DDT Vol. 9, No. 12 June 2004

Compound lipophilicity for substrate binding to human P450s in drug metabolism

David F.V. Lewis, Miriam N. Jacobs and Maurice Dickins

Compound lipophilicity is of key importance to P450 binding affinity and enzyme selectivity. Here, lipophilicity is discussed with reference to the human drug-metabolizing P450 enzymes of families CYP1, CYP2 and CYP3. From an extensive compilation of log P values for P450 substrates, and by analysis of relationships between partitioning energy and substrate-binding free energy, the relevance of lipophilicity and other factors pertaining to P450 binding affinity is explained, leading to the formulation of lipophilicity relationships within substrates of each human P450 enzyme involved in drug metabolism. Furthermore, log P values for P450 substrates appear to represent markers for enzyme selectivity. Together with the important roles of hydrogen bonding and π–π stacking interaction energies, the desolvation of the P450 active site makes a major contribution to the overall substrate-binding energy and, consequently, a good agreement with experimental information is reported based on this analysis.

David F.V. Lewis*
Miriam N. Jacobs
School of Biomedical and Molecular Sciences
University of Surrey
Guildford
Surrey
UK GU2 7XH
*e-mail: d.lewis@surrey.ac.uk
Maurice Dickins
Pfizer Central Research
Ramsgate Road
Sandwich
Kent
UK CT13 9NJ

▼ Enzymes of the cytochrome P450 (CYP) superfamily represent the major catalysts for the metabolism of drugs and other xenobiotics in *Homo Sapiens*. In addition, they have endogenous roles in, for example, the biosynthesis of steroid hormones and the oxidative metabolism of fatty acids [1–3]. In particular, it is enzymes of the CYP1, CYP2 and CYP3 families that are primarily involved in the phase 1 metabolism of foreign compounds [4,5] and the relative percentages of these enzymes have been estimated by various methods [6,7]. However, it is apparent that inter-individual variations exist, in addition to established pharmacogenetic factors encountered in ethnogeographical populations, which have important ramifications for drug development [8].

The P450 enzyme families

Each P450 enzyme family and subfamily exhibits selectivity for substrates of a particular type, although there can be a certain degree of overlap depending on the chemical class of compound in question [9]. Nevertheless, it is apparent that lipophilicity is fundamentally important to P450 substrate selectivity and binding affinity or clearance [10–13], together with showing some influence on relative catalytic rate [14,15]. Hansch and colleagues have published the details of several quantitative SAR (QSAR) studies on P450 substrates, where log P (the lipophilicity parameter, where P is the octanol–water partition coefficient) is either the main or sole descriptor [16,17]; log P is also important in estimating metabolic half-lives in man based on rat data [18]. Consequently, from these and other reported findings, it is apparent that the primary component of substrate-binding to P450 enzymes is one of lipophilicity, probably due to desolvation, where the compound log P (or log $D_{7.4}$) represents a useful marker of their overall lipophilic character [19]. Lipophilicity is one of the key parameters for evaluating the ADME/Tox properties of new chemical entities [20–23] and, therefore, it is interesting to note that this quantity is also of considerable relevance to the ability of compounds to act as substrates for P450 enzymes.

Lipophilicity relationships in human P450 substrates

The role of compound lipophilicity in biological activity and related interactions [24–28] has been well established, particularly over the past 30 years, following the pioneering work of Corwin Hansch and Al Leo [29,30]. More recently, the importance of lipophilic character to the success of drug discovery has been underscored in the well-known Lipinski 'rule of five' concept [31].

1359-6446/04/$ – see front matter PII: S1359-6446(04)03115-0

There have been several reports in the literature that point to the relevance of P450 substrate lipophilic character to enzyme binding, inhibition and metabolic clearance [16,17,32–36]. For example, for a series of statins acting as competitive inhibitors of CYP3A4, there is a good correlation (R = 0.96) between log K_i (where K_i is the inhibition constant) and log $D_{7.0}$ where log $D_{7.0}$ is a measure of compound lipophilicity, this being the ionization-corrected log P value at pH 7.0 [33]. A similar relationship is apparent for a more structurally diverse series of CYP3A4 inhibitors, some of which are also substrates [35]. CYP2D6 substrates also display a broadly linear correlation between binding to the enzyme and log $D_{7.4}$ value [37]. Furthermore, in an investigation on 7-alkoxycoumarins [38] the logarithm of O-dealkylation mediated by CYP2B4 is found to be directly proportional to log P. It would appear that basic substrates tend to bind to CYP2B enzymes in the unionized state [39], as opposed to the situation encountered in CYP2D6. From these findings, and also those reported by Hansch and co-workers [16,17], we have made a detailed investigation of the possible relationships between lipophilicity, in the form of either log P or log $D_{7.4}$ values, and binding to human P450 enzymes that are associated with the metabolism of drugs and other xenobiotics [14,15]. The results of this work have shown clearly that substrate binding affinity obtained from K_m or (K_D) data exhibit linear correlations with log P, either expressed as the corresponding partitioning energy or as the raw log P values versus –log K_m. For example, when considering a set of 16 structurally diverse substrates of CYP2B6, there is a simple linear relationship between –log K_m and log P of the following form [Equation 1]:

