J Comput Aided Mol Des (2009) 23:343–354
DOI 10.1007/s10822-009-9265-4

DFT-based QSAR and QSPR models of several *cis*-platinum complexes: solvent effect

Pubalee Sarmah · Ramesh C. Deka

Received: 6 January 2009 / Accepted: 1 March 2009 / Published online: 24 March 2009

Abstract Cytotoxic activities of *cis*-platinum complexes against parental and resistant ovarian cancer cell lines were investigated by quantitative structure-activity relationship (QSAR) analysis using density functional theory (DFT) based descriptors. The calculated parameters were found to increase the predictability of each QSAR model with incorporation of solvent effects indicating its importance in studying biological activity. Given the importance of logarithmic *n*-octanol/water partition coefficient (log $P_{o/w}$) in drug metabolism and cellular uptake, we modeled the log $P_{o/w}$ of 24 platinum complexes with different leaving and carrier ligands by the quantitative structure-property relationship (QSPR) analysis against five different concentrations of MeOH using DFT and molecular mechanics derived descriptors. The log $P_{o/w}$ values of an additional set of 20 platinum complexes were also modeled with the same descriptors. We investigated the predictability of the model by calculating log $P_{o/w}$ of four compounds in the test set and found their predicted values to be in good agreement with the experimental values. The QSPR analyses performed on 24 complexes, combining the training and test sets, also provided significant values for the statistical parameters. The solvent medium played an important role in QSPR analysis by increasing the internal predictive ability of the models.

Electronic supplementary material The online version of this article (doi:10.1007/s10822-009-9265-4) contains supplementary material, which is available to authorized users.

P. Sarmah · R. C. Deka (✉)
Department of Chemical Sciences, Tezpur University, Napaam, Tezpur, Assam 784028, India
e-mail: ramesh@tezu.ernet.in

Keywords QSAR · QSPR · *cis*-Platinum complexes · DFT · Solvent effect

Introduction

There is a long-standing interest in platinum complexes because of their well established anticancer activity. The *cis*-diamminedichloroplatinum(II), clinically known as cisplatin, was first recognized as an anti-tumor agent in the early 1970s [1]. Cisplatin has since been a paradigm for the treatment of testicular and ovarian cancers [2–5]. The limitations of usefulness of cisplatin by the development of resistance after continued treatment and high toxicity to some normal cells have stimulated research toward developing analogs of cisplatin with lesser toxic effects. One of them is the carboplatin [6], a second generation drug, which presents lower toxicity than cisplatin. The pharmacokinetic difference between cisplatin and carboplatin is due to the slower rate of conversion of carboplatin to the reactive species. In continued search of new platinum-based drugs of improved anticancer activity, more than 3,000 platinum compounds have been prepared and tested against several tumor cell lines. The cytotoxicity of platinum complexes depends on the nature of carrier and leaving ligands. Monti et al. [7] studied the cytotoxicities of 16 platinum complexes with different leaving and carrier groups in two cancer cell lines. Their results confirm the Cleare and Hoeschele's empirical rules that the presence of NH_3 and DACH [(1R,2R)-1,2-diaminocyclohexane] as carrier groups and chloride and oxalate as leaving ligands yield the highest cytotoxic effects.

Hydrophobicity, measured as logarithm of the 1-octanol/water partition coefficient (log $P_{o/w}$) is a very important property owing to its usefulness to assess biological effects

relevant to drug action, such as lipid solubility, tissue distribution, receptor binding, cellular uptake, metabolism, and bioavailability. Several experimental and theoretical studies have been devoted in determining hydrophobicity of different organic molecules of pharmacological and toxicological importance. However, limited studies have been carried out on log $P_{o/w}$ of platinum complexes. Screnci et al. [8] reported log $P_{o/w}$ of eight platinum drugs, including four platinum(IV) drug molecules using shake-flask method. They also derived another hydrophobicity parameter, log k_w and observed a weak correlation between log $P_{o/w}$ and log k_w. Platts et al. [9] calculated hydrophobicity of a series of 24 platinum complexes with the help of RP-HPLC technique and found a good correlation of these values with that derived from DFT calculations.

The ultimate goal of quantitative structure-activity and structure-property relationship (QSAR/QSPR) studies is to correlate the biological activity/property of a series of compounds with some appropriate descriptors. Among different descriptors for describing the electronic properties of molecules, the quantum chemical descriptors based on density functional theory (DFT) and semi-empirical methods have been found useful in several QSAR studies [10, 11]. In particular, net atomic charges, HOMO–LUMO energies, frontier orbital electron densities, and superdelocalizabilities have shown to correlate with various biological activities [12].

In recent years, DFT based reactivity descriptors namely, global hardness (η), electronegativity (χ), chemical potential (μ), electrophilicity index (ω), Fukui functions ($f(r)$), philicity (ω_k^α), etc. [13–17] have attracted considerable interests to describe reactivity and site selectivity of various bio-molecules [18, 19]. The electrophilicity and philicity indices have successfully been used to predict the biological activity/toxicity/property of different organic systems [20–22]. Although numerous theoretical calculations have been performed to understand the structure and binding mechanism [23–32] of platinum drugs with DNA, very few studies have paid attention on QSAR/QSPR analyses of these molecules [7, 9, 33]. In a recent paper, we have calculated cytotoxicity of platinum complexes using electrophilicity index in solvent phase which was found to correlate well with the experimental values [34]. However, this single QSAR parameter failed to reproduce the experimental cytotoxicity values in gas phase. In the present study, we have found that DFT derived reactivity descriptors, in particular electrophilicity and philicity in combination with energy of next LUMO orbital can correlate drug activity of *cis*-platinum complexes remarkably in both gas and solvent phases. In addition, we have calculated hydrophobicity of those complexes, by noting its importance in drug action, metabolism and receptor binding. We found from QSPR analysis that DFT based reactivity descriptors in combination with molar refractivity and surface area can be used for prediction of hydrophobicity of platinum complexes.

Methods

Theoretical background

The global electrophilicity index (ω) introduced by Parr et al. [16] is expressed in terms of chemical potential and hardness as:

$$\omega = \frac{\mu^2}{2\eta} \tag{1}$$

where the chemical potential (μ) and hardness (η) are the partial derivatives of the system's energy E expressed as a functional of an external potential, $v(\vec{r})$, and a function of the number of electrons N:

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v(\vec{r})} \tag{2}$$

$$\eta = \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} \tag{3}$$

In finite difference approach, global hardness and chemical potential can be approximated as

$$\eta = \frac{\mathrm{IP} - \mathrm{EA}}{2} \tag{4}$$

$$\mu = -\left(\frac{\mathrm{IP} + \mathrm{EA}}{2}\right) \tag{5}$$

where, IP and EA are the first vertical ionization potential and electron affinity, respectively, of the chemical system.

Further approximation using Koopmans' theorem [35], the above parameters can be expressed by taking IP and EA as negative of the HOMO and LUMO energies:

$$\mu = \frac{E_{\mathrm{LUMO}} + E_{\mathrm{HOMO}}}{2} \tag{6}$$

and

$$\eta = \frac{E_{\mathrm{LUMO}} - E_{\mathrm{HOMO}}}{2} \tag{7}$$

where E_{LUMO} is the energy of the lowest unoccupied molecular orbital and E_{HOMO} is the energy of the highest occupied molecular orbital.

