

**SORPTION AND PARTITIONING OF
NEUTRAL AND CHARGED ORGANIC SPECIES:
A LOG-NORMAL LANGMUIR ISOTHERM MODEL & APPLICATION OF
QUANTUM-CHEMICALLY ESTIMATED ABRAHAM SOLUTE
PARAMETERS**

by

Craig Warren Davis

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering

Spring 2016

© 2016 Davis
All Rights Reserved

ProQuest Number: 10156570

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10156570

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

**SORPTION AND PARTITIONING OF
NEUTRAL AND CHARGED ORGANIC SPECIES:
A LOG-NORMAL LANGMUIR ISOTHERM MODEL & APPLICATION OF
QUANTUM-CHEMICALLY ESTIMATED ABRAHAM SOLUTE
PARAMETERS**

by

Craig Warren Davis

Approved: _____
Harry W. Shenton III, Ph.D.
Chair of the Department of Civil & Environmental Engineering

Approved: _____
Babatunde Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Dominic M. Di Toro, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Herbert E. Allen, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Stanley I. Sandler, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Pei C. Chiu, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I'd like to first and foremost thank my advisor and mentor Dr. Dominic Di Toro, without whom none of this would be possible. I would also like to thank my committee members, Dr. Herbert Allen, Dr. Stanley Sandler, and Dr. Pei Chiu for their valuable input, helpful discussion, and their guidance over the course of my graduate work. I'd like to thank my coworkers, both past and present, for their advice and support, particularly Dr. Andrew Miglino and Cory Daviau. Finally, I'd like to thank my family and friends for their support and love throughout my journey: My parents, Frances and George, my sister, Alexis, and my girlfriend, Megan.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	xv
ABSTRACT	xxi

Chapter

1	INTRODUCTION	1
1.1	Motivation	1
1.2	Research Goals	2
2	MODELING NON-LINEAR ADSORPTION TO CARBON WITH A SINGLE CHEMICAL PARAMETER: A LOG-NORMAL LANGMUIR ISOTHERM.....	3
2.1	Introduction	4
2.2	Modeling and Experimental Data.....	5
2.2.1	Log-normal Langmuir (LNL) Isotherm.....	5
2.2.2	Experimental Data	8
2.3	Results	10
2.4	Discussion.....	17
2.4.1	Normalized LNL Isotherm	17
2.4.2	Residual Analysis	20
2.5	Implications	25
3	MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: PREDICTING CHEMICAL MEDIAN LANGMUIR BINDING CONSTANTS	26
3.1	Introduction	27
3.2	Materials and Data.....	28
3.3	Results & Discussion.....	29

3.3.1	Abraham Poly-parameter Linear Free Energy Relationship (pp-LFER)	29
3.3.2	Graphite	32
3.3.3	Charcoal.....	32
3.3.4	Darco GAC.....	33
3.3.5	F400 GAC	34
3.3.6	Separation of Chemicals.....	35
3.3.7	Separation of Large Molecular Volume Sorbates	36
3.3.8	Residual Analysis	40
3.4	Implications	42
4	PREDICTING SOLVENT-WATER PARTITIONING OF CHARGED ORGANIC SPECIES USING QUANTUM-CHEMICALLY ESTIMATED ABRAHAM PP-LFER SOLUTE PARAMETERS	44
4.1	Introduction	45
4.2	Modeling & Experimental Data	47
4.2.1	Abraham Poly-parameter Linear Free Energy Relationship (pp-LFER)	47
4.2.2	Quantum-Chemically Estimated Abraham Solute Parameters (QCAP).....	50
4.2.3	Quantum Chemical Computations	52
4.2.4	Experimental Data	54
4.3	Results & Discussion.....	55
4.3.1	Carboxylate Anions	58
4.3.2	Cations	61
4.3.3	Fitted Abraham Solvent – Water Parameters	64
4.3.4	Residual Analysis	66
4.4	Implications	68
5	PREDICTING NON-LINEAR PARTITIONING OF CHARGED ORGANIC SPECIES TO NATURAL ORGANIC MATTER USING QUANTUM-CHEMICALLY ESTIMATED ABRAHAM PP-LFER SOLUTE PARAMETERS	70
5.1	Introduction	71
5.2	Modeling & Experimental Data	73
5.2.1	Abraham Poly-parameter Linear Free Energy Relationship	73

5.2.2	Quantum-chemically Estimated Abraham Parameters (QCAP) .	74
5.3	Results & Discussion.....	78
5.3.1	Isotherm Model	78
5.3.2	Predicting Median Langmuir Binding Constants for Ionic Species	83
5.3.3	Residual Analysis	86
5.4	Implications	88
6	CONCLUSIONS & RECOMMENDATIONS	89
	REFERENCES	93
Appendix		
A	MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: A LOG-NORMAL LANGMUIR ISOTHERM.....	105
A.1	Log-normal Langmuir Visual Basic Code	106
A.2	Log-normal Langmuir Fitted Isotherm Parameters	108
A.3	Log-normal Langmuir Estimated Isotherm Parameters – F400 GAC... ..	110
A.4	Discussion of Standard Errors of Estimated Parameters – Model (4)... ..	118
A.5	Estimated ss-LNL Isotherms for F400 GAC.....	120
A.6	Comparison of Freundlich and LNL Isotherm Residuals.....	121
A.7	Tables of Estimated Freundlich Isotherm Parameters and Associated Standard Errors.....	126
B	MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: PREDICTING CHEMICAL MEDIAN LANGMUIR BINDING CONSTANTS	131
B.1	Abraham Solute Parameters, Aqueous Solubility, and Predicted vs. LNL-Estimated $\log(\tilde{K}_L)$	132
B.2	Estimated Abraham Sorbent-water Parameters (with Polarizability) for Graphite, Charcoal, Darco GAC, and F400 GAC	136
B.3	Comparison of Abraham Sorbent-water Parameters (with and without Polarizability, sS, Term) for F400 GAC	138
B.4	Calculation of Cook's Distance for Oxamyl in Solubility vs. $\log(\tilde{K}_L)$ Regression.	139

B.5	Comparison of Abraham Sorbent-water Parameters (with and without Hydrogen-bond Basicity, bB, Term) for F400 GAC sorbates with V < 110.	140
B.6	Comparison of LNL pp-LFER vs. Shih & Gschwend pp-LFER Predicted Isotherm Residuals.	141
C	PREDICTING SOLVENT-WATER PARTITIONING OF CHARGED ORGANIC SPECIES USING QUANTUM CHEMICALLY-DERIVED ABRAHAM PP-LFER SOLUTE PARAMETERS	143
C.1	Direct Quantum Chemically Computed Solvent-water Partition Coefficients of Ions in Acetonitrile-, DMSO-, and Methanol-water Systems.....	144
C.2	Calculation of Excess Molar Refraction from McGowan Volume and QC-Computed Polarizability	148
C.2.1	R Code – QCAP Solute / Solvent Parameter Optimization	149
C.3	Experimental Neutral Solvent-water System Parameters.....	151
C.4	Experimental Solvent-water Partition Coefficients	158
C.5	Absolv-estimated (AAP) and Quantum-chemically Derived (QCAP) Abraham Solute Parameters	172
C.6	Quantum-chemically Estimated Solvent Parameters	184
C.7	Complete AAP & QCAP Abraham pp-LFERs	188
C.8	Residual Plots of A Priori QC Computed and AAP –predicted Partition Coefficients.....	191
D	PREDICTING PARTITIONING OF CHARGED ORGANIC SPECIES TO NATURAL ORGANIC MATTER USING QUANTUM CHEMICALLY-DERIVED ABRAHAM PP-LFER SOLUTE PARAMETERS.....	193
D.1	Complete List of Solute Names and ID Abbreviations	194
D.2	LNL Isotherm Visual Basic Code & Abraham Solute / Solvent Parameter Optimization (QCAP) R Code	197
D.3	Estimated Log-normal Langmuir (LNL) Isotherm Parameters for Primary through Quaternary Amine Cations.....	200
D.4	Legend for Estimated Isotherm (Eq. (5-15) for Figure 5-1	203
D.5	Absolv (AAP) and Quantum-chemically Estimated (QCAP) Solute Parameters for Primary through Quaternary Amine Cations	204
D.6	Quantum-chemically Estimated (QCAP) Solvent Parameters for Primary through Quaternary Amine Cations.....	206
E	JOURNAL PERMISSIONS	208

LIST OF TABLES

Table 2-1.	Summary of experimental data used in construction of the log-normal Langmuir isotherm model (graphite ³⁶ , charcoal ³⁶ , Darco granular activated carbon (GAC) ³⁴ , and F400 GAC ³⁵).	8
Table 2-2.	Summary of LNL isotherm models, sorbent and sorbate-specific parameters, RMS errors, and median standard errors (SE) of the estimated parameters.	13
Table 2-3.	Summary of $\log(\tilde{K}_L)$ standard errors, individual, and combined RMS errors (Eq. (2-15)) for the ss-LNL model fits for graphite, charcoal, Darco GAC, and F400 GAC. Units of $\tilde{K}_L$ are $(\text{mM})^{-1}$	16
Table 2-4.	Comparison of Freundlich and LNL isotherm RMS errors (Eq. (2-15)) for graphite, charcoal, Darco GAC, and F400 GAC.	24
Table 3-1.	Summary of the single sorbate-specific log-normal Langmuir (ss-LNL) estimated isotherm parameters and RMS errors for graphite ³⁶ , charcoal ³⁶ , Darco GAC ³⁴ , and F400 GAC ³⁵	28
Table 3-2.	Summary of estimated sorbent-water system Abraham parameters for graphite, charcoal, Darco GAC, and F400 GAC. Separate RMS errors are indicated for the separate equations in the revised F400 model. The notation “All”, “V < 110” and “V > 110” refers to the data used for the estimates. The standard errors (SE) and coefficient of variation (CV) of the parameters for each model are reported. Units of $\tilde{K}_L$ and solubility (S) are $(\text{mM})^{-1}$	31
Table 4-1.	Summary of anion and cation solvent-water partitioning data.....	55
Table 4-2.	Summary of RMS errors for direct QC computed, Absolv Abraham solute parameter (AAP), and quantum-chemically estimated Abraham solute parameter (QCAP) methods.....	61

Table 4-3.	Summary of corrected Abraham solvent-water system parameters (using QCAP solute descriptors), their respective coefficients of variation (%), number of experimental data points (N), and root mean square (RMS) errors for the carboxylic acid anions (N = 180) and quaternary amine cations (N = 217) experimental partitioning data.	65
Table 5-1.	Summary of experimental sorption data for a set of primary through quaternary amines onto standard Pahokee Peat soil ¹³	73
Table 5-2.	Summary of RMS errors and Tukey R ² correlation coefficients for AAP- and QCAP-predicted median Langmuir binding constants.	85
Table A-1a.	Estimated LNL isotherm parameters and associated standard errors for graphite and charcoal. Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless. Coefficients of variation (CV) are expressed as a % of the estimated parameter.	108
Table A-1b.	Estimated LNL isotherm parameters and associated standard errors for Darco GAC. Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless. Coefficients of variation (CV) are expressed as a % of the estimated parameter.	109
Table A-2a.	Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (1). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless.	110
Table A-2b.	Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (2). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless.	112
Table A-2c.	Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (3). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless.	114
Table A-2d.	Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (4). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM) ⁻¹ , respectively, and σ_κ is dimensionless.	116
Table A-3a.	Estimated Freundlich parameters and standard errors for graphite. Units of Freundlich K_f are (mmol/kg)(mmol/L) ^{1/v} . v is dimensionless.	126

Table A-3b. Estimated Freundlich parameters and standard errors for charcoal. Units of Freundlich K_f are $(mmol/kg)(mmol/L)^{1/v}$. v is dimensionless.	127
Table A-3c. Estimated Freundlich parameters and standard errors for Darco GAC. Units of Freundlich K_f are $(mmol/kg)(mmol/L)^{1/v}$. v is dimensionless.	128
Table A-3d. Estimated Freundlich parameters and standard errors for F400 GAC. Units of Freundlich K_f are $(mmol/kg)(mmol/L)^{1/v}$. v is dimensionless.	129
Table B-1a. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL-estimated $\log(\tilde{K}_L)$ for F400 GAC. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$	132
Table B-1b. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL-estimated $\log(\tilde{K}_L)$ for graphite and charcoal. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$	134
Table B-1c. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL-estimated $\log(\tilde{K}_L)$ for Darco GAC. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$	135
Table B-2. Summary of estimated Abraham sorbent-water pp-LFER coefficients for graphite, charcoal, Darco GAC, and F400 GAC including solvent hydrogen bond acidity (a) and polarizability (s) terms. Separate RMS errors are indicated for the separate equations in the revised F400 model. The notation “All”, “ $V > 110$ ”, and “ $V < 110$ ” refer to the data used to make the estimates of the parameters. The standard errors of the parameters for each model are given in parentheses below their respective parameters. Units correspond to $\tilde{K}_L$ and S with units $(mM)^{-1}$	136
Table B-3. Comparison of reduced Abraham pp-LFER model for F400 GAC, with (top) and without (bottom) the polarizability (sS) interaction term. Units correspond to $\tilde{K}_L$ with units $(mM)^{-1}$	138
Table B-4. Comparison of reduced Abraham pp-LFER model for F400 GAC ($V < 110$), with (top) and without (bottom) the hydrogen bond basicity (bB) interaction term. Units correspond to $\tilde{K}_L$ with units $(mM)^{-1}$	140

Table C-1a. Summary of RMS errors for solvent-water partitioning of anion, cation, and combined ionic data sets presented by Marenich et al ¹⁵ . QC computations were all performed at the M062X/MG3S level of theory.	144
Table C-1b. Experimental vs. QC computed solvent-water partition coefficients for anions and cations obtained from Marenich et al ¹⁵ . QC computations were all performed at the M062X/MG3S level of theory.	145
Table C-2. Literature values for solvent-water Abraham pp-LFER system parameters. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). All system parameters come from the UFZ LSER database ⁷¹ , the REF column indicates the original experimental source of the data.	151
Table C-3a. Table of experimental acetone-water partition coefficients ($\log(K_{\text{acetone-water}})$) for carboxylate anions in acetone. $K_{\text{acetone-water}}$ has units (L water / kg acetone).	158
Table C-3b. Table of experimental acetonitrile-water partition coefficients ($\log(K_{\text{acetonitrile-water}})$) for carboxylate anions in acetonitrile. $K_{\text{acetonitrile-water}}$ has units (L water / kg acetonitrile).	160
Table C-3c. Table of experimental DMSO-water partition coefficients ($\log(K_{\text{DMSO-water}})$) for carboxylate anions in dimethylsulfoxide (DMSO). $K_{\text{DMSO-water}}$ has units (L water / kg DMSO).	162
Table C-3d. Table of experimental methanol-water partition coefficients ($\log(K_{\text{methanol-water}})$) for carboxylate anions in methanol. $K_{\text{methanol-water}}$ has units (L water / kg methanol).	164
Table C-3-e. Table of experimental octanol-water partition coefficients ($\log(K_{\text{octanol-water}})$) for quaternary amine cations in octanol. $K_{\text{octanol-water}}$ has units (L water / kg octanol). Solute and Group Identification Numbers refer to the original classifications in the primary reference (Abraham & Zhao, 2008) ¹⁴ . SMILES strings are for the charged (cationic) species.	166

Table C-4a. Absolv-estimated and quantum-chemically estimated solute descriptors for carboxylate anions. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	172
Table C-4b. Absolv-estimated and quantum-chemically estimated solute descriptors for quaternary amine cations. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	175
Table C-5a. Quantum-chemically estimated Abraham pp-LFER solvent parameters for carboxylate anions. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	184
Table C-5b. Quantum-chemically estimated Abraham pp-LFER solvent parameters for quaternary amine cations. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	186
Table C-6a. Full AAP Abraham pp-LFER including all solvent parameters for carboxylate anions estimated with experimental solvent-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	188
Table C-6b. Full QCAP Abraham pp-LFER including all solvent parameters for carboxylate anions estimated with experimental solvent-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).	189
Table C-6c. Full AAP Abraham pp-LFER including all solvent parameters for quaternary amine cations estimated with experimental octanol-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{octanol-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{octanol-water}}$ has units of (L water / kg octanol).	190

Table D-1.	Complete list of solute names and ID abbreviations.	194
Table D-2.	Estimated log-normal Langmuir (LNL) isotherm parameters and coefficients of variation (CV %) and linear partition coefficients ($\log(K_p)$) and associated standard errors for primary through quaternary amine cations. Units of $\log(\tilde{K}_L)$ are (L/mmol). For P15 and S14, coefficients of variation were not calculated as the standard errors are orders of magnitude larger than the estimated values. $\log(K_p)$ values are obtained via the pp-LFER equation for particulate organic carbon ⁷ (Eq. (5-13))......	200
Table D-3.	Absolv-estimated (AAP) and quantum-chemically estimated (QCAP) solute descriptors for quaternary amine cations. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eEi + sSi + aAi + bBi + vVi + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eEi + sSi + aAi + bBi + vVi + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent)......	204
Table D-4.	Quantum-chemically estimated Abraham p-LFER solvent parameters for quaternary amine cations. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eEi + sSi + aAi + bBi + vVi + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent)......	206

LIST OF FIGURES

- Figure 2-1.** Equilibrium adsorption isotherm data for graphite (A), charcoal (B), and Darco GAC (C). The legend on the top right corresponds to panels (A) and (B). The legend on the bottom right corresponds to panel (C). 9
- Figure 2-2.** (A) Residual box plots of $\log(q_{pred}) - \log(q_{obs})$ for the LNL isotherm models, 1 through 4, on F400 GAC, where (1) $\tilde{K}_L$, σ_κ , and q_{max} , (2) $\tilde{K}_L$ and q_{max} , (3) $\tilde{K}_L$ and σ_κ , or (4) $\tilde{K}_L$ are the sorbate-specific parameters within the models. Semi-dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations, dotted lines represent ± 1.0 log-units difference between predicted and observed adsorbed concentrations. RMS errors for models 1 through 4 are 0.0659, 0.0680, 0.0690, and 0.113, respectively. (B) Box plots of the logarithmic standard errors of the estimated parameters for models 1 through 4. Dotted line represents 1 order of magnitude error in the estimated parameters, semi-dashed line represents 2 order of magnitude error in the estimated parameters, solid blue line represents 0.4 log-units error in the estimated parameters (a factor of 3). (C) Bar plot of the total number of estimated parameters for models 1 through 4. For the box plots, the interquartile range (IQR) contains 50% of the data, whiskers represent $\pm 1.5\text{IQR}$, and points represent outliers ($> 1.5\text{IQR}$). 12
- Figure 2-3.** Adsorption isotherms for the single sorbate-specific parameter log-normal Langmuir (ss-LNL) model for graphite, charcoal, and Darco GAC. The sorbate symbols are the same for Figures 1 and 3. Solid lines represent the fit to the LNL isotherm model. Estimated LNL parameters are presented in Appendix A (Tables A-1a,b). 15
- Figure 2-4.** (A) Effect of systematically varying the standard deviation of the median binding constant, σ_κ , on the shape of the normalized ss-LNL isotherm. (B),(C) Normalized isotherms for graphite, charcoal, and Darco GAC (B) and F400 GAC (C). Estimated LNL isotherm parameters and associated standard errors can be found in Tables A-1a,b and A-2a-d in Appendix A. 18

- Figure 2-5.** (A), (B), (C) Plots of LNL median binding constants, sorted from smallest to largest for (A) graphite, (B) charcoal, and (C) Darco GAC. (D), (E), (F) Box plots of the LNL isotherm model residuals $\log(q_{pred}) - \log(q_{obs})$ for (D) graphite, (E) charcoal, and (F) Darco GAC. The solid lines represents 1:1 agreement between q_{pred} and q_{obs} adsorbed concentrations, dashed lines represent ± 0.3 log-units difference between observed and modeled adsorbed. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and points represent outliers (> 1.5 IQR). 22
- Figure 2-6.** (A) Plot of LNL median binding constants, sorted from smallest to largest for F400 GAC. (B) Box plot of the LNL isotherm model residuals $\log(q_{pred}) - \log(q_{obs})$ for F400 GAC. The solid line represents 1:1 agreement between q_{pred} and q_{obs} adsorbed concentrations, dashed lines represent ± 0.3 log-units difference between observed and modeled adsorbed. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and points represent outliers (> 1.5 IQR). 23
- Figure 3-1.** Plot of Abraham-predicted vs. LNL-estimated median binding constants ($\tilde{K}_L$) for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC. Solid lines represent 1:1 agreement; dashed lines represent ± 1 order of magnitude. RMSE for the four sorbent data sets are 0.129, 0.307, 0.407, and 0.709, respectively. 30
- Figure 3-2.** Plot of LNL-estimated median binding constants ($\tilde{K}_L$) vs. the Absolv-estimated solute McGowan volumes for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC. Dashed lines represent the observed volumetric cutoff of 110 Å³/molecule. 36
- Figure 3-3.** Plot of LNL-estimated median binding constants ($\tilde{K}_L$) vs. experimental aqueous solubility (S) for solutes above McGowan cutoff in F400 GAC. The dashed line represents the linear regression $\log \tilde{K}_L = -0.508 \log S - 0.202$. Oxamyl was not included in the regression due to its large Cook distance (see discussion in Appendix B.4). 38

- Figure 3-4.** Plot of Abraham-predicted vs. LNL-estimated median binding constants ($\tilde{K}_L$) for F400 GAC. Solid lines represent 1:1 agreement; dashed lines represent ± 1 order of magnitude. Open circles represent sorbates below the volumetric cutoff (Eq. (3-2)), filled circles represent sorbates above the volumetric cutoff (Eq. (3-4)). Combined RMSE for all sorbates is 0.424. The RMS error excluding oxamyl is 0.345. 39
- Figure 3-5.** Box plots of the Abraham-predicted isotherm residuals $\log(q_{pred}) - \log(q_{obs})$ for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC, ordered by increasing median Langmuir binding constant. The solid line represents 1:1 agreement between q_{pred} and q_{obs} sorbed concentrations, dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and filled points represent outliers (> 1.5 IQR). 41
- Figure 4-1.** Box plot of experimental solvent-water partition coefficients for cationic (octanol) and anionic (acetone, acetonitrile, dimethylsulfoxide (DMSO), and methanol) species by solvent. The interquartile range (IQR) represents 50% of the data, whiskers represent 1.5IQR, and filled points represent outliers (> 1.5 IQR). Box widths are proportional to the square root of the number of data points. The number of solutes within each solvent data set is shown below the solvent names. 56
- Figure 4-2.** (A) Direct QC computed (SMD, M062X/6-31++G**), (B) Absolv-predicted, (C) QCAP-predicted vs. experimental solvent-water partition coefficients for carboxylic acid anions for acetone (ACET), acetonitrile (ACN), DMSO, and methanol (MEOH). Solid lines represent 1:1 agreement, dashed lines represent $\pm$ an order of magnitude errors. RMS errors are 2.48, 0.740, and 0.462, respectively (N = 180). 60
- Figure 4-3.** (A) Direct QC-computed (SMD, M062X/6-31++G**), (B) Absolv-predicted, (C) QCAP-predicted vs. experimental octanol-water partition coefficients for quaternary amine cations. (Solid lines represent 1:1 agreement, dashed lines represent $\pm$ an order of magnitude errors. RMS errors are 2.82, 0.997, and 1.16, respectively (N = 217). 63

- Figure 4-4.** Logarithmic residual box plots ordered by increasing solvent-water median residuals. Boxes correspond to the interquartile range (IQR), where $25\% < \text{IQR} < 75\%$. Whiskers correspond to $\pm 1.5\text{IQR}$. Points correspond to outliers ($> 1.5\text{IQR}$). Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10). Box width is proportional to the square root of the number of data points. 67
- Figure 5-1.** Adsorption isotherms for a selection of 34 primary through quaternary amine cations. Points represent the experimental adsorption data (see Figure D-1 in the Appendix for a legend of solutes). Solid lines represent the isotherm model (Eq. (5-15)). Model parameters are presented in Table D-2 in the Appendix. 80
- Figure 5-2.** Normalized isotherms for a selection of 63 primary through quaternary amine cations (where $f_{q+} \geq 99.9\%$). Points represent the experimental adsorption data. Solid lines represent the fit to the LNL isotherm model (Eq. (5-15)). Dashed lines represent ± 0.3 log-units (factor of 2) error in the predicted isotherms. The RMS error for the $f_{q+} \geq 99.9\%$ subset of chemicals is 0.139 (N = 69, with 567 data points). Estimated isotherm parameters and associated standard errors can be found in Table D-2 in the Appendix. 82
- Figure 5-3.** (A) AAP-predicted vs. LNL-estimated median binding constant, $(\log(K_{L,+}))$. (B) Logarithmic residuals $\log(K_{L,+ \text{ AAP}}) - \log(K_{L,+ \text{ LNL}})$ vs. Tukey mean median binding constant $(\log(K_{L,+ \text{ AAP}}) + \log(K_{L,+ \text{ LNL}}))/2$. (C) QCAP-predicted vs. LNL-estimated median binding constant, $(\log(K_{L,+}))$. (D) Logarithmic residuals $\log(K_{L,+ \text{ QCAP}}) - \log(K_{L,+ \text{ LNL}})$ vs. Tukey mean median binding constant $(\log(K_{L,+ \text{ QCAP}}) + \log(K_{L,+ \text{ LNL}}))/2$ 84

Figure 5-4. Box plots of the (A) AAP-predicted LNL isotherm residuals $\log(q_{AAP}) - \log(q_{expt})$ and (B) QCAP-predicted LNL isotherm residuals $\log(q_{QCAP}) - \log(q_{expt})$, ordered by increasing predicted LNL median binding constant within each type of amine. Vertical lines separate the types of amines. Horizontal black solid lines represents 1:1 agreement between predicted and experimental sorbed concentrations, dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations and semi-dashed lines represent ± 1.0 log-units difference between predicted and observed adsorbed concentrations. IQR contains 50% of the data, whiskers represent ± 1.5 IQR and filled points represent outliers > 1.5 IQR).....	87
Figure A-1. Adsorption isotherms for the single sorbate-specific parameter LNL (ss-LNL) isotherm model. Points represent experimental adsorption data, solid lines represent the ss-LNL model. Estimated LNL parameters are presented in Table A-3d.	120
Figure A-2a. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Graphite. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.....	122
Figure A-2b. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Charcoal. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.....	123
Figure A-2c. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Darco GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.	124
Figure A-2d. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on F400 GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.....	125

Figure B-1. Box plots comparing the residuals ($\log(q_{pred}) - \log(q_{obs})$) for the Shih & Gschwend Freundlich pp-LFER (left) and LNL pp-LFER (Eq. (3-2)) (right) models on Darco GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences between predicted and observed sorbed concentrations. Solid line represents 1:1 agreement between the predicted and observed sorbed concentrations. 142

Figure C-1a. Logarithmic residual box plots for a priori QC computed solvent-water partition coefficients. Boxes correspond to the interquartile range (IQR), where $25\% < IQR < 75\%$. Whiskers correspond to $\pm 1.5IQR$. Points correspond to values $> 1.5IQR$. Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10). 191

Figure C-1b. Logarithmic residual box plots for AAP-predicted solvent-water partition coefficients. Boxes correspond to the interquartile range (IQR), where $25\% < IQR < 75\%$. Whiskers correspond to $\pm 1.5IQR$. Points correspond to values $> 1.5IQR$. Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10). 192

Figure D-1. Legend of sorbates for adsorption data in Figure 5-1. Point type and color correspond to the matching isotherms in Figure 5-1. 203

ABSTRACT

The overall goal of this doctoral dissertation is to develop methods to accurately predict the partitioning and sorption behavior of neutral and ionic species from water into liquid organic phases as well as onto various forms of black and organic carbon. For neutral solutes, this includes the development of a novel sorption isotherm – the log-normal Langmuir model – that describes nonlinear partitioning onto carbonaceous sorbents with one sorbate-specific parameter. For charged species, a model has been developed that can generate the Abraham parameters that are the physical-chemical descriptors for charged species using quantum chemical computations. These solute descriptors are then used to predict both the solvent-water partitioning of the ionic species, as well as the sorption of ionic species onto soil organic carbon, using the log-normal Langmuir isotherm model developed for neutral species.

To better understand the nature of non-linear adsorption of organic solutes onto black carbon, a non-linear sorption model – the log-normal Langmuir model – was developed in Chapter 2 that utilizes Langmuir isotherms with a log-normal distribution of binding constants and a single maximum sorption capacity. The model has two sorbent-specific parameters: the maximum sorption capacity, q_{max} , and the standard deviation, σ_K , of the log of the Langmuir binding constants; and one sorbate-specific parameter, the median Langmuir binding constant, $\tilde{K}_L$. This is an important advance over previously available sorption isotherm models, for example the Freundlich isotherm, which has two sorbate-specific parameters. The reduction to a single

sorbate-specific parameter which has chemical meaning – the median Langmuir binding constant, $\tilde{K}_L$ – is an important advance. In particular it allows quantitative prediction of the isotherm for a new chemical if the single sorbate-specific parameter can be predicted.

In Chapter 3, the median Langmuir binding constants are predicted using an Abraham poly-parameter linear free energy relationship (pp-LFERs). For sorption of neutral organic solutes onto graphite, charcoal, Darco GAC, and F400 GAC ($n = 13, 11, 14, 44$ sorbates, respectively), RMS errors of predicted median binding constants, $\log(\tilde{K}_L)$, of 0.129, 0.307, 0.407, and 0.424 were obtained. Predicted isotherms were constructed with RMS errors of the predicted sorbed concentrations, $\log(q(c))$, of 0.0820, 0.1809, 0.183, and 0.220, respectively. This demonstrates that using the LNL isotherm and Abraham pp-LFER models, it is possible to predict the sorption isotherm of a new sorbate from only its molecular structure.

In Chapter 4, Abraham parameters for ionic species are estimated directly from quantum chemical (QC) computations of solvent-water partition coefficients and molecular polarizability by extending a method developed by Liang & Di Toro for neutral species. Quantum-chemically estimated Abraham solute (QCAP) parameters are determined for the solvent-water partitioning of a suite of carboxylic acid anions ($n = 60$) in acetone-, acetonitrile-, dimethylsulfoxide-, and methanol-water systems, as well as a suite of quaternary amine cations ($n = 217$) in an octanol-water system. Using these QCAP solute parameters, predictions of experimental solvent-water partition coefficients are made for the carboxylate anions with RMS errors of 0.475, 0.512, 0.460, and 0.393 for the four solvent-water systems, respectively. This is an improvement over both direct a priori QC calculations (RMSE = 3.43, 3.71, 0.698,

and 2.14, respectively) and predictions made using Absolv-estimated Abraham solute descriptors (AAP) (RMSE = 0.636, 0.59, 1.11, and 0.389, respectively) for the four solvent-water systems.

For the quaternary amine cations, the QCAP and AAP methods showed comparable improvements over direct QC computations of the octanol-water partition coefficients (RMSE = 1.16, 2.82, and 0.997, for the QCAP, direct QC, and AAP methods, respectively).

In Chapter 5, the log-normal Langmuir (LNL) isotherm is used to model non-linear sorption of a suite of primary through quaternary amines ($n = 80$) onto natural organic carbon (Pahokee peat). The LNL model can reproduce the sorption data (RMSE = 0.272, $N = 80$) for both fully ionized, as well as partially ionized, species. The latter are modeled as the sum of a linear isotherm for the neutral species, the usual model, and a LNL model for the charged species, weighted by the fractions neutral and ionized species present at the experimental pH = 4.5 and 6.8.

The median Langmuir binding constants for the ionic species are predicted using QCAP solute descriptors (RMSE = 0.526, $N = 60$) with accuracies comparable to those of linear partition coefficients for neutral species. The parameter for the neutral species linear isotherm model, K_{OC} , is predicted using a previously developed Abraham model. The predicted isotherms, constructed using the QCAP-predicted median binding constants and Abraham-predicted K_{OC} , demonstrate good agreement with the experimental data (RMSE = 0.457, $N = 60$).

The results of this dissertation research are models that can reproduce non-linear sorption isotherm data of neutral and ionic species with only one sorbate-specific parameter. For the cases considered: samples of graphite, powered charcoal,

and activated carbon, and for a natural organic carbon (Pahokee peat), the sorbate specific parameter can be predicted from its molecular structure only. The QCAP parameters can also be used to predict solvent-water partitioning of ionic compounds. This greatly expands the range of compounds that can be analyzed for physical/chemical properties that are used to evaluate the environmental risk posed by new compounds for which only the molecular structure is known.

Chapter 1

INTRODUCTION

1.1 Motivation

Modeling the partitioning of organic pollutants in the environment has been an area of concern for many years. Early work has focused primarily on the partitioning of neutral organic species between water and organic phases (e.g., octanol, natural organic matter (NOM) in soils, lipid in fish and other organisms)¹⁻⁷. This work has focused primarily on the linear partitioning of these species between environmental phases, largely ignoring non-linear partitioning behavior.

Models exist for predicting the linear partitioning of neutral organic species to soil organic carbon from the chemical structure or the physical chemical properties of the sorbates^{4,8}. However, similar models do not exist for ionizable organic species. Recent work by Franco et al.⁹ has demonstrated that for a subset of the 117,000 organic chemicals registered in the European REACH database, approximately 33% have been shown to be “significantly ionized” at environmentally relevant pH values (pH ~ 7.0). Consequently, 1/3 of the chemical database lacks a predictive model for accurately determining the partitioning, and ultimately the fate and transport, of these chemicals in the environment.

Previous work has attempted to extend existing predictive models for neutral species to predict the partitioning and sorption of ionizable species¹⁰⁻¹⁴. These methods, however, often exhibit significantly larger predictive errors^{14,15}, or require

large numbers of regressions and fitting parameters to make predictions of partitioning for the ionized species.¹⁰⁻¹².

Consequently, it would be a significant improvement to develop a non-linear partitioning model for both neutral and ionizable pollutants onto various carbonaceous sorbents with the ability to predict partitioning directly from the chemical structure or physical chemical properties of the sorbates.

1.2 Research Goals

The overall goal of this doctoral dissertation is to develop methods to accurately predict the sorption and partitioning of neutral and ionic species into liquid organic phases as well as onto various forms of carbonaceous sorbents. For neutral solutes, this includes the development of a sorption isotherm – the log-normal Langmuir (LNL) isotherm model – to describe non-linear partitioning onto carbonaceous sorbents. This model is presented and tested in Chapter 2. Further, it is important that the parameters of the isotherm model can be predicted from the physical chemical properties of the sorbates, a model for which is presented in Chapter 3.

For charged species, this includes the development of a model that can accurately generate physical-chemical descriptors for charged species independent of their neutral counterparts. These solute descriptors are then used to predict both the solvent-water partitioning of ionic species (Chapter 4), as well as the sorption of ionic species onto soil organic carbon (Chapter 5), using the LNL model developed for neutral species.

Chapter 2

MODELING NON-LINEAR ADSORPTION TO CARBON WITH A SINGLE CHEMICAL PARAMETER: A LOG-NORMAL LANGMUIR ISOTHERM¹

Predictive models for linear sorption of solutes onto various media (e.g., soil organic carbon) are well-established. However, methods for predicting parameters for non-linear isotherm models (e.g., Freundlich and Langmuir models) are not. Predicting non-linear partition coefficients is complicated by the number of model parameters to fit each isotherm (e.g., Freundlich (2), Langmuir (2), or Polanyi-Manes (3)). For a data set with n sorbates and a single sorbent, the number of parameters to fit the entire data set is $2n$ (Freundlich, Langmuir) or $3n$ (Polanyi-Manes). The purpose of this chapter is to present a non-linear adsorption model with only one sorbate-specific parameter. To accomplish this, several simplifications to a log-normal Langmuir (LNL) isotherm model were explored. A single sorbate-specific binding constant, the median Langmuir binding constant, and two sorbent-specific parameters; the total site density, q_{max} , and the standard deviation of the Langmuir binding constant, σ_K , were employed. This single sorbate-specific (ss-LNL) model ($2 + n$ parameters) was demonstrated to fit adsorption data as well as the $2n$ parameter Freundlich model. The LNL isotherm model is fit to four data sets comprised of various chemicals sorbed to graphite, charcoal, and two types of activated carbon. The RMS errors of sorbed concentration (q) for the 3, 2, 2, and 1 sorbate-specific parameter LNL models were

¹ Appeared in Environmental Science & Technology 2015, 49, 7810-7817.

0.066, 0.068, 0.069, and 0.113, respectively. The median logarithmic parameter standard errors for the four models were 1.070, 0.454, 0.382, and 0.201 respectively. Further, the single sorbate-specific parameter model was the only model for which there were no standard errors of estimated parameters greater than a factor of 3. The surprising result is that very little increase in RMSE occurs when two of the three parameters, the standard deviation of the log of the median binding constant, σ_κ , and the maximum sorption capacity, q_{max} , are sorbate independent. However, the large standard errors present in the other models are significantly reduced. This remarkable simplification yields the single sorbate-specific parameter log-normal Langmuir (ss-LNL) model.

2.1 Introduction

Many models have been developed for adsorption of organic chemicals onto soils and carbonaceous materials, including black carbon^{16,17}. These models can be separated into two major classes: empirical and mechanistic. Although empirical models (e.g., the Freundlich isotherm) typically offer better fits to experimental data, they are not based on a mechanistic representation of the sorption process. For example, the Freundlich isotherm lacks a total site density parameter, never saturating (even at large aqueous concentrations), which is unrealistic. Conversely, mechanistic models (e.g., the Langmuir isotherm) employ a simplified model of the adsorption processes, but generally offer poor fits for the adsorption of organic sorbates onto black carbon¹⁶.

Statistically derived isotherm models, while less popular than Langmuir and Freundlich models, offer a compromise between empirical and mechanistic models. The Langmuir-Freundlich (LF) isotherm, proposed in 1948 by Sips, is one such

model^{18,19}. The LF isotherm offers the flexibility of the empirical models with the mechanistic basis of a Langmuir isotherm. It has been utilized extensively in the modeling of metal-ligand interactions for humic ligands²⁰⁻²⁴ and to model the adsorption energy distributions for organic compounds onto various types of carbon²⁵⁻²⁹.

The purpose of this chapter is to develop and apply a statistical isotherm model that accurately reproduces non-linear adsorption with only one sorbate-specific parameter. The model employs the basic principles of the LF model. However, the distribution of Langmuir binding constants is assumed to be log-normal²⁸⁻³⁰.

2.2 Modeling and Experimental Data

2.2.1 Log-normal Langmuir (LNL) Isotherm

The Langmuir isotherm³¹ is conveniently expressed as:

$$q(c) = \frac{q_{max} K_L c}{1 + K_L c} \quad (2-1)$$

where $q(c)$ is the sorbed concentration (mmol/kg sorbent) at aqueous concentration c (mM), K_L is the Langmuir binding constant (L/mmol sorbate) and q_{max} is the saturated monolayer sorption capacity (mmol/kg sorbent). The Freundlich isotherm³¹ is:

$$q(c) = K_F c^v \quad (2-2)$$

where K_F is the Freundlich binding constant (mmol/kg)(mM)^{1/v}, c is the aqueous concentration (mM), and v is the Freundlich exponent. The Langmuir-Freundlich isotherm¹⁸ is a superposition of Langmuir isotherms

$$q(c) = q_{max} \int_0^{\infty} \frac{cK_L}{1 + cK_L} f(K_L) dK_L = \frac{q_{max}(K_{LF}c)^v}{1 + (K_{LF}c)^v} \quad (2-3)$$

where q_{max} is the total site density (mmol/kg sorbent), K_L is the local Langmuir binding constant (L/mmol sorbate), and $f(K_L)$ is the probability density function (pdf) of the Langmuir binding constants. The approach that Sips employed was to derive a $f(K_L)$ which produces the LF isotherm, the right-hand side of Eq. (2-3). K_{LF} is the LF binding constant (L/mmol sorbate), and v is the LF exponent. When $v = 1$ the isotherm reduces to the Langmuir equation, and for $K_{LF}c \ll 1$, the isotherm reduces to the Freundlich isotherm with Freundlich constant, $K_F = q_{max}(K_{LF})^v$, since in Eq. (2-3) K_{LF} is inside the parentheses. This is to maintain consistent units of K_{LF} which do not depend on the exponent, v .

Sips observed that $f(K_L)$ is nearly a log-normal distribution¹⁸. Since the resulting mathematics is more straightforward and, as shown below, a more intelligible equation results; it is assumed that K_L is a log-normally distributed random variable^{27,29,30}. Defining κ as the natural logarithm of K_L :

$$\kappa = \ln(K_L) \quad (2-4)$$

Eq. (2-1) then becomes:

$$q(c) = \frac{q_{max}e^{\kappa}c}{1 + e^{\kappa}c} \quad (2-5)$$

The normal probability density function for κ is:

$$f(\kappa, \mu_{\kappa}, \sigma_{\kappa}^2) = \frac{1}{\sigma_{\kappa}\sqrt{2\pi}} \exp\left(\frac{-(\kappa - \mu_{\kappa})^2}{2\sigma_{\kappa}^2}\right) \quad (2-6)$$

where μ_{κ} and σ_{κ} are the mean and standard deviation of $\kappa = \ln(K_L)$. The LNL isotherm, Eq. (2-7), is the superposition of the Langmuir isotherm Eq. (2-5) weighted by the probability density function Eq. (2-6):

$$q(c) = \int_{-\infty}^{\infty} \left(\frac{q_{max} e^{\kappa} c}{1 + e^{\kappa} c} \right) \frac{\exp\left(-\left(\frac{\kappa - \mu_{\kappa}}{2\sigma_{\kappa}^2}\right)^2\right)}{\sigma_{\kappa} \sqrt{2\pi}} d\kappa \quad (2-7)$$

It is useful to transform Eq. (2-7) using a standard normal random variable, z :

$$z = \frac{\mu_{\kappa} - \kappa}{\sigma_{\kappa}} \quad (2-8)$$

so that Eq.(2-7) becomes:

$$q(c) = \frac{q_{max}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c e^{\mu_{\kappa} + z\sigma_{\kappa}}}{1 + c e^{\mu_{\kappa} + z\sigma_{\kappa}}} \right) \exp\left(-\frac{z^2}{2}\right) dz \quad (2-9)$$

For log-normal random variables, the relationship between the median $\tilde{K}_L$ and the log mean μ_{κ} is³²:

$$\tilde{K}_L = e^{\mu_{\kappa}} \quad (2-10)$$

Using this relationship, the log-normal Langmuir (LNL) isotherm is:

$$q(c) = \frac{q_{max}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c \tilde{K}_L e^{z\sigma_{\kappa}}}{1 + c \tilde{K}_L e^{z\sigma_{\kappa}}} \right) \exp\left(-\frac{z^2}{2}\right) dz \quad (2-11)$$

The isotherm has three parameters: $\tilde{K}_L$, the median Langmuir binding constant, σ_{κ} the standard deviation of $\kappa = \ln(\tilde{K}_L)$, and q_{max} , the total sorption capacity. Note that, as in the case of the Langmuir isotherm (Eq. (2-1)), the LNL isotherm has a dimensionless normalized concentration, $c\tilde{K}_L$, which, as shown below, is useful in data analysis and display.

The integral in the isotherm (Eq. (2-11)) can be evaluated numerically, using a Gauss-Hermite quadrature method, with a transformation of variables ($x = z\sqrt{2}$)³³:

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \cong \sum_{i=1}^n w_i f(x_i) \quad (2-12)$$

where $f(x)$ is the argument inside the integral (Eq. (2-11)), excluding the exponential term ($\exp(-z^2/2)$), and w_i and x_i are the weighting factors and roots of the n^{th} order Hermite polynomial:

$$w_i = \frac{2^n n! \sqrt{\pi}}{n^2 (H_{n-1}(x_i))^2} \quad (2-13)$$

that are listed in Appendix A.1. The final functional form of the isotherm is:

$$q(c) = \frac{q_{max}}{\sqrt{\pi}} \sum_{i=1}^n w_i \frac{c \tilde{K}_L \exp(\sqrt{2} \sigma_K x_i)}{1 + c \tilde{K}_L \exp(\sqrt{2} \sigma_K x_i)} \quad (2-14)$$

A computer program to evaluate this equation is presented in Appendix A.1.

2.2.2 Experimental Data

The experimental data used to develop the model was gathered from several different sources³⁴⁻³⁶ to incorporate various types of black carbon sorbents and sorbates from multiple chemical classes. Table 2-1 presents a summary and Figure 2-1 presents sample isotherm plots to illustrate the range and diversity of the data to be analyzed using the LNL isotherm model.

Table 2-1. Summary of experimental data used in construction of the log-normal Langmuir isotherm model (graphite³⁶, charcoal³⁶, Darco granular activated carbon (GAC)³⁴, and F400 GAC³⁵).

Sorbent	# of Sorbates	Sorbate Chemical Classes	Reference
Graphite	13	PAHs, Nitroaromatics, Chlorinated Aromatics	36
Charcoal	11	PAHs, Nitroaromatics, Chlorinated Aromatics	36
Darco GAC	14	Hydrocarbons, Ketones, Ethers, Mostly Non-aromatic	34
F400 GAC	44	PAHs, Nitroaromatics, Chlorinated Aromatics, Pesticides	35

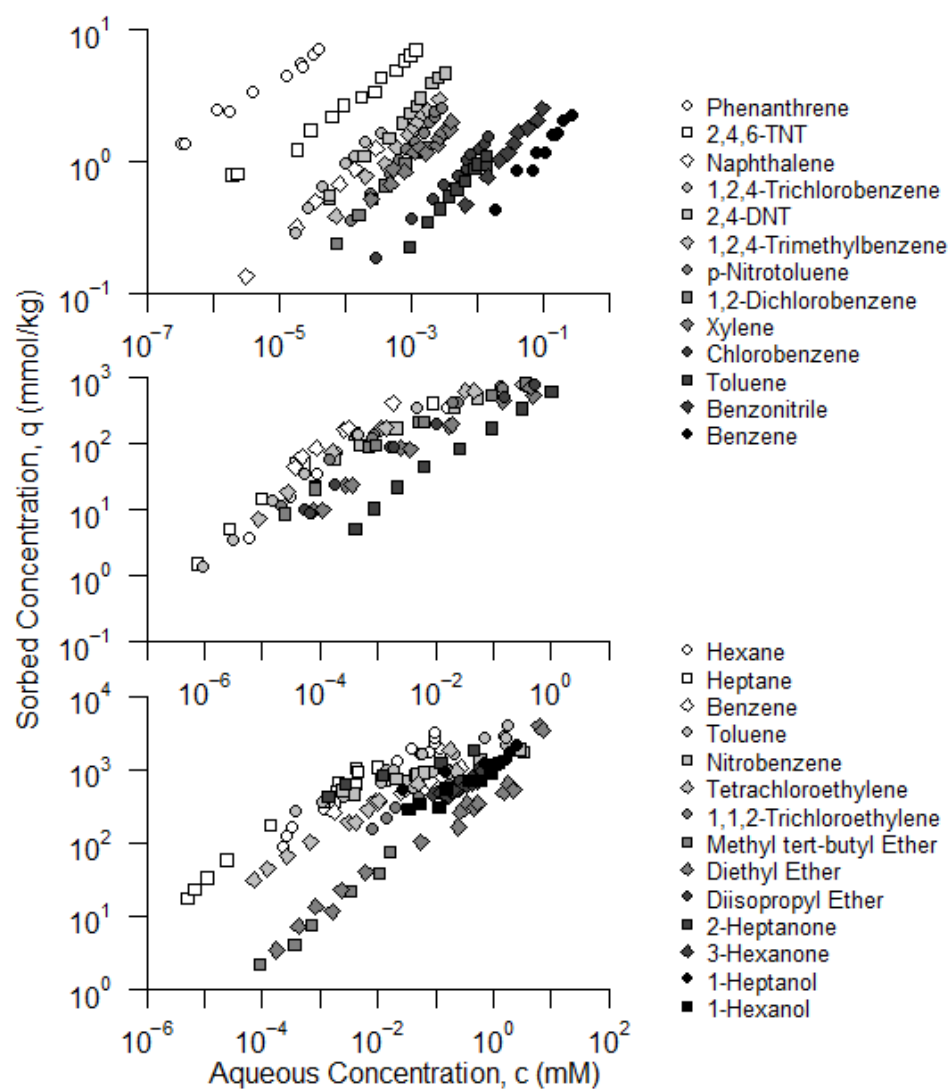


Figure 2-1. Equilibrium adsorption isotherm data for graphite (A), charcoal (B), and Darco GAC (C). The legend on the top right corresponds to panels (A) and (B). The legend on the bottom right corresponds to panel (C).

There are large differences in the apparent adsorption capacity and behavior of the sorbents. The isotherms span nine orders of magnitude in aqueous concentration and five orders of magnitude in sorbed concentration. For graphite (Figure 2-1A), the

isotherms are nearly linear on the logarithmic axes, while isotherms in the charcoal (Figure 2-1B) and Darco GAC (Figure 2-1C) data sets exhibit curvature. The F400 data set, which exhibits similar curvature, is omitted from this figure due to the size of the data set, but is analyzed subsequently. The curvature suggests that the aqueous concentrations are approaching the maximum sorption capacity much more rapidly than the Freundlich model would predict. The LNL isotherm has a maximum sorption capacity, q_{max} , which can be estimated for each of the sorbents, including graphite. This result is unexpected, since the individual isotherms do not appear to saturate. The reason this is possible is presented below.

2.3 Results

The log-normal Langmuir model is fit to the isotherm data for the three sorbent data sets discussed above as well as the larger F400 GAC data set. For each sorbent-sorbate isotherm pair there are three LNL isotherm parameters to be estimated: the maximum site density, q_{max} , the standard deviation of the site energy distribution, σ_κ , and the median Langmuir binding constant, $\tilde{K}_L$. The parameters were determined by minimizing the root mean square of the residuals using Excel's Solver package, with Eq. (2-15) as the prediction:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(q_{pred,i}) - \log(q_{obs,i}))^2} \quad (2-15)$$

where $q_{pred,i}$ is the LNL-computed sorbed concentration (mmol/kg) and $q_{obs,i}$ is the experimental sorbed concentration (mmol/kg) for the i^{th} observation. The logarithms of the median binding constants were estimated to ensure positivity of the estimated $\tilde{K}_L$. The standard errors of $\log(\tilde{K}_L)$, q_{max} , and σ_κ were estimated using SolverAid³⁷.

When all three parameters are allowed to vary for each of the n sorbates, there are $3n$ parameters, since each sorbate is fit with its own individual LNL isotherm. However, there are not enough data points for each sorbate to accurately estimate three parameters, which leads to large standard errors of the estimated parameters and large correlation coefficients between the fitted parameters.

Consequently, several simplifications to the LNL model were investigated, with the aim of reducing the number of sorbate-specific parameters. Three alternative models were investigated: either the standard deviation of the median Langmuir binding constants, σ_K , the total site density, q_{max} , or both were made sorbate-independent for each sorbent data set. The RMS errors and standard errors of the estimated LNL parameters were then used to compare the accuracy and potential predictive capabilities of the various models.