$$\log K_m = 0.881 \log P\ (\pm 0.058) + 1.676 \quad \text{[Eqn 1]}$$

where n = 16; s = 0.2378; R = 0.97; F = 233.

Figure 1 shows a plot of this correlation for the CYP2B6 substrates listed in Table 1. The K_m values have been taken from a review of the relevant literature [40], whereas the log P data have been collated from several sources [41,42] or calculated using the Pallas software [19] from CompuDrug (http://www.compudrug.com). Relations of the above type can be rewritten in terms

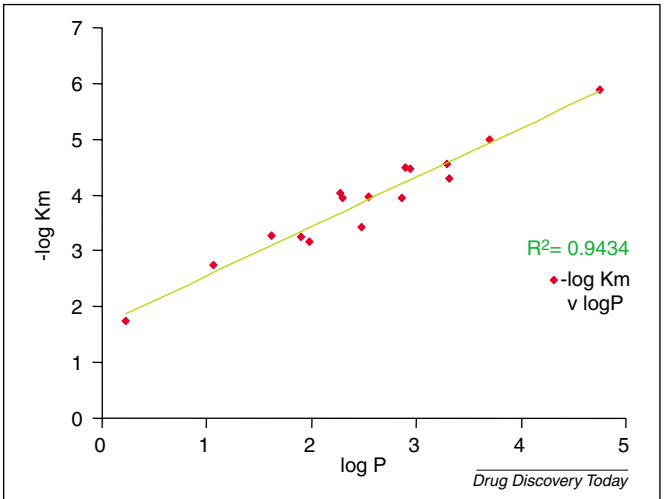

Figure 1. The correlation between -log K_m (*x*-axis) and log P (*y*-axis) for 16 structurally diverse substrates of CYP2B6, where P is the octanol-water partition coefficient and K_m is the apparent Michaelis constant (data from Table 1).

Table 1. Dataset for CYP2B6 substrate lipophilicity relationship

Compound	log P	K_m (μM)	–log K_m	ΔG_{part}	ΔG_{bind}
7-Benzyloxyresorufin	4.75[c]	1.28	5.8928	–6.7377	–8.3587
Testosterone	3.32	50.5	4.2967	–4.7093	–6.0947
Benzphetamine	2.27[c]	93.4	4.0297	–3.2199	–5.7159
7-Ethoxycoumarin (EC)	2.30	115	3.9393	–3.2625	–5.5878
Diazepam	2.86	113	3.9469	–4.0568	–5.5986
Bupropion	2.54	107.5	3.9686	–3.6029	–5.6293
S-Mephenytoin	1.90	564	3.2487	–2.6951	–4.6082
SM-12502	1.06	1767	2.7528	–1.5036	–3.9047
Antipyrine	0.23	17700	1.7520	–0.3262	–2.4852
4-Chloromethyl EC	2.94[c]	33.7	4.4724	–4.1703	–6.3439
R-Deprenyl	2.90	33	4.4815	–4.1136	–6.3568
Propofol	3.70	10	5.0000	–5.2483	–7.0923
Lidocaine	1.62	537.6	3.2695	–2.2979	–4.6377
Carbamazepine	1.98	700	3.1549	–2.8086	–4.4751
Imipramine	2.48	383	3.4168	–3.5178	–4.8466
Arteether	3.29	28	4.5528	–4.6668	–6.4581

Regression Equation	n	s	R	F
-log K_m = 0.881 log P (±0.058) + 1.676	16	0.2378	0.9713	233.45
ΔG_{bind} = 0.881 ΔG_{part} (±0.058) – 2.377	16	0.3373	0.9713	233.45

ΔG_{bind} = RTln K_m
ΔG_{part} = Rln P
R = gas constant (1.9872 cal $deg^{-1}mol^{-1}$)
T = absolute temperature (310K)
c = calculated value (Pallas Software, CompuDrug, http://www.compudrug.com)
Reference to K_m data in reference [40].

