DDT Vol. 7, No. 17 September 2002

Substrate SARs in human P450s

David F.V. Lewis and Maurice Dickins

Drug metabolism is now an integral part of the drug discovery process, and the cytochromes P450 (CYPs) are the most important family of enzymes involved in human drug metabolism. An increased understanding of the properties of the substrates for the major human CYPs is thus highly desirable. This article shows how key characteristics of CYP substrates, such as lipophilicity, molecular mass and hydrogen-bonding potential, govern selectivity towards individual CYPs. Importantly, the variation in binding affinities of 60 human CYP substrates can be explained by understanding the key physicochemical, structural and electronic characteristics that govern substrate binding to each isozyme.

David F.V. Lewis*
School of Biomedical and Life Sciences
University of Surrey
Guildford, Surrey, UK GU2 7XH
*tel: +44 1483 876 477
fax: +44 1483 300 803
e-mail: d.lewis@surrey.ac.uk
Maurice Dickins
GlaxoSmithKline Research and Development
Park Road, Ware
Hertfordshire, UK SG12 0DP

▼ With poor pharmacokinetics (PK) accounting for over 50% of lead compound failure in drug discovery and development, it is important to identify as early as possible the likely metabolic fate in *Homo sapiens* of new chemical entities (NCEs). Consequently, the prediction of human PK represents an important challenge to those engaged in drug discovery and NCE development [1]. Of the four drug properties that contribute to its overall PK profile (i.e. absorption, distribution, metabolism and excretion, or ADME) it is metabolism that is possibly the most complicated because of the number of enzymes involved, including those of both Phase 1 and Phase 2 metabolism [2].

ADME: metabolism

This article focuses on the metabolism aspect of ADME, although NCE failure can involve toxicity and other adverse effects, both of which could be related to P450 interactions [3]. In fact, a balance is usually required between efficacy and overall PK in the development of a lead compound, such that a high therapeutic activity often has to be sacrificed in favour of an optimal metabolic profile. Consequently, application of the 'rule of 5' to HTS programmes is likely to result in more promising candidate drug structures because high molecular weight (M_r >500) and highly lipophilic (log P>5) compounds, together with those containing more than five hydrogen-bond donors and more than 10 hydrogen-bond acceptors, will give rise to poor permeability and poor dissolution properties [4]. Furthermore, a knowledge and understanding of the P450 system can offer a means of accelerating the drug discovery process because the metabolic profile of a lead compound is extremely important to its probable success in clinical trials.

P450s represent the major Phase 1 enzymes and the occurrence of polymorphic drug metabolism [5,6] in human ethnogeographical populations, where genetic defects in CYP2D6 and CYP2C19 appear to have dominant roles, implies that careful design or redesign of drug candidates is required to avoid compound lead failure caused by undesirable P450 isoform selectivity. Consequently, it is of great relevance to drug discovery to identify at an early stage in NCE development the likely P450 enzymes involved in compound metabolism, together with possible sites of metabolism in the relevant chemical concerned. Here, we will outline the structural criteria that govern substrate selectivity toward human P450s, and demonstrate how substrate characteristics and mode of binding to the enzymes can facilitate estimation of binding affinity, which can be employed in drug screening programmes.

The P450 superfamily

It is established that ~90% of the Phase 1 metabolism of drug substrates is mediated by one or more of the cytochromes P450; these constitute a superfamily of heme-thiolate enzymes, which catalyze primarily monooxygenase reactions involving a two-stage reduction of molecular oxygen and subsequent single-oxygen atom insertion into substrate molecules, although reductive metabolism is also known [7–15]. Human P450 enzymes of families CYP1, CYP2 and CYP3 are associated

PII: S1359-6446(02)02412-1

Table 1. Typical substrates and inhibitors of human cytochromes P450

CYP	Involvement in drug oxidations (%)	Typical marker substrate	K_m (μM)	Selective inhibitor	K_i or K_m (μM)
1A2	8.2	Caffeine	180	Furafylline	0.7
2A6	2.5	Coumarin	2.1	8-Methoxypsoralen	0.5
2B6	3.4	4-Trifluoromethyl 7-ethoxycoumarin	2.9	Orphenadrine	30
2C8	15.8	Rosiglitazone	10	Sulfinpyrazone	17
2C9		Tolbutamide	132	Sulfaphenazole	0.2
2C19	8.3	Omeprazole	8.6	Fluconazole	2
2D6	18.8	Propranolol	2.73 (K_D)	Quinidine	0.06
2E1	4.1	Benzene	25	Pyridine*	0.4
3A4	34.1	Erythromycin	1.8 (K_D)	Ketoconazole	0.1

