Applied Clay Science xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Applied Clay Science

journal homepage: www.elsevier.com/locate/clay

Research paper

Modeling binding of organic pollutants to a clay–polycation adsorbent using quantitative structural–activity relationships (QSARs)

Adi Radian [a,*], Merav Fichman [b], Yael Mishael [a]

[a] Dept. of Soil and Water Sciences, Robert H. Smith Faculty of Agriculture, The Hebrew University in Jerusalem, Rehovot, Israel
[b] Institute of Biochemistry, Food Science and Nutrition, Robert H. Smith Faculty of Agriculture, The Hebrew University in Jerusalem, Rehovot Israel

ARTICLE INFO

Article history:
Received 22 October 2014
Received in revised form 1 March 2015
Accepted 3 March 2015
Available online xxxx

Keywords:
Clay–polymer nanocomposites (CPNs)
Adsorption
Pollutant removal
QSAR

ABSTRACT

The adsorption of organic pollutants to a novel adsorbent–polyvinyl-pyridine-*co*-styrene–montmorillonite nanocomposite was quantified and modeled. To elucidate the adsorption mechanisms, experimental methods and QSAR analysis were combined, searching for correlations between the pollutant-nanocomposite adsorption coefficient (k_d) and pollutant chemical–physical properties. The adsorption isotherms at a wide range of concentrations were fitted to the Freundlich equation and the logk_d values were extracted at a low, environmentally significant, concentration. A significant regression was achieved with QSAR, predicting adsorption affinity by four meaningful descriptors: adsorption was positively correlated to heat of formation, number of hydrogen acceptor groups and the partitioning coefficient, and was negatively correlated to molecular mass. The resulting model predicted logk_d for test pollutants with an average deviation of only 0.77 log units from the experimental values. Consequently, this method could be applied to better understand adsorption mechanisms and to screen for compatibility between pollutants and a variety of novel and commonly used adsorbents.

1. Introduction

Developing advanced materials for enhanced adsorption of organic pollutants has become a major research focus due to the depletion of water sources worldwide (Cerejeira et al., 2003; Falta, 2004; Deblonde et al., 2011). Clay–polyelectrolyte nanocomposites are one example of hybrid materials specially tailored for the adsorption of a wide range of pollutants (Breen and Watson, 1998; Churchman, 2002; Zadaka et al., 2009; Ganigar et al., 2010; Radian et al., 2010). The advantage of these materials lies in the large surface area of the clay mineral coupled with a polymer coating that has a specific affinity for the target pollutants. Recently, the efficient removal of atrazine (Zadaka et al., 2009) and pyrene (Radian and Mishael, 2012) by a montmorillonite–poly-4-vinylpyridine-co-styrene (Mt–HPVPcoS) nanocomposite in the presence of dissolved organic matter (DOM), was reported. These studies demonstrated the advantage of the Mt–HPVP*co*S nanocomposite, as an adsorbent, over Mt–Polydiallyldimethylammonium chloride (PDADMAC) nanocomposites and granulated activated carbon (GAC) resulting from specific chemical interactions with the functionalized surface of the nanocomposite.

One of the major problems in the development of such tailored materials is choosing the optimal modifier (in our case polymers) for maximal pollutant adsorption. This requires an in depth understanding of the adsorption mechanism of organic pollutants to the new heterogeneous clay–polymer surface. This adsorption process can be driven by a number of different mechanisms: the capacity of the molecule to leave the aquatic medium, physical access to the adsorption sites ("cavity size") and the ability to form specific bonds with the surface (Kowalska et al., 1994; Goss and Schwarzenbach, 2000). Previous studies suggested that the multifunctional structure of the chosen co-polymer, HPVP*co*S, enabled van der Waals, hydrogen and π–π bond formation leading to rapid adsorption with high efficiency (Zadaka et al., 2009; Ganigar et al., 2010; Radian and Mishael, 2012).

In the current study we aimed to better understand the adsorption of micro-organic pollutants to the Mt–HPVPcoS nanocomposite and to develop a predictive adsorption model. For that purpose, the adsorption of 30 organic pollutants (from industrial, agricultural and urban origins) to the clay polymer nanocomposite (CPN) was examined. A statistical method — Quantitative Structural Activity Relationships (QSARs), was employed to correlate between the experimental adsorption behavior (k_d coefficients) and a variety of computationally calculated physical–chemical pollutant structural properties.

A specific form of QSAR, Linear Solvation Energy Relationship (LSER), is commonly used to describe the partitioning of organic compounds between two phases in terms of the main energy contributions of intermolecular interactions. This approach is unique in the QSAR discipline due to the built in mechanistic interpretation of five experimentally derived descriptors (Abraham descriptors). The LSER method has been widely employed by environmental chemists to predict pollutant

* Corresponding author at: The Robert H. Smith Faculty of Agriculture, Food and Environment, P.O. Box 12, Rehovot 76100, Israel. Tel.: +972 8 9489171; fax: +972 8 9475181.
E-mail address: adi.radian@mail.huji.ac.il (A. Radian).

http://dx.doi.org/10.1016/j.clay.2015.03.021
0169-1317/

partitioning parameters such as the soil organic content partitioning coefficient-k_{oc}, the octanol water partitioning coefficient-k_{ow}, and solubility (Tülp et al., 2008; Satoshi et al., 2009; Bronner and Goss, 2010). Additionally, LSER has been employed to improve the mechanistic understanding of pollutant adsorption to the commonly applied adsorbent, GAC (Shih and Gschwend, 2009).

Even though LSER is commonly used and Abraham descriptors for many compounds are available (Abraham et al., 1994; Tülp et al., 2008; Bronner and Goss, 2010), there is a significant lack of descriptors for complex compounds, their values from different sources do not agree and their use in describing complex systems is questionable (Platts et al., 1999; Endo and Schmidt, 2006). General QSAR has no predetermined mechanistic interpretation and can be applied to diverse systems and molecules.

