QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSARs) WITHIN SUBSTRATES OF HUMAN CYTOCHROMES P450 INVOLVED IN DRUG METABOLISM

David F.V. Lewis[1]*, Sandeep Modi[2] and Maurice Dickins[2]

[1]*Molecular Toxicology Group, School of Biomedical and Life Sciences, University of Surrey, Guildford, Surrey, GU2 7XH*
[2]*GlaxoWellcome Research and Development, Park Road, Ware, Herts SG12 0DP, UK*

SUMMARY

The results of quantitative structure-activity relationship (QSAR) analyses are reported for structurally diverse series of chemicals which act as substrates or inhibitors for human hepatic microsomal cytochromes P450 (CYP). In particular, this study focuses on the major catalysts of drug metabolism in man, namely CYP1A2, CYP2B6, CYP2C9, CYP2C19, CYP2D6 and CYP3A4. It is found that good correlations (with correlation coefficients ranging from R = 0.94 to 0.99) with P450 binding affinity (K_m and K_D) or competitive inhibition (K_i) values are obtained in each case, especially when consideration of hydrogen bonding parameters are included in the QSAR analysis, together with the number of π-π stacking interactions.

KEY WORDS

quantitative structure-activity relationships (QSARs), cytochromes P450, homology modelling, drug metabolism, substrate selectivity

* Author for correspondence:
e-mail: d.lewis@surrey.ac.uk

INTRODUCTION

Enzymes of the cytochrome P450 (CYP) superfamily are present in all biological kingdoms /1/ and in most species investigated to date /2,3/. In *Homo sapiens*, it is apparent that hepatic microsomal P450s are associated with the majority (over 90%) of drug oxidations; the main human isoforms mediating Phase 1 metabolism include CYP1A2, CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4 /4/. An understanding of the substrate selectivity of these eight enzymes /5,6/ is important for the pre-screening of pharmaceutical agents, whereby an early assessment of potential routes of metabolism in human subjects can be advantageous for future development of new chemical entities (NCEs).

Table 1 shows the percentages of various P450 enzymes which make up the human hepatic P450 complement, together with their relative involvement in drug oxidations /4/. From this table, it can be appreciated that enzymes of the CYP3A (34%), CYP2D (19%) and CYP2C (24%) subfamilies represent the most extensive catalysts of human Phase 1 metabolism, with those of the CYP1A (10%), CYP2E (4%) and CYP2B (3%) subfamilies being of lesser importance. The CYP2A (2%) and CYP2F (1%) subfamilies can be regarded as of relatively minor consequence with respect to drug oxidations in man, although CYP2A6 is relevant to the metabolism of halothanes, cyclophosphamide and AZT (3′-azido-3′-deoxythymidine) /4/. With the exception of the major polymorphic forms, CYP2D6 and CYP2C19, there is overall a very good correlation ($R = 0.99$) between the average relative percentages of six human liver P450s and their percentage participation in drug metabolism /7/. However, it should be appreciated that there is also a considerable individual variation in human hepatic P450 complement, with the data presented in Table 1 corresponding to average values.

It is apparent that some degree of substrate selectivity exists in human hepatic P450 enzymes /5,6,8/. It is known, for example, that CYP2D6 substrates are all basic with at least one protonatable nitrogen lying at a distance of between 4Å and 7Å from the site of metabolism, although some exceptions exist /9/. Tables 2 to 7 provide summaries of substrate selectivity for the major human P450 enzymes /6,8/ based on extensive analyses of many compounds which act as substrates for individual P450s. However, in order to distinguish

TABLE 1

Human hepatic P450s and drug oxidations /4,27/

CYP	% P450 Complement*	% Drug oxidations
1A1	<1	2.5
1A2	~13	8.2
1B1	<1	N/A
2A6	~4	2.5
2B6	<1	3.4
2C8, 2C9	~18	15.8
2C18, 2C19	~1	8.3
2D6	≤2.5	18.8
2E1	≤7	4.1
2F1	≤1	~1.3
3A4, 3A5	≤28	34.1
4A11	~14	N/A

* There is extensive variability in the levels of some P450 enzymes and the figures shown here represent average values.
N/A = data not available.

between substrates of different human P450s it is usually necessary to consider several parameters which describe various properties of the compounds concerned, which may be both physicochemical and structural in nature /10,11/. In this study, we focus on the major human hepatic P450 enzymes involved in drug metabolism such that QSARs describing the variations in binding affinity or inhibitory activity can be formulated for future screening programmes. Ekins and co-workers have employed three-dimensional QSAR methodology to investigate inhibitors of CYP2D6 /12/, CYP3A4 /13/, CYP2C9 /14/ and CYP2B6 /15/ substrates. Although these studies involved the

consideration of different sets of compounds to those described in this one, it would appear that there is generally good agreement with results of homology modelling of the enzymes themselves for both approaches. The methodology involved in our own studies is now described.

METHODS

Biological data in the form of dissociation constants (K_D values) for substrates binding to human P450s were determined by UV-visible spectroscopy. For this procedure, human P450s were purified by methods that were very similar to those published previously /16/. Changes in absorbance at 418 nm were measured keeping the enzyme concentration constant (5-8 μM) in 0.1 M phosphate buffer (pH 7.5). Substrate concentration was varied from 30 to 400 μM, whereas K_D values were calculated according to the method described by He and coworkers /17/. Michaelis constants (K_m values) for substrate interactions with individual P450s were collated from the literature, as was the inhibition data (K_i values) in some cases (i.e. for CYP2C19 and CYP2D6 inhibitors). The free energy of binding (ΔG_{bind}) for the enzyme-substrate interaction /18/ was calculated using the following equation:

$$\Delta G_{bind} = RT \ln K_D \text{ (or } RT \ln K_m)$$

where R is the gas constant and T is the absolute temperature (a value of 310K corresponding to 37°C has been used throughout). Physicochemical properties of various substrates and inhibitors were either obtained from the literature /19-22/ or, in a few cases, calculated using the Pallas system (CompuDrug Ltd., Budapest).

