Calculation of Lipophilicity of a Large, Diverse Dataset of Anticancer Platinum Complexes and the Relation to Cellular Uptake

Steven P. Oldfield,[†] Matthew D. Hall,[‡] and James A. Platts*,[†]

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, U.K., and Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received July 9, 2007

A quantitative structure−property relationship (QPSR) for the octanol−water partition of platinum complexes was constructed using molecular descriptors derived from density functional (DFT) calculations. A dataset of partition data for 64 complexes, consisting of 43 square-planar platinum(II) and 21 octahedral platinum(IV) complexes, was drawn from literature sources. Not only does this dataset include considerable structural diversity of complexes considered but also a variety of techniques for the measurement of partition coefficients. These data were modeled using descriptors drawn from electrostatic potentials and hardness/softness indices projected onto molecular surfaces. This required initial descriptor selection using a genetic algorithm approach, followed by partial least-squares regression against log $P_{o/w}$ data. In this way, a statistically robust model was constructed, with errors of similar size to the variation in log $P_{o/w}$ from multiple experimental measurements. Implications of lipophilicity for cellular accumulation of Pt-based drugs, and hence for design of new drugs, are discussed, as is the uptake of metabolites of cisplatin.

Introduction

Hydrophobicity or lipophilicity parameters have found extensive use as tools in the field of quantitative structure activity relationships (QSARs[a]) for numerous biological effects. Among other factors, hydrophobicity has been shown to affect absorption, transmembrane transport, bioavailability, cellular drug accumulation, hydrophobic drug-receptor interactions, metabolism, pharmacological activity, as well as the toxicity of molecules[1]. The lipophilicity of a molecule or drug has been related to its ability to cross cell membranes by means of passive diffusion and reflects the relative solubility of the drug in lipid-like (e.g., lipid bilayers of a cell membrane) and aqueous (the fluid in and out of cells) environments.

Following the seminal work of Hansch and Leo, lipophilicity has become almost synonymous with the *n*-octanol/water partition coefficient log $P_{o/w}$, where $P_{o/w}$ is the ratio of concentrations of a compound in *n*-octanol and in water. Traditionally, measurement of partition coefficients has been performed using the "shake flask" method in which a compound is dissolved in water, layered with an equivalent volume of *n*-octanol, and after a period of shaking, the concentration of drug in both phases is measured. Because log $P_{o/w}$ has become an important hydrophobic parameter, many experimental and theoretical methods have been developed to measure or calculate log $P_{o/w}$ for organic molecules. Nowadays, chromatography[2,3] is often used to replace slow and expensive shake flask methods. For instance, RP-HPLC retention parameters have been correlated to the hydrophobicity of compounds.[4]

Platinum complexes such as cisplatin and carboplatin (Figure 1) are well established as effective anticancer drugs. However, clinical challenges arising from toxicities and intrinsic and acquired resistance have prompted the search for novel analogues and the synthesis and testing of thousands of platinum complexes.[5,6] For platinum complexes, log $P_{o/w}$ is an important parameter related to the accumulation of the drug into cells; an exponential relationship between log $P_{o/w}$ and Pt accumulation in cells[7,8] has been reported previously. This is a significant observation, particularly given the reduced cellular accumulation associated with acquired resistance of clinical platinum drugs. Platinum(II) complexes are generally hydrophilic, with generally negative log $P_{o/w}$ values. However, addition of large nonpolar ligands such as 1,2-diaminocyclohexane (dach) can increase lipophilicity and hence uptake into cells. The platinum(IV) drug satraplatin (Figure 1, **22**), currently in clinical trials, was synthesized as a lipophilic platinum drug with two acetate axial ligands and retains activity in cisplatin resistant cell lines in part due to its increased cellular accumulation.[9,10]

Although log $P_{o/w}$ of platinum complexes can be measured via HPLC methods, the experimental process is still relatively time-consuming and expensive, limiting its use in screening as part of a drug design procedure, and prediction of platinum drug lipophilicity is useful for both retrospective analysis of data and for predicting the features of drug design candidates. Thus, reliable estimations of log $P_{o/w}$ based on chemical structure are of considerable interest. Hansch and co-workers[11] set out a group-additive method for calculating log $P_{o/w}$, quantified by a substituent constant, so that features of compounds could be predicted prior to their syntheses. Subsequently, many groups have established similar group contribution methods for estimation of log $P_{o/w}$; these work well when all group constants are known, but unfortunately, these are not available for many ligands of metal complexes, including platinum. In addition, the polarizing effect of the metal center alters the properties and interactions of organic ligands,[12] such that their properties cannot be simply transferred to metal complexes.

* To whom correspondence should be addressed. Tel: +44 2920 874950. Fax: +44 2920 874030. E-mail: plats@cardiff.ac.uk.

[†] Cardiff University.

[‡] National Institutes of Health.

[a] Abbreviations: QSPR, quantitative structural property relationship; QSAR, quantitative structure−activity relationship; DFT, density functional theory; $P_{o/w}$, *n*-octanol−water partition coefficient; HPLC, high performance liquid chromatography; RP-HPLC, reverse phase high performance liquid chromatography; dach, 1,2-diaminocyclohexane; $V(\mathbf{r})$, molecular electrostatic potential; ESP, electrostatic potential; IP, ionization potential; MLRA, multilinear regression analysis; PLS, partial least-squares; GA, genetic algorithm; RMSE, root-mean-square error; PRMSE, prediction root-mean-square error; VIP, variable importance parameter; GFAAS, graphite furnace atomic absorption spectroscopy; ICP-OES, inductively coupled plasma-optical emission spectroscopy; GS-X, glutathione conjugate