QSAR modeling of k_{oc} for pollutants in various soils has been reported (Gramatica et al., 2000; John, 2004). Gramatica et al. (2000) correlated the properties of 185 non-ionic organic pesticides to experimental logk_{oc} which were extracted from the literature. These QSAR calculations provided models with good predictive performances, giving an R^2 of 0.84 for heterogeneous pesticides and higher still for class-specific models. Size and electronic-descriptors were shown to be most significant; generally, adsorption of high molecular mass and low polarity compounds was favorable.

The current study employs experimental methods combined with the statistical methods described above (LSER and QSAR), to produce a descriptive and predictive adsorption model for organic pollutants to the CPN. Adsorption data for a large and diverse set of chemicals was collected, which made building a comprehensive model much more complex. To overcome this, a uniform activity adhering to low concentration (relevant in environmental systems) of pollutants was chosen. The results suggest that such QSAR modeling can aid in understanding pollutant–adsorbent compatibility which is key in any adsorbent material development. Although this QSAR study was carried out for a specially designed adsorbent under specific conditions, this methodology can be adapted for a variety of adsorbents and solutions with different pH, ionic strength, temperatures etc.

2. Experimental

2.1. Materials

Wyoming sodium-montmorillonite (Mt SWy-2) was purchased from the Source Clays Repository of the Clay Mineral Society (Columbia, MO). Poly (4-vinylpyridine-co-styrene) (PVP-co-S: MM—1,200,000–1,500,000; PVP-to-S ratio 9:1) was purchased from Sigma-Aldrich (Steinheim, Germany).

Organic pollutants: *Industrial pollutants*: picric acid (1.2%) was obtained from Ridley (Riedel-de Haën, Germany). Toluene acid and benzoic acid were purchased from Sigma Aldrich (Steinheim, Germany). Eosin blue dye was purchased from The National Aniline Division, Allied Chemicals NY, USA. *PAHs*: Pyrene, phenanthrene, naphthalene and acenaphthene were purchased from Sigma Aldrich (Steinheim, Germany). *Pesticides*; atrazine, ametryn, prometryn, terbuthylazine, simazine, imazapyr, imazaquin, metaldehyde, diuron, diazinon, bromacil, metolachlor and metazachlor (all compounds of 98% purity) were supplied by Agan Makhteshim, Israel. Technical sulfentrazone (purity 91.3%) was obtained from FMC (Princeton, NJ) and pyrithiobac was obtained from Godrej Agrovet, Mumbai, India. *Pharmaceuticals*: Carbamazepine, diclofenac, ketoprofen, ibuprofen, gemfibrozil, paracetamol and clofibric acid were purchased from Sigma Aldrich (Steinheim, Germany).

2.2. Methods

2.2.1. Overview

The models that were built are a linear relation between a set of pollutant molecular descriptors and the experimentally measured log k_d values for the molecules. The derived coefficient values describe the effect of each descriptor on the adsorption affinity: A positive coefficient means the descriptor adds to pollutant–CPN adsorption affinity whereas a negative coefficient means the descriptor inhibits adsorption.

2.2.2. Preparation of Mt–HPVPcoS nanocomposites

The preparation of Mt–HPVPcoS nanocomposites has been previously described in detail (Zadaka et al., 2009; Radian and Mishael, 2012). Briefly, the PVPcoS polymer was dissolved in an acidified solution by adding stoichiometric concentrations of H_2SO_4 (95%) to protonate the pyridine rings (H-PVPcoS). Mt dispersion of 0.05% w/w (0.5 g/L, 4 mL) was added to 1 g/L H-PVPcoS solution (8 mL) and agitated for 2 h. The resulting CPN has 0.2 g H-PVPcoS adsorbed per g Mt.

2.2.3. X-ray diffraction measurements

X-ray diffraction measurements were done to evaluate the change in basal spacing of Mt as a function of polymer loading (0.1 and 0.2 g/g). The CPNs were measured by X-ray powder diffraction on air-dried, oriented samples using a Philips PW1830/3710/3020 diffractometer (Cu Kα radiation, $\lambda = 1.526$). Oriented samples were prepared by placing 1 mL of the dispersion onto a glass slide and left to settle for 1 day.

2.2.4. Pollutant adsorption to the CPN and calculation of k_d values

A 12 mL solution of each pollutant, ranging in concentrations from 0.05 to 50 mg/L (or up to solubility), was added to 1.85 g/L Mt–HPVPcoS nanocomposite in glass centrifuge tubes and Teflon caps and agitated for 72 h. To determine pollutant adsorption, supernatants were separated by centrifugation (4000 rpm for 20 min) and pollutant concentration was measured. The initial pH of the dispersion was measured and found to be on average 3.5 (due to the acidic nature of the polymer). The detection method was chosen depending on the compound (Supporting information Table A1): UV–Vis spectroscopy (Thermo Scientific, Evolution 300, Waltham, MA, USA), fluorescence measurements (Cary Eclipse, Varian fluorescence spectrophotometer, Agilent technologies), GC–MS (Finnigan Polaris/GCQ Plus) or HPLC measurements (Agilent Technologies 1200 series equipped with a diode-array detector, HPLC column was LiChroCART 250-4 PurosphereR STAR RP-18 (5 mm)).