Recently, Chattaraj et al. [18] have defined a generalized concept of philicity associated with a site k in a molecule as:

$$\omega_k^\alpha = \omega f_k^\alpha \tag{8}$$

where $\alpha = +, -,$ and 0 represent nucleophilic, electrophilic, and radical attacks, respectively, and f_k^α, the

Fukui function (FF), [16] is by far the most important local reactivity index and defined as:

$$f_k^{\alpha} = \left[\frac{\delta\mu}{\delta v(r)}\right]_N = \left[\frac{\delta\rho(r)}{\delta N}\right]_{v(r)} \quad (9)$$

Mendez and Gazquez [36] and Yang and Mortier [37] introduced a procedure to obtain information about f_k^{α}. This procedure condenses the values around each atomic site into a single value that characterizes the atom in the molecule. With this approximation, the condensed Fukui function becomes

$$f_k^{+} = [q_k(N+1) - q_k(N)] \quad \text{(for nucleophilic attack on the system)} \quad (10a)$$

$$f_k^{-} = [q_k(N) - q_k(N-1)] \quad \text{(for electrophilic attack on the system)} \quad (10b)$$

$$f_k^{0} = \frac{1}{2}[q_k(N+1) - q_k(N-1)] \quad \text{(for radical attack on the system)} \quad (10c)$$

where $q_k(N)$, $q_k(N+1)$, and $q_k(N-1)$ are the charges of the kth atom for N, $N+1$ and $N-1$ electron systems, respectively.

Computational details

Full unconstrained geometry optimizations of all complexes were carried out at gradient corrected DFT using the DMol^3 program [38]. The most widely used exchange-correlation functional suggested exchange potential by Beck [39] with gradient corrected correlation provided by Lee, Yang and Parr [40] (BLYP) was used in combination with double numerical with polarization (DNP) basis set to study all the complexes in both gas and solvent phases. The BLYP/DNP level was adopted as it can predict comparatively better geometry of platinum complexes as found in our previous study [34]. The size of this DNP basis set is comparable to the 6-31G** basis of Hehre et al. [41]. However, it is believed to be much more accurate than a Gaussian basis set of the same size. We performed all electron calculations, including relativistic effects for all complexes, as available in DMol^3. All complexes were characterized as minima (no imaginary frequency) in their potential energy surface through harmonic frequency analysis. The reactivity descriptors electrophilicity index (ω) and local philicity $\left(\omega_k^{+}\right)$ were calculated for all the systems using Eqs. 1 and 8, respectively. The Hirshfeld [42] population analysis (HPA) was used to calculate the FF. The conductor-like screening model (COSMO) [43] as incorporated into the DMol^3 program with dielectric constant of 78.4 was adopted to study the solvent (water) effect. The molar refractivity parameter of carrier ligands and surface area of each complex were obtained from the MM+ computations with Hyperchem software [44]. The predictive ability of models was determined using the "leave one out" (LOO) cross-validation method.

QSAR/QSPR modeling

From the results of DFT calculations, different descriptors were selected for QSAR and QSPR modeling such as, the energy of highest occupied molecular orbital (E_{HOMO}), energy of lowest unoccupied molecular orbital (E_{LUMO}), energy of the next lowest unoccupied molecular orbital (E_{NL}), energy difference between LUMO and HOMO ($\Delta_{\mathrm{L-H}}$), dipole moments, electrophilicity (ω), hardness (η), philicity (ω^{+}), etc. In addition, the molecular mechanics (MM) parameters such as molar refractivity of carrier ligand ($\mathrm{MR_{CL}}$), van der Waals surface area (SA), molecular volume, hydrophobicity of carrier ligand ($\log P_{\mathrm{CL}}$) were also selected.

The anticancer activity data of complexes (1–16) against the A2780 human ovarian adenocarcinoma cell line and its cisplatin resistant subline (A2780Cp8) were taken from the results reported by Osella et al. [7]. These values were conventionally transformed to $\log \mathrm{IC_{50}^{-1}}$ in QSAR studies. The analyses were performed in both gas and solvent media for the 16 platinum complexes. We carried out QSPR studies for analyzing the $\log P_{\mathrm{o/w}}$ values of platinum complexes (1–24) for 0% (extrapolated), 20, 30, 40 and 50% MeOH [7]. Since the partition behavior markedly depends on the solvent, we also performed multiple regression analysis using solvent phase predicted molecular properties. The $\log P_{\mathrm{o/w}}$ values of a training set of 20 platinum complexes were also modeled. Further we investigated predictability of the models by calculating $\log P_{\mathrm{o/w}}$ of four compounds in the test set. The regression analyses were also performed on 24 complexes obtained from combination of training set and test set. The descriptors having greater correlation to $\log \mathrm{IC_{50}^{-1}}$ and $\log P_{\mathrm{o/w}}$ with smaller autocorrelation were selected out to perform the stepwise multiple linear regression. Three parameter QSAR and four parameter QSPR [45] were performed using least square error estimation method [46] to calculate and compare the cytotoxicity ($\log \mathrm{IC_{50}^{-1}}$) and hydrophobicity ($\log P_{\mathrm{o/w}}$) of the complexes, respectively. The predictive ability of models was determined using the "leave one out" (LOO) cross-validation method.

Results and discussion

All studied platinum complexes are presented in Fig. 1. The optimized geometries of the complexes have square planar configuration with angles close to the ideal values of 90° and 180°. The optimized geometry of cisplatin (1) is in