10.1021/jm0708275 CCC: $37.00
Published on Web 09/27/2007

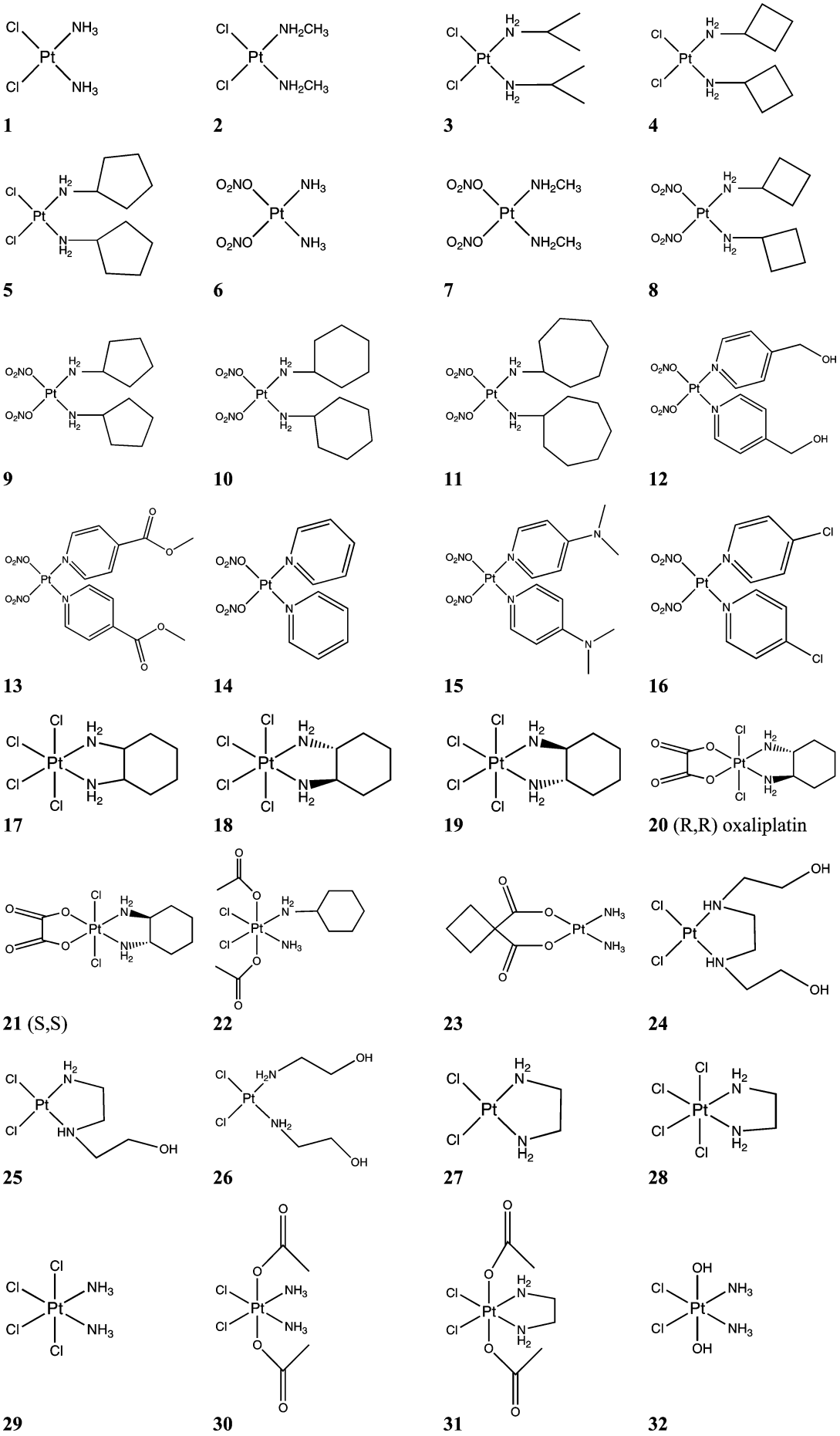


Figure 1. Structures of complexes included in the dataset.

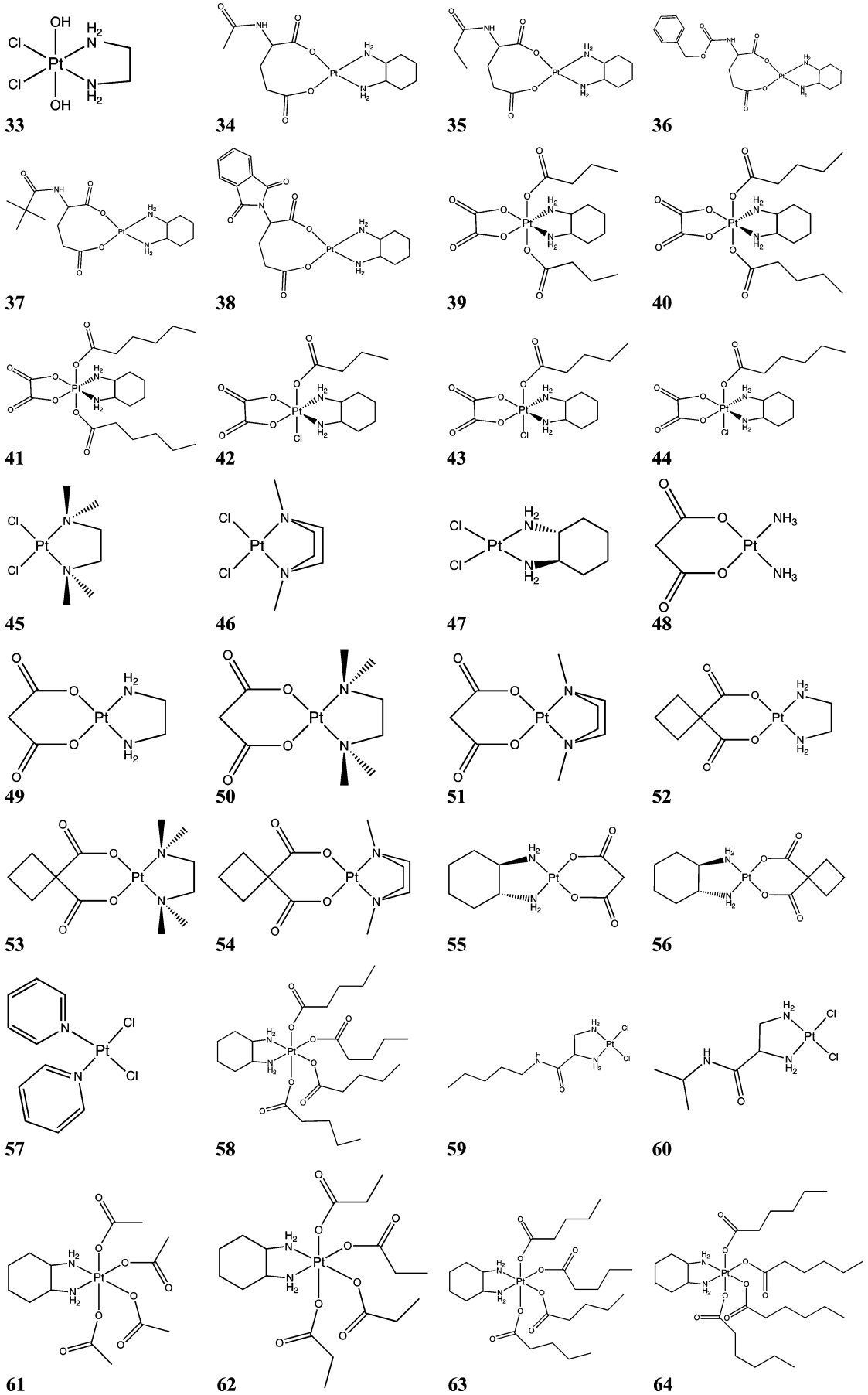


Figure 1. Continued.

Instead, we turn to quantum mechanical calculations, the results of which can be condensed into three-dimensional (3D) descriptors for a whole molecule and used to construct biologically relevant quantitative structure−property relationship (QSPR) models. The molecular electrostatic potential $V(\mathbf{r})$ created around molecules has been shown by Politzer[13] to describe the ability to form noncovalent interactions, which in turn determine a molecule's log $P_{o/w}$. $V(\mathbf{r})$ is given by the sum of nuclear and electronic contributions

$$V(\mathbf{r}) = \sum_A \frac{Z_A}{|R_A - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|} \quad (1)$$

where Z_A is the charge on nucleus A, located at R_A, and $\rho(\mathbf{r})$ is the electron density. The sign of $V(\mathbf{r})$ at position $\mathbf{r}$ depends upon whether the effects of the nuclei or the electrons are dominant.

Politzer has shown that molecular descriptors based on the electrostatic potential measured on isosurfaces of the electron density, $V_S(\mathbf{r})$, can be used to describe and predict hydrogen bonding and related phenomena,[14] including those in biological systems.[15] These descriptors model "hard" Coulombic interactions, long-range forces that often dominate intermolecular interactions. It has been suggested that "softer" short-range forces can also be represented by local surface properties. QSPR descriptors seeking to model these rely, for instance, on the local ionization energy[16] or the local electron affinity,[17] which are well suited for the prediction of chemical reactivity since they give information about donor/acceptor interactions.[18] The local ionization energy may be defined as

$$\bar{I}(\mathbf{r}) = \sum_i^{\text{occMO}} \frac{\rho_i(\mathbf{r})|\epsilon_i|}{\rho(\mathbf{r})} \quad (2)$$

where $\rho_i(\mathbf{r})$ are the electronic densities of the occupied molecular orbitals with eigenvalues ϵ_i. An alternative definition of local hardness was proposed by Geerlings et al. (DFT) based on the electronic-only contribution to the electrostatic potential, V_{el}, scaled by the total number of electrons in the molecule, N.[19−21] This definition was tested in this work but found to be inferior for our purposes to the local ionization potential, and therefore, it was not pursued any further.