CYPs 3A4, 2D6 and the CYP2C subfamily are responsible for the majority of drug oxidations mediated by human CYPs. Several widely used marker substrates and inhibitors for these enzymes are shown in this table and reflect the range of K_m and K_i values which have been reported.
*It has been reported [48] that 3-amino-1,2,4-triazole is a selective inhibitor for CYP2E1 but its K_i value of 10mM suggests that it does not exhibit particularly high binding affinity towards this enzyme. K_m = Michaelis constant for substrate metabolism (μM). See references: [7,10,45,49,50].

with the majority of drug oxidations in man [7]. Table 1 shows that the CYP3A, CYP2D and CYP2C subfamilies catalyze most of these oxidations, with those of CYP1A, CYP2A, CYP2B and CYP2E being of lesser importance [7,16]. There is a good correlation between the percentage involvement in drug oxidations and the average percentage of total P450 complement for human P450 enzymes, provided that the major polymorphic forms CYP2D6 and CYP2C19 are excluded [17]. The reason why CYP2D6 is an outlier, in this respect, is because, although it represents a relatively minor component of the total hepatic P450 complement, many drugs contain a basic nitrogen atom that is protonatable at physiological pH and, consequently, are liable to act as CYP2D6 substrates.

To identify substrate SARs within the human P450s [10,18–23], which oxidatively metabolize drug substrates, we have collated extensive information from the literature [7]. This is in the form of K_m values for selective P450 substrates, together with assimilating physicochemical and other data, such as hydrogen bond-forming potential, to analyze the likely aspects of enzyme selectivity so that P450 substrates can be readily identified from their structural characteristics. In the case of the substrates for CYP2D6 and CYP3A4, we have generated K_D values spectroscopically [24] and these are preferable to K_m values in obtaining measurements of substrate-binding affinity; however, Bauer *et al.* [25] have shown that ΔG_{bind} can be calculated directly from K_m. Moreover, in many examples of P450 substrates $K_D \cong K_m$, albeit with some exceptions, and K_m values are widely reported in the literature [7]. Furthermore, we have used these datasets of potential descriptors to formulate quantitative SARs (QSARs), which indicate the main contributions to substrate binding for each major human P450 isoform [24,26–28]. In some cases, the number of active site hydrogen bonds formed between substrate and enzyme has been evaluated from homology modelling of the relevant P450 from the CYP102 structural template [28]. More recently, human P450s have been modelled using the CYP2C5 crystal structure [29] and this represents an advance on previous studies, which employed bacterial forms of the enzyme because of the improved sequence homologies between CYP2C5 and the drug-metabolizing P450s (30–80%) [30,31]. In addition, we have also investigated the factors determining the rates of P450-mediated reactions and the QSARs generated from these studies have been reported recently [32]. Further improvements to the QSAR analyses could be achieved by consideration of the possibility of alternative mechanisms for P450-mediated metabolism, whereas molecular dynamics simulation of P450-substrate binding could prove important for elucidating the processes involved in molecular recognition and selectivity.

Characteristics of human P450 substrates

Table 2 presents information on various P450 substrate characteristics, including their log P ranges and average log P values, although the extent of compound ionization at physiological pH will modify these, such that the distribution coefficients at pH 7.4 (log $D_{7.4}$ values) are also important physicochemical properties [23]. Compilation of various parameters for substrates of human P450 enzymes associated with drug metabolism [24] enables the formulation of structural criteria governing human P450 selectivity [33] and, furthermore, aids the analysis of QSARs for

Table 2. A summary of physicochemical and other characteristics of human P450 substrates

CYP	log P range	Average log P	Other characteristics
1A2	0.01 to 3.61	2.146 (14)	Planar molecules, neutral or basic in character
2A6	0.07 to 2.88	1.717 (10)	Diverse, relatively small neutral or basic molecules usually containing one aromatic ring
2B6	0.23 to 4.75	2.423 (10)	Angular, medium-sized neutral or basic molecules with 1–2 hydrogen bond donor/acceptor atoms
2C8	–0.36 to 4.75	3.688 (12)	Relatively large, elongated molecules, mostly either acidic or neutral in character
2C9	1.58 to 5.18	3.345 (10)	Medium-sized acidic molecules with 1–2 hydrogen bond acceptors
2C19	1.49 to 3.37	2.357 (12)	Medium-sized molecules, mostly basic with 2–3 hydrogen bond acceptors
2D6	0.75 to 5.04	3.031 (12)	Medium-sized basic molecules with protonatable nitrogen 5–7Å from site of metabolism
2E1	–1.35 to 2.83	1.103 (14)	Relatively small neutral molecules, structurally diverse in character
3A4	1.00 to 5.04	2.783 (10)	Relatively large, structurally diverse molecules

CYP substrates are typically lipophilic as shown by their log P values. However, typical substrates of the various CYP isozymes have other characteristics including other physicochemical properties, size and shape which discriminate between the CYPs. These properties can be used as a guide to indicate which CYP(s) might be involved in the metabolism of these substrates. The numbers in parentheses refer to the number of compounds evaluated in each class of P450 substrates. The average log P value is frequently found to correspond with a selective substrate or inhibitor of the enzyme concerned. log P is the logarithm of the octanol–water partition coefficient. See references [23,28].

substrate binding affinities towards individual P450 enzymes [24,26,34,35].