The results were plotted as isotherms and fitted to the Freundlich equation. The Freundlich equation (Eq. (2.1)) relates the concentration of solute adsorbed on the surface (Y axis : C_{ads} (mmol/Kg)) to the concentration remaining in the solution (X axis : C_{eq} (mg/L)).

$$\begin{aligned}&\text{Freundlich equation}: C_{ads} = k_f * C_{eq}^{n}\\&\text{Adsorption coefficient at a specific concentration}: k_d = k_f * C_{eq}^{(n-1)}\end{aligned} \tag{2.1}$$

The adsorption isotherms were measured at a wide range of concentrations (reaching high concentrations not likely to be found in the environment) and yielded non-linear isotherms. Presenting a wide range of concentrations emphasizes the different adsorption mechanisms and behavioral trends of the chosen pollutants. Upon restricting the isotherms to low concentrations the linear equation is obtained with n ~ 1. To adequately compare the relevant adsorption affinity, the molecule's adsorption coefficients (k_d) were calculated at low, environmentally significant, concentrations (0.001 activity) (chemical activity was calculated by dividing the concentration (C_{eq}) by the compound's water solubility (s_w)). At the low concentration tested (0.001 activity), logk_d tends to become a constant and the isotherms approach linearity (McGuire and Suffet, 1978; Suffet, 1980).

Statistical analysis (ANOVA) was performed on the fit to the Freundlich model and only statistically significant results ($p < 0.05$) were used (Table 1).

2.2.5. LSER model

Correlation between the extracted logk_d values and the adsorbent's properties were modeled using LSER. The Abraham descriptors for the

Please cite this article as: Radian, A., et al., Modeling binding of organic pollutants to a clay–polycation adsorbent using quantitative structural–activity relationships (QSARs), Appl. Clay Sci. (2015), http://dx.doi.org/10.1016/j.clay.2015.03.021

Table 1
Freundlich coefficients (Eq. (2.1)) of pollutants on Mt–HPVPcoS nanocomposite (0.18 g/g).

Name	log k_f	n	log k_d (activity 0.001)	R^2	P value (ANOVA)
Acenaphthylene	4.00	0.96	4.19	1.000	<0.0001
Ametryn	2.67	0.39	3.94	0.958	<0.0001
Atrazine	2.85	0.56	4.52	0.998	<0.0001
Benzoic acid	1.87	0.30	2.14	0.980	0.0004
Bromacil	1.64	0.77	2.22	0.999	<0.0001
Carbamazepine	2.73	1.20	1.89	0.980	<0.0001
Clofibric acid	2.07	0.71	2.80	0.999	<0.0001
Diazinon	3.43	0.66	4.76	0.954	0.0027
Diclofenac	1.81	0.20	5.88	0.864	0.0481
Diuron	2.24	0.87	2.74	0.999	<0.0001
Eosin B	2.84	0.43	3.45	0.999	0.0077
Gemfibrozil	1.87	0.43	4.14	0.952	0.0029
Ibuprofen	2.70	0.84	3.33	0.9945	0.0002
Imazapyr	1.73	0.84	1.96	0.998	<0.0001
Imazaquin	3.08	0.94	3.30	0.999	<0.0001
ketoprofen	2.54	0.82	3.20	0.992	0.0003
Metaldehyde	2.93	1.55	1.32	0.999	<0.0001
Metazachlor	2.92	1.25	2.22	0.999	<0.0001
Metolachlor	2.91	1.12	2.58	1.000	<0.0001
Naphthalene	3.26	1.25	2.37	0.997	0.0208
Paracetamol	1.63	0.70	1.96	0.998	<0.0001
Phenanthrene	5.15	1.12	4.54	0.998	<0.0001
Picric acid	4.49	0.25	5.40	0.960	0.0018
Prometryn	5.12	1.24	4.19	1.000	0.0004
Pyrene	3.20	0.77	4.64	0.9991	<0.0001
Pyrithiobac	3.02	0.74	2.95	0.997	<0.0001
Simazine	2.08	0.33	5.17	0.980	0.0002
Sulfentrazone	1.87	0.76	2.53	0.999	<0.0001
Terbuthylazine	2.83	0.94	3.10	0.997	<0.0001
Toluene	2.81	0.89	3.05	0.980	<0.0001

compounds used were collected from literature (Supporting Information, Table A2). For six of the compounds, experimental descriptors could not be found and were therefore calculated by a fragment based method (Andrew, 2012). It is worth noting that the ionic descriptors (j−, j+) were not used because the low pH rendered the bulk of the compounds neutral (Abraham and Acree, 2010; Stephens et al., 2011).

The LSER model was developed by Multiple Linear Regression (MLR) using Accelrys Materials Studio software. The t-probability calculated by the software represents the confidence of the derived coefficients (a t-probability of 0.05 indicates that a variable is significant at the 95% level). R^2 was used as a measure of the total variance of the response explained by the regression model (i.e. goodness-of-fit) and F probability represents the significance of the regression model (an F-probability of 0.05 indicates that a variable is significant at the 95% level).

2.2.6. QSAR model

2.2.6.1. QSAR descriptor calculations. To calculate the QSAR descriptors the chemical structures of all 30 molecules were drawn manually in Accelrys Materials Studio software and minimized to their lowest energy conformations using the Forcite classical simulation engine and the COMPAS force-field. It should be noted the all molecules are drawn with the respective hydrogens (not dissociated), which allows even-handed molecular calculations and is in agreement with the low pH of our given system. A set of 86 theoretical molecular descriptors for the 30 molecules were calculated by Accelrys Materials Studio and Discovery Studio (Table A3). Only theoretically calculated descriptors were chosen because they are produced in the same manner and pose an advantage over experimentally derived descriptors that rely on diverse experimental methods. Furthermore, to minimize the importance of pollutant conformation, 2-D descriptors were preferred (e.g. MM vs volume). The descriptor values across the 30 molecules were inter-correlated and ~~in~~ pairs of descriptors found to be highly correlated (Pearson $\rho > 0.9$), one was excluded from the set to reduce redundancy, leaving a set of 48 descriptors.

