.018	−12.8	2.3	1.8939	1.6576	1.6367	1.5976
2	Terconazole	−0.025	−11.8	2.2	2.0572	1.4949	1.3930	1.4168
3	Parconazole	−0.015	−18.5	2.3	2.0572	1.4685	1.4511	1.5235
4	Propiconazole	−0.019	−17.0	2.1	1.9592	1.3979	1.4461	1.4469
5	Ketoconazole	−0.018	−26.4	2.0	1.9783	1.2596	1.3662	1.2065
6	Miconazole	−0.028	−15.3	1.3	1.8994	1.0969	1.0486	1.2263
7	Imazalil	−0.033	−17.5	1.8	2.1198	1.0555	1.0891	1.0134
	Regression equation	*n*	*s*	*R*	*F*			
(1)	$pIC_{50} = 0.58a/d^2 - 1.14\ E_{LUMO} + 2.45$ (± 0.10) (± 0.41)	7	0.084	0.95	18.6			
(2)	$pIC_{50} = 29.94Q_4 + 0.03E_{min} + 2.50$ (± 0.10) (± 0.41)	7	0.095	0.94	14.0			

pIC_{50} = −log concentration required for 50% inhibition of *S. cerevisiae* growth;[11]
Q_4 = net atomic charge on nitrogen atom 4 in the azole ring;
E_{min} = electrostatic potential energy minimum (kcal.mole^{-1});
E_{LUMO} = energy of the lowest unoccupied molecular orbital (eV.);
a/d^2 = ratio of molecular area to the square of molecular depth.

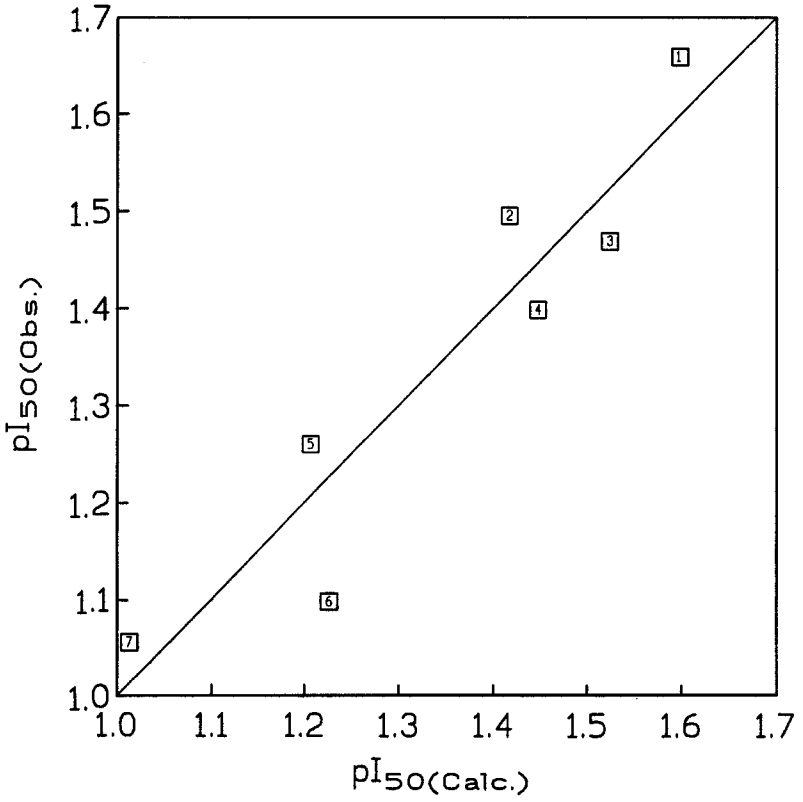

FIGURE 5 Comparison between observed and calculated pIC_{50} values for 7 compounds based on the data presented in Table II, Eq. (1).

the statistics, although comparisons between the experimental pIC_{50} values and those calculated from the two QSAR expressions show that Eq. (2) (Table II) gives a better fit for terconazole, propiconazole and ketoconazole. The two equations in (Table II) contain quite different pairs of structural descriptors: Eq. (1) represents a combination of molecular planarity (area/depth^2) and the LUMO energy value (E_{LUMO}), whereas Eq. (2) involves the net atomic charge (Q_4) on the substituted nitrogen of the azole ring and the energy of the electrostatic potential minimum (E_{min}) which lies in the vicinity of the azole nitrogen which ligates the haem iron in CYP51.