Table 2. Lipophilicity relationships in human P450 substrates

CYP	No. of Compounds	Correlation	Slope	Intercept	Common interactions
1A1	5	0.964	0.694	–4.263	One hydrogen bond and two π–π stacks
1A2	10	0.975	0.595	–4.832	Two hydrogen bonds and one π–π stack
1B1	7	0.937	0.398	–5.857	Two hydrogen bonds and two π–π stacks
2A6	5	0.973	1.617	–3.213	One hydrogen bond and one π–π stack
2B6	16	0.971	0.881	–2.377	One hydrogen bond and one π–π stack
2C8	5	0.963	1.859	–1.820	One hydrogen bond
2C9	9 [6]	0.947	0.558	–3.366	One hydrogen bond and one π–π stack
2C19	6 [5]	0.963	1.119	–1.925	One hydrogen bond
2D6	6 [5]	0.908	1.004	–4.027	One ion-pair
2E1	7 [7]	0.998	1.102	–1.645	One hydrogen bond or two π-stacks
3A4	40+	Several	Several	Several	Different combinations of hydrogen bonding and π–π stacking

Numbers in parentheses refer to secondary parallel lipophilicity relationships in other substrates towards the same enzyme. These correspond to additional hydrogen bond or π-π stacking interactions within the P450 active site region, although the statistical data presented only refers to the primary relationship. Data obtained from reference [15]

of the corresponding free energy changes, ΔG_{bind} and ΔG_{part} [15] where [Equation 2]:

$$\Delta G_{bind} = RTlnK_m \text{ and } \Delta G_{part} = -RTlnP \qquad \text{[Eqn 2]}$$

where R is the gas constant and T is the absolute temperature, usually taken as 310K

The utility of this transformation is that one can readily obtain an estimation of the non-desolvation related contributions to the enzyme–substrate binding energy from the y intercept of the relationship expressed either graphically or via regression, together with a quantification of the hydrophobic character of the P450 active-site environment [15]. Table 2 presents a summary of these lipophilicity relationships for many P450 substrates.

However, the following points should be borne in mind in relation to these findings:

- The range of hydrogen bond energies for biological systems is from –1.5 to –2.5 kcal mol^{-1} with –2.0 kcal mol^{-1} representing an average value.
- The value of π–π stacking energies is thought to range from –0.8 to –1.0 kcal mol^{-1} with –0.9 kcal mol^{-1} representing an average value.
- It is common for there to be more than one linear lipophilicity relationship for each P450 enzyme.

Contributory factors

There are several contributory factors that give rise to the total enzyme–substrate binding energy, therefore, one of the useful aspects of lipophilicity relationships (Table 2) is that one tends to observe several parallel lines for larger groups of P450 substrates, thus indicating that the number of hydrogen bond and π–π stacking interactions might vary for a given P450, as is the case with CYP3A4 substrates [15]. In fact, the intercept of the lipophilicity relationship directly provides a measure of the average non-desolvation component, and this appears to be a simple linear combination of the hydrogen bond and π–π stacking energies, that have been reported as –2 kcal mol^{-1} and –0.9 kcal mol^{-1}, respectively. Furthermore, the ranges and averages of log P values for P450 substrates might also be indicative of their enzyme selectivity, as shown in Table 3.

Some points can be recognized from these and related findings:

- Binding affinity is determined by log P, number of hydrogen bond acceptors and/or donors and π–π stacking interactions (N) and number of rotatable bonds restricted on binding [19].
- The site of substrate metabolism is determined by hydrogen bond acceptors and/or donors in the molecule [19].
- The log P values of typical substrates are close to the average value for the P450s in question (see Table 3).
- Lipophilicity relationships exist with ΔG_{bind} for substrates of all P450s investigated to date [15].

Combining this information can prove useful for evaluating the characteristic structural determinants for P450 binding and selectivity and, in addition to lipophilicity,

Table 3. Ranges and average log P values for human P450 substrates of families CYP1, CYP2 and CYP3

% Drugs[†]	CYP	Range of log P values	Average log P value (n)	Other characteristics	Typical substrate and log P
3	1A1*	1.39 to 6.35	3.41 [16]	Planar PAHs and their diols	DMBA-3,4-diol 3.42
10	1A2	0.08 to 3.61	2.01 [18]	Planar amines and amides	MeIQ 1.98
1	1B1	1.40 to 6.35	3.73 [12]	Planar PAHs and their diols	BP-7,8-diol 3.87[c]
3	2A6	0.07 to 2.79	1.44 [18]	Fairly small molecules	Losigamone 1.46
4	2 B6	0.23 to 4.89	2.54 [16]	Basic (Unionized)	Buproprion 2.54
	2C8	0.06 to 6.98	3.38 [12]	Acidic (Ionized)	Rosiglitazone 3.20
25	2C9	0.89 to 5.18	3.20 [18]	Acidic (Unionized)	Naproxen 3.18
	2C19	1.49 to 4.42	2.56 [16]	Amides and amines	Proguanil 2.53
15	2D6	0.75 to 5.04	3.08 [16]	Basic (Ionized)	Propranolol 3.09
3	2E1	-1.35 to 3.63	2.07 [20]	Small molecules	4-Nitrophenol 2.04
N/A	–2F1*	0.37 to 5.14	2.63 [12]	Fairly small molecules	3-Methylindole 2.72[c]
36	3A4	0.97 to 7.54	3.10 [50]	Large molecules	Nifedipine 3.17