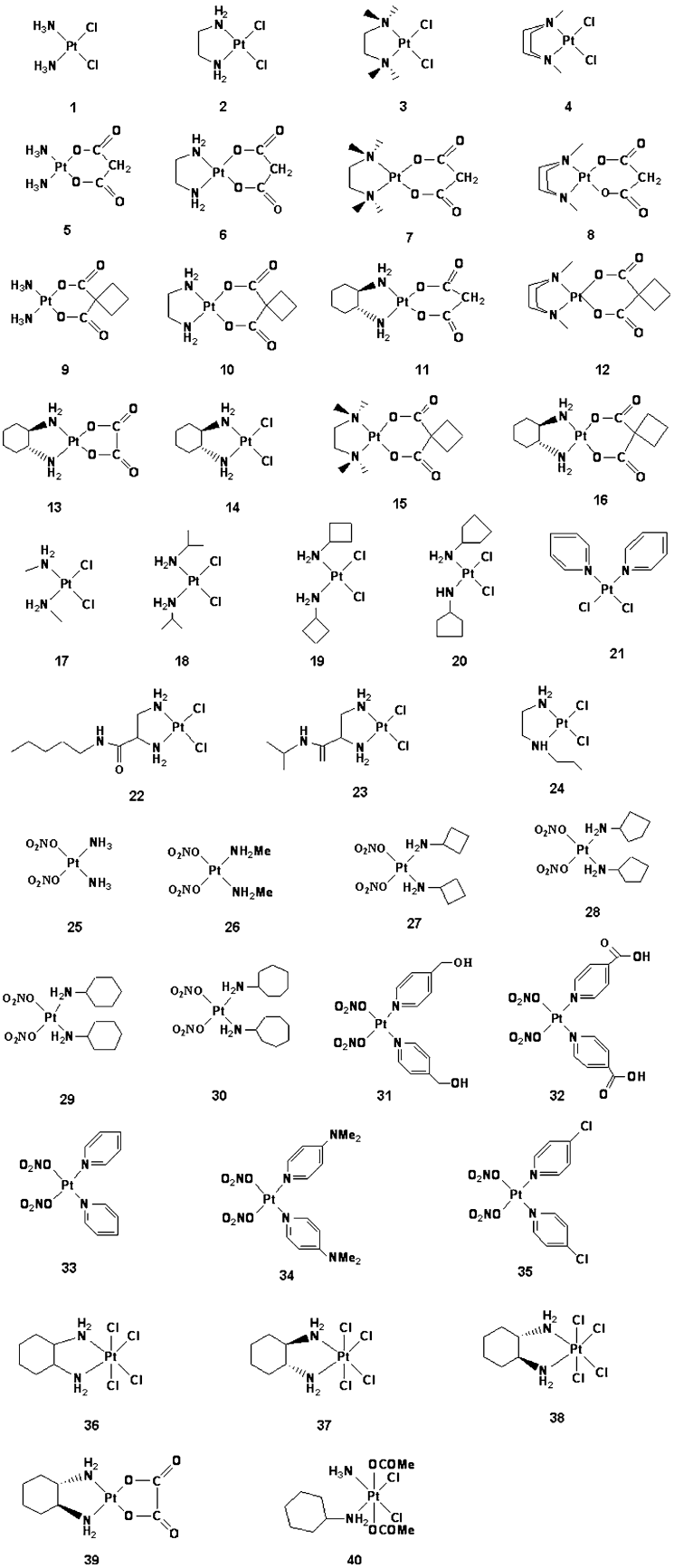

Fig. 1 Sketch of the platinum complexes used to build QSAR and QSPR models

good agreement with X-ray crystal structure reported by Milbum and Truter [47]. The calculated Pt–Cl and Pt–N bond lengths are 2.32 and 2.11 Å, respectively, in accordance with their experimental values. The N–Pt–N angle (97.1°) and Cl–Pt–Cl (96.8°) angle are larger by about 5–7° from their experimental values, 87 ± 1.5° and 91.9 ± 0.4°, respectively. Similar deviation of bond angles in cisplatin was reported in theoretical studies performed by Wysokinske and Michalska [48]. The planar environment of the platinum atom and the boat conformation for the six-membered chelate ring obtained for carboplatin (9) are in agreement with the X-ray diffraction data [49, 50]. The Pt–N (2.10 Å) and Pt–O (2.00 Å) bond lengths and N–Pt–N angle (96.4°) and O–Pt–O angle (84.3°) of oxaliplatin (13) are close to its X-ray crystal structure [51]. Chair configuration of the cyclohexane ring with two amino groups in equatorial positions is found in accordance with the experimental results. The geometrical parameters of other complexes, for which X-ray data are not available, are compared with geometries of their similar analogue. We found that our calculated geometries for all complexes are in good agreement with the available experimental data.

QSAR analysis on A2780 cell line

The QSAR equations with absolute values of statistical parameters in both gas and solvent phases for 16 platinum complexes against A2780 cell line are represented by Eqs. 11 and 12. The values were calculated by considering the cytotoxicity (log IC_{50}^{-1}) as a dependent variable and electrophilicity (ω), philicity (ω^+), and energy of the next LUMO orbital (E_{NL}) as independent variables. The descriptors used to build the QSAR model for both gas and solvent phases are presented in Table 1.

$$\begin{aligned}\text{Gas phase: } \log\left(IC_{50}^{-1}\right) &= -9.942 + 3.25\omega - 11.397\omega^+ \\ &\quad - 3.71E_{NL} \\ n = 16, \quad r^2 &= 0.706, \quad r_{CV}^2 = 0.430, \\ SD = 1.147, \quad F &= 9.63, \quad p < 0.05\end{aligned} \tag{11}$$

$$\begin{aligned}\text{Solvent phase: } \log\left(IC_{50}^{-1}\right) &= -22.437 + 2.341\omega \\ &\quad + 10.562\omega^+ - 5.019E_{NL} \\ n = 16, \quad r^2 &= 0.710, \quad r_{CV}^2 = 0.637, \\ SD = 1.141, \quad F &= 9.78, \quad p < 0.05\end{aligned} \tag{12}$$

Here, r^2 is the square of correlation coefficient, r_{CV}^2 is the leave-one-out (LOO) cross validated squared correlation coefficient, F is the overall F-statistics for the addition of each successive term, p is the p values using the F statistics, and SD is the standard deviations of regression. We found that the gas phase r^2 value (0.706) increases slightly (0.710) with the inclusion of solvent. However, for this case r_{CV}^2 value (0.441) increases to an acceptable value

Table 1 Parameters used to build the QSAR models with the jackknife results for gas and solvent phases against two cancer cell lines

Complex	log IC_{50}^{-1}		Gas phase					Solvent phase				
	A2780	A2780Cp8	ω	ω^+	E_{NL}	r_j^2		ω	ω^+	E_{NL}	r_j^2	
						A2780	A2780Cp8				A2780	A2780Cp8
1	−0.315	−3.807	4.134	0.868	−1.264	0.750	0.723	3.884	1.056	−0.223	0.698	0.826
2	−1.714	−4.770	3.468	0.697	−1.212	0.707	0.705	3.873	1.026	0.032	0.728	0.812
3	−3.305	−4.843	3.353	0.624	−0.616	0.709	0.719	4.107	0.920	−0.197	0.735	0.825
4	−3.660	−4.607	3.503	0.648	−0.565	0.698	**0.745**	4.188	0.955	−0.097	0.747	0.816
5	−1.991	−3.330	3.585	0.649	−1.523	**0.794**	0.713	2.956	0.834	−1.004	0.711	0.809
6	−2.179	−4.172	3.131	0.517	−1.312	**0.744**	**0.748**	2.978	0.822	−1.008	0.713	0.824
7	−5.171	−6.859	2.819	0.555	−0.501	0.658	0.647	3.078	0.605	−1.008	0.670	0.784
8	−5.120	−5.956	3.001	0.582	−0.847	0.708	0.691	3.230	0.707	−0.704	0.699	0.801
9	−2.036	−4.814	3.368	0.546	−0.907	0.706	0.726	2.820	0.663	−1.240	0.715	0.818
10	−2.220	−4.190	2.964	0.430	−0.959	0.713	0.720	2.666	0.645	−1.196	0.738	**0.832**
11	−4.912	−6.467	2.673	0.513	−0.469	0.673	0.661	2.907	0.610	−0.956	0.678	0.793
12	−5.121	−6.422	2.879	0.538	−0.825	0.709	0.709	3.118	0.692	−0.715	0.682	0.795
13	0.734	−1.890	2.753	0.347	−1.222	0.681	0.658	4.357	0.388	−1.607	**0.787**	0.798
14	−0.030	−2.111	3.072	0.602	−1.49	0.724	0.732	3.815	0.992	−0.139	**0.798**	**0.936**
15	−0.798	−2.271	2.749	0.412	−1.133	0.702	0.693	2.966	0.813	−1.209	0.695	0.798
16	−0.419	−1.319	2.608	0.344	−1.111	0.701	0.704	3.748	0.885	−0.987	0.716	0.764

Complexes having bold values are outliers

(0.637) with the change of gas phase to solvent phase indicating the importance of the solvent model. In general, a regression model is significant at p value <0.05 using the F statistics [52] and so these models are statistically significant. However, according to the generally statistical standards, a model with $r^2 > 0.80$ [53] and $r^2_{CV} > 0.60$ [54] is acceptable. Therefore, these QSAR equations should be further improved to become a statistically significant model.