There are several factors that appear to contribute to substrate-binding affinity towards P450 enzymes, and these are listed in Table 3 [15,28]. Based on analyses of many series of P450 substrates, it can be shown that desolvation of the active site and hydrogen-bond formation with key amino acid residues lining the heme pocket represent the major factors governing P450 substrate-binding affinity [34–36]. In most cases, the results of QSAR analysis (summarized in Table 4) tend to agree with information derived from molecular modelling of enzyme–substrate interactions for the human P450s concerned [34]. Furthermore, several key amino acid residues present in the putative active sites of these enzymes have been the subject of site-directed mutagenesis (SDM) experiments, the results of which are largely consistent with the homology modelling studies [15,30]. In some cases, the number of compounds analyzed in a set is somewhat less than ideal. However, fewer variables are nevertheless able to provide satisfactory correlations with affinity. Also, we have found that it is possible to combine these datasets to encompass, for example, an entire CYP2 family. Interestingly, log P and the numbers of hydrogen bonds and π–π stacking interactions give good correlations with binding affinity for 70–90 compounds, especially when the number of rotatable bonds is included in the analysis.

Tables 5–7 present the relevant physicochemical properties and binding data for substrates of CYP1A2, CYP2B6 and CYP3A4, whereas Fig. 1 indicates how a decision tree approach [33] might be employed in the discrimination analysis of various P450

Table 3. Contributions to binding affinity (ΔG_{bind})

Type of process	Energy (kcal. mole^{-1})
Ionic interactions	–4.5 (average or typical value)
Hydrogen bond interactions	–2.0 (average or typical value)
π–π stacking interactions	–0.9 (based on six-membered ring overlaps)
Desolvation processes during binding	–0.025SA (or -0.425 Vol)
Loss in T and R freedom on binding	+log M_r
Loss in bond rotational freedom on binding	+0.6 per bond

Several factors contribute to the binding affinity of compounds for different CYPs. The binding energy can be calculated by summing the energy terms described in the equation for each CYP – substrate pairing. The total expression for overall binding energy can be presented as follows:
$\Delta G_{bind} = \Delta G_{ionic} + \Delta G_{hbond} + \Delta G_{\pi-\pi} + \Delta G_{desolv} + \Delta G_{T+R} + \Delta G_{rotors}$
ΔG_{ionic} = Ionic interaction energy
ΔG_{hb} = Hydrogen bond energy
$\Delta G_{\pi-\pi}$ = π–π Stacking energy
ΔG_{desolv} = Desolvation energy
ΔG_{rotors} = Loss in rotational bond energy on binding
SA = Surface area of solvent accessible (Connolly) surface
Vol = Volume of solvent accessible (Connolly) surface
ΔG_{T+R} = Translational and rotational energy
M_r = Relative molecular mass
The first three terms are enthalpic in nature, whereas the latter three represent entropic terms.
See references [15,28,34].

substrates from families CYP1, CYP2 and CYP3. We have also found that a minimum of five structural descriptors are required for differentiating (with a 92% correlation) between 48 P450 substrates from subfamilies CYP1A, CYP2A–E and CYP3A [33] containing six compounds in each group. In particular, the variables employed for producing a satisfactory discrimination (there were only 8% compounds incorrectly assigned) between the substrates investigated were: area/depth^2, surface area, log $D_{7.4}$, pK_a and E_{LUMO}; the last quantity being able to differentiate CYP2C19 substrates from those of CYP2D6.

In addition, QSAR analyses can be conducted on human P450 substrates [24,26,27,32,37–40] such that the activity of untested compounds can be predicted from the relevant QSAR equation. In this respect, there is usually close agreement between modelling of the enzymes and QSAR findings [30,37]. However, it is also possible to derive values for the binding affinity of P450 substrates based on a combination of the various contributions to the overall binding to the enzyme involved [30]. This involves the use of log P to estimate the desolvation energy, which accompanies substrate binding, together with consideration of the number of active site contacts in the form of hydrogen bonding and π–π stacking interactions. For >20 substrates of P450s from the CYP2 family it is possible to estimate values for their binding affinity, which gives a 98% agreement with experimental data [30], thus demonstrating the use of modelling P450–substrate interactions for evaluating Phase 1 drug metabolism in man.