n = number of compounds investigated c = calculated value (Pallas Software) N/A = data not available * = extrahepatic enzymes
DMBA = dimethylbenzanthracene MeIQ = 2-amino-3,4-dimethylimidazo[4,5-f]quinoline BP = benzo(a)pyrene
[†] = percentage of total drug oxidations catalyzed by a given P450 enzyme based on information from reference 4 for a total of 2112 compounds.
Potentially important discriminants of P450 selectivity include:
Size (diameter), molecular mass or length of the molecule
Planarity (a/d^2) or Rectangularity (l/w) of the molecule
Lipophilicity (log P or log $D_{7.4}$)
Basicity/Acidity of the compound (pK_a)
Polarity (Dipole moment) of the molecule

the role of hydrogen bonding [43] should be emphasized because this is a key component of enzyme–substrate interactions [44,45]. Molecular modelling of human P450-substrate complexes tends to be complementary to the lipophilicity relationship analysis, although the dynamics of substrate binding [46] might suggest the operation of an entropy-driven desolvation process, which could be related to compound lipophilicity. In addition, pharmacophore modelling and QSAR analysis [47,48] in the P450 field tend to support the molecular modelling approaches [12,13] where substrate properties are taken into consideration, with structural similarity representing another useful characteristic [49]. Therefore, although important, compound lipophilicty represents one factor that would need to be augmented by other information to derive a full profile of the likely P450 selectivity and metabolism of a chemical.

Calculation of P450 substrate binding affinity and log P data

The general expression for estimating the binding affinity (ΔG_{bind}) of P450 substrates can be formulated as follows [Equation 3]:

$$\Delta G_{bind} = \Delta G_{part} + \Delta G_{hb} + \Delta G_{\pi\text{-}\pi} + \Delta G_{rot} + \text{constant} \quad \text{[Eqn 3]}$$

ΔG_{part} relates to the desolvation component obtained from the lipophilicity parameter log P, described previously; ΔG_{hb} is the hydrogen bond component; $\Delta G_{\pi\text{-}\pi}$ is the π–π stacking interaction energy, and ΔG_{rot} is the loss in individual bond rotational energy when a small substrate molecule binds to the P450 active site [19].

Using Equation 3, it is possible to make a calculation of the overall binding energy in terms of the number of active site interactions observed from molecular modelling studies [19]. Each of these interactions are multiplied by the average energy quoted in the literature, and the desolvation component is calculated directly from the experimental log P value using Equation 2. For a total of 90 P450 substrate interactions, the correlation with experimental binding affinity has been shown to be as high as 98% [19] and the constant term is effectively zero. These findings tend to suggest that the loss in substrate translational and rotational energy upon binding to the P450 enzyme is negligible, possibly due to some degree of translational and rotational energy evident in the P450 itself, as it is embedded in a relatively fluid phospholipid bilayer, which constitutes the endoplasmic reticular membrane. Subsequent use of this approach, using Equation 3, enabled calculation of CYP3A4 substrate binding affinities, which gave good agreement with experimental values [19].

Table 4. Comparison between Pallas and ClogP systems for calculating log P values of selected P450 substrates

Compound	Pallas	ClogP	log P_{expt}
1,4-Cineole	2.53	2.83	2.50
1-Nitropyrene	4.94	4.69	5.06
3-Amino-1,2,4-triazole	–0.83	–0.53	–0.87
7-Ethoxycoumarin	2.15	2.27	2.30
7-Pentoxycoumarin	3.68	3.86	3.92
7-Propoxycoumarin	2.66	2.80	2.86
Artemether	2.89	2.91	2.86
Artemisinin	2.86	2.72	2.90
Camphor	1.82	2.18	2.38
Carbamazepine	3.52	1.98	2.19
Clarithromycin	3.03	2.09	3.16
Clofibric acid	2.73	3.26	2.57
Coumarin	1.75	1.41	1.39
Deprenyl (Selegiline)	2.85	2.52	2.90
Diltiazem	2.58	3.65	2.80
Hexamethylphosphoramide (HMPA)	0.28	0.11	0.28
Isoflurane	3.07	3.00	2.06
Lornoxicam	0.80	2.33	2.62
Melatonin	1.45	1.03	1.01
Methoxsalen	2.37	2.30	1.93
MPTP	2.58	2.81	2.71
Nifedipine	2.82	3.41	2.17

The correlation between ClogP and log P_{expt} is 0.91 for the 22 compounds listed above although consideration of a larger compound dataset (>100) improves this to 0.98.
The correlation between Pallas and log P_{expt} is 0.87 for the 22 compounds listed above although consideration of a larger dataset (>100) improves this to 0.98.