Several years ago, we reported[22] one of the first attempts to construct a QSPR for log $P_{o/w}$, based on 23 compounds whose experimental log $P_{o/w}$ values had been previously reported, using molecular descriptors based around polar surface areas. This yielded an R^2 value of 0.90 and a standard deviation of 0.35 log units. Subsequently, we reported[23] a QSPR model for 24 Pt(II) complexes (complexes **1**−**5**, **20**, **23**, **25**, **27**, **47**−**59**, **61**, and **62** in Table 1 and Figure 1), whose log $P_{o/w}$ values were determined using reverse-phase high-performance liquid chromatography (RP-HPLC). This model was based on four molecular descriptors derived from DFT,[24] including surface electrostatic and surface ionization potentials, $V_S(\mathbf{r})$, and $\bar{I}(\mathbf{r})$, as well as the dipole moment and polarizability. This model was found to correlate log $P_{o/w}$ values with an R^2 value of 0.95 and an rms error of 0.21 log units, that is, significantly improved over our first paper. This was, to our knowledge, the first attempt to use descriptors derived from quantum chemical calculations to predict log $P_{o/w}$ for platinum complexes. Despite the promise of this approach, however, the structural variety of the complexes considered was somewhat limited. In this work, therefore, we attempt to extend this methodology to a much larger and more structurally diverse dataset. In addition, we also attempted to combine log $P_{o/w}$ data from a variety of literature sources based on different measurement protocols, which should provide a greater challenge to the QSPR methodology and ultimately a more reliable and generally applicable predictive model.

Results

The dataset was collected from literature data published between 1991 and 2006[4,23,25−31] and contains 64 platinum complexes with associated experimentally determined log $P_{o/w}$ values. Most data stem from shake flask methods, some using saline solution to prevent equilibration into aquated species, and drug detection employed atomic absorption spectroscopy. A significant minority of data is from RP-HPLC studies. In most cases, the log $P_{o/w}$ values were provided in the literature; however, values in Kidani et al.[30] were converted to log $P_{o/w}$ from the $P_{o/w}$ data provided. While some earlier studies report the partition coefficient between chloroform and water,[32] these were not incorporated into the dataset. Enantiomers of the same complex with different experimental values were treated as separate data points with the same set of descriptors. The complexes studied are shown in Figure 1 and Table 1, also giving the original literature reference of the experimental log $P_{o/w}$ measurements. Of the 64 compounds selected, 21 are platinum(IV) complexes, and while not all conform to classical Pt structure−activity rules, all complexes reported contained *cis*-am(m)ine ligands—a fundamental requirement for activity in conventional platinum drug design.[5]

Three molecules within this dataset have multiple literature values, that is, cisplatin (four values reported), oxaliplatin (three values), and carboplatin (three values). Several other molecules have two data values. In such cases, log $P_{o/w}$ values were transformed into the corresponding $P_{o/w}$ data before averaging, and the log of this average was taken as a single value in the QSPR. For the three molecules mentioned above, the availability of multiple data from different laboratories and/or methods allows us to monitor the reproducibility of log $P_{o/w}$ data. For cisplatin and oxaliplatin, a range of approximately ±0.15 log units results from this analysis, but for carboplatin, a larger spread of ±0.44 log units is observed. These values provide a useful context for the statistical accuracy of any QSPR models made from this data.

Using the same set of descriptors as those in our previous study of 24 Pt(II) complexes,[23] a model constructed using multiple linear regression analysis (MLRA) gives $R^2 = 0.767$ and RMSE = 0.671 for all 64 complexes, that is, considerably worse performance results from application to this larger dataset. This could be ascribed to the much greater structural diversity of the current, larger dataset or may stem from the inherent variability of values derived from several sources using different experimental methods. In order to test these hypotheses, a more sophisticated statistical analysis is required, using a better combination of descriptors to incorporate the wider range of physical characteristics present in the more diverse dataset.

In total, we calculated 39 descriptors for each molecule in the dataset (see Experimental Section), many of which were highly intercorrelated. Because of this, conventional MLRA is not an appropriate method to use due to its sensitivity to intercorrelation. Partial least-squares (PLS) offers a more statistically robust method of handling descriptors that are correlated, balancing the two objectives of fitting to response data (log $P_{o/w}$ in this case), and predicting the external data, based on the assumption that directions in the predictor space that are well sampled should provide a better prediction for new observations when the predictors are highly correlated. This focus on prediction, rather than fitting, matches our aim of producing a model to predict log $P_{o/w}$ for new complexes.

Table 1. Log $P_{o/w}$ of Platinum Complexes[a]

structure no.	chemical name	log $P_{o/w}$ (average)	primary source[b]
1	*cis*-[$PtCl_2(NH_3)_2$] [cisplatin]	−2.30 (−2.19, −2.53, −2.28, −2.27)	a, b, d, h
2	*cis*-[$PtCl_2(NH_2CH_3)_2$]	−1.80 (−1.68, −1.97)	a, h
3	*cis*-[$PtCl_2(NH_2CH(CH_3)_2)_2$]	−0.44 (−0.32, −0.61)	a, h
4	*cis*-[$PtCl_2(NH_2Bu^c)_2$]	0.25 (0.36, 0.09)	a, h
5	*cis*-[$PtCl_2(NH_2Pe^c)_2$]	0.95 (0.81, 1.06)	a, h
6	*cis*-[$Pt(NO_3)_2(NH_3)_2$]	−3.36	a
7	*cis*-[$Pt(NO_3)_2(NH_2CH_3)_2$]	−3.28	a
8	*cis*-[$Pt(NO_3)_2(NH_2Bu^c)_2$]	−1.71	a
9	*cis*-[$Pt(NO_3)_2(NH_2Pe^c)_2$]	−1.14	a
10	*cis*-[$Pt(NO_3)_2(cha)_2$]	−0.91	a
11	*cis*-[$Pt(NO_3)_2(NH_2Hp^c)_2$]	−0.35	a
12	*cis*-[$Pt(NO_3)_2(4\text{-}HOCH_2Py)_2$]	−2.13	a
13	*cis*-[$Pt(NO_3)_2(4\text{-}CH_3COOPy)_2$]	−1.41	a
14	*cis*-[$Pt(NO_3)_2(Py)_2$]	−1.59	a
15	*cis*-[$Pt(NO_3)_2(4\text{-}(CH_2)_2NPy)_2$]	−0.83	a
16	*cis*-[$Pt(NO_3)_2(4\text{-}ClPy)_2$]	−1.06	a
17	*rac*-[$PtCl_4$(dach)] [ormaplatin]	−1.17	b
18	*R,R*-[$PtCl_4$(dach)]	−1.18	b
19	*S,S*-[$PtCl_4$(dach)]	−1.03	b
20	*R,R*-[Pt(ox)(dach)] [oxaliplatin]	−1.54 (−1.65, −1.64, −1.39)	b, g, h
21	*S,S*-[Pt(ox)(dach)]	−1.59	b
22	*cis*−*trans*-[$Pt_2Cl_2(OCOCH_3)_2(NH_3)$(cha)] [JM216, satraplatin]	0.10 (−0.16, 0.26)	b, g
23	[$Pt(CBDC)(NH_3)_2$] [carboplatin]	−1.64 (−2.30, −1.398, −1.63)	b, e, h
24	[$PtCl_2$(*N,N*′-$(C_2H_5OH)_2$-en)]	−2.136	c
25	[$PtCl_2$(*N*-(C_2H_5OH)-en)]	−2.16 (−2.214, −2.12)	c, h
26	*cis*-[$PtCl_2(NH_2C_2H_5OH)_2$]	−2.223	c
27	[$PtCl_2$(en)]	−2.22 (−2.30, −2.16)	d, h
28	[$PtCl_4$(en)]	−1.91	d
29	*cis*-[$PtCl_4(NH_3)_2$]	−2.06	d
30	*cis*,*trans*,*cis*-[$PtCl_2(OCOCH_3)_2(NH_3)_2$]	−2.20	d
31	*cis*,*trans*-[$PtCl_2(OCOCH_3)_2$(en)]	−1.88	d
32	*cis*,*trans*,*cis*-[$PtCl_2(OH)_2(NH_3)_2$]	−2.81	d
33	*cis*,*trans*-[$PtCl_2(OH)_2$(en)]	−2.78	d
34	[Pt(acetyl-Glu)(dach)]	−0.903	e
35	[Pt(propionly-Glu)(dach)]	−0.453	e
36	[Pt(carbobenzyloxy-Glu)(dach)]	1.197	e
37	[Pt(pivaloyl-Glu)(dach)]	0.179	e
38	[Pt(phthaloyl-Glu)(dach)]	2.067	e
39	*trans*,*cis*,*cis*-[$Pt(OCOC_3H_7)_2$(ox)(dach)]	0	g
40	*trans*,*cis*,*cis*-[$Pt(OCOC_4H_9)_2$(ox)(dach)]	0.93	g
41	*trans*,*cis*,*cis*-[$Pt(OCOC_5H_{11})_2$(ox)(dach)]	2.00	g
42	*trans*,*cis*,*cis*-[$Pt(OCOC_3H_7)Cl$(ox)(dach)]	−1.25	g
43	*trans*,*cis*,*cis*-[$Pt(OCOC_4H_9)Cl$(ox)(dach)]	−0.62	g
44	*trans*,*cis*,*cis*-[$Pt(OCOC_5H_{11})Cl$(ox)(dach)]	−0.076	g
45	[$PtCl_2$(tmen)]	−0.85	h
46	[$PtCl_2$(dmppz)]	−1.23	h
47	[$PtCl_2$(dach)]	−1.40	h
48	[Pt(mal)$(NH_3)_2$]	−2.32	h
49	[Pt(mal)(en)]	−2.19	h
50	[Pt(mal)(tmen)]	−1.17	h
51	[Pt(mal)(dmppz)]	−1.47	h
52	[Pt(CBCD)(en)]	−1.70	h
53	[Pt(CBDC)(tmen)]	−0.47	h
54	[Pt(CBDC)(dmppz)]	−0.79	h
55	[Pt(mal)(dach)]	−1.37	h
56	[Pt(CBDC)(dach)]	−0.85	h
57	*cis*-[$PtCl_2(Py)_2$]	−0.04	h
58	[Pt(dach)$(OCOC_5H_{11})_4$]	3.03	i
59	[$PtCl_2$(en$CONHC_5H_{11}$)]	−0.04	h
60	[$PtCl_2$(en$CONHCH(CH_3)_2$))]	−0.46	h
61	[$Pt(OCOCH_3)_4$(dach)]	−1.59	i
62	[$Pt(OCOC_2H_5)_4$(dach)]	0.18	i
63	[$Pt(OCOC_3H_7)_4$(dach)]	1.54	i
64	[$Pt(OCOC_4H_9)_4$(dach)]	2.54	i