Table 4. A summary of QSARs for human P450 substrates

CYP	Descriptors	Number of compounds	Correlation	Binding data
1A2	N_{HB}^{Site}, μ, ΔE, I/w	11	0.97	K_m
2A6	N_{HB}^{Site}, log P	6	0.97	K_m
2B6	N_{HB}^{Site}, HB_D, HB_A	10	0.97	K_m
2C9	HB_D, pK_a, log P, log $D_{7.4}$	8	0.99	K_m
2C19	HB_D, log P, M_r, log M_r	8	0.99	K_m
2D6	N_{HB}^{Site}, $N_{\pi-\pi}^{Site}$, M_r, log M_r	10	0.94	K_D
2E1	N_{HB}^{Site}, $N_{\pi-\pi}^{Site}$, M_r, log M_r	10	0.99	K_m
3A4	N_{HB}^{Site}, $N_{\pi-\pi}^{Site}$, E_L, E_H	10	0.96	K_D

The electrostatic, electronic and physicochemical descriptors which best describe the binding data to human CYPs using quantitative SAR (QSAR) analysis are shown in this table. The parameters which occur most frequently across the binding interactions are log P, log M_r and the number of hydrogen bonds, although other parameters are also seen to be important. There is a high degree of correlation between experimentally derived values for binding affinity and those calculated using the relevant QSAR equation using the terms for the relevant CYP indicated here [26].

$N_{\pi-\pi}^{Site}$ = Number of π–π stacking interactions in the active site

N_{HB}^{Site} = Number of hydrogen bond interactions in the active site

μ = Dipole moment

$E_{L,H}$ = Energy of the lowest unoccupied and highest unoccupied MOs, respectively

$\Delta E = E_L - E_H$

I/w = Ratio of molecular length and width

log P = Logarithm of the octanol–water partition coefficient

$HB_{D,A}$ = Number of hydrogen bond donors or acceptors in the molecule, respectively

pK_a = Negative logarithm of the acid–base dissociation constant

log $D_{7.4}$ = Logarithm of the distribution coefficient at pH7.4

K_m = Michaelis constant for substrate interaction with the relevant P450 enzyme (μM)

K_D = Enzyme-substrate dissociation constant (μM)

See references [23,24,26].

Modelling and HTS

It is apparent from the results of extensive QSAR analyses of many substrates of human cytochromes P450 that certain substrate structural descriptors are relevant to particular P450 enzyme binding characteristics. In addition, it would appear that hydrogen bonding within the P450 active site region is important in all cases, irrespective of the human P450 enzyme concerned. However, it is the location and disposition of these hydrogen-bond formations that are likely to govern substrate selectivity towards human P450s, and further work should elucidate these factors in more detail. In particular, homology modelling of human P450s involved in drug metabolism can assist in rationalizing the structural criteria for enzyme selectivity, and the recently available crystallographic coordinates of the mammalian P450, CYP2C5, is proving to be particularly useful in this respect [30]. It would appear that P450s discriminate between substrates on the grounds of various physicochemical properties and structural factors. For example, CYP1A2 substrates (Table 5) tend to be relatively planar molecules of medium size, usually with ~2 or 3 fused aromatic rings in the structure and potential hydrogen-bond acceptor atoms (2–3 acceptors are common) located at a key distance from the observed sites of metabolism. By contrast, CYP2B6 substrates (Table 6) possess generally non-planar molecules, whereas those of CYP3A4 (Table 7) are structurally diverse.

Generally, a compound is not specifically designed to be a substrate of a particular P450 unless it is desirable to produce a diagnostic probe chemical for analytical purposes. However, it might be important to 'design-out' selectivity towards CYP2D6 or CYP2C19, for example, as these enzymes are associated with polymorphic drug metabolism.