Information from molecular modelling

In the absence of information from molecular modelling, one has to rely on the lipophilicity relationship solely because this is essentially the major contribution to the binding energy in most cases. In some examples, this is quite satisfactory as shown by the study of CYP2B6 substrates outlined previously. Clearly, there is a need for accurate log P values in such calculations [50] and, in the absence of experimental values for some compounds, one has to resort to computational methods of log P calculation, of which several systems are currently available [42,51–54]. For example, the ClogP software package (Biobyte Corporation; http://www.biobyte.com) and the Pallas system are able to provide calculated values of log P for structural input in the form of either a SMILES (Simplified Molecular Input Line Entry System) string (ClogP) or via a graphical 2D representation (Pallas), where the latter system employs the Rekker fragmental methodology, together with an ability to calculate pK_a values and, hence, log D for any pH required.

Table 4 shows a comparison between these two methods for several P450 substrates and, in fact, there is a satisfactory agreement with experimental log P data for each method, particularly when relatively large numbers of compounds are examined. The excellent work by Pliska and coworkers [55] contains further comparisons between a greater number of log P calculators, but the results presented in this review are of a similar order. Possibly the ClogP system is somewhat superior to Pallas for drug structures, although the latter is preferred for compounds such as cyclosporin (where the Pallas value is 3.27, the ClogP value is 14.36 and log P_{expt} is 2.92) and other types of macrocycles. Consequently, it is of benefit to have access to both systems for the calculation of log P data in the absence of experimental values. However, there are other methods that have been developed for the estimation of log P data, and some of these are available online. For occasions where experimental log P data are difficult or time-consuming to generate, many procedures for calculating these values can be used, which, in general, exhibit satisfactory agreement with measured data. The generally accepted methods for obtaining experimental values of log P for various compounds include the 'shake flask' procedure and from HPLC measurements [56] with the latter system becoming the usual approach adopted by most workers [36,55,57]. However, it should be noted that HPLC methods generally give rise to log $D_{7.4}$ values [34], which would require information on the relevant pK_a data [58,59] before the log P value can be calculated using Equations 4 and 5 [55]:

$$\log D = \log P - \log (1 + 10^{pH-pKa}) \text{ for weak acids} \quad [Eqn\ 4]$$

$$\log D = \log P - \log (1 + 10^{pKa-pH}) \text{ for weak bases} \quad [Eqn\ 5]$$

Consequently, it would appear that experimental log P values are able to provide useful determinants of the desolvation

contribution to binding affinity both for P450 substrates and in other systems. As such, simple correlation analyses between activity and log P can provide useful insights into the main contributions to overall binding affinity, especially when this is performed on the associated energy values.

According to the theoretical derivation [15], $\Delta G_{bind} = a \Delta G_{part} + b$ where a is the hydrophobicity of the P450 active site [19] and b is the combined contributions from hydrogen bonding, π–π stacking and other components of the binding energy. Therefore, a plot of $-\log K_m$ versus log P (see Figure 1) should give a straight line of slope 'a' and a y intercept of –0.705b.

For example, with structurally diverse CYP2B6 substrates we get Equation 6:

$$\log K_m = 0.881 \log P \ (\pm 0.053) + 1.676 \quad \text{[Eqn 6]}$$
$$n = 16;\ s = 0.2378;\ R = 0.9713;\ F = 233.45$$

Such that a = 0.881, indicating a fairly hydrophobic active site, and b = –2.377 (–1.676/0.705), which implies that the binding site interaction is likely to be a hydrogen bond, as this is close to the average value of –2 kcal mol^{-1} although one can not rule out the possibility of some contribution from π–π stacking. The degree of standard deviation and error limits in the above expression satisfactory encompass the likely variations in hydrogen bond energies for O—O and O—N hydrogen bonds [44] and this accords with the typical active site interactions shown from molecular modelling studies of CYP2B6 [60]. However, it should also be recognized that weak hydrogen bonds can form between aromatic ring systems and hydrogen bond donors [61].