Figure 2-2 summarizes the model performances when the simplifications discussed above were applied to the largest data set, F400 GAC ($n = 44$). A summary of the model results is presented in Table 2-2. Figure 2-2A, a box plot of the residuals of the sorbed concentrations, $\log(q)$, demonstrates that while the RMS errors for the sorbate-specific σ_K and q_{max} models, 1-3, are smaller than those for the single sorbate-specific parameter LNL model, 4, most of the improvement is a reduction in the number of statistical outliers. Since model 4 has only one sorbate-specific parameter,

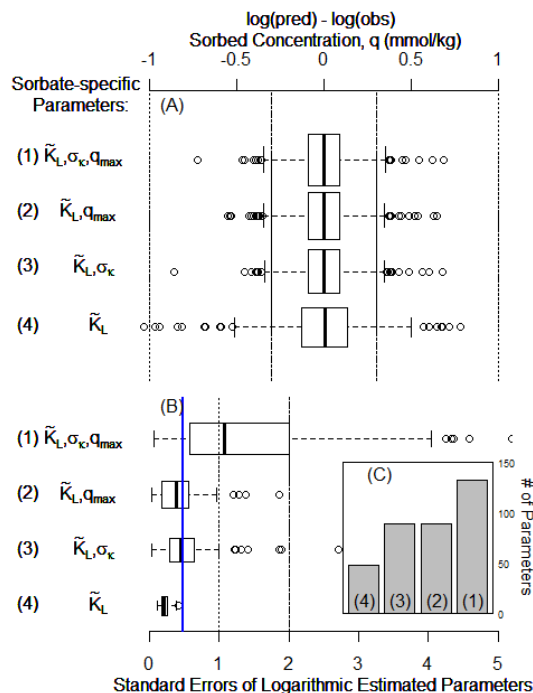


Figure 2-2. (A) Residual box plots of $\log(q_{pred}) - \log(q_{obs})$ for the LNL isotherm models, 1 through 4, on F400 GAC, where (1) $\tilde{K}_L$, σ_K , and q_{max} , (2) $\tilde{K}_L$ and q_{max} , (3) $\tilde{K}_L$ and σ_K , or (4) $\tilde{K}_L$ are the sorbate-specific parameters within the models. Semi-dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations, dotted lines represent ± 1.0 log-units difference between predicted and observed adsorbed concentrations. RMS errors for models 1 through 4 are 0.0659, 0.0680, 0.0690, and 0.113, respectively. (B) Box plots of the logarithmic standard errors of the estimated parameters for models 1 through 4. Dotted line represents 1 order of magnitude error in the estimated parameters, semi-dashed line represents 2 order of magnitude error in the estimated parameters, solid blue line represents 0.4 log-units error in the estimated parameters (a factor of 3). (C) Bar plot of the total number of estimated parameters for models 1 through 4. For the box plots, the interquartile range (IQR) contains 50% of the data, whiskers represent $\pm 1.5\text{IQR}$, and points represent outliers ($> 1.5\text{IQR}$).

it has significantly fewer degrees of freedom, compared to models 1-3. Consequently, there are more statistical outliers in model 4. However, the median residuals and the interquartile ranges (50% of the residuals) are not significantly different between the four models.

Table 2-2. Summary of LNL isotherm models, sorbent and sorbate-specific parameters, RMS errors, and median standard errors (SE) of the estimated parameters.

Model	Sorbent-specific Parameters	Sorbate-specific Parameters	Number of Estimated Parameters	RMSE	Median SEs of Estimated Parameters
(1)		$\tilde{K}_L, \sigma_\kappa, q_{max}$	132	0.066	1.07
(2)	σ_κ	$\tilde{K}_L, q_{max}$	89	0.068	0.454
(3)	q_{max}	$\tilde{K}_L, \sigma_\kappa$	89	0.069	0.382
(4)	σ_κ, q_{max}	$\tilde{K}_L$	46	0.113	0.201

By contrast, the reduction in the number of estimated parameters from $3n$ in model 1 to $2n + 1$ in models 2 and 3 to $n + 2$ in model 4 has a dramatic effect on the standard errors of the estimated parameters. The sorbate-specific $\tilde{K}_L$, σ_κ , and q_{max} model 1 has a median standard error of the estimated parameters of 1.070 with an interquartile range of 0.58 to 2.00 log-units. Half of the parameters have standard errors in excess of one log-unit. The standard errors are smaller for the individual σ_κ and q_{max} sorbate-specific models 2 and 3, with median standard errors of 0.454 and 0.382 and interquartile ranges of 0.287-0.644 and 0.175-0.562 log-units respectively.

While the median and interquartile ranges (IQR) of standard errors are significantly smaller for the individual σ_κ and q_{max} sorbate-specific models, they still contain a significant number of parameters with standard errors in excess of one log-unit (N = 4 and N = 7, respectively). The sorbent-specific σ_κ and q_{max} model 4 offers a similar goodness of fit to models 1-3 (Figure 2-2). However, the standard errors of the

estimated parameters are significantly smaller than the other three models, with a median standard error of 0.201 and an interquartile range of 0.167-0.261 log-units. Further, the model has no standard errors in excess of 0.5 log-units. By eliminating the sorbate-specificity of the σ_K and q_{max} parameters, the remaining parameter, $\tilde{K}_L$, is the only sorbate-specific parameter.

Others have proposed methods for reducing the number of sorbate-specific parameters; for example, the use of a fixed activity linear partition coefficient as a predictive parameter for non-linear adsorption onto heterogeneous sorbents³⁴. While this method offers a single sorbate-specific parameter, the selection of the fixed activity, as well as whether the fixed activity should be the sorbed or aqueous activity, is unclear. Further, sorption isotherms cannot be recreated directly from a single fixed activity partition coefficient. These fixed activity coefficients must then be related back to isotherm parameters (e.g., Freundlich K_f and ν)³⁴. Consequently, the single sorbate-specific parameter of the LNL isotherm model greatly simplifies constructing a predictive model, as shown in consequent chapters.

It is remarkable that σ_K and q_{max} can be assumed to be sorbate independent and still achieve a fit to the data that is nearly as good as with sorbate-specific σ_K and q_{max} . The median Langmuir binding constant, $\tilde{K}_L$, is unique for each chemical. Thus the isotherm model analyzed subsequently has two sorbent-specific parameters: the maximum site density, q_{max} , and the standard deviation of the Langmuir binding constants, σ_K ; and one sorbate-specific parameter, $\tilde{K}_L$. It should be noted that this significant reduction in the number of estimated parameters also significantly reduces the number of sorbates needed to describe the sorbent-specific parameters, q_{max} and σ_K . To control standard errors of the estimated parameters, it is suggested that a

minimum of 10 sorbates be included when estimating q_{max} and σ_k for a new sorbent, giving a redundancy in the number of data points per parameter of approximately 8:1 (with 5 isotherm data points per sorbate). Estimated isotherm parameters for the four sorbent data sets are presented in Tables A-1a,b and A-2a-d in Appendix A. A discussion of the method for estimating the standard errors of the median Langmuir binding constants, $\tilde{K}_L$, separately can be found in section A.4 of the Appendix.

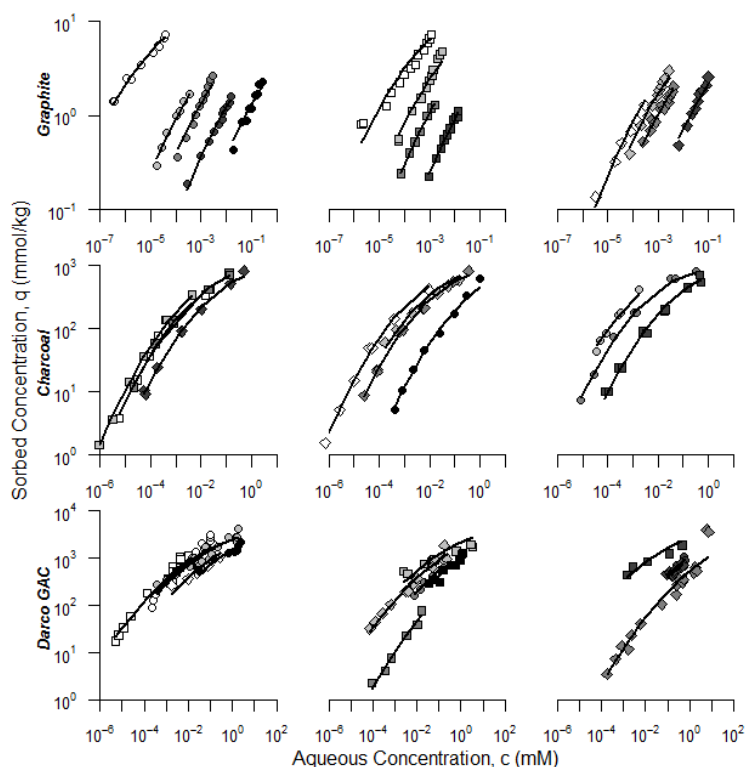


Figure 2-3. Adsorption isotherms for the single sorbate-specific parameter log-normal Langmuir (ss-LNL) model for graphite, charcoal, and Darco GAC. The sorbate symbols are the same for Figures 1 and 3. Solid lines represent the fit to the LNL isotherm model. Estimated LNL parameters are presented in Appendix A (Tables A-1a,b).

Figure 2-3 illustrates the single sorbate-specific parameter (ss-LNL) isotherm fits for three data sets (charcoal, graphite, and Darco GAC). The ss-LNL model provides excellent fits to the experimental sorption data. The model is able to fit both the non-saturating behavior observed in the graphite data set as well as the curvature apparent in the charcoal and GAC data sets. Fits for the F400 data set are comparable, and are available in Appendix A as Figure A-1. For the four data sets, root mean square errors (Eq. (2-15)) are calculated to evaluate the overall goodness of fit of the model and are presented in Table 2-3.

Table 2-3. Summary of $\log(\tilde{K}_L)$ standard errors, individual, and combined RMS errors (Eq. (2-15)) for the ss-LNL model fits for graphite, charcoal, Darco GAC, and F400 GAC. Units of $\tilde{K}_L$ are $(mM)^{-1}$

Sorbent	Parameters	Total RMSE ¹	RMSE Range ²	Range of SE($\log(\tilde{K}_L)$)
Graphite	15	0.0530	0.00912 - 0.0937	0.384 – 0.404
Charcoal	13	0.0515	0.0293 – 0.0699	0.121 – 0.132
Darco GAC	16	0.121	0.0336 – 0.214	0.423 – 0.477
F400 GAC	46	0.113	0.0258 – 0.441	0.112 – 0.416

$$^1 RMSE_{total} = \sqrt{\frac{\sum (\log(q_{pred}) - \log(q_{obs}))^2}{n}}$$

²RMS Errors for each sorbate isotherm, i:

$$RMSE_i = \sqrt{\frac{\sum_i (\log(q_{pred}) - \log(q_{obs}))^2}{n_i}}$$

The RMS errors (Eq. (2-15)) range from RMSE = 0.053 (graphite) and 0.051 (charcoal) to RMSE = 0.121 (Darco GAC) and 0.113 (F400 GAC), suggesting that the fits for the GACs are slightly worse. Since these are logarithmic RMS errors, they measure the ratio of q_{pred} to q_{obs} . For RMSE = 0.05 and 0.10 the error in the ratio is $10^{\pm RMSE}$, which are 0.89-1.12 and 0.79-1.26, or approximately a 10-25% error

respectively. Although the RMS errors for the GACs are larger, they are still quite small. The range of individual sorbate RMS errors is presented in Table 2-3.

2.4 Discussion

2.4.1 Normalized LNL Isotherm

A useful feature of the single sorbate-specific LNL model is that isotherms for multiple sorbates can be presented in a single graph of $q(c)$ versus $c\tilde{K}_L$ (Figure 2-4A). The result is a single normalized isotherm for all sorbed chemicals whose saturation limit is q_{max} and standard deviation is σ_κ . This is useful for both data comparison and to illustrate the behavior of the isotherm.

Figure 2-4A illustrates the effect of varying σ_κ on the resulting normalized isotherm. For small values of σ_κ ($\sigma_\kappa = 0 - 1.0$), the isotherm approaches the Langmuir isotherm since the distribution of K_L is very narrow. As σ_κ increases, the isotherm slope decreases over the range of $c\tilde{K}_L$. Larger values of σ_κ ($\sigma_\kappa = 5.0$) represent sorbents with broader distributions of site energies. For the same q_{max} , a larger σ_κ results in the same number of total sites, but a larger distribution of site binding constants each with a smaller sorption capacity. Therefore for the isotherm with the larger σ_κ the individual sites with the largest binding constants saturate at lower concentrations, raising the sorbed concentration over an isotherm with a smaller σ_κ . The result is a shallower slope for the normalized isotherm. Large values of σ_κ should be found for very heterogeneous sorbents (e.g., GAC), while one might expect smaller values of σ_κ for more homogeneous sorbents.

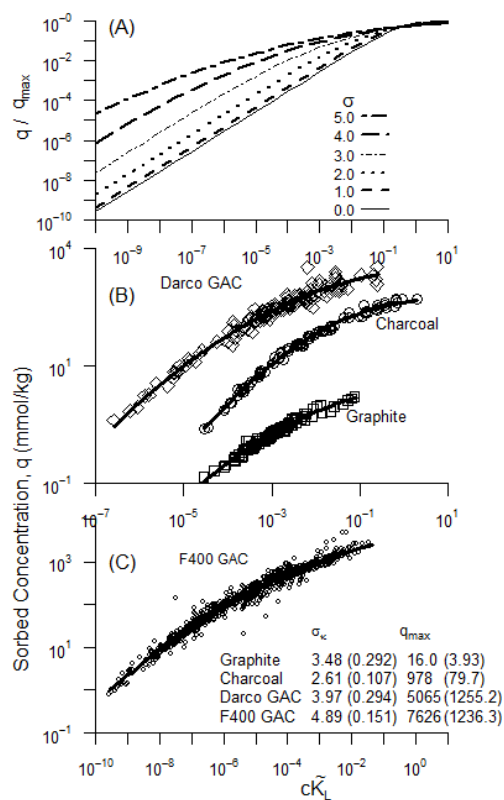


Figure 2-4. (A) Effect of systematically varying the standard deviation of the median binding constant, σ_{κ} , on the shape of the normalized ss-LNL isotherm. (B),(C) Normalized isotherms for graphite, charcoal, and Darco GAC (B) and F400 GAC (C). Estimated LNL isotherm parameters and associated standard errors can be found in Tables A-1a,b and A-2a-d in Appendix A.

Note that as the concentration moves significantly far from saturation, the LNL isotherm slope becomes a constant, approaching linearity, similar to that of an individual Langmuir isotherm.

The normalized isotherm plots for the data are presented in Figures 2-4B and 2-4C. The normalized isotherm illustrates the ability of the LNL isotherm to fit a wide

variety of isotherm shapes as well as sorbents of varying site heterogeneity. The robustness of the model is most easily observed in the F400 data set. With 44 chemicals, the model is capable of obtaining fits comparable to or better than the conventional Freundlich isotherm³⁵ with only one parameter per sorbate (see Figures A-2a-d in Appendix A). Estimated Freundlich constants and their associated standard errors can be found in Tables A-3a-d in the Appendix.

Additionally, it can be seen that the more heterogeneous, highly-functionalized GACs have larger σ_K values than the more uniform graphite and charcoal. Interestingly, the σ_K for graphite is larger than that of charcoal, whereas one might expect the reverse. This might indicate that the differences in energy between edge, surface, or corner sites in graphite are larger than the energy differences between the sites in charcoal. However, without further quantitative descriptions of the surfaces it is difficult to draw firm conclusions. The trend in the fitted q_{max} values was less surprising, as the non-porous graphite had the smallest maximum site density (q_{max} = 15.99 mmol/kg-carbon), while the highly heterogeneous and porous Darco and F400 activated carbons had the largest (q_{max} = 5065 and 10199 mmol/kg-carbon, respectively).

Of the four data sets, F400 GAC has the smallest coefficient of variation (CV) for the estimated q_{max} (approximately 1.18%, compared to 24.6, 8.15%, and 24.8% for graphite, charcoal, and Darco GAC, respectively). The coefficients of variation for σ_K and q_{max} are listed in Tables A-1a,b and A-2a,d in the Appendix. Note in Figure 2-3 that the individual isotherms on charcoal exhibit curvature as the aqueous concentration increases, tending towards saturation. This results in smaller standard errors for the predicted q_{max} compared to graphite and Darco GAC.

This non-linearity is not clear, however, for the graphite data set. Nevertheless, a q_{max} can be estimated with a comparable standard error. For graphite, it is the differences in the individual isotherm slopes that are used to estimate q_{max} accurately. The open circles (phenanthrene) and open diamonds (naphthalene) in Figure 2-3 illustrate this point. Because phenanthrene sorbs more strongly over the same aqueous concentration range, it has a much larger median binding constant. However, it has a shallower slope. By contrast, the slope for naphthalene is larger, but the median binding strength is smaller (~ 3 orders of magnitude). When these isotherms are superimposed by estimating $\tilde{K}_L$ for each compound, the steeper slope at the lower normalized concentration merges with the shallower slope at the higher normalized concentration such that they trend towards a saturation limit, as can be seen in Figure 2-4C. This is the reason that the estimate for q_{max} has a comparable standard error. It is the shape of the composite isotherm (Figure 2-4C) that determines the accuracy with which q_{max} can be estimated.

Finally, the ranges of the median Langmuir binding constants, $\log(\tilde{K}_L)$, were similar for the various sorbents (see Tables A-1a,b and A-2a-d in Appendix A). The range of binding constants for graphite, charcoal, and Darco GAC were approximately 4 orders of magnitude, while the range of the binding constants for F400 GAC was slightly larger, approximately 6 orders of magnitude.

2.4.2 Residual Analysis

An analysis of the residuals for the four data sets using box plots is presented in Figures 2-5 and 2-6. With the exception of two chemicals (oxamyl and simazine onto F400 GAC) the interquartile ranges (IQR) of the residuals of $\log(q_{pred})$ do not exceed 0.3 log-units for the LNL model. Further, the residual median values are

centered about zero, indicating a lack of bias in the predicted isotherms, with the exception of oxamyl and simazine. However, it was determined that adequate methods for determining the adsorption of these chemicals were unavailable at the time of the experiments and instead total organic carbon (TOC) analysis was used as a surrogate for a more direct measure (Speth, personal communication)³⁵. Consequently, it is believed that the lack of agreement to the LNL model for these chemicals is a result of experimental error.

Since the conventional model for non-linear adsorption data is a Freundlich isotherm, the LNL isotherm was compared to individual Freundlich isotherms for the sorbates in all four sorbent systems. Further, the Freundlich sorption isotherm is the basis for the poly-parameter predictive models proposed by Shih & Gschwend³⁴ for Darco GAC. Consequently, the underlying ability of the sorption isotherms to accurately reproduce the data must be examined. The accuracy of the resulting LNL and Shih & Gschwend predictive models is compared in Chapter 3 of this dissertation. Table 2-4 summarizes the resulting RMS errors for the two isotherm models.

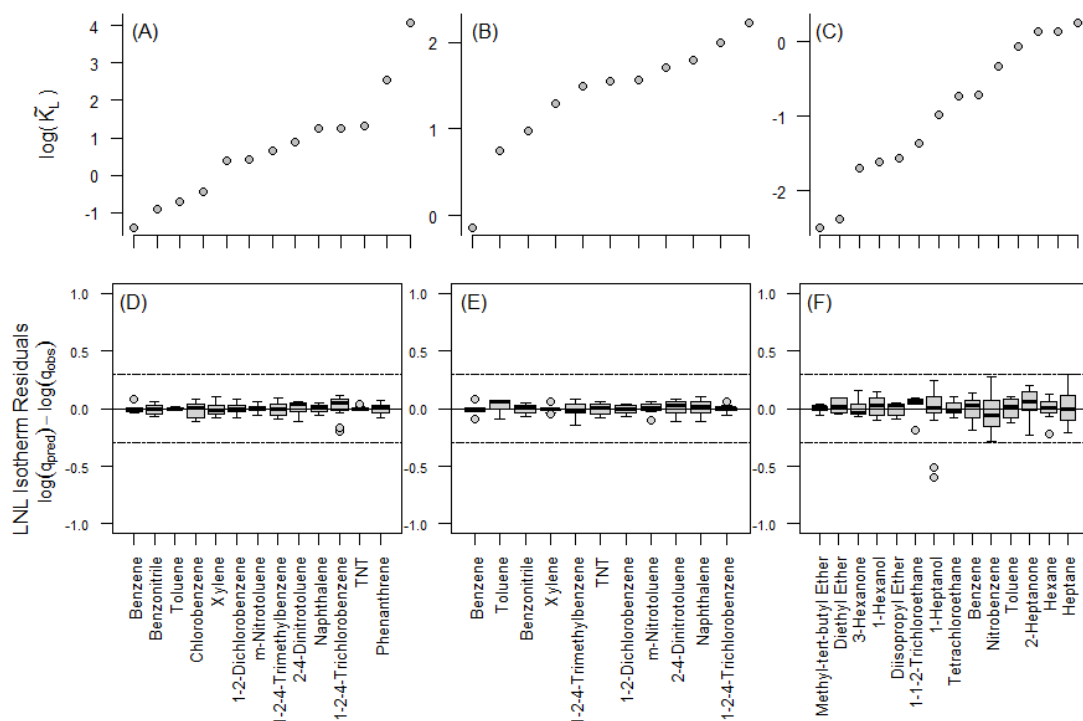


Figure 2-5. (A), (B), (C) Plots of LNL median binding constants, sorted from smallest to largest for (A) graphite, (B) charcoal, and (C) Darco GAC. (D), (E), (F) Box plots of the LNL isotherm model residuals ($\log(q_{pred}) - \log(q_{obs})$) for (D) graphite, (E) charcoal, and (F) Darco GAC. The solid lines represents 1:1 agreement between q_{pred} and q_{obs} adsorbed concentrations, dashed lines represent ± 0.3 log-units difference between observed and modeled adsorbed. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and points represent outliers (> 1.5 IQR).

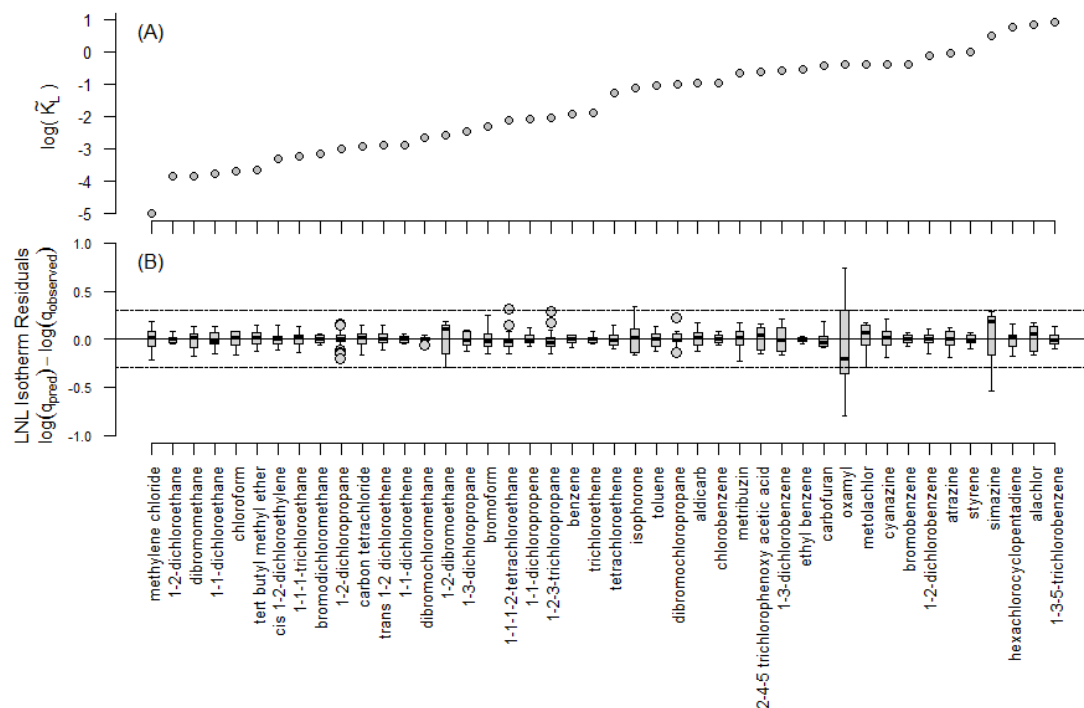


Figure 2-6. (A) Plot of LNL median binding constants, sorted from smallest to largest for F400 GAC. (B) Box plot of the LNL isotherm model residuals ($\log(q_{pred}) - \log(q_{obs})$) for F400 GAC. The solid line represents 1:1 agreement between q_{pred} and q_{obs} adsorbed concentrations, dashed lines represent ± 0.3 log-units difference between observed and modeled adsorbed. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and points represent outliers (> 1.5 IQR).

Table 2-4. Comparison of Freundlich and LNL isotherm RMS errors (Eq. (2-15)) for graphite, charcoal, Darco GAC, and F400 GAC.

Sorbent	N	RMSE	
		Freundlich	LNL
Graphite	13	0.142	0.053
Charcoal	11	0.118	0.0515
Darco GAC	14	0.0941	0.121
F400 GAC	44	0.0674	0.113
Combined	82	0.0895	0.0985

For graphite and charcoal, the LNL model outperforms the traditional Freundlich model. This is surprising since the data for graphite (and to a lesser degree, charcoal) appears to be mostly log-linear (Figure 2-1). However, by fitting a smaller median binding constant for the more log-linear chemicals, the LNL model is able to fit non-saturating isotherms. The Darco and F400 GAC data sets, however, show the Freundlich model performing better than the LNL model. For Darco GAC this difference in RMS error is marginal (corresponding to $\sim 7.9\%$ difference in the ratio of q_{pred}/q_{obs} for the two models). For F400 GAC, some of this can be explained by the two outliers in the data set, oxamyl and simazine, discussed previously. Because there is a single sorbent-specific value of q_{max} and σ_{κ} , experimental errors in a single chemical can result in significant error in the LNL isotherm. This is not the case for the Freundlich isotherm, where each sorbate is modeled independently. When these two chemicals are removed, the difference in the RMS errors between the Freundlich and LNL models becomes smaller (from 0.0674 and 0.113 to 0.0647 and 0.0878, respectively), corresponding to a 6.34% difference in the ratio of q_{pred}/q_{obs} for the two models, similar to that of Darco GAC. These differences are small, however, compared to the significant reductions in degrees of freedom for the LNL model. Box

plots comparing the residuals were also constructed for the Freundlich isotherm model (see Figures A-2a-d in Appendix A) for comparison.

2.5 Implications

The log-normal Langmuir model is able to model adsorption onto a range of sorbents with the sorbate contribution described with a single sorbate-specific parameter, $\tilde{K}_L$. This is a significant reduction in degrees of freedom over the conventional Freundlich model with no significant increase in RMS error. The single parameter is directly related to the sorption strength, which suggests that it can be predicted using chemical descriptors, as demonstrated in Chapter 3 of this dissertation. The sorbent-specific q_{max} and σ_K approximation was found to be satisfactory for the four sorbent systems, indicating that there is a comparable accessibility and relative affinity for the various sorbent sites independent of the sorbate, and that only the median binding strength, $\tilde{K}_L$, is characteristic of the sorbate. It is expected that for neutral organic compounds (e.g., pesticides, PAHs, chlorinated compounds), the LNL model can work well in modeling adsorption behavior onto similar carbonaceous sorbents (e.g., GAC, graphite, and chars/charcoals). Further, it is expected that the model will be applicable to additional sorbents which exhibit non-linear sorption behavior (e.g., biochar, carbon nanotubes (CNTs)). However, it should be noted that most natural systems contain natural organic matter (NOM) and complex mixtures of sorbates. To be applicable for this situation, the LNL model needs to be modified. Since the LNL model is constructed from a distribution of individual Langmuir isotherms, it can be modified for competitive sorption systems.

Chapter 3

MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: PREDICTING CHEMICAL MEDIAN LANGMUIR BINDING CONSTANTS²

Procedures for accurately predicting linear partition coefficients onto various sorbents (e.g., organic carbon, soils, clay) are reliable and well-established. However, similar procedures for the prediction of sorption parameters of non-linear isotherm models are not. The purpose of this chapter is to present a procedure for predicting non-linear isotherm parameters, specifically the median Langmuir binding constant, $\tilde{K}_L$, obtained utilizing the single sorbate-specific parameter log-normal Langmuir isotherm developed in Chapter 2. A reduced poly-parameter linear free energy relationship (pp-LFER) is able to predict median Langmuir binding constants for graphite, charcoal and Darco granular activated carbon (GAC) sorption data. For the larger F400 GAC data set, a single pp-LFER model was insufficient, as a plateau is observed for the median Langmuir binding constants of larger molecular volume sorbates. This volumetric cutoff occurs in proximity to the median pore diameter for F400 GAC. A log-linear relationship exists between the aqueous solubility of these large compounds and their median Langmuir binding constants. Using this relationship for the chemicals above the volumetric cutoff and the pp-LFER below the cutoff, the median Langmuir binding constants can be predicted with root mean square

² Appeared in Environmental Science & Technology 2015, 49, 7818-7824

errors for graphite ($n = 13$), charcoal ($n = 11$), Darco GAC ($n = 14$), and F400 GAC ($n = 44$) of 0.129, 0.307, 0.407, and 0.424 respectively.

3.1 Introduction

Linear free energy relationships (LFERs) are widely used to predict aqueous solubility, solvent-water partitioning, soil-water partitioning, and chemical toxicity to aquatic organisms^{2,5,38-44}. Predicting linear partition coefficients (e.g., organic carbon-water partition coefficients) is generally straightforward and such models are typically accurate to within an order of magnitude in predicted sorbed concentrations⁸. This is not the case, however, for parameters in non-linear isotherm models (e.g., Langmuir and Freundlich binding constants). Attempts have been made to improve the prediction of these non-linear parameters by including sorbate activity terms³⁴. However, these attempts do not match the accuracy and simplicity of the predictive methods for linear parameters. The problem is complicated by the fact that the non-linear isotherm models commonly used (e.g., Freundlich, Langmuir, or Polanyi-Manes) have two or three parameters for each chemical. Therefore, predicting the isotherm for a new chemical requires models for each of the sorbate-specific parameters (e.g., for the Freundlich constant and the exponent).

The model presented in Chapter 2 fits the isotherm data with two sorbent-specific parameters and one sorbate-specific parameter with little to no increase in RMS error when compared to the Freundlich isotherm model (2 sorbate-specific parameters). Once the sorbent-specific parameters are estimated, the isotherm for a new chemical sorbing to the same sorbent requires a model for only the one sorbate-specific parameter. The purpose of this chapter is to present such a model.

3.2 Materials and Data

Adsorption data was gathered from several different sources³⁴⁻³⁶ to incorporate various carbon and sorbate types. The single sorbate-specific parameter log-normal Langmuir isotherm (ss-LNL) was fit to these four data sets. The isotherm is a linear superposition of Langmuir isotherms:

$$q(c) = \frac{q_{max}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c\tilde{K}_L e^{z\sigma_\kappa}}{1 + c\tilde{K}_L e^{z\sigma_\kappa}} \right) \exp\left(-\frac{z^2}{2}\right) dz \quad (3-1)$$

where q_{max} is the maximum site density (mmol/kg sorbent), $\tilde{K}_L$ is the median Langmuir binding constant for the distribution of sites (mM)⁻¹, σ_κ is the standard deviation of κ , the natural logarithm of $\tilde{K}_L$, c is the aqueous concentration (mM), and z is the variable of integration. The Langmuir binding constants are assumed to be log-normally distributed, with a median binding constant, $\tilde{K}_L$, and a standard deviation, σ_κ , of $\kappa = \ln(\tilde{K}_L)$. It has been found that the isotherm fits are not significantly degraded if q_{max} and σ_κ are assumed to be sorbate-independent. Only $\tilde{K}_L$ is sorbate-specific. Table 3-1 presents a summary of the sorbent parameters for four black carbons.

Table 3-1. Summary of the single sorbate-specific log-normal Langmuir (ss-LNL) estimated isotherm parameters and RMS errors for graphite³⁶, charcoal³⁶, Darco GAC³⁴, and F400 GAC³⁵.

Sorbent	N	Log(q_{max}) ($\frac{mmol}{kg}$)	σ_κ	RMSE
Graphite	13	1.203	3.481	0.0530
Charcoal	11	2.991	2.610	0.0515
Darco GAC	14	3.704	3.969	0.121
F400 GAC	44	4.008	4.962	0.113

Abraham pp-LFER parameters for the solutes were obtained using the Absolv⁴⁵ software package. Experimental Abraham solute parameters are available for

many neutral organic compounds⁴⁶. However, experimental solute parameters were unavailable for a subset of chemicals in the F400 data set. Further, it is often unclear whether reported sorbent-water system parameters (e , s , a , b , v , c in Eq. (3-2) below) are obtained using strictly experimental or Absolv-estimated sorbate parameters. The sorbent-water system parameters can vary considerably (without significantly affecting the RMS error of the pp-LFER model) depending on the original source of Abraham sorbate parameters. Consequently, Absolv-estimated sorbate parameters were used exclusively in the development of the pp-LFER models to ensure consistency. Other physical chemical parameters: aqueous solubility and molecular weight were obtained from ChemSpider⁴⁷ using the chemical CAS numbers. A complete table of Absolv-estimated Abraham parameters and aqueous solubilities can be found in Tables B-1a-c in the Appendix.

3.3 Results & Discussion

3.3.1 Abraham Poly-parameter Linear Free Energy Relationship (pp-LFER)

An Abraham pp-LFER model for the LNL median Langmuir binding constant is expressed:

$$\log(\tilde{K}_{L,i}) = e_c E_i + s_c S_i + a_c A_i + b_c B_i + v_c V_i + c_c \quad (3-2)$$

where the upper case letters represent the solute (i), contributions to the binding energy; E , molar refractivity, S , polarizability, A , hydrogen bond donating; and B , hydrogen bond accepting capacities, and V , the energy required to form a cavity. The lower case letters represent the complementary sorbent (c) relative to water; excess molar refractivity, polarizability, hydrogen bond accepting and donating, and cavitation energy terms, respectively. The c_c term accounts for any non-specific

binding interactions not predicted by the Abraham solute parameter product terms ($e_c E_i$, $s_c S_i$, $a_c A_i$, $b_c B_i$, and $v_c V_i$) and varies with the selected units of $\tilde{K}_L$. Finally, sorbent-water parameters (Eq. (3-2), $e_c \dots v_c$) were estimated using a multiple linear regression and the solver add-in for Excel for each of the four data sets. The resulting predicted median Langmuir binding constants are presented in Figure 3-1 and summarized in Table 3-2.

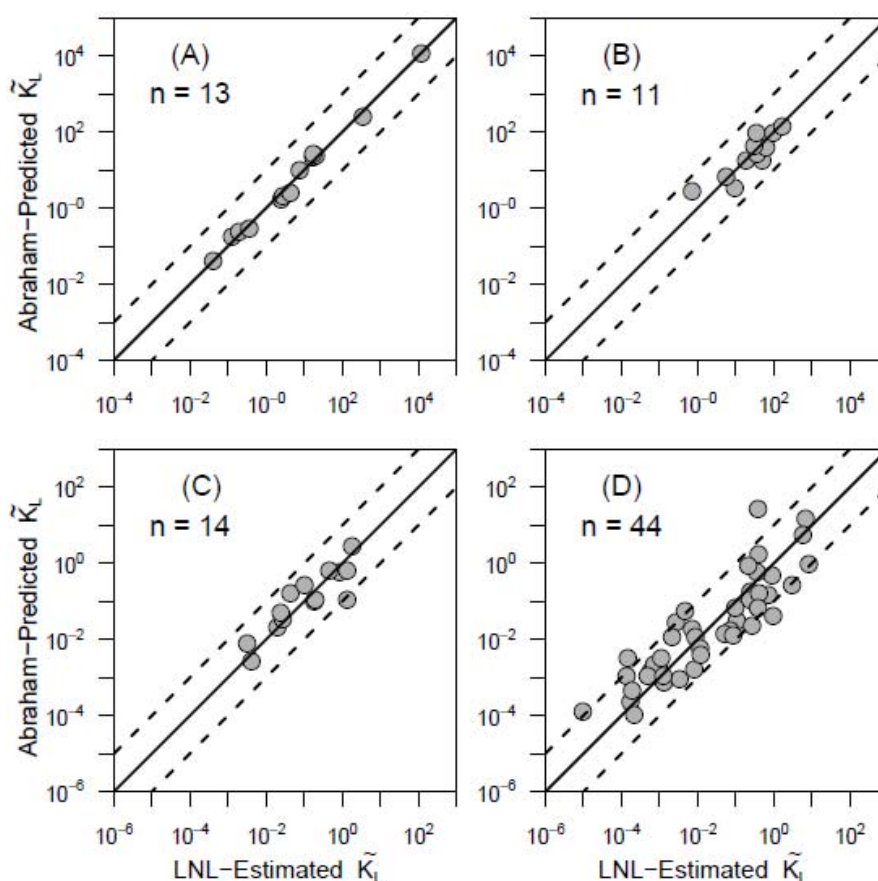


Figure 3-1. Plot of Abraham-predicted vs. LNL-estimated median binding constants ($\tilde{K}_L$) for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC. Solid lines represent 1:1 agreement; dashed lines represent ± 1 order of magnitude. RMSE for the four sorbent data sets are 0.129, 0.307, 0.407, and 0.709, respectively.

Table 3-2. Summary of estimated sorbent-water system Abraham parameters for graphite, charcoal, Darco GAC, and F400 GAC. Separate RMS errors are indicated for the separate equations in the revised F400 model. The notation “All”, “V < 110” and “V > 110” refers to the data used for the estimates. The standard errors (SE) and coefficient of variation (CV) of the parameters for each model are reported. Units of $\tilde{K}_L$ and solubility (S) are $(mM)^{-1}$

Sorbent	Solutes	e	s	a	b	v	c	m	b	N ¹	RMSE ²
<i>Graphite</i>	All	1.184	[-]	[-]	-2.91	5.543	-5.681	[-]	[-]	13	0.129
	SE	0.207	[-]	[-]	0.523	0.417	0.281	[-]	[-]		
	CV (%)	17.5%			18.0%	7.52%	4.95%				
<i>Charcoal</i>	All	1.032	[-]	[-]	-3.607	2.577	-1.519	[-]	[-]	11	0.307
	SE	0.714	[-]	[-]	1.419	1.084	0.765	[-]	[-]		
	CV (%)	69.2%			39.3%	42.1%	50.4%				
<i>Darco GAC</i>	All	1.669	[-]	[-]	-6.163	5.122	-4.823	[-]	[-]	14	0.407
	SE	0.564	[-]	[-]	1.281	1.067	1.006	[-]	[-]		
	CV (%)	33.8%			20.8%	20.8%	20.9%				
<i>F400 GAC (all sorbates)</i>	All	3.387	[-]	-3.012	-1.351	1.851	-5.271	[-]	[-]	44	0.709
	SE	0.1092	[-]	0.2077	0.1046	0.1108	0.0735	[-]	[-]		
	CV (%)	3.22%	[-]	6.90%	7.75%	5.99%	1.40%				
<i>F400 GAC (with cutoff)</i>	V < 110	-1.503	6.744	-3.138	[-]	4.680	-9.025	[-]	[-]	31	0.402
	SE	0.2996	0.4370	0.5389	[-]	0.1922	0.1853	[-]	[-]		
	CV (%)	19.93%	6.48%	17.17%	[-]	4.11%	2.05%				
	V > 110	[-]	[-]	[-]	[-]	[-]	[-]	-0.5082	-0.2024	13	0.513
	SE	[-]	[-]	[-]	[-]	[-]	[-]	0.03614	0.05123		
	CV (%)							7.11%	25.3%		

0.424³

3.3.2 Graphite

For the graphite data set³⁶, there were no solutes capable of donating hydrogen bonds. Consequently, the aA term was omitted from the pp-LFER regression (Eq. (3-2)). Further, previous studies have shown a cross-correlation between the E and S sorbate parameters⁴⁸. Consequently, the sS term was omitted and the excess molar refractivity term (eE) was allowed to absorb these correlation effects.

With the reduction in the degrees of freedom in the regression, the standard errors of the remaining parameters were reduced without significantly affecting the RMS error (see Table 3-2 for RMSE equation) of the model (RMSE = 0.129). Figure 3-1A compares the Abraham-predicted and LNL-estimated median binding constants. The fit is remarkably good, partly because 4 parameters are being fit to 13 data points. However, the standard errors are not excessively large. Table 3-2 contains a complete description of the estimated parameters for the four sorbents and pertinent information for the models. Table B-2 in the Appendix contains the estimated parameters, standard errors, and RMS errors for the full Abraham regressions for graphite, charcoal, and Darco GAC.

3.3.3 Charcoal

For the charcoal data set³⁶, reductions similar to the graphite Abraham model were observed. The resulting model for the charcoal-water median Langmuir binding constants was identical to that for graphite. The result of the regression was an RMS error of 0.307 with 4 parameters fit to 11 chemicals. Figure 3-1B compares the Abraham-predicted and LNL-estimated median binding constants. There is more scatter and the standard errors in the parameters are quite large, within one-half of the

parameter estimates, indicating that there is an insufficient number of chemicals to produce a good model fit for charcoal. The charcoal data set has the same sorbates (with the exception of phenanthrene and chlorobenzene) as the graphite data set. While some of the increase in standard errors of the estimated parameters can be attributed to the decrease in degrees of freedom (11 vs. 13 chemicals), the range of sorbate parameter values may also play a significant role. Phenanthrene and chlorobenzene have values of E and V that are at the high and low ends of the range of these parameters for the charcoal data set. Removing these sorbate parameters significantly increases the standard errors of the corresponding sorbent Abraham parameters, since the range of the estimators (in this case E and V) are significantly reduced. Therefore, the standard error reflects both the amount of data and the range of the sorbate parameters.

3.3.4 Darco GAC

Constructing the pp-LFER model for the Darco GAC data set³⁴ was more difficult, since several of the chemicals exhibit hydrogen-bond donating capability. However, there were not enough of these chemicals to accurately fit a GAC-water hydrogen bond accepting parameter (a), since only 3 sorbates are capable of donating hydrogen bonds. Consequently, the aA term was omitted in the formulation of the GAC-water median binding constant pp-LFER. Figure 3-1C illustrates the fit obtained for Darco GAC. There is one chemical that appears to be an outlier. The standard errors of the estimated parameters are approximately 20% of the estimated parameter values. The result of the regression was an RMS error of 0.407 with 4 parameters estimated for 15 chemicals.

3.3.5 F400 GAC

The F400 GAC data set³⁵ was considerably larger than the other data sets and consequently all of the Abraham parameters can be estimated. Unlike the other data sets, there were a significant number of solutes that exhibit hydrogen bond donating capability. The result was an RMS error of 0.709 with 6 parameters estimated for 44 chemicals. The RMS error is considerably larger than those for graphite (0.129), charcoal (0.307), and Darco GAC (0.407) and suggests that additional factors (e.g., sorbent-site accessibility) may be involved in the sorption to F400 GAC. While the standard errors on the estimated e , a , b , v , and c parameters are significantly smaller than those for the other data sets, the standard error for s is approximately 7 times the estimated value. Since there is a strong correlation between the solute E and S parameters⁴⁸ and the estimated value for s is very nearly zero, it was removed from the model and the remaining parameters were refit with little change to the RMS error or the remaining estimated parameters. A summary of the removal of the polarizability term from the F400 analysis can be found in Table B-3 in Appendix B.

Figure 3-1D shows that a single pp-LFER model (Eq. (3-2)) generally overestimates compounds with small binding constants (weakly sorbing) and underestimates compounds with large binding constants (strongly sorbing). This suggests one of two possibilities: either the solute Abraham parameters for some of these larger more structurally complicated pesticide compounds are poorly-characterized, or that chemicals with certain properties are interacting with the sorbent differently, and a single pp-LFER model applied to the entire suite of solutes is not capable of capturing these different interactions.

3.3.6 Separation of Chemicals

Sorption to GAC occurs predominantly within the pore structure. It was thought that the larger chemicals may be excluded from binding to certain pore sites⁴⁹. A plot of the LNL-estimated median binding constants, $\tilde{K}_L$, versus the solute McGowan volumes for the four data sets is shown in Figure 3-2. For the nonporous graphite, no deviation from a linear relationship is evident. However, for the porous black carbons there appears to be a plateau value of $\tilde{K}_L$ at a molecular volume of approximately 110 Å³/molecule. This cutoff is clearly observed in the F400 data set (and to a lesser degree in the charcoal data). It is difficult to draw conclusions for the Darco GAC data set, however, since few compounds are above the molecular volume cutoff.

One possible explanation for the cutoff is that the larger molecules are unable to interact with a certain subset of sorption sites due to their size. Since the cutoff is not observed in the non-porous graphite data, it is likely that there are sites in these porous sorbents with which these large molecules cannot interact. The McGowan volume of a molecule can be related to a representative diameter by assuming a representative shape or cross-sectional area. Assuming that the solute molecules are spherical, the observed McGowan volume cutoff of 110 Å³/molecule yields a cutoff diameter, d_{eff} , of approximately 4.8 Å. This value was found to be within the range of median pore diameters (~ 4.0 to 6.0 Å)⁵⁰ for F400 GAC⁵⁰⁻⁵². While most of these molecules are not spherical, the magnitude of the representative diameters suggests that an inability to access certain sites is a plausible explanation for the observed cutoff.

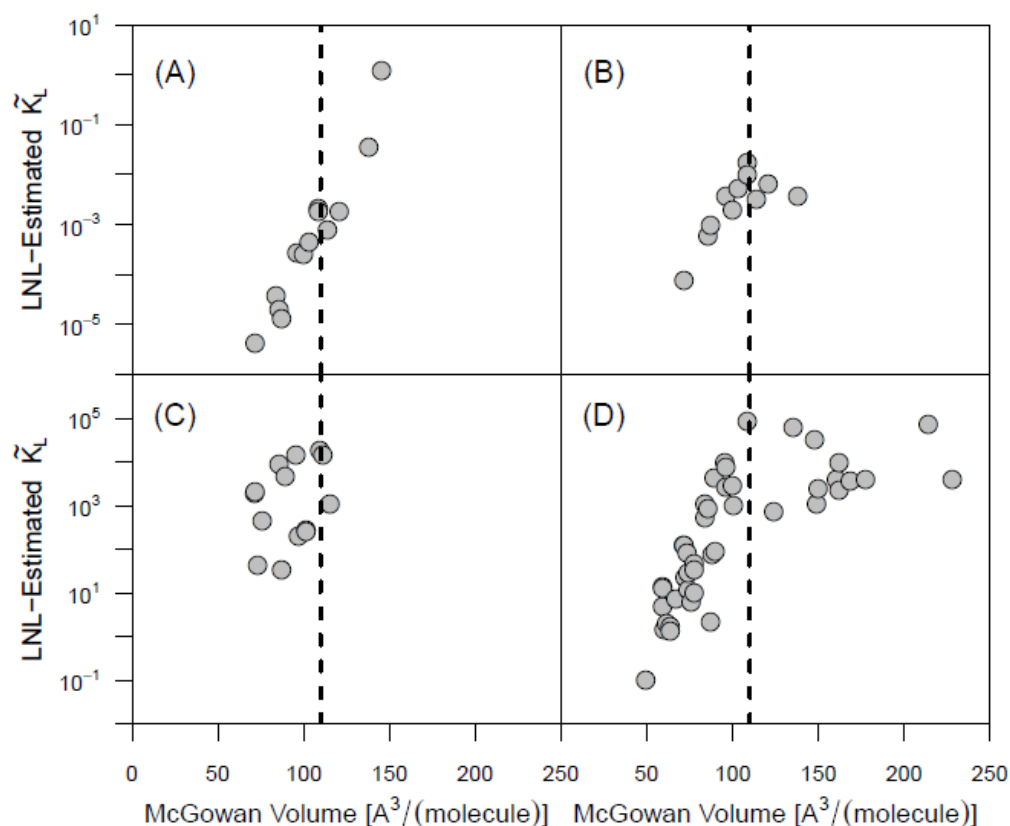


Figure 3-2. Plot of LNL-estimated median binding constants ($\tilde{K}_L$) vs. the Absolv-estimated solute McGowan volumes for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC. Dashed lines represent the observed volumetric cutoff of $110 \text{ \AA}^3/\text{molecule}$.

3.3.7 Separation of Large Molecular Volume Sorbates

The simplest modification to the Abraham model was to employ a separate molecular volume sorbent parameter (v) for small and large chemicals respectively.

$$\log(\tilde{K}_{L,i}) = e_c E_i + s_c S_i + a_c A_i + b_c B_i + v_{c,j} V_i + c_c \quad (3-3)$$

where the subscript j in the volume coefficient $v_{c,j}$ indicates whether the chemical is above or below the volumetric cutoff value. While the RMS error was reduced slightly, the fit was not significantly improved. There remained approximately 3

orders of magnitude spread in the predictions of $\log(\tilde{K}_L)$ for the solutes with McGowan volumes above the cutoff value.

A model with separate Abraham sorbent-water system parameters for molecules above and below the McGowan volume cutoff was also explored. While this model offered a reduction in the RMSE, it had twice as many degrees of freedom. Due to this significant increase in model parameters, the standard errors of the estimated sorbent-water system parameters increased significantly. Consequently, this model was rejected in favor of one that offered an improved fit without significantly increasing the number of parameters or their standard errors.

Previous work by Chiou et al.⁵³, Endo et al.⁵, and Razzaque & Grathwohl⁵⁴ have shown strong inverse correlations between octanol-water and organic carbon-water partition coefficients and aqueous solubility. A similar inverse correlation between the experimental aqueous solubility, S , and the LNL-estimated binding constants, $\tilde{K}_L$ was observed for compounds above the molecular volume cutoff in the F400 data set. Figure 3-3 illustrates this relationship.

Utilizing this relationship for molecules above the volumetric cutoff, the binding constant was expressed as a log-linear function of the aqueous solubility⁵⁵:

$$\log\left(\tilde{K}_{L,V \geq V_{cutoff}}\right) = m \log(S) + b \quad (3-4)$$

where m and b are the slope and intercept, and are functions of the sorbent only, S is the aqueous solubility (mM). It was observed that there was a single outlier, oxamyl, which significantly altered the regression slope and intercept estimates. To determine whether or not oxamyl should be excluded from the regression, a Cook's distance calculation was performed^{56,57} (Appendix B.4). The Cook's distance, D_C , for oxamyl was determined to be 3.12, well above the standard operating cutoff of $4/n$ (where n is

the number of parameters in the regression). While Cook's distance is not a test for statistical outliers, it does consider data points that will have disproportionately larger leverage in determining regression parameters, which can skew regression results. Consequently, oxamyl was not included in determining the slope and intercept for Eq. (3-4).