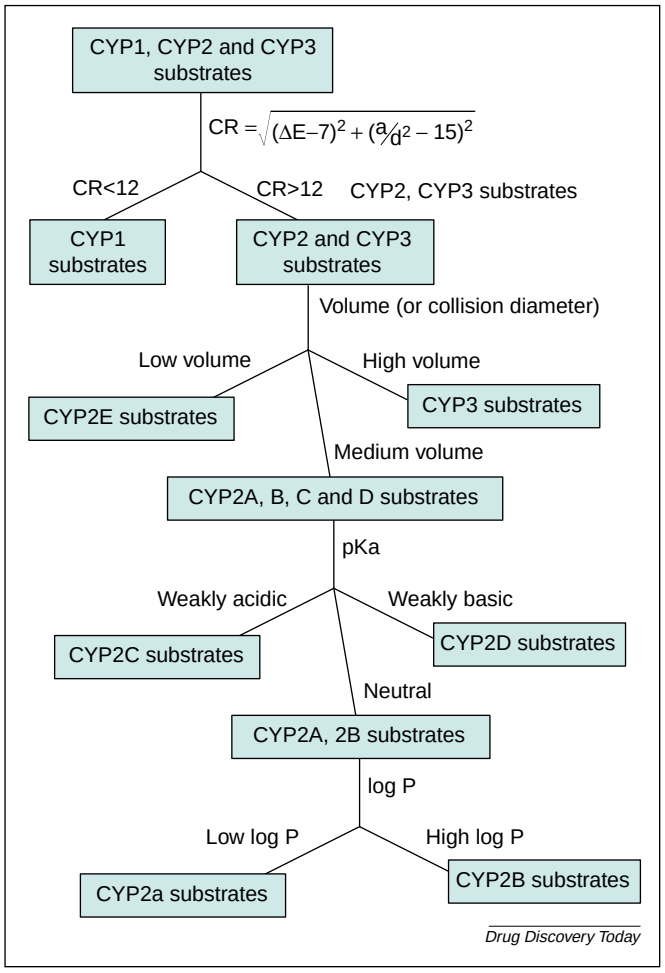

Figure 1. A decision tree approach for evaluating P450 substrate selectivity. This approach is a method for predicting the likely cytochrome P450 (CYP) specificity for potential substrates of these enzymes. The overall size and shape of a molecule is important in defining CYP substrate selectivity. A combination of area/depth^2 and ΔE (the COMPACT or CR ratio; CR = $[(a/d^2 - 15)^2 + \Delta E - 7)^2]$) will differentiate between CYP1A substrates and those of other CYPs. The volume (or collision diameter) can then separate CYP2E substrates (typically small volume) and CYP3A substrates (typically large volume) from those that interact preferentially with CYPs 2A–D. Physicochemical properties, such as log P, pK_a and acidity/basicity, are important for discriminating between substrates of CYPs 2A–D.

Usually, CYP3A4-mediated reactions are regarded as more acceptable for a lead compound than either CYP2D6 or CYP2C19, for example, although potent inhibition or induction of CYP3A are generally regarded as undesirable properties because of the possibility of adverse reactions including toxic effects. Sometimes, a lead compound can be readily redesigned to avoid selectivity for CYP2D6 or CYP2C19, and it is probably beneficial for the candidate drug to be metabolized by more than one P450 so that saturation of a particular pathway is unlikely to occur, as this could lead to adverse drug reactions (ADR). Although the derivation of separate algorithms for therapeutic target properties and P450 selectivity in HTS could be used to drive novel compound design by facilitating parallel combinatorial synthesis, it is also important to emphasize that P450 enzymes themselves can be regarded as therapeutic targets, such as the aromatase enzyme (CYP19) for the

Table 5. Data for selected CYP1A2 substrates

	Compound	log P	pK_a	log $D_{7.4}$	M_r	K_m (μM)	ΔG_{bind}
1	Caffeine	0.01	neutral	0.01	194.22	180	−5.3118
2	PhIP	2.23	8.69[b]	0.92	224.29	55	−6.0422
3	7-Methoxyresorufin	3.15[c]	neutral	3.15[c]	227.23	0.21	−9.4577
4	Phenacetin	1.57	neutral	1.57	179.24	48	−6.1260
5	IQ	1.84	9.18[b]	0.06	198.25	33	−6.3568
6	MeIQ	2.40[c]	9.77[b]	0.06[c]	212.28	13	−6.9307
7	4-Aminobiphenyl	2.86	4.61[b]	2.86	169.24	30	−6.4156
8	7-Ethoxyresorufin	3.61[c]	9.85[b]	1.19[c]	241.26	1.7	−8.1839
9	Aflatoxin	2.20[c]	neutral	2.20[c]	312.29	31	−6.2231
10	Theophylline	1.40[c]	8.8[b]	-0.02	180.2	455	−4.7405
11	Tacrine	2.71	9.8[b]	0.46	198.27	14	−6.8851
12	Estradiol	2.69	neutral	2.69	272.37	20	−6.6653
13	Acetanilide	1.16	neutral	1.16	135.16	N/A	N/A
14	Zoxazolamine	2.21	neutral	2.21	168.58	N/A	N/A

The relevant physicochemical properties and binding data for several CYP1A2 substrates are shown. The data cover a range of substrates with differing affinities for CYP1A2 and the binding energy (ΔG_{bind}) values are calculated from the experimental K_m data. CYP1A2 substrates are typically relatively planar molecules of intermediate size. b = basic; c = calculated value N/A = data not available; IQ = 2-amino-3-methylimidazo-[4,5-f]quinoline; MeIQ = 4-methyl IQ; PhIP = 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine; $\Delta G_{bind} = RT\ln K_m$; K_m = Michaelis constant for substrate metabolism (μM).

design of antitumour agents [41]. Moreover, the same physicochemical characteristics (e.g. lipophilicity) could be relevant for increasing compound potency towards the target and for binding to P450 enzymes [42].

Conclusions

As far as human P450s are concerned, we point to the finding that crucial factors for substrate binding include molecular size and shape, number and disposition of hydrogen bond donor/acceptor atoms and aromatic rings in the molecule and, in some cases, the possession of a basic or acidic grouping. These give rise to P450 enzyme selectivity and determine the route of metabolism, as there is usually a key distance in the substrate molecule between a site of interaction with the enzyme (e.g. hydrogen bond donor/acceptor atom or charged grouping) and the position where metabolism, usually oxidation, occurs. Molecular modelling of substrate–inhibitor interactions with P450 active sites can help to explain selectivity, route of metabolism and, in some cases, even enantioselectivity. The results of molecular modelling of human P450s usually agree closely with the reported findings of SDM experiments, and this adds confidence to the model-building process, together with providing a degree of validation for the models themselves [30].