Quantitative structure-activity relationships (QSARs) were generated via stepwise linear multiple regression analysis (MRA) of the biological and structural data using standard statistical procedures. The variables employed are listed in Tables 2 to 8 and, in some cases, the number of hydrogen bond donors and acceptors were included, together with the numbers of hydrogen-bonded and π-π stacking interactions present in the enzyme active site. For CYP1A2 substrates, molecular orbital (MO)-calculated frontier orbital energies were included, together with molecular shape parameters, area/depth^2 and length/width. In the case of CYP3A4 substrates, the frontier orbital

energy values of the highest occupied MO (HOMO) and lowest unoccupied MO (LUMO) were also included in the analysis. The total number of variables utilized in each QSAR analysis varied from 8 to 16, depending on the P450 enzyme and numbers of substrates employed for which binding or inhibition data were available. The MRA results were examined for possible cross-correlations and the most statistically valid expressions were then selected, following the elimination of covariant equations, as being sufficiently robust QSARs for further study.

RESULTS AND DISCUSSION

Table 2 presents the relevant physicochemical and biological data for a number of human CYP1A2 substrates, and Table 3 lists similar information for CYP2B6 substrates. Tables 4 and 5 contain data for CYP2C9 and CYP2C19 substrates, and those for CYP2D6 substrates are presented in Table 6. Table 7 provides the physicochemical and biological data for substrates of CYP3A4, and the relevant QSARs produced from correlation analysis of these P450 substrates are summarized in Table 8. The results of QSAR analysis are now discussed under each P450 type.

1. CYP1A2 substrates

Table 2 provides information on 14 selective substrates of human CYP1A2, and binding data in the form of K_m values are presented for 11 compounds. A general structural characteristic of these substrates is their molecular planarity, although binding affinity does not appear to correlate with indices of planarity, such as area/depth^2, probably due to the fact that the variation is relatively small. However, a second shape parameter, namely rectangularity, as measured by length/width ratio (l/w), occurs in the QSAR expressions governing binding free energy (ΔG_{bind}), as shown in Table 8, Section 1. Other factors contributing to the binding affinity include dipole moment (μ) and energy of lowest unoccupied molecular orbital (E_L). Equation 1a in Table 8 represents a QSAR with these three descriptors, and this gives a good correlation (R = 0.95) with experimentally-derived values for CYP1A2 binding affinity. However, this correlation can be improved

TABLE 2

CYP1A2 substrates

Compound	log P	pK_a	log $D_{7.4}$	M_r	log M_r	K_{m1} (μM)	ΔG_{bind} (kcal/mol^{-1})
1. Caffeine	0.01	neutral	0.01	194.22	2.2883	180	-5.3118
2. PhIP	2.23	8.69[b]	0.92	224.29	2.3508	55	-6.0422
3. 7-Methoxyresorufin	3.15[c]	neutral	3.15[c]	227.23	2.3565	0.21	-9.4577
4. Phenacetin	1.57	neutral	1.57	179.24	2.2534	48	-6.1260
5. IQ	1.84	9.18[b]	0.06	198.25	2.2972	33	-6.3568
6. MeIQ	2.40[c]	9.77[b]	0.06[c]	212.28	2.3269	13	-6.9307
7. 4-Aminobiphenyl	2.86	4.61[b]	2.86	169.24	2.2285	30	-6.4156
8. 7-Ethoxyresorufin	3.61[c]	9.85[b]	1.19[c]	241.26	2.3825	1.7	-8.1839
9. Aflatoxin	2.20[c]	neutral	2.20[c]	312.29	2.4946	31	-6.2231
10. Theophylline	1.40[c]	8.8[3]	-0.02	180.2	2.2558	455	-4.7405
11. Tacrine	2.71	9.8[3]	0.46	198.27	2.2973	14	-6.8851
12. Estradiol	2.69	neutral	2.69	272.37	2.4351	N/A	N/A
13. Acetanilide	1.16	neutral	1.16	135.16	2.1308	N/A	N/A
14. Zoxazolamine	2.21	neutral	2.21	168.58	2.2268	N/A	N/A

b = basic; c = calculated value; N/A = data not available.
PhIP = 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine; IQ = 2-amino-3-methylimidazo[4,5-f]quinoline; MeIQ = 2-amino-3,4-dimethylimidazo[4,5-f]quinoline.

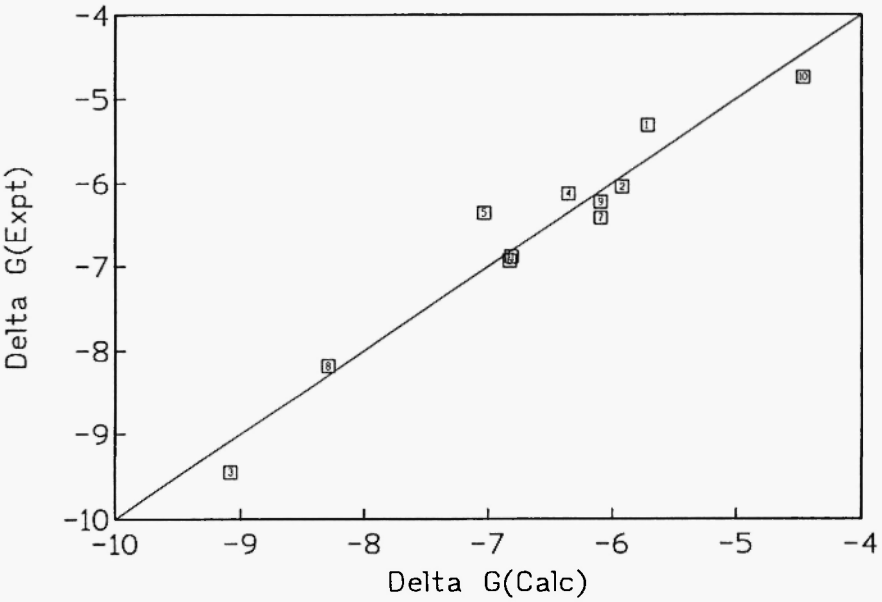

Fig. 1: A comparison between observed and calculated ΔG_{bind} values for several CYP1A2 substrates based on data presented in Tables 2 and 8. Points 1 to 11 refer to compounds 1 to 11, respectively, in Table 2.

somewhat (R = 0.97) by the inclusion of another variable, this being the number of hydrogen bonds formed (N_{HB}) between each substrate and the putative active site of CYP1A2 /23/. The relevant QSAR expression is shown as Equation 1b in Table 8, and this gives a very good agreement (see Fig. 1) between experimentally observed values for ΔG_{bind} and those calculated using the relevant equation (1b in Table 8). Consequently, it appears that the avidity of binding to human CYP1A2 is determined by substrate polarity (μ term), rectangularity (l/w term) and electron-accepting ability (E_L), with active site hydrogen bonding playing a role in the enzyme-substrate interaction.