2.2.6.2. Model preparation. The model was prepared in two steps, initially all 30 molecules were used to find a subset of descriptors out of the initial set (86 descriptors). This subset selection of the descriptors was carried out by running a multiple linear regression and restricting the program to choose a subset of five or less descriptors for the model (≤Abraham descriptors) to avoid over fitting (the fewer the descriptors, the more significant the model). All MLR runs were carried out by the Materials Studio built in Genetic Function Algorithm. This initial model was created many times and the prevalent descriptors (MM, AlogP98, H-acceptor and heat of formation) were noted. These four descriptors create a highly significant model with very low t probability of the coefficients (t probability between $0.03–1.1 * 10^{-6}$).

To improve the initial QSAR model, the 30 molecules were split into training and test sets with 70/30 proportion respectively. 10 splits were performed resulting in 10 training and test sets: 5 splits relied on sorting the molecules according to logk_d and placing high and low values in both sets. 5 splits were sorted according to structural similarity (Gramatica et al., 2012). Training sets were used for model development, and test sets for model internal validation.

MLR was performed separately on each training set, with the predefined subset of descriptors from the initial model. This resulted in 10 models, i.e., 10 sets of 4 coefficients, each with R^2 and R^2_{cv} values.

Each coefficient in the final QSAR model is the average of the respective coefficient from the 10 training models. Final model R^2 and R^2_{cv} are the average of R^2 and R^2_{cv} from the 10 training models. It is worth noting that an external validation, using a larger data set which was not involved in building the model, was not possible since this is a unique adsorbent and no external data from literature is available.

3. Results and discussion

3.1. Experimental adsorption isotherms

The Mt–HPVPcoS nanocomposite was chosen as a model adsorbent due to previous data suggesting high adsorption capacity for a range of environmentally relevant pollutants (Zadaka et al., 2009; Ganigar et al., 2010; Radian and Mishael, 2012). The CPN has a loading of 0.2 g polymer per g clay – which covers roughly 70% of the specific clay surface area, and a basal spacing of 1.48 nm – which is equivalent to a thickness of one polymer chain adsorbed in the interlamellar region (Figure A1). A further increase in the basal spacing was not observed upon pollutant (atrazine) adsorption suggesting that the pollutants adsorb on the external surfaces of the composite.

The adsorption of 30 organic pollutants to the Mt–HPVPcoS nanocomposites was measured and the results were fitted to the Freundlich equation. The isotherm trends (Figure A2) evolve from S curves for the low affinity pollutants, through linear C curves for the moderate affinities and L and H curves for the high affinities (Freundlich n coefficient range 0.2 to 1.55).

To adequately compare relevant adsorption affinities logk_d values for all the compounds were calculated at the very low specific activity of 0.001 (Table 1). The low activity was chosen to be within the linear range of the Freundlich isotherms where n ~ 1 (McGuire and Suffet, 1978; Suffet, 1980).

The results show the diverse pollutant behaviors, ranging from relatively low adsorption (methaldehyde logk_d 1.3 and carbamazepine logk_d 1.9) to high affinity and compatibility (picric acid logk_d 5.4 and diclofenac logk_d 5.9). Despite the differences, most of the pollutants showed high adsorption to the CPN (22 pollutants have a log$k_d > 2.5$) compared to previous studies involving organo-clays and activated carbon (Borisover et al., 2001; Qu et al., 2008; Kong et al., 2011).

Please cite this article as: Radian, A., et al., Modeling binding of organic pollutants to a clay–polycation adsorbent using quantitative structural–activity relationships (QSARs), Appl. Clay Sci. (2015), http://dx.doi.org/10.1016/j.clay.2015.03.021

The adsorption isotherms did not correlate to any one specific chemical or physical trait of the pollutants; therefore, statistical methods correlating the adsorption behavior to numerous pollutant physical–chemical properties were employed.

3.2. Linear solvation energy relationship

The classic LSER method is described by Eq. (3.1): log $k_{i1,2}$ is the partitioning coefficient and the five well known Abraham descriptors quantify the molecular interactions: E is the excess molar refraction, S is the solute polarizability, A is the effective hydrogen bond acidity, B is the hydrogen bond basicity and V is the McGowan solute volume. The corresponding phase descriptors describe the difference in capacity between the phases where a positive component suggests a stronger affinity to the second phase (log$k_{i1,2}$). The ionic descriptors (j−, j+) were not included since the pH of the solution was approximately 3.5 which is below the pKa of most of the pollutants.

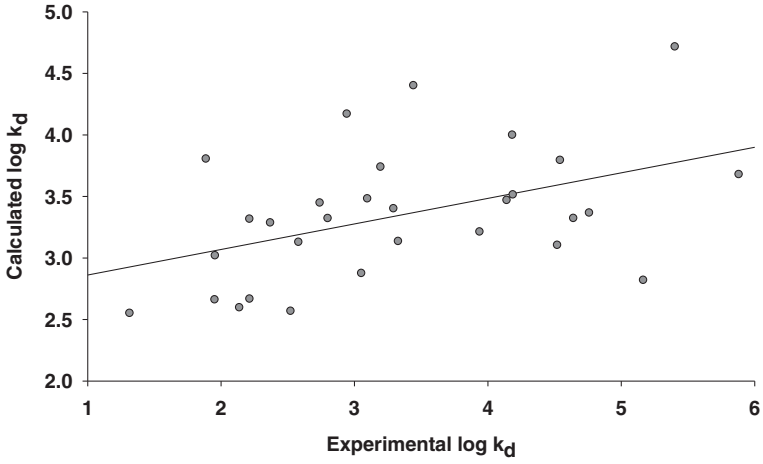
$$\log k_{i1,2} = e_{1,2}E_i + s_{1,2}S_i + a_{1,2}A_i + b_{1,2}B_i + v_{1,2}V_i + C_{1,2} \quad (3.1)$$