Lipophilicity and P450 selectivity

The lipophilicity of the potential substrate is crucial for P450-mediated metabolism. In general, mammalian P450 substrates tend to possess log P values greater then zero and, in fact, there are few that have a negative log P. The likely reasons for this relate to membrane transport and substrate binding to the P450 active site, where desolvation of the haem environment hydrophobic 'pocket' usually represents the major contribution to binding affinity.

Human P450 enzymes display certain substrate selectivities, which are reflected, to some extent, in substrate log P ranges and average log P values that are close to those of highly selective substrates of the particular enzymes in question. Indeed, the finding that some compounds might act as substrates for more than one P450 can be explained, in part, by their log P value lying within the ranges encountered for those enzymes. Particular ranges of log P for individual P450s can be rationalized in terms of variations in hydrophobicity and local dielectric constant of each P450 active site, and the slope of the lipophilicity relationship for substrates of a particular P450 can provide a measure of the degree of hydrophobicity in the environment of the haem.

Apparently, the log P value of a substrate can give an indication of its P450 selectivity (although, in some cases, a substrate might be metabolized by more than one enzyme) and provides a means of estimating its binding affinity if the lipophilicity relationship is already established, especially for a congeneric series, and also provided that the potential active site interactions (i.e. hydrogen bonding and π–π stacking) are well characterized from molecular modelling. However, it is important to emphasize that multiple parallel linear relationships are apparent when a fairly large number of structurally diverse substrates are examined for a given P450. Furthermore, other factors should be taken into consideration when analyzing P450 selectivity, although the lipophilicity relationships facilitate a further analysis of binding affinity as the hydrophobicity of the enzyme active site and substrate interactions can be inferred.

In effect, the compound log P is crucial for the evaluation of new chemical entities (NCEs) and, for ionizable substrates, the log $D_{7.4}$ can be more relevant, which effectively means that consideration of the appropriate pK_a values is essential. Procedures for calculating log P are useful for HTS as the established experimental methods can be time-consuming and obviously require chemical synthesis. In fact, the most widely used software systems give good correlations (with R values between 0.95 and 0.98) with experimental values for log P and log $D_{7.4}$, although some compounds are notable outliers. Cyclosporin, for example, is not well estimated by ClogP but the Pallas system provides a calculated value closed to the experimental one. Consequently, it is advisable to make use of at least two log P calculators in the evaluation of compound lipophilicity for drug development.

Conclusions

In conclusion, therefore, it can be assumed that compound lipophilicity will play a major role in substrate binding to cytochromes P450 in mammalian systems and this will have an effect on the overall clearance of such compounds. By and large, the greater the value of log P for a chemical, the more avidly it will bind to a given P450 enzyme, although, after a certain point, binding affinity (and probably clearance) will start to fall with increasing log P, as the compound will become too hydrophobic. There is likely to be an optimum log P value for substrates

of a given P450, with the majority of such selective compounds falling within a certain range of log P values (see Table 3). The precise lipophilic nature of the substrate will therefore determine, at least to some extent, the P450 enzyme most likely to metabolize it and an approximate binding affinity can be estimated from substrate log P, although it is important to emphasize that other factors might also be involved, such as polarity, hydrogen bond donor and/or acceptor properties and the number of aromatic rings present in the molecule. Nevertheless, such considerations could facilitate the process of NCE candidate selection during drug discovery and thus aid in the overall development process.

Acknowledgements

The financial support of GlaxoSmithKline Research and Development Limited, Merck Sharp and Dohme Limited and the University of Surrey Foundation Fund is gratefully acknowledged by DFVL. Miriam Jacobs would like to thank GlaxoSmithKline and BBSRC for their generous funding of a PhD studentship.