2. CYP2B6 substrates

Table 3 indicates physicochemical and binding data on various CYP2B6 substrates, some of which show selectivity towards this

TABLE 3

CYP2B6 substrates

	Compound	log P	pK_a	log $D_{7.4}$	M_r	log M_r	K_m (μM)	ΔG_{ind} (kcal·mol^{-1})
1.	**7-Benzyloxyresorufin**	4.75[c]	10.22[b]	1.98[c]	303.33	2.4819	1.28	-8.3587
2.	**4-Trifluoromethyl 7-ethoxycoumarin**	3.31	neutral	3.31	242.21	2.3842	2.9	-7.8549
3.	**Testosterone**	3.32	neutral	3.32	288.43	2.4600	50.5	-6.0947
4.	**Benzphetamine**	2.27[c]	6.6[b]	2.21[c]	239.35	2.3790	93.4	-5.7159
5.	**7-Ethoxycoumarin**	2.15[c]	neutral	2.15	190.21	2.2792	115	-5.5878
6.	**Diazepam**	2.86	3.7[b]	2.82	284.8	2.4545	113	-5.5986
7.	**Bupropion**	2.54	8.35[b]	1.54	239.77	2.3798	107.5	-5.6293
8	**S-Mephenytoin**	1.74	8.1[b]	0.96	218.28	2.3390	564	-4.6082
9	**SM-12502**	1.06	4.46[b]	1.02	208.30	2.3187	1767	-3.9047
10.	**Antipyrine**	0.23	neutral	0.23	188.25	2.2747	17100	-2.4852

b = basic; c = calculated value.
SM-12502 = (+)-cis-3,5-dimethyl-2-(3-pyridyl)thiazolidin-4-one.

isoform. It can be appreciated from inspection of Table 3 that typical CYP2B substrates are either weakly basic or neutral in character, and with a medium lipophilic nature as evidenced by their log P and log $D_{7.4}$ values. These substrates of CYP2B6 tend to possess non-planar molecules of medium relative molecular mass, and there are usually several hydrogen bond donor/acceptor groups in the molecule. It appears that the number of hydrogen bonds formed in the CYP2B6 putative active site is of importance to binding affinity when combined with the numbers of hydrogen bond donors and acceptors in the molecule. Equation 2a in Table 8 shows that these three parameters provide a good description of binding energy for substrates of CYP2B6 with high correlation (R = 0.97) and Figure 2 provides a comparison between experimental and QSAR-calculated values. Homology modelling of CYP2B6 and subsequent interactive docking of various substrates of the enzyme within the putative active site

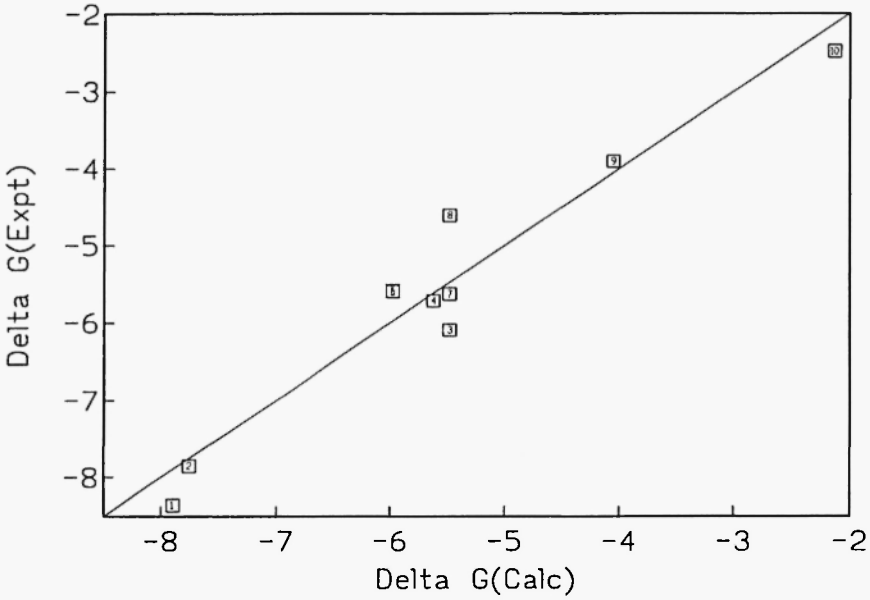

Fig. 2: A comparison between observed and calculated ΔG_{bind} values for several CYP2B6 substrates based on data presented in Tables 3 and 8. Points 1 to 10 refer to compounds 1 to 10, respectively, in Table 3.

indicates that hydrogen-bonded interactions with amino acid residues are important for orientating the substrate molecule for metabolism at the experimentally observed position /24/. Consequently, there is reinforcement of the modelling studies by the results of QSAR evaluations on CYP2B6 substrates. Furthermore, several key amino acid residues lining the haem pocket have been shown to have an importance for substrate binding and selectivity based on the results of site-directed mutagenesis experiments on CYP2B subfamily enzymes /24/.