Although LSER calculations are commonly applied and have shown good results in the past, the resulting regression was insignificant and did not describe the adsorption phenomenon in this case (Fig. 1: F-probability for the regression = 0.31 and a correlation factor of $R^2 = 0.2$). Several factors may explain this result: First, the set of molecules is very diverse, making it difficult to get a clear trend (for specific groups different effects may contradict, for one group the size seems to enhance adsorption yet for another group it hinders it). Second, the Abraham descriptors for large compounds with several functional groups may be problematic since they do not encompass intramolecular interactions (Platts et al., 1999). Third, the LSER method was initially created for adsorption by a partitioning mechanism where a linear Freundlich trend is observed ($n \approx 1$). This is not the case here; the adsorption to the CPN showed a multitude of trends ($0.2 < n < 1.55$) and is probably a multi-mechanism system. Consequently, the more general QSAR statistical tool was used, it has no predetermined mechanistic interpretation and can encompass a wider range of descriptors to generate a more relevant model.

3.3. Quantitative structural activity relationships (QSARs)

The model building method is described in detail in the Materials section and Methods section. The resulting model (Eq. (3.2)) is highly significant with F-probability of $1.2 * 10^{-5}$, an average R^2 of 0.7 and a cross validation R^2_{cv} of 0.6 (Fig. 2). Some previous LSER and QSAR studies (k_{oc} in various soils and k_d on activated carbons) have presented models with somewhat higher R^2, but those models were mostly class-specific, and often obtained from a smaller set of molecules (Gramatica et al., 2000; Clarke, 2009; Shih and Gschwend, 2009). In the model suggested in this study the average deviation from the experimental logk_d values is only 0.77 log units, which indicates that the basic affinity for an untested pollutant can be established and that the model

A



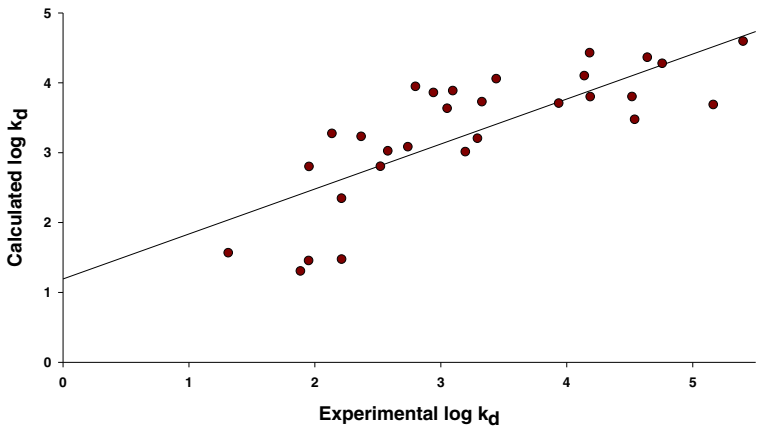
B

Abraham descriptors	Coefficient	Standard Error	t-probability of coefficients	Regression Validation
E	-0.17	0.24	0.69	R^2=0.21
S	1.01	0.03	0.08	F probability =0.31
A	-0.71	0.45	0.35	
B	-1.10	0.69	0.08	
V	0.88	1.20	0.12	
Constant	1.85	1.00		

Fig. 1. (A) plot of experimental logk_d vs the calculated logk_d from the MLR correlating to the five Abraham descriptors. (B) Statistical analysis of regression coefficients.

Please cite this article as: Radian, A., et al., Modeling binding of organic pollutants to a clay–polycation adsorbent using quantitative structural–activity relationships (QSARs), Appl. Clay Sci. (2015), http://dx.doi.org/10.1016/j.clay.2015.03.021

A



B

Descriptors	Coefficient	Standard Error	t-probability of coefficients	Regression validation
Total molecular mass	-2.43	0.17	0.000010	R^2=0.69
Hydrogen bond acceptor	0.75	0.03	0.000007	R_{cv}^2= 0.58
AlogP98	1.25	0.07	0.000001	F probability =1.2*10^5
Heat of Formation	0.39	0.17	0.034470	
Constant	2.75	0.72		

Fig. 2. (A) Plot of experimental logk_d vs the calculated logk_d from the MLR correlating relation to four QSAR descriptors. (B) Statistical analysis of regression coefficients.

can be used to screen for the pollutants that will be adequately removed by this novel adsorbent.

$$\log k_d = -2.43(\pm 0.31) * \mathrm{MM} + 0.75(\pm 0.08) * \mathrm{H}_{\mathrm{acceptor}} + 1.25(\pm 0.17) * \mathrm{AlogP98} + 0.39 * (\pm 0.1)\mathrm{Heat\ of\ Formation} + 2.75(\pm 0.6) \quad (3.2)$$

The chosen subset of physical/chemical descriptors: Molecular mass (MM), AlogP98 (calculated partition coefficient), number of H-acceptor groups and the heat of formation; represent the most dominant properties of the compound that affect adsorption.

High MM (coefficient = −2.34) was found to decrease adsorption affinity. This is probably linked to the configuration of the polymer on the clay mineral surface (and perhaps interlayer) which is more easily penetrated by smaller, less bulky molecules. For example, the s-triazine herbecides; simazine, atrazine and terbuthylazine, (differing by one methyl group respectively) showed a high adsorption affinity to the Mt–HPVPcoS nanocomposite (logk_d: simazine 5.1 > atrazine 4.92 > terbuthylazine 4.06). The speculated adsorption mechanism according to previous studies was primarily hydrogen and VdW bonding (Celis et al., 1997). Indeed all three herbicides have the same number of hydrogen-bond acceptor groups and H-bonding was found significant (Table 2). However, adsorption coefficients of the herbicides decreased with the increase of methyl groups and the calculated partitioning coefficient (log k_{ow}) negatively correlated to the adsorption coefficient. In this case, the MM hindered adsorption, this suggests that cavity formation and size on the surface play an important role; simazine which is the smallest molecule, is the least bulky and rigid and therefore has more adsorption sites accessible to it.