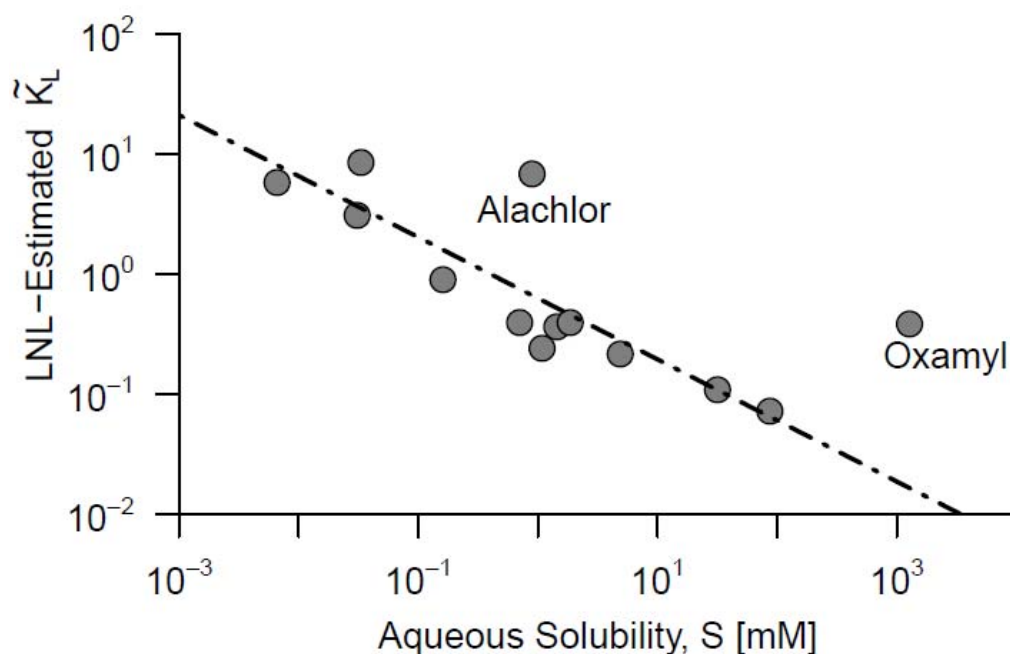


Figure 3-3. Plot of LNL-estimated median binding constants ($\tilde{K}_L$) vs. experimental aqueous solubility (S) for solutes above McGowan cutoff in F400 GAC. The dashed line represents the linear regression $\log(\tilde{K}_L) = -0.508 \log(S) - 0.202$. Oxamyl was not included in the regression due to its large Cook distance (see discussion in Appendix B.4).

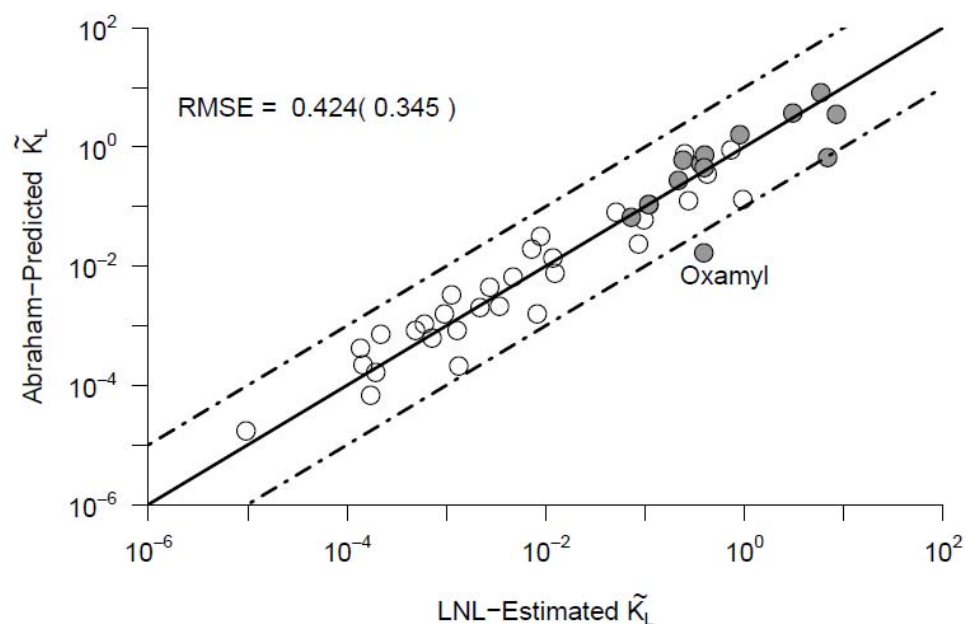


Figure 3-4. Plot of Abraham-predicted vs. LNL-estimated median binding constants ($\tilde{K}_L$) for F400 GAC. Solid lines represent 1:1 agreement; dashed lines represent ± 1 order of magnitude. Open circles represent sorbates below the volumetric cutoff (Eq. (3-2)), filled circles represent sorbates above the volumetric cutoff (Eq. (3-4)). Combined RMSE for all sorbates is 0.424. The RMS error excluding oxamyl is 0.345.

Figure 3-4 shows the improved model fit for all solutes in the F400 data set, utilizing a pp-LFER model below the molecular volume cutoff (Eq. (3-2)) and the solubility relationship (Eq. (3-4)) for those above. The bias in the lower and upper ranges of binding strengths no longer exists. For the F400 data set the RMSE was reduced significantly (from 0.709 to 0.424). The standard errors of the estimated e , s , a , v , and c parameters for sorbates below the volumetric cutoff were not significantly larger than those of the previous model. However, the standard error of b was significantly large. This is likely because the range of solute B parameters below the cutoff is very small, leading to large uncertainty in the estimated sorbent-water b

parameter. Consequently, the F400 sorbent parameters were refit without the hydrogen bond basicity (bB) term. The remaining parameters and the RMS error did not change significantly when the bB term was removed (see Table B-4 in the Appendix).

3.3.8 Residual Analysis

For the four sorption data sets, predicted sorbed concentrations (q_{pred}) were computed using the predicted $\log(\tilde{K}_L)$. Logarithmic residual values were then calculated for each experimental sorbed concentration (q_{obs}). Box plots of the residual values were constructed for each data set.

Figure 3-5 shows that with the exception of several sorbates in the F400 set ($n = 6$), that the interquartile ranges (IQR) of the residuals do not exceed 0.3 log-units. Further, it can be seen that nearly all of the median residual values are centered about 0, suggesting a lack of bias in the predicted isotherms. It is worth noting that some of the same chemicals that exhibit large residual ranges also exhibit median residual values that are not centered at zero. This suggests that this deviation might be an experimental artifact, rather than a failure of the model. Two of these chemicals (oxamyl and simazine) are the same chemicals discussed in Chapter 2 (section 2.4.2) for which accurate analytical methods were unavailable. This is likely the source of the bias in the isotherm predictions for these compounds. The RMS errors for the predicted isotherms were 0.0820, 0.1809, 0.183, and 0.220 for graphite, charcoal, Darco GAC, and F400 GAC, respectively.

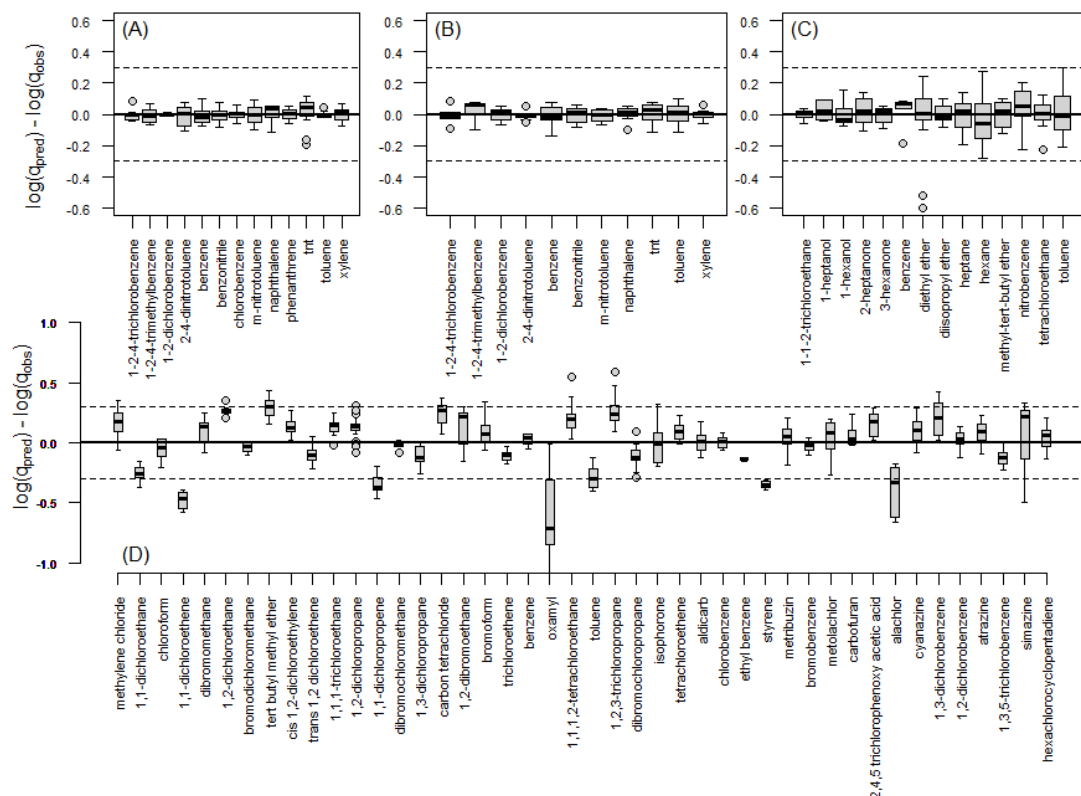


Figure 3-5. Box plots of the Abraham-predicted isotherm residuals ($\log(q_{pred}) - \log(q_{obs})$) for (A) graphite, (B) charcoal, (C) Darco GAC, and (D) F400 GAC, ordered by increasing median Langmuir binding constant. The solid line represents 1:1 agreement between q_{pred} and q_{obs} sorbed concentrations, dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations. IQR contains 50% of the data, whiskers represent ± 1.5 IQR, and filled points represent outliers (> 1.5 IQR).

For Darco GAC, predicted isotherms for the 11 solutes can be compared directly to predicted isotherms obtained by Shih & Gschwend³⁴ using Abraham pp-LFER models for the Freundlich isotherm parameters, K_f and v . A comparison of the two models is presented in Appendix B.6 together with a box plot of the isotherm residuals (Figure B-1). The RMS errors for the two models were similar (0.183 and

0.207 for the ss-LNL median binding constant pp-LFER (Eq. (3-2)) and the Freundlich pp-LFER models of Shih & Gschwend, respectively). While the RMS errors are similar, the number of parameters required to make a new isotherm prediction are not. Whereas the pp-LFER for Darco GAC presented in this work has 4 parameters, the combined pp-LFERs for the Freundlich K_f and ν have 12 (8 for K_f , 4 for ν), which approximately corresponds to 1 parameter per solute. Consequently, the prediction of the two Freundlich isotherm parameters requires four times the number of parameters to obtain comparable fits to those obtained with the ss-LNL model and the median binding constants predicted using Eq. (3-2).

3.4 Implications

It has been shown that the median binding constants, $\tilde{K}_L$, for the log-normal Langmuir model can be predicted utilizing modified versions of the Abraham poly-parameter linear free energy relationships. For the case of F400 GAC, there are sites that are inaccessible to compounds with a McGowan volume greater than the observed cutoff, $V = 110 \text{ \AA}^3 / \text{molecule}$ ($d_{eff} = 4.8 \text{ \AA}$). This corresponds approximately to the median pore diameter of F400 GAC ($\sim 4\text{-}6 \text{ \AA}$). This cutoff, along with poor correlations to the other Abraham solute parameters, suggests that these large molecules do not interact with the sorbent through specific interactions included in the pp-LFER model (Eq. (3-2)). Rather, for these large compounds, their solubility in water (Eq. (3-4)) better describes their partitioning between the bulk water phase and the GAC matrix.

It is expected that for neutral organic compounds (e.g., pesticides, PAHs, chlorinated compounds) the median Langmuir binding constant, $\tilde{K}_L$, can be predicted accurately using the pp-LFER and solubility relationships described previously.

Further, it is expected that these models can be easily adapted for additional sorbents that exhibit similar non-linear sorption behavior (e.g., biochar, carbon nanotubes (CNTs)). For sorbents for which the sorbent-specific parameters (q_{max} and σ_{κ}) are known, sorbent-water Abraham system parameters ($e_c \dots c_c$) can be estimated using a small number of sorbates (approximately 10), given sufficient range in sorbate Abraham parameters ($E_i \dots V_i$).

Finally, it has been shown that the resulting predicted isotherms reproduce the experimental data very well with little increase in RMS error. The combination of the log-normal Langmuir and pp-LFER/solubility models provide predictive capabilities over a wide range of solutes and black carbon sorbents utilizing readily available physical chemical parameters.

Chapter 4

PREDICTING SOLVENT-WATER PARTITIONING OF CHARGED ORGANIC SPECIES USING QUANTUM-CHEMICALLY ESTIMATED ABRAHAM PP-LFER SOLUTE PARAMETERS

Methods for obtaining accurate predictions of solvent-water partitioning for neutral organic chemicals (e.g., K_{ow}) are well-established. However, methods that provide comparable accuracy are not available for predicting the solvent-water partitioning of ionic species. The purpose of this chapter is to present a method for predicting solvent-water partitioning of charged species using quantum-chemically estimated Abraham (QCAP) solute parameters. For a suite of carboxylic acid anions, solvent-water partition coefficients for 4 solvent-water systems: acetonitrile-, acetone-, methanol-, and dimethylsulfoxide-water (computed from experimental ionization constants in the solvents and water) are predicted with root mean square (RMS) errors of 0.475, 0.512, 0.460, and 0.393, respectively ($n = 44, 48, 47$, and 41). For a larger set of substituted quaternary amine cations ($n = 217$), experimentally determined octanol-water partition coefficients were predicted with an RMS error of 1.16.

Predictions made using QCAPs are shown to provide improved accuracy in predicting solvent-water partition coefficients, compared to predictions of solvent-water partition coefficients, $\log(K_{solvent-water})$ made using existing Absolv-estimated Abraham solute descriptors derived from the neutral species. For partitioning of anionic solutes in the 4 organic solvent-water systems the overall $\log(K_{solvent-water})$ RMS errors were 0.740 and 0.462 for the Absolv and QCAP

methods, respectively. For cations partitioning into octanol the $\log(K_{\text{solvent-water}})$ overall RMS errors were 0.997 and 1.16, respectively.

The QCAP method also provides significantly improved accuracy compared to directly-calculated ab initio quantum chemical partition coefficients at comparable levels of theory (M062X/6-31++G**) for both anions partitioning into the 4 organic solvents (RMSE = 0.462 vs. 2.48 for QCAP-predicted vs. direct QC computed, respectively) and cations partitioning into octanol (RMSE = 1.16 vs. 2.82 for QCAP-predicted vs. direct QC computed, respectively).

4.1 Introduction

Methods for obtaining accurate predictions of solvent-water partition coefficients for neutral species (e.g., K_{ow}) are well-established^{1,4,8,53,55,58,59}. Several approaches exist that can be separated into two groups: (1) direct a priori quantum-chemical (QC) calculation of the solvent-water partition coefficient⁶⁰⁻⁶³ (e.g., using the programs COSMO-SAC, COSMO-RS, SMD) and (2) poly-parameter linear free energy relationships (pp-LFERs)^{3,6,12-14} or fragment-based⁸ models that rely on physical-chemical descriptors and group additive contributions.

For neutral species, the accuracies of these methods are comparable. Previous results for solvent-water QC predictions of octanol-water partition coefficients for a set of 103 neutral organic solutes have shown excellent agreement with experimental partition coefficients (RMSE = 0.566)¹⁵ using the SMD⁶³ solvation model in Gaussian 09⁶⁴. Similarly, a larger octanol-water partitioning data set (n = 992) was tested using the COSMO-SAC 2007/2010 (conductor screening model – segment activity coefficient) model. The resulting accuracies were comparable for both the 2007 and 2010 models (RMSE = 0.69 and 0.72, respectively)⁶⁵.

Similar results were obtained from fragment/pp-LFER models for octanol-water partition coefficients. Abraham pp-LFER models for predicting octanol-water partition coefficients have been shown to produce similar RMS errors using the same data set⁶⁵ (RMSE = 0.560, n = 992)⁶⁶. The EPI Suite model, KOWWIN, predicts octanol-water partition coefficients for a large validation data set (n = 10,946) with an RMS error of 0.476⁸.

Similar predictive accuracies, however, have not been observed for either group of methods for charged species. Direct quantum-chemical computations of solvent-water partition coefficients for both anions and cations¹⁵ have been shown to result in significantly larger errors. For SMD⁶³ an overall RMS error of 4.35 was observed for a set of organic anions and cations (N = 88) in acetonitrile-, methanol-, and dimethylsulfoxide-water systems (Table C-1a in the Appendix). Even within the same QC method, errors have been shown to range over several orders of magnitude between different solvent-water systems (e.g., acetone-water vs. methanol-water) with significant evidence of a solvent-based prediction bias (Table C-1a,b in the Appendix).

Attempts have been made to extend pp-LFER methods developed for neutral species to predict the solvent-water partitioning of charged species. Abraham et al.^{10,11} have developed a method to obtain new solute descriptors based on linear regressions of the existing experimental solute descriptors for the neutral species. However, these relationships are not descriptor-specific (e.g., the polarizability (*S*) for the ion being a function only of the neutral species polarizability). While these regressions to neutral solute Abraham parameters offer good agreement to the experimental ionic partition coefficients, the equations and the fitted coefficients are chemical-class specific^{10,11}, and are not universally applicable to wider ranges of charged solutes.

Alternatively, Zissimos et al.⁶⁷ have shown that for neutral species, Abraham solute descriptors can be obtained through individual regressions to the sigma profile moments computed from the COSMOTerm^{60,61} program, with a separate regression for each Abraham solute descriptor (E , S , A , B , and V). Cho et al.¹² have extended this method to include charged species. However, predictions of the Abraham solute descriptors for S , A , and two new sorbate descriptors, J^+ , and J^- (charged descriptors for cations and anions that describe the ion-specific hydrogen bond donating and accepting capabilities, respectively)⁶⁸ were shown to have significantly larger errors than those of the neutral species, despite a large number of fitting parameters ($N = 5, 6, 8, 10$, for S , A , J^+ , and J^- respectively)¹². In total, 21 (4 structural/bond, and 17 quantum chemical) descriptors with 50 fitted equation constants are required to predict the charged species Abraham solute descriptors (E , S , A , B , V , J^+ , and J^-). Predicted solvent-water partition coefficients vary significantly by solvent and range from $RMSE = 0.46$ to 1.01 ¹².

This chapter will compare several methods for predicting solvent-water partition coefficients of charged species: (A) direct quantum chemical computation; (B) Abraham pp-LFER predictions using existing Absolv-estimated Abraham solute descriptors; and (C) a new approach using quantum-chemically estimated solute Abraham descriptors for charged species.

4.2 Modeling & Experimental Data

4.2.1 Abraham Poly-parameter Linear Free Energy Relationship (pp-LFER)

A popular model for predicting partitioning of neutral organic compounds between organic phases is the Abraham poly-parameter linear free energy

relationship^{1,2,4,55,68,69} (pp-LFER). The model relies directly on the physical chemical properties of the solute and its individual interactions with the solvent phases (e.g., dispersion forces, hydrogen bond acidity, and cavitation energy). For neutral organics, the Abraham pp-LFER can be expressed as:

$$\log(K_{i,j}) = e_j E_i + s_j S_i + a_j A_i + b_j B_i + v_j V_i + c_j \quad (4-1)$$

where the upper case letters represent the solute (*i*) contributions to the binding energy; *E*, molar refractivity, *S*, polarizability, *A*, hydrogen bond donating; and *B*, hydrogen bond accepting capacities, and *V*, the energy required to form a cavity. The lower case letters represent the complementary solvent (*j*) constants relative to water; excess molar refractivity, polarizability, hydrogen bond accepting and donating, and cavitation energy terms, respectively. The *c_c* term accounts for any non-specific binding interactions not predicted by the Abraham solute parameter product terms (*e_jE_i*, *s_jS_i*, *a_jA_i*, *b_jB_i*, and *v_jV_i*) and varies with the units of *K_{i,j}*.

Previous work by Zhao & Abraham^{14,68} have attempted to predict the octanol-water partition coefficients of a set of quaternary amine cations by estimating the Abraham solute descriptors for a comparable neutral species. They accomplish this by neutralizing the molecule, replacing the charged nitrogen atom, [*N*⁺] with a neutral carbon atom, C. The solute descriptors are then estimated using the method of Platts⁴ with the experimentally derived octanol-water solvent-system descriptors⁷⁰. While a clear trend was demonstrated within homologous series and structurally similar solutes, the RMS error of the predictions was large, RMSE = 2.66.

Alternatively, Abraham & Zhao⁶⁸ propose a modification to the Abraham pp-LFER in which specific ion descriptors for solutes (*J*⁺ and *J*[−]) are used to describe the additional charge-based hydrogen bond acidity/basicity of the ionic species with

respect to the neutral solute species. In a recent publication, Abraham & Acree¹¹ proposed a similar modification to the Abraham pp-LFER and applied it to charged organic species. The modified pp-LFER for charged organic solute partitioning is:

$$\log(K_{i,j}) = e_j E_i^* + s_j S_i^* + a_j A_i^* + b_j B_i^* + v_j V_i^* + j_j^+ J_i^+ + j_j^- J_i^- + c_j \quad (4-2)$$

where the asterisks denote the Abraham solute parameters for the charged species, which are computed from those of the neutral species.

The new product terms, $j_j^+ J_i^+$ and $j_j^- J_i^-$, are included to account for ionic solute-solvent charge interactions for positive and negatively charged ionic solutes, respectively. Abraham et al. propose that the solute parameters for the charged compounds (including the new J^+ and J^- terms) can be obtained directly from the Abraham solute parameters for the neutral species¹¹. However, the relevant parameters and fitted equation coefficients for relating the neutral and charged Abraham solute parameters vary significantly for different types of charged species¹¹. Consequently, these relationships are not universally applicable to new organic species. Further, the number of fitting parameters is large, relative to the number of data points used to obtain the regression parameters. For 78 carboxylate ions and 26 amines, there are 20 and 13 fitted equation coefficients required, respectively, to obtain the charged solute parameters. While the models perform well within their chemical class, they cannot be applied universally. Therefore, a method for obtaining the charged Abraham solute parameters which is independent of the chemical class and has universal applicability is required.

4.2.2 Quantum-Chemically Estimated Abraham Solute Parameters (QCAP)

The method presented in this chapter can be used to generate new Abraham solute parameters for charged species independent of their neutral counterparts. The standard Abraham pp-LFER equation (Eq. (4-1)) is a linear combination of the paired solvent-solute interaction terms together with a constant which carries units and any non-specific binding. As discussed previously, the solvent parameters are solute independent and available⁴⁶ for common solvents and the solute parameters are independent of the solvent system. Therefore it is possible to estimate the solute parameters if the solvent parameters are known and vice versa.

For example, suppose the 5 unknown Abraham parameters ($E_i \dots V_i$) are required for a new solute. Since there are 5 unknowns, at least 5 solvent-water partition coefficients are required to estimate the five solute parameters. Using more than 5 solvent-water partition coefficients adds redundancy to the estimate. With 5 unknowns, the more than 5 equations of the form of Eq. (4-1) can be used to estimate the 5 unknown Abraham parameters ($E_i \dots V_i$) by minimizing the sum of the squares of the residuals using multiple linear regression:

$$\hat{E} = \sum_{i=1}^n (\log K_{i,j}^{expt} - e_j E_i - s_j S_i - a_j A_i - b_j B_i - v_j V_i - c_j)^2 \quad (4-3)$$

where n is the number of solvent-water system partition coefficients, $\log K_{i,j}^{expt}$ is the experimental solvent-water partition coefficient (L water / kg solvent) for solute (i) in the solvent (j)-water system. $\hat{E}$ is the error function (L water / kg solvent)². This method is described in detail in Platts et al.⁴

Similarly, if it is necessary to estimate the 6 solvent parameters ($e_j \dots c_j$) for a new solvent, then $N \geq 6$ experimental solvent-water partition coefficients can be measured for N solutes in the solvent-water system. The experimental partition

coefficients and the known Abraham solute parameters can then be used to estimate the solvent parameters ($e_j \dots c_j$) using a similar multiple linear regression as discussed above.

Previous work by Liang & Di Toro⁶⁶ has shown that fitting all 5 Abraham solute parameters simultaneously using multiple linear regression results in large standard errors and significant cross-correlation of estimated parameters, particularly for E and S . To reduce the number of estimated parameters, the McGowan volume (V) and the excess molar refractivity (E) were estimated independently using quantum chemical methods⁶⁶. Consequently, only 3 solute parameters, S , A , and B need to be estimated for each solute via multiple regression. A full description of the method including the R code for the multiple linear regressions is included in Appendix C.2.

In order to apply this method to estimating Abraham solute parameters, solvent-water partition coefficients are required for at least 3 solvent-water systems. Liang & Di Toro⁶⁶ propose to use quantum chemically computed solvent-water partition coefficients instead of experimental data. The key idea in their method is to use a large number of solvents and to compute the solvent-water partition coefficients for all the solvents. Although the individual quantum chemical computations of solvent-water partition coefficients have large errors, the use of multiple linear regressions reduces the error in the estimated Abraham solute parameters⁶⁶.

For neutral solutes, values for solvent-water system parameters ($e_j \dots c_j$) (Eq. (4-2)) are well-established⁷¹. Consequently, these are used for new solutes for which Abraham solute descriptors are to be estimated⁶⁶. This is not the case, however, for ionic solutes. For new ionic solutes, methods for obtaining the McGowan volume (V) and the excess molar refractivity (E) developed for neutral species⁶⁶ can be applied.

However, the remaining Abraham solute (S , A , B) as well as the solvent-water system parameters (e , s , a , b , v , c) are unknown and must be estimated. The method employed is iterative. First, the neutral solvent-water parameters ($e_j \dots c_j$) are used as initial guesses with QC computed $\log(K_{i,j})$ for a large number of solvent-water systems to estimate charged solute Abraham parameters (S , A , and B) using a multiple linear regression and minimizing the sum of squares of the errors (Eq. (4-3)). The estimated solute Abraham parameters are then used to estimate new charged solvent-water descriptors ($e_j \dots c_j$). This process is repeated until the sum of squares of the residuals (Eq. (4-3)) is minimized. The result is quantum-chemically estimated solute and solvent Abraham parameters for charged species.

The QC computational accuracy of solvent-water partition coefficients for charged compounds is very poor, with root mean square errors of approximately 2.3 – 7.1 log-units¹⁵ (see Table C-1a in Appendix C). However, much of this error is due to systematic solvent-specific biases in the QC computed partition coefficients. In any case, it is clear that using the maximum number of solvent-water partition coefficients is necessary. The SMD solvation model⁶³ can estimate solvent-water partitioning of charged solutes for 49 solvent-water systems for which literature values for neutral solvent-water Abraham parameters exist (Table C-2 in the Appendix)⁷¹. Therefore it was chosen for this application.

4.2.3 Quantum Chemical Computations

All quantum chemical computations were performed in Gaussian 09⁶⁴ using the SMD⁶³ solvation model. M062X was chosen as the density functional, as it has improved accuracies over other M06 Minnesota functionals in computing main group thermochemistry⁷². Since the solutes of interest are charged, long-range electrostatic

interactions were of particular concern. Consequently, the 6-31++G** basis set was chosen which includes diffuse functions. This basis set is similar in size (and offers comparable performance for main group thermochemistry) to the correlation-consistent double zeta (aug-cc-pVDZ) Dunning basis set, but with slightly reduced computational times⁷³. For the quaternary amine cations, larger errors were observed for halogenated species. Consequently, a mixed basis set was constructed for these species, with the 6-31++G** basis set for C, N, H, O, and S and a larger basis set (aug-cc-pVTZ) for the halogen atoms (Cl, F).

The 3-D structures for the ions were constructed in ArgusLab⁷⁴ and were pre-optimized in the gas phase using a built-in semi-empirical quantum mechanical computation (AM1)⁷⁵. These structures were then optimized in the gas phase using the appropriate basis set (6-31++G** or 6-31++G**/aug-cc-pVTZ) in Gaussian 09⁶⁴. Frequency calculations were performed to ensure the structures were at an energy minimum and no imaginary frequencies were present. Finally, single point energy calculations were performed for the solutes in 50 solvents (including water), and 49 solvent-water partition coefficients were calculated for each solute using the SMD-computed solvation free energies as follows:

$$\Delta G_{\text{solvent-water}} = G_{\text{solvent}} - G_{\text{water}} \quad (4-4)$$

$$\log(K_{\text{solvent-water}}) = \frac{\Delta G_{\text{solvent-water}}}{-2.303RT} \quad (4-4a)$$

where G_{solvent} and G_{water} are the SMD-computed Gibbs free energies in the solvent and water phases, respectively. $\Delta G_{\text{solvent-water}}$ is the Gibbs free energy of transfer between the solvent and water phases, R is the universal gas constant (kcal/(mol °K)),

T is the temperature ($^{\circ}\text{K}$), and $\log(K_{\text{solvent-water}})$ is the base-10 logarithm of the dimensionless (mole-fraction) solvent-water partition coefficient.

4.2.4 Experimental Data

The experimental data used to develop the model was obtained from several sources^{11,14} in order to include anions and cations, as well as various organic solvents. For the anionic species, direct experimental measurements were not made for the partition coefficients, $\log(K_{\text{solvent-water}})$, for the four solvent-water systems (acetone-, acetonitrile-, dimethylsulfoxide-, and methanol-water) since the solvents are miscible in water. Instead, existing Abraham pp-LFER solvent-water system parameters were used, together with the solute descriptors for the neutral species to compute partition coefficients for the neutral species. These values were used together with experimental values of the pKa in solvent and water as well as experimental values for solvent-water partitioning of the proton to compute solvent-water partition coefficients for the ionic species¹¹.

For the quaternary amine cations, octanol-water partition coefficients are computed from the octanol-water partition coefficients of halide-salts of the cation (e.g., Cat-Cl, Cat-Br) and experimental partition coefficients of the bare halide ions (Cl^- , Br^-)¹⁴.

Figure 4-1 and Table 4-1 present summaries of the experimental solvent-water partitioning data used in developing and validating the QCAP Abraham method.

Table 4-1. Summary of anion and cation solvent-water partitioning data.

Solutes	Solvents	N	Notes	Ref
Carboxylic Acids	Acetone	44	Aliphatic, aromatic, chlorine-, and nitro-substituted	11
	Acetonitrile	48		
	Methanol	47		
	DMSO	41		
Quaternary Amines	Octanol	217	Aliphatic, aromatic, chlorine-, fluorine-, amino-, imidazolium-, and pyridinium- substituted	14

Figure 4-1 illustrates the range of the partition coefficients for the 5 solvent-water systems. There are significant variations in partition coefficients both within each system as well as between the systems. Individual measurements of partition coefficients span approximately 4 to 5 orders of magnitude for the anionic compounds and approximately 12 orders of magnitude for the cationic compounds. Additionally, the solutes vary considerably both in size and structural moieties. A complete table of the solutes and their experimental solvent-water partition coefficients can be found in Table C-3a-e in Appendix C.

4.3 Results & Discussion

Previous work has shown that within solvents, there can be systematic biases in QC computed vs. experimental solvent-water partition coefficients for both neutral and ionic species¹⁵. Solvent-specific linear corrections are used in quantum chemical predictions of acid-dissociation constants (pKa) and have been shown to correct for solvent-specific biases⁷⁶⁻⁷⁸.

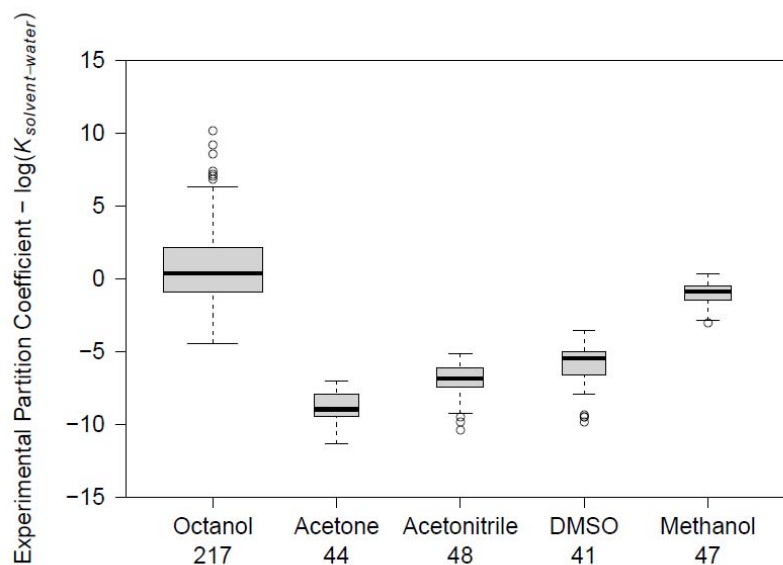


Figure 4-1. Box plot of experimental solvent-water partition coefficients for cationic (octanol) and anionic (acetone, acetonitrile, dimethylsulfoxide (DMSO), and methanol) species by solvent. The interquartile range (IQR) represents 50% of the data, whiskers represent 1.5IQR, and filled points represent outliers ($> 1.5\text{IQR}$). Box widths are proportional to the square root of the number of data points. The number of solutes within each solvent data set is shown below the solvent names.

A similar linear correction to the QCAP Abraham-predicted solvent-water partition coefficients can be expressed:

$$\log(K_{i,j}') = m_j \log(K_{i,j}) + b_j \quad (4-5)$$

where $\log(K_{i,j})$ is the solvent (j) - water partition coefficient for a solute (i) computed using the solute and solvent-water system parameters obtained with the QCAP method described previously. The parameters m_j and b_j are the solvent-specific slope and intercept corrections, and $\log(K_{i,j}')$ is the corrected solvent (j) - water partition coefficient for the solute (i). A linear correction (Eq. (4-5)) to the QCAP Abraham-predicted (Eq. (4-1)) $\log(K_{i,j})$ can be written as follows:

$$\log(K_{i,j}') = m_j(e_j E_i + s_j S_i + a_j A_i + b_j B_i + v_j V_i + c_j) + b_j \quad (4-6)$$

where the slope, m_j , multiplies the terms of the uncorrected Abraham pp-LFER, and the intercept term, b_j , is added to the slope-corrected terms. Since the slope and the lower case solvent-water system parameters are both solute-independent, the product terms ($m_j e_j \dots m_j v_j$) can be grouped together into new corrected solvent-water system parameters, with the intercept term, b_j , being absorbed into the constant term, $m_j c_j$, as follows:

$$\log(K_{i,j}') = e'_j E_i + s'_j S_i + a'_j A_i + b'_j B_i + v'_j V_i + c'_j \quad (4-7)$$

where the ' denotes the linearly corrected solvent-water system parameters (e.g., $e'_j = m_j e_j$, $c'_j = m_j c_j + b_j$). This is the equation that will be used to predict the solvent-water partition coefficients for ionic species. It is formally identical to Eq. (4-1).

The QC solvent-water partition coefficients were computed using SMD for the set of anions and cations in the 49 solvent-water systems. QCAP solute and solvent-water system parameters were then obtained via the method discussed above by minimizing the sum of squares of the residual errors (Eq. (4-3)). This produces the QCAP estimates for the solute parameters. The final step is to produce the corrected solvent-water system parameters (Eq. (4-7), lower case parameters). These were obtained by multiple linear regressions, using the QCAP estimates for the solute parameters for the ionic species and the experimental values for the respective solvent-water partition coefficients.

The results are compared below to (A) a priori QC calculations (at the same level of theory: SMD, M062X/6-311++G**), and (B) predictions made using the existing neutral species Absolv Abraham solute parameters (AAPs). Complete tables

of Absolv and QCAP solute descriptors and QCAP solvent descriptors can be found in Tables C-4a,b and Tables C-5a,b in the Appendix.

4.3.1 Carboxylate Anions

For the set of carboxylate anions, the results of the three methods (a priori QC, Absolv estimates, and QCAP) for estimating the solvent-water partition coefficients are summarized in Figure 4-2. It should be noted that for ionizable species, Absolv returns estimates for the neutral species, and will not compute different Abraham descriptors for the anions. The RMS errors for the three methods are summarized in Table 4-2 below.

The direct quantum chemical computations (Figure 4-2A) have large systematic biases which are different for the different solvent-water systems. Additionally, large errors can be seen within single solvent-water systems (~ 4 orders of magnitude for acetonitrile-water). While the predictions using Absolv-estimated solute descriptors show significant improvement over direct QC predictions (Figure 4-2B), there is still significant variability in the accuracies between solvsolvent-water systems. The errors in prediction for DMSO-water using Absolv solute descriptors are actually larger than those computed quantum-chemically. However, there is no apparent systematic solvent-based bias for the partition coefficients predicted using QCAP descriptors (Figure 4-2C). It is interesting to note that the errors increase at lower solvent-water partition coefficients for both the direct QC predictions and Absolv-estimated predictions (Figure 4-2A,B). This suggests that the average errors increase for increasingly hydrophilic anions for both the direct QC and Absolv-estimated AP methods.

The estimated solvent-water system parameters (using QCAP descriptors) and their associated standard errors as well as the individual solvent RMS errors are summarized in Table 4-3 below. Estimated solvent-water system parameters using Absolv-estimated Abraham parameters (and associated standard errors) can be found in Tables C-6a-c in the Appendix.

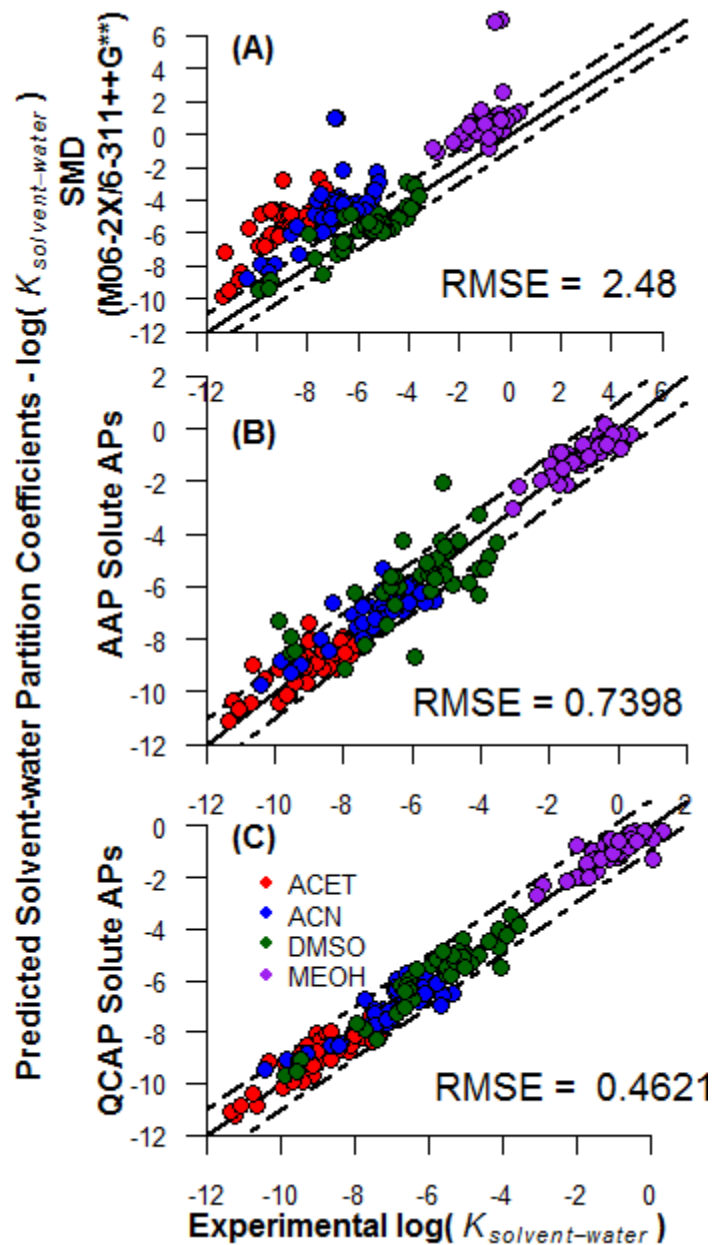


Figure 4-2. (A) Direct QC computed (SMD, M062X/6-31++G**), (B) Absolv-predicted, (C) QCAP-predicted vs. experimental solvent-water partition coefficients for carboxylic acid anions for acetone (ACET), acetonitrile (ACN), DMSO, and methanol (MEOH). Solid lines represent 1:1 agreement, dashed lines represent $\pm$ an order of magnitude errors. RMS errors are 2.48, 0.740, and 0.462, respectively (N = 180).

Table 4-2. Summary of RMS errors for direct QC computed, Absolv Abraham solute parameter (AAP), and quantum-chemically estimated Abraham solute parameter (QCAP) methods.

Species	Solvent	Method	N	RMSE
Anion	Acetone	QC	44	3.43
		AAP	44	0.636
		QCAP	44	0.475
	Acetonitrile	QC	48	3.71
		AAP	48	0.59
		QCAP	48	0.512
	DMSO	QC	47	0.698
		AAP	47	1.11
		QCAP	47	0.460
	Methanol	QC	41	2.14
		AAP	41	0.389
		QCAP	41	0.393
	Total	QC	180	2.48
		AAP	180	0.74
		QCAP	180	0.462
Cation	Octanol	QC	217	2.82
		AAP	217	0.997
		QCAP	217	1.16

4.3.2 Cations

For the set of quaternary amine cations, corrected solvent-water parameters ($e'_j \dots c'_j$) were estimated for the octanol-water system, using the solute QCAPs and experimental solvent-water partition coefficients. The QCAP-predicted partition coefficients (Eq. (4-7)) were then compared to predictions made using existing Absolv-estimated Abraham (AAP) solute descriptors and the direct QC computed partition coefficients (SMD, M062X/6-311++G**). It should be noted that for

permanently charged species, Absolv returns estimates for the charged species directly. The results are summarized in Figure 4-3.

There are two large groupings of solutes for the direct QC computations that show a significant systematic bias. There is also significant scatter among the two distinct groupings (as much as 5 orders of magnitude). The RMS errors are similar between the direct QC computations for both anions and cations (RMS errors of 2.48 and 2.82, respectively). However, the RMS error for both the Absolv-predicted and QCAP-predicted partition coefficients for the cations are larger (RMSE = 0.997 and 1.16, respectively) than those for the anions (RMSE= 0.740 and 0.426). This is likely due to differences in size and structural complexity of the solutes between the two data sets. The anions are generally small (fewer than 10 heavy atoms) and either aliphatic or aromatic carboxylic acids with a small set of structural moieties (predominantly chlorine and nitrate functional groups). The quaternary amines however have flexible aliphatic chains of up to 16 carbons and more structurally diverse functions groups (e.g., $-\text{SO}_2\text{-CF}_3$) as well as nitro-substituted aromatic ring structures.

The RMS error for the QCAP-predicted partition coefficients for the quaternary amine cations is significantly smaller than that of the direct a priori QC calculations (compare Figure 4-3A and 4-3C) and is nearly identical to that obtained using Absolv-estimated solute parameters (Figure 4-3B). The estimated octanol-water solvent-water system parameters (using QCAP solute descriptors) as well as the individual solvent RMS errors are summarized in Table 4-3. Estimated solvent-water system parameters using Absolv-estimated solute parameters (and associated standard errors) can be found in Tables C-6a-c in the Appendix.

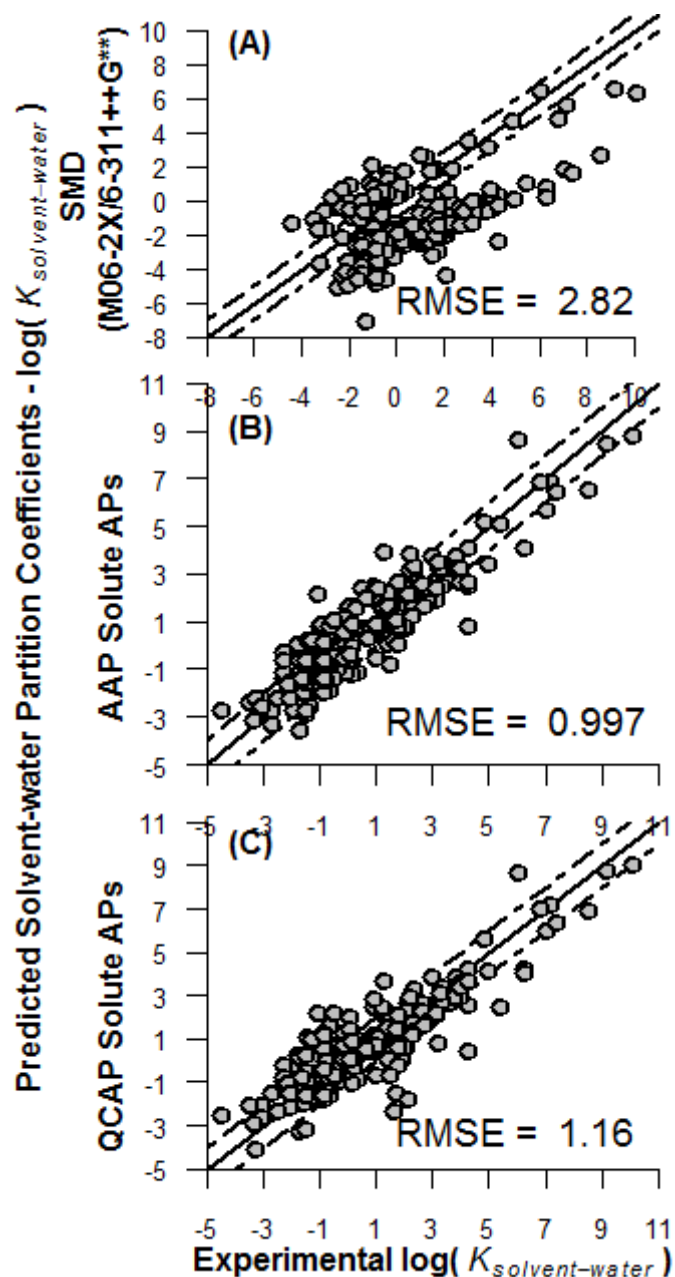


Figure 4-3. (A) Direct QC-computed (SMD, M062X/6-31++G**), (B) Absolv-predicted, (C) QCAP-predicted vs. experimental octanol-water partition coefficients for quaternary amine cations. (Solid lines represent 1:1 agreement, dashed lines represent $\pm$ an order of magnitude errors. RMS errors are 2.82, 0.997, and 1.16, respectively (N = 217).

4.3.3 Fitted Abraham Solvent – Water Parameters

For the carboxylate anion data set, there were significant variations in the statistical significance and the magnitude of the standard errors for the Abraham solvent-water system parameters. To reduce the standard errors of the estimated parameters all 6 solvent-water parameters were initially included in the multiple linear regressions (Eq. (4-7)). The parameters with the largest p-values were then selected one at a time, removed from the regression equation, and the remaining parameters were refit. This was done using the “step()” linear regression function in R⁷⁹. It was noted that despite the reduction in degrees of freedom, the RMS errors of the reduced Abraham pp-LFER predicted solvent-water partition coefficients were not significantly different from those of the full Abraham pp-LFER model (Eq. (4-7)). For the cations, there were no parameters with $p > 0.05$. Consequently, the full Abraham pp-LFER model was retained. The estimated solvent-water parameters for both data sets as well as the coefficients of variation (%) and the associated RMS errors are summarized in table 4-3. The regression results which include all Abraham pp-LFER parameters for both the AAP and QCAP methods can be found in Tables C-6a-c in Appendix C. For the anion and cation solvent-water systems, the coefficients of variation (CV) of the estimated parameters are generally small, with a median CV of 24.0%. However, it is clear that the coefficients of variation for the estimated intercepts (c) are large. This is due in part to the fact that the intercepts are all small, indicating that the molecular interaction terms (eE , sS , ..., vV) describe nearly all of the partitioning behavior for the solute species.

Table 4-3. Summary of corrected Abraham solvent-water system parameters (using QCAP solute descriptors), their respective coefficients of variation (%), number of experimental data points (N), and root mean square (RMS) errors for the carboxylic acid anions (N = 180) and quaternary amine cations (N = 217) experimental partitioning data.

Species	Solvent	e	s	a	b	v	c	N	RMSE
Anions	Acetone	1.04	[-]	1718.1	-2.56	1.98	0.812	44	0.475
		65.38%	[-]	17.47%	12.69%	36.58%	178.7%		
	Acetonitrile	2.94	-1.70	-676.4	-1.20	[-]	0.257	48	0.512
		11.05%	16.69%	53.82%	19.44%	[-]	559.1%		
	DMSO	2.38	[-]	567.4	-3.41	2.02	.432	47	0.460
		27.00%	[-]	51.34%	9.14%	36.44%	318.64%		
	Methanol	2.37	[-]	1202.3	[-]	[-]	-0.133	41	0.393
		10.31%	[-]	16.94%	[-]	[-]	514.43%		
Cations	Octanol	0.6793	-2.21	-4498.2	0.201	5.39	-4.10	217	1.16
		19.21%	20.89%	12.65%	90.09%	7.97%	32.55%		

It should also be noted that for the anions, particularly in the acetonitrile- and DMSO-water systems, the CV of the hydrogen bond accepting term, a , conjugate to the solute hydrogen bond donating term (A), is quite large. This is not surprising, since the anions are all acids, which exhibit little to no ability to donate a hydrogen bond (A). Similarly, the largest CV for the quaternary amine cations is for the hydrogen bond donating term, b , conjugate to the solute hydrogen bond accepting term, B . This is also expected, since they are positively charged species with a large partial positive charge at the nitrogen site (i.e., excellent hydrogen bond donors). For the remaining terms, however, the coefficients of variation are similar both between the solvent-water systems and for both cations and anions. This indicates that there is sufficient data to estimate all of the parameters used in the regressions. Further, there do not appear to be any solvent or species-specific (anion vs. cation) biases in the coefficients of variation for the estimated parameters. The accuracies of the estimated Abraham parameters appear to be independent of both the solvent-water system and the charge of the solutes.