A clear relationship bet ween compound lipophilicity and binding affinity can sometimes be discerned for substrates of a given P450. For example, 10 structurally diverse CYP2B6-selective compounds (see Table 6) covering a wide range of affinity towards the enzyme exhibit a good correlation between partitioning free energy, ΔG_{part}, and binding free energy, ΔG_{bind}, according to Eqn 1:

$$\Delta G_{bind} = 0.871\ \Delta G_{part} - 2.447\ (\pm 0.062) \qquad \text{[Eqn 1]}$$

$n = 10$; $s = 0.3256$; $R = 0.98$; $R^2 = 0.96$; $F = 197.1$

where n is the number of points, s is the standard error, R is the correlation coefficient, R^2 is the determination coefficient and F is the variance ratio.

Table 6. Data for selected CYP2B6 substrates

	Compound	log P	pK_a	log $D_{7.4}$	M_r	K_m (μM)	ΔG_{bind}
1	7-Benzyloxyresorufin	4.75[c]	10.22[b]	1.98[c]	303.33	1.28	−8.3587
2	4-Trifluoromethyl 7-ethoxycoumarin	3.31	neutral	3.31	242.21	2.9	−7.8549
3	Testosterone	3.32	neutral	3.32	288.43	50.5	−6.0947
4	Benzphetamine	2.27[e]	6.6[b]	2.21[c]	239.35	93.4	−5.7159
5	7-Ethoxycoumarin	2.30	neutral	2.30	190.21	115	−5.5878
6	Diazepam	2.86	3.7[b]	2.82	284.80	113	−5.5986
7	Bupropion	2.54	8.35[b]	1.54	239.77	107.5	−5.6293
8	S-Mephenytoin	1.74	8.1[b]	0.96	218.28	564	−4.6082
9	SM-12502	1.06	4.46[b]	1.02	208.30	1767	−3.9047
10	Antipyrine	0.23	neutral	0.23	188.25	17700	−2.4852

The relevant physicochemical properties and binding data for several 2B6substrates are shown. The data cover a range of substrates with differing affinities for CYP2B6 and the binding energy (ΔG_{bind}) values are calculated from the experimental K_m data. CYP2B6 substrates are typically non-planar molecules. b = basic; c = calculated value; e = estimated value; $\Delta G_{bind} = RT\ln K_m$; K_m = Michaelis constant for substrate metabolism (μM).

Table 7. Data for selected CYP3A4 substrates

	Compound	log P	pK_a	log $D_{7.4}$	M_r	K_D (μM)	ΔG_{bind}
1	Amitriptyline	5.04	9.4[b]	2.50	227.30	348	−4.7474
2	Budesonide	3.28	neutral	3.28	430.60	6.4	−7.1296
3	Etoposide	1.00	neutral	1.00	588.60	23.2	−6.3618
4	Midazolam	1.53	neutral	1.53	325.80	8.8	−6.9398
5	Phenacetin	1.57	neutral	1.57	179.24	9.1	−6.9198
6	Verapamil	3.79	8.29[b]	2.85	454.60	3.5	−7.4894
7	Ethynylestradiol	2.11	neutral	2.11	296.40	6.5	−7.1204
8	Testosterone	3.32	neutral	3.32	288.43	21.7	−6.4017
9	Erythromycin	2.48	8.8[b]	1.06	733.90	1.8	−7.8858
10	Salmeterol	3.71	9.3[b]	1.80	415.60	165	−5.1923

The relevant physicochemical properties and binding data for several 3A4substrates are shown. The data cover a range of substrates with differing affinities for CYP3A4 and the binding energy (ΔG_{bind}) values are calculated from the experimental K_D data. CYP3A4 substrates show the greatest diversity in their properties and are typically large, lipophilic molecules. b = basic; $\Delta G_{bind} = RT\ln K_D$; K_D = Dissociation constant (μM) for substrate binding to the enzyme.

Outliers to such lipophilicity relationships tend to show additional hydrogen bond-forming capabilities or possess a greater number of rotatable bonds than those of most substrates. Consequently, these factors all need to be taken into account for a full description of P450 binding affinity. It can be demonstrated that a combination of log P, number of hydrogen bonds formed on binding, number of π–π stacking interactions formed and number of rotatable bonds restricted on binding correlates closely with ΔG_{bind} for >20 P450 substrates, within the CYP2 family of enzymes, with an $R^2 = 0.96$ [30]. These considerations might also be

applicable to the therapeutic target enzyme or protein, thus leading to a parallel methodology for HTS of drug candidates for optimum efficacy and drug metabolism and PK (DMPK) profiles. However, the rate of a compound's metabolic clearance could depend on further factors, which have not been considered so far, although it is likely that lipophilic character will have an important role, possibly in combination with molecular electronic properties [32].