3. CYP2C9 substrates

The physicochemical and binding data for several substrates of CYP2C9 are presented in Table 4, from which it can be appreciated that all compounds listed are weakly acidic in character. Table 8 (Section 3) provides details of QSARs conducted on several CYP2C9 substrates. Equation 3a (Table 8) refers to analyses carried out on compounds 1-7 in Table 4, whereas Equation 3b (Table 8) gives information on a QSAR produced when compound 8 (naproxen) is included in the analysis (see Fig. 3 for a comparison between experimental and QSAR-calculated values). The latter compound was included in the analysis following the more recent availability of K_m information. The two equations, however, are very similar in terms of the descriptors involved and their relative coefficients, although the variance ratio (F value) is somewhat lower for eight compounds (Equation 3b, Table 8). These expressions, nevertheless, indicate the importance of lipophilicity (log P, log D) and acidic character (pKa) to the overall binding affinity of CYP2C9 substrates for the enzyme. Furthermore, the number of hydrogen bond donor atoms in the substrate clearly has a bearing on binding energy, presumably due to the extensive hydrogen-bonded interactions thought to be present in the CYP2C9 active site when a typical substrate binds to the enzyme /25/. Although the appearance of both log P and log D terms is unusual, there is little correlation between these variables, and it is possible that they could be describing different aspects of the substrate binding affinity.

TABLE 4

CYP2C9 substrates

Compound	log P	pK_a	log $D_{7.4}$	M_r	log M_r	K_m (μM)	ΔG_{bind} (kcal mol^{-1})
1. **Phenytoin**	2.47	8.1[a]	2.39	252.27	2.4019	45	-6.1658
2. **Tolbutamide**	2.34	5.43[a]	0.37	270.35	2.4319	132	-5.5028
3. **Ibuprofen**	3.51	5.2[1]	1.31	206.28	2.3145	52	-6.0757
4. **Diclofenac**	4.40	4.22[a]	1.22	296.15	2.4715	6	-7.4070
5. **Warfarin**	2.52	5.1[a]	0.12	308.33	2.4890	4	-7.5568
6. **Tienilic acid**	3.15[2]	4.8[1]	0.55[c]	331.17	2.5201	6	-7.4070
7. **58C80**	5.18[2]	5.0[1]	2.78[2]	312.44	2.4948	141	-5.4622
8. **Naproxen**	3.18	4.15[a]	0.33	230.263	2.3622	125	-5.5315
9. **Piroxicam**	1.58	6.3[a]	0.45	331.345	2.5203	40	-6.2383
10. **Mefenamic acid**	5.12	4.2[1]	2.0	241.289	2.3825	7	-7.6568

a = acidic; c = calculated value.

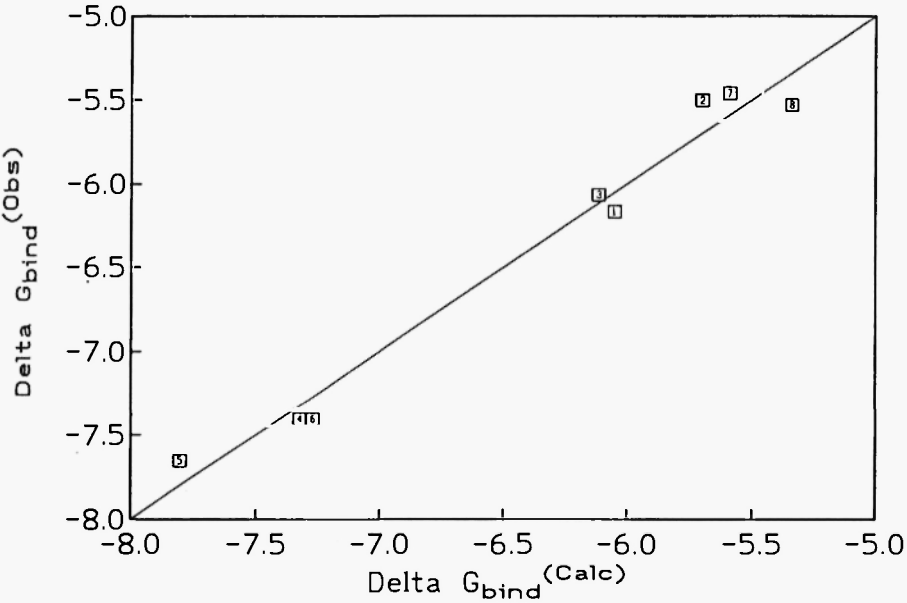

Fig. 3: A comparison between observed and calculated ΔG_{bind} values for several CYP2C9 substrates based on data presented in Tables 4 and 8. Points 1 to 8 refer to compounds 1 to 8, respectively, in Table 4.

4. CYP2C19 substrates

Table 5 presents information on CYP2C19 substrates and inhibitors in terms of binding affinities and physicochemical properties. The relevant QSAR equations governing binding and inhibition are shown in Table 8, Section 4, for 8 and 10 compounds, respectively. It can be appreciated from inspection of Equations 4a, b and c in Table 8 that similar descriptors are associated with both enzyme inhibition and substrate binding, although somewhat different groups of compounds were employed. These include relative molecular mass (M_r) and the number of hydrogen bond donors/acceptors (Equations 4b and c, Table 8) in the case of inhibition of CYP2C19, whereas for substrate