4.3.4 Residual Analysis

For each solvent-water system, residuals were computed for each experimental solvent-water partition coefficient ($\log(K_{i,j}^{exp})$). Box plots of the residuals were constructed for each of the solvents. Figure 4-4 shows that there is no significant bias by solvent as the median $\log(K_{i,j}^{exp})$ values increase, indicating that the model works equally well for both hydrophobic and hydrophilic solvents. Further, it can be seen that for acetone, DMSO, acetonitrile, and methanol, approximately 50% of the residuals are less than 0.3 log-units (a factor of 2 in linear space) and > 95% of the

residuals are less than 1.0 log-unit (an order of magnitude). This indicates excellent agreement between the QCAP-predicted and the experimental solvent-water partition

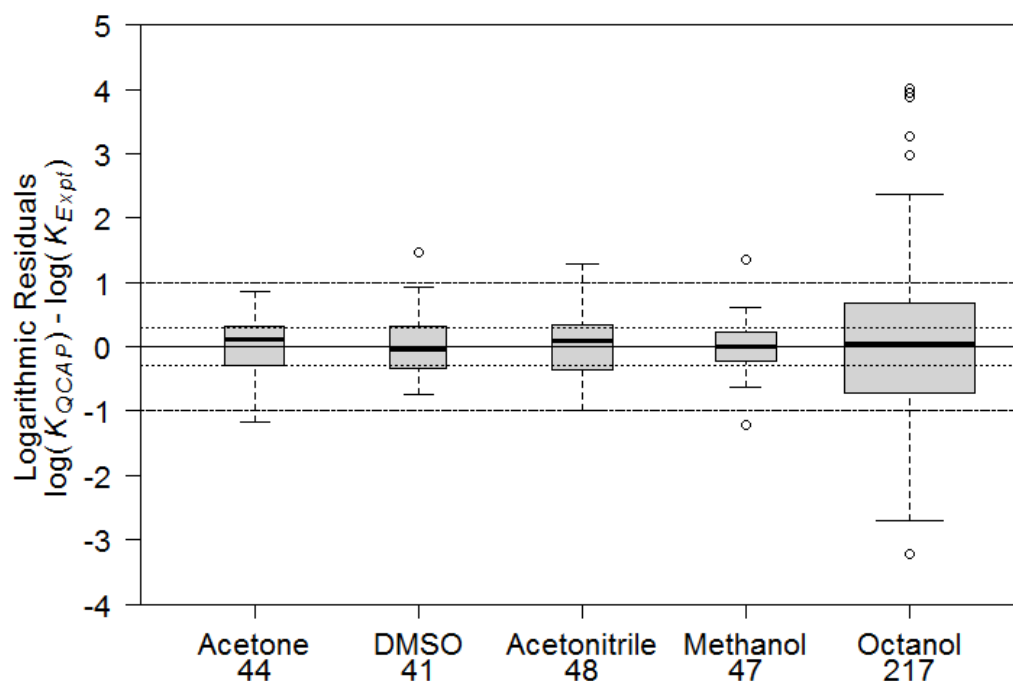


Figure 4-4. Logarithmic residual box plots ordered by increasing solvent-water median residuals. Boxes correspond to the interquartile range (IQR), where $25\% < \text{IQR} < 75\%$. Whiskers correspond to $\pm 1.5\text{IQR}$. Points correspond to outliers ($> 1.5\text{IQR}$). Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10). Box width is proportional to the square root of the number of data points.

coefficients. While the range is significantly larger for the quaternary amine cation residuals, $> 50\%$ of the residuals are within 1.0 log-units. While these residuals are

larger than those for the anionic solutes, it is a significant improvement over the direct quantum chemical computations (see Figures C-1a, b in Appendix C).

4.4 Implications

It has been shown that for charged species, Abraham-predicted solvent-water partition coefficients obtained using quantum-chemically estimated solute parameters (QCAP) offer significant improvement over direct, a priori quantum chemical computations for all solvents. Further, it has been shown that using QCAP solute parameters for charged species results in comparable predictions for the octanol-water partition coefficients of the quaternary amine cations as well as improved predictions for the acetone-, acetonitrile-, and methanol-water partition coefficients for the carboxylic acid anions. For both the anion and cation data sets, the reduction in RMS error and the removal of solvent-based bias in the predictions were significant (from RMSE = 2.48 and 2.82 to RMSE=0.462 and 1.16, respectively). For anionic partitioning, the biases for the different solvents were eliminated, and the overall RMS error was reduced to within a factor of 3 (RMSE = 0.462). This is comparable to the RMS errors for neutral species obtained from both COSMO-SAC⁶⁵ (RMSE = 0.690) as well as Absolv⁶⁶ (RMSE = 0.560). While the RMS error for the octanol-water partitioning of the cationic quaternary amines was larger (RMSE = 1.16), this is a significant improvement over existing quantum chemical methods (RMSE = 2.82) and is comparable to the predictions made using Absolv-estimated solute descriptors (RMSE = 0.997). Further, the QCAP solute descriptor method does not rely on fragment values or chemical class-specific corrections, and has only 3 estimated parameters per solute (*S*, *A*, and *B*), compared to 6 for the alternative Abraham pp-LFER methods for charged species¹⁰⁻¹². Finally, QCAP solute descriptors for the ions

are obtained directly from molecular structure and quantum chemical computations and do not rely on structural/bond parameters¹² or physical chemical descriptors of the neutral species¹¹.

It is expected that for additional solvent-water systems as well as additional ionizable solutes, the QCAP method can predict solvent-water partition coefficients to a similar level of accuracy for both anionic and cationic species. For new solvents, solvent-water system parameters can be obtained with existing Abraham charged solute parameters. Similarly, for new solutes in existing solvents, Abraham parameters can be obtained by computing QC solvent-water partition coefficients in the 49 solvent-water systems and estimating the charged QCAP solute parameters.

Chapter 5

PREDICTING NON-LINEAR PARTITIONING OF CHARGED ORGANIC SPECIES TO NATURAL ORGANIC MATTER USING QUANTUM-CHEMICALLY ESTIMATED ABRAHAM PP-LFER SOLUTE PARAMETERS

Procedures for accurately predicting linear partition coefficients for neutral chemicals onto various sorbents (e.g., black carbon, natural organic matter, clay) are reliable and well-established. In Chapter 2 an accurate method for modeling non-linear adsorption of neutral species onto various black carbons with a single, sorbate-specific, binding parameter is presented. However, similar methods for the prediction of non-linear sorption of charged species have not been successful. The purpose of this chapter is to present a procedure for predicting non-linear isotherm parameters of charged species, specifically the median Langmuir binding constants, $\tilde{K}_L$, obtained utilizing the log-normal Langmuir isotherm discussed in Chapters 2 and 3. An Abraham poly-parameter linear free energy relationship (pp-LFER) is able to predict median Langmuir binding constants for a set of primary, secondary, tertiary, and quaternary amines sorbed onto soil organic carbon using quantum-chemically estimated Abraham (QCAP) solute descriptors for the charged species. Predictions made using the QCAP solute parameters were then compared to predictions made using Absolv-estimated Abraham solute parameters (AAPs). For the set of 60 amine cations, QCAP predictions of organic carbon – water partition coefficients had an overall RMS error of 0.526, while predictions made using the existing AAPs for the neutral species had an overall RMS error of 0.569. Clear biases present for secondary

and quaternary amine species predicted using Absolv-estimated Abraham solute descriptors were reduced significantly in the predictions made using the quantum-chemically estimated solute descriptors.

5.1 Introduction

Methods for predicting linear partitioning of neutral solutes between organic carbon and water are well-established^{1,2,5-7,38-42}. The most popular method for predicting organic carbon – water linear partition coefficients (K_{oc}) are Abraham poly-parameter linear free energy relationships (pp-LFERs)^{1,2,46}. These models are generally accurate to within an order of magnitude across various organic solvent – water systems. For organic carbon – water partition coefficients, pp-LFER predictions are accurate to within a factor of 2-3 for three data sets^{3,6,7} (with N = 1375, 138, and 470 solutes, and $\log(K_{oc})$ RMSE = 0.65, 0.55, and 0.39, respectively).

This is not the case, however, for ionized species. Attempts have been made to predict the organic carbon – water partition coefficients of ionizable species using linear solvation energy relationships (LSERs) between the octanol – water and organic carbon – water partition coefficients for the ionic species. These methods have demonstrated some success (N = 62, RMSE = 0.47 and N = 43, RMSE = 0.49) for a set of ionized organic acids and ionized organic bases, respectively⁸⁰. Attempts have also been made to predict the partitioning of cationic amines onto soil organic carbon using Abraham solute descriptors for the charged species computed using linear relationships of the Abraham solute descriptors for the corresponding neutral species^{11,68}. While this method offers accurate predictions for a set of 66 primary through tertiary amines (RMSE = 0.42)¹³, there are a large number of coefficients (N = 20) required to obtain the charged species Abraham solute parameters from the

uncharged Abraham parameters¹³. Further, the regressions between the charged and neutral species Abraham solute parameters are not universal. There are varying equations with unique equation coefficients for different classes of chemical compounds^{10,11} (e.g., carboxylate anions, phenolate anions). Consequently, this method is not universally applicable for new solutes.

While both of these methods are reasonably accurate for predicting the linear organic carbon – water partition coefficients for ionizable species, previous work has shown non-linear sorption behavior for ionizable species, with Freundlich exponents ranging from 0.41 to 0.93, indicating highly non-linear sorption isotherms¹³. Consequently, a model is required that can accurately model non-linear adsorption of ionized species, and whose parameters can be predicted using physical chemical descriptors of either the neutral or ionized species.

The log-normal Langmuir (LNL) isotherm model, presented in Chapter 2, uses a log-normal distribution of Langmuir isotherms with 2 sorbent-specific parameters: the overall site density, q_{max} , and standard deviation of binding sites, σ_K ; and a single sorbate-specific median Langmuir binding constant, $\tilde{K}_L$, for each species⁸¹. The LNL model reproduces non-linear isotherms for neutral species to a high degree of accuracy ($N = 82$, $RMSE = 0.0985$)⁸¹. Further, as shown in Chapter 3, the resulting median binding constants, $\tilde{K}_L$ are predicted ($N = 82$, $RMSE = 0.364$)⁸² with accuracies comparable to those for the linear partition coefficients^{3,6,7}.

The purpose of this chapter is to extend the application of the LNL isotherm model to the non-linear partitioning of ionizable species and to predict the resulting chemical-specific median Langmuir binding constant, $\tilde{K}_L$, with accuracies comparable to those for linear partition coefficients of neutral sorbates.

5.2 Modeling & Experimental Data

Adsorption data for a suite of primary through quaternary amines onto standard Pahokee Peat soil were obtained from the previous work of Droge & Goss¹³. The compounds ranged significantly in size and structural moieties (Table 5-1). A complete table of the compounds analyzed can be found in Table D-1 in the Appendix.

Table 5-1. Summary of experimental sorption data for a set of primary through quaternary amines onto standard Pahokee Peat soil¹³.

Solute	Class	N	Notes
Amine Cations	Primary	18	aliphatic, substituted benzyl-aromatic, chloro-substituted
	Secondary	20	pharmaceuticals
	Tertiary	28	pharmaceuticals
	Quaternary	14	benzyl/phenyl-substituted, aliphatic-substituted
	Total	80	

In the original paper, Freundlich isotherms were fit to the experimental data. In this work, the log-normal Langmuir isotherm was fit to the same experimental sorption data. The isotherm is a linear superposition of Langmuir isotherms with a log-normal distribution of binding constants, K_L :

$$q(c) = \frac{q_{max}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c \tilde{K}_L e^{z\sigma_\kappa}}{1 + c \tilde{K}_L e^{z\sigma_\kappa}} \right) \exp\left(-\frac{z^2}{2}\right) dz \quad (5-1)$$

where q_{max} is the maximum site density (mmol/kg sorbent), $\tilde{K}_L$ is the median Langmuir binding constant for the distribution of sites (mM)⁻¹, σ_κ is the standard deviation of κ , the natural logarithm of $\tilde{K}_L$, c is the aqueous concentration (mM), and z is the normal random variable of integration ($z = \mu_\kappa - \kappa/\sigma_\kappa$).

5.2.1 Abraham Poly-parameter Linear Free Energy Relationship

The Abraham pp-LFER for predicting the median Langmuir binding constant, $\tilde{K}_{L,i}$, of a solute, i , is⁸²:

$$\log(\tilde{K}_{L,i}) = e_c E_i + s_c S_i + a_c A_i + b_c B_i + v_c V_i + c_c \quad (5-2)$$

where the lower case parameters represent the relative sorbent (*c*) – water; hydrogen bond accepting (*a_c*) and donating (*b_c*) capabilities, excess molar refraction (*e_c*), polarizability (*s_c*), and cavitation contributions (*v_c*), respectively. The upper case parameters represent the complimentary interactions of the solutes (*i*); hydrogen bond donating capacity (*A_i*), accepting capacity (*B_i*), excess molar refractivity (*E_i*), polarizability (*S_i*), and McGowan Volume (*V_i*) such that the product terms represent the sorbent-sorbate combined contributions for each type of inter-molecular interaction energy. The *c_c* term accounts for any non-specific binding interactions and has the units of $\tilde{K}_{L,i}$.

Absolv-estimated Abraham solute parameters (AAPs) were obtained using the Absolv software package⁴⁵. For the primary through tertiary amines, Absolv estimates were not available for the ionized species. Consequently, the Absolv Abraham parameters for the neutral species were obtained. For the quaternary amines Absolv estimates were available for the charged species directly.

5.2.2 Quantum-chemically Estimated Abraham Parameters (QCAP)

A method for obtaining QCAP solute descriptors has been developed by Liang & Di Toro⁶⁶. In the QCAP method, McGowan volumes (*V*) for the solutes are obtained directly from Absolv⁴⁵. Previous work by Zhao et al.⁸³ has shown excellent agreement between the McGowan volumes and van der Waals volumes for both neutral and ionic species. For convenience, the estimated McGowan volumes from Absolv are utilized.

The excess molar refractivity, E , defined as the molar refractivity of the given solute (MR_i), in excess of an n-alkane (MR_i^*) with the same characteristic volume is^{84,85}:

$$E = MR_i - MR_i^* \quad (5-3)$$

The total molar refraction, MR_i , can be computed as a function of the index of refraction (η) and the McGowan volume (V), of the solute, as follows⁷⁰:

$$MR_i = 10 \left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] V \quad (5-4)$$

Additionally, previous work has shown that molar refractivity of the alkane, MR_i^* , can be estimated from the McGowan volume⁸⁴:

$$MR_A^* = 2.832V - 0.526 \quad (5-5)$$

Combining Eqs. (5-3) through (5-5), an expression for the excess molar refractivity is obtained:

$$E = 10 \left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] V - 2.832V + 0.526 \quad (5-6)$$

The index of refraction, η , can be related to the molecular polarizability, α , using the Clausius-Mossotti equation⁸⁶:

$$\left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] = \frac{4}{3} \pi \cdot N_A \cdot \frac{\alpha}{V} \quad (5-7)$$

Combining Eqs. (5-6) and (5-7), an equation for the excess molar refractivity, E , results⁶⁶ which is a function of the molecular polarizability, α , and the McGowan volume, V :

$$E = 10 \cdot \frac{4}{3} \pi \cdot N_A \cdot \alpha - 2.832V + 0.526 \quad (5-8)$$

where N_A is Avogadro's constant. The result is an equation for E that can be computed using quantum chemical computations only. The molecular polarizability is computed using Gaussian 09⁶⁴ at the M062X/6-311++G** level of theory, which has been

shown in previous work to provide accurate computations of molecular polarizability^{66,87}.

The Abraham solute descriptors in Eq. (5-2), S , A , and B , remain to be determined. For neutral solutes, values for solvent-water system parameters ($e_c \dots c_c$) (Eq. (5-2)) are well-established⁴⁶. Consequently, these are used for new solutes for which Abraham solute descriptors are to be estimated⁶⁶. This is not the case, however, for ionic solutes. For new ionic solutes, both Abraham solute (S , A , B) as well as solvent-system parameters (e , s , a , b , v , c) are unknown and must be estimated. This can be accomplished by alternating multiple linear regressions, in which the error function, defined below, is minimized.

$$\hat{E} = \sum_{i=1}^n (\log K_{j,QC} - a_j A_i - b_j B_i - e_j E_i - s_j S_i - v_j V_i - c_j)^2 \quad (5-9)$$

where $\log K_{j,QC}$ is the quantum-chemically computed solvent (j) – water partition coefficient (L water / kg solvent) for solute (i). $\hat{E}$ is the error function (L water / kg solvent)² to be minimized.

In the first step, the solvent-system parameters are held constant. The neutral solvent-system descriptors⁴⁶ are used as initial values. A multiple linear regression is performed using QC computed partition coefficients for 49 solvent-water systems and the ionic species solute Abraham descriptors (S , A , and B) are estimated. In the second step, the estimated solute descriptors (S , A , B) are held constant and the solvent-system parameters (e , s , a , b , v , c) are estimated for each solvent using multiple linear regressions. This cycle is then repeated until the error function (Eq. (5-9)) is minimized. This results in both new quantum-chemically estimated (QCAP) solute and solvent-system parameters for the ionic solutes and the 49 solvent-water systems.

A computer program for estimating the QCAP solute and solvent parameters is presented in Appendix D.2a.

Quantum chemical calculations were performed in Gaussian 09⁶⁴ using the SMD⁶³ solvation model. M062X was chosen as the density functional, as it has demonstrated improved accuracies over similar M06 functionals in computing main group thermochemistry⁷². Since the solutes of interest are cations, long-range electrostatic interactions are important. Consequently, the 6-311++G** basis set was chosen, that includes diffuse functions. This basis set is similar in size and performance to the correlation-consistent double zeta (aug-cc-pVDZ) Dunning basis set with a slight improvement in computational efficiency⁷³.

Three-dimensional structures for the cations were constructed in ArgusLab⁷⁴ and were first optimized in the gas phase using a semi-empirical (AM1)⁷⁵ level of theory. The structures were then optimized in the gas phase first at the PM3⁸⁸ level of theory with gradients calculated at each iteration, then using M062X/6-311++G** in Gaussian 09⁶⁴. Frequency calculations were performed to ensure the structures were at an energy minimum with no imaginary frequencies. Finally, single point energy calculations were performed for the 49 solvent-water systems. Partition coefficients were calculated for each solute using the SMD-computed solvation free energies as follows:

$$\Delta G_{\text{solvent-water}} = G_{\text{solvent}} - G_{\text{water}} \quad (5-10)$$

$$\log(K_{\text{solvent-water}}) = \frac{\Delta G_{\text{solvent-water}}}{-2.303RT} \quad (5-10a)$$

where G_{solvent} and G_{water} are the Gibbs free energies in the solvent and water phases, respectively. $\Delta G_{\text{solvent-water}}$ is the Gibbs free energy of transfer between the two

phases, R is the universal gas constant (kcal/mol °K), T is the temperature (°K), and $\log(K_{\text{solvent-water}})$ is the base-10 logarithm of the solvent-water partition coefficient.

Of the 80 amines in the data set, successful optimized structures with no imaginary frequencies ($\text{NImag} = 0$) were obtained for 61 of the sorbates. Optimized structures (with $\text{NImag} > 0$) for the remaining sorbates were not obtained at the level of theory used in this analysis. This is likely due to structural conformers or a starting geometry too far from the optimum geometry for some of the more structurally-complex molecules.

5.3 Results & Discussion

5.3.1 Isotherm Model

Since some species were partially ionized at the experimental pH values, sorption isotherms for both the neutral and ionized species are required. For the set of amines ($n = 80$) 67 species are fully ionized, while the remaining species ($n = 13$) are mostly ionized with an average ionized fraction, $\bar{f}_+ = 0.72$ for both experimental pH's. For the neutral species a conventional linear isotherm model was selected since these are known to provide reliable estimates of partitioning to organic carbon for neutral chemicals. The log-normal Langmuir model (Eq. (5-1)) was chosen for the charged species since the total sorption data exhibits strongly non-linear behavior. This results in a combined isotherm equation, which is a superposition of the individual charged and neutral species sorption models:

$$q_{\text{total}}(c) = f_+ \frac{q_{\text{max}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c \tilde{K}_{L,+} e^{z\sigma_{\kappa,+}}}{1 + c \tilde{K}_{L,+} e^{z\sigma_{\kappa,+}}} \right) \exp\left(-\frac{z^2}{2}\right) dz + f_0 K_P c \quad (5-11)$$

where f_+ and f_0 are the fractions of ionized and neutral species, respectively:

$$f_+ = \frac{[H^+]}{[H^+] + K_A} \quad (5-12a)$$

$$f_0 = \frac{K_A}{[H^+] + K_A} \quad (5-12b)$$

and where K_A is the acid-dissociation constant, and $[H^+]$ is the aqueous proton concentration (mol/L).

For neutral species, a reliable pp-LFER model for the linear organic carbon – water partition coefficients ($\log(K_{OC})$) exists^{6,7}. The K_{OC} for neutral species can be computed using the organic carbon – water Abraham pp-LFER system parameters obtained by Kipka and Di Toro⁷ and the Absolv-estimated solute descriptors for the neutral species:

$$\log(K_{OC}) = 1.198E - 0.08S - 0.192A - 1.807B + 1.155V + 0.724 \quad (5-13)$$

The organic carbon – water partition coefficient (Eq. (5-13)) can be converted to an overall partition coefficient by multiplying by the fraction organic carbon present in the soil (f_{OC}):

$$K_P = f_{OC} K_{OC} \quad (5-14)$$

Combining Eqs. (5-14) and (5-11), an expression for the total sorption of neutral and charged sorbate species is obtained:

$$q_{total}(c) = f_+ \frac{q_{max}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{c \tilde{K}_{L,+} e^{z\sigma_{\kappa,+}}}{1 + c \tilde{K}_{L,+} e^{z\sigma_{\kappa,+}}} \right) \exp\left(\frac{-z^2}{2}\right) dz \quad (5-15)$$

$$+ f_0 f_{OC} K_{OC} c$$

The remaining parameters in Eq. (5-15): q_{max} , $\sigma_{\kappa,+}$, and $\tilde{K}_{L,+}$ were estimated by minimizing the sum of squares of the residuals using the Excel Solver package with Eq. (5-15) as the prediction. The resulting isotherms for a subset of the primary through quaternary amines ($n = 34$) is illustrated in Figure 5-1. A computer program to

evaluate Eq. (5-15) is presented in Appendix D.2b. A complete table of estimated isotherm parameters is presented in Table D-2 in the Appendix.

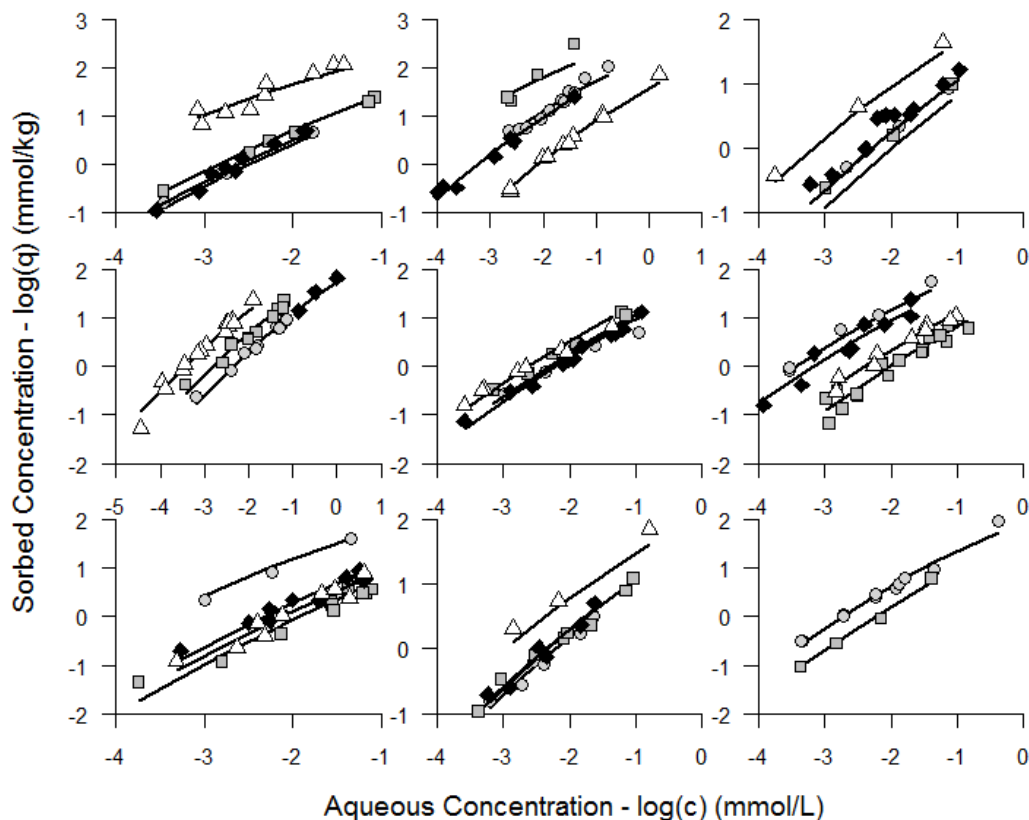


Figure 5-1. Adsorption isotherms for a selection of 34 primary through quaternary amine cations. Points represent the experimental adsorption data (see Figure D-1 in the Appendix for a legend of solutes). Solid lines represent the isotherm model (Eq. (5-15)). Model parameters are presented in Table D-2 in the Appendix.

The isotherm model (Eq. (5-15)) does an excellent job of reproducing the experimental sorption data for the primary through quaternary amine species. The RMS error for the model-predicted sorbed concentrations, $\log(q(c))$, is 0.272 for the

entire set of solutes ($N = 80$). For species which were fully ionized at the experimental pH the RMS error was slightly smaller ($RMSE = 0.137$, $N = 67$) indicating that there may larger errors in the linear sorption model for the neutral species. It should be noted that for partially ionized species there are no estimated parameters for the sorption of the neutral fraction, it is a direct prediction. This is likely the cause of the larger RMS errors for the partially ionized species.

As demonstrated for neutral solutes⁸¹, when a single sorbent-specific sorption capacity, q_{max} , and sorbent-specific distribution of binding sites, σ_k , are used, adsorption isotherms for multiple solutes can be displayed as a single isotherm, when the sorbed concentration, q , is plotted against the product of the aqueous concentration, c , and the median Langmuir binding constants for the individual solutes, $\tilde{K}_L$, as demonstrated in Figure 2-4B,C. This is a useful tool for more easily comparing the individual isotherms as well as illustrating the behavior of the LNL isotherm model. Since there is both sorption of the charged and neutral species when a sorbate is partially ionized, an analysis of the normalized LNL isotherm is only possible for the subset of species where the majority of the sorption occurring is that of the charged species. Linear partition coefficients for neutral species can span orders of magnitude. Consequently, the f_+ and f_0 terms in Eq. (5-15) cannot singularly determine if sorption of the neutral species is significant (even at very low values of f_0). It is necessary to compute the contributions of the neutral and charged species to the total sorbed concentration, $q_{total}(c)$. A cutoff fraction of total sorption, f_{q+} , for the charged species of 99.9% was selected, above which the sorption of neutral species was considered to be negligible. The resulting normalized LNL isotherm for the sorbates above the cutoff is shown in Figure 5-2.

Figure 5-2 demonstrates that there is no significant bias in the predicted isotherms as the LNL isotherm approaches saturation. This suggests that the LNL isotherm works equally well for the non-linear region of adsorption as well as the nearly linear region. Further, nearly all of the experimental sorption points for the ionized species are predicted to within a factor of 2, indicating a high level of accuracy for the LNL model. This illustrates that the LNL isotherm can be used to model charged sorption data with accuracies comparable to those for neutral species (RMSE = 0.0985, N = 82)⁸¹.

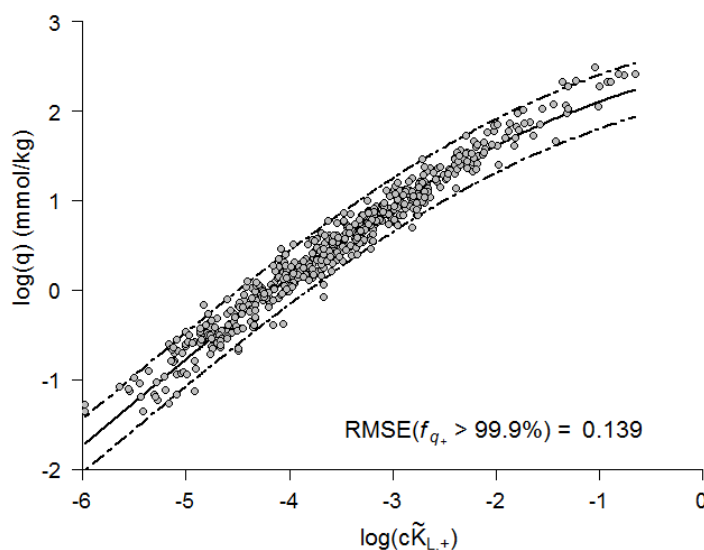


Figure 5-2. Normalized isotherms for a selection of 63 primary through quaternary amine cations (where $f_{q+} \geq 99.9\%$). Points represent the experimental adsorption data. Solid lines represent the fit to the LNL isotherm model (Eq. (5-15)). Dashed lines represent ± 0.3 log-units (factor of 2) error in the predicted isotherms. The RMS error for the $f_{q+} \geq 99.9\%$ subset of chemicals is 0.139 (N = 69, with 567 data points). Estimated isotherm parameters and associated standard errors can be found in Table D-2 in the Appendix.

5.3.2 Predicting Median Langmuir Binding Constants for Ionic Species

The median binding constants, $\tilde{K}_{L,+}$, are the only sorbate-specific parameter for the LNL isotherm. Consequently they need to be predicted in order to make isotherm predictions for the sorption of new charged species. Two methods were tested for predicting the LNL-estimated median Langmuir binding constants for the ionized species, $\tilde{K}_{L,+}$, obtained from fitting Eq. (5-15) to the sorption isotherms. Both were based on the Abraham pp-LFER (Eq. 5-2). The solute parameters used were either the Absolv-estimated (AAP) solute parameters or the quantum-chemically estimated (QCAP) solute parameters. For each method, plots of the logarithmic residuals vs. the Tukey mean⁸⁹ of the observed and predicted median binding constants were constructed (Figure 5-3). A complete table of Absolv-estimated solute parameters, QCAP solute parameters, and QCAP solvent parameters are presented in Tables D-3 and D-4 in the Appendix.

Both models provide accurate predictions of the median Langmuir binding constants, with RMS errors of 0.526 and 0.569 for the QCAP and AAP solute descriptors, respectively. However, there is a bias with increasing median binding constant when the Absolv solute descriptors are used (Figure 5-2A). This bias is not observed for the QCAP model (Figure 5-2C). This bias can be quantified by computing a Pearson R-squared (R^2) correlation coefficient between the logarithmic residuals of the median binding constants and the Tukey mean values of the predicted and LNL-estimated median binding constants (Figures 5-2B, 5-2D). Large R^2 values indicate a strong correlation between the magnitude of the binding constant and the error in the predicted values of the median binding constant which indicates the presence of a bias.

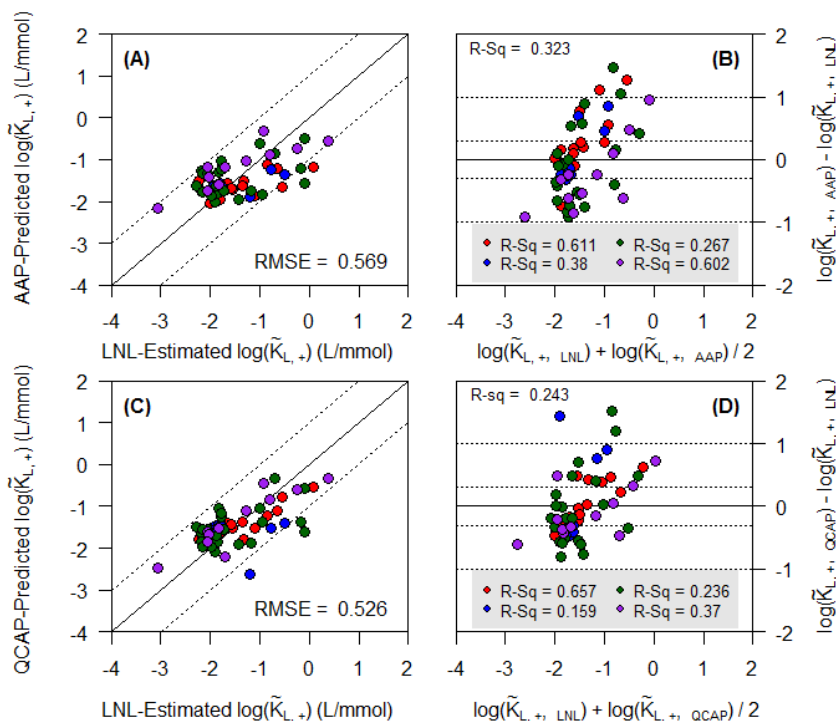


Figure 5-3. (A) AAP-predicted vs. LNL-estimated median binding constant, ($\log(\tilde{K}_{L,+})$). (B) Logarithmic residuals $\log(\tilde{K}_{L,+ \text{ AAP}}) - \log(\tilde{K}_{L,+ \text{ LNL}})$ vs. Tukey mean median binding constant $(\log(\tilde{K}_{L,+ \text{ AAP}}) + \log(\tilde{K}_{L,+ \text{ LNL}})) / 2$. (C) QCAP-predicted vs. LNL-estimated median binding constant, ($\log(\tilde{K}_{L,+})$). (D) Logarithmic residuals $\log(\tilde{K}_{L,+ \text{ QCAP}}) - \log(\tilde{K}_{L,+ \text{ LNL}})$ vs. Tukey mean median binding constant $(\log(\tilde{K}_{L,+ \text{ QCAP}}) + \log(\tilde{K}_{L,+ \text{ LNL}})) / 2$.

While the overall Tukey residual R^2 for the 60 amines is small for both the AAP and QCAP methods ($R^2 = 0.323$ and 0.243 , respectively), the individual amine-type R^2 values vary significantly between the two methods. The Tukey residual R^2 for the primary and tertiary amines are comparable for both the AAP- and QCAP-predicted median binding constants. However, the AAP-predicted median binding constants show significantly increased bias in the residuals for the secondary and

quaternary amines ($R^2 = 0.380$ and 0.602 for AAP vs. 0.159 and 0.370 for QCAP), respectively. Estimated sorbent-water system parameters and associated standard errors for the AAP and QCAP methods can be seen in Eqs. (5-16) and (5-17). A complete description of the RMS errors and Tukey R^2 correlation coefficients for the AAP and QCAP methods can be found in Table 5-2.

$$\log(\tilde{K}_{L,+ AAP}) = 0.84(\pm 0.433)E - 0.37(\pm 0.714)S + 0.91(\pm 0.479)A - 1.05(\pm 0.443)B + 0.60(\pm 0.172)V - 2.21(\pm 0.290) \quad (5-16)$$

$$N = 60, RMSE = 0.569$$

$$\log(\tilde{K}_{L,+ QCAP}) = 0.64(\pm 0.180)E + 0.91(\pm 0.314)S + 1313(\pm 530)A - 0.17(\pm 0.299)B + 0.56(\pm 0.201)V - 4.14(\pm 0.861) \quad (5-17)$$

$$N = 60, RMSE = 0.526$$

Table 5-2. Summary of RMS errors and Tukey R^2 correlation coefficients for AAP- and QCAP-predicted median Langmuir binding constants.

Class	AAP Solute Parameters			QCAP Solute Parameters		
	N	RMSE	Tukey R^2	N	RMSE	Tukey R^2
Primary	16	0.501	0.61	16	0.340	0.66
Secondary	10	0.341	0.38	10	0.615	0.16
Tertiary	23	0.651	0.27	23	0.623	0.24
Quaternary	11	0.637	0.60	11	0.433	0.37
Total	60	0.569	0.32	60	0.526	0.24

5.3.3 Residual Analysis

A box plot of the logarithmic residual values is shown in Figure 5-4, ordered by increasing predicted LNL median binding constant within each type of amine. For the primary and tertiary amines, there is no significant bias in the predicted isotherm residuals with increasing median binding strength. The residuals are smaller, in most cases, for the primary amines. This is reflected in their respective RMS errors (RMSE = 0.340 vs. 0.501, Table 5-2). Figure 5-3 shows that for secondary and quaternary amines there are significantly larger R^2 correlations between the Tukey mean median binding constants and the residuals. For the secondary amines, because the median binding constants are very close in magnitude, this bias is not visible in the predicted median binding constants (Figure 5-3A,B). However, for both AAP and QCAP-predicted isotherms, there is a systematic over prediction of sorbed concentrations (q), with the AAP model performing slightly better for the 9 secondary amines (Figure 5-4). This is not the case for quaternary amines, where there is significant inverse correlation between the residuals and the median binding constant. This correlation is reduced in the QCAP predictions of the isotherms (Figure 5-4B). Overall, predicting sorption isotherms using the QCAP-predicted median binding constants results in a slightly smaller RMS error (RMSE = 0.457 vs. 0.478) compared to predictions made using the AAP-predicted median binding constants.

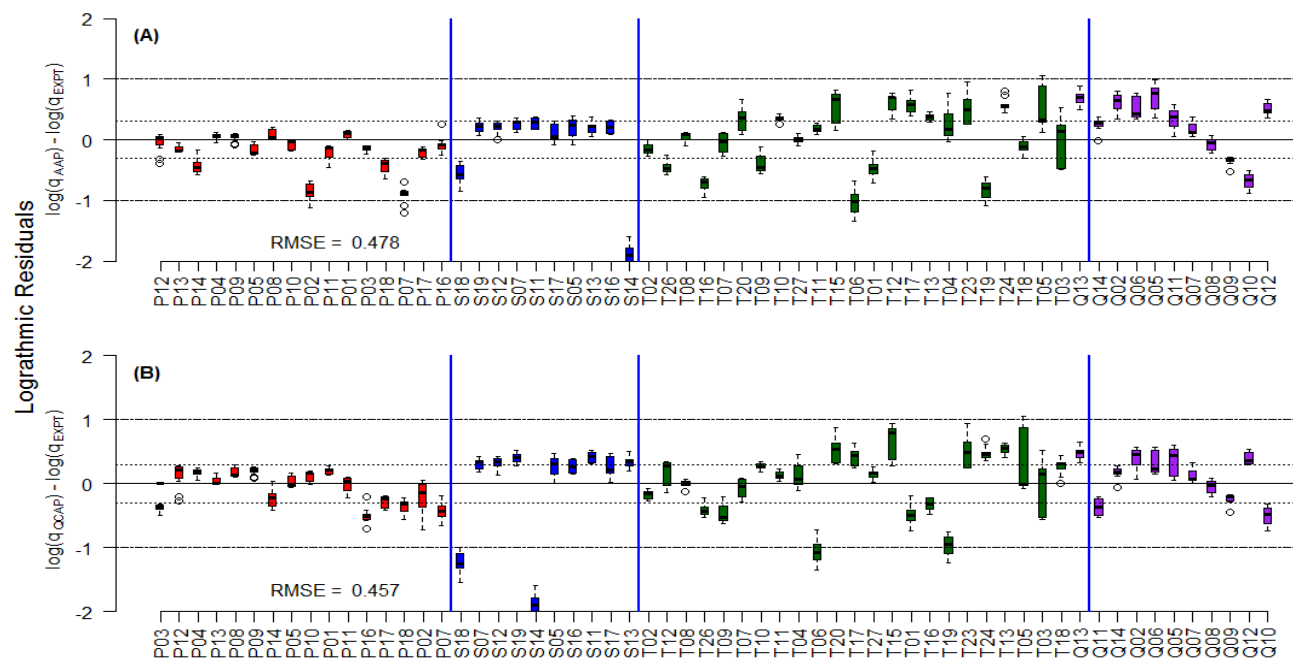


Figure 5-4. Box plots of the (A) AAP-predicted LNL isotherm residuals $\log(q_{AAP}) - \log(q_{expt})$ and (B) QCAP-predicted LNL isotherm residuals $\log(q_{QCAP}) - \log(q_{expt})$, ordered by increasing predicted LNL median binding constant within each type of amine. Vertical lines separate the types of amines. Horizontal black solid lines represent 1:1 agreement between predicted and experimental sorbed concentrations, dashed lines represent ± 0.3 log-units difference between predicted and observed adsorbed concentrations and semi-dashed lines represent ± 1.0 log-units difference between predicted and observed adsorbed concentrations. IQR contains 50% of the data, whiskers represent ± 1.5 IQR and filled points represent outliers > 1.5 IQR).

5.4 Implications

It has been shown that the LNL isotherm model can accurately reproduce non-linear adsorption of the primary through quaternary amine ionic species onto natural organic matter. The median binding constants, $\tilde{K}_{L,+}$, for the ionized species can be predicted with an Abraham pp-LFER model using quantum-chemically estimated Abraham solute descriptors. The accuracy of the predictions made using the QCAP solute descriptors are an improvement over those made using the existing AAP solute descriptors. The significant biases in the AAP-predicted binding constants for the secondary and quaternary amines are reduced when the QCAP descriptors are used. The predicted isotherms using QCAP-predicted median Langmuir binding constants accurately reproduce the experimental data with decreased RMS error compared to the isotherms predicted using AAP-predicted median Langmuir binding constants.

The combination of the LNL and QCAP Abraham models provide prediction capabilities for the sorption of charged solutes without chemical-class specific regressions. The model is independent of the chemical class of the sorbate and is applicable for ionic species. It is expected that for other classes of ionizable organic compounds (e.g., drug-like compounds, pesticides, ionic liquids) the median Langmuir binding constant, $\tilde{K}_{L,+}$, can be predicted using the QCAP pp-LFER method. The LNL model can also be utilized for other sorbents that exhibit non-linear sorption behavior (e.g. activated carbon, graphite, biochar).

Chapter 6

CONCLUSIONS & RECOMMENDATIONS

In this doctoral thesis, a non-linear isotherm model was developed for both neutral and ionizable organic species. This model utilizes a single sorbate-specific parameter, the median Langmuir binding constant, $\log(\tilde{K}_L)$, which can be predicted for both neutral and ionizable species using an Abraham pp-LFER model.

The non-linear sorption of neutral organic species onto various forms of black carbon is accurately modeled using the log-normal Langmuir (LNL) isotherm model presented in Chapter 2. The results suggest that for neutral species, the LNL isotherm model can accurately reproduce sorption isotherms for various sorbates onto graphite and wood char as well as Darco and F400 granular activated carbons. Further, the LNL isotherm model accomplishes this with two sorbent-specific parameters: the maximum sorption capacity, q_{max} , and the standard deviation of the logarithm of the Langmuir binding constants, σ_K ; and a single sorbate-specific isotherm parameter, the median Langmuir binding constant, $\tilde{K}_L$. Further, the model can predict the maximum sorption capacity, q_{max} , even when isotherm curvature is not apparent in the individual isotherms, as illustrated by the application to graphite. It accomplishes this using the differences in isotherm slopes of the various sorbates to extrapolate to saturation.

Future work should concentrate on understanding the sorbent-specific parameters, q_{max} , and, in particular, the standard deviation of the logarithm of the Langmuir binding constants, σ_K . It is only the median Langmuir binding constant, $\tilde{K}_L$,

that is sorbate specific. The remarkable finding is that the same σ_k can be applied to all the sorbates. This should be tested for additional sorbents. It would also be useful to have a more mechanistic understanding of the underlying chemical properties that result in this behavior. Examining molecular models of simple sorbents, for example graphite, may contribute to a more in-depth understanding.

In Chapter 3, the median Langmuir binding constants, $\log(\tilde{K}_L)$, were predicted for the four carbon sorbents presented in Chapter 2, using the Abraham pp-LFER model and Absolv-estimated Abraham solute descriptors. The results suggested that for the porous sorbents, particularly F400 GAC, a volumetric cutoff is present. The molecular volume cutoff is consistent with the median pore volume of F400 GAC. It appears that the large molecular volume sorbates do not interact with the sorbent sites accessed by the smaller molecules. This conclusion is suggested by observing that for these large sorbates, there was weak correlation between the observed median Langmuir binding constants and the Absolv-estimated Abraham solute descriptors (E , S , A , B , V). Instead, for these sorbates above the molecular volume cutoff, a strong inverse correlation to the experimental aqueous solubility, $\log(S)$, was observed. Further, the molecular volume cutoff was not observed for the graphite data set, the only non-porous sorbent. It would be useful to include isotherm data in future work which directly compares identical sets of sorbates onto the different sorbents to determine if the volumetric cutoff observed varies for different sorbents. Also it would be useful to have a more mechanistic understanding of the molecular properties of both the sorbates and sorbents that exhibit this behavior.

In Chapter 4, the QCAP method developed by Liang & Di Toro was extended to estimate solute and solvent-water Abraham descriptors for ionizable species. The

QCAP solute descriptors were then used to predict solvent-water partition coefficients for a set of carboxylate anions and quaternary amine cations in various solvent-water systems. For the carboxylate anions in acetone-, acetonitrile-, DMSO-, and methanol-water systems, the QCAP Abraham solute descriptors offered significant improvements over both direct a priori quantum chemical computations and predictions made using the existing Absolv Abraham solute descriptors. For octanol-water partitioning of quaternary amines, the QCAP Abraham predictions were a significant improvement over the direct QC computations; however, they were not an improvement over predictions made using the existing Absolv-estimated solute descriptors. It is difficult to compare the performance of the QCAP method between cations and anions, as there is variation both in the solvent-water systems examined, as well as significant differences in the size and complexity of the solute molecules in the two data sets. It would be useful in future work to directly compare anions and cations of similar size and complexity in the same solvent-water systems. This would help to determine if the performance of the QCAP method is similar for cations and anions, or if the increased RMS error observed for the set of quaternary amine cations is representative of increased predictive errors for cations in general.

In Chapter 5, the LNL isotherm model developed in Chapter 2 was extended to model the non-linear partitioning of ionizable organic species onto natural organic carbon (Pahokee peat). The model was modified to include a conventional linear isotherm for the neutral component present and the LNL model for the ionized component. For a set of primary through quaternary amines, QCAP solute descriptors were estimated and used to predict median Langmuir binding constants for the non-linear sorption of the ionic component onto organic carbon. It was demonstrated that

using the QCAP Abraham solute descriptors offered slight improvements in predicted median Langmuir binding constants compared to the existing Absolv-estimated descriptors. Further, it was demonstrated that using the QCAP solute descriptors significantly reduced the prediction bias for the secondary and quaternary amine sorbates. It would be useful for future work to include more varied sorbates, particularly anionic sorbates, as well as additional sorbents to test the applicability and range of the LNL sorption and QCAP pp-LFER models.

REFERENCES

- (1) Endo, S.; Goss, K.-U. Applications of Polyparameter Linear Free Energy Relationships in Environmental Chemistry. *Environmental Science & Technology* **2014**, *48*, 12477-12491.
- (2) Goss, K. U.; Schwarzenbach, R. P. Linear free energy relationships used to evaluate equilibrium partitioning of organic compounds. *Environmental Science & Technology* **2001**, *35*, 1-9.
- (3) Poole, S. K.; Poole, C. F. Chromatographic models for the sorption of neutral organic compounds by soil from water and air. *Journal of Chromatography A* **1999**, *845*, 381-400.
- (4) Platts, J. A.; Abraham, M. H.; Butina, D.; Hersey, A. Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients. *Journal of Chemical Information and Computer Sciences* **2000**, *40*, 71-80.
- (5) Endo, S.; Grathwohl, P.; Haderlein, S. B.; Schmidt, T. C. LFERs for Soil Organic Carbon-Water Distribution Coefficients (K-OC) at Environmentally Relevant Sorbate Concentrations. *Environmental Science & Technology* **2009**, *43*, 3094-3100.
- (6) Kipka, U.; Di Toro, D. M. A linear solvation energy relationship model of organic chemical partitioning to dissolved organic carbon. *Environmental Toxicology and Chemistry* **2011**, *30*, 2023-2029.

- (7) Kipka, U.; Di Toro, D. M. A linear solvation energy relationship model of organic chemical partitioning to particulate organic carbon in soils and sediments. *Environmental Toxicology and Chemistry* **2011**, *30*, 2013-2022.
- (8) EPA Estimation Programs Interface Suite™ for Microsoft® Windows. v4.11 ed.; US EPA: Washington, D.C., **2014**.
- (9) Franco, A.; Ferranti, A.; Davidsen, C.; Trapp, S. An unexpected challenge: ionizable compounds in the REACH chemical space. *The International Journal of Life Cycle Assessment* **2010**, *15*, 321-325.
- (10) Abraham, M. H.; Acree Jr, W. E. The transfer of neutral molecules, ions and ionic species from water to ethylene glycol and to propylene carbonate; descriptors for pyridinium cations. *New Journal of Chemistry* **2010**, *34*, 2298-2305.
- (11) Abraham, M. H.; Acree, W. E., Jr. Equations for the Transfer of Neutral Molecules and Ionic Species from Water to Organic phases. *Journal of Organic Chemistry* **2010**, *75*, 1006-1015.
- (12) Cho, C.-W.; Stolte, S.; Yun, Y.-S.; Krossing, I.; Thöming, J. In silico prediction of linear free energy relationship descriptors of neutral and ionic compounds. *RSC Advances* **2015**, *5*, 80634-80642.
- (13) Droge, S. T. J.; Goss, K.-U. Ion-Exchange Affinity of Organic Cations to Natural Organic Matter: Influence of Amine Type and Nonionic Interactions at Two Different pHs. *Environmental Science & Technology* **2013**, *47*, 798-806.
- (14) Zhao, Y. H.; Abraham, M. H. Octanol/Water Partition of Ionic Species, Including 544 Cations. *The Journal of Organic Chemistry* **2005**, *70*, 2633-2640.

- (15) Marenich, A. V.; Kelly, C. P.; Thompson, J. D.; Hawkins, G. D.; Chambers, C. C.; Giesen, D. J.; Winget, P. C., C. J.; Truhlar, D. G., Minnesota Solvation Database – version **2012**
- (16) Limousin, G.; Gaudet, J. P.; Charlet, L.; Szenknect, S.; Barthès, V.; Krimissa, M. Sorption isotherms: A review on physical bases, modeling and measurement. *Applied Geochemistry* **2007**, 22, 249-275.
- (17) Cornelissen, G.; Gustafsson, Ö.; Bucheli, T. D.; Jonker, M. T. O.; Koelmans, A. A.; van Noort, P. C. M. Extensive Sorption of Organic Compounds to Black Carbon, Coal, and Kerogen in Sediments and Soils: Mechanisms and Consequences for Distribution, Bioaccumulation, and Biodegradation. *Environmental Science & Technology* **2005**, 39, 6881-6895.
- (18) Sips, R. On the Structure of a Catalyst Surface. *Journal of Chemical Physics* **1948**, 16, 490-495.
- (19) Sips, R. On the Structure of a Catalyst Surface .2. *Journal of Chemical Physics* **1950**, 18, 1024-1026.
- (20) Grimm, D. M.; Azarraga, L. V.; Carreira, L. A.; Susetyo, W. Continuous Multiligand Distribution Model used to Predict the Stability Constant of CU(II) Metal Complexation with Humic Material from Fluorescence Quenching Data. *Environmental Science & Technology* **1991**, 25, 1427-1431.
- (21) Nederlof, M. M.; Dewit, J. C. M.; Vanriemsdijk, W. H.; Koopal, L. K. Determination of Proton Affinity Distributions for Humic Substances. *Environmental Science & Technology* **1993**, 27, 846-856.