To improve r^2, one scheme was suggested by Dietrich et al. [55] and Cornish-Bowden and Wang [56] in which a compound is considered as outlier if its corresponding r^2, called jackknife r^2 (r^2_j) value obtained from the regression analysis after deleting the compound, is comparatively higher than the other r^2_j values. We applied this method to increase overall quality of the regression models. The r^2_j values calculated in gas and solvent phases for the cell lines are presented in Table 1. Since the independent variables are different in both gas and solvent phases, a particular complex has quite different values of r^2_j in gas phase than that calculated in solvent phase. Thus the outliers are different for both the phases. We observed that the complexes 1, 5 and 6 exhibited unduly high r^2_j values (0.75, 0.794, and 0.744, respectively) in the gas phase; whereas, in the solvent phase the complexes 13, and 14 possessed higher r^2_j values (0.787, and 0.798, respectively) and thus these complexes may be considered as outliers. However, it is seen that when complexes 5 and 6 were deleted from the data set, a significant improvement of the statistical parameters were observed compared to that obtained by deleting complexes 1 and 5.

The QSAR equations after deleting these complexes (5, and 6, in gas phase) and (13, and 14, in solvent phase) with significant statistical quality are presented in Table 2. We observed that r^2 values increased from 0.706 to 0.859 and r^2_{CV} values from 0.430 to 0.748 in the gas phase. The solvent model did not show any influence for this cell line. However, the solvent phase predicted r^2 and r^2_{CV}values after applying jackknife test increased from 0.710 to 0.844 and 0.637 to 0.695, respectively.

QSAR analysis on A2780Cp8 cell line

Multi-linear regression analysis between log IC_{50}^{-1} of platinum complexes (1–16) against A2780Cp8 cell line and the combination of three DFT derived descriptors yielded the following QSAR equations

$$\text{Gas phase: } \log\left(IC_{50}^{-1}\right) = -8.510 + 2.313\omega - 10.984\omega^{+} - 3.13E_{NL}$$
$$n = 16, \quad r^2 = 0.702, \quad r^2_{CV} = 0.450,$$
$$SD = 1.00, \quad F = 9.43, \quad p < 0.05 \tag{13}$$

$$\text{Solvent phase: } \log\left(IC_{50}^{-1}\right) = -22.766 + 2.06\omega + 10.016\omega^{+} - 4.846E_{NL}$$
$$n = 16, \quad r^2 = 0.813, \quad r^2_{CV} = 0.595,$$
$$SD = 0.796, \quad F = 17.39, \quad p < 0.05 \tag{14}$$

The influence of solvent effect was very much prominent for this cell line. The r^2 and r^2_{CV} values (0.702 and 0.45, respectively) obtained in the gas phase increased to 0.813 and 0.595, respectively, with the inclusion of the solvent. Although the model in the solvent medium displayed acceptable statistical quality revealing the importance of the descriptors in the determination of biological activity of platinum complexes, the jackknife test may provide more insight in building more significant models for the cell line in both the media.

The complexes 4 and 6 with higher r^2_j values (0.745 and 0.748, respectively) in the gas phase and complexes 10 and 14 indicating higher r^2_j values (0.832 and 0.936, respectively) in the solvent phase could be considered as outliers (Table 1). The QSAR equations obtained after deleting these complexes are given in Table 2 along with statistically significant quantities. Importantly, for this cell line, the r^2 value 0.794 obtained after applying the jackknife test increased to 0.954 and r^2_{CV} value increased from 0.568 to a very acceptable value of 0.908 in solvent medium demonstrating the importance of the selected descriptors in the determination of log IC_{50}^{-1} values of platinum complexes. Autocorrelation coefficients among the descriptors of QSAR models (Table 2) are reasonable. We found that in gas phase E_{NL} had very low correlations with ω and ω^+ (<0.3) for both cell lines. Similarly, in solvent phase, ω had low correlations with E_{NL} and ω^+ (<0.5). We found slightly higher autocorrelation between ω and ω^+ in gas phase and E_{NL} and ω^+ in solvent phase. However, models

Table 2 QSAR models with the statistical parameters for two cancer cell lines in gas and solvent media

Cell line		QSAR equations	r^2	r^2_{CV}	SD	F
A2780	Gas phase	$\log\left(IC_{50}^{-1}\right) = -12.063 + 3.852\omega - 12.442\omega^{+} - 4.855E_{NL}$	0.859	0.748	0.863	20.30
	Solvent phase	$\log\left(IC_{50}^{-1}\right) = -16.087 - 0.333\omega + 14.742\omega^{+} - 3.566E_{NL}$	0.844	0.695	0.760	18.05
A2780Cp8	Gas phase	$\log\left(IC_{50}^{-1}\right) = -7.593 + 1.493\omega - 9.743\omega^{+} - 4.083E_{NL}$	0.794	0.568	0.905	12.91
	Solvent phase	$\log\left(IC_{50}^{-1}\right) = -23.184 + 2.129\omega + 9.717\omega^{+} - 5.11E_{NL}$	0.954	0.908	0.403	69.66

having descriptors with autocorrelation of about 0.8 have been reported for QSAR analyses [53].

In platinum drug-DNA binding, the DNA molecule acts as an electron donor whereas the complex is an electron acceptor and the mechanism involves the nucleophilic attack at Pt atom. In this type of interaction E_{LUMO} and E_{NL} play an important role. The lower values of these parameters increase the capability of the molecules to accept electrons from DNA making the system stable. We found that the coefficients of E_{NL} in all the QSAR equations (Table 2) were negative suggesting highly favorable intermolecular interactions between DNA molecule and the complex and an enhanced cytotoxic activity of the complex. The coefficients of other two independent factors (ω and ω^+), however, were not consistent in all equations. Importantly, the most significant model $\left(r^2 = 0.954 \text{ and } r^2_{CV} = 0.908\right)$ had positive coefficients for ω and ω^+. Thus, increasing their values can improve the anticancer activity. Although, all QSAR models are statistically significant, we found solvent phase derived model with $r^2 = 0.954$ and $r^2_{CV} = 0.908$, and gas phase derived model with $r^2 = 0.859$ and $r^2_{CV} = 0.748$ as the best models. The standard errors of regression coefficients (S_β) for two cancer cell lines in gas and solvent phases were calculated and provided as Supplementary Table a. We found that the best two models have lower values of S_β than that of other two models. The correlation plots between experimental and calculated log IC_{50}^{-1} values of the platinum complexes derived from these two QSAR models are shown in Fig. 2 which indicates that these descriptors can be effectively used in the prediction of cytotoxicity of platinum complexes.