TABLE 5

CYP2C19 substrates

Compound	log P	pKa	log $D_{7.4}$	M_r	log M_r	K_i (μM)	pKi	K_m (μM)	ΔG_{bind} (kcal.mol^{-1})
1. **Hexobarbital**	1.49	8.2[b]	0.63	236.37	2.3734	N/A	N/A	740	-4.4409
2. **Moclobemide**	2.13	7.09[b]	1.96	268.77	2.4274	N/A	N/A	N/A	N/A
3. **R-Mephobarbital**	1.86	7.8[b]	1.31	246.26	2.3914	N/A	N/A	N/A	N/A
4. **S-Mephenytoin**	1.74[c]	8.1[b]	0.96	218.28	2.3390	60	4.2218	38	-6.2699
5. **Omeprazole**	2.23[d]	8.7[b]	0.91	345.45	2.5384	4.1	5.3872	8.6	-7.1852
6. **Proguanil**	2.53[e]	10.6[b]	-0.47	253.73	2.4044	N/A	N/A	96	-5.6990
7. **Propranolol**	3.37	9.5[b]	1.18	259.35	2.4139	112	3.9508	N/A	N/A
8. **Diazepam**	2.84	4.75[b]	2.86	284.745	2.4545	100	4.0000	20	-6.6653
9. **R-Warfarin**	2.52	5.1[a]	0.22	308.33	2.4890	32	4.49485	60	-5.9885
10. **Phenytoin**	2.47	8.1[a]	2.44	252.27	2.4019	280	3.5528	70	-5.8936
11. **LY307640**	1.79[c]	5.40[b]	1.79[c]	359.48	2.5557	9.2	5.0362	19	-6.6969
12. **Fluconazole**	-1.43[c]	2.03[b]	-1.43[c]	306.3	2.4861	2	5.6990	N/A	N/A
13. **Tranylcypromine**	1.52[c]	8.2[b]	0.66[c]	133.2	2.1245	8	5.0969	N/A	N/A
14. **Demethyldiazepam**	3.32[c]	4.58[b]	3.32[c]	270.73	2.4325	115	3.9393	N/A	N/A

a = acidic; b = basic; c = calculated value; d = experimental value at pH 6.4; e = experimental value at pH 13.0.
N/A = data not available.

binding affinity, the log P value becomes more important than the number of hydrogen bond acceptors (Equation 4a, Table 8). The presence of log M_r as a descriptor variable suggests that the loss of translational and rotational energy in the substrate molecule is an important factor in explaining either the binding or inhibition data. However, hydrogen-bonded interactions are of relevance to substrate binding within the CYP2C19 active site, as has been shown for derived molecular models of the enzyme-substrate complexes based on homology with the bacterial P450 enzyme CYP102 /25/. A comparison between experimental and QSAR-calculated values for substrate binding affinity is shown in Figure 4.

5. CYP2D6 substrates

Structural, physicochemical and biological data for selected CYP2D6 substrates are shown in Table 6. As provided in Section 5 of

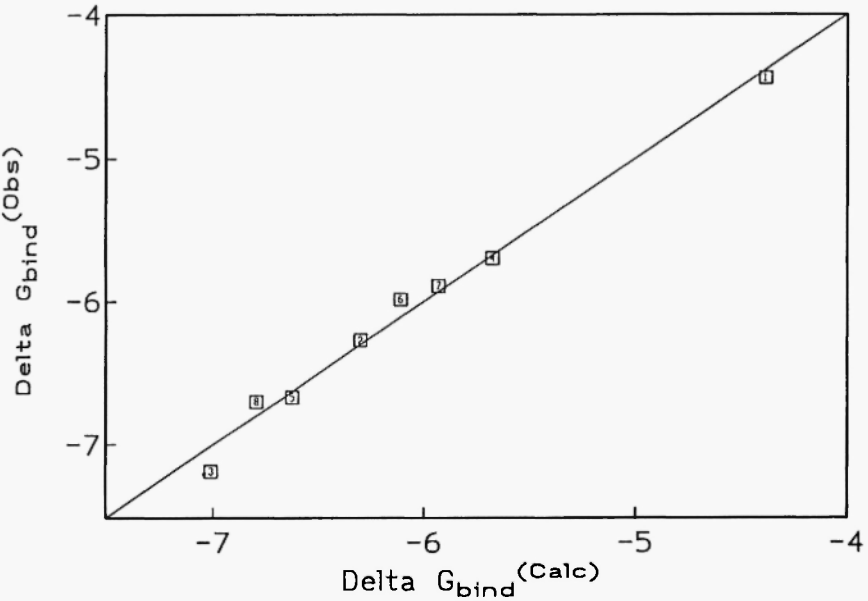

Fig. 4: A comparison between observed and calculated ΔG_{bind} values for several CYP2C19 substrates based on data presented in Tables 5 and 8. Points 1 to 8 refer to compounds 1, 4, 5, 6, 8 to 11, respectively, in Table 5.

TABLE 6

CYP2D6 substrates

Compound	log P	pK_a	log $D_{7.4}$	M_r	log M_r	K_D (μM)	ΔG_{bind} (kcal.mol^{-1})
1. **Bufuralol**	3.50	9.0[b]	1.89	261.36	2.4172	8.6	-6.9535
2. **Codeine**	1.07	8.2[b]	0.23	299.37	2.4762	15	-6.6218
3. **Ondansetron**	2.14	7.7[b]	1.30	293.40	2.4675	102	-5.4790
4. **Imipramine**	4.42	9.5[b]	2.52	280.41	2.4478	2.4	-7.7143
5. **Desipramine**	4.05	10.0[b]	1.45	266.39	2.4255	20	-6.4503
6. **Nortriptyline**	4.04	9.73[b]	1.71	263.38	2.4206	47	-5.9410
7. **Amitriptyline**	5.04	9.4[b]	2.50	263.38	2.4206	74	-5.6703
8. **Debrisoquine**	0.75	13.01[b]	0.75	175.33	2.2436	13	-6.7071
9. **Propranolol**	3.37	9.5[b]	1.18	259.35	2.4139	2.73	-7.6375
10. **Dextromethorphan**	3.36	8.3[b]	0.91	271.4	2.4336	2.76	-7.6310
11. **Metoprolol**	2.35	9.68[b]	0.07	267.4	2.4272	46 (K_S)	-6.4515
12. **MDMA**	2.28	10.04[b]	-0.27	193.27	2.2862	1.72 (K_m)	-7.913

b = basic.

MDMA = 3,4-methylenedioxymethylamphetamine.