- (22) Perdue, E. M.; Lytle, C. R. Distribution Model for Binding of Protons and Metal-Ions by Humic Substances. *Environmental Science & Technology* **1983**, *17*, 654-660.
- (23) Rate, A. W.; McLaren, R. G.; Swift, R. S. Evaluation of a Log-Normal Distribution 1st-Order Kinetic-Model for Copper(II)-Humic Acid Complex Dissociation. *Environmental Science & Technology* **1992**, *26*, 2477-2483.
- (24) Susetyo, W.; Dobbs, J. C.; Carreira, L. A.; Azarraga, L. V.; Grimm, D. M. Development of a Statistical-Model for Metal Humic Interactions. *Analytical Chemistry* **1990**, *62*, 1215-1221.
- (25) Calleja, G.; Coto, B.; Morales-Cas, A. M. Adsorption energy distribution in activated carbon from grand canonical Monte Carlo calculation. *Applied Surface Science* **2006**, *252*, 4345-4352.
- (26) Dabrowski, A.; Podkoscielny, P.; Bulow, M. Comparison of energy-distribution functions calculated for gas-solid and liquid-solid adsorption data. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **2003**, *212*, 109-114.
- (27) de Torre, L. E. C.; Bottani, E. J. Adsorption energy distribution functions. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **1996**, *116*, 285-291.
- (28) Jagiello, J.; Schwarz, J. A. Local Exact and Approximate Solutions of the Adsorption Integral-Equation with a Kernel of a Langmuir-like Isotherm - Determination of Adsorption Energy-Distribution. *Journal of Colloid and Interface Science* **1991**, *146*, 415-424.

- (29) Wu, C.-H.; Shr, J.-F.; Wu, C.-F.; Hsieh, C.-T. Adsorption energy distribution of carbon tetrachloride on carbon nanofiber arrays prepared by template synthesis. *Physica E-Low-Dimensional Systems & Nanostructures* **2008**, *40*, 814-821.
- (30) Podkościelny, P.; Nieszporek, K. Adsorption of phenols from aqueous solutions: Equilibria, calorimetry and kinetics of adsorption. *Journal of Colloid and Interface Science* **2011**, *354*, 282-291.
- (31) Schwartzbach, R. P. G., P. M.; Imboden, D. *Environmental Organic Chemistry*; John Wiley & Sons: Hoboken, **2003**.
- (32) Aitchison, J. B., A. A. C.; *The Lognormal Distribution*; University Press: Cambridge, **1957**.
- (33) Press, W. H. T., S.; Vetterling, W.; Flannery, B. *Numerical Recipes in Fortran 77*; Cambridge University Press: Melbourne, **1992**.
- (34) Shih, Y.-H.; Gschwend, P. M. Evaluating Activated Carbon-Water Sorption Coefficients of Organic Compounds Using a Linear Solvation Energy Relationship Approach and Sorbate Chemical Activities. *Environmental Science & Technology* **2009**, *43*, 851-857.
- (35) Speth, T. F.; Miltner, R. J. Technical Note - Adsorption Capacity of GAC for Synthetic Organics. *Journal American Water Works Association* **1990**, *82*, 72-75.
- (36) Zhu, D. Q.; Pignatello, J. J. Characterization of aromatic compound sorptive interactions with black carbon (charcoal) assisted by graphite as a model. *Environmental Science & Technology* **2005**, *39*, 2033-2041.
- (37) de Levie, R. *Advanced Excel for Scientific Data Analysis*; Oxford University Press: New York, **2004**.

- (38) Schenzel, J.; Goss, K.-U.; Schwarzenbach, R. P.; Bucheli, T. D.; Droge, S. T. J. Experimentally Determined Soil Organic Matter-Water Sorption Coefficients for Different Classes of Natural Toxins and Comparison with Estimated Numbers. *Environmental Science & Technology* **2012**, *46*, 6118-6126.
- (39) Neale, P. A.; Escher, B. I.; Goss, K.-U.; Endo, S. Evaluating dissolved organic carbon-water partitioning using polyparameter linear free energy relationships: Implications for the fate of disinfection by-products. *Water Research* **2012**, *46*, 3637-3645.
- (40) Card, M. L.; Chin, Y.-P.; Lee, L. S.; Khan, B. Prediction and Experimental Evaluation of Soil Sorption by Natural Hormones and Hormone Mimics. *Journal of Agricultural and Food Chemistry* **2012**, *60*, 1480-1487.
- (41) Hawthorne, S. B.; Grabanski, C. B.; Miller, D. J.; Arp, H. P. H. Improving Predictability of Sediment-Porewater Partitioning Models using Trends Observed with PCB-Contaminated Field Sediments. *Environmental Science & Technology* **2011**, *45*, 7365-7371.
- (42) Bronner, G.; Goss, K.-U. Predicting Sorption of Pesticides and Other Multifunctional Organic Chemicals to Soil Organic Carbon. *Environmental Science & Technology* **2011**, *45*, 1313-1319.
- (43) de Ridder, D. J.; Villacorte, L.; Verliefde, A. R. D.; Verberk, J. Q. J. C.; Heijman, S. G. J.; Amy, G. L.; van Dijk, J. C. Modeling equilibrium adsorption of organic micropollutants onto activated carbon. *Water Research* **2010**, *44*, 3077-3086.
- (44) Sun, H.; Zhou, Z. Impacts of charcoal characteristics on sorption of polycyclic aromatic hydrocarbons. *Chemosphere* **2008**, *71*, 2113-2120.

- (45) ACD/ADME Suite. 5.0 ed.; Advanced Chemistry Development, Inc.: Toronto, ON, **2010**.
- (46) Endo, S; Ulrich, N.; Bronner, G.; Goss, K.-U. UFZ-LSER Database. 2.1 ed.; Helmholtz Centre for Environmental Research: Leipzig, Germany, **2014**.
- (47) ChemSpider, Chemistry, R. S. o., **2014**
- (48) Arey, J. S.; Green, W. H.; Gschwend, P. M. The Electrostatic Origin of Abraham's Solute Polarity Parameter. *The Journal of Physical Chemistry B* **2005**, *109*, 7564-7573.
- (49) Nguyen, T. H.; Cho, H.-H.; Poster, D. L.; Ball, W. P. Evidence for a Pore-Filling Mechanism in the Adsorption of Aromatic Hydrocarbons to a Natural Wood Char. *Environmental Science & Technology* **2007**, *41*, 1212-1217.
- (50) Lu, Q. L.; Sorial, G. A. The role of adsorbent pore size distribution in multicomponent adsorption on activated carbon. *Carbon* **2004**, *42*, 3133-3142.
- (51) Li, L.; Quinlivan, P. A.; Knappe, D. R. U. Effects of activated carbon surface chemistry and pore structure on the adsorption of organic contaminants from aqueous solution. *Carbon* **2002**, *40*, 2085-2100.
- (52) Walker, G. M.; Weatherley, L. R. Adsorption of dyes from aqueous solution - the effect of adsorbent pore size distribution and dye aggregation. *Chemical Engineering Journal* **2001**, *83*, 201-206.
- (53) Chiou, C. T.; Schmedding, D. W.; Manes, M. Improved Prediction of Octanol–Water Partition Coefficients from Liquid–Solute Water Solubilities and Molar Volumes. *Environmental Science & Technology* **2005**, *39*, 8840-8846.

- (54) Razzaque, M. M.; Grathwohl, P. Predicting organic carbon–water partitioning of hydrophobic organic chemicals in soils and sediments based on water solubility. *Water Research* **2008**, *42*, 3775-3780.
- (55) Hansch, C. Q., J.; Lawrence, G. The Linear Free-Energy Relationship between Partition Coefficients and the Aqueous Solubility of Organic Liquids. *Journal of Organic Chemistry* **1968**, *33*, 347-350.
- (56) Cook, R. D.; Weisberg, S. Residuals and influence in regression. **1982**.
- (57) Mendenhall, W.; Sincich, T.; Boudreau, N. S. *A second course in statistics: regression analysis*; Prentice Hall Upper Saddle River, New Jersey, **1996**; Vol. 5.
- (58) Leo, A.; Weininger, D. Pomona MedChem CLogP Database and Software Manual. *Daylight Chemical Information Systems, Irvine, CA, USA* **1989**.
- (59) Schüürmann, G.; Ebert, R.-U.; Kühne, R. Prediction of Physicochemical Properties of Organic Compounds from 2D Molecular Structure – Fragment Methods vs. LFER Models. *CHIMIA International Journal for Chemistry* **2006**, *60*, 691-698.
- (60) Klamt, A. The COSMO and COSMO-RS solvation models. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 699-709.
- (61) Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS: An Alternative to Simulation for Calculating Thermodynamic Properties of Liquid Mixtures. *Annual Review of Chemical and Biomolecular Engineering* **2010**, *1*, 101-122.
- (62) Lin, S.-T.; Sandler, S. I. A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. *Industrial & Engineering Chemistry Research* **2002**, *41*, 899-913.

- (63) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on the Generalized Born Approximation with Asymmetric Descreening. *Journal of Chemical Theory and Computation* **2009**, *5*, 2447-2464.
- (64) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09. Gaussian, Inc.: Wallingford, CT, USA, **2009**.
- (65) Xiong, R.; Sandler, S. I.; Burnett, R. I. An Improvement to COSMO-SAC for Predicting Thermodynamic Properties. *Industrial & Engineering Chemistry Research* **2014**, *53*, 8265-8278.
- (66) Liang, Y. Z. Determining Abraham Solute and System Parameters for Neutral Organic Compounds using Quantum Chemical Methods. *Dissertation in Civil & Environmental Engineering*; University of Delaware: Newark, **2016**.
- (67) Zissimos, A. M.; Abraham, M. H.; Klamt, A.; Eckert, F.; Wood, J. A. Comparison between the Two General Sets of Linear Free Energy Descriptors of Abraham and Klamt. *Journal of Chemical Information and Computer Sciences* **2002**, *42*, 1320-1331.
- (68) Abraham, M. H.; Zhao, Y. H. Determination of Solvation Descriptors for Ionic Species: Hydrogen Bond Acidity and Basicity. *The Journal of Organic Chemistry* **2004**, *69*, 4677-4685.
- (69) Goss, K.-U. Prediction of the temperature dependency of Henry's law constant using poly-parameter linear free energy relationships. *Chemosphere* **2006**, *64*, 1369-1374.

- (70) Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. Determination of sets of solute descriptors from chromatographic measurements. *Journal of Chromatography A* **2004**, *1037*, 29-47.
- (71) UFZ-LSER database v 2.1, Endo, S., Watanabe, N., Ulrich, N., Bronner, G., Goss, K.-U., **2014**
- (72) Zhao, Y.; Truhlar, D. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Account* **2008**, *120*, 215-241.
- (73) Plumley, J. A.; Dannenberg, J. J. A Comparison of the Behavior of Functional/Basis Set Combinations for Hydrogen-Bonding in the Water Dimer with Emphasis on Basis Set Superposition Error. *Journal of Computational Chemistry* **2011**, *32*, 1519-1527.
- (74) Thompson, M. ArgusLab. 4.0.1 ed., **2004**.
- (75) Cramer, C. J.; Truhlar, D. G. AM1-SM2 and PM3-SM3 Parameterized SCF Solvation Models for Free-Energies in Aqueous-Solution. *Journal of Computer-Aided Molecular Design* **1992**, *6*, 629-666.
- (76) Yu, H.-Z.; Yang, Y.-M.; Zhang, L.; Dang, Z.-M.; Hu, G.-H. Quantum-Chemical Predictions of p K a's of Thiols in DMSO. *The Journal of Physical Chemistry A* **2014**, *118*, 606-622.
- (77) Miguel, E. L.; Silva, P. L.; Pliego, J. R. Theoretical Prediction of pKa in Methanol: Testing SM8 and SMD Models for Carboxylic Acids, Phenols and Amines. *The Journal of Physical Chemistry B* **2014**.

- (78) Klicic, J. J.; Friesner, R. A.; Liu, S. Y.; Guida, W. C. Accurate prediction of acidity constants in aqueous solution via density functional theory and self-consistent reaction field methods. *Journal of Physical Chemistry A* **2002**, *106*, 1327-1335.
- (79) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, **2013**.
- (80) Franco, A.; Trapp, S. Estimation of the soil-water partition coefficient normalized to organic carbon for ionizable organic chemicals. *Environmental Toxicology and Chemistry* **2008**, *27*, 1995-2004.
- (81) Davis, C. W.; Di Toro, D. M. Modeling Nonlinear Adsorption to Carbon with a Single Chemical Parameter: A Lognormal Langmuir Isotherm. *Environmental Science & Technology* **2015**, *49*, 7810-7817.
- (82) Davis, C. W.; Di Toro, D. M. Modeling Nonlinear Adsorption with a Single Chemical Parameter: Predicting Chemical Median Langmuir Binding Constants. *Environmental Science & Technology* **2015**, *49*, 7818-7824.
- (83) Zhao, Y. H.; Abraham, M. H.; Zissimos, A. M. Determination of McGowan Volumes for Ions and Correlation with van der Waals Volumes. *Journal of Chemical Information and Computer Sciences* **2003**, *43*, 1848-1854.
- (84) Abraham, M. H.; Whiting, G. S.; Doherty, R. M.; Shuely, W. J. Hydrogen bonding. Part 13. A new method for the characterisation of GLC stationary phases—the laffort data set. *Journal of the Chemical Society, Perkin Transactions 2* **1990**, 1451-1460.

- (85) Poole, C. F.; Atapattu, S. N.; Poole, S. K.; Bell, A. K. Determination of solute descriptors by chromatographic methods. *Analytica Chimica Acta* **2009**, *652*, 32-53.
- (86) Haynes, W. M. *CRC handbook of chemistry and physics*; CRC press, **2014**.
- (87) Hickey, A. L.; Rowley, C. N. Benchmarking Quantum Chemical Methods for the Calculation of Molecular Dipole Moments and Polarizabilities. *The Journal of Physical Chemistry A* **2014**, *118*, 3678-3687.
- (88) Cramer, C. J.; Truhlar, D. G. PM3-SM3 - A General Parameterization for Including Aqueous Solvation Effects in the PM3 Molecular-Orbital Model. *Journal of Computational Chemistry* **1992**, *13*, 1089-1097.
- (89) Anscombe, F. J.; Tukey, J. W. The Examination and Analysis of Residuals. *Technometrics* **1963**, *5*, 141-160.

Appendix A

MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: A LOG-NORMAL LANGMUIR ISOTHERM

A.1. Log-normal Langmuir Visual Basic Code

A.2. Log-normal Langmuir Estimated Isotherm Parameters

- a. Graphite/Charcoal
- b. Darco GAC

A.3. Log-normal Langmuir Estimated Isotherm Parameters – F400 GAC

- a. Sorbate-specific $\log(\tilde{K}_L)$, σ_κ , and q_{max} – Model (1)
- b. Sorbate-specific $\log(\tilde{K}_L)$ and q_{max} – Model (2)
- c. Sorbate-specific $\log(\tilde{K}_L)$ and σ_κ – Model (3)
- d. Sorbate-specific $\log(\tilde{K}_L)$ – Model (4)

A.4. Discussion of Standard Errors of Estimated Parameters – Model (4)

A.5. Estimated ss-LNL Isotherms for F400 GAC

A.6. Comparison of Freundlich and LNL Isotherm Residuals

A.7. Tables of Estimated Freundlich Isotherm Parameters and Associated Standard Errors

A.1 Log-normal Langmuir Visual Basic Code

```
Function LNLfunc(k, qtot, sigma, conc)
' k is  $\ln(\tilde{K}_L)$  with units [mM]-1
' qtot is the maximum sorption capacity with units mmol/kg
' sigma is the standard deviation of k, dimensionless
' conc is the aqueous concentration in mM
Pi = 3.141592654

Dim weights
weights = Array(2.22939364553415E-13, 4.39934099227318E-10,
1.08606937076928E-07, 7.80255647853206E-06, 2.28338636016354E-04,
3.24377334223786E-03, 2.48105208874636E-02, 0.109017206020023,
0.286675505362834, 0.46224366960061, 0.46224366960061,
0.286675505362834, 0.109017206020023, 2.48105208874636E-02,
3.24377334223786E-03, 2.28338636016354E-04, 7.80255647853206E-06,
1.08606937076928E-07, 4.39934099227318E-10, 2.22939364553415E-13)

Dim nodes
nodes = Array(-5.38748089001123, -4.60368244955074, -3.94476404011563, -
3.34785456738322, -2.78880605842813, -2.25497400208928, -
1.73853771211659, -1.23407621539532, -0.737473728545394, -
0.245340708300901, 0.245340708300901, 0.737473728545394,
1.23407621539532, 1.73853771211659, 2.25497400208928, 2.78880605842813,
3.34785456738322, 3.94476404011563, 4.60368244955074, 5.38748089001123)

Dim func(0 To 19)

For i = 0 To 19
func(i) = ((conc * Exp(k + sigma * nodes(i) * (Sqr(2)))) / (1 + conc * Exp(k +
sigma * nodes(i) * (Sqr(2)))))
Next i

Dim runsum
runsum = 0

For i = 0 To 19
runsum = runsum + func(i) * weights(i)
Next i
```


A.2 Log-normal Langmuir Fitted Isotherm Parameters

Table A-1a. Estimated LNL isotherm parameters and associated standard errors for graphite and charcoal. Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless. Coefficients of variation (CV) are expressed as a % of the estimated parameter.

Chemical	Graphite		Charcoal		
	$\log(\tilde{K}_L)$	SE	$\log(\tilde{K}_L)$	SE	
Phenanthrene	4.078	0.0221	-	-	
2,4,6-Trinitrotoluene	2.544	0.0179	1.555	0.0162	
Naphthalene	1.247	0.0177	1.993	0.0134	
1,2,4-Trichlorobenzene	1.313	0.0177	2.227	0.0188	
2,4-Dinitrotoluene	1.241	0.0171	1.801	0.0120	
1,2,4-Trimethylbenzene	0.883	0.0167	1.495	0.0223	
p-Nitrotoluene	0.651	0.0155	1.710	0.0168	
1,2-Dichlorobenzene	0.439	0.0170	1.566	0.0184	
Xylene	0.404	0.0146	1.285	0.0146	
Chlorobenzene	-0.431	0.0140	-	-	
Toluene	-0.701	0.0141	0.747	0.0130	
Benzonitrile	-0.886	0.0172	0.977	0.0148	
Benzene	-1.390	0.0157	-0.137	0.0140	
σ_κ	SE(σ_κ)	CV (%)	σ_κ	SE(σ_κ)	CV (%)
3.481	0.292	8.39%	2.6098	0.107	4.10%
q_{max}	SE(q_{max})	CV (%)	q_{max}	SE(q_{max})	CV (%)
15.99	3.932	24.6%	978.4	79.75	8.15%

Table A-1b. Estimated LNL isotherm parameters and associated standard errors for Darco GAC. Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless. Coefficients of variation (CV) are expressed as a % of the estimated parameter.

Chemical	$\log(\tilde{K}_L)$	SE
Hexane	0.1402	0.0437
Heptane	0.261	0.0300
Benzene	-0.705	0.0624
Toluene	-0.0599	0.0419
Nitrobenzene	-0.333	0.0571
Trichloroethene	-0.726	0.0322
1,1,2-Tetrachloroethane	-1.360	0.0545
Methyl tert-butyl Ether	-2.491	0.0330
Diethyl Ether	-2.382	0.0256
Diisopropyl Ether	-1.556	0.0610
2-Heptanone	0.136	0.0746
3-Hexanone	-1.695	0.0448
1-Heptanol	-0.984	0.0711
1-Hexanol	-1.618	0.0485
σ_κ	$SE(\sigma_\kappa)$	CV (%)
3.969	0.294	7.41%
q_{max}	$SE(q_{max})$	CV (%)
5064.	1255.	24.8%

A.3 Log-normal Langmuir Estimated Isotherm Parameters – F400 GAC.

Table A-2a. Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (1). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless.

Chemical	$\log(\tilde{K}_L)$	SE	σ_κ	SE	$\log(q_{max})$	SE
1,1,1,2-tetrachloroethane	-1.645	2.242	4.376	1.029	4.068	0.717
1,1,1-trichloroethane	-2.951	1.374	5.112	0.721	3.775	0.376
1,1-Dichloroethane	-4.885	1.302	4.777	0.402	4.904	0.580
1,1-Dichloroethene	-4.113	1.366	5.658	0.701	4.249	0.341
1,1-Dichloropropene	-1.916	4.038	6.126	4.360	3.568	0.349
1,2,3-trichloropropane	-2.122	2.149	4.433	0.974	4.362	0.682
1,2-dibromoethane	-3.552	0.809	6.349	0.366	3.658	0.230
1,2-dichlorobenzene	0.142	2.781	4.517	1.935	4.007	0.647
1,2-Dichloroethane	-2.953	3.124	5.071	2.018	3.443	0.634
1,2-Dichloropropane	-4.008	2.042	5.438	1.101	4.247	0.389
1,3,5-trichlorobenzene	0.973	1.643	4.682	1.271	4.039	0.346
1,3-dichlorobenzene	-0.216	6.415	3.586	3.256	4.440	2.268
1,3-Dichloropropane	-3.734	0.876	5.654	0.454	4.283	0.205
2,4,5 Trichlorophenoxyacetic Acid	-0.216	3.386	7.712	6.336	3.489	0.143
Alachlor	0.001	1.102	7.002	0.998	3.858	0.158
Aldicarb	-0.526	1.206	5.253	0.691	3.685	0.295
Atrazine	0.644	1.001	5.620	1.040	3.676	0.171
Benzene	-2.758	1.875	5.091	0.846	4.372	0.525
Bromobenzene	0.027	2.203	4.583	1.640	3.940	0.481
Bromodichloromethane	-2.506	3.799	4.330	1.599	4.067	1.326
Bromoform	-2.980	1.514	4.825	0.671	4.524	0.469
Carbofuran	-0.068	2.548	5.516	2.244	3.688	0.379
Carbon tetrachloride	-2.355	1.759	4.625	0.712	3.878	0.615
Chlorobenzene	-0.911	2.218	5.343	1.915	3.887	0.350
Chloroform	-5.082	0.989	5.375	1.073	4.541	1.051
cis 1,2-dichloroethylene	-2.301	1.080	4.581	0.493	3.663	0.354
Cyanazine	0.011	1.466	7.994	2.280	3.616	0.110
Dibromochloromethane	-1.753	4.326	4.335	1.730	3.893	1.578
Dibromochloropropane	-1.030	1.118	5.216	0.608	3.885	0.284
Dibromomethane	-5.034	1.567	4.832	0.556	4.923	0.579

Ethyl Benzene	-0.135	6.354	4.235	4.595	4.021	1.503
Hexachlorocyclopentadiene	0.097	3.589	4.380	1.841	4.464	1.067
Isophorone	-1.046	1.275	6.662	1.157	3.492	0.179
Methylene Chloride	-5.765	2.431	4.628	0.867	4.934	0.904
Metolachlor	-0.011	1.258	8.865	2.245	3.502	0.066
Metribuzin	-0.266	2.247	7.303	3.934	3.606	0.129
Oxamyl	-0.043	1.691	2.473	0.946	4.767	0.744
Simazine	0.675	0.833	8.320	0.811	3.180	0.128
Styrene	0.370	7.492	3.851	5.201	4.146	1.983
Tert-butyl methyl ether	-2.573	6.463	4.939	4.253	3.438	1.225
Tetrachloroethene	-1.114	1.566	4.599	0.804	4.088	0.444
Toluene	0.294	2.567	4.520	1.750	3.558	0.604
Trans 1,2 Dichloroethene	-2.571	2.136	5.432	1.430	3.606	0.369
Trichloroethene	-0.597	2.067	4.382	1.101	3.605	0.590

Table A-2b. Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (2). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless.

Chemical	$\log(\tilde{K}_L)$	SE	$\log(q_{max})$	SE
1,1,1,2-tetrachloroethane	-4.070	0.387	4.958	0.197
1,1,1-trichloroethane	-3.064	0.574	3.787	0.279
1,1-Dichloroethane	-6.033	0.357	5.285	0.184
1,1-Dichloroethene	-3.213	0.809	4.050	0.429
1,1-Dichloropropene	0.172	1.378	3.091	0.421
1,2,3-trichloropropane	-3.735	0.445	4.803	0.227
1,2-dibromoethane	-0.563	0.368	2.839	0.148
1,2-dichlorobenzene	-0.613	0.359	4.133	0.124
1,2-Dichloroethane	-3.100	0.634	3.454	0.319
1,2-Dichloropropane	-2.260	0.371	3.448	0.174
1,3,5-trichlorobenzene	0.502	0.452	4.106	0.138
1,3-dichlorobenzene	-5.065	0.447	6.414	0.263
1,3-Dichloropropane	-2.919	0.350	4.120	0.147
2,4,5 Trichlorophenoxy Acetic Acid	1.746	0.647	3.264	0.140
Alachlor	2.270	0.467	3.510	0.116
Aldicarb	-0.435	0.469	3.666	0.186
Atrazine	0.944	0.461	3.645	0.133
Benzene	-2.943	0.496	4.408	0.239
Bromobenzene	-0.718	0.412	4.082	0.139
Bromodichloromethane	-4.408	0.696	4.605	0.425
Bromoform	-3.615	0.470	4.621	0.268
Carbofuran	-1.187	0.445	4.262	0.202
Carbon tetrachloride	-3.028	0.790	3.913	0.407
Chlorobenzene	-0.776	0.382	3.873	0.134
Chloroform	-2.985	1.371	3.451	0.740
cis 1,2-dichloroethylene	-4.232	0.306	4.381	0.141
Cyanazine	2.748	0.680	3.265	0.099
Dibromochloromethane	-4.544	0.961	5.014	0.619
Dibromochloropropane	-1.008	0.391	3.883	0.176
Dibromomethane	-5.965	0.431	5.205	0.249
Ethyl Benzene	-1.424	0.967	4.284	0.393
Hexachlorocyclopentadiene	-2.229	1.286	5.264	0.679
Isophorone	1.410	0.515	3.028	0.144
Methylene Chloride	-6.980	0.782	5.176	0.548
Metolachlor	3.741	0.723	3.060	0.080
Metribuzin	1.864	0.448	3.313	0.080
Oxamyl	-7.891	0.562	8.263	0.362
Simazine	4.129	0.400	2.732	0.087
Styrene	-5.120	1.206	6.550	0.779
Tert-butyl methyl ether	-1.822	1.852	3.018	0.785

Tetrachloroethene	-2.321	0.334	4.410	0.149
Toluene	-0.312	0.459	3.612	0.175
Trans 1,2 Dichloroethene	-2.224	0.357	3.548	0.152
Trichloroethene	-2.269	0.347	4.076	0.154
		σ_{κ}	SE (σ_{κ})	
		5.197	0.108	

Table A-2c. Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (3). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless.

Chemical	$\log(\tilde{K}_L)$	SE	σ_κ	SE
1,1,1,2-tetrachloroethane	-1.442	0.259	4.282	0.249
1,1,1-trichloroethane	-3.549	0.481	5.265	0.419
1,1-Dichloroethane	-2.632	0.287	4.001	0.247
1,1-Dichloroethene	-3.309	0.629	5.360	0.558
1,1-Dichloropropene	-2.785	0.644	5.766	0.707
1,2,3-trichloropropane	-1.003	0.216	3.905	0.211
1,2-dibromoethane	-4.800	0.412	6.811	0.319
1,2-dichlorobenzene	0.037	0.286	4.704	0.441
1,2-Dichloroethane	-4.174	0.696	5.236	0.588
1,2-Dichloropropane	-2.893	0.322	4.864	0.263
1,3,5-trichlorobenzene	1.131	0.272	4.582	0.509
1,3-dichlorobenzene	1.035	0.226	2.864	0.296
1,3-Dichloropropane	-2.597	0.259	5.109	0.247
2,4,5 Trichlorophenoxyacetic Acid	-5.331	12.538	11.525	16.528
Alachlor	-0.804	0.460	7.512	0.654
Aldicarb	-1.980	0.454	6.048	0.456
Atrazine	-0.850	0.457	6.207	0.648
Benzene	-1.467	0.350	4.503	0.351
Bromobenzene	-0.227	0.277	4.738	0.437
Bromodichloromethane	-2.326	0.486	4.257	0.427
Bromoform	-1.393	0.327	4.155	0.285
Carbofuran	-0.350	0.483	4.880	0.563
Carbon tetrachloride	-2.676	0.359	4.726	0.321
Chlorobenzene	-2.275	1.413	6.797	1.888
Chloroform	-3.035	0.613	4.416	0.505
cis 1,2-dichloroethylene	-3.209	0.269	4.883	0.228
Cyanazine	-2.954	1.226	9.894	2.713
Dibromochloromethane	-2.034	0.605	4.434	0.519
Dibromochloropropane	-1.417	0.408	5.367	0.383
Dibromomethane	-2.597	0.274	3.937	0.229
Ethyl Benzene	0.109	0.486	3.908	0.787
Hexachlorocyclopentadiene	1.744	0.319	3.332	0.535
Isophorone	-3.499	0.545	7.573	0.607
Methylene Chloride	-3.489	0.335	3.834	0.255
Metolachlor	-3.857	1.415	10.486	1.860

Metribuzin	-2.320	0.404	7.765	0.588
Oxamyl	1.665	0.087	1.212	0.182
Simazine	-3.826	0.526	9.889	0.546
Styrene	0.866	0.404	3.531	0.674
Tert-butyl methyl ether	-4.480	1.314	5.730	1.217
Tetrachloroethene	-0.829	0.234	4.464	0.245
Toluene	-1.233	0.437	5.167	0.479
Trans 1,2 Dichloroethene	-4.265	0.998	6.261	0.936
Trichloroethene	-1.985	0.396	5.056	0.396

$\log(q_{max})$	$SE(\log(q_{max}))$
4.002	0.025

Table A-2d. Estimated LNL isotherm parameters and associated standard errors for F400 GAC – Model (4). Units for q_{max} and $\tilde{K}_L$ are (mmol/kg carbon) and (mM)⁻¹, respectively, and σ_κ is dimensionless.

Chemical	$\log(\tilde{K}_L)$	SE
1,1,1,2-tetrachloroethane	-2.022	0.2333
1,1,1-trichloroethane	-3.098	0.2205
1,1-Dichloroethane	-3.641	0.3515
1,1-Dichloroethene	-2.757	0.1622
1,1-Dichloropropene	-1.968	0.1861
1,2,3-trichloropropane	-1.933	0.2321
1,2-dibromoethane	-2.443	0.2883
1,2-dichlorobenzene	-0.034	0.1691
1,2-Dichloroethane	-3.744	0.2035
1,2-Dichloropropane	-2.900	0.3793
1,3,5-trichlorobenzene	1.058	0.1418
1,3-dichlorobenzene	-0.481	0.2012
1,3-Dichloropropane	-2.344	0.2390
2,4,5 Trichlorophenoxy Acetic Acid	-0.495	0.1684
Alachlor	0.964	0.1278
Aldicarb	-0.848	0.2009
Atrazine	0.085	0.1723
Benzene	-1.811	0.1931
Bromobenzene	-0.253	0.1536
Bromodichloromethane	-3.021	0.2060
Bromoform	-2.208	0.2763
Carbofuran	-0.307	0.1651
Carbon tetrachloride	-2.827	0.2811
Chlorobenzene	-0.841	0.1896
Chloroform	-3.590	0.2699
cis 1,2-dichloroethylene	-3.190	0.4160
Cyanazine	-0.272	0.1561
Dibromochloromethane	-2.540	0.1842
Dibromochloropropane	-0.890	0.2592
Dibromomethane	-3.718	0.3387
Ethyl Benzene	-0.440	0.1121
Hexachlorocyclopentadiene	0.896	0.1169
Isophorone	-1.020	0.1789
Methylene Chloride	-4.890	0.3528
Metolachlor	-0.281	0.1669

Metribuzin	-0.546	0.2046
Oxamyl	-0.296	0.1569
Simazine	0.612	0.1531
Styrene	0.112	0.1179
Tert-butyl methyl ether	-3.543	0.2359
Tetrachloroethene	-1.177	0.3058
Toluene	-0.947	0.1904
Trans 1,2 Dichloroethene	-2.772	0.2649
Trichloroethene	-1.791	0.2497

q_{max}	SE(q_{max})	CV(%)
10199.	120.4	1.18%
σ_{κ}	SE(σ_{κ})	CV(%)
4.962	0.009807	0.198%

A.4 Discussion of Standard Errors of Estimated Parameters – Model (4)

Initially the range of standard errors for the predicted $\log(\tilde{K}_L)$ were found to be nearly identical within each data set, which was not realistic. It was determined that this was due to the large inter-correlation between σ_κ and the individual chemical $\log(\tilde{K}_L)$ values. The standard errors were recalculated by re-fitting the $\log(\tilde{K}_L)$ parameters while holding σ_κ and q_{max} constant.

The graphite standard errors were approximately 0.3 log-units larger than those for charcoal. This is likely because powdered graphite is more homogeneous and lacks the broad site distribution of other sorbents. In the model, the only parameter which is chemical-specific is the $\log(\tilde{K}_L)$. Consequently if the isotherm slopes are very close, it is more difficult for the model to accurately predict these binding constants when they occur over significantly different concentration ranges.

The range of standard errors for the predicted binding constants to Darco GAC are similar to that of graphite. However, this is more likely due to variation in the sorption data, as the overall RMSE for the Darco data set was larger. It can be seen that there is more variability in the Darco GAC adsorption data (Figure 2-1C) than in the graphite data (Figure 2-1A).

Finally, the RMS error for the larger F400 GAC data set (0.113) was similar to that of the Darco GAC set (0.121). However, the range of RMS errors for the individual isotherm fits was slightly larger (0.0258 – 0.441). This was due primarily to several outlier pesticides (oxamyl and simazine) that had uncharacteristically large residual errors, which is discussed below. When these chemicals were excluded, the range of individual predicted isotherm RMS errors was comparable to those for graphite and charcoal (RMSE = 0.0258 – 0.156). Further, the range of standard errors

for the predicted $\log(\tilde{K}_L)$ values fell between those of the graphite and charcoal and the Darco GAC sets.

A.5 Estimated ss-LNL Isotherms for F400 GAC

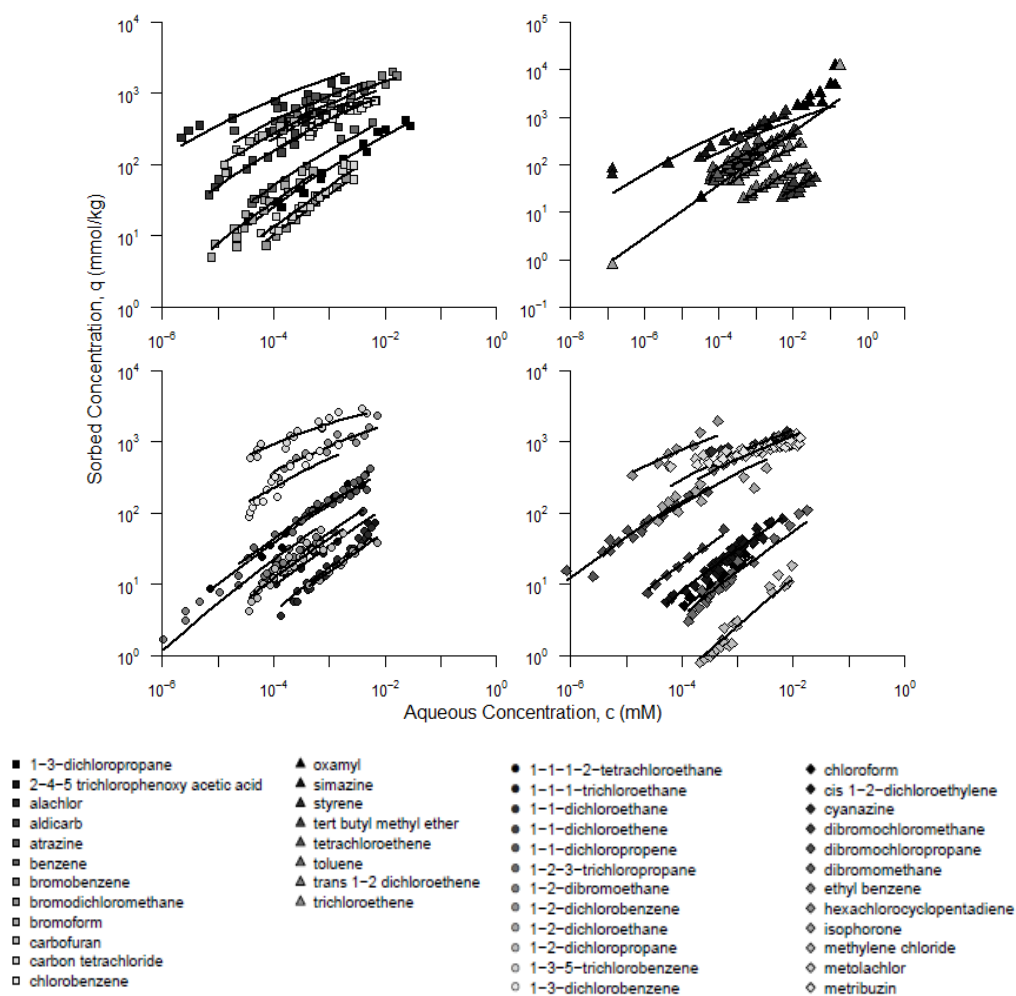


Figure A-1. Adsorption isotherms for the single sorbate-specific parameter LNL (ss-LNL) isotherm model. Points represent experimental adsorption data, solid lines represent the ss-LNL model. Estimated LNL parameters are presented in Table A-2-d.

A.6 Comparison of Freundlich and LNL Isotherm Residuals

Since the conventional Freundlich non-linear isotherm is common in adsorption isotherm modeling, the residuals of the single-chemical parameter LNL isotherm (Model (4)) were compared to those obtained from a traditional 2-parameter Freundlich model. The results for the four data sets are summarized in the box plots below.

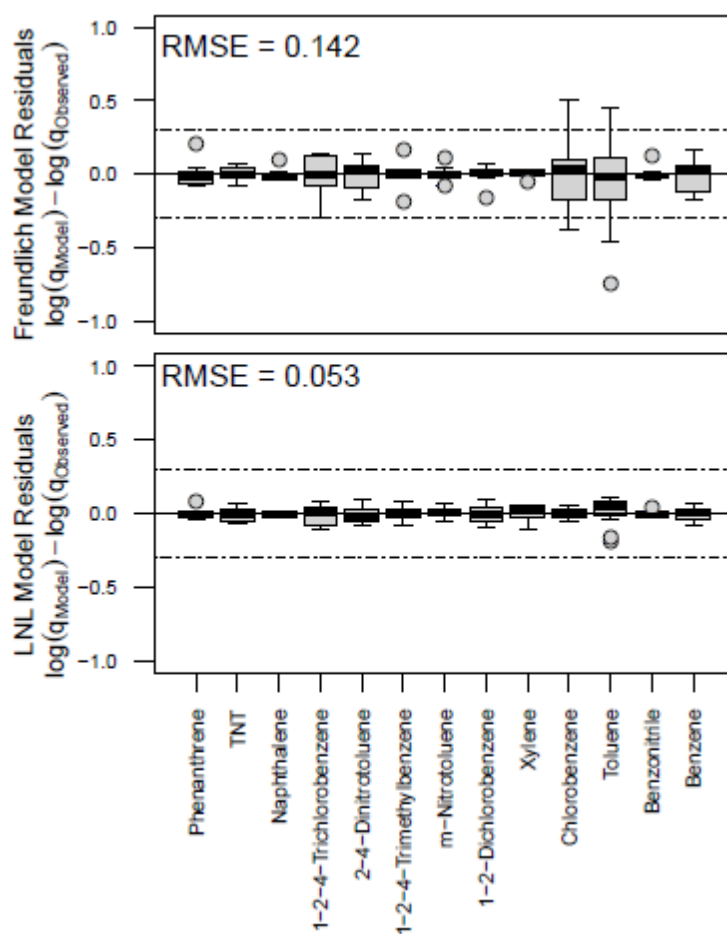


Figure A-2a. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Graphite. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.

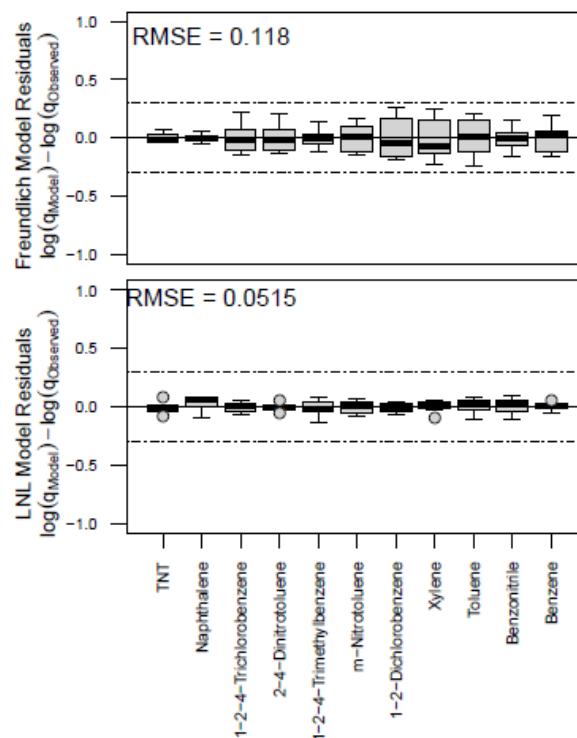


Figure A-2b. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Charcoal. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.

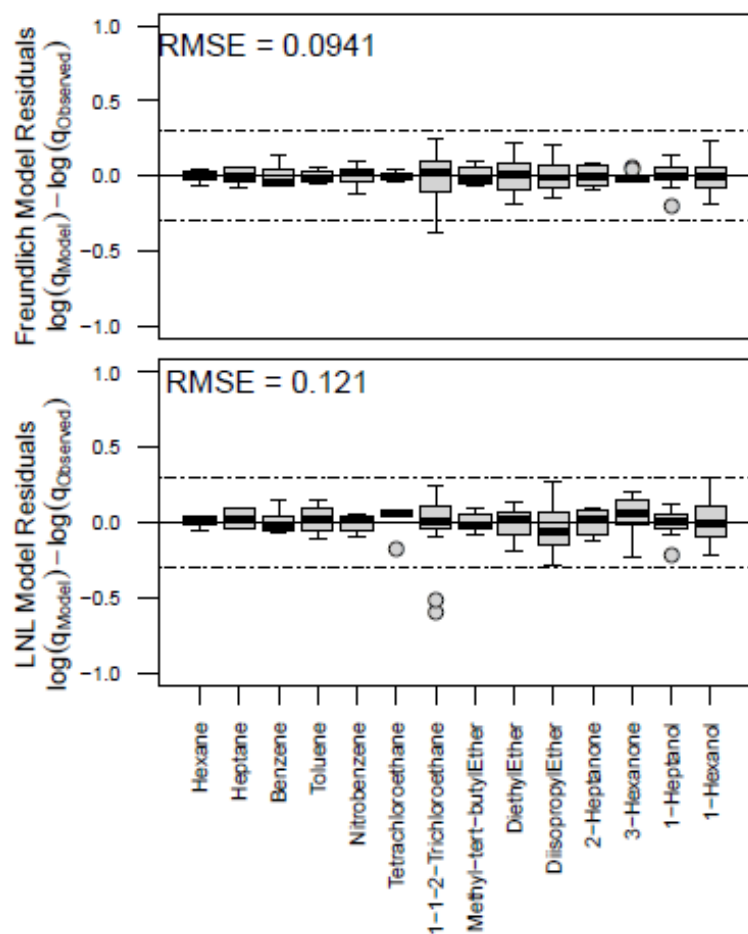


Figure A-2c. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on Darco GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.

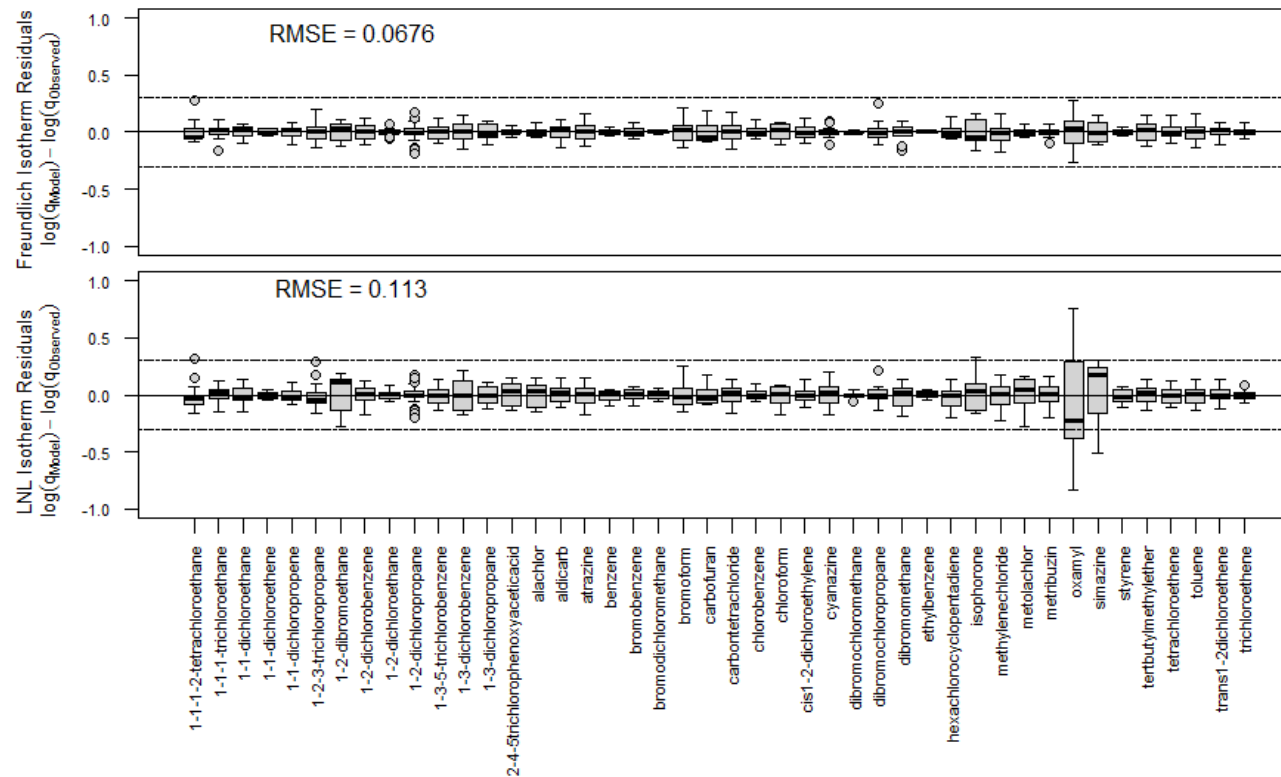


Figure A-2d. Box plots comparing the residuals ($\log(q_{Model}) - \log(q_{Observed})$) for the Freundlich (top) and LNL (bottom) models on F400 GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences. Solid line represents 1:1 agreement between the model and observed sorbed concentrations.

A.7 Tables of Estimated Freundlich Isotherm Parameters and Associated Standard Errors

Table A-3a. Estimated Freundlich parameters and standard errors for graphite. Units of Freundlich K_f are $(\frac{mmol}{kg})(\frac{mmol}{L})^{1/v}$. v is dimensionless.

Chemical	$\log(K_f)$	SE	v	SE
Phenanthrene	2.2889	0.0738	0.3299	0.0138
TNT	1.8101	0.0566	0.3441	0.0139
Naphthalene	1.8519	0.0728	0.4926	0.0166
1-2-4-Trichlorobenzene	2.2610	0.1665	0.5756	0.0404
2-4-Dinitrotoluene	2.0195	0.0431	0.5412	0.0133
1-2-4-Trimethylbenzene	1.9243	0.0454	0.5619	0.0143
m-Nitrotoluene	1.9948	0.0345	0.6228	0.0113
1-2-Dichlorobenzene	1.6632	0.0589	0.5461	0.0172
Xylene	1.3205	0.0905	0.4389	0.0314
Chlorobenzene	1.1311	0.0339	0.5249	0.0140
Toluene	1.0840	0.0484	0.5632	0.0208
Benzonitrile	1.0170	0.0364	0.6094	0.0239
Benzene	0.7043	0.0359	0.6062	0.0352

Table A-3b. Estimated Freundlich parameters and standard errors for charcoal. Units of Freundlich K_f are $(\frac{mmol}{kg})(\frac{mmol}{L})^{1/v}$. v is dimensionless.

Chemical	$\log(K_f)$	SE	v	SE
TNT	3.6935	0.3371	0.5548	0.0836
Naphthalene	4.0848	0.2885	0.5951	0.0630
1-2-4-Trichlorobenzene	4.1859	0.1291	0.5575	0.0338
2-4-Dinitrotoluene	4.3447	0.1652	0.6819	0.0356
1-2-4-Trimethylbenzene	3.1103	0.0387	0.3399	0.0161
m-Nitrotoluene	3.3902	0.1446	0.4480	0.0452
1-2-Dichlorobenzene	3.3245	0.0991	0.4342	0.0355
Xylene	3.2750	0.1042	0.4658	0.0339
Toluene	3.0493	0.0605	0.4815	0.0229
Benzonitrile	3.1671	0.0985	0.4866	0.0334
Benzene	2.8506	0.0543	0.5906	0.0253

Table A-3c. Estimated Freundlich parameters and standard errors for Darco GAC.
Units of Freundlich K_f are $(\frac{mmol}{kg})(\frac{mmol}{L})^{1/v}$. v is dimensionless.

Chemical	$\log(K_f)$	SE	v	SE
Hexane	3.8447	0.0620	0.4854	0.0291
Heptane	4.0339	0.0850	0.5051	0.0253
Benzene	3.1592	0.0399	0.2832	0.0225
Toluene	3.3863	0.0399	0.2859	0.0245
Nitrobenzene	3.1732	0.0141	0.1864	0.0093
Tetrachloroethane	3.4011	0.0477	0.4458	0.0195
1-1-2-Trichloroethane	3.1749	0.0844	0.4433	0.0524
Methyl-tert-butylEther	2.9745	0.1151	0.6635	0.0393
DiethylEther	2.7675	0.0520	0.5761	0.0280
DiisopropylEther	3.0575	0.0657	0.4268	0.1151
2-Heptanone	3.3323	0.0451	0.2257	0.0229
3-Hexanone	2.8278	0.0508	0.1275	0.0630
1-Heptanol	3.1507	0.0243	0.2505	0.0595
1-Hexanol	3.0088	0.0363	0.3959	0.0460

Table A-3d. Estimated Freundlich parameters and standard errors for F400 GAC.

Units of Freundlich K_f are $(\frac{mmol}{kg})(\frac{mmol}{L})^{1/v}$. v is dimensionless.