QSPR analysis

We carried out QSPR studies for analyzing the log $P_{o/w}$ values of platinum complexes (1–24) for 0% (extrapolated), 20, 30, 40 and 50% MeOH. These log $P_{o/w}$ values were estimated by reversed-phase high performance liquid chromatography (RP-HPCL) technique [9]. Since the partition behavior markedly depends on the solvent, we also performed multiple regression analysis using solvent phase predicted molecular properties.

Multi-linear regression analyses were performed using the experimental log $P_{o/w}$ values for 0, 20, 30, 40, and 50% MeOH as a dependent variable and combination of four descriptors, namely electrophilicity (ω), philicity (ω^+), molar refractivity of carrier ligands (MR_{CL}), and surface area (SA) of the complexes as independent variables in gas and solvent models (Table 3). The QSPR equations obtained in both gas and solvent phases with r^2, r^2_{CV}, SD and F-values are listed in Table 4. As expected the solvent phase models displayed higher predictive power with r^2_{CV} values ranging from 0.914 for 0% MeOH to 0.795 for 50% MeOH. The statistical significance of the models of log $P_{o/w}$ in the solvent phase was similar for 0 and 20% MeOH but different in the gas phase. Although predictability of the models decreases for higher concentration, the solvent phase derived QSPR equations for 30–50% MeOH could predict the partition coefficient with the r^2_{CV} values in the range 0.867–0.796, respectively. Also, the standard errors of regression coefficients for QSPR models in all concentrations are lower in solvent phase than that in gas phase (Supplementary Table b). The sign of all descriptors in gas phase is consistent with that obtained in solvent phase, except for MR_{CL}. But we found positive correlation of log $P_{o/w}$ values of the complexes at all concentrations of MeOH with MR_{CL} while taking it as a single descriptor. The steric factor of the amine carrier ligands can be expressed by MR_{CL}. Thus, greater the steric effect of carrier amine ligands, greater might be the hydrophobicity of the complexes, in agreement with the observations by Platts et al. [9]. Importantly, the complexes (19, 20, and 22) with bulkiest amine carrier ligands and higher values of molar refractivity (38.22, 47.42, and 40.73, respectively) exhibited higher values of log $P_{o/w}$ (in all concentrations of MeOH). The plots between experimental and calculated values of log $P_{o/w}$ for 0, 20, 30, 40 and 50% MeOH predicted by gas and solvent phases presented in Fig. 3 suggest that the

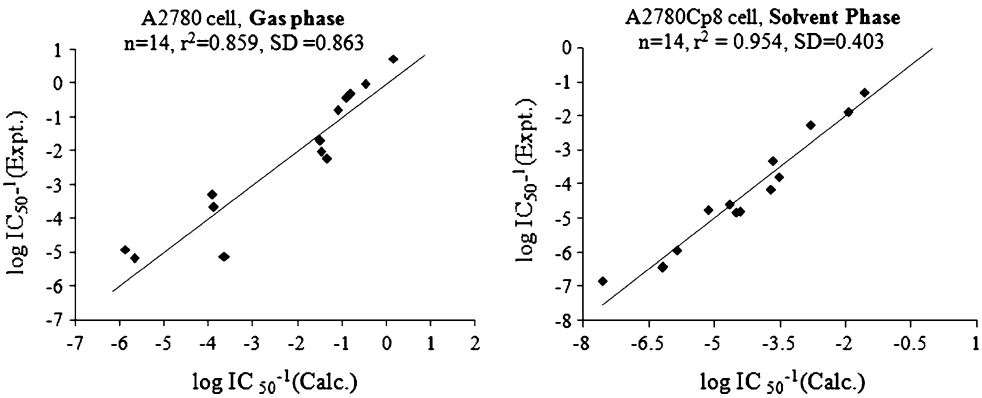

Fig. 2 Plots of experimental versus calculated values of cytotoxicity (log IC_{50}^{-1}) for two best models

Table 3 Parameters used to build the QSPR models for 24 platinum complexes in both gas and solvent phases

Complex	log $P_{o/w}$ (% MeOH)					Gas phase				Solvent phase			
	0	20	30	40	50	ω	ω^+	MR	SA	ω	ω^+	MR	SA
1	−2.27	−2.28	−2.08	−2.15	−2.09	4.134	0.868	4.8	142.2	3.884	1.056	4.8	144.64
2	−2.16	−2.20	−2.04	−2.08	−2.19	3.468	0.697	12.62	167.56	3.873	1.056	12.62	168.86
3	−0.85	−0.87	−0.90	−0.98	−0.97	3.353	0.624	32.21	242.64	4.107	0.920	32.21	243.6
4	−1.23	−1.24	−1.22	−1.32	−1.27	3.503	0.648	30.24	228.14	4.188	0.955	30.24	228.66
5	−2.32	−2.35	−2.13	−2.19	−2.20	3.585	0.649	4.8	172.88	2.956	0.834	4.8	176.75
6	−2.19	−2.25	−2.04	−2.15	−2.23	3.131	0.517	12.62	197.29	2.978	0.822	12.62	197.9
7	−1.17	−1.16	−1.26	−1.31	−1.27	2.819	0.555	32.21	275.92	3.078	0.605	32.21	277.82
8	−1.47	−1.44	−1.47	−1.46	−1.37	3.001	0.582	30.24	260.17	3.230	0.707	30.24	259.95
9	−1.63	−1.69	−1.60	−1.72	−1.80	3.368	0.546	4.8	217.47	2.820	0.663	4.8	223.54
10	−1.70	−1.70	−1.66	−1.67	−1.64	2.964	0.430	12.62	241.82	2.666	0.645	12.62	243.55
11	−0.47	−0.49	−0.63	−0.78	−0.75	2.673	0.513	32.21	320.62	2.907	0.610	32.21	323.19
12	−0.79	−0.81	−0.95	−1.05	−1.06	2.879	0.538	30.24	304.23	3.118	0.692	30.24	304.97
13	−1.39	−1.41	−1.37	−1.46	−1.43	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
14	−1.40	−1.41	−1.34	−1.41	−1.34	3.072	0.602	28.71	226.24	3.815	0.992	28.71	228.19
15	−1.37	−1.40	−1.35	−1.44	−1.43	2.749	0.412	28.71	253.08	2.966	0.813	28.71	257.17
16	−0.85	−0.88	−0.92	−1.02	−1.04	2.608	0.344	28.71	301.85	3.748	0.885	28.71	301.94
17	−1.94	−1.90	−1.93	−1.86	−1.79	3.654	0.735	14.59	187.72	3.289	0.783	14.59	190.17
18	−0.61	−0.54	−0.91	−0.65	−0.72	3.382	0.643	32.13	253.83	4.911	0.783	32.13	271.44
19	0.09	0.06	0.06	0.00	−0.05	3.500	0.668	38.22	283.27	5.621	1.360	38.22	283.88
20	1.06	1.07	1.14	1.16	1.19	3.411	0.645	47.42	315.4	5.431	1.350	47.42	315.18
21	−0.04	−0.18	−0.41	−0.78	−1.06	5.425	0.705	13.69	263.8	4.876	0.880	13.69	260.75
22	−0.04	−0.06	0.12	0.12	0.13	3.778	0.362	40.73	310.73	4.227	0.980	40.73	312.54
23	−0.46	−0.55	−0.64	−0.86	−1.01	3.934	0.377	31.42	267.82	4.137	1.170	31.42	268.01
24	−2.12	−2.25	−2.22	−2.38	−2.55	3.265	0.580	23.81	214.73	3.169	0.815	23.81	209.75