Table 8, QSAR expressions governing inhibition and binding affinity towards the CYP2D6 enzyme include relative molecular mass, log P, hydrogen bond properties and π-stacking interactions within the putative active site of CYP2D6 /9/. For 10 compounds for which K_D values have been measured, there is a very good correlation (R = 0.98) between pK_i (Equation 5a, Table 8) and a combination of log P, M_r, log M_r, number of hydrogen bonds in the putative active site and number of π-stacking interactions between substrate and enzyme (Equation 5b, Table 8). Binding to CYP2D6 appears to be explained by a combination of hydrogen bond formation, π-π interaction, desolvation (M_r term) and loss in translational/rotational energy (log M_r term), whereas the number of basic nitrogen atoms in the molecule is of relevance to inhibition of the enzyme, together with a consideration of compound lipophilicity (log P term). Consequently, there is a satisfactory explanation (see Fig. 5 for a comparison between experimental and QSAR-calculated values) of the binding

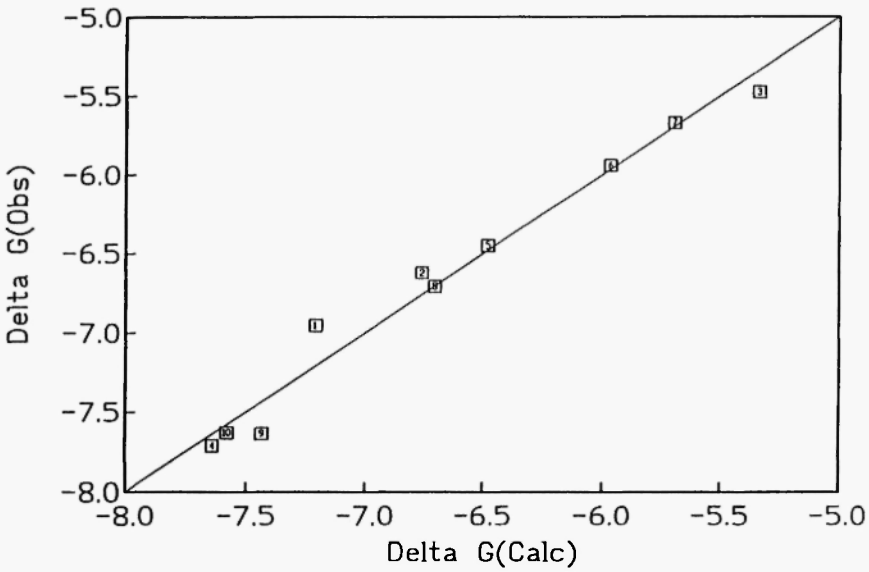

Fig. 5: A comparison between observed and calculated ΔG_{bind} values for several CYP2D6 substrates based on data presented in Tables 6 and 8. Points 1 to 10 refer to compounds 1 to 10, respectively, in Table 6.

affinity data (R = 0.94) for CYP2D6 substrates, although the correlation for inhibition is better (R = 0.98).

6. CYP3A4 substrates

The relevant biological and physicochemical data for several CYP3A4 substrates are presented in Table 7, and the corresponding QSAR expression governing binding affinity is provided in Table 8, Section 6. In this case, a four-variable expression correlates well (R = 0.96) with K_D-derived ΔG values; Figure 6 shows a plot which compares experimental and QSAR-derived binding affinity. The structural descriptors employed in this correlation include the frontier orbital energies (E_{HOMO} and E_{LUMO} values), the number of CYP3A4 active site hydrogen bonds and π-π stacking interactions between enzyme and substrate, as determined from molecular modelling of CYP3A4 /26/. It is noteworthy that two of these parameters relating to active site contacts also appear in the QSAR equation governing substrate binding to CYP2D6, whereas the number of hydrogen bonds formed between enzyme and substrate are important features in expressions for substrate binding affinity towards CYP1A2 and CYP2B6, as mentioned above.

CONCLUSIONS

The factors which contribute to the free energy of binding for substrates of human P450s involved in drug metabolism include, primarily, hydrogen bond interactions, frontier orbital energies and lipophilicities of the compounds themselves. However the QSAR expressions differ somewhat between the various enzymes involved due to specific properties of the P450 substrates themselves, such as the acidity of CYP2C9-selective compounds and the rectangularity of CYP1A2 substrates. The particular expressions presented in Table 8 indicate that binding to human P450s involves a different set of substrate properties in each case, despite the common occurrence of hydrogen bonding terms. These equations may be applicable for the estimation of binding affinities for novel development compounds, although it is also useful to consider the likely mode of substrate interaction within the particular P450 active site. It is hoped that

TABLE 7

CYP3A4 substrates (K_D data)

Compound	log P	pK_a	log $D_{7.4}$	M_r	log M_r	K_s (μM)	ΔG_{bind} (kcal mol^{-1})
1. Amitriptyline	5.04	9.4[b]	2.50	227.3	2.4429	348	-4.7474
2. Budesonide	3.28	neutral	3.28	430.6	2.6341	6.4	-7.1296
3. Etoposide	1.00	neutral	1.00	588.6	2.7698	23.2	-6.3618
4. Midazolam	1.53	neutral	1.53	325.8	2.5130	8.8	-6.9398
5. Phenacetin	1.57	neutral	1.57	179.24	2.2534	9.1	-6.9198
6. Verapamil	3.79	8.29[3]	2.85	454.6	2.6576	3.5	-7.4894
7. Ethynylestradiol	2.11	neutral	2.11	296.4	2.4719	6.5	-7.1204
8. Testosterone	3.32	neutral	3.32	288.43	2.4600	21.7	-6.4017
9. Erythromycin	2.48	8.8[3]	1.06	733.9	2.8656	1.8	-7.8858
10. Salmeterol	3.71	9.3[3]	1.80	415.6	2.6187	165	-5.1923

b = basic.