Chemical	$\log(K_f)$	SE	v	SE
1-1-1-2-Tetrachloroethane	3.9569	0.1354	0.6036	0.0377
1-1-1-Trichloroethane	3.1243	0.1253	0.5314	0.0393
1-1-Dichloroethane	3.3385	0.0660	0.7053	0.0219
1-1-Dichloroethene	3.2541	0.0621	0.5151	0.0182
1-1-Dichloropropene	3.2698	0.1254	0.3743	0.0438
1-2-3-Trichloropropane	4.0329	0.1045	0.6129	0.0293
1-2-Dibromoethane	3.1563	0.1002	0.4706	0.0231
1-2-Dichlorobenzene	4.0713	0.0972	0.3781	0.0308
1-2-Dichloroethane	2.7783	0.1030	0.5329	0.0371
1-2-Dichloropropane	3.4580	0.1322	0.5969	0.0351
1-3-5-Trichlorobenzene	4.2477	0.0959	0.3237	0.0271
1-3-Dichlorobenzene	4.8575	0.2000	0.6296	0.0508
1-3-Dichloropropane	3.4114	0.0593	0.4971	0.0191
2-4-5 Trichlorophenoxy Acetic Acid	3.3613	0.0393	0.2098	0.0126
Alachlor	3.8762	0.0585	0.2569	0.0134
Aldicarb	3.7607	0.1222	0.4021	0.0288
Atrazine	3.8046	0.1240	0.2907	0.0354
Benzene	3.8157	0.0394	0.5334	0.0109
Bromobenzene	3.9298	0.0526	0.3637	0.0188
Bromodichloromethane	3.5843	0.0290	0.6553	0.0084
Bromoform	4.1586	0.1865	0.6650	0.0436
Carbofuran	4.0518	0.2878	0.4083	0.0669
Carbon tetrachloride	3.4786	0.1301	0.5935	0.0387
Chlorobenzene	3.6679	0.0696	0.3475	0.0233
Chloroform	3.2857	0.1799	0.6690	0.0544
Cis 1-2-dichloroethylene	3.2469	0.0620	0.5871	0.0193
Cyanazine	3.3050	0.0665	0.1257	0.0256
Dibromochloromethane	3.8308	0.0506	0.6360	0.0125
Dibromochloropropane	4.1560	0.1592	0.5006	0.0329
Dibromomethane	3.2885	0.0730	0.7004	0.0230
Ethyl Benzene	4.0285	0.0163	0.4153	0.0056
Hexachlorocyclopentadiene	4.9363	0.2333	0.5039	0.0566
Isophorone	3.2434	0.2036	0.2714	0.0546
Methylene Chloride	2.8162	0.1212	0.8012	0.0402
Metolachlor	3.2211	0.0455	0.1251	0.0155

Metribuzin	3.3850	0.0351	0.1929	0.0135
Oxamyl	5.1485	0.1218	0.7931	0.0420
Simazine	3.3975	0.1835	0.2274	0.0339
Styrene	4.4720	0.0940	0.4793	0.0277
Tert-butyl-methyl Ether	2.7616	0.2497	0.4788	0.0971
Tetrachloroethene	4.0814	0.0713	0.5160	0.0188
Toluene	3.8545	0.1950	0.4266	0.0503
Trans 1-2 Dichloroethene	3.0589	0.0876	0.4521	0.0276
Trichloroethene	3.6518	0.0597	0.4826	0.0169

Appendix B

MODELING NON-LINEAR ADSORPTION WITH A SINGLE CHEMICAL PARAMETER: PREDICTING CHEMICAL MEDIAN LANGMUIR BINDING CONSTANTS

B.1. Abraham Solute Parameters, Aqueous Solubility, and Predicted vs. LNL-

Estimated $\log(\tilde{K}_L)$

B.2. Estimated Abraham Sorbent-water Parameters (with Polarizability) for Graphite, Charcoal, Darco GAC, and F400 GAC

B.3. Comparison of Abraham Sorbent-water Parameters (with and without Polarizability, sS, Term) for F400 GAC

B.4. Calculation of Cook's Distance for Oxamyl in Solubility vs. $\log(\tilde{K}_L)$ Regression.

B.5. Comparison of Abraham Sorbent-water Parameters (with and without Hydrogen-bond Basicity, bB, Term) for F400 GAC sorbates with $V < 110$.

B.6. Comparison of LNL pp-LFER vs. Shih & Gschwend pp-LFER Predicted Isotherm Residuals.

B.1 Abraham Solute Parameters, Aqueous Solubility, and Predicted vs. LNL-Estimated $\log(\tilde{K}_L)$

Table B-1a. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL_estimated $\log(\tilde{K}_L)$ for F400 GAC. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$.

Chemical	E	S	A	B	V	Solubility	$\log(\tilde{K}_L)$ LNL-estimated	$\log(\tilde{K}_L)$ Predicted
Methylene Chloride	0.22	0.38	0.09	0	0.49		-5.015	-4.762
1,1-Dichloroethene	0.35	0.46	0	0.1	0.59		-2.876	-3.676
Cis 1,2-dichloroethylene	0.4	0.56	0	0.09	0.59		-3.310	-3.077
Trans 1,2 Dichloroethene	0.4	0.56	0	0.09	0.59		-2.891	-3.077
Dibromomethane	0.58	0.55	0.09	0	0.60		-3.838	-3.664
Chloroform	0.34	0.48	0.12	0	0.62		-3.713	-3.789
1,1-Dichloroethane	0.22	0.37	0.09	0.02	0.64		-3.763	-4.170
1,2-Dichloroethane	0.38	0.48	0	0.1	0.64		-3.864	-3.385
Bromodichloromethane	0.52	0.57	0.12	0.01	0.67		-3.142	-3.206
Trichloroethene	0.5	0.64	0	0.11	0.71		-1.907	-2.114
Benzene	0.56	0.69	0	0.12	0.72		-1.926	-1.858
Dibromochloromethane	0.7	0.65	0.12	0.01	0.72		-2.662	-2.690
1,1-Dichloropropene	0.38	0.5	0	0.11	0.73		-2.085	-2.792
Carbon tetrachloride	0.42	0.55	0	0	0.74		-2.946	-2.491
1,2-dibromoethane	0.74	0.64	0	0.1	0.74		-2.564	-2.354
1,1,1-trichloroethane	0.31	0.44	0	0.01	0.76		-3.218	-2.981
Bromoform	0.88	0.73	0.12	0.01	0.77		-2.328	-2.175
1,2-Dichloropropane	0.38	0.47	0	0.13	0.78		-3.021	-2.793
1,3-Dichloropropane	0.25	0.46	0	0.03	0.78		-2.462	-2.669

Tetrachloroethene	0.6	0.73	0	0.12	0.84		-1.293	-1.085
Chlorobenzene	0.7	0.77	0	0.11	0.84		-0.960	-0.957
Toluene	0.58	0.63	0	0.12	0.86		-1.063	-1.634
Tert butyl methyl ether	0.02	0.27	0	0.29	0.87		-3.661	-3.148
1,1,1,2-tetrachloroethane	0.62	0.61	0	0.12	0.88		-2.138	-1.723
Bromobenzene	0.89	0.85	0	0.11	0.89		-0.373	-0.457
1,2,3-trichloropropane	0.57	0.62	0	0.17	0.90		-2.052	-1.492
Styrene	0.7	0.7	0	0.17	0.96		-0.014	-0.883
1,2-dichlorobenzene	0.83	0.85	0	0.1	0.96		-0.132	-0.041
1,3-dichlorobenzene	0.83	0.84	0	0.05	0.96		-0.599	-0.110
Ethyl Benzene	0.58	0.64	0	0.12	1.00		-0.561	-0.908
Dibromochloropropane	0.67	0.61	0	0.02	1.00		-1.008	-1.224
1,3,5-trichlorobenzene	0.97	0.91	0	0.02	1.08	0.0331	0.931	0.550
Isophorone	0.54	0.76	0	0.45	1.24	86.8307	-1.137	-1.188
Hexachlorocyclopentadiene	1.12	0.96	0	0.2	1.35	0.0066	0.769	0.906
Simazine	1.25	1.25	0.36	0.86	1.48	0.0307	0.488	0.566
Aldicarb	0.78	0.91	0.21	0.87	1.49	31.6935	-0.964	-0.965
2,4,5 Trichlorophenoxy Acetic Acid	1.16	1.51	0.57	0.5	1.50	1.0881	-0.615	-0.221
Oxamyl	1.1	1.79	0.21	1.46	1.60	1277.02	-0.407	-1.781
Atrazine	1.26	1.24	0.36	0.89	1.62	0.161	-0.042	0.201
Metribuzin	1.46	1.21	0.21	1.53	1.62	4.900	-0.665	-0.553
Carbofuran	1.1	1.41	0.21	0.87	1.69	1.446	-0.435	-0.284
Cyanazine	1.41	1.74	0.36	1.07	1.77	0.706	-0.396	-0.126
Alachlor	1.11	1.63	0	0.94	2.14	0.890	0.838	-0.177
Metolachlor	1.12	1.62	0	0.98	2.28	1.868	-0.401	-0.340

$$V \leq 1.1 : \log(\tilde{K}_{L,F400}) = -1.503(E) + 6.744(S) - 3.138(A) + 4.680(V) - 9.025$$

$$V > 1.1 : \log(K_{L,F400}) = -0.508 * \log(S) - 0.202$$

Table B-1b. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL_{est} estimated $\log(\tilde{K}_L)$ for graphite and charcoal. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$.

Chemical	E	B	V	Graphite		Charcoal	
				$\log(\tilde{K}_L)$	$\log(\tilde{K}_L)$	$\log(\tilde{K}_L)$	$\log(\tilde{K}_L)$
				Estimated	Predicted	Estimated	Predicted
Phenanthrene	1.99	0.23	1.454	4.078	4.067		
2,4,6-Trinitrotoluene	1.39	0.41	1.380	2.544	2.420	1.555	1.993
Naphthalene	1.27	0.17	1.085	1.247	1.344	1.993	1.976
1,2,4-Trichlorobenzene	0.96	0.03	1.084	1.313	1.374	2.227	2.156
2,4-Dinitrotoluene	1.12	0.31	1.206	1.241	1.425	1.801	1.626
1,2,4-Trimethylbenzene	0.63	0.12	1.139	0.883	1.029	1.495	1.634
p-Nitrotoluene	0.85	0.21	1.032	0.651	0.431	1.710	1.259
1,2-Dichlorobenzene	0.83	0.1	0.961	0.439	0.338	1.566	1.454
Xylene	0.61	0.12	0.998	0.404	0.224	1.285	1.250
Chlorobenzene	0.70	0.11	0.839	-0.431	-0.524		
Toluene	0.58	0.12	0.857	-0.701	-0.592	0.747	0.856
Benzonitrile	0.73	0.26	0.871	-0.886	-0.746	0.977	0.542
Benzene	0.56	0.12	0.716	-1.390	-1.397	-0.137	0.473

$$\log(\tilde{K}_{L,graphite}) = 1.184(E) - 2.910(B) + 5.543(V) - 5.681$$

$$\log(\tilde{K}_{L,charcoal}) = 1.032(E) - 3.607(B) + 2.577(V) - 1.519$$

Table B-1c. Table of Abraham solute parameters, aqueous solubility (S), and predicted vs. LNL_estimated $\log(\tilde{K}_L)$ for Darco GAC. Units of $\tilde{K}_L$ and S are $(mM)^{-1}$.

Chemical	E	B	V	$\log(\tilde{K}_L)$ Estimated	$\log(\tilde{K}_L)$ Predicted
Hexane	0	0.04	0.954	0.140	-0.183
Heptane	0	0.05	1.095	0.261	0.477
Benzene	0.56	0.12	0.716	-0.705	-0.958
Toluene	0.58	0.12	0.857	-0.060	-0.203
Nitrobenzene	0.83	0.21	0.891	-0.333	-0.170
Tetrachloroethylene	0.5	0.11	0.715	-0.726	-1.006
1,1,2-Trichloroethane	0.5	0.11	0.758	-1.360	-0.786
Methyl t-butyl Ether	0.02	0.29	0.872	-2.491	-2.111
Diethyl Ether	0.04	0.25	0.731	-2.382	-2.553
Diisopropyl Ether	0.06	0.31	1.013	-1.556	-1.446
2-Heptanone	0.21	0.35	1.111	0.136	-0.941
3-Hexanone	0.21	0.35	0.970	-1.695	-1.663
1-Heptanol	0.2	0.32	1.154	-0.984	-0.553
1-Hexanol	0.2	0.32	1.013	-1.618	-1.274

$$\log(K_{L,Darco\ GAC}) = 1.669(E) - 6.163(B) + 5.122(V) - 4.823$$

B.2 Estimated Abraham Sorbent-water Parameters (with Polarizability) for Graphite, Charcoal, Darco GAC, and F400 GAC

¹ The number of sorbates utilized in the model regressions.

$$^2 RMSE = \sqrt{\frac{\sum (\log(K_{LPredicted}) - \log(K_{LLNL-estimated}))^2}{N}}$$

³The overall RMS error for the full F400 GAC data set, including both the sorbates above (Eq. (3-4)) and below (Eq. (3-2))

the volumetric cutoffs

B.3 Comparison of Abraham Sorbent-water Parameters (with and without Polarizability, sS, Term) for F400 GAC

B.4 Calculation of Cook's Distance for Oxamyl in Solubility vs. $\log(\tilde{K}_L)$ Regression.

To determine whether or not including oxamyl in fitting regression coefficients for the relationship between the log median binding constant, $\log(\tilde{K}_L)$ and the experimental aqueous solubility, $\log(S)$, the Cook's distance, D_C , was computed. The Cook's distance is defined as:

$$D_C = \frac{\sum_{j=1}^N (\hat{Y}_j - \hat{Y}_{j,i})^2}{pMSE} \quad (B-1)$$

where $\hat{Y}_j$ is the predicted logarithmic median binding constant of chemical "j" when chemical "i" is not excluded from the regression, $\hat{Y}_{j,i}$ is the predicted logarithmic median binding constant of chemical "j" when chemical "i" is excluded from the regression, p is the number of parameters in the model, and MSE is the mean square error (Eq. (2-15), squared) of the regression, finally, D_C is Cook's distance. A standard operating guideline is that if the Cook's distance is greater than $4/p$, then the data point "i" has significant leverage over the regression and may skew the resulting estimated parameters. For oxamyl the computed Cook's distance was 3.12, significantly larger than $4/p$ ($p=2$ for Eq. (3-4)). Consequently, oxamyl was not included in estimating the slope and intercept parameters for Eq. (3-4).

B.5 Comparison of Abraham Sorbent-water Parameters (with and without Hydrogen-bond Basicity, bB, Term) for F400 GAC sorbates with $V < 110$.

Table B-4. Comparison of reduced Abraham pp-LFER model for F400 GAC ($V < 110$), with (top) and without (bottom) the hydrogen bond basicity (bB) interaction term. Units correspond to $\tilde{K}_L$ with units $(mM)^{-1}$.

Sorbent	Subset	e	s	a	b	v	c	N	RMSE
<i>F400 GAC</i>	V < 110	-1.499	6.743	-3.114	0.038	4.674	-9.025	31	0.402
	SE	0.3065	0.4433	0.6008	0.3902	0.2055	0.1879		
	CV (%)	20.45%	6.57%	19.29%	1017.87%	4.40%	2.08%		
	V < 110	-1.503	6.744	-3.138	[-]	4.680	-9.025	31	0.402
	SE	0.2996	0.4370	0.5389	[-]	0.1922	0.1853		
	CV (%)	19.93%	6.48%	17.17%	[-]	4.11%	2.05%		

B.6 Comparison of LNL pp-LFER vs. Shih & Gschwend pp-LFER Predicted Isotherm Residuals.

Using the pp-LFER models presented in Shih & Gschwend³⁴, Freundlich isotherm parameters, K_F and ν , were calculated. The pp-LFER for the Freundlich constant, K_f , is as follows:

$$\log(K_f) = 3.76V - 4.47B - 0.80S + 0.73 + (0.20V - 0.16B + 0.48S + 0.24)\log(C_{sat}) \quad (\text{B-2})$$

where the units of K_f are $((\text{mg/kg})/(\text{mg/L})^\nu)$ and the units of C_{sat} , the aqueous solubility, is (mol/L) . The pp-LFER for the Freundlich exponent, ν , is as follows:

$$\nu = -0.20V + 0.16B - 0.48S + 0.76 \quad (\text{B-3})$$

The predicted Freundlich parameters (Eqs. (B-2) and (B-3)) were then used to compute predicted isotherm sorbed concentrations, $q_{pred,fr}$, for the Darco GAC data set using the Freundlich isotherm model:

$$q_{pred-fr} = K_f(c^\nu) \quad (\text{B-4})$$

By comparison, predicted isotherm sorbed concentrations were computed using the ss-LNL isotherm (Eq. (2-1)) with the Abraham pp-LFER (Eq. (3-2)) predicted median binding constants and the fitted sorbent-specific parameters (Table 3-1) for Darco GAC. The Darco GAC data set was selected as it was the data set used to calibrate the pp-LFERs (Eq. (B-2) and (B-3)). RMS errors were computed for both the LNL and Shih & Gschwend pp-LFER residual analysis, the results of which are summarized in Figure B6, below.

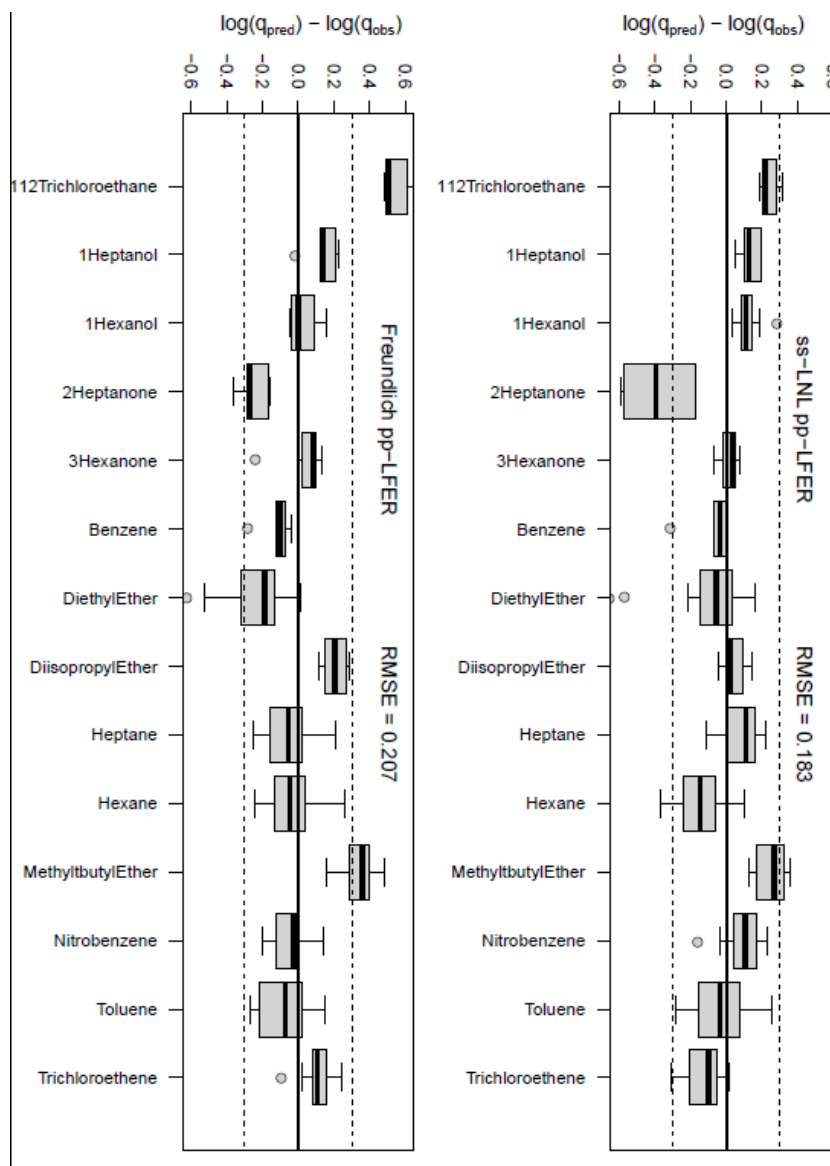


Figure B-1. Box plots comparing the residuals ($\log q_{pred} - \log(q_{obs})$) for the Shih & Gschwend Freundlich pp-LFER (left) and LNL pp-LFER (Eq. (3-2)) (right) models on Darco GAC. Dashed lines represent ± 0.30 log-units (factor of two) differences between predicted and observed sorbed concentrations. Solid line represents 1:1 agreement between the predicted and observed sorbed concentrations.

Appendix C

PREDICTING SOLVENT-WATER PARTITIONING OF CHARGED ORGANIC SPECIES USING QUANTUM CHEMICALLY-DERIVED ABRAHAM PP-LFER SOLUTE PARAMETERS

- C.1.** Direct Quantum Chemically Computed Solvent-water Partition Coefficients for Ions in Acetonitrile-, DMSO-, and Methanol-water Systems
- C.2.** Calculation of Excess Molar Refraction from McGowan Volume and QC-Computed Polarizability
- C.3.** Experimental Neutral Solvent-water System Parameters
- C.4.** Experimental Solvent-water Partition Coefficients
- C.5.** Absolv-Estimated (AAP) & Quantum-chemically Estimated (QCAP) Solute Parameters
- C.6.** QCAP Estimated Solvent Parameters
- C.7.** Complete AAP & QCAP Abraham pp-LFERs
- C.8.** Residual Plots of A Priori Quantum Chemically Computed and AAP-predicted Partition Coefficients

C.1 Direct Quantum Chemically Computed Solvent-water Partition Coefficients of Ions in Acetonitrile-, DMSO-, and Methanol-water Systems.

Table C-1a. Summary of RMS errors for solvent-water partitioning of anion, cation, and combined ionic data sets presented by Marenich et al¹⁵. QC computations were all performed at the M062X/MG3S level of theory.

Solvent	Subset	RMSE	N
Acetonitrile	All	4.263	30
	Anions	7.158	10
	Cations	2.002	20
Methanol	All	2.347	29
	Anions	2.721	12
	Cations	2.042	17
DMSO	All	7.134	29
	Anions	7.499	25
	Cations	4.188	4
All	Anions	6.206	47
All	Cations	2.232	41
All	All	4.355	88

Table C-1b. Experimental vs. QC computed solvent-water partition coefficients for anions and cations obtained from Marenich et al¹⁵. QC computations were all performed at the M062X/MG3S level of theory.

SoluteName	Charge	Solvent	Expt $\log(K_{\text{solvent}-\text{water}})$	QC Computed $\log(K_{\text{solvent}-\text{water}})$	Residuals
aceticacid	-1	acetonitrile	13.78	8.52	-5.26
benzoicacid	-1	acetonitrile	11.58	5.24	-6.34
phenol	-1	acetonitrile	12.31	3.27	-9.05
3-nitrophenol	-1	acetonitrile	11.14	2.17	-8.97
4-nitrophenol	-1	acetonitrile	9.38	1.91	-7.47
hydrochloricacid	-1	acetonitrile	8.87	-0.44	-9.31
hydrobromicacid	-1	acetonitrile	6.60	-0.54	-7.14
chloroaceticacid	-1	acetonitrile	11.07	6.19	-4.87
dichloroaceticacid	-1	acetonitrile	8.13	4.43	-3.70
methylamine	1	acetonitrile	-2.78	-0.95	1.83
n-propylamine	1	acetonitrile	-3.08	-1.82	1.26
tert-butylamine	1	acetonitrile	-2.64	-2.15	0.49
diethylamine	1	acetonitrile	-3.00	-2.82	0.19
trimethylamine	1	acetonitrile	-3.66	-2.80	0.86
triethylamine	1	acetonitrile	-3.81	-3.65	0.16
tri-n-propylamine	1	acetonitrile	-5.13	-4.91	0.22
aniline	1	acetonitrile	-3.22	-2.86	0.36
4-methylaniline	1	acetonitrile	-3.37	-3.33	0.04
pyrrolidine	1	acetonitrile	-3.66	-2.21	1.45
piperidine	1	acetonitrile	-3.66	-2.58	1.09
pyridine	1	acetonitrile	-4.10	-2.56	1.55
ammonia	1	acetonitrile	-3.00	-0.23	2.77
hydrazine	1	acetonitrile	-2.27	-0.07	2.20
methanol	1	acetonitrile	0.66	6.17	5.51
diethylether	1	acetonitrile	0.37	0.36	-0.01
acetone	1	acetonitrile	-0.29	0.97	1.27
acetophenone	1	acetonitrile	-3.44	-1.21	2.24
morpholine	1	acetonitrile	-3.30	-1.19	2.11
benzamide	1	acetonitrile	-1.83	-0.41	1.42
dimethylsulfoxide	1	acetonitrile	-4.25	-0.21	4.04
acetonitrile	-1	dimethylsulfoxide	8.57	-0.64	-9.21
cyanamide	-1	dimethylsulfoxide	9.31	-2.53	-11.84
aniline	-1	dimethylsulfoxide	7.33	-0.03	-7.36
diphenylamine	-1	dimethylsulfoxide	5.06	-1.50	-6.56

hydrogencyanide	-1	dimethylsulfoxide	11.87	-1.25	-13.13
aceticacid	-1	dimethylsulfoxide	13.48	9.23	-4.25
benzoicacid	-1	dimethylsulfoxide	11.43	6.49	-4.94
methanol	-1	dimethylsulfoxide	20.08	12.48	-7.60
ethanol	-1	dimethylsulfoxide	19.13	11.84	-7.29
isopropanol	-1	dimethylsulfoxide	19.27	11.51	-7.76
t-butanol	-1	dimethylsulfoxide	19.20	11.43	-7.77
phenol	-1	dimethylsulfoxide	12.97	4.43	-8.54
acetone	-1	dimethylsulfoxide	11.58	5.79	-5.79
3-pentanone	-1	dimethylsulfoxide	10.77	5.67	-5.11
water	-1	dimethylsulfoxide	20.30	17.85	-2.46
4-nitrophenol	-1	dimethylsulfoxide	9.09	2.92	-6.17
nitromethane	-1	dimethylsulfoxide	11.14	7.95	-3.19
4-nitroaniline	-1	dimethylsulfoxide	7.33	2.01	-5.32
acetamide	-1	dimethylsulfoxide	16.27	6.25	-10.02
thiophenol	-1	dimethylsulfoxide	7.04	-1.78	-8.81
dimethylsulfoxide	-1	dimethylsulfoxide	9.01	4.63	-4.39
hydrochloricacid	-1	dimethylsulfoxide	8.65	-0.54	-9.19
hydrobromicacid	-1	dimethylsulfoxide	7.69	-0.59	-8.28
dichloroaceticacid	-1	dimethylsulfoxide	9.60	4.66	-4.95
2,2,2-trifluoroethanol	-1	dimethylsulfoxide	15.68	8.80	-6.88
methylamine	1	dimethylsulfoxide	-4.40	-0.76	3.63
aniline	1	dimethylsulfoxide	-5.42	-1.92	3.50
pyridine	1	dimethylsulfoxide	-4.47	-1.72	2.75
ammonia	1	dimethylsulfoxide	-6.38	-0.28	6.09
aceticacid	-1	methanol	3.44	1.02	-2.42
propanoicacid	-1	methanol	3.08	0.45	-2.63
acrylicacid	-1	methanol	3.81	0.80	-3.01
benzoicacid	-1	methanol	2.71	-0.27	-2.98
phenol	-1	methanol	1.91	-0.67	-2.58
3-nitrophenol	-1	methanol	2.20	-0.21	-2.41
4-nitrophenol	-1	methanol	2.71	-0.28	-2.99
hydrochloricacid	-1	methanol	2.20	-0.44	-2.64
hydrobromicacid	-1	methanol	1.83	-0.57	-2.40
chloroaceticacid	-1	methanol	3.66	0.33	-3.33
dichloroaceticacid	-1	methanol	2.20	-0.26	-2.46
2-chlorophenol	-1	methanol	1.39	-1.23	-2.62
methylamine	1	methanol	1.54	-0.80	-2.34
trimethylamine	1	methanol	1.17	-2.17	-3.34

triethylamine	1	methanol	1.17	-3.25	-4.42
aniline	1	methanol	-1.25	-2.03	-0.79
2-methylaniline	1	methanol	-1.32	-2.38	-1.06
3-methylaniline	1	methanol	-1.25	-2.53	-1.28
4-methylaniline	1	methanol	-1.39	-2.52	-1.13
N-methylaniline	1	methanol	-0.66	-2.60	-1.94
1-aminonaphthalene	1	methanol	-1.61	-2.80	-1.18
piperidine	1	methanol	0.95	-2.43	-3.38
pyridine	1	methanol	0.22	-1.51	-1.73
quinoline	1	methanol	-1.10	-2.37	-1.27
ammonia	1	methanol	-0.29	-0.23	0.06
4-methoxyaniline	1	methanol	-0.66	-2.05	-1.39
4-nitroaniline	1	methanol	0.44	-1.18	-1.62
3-chloroaniline	1	methanol	-0.95	-2.46	-1.51
4-chloroaniline	1	methanol	-1.03	-2.46	-1.44

C.2 Calculation of Excess Molar Refraction from McGowan Volume and QC-Computed Polarizability

In the QCAP method, McGowan volumes for the ionized species (V) were obtained directly from Absolv⁴⁵. The excess molar refractivity, E , defined as the molar refractivity of the given solute (MR_i), in excess of an alkane (MR_i^*) of the same characteristic volume is^{84,85}:

$$E = MR_i - MR_i^* \quad (\text{C-1})$$

The total molar refraction, MR_i , can be computed as a function of the index of refraction (η) and the McGowan volume (V), of the solute, as follows⁷⁰:

$$MR_i = 10 \left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] V \quad (\text{C-2})$$

Additionally, previous work has shown that molar refractivity of the alkane, MR_i^* , can be estimated from the McGowan volume⁸⁴:

$$MR_A^* = 2.832V - 0.526 \quad (\text{C-3})$$

Combining Eqs. (C-1) – (C-3), an expression for the excess molar refractivity is obtained:

$$E = 10 \left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] V - 2.832V + 0.526 \quad (\text{C-4})$$

Finally, the index of refraction, η , can be related to the molecular polarizability, α , which can be obtained via quantum-chemical computation, by the Clausius-Mossotti equation⁸⁶:

$$\left[\frac{(\eta^2 - 1)}{(\eta^2 + 2)} \right] = \frac{4}{3} \pi \cdot N_A \cdot \frac{\alpha}{V} \quad (\text{C-5})$$

Combining Eqs. (C-4) and (C-5), an equation for the excess molar refractivity, E , results which is a function of the molecular polarizability and the McGowan volume:

$$E = 10 \cdot \frac{4}{3} \pi \cdot N_A \cdot \alpha - 2.832V + 0.526 \quad (\text{C-6})$$

where N_A is Avogadro's constant. The molecular polarizability is computed using Gaussian 09⁶⁴ at the M062X/6-311++G** level of theory, which has been shown in previous work to provide accurate computations of molecular polarizability⁸⁷.

C.2.1 R Code – QCAP Solute / Solvent Parameter Optimization

```
#define RMSE
RMSE=function(bx){
  ss=cumsum(bx^2)
  sl=length(bx)
  RMSE=(ss[sl]/sl)^0.5
}

#read solvents and solutes files
rawsolv <- read.csv(file="Solvents.csv")
dat <- data.frame(rawsolv)
#dat
solv <- cbind(dat$c,dat$e,dat$s,dat$a,dat$b,dat$v)
#solv
solv1 <- cbind(dat$e,dat$s,dat$a,dat$b,dat$v)
#solv1
solv2 <- cbind(dat$s,dat$a,dat$b)
#solv2
rawsolu <- read.csv(file="Solutes.csv")
dat1 <- data.frame(rawsolu)
#dat1
solu<-cbind(dat1$E_Gaussian,dat1$S,dat1$A,dat1$B,dat1$V_McGowan)
#solu
solu1<-cbind(1,dat1$E_Gaussian,dat1$S,dat1$A,dat1$B,dat1$V_McGowan)
#solu1
solu2<-cbind(dat1$S,dat1$A,dat1$B)
#solu2

#logKij
logKij<-solv%*%t(solu1)
logKij

#logK_Absolv
```

```

SMD = read.csv(file="Mixed_Basis_Eout.csv")
solvents = unique(SMD$Solvent)
solutes = unique(SMD$Solute)
logp = matrix(nrow=length(solvents), ncol=length(solutes))

for (i in 1:length(dat$Solvent)){
  logp[i,] = t(SMD[SMD$Solvent==dat$SOL[[i]],6])
}

#Initiate
b=logp-logKij
RMSEb=NULL

#loop
ybold=RMSE(b)
for(i in 1:200){
xu<-solve(t(solv2)%*%solv2,t(solv2)%*%b )

solu2=solu2+t(xu)
solu1<-cbind(1, dat1$E_Gaussian, solu2[,1:3],dat1$V_McGowan)
logKnew<-solv%*%t(solu1)
b=logp-logKnew

xv<-solve(t(solu1)%*%solu1,t(solu1)%*%t(b))

solv=solv+t(xv)
solv2<-solv[,3:5]
solv3<-cbind(solv[,3:5])
logKnew<-solv%*%t(solu1)
b=logp-logKnew

ybnew=RMSE(b)
print(c(ybold,ybnew),digits=22)
if(abs(ybold-ybnew)<10^-30)break()
ybold=ybnew
RMSEb=c(RMSEb,ybnew)
}
#Output QC-optimized solute and solvent parameters
write.table(solu1,file='Optimized_Solute_Parameters.csv',sep=',',row.names=TRUE,col.names=N
A)

write.table(solv,file='Optimized_Solvent_Parameters.csv',sep=',',row.names=TRUE,col.names=N
A)

```

C.3 Experimental Neutral Solvent-water System Parameters

Table C-2. Literature values for solvent-water Abraham pp-LFER system parameters. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). All system parameters come from the UFZ LSER database⁷¹, the REF column indicates the original experimental source of the data.

SYSTEM	E	S	A	B	V	C	REF
Ethyleneglycol	0.58	-0.51	0.72	-2.62	2.73	-0.27	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
Benzonitrile	0.03	0.06	-1.61	-4.56	4.03	0.10	Abraham, M. H., Acree, W. E., Jr. (2011) Thermochimica Acta 526(1-2), 22-28.
Acetone	0.31	-0.12	-0.61	-4.75	3.94	0.31	Abraham, M. H.; Acree, W. E.; Leo, A. J.; Hoekman, D. New J. Chem. 2009, 33 (3), 568-573.
Diethylether	0.36	-0.82	-0.59	-4.96	4.35	0.35	Grubbs, L. M.; Saifullah, M.; De La Rosa, N. E.; Ye, S. L.; Achi, S. S.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 298 (1), 48-53.
N-methylpyrrolidin-2-one	0.53	0.23	0.84	-4.79	3.67	0.15	Abraham, M. H.; Acree, W. E.; Cometto-Muniz, J. E. New J. Chem. 2009, 33(10), 2034-2043.
N,N-Dimethylformamide	-0.06	0.34	0.36	-4.87	4.49	-0.31	Abraham, M. H.; Acree, W. E.; Cometto-Muniz, J. E. New J. Chem. 2009, 33(10), 2034-2043.
Propylene carbonate	0.17	0.50	-1.28	-4.41	3.42	0.00	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
tetrahydrofuran	0.36	-0.38	-0.24	-4.93	4.45	0.22	Saifullah, M.; Ye, S.; Grubbs, L. M.; La Rosa, N. E.; Acree, W. E., Jr.; Abraham, M. H. J. Solution Chem. 2011, 40 (12), 2082-2094.
1,4-dioxane	0.35	-0.03	-0.58	-4.81	4.11	0.12	Saifullah, M.; Ye, S.; Grubbs, L. M.; La Rosa, N. E.; Acree, W. E., Jr.; Abraham, M. H. J. Solution Chem. 2011, 40 (12), 2082-2094.

2-butanol	0.25	-0.98	0.16	-3.88	4.11	0.13	Sprunger, L. M.; Achi, S. S.; Pointer, R.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 288 (1-2), 121-127.
2-propanol	0.34	-1.05	0.41	-3.83	4.03	0.10	Sprunger, L. M.; Achi, S. S.; Pointer, R.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 288 (1-2), 121-127.
n-hexane	0.56	-1.71	-3.58	-4.94	4.46	0.33	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
bromobenzene	0.44	-0.42	-3.17	-4.56	4.45	-0.02	Abraham, M. H.; Acree, W. E.; Leo, A. J.; Hoekman, D. New J. Chem. 2009, 33 (8), 1685-1692.
chlorobenzene	0.38	-0.52	-3.18	-4.70	4.61	0.07	Abraham, M. H.; Acree, W. E.; Leo, A. J.; Hoekman, D. New J. Chem. 2009, 33 (8), 1685-1692.
iodobenzene	0.30	-0.31	-3.21	-4.65	4.59	-0.19	Abraham, M. H.; Acree, W. E.; Leo, A. J.; Hoekman, D. New J. Chem. 2009, 33 (8), 1685-1692.
2-Methyl-1-propanol	0.35	-1.13	0.02	-3.57	3.97	0.19	Sprunger, L. M.; Achi, S. S.; Pointer, R.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 288 (1-2), 121-127.
Butanone	0.26	-0.08	-767	-4.86	4.18	0.25	Abraham, M. H.; Acree, W. E.; Leo, A. J.; Hoekman, D. New J. Chem. 2009, 33 (3), 568-573.
Dibutylether	0.39	-0.99	-1.41	-5.36	4.52	0.18	Grubbs, L. M.; Saifullah, M.; De La Rosa, N. E.; Ye, S. L.; Achi, S. S.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 298 (1), 48-53.
Ethylacetate	0.37	-0.45	-0.70	-4.90	4.15	0.33	Sprunger, L. M.; Proctor, A.; Acree Jr, W. E.; Abraham, M. H.; Benjelloun-Dakhama, N. Fluid Phase Equilib. 2008, 270 (1-2), 30-44.
Methylacetate	0.22	-0.15	-1.04	-4.53	3.97	0.35	Sprunger, L. M.; Proctor, A.; Acree Jr, W. E.; Abraham, M. H.; Benjelloun-Dakhama, N. Fluid Phase Equilib. 2008, 270 (1-2), 30-44.
Methyltbutylether	0.31	-0.82	-0.62	-5.10	4.43	0.34	Grubbs, L. M.; Saifullah, M.; De La Rosa, N. E.; Ye, S. L.; Achi, S. S.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 298 (1), 48-53.
N-methylformamide	0.41	-0.29	0.54	-4.09	3.47	0.11	Abraham, M. H.; Acree, W. E.; Cometto-Muniz, J. E. New J. Chem. 2009, 33(10), 2034-2043.
p-xylene	0.48	-0.81	-2.94	-4.87	4.53	0.17	Stephens, T. W.; De La Rosa, N. E.; Saifullah, M.; Ye, S.; Chou, V.; Quay, A. N.; Acree Jr, W. E.; Abraham, M. H.

							Fluid Phase Equilib. 2011, 309 (1), 30-35.
n-decane	0.72	-1.74	-3.45	-4.97	4.48	0.19	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
n-heptane	0.63	-1.76	-3.57	-4.95	4.49	0.30	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
n-octane	0.74	-1.84	-3.69	-4.91	4.50	0.23	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
ethylbenzene	0.47	-0.72	-3.00	-4.84	4.51	0.09	Stephens, T. W.; Loera, M.; Quay, A. N.; Chou, V.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Open Thermodyn. J. 2011, 5, 104-121.
toluene	0.43	-0.64	-3.00	-4.75	4.52	0.13	Stephens, T. W.; Loera, M.; Quay, A. N.; Chou, V.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Open Thermodyn. J. 2011, 5, 104-121.
2-Methyl-2-propanol	0.17	-0.95	0.33	-4.09	4.11	0.21	Sprunger, L. M.; Achi, S. S.; Pointer, R.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 288 (1-2), 121-127.
3-Methyl-1-butanol	0.36	-1.27	0.09	-3.77	4.27	0.07	Sprunger, L. M.; Achi, S. S.; Pointer, R.; Acree, W. E.; Abraham, M. H. Fluid Phase Equilibria 2010, 288 (1-2), 121-127.
Acetonitrile	0.08	0.33	-1.57	-4.39	3.36	0.41	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
Butylacetate	0.36	-0.50	-0.87	-4.97	4.28	0.25	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Cyclohexanone	0.23	0.06	-0.98	-4.84	4.32	0.04	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Decanol	0.75	-1.46	0.06	-4.05	4.29	-0.06	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Dubutylformamide	0.30	-0.44	0.36	-4.90	3.95	0.33	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3),

							1500-1627
Diethylacetamide	0.03	0.09	1.34	-5.08	4.09	0.21	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Dimethylacetamide	0.08	0.21	0.92	-5.00	4.56	-0.27	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Dimethylsulfoxide	0.33	0.79	1.26	-4.54	3.36	-0.19	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
Ethanol	0.47	-1.04	0.33	-3.60	3.86	0.22	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
Formamide	0.07	0.31	0.59	-3.15	2.43	-0.17	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Heptanol	0.49	-1.26	0.04	-4.16	4.42	-0.03	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Isobutanol	0.31	-1.07	0.18	-3.77	4.04	0.16	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Methanol	0.33	-0.71	0.24	-3.32	3.55	0.28	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
N-Ethylacetamide	0.13	-0.44	1.18	-4.73	3.86	0.28	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
N-Ethylformamide	0.03	-0.17	0.94	-4.59	3.73	0.22	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
N-Formylmorpholine	0.70	-0.06	0.01	-4.09	3.41	-0.03	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Nitromethane	-0.09	0.79	-1.46	-4.36	3.46	0.02	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627

N-Methylacetamide	0.21	-0.17	1.31	-4.59	3.83	0.09	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
N-Methylpyrrolidinone	0.53	0.28	0.84	-4.79	3.67	0.15	Abraham, M. H.; Acree, W. E. <i>New J. Chem.</i> 2010, 34 2298-2305.
N-Methyl-2-piperidone	0.33	0.26	1.56	-5.04	3.98	0.06	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Octanol	0.49	-1.04	-0.02	-4.24	4.22	-0.03	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Pentanol	0.52	-1.29	0.21	-3.91	4.21	0.08	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Propanone	0.31	-0.12	-0.61	-4.75	3.94	0.31	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Isopropanol	0.32	-1.02	0.53	-3.87	4.02	0.10	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
s-Butanol	0.38	-0.96	0.13	-3.61	3.83	0.19	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
t-Butanol	0.14	-0.92	0.32	-4.03	4.11	0.20	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Tributylphosphate	0.57	-0.84	-1.07	-4.33	3.92	0.33	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Trifluoroethanol	-0.09	-0.59	-1.28	-1.27	3.09	0.40	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
butanol	0.44	-1.18	0.10	-3.92	4.12	0.15	Abraham, M. H., Acree, W. E. (2010) <i>New J. Chem.</i> 34, 2298-2305.
hexanol	0.49	-1.16	0.05	-3.97	4.13	0.12	Abraham, M. H., Acree, W. E. (2010) <i>New J. Chem.</i> 34, 2298-2305.

propanol	0.41	-1.03	0.25	-3.77	3.99	0.14	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
2-Pentanol	0.46	-1.33	0.21	-3.75	4.20	0.12	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
n-dodecane	0.67	-1.64	-3.55	-5.01	4.46	0.11	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
n-hexadecane	0.67	-1.62	-3.59	-4.87	4.43	0.09	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
N-Methylpyrrolidinone	0.53	0.28	0.84	-4.79	3.67	0.15	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
n-undecane	0.60	-1.66	-3.42	-5.12	4.62	0.06	Stephens, T. W.; Quay, A. N.; Chou, V.; Loera, M.; Shen, C.; Wilson, A.; Acree, W. E.; Abraham, M. H. Global J. Phys. Chem. 2012, 3, 1-12.
Propanone	0.31	-0.12	-0.61	-4.75	3.94	0.31	Abraham, M. H., Acree, W. E. (2010) New J. Chem. 34, 2298-2305.
benzene	0.46	-0.59	-3.10	-4.63	4.49	0.14	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
butane	-0.01	-1.58	-3.19	-4.57	4.56	0.30	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Carbondisulfide	0.69	-0.94	-3.60	-5.82	4.92	0.05	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Cyclohexane	0.78	-1.68	-3.74	-4.93	4.58	0.16	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Dichloromethane	0.10	-0.19	-3.06	-4.09	4.32	0.32	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627
Fluorobenzene	0.15	-0.37	-3.03	-4.60	4.54	0.14	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. Pharma. Sci. 2010, 99 (3), 1500-1627

Isooctane	0.56	-1.74	-3.68	-4.86	4.42	0.32	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Isopropylmyristate	0.93	-1.15	-1.68	-4.09	4.25	-0.61	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Methylcyclohexane	0.78	-1.98	-3.52	-4.29	4.53	0.25	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Nitrobenzene	0.54	0.04	-2.33	-4.61	4.31	-0.20	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Oleyl alcohol	0.15	-0.84	-0.44	-4.04	4.13	-0.10	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Pentane	0.39	-1.57	-3.54	-5.22	4.51	0.37	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Tetrachloromethane	0.52	-1.16	-3.56	-4.59	4.62	0.20	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Toluene	0.53	-0.72	-3.01	-4.82	4.55	0.14	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Trichloromethane	0.11	-0.40	-3.11	-3.51	4.40	0.19	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Triolein	0.98	-2.08	-2.01	-3.45	4.07	0.39	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
Chlorobutane	0.27	-0.57	-2.92	-4.88	4.46	0.22	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627
1,2-Dichloroethane	0.29	-0.13	-2.80	-4.29	4.18	0.18	Abraham, M. H.; Smith, R. E.; Luchtefeld, R.; Boorem, A. J.; Luo, R. S.; Acree, W. E. J. <i>Pharma. Sci.</i> 2010, 99 (3), 1500-1627

C.4 Experimental Solvent-water Partition Coefficients

Table C-3a. Table of experimental acetone-water partition coefficients ($\log(K_{\text{acetone-water}})$) for carboxylate anions in acetone. $K_{\text{acetone-water}}$ has units (L water / kg acetone).

Solute	Expt $\log(K_{\text{acetone-water}})$
26DibromobenzoicAcid	-9.02
26DichlorobenzoicAcid	-8.98
26DiethoxybenzoicAcid	-10.3
26DifluorobenzoicAcid	-9.86
26DimethoxybenzoicAcid	-11.23
26DimethylbenzoicAcid	-9.1
26DinitrobenzoicAcid	-8.99
26DipropoxybenzoicAcid	-9.28
2CyanobenzoicAcid	-9.23
2NitrobenzoicAcid	-9.47
34DichlorobenzoicAcid	-7.01
34DimethoxybenzoicAcid	-9.91
34DinitrobenzoicAcid	-7.57
35DinitrobenzoicAcid	-7.32
3Bromo4methoxybenzoicAcid	-8.12
3Bromo4methylbenzoicAcid	-8.04
3BromobenzoicAcid	-7.43
3Chloro4nitrobenzoicAcid	-7.41
3ChlorobenzoicAcid	-7.77
3CyanobenzoicAcid	-8.9
3FluorobenzoicAcid	-9.02
3Methoxy4chlorobenzoicAcid	-8.17
3Methoxy4methylbenzoicAcid	-9.1
3Methoxy4nitrobenzoicAcid	-8.45
3MethoxybenzoicAcid	-9.57
3Methyl4chlorobenzoicAcid	-7.36
3Methyl4methoxybenzoicAcid	-9.41
3Methyl4nitrobenzoicAcid	-7.96

3MethylbenzoicAcid	-9.17
3Nitro4chlorobenzoicAcid	-7.28
3Nitro4methoxybenzoicAcid	-8.62
3Nitro4methylbenzoicAcid	-8.08
3TrifluoromethylbenzoicAcid	-7.49
4BromobenzoicAcid	-7.67
4ChlorobenzoicAcid	-7.97
4CyanobenzoicAcid	-8.76
4FluorobenzoicAcid	-9.44
4HydroxybenzoicAcid	-10.65
4MethoxybenzoicAcid	-9.67
4MethylbenzoicAcid	-9.14
4NitrobenzoicAcid	-8.62
AceticAcid	-11.34
IsobutanoicAcid	-10.72
PropanoicAcid	-11.07

Table C-3b. Table of experimental acetonitrile-water partition coefficients ($\log(K_{\text{acetonitrile-water}})$) for carboxylate anions in acetonitrile. $K_{\text{acetonitrile-water}}$ has units (L water / kg acetonitrile).

Solute	Expt $\log(K_{\text{acetonitrile-water}})$
23DibromopropanoicAcid	-5.32
26DibromobenzoicAcid	-6.86
26DichlorobenzoicAcid	-6.75
26DiethoxybenzoicAcid	-7.74
26DifluorobenzoicAcid	-7.61
26DimethoxybenzoicAcid	-8.65
26DimethylbenzoicAcid	-7.21
26DinitrobenzoicAcid	-6.64
26DipropoxybenzoicAcid	-6.89
2CyanobenzoicAcid	-7
2NitrobenzoicAcid	-7.44
34DichlorobenzoicAcid	-5.38
34DimethoxybenzoicAcid	-7.44
34DimethylbenzoicAcid	-6.84
34DinitrobenzoicAcid	-5.29
35DimethylbenzoicAcid	-6.9
35DinitrobenzoicAcid	-5.18
3Bromo4methoxybenzoicAcid	-6.13
3Bromo4methylbenzoicAcid	-6.13
3Chloro4nitrobenzoicAcid	-5.4
3ChlorobenzoicAcid	-6.09
3CyanobenzoicAcid	-6.86
3FluorobenzoicAcid	-7.08
3Methoxy4chlorobenzoicAcid	-6.27
3Methoxy4methylbenzoicAcid	-7.02
3Methoxy4nitrobenzoicAcid	-6.46
3MethoxybenzoicAcid	-7.37
3Methyl4chlorobenzoicAcid	-5.59
3Methyl4methoxybenzoicAcid	-7.13
3Methyl4nitrobenzoicAcid	-6.09
3MethylbenzoicAcid	-6.93
3Nitro4chlorobenzoicAcid	-5.34
3Nitro4methoxybenzoicAcid	-6.45
3Nitro4methylbenzoicAcid	-6.16

3TrifluoromethylbenzoicAcid	-5.64
4BromobenzoicAcid	-5.75
4ChlorobenzoicAcid	-6.09
4CyanobenzoicAcid	-6.55
4FluorobenzoicAcid	-7.42
4HydroxybenzoicAcid	-8.34
4MethoxybenzoicAcid	-7.43
4MethylbenzoicAcid	-7.1
4NitrobenzoicAcid	-6.59
AceticAcid	-10.4
ButanoicAcid	-9.84
CyanoaceticAcid	-8.41
IsobutanoicAcid	-9.27
PropanoicAcid	-9.52

Table C-3c. Table of experimental DMSO-water partition coefficients ($\log(K_{DMSO-water})$) for carboxylate anions in dimethylsulfoxide (DMSO). $K_{DMSO-water}$ has units (L water / kg DMSO)

Solute	Expt $\log(K_{DMSO-water})$
23DibromopropanoicAcid	-4.04
26DibromobenzoicAcid	-5.09
26DichlorobenzoicAcid	-5.17
26DiethoxybenzoicAcid	-6.76
26DifluorobenzoicAcid	-6.5
26DimethoxybenzoicAcid	-7.69
26DimethylbenzoicAcid	-6.56
26DinitrobenzoicAcid	-4.07
26DipropoxybenzoicAcid	-5.95
2NitrobenzoicAcid	-6.3
34DichlorobenzoicAcid	-4.37
34DimethoxybenzoicAcid	-6.87
34DinitrobenzoicAcid	-3.75
35DinitrobenzoicAcid	-3.57
3Bromo4methylbenzoicAcid	-5.01
3BromobenzoicAcid	-4.64
3Chloro4nitrobenzoicAcid	-4.07
3ChlorobenzoicAcid	-5.18
3CyanobenzoicAcid	-5.76
3Methoxy4chlorobenzoicAcid	-5.49
3Methoxy4methylbenzoicAcid	-6.42
3Methoxy4nitrobenzoicAcid	-5.5
3Methyl4chlorobenzoicAcid	-4.84
3Methyl4methoxybenzoicAcid	-6.63
3Methyl4nitrobenzoicAcid	-5.15
3MethylbenzoicAcid	-6.5
3Nitro4chlorobenzoicAcid	-3.9
3Nitro4methoxybenzoicAcid	-5.41
3Nitro4methylbenzoicAcid	-5.04
3NitrobenzoicAcid	-5.29
4BromobenzoicAcid	-5.04
4ChlorobenzoicAcid	-5.03
4CyanobenzoicAcid	-5.33
4HydroxybenzoicAcid	-7.39
4MethylbenzoicAcid	-6.63

4NitrobenzoicAcid	-5.58
AceticAcid	-9.9
ButanoicAcid	-9.51
CyanoaceticAcid	-7.95
IsobutanoicAcid	-9.42
PropanoicAcid	-9.53

Table C-3d. Table of experimental methanol-water partition coefficients ($\log(K_{\text{methanol-water}})$) for carboxylate anions in methanol. $K_{\text{methanol-water}}$ has units (L water / kg methanol).