[a] Abbreviations: Bu^c, cyclobutyl; CBDC, 1,1′-cyclobutanedicarboxylato; cha, cyclohexylamine; dach, 1,2-diaminocyclohexane; dmppz, *N,N*′-dimethylpiperazine; en, ethane-1,2-diamine; Glu, L-glutamate; Hp^c, cylcoheptyl; Hx^c, cyclohexyl; mal, malonato; ox, oxalato; Pe^c, cyclopentyl; Py, pyridine; tmen, *N,N,N*′,*N*′-tetramethylethane-1,2-diamine. [b] References: (a) Ref 25; (b) ref 4; (c) ref 26; (d) ref 27; (e) ref 28; (f) ref 29; (g) ref 30; (h) ref 23; (i) ref 31.

It is known that the performance of PLS degrades if descriptors not relevant to the property of interest are employed, that is, there is an optimal set of input descriptors for a given PLS model. In order to select these, we employed a genetic algorithm approach, in which populations of descriptor sets were constructed randomly and their "fitness" to explain log $P_{o/w}$ via PLS was tested. The fittest populations were retained and allowed to mutate and interchange the descriptors included, until an optimum set was reached.

This combination of genetic algorithm descriptor selection and partial least-squares regression (GA-PLS) was therefore employed for all 64 complexes, selecting DFT-based descriptors

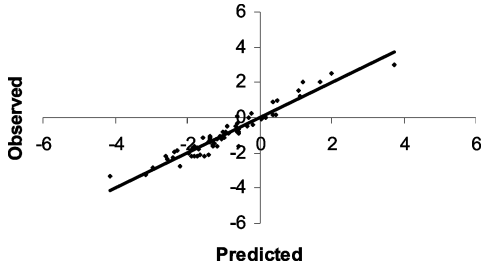


Figure 2. Observed versus predicted log $P_{o/w}$ values from eq 4, incorporating a line of best fit with $R^2 = 0.92$

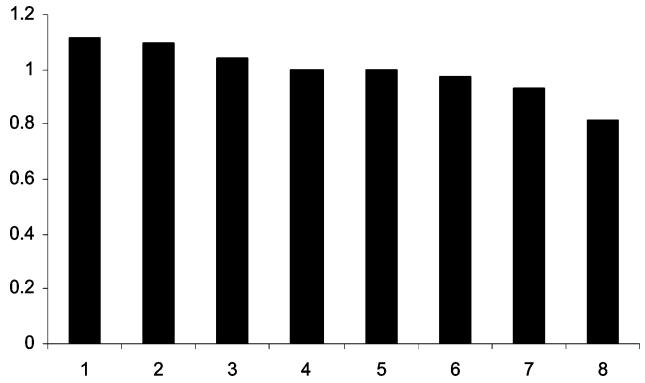


Figure 3. Variable importance values for the best GA-PLS model.

as required from the pool of 39. The best model found in this manner was based on eight descriptors: polarizability; the number of ESP local minima; the IP minima and variance; along with the number and sums of the IP local maxima and minima. Seven latent vectors were used, giving the final eq 3 and statistics shown below. Randomization tests were carried out 500 times, yielding intercepts for $R^2 = 0.04$ and for $Q^2 = -0.08$, indicating that overfitting has not occurred in this case

$$\log P_{o/w} = 1.102 + 0.00550\alpha - 0.0497N[V_S^-] - 11.936I_{S,min} + 0.654N[I_S^+] - 1.461\Sigma I_S^+ + 0.184N[I_S^-] - 0.363\Sigma I_S^- - 183.128\sigma^2(I_S)$$

$$n = 64 \quad R^2 = 0.919 \quad \text{RMSE} = 0.386 \quad Q^2 = 0.871 \qquad (3)$$

Descriptors based on local hardness were included in the GA procedure but did not appear in the best models found using GA-PLS. All such descriptors constructed were closely related to polarizability and volume descriptors, suggesting they represent similar physical properties. The maximum hardness value was alone in having no cross-correlations with other descriptors considered but was not significant in the PLS models of log $P_{o/w}$. Figure 2 shows the observed against predicted log $P_{o/w}$ using the model above and shows good accuracy across the full range of observed values traversing more than 6 log units. The average log $P_{o/w}$ value for all complexes in the set is negative (mean $= -0.92$), as one might expect for hydrophilic complexes, though clearly complexes with lipophilicity similar to that of organic molecules and drugs can be achieved.

Since the descriptors used in constructing eq 3 are not on the same scale, coefficients cannot be directly compared to evaluate which properties dominate the model. Instead, we calculated the variable importance parameter (VIP) of each descriptor in the final model. These are reported in Figure 3 and Table 2. This analysis reveals that polarizability is the dominant descriptor, as was found to be the case in our previous work.[23] The positive coefficient for polarizability shows that more polarizable compounds prefer octanol, as might be anticipated on the basis of the physical properties of water and octanol. Polarizability is also closely correlated to molecular size, thereby reflecting the relative ease of forming cavities in octanol compared to water. Other measures of molecular size, such as the volume enclosed by the 0.001 au isodensity surface, were found to be less significant than polarizability in the GA-PLS model.