Hydrogen bond acceptor groups enhance adsorption (coefficient = 0.75) due to the strong hydrogen donor properties of the polymer's pyridine ring. Looking at the different descriptors of the pollutants (Table 2), the number of hydrogen bond acceptor groups is generally linked to high log k_d values (e.g. picric acid and Diazinon). Generally, compounds with more than 4 hydrogen bond acceptors can be characterized as high affinity compounds. Indeed, the adsorption of the s-triazines to the Mt–HPVPcoS nanocomposite was consistently higher compared to previous studies employing different organo-clays without the ability to form such H-bonds (e.g. Atrazine uptake by a dye organo-clay has a log k_d of 2.9, and under comparable conditions the uptake by the Mt–HPVPcoS nanocomposite showed a log k_d of 4.8) (Borisover et al., 2001).

The partitioning coefficient AlogP98 is an improved Ghose–Crippen additive atom-based method for calculating log k_{ow} (Wildman and Crippen, 1999; Mannhold et al., 2009). It is indicative of hydrophobic VdW and π–π bond formation. The positive correlation to this descriptor suggests that adsorption is favorable for molecules with lower affinity for the aquatic medium and that hydrophobic interactions are important (coefficient = 1.25). For example, looking at the model descriptors for PAHs, this group of compounds has no hydrogen-bond acceptor groups, yet adsorption coefficients were high (between 2.4 and 4.5) and three orders of magnitude higher than log k_d found for PAHs adsorbed on commercial activated carbon (Kong et al., 2011). Similar to previous reports involving organo-clays, the order of adsorption affinities to the CPN was related to log k_{ow}. This means the

Table 2
Descriptors from initial QSAR model.

Compound	Experimental log k_d (0.001 activity)	MM (cg/mol)	Hydrogen bond acceptor groups	Alogp98 (calculated k_{ow})	Heat of formation
Acenaphthylene	4.191	1.54	0	3.34	1.59
Ametryn	3.941	2.27	4	2.74	0.09
Atrazine	4.523	2.16	4	2.54	0.29
Benzoic acid	2.14	1.22	2	1.46	0.4
Bromacil	2.217	2.61	3	1.94	0.99
Carbamazepine	1.89	2.48	1	2.68	1.24
Clofibric acid	2.803	2.15	4	2.76	−0.11
Diazinon	4.763	3.04	6	3.51	0.02
Diclofenac	5.884	2.96	4	4.37	1.92
Diuron	2.743	2.33	3	2.48	1.64
Eosin B	3.445	5.8	11	5.5	0.63
Gemfibrozil	4.144	2.5	3	4.17	−0.12
Ibuprofen	3.331	2.06	2	3.61	−0.09
Imazapyr	1.958	2.61	5	1.32	2.52
Imazaquin	3.296	3.11	5	2.52	2.83
Ketoprofen	3.199	2.54	3	3.36	−0.06
Metaldehyde	1.317	1.76	4	0.12	−0.18
Metazaclor	2.216	2.78	3	2.94	1.08
Metolachlor	2.584	2.84	3	3.73	0.66
Naphthalene	2.372	1.28	0	2.74	0.43
Paracetamol	1.955	1.51	2	0.71	−0.05
Phenanthrene	4.544	1.78	0	3.65	1.26
Picric acid	5.405	2.29	7	1.27	1.41
Prometryn	4.644	2.41	4	3.12	1.44
Pyrene	2.946	2.02	0	3.95	2.79
Pyrithiobac	4.186	3.27	8	2.56	1.03
Simazine	5.168	2.02	4	2.16	0.34
Sulfentrazone	2.525	3.87	8	2.26	1.57
Terbuthylazine	3.1	2.3	4	2.74	0.74
Toluene	3.055	0.92	0	2.32	0.56

hydrophobic nature of the compounds facilitates their transfer from the aquatic medium to the surface, enabling the formation of more specific π–π bonds (Kowalska et al., 1994; Goss and Schwarzenbach, 2000; Lemić et al., 2007; Qu et al., 2008).

The adsorption affinity of the pharmaceuticals with carboxylic groups was also found to be linked to their log k_{ow}; both log k_d and log k_{ow} are in the order of benzoic acid < clofibric acid < ketoprofen < ibuprofen < diclofenac. Benzoic acid is a small molecule with 2 H-acceptor groups, but its hydrophilicity (log k_{ow} 1.43) and solubility are so high it does not leave the aquatic medium, unlike Ibuprofen, which also has 2 H-acceptor groups and a larger MM , but is much more hydrophobic (log k_{ow} of 4.37). This indicates that the predominant trait governing adsorption affinity for this group of molecules is their propensity to leave the aquatic medium and create hydrophobic bonds with the surface.

In previous studies, the governing hypothesis regarding the adsorption mechanism of the pollutant to the composite was dependent mostly on polymer functionality; size of adsorption sites on the CPN and hydrogen bonding or hydrophobic interaction with the polymers functional groups (Zadaka et al., 2009; Ganigar et al., 2010; Radian and Mishael, 2012). The first three descriptors mentioned in the current study, MM, hydrogen bonds and log k_{ow}, coincide with these previously hypothesized mechanisms with the exception of electrostatic binding which is not represented in this model. Although many electrostatic descriptors were present in the original descriptor set (Table A3), they were not significant in the MLR result, which means that the presence or absence of electrostatic charge does not seem to affect adsorption. This, as previously mentioned, is due to the low pH in dispersion which renders the molecules non-ionic, minimizing the effect of electrostatic interactions.