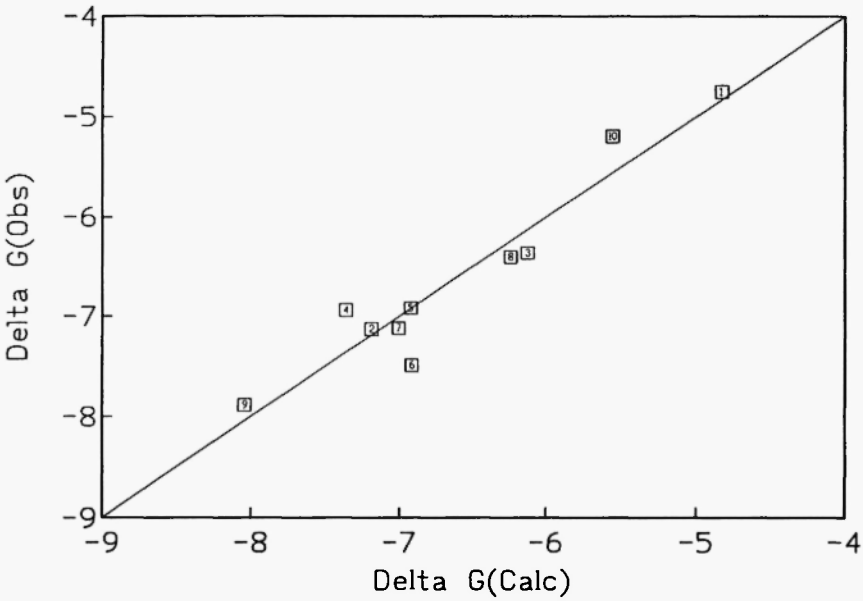

Fig. 6: A comparison between observed and calculated ΔG_{bind} values for several CYP3A4 substrates based on data presented in Tables 7 and 8. Points 1 to 10 refer to compounds 1 to 10, respectively, in Table 7.

further analyses of human P450 enzyme substrate interactions will elucidate additional features governing binding affinity and selectivity.

ACKNOWLEDGEMENTS

The financial support of GlaxoWellcome Research and Development Ltd., Merck, Sharp and Dohme Ltd. and the University of Surrey Foundation Fund is gratefully acknowledged by one of us (DFVL).

TABLE 8: QSARs for human P450 substrates and inhibitors

QSAR Equation	n	s	R	F
1. CYP1A2 substrates				
a) $\Delta G_{bind} = 0.63\mu\ (\pm 0.16) - 1.5\ l/w\ (\pm 0.41) + 2.32\ E_{-}\ (\pm 0.31) - 5.50$	11	0.474	0.95	22.47
b) $\Delta G_{bind} = 0.30\mu\ (\pm 0.12) - 0.89\ l/w\ (\pm 0.36) + 2.11\ \Delta E\ (\pm 0.26) - 0.50\ N_{HB}\ (\pm 0.17) - 22.41$	11	0.419	0.97	22.34
2. CYP2B6 substrates				
a) $\Delta G_{bind} = 3.98\ N_{HB}\ (\pm 0.76) - 5.41\ HB_D\ (\pm 0.86) - 1.92\ HB_A\ (\pm 0.23) - 4.19$	10	0.552	0.97	26.96
3. CYP2C9 substrates				
a) $\Delta G_{bind} = 9.59\ log\ D\ (\pm 0.83) - 9.00\ log\ P\ (\pm 0.82) - 7.05\ pK_a\ (\pm 0.65) + 0.58\ HB_D\ (\pm 0.15) + 49.08$	7	0.173	0.99	43.97
b) $\Delta G_{bind} = 8.62\ log\ D\ (\pm 0.82) - 8.02\ log\ P\ (\pm 0.79) - 6.26\ pK_a\ (\pm 0.62) + 0.57\ HB_D\ (\pm 0.19) + 42.74$	8	0.227	0.99	29.47
4. CYP2C19 substrates				
a) $\Delta G_{bind} = 201.69\ log\ M_r\ (\pm 19.87) - 0.31\ M_r\ (\pm 0.03) - 2.50\ log\ P\ (\pm 0.23) + 0.12\ HB_D\ (\pm 0.04) - 405.68$	8	0.146	0.99	56.16
b) $pK_i = 0.03\ M_r\ (\pm 0.01) - 18.76\ log\ M_r\ (\pm 4.07) + 0.35\ HB_A\ (\pm 0.07) + 40.62$	10	0.239	0.96	25.37
c) $pK_i = 0.03\ M_r\ (\pm 0.01) - 19.13\ log\ M_r\ (\pm 2.93) + 0.38\ HB_A\ (\pm 0.05) - 0.28\ HB_D\ (\pm 0.11) + 42.09$	10	0.172	0.98	38.42
5. CYP2D6 substrates				
a) $pK_i = 0.01\ M_r\ (\pm 0.002) - 0.48\ log\ P\ (\pm 0.13) - 0.57\ HB_A\ (\pm 0.15) + 1.79\ N_B\ (\pm 0.21) - 3.56$	10	0.187	0.98	35.17
b) $\Delta G_{bind} = 492.0\ log\ M_r\ (\pm 100.0) - 5.03\ N_{HB}\ (\pm 0.20) - 0.76\ N_{\pi\text{-}\pi}\ (\pm 0.82) - 0.88\ M_r\ (\pm 0.18) - 947.7$	10	0.373	0.94	9.61
6. CYP3A4 substrates				
a) $\Delta G_{bind} = 1.98\ E_{-}\ (\pm 0.45) - 2.98\ E_H\ (\pm 0.61) - 0.52\ N_{HB}\ (\pm 0.14) + 3.55\ N_{\pi\text{-}\pi}\ (\pm 0.71) - 35.10$	10	0.391	0.96	13.01

n = number of observations; s = standard error; R = correlation coefficient; F = variance ratio.

$\Delta G_{bind} = RT\ln K_m$ (or $RT\ln K_D$ in the case of Sb and 61 data); errors limits are shown in parentheses.

$pK_i = \log K_i$ where K_i is the concentration (μM) required for competitive inhibition of a selective marker substrate for the enzyme;

M_r = relative molecular mass; log P = logarithm of the octanol-water partition coefficient;

HB_A, HB_D = number of hydrogen bond acceptors or donors; $pK_a = -\log K_a$ where K_a is the acid dissociation constant;

log D = logarithm of the distribution coefficient at pH 7.4; N_{HB} = number of active site hydrogen bonds; N_B = number of basic nitrogens in the molecule;

$N\pi\text{-}\pi$ = number of active site π-π stacking interactions; μ = dipole moment; l/w = ratio of length to width;

$\Delta E = E_L - E_H$, where E_L and E_H are the LUMO and HOMO energies, respectively.