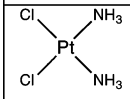
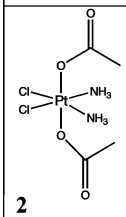
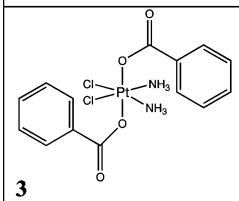
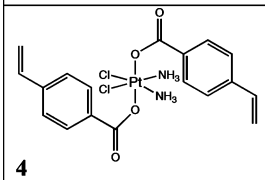
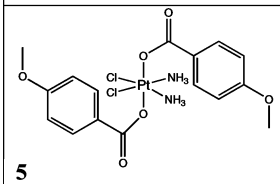
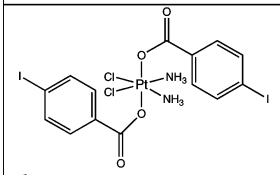
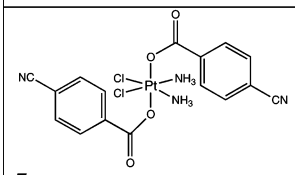
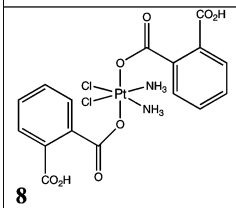
Table 2. Variable Importance for the Best Ga-PLS Model

	descriptor	VIP
1	α	1.11259
2	$I_{S,min}$	1.09882
3	ΣI_S^+	1.04137
4	$N[I_S^-]$	0.998374
5	$N[I_S^+]$	0.995512
6	$N[V_S^-]$	0.97517
7	ΣI_S^-	0.932499
8	$\sigma^2(I_S)$	0.813634

The number of minima on the surface electrostatic potential, $N[V_S^-]$, is in essence a measure of the hydrogen bond basicity of a molecule. Its importance is in agreement with the known physical properties of the octanol/water system since water is a more Lewis acidic medium than octanol. V_S^+ and related descriptors for H-bond acidity are not featured in any model, which can be attributed to the similarity in the hydrogen bond basicity of water and octanol. That the number of such minima, rather than their depth, emerges as a significant descriptor from the GA-PLS procedure is somewhat surprising. This may reflect the importance of multiple hydrogen bonds, that is, several water or octanol molecules surrounding a single complex, which is likely to be particularly relevant to larger complexes. Several descriptors based on the surface ionization potential were found to be significant in this QSAR, including the global minimum IP, along with the number and sum of minima and maxima on the surface. The significance of these terms was assigned[23] to the "harder" nature of water over octanol, leading to a lower log $P_{o/w}$ for harder molecules.

Of the 64 complexes studied, 21 are octahedral platinum-(IV), and 43 are square-planar platinum(II) complexes, or roughly 1:2, reflecting the Pt(II) compounds' greater heritage as potential successors to cisplatin. The overall rms error of 0.39 for all data can be decomposed into values of 0.35 for the 43 Pt(II) and 0.43 for the 21 Pt(IV) complexes, that is, good results for both coordination types, with slightly better performance for Pt(II). This may reflect the greater structural diversity of Pt(IV) complexes, or it may simply be due to the greater number of Pt(II) complexes in the dataset used to build the model. It is also possible to develop separate QSPR models for Pt(II) and Pt(IV) complexes, although the relatively small number of Pt(IV) data hinders this somewhat. Stepwise linear regression for the Pt(II) data shows that a QSPR model similar to that previously published for a subset of this data, incorporating polarizability, dipole moment, $V_{S,Min}$, and $I_{S,Max}$ performs reasonably well, with $R^2 = 0.85$ and RMSE $= 0.46$. For the Pt(IV) complexes, it is interesting to note that polarizability alone correlates with log $P_{o/w}$, with $R^2 = 0.90$ and RMSE $=$ 0.55. The dipole moment is not significant here, perhaps reflecting the higher symmetry of these complexes, while adding V_S and I_S terms provides only minor improvements, $R^2 = 0.93$ and RMSE $= 0.47$.

Table 3. Cellular Accumulation Values Reported for Complexes **1**−**8** in Ang et al.[37] and Calculated Log $P_{o/w}$ Values for the Complexes; Complexes **2**−**8** are Platinum(IV) Complexes Possessing a Common Equatorial Moiety (Based on Cisplatin, **1**) and Varying in the Axial Ligands

Complex	Pt uptake[a]		log $P_{o/w}$	
	A549	HT29	Axial ligand [a]	Complex [b]
 cisplatin, **1**	-	0.01	-	-2.4
 **2**	0.35	0.01	-	-1.7
 **3**	2.27	2.27	1.87	0.46
 **4**	5.88	3.85	2.61	2.45
 **5**	3.23	6.14	3.02	0.74
 **6**	3.16	3.75	1.96	2.45
 **7**	5.93	3.62	1.56	0.70
 **8**	0.10	0.02	0.70	0.61

[a] Experimental data from Ang et al.[37] [b] As predicted from eq 3.

Discussion

The cellular accumulation of platinum drugs is believed to be dependent on both passive diffusion and active transport. Active transport can involve mediated entry into the cell via plasma membrane transporters such as the copper transporter CTR1 (SLC33A1) and organic cation transporters 1 and 2 (SLC22A1 and SLC22A2), as well as active efflux of the drug by the copper efflux transporters ATP7A and ATP7B or efflux of Pt−glutathione conjugates by the GS-X efflux transporters.[33] The relative importance of passive versus active transport

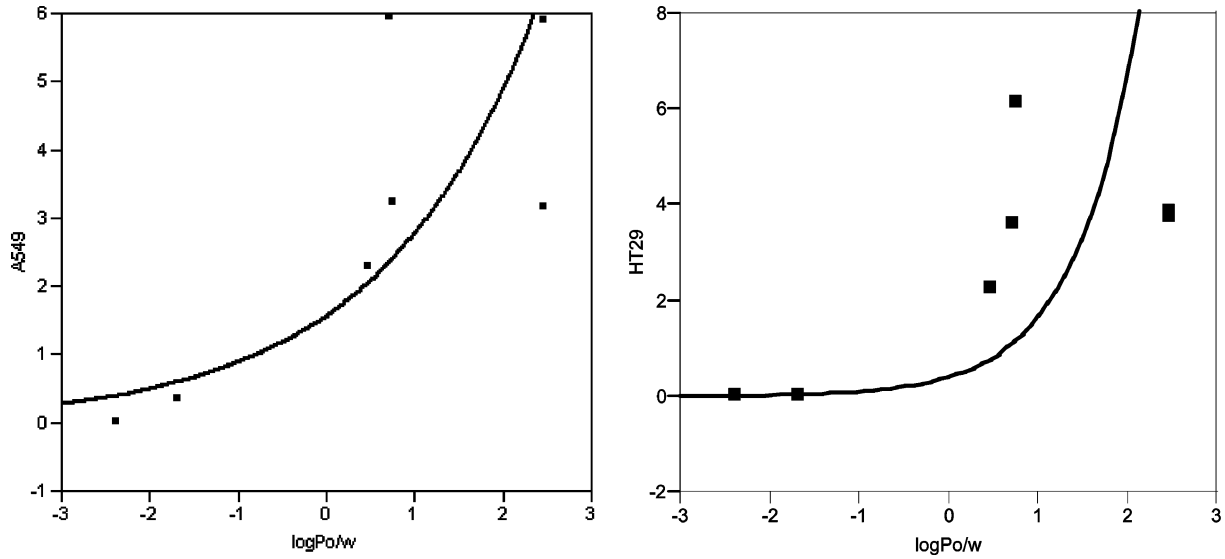


Figure 4. Plot of calculated log $P_{o/w}$ for complexes **3**−**7** in Table 1 versus the reported cellular accumulation in A549 and HT29 cells. Values for complex **8** are excluded; see text for details.

pathways for drug uptake is yet to be determined but is probably cell-type specific. Platinum drugs display the hallmarks of passive diffusion: linear, nonsaturable uptake kinetics, a lack of competitive uptake inhibition by analogues, and accumulation that correlates with lipophilicity.[34] Yet, uptake can be lowered with metabolic inhibitors, which is indicative of energy-dependent uptake, and most importantly, cell lines that are selected for resistance to cisplatin demonstrate lowered accumulation, probably due to loss of the energy-dependent transport function, as little change in the cell membrane composition has been reported in resistant cells.[34]