Solute	Expt $\log(K_{\text{methanol-water}})$
23DibromopropanoicAcid	-1.57
26DiethoxybenzoicAcid	-1.13
26DifluorobenzoicAcid	-1.44
26DimethoxybenzoicAcid	-1.71
26DimethylbenzoicAcid	-0.53
26DinitrobenzoicAcid	-1.81
26DipropoxybenzoicAcid	-0.26
2CyanobenzoicAcid	-1.32
2NitrobenzoicAcid	-1.69
34DichlorobenzoicAcid	0.34
34DimethoxybenzoicAcid	-1.49
34DimethylbenzoicAcid	-0.4
34DinitrobenzoicAcid	-0.8
35DimethylbenzoicAcid	-0.6
35DinitrobenzoicAcid	-0.87
3Bromo4methoxybenzoicAcid	-0.43
3Bromo4methylbenzoicAcid	-0.42
3BromobenzoicAcid	-0.09
3Chloro4nitrobenzoicAcid	-0.55
3ChlorobenzoicAcid	-0.48
3CyanobenzoicAcid	-1.22
3FluorobenzoicAcid	-0.87
3Methoxy4chlorobenzoicAcid	-0.5
3Methoxy4methylbenzoicAcid	-0.82
3Methoxy4nitrobenzoicAcid	-1.96
3MethoxybenzoicAcid	-1.11
3Methyl4chlorobenzoicAcid	0.06
3Methyl4methoxybenzoicAcid	-0.9
3Methyl4nitrobenzoicAcid	-0.55
3MethylbenzoicAcid	-0.87
3Nitro4chlorobenzoicAcid	-0.28
3Nitro4methoxybenzoicAcid	-1.07
3Nitro4methylbenzoicAcid	-0.56
3TrifluoromethylbenzoicAcid	0.08
4BromobenzoicAcid	-0.14

4ChlorobenzoicAcid	-0.36
4CyanobenzoicAcid	-1.08
4FluorobenzoicAcid	-1.11
4HydroxybenzoicAcid	-1.72
4MethoxybenzoicAcid	-1.31
4MethylbenzoicAcid	-1.02
4NitrobenzoicAcid	-0.87
AceticAcid	-2.88
CyanoaceticAcid	-3.08
IsobutanoicAcid	-1.98
PentanoicAcid	-1.64
PropanoicAcid	-2.28

Table C-3e. Table of experimental octanol-water partition coefficients ($\log(K_{\text{octanol-water}})$) for quaternary amine cations in octanol. $K_{\text{octanol-water}}$ has units (L water / kg octanol). Solute and Group Identification Numbers refer to the original classifications in the primary reference (Abraham & Zhao, 2008)¹⁴. SMILES strings are for the charged (cationic) species.

Solute	SMILES	Group	$\log(K_{\text{octanol-water}})$
1	<chem>C[N+](C)(C)C</chem>	1	-1.71
2	<chem>C[N+](C)(C)CC</chem>	1	-2.67
3	<chem>C[N+](C)(C)CCCC</chem>	1	0.35
4	<chem>C[N+](C)(C)CCCCC</chem>	1	-1.11
5	<chem>C[N+](C)(C)CCCCC</chem>	1	-0.35
6	<chem>C[N+](C)(C)CCCCCCCC</chem>	1	1.19
7	<chem>C[N+](C)(C)CCCCCCCCC</chem>	1	3.01
8	<chem>C[N+](C)(C)CCCCCCCCCCCCC</chem>	1	6.05
9	<chem>C[N+](C)(C)CCCI</chem>	1	-3.09
10	<chem>C[N+](C)(C)CC1CCCC1</chem>	1	-0.21
12	<chem>C[N+](C)(C)C(C)(C)C</chem>	1	-0.83
13	<chem>C[N+](C)(C)C(C)CC</chem>	1	0.11
14	<chem>C[N+](C)(C)CC1CC(C)CC1</chem>	1	0.33
15	<chem>C[N+](C)(C)CC1C=C(C)CC1</chem>	1	-0.37
17	<chem>C[N+](C)(C)CCOC(=O)C</chem>	1	-4.45
18	<chem>C[N+](C)(C)CCOC(=O)C=C</chem>	1	-3.47
19	<chem>C[N+](C)(C)CCOC(=O)C=N#N</chem>	1	-3.25
20	<chem>C[N+](C)(C)CCOC(=O)CBr</chem>	1	-1.83
22	<chem>C[N+](C)(C)CCOC(=O)CF</chem>	1	-3.03
24	<chem>C[N+](C)(C)CCOC(=O)OC</chem>	1	-3.37
25	<chem>C[N+](C)(C)CC1CCC01</chem>	1	-1.27
26	<chem>C[N+](C)(C)CC1CCOC1</chem>	1	-1.97
27	<chem>C[N+](C)(C)CC1CC(C)OC1</chem>	1	-0.97
28	<chem>C[N+](C)(C)CC1CCC(C)O1</chem>	1	-0.27
29	<chem>C[N+](C)(C)CC1COC(C)O1</chem>	1	-1.33
30	<chem>C[N+](C)(C)CC1COCO1</chem>	1	-2.73
31	<chem>C[N+](C)(C)CC1CSCO1</chem>	1	-2.07
33	<chem>C[N+](C)(C)CCOC</chem>	1	-1.91
34	<chem>C[N+](C)(C)CCOC(C)C</chem>	1	-0.11
35	<chem>C[N+](C)(C)CCOCC</chem>	1	-1.11
36	<chem>C[N+](C)(C)CCOCCC</chem>	1	-0.61
39	<chem>C[N+](C)(C)Cc1cccc(CCC)c1</chem>	1	1.73

40	<chem>C[N+](C)(C)Cc1cccc(C(C)C)c1</chem>	1	1.53
46	<chem>C[N+](C)(C)Cc1cccc(C(=O)OCC)c1</chem>	1	-0.17
48	<chem>C[N+](C)(C)Cc1cccc(C(=O)OC)c1</chem>	1	-1.17
50	<chem>C[N+](C)(C)Cc1cccc(OCC)c1</chem>	1	0.13
51	<chem>C[N+](C)(C)Cc1cccc(OC)c1</chem>	1	-0.01
56	<chem>C[N+](C)(C)Cc1cccc(C#N)c1</chem>	1	-2.31
64	<chem>C[N+](C)(C)Cc1cccc(C)n1</chem>	1	-2.01
68	<chem>C[N+](C)(C)Cc1ccc(C)o1</chem>	1	-0.37
70	<chem>C[N+](C)(C)Cc1ccoc1</chem>	1	-1.47
71	<chem>C[N+](C)(C)Cc1cc(C)oc1</chem>	1	-0.67
72	<chem>C[N+](C)(C)Cc1ccc(CCl)o1</chem>	1	-0.67
81	<chem>C[N+](CC)(CC)CC</chem>	1	-0.93
82	<chem>C[N+](CCC)(CCC)CCC</chem>	1	1.37
83	<chem>C[N+](CCCC)(CCCC)CCCC</chem>	1	3.85
84	<chem>CC[N+](CC)(CC)CC</chem>	1	-2.31
85	<chem>CC[N+](CCC)(CCC)CCC</chem>	1	-1.05
116	<chem>CCCCCCCCCCCC[N+](C)(C)Cc1ccc(Cl)cc1</chem>	1	10.11
121	<chem>CCCC[n+]1cccc1</chem>	1	-1.59
124	<chem>CCCCCCCCCCCC[n+]1cccc1</chem>	1	2.35
136	<chem>CCCCCC[n+]1cc(C)cc(C)c1</chem>	1	0.95
138	<chem>CCCCCCCCCCCC[n+]1cc(C)cc(C)c1</chem>	1	4.85
139	<chem>CCCCCCCCCCCC[n+]1cc(C)cc(C)c1</chem>	1	7.15
140	<chem>CCCCCCCCCCCCCCCC[n+]1cc(C)cc(C)c1</chem>	1	9.13
146	<chem>CCCCCCCCCCCCCCCC[n+]1ccc(C)c(C)c1</chem>	1	6.83
165	<chem>c1cccc1C[n+]2ccccc2</chem>	1	-1.45
203	<chem>Cn1cc[n+](COCC(C)(C)N(=O)=O)c1</chem>	1	-1.47
204	<chem>Cn1cc[n+](COCC(C)(C)N(=O)=O)c1C</chem>	1	-0.67
205	<chem>Cn1cc[n+](COC(C)C(C)(C)N(=O)=O)c1C</chem>	1	-2.29
212	<chem>Cc1n(C)cc[n+]1CCCN(C)S(=O)(=O)C(F)(F)F</chem>	3	-1.57
213	<chem>Cc1n(C)cc[n+]1CCN(C)S(=O)(=O)C(F)(F)F</chem>	3	-1.51
215	<chem>CCc1n(C)cc[n+]1CCN(C)S(=O)(=O)C(F)(F)F</chem>	3	-0.83
220	<chem>Cn1cc[n+](CCC#N)c1C=NO</chem>	4	-2.09
221	<chem>Cn1cc[n+](CCCC#N)c1C=NO</chem>	4	-1.49
222	<chem>Cn1cc[n+](CCCCC#N)c1C=NO</chem>	4	-1.09
223	<chem>Cn1cc[n+](COCC)c1C=NO</chem>	4	-0.89
224	<chem>Cn1cc[n+](COCCCC)c1C=NO</chem>	4	1.73
225	<chem>Cn1cc[n+](COCCCCCCC)c1C=NO</chem>	4	4.29
226	<chem>Cn1cc[n+](COCCCCCCCCC)c1C=NO</chem>	4	6.27
228	<chem>Cn1cc[n+](COC(C)C#C)c1C=NO</chem>	4	0.23

229	Cn1cc[n+](COC(CC)C#C)c1C=NO	4	1.53
230	Cn1cc[n+](COC(CCC)C#C)c1C=NO	4	2.31
232	Cn1cc[n+](COC2CCCC2)c1C=NO	4	1.59
233	Cn1cc[n+](COC2CCCCC2)c1C=NO	4	2.11
237	Cn1cc[n+](CCOCC#C)c1C=NO	4	-0.99
238	Cn1cc[n+](CCCOCC#C)c1C=NO	4	-0.03
240	Cn1cc[n+](COCCCCc2ccccc2)c1C=NO	4	3.61
242	Cn1cc[n+](CCCC(C)(C)N(=O)=O)c1C=NO	4	0.15
244	CCn1cc[n+](CCCC(C)(C)N(=O)=O)c1C=NO	4	0.91
245	Cn1cc[n+](CCF)c1C=NO	4	-0.65
246	Cn1cc[n+](CCCl)c1C=NO	4	-1.51
247	Cn1cc[n+](CCCCCCCCl)c1C=NO	4	2.33
248	Cn1cc[n+](CCCCC#C)c1C=NO	4	0.47
249	Cn1cc[n+](CCCBrc1C=NO	4	1.95
250	Cn1cc[n+](CCBr)c1C=NO	4	-0.37
252	CCn1cc[n+](CCN(C)C)c1C=NO	4	-0.95
253	CCn1cc[n+](CCN(C)C)c1C=NO	4	-0.91
254	Cn1cc[n+](CC(C)N(C)C)c1C=NO	4	-1.11
255	Cn1cc[n+](CC(C)N(CC)CC)c1C=NO	4	0.09
256	Cn1cc[n+](CC(N(C)C)C(C)C)c1C=NO	4	1.17
257	Cn1cc[n+](CC(N(C)C)CC)c1C=NO	4	-0.63
258	Cn1cc[n+](CCCN(C)C)c1C=NO	4	-1.77
259	Cn1cc[n+](CCN(C(C)C)C(C)C)c1C=NO	4	0.93
261	Cn1cc[n+](CCN2CCOCC2)c1C=NO	4	-1.31
262	Cn1cc[n+](CCN2CCCCC2)c1C=NO	4	-0.13
263	Cn1cc[n+](CCN2CCCC2)c1C=NO	4	-0.95
264	Cn1cc[n+](CC2CCCN2C)c1C=NO	4	-1.43
271	Cn1cc[n+](COCCF)c1C=NO	4	-1.17
272	Cn1cc[n+](COCCCCl)c1C=NO	4	0.65
273	Cn1cc[n+](COCCCCCl)c1C=NO	4	1.57
274	Cn1cc[n+](COCCCC#C)c1C=NO	4	0.31
275	Cn1cc[n+](COCCCCBr)c1C=NO	4	0.99
276	Cn1cc[n+](COCCC=CCC)c1C=NO	4	2.99
277	Cn1cc[n+](COCCC=C)c1C=NO	4	0.83
278	Cn1cc[n+](COCCC(C)OC)c1C=NO	4	0.07
280	Cn1cc[n+](COCCC(C)=C)c1C=NO	4	1.69
281	Cn1cc[n+](COCCC#CC)c1C=NO	4	0.53
283	Cn1cc[n+](COCCBr)c1C=NO	4	-0.09
284	Cn1cc[n+](COCC=CCCC)c1C=NO	4	3.11

285	Cn1cc[n+](COCC=CCC)c1C=NO	4	1.91
286	Cn1cc[n+](COCC=CC)c1C=NO	4	1.43
287	Cn1cc[n+](COCC=CC(C)(C)C)c1C=NO	4	3.57
288	Cn1cc[n+](COCC=C)c1C=NO	4	-0.49
289	Cn1cc[n+](COCC=C(C)C)c1C=NO	4	1.71
290	Cn1cc[n+](COCC(C)CBr)c1C=NO	4	2.05
291	Cn1cc[n+](COCC(C)=C)c1C=NO	4	0.85
292	Cn1cc[n+](COCC(C)(C)CCl)c1C=NO	4	2.51
293	Cn1cc[n+](COCC(C)(C)CBr)c1C=NO	4	2.89
294	Cn1cc[n+](COCC(C)(C)C)c1C=NO	4	2.69
295	Cn1cc[n+](COCC(C)(C)C#C)c1C=NO	4	2.01
296	Cn1cc[n+](COC)c1C=NO	4	-1.49
297	Cn1cc[n+](COC(CCCl)CC)c1C=NO	4	2.27
298	Cn1cc[n+](COC(CCC)CC#C)c1C=NO	4	2.17
299	Cn1cc[n+](COC(CC)CCC)c1C=NO	4	2.79
300	Cn1cc[n+](COC(C=C)CCC)c1C=NO	4	2.77
301	Cn1cc[n+](COC(C)CCCCl)c1C=NO	4	2.05
302	Cn1cc[n+](COC(C)CCC)c1C=NO	4	2.27
303	Cn1cc[n+](COC(C)CC(C)(C)C)c1C=NO	4	3.95
304	Cn1cc[n+](COC(C)CC#C)c1C=NO	4	0.31
305	Cn1cc[n+](COC(C)C=C)c1C=NO	4	0.71
306	Cn1cc[n+](COC(C)C)c1C=NO	4	0.17
308	Cn1cc[n+](COC(C)C(C)C#C)c1C=NO	4	1.41
309	Cn1cc[n+](COC(C)C(C)(C)C)c1C=NO	4	3.15
311	Cn1cc[n+](COC(C)(C)C=C)c1C=NO	4	1.43
312	Cn1cc[n+](COC(C)(C)C(C)(C)C)c1C=NO	4	4.01
313	Cn1cc[n+](COC(C)(C)C#CC)c1C=NO	4	1.79
314	Cn1cc[n+](COC(C)(C)C#C)c1C=NO	4	1.27
315	Cn1cc[n+](COC(C(C)=C)CC(C)=C)c1C=NO	4	3.27
316	Cn1cc[n+](COC(C(C)(C)C)CC)c1C=NO	4	3.79
317	Cn1cc[n+](COC(=O)C=CC)c1C=NO	4	0.99
318	Cn1cc[n+](COC(=O)C(C)CC)c1C=NO	4	2.09
319	CCn1cc[n+](COCCCCl)c1C=NO	4	1.39
320	CCn1cc[n+](COCC(C)CBr)c1C=NO	4	2.63
321	CCn1cc[n+](COCC(C)(C)CBr)c1C=NO	4	3.37
322	CCCN1cc[n+](COCC(C)(C)CBr)c1C=NO	4	4.29
323	CCCN1cc[n+](COCC(C)(C)C)c1C=NO	4	3.95
325	CCCCCN1cc[n+](COCCCCCC)c1C=NO	4	8.55
326	CCCCCN1cc[n+](COCC(C)(C)C)c1C=NO	4	7.01

327	CCCCCn1cc[n+](COC)c1C=NO	4	3.35
328	CCCCCn1cc[n+](COC(C)C(C)(C)C)c1C=NO	4	7.37
329	CC=Cn1cc[n+](COCCCC#C)c1C=NO	4	1.77
330	CC=Cn1cc[n+](COCC(C)(C)C)c1C=NO	4	3.95
331	CC=Cn1cc[n+](COC)c1C=NO	4	-0.13
332	CC=Cn1cc[n+](COC(CC)C#C)c1C=NO	4	2.63
335	C#CCCCCn1cc[n+](COC)c1C=NO	4	1.09
336	Cn1cc[n+](COCCCN(=O)=O)c1C=NO	4	-0.79
337	Cn1cc[n+](COC(CC)C(C)N(=O)=O)c1C=NO	4	4.29
338	Cn1cc[n+](COC(CC)C(C)(C)N(=O)=O)c1C=NO	4	1.59
339	Cn1cc[n+](COC(C)CCCN(=O)(=O)c1C=NO	4	0.29
340	Cn1cc[n+](COC(C)C(CC)N(=O)=O)c1C=NO	4	1.09
341	Cn1cc[n+](COC(C)C(C)(C)N(=O)=O)c1C=NO	4	1.01
342	CCn1cc[n+](COCC(C)(C)N(=O)=O)c1C=NO	4	0.97
344	CCn1cc[n+](COC(C)C(C)(C)N(=O)=O)c1C=NO	4	1.43
345	CCn1cc[n+](COC(CC)C(C)(C)N(=O)=O)c1C=NO	4	3.25
349	Cn1cc[n+](COC(C)C2CCCCC2)c1C=NO	4	4.95
352	Cn1cc[n+](COC2CCC=CC2)c1C=NO	4	1.67
353	Cn1cc[n+](COC2CCCCC2C#C)c1C=NO	4	2.37
355	Cn1cc[n+](COCC2CCCC2)c1C=NO	4	2.87
356	Cn1cc[n+](COCC2CCCCC2)c1C=NO	4	4.27
375	Cn1cc[n+](COCc2ccc(F)cc2)c1C=NO	4	2.27
379	Cn1cc[n+](CCOCCF)c1C=NO	4	-0.73
380	Cn1cc[n+](CCOCCCl)c1C=NO	4	-0.13
383	Cn1cc[n+](CCOCCBr)c1C=NO	4	0.11
384	Cn1cc[n+](CCOC)c1C=NO	4	-1.45
386	CC(C)(C)COCn1cc[n+](C)c1C=NO	4	2.69
387	CC(OCn1cc[n+](C)c1C=NO)C(C)(C)C	4	3.15
388	CC(C)OCn1cc[n+](C)c1C=NO	4	0.17
389	CCCCOCn1cc[n+](C)c1C=NO	4	1.73
390	CCCCCCCCOCn1cc[n+](C)c1C=NO	4	6.27
393	Cn1cc[n+](CC(=O)C)c1C=NO	4	-0.89
394	CCn1cc[n+](CC(=O)C)c1C=NO	4	0.07
396	c1ccccc1Cn2cc[n+](CC(=O)C)c2C=NO	4	1.99
397	Cn1cc[n+](COC(=O)C(C)(C)C)c1C=NO	4	1.79
398	Cn1cc[n+](COC(=O)CCC)c1C=NO	4	1.45
400	Cn1cc[n+](CSCCC#C)c1C=NO	4	0.63
408	Cn1cc[n+](CCS(=O)(=O)C)c1C=NO	5	-2.49
409	Cn1cc[n+](CCCS(=O)(=O)C)c1C=NO	5	-0.89

410	Cn1cc[n+](CCCCS(=O)(=O)C)c1C=NO	5	-2.29
413	CCn1cc[n+](CCN(C)S(=O)(=O)C)c1C=NO	5	-0.71
414	Cn1cc[n+](CCCN(C)S(=O)(=O)C)c1C=NO	5	-0.75
415	Cn1cc[n+](CCCN(CC)S(=O)(=O)C)c1C=NO	5	-0.87
416	Cn1cc[n+](CCN(C)S(=O)(=O)C)c1C=NO	5	-2.11
417	Cn1cc[n+](CCN(CC)S(=O)(=O)C)c1C=NO	5	-0.93
450	Cn1cc[n+](COC(C)(C)CS(=O)(=O)C)c1C=NO	5	-1.29
451	CCn1cc[n+](COC(C)(C)CS(=O)(=O)C)c1C=NO	5	-1.29
452	CCCN1cc[n+](COC(C)(C)CS(=O)(=O)C)c1C=NO	5	-0.49
453	Cn1cc[n+](COCCS(=O)(=O)C)c1C=NO	5	-2.09
454	Cn1cc[n+](COCCCS(=O)(=O)C)c1C=NO	5	-0.49
455	Cn1cc[n+](COCCCCS(=O)(=O)C)c1C=NO	5	-1.71
456	CCCN1cc[n+](COCCS(=O)(=O)C)c1C=NO	5	-1.49
457	CCn1cc[n+](COC(C)C(C)S(=O)(=O)C)c1C=NO	5	-0.85
458	CCn1cc[n+](COCCS(=O)(=O)C)c1C=NO	5	-1.49
459	CCn1cc[n+](COCCS(=O)(=O)CC)c1C=NO	5	-0.83
460	Cn1cc[n+](CCOCCS(=O)(=O)C)c1C=NO	5	-1.63
462	Cn1cc[n+](COC(C)CS(=O)(=O)C)c1C=NO	5	-0.81
463	Cn1cc[n+](COC(C)CS(=O)(=O)CC)c1C=NO	5	-0.89
464	Cn1cc[n+](COCCS(=O)(=O)CC)c1C=NO	5	-0.89
470	CCC[n+]1cccc(c1)C(=O)N	6	1.65
471	CCCC[n+]1cccc(c1)C(=O)N	6	1.73
473	CCCCCCC[n+]1cccc(c1)C(=O)N	6	3.17
475	CCCCCCCC[n+]1cccc(c1)C(=O)N	6	5.43
481	Cn1cc[n+](c1C=NO)CCCN(C)S(=O)(=O)C(F)(F)F	7	1.77
484	Cn1cc[n+](c1C=NO)CCN(C)S(=O)(=O)C(F)(F)F	7	1.51
487	Cn1cc[n+](c1C=NO)CCNS(=O)(=O)C(F)(F)F	7	2.11
541	CC[N+](C)(CC)CCSC(=NO)C(=O)c1ccc(OC)cc1	9	1.31
542	CC[N+](C)(CC)CCSC(=NO)C(=O)c1ccc(F)cc1	9	2.17

C.5 Absolv-estimated (AAP) and Quantum-chemically Derived (QCAP) Abraham Solute Parameters

Table C-4a. Absolv-estimated and quantum-chemically estimated solute descriptors for carboxylate anions. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solute	Absolv Abraham Solute Parameters (AAPs)					Quantum Chemical Abraham Solute Parameters (QCAPs)				
	E	S	A	B	V	E	S	A	B	V
23DibromopropanoicAcid	0.92	1.07	0.79	0.46	0.96	1.12	5.22	-0.00289	2.61	0.96
26DibromobenzoicAcid	1.39	1.39	0.57	0.38	1.28	1.86	6.20	-0.00343	2.91	1.28
26DichlorobenzoicAcid	1.02	1.22	0.57	0.38	1.18	1.56	6.34	-0.00357	2.80	1.18
26DiethoxybenzoicAcid	0.85	1.26	0.57	0.82	1.61	1.57	5.84	-0.00353	3.41	1.61
26DifluorobenzoicAcid	0.56	1.00	0.57	0.39	0.97	1.18	6.34	-0.00370	2.66	0.97
26DimethoxybenzoicAcid	0.86	1.26	0.57	0.81	1.33	1.47	6.82	-0.00410	3.58	1.33
26DimethylbenzoicAcid	0.80	0.96	0.57	0.44	1.21	1.43	5.98	-0.00336	3.10	1.21
26DinitrobenzoicAcid	1.29	2.22	0.57	0.64	1.28	1.45	6.61	-0.00441	2.54	1.28
26DipropoxybenzoicAcid	0.85	1.27	0.57	0.83	1.89	1.67	5.74	-0.00353	3.42	1.89
2CyanobenzoicAcid	0.92	1.56	0.57	0.59	1.09	1.41	6.30	-0.00372	2.78	1.09
2NitrobenzoicAcid	1.02	1.65	0.57	0.54	1.11	1.33	6.23	-0.00382	2.71	1.11
34DichlorobenzoicAcid	1.02	1.24	0.73	0.43	1.18	1.49	5.41	-0.00297	2.68	1.18
34DimethoxybenzoicAcid	0.89	1.68	0.57	0.90	1.33	1.38	6.11	-0.00369	3.41	1.33

34DimethylbenzoicAcid	0.80	0.96	0.57	0.44	1.21	1.41	6.35	-0.00296	1.79	1.21
34DinitrobenzoicAcid	1.29	2.22	0.73	0.64	1.28	1.54	5.47	-0.00364	2.38	1.28
35DimethylbenzoicAcid	0.80	0.96	0.57	0.44	1.21	1.43	6.30	-0.00293	1.78	1.21
35DinitrobenzoicAcid	1.29	2.22	0.68	0.64	1.28	1.52	5.32	-0.00357	2.49	1.28
3Bromo4methoxybenzoicAcid	1.12	1.34	0.64	0.64	1.31	1.60	5.86	-0.00338	3.06	1.31
3Bromo4methylbenzoicAcid	1.10	1.18	0.64	0.44	1.25	1.62	5.83	-0.00324	2.92	1.25
3BromobenzoicAcid	1.08	1.24	0.64	0.44	1.11	1.54	5.69	-0.00312	2.81	1.11
3Chloro4nitrobenzoicAcid	1.17	1.73	0.73	0.54	1.23	1.57	5.34	-0.00322	2.58	1.23
3ChlorobenzoicAcid	0.90	1.16	0.64	0.44	1.05	1.38	5.70	-0.00316	2.77	1.05
3CyanobenzoicAcid	0.92	1.56	0.64	0.59	1.09	1.44	5.77	-0.00333	2.75	1.09
3FluorobenzoicAcid	0.67	1.04	0.64	0.45	0.95	1.17	5.74	-0.00324	2.78	0.95
3Methoxy4chlorobenzoicAcid	0.94	1.25	0.66	0.64	1.25	1.46	5.90	-0.00342	3.02	1.25
3Methoxy4methylbenzoicAcid	0.84	1.11	0.57	0.66	1.27	1.42	6.23	-0.00364	3.23	1.27
3Methoxy4nitrobenzoicAcid	1.08	1.74	0.66	0.75	1.31	1.55	5.75	-0.00356	3.00	1.31
3MethoxybenzoicAcid	0.81	1.17	0.57	0.66	1.13	1.32	6.15	-0.00360	3.14	1.13
3Methyl4chlorobenzoicAcid	0.92	1.10	0.66	0.44	1.20	1.45	5.76	-0.00321	2.89	1.20
3Methyl4methoxybenzoicAcid	0.84	1.11	0.57	0.66	1.27	1.39	6.20	-0.00363	3.31	1.27
3Methyl4nitrobenzoicAcid	1.04	1.59	0.66	0.54	1.25	1.54	5.54	-0.00334	2.80	1.25
3MethylbenzoicAcid	0.77	1.02	0.57	0.44	1.07	1.35	6.13	-0.00345	2.99	1.07
3Nitro4chlorobenzoicAcid	1.17	1.73	0.73	0.54	1.23	1.52	5.47	-0.00331	2.62	1.23
3Nitro4methoxybenzoicAcid	1.08	1.74	0.64	0.75	1.31	1.48	5.85	-0.00366	3.02	1.31
3Nitro4methylbenzoicAcid	1.04	1.59	0.64	0.54	1.25	1.49	5.78	-0.00351	2.87	1.25
3NitrobenzoicAcid	1.02	1.65	0.64	0.54	1.11	1.40	5.61	-0.00340	2.73	1.11
3TrifluoromethylbenzoicAcid	0.41	0.85	0.64	0.33	1.13	1.14	5.52	-0.00320	2.76	1.13
4BromobenzoicAcid	1.08	1.24	0.66	0.44	1.11	1.53	5.63	-0.00307	2.83	1.11
4ChlorobenzoicAcid	0.90	1.16	0.66	0.44	1.05	1.38	5.64	-0.00311	2.80	1.05
4CyanobenzoicAcid	0.92	1.56	0.66	0.59	1.09	1.48	5.68	-0.00325	2.73	1.09

4FluorobenzoicAcid	0.67	1.04	0.66	0.45	0.95	1.15	5.70	-0.00321	2.83	0.95
4HydroxybenzoicAcid	0.98	1.29	1.00	0.72	0.99	1.25	6.27	-0.00355	3.44	0.99
4MethoxybenzoicAcid	0.81	1.17	0.57	0.66	1.13	1.32	6.12	-0.00357	3.22	1.13
4MethylbenzoicAcid	0.77	1.02	0.57	0.44	1.07	1.36	6.16	-0.00346	3.01	1.07
4NitrobenzoicAcid	1.02	1.65	0.66	0.54	1.11	1.46	5.43	-0.00326	2.68	1.11
AceticAcid	0.17	0.61	0.57	0.36	0.46	0.67	5.85	-0.00311	3.21	0.46
ButanoicAcid	0.17	0.62	0.57	0.36	0.75	0.76	5.71	-0.00309	3.25	0.75
CyanoaceticAcid	0.36	1.17	0.79	0.53	0.62	0.67	5.85	-0.00346	2.64	0.62
IsobutanoicAcid	0.18	0.61	0.57	0.39	0.75	0.75	5.57	-0.00302	3.25	0.75
PentanoicAcid	0.17	0.63	0.57	0.37	0.89	0.81	5.72	-0.00314	3.27	0.89
PropanoicAcid	0.17	0.62	0.57	0.36	0.61	0.70	5.70	-0.00305	3.26	0.61

Table C-4b. Absolv-estimated and quantum-chemically estimated solute descriptors for quaternary amine cations. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solute	Absolv Abraham Solute Parameters (AAPs)					Quantum Chemical Abraham Solute Parameters (QCAPs)				
	E	S	A	B	V	E	S	A	B	V
1	-0.2083	0.0955	0	0.0596	0.7935	0.184556	-6.49476	0.004011	0.624775	0.7935
2	-0.2101	0.0999	0	0.0627	0.9344	0.201055	-6.18583	0.003864	0.818594	0.9344
3	-0.2136	0.1086	0	0.0689	1.2162	0.298326	-5.9187	0.003749	1.28615	1.2162
4	-0.2154	0.113	0	0.072	1.3571	0.356677	-5.8177	0.003712	1.523929	1.3571
5	-0.2172	0.1174	0	0.0751	1.498	0.412785	-5.72488	0.003677	1.76162	1.498
6	-0.2207	0.1261	0	0.0813	1.7798	0.52874	-5.54673	0.003613	2.236354	1.7798
7	-0.2243	0.1348	0	0.0875	2.0616	0.645067	-5.37382	0.003552	2.710224	2.0616
8	-0.2349	0.1611	0	0.1061	2.907	0.996293	-4.86299	0.003374	4.131715	2.907
9	-0.0888	0.2419	0.0007	0.0622	1.0568	0.320766	-6.98709	0.004342	0.858737	1.0568
10	0.0036	0.1885	0	0.0791	1.3894	0.678862	-5.8962	0.003834	1.640474	1.3894
12	-0.2291	0.042	0	0.1033	1.2162	0.21948	-5.71487	0.00369	1.181486	1.2162
13	-0.2039	0.0928	0	0.0965	1.2162	0.248253	-5.76366	0.003695	1.225195	1.2162
14	0.0116	0.1771	0	0.1098	1.5303	0.823532	-5.48756	0.003617	1.892155	1.5303
15	0.1366	0.2347	0	0.1615	1.4873	0.905325	-5.70799	0.003817	1.787527	1.4873
17	-0.1496	0.5042	0	0.3964	1.2906	0.275958	-6.8866	0.004665	0.995877	1.2906
18	-0.0274	0.5658	0	0.4504	1.3885	0.472526	-6.68618	0.004631	1.17666	1.3885
19	0.0339	0.597	0	0.4162	1.4042	0.645543	-6.74033	0.005111	0.572549	1.4042
20	0.1546	0.728	0.0007	0.3961	1.4656	0.475769	-7.29136	0.004951	1.230336	1.4656
22	-0.2497	0.5013	0.0007	0.3625	1.3082	0.258251	-7.60917	0.005069	0.917893	1.3082

24	-0.1727	0.575	0	0.5458	1.3493	0.258442	-6.98221	0.004863	0.999859	1.3493
25	0.0449	0.3386	0	0.2884	1.3072	0.472957	-6.67077	0.004436	1.24331	1.3072
26	0.0449	0.3386	0	0.2884	1.3072	0.586181	-4.99689	0.00336	1.491826	1.3072
27	0.0528	0.3271	0	0.3191	1.4481	0.870606	-5.58858	0.003785	1.63235	1.4481
28	0.0528	0.3271	0	0.3191	1.4481	0.51225	-5.90048	0.003886	1.334503	1.4481
29	0.094	0.4772	0	0.5283	1.3659	0.444605	-6.50776	0.004363	0.907533	1.3659
30	0.0861	0.4887	0	0.4976	1.225	1.243841	-5.04426	0.003626	1.38664	1.225
31	0.3905	0.4859	0	0.4118	1.3298	0.682111	-5.91098	0.00391	1.234502	1.3298
33	-0.1724	0.2587	0	0.2781	1.134	0.251233	-6.31412	0.004079	0.806675	1.134
34	-0.1663	0.2516	0	0.312	1.4158	0.346635	-6.01791	0.003898	1.291879	1.4158
35	-0.1742	0.2631	0	0.2812	1.2749	0.314068	-6.15265	0.003959	1.078343	1.2749
36	-0.176	0.2674	0	0.2843	1.4158	0.362329	-6.05446	0.003916	1.321986	1.4158
39	0.3852	0.5117	0	0.1557	1.824	1.241347	-5.43446	0.003879	2.27478	1.824
40	0.3949	0.4959	0	0.1833	1.824	1.221169	-5.44086	0.003891	2.242893	1.824
46	0.4492	0.9072	0	0.4832	1.8984	1.323982	-5.74396	0.004363	2.182082	1.8984
48	0.451	0.9029	0	0.4801	1.7575	1.281326	-5.85284	0.004456	1.942782	1.7575
50	0.4264	0.6618	0	0.365	1.7418	1.225269	-5.71508	0.004154	2.017509	1.7418
51	0.4282	0.6574	0	0.3619	1.6009	1.12843	-5.88444	0.004264	1.759683	1.6009
56	0.537	1.043	0	0.2936	1.556	1.262686	-6.88153	0.004856	1.635489	1.556
64	0.4323	0.633	0	0.4334	1.5011	1.030661	-5.98217	0.004268	1.438393	1.5011
68	0.171	0.3806	0	0.1847	1.3621	0.751318	-5.72701	0.003921	1.526858	1.3621
70	0.1467	0.4382	0	0.184	1.2212	2.653273	-5.26667	0.003958	1.336688	1.2212
71	0.171	0.3806	0	0.1847	1.3621	0.877995	-5.81002	0.00408	1.620897	1.3621
72	0.2923	0.5226	0.0007	0.1842	1.4845	0.84263	-6.17124	0.004172	1.63206	1.4845
81	-0.2136	0.1086	0	0.0689	1.2162	0.234054	-5.60783	0.003599	1.208228	1.2162
82	-0.219	0.1217	0	0.0782	1.6389	0.390048	-5.16852	0.003413	1.892367	1.6389
83	-0.2243	0.1348	0	0.0875	2.0616	0.55389	-4.83764	0.003282	2.597638	2.0616

84	-0.2154	0.113	0	0.072	1.3571	0.249432	-5.34265	0.003483	1.394988	1.3571
85	-0.2207	0.1261	0	0.0813	1.7798	0.4058	-4.91604	0.003309	2.071849	1.7798
116	0.4922	0.6886	0	0.18	3.0736	2.253855	-4.44794	0.003493	4.434574	3.0736
121	0.3879	0.5926	0	0.1202	1.2604	2.454071	-4.93623	0.003785	1.048036	1.2604
124	0.3772	0.6188	0	0.1388	2.1058	3.213351	-4.31975	0.003597	2.497943	2.1058
136	0.433	0.4862	0	0.1278	1.824	1.128871	-4.58685	0.003265	2.419792	1.824
138	0.426	0.5037	0	0.1402	2.3876	1.386563	-4.23215	0.003142	3.362684	2.3876
139	0.4224	0.5124	0	0.1464	2.6694	1.508869	-4.06681	0.003086	3.832051	2.6694
140	0.4189	0.5212	0	0.1526	2.9512	1.632671	-3.90541	0.003039	4.298865	2.9512
146	0.4224	0.5124	0	0.1464	2.6694	2.369819	-3.97948	0.00321	3.597726	2.6694
165	0.9659	1.0445	0	0.2002	1.4455	15.18341	-3.21226	0.005011	1.542758	1.4455
203	0.5884	1.2282	0	0.7162	1.6361	0.750355	-6.70119	0.004889	1.905282	1.6361
204	0.6127	1.1707	0	0.7169	1.777	0.883441	-6.51196	0.004968	1.3213	1.777
205	0.6207	1.1592	0	0.7476	1.9179	0.834546	-6.50681	0.004984	1.378539	1.9179
212	0.4903	1.2514	0	0.9668	1.9776	1.523063	-6.64422	0.005448	1.36427	1.9776
213	0.4921	1.2471	0	0.9637	1.8367	1.698259	-7.17284	0.005825	1.080282	1.8367
215	0.4903	1.2514	0	0.9668	1.9776	1.835082	-5.81677	0.004902	1.532367	1.9776
220	0.9491	1.2876	0.3133	0.8335	1.3916	1.130759	-7.12002	0.004928	0.318739	1.3916
221	0.9473	1.2919	0.3133	0.8366	1.5325	1.179768	-6.90556	0.004794	0.605941	1.5325
222	0.9456	1.2963	0.3133	0.8397	1.6734	1.288191	-6.71805	0.00472	0.865671	1.6734
223	0.8058	0.8939	0.3126	0.8739	1.4365	1.169141	-5.16655	0.00367	0.555876	1.4365
224	0.8022	0.9026	0.3126	0.8801	1.7183	1.287711	-4.96643	0.003596	1.043543	1.7183
225	0.7987	0.9114	0.3126	0.8863	2.0001	1.402918	-4.79349	0.003544	1.508746	2.0001
226	0.7951	0.9201	0.3126	0.8925	2.2819	1.52074	-4.62272	0.003483	1.980247	2.2819
228	1.0646	1.0604	0.3978	0.9641	1.6323	1.312289	-5.4838	0.004049	0.771221	1.6323
229	1.0628	1.0648	0.3978	0.9672	1.7732	1.183054	-5.20106	0.003821	0.921361	1.7732
230	1.061	1.0692	0.3978	0.9703	1.9141	1.255604	-5.07506	0.003761	1.183277	1.9141

232	1.0213	0.9781	0.3126	0.8872	1.7506	1.654364	-5.3825	0.003963	1.145833	1.7506
233	1.0195	0.9825	0.3126	0.8903	1.8915	12.6136	-2.87341	0.004168	1.723605	1.8915
237	1.0549	1.0763	0.3978	0.9365	1.6323	1.330972	-5.89084	0.004291	0.73557	1.6323
238	1.0531	1.0806	0.3978	0.9396	1.7732	1.380355	-5.85646	0.004302	0.908883	1.7732
240	1.3767	1.3633	0.3126	0.9663	2.1852	2.783342	-5.26247	0.004207	1.830115	2.1852
242	0.9657	1.2235	0.3126	0.9201	1.9747	1.295112	-6.29054	0.004762	1.219855	1.9747
244	0.964	1.2279	0.3126	0.9232	2.1156	1.407272	-6.15364	0.00467	1.53988	2.1156
245	0.6679	0.7321	0.3133	0.6246	1.2545	1.046326	-6.214	0.004169	0.301937	1.2545
246	0.8893	0.877	0.3133	0.658	1.3593	1.20878	-6.08804	0.004116	0.485038	1.3593
247	0.8822	0.8945	0.3133	0.6704	1.9229	1.454888	-5.56565	0.003898	1.466639	1.9229
248	1.0136	0.9262	0.3978	0.7273	1.7145	1.404223	-5.70316	0.004121	1.095709	1.7145
249	1.0705	0.9633	0.3133	0.6613	1.5528	1.390951	-5.91295	0.004088	0.777966	1.5528
250	1.0723	0.9589	0.3133	0.6582	1.4119	1.332227	-5.94423	0.004081	0.540242	1.4119
252	0.9295	0.9305	0.3126	1.1632	1.9003	1.432946	-5.09432	0.003891	1.110886	1.9003
253	0.9313	0.9262	0.3126	1.1601	1.7594	1.362638	-5.18925	0.003894	0.807652	1.7594
254	0.941	0.9103	0.3126	1.1877	1.7594	1.334238	-5.14484	0.0039	0.653741	1.7594
255	0.9374	0.9191	0.3126	1.1939	2.0412	1.455424	-4.95849	0.003802	1.242484	2.0412
256	0.9471	0.9032	0.3126	1.2215	2.0412	1.398999	-4.6651	0.003559	1.583364	2.0412
257	0.9392	0.9147	0.3126	1.1908	1.9003	1.359706	-4.86018	0.003648	1.33582	1.9003
258	0.9313	0.9262	0.3126	1.1601	1.7594	1.349185	-5.22205	0.003942	0.732181	1.7594
259	0.9454	0.9076	0.3126	1.2246	2.1821	1.38822	-4.86418	0.003753	1.572036	2.1821
261	1.18	1.172	0.3126	1.3519	1.8504	8.615069	-5.12845	0.005238	0.137449	1.8504
262	1.1388	1.0219	0.3126	1.1427	1.9326	6.666358	-4.52012	0.004408	1.045819	1.9326
263	1.1406	1.0175	0.3126	1.1396	1.7917	1.660535	-5.23926	0.004049	1.007936	1.7917
264	1.1503	1.0017	0.3126	1.1672	1.7917	2.001327	-4.90194	0.003771	1.460681	1.7917
271	0.7056	0.891	0.3133	0.8401	1.4541	1.145081	-5.82782	0.004034	0.497285	1.4541
272	0.9252	1.0402	0.3133	0.8765	1.6998	1.34795	-5.67321	0.003989	0.910166	1.6998

273	0.9235	1.0446	0.3133	0.8796	1.8407	1.407422	-5.54669	0.003931	1.138461	1.8407
274	1.0531	1.0806	0.3978	0.9396	1.7732	1.425196	-5.52384	0.004079	1.024628	1.7732
275	1.1082	1.1221	0.3133	0.8767	1.7524	1.46056	-5.55565	0.003963	0.970279	1.7524
276	0.9545	1.0101	0.3126	0.9496	1.9571	1.576635	-4.89358	0.003665	1.454787	1.9571
277	0.9261	0.9599	0.3126	0.931	1.6753	1.403134	-5.16751	0.003794	0.94325	1.6753
278	0.8496	1.0456	0.3126	1.1232	1.9179	1.313225	-5.10909	0.00381	0.92355	1.9179
280	0.9032	0.9074	0.3126	0.9501	1.8162	1.441307	-5.10586	0.003816	1.11026	1.8162
281	0.9475	1.0371	0.3126	0.9514	1.7732	1.51002	-5.23352	0.003901	1.117576	1.7732
283	1.11	1.1177	0.3133	0.8736	1.6115	1.416409	-5.58413	0.003969	0.737684	1.6115
284	0.9545	1.0101	0.3126	0.9496	1.9571	1.61475	-4.83272	0.003644	1.455513	1.9571
285	0.9563	1.0057	0.3126	0.9465	1.8162	1.545562	-4.92647	0.00369	1.197426	1.8162
286	0.9581	1.0013	0.3126	0.9434	1.6753	1.47899	-5.03381	0.003726	0.959573	1.6753
287	0.9373	0.9478	0.3126	0.9871	2.098	1.648438	-4.77875	0.003682	1.617738	2.098
288	0.9279	0.9555	0.3126	0.9279	1.5344	1.350762	-5.26909	0.003847	0.671904	1.5344
289	0.9351	0.9488	0.3126	0.9625	1.8162	1.541078	-4.90597	0.003693	1.153083	1.8162
290	1.1161	1.1106	0.3133	0.9075	1.8933	1.493501	-5.44106	0.003952	1.148798	1.8933
291	0.905	0.9031	0.3126	0.947	1.6753	1.397529	-5.11016	0.003757	0.919892	1.6753
292	0.9062	0.9823	0.3133	0.9172	1.9816	1.402995	-5.45539	0.003947	1.259942	1.9816
293	1.0892	1.0642	0.3133	0.9174	2.0342	1.526815	-5.35211	0.003925	1.323687	2.0342
294	0.785	0.8404	0.3126	0.9177	1.8592	1.304583	-4.88526	0.00359	1.180654	1.8592
295	1.0358	1.0184	0.3978	0.9771	1.9141	1.457763	-5.2897	0.003975	1.208093	1.9141
296	0.8075	0.8895	0.3126	0.8708	1.2956	1.108922	-5.34429	0.003796	0.274052	1.2956
297	0.9314	1.0331	0.3133	0.9103	1.9816	1.413832	-5.44964	0.003919	1.334341	1.9816
298	1.0593	1.0735	0.3978	0.9734	2.055	1.342355	-5.38791	0.004046	1.365549	2.055
299	0.8084	0.8955	0.3126	0.9139	2.0001	1.243732	-4.91812	0.003619	1.421281	2.0001
300	0.9323	0.9528	0.3126	0.9649	1.9571	1.458179	-4.74863	0.003581	1.401972	1.9571
301	0.9314	1.0331	0.3133	0.9103	1.9816	1.449704	-5.42689	0.003908	1.345396	1.9816

302	0.8101	0.8912	0.3126	0.9108	1.8592	1.321025	-4.84956	0.003571	1.238022	1.8592
303	0.7911	0.8333	0.3126	0.9515	2.141	1.374949	-4.68418	0.003572	1.622458	2.141
304	1.0628	1.0648	0.3978	0.9672	1.7732	1.391565	-5.38906	0.004018	1.01773	1.7732
305	0.9359	0.9441	0.3126	0.9587	1.6753	1.39454	-5.0656	0.00375	0.939573	1.6753
306	0.8137	0.8824	0.3126	0.9046	1.5774	1.142667	-5.07644	0.00365	0.766354	1.5774
308	1.0707	1.0533	0.3978	0.9979	1.9141	1.452532	-5.28701	0.004009	1.211013	1.9141
309	0.7929	0.8289	0.3126	0.9484	2.0001	1.324072	-4.76908	0.003589	1.340157	2.0001
311	0.9089	0.8977	0.3126	0.9686	1.8162	1.392355	-4.95679	0.003716	1.09858	1.8162
312	0.7659	0.7825	0.3126	0.9583	2.141	1.318898	-4.62549	0.003551	1.555234	2.141
313	0.9303	0.9749	0.3126	0.9889	1.9141	1.513815	-4.97521	0.003779	1.33517	1.9141
314	1.0376	1.014	0.3978	0.974	1.7732	1.365782	-5.30845	0.003977	0.999779	1.7732
315	1.0103	0.9052	0.3126	1.0539	2.1959	1.581191	-4.78238	0.003766	1.668231	2.1959
316	0.7911	0.8333	0.3126	0.9515	2.141	1.235194	-4.88402	0.003666	1.498434	2.141
317	0.9827	1.2424	0.3126	1.0585	1.691	1.50179	-5.42199	0.00419	0.859868	1.691
318	0.8348	1.1323	0.3126	1.026	1.8749	1.260121	-5.46036	0.004124	1.082182	1.8749
319	0.9235	1.0446	0.3133	0.8796	1.8407	1.456747	-5.52265	0.003884	1.209116	1.8407
320	1.1144	1.115	0.3133	0.9106	2.0342	1.627334	-5.31286	0.003861	1.459657	2.0342
321	1.0874	1.0686	0.3133	0.9205	2.1751	1.653549	-5.1981	0.003827	1.653003	2.1751
322	1.0856	1.073	0.3133	0.9236	2.316	1.710779	-5.09295	0.003774	1.900379	2.316
323	0.7814	0.8491	0.3126	0.9239	2.141	1.484063	-4.63448	0.003441	1.764803	2.141
325	0.7898	0.9332	0.3126	0.9018	2.7046	1.826953	-4.26443	0.003294	2.81416	2.7046
326	0.7761	0.8622	0.3126	0.9331	2.5637	1.715166	-4.34555	0.003326	2.464671	2.5637
327	0.7987	0.9114	0.3126	0.8863	2.0001	1.507173	-4.81754	0.003548	1.561301	2.0001
328	0.784	0.8508	0.3126	0.9639	2.7046	1.622926	-4.44632	0.003458	2.525216	2.7046
329	1.2054	1.1881	0.3978	1.0091	2.012	2.42971	-5.08982	0.003971	1.759007	2.012
330	0.9373	0.9478	0.3126	0.9871	2.098	2.311713	-4.47097	0.003478	1.928037	2.098
331	0.9599	0.9969	0.3126	0.9403	1.5344	2.125019	-4.91228	0.003688	1.00384	1.5344