REFERENCES

1. Lewis DFV. Cytochromes P450: Structure, Function and Mechanism. London: Taylor & Francis, 1996.
2. Nelson DR. Metazoan cytochrome P450 evolution, Comp Biochem Physiol C 1998; 121: 15-22.
3. Nelson DR. Cytochrome P450 and the individuality of species. Arch Biochem Biophys 1999; 369: 1-10.
4. Rendic S, DiCarlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. Drug Metab Rev 1997; 29: 413-580.
5. Lewis DFV. Structural characteristics of human P450s involved in drug metabolism: QSARs and lipophilicity profiles. Toxicology 2000; 144: 197-203.
6. Lewis DFV. On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics. Biochem Pharmacol 2000; 60: 293-306.
7. Lewis DFV. The CYP2 family: models, mutants and interactions. Xenobiotica 1998; 28: 617-661.
8. Lewis DFV, Dickins M, Eddershaw PJ, Tarbit MH, Goldfarb PS. Cytochrome P450 substrate specificities, substrate structural templates and enzyme active site geometries. Drug Metab Drug Interact 1999; 15: 1-49.
9. Lewis DFV, Eddershaw PJ, Goldfarb PS, Tarbit MH. Molecular modelling of cytochrome P4502D6 (CYP2D6) based on an alignment with CYP102: structural studies on specific CYP2D6 substrate metabolism. Xenobiotica 1997; 27: 319-340.
10. Smith DA, Ackland MJ, Jones BC. Properties of cytochrome P450 isoenzymes and their substrates. Part 1: Active site characteristics. Drug Discovery Today 1997; 2: 406-414.
11. Smith DA, Ackland MJ, Jones BC. Properties of cytochrome P450 isoenzymes and their substrates. Part 2: Properties of cytochrome P450 substrates. Drug Discovery Today 1997; 2: 479-486.
12. Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH, Wrighton SA. Three and four-dimensional quantitative structure-activity relationship (3D/4D-QSAR) analyses of CYP2D6 inhibitors. Pharmacogenetics 1999; 9: 477-489.
13. Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH, Wrighton SA. Three-dimensional quantitative structure-activity relationship (3D-QSAR) analyses of inhibitors for CYP3A4, J Pharmacol Exp Ther 1999; 290: 429-438.
14. Ekins S, Bravi G, Ring BJ, Gillespie TA, Gillespie JS, Van den Branden M, Wrighton SA, Wikel JH. Three-dimensional quantitative structure-activity relationship (3D-QSAR) of substrates for CYP2B6. J Pharmacol Exp Ther 1999; 288: 21-29.
15. Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH, Wrighton SA. Three and four-dimensional quantitative structure-activity relationship (3D/

4D-QSAR) analyses of CYP2C9 inhibitors. Drug Metab Dispos 2000; 28: 994-1002.
16. Modi S, Paine MJ, Sutcliffe MJ, Lian LY, Primrose WH, Wolf CR, Roberts GC. A model for human cytochrome P450 2D6 based on homology modeling and NMR studies of substrate binding. Biochemistry 1996; 35: 4540-4550.
17. He S, Modi S, Bendall DS, Gray JC. The surface-exposed tyrosine residue Tyr83 of pea plastocyanin is involved in both binding and electron transfer reactions with cytochrome f. EMBO J 1991; 10: 4011-4016.
18. Lewis DFV, Eddershaw PJ, Dickins M, Tarbit MH, Goldfarb PS. Structural determinants of cytochrome P450 substrate specificity, binding affinity and catalytic rate. Chem-Biol Inter 1998; 115: 175-199.
19. Sangster J. Octanol-water partition coefficients of simple organic compounds, J Phys Chem Ref Data 1989; 18: 1111-1229.
20. Hansch C, Leo AJ. Substituent Constants for Correlation Analysis in Chemistry and Biology. New York: Wiley, 1979.
21. Rekker RF, ter Laak AM, Mannhold R. On the reliability of calculated log P values: Rekker, Hansch/Leo and Suzuki approach. Quant Struct-Act Relat 1993; 12: 152-157.
22. Suzuki T, Kudo Y. Automatic log P estimation based on combined additive modeling methods, J Comput Aided Molec Design 1990; 4: 155-198.
23. Lewis DFV, Lake BG, George SG, Dickins N, Eddershaw PJ, Tarbit MH, Beresford AP, Goldfarb PS, Guengerich FP. Molecular modelling of CYP1 family enzymes CYP1A1, CYP1A2, CYP1A6 and CYP1B1 based on sequence homology with CYP102. Toxicology 1999; 139: 53-79.
24. Lewis DFV, Lake BG, Dickins M, Eddershaw PJ, Tarbit MH, Goldfarb PS. Molecular modelling of CYP2B6, the human CYP2B isoform, by homology with the substrate-bound CYP102 crystal strcuture: evaluation of CYP2B6 substrate characteristics, the cytochrome b_5 binding site and comparisons with CYP2B1 and CYP2B4. Xenobiotica 1999; 29: 361-393.
25. Lewis DFV, Dickins M, Weaver RJ, Eddershaw PJ, Goldfarb PS, Tarbit MH. Molecular modelling of human CYP2C subfamily enzymes CYP2C9 and CYP2C19: rationalization in the CYP2C subfamily. Xenobiotica 1998; 28: 235-268.
26. Lewis DFV, Eddershaw PJ, Goldfarb PS, Tarbit MH. Molecular modelling of CYP3A4 from an alignment with CYP102: identification of key interactions between putative active site residues and CYP3A-specific chemicals. Xenobiotica 1996; 26: 1067-1086.
27. Edwards RJ, Adams DA, Watts PS, Davies DS, Boobis AR. Development of a comprehensive panel of antibodies against the major xenobiotic metabolizing forms of cytochromes P450 in humans. Biochem Pharmacol 1998; 56: 377-387.