However, both equations appear to suggest that there may be interactions between the azole ring and the haem of the fungal P450, which could be synergistic in nature in that there is a possbility of both electron donation from the azole ring to the iron and electron acceptance by the azole system

brought about by back-donation from the haem moiety. Consequently, the QSAR analyses performed on the two series of antifungals indicate that there are strong reinforcing interactions between the azole ring and the CYP51 haem moiety which help to explain the potency differences in the various compounds. In conclusion, therefore, we feel that the results of QSAR analysis are consistent with the three-dimensional modelling studies of CYP51, because active site interactions help to explain some of the variations in antifungal activity shown by the compounds investigated in yeasts. It is hoped that these parallel studies will assist in the rationalization and prediction of antifungal potency for other structurally related chemicals in both clinical and industrial applications. Future utilization of related designer chemicals may also prove to be valuable as antifungals in both soil and aquatic environments.

Acknowledgements

The financial support of Glaxo Wellcome Research & Development Limited, Merck, Sharp & Dohine Limited, the European Union Biomed II programme and the University of Surrey Foundation Fund is gratefully acknowledged by one of us (D.F.V.L).

References

[1] Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C. and Nebert, D.W. (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, **6**, 1–42.

[2] Lewis, D.F.V. (1996) *Cytochromes P450: Structure, Function and Mechanism*. Taylor & Francis: London.

[3] Ortiz de Montellano, P.R. (ed.) *Cytochrome P450*: Plenum, New York; 1995.

[4] Vanden Bossche, H., Willemsens, G., Marichal, P., Cools, W. and Lauwers, W. (1984) The molecular basis for the antifungal activities of N-substituted azole derivatives: focus on R51211. In. *Mode of Action of Antifungal Agents* (A.P.J. Trinci and J.F., Ryley, eds.) Ch. 5, pp. 321–341. British Mycological Society; London.

[5] Janssen, P.A.J. and Vanden Bossche, H. (1987) Mode of action of cytochrome P450 monooxygenase inhibitors: focus on azole derivatives. *Arch. Pharm. Chemi. (Sci. ed.)*, **15**, 23–40.

[6] Yoshida, Y. and Aoyama, Y. (1991) Cytochromes P450 in the ergosterol biosynthesis. *Frontiers in Biotransformation*, **4**, 127–148.

[7] Schuster, I. (1993) Ergosterol synthesis: the major target in antimycotic therapy. *Frontiers in Biotransformation*, **8**, 147–185.

[8] Vanden Bossche, H. (1988) Mode of action of pyridine, pyrimidine and azole antifungals. In. *Sterol Biosynthesis Inhibitors* (D. Berg and M. Plempel, eds.), pp. 79–119. Ellis Horwood; Chichester.

[9] Yoshida, Y. (1993) P450 oxygenase systems of animal tissues. I. Sterol biosynthesis. In. *Cytochrome P450* (T. Omura, Y. Ishimura and Y. Fujii-Kuriyama, eds.), pp. 93–101. Kodansha; Tokyo.

[10] Yoshida, Y. (1993) P450 monooxygenase system in microorganisms. In. *Cytochrome P450* (T. Omura, Y. Ishimua and, Y. Fujii-Kuriyama, eds.), pp. 171–185. Kodansha; Tokyo.

[11] Vanden Bossche, H., Lauwers, W., Willemsens, G., Marichal, P., Cornelissen, F. and Cools, W. (1984) Molecular basis for the antimycotic and antibacterial activity of N-substituted imidazoles and triazoles: the inhibition of isoprenoid biosynthesis, *Pest. Sci.*, **15**, 188–198.