References

1 Lewis, D.F.V. (1996) *Cytochromes P450: Structure, Function and Mechanism*, Taylor & Francis, London
2 Lewis, D.F.V. (2001) *Guide to Cytochromes P450 Structure and Function*, Taylor & Francis, London
3 Stegeman, J.J. and Livingstone, D.R. (1998) Forms and functions of cytochrome P450. *Comp. Biochem. Physiol. C* 121, 1–3
4 Rendic, S. (2002) Summary of information on human CYP enzymes: human P450 metabolism data. *Drug Metab. Rev.* 34, 83–448
5 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
6 Shimada, T. *et al.* (1994) Inter-individual variations in human liver cytochrome P450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. *J. Pharmacol. Exp. Ther.* 270, 414–423
7 Edwards, R.J. *et al.* (1998) Development of a comprehensive panel of antibodies against the major xenobiotic metabolizing forms of cytochrome P450 in humans. *Biochem. Pharmacol.* 56, 377–387
8 Rodrigues, A.D. and Rushmore, T.H. (2002) Cytochrome P450 pharmacogenetics in drug development: *in vitro* studies and clinical consequences. *Curr. Drug Metab.* 3, 289–309
9 Lewis, D.F.V. (2003) Human cytochromes P450 associated with the Phase 1 metabolism of drugs and other xenobiotics: a compilation of substrates and inhibitors of the CYP1, CYP2 and CYP3 families. *Curr. Med. Chem.* 10, 1955–1972
10 Smith, D.A. (1994) Chemistry and enzymology: their use in the prediction of human drug metabolism. *Eur. J. Pharm. Sci.* 2, 69–71
11 Smith, D.A. and van de Waterbeemd, H. (1999) Pharmacokinetics and metabolism in early drug discovery. *Curr. Opin. Chem. Biol.* 3, 373–378
12 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
13 Smith, D.A. *et al.* (1997) Properties of cytochrome P450 isoenzymes and the substrates. Part 2: properties of cytochrome P450 substrates. *Drug Discov. Today* 2, 479–486
14 Lewis, D.F.V. and Dickins, M. (2002) Substrate SARs in human P450s. *Drug Discov. Today* 7, 918–925
15 Lewis, D.F.V. and Dickins, M. (2003) Baseline lipophilicity relationships in human cytochromes P450 associated with drug metabolism. *Drug Metab. Rev.* 35, 1–18
16 Hansch, C. and Zhang, L. (1993) Quantitative structure-activity relationships of cytochrome P450. *Drug Metab. Rev.* 25, 1–48
17 Gao, H. and Hansch, C. (1996) QSAR of P450 oxidation: on the value of comparing k_{cat} and K_m with k_{cat}/K_m. *Drug Metab. Rev.* 28, 513–526
18 Sarver, J.G. *et al.* (1997) Estimating xenobiotic half-lives in humans from rat data: influence of log P. *Environ. Health Perspect.* 105, 1204–1209
19 Lewis, D.F.V. (2003) On the estimation of binding affinity (ΔG_{bind}) for human P450 substrates. *Curr. Drug Metab.* 4, 331–340
20 Hansch, C. (1969) A quantitative approach to biochemical structure-activity relationships. *Acc. Chem. Res.* 2, 232–239
21 Modi, S. (2003) Computational approaches to the understanding of ADMET properties and problems. *Drug Discov. Today* 8, 621–623
22 Leo, A. *et al.* (1971) Partition coefficients and their uses. *Chem. Rev.* 71, 525–616
23 van de Waterbeemd, H. *et al.* (2001) Property-based design: optimization of drug absorption and pharmacokinetics. *J. Med. Chem.* 44, 1313–1333
24 Kerns, E.H. and Di, L. (2003) Pharmaceutical profiling in drug discovery. *Drug Discov. Today* 8, 316–323
25 Kennedy, T. (1997) Managing the drug discovery/development interface. *Drug Discov. Today* 2, 436–444
26 Livingstone, D.J. (2000) The characterization of chemical structures using molecular properties: a survey. *J. Chem. Inf. Comput. Sci.* 40, 195–209
27 van de Waterbeemd, H. and Testa, B. (1987) The parametrization of lipophilicity and other structural properties in drug design. *Adv. Drug Res.* 16, 85–225
28 Kubinyi, H. (1997) QSAR and 3D QSAR in drug design. Part 2: applications and problems. *Drug Discov. Today* 2, 538–546
29 Clark, D.E. (2001) Prediction of intestinal absorption and blood-brain barrier penetration by computational methods. *Comb. Chem. High Throughput Screen.* 4, 477–496
30 Cronin, M.T.D. *et al.* (2001) Parametrization of electrophilicity for the prediction of the toxicity of aromatic compounds. *Chem. Res. Toxicol.* 14, 1498–1505
31 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Adv. Drug Deliv. Rev.* 23, 3–25
32 Al-Gailany, K.A.S. *et al.* (1978) The role of substrate lipophilicity in determining type 1 microsomal P450 binding characteristics. *Biochem. Pharmacol.* 27, 783–788
33 Ishigami, M. *et al.* (2001) A comparison of the effectgs of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibitors on the CYP3A4-dependent oxidation of mexazolam *in vitro*. *Drug Metab. Dispos.* 29, 282–288
34 Austin, R.P. *et al.* (2002) The influence of nonspecific microsomal binding on apparent intrinsic clearance, and its prediction from physicochemical properties. *Drug Metab. Dispos.* 30, 1497–1503
35 Riley, R.J. *et al.* (2001) Development of a generalized, quantitative physicochemical model of CYP3A4 inhibition for use in early drug discovery. *Pharm. Res.* 18, 652–655
36 Ramos-Nino, M.E. *et al.* (1996) Quantitative structure-activity relationship for the effect of benzoic acids, cinnamic acids and benzaldehydes on *Listeria monocytogenes*. *J. Appl. Bacteriol.* 80, 303–310
37 Ferrari, S. *et al.* (1991) The role of lipophilicity in the inhibition of polymorphic cytochrome P450IID6 oxidation by beta-blocking agents *in vitro*. *Life Sci.* 48, 2259–2265
38 Vermeir, M. *et al.* (1992) Interaction of 7-n-alkoxycoumarins with cytochrome $P450_2$ and their partitioning into liposomal membranes. *Biochem. J.* 284, 483–490