It is believed that an application of these principles could lead to improved drug discovery programmes, resulting in better products that are likely to be tailored to meet the requirements of genetically diverse human populations. Consequently, an understanding of the complexities of human P450 enzymes [43,44] can provide accelerated lead times for the design of safer medicines in the future, provided that such considerations become successfully incorporated into HTS programmes [45–47].

Acknowledgements

The financial support of GlaxoSmithKline R&D (http://www.gsk.com), Merck Sharp & Dohme (http://www.merck.com), the European Union and the University of Surrey Foundation Fund (http://www.surrey.ac.uk) is gratefully acknowledged by DFVL.

References

1 Pelkonen, O. *et al.* (2002) *Pharmacokinetic Challenges in Drug Discovery*, Springer–Verlag, Berlin
2 Gibson, G.G. and Skett, P. (2001) *Introduction to Drug Metabolism*, Nelson Thorns, Cheltenham
3 Rodrigues, A.D. and Lin, J.H. (2001) Screening of drug candidates for their drug–drug interaction potential. *Curr. Opin. Chem. Biol.* 5, 396–401
4 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3–25
5 Evans, W.E. and Relling, M.V. (1999) Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* 286, 487–491
6 Ingelman-Sundberg, M. (2001) Implications of polymorphic cytochrome P450-dependent drug metabolism for drug development. *Drug Metab. Dispos.* 29, 570–573
7 Rendic, S. and DiCarlo, F.J. (1997) Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. *Drug Metab. Rev.* 29, 413–580
8 Anzenbacher, P. and Anzenbacherova, E. (2001) Cytochromes P450 and metabolism of xenobiotics. *Cell. Mol. Life Sci.* 58, 737–747
9 Lewis, D.F.V. (1996) *Cytochromes P450: Structure, Function and Mechanism*, Taylor & Francis, London
10 Lewis, D.F.V. (2001) *Guide to Cytochromes P450 Structure and Function*, Taylor & Francis, London
11 Ortiz de Montellano, P.R. (1995) *Cytochrome P450*, Plenum, New York
12 Guengerich, F.P. (2001) Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* 14, 611–650
13 Guengerich, F.P. (2002) Cytochrome P450. In *Enzyme Systems that Metabolize Drugs and other Xenobiotics* (Ioannides, C., ed.), Chapter 2, pp.33–65, Wiley, New York
14 Lewis, D.F.V. and Hlavica, P. (2000) Interactions between redox partners in various cytochrome P450 systems: functional and structural aspects. *Biochim. Biophys. Acta* 1460, 353–374
15 Lewis, D.F.V. *et al.* (1998) Structural determinants of P450 substrate specificity, binding affinity and catalytic rate. *Chem. Biol. Interact.* 115, 175–199
16 Ioannides, C. (1996) *Cytochromes P450: Metabolic and Toxicological Aspects*, CRC Press, Boca Raton, Florida
17 Lewis, D.F.V. (1998) The CYP2 family: models, mutants and interactions. *Xenobiotica* 28, 617–661
18 Smith, D.A. and Jones, B.C. (1992) Speculations on the substrate structure-activity relationship (SSAR) of cytochrome P450 enzymes. *Biochem. Pharmacol.* 44, 2089–2098
19 Smith, D.A. *et al.* (1997) Properties of cytochrome P450 isoenzymes and their substrates. Part 1: active site characteristics. *Drug Discov. Today* 2, 406–414
20 Smith, D.A. *et al.* (1997) Properties of cytochromes P450 isoenzymes and their substrates. Part 2: properties of cytochrome P450 substrates. *Drug Discov. Today* 2, 479–486
21 Lewis, D.F.V. (1999) Molecular modelling of human cytochromes P450 involved in xenobiotic metabolism and rationalization of substrate specificity. *Exp. Toxicol. Pathol.* 51, 369–374
22 Lewis, D.F.V. Quantitative structure-activity relationships (QSARs) within the cytochrome P450 system: QSARs describing substrate binding, inhibition and induction of P450s. *Inflammopharmacology* (in press)
23 Lewis, D.F.V. (2000) Structural characteristics of human P450s involved in drug metabolism: QSARs and lipophilicity profiles. *Toxicology* 144, 197–203
24 Lewis, D.F.V. *et al.* (2001) Quantitative structure-activity relationships (QSARs) within substrates of human cytochromes P450 involved in drug metabolism. *Rev. Drug Metab. Drug Interact.* 18, 221–241
25 Bauer, C. *et al.* (2001) A unified theory of enzyme kinetics based upon the systematic analysis of the variations of k_{cat}, K_m and k_{cat}/K_m, and the relevant ΔG^o values – possible implications in chemotherapy and biotechnology. *Biochem. Pharmacol.* 61, 1049–1055
26 Lewis, D.F.V. *et al.* (2002) Structure-activity relationships for human cytochrome P450 substrates and inhibitors. *Drug Metab. Rev.* 34, 69–82
27 Lewis, D.F.V. and Dickins, M. (2001) Quantitative structure-activity relationships (QSARs) within series of inhibitors for mammalian cytochromes P450 (CYPs). *J. Enzyme Inhib.* 16, 321–330
28 Lewis, D.F.V. *et al.* (1999) Cytochrome P450 substrate specificities, substrate structural templates and enzyme active site geometries. *Rev. Drug Metab. Drug Interact.* 15, 1–49
29 Williams, P.A. *et al.* (2000) Mammalian cytochrome P450 monooxygenase: structure adaptations for membrane binding and functional diversity. *Mol. Cell* 5, 121–131
30 Lewis, D.F.V. (2002) Homology modelling of human CYP2 family enzymes based on the CYP2C5 crystal structure. *Xenobiotica* 32, 305–323
31 Wang, Q. and Halpert, J.R. (2002) Combined three-dimensional quantitative structure-activity relationship analysis of cytochrome P450 2B6 substrates and protein homology modeling. *Drug Metab. Dispos.* 30, 86–95
32 Lewis, D.F.V. and Dickins, M. (2002) Factors influencing rates and clearance in P450-mediated reactions: QSARs for substrates of the xenobiotic-metabolizing hepatic microsomal P450s. *Toxicology* 170, 45–53
33 Lewis, D.F.V. (2000) On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics: towards the prediction of human P450 substrate specificity. *Biochem. Pharmacol.* 60, 293–306
34 Lewis, D.F.V. (2000) Modelling human cytochromes P450 for evaluating drug metabolism. *Rev. Drug Metab. Drug Interact.* 16, 307–324
35 Lewis, D.F.V. (2001) COMPACT: a structural approach to the modelling of cytochromes P450 and their interactions with xenobiotics. *J. Chem. Technol. Biotechnol.* 76, 1–8
36 Lewis, D.F.V. and Lake, B.G. Species differences in coumarin metabolism: a molecular modelling evaluation of CYP2A interactions. *Xenobiotica* (in press)
37 Ekins, S. *et al.* (2001) Pharmacophore and three-dimensional quantitative structure-activity relationship methods for modeling cytochrome P450 active sites. *Drug Metab. Dispos.* 29, 936–944