Table 4 QSPR models for 24 platinum complexes with the statistical parameters in gas and solvent media

	MeOH (%)	QSPR equations	r^2	r^2_{CV}	SD	F
Gas phase	0	$\log P_{o/w} = -7.253 + 0.559\omega + 0.812\omega^+ + 0.025MR_{CL} + 0.012SA$	0.914	0.798	0.274	50.63
	20	$\log P_{o/w} = -7.234 + 0.512\omega + 0.986\omega^+ + 0.026MR_{CL} + 0.012SA$	0.908	0.845	0.284	47.10
	30	$\log P_{o/w} = -6.473 + 0.502\omega + 0.648\omega^+ + 0.029MR_{CL} + 0.010SA$	0.86	0.734	0.329	29.17
	40	$\log P_{o/w} = -6.304 + 0.414\omega + 0.925\omega^+ + 0.033MR_{CL} + 0.009SA$	0.833	0.656	0.362	23.71
	50	$\log P_{o/w} = -6.071 + 0.312\omega + 1.140\omega^+ + 0.035MR_{CL} + 0.009SA$	0.812	0.602	0.388	20.53
Solvent phase	0	$\log P_{o/w} = -6.859 + 0.449\omega + 0.651\omega^+ - 0.013MR_{CL} + 0.015SA$	0.952	0.914	0.203	95.33
	20	$\log P_{o/w} = -6.802 + 0.449\omega + 0.591\omega^+ - 0.010MR_{CL} + 0.015SA$	0.951	0.917	0.206	93.63
	30	$\log P_{o/w} = -6.375 + 0.363\omega + 0.894\omega^+ - 0.008MR_{CL} + 0.013SA$	0.927	0.867	0.236	60.77
	40	$\log P_{o/w} = -6.232 + 0.369\omega + 0.790\omega^+ - 0.001MR_{CL} + 0.012SA$	0.916	0.855	0.257	51.82
	50	$\log P_{o/w} = -5.947 + 0.321\omega + 0.743\omega^+ + 0.005MR_{CL} + 0.011SA$	0.882	0.795	0.307	35.68

selected descriptors can be effectively used in the determination of log $P_{o/w}$ of the complexes.

We have also carried out QSPR analysis for an additional set of 20 complexes, whose log $P_{o/w}$ values were calculated by standard shake-flask method. The log $P_{o/w}$ values of complexes 1, 9, 13, 17, 19, 20, 21 and 25–40 (Fig. 1) were taken from results reported by Screnci et al. [8] and Souchard et al. [57] (both these papers reported values for 1). The complexes 1 (from reference [8]), 9, 13, 20 and 25–40 (a set of 20 compounds) were considered as a training set and the other four compounds (1, 17, 19, and 21) were treated as a test set. Table 5 lists the values of log $P_{o/w}$ and other descriptors derived from gas and solvent phases for the training set. The multi-linear regression analysis between log $P_{o/w}$ values of these 20 complexes and four descriptors yielded the QSPR equations as shown below

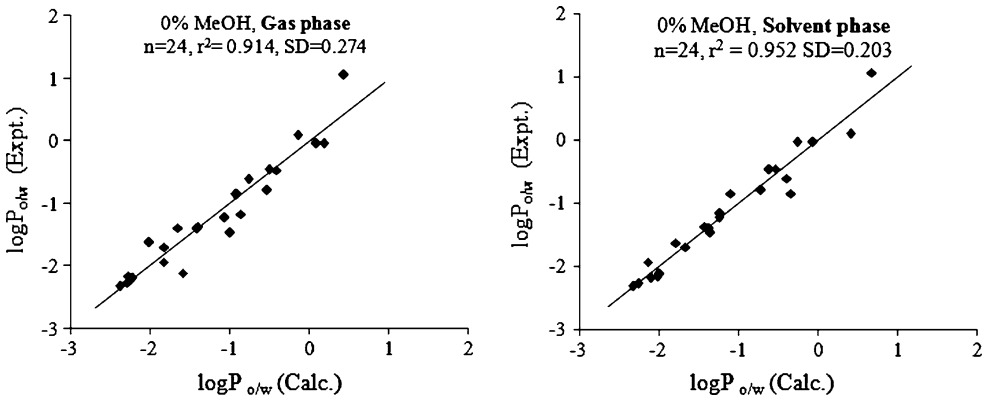

Fig. 3 Experimental versus calculated log $P_{o/w}$ values of 24 platinum complexes at 0% MeOH in gas and solvent media

$$\text{Gas phase: } \log P_{o/w} = -6.849 - 0.623\omega + 4.864\omega^{+} + 0.014\text{MR}_{\text{CL}} + 0.018\text{SA}$$
$$n = 20, \quad r^2 = 0.946, \quad r^2_{\text{CV}} = 0.913, \quad \text{SD} = 0.251, \quad F = 65.79 \qquad (15)$$

$$\text{Solvent phase: } \log P_{o/w} = -4.947 - 0.230\omega + 1.884\omega^{+} + 0.027\text{MR}_{\text{CL}} + 0.010\text{SA}$$
$$n = 20, \quad r^2 = 0.955, \quad r^2_{\text{CV}} = 0.920, \quad \text{SD} = 0.228, \quad F = 80.67 \qquad (16)$$

As expected, the solvent phase played an important role in improving the statistical quality of the model. The correlation plot between experimental and calculated log $P_{o/w}$ values in gas and solvent media presented in Fig. 4 indicates that the selected parameters can predict the hydrophobicity of platinum complexes with greater predictability in the solvent phase. The predicted log $P_{o/w}$ values of the compounds in the test set are presented in Table 6. The training set and test set were then combined and multi-linear regression analysis was performed on this data set of 24 complexes. The QSPR equations obtained in gas and solvent phases are reported below

Table 5 Parameters used to build the QSPR models for additional 20 platinum complexes in both gas and solvent phases