The relation of the heat of formation of a compound to the adsorption affinity is less obvious, however, removing this descriptor decreases the quality of regression considerably (coefficient = 0.39, R^2 = 0.57). No other descriptor had such an impact on the model. The importance of this trait in predicting adsorption affinity may be linked to the influence of the different functional groups in the molecules: the heat of formation is a measure of the internal energy; it describes the ability to form stable intra-molecular bonds in respect to the functional groups available. Therefore, the heat of formation may give a partial solution for the diversity of the set by representing the change in internal energy in respect to the molecules structure.

Very similar descriptors were found when correlating the adsorption of non-ionic pesticides to soil carbon content (k_{oc}) (Gramatica et al., 2000). Strong dependence on the compound size was found; however, unlike the model presented in this study, an increase in MM leads to an increase in adsorption. Also in contrast to our findings, it was established that the number of H-acceptor groups hindered pesticide adsorption to the soil (Gramatica et al., 2000). These differences found are related to the differences in adsorption mechanisms between the two adsorbents; the adsorption to soil carbon is mostly governed by non-specific, hydrophobic interactions where the size of the molecule enables more VdW bonding and the H-acceptors enable formation of hydrogen-bonds between pesticides and soil solution. In contrast, the Mt–HPVPcoS nanocomposite is a strong H-donor as well as a π-electron donor enabling the formation of specific intermolecular chemical bonds. Furthermore, the configuration of the polymer on the surface plays a role, establishing preferential adsorption to small molecules. These differences in the models imply that this method can indeed be used to elucidate the difference in governing adsorption mechanisms between adsorbents or even within a given system under different conditions (pH, salinity, temperature etc.).

Within the applicability domain of the selected molecule an informed prediction of the binding affinity can be made. External validation was not possible due to the fact that this is a novel adsorbent that we developed and no literary data is available. However, a significant cross-validation for log k_d prediction was found and the standard deviation of the predicted log k_d was only 0.77 log units. This implies that this reliable model could be used for initial screening of pollutant adsorption to the CPN.

4 . Conclusions

The experimental adsorption isotherms revealed high adsorption affinity to the novel Mt–HPVPcoS nanocomposite for most of the pollutants (22 out of 30 pollutants had a log k_d > 2.5). QSAR calculations resulted in a significant regression describing the binding of the diverse set of molecules using four descriptors. The results emphasize the main interactions which govern the adsorption and a significant cross-validation for log k_d prediction was found. One of the main advantages of this model is the simplicity of its descriptors, which are universal and easily calculated, eliminating experimental errors of literary data. These descriptor calculations were conducted for the molecules under basal conditions (in vacuum) and were fitted to the specific experimental adsorption coefficients (under acidic conditions of this study). Therefore, the results and conclusions presented are based on these conditions only; under different conditions (surface modification, pH, salinity, temperature etc.) the adsorption mechanism may change which will result in a different correlation equation. Each equation, derived from different conditions, should express the relevant governing adsorption mechanisms of that system. We suggest that the above method for QSAR calculations can not only explain complex adsorption mechanisms of heterogeneous surfaces but also enable smart, inexpensive and quick decision-making and help the design of improved adsorbents.

Acknowledgments

This research was supported by The Bi-National (Israel–US) Foundation (application # 2010438). In addition, we would like to thank

Mr. Oren Ish-Am for his noteworthy contribution regarding the statistical analysis of the results.

Appendix A

This document contains three tables and two figures: X-ray diffraction patterns of the CPN are depicted in Figure A1. The pollutant isotherms are shown in Figure A2. Detection methods for the different pollutants are detailed in Table A1, the Abraham descriptors for 30 pollutants are shown in Table A2 and the QSAR descriptors calculated from Materials Studio and Discovery Studio (Accelrys) are listed in Table A3.

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.clay.2015.03.021.

References

Abraham, M.H., Acree Jr., W.E., 2010. Equations for the transfer of neutral molecules and ionic species from water to organic phases. J. Org. Chem. 75, 1006–1015.

Abraham, M.H., Andonian-Haftvan, J., Whiting, G.S., Leo, A., Taft, R.S., 1994. Hydrogen bonding. Part 34. The factors that influence the solubility of gases and vapours in water at 298 K, and a new method for its determination. J. Chem. Soc. Perkin Trans. 20, 1777–1791.

Andrew, L., 2012. Predicting Abraham Descriptors from SMILES. http://onschallenge.wikispaces.com/AbrahamDescriptorsModel003 (ONSchallenge).

Borisover, M., Graber, E.R., Bercovich, F., Gerstel, Z., 2001. Suitability of dye–clay complexes for removal of non-ionic organic compounds from aqueous solutions. Chemosphere 44, 1033–1040.

Breen, C., Watson, R., 1998. Polycation–exchanged clays as sorbents for organic pollutants: influence of layer charge on pollutant sorption capacity. J. Colloid Interface Sci. 208, 422–429.

Bronner, G., Goss, K.-U., 2010. Predicting sorption of pesticides and other multifunctional organic chemicals to soil organic carbon. Environ. Sci. Technol. 45, 1313–1319.

Celis, R., Cornejo, J., Hermosín, M.C., Koskinen, W.C., 1997. Sorption–desorption of atrazine and simazine by model soil colloidal components. Soil Sci. Soc. Am. J. 61, 436–443.

Cerejeira, M.J., Viana, P., Batista, S., Pereira, T., Silva, E., Valério, M.J., Silva, A., Ferreira, M., Silva-Fernandes, A.M., 2003. Pesticides in Portuguese surface and ground waters. Water Res. 37, 1055–1063.