Given the loss of accumulation due to the active transport function in resistant cells, complexes with increased lipophilicity may increase passage across the lipid bilayer without relying on active/facilitated transport, potentially circumventing resistance. As such, an understanding of the potential for platinum-derived chemotherapeutics to diffuse across lipid bilayers is paramount. For instance, satraplatin was developed as an orally available lipophilic platinum(IV) drug and was found to circumvent cisplatin resistance in cells due to increased cellular accumulation.[35] While much information exists in the literature on the cytotoxicity and DNA binding (the key cytotoxic mechanism of action) of platinum complexes, there is frustratingly little comparative data on cellular accumulation of diverse platinum complexes.[5,33] This is probably partly due to the techniques required for platinum drug accumulation (GFAAS or ICP-OES) not regularly being present in cell biological laboratories. While one may find in the literature a number of papers comparing accumulation of one or two novel compounds to cisplatin, the range of cell lines, conditions, and concentration or time dependence of the accumulation studies render comparison among studies difficult at best, and no accepted standards for drug accumulation assays exist in the field that might facilitate this.

A small number of publications have been identified by us where accumulation for four or more agents in a single cell line is reported. We have previously demonstrated[8] an exponential relationship between lipophilicity and the uptake of five complexes reported by Loh et al.[36] in 41M human ovarian cancer cells, and a similar exponential relationship ($R^2 = 0.94$) is found using the method for calculation of log $P_{o/w}$ set out in this work. For eight complexes with a narrow range of experimentally determined log $P_{o/w}$ values (−2.8 to −1.8), one of us showed a positive correlation between log $P_{o/w}$ in A2780 ovarian cancer cells and A2780cisR cisplatin resistant cells (accumulation being lower in the resistant line).[27]

Dyson and co-workers[37] recently reported a series of platinum(IV) complexes with substituted benzoates coordinated in the axial positions and examined their accumulation in A549 lung and HT29 colon carcinoma cell lines. In order to rationalize their activity, log $P_{o/w}$ values for the corresponding benzoic acids were reported, but no clear relation to accumulation was observed (Table 3). We have revisited this data using the method refined above for prediction of log $P_{o/w}$ values for platinum complexes. This approach has the key advantage that log $P_{o/w}$ is predicted for the entire complex rather than for a single ligand, thus allowing us to compare structurally unrelated molecules and, hence, to explore relationships between log $P_{o/w}$ and accumulation in different cell lines. This is of particular interest as cell lines express a range of influx and efflux transporters that complement passive diffusion into the cell. Figure 4 plots accumulation into two different cell types against log $P_{o/w}$, along with an exponential line of best fit. These data further support our previous finding of an exponential relation between lipophilicity and accumulation in cells, with R^2 values of 0.681 and 0.799 for A549 and HT29, respectively.

Comparison of log $P_{o/w}$ values of the axial benzoate ligands incorporated in **3**−**8** (Table 3) with predicted values for the entire complexes reveals little or no relationship between these two estimates of lipophilicity ($R^2 = 0.238$). This may be due to differing ligand orientation among the complexes resulting in varying ligand surface exposure to solvent, suggesting that simply incorporating a lipophilic moiety into a platinum complex does not necessarily confer increased lipophilicity. Given that cellular accumulation is a key determinant of cellular sensitivity, we suggest that prediction of log $P_{o/w}$ for new complexes should be incorporated into presynthetic drug design strategies for platinum complexes.

While complex **7** has a similar log $P_{o/w}$ value to that of **8** (0.70 and 0.61, respectively), **8** demonstrates poor cellular accumulation (Table 3). This is probably due to the pendant acetic acid moieties present in **8** being predominately deprotonated at physiological pH, giving singly and doubly negatively charged species. While at equilibrium some of **8** would remain uncharged, this amount is clearly low; compound **8** has

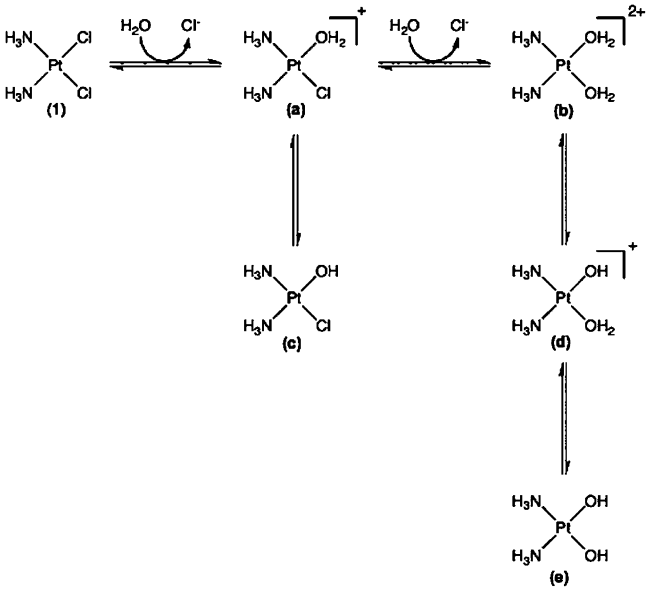


Figure 5. Aquation scheme for cisplatin in water, showing singly (**a**) and doubly (**b**) aquated species and their hydrolysis products. Note that the chlorohydroxo species (**c**) is neutral.

a calculated lipophilicity similar to those of **3** and **7** (log $P_{o/w}$: **3** = 0.46, **7** = 0.70) that demonstrate uptake 2−3 orders of magnitude greater than **8** in vitro. This is not unexpected, as no transporters have yet been identified that can translocate negatively charged platinum complexes (though cationic Pt complexes demonstrate avid uptake by an as-yet undetermined mechanism).[38] Given the anomalous cellular accumulation in both cell lines due to charge, complex **8** was not incorporated into Figure 4. The data demonstrate an upward trend for cellular accumulation with increasing lipophilicity and little difference in accumulation patterns between the two cell lines.

Recent work examining the biotransformation products of the clinical drugs carboplatin and cisplatin in biological media have shown that a number of charged species form with endogenous biomolecules such as carbonate and phosphate, and the paradigm of cisplatin as a naive molecule which exclusively enters cells via passive diffusion has been discarded.[39] While these charged species are unlikely to enter cells via passive diffusion, there is little consensus in the literature on whether aquated species are able to enter cells. It is usually stated that when a chloro ligand on cisplatin is displaced by a labile water ligand outside of the cell, a process known as aquation, the drug cross-reacts with biomolecules, including proteins that deactivate the drug.[33] Cisplatin forms a number of aquated species; however, the primary species formed (Figure 5) is the neutral chlorohydroxo species (**c** in Figure 5), which accounts for 26% of species at equilibrium in high chloride concentration. Not only is **c** present in the extracellular media at relatively high levels, the platinum−hydroxo bond is relatively unreactive; however, experimental determination of its log $P_{o/w}$ is not possible given the complicated equilibria at play. Table 4 reports the proportions of each species present in 104 mM chloride,[40] the approximate chloride concentration present in the extracellular media, the charge of each of the aquation species shown in the scheme, and the log $P_{o/w}$ calculated for cisplatin (**1**) and the neutral chlorohydroxo species. The calculated log $P_{o/w}$ for *cis*-[PtCl(OH)$(NH_3)_2$] (c) is −2.7, only slightly lower than that for cisplatin (−2.4). Given the prospect of hydrogen bonding with phosphates on the cell membrane fostering preassociation with the cell membrane,[41] this result suggests that the chlorohydroxo species is capable of passively diffusing into the cell and that "aquation" of cisplatin does not necessarily represent deactivation.