Complex	log $P_{o/w}$[a]	Gas phase				Solvent phase			
		ω	ω^{+}	MR_{CL}	SA	ω	ω^{+}	MR_{CL}	SA
1	−2.53	4.134	0.868	4.8	142.2	3.884	1.056	4.8	144.64
9	−2.30	3.368	0.546	4.8	217.47	2.820	0.663	4.8	223.54
13	−1.65	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
20	0.81	3.411	0.645	47.42	315.4	5.431	1.350	47.42	315.18
25	−3.36	5.283	0.718	4.8	183.13	7.948	0.540	4.8	224.13
26	−3.28	5.263	0.511	14.59	224.72	6.946	0.392	14.59	225.07
27	−1.71	4.727	0.397	38.22	323.12	6.967	0.385	38.22	318.91
28	−1.14	6.967	0.507	47.42	351.86	6.934	0.362	47.42	353.65
29	−0.91	6.649	0.390	56.62	386.53	8.755	0.359	56.62	397.07
30	−0.35	7.759	0.698	62.62	382.13	7.957	0.607	62.62	378.95
31	−2.13	8.821	0.547	26.46	372.26	8.083	0.225	26.46	375.7
32	−1.41	8.496	0.509	30.43	418.04	9.617	0.326	30.43	419.56
33	−1.59	6.447	0.621	13.69	308.75	7.304	0.727	13.69	315.54
34	−0.83	6.596	0.387	40.49	417.54	7.547	0.237	40.49	425.23
35	−1.06	4.818	0.384	22.68	346.64	6.933	0.643	22.68	349.61
36	−1.17	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
37	−1.18	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
38	−1.03	7.430	0.952	28.71	276.97	9.835	1.321	28.71	278.99
39	−1.59	2.753	0.347	28.71	237.77	4.357	0.388	28.71	237.18
40	−0.16	7.191	0.683	30.71	371.37	8.942	1.096	30.71	374.59

[a] Experimental log $P_{o/w}$ values obtained from Ref. [8, 57]

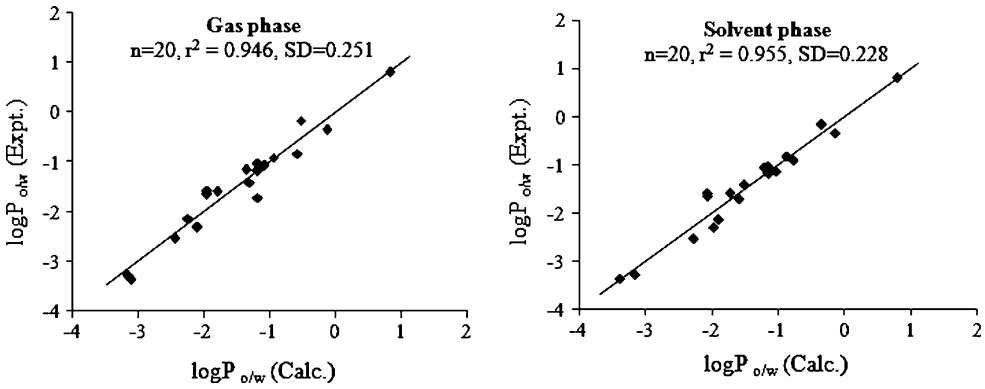

Fig. 4 Experimental versus calculated log $P_{o/w}$ values for training set of platinum complexes in both gas and solvent phases

Table 6 Experimental and predicted log $P_{o/w}$ values of four complexes in the test set

Complex	log $P_{o/w}$ (expt.)[a]	Gas phase		Solvent phase	
		log $P_{o/w}$ (calc.)[b]	Residual[c]	log $P_{o/w}$ (calc.)[d]	Residual[c]
1	−2.19	−2.443	0.253	−2.264	0.0742
17	−1.68	−1.785	0.105	−1.917	0.237
19	0.36	0.142	0.217	0.216	0.143
21	−0.32	−1.610	1.290	−1.412	1.092

[a] Experimental log $P_{o/w}$ values obtained from Ref. [8, 57]

[b] Predicted log $P_{o/w}$ values calculated using Eq. 15

[c] Difference between the experimental and calculated values of log $P_{o/w}$

[d] Predicted log $P_{o/w}$ values calculated using Eq. 16

$$\text{Gas phase: } \log P_{o/w} = -7.057 - 0.677\omega + 5.320\omega^{+} + 0.008\text{MR}_{CL} + 0.020\text{SA}$$
$$n = 24, \quad r^2 = 0.895, \quad r^2_{CV} = 0.842, \quad \text{SD} = 0.393, \quad F = 32.05 \qquad (17)$$

$$\text{Solvent phase: } \log P_{o/w} = -4.917 - 0.275\omega + 2.031\omega^{+} + 0.022\text{MR}_{CL} + 0.011\text{SA}$$
$$n = 24, \quad r^2 = 0.922, \quad r^2_{CV} = 0.882, \quad \text{SD} = 0.339, \quad F = 44.32 \qquad (17)$$

The sign of coefficients of all descriptors are same for both training set and data set in gas phase as well as in solvent phase. However, their signs for ω and MR_{CL} are different from those obtained in QSPR models for five different concentrations of MeOH. This inconsistency may be due to their log $P_{o/w}$ values calculated from different experimental techniques. The plots between experimental and calculated log $P_{o/w}$ values of the data set displayed a good correlation among them (Fig. 5). The correlation in the solvent phase was better than that of gas phase as expected. The standard errors of regression coefficients for QSPR models of both training and data set are lower in solvent phase than that in gas phase (Supplementary Table b). Together, these results demonstrate that the four descriptors (ω, ω^{+}, MR_{CL} and SA) can be satisfactorily used in the prediction of hydrophobicity of platinum complexes and the

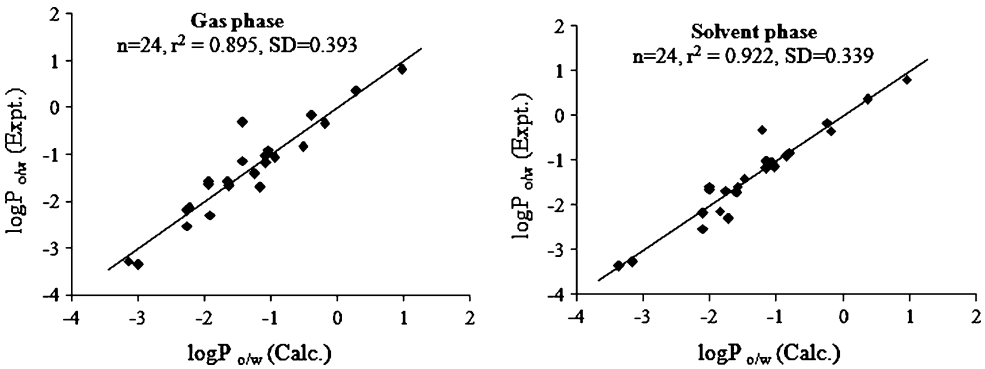

Fig. 5 Correlation plots between experimental and calculated log $P_{o/w}$ values of the data set obtained in gas and solvent phases

solvent model derived descriptors provide a better correlation.

Conclusions

The QSAR approach with three parameters, i.e., ω, ω^+ and E_{NL} provides regression models capable of predicting log IC_{50}^{-1} of *cis*-platinum complexes against A2780 and A2780Cp8 cancer cell lines. The jackknife test applied on the QSAR studies improved the statistical quality of the models in both gas and solvent phases. The inclusion of solvent medium increases the correlation coefficient ($r^2 = 0.954$) and cross-validated squared correlation $\left(r^2_{CV} = 0.908\right)$ for the A2780Cp8 cancer cell line suggesting the importance of solvent effect and significance of the selected descriptors. The QSPR equations modeled by ω, ω^+, MR_{CL}, and SA parameters against five different concentrations of MeOH (0–50%) are capable of predicting log $P_{o/w}$ values of 24 platinum complexes with r^2 values in the range of 0.914–0.812 (gas phase) and 0.952–0.882 (solvent phase) and r^2_{CV} values in the range of 0.798–0.602 (gas phase) and 0.914–0.795 (solvent phase), respectively. The solvent effect influences the QSPR model developed for 20 platinum complexes (training set) where calculated log $P_{o/w}$ values are in close proximity to their experimental values with $r^2 = 0.946$ (0.955) and $r^2_{CV} = 0.913(0.920)$ for gas (solvent) media. The predicted log $P_{o/w}$ values of four complexes in test set derived from the models are near to their corresponding experimental values, indicating significance of the selected descriptors in determination of hydrophobicity of platinum complexes. QSPR models for the data set also show good statistic qualities in both gas and solvent phases. Thus these descriptors emerged from DFT and MM+ methods can successfully be used to predict activity and hydrophobicity of platinum complexes. In summary, the current work clearly shows the importance of the selected parameters as well as solvent effect in the QSAR and QSPR analyses of *cis*-platinum complexes.