Churchman, G.J., 2002. Formation of complexes between bentonite and different cationic polyelectrolytes and their use as sorbents for non-ionic and anionic pollutants. Appl. Clay Sci. 21, 177–189.

Clarke, E.D., 2009. Beyond physical properties—application of Abraham descriptors and LFER analysis in agrochemical research. Bioorg. Med. Chem. 17, 4153–4159.

Deblonde, T., Cossu-Leguille, C., Hartemann, P., 2011. Emerging pollutants in wastewater: a review of the literature. Int. J. Hyg. Environ. Health 214, 442–448.

Endo, S., Schmidt, T.C., 2006. Partitioning properties of linear and branched ethers: determination of linear solvation energy relationship (LSER) descriptors. Fluid Phase Equilib. 246, 143–152.

Falta, R.W., 2004. The potential for ground water contamination by the gasoline lead scavengers ethylene dibromide and 1,2-dichloroethane. Ground Water Monit. Rem. 24, 76–87.

Ganigar, R., Rytwo, G., Gonen, Y., Radian, A., Mishael, Y.G., 2010. Polymer–clay nanocomposites for the removal of trichlorophenol and trinitrophenol from water. Appl. Clay Sci. 49, 311–316.

Goss, K.-U., Schwarzenbach, R.P., 2000. Linear free energy relationships used to evaluate equilibrium partitioning of organic compounds. Environ. Sci. Technol. 35, 1–9.

Gramatica, P., Corradi, M., Consonni, V., 2000. Modeling and prediction of soil sorption coefficients of non-ionic organic pesticides by molecular descriptors. Chemosphere 41, 763–777.

Gramatica, P., Cassani, S., Roy, P.P., Kovarich, S., Yap, C.W., Papa, E., 2012. QSAR modeling is not "push a button and find a correlation": a case study of toxicity of (benzo-) triazoles on algae. Mol. Inf. 31, 817–835.

John, D., 2004. QSAR modeling of soil sorption. Predicting Chemical Toxicity and Fate. CRC Press.

Kong, H., He, J., Gao, Y., Han, J., Zhu, X., 2011. Removal of polycyclic aromatic hydrocarbons from aqueous solution on soybean stalk-based carbon. J. Environ. Qual. 40, 1737–1744.

Kowalska, M., Güler, H., Cocke, D.L., 1994. Interactions of clay minerals with organic pollutants. Sci. Total Environ. 141, 223–240.

Lemić, J., Tomašević-Čanović, M., Adamović, M., Kovačević, D., Milićević, S., 2007. Competitive adsorption of polycyclic aromatic hydrocarbons on organo-zeolites. Microporous Mesoporous Mater. 105, 317–323.

Mannhold, R., Poda, G.I., Ostermann, C., Tetko, I.V., 2009. Calculation of molecular lipophilicity: state-of-the-art and comparison of log P methods on more than 96,000 compounds. J. Pharm. Sci. 98, 861–893.

McGuire, M.J., Suffet, I.H., 1978. Adsorption of organics from domestic water supplies. Am. Water Works Assoc. J. 70, 621–636.

Platts, J.A., Abraham, M.H., Butina, D., Hersey, A., 1999. Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients. J. Chem. Inf. Comput. Sci. 40, 71–80.

Qu, X.L., Liu, P., Zhu, D.Q., 2008. Enhanced sorption of polycyclic aromatic hydrocarbons to tetra-alkyl ammonium modified smectites via cation–pi interactions. Environ. Sci. Technol. 42, 1109–1116.

Radian, A., Mishael, Y., 2012. Effect of humic acid on pyrene removal from water by polycation–clay mineral composites and activated carbon. Environ. Sci. Technol. 46, 6228–6235.

Radian, A., Michaeli, D., Serban, C., Nechushtai, R., Mishael, Y.G., 2010. Bioactive apoferredoxin–polycation–clay composites for iron binding. J. Mater. Chem. 20, 4361–4365.

Satoshi, E., Peter, G., Stefan, H., Torsten, S., 2009. LFERs for soil organic carbon–water distribution coefficients (KOC) at environmentally relevant sorbate concentrations. Environ. Sci. Technol. 43, 3094–3100.

Shih, Y.-h., Gschwend, P.M., 2009. Evaluating activated carbon–water sorption coefficients of organic compounds using a linear solvation energy relationship approach and sorbate chemical activities. Environ. Sci. Technol. 43, 851–857.

Stephens, T.W., Rosa, N.E.D.L., Saifullah, M., Ye, S., Chou, V., Quay, A.N., Acree Jr., W.E., Abraham, M.H., 2011. Abraham model correlations for transfer of neutral molecules and ions to sulfolane. Fluid Phase Equilib. 309, 30–35.

Suffet, I.H., 1980. An evaluation of activated carbon for drinking water treatment: a National Academy of Science report. J. Am. Water Works Assoc. 41–50.

Tülp, H.C., Goss, K.-U., Schwarzenbach, R.P., Fenner, K., 2008. Experimental determination of LSER parameters for a set of 76 diverse pesticides and pharmaceuticals. Environ. Sci. Technol. 42, 2034–2040.

Wildman, S.A., Crippen, G.M., 1999. Prediction of physicochemical parameters by atomic contributions. J. Chem. Inf. Comput. Sci. 39, 868–873.

Zadaka, D., Nir, S., Radian, A., Mishael, Y.G., 2009. Atrazine removal from water by polycation–clay composites: effect of dissolved organic matter and comparison to activated carbon. Water Res. 43, 677–683.

Please cite this article as: Radian, A., et al., Modeling binding of organic pollutants to a clay–polycation adsorbent using quantitative structural–activity relationships (QSARs), Appl. Clay Sci. (2015), http://dx.doi.org/10.1016/j.clay.2015.03.021