38 Ekins, S. *et al.* (1999) Three- and four-dimensional quantitative structure-activity relationship (3D/4D-QSAR) analyses for CYP2C9 inhibitors. *Drug Metab. Dispos.* 28, 994–1002
39 Ekins, S. *et al.* (1999) Three-dimensional quantitative structure-activity relationship analyses of substrates for CYP2B6. *J. Pharmacol. Exp. Ther.* 288, 21–29
40 Ekins, S. *et al.* (1999) Three-dimensional quantitative structure-activity relationship analysis of cytochrome P450 3A4 substrates. *J. Pharmacol. Exp. Ther.* 291, 424–433
41 Guengerich, F.P. (2002) Cytochrome P450 enzymes in the generation of commercial products. *Nat. Rev* 1, 359–366
42 Smith, D.A. and Van de Waterbeemd, H. (1999) Pharmacokinetics and metabolism in early drug discovery. *Curr. Opin. Chem. Biol.* 3, 373–379
43 Lewis, D.F.V. and Sheridan, G. (2001) Cytochromes P450, oxygen and evolution. *The Scientific World* 1, 151–167
44 Lewis, D.F.V. *et al.* (1998) Evolution of the cytochrome P450 superfamily: sequence alignments and pharmacogenetics. *Mutat. Res.* 410, 245–270
45 Clarke, S.E. (1998) *In vitro* assessment of human cytochrome P450. *Xenobiotica* 28, 1167–1202
46 Gedeck, P. and Willett, P. (2001) Visual and computational analysis of structure-activity relationships in high-throughput screening data. *Curr. Opin. Chem. Biol.* 5, 389–395
47 Eddershaw, P.J. and Dickins, M. (1999) Advances in *in vitro* drug metabolism screening. *Pharm. Sci. Technol. Today* 2, 13–19
48 Koop, D.R. (1990) Inhibition of ethanol-inducible cytochrome P450IIE1 by 3-amino-1,2,4-triazole. *Chem. Res. Toxicol.* 3, 377–383
49 Hargreaves, M.B. *et al.* (1994) Inhibition of *p*-nitrophenol hydroxylase in rat liver microsomes by small aromatic and heterocyclic molecules. *Drug Metab. Dispos.* 22, 806–810
50 Pelkonen, O. *et al.* (2000) CYP2A6: a human coumarin 7-hydroxylase. *Toxicology* 144, 139–147