Acknowledgments The authors thank Council of Scientific and Industrial Research (CSIR), New Delhi and Department of Science and Technology (DST), New Delhi for some financial support. The authors also thank Mrs Surobhi Deka, Department of Mathematical Sciences, Tezpur University for fruitful discussion.

References

1. Rosenberg B, VanCamp L, Trosko JE, Mansour VH (1969) Nature 222:385. doi:10.1038/222385a0
2. Reedijk J (2003) Proc Natl Acad Sci USA 100:3611. doi:10.1073/pnas.0737293100
3. Wang D, Lippard SJ (2005) Nat Rev Drug Discovery 4:307
4. Jamieson ER, Lippard SJ (1999) Chem Rev 99:2467
5. Wong E, Giandomenico CM (1999) Chem Rev 99:2451. doi:10.1021/cr980420v
6. Lebwohl D, Canetta R (1998) Eur J Cancer 34:1522. doi:10.1016/S0959-8049(98)00224-X
7. Monti E, Gariboldi M, Maiocchi A, Marengo E, Cassino C, Gabano E, Osella D (2005) J Med Chem 48:857. doi:10.1021/jm049508t
8. Screnci D, McKeage MJ, Galettis P, Hambley TW, Palmer BD, Baguley BC (2000) Br J Cancer 82:966. doi:10.1054/bjoc.1999.1026
9. Platts JA, Oldfield SP, Reif MM, Palmucci A, Gabano E, Osella D (2006) J Inorg Biochem 100:1199. doi:10.1016/j.jinorgbio.2006.01.035
10. Wan J, Zhang L, Yang GF (2004) J Comput Chem 25:1827. doi:10.1002/jcc.20122
11. Srivastava HK, Pasha FA, Singh PP (2005) Int J Quantum Chem 103:237. doi:10.1002/qua.20506
12. Karelson M, Lobanov VS (1996) Chem Rev 96:1027. doi:10.1021/cr950202r
13. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512
14. Parr RG, Donnelly RA, Levy M, Palke WE (1978) J Chem Phys 68:3801
15. Parr RG, Szentpaly LV, Liu S (1999) J Am Chem Soc 121:1922
16. Parr RG, Yang W (1984) J Am Chem Soc 106:4049
17. Chattaraj PK, Maiti B, Sarkar U (2003) J Phys Chem A 107:4973
18. Chatterjee A, Balaji T, Matsunaga H, Mizukami F (2006) J Mol Graph Model 25:208
19. Roos G, Loverix S, De Proft F, Wyns L, Geerlings P (2003) J Phys Chem A 107:6828
20. Parthasarathi R, Subramanian V, Roy DR, Chattaraj PK (2004) Bioorg Med Chem 12:5533
21. Padmanabhan J, Parthasarathi R, Subramanian V, Chattaraj PK (2006) Bioorg Med Chem 14:1021
22. Padmanabhan J, Parthasarathi R, Subramanian V, Chattaraj PK (2006) Chem Res Toxicol 19:356
23. Wysokiński R, Michalska D (2001) J Comput Chem 22:901
24. Michalska D, Wysokiński R (2005) Chem Phys Lett 403:211
25. Zhang Y, Guo Z, You X-Z (2001) J Am Chem Soc 123:9378
26. Costa LAS, Rocha WR, De Almeida WB, Dos Santos HF (2003) J Chem Phys 11:10584
27. Matsui T, Shigeta Y, Hirao K (2006) Chem Phys Lett 423:331
28. Robertazzi A, Platts JA (2006) Chem Eur J 12:5747
29. Burda JV, Leszczynski J (2003) Inorg Chem 42:7162
30. Raber J, Zhu CB, Eriksson LA (2005) J Phys Chem B 109:11006
31. Magistrato A, Ruggerone P, Spiegel K, Carloni P, Reedijk J (2006) J Phys Chem B 110:3604
32. Mantri Y, Lippard SJ, Baik M-H (2007) J Am Chem Soc 129:5023
33. Platts JA, Hibbs DE, Hambley TW, Hall MD (2001) J Med Chem 44:472
34. Sarmah P, Deka RC (2008) Int J Quantum Chem 108:1400
35. Koopmans TA (1933) Physica 1:104
36. Mendez F, Gazquez JL (1994) J Am Chem Soc 116:9298
37. Yang W, Mortier WJ (1986) J Am Chem Soc 108:5708
38. Delley B (1990) J Chem Phys 92:508
39. Becke AD (1988) Phys Rev A 38:3098
40. Lee C, Yang W, Parr RG (1988) Phys Rev 37:785
41. Hehre WJ, Radom L, Schlyer PVR, Pople JA (1986) Ab Initio molecular orbital theory. Wiley, New York
42. Hirshfeld FL (1977) Theor Chim Acta 44:129
43. Andzelm J, Koelmel C, Klamt A (1995) J Chem Phys 103:9312
44. HyperChem Release 7 (2002) Hypercube; http://www.hyper.com
45. MATLAB (1999) The Math Works, Inc., Natick, USA
46. Penrose R (1955) Proc Cambridge Philos Soc 51:406
47. Milburn GHW, Truter MR (1966) J Chem Soc A 1609. doi:10.1039/J19660001609

48. Wysokiński R, Michalska D (2001) J Comput Chem 9:901
49. Soltzberg L, Margulis TM (1971) J Chem Phys 55:4907
50. Beagley B, Cruickshank DWJ, McAuliffe CA, Pritchard RG, Zaki AM, Beddoes RL, Cernik RJ, Mills OS (1985) J Mol Struct 130:97
51. Bruck MA, Bau R (1984) Inorg Chim Acta 92:279
52. Cho DH, Lee SK, Kim BT, No KT (2001) Bull Korean Chem Soc 22:388
53. Yao SW, Lopes VHC, Fernandez F, Garcia-Mera X, Morales M, Rodriguez-Borges JE, Cordeiroa MNDS (2003) Bioorg Med Chem 11:4999
54. Wold S (1991) Quantum Struct-Act Relat 10:191
55. Dietrich SW, Dreyer ND, Hansch C, Bentley DL (1980) J Med Chem 23:1201
56. Cornish-Bowden A, Wong JT (1978) Biochem J 175:969
57. Souchard JP, Ha TTB, Cros S, Johnson NP (1991) J Med Chem 34:863