[12] Vanden Bossche, H., Willemsens, G., Bellens, D., Roels, I. and Janssen, P.A.J. (1990) From 14α-demethylase inhibitors in fungal cells to androgen and oestrogen biosynthesis inhibitors in mammalian cells. *Biochem. Soc. Trans.*, **18**, 10–13.

[13] Vanden Bossche, H., Bellens, D., Cools, W., Gorrens, J., Marichal, P., Verhoeven, H., Willemsens, G., De Coster, R., Beerens, D., Haelterman, C., Coene, M.-C., Lauwers, W. and Le Jeune, L. (1986) Cytochrome P450: target for itraconazole. *Drug Devel. Res.*, **8**, 287–298.

[14] Coulson, C.J., King, D.J. and Wiseman, A. (1984) Chemotherapeutic and agrochemical applications of cytochrome P450 ligands. *Trends Biochem. Sci.*, **9**, 446–449.

[15] Muller, H.-G., Schunck, W.-H. and Kargel, E. (1991) Cytochromes P450 in alkane-assimilating yeasts. *Frontiers in Biotransformation*, **4**, 87–126.

[16] Seghezzi, W., Meili, C., Ruffiner, R., Kuenzi, R., Sanglard, D. and Fiechter, A. (1992) Identification and characterization of additional members of the cytochrome P450 multigene family *CYP52 of Candida tropicalis*. *DNA Cell Biol.*, **11**, 767–780.

[17] Chen, C., Kalb, V.F., Turi, T.G. and Loper, J.C. (1988) Primary structure of the cytochrome P450 lanosterol 14-alpha demethylase gene from *Candida tropicalis*. *DNA*, **7**, 617–626.

[18] Gotoh, O. (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.*, **267**, 83–90.

[19] Ishida, N., Aoyama, Y., Hatanaka, R., Oyama, Y., Imajo, S., Ishiguro, M., Oshima, T., Nakazato, H., Noguchi, T., Maitra, U.S., Mohan, V.P., Sprinson, D.B. and Yoshida, Y. (1988) A single amino acid substitution converts cytochrome $P450_{14DM}$ to an inactive form, cytochrome $P450_{SGI}$: Complete primary structures deduced from cloned DNAs. *Biochem. Biophys. Res. Commun.*, **155**, 317–323.

[20] Raag, R., Li, H., Jones, B.C. and Poulos, T.L. (1993) Inhibitor-induced conformational change in cytochrome $P450_{cam}$. *Biochemistry*, **32**, 4571–4578.

[21] Morris, G.M. and Richards, W.G. (1991) Molecular modelling of the sterol C-14 demethylase of *Saccharomyces cerevisiae*. *Biochem. Soc. Trans.*, **19**, 793–795.

[22] Lewis, D.F.V. and Moereels, H. (1992) The sequence homologies of cytochrome P450 and active site geometries. *J. Computer-Aided Mol. Des.*, **6**, 235–252.

[23] Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. and Deisenhofer, J. (1993) Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450s. *Science*, **261**, 731–736.

[24] Kalb, V.F., Woods, C.W., Turi, T.G., Dey, C.R., Sutter, T.R. and Loper, J.C. (1987) Primary structure of the P450 lanosterol demethylase gene from *Saccharomyces cerevisiae*. *DNA*, **6**, 529–537.

[25] Lewis, D.F.V. and Lee-Robichaud, P. (1998) Molecular modelling of steroidogenic cytochromes P450 from families CYP11, CYP17, CYP19 and CYP21 based on the CYP102 crystal structure. *J. Steroid Biochem. Mol. Biol.*, **66**, 217–233.

[26] Peeters, O.M., Blaton, N.M. and De Ranter, C.J. (1979) The crystal structure of miconazole. *Bull. Soc. Chim. Belg.*, **88**, 265–272.

[27] Peeters, O.M., Blaton, N.M. and De Ranter, C.J. (1979) The crystal structure of ketoconazole. *Acta Crystallographica*, **B35**, 2461–2464.

[28] Skrzypczak-Jankun, E. and Kurumbail, R.G. (1996) 1-Trityl-4-nitroimidazole. *Acta Crystallographica*, **C52**, 189–191.