39 Vermeir, M. *et al.* (1992) Modelling of interaction of basic lipophilic ligands with cytochrome P450 reconstituted in liposomes. Determination of membrane partition coefficients of S(-)-nicotine and N,N-diethylaniline from spectral binding studies and fluorescence quenching. *Biochim. Biophys. Acta* 1107, 93–104
40 Ekins, S. and Wrighton, S.A. (1999) The role of CYP2B6 in human xenobiotic metabolism. *Drug Metab. Rev.* 31, 719–754
41 Hansch, C. and Leo, A.J. (1979) *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York
42 Hansch, C. *et al.* (1995) *Exploring QSAR 2. Hydrophobic, Electronic and Steric Constants*, American Chemical Society, Washington
43 Abraham, M.H. *et al.* (2002) Application of hydrogen bonding calculations in property based drug design. *Drug Discov. Today* 7, 1056–1063
44 Lewis, D.F.V. Hydrogen bonding in human P450-substrate interactions: a major contribution to binding affinity. *Scientific World J.* (in press)
45 Lewis, D.F.V. and Broughton, H.B. (2002) Molecular binding interactions: their estimation and rationalization of QSARs in terms of theoretically derived parameters. *Scientific World J.* 2, 1776–1802
46 Winn, P.J. *et al.* (2002) Comparison of the dynamics of substrate access in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5361–5366
47 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
48 de Groot, M.J. and Ekins, S. (2002) Pharmacophore modeling of cytochromes P450. *Adv. Drug Deliv. Rev.* 54, 367–383
49 Clark, D.E. and Pickett, S.D. (2000) Computational methods for the prediction of 'drug-likeness'. *Drug Discov. Today* 5, 49–57
50 Vaes, W.H.J. *et al.* (1998) Understanding and estimating membrane/water partition coefficients: approaches to derive quantitative structure property relationships. *Chem. Res. Toxicol.* 11, 847–854
51 Meylan, W.M. and Howard, P.H. (1995) Atom/fragment contribution method for estimating octanol-water partition coefficients. *J. Pharm. Sci.* 84, 83–92
52 Moriguchi, I. *et al.* (1992) Simple method of calculating octanol/water partition coefficient. *Chem. Pharm. Bull. (Tokyo)* 40, 127–130
53 Moriguchi, I. *et al.* (1994) Comparison of reliability of log P values for drugs calculated by several methods. *Chem. Pharm. Bull. (Tokyo)* 42, 976–978
54 Lombardo, F. *et al.* (2000) $ElogP_{oct}$: a tool for lipophilicity determination in drug discovery. *J. Med. Chem.* 43, 2922–2928
55 Pliska, V. *et al.* (eds.) (1996) *Lipophilicity in Drug Action and Toxicology*, VCH, Weinheim
56 Sangster, J. (1989) Octanol-water partition coefficients of simple organic compounds. *J. Phys. Chem. Ref. Data* 18, 1111–1229
57 Dias, N.C. *et al.* (2003) Evaluation of a reversed-phase column (Supercosil LC-ABZ) under isocratic and gradient elution conditions for estimating octanol-water partition coefficients. *Analyst* 128, 427–433
58 Newton, D.W. and Kluza, R.B. (1978) pK_a values of medicinal compounds in pharmacy practice. *Drug Intell. Clin. Pharm.* 12, 546–554
59 Dean, J.A. (ed.) (1992) *Lange's Handbook of Chemistry* (14th edn), McGraw-Hill
60 Lewis, D.F.V. *et al.* (2003) Molecular modelling of CYP2B6 based on homology with the CYP2C5 crystal structure: analysis of enzyme–substrate interactions. *Drug Metabol. Drug Interact.* 19, 115–135
61 Levitt, M. and Perutz, M.F. (1988) Aromatic rings act as hydrogen bond donors. *J. Mol. Biol.* 201, 751–754