Table 4. Proportions of Each Species Present in 104 mM Chloride

compound	104 mM	charge	log $P_{o/w}$ (calc)[a]
1	67	0	−2.4
a	4	1+	*a*
b	<1	2+	*a*
c	26	0	−2.7
d	<1	1+	*a*
e	1	0	

[a] Log $P_{o/w}$ values were not calculated for charged species and **e**, which is not detected in biological reactions such as DNA binding experiments.[41]

Conclusions

We present a QSPR for the lipophilicity of platinum complexes based on 64 literature octanol−water partition data. Using a combination of genetic algorithm descriptor selection and partial least-squares regression (GA-PLS), we can identify eight descriptors that form a statistically valid and robust model of log $P_{o/w}$. This model explains 92% of the variance in the log $P_{o/w}$ dataset, while cross-validation gives a Q^2 value of 87%. The root-mean-square error of 0.39 log units can be compared to the spread of data found from multiple measurements, which is in the range of 0.15 to 0.45 log units. The QSPR model allows us to predict whole-molecule lipophilicity for unmeasured complexes, which shows significant differences from attempts to use substituent constants alone. A broad upward trend of lipohilicity against uptake into cells is observed, and the implications of this for design of new drugs that may circumvent resistance to cisplatin is discussed. We also predict the log $P_{o/w}$ value for an activated metabolite of cisplatin, that is, *cis*-[PtCl(OH)$(NH_3)_2$], which is only slightly less lipophilic than cisplatin itself, demonstrating the ability for prediction to impact on our understanding of cellular accumulation mechanisms. This work should have great significance to the pharmaceutical industry allowing in silico screening of potential synthetic targets using calculated molecular properties alone.

Experimental Section

Descriptor Generation. All calculations were performed with Gaussian03[42] on the "Helix" Beowulf cluster[43] Computation and Visualization Facility of Cardiff University. Following previous work,[23] geometries of the gas-phase complexes were fully optimized without any symmetry constraint using the B3LYP hybrid functional with the LANL2DZ basis set; then, electron density, $V(\mathbf{r})$, and $I(\mathbf{r})$ were calculated using B3LYP with the 6-31+G(d,p) basis set on all light atoms, and the CEP-121G basis set and core potential were used for Pt. The molecular dipole moment and dipole polarizability were also calculated at this level.

Properties were calculated individually at every lattice point within an approximately cubic grid centered on each molecule using $80^3 = 512000$ lattice points. An in-house C-program then extracted the isodensity surface, located local $V(\mathbf{r})$ and $I(\mathbf{r})$ extrema, and calculated statistical descriptors from these surface properties. As in our previous work,[23] surface descriptors were computed on the 0.001 and 0.0005 au electron density isocontour for $V_S(\mathbf{r})$ and $I(\mathbf{r})$, respectively. On the basis of previous experience with organic molecules,[44] an empirical cutoff for minima in $V_S(\mathbf{r})$ of −0.03 au was used for two additional descriptors ($\Sigma V_S^-{}_{cutoff}$ and Ave[$V_S^-{}_{cutoff}$]), attempting to eliminate spurious minima that may distort descriptor values for large compounds; however, these cutoff descriptors did not enter the final QSPR model. An approximate measure of molecular volume was constructed from the number grid points within the 0.001 au density surface. Statistical descriptors[23] were produced from this data, including global maximum and minimum values, the number and sum total of minima and maxima, averages,

and standard deviations (see Supporting Information for full definitions and numerical data).

Statistical Methods. For the genetic algorithm (GA), a population of 256 possible combinations of descriptors were generated at random and allowed to mutate, crossover, and add or subtract descriptors to find populations with better PLS statistics, over a maximum of 200 generations. All descriptors were scaled to a standard deviation = 1 and centered to have a mean of 0 before GA methods were applied.

The quality of fit was measured by the squared correlation coefficient, R^2, the fraction of total variance in log $P_{o/w}$ explained by the regression equation, and by the root-mean-square error (RMSE). PLS is capable of overfitting, that is, tailoring the model too much to the current data, to the detriment of future predictions. Cross-validation, in which the data set is divided into groups, is therefore important. The model is fitted to each group except one and then used to predict the omitted group, yielding the Q^2 and prediction root-mean-square error (PRMSE) values.[45] A further test for overfitting was the randomization procedure available in SIMCA-P[46], which randomly permutes the log $P_{o/w}$ values and fits a model to the permuted data. Only if overfitting is present should a significant model result; this process was repeated, and R^2 and Q^2 values were plotted against the correlation between original and permuted log $P_{o/w}$ and extrapolated to the intercept. High values indicate an overfitted model, with a rule of thumb of 0.05 as the maximum allowable value for R^2.

Acknowledgment. Thanks to Dr. Michael M. Gottesman for providing a stimulating research environment and valuable discussion. All DFT calculations were carried out using Cardiff University's Helix computing cluster.[43]