[29] Pye, G.W. and Marriott, M.S. (1982) Inhibition of sterol C14 demethylation by imidazole-containing antifungals. *Sabouradia*, **20**, 325–329.

[30] Pople, J.A., Santry, D.P. and Segal, G.A. (1965) Approximate self-consistent MO theory. I. Invariant procedures. *J. Chem. Phys.*, **43** (suppl.), 129–135.

[31] Vinter, J.G., Davis, A. and Saunders, M.R. (1987) Strategic approaches to drug design. I. An integrated software framework for molecular modelling, *J. Computer-Aided Mol. Des.*, **1**, 35–51.

[32] Tuck, S.F., Aoyama, Y., Toshida, Y. and Ortiz de Montellano, P.R. (1992) Active site topology of *Saccharomyces cerevisiae* lanosterol 14α-demethylase (CYP51) and its G310D mutant (cytochrome $P450_{SG1}$). *J. Biol. Chem.*, **267**, 13175–13179.
[33] Marichal, P., Gorrens, J. and Vanden Bossche, H. (1985) The action of itraconazole and ketoconazole on growth and sterol synthesis in *Aspergillus fumigatus* and *Aspergillus niger*. *Sabouradia*, **23**, 13–21.
[34] James, M.O. and Sloan, K.B. (1985) Structural features of imidazole derivatives that enhance styrene oxide hydroxylase activity in rat hepatic microsomes. *J. Med. Chem.*, **28**, 1120–1124.
[35] Ballard, S.A., Lodola, A. and Tarbit, M.H. (1988) A comparative study of 1-substituted imidazole and 1,2,4-triazole antifungal compounds as inhibitors of testosterone hydroxylations catalyzed by mouse hepatic microsomal cytochromes P450. *Biochem. Pharmacol.*, **37**, 4643–4651.
[36] Rodrigues, A.D., Gibson, G.G., Ioannides, C. and Parke, D.V. (1987) Interactions of imidazole antifungal agents with purified cytochrome P450 proteins. *Biochem. Pharmacol.*, **36**, 4277–4281.
[37] Rodrigues, A.D., Lewis, D.F.V., Ioannides, C. and Parke, D.V. (1987) Spectral and kinetics studies of the interaction of imidazole antifungal agents with microsomal cytochrome P450. *Xenobiotica*, **17**, 1315–1327.
[38] Lewis, D.F.V., Rodrigues, A.D., Ioannides, C. and Parke, D.V. (1989) Adverse reactions of imidazole antifungal agents: computer graphic studies of cytochrome P450 interactions. *J. Biochem. Toxicol.*, **4**, 231–234.
[39] Ortiz de Montellano, P.R. and Correia, M.A. (1995) Inhibition of cytochrome P450 enzymes. In. *Cytochrome P450* (P.R Ortiz de Montellano, ed.), Chapter 9, pp. 305–364. Plenum; New York.
[40] Correia, M.A. and Ortiz de Montellano, P.R. (1993) Inhibitors of cytochrome P450 and possibilities for their therapeutic application. *Frontiers in Biotransformation*, **8**, 74–146.
[41] Atkins, P.W. and Friedman, R.S. (1997) *Molecular Quantum Mechanics*. Oxford University Press; Oxford.
[42] Pople, J.A. and Beveridge, D.L. (1970) *Approximate Molecular Orbital Theory*. McGraw-Hill; New York.
[43] Lewis, D.F.V. (1989) The calculation of molar polarizabilities by the CNDO/2 method: correlation with the hydrophobic parameter, log, P. *J. Comput. Chem.*, **10**, 145–151.
[44] Lewis, D.F.V. (1990) MO-QSARs: a review of molecular orbital-generated quantitative structure–activity relationships, *Prog. Drug Metab.*, **12**, 205–255.
[45] Avdeef, A. (1996) Assessment of distribution-pH profiles. In. *Lipophilicity in Drug Action and Toxicology* (V. Pliska, B. Testa and H. van de Waterbeemd, eds.), pp. 109–139. VCH; Weinheim.