Supporting Information Available: Definitions and numerical values of all descriptors tested in the GA-PLS optimization. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hansch, C.; Bjorkroth, J. P.; Leo, A. Hydrophobicity and Central-Nervous-System Agents — On the Principle of Minimal Hydrophobicity in Drug Design. *J. Pharm. Sci.* **1987**, *76*, 663−687.
(2) Lambert, W. J. Modeling Oil−Water Partitioning and Membrane Permeation using Reversed-Phase Chromatography. *J. Chromatogr., A* **1993**, *656*, 469−484.
(3) Braumann, T. Determination of Hydrophobic Parameters by Reversed-Phase Liquid-Chromatography — Theory, Experimental-Techniques, and Application in Studies on Quantitative Structure-Activity-Relationships. *J. Chromatogr.* **1986**, *373*, 191−225.
(4) Screnci, D.; McKeage, M. J.; Galettis, P.; Hambley, T. W.; Palmer, B. D.; et al. Relationships between Hydrophobicity, Reactivity, Accumulation and Peripheral Nerve Toxicity of a Series of Platinum Drugs. *Br. J. Cancer* **2000**, *82*, 966−972.
(5) Hambley, T. W. The Influence of Structure on the Activity and Toxicity of Pt Anticancer Drugs. *Coord. Chem. Rev.* **1997**, *166*, 181−223.
(6) Muggia, F. M.; Fojo, T. Platinums: Extending their Therapeutic Spectrum. *J. Chemother.* **2004**, *16*, 77−82.
(7) Ghezzi, A.; Aceto, M.; Cassino, C.; Gabano, E.; Osella, D. Uptake of Antitumor Platinum(II)-Complexes by Cancer Cells, Assayed by Inductively Coupled Plasma Mass Spectrometry (ICP-MS). *J. Inorg. Biochem.* **2004**, *98*, 73−78.
(8) Platts, J. A.; Hibbs, D. E.; Hambley, T. W.; Hall, M. D. Calculation of the Hydrophobicity of Platinum Drugs. *J. Med. Chem.* **2001**, *44*, 472−474.
(9) Mellish, K. J.; Kelland, L. R.; Harrap, K. R. In-Vitro Platinum Drug Chemosensitivity of Human Cervical Squamous-Cell Carcinoma Cell-Lines with Intrinsic and Acquired-Resistance to Cisplatin. *Br. J. Cancer* **1993**, *68*, 240−250.
(10) Choy, H. Satraplatin: An Orally Available Platinum Analog for the Treatment of Cancer. *Expert Rev. Anticancer Ther.* **2006**, *6*, 973−982.
(11) Hansch, C.; Fujita, T. $\rho-\sigma-\pi$ Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **1964**, *86*, 1616−1626.
(12) Robertazzi, A.; Platts, J. A. Hydrogen Bonding and Covalent Effects in Binding of Cisplatin to Purine Bases: Ab initio and Atoms in Molecules Studies. *Inorg. Chem.* **2005**, *44*, 267−274.
(13) Politzer, P.; Truhlar, D. G. *Chemical Applications of Atomic and Molecular Electrostatic Potentials. Reactivity, Structure, Scattering and Energetics of Organic, Inorganic and Biological Systems*; Plenum Press: New York, 1981.
(14) Sattelle, B. M. Properties of Molecular Surfaces: Electrostatics and Reactivity. M.S. Thesis, Cardiff University, 2004.
(15) Politzer, P.; Murray, J. S.; Peralta-Inga, Z. Molecular Surface Electrostatic Potentials in Relation to Noncovalent Interactions in Biological Systems. *Int. J. Quantum Chem.* **2001**, *85*, 676−684.
(16) Sjoberg, P.; Murray, J. S.; Brinck, T.; Politzer, P. Average Local Ionization Energies on the Molecular-Surfaces of Aromatic Systems as Guides to Chemical-Reactivity. *Can. J. Chem.* **1990**, *68*, 1440−1443.
(17) Ehresmann, B.; Martin, B.; Horn, A. H. C.; Clark, T. Local Molecular Properties and their Use in Predicting Reactivity. *J. Mol. Model.* **2003**, *9*, 342−347.
(18) Ehresmann, B.; de Groot, M. J.; Alex, A.; Clark, T. New Molecular Descriptors Based on Local Properties at the Molecular Surface and a Boiling-point Model Derived from Them. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 658−668.
(19) Langenaeker, W.; Deproft, F.; Geerlings, P. Development of Local Hardness Related Reactivity Indexes — Their Application in a Study of the Se at Monosubstituted Benzenes within the Hsab Context. *J. Phys. Chem.* **1995**, *99*, 6424−6431.
(20) Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P. Influence of Stacking on Hydrogen Bonding: Quantum Chemical Study on Pyridine−Benzene Model Complexes. *J. Phys. Chem. A* **2004**, *108*, 6038−6044.
(21) Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P. Influence of the pi−pi Interaction on the Hydrogen Bonding Capacity of Stacked DNA/RNA Bases. *Nucleic Acids Res.* **2005**, *33*, 1779−1789.
(22) Saunders, R. A.; Platts, J. A. Scaled Polar Surface Area Descriptors: Development and Application to Three Sets of Partition Coefficients. *New J. Chem.* **2004**, *28*, 166−172.
(23) Platts, J. A.; Oldfield, S. P.; Reif, M. M.; Palmucci, A.; Gabano, E.; et al. The RP-HPLC Measurement and QSPR Analysis of log $P_{o/w}$ Values of Several Pt(II) Complexes. *J. Inorg. Biochem.* **2006**, *100*, 1199−1207.
(24) Parr, R. G.; Weitao, Y. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 2003; 352.
(25) Souchard, J. P.; Ha, T. T. B.; Cros, S.; Johnson, N. P. Hydrophobicity Parameters for Platinum Complexes. *J. Med. Chem.* **1991**, *34*, 863−864.
(26) Robillard, M. S.; Galanski, M.; Zimmermann, W.; Keppler, B. K.; Reedijk, J. (Aminoethanol)dichloroplatinum(II) Complexes: Influence of the Hydroxyethyl Moiety on 5′-GMP and DNA Binding, Intramolecular Stability, the Partition Coefficient and Anticancer Activity. *J. Inorg. Biochem.* **2002**, *88*, 254−259.
(27) Hall, M. D.; Amjadi, S.; Zhang, M.; Beale, P. J.; Hambley, T. W. The Mechanism of Action of Platinum(IV) Complexes in Ovarian Cancer Cell Lines. *J. Inorg. Biochem.* **2004**, *98*, 1614−1624.
(28) Kim, Y.-S.; Song, R.; Chung, H. C.; Jun, M. J.; Sohn, Y. S. Coordination Modes vs. Antitumor Activity: Synthesis and Antitumor Activity of Novel Platinum(II) Complexes of N-Substituted Amino Dicarboxylic Acids. *J. Inorg. Biochem.* **2004**, *98*, 98−104.
(29) Coluccia, M.; Nassi, A.; Boccarelli, A.; Giordano, D.; Cardellicchio, N.; et al. In Vitro and In Vivo Antitumour Activity and Cellular Pharmacological Properties of New Platinum−iminoether Complexes with Different Configuration at the Iminoether Ligands. *J. Inorg. Biochem.* **1999**, *77*, 31−35.
(30) Kidani, Y. *Platinum and Other Metal Coordination Compounds in Cancer Chemotherapy 2*; Plenum Press: New York, 1996; p 43.
(31) Lee, Y.-A.; Lee, S. S.; Kim, K. M.; Lee, C. O.; Sohn, Y. S. Synthesis and Oral Antitumor Activity of Tetrakis(carboxylato)platinum(IV) Complexes. *J. Med. Chem.* **2000**, *43*, 1409−1412.
(32) Braddock, P. D.; Connors, T. A.; Jones, M.; Khokhar, A. R.; Melzack, D. H.; et al. Structure and Activity Relationships of Platinum Complexes with Anti-Tumour Activity. *Chem.-Biol. Interact.* **1975**, *11*, 145−161.
(33) Jung, Y.; Lippard, S. J. Direct Cellular Responses to Platinum-Induced DNA Damage. *Chem. Rev.* **2007**, *107*, 1387−1407.
(34) Gately, D. P.; Howell, S. B. Cellular Accumulation of the Anticancer Agent Cisplatin — A Review. *Br. J. Cancer* **1993**, *67*, 1171−1176.
(35) Mellish, K. J.; Kelland, L. R. Mechanisms of Acquired-Resistance to the Orally-Active Platinum-Based Anticancer Drug bis-Aceto-Amine-Dichloro-Cyclohexylamine Platinum(IV) (JM216) in 2 Human Ovarian-Carcinoma Cell-Lines. *Cancer Res.* **1994**, *54*, 6194−6200.
(36) Loh, S. Y.; Mistry, P.; Kelland, L. R.; Abel, G.; Harrap, K. R. Reduced Drug Accumulation as a Major Mechanism of Acquired-Resistance to Cisplatin in a Human Ovarian-Carcinoma Cell-Line — Circumvention Studies Using Novel Platinum(II) and (IV) Ammine Amine Complexes. *Br. J. Cancer* **1992**, *66*, 1109−1115.

(37) Ang, W. H.; Pilet, S.; Scopelliti, R.; Bussy, F.; Juillerat-Jeanneret, L.; et al. Synthesis and Characterization of Platinum(IV) Anticancer Drugs with Functionalized Aromatic Carboxylate Ligands: Influence of the Ligands on Drug Efficacies and Uptake. *J. Med. Chem.* **2005**, *48*, 8060−8069.
(38) Hall, M. D.; Okabe, M.; Shen, D.-W.; Liang, X.-J.; Gottesman, M. M. The Role of Cellular Accumulation in Determining Sensitivity to Platinum-Based Chemotherapy. *Ann. Rev. Pharmacol. Toxicol.* **2008**, in press.
(39) Centerwall, C. R.; Goodisman, J.; Kerwood, D. J.; Dabrowiak, J. C. Cisplatin Carbonato Complexes. Implications for Uptake, Antitumor Properties, and Toxicity. *J. Am. Chem. Soc.* **2005**, *127*, 12768−12769.
(40) Martin, R. B. *Platinum Complexes: Hydrolysis and Binding to N(7) and N(1) of Purines*; Wiley-VCH: New York, 1999; pp 183−205.
(41) Davies, M. S.; Berners-Price, S. J.; Hambley, T. W. Rates of Platination of AG and GA Containing Double-Stranded Oligonucleotides: Insights into Why Cisplatin Binds to GG and AG but not GA Sequences in DNA. *J. Am. Chem. Soc.* **1998**, *120*, 11380−11390.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(43) Helix. www.helix.cf.ac.uk/.
(44) Lamarche, O.; Platts, J. A. Atoms in Molecules Investigation of the p*K*(HB) Basicity Scale: Electrostatic and Covalent Effects in Hydrogen Bonding. *Chem. Phys. Lett.* **2003**, *367*, 123−128.
(45) JMP, 5.0 ed.. www.jmpdiscovery.com. SAS Institute Inc: Cary, NC.
(46) *Simca-P*, 8.0 ed.; Umetrics AB: Kinnelon, NJ.

JM0708275