332	1.2151	1.1722	0.3978	1.0367	2.012	2.202141	-4.80501	0.003752	1.662444	2.012
335	1.0513	1.085	0.3978	0.9427	1.9141	1.582945	-5.54038	0.004152	1.23599	1.9141
336	1.0242	1.4359	0.3133	1.0918	1.7516	1.282714	-6.42643	0.004875	0.709873	1.7516
337	1.0401	1.4129	0.3133	1.1532	2.0334	1.295533	-6.05417	0.004657	1.183319	2.0334
338	1.0131	1.3665	0.3126	1.1632	2.1743	1.30979	-5.82096	0.004541	1.419778	2.1743
339	1.0304	1.4288	0.3133	1.1256	2.0334	1.379237	-6.11715	0.004752	1.153689	2.0334
340	1.0401	1.4129	0.3133	1.1532	2.0334	1.316459	-6.07788	0.004676	1.180508	2.0334
341	1.0149	1.3621	0.3126	1.1601	2.0334	1.264892	-5.95502	0.004564	1.239035	2.0334
342	1.0052	1.378	0.3126	1.1325	2.0334	1.380732	-5.99234	0.004579	1.320033	2.0334
344	1.0131	1.3665	0.3126	1.1632	2.1743	1.397231	-5.83772	0.004505	1.522904	2.1743
345	1.0096	1.3753	0.3126	1.1694	2.4561	1.488149	-5.58742	0.004402	1.93875	2.4561
349	1.0257	0.9754	0.3126	0.9241	2.1733	7.92669	-2.83726	0.003457	2.125354	2.1733
352	1.1754	1.0812	0.3126	0.9536	1.8485	1.37188	-5.00926	0.0037	1.187565	1.8485
353	1.2783	1.1491	0.3978	0.9805	2.0873	1.8312	-5.27732	0.004239	1.576982	2.0873
355	1.0195	0.9825	0.3126	0.8903	1.8915	1.698889	-5.11642	0.00379	1.408561	1.8915
356	1.0177	0.9869	0.3126	0.8934	2.0324	8.186711	-3.0734	0.00353	1.89307	2.0324
375	1.2975	1.3234	0.3126	0.9632	1.921	2.240505	-5.28483	0.004119	1.366775	1.921
379	0.7038	0.8953	0.3133	0.8432	1.595	1.172791	-5.92335	0.004116	0.702339	1.595
380	0.9252	1.0402	0.3133	0.8765	1.6998	1.318803	-5.81068	0.004067	0.881706	1.6998
383	1.1082	1.1221	0.3133	0.8767	1.7524	1.435524	-5.68647	0.004043	0.93991	1.7524
384	0.8058	0.8939	0.3126	0.8739	1.4365	1.144104	-5.48828	0.003893	0.448352	1.4365
386	0.785	0.8404	0.3126	0.9177	1.8592	1.452559	-5.56649	0.003966	1.068587	1.8592
387	0.7929	0.8289	0.3126	0.9484	2.0001	1.467937	-5.48147	0.003967	1.210132	2.0001
388	0.8137	0.8824	0.3126	0.9046	1.5774	1.335484	-5.77359	0.004042	0.609374	1.5774
389	0.8022	0.9026	0.3126	0.8801	1.7183	1.413266	-5.71987	0.004034	0.835121	1.7183
390	0.7951	0.9201	0.3126	0.8925	2.2819	1.664232	-5.36828	0.003906	1.841848	2.2819
393	0.9785	1.2314	0.3126	0.9654	1.3935	1.045038	-5.69228	0.004119	0.356569	1.3935

394	0.975	1.2401	0.3126	0.9716	1.6753	1.236848	-5.46467	0.004007	0.909349	1.6753
396	1.5512	1.6964	0.3126	1.0546	2.0013	2.994743	-5.98135	0.004809	1.433298	2.0013
397	0.8096	1.0815	0.3126	1.0328	1.8749	1.236579	-5.46069	0.004116	1.059364	1.8749
398	0.8268	1.1437	0.3126	0.9952	1.734	1.04931	-5.29358	0.003946	0.744721	1.734
400	1.3593	1.0735	0.3978	0.8507	1.7371	1.870092	-5.54579	0.004113	1.163763	1.7371
408	0.8661	1.6513	0.3126	1.1975	1.6587	1.302543	-6.71663	0.005106	0.146486	1.6587
409	0.8644	1.6557	0.3126	1.2006	1.7996	1.34931	-6.569	0.005051	0.314891	1.7996
410	0.8626	1.6601	0.3126	1.2037	1.9405	1.458855	-6.66205	0.005104	0.525192	1.9405
413	1.2469	1.6252	0.3126	1.4944	2.0403	1.139227	-5.70692	0.004732	1.137208	2.0403
414	1.2469	1.6252	0.3126	1.4944	2.0403	1.191542	-4.92667	0.004068	0.664357	2.0403
415	1.2451	1.6296	0.3126	1.4975	2.1812	1.417299	-5.8373	0.004715	0.943273	2.1812
416	1.2487	1.6208	0.3126	1.4913	1.8994	1.022209	-5.66214	0.004556	0.546973	1.8994
417	1.2469	1.6252	0.3126	1.4944	2.0403	1.417616	-5.4896	0.004496	0.843951	2.0403
450	0.8848	1.7523	0.3126	1.4536	2.1401	1.106209	-5.17296	0.004394	1.002621	2.1401
451	0.883	1.7566	0.3126	1.4567	2.281	1.557294	-6.08001	0.004864	0.97562	2.281
452	0.8812	1.761	0.3126	1.4598	2.4219	1.434411	-5.61728	0.004529	1.139602	2.4219
453	0.9038	1.8101	0.3126	1.413	1.8583	1.430072	-6.53898	0.005096	0.286583	1.8583
454	0.902	1.8145	0.3126	1.4161	1.9992	1.463013	-6.40869	0.005019	0.454557	1.9992
455	0.9003	1.8189	0.3126	1.4192	2.1401	1.532947	-6.32061	0.004992	0.661615	2.1401
456	0.9003	1.8189	0.3126	1.4192	2.1401	1.588999	-6.28515	0.004972	0.818537	2.1401
457	0.9179	1.7916	0.3126	1.4775	2.281	1.455654	-5.95806	0.00472	0.905077	2.281
458	0.902	1.8145	0.3126	1.4161	1.9992	1.517196	-6.42183	0.005018	0.614951	1.9992
459	0.9003	1.8189	0.3126	1.4192	2.1401	1.579283	-6.1629	0.004838	0.835766	2.1401
460	0.902	1.8145	0.3126	1.4161	1.9992	1.433492	-6.57704	0.005122	0.474282	1.9992
462	0.9118	1.7987	0.3126	1.4437	1.9992	1.405093	-5.20399	0.00429	0.739843	1.9992
463	0.91	1.803	0.3126	1.4468	2.1401	1.460454	-4.93301	0.004093	1.017009	2.1401
464	0.902	1.8145	0.3126	1.4161	1.9992	1.466376	-6.27672	0.00491	0.474784	1.9992

470	0.8301	1.4431	0.4867	0.6626	1.3759	1.255935	-6.19054	0.00454	0.955789	1.3759
471	0.8283	1.4475	0.4867	0.6657	1.5168	1.325123	-6.07924	0.004492	1.204833	1.5168
473	0.823	1.4606	0.4867	0.675	1.9395	1.524091	-5.79609	0.004398	1.90468	1.9395
475	0.8194	1.4694	0.4867	0.6812	2.2213	1.657981	-5.60961	0.004318	2.388682	2.2213
481	0.8845	1.4544	0.3126	1.3792	2.0931	1.256502	-6.15785	0.004915	1.146397	2.0931
484	0.8863	1.45	0.3126	1.3761	1.9522	1.290823	-5.83948	0.004693	0.933882	1.9522
487	0.8696	1.4386	0.6777	1.3213	1.8113	1.191368	-6.58724	0.005107	0.52496	1.8113
541	1.3484	1.4689	0.3126	1.2361	2.6001	2.852686	-5.01208	0.004404	2.596906	2.6001
542	1.2018	1.3409	0.3126	1.0262	2.4181	2.399542	-5.65737	0.004579	2.474822	2.4181

C.6 Quantum-chemically Estimated Solvent Parameters

Table C-5a. Quantum-chemically estimated Abraham pp-LFER solvent parameters for carboxylate anions. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solvent	c	e	s	a	b	v
1-Butanol	-0.32044	-2.17634	2.624035	3936.017	-2.4174	5.848432
1-Decanol	-0.04756	-1.58809	1.858492	3793.296	-2.65493	5.054496
1-Heptanol	-0.14345	-1.9145	2.303832	3886.482	-2.51556	5.495225
1-Hexanol	-0.1759	-1.98897	2.390087	3899.59	-2.48916	5.615172
1-Pentanol	-0.25729	-2.0988	2.535157	3924.65	-2.44464	5.757557
1-Propanol	-0.37914	-2.27243	2.719263	3944.1	-2.38895	6.01365
2-butanol	-0.18589	-1.98634	2.282172	3353.801	-2.68917	5.687153
2-Methyl-1-propanol	-0.26963	-2.18668	2.601927	3925.266	-2.42536	5.971466
2-Methyl-2-propanol	-0.00761	-1.82237	1.957743	3041.366	-2.8653	5.56581
2-propanol	-0.27272	-2.09668	2.396394	3362.691	-2.65719	5.851656
Acetone	-0.16295	-0.14102	0.452425	-112.951	-4.54624	4.415176
Acetonitrile	-0.89742	-0.24593	0.896735	268.2544	-4.3808	4.332818
benzene	0.647501	1.966997	-3.32413	-1954.7	-5.64639	3.856019
Benzonitrile	-0.09946	0.510923	0.276115	-446.538	-4.81535	3.99372
bromobenzene	1.12777	1.229664	-1.20804	-992.744	-5.18902	3.808773
Butanone	0.250861	0.131948	0.092935	-578.114	-4.78545	4.261356
ButylEthanoate	1.330472	0.730527	-1.35219	-1097.4	-5.08638	3.673036
Carbondisulfide	0.926093	2.089769	-2.87846	-1667.62	-5.56254	3.385737
chlorobenzene	1.06549	1.140638	-1.11953	-982.114	-5.17853	3.932716
Cyclohexane	-0.10573	2.205864	-3.72687	-2196.4	-5.79045	3.502487
Cyclohexanone	0.31523	0.311475	0.004256	-547.742	-4.77909	3.656055
Dibutylether	1.29667	1.202219	-2.45456	-1557.34	-5.32472	3.413708
Dichloromethane	-0.16005	0.195904	0.225912	358.4059	-4.47331	4.590186
Diethylether	1.460332	0.76748	-1.68037	-1268.13	-5.17309	4.034315
DiMethylSulfoxide	-0.221	-0.07943	0.479649	-360.456	-4.57249	3.249277
Ethanol	-0.45776	-2.3781	2.809941	3946.431	-2.36153	6.149062
Ethylbenzene	0.784377	1.873502	-3.10119	-1842.12	-5.58647	3.61028
EthylEthanoate	1.197113	0.568034	-1.04432	-991.51	-5.02405	3.833404
Fluorobenzene	1.166944	1.020402	-1.20967	-1051.84	-5.18819	4.202454
Formamide	-5.60132	-2.81724	5.171808	7865.366	-0.89417	4.868627
Heptane	-0.33214	2.212819	-3.92243	-2322.5	-5.84138	3.754099

iodobenzene	1.261387	1.447731	-1.51897	-1076.4	-5.24202	3.559617
Methanol	-0.86446	-2.81857	3.433221	4979.78	-1.86987	6.579922
Methylcyclohexane	-0.06967	2.183776	-3.71488	-2194.38	-5.78815	3.579812
MethylEthanoate	1.040372	0.485389	-0.83948	-917.636	-4.98218	3.811248
n-Decane	-0.17895	2.201715	-3.78455	-2235.49	-5.80512	3.557708
n-dodecane	-0.14058	2.209999	-3.74527	-2207.63	-5.79474	3.480358
n-hexadecane	-0.0842	2.213901	-3.68429	-2167.19	-5.77867	3.368941
n-hexane	-0.40133	2.211126	-3.98044	-2361.35	-5.85673	3.844322
Nitrobenzene	-0.29823	0.568191	0.393057	-397.454	-4.81199	3.975403
Nitromethane	-0.95028	-0.06522	0.838688	191.912	-4.43806	3.884816
n-Octane	-0.26721	2.208727	-3.86629	-2286.79	-5.82661	3.675821
n-Octanol	-0.09431	-1.81723	2.170483	3860.088	-2.557	5.370939
n-Pentane	-0.51768	2.21588	-4.07133	-2420.52	-5.88079	3.968854
n-undecane	-0.14136	2.245676	-3.77122	-2213.09	-5.8023	3.546668
p-Xylene	0.597722	1.969208	-3.32226	-1948.35	-5.6365	3.636256
tetrahydrofuran	0.998286	0.506729	-0.72717	-847.027	-4.94909	3.740029
Toluene	0.746269	1.902983	-3.1805	-1884.72	-5.6104	3.742993
Tributylphosphate	1.654517	-0.13791	-0.60057	-727.298	-4.66801	4.084365

Table C-5b. Quantum-chemically estimated Abraham pp-LFER solvent parameters for quaternary amine cations. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solvents	c	e	s	a	b	v
1-Butanol	0.086394	-0.2516	1.57019	2537.21	-2.33653	0.061158
1-Decanol	-0.54912	0.005477	0.87728	2300.591	-2.89266	1.487473
1-Heptanol	-0.20416	-0.14101	1.292654	2459.574	-2.56946	0.643944
1-Hexanol	-0.13148	-0.17106	1.369663	2482.79	-2.50765	0.460951
1-Pentanol	0.007828	-0.22178	1.499145	2524.569	-2.40068	0.204663
1-Propanol	0.170124	-0.286	1.648011	2550.036	-2.26964	-0.1516
2-butanol	0.314065	-0.199	1.442411	2294.183	-2.46411	0.157443
2-Methyl-1-propanol	0.055304	-0.24793	1.556461	2532.929	-2.35715	-0.03236
2-Methyl-2-propanol	0.232765	-0.11346	1.2173	2092.251	-2.65805	0.45242
2-propanol	0.399875	-0.2382	1.53062	2303.919	-2.38726	-0.06158
Acetone	0.239496	0.120831	0.357818	-122.015	-4.39689	3.343337
Acetonitrile	-0.09991	-0.03907	0.717218	381.6149	-4.04031	2.937232
Benzene	0.476013	0.898975	-2.64718	-1101.17	-5.67287	6.224204
Benzonitrile	0.273089	0.083948	0.433227	-154.706	-4.70817	3.815221
bromobenzene	-0.8916	0.463818	-0.75664	-327.582	-5.21239	4.80049
Butanone	0.333222	0.216264	0.095644	-586.571	-4.77604	3.972227
ButylEthanoate	-0.22375	0.619371	-1.12848	-926.561	-5.20764	5.347741
Carbondisulfide	0.070398	0.781524	-2.1175	-688.464	-5.56461	6.096593
chlorobenzene	-1.02279	0.447557	-0.71894	-355.614	-5.20684	4.71857
Cyclohexane	0.481943	0.984845	-3.12831	-1335.23	-5.77731	7.054959
Cyclohexanone	0.549612	0.254114	0.063548	-538.218	-4.75707	4.425213
Dibutylether	0.268865	0.848345	-2.07754	-1253.35	-5.46001	6.285707
Dichloromethane	-1.34518	0.125556	0.243112	785.1812	-4.22351	2.812261
Diethylether	-0.34654	0.682415	-1.42528	-1048.5	-5.32387	5.305723
DiMethylSulfoxide	1.726988	0.174007	0.356415	-722.886	-4.51685	4.601029
Ethanol	0.233298	-0.31491	1.706034	2538.565	-2.20808	-0.30473
Ethylbenzene	0.234709	0.873261	-2.47882	-1065.94	-5.63035	6.336789
EthylEthanoate	-0.28425	0.550279	-0.88819	-877.121	-5.13394	5.057452
Fluorobenzene	-1.05117	0.481841	-0.84649	-488.718	-5.24761	4.636869
Formamide	-0.09859	-0.57443	2.574091	3947.219	-2.05678	1.508407
Heptane	0.697035	1.021513	-3.33423	-1459.92	-5.81915	7.026069
iodobenzene	-0.63843	0.525804	-0.95851	-325.595	-5.26289	5.079513
Methanol	0.073012	-0.38393	1.88728	2815.634	-2.03348	-0.47659
Methylcyclohexane	0.455576	0.980981	-3.11552	-1331.92	-5.77894	6.979229
MethylEthanoate	-0.28674	0.504704	-0.73536	-849.471	-5.07816	4.987058
n-Decane	0.532625	0.998187	-3.19715	-1383.65	-5.79094	7.077838
n-dodecane	0.498988	0.990597	-3.15339	-1353.51	-5.78056	7.101234

n-hexadecane	0.438172	0.980132	-3.09114	-1315.89	-5.76597	7.140974
n-hexane	0.762119	1.032203	-3.39601	-1498.61	-5.83194	7.005409
Nitrobenzene	0.284937	0.026315	0.566229	-46.1707	-4.6724	3.721904
Nitromethane	-0.01634	-0.01497	0.664663	268.8075	-4.15242	3.477575
n-Octane	0.629704	1.0118	-3.27759	-1427.69	-5.80771	7.044288
n-Octanol	-0.31963	-0.096	1.170156	2415.089	-2.66726	0.883851
n-Pentane	0.874585	1.048574	-3.49068	-1555.02	-5.85022	6.982626
n-undecane	0.575902	0.985959	-3.14681	-1319.45	-5.77985	7.006029
p-Xylene	0.520526	0.912563	-2.6726	-1141.37	-5.66346	6.44986
tetrahydrofuran	-0.09418	0.468548	-0.59229	-755.406	-5.02295	4.876178
Toluene	0.295402	0.880183	-2.53649	-1074.19	-5.64914	6.257485
Tributylphosphate	1.986488	0.537025	-0.47493	-1084.83	-4.80597	4.23261

C.7 Complete AAP & QCAP Abraham pp-LFERs

Table C-6a. Full AAP Abraham pp-LFER including all solvent parameters for carboxylate anions estimated with experimental solvent-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solvent	Parameter	Value	SE	CV (%)
Acetone	c	-13.36	1.132	8.48%
	e	0.33	0.870	261.00%
	s	1.19	0.551	46.21%
	a	4.07	1.529	37.56%
	b	-6.59	1.137	17.25%
	v	3.19	0.841	26.39%
Acetonitrile	c	-12.73	1.020	8.01%
	e	0.89	0.762	85.51%
	s	1.00	0.489	49.02%
	a	4.45	1.300	29.24%
	b	-5.31	1.028	19.36%
	v	3.37	0.763	22.65%
DMSO	c	-0.54	2.333	434.98%
	e	7.48	1.926	25.74%
	s	-0.52	1.130	218.97%
	a	-10.10	3.041	30.09%
	b	-0.08	2.745	3633.36%
	v	-4.21	1.994	47.41%
Methanol	c	-0.79	0.707	89.26%
	e	2.06	0.521	25.27%
	s	-0.93	0.328	35.13%
	a	-1.45	0.895	61.73%
	b	-3.48	0.686	19.70%
	v	1.82	0.500	27.43%

Table C-6b. Full QCAP Abraham pp-LFER including all solvent parameters for carboxylate anions estimated with experimental solvent-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solvent	Parameter	Value	SE	CV (%)
Acetonitrile	c	1.01	1.504	148%
	e	1.32	0.838	63%
	s	-0.32	0.549	171%
	a	1382.89	648.695	47%
	b	-2.35	0.487	21%
	v	1.57	1.018	65%
Acetone	c	0.24	1.468	607%
	e	2.98	0.631	21%
	s	-1.70	0.289	17%
	a	-682.84	377.611	55%
	b	-1.20	0.245	21%
	v	-0.05	0.710	1298%
DMSO	c	0.56	1.424	253%
	e	2.59	0.812	31%
	s	-0.24	0.542	228%
	a	319.73	637.693	199%
	b	-3.24	0.486	15%
	v	1.70	1.031	61%
Methanol	c	0.07	1.175	1776%
	e	1.89	0.509	27%
	s	0.03	0.226	892%
	a	1280.85	286.283	22%
	b	-0.05	0.188	390%
	v	0.63	0.552	88%

Table C-6c. Full AAP Abraham pp-LFER including all solvent parameters for quaternary amine cations estimated with experimental octanol-water partition coefficients. Parameters are for the Abraham equation of the form $\log(K_{\text{octanol-water}}) = aA + bB + eE + sS + vV + c$, where $K_{\text{octanol-water}}$ has units of (L water / kg octanol).

Solvent	Parameter	Value	SE	CV(%)
Octanol	c	-7.6711	0.3641	4.75%
	e	0.7746	0.3719	48.01%
	s	-2.0915	0.3509	16.78%
	a	9.6952	0.9294	9.59%
	b	-3.8985	0.4285	10.99%
	v	5.9321	0.2174	3.66%

C.8 Residual Plots of A Priori QC Computed and AAP –predicted Partition Coefficients

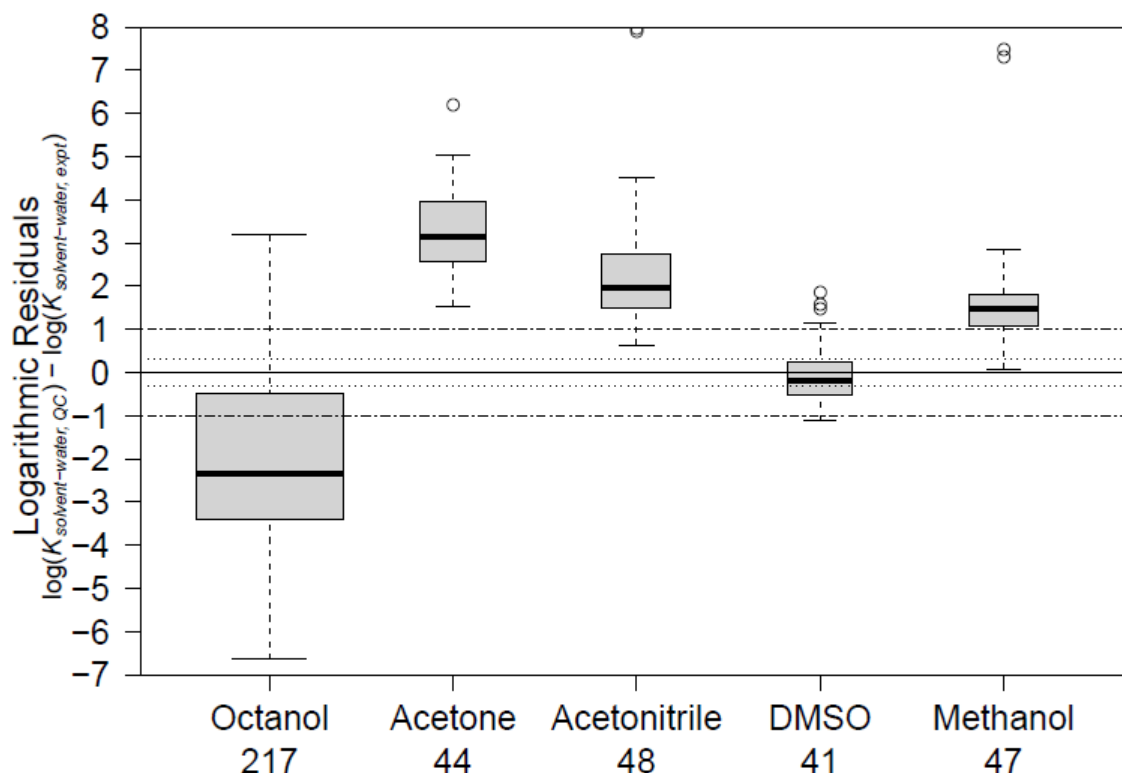


Figure C-1a. Logarithmic residual box plots for a priori QC computed solvent-water partition coefficients. Boxes correspond to the interquartile range (IQR), where 25% < IQR < 75%. Whiskers correspond to ± 1.5 IQR. Points correspond to values > 1.5 IQR. Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10).

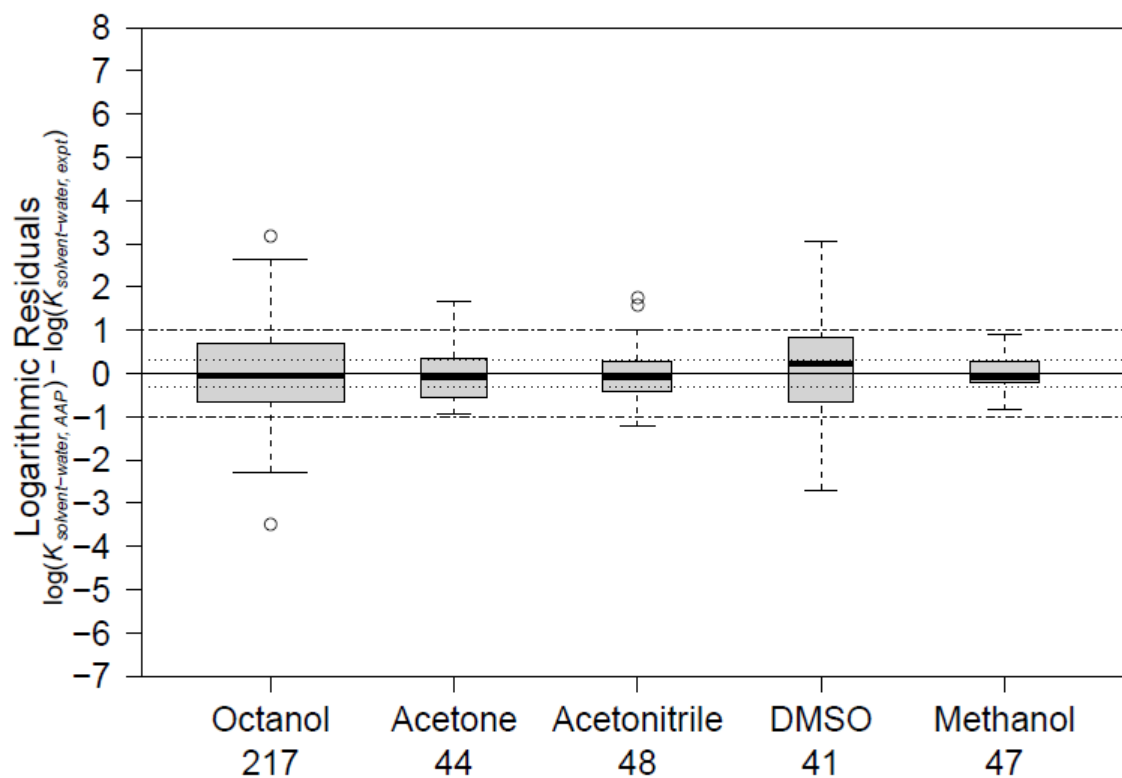


Figure C-1b. Logarithmic residual box plots for AAP-predicted solvent-water partition coefficients. Boxes correspond to the interquartile range (IQR), where $25\% < \text{IQR} < 75\%$. Whiskers correspond to $\pm 1.5\text{IQR}$. Points correspond to values $> 1.5\text{IQR}$. Solid line indicates perfect agreement between predicted and experimental values (1:1), dashed lines correspond to ± 0.3 log-units (a factor of 2), semi-dashed lines correspond to ± 1.0 log-units (a factor of 10).

Appendix D

PREDICTING PARTITIONING OF CHARGED ORGANIC SPECIES TO NATURAL ORGANIC MATTER USING QUANTUM CHEMICALLY-DERIVED ABRAHAM PP-LFER SOLUTE PARAMETERS

D.1. Complete List of Solute Names and ID Abbreviations

D.2. LNL Isotherm Visual Basic Code & Abraham Solute / Solvent Parameter Optimization (QCAP) R Code

D.3. Estimated Log-normal Langmuir (LNL) Isotherm Parameters for Primary through Quaternary Amine Cations

D.4. Legend for Isotherms (Eq. (5-15) for Figure 5-1

D.5. Absolv (AAP) and Quantum-chemically Estimated (QCAP) Solute Parameters for Primary through Quaternary Amine Cations

D.6. Quantum-chemically Estimated (QCAP) Solvent-water System Parameters for Primary through Quaternary Amine Cations

D.1 Complete List of Solute Names and ID Abbreviations

Table D-1. Complete list of solute names and ID abbreviations.

Solute ID	Solute Name
P01	(±)-1-Aminoindane
P02	3,4-Dichlorobenzylamine
P03	4-Chlorobenzylamine
P04	Benzylamine
P05	4-Methylbenzylamine
P06	4-Butylbenzylamine
P07	4-Octylbenzylamine
P08	D-Amphetamine sulphate
P09	2-Phenylethylamine
P10	3-Phenylpropylamine
P11	4-Phenylbutylamine
P12	1-Hexylamine
P13	1-Heptylamine
P14	1-Octylamine
P15	1-Decylamine
P16	Serotonin.HCl
P17	Tryptamine.HCl
P18	1-Naphthylmethylamine
S01	N-ethyl-M-toluidine
S02	Prilocaine.HCl
S03	(±)-Metoprolol (+)-tartrate
S04	Propranolol.HCl
S05	R-Atenolol
S06	Alprenolol.HCl
S07	N-Benzyl-N-ethanolamine
S08	N-Benzyl-N-octylamine
S09	N-Benzyl-N-hexylamine.HCl
S10	N-Benzyl-N-butylamine
S11	N-Benzyl-N-ethylamine
S12	N-Benzyl-N-methylamine
S13	3-Meth-N-Meth.Benz.amine
S14	L-Adrenaline.HCl
S15	Fluoxetine.HCl
S16	(+)-Methamphetamine.HCl
S17	N-Methyl-Phenethylamine

S18	Dibenzylamine
S19	N-Benzylaminoacetaldehydediethylacetal
S20	N-Benz--alanine-eth.ester
T01	Quinoline
T02	Pyridine
T03	2-Phenylbenzimidazole
T04	Benzimidazole
T05	Thiabendazole
T06	N,N-Diethylaniline
T07	2-Ethylpyridine
T08	2-Methylpyridine
T09	3,4-Dimethylpyridine
T10	2,6-Dimethylpyridine
T11	2,4,6-Trimethylpyridine
T12	Scopolamine.HCl
T13	Ropivacaine.HCl
T14	Bupivacaine.HCl
T15	Lidocaine
T16	Procaine.HCl
T17	Atropine
T18	Imipramine.HCl
T19	4-Amino-2-methylquinoline
T20	N-Benzyltrimethylamine
T21	Methylephedrine
T22	(±)-Verapamil.HCl
T23	Codeine.HCl
T24	Clonidine.HCl
T25	(S)-(-)Nicotine
T26	3-Dimethylaminopropiophenone.HCl
T27	S-(-)-1-benzyl-3-acetamidopyrrolidine
T28	2,2'-(benzylimino)-diethanol
Q01	Octyltrimethylammonium.Cl
Q02	Phenyltrimethylammonium.Cl
Q03	Benzyltributylammonium.Cl
Q04	Benzyltripropylammonium.Cl
Q05	Benzyltriethylammonium.Cl
Q06	Benzyltrimethylammonium.Cl
Q07	Benzyltrimethylhexylamm.Cl
Q08	Benzyltrimethyloctylamm.Cl
Q09	Benzyltrimethyldecylamm.Cl
Q10	Benzyltrimethyldodecylamm.Cl
Q11	Benzyl(2-hydroxyethyl)dimethylamm.Cl
Q12	Difenzoquatmethylsulfate

Q13	Acetylcholine.Cl
Q14	Butyrylthiocholine.I

D.2 LNL Isotherm Visual Basic Code & Abraham Solute / Solvent Parameter Optimization (QCAP) R Code

D2a. Abraham QCAP solute / solvent optimization code (R)

```
#define RMSE
RMSE=function(bx){
  ss=cumsum(bx^2)
  sl=length(bx)
  RMSE=(ss[sl]/sl)^0.5
}

#read solvents and solutes files
rawsolv <- read.csv(file="Solvents.csv")
dat <- data.frame(rawsolv)
#dat
solv <- cbind(dat$c,dat$e,dat$s,dat$a,dat$b,dat$v)
#solv
solv1 <- cbind(dat$e,dat$s,dat$a,dat$b,dat$v)
#solv1
solv2 <- cbind(dat$s,dat$a,dat$b)
#solv2
rawsolu <- read.csv(file="Solutes.csv")
dat1 <- data.frame(rawsolu)
#dat1
solu<-cbind(dat1$E_Gaussian,dat1$S,dat1$A,dat1$B,dat1$V_McGowan)
#solu
solu1<-cbind(1,dat1$E_Gaussian,dat1$S,dat1$A,dat1$B,dat1$V_McGowan)
#solu1
solu2<-cbind(dat1$S,dat1$A,dat1$B)
#solu2

#logKij
logKij<-solv%>%t(solu1)
logKij

#logK_Absolv
SMD = read.csv(file="Eout.csv")
solvents = unique(SMD$Solvent)
solutes = unique(SMD$Solute)
logp = matrix(nrow=length(solvents), ncol=length(solutes))

for (i in 1:length(dat$Solvent)){
  logp[i,] = t(SMD[SMD$Solvent==dat$SOL[[i]],6])
}

#Initiate
b=logp-logKij
RMSEb=NULL

#loop
ybold=RMSE(b)
for(i in 1:200){
  xu<-solve(t(solv2)%>%solv2,t(solv2)%>%b )
```

```

solu2=solu2+t(xu)
solu1<-cbind(1, dat1$E_Gaussian, solu2[,1:3],dat1$V_McGowan)
logKnew<-solv%%t(solu1)
b=logp-logKnew

xv<-solve(t(solu1)%t(solu1),t(solu1)%t(b))

solv=solv+t(xv)
solv2<-solv[,3:5]
solv3<-cbind(solv[,3:5])
logKnew<-solv%%t(solu1)
b=logp-logKnew

ybnew=RMSE(b)
print(c(ybold,ybnew),digits=22)
if(abs(ybold-ybnew)<10^-30)break()
ybold=ybnew
RMSEb=c(RMSEb,ybnew)
}
write.table(solu1,file='Optimized_Solute_Parameters.csv',sep=',',row.names=TRUE,col.names=NA)
write.table(solv,file='Optimized_Solvent_Parameters.csv',sep=',',row.names=TRUE,col.names=NA)

```

D2b. Log-normal Langmuir isotherm code (Visual Basic)⁸¹

```
Function LNLfunc(k, qtot, sigma, conc)

Dim weights
weights = Array(2.22939364553415E-13, 4.39934099227318E-10, 1.08606937076928E-07,
7.80255647853206E-06, 2.28338636016354E-04, 3.24377334223786E-03,
2.48105208874636E-02, 0.109017206020023, 0.286675505362834, 0.46224366960061,
0.46224366960061, 0.286675505362834, 0.109017206020023, 2.48105208874636E-02,
3.24377334223786E-03, 2.28338636016354E-04, 7.80255647853206E-06,
1.08606937076928E-07, 4.39934099227318E-10, 2.22939364553415E-13)

Dim nodes
nodes = Array(-5.38748089001123, -4.60368244955074, -3.94476404011563, -
3.34785456738322, -2.78880605842813, -2.25497400208928, -1.73853771211659, -
1.23407621539532, -0.737473728545394, -0.245340708300901, 0.245340708300901,
0.737473728545394, 1.23407621539532, 1.73853771211659, 2.25497400208928,
2.78880605842813, 3.34785456738322, 3.94476404011563, 4.60368244955074,
5.38748089001123)

Dim func(0 To 19)

For i = 0 To 19
func(i) = ((conc * Exp(k + sigma * nodes(i) * (Sqr(2)))) / (1 + conc * Exp(k + sigma * nodes(i) *
(Sqr(2)))))
Next i

Dim runsum
runsum = 0

For i = 0 To 19
runsum = runsum + func(i) * weights(i)
Next i

LNLfunc = (10^(qtot) / (Sqr(3.141592654))) * runsum
End Function
```

D.3 Estimated Log-normal Langmuir (LNL) Isotherm Parameters for Primary through Quaternary Amine Cations

Table D-2. Estimated log-normal Langmuir (LNL) isotherm parameters and coefficients of variation (CV %) and linear partition coefficients ($\log(K_p)$) and associated standard errors for primary through quaternary amine cations. Units of $\log(\tilde{K}_L)$ are (L/mmol). For P15 and S14, coefficients of variation were not calculated as the standard errors are orders of magnitude larger than the estimated values. $\log(K_p)$ values are obtained via the pp-LFER equation for particulate organic carbon⁷ (Eq. (5-13)).

Solute	$\log(\tilde{K}_{L,+})$	SE($\log(\tilde{K}_{L,+})$)	CV (%)	$\log(K_p)$
P01	-1.65	0.76	46.3%	0.89
P02	-0.56	0.81	144.9%	0.83
P03	-1.32	0.79	60.1%	0.97
P04	-1.91	0.80	41.7%	0.68
P05	-1.55	0.77	50.0%	0.76
P06	-0.86	0.77	88.7%	0.98
P07	0.09	0.78	857.9%	1.27
P08	-1.79	0.79	44.5%	0.81
P09	-1.82	0.80	43.7%	0.75
P10	-1.57	0.76	48.5%	0.83
P11	-1.35	0.76	56.4%	0.89
P12	-2.00	0.77	38.2%	0.48
P13	-1.80	0.83	46.3%	0.56
P14	-1.11	0.86	77.7%	0.63
P15	4.92	[-]	[-]	0.78
P16	-0.73	0.75	103.0%	0.92
P17	-0.85	0.77	90.6%	1.05
P18	-0.64	0.76	119.0%	1.22
S01	-1.64	0.77	46.8%	1.06
S02	-2.10	0.77	36.5%	0.88
S03	-1.86	0.78	41.8%	1.38
S04	-0.54	0.76	141.1%	1.27
S05	-1.84	0.75	41.1%	0.66
S06	-1.33	0.77	57.6%	0.96
S07	-1.95	0.76	38.9%	0.72
S08	-0.80	0.79	98.7%	1.33
S09	-1.53	0.78	51.1%	1.18
S10	-1.80	0.77	42.9%	1.03

S11	-1.94	0.76	39.4%	0.90
S12	-1.95	0.76	39.1%	0.82
S13	-1.83	0.77	42.0%	0.92
S14	-0.78	[-]	[-]	0.46
S15	-0.67	0.76	113.7%	1.38
S16	-1.80	0.80	44.3%	0.95
S17	-1.77	0.78	43.8%	0.90
S18	-1.19	0.76	63.5%	0.91
S19	-2.00	0.78	39.2%	0.88
S20	-1.85	0.79	42.5%	0.90
T01	-0.50	0.78	155.5%	1.23
T02	-1.89	0.77	40.9%	0.69
T03	-0.08	0.79	951.2%	1.75
T04	-2.19	0.78	35.5%	1.06
T05	-1.01	0.77	75.8%	1.67
T06	-0.10	0.77	769.0%	1.09
T07	-1.87	0.80	42.7%	0.86
T08	-1.97	0.78	39.4%	0.78
T09	-1.17	0.78	66.5%	0.88
T10	-2.17	0.78	35.7%	0.88
T11	-1.88	0.76	40.3%	0.97
T12	-2.17	0.77	35.6%	1.18
T13	-1.82	0.76	42.0%	1.26
T14	-1.88	0.76	40.5%	1.33
T15	-2.28	0.81	35.4%	0.99
T16	-0.96	0.78	81.3%	0.85
T17	-2.06	0.76	36.9%	1.18
T18	-0.70	0.75	108.0%	1.77
T19	-0.16	0.76	466.6%	1.23
T20	-2.16	0.75	35.0%	0.88
T21	-2.10	0.77	36.8%	0.71
T22	-1.00	0.76	76.4%	1.68
T23	-1.78	0.78	43.5%	1.32
T24	-1.77	0.78	43.8%	1.20
T25	-1.93	0.79	41.0%	0.85
T26	-1.42	0.76	53.1%	0.87
T27	-1.74	0.82	47.1%	0.92
T28	-1.92	0.80	41.5%	0.69
Q01	-2.00	0.76	37.8%	
Q02	-2.24	0.77	34.2%	
Q03	-1.83	0.81	44.0%	
Q04	-1.84	0.76	41.2%	
Q05	-2.05	0.80	39.0%	

Q06	-2.04	0.79	38.7%
Q07	-1.27	0.76	59.6%
Q08	-0.79	0.76	96.7%
Q09	-0.25	0.78	310.9%
Q10	0.39	0.83	216.3%
Q11	-1.71	0.77	44.7%
Q12	-0.93	0.78	84.0%
Q13	-3.07	0.76	24.6%
Q14	-2.05	0.76	36.9%
q_{max}	2.74	0.475	15.1%
$\sigma_{\kappa,+}$	2.65	0.414	17.9%

D.4 Legend for Estimated Isotherm (Eq. (5-15) for Figure 5-1

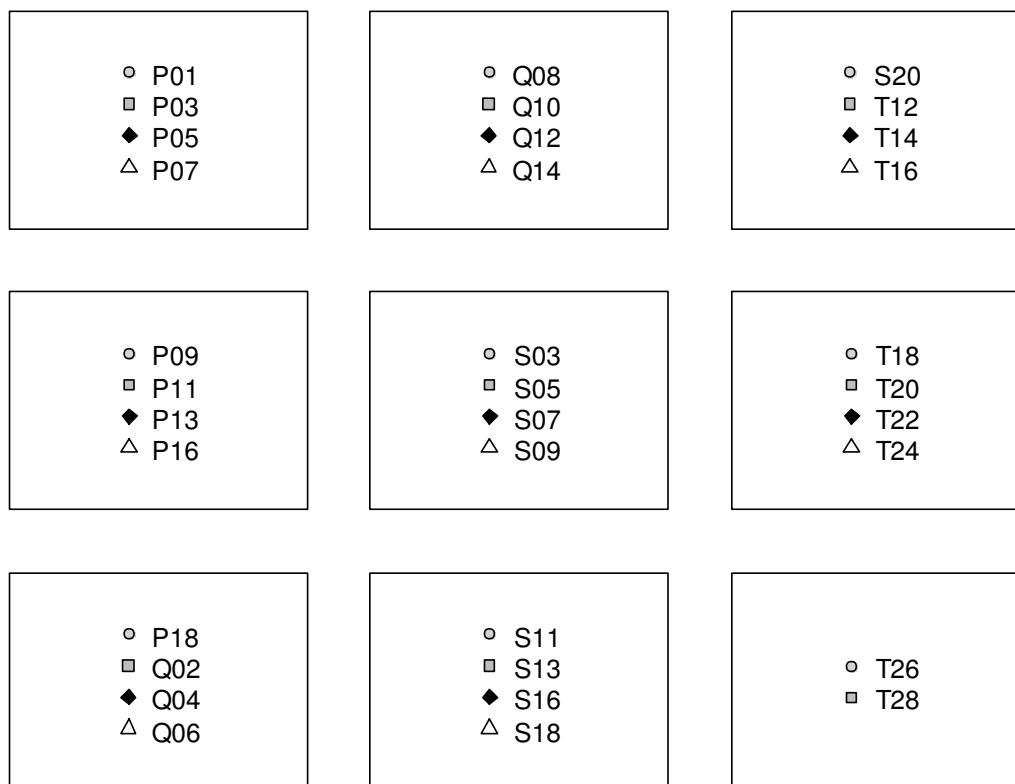


Figure D-1. Legend of sorbates for adsorption data in Figure 5-1. Point type and color correspond to the matching isotherms in Figure 5-1.

D.5 Absolv (AAP) and Quantum-chemically Estimated (QCAP) Solute Parameters for Primary through Quaternary Amine Cations

Table D-3. Absolv-estimated (AAP) and quantum-chemically estimated (QCAP) solute descriptors for quaternary amine cations. AAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eE_i + sS_i + aA_i + bB_i + vV_i + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent). QCAP parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eE_i + sS_i + aA_i + bB_i + vV_i + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solutes	QCAP Estimated Solute Descriptors					AAP Estimated Solute Descriptors				
	E	S	A	B	V	E	S	A	B	V
P01	1.10	8.41	0.00	1.29	1.13	1.02	0.97	0.21	0.68	1.13
P02	1.57	9.57	-0.01	1.40	1.08	0.93	1.03	0.21	0.66	1.08
P03	0.75	9.07	-0.01	1.49	1.20	1.05	1.12	0.21	0.65	1.20
P04	0.88	8.66	0.00	1.30	0.96	0.78	0.95	0.21	0.67	0.96
P05	1.02	8.55	0.00	1.31	1.10	0.80	0.90	0.21	0.67	1.10
P07	1.55	8.76	0.00	1.42	2.08	0.79	0.93	0.21	0.69	2.08
P08	0.94	8.20	0.00	1.33	1.24	0.79	0.95	0.21	0.70	1.24
P09	0.94	9.07	-0.01	1.38	1.10	0.78	0.96	0.21	0.67	1.10
P10	1.00	9.16	-0.01	1.39	1.24	0.78	0.96	0.21	0.67	1.24
P11	1.03	9.01	-0.01	1.39	1.38	0.77	0.97	0.21	0.68	1.38
P12	0.48	8.51	0.00	1.26	1.05	0.20	0.51	0.21	0.59	1.05
P13	0.53	8.54	0.00	1.27	1.19	0.20	0.52	0.21	0.59	1.19
P14	0.56	8.56	0.00	1.28	1.34	0.19	0.52	0.21	0.60	1.34
P15	0.66	8.61	0.00	1.33	1.62	0.19	0.53	0.21	0.60	1.62
P16	1.47	9.60	-0.01	1.99	1.39	1.67	1.58	1.02	1.13	1.39
P17	1.43	9.52	-0.01	1.55	1.33	1.44	1.37	0.52	0.85	1.33
P18	1.59	8.71	0.00	1.42	1.33	1.50	1.28	0.21	0.73	1.33
Q02	0.75	6.98	0.00	0.95	1.26	0.37	0.56	0.00	0.15	1.26
Q05	0.91	6.64	0.00	1.20	1.82	0.36	0.57	0.00	0.16	1.82
Q06	0.83	7.13	0.00	1.02	1.40	0.36	0.56	0.00	0.15	1.40
Q07	1.17	6.96	0.00	1.16	2.11	0.36	0.58	0.00	0.16	2.11
Q08	1.30	7.04	0.00	1.19	2.39	0.35	0.59	0.00	0.17	2.39
Q09	1.43	7.10	0.00	1.23	2.67	0.35	0.60	0.00	0.18	2.67
Q10	1.55	7.16	0.00	1.26	2.95	0.34	0.61	0.00	0.18	2.95
Q11	0.50	7.37	0.00	1.81	1.74	0.56	0.84	0.31	0.43	1.74
Q12	2.46	6.64	0.00	1.29	2.06	1.72	1.30	0.00	0.28	2.06
Q13	0.26	7.80	0.00	1.38	1.29	-0.15	0.50	0.00	0.40	1.29
Q14	0.77	8.63	-0.01	1.46	1.68	0.34	0.76	0.00	0.51	1.68

S05	1.32	8.60	-0.01	2.61	2.18	1.48	1.97	0.78	1.83	2.18
S07	0.97	8.29	0.00	1.71	1.30	0.95	1.04	0.38	0.93	1.30
S11	0.97	7.91	0.00	1.14	1.24	0.73	0.81	0.13	0.57	1.24
S12	0.92	8.10	0.00	1.12	1.10	0.73	0.81	0.13	0.57	1.10
S13	1.02	7.99	0.00	1.12	1.24	0.75	0.75	0.13	0.56	1.24
S14	1.06	8.26	0.00	2.52	1.42	1.35	1.34	1.15	1.48	1.42
S16	0.96	8.25	0.00	1.26	1.38	0.74	0.80	0.13	0.60	1.38
S17	0.96	8.48	0.00	1.19	1.24	0.73	0.81	0.13	0.57	1.24
S18	0.21	7.77	-0.01	1.65	2.22	1.04	1.56	0.13	1.35	2.22
S19	1.11	7.60	0.00	1.73	1.92	0.78	1.14	0.16	1.04	1.92
T01	1.38	7.05	0.00	1.00	1.04	1.32	1.15	0.00	0.44	1.04
T02	0.61	7.34	0.00	0.93	0.68	0.60	0.82	0.00	0.38	0.68
T03	2.31	7.40	0.00	1.12	1.51	2.17	1.84	0.35	0.62	1.51
T04	1.05	7.79	0.00	0.82	0.91	1.34	1.37	0.35	0.52	0.91
T05	2.27	7.28	0.00	1.48	1.40	2.22	1.94	0.35	0.67	1.40
T06	0.85	7.09	0.00	1.09	1.38	0.79	0.92	0.00	0.48	1.38
T07	0.77	6.98	0.00	0.98	0.96	0.62	0.77	0.00	0.38	0.96
T08	0.71	7.04	0.00	0.97	0.82	0.62	0.76	0.00	0.38	0.82
T09	0.77	6.93	0.00	0.96	0.96	0.65	0.71	0.00	0.38	0.96
T10	0.81	6.80	0.00	1.00	0.96	0.65	0.71	0.00	0.38	0.96
T11	0.91	6.61	0.00	1.00	1.10	0.67	0.65	0.00	0.38	1.10
T12	1.20	8.33	-0.01	2.53	2.28	1.46	1.63	0.31	1.33	2.28
T13	1.50	7.62	0.00	1.75	2.37	1.32	1.59	0.26	1.21	2.37
T15	1.12	7.33	0.00	1.53	2.06	1.10	1.50	0.26	1.19	2.06
T16	1.54	8.50	-0.01	1.84	1.98	1.11	1.62	0.23	1.30	1.98
T17	1.38	8.43	-0.01	2.28	2.28	1.46	1.63	0.31	1.33	2.28
T18	2.09	8.67	0.00	1.43	2.40	1.81	1.59	0.00	0.97	2.40
T19	1.57	7.01	0.00	1.28	1.29	1.65	1.48	0.23	0.78	1.29
T20	0.90	7.58	0.00	1.02	1.24	0.75	0.82	0.00	0.62	1.24
T23	1.78	8.49	-0.01	2.44	2.21	2.16	1.92	0.23	1.58	2.21
T24	1.46	7.21	0.00	1.15	1.53	1.55	1.19	0.39	0.90	1.53
T26	1.05	8.06	0.00	1.40	1.54	0.95	1.32	0.00	0.93	1.54
T27	1.24	8.54	-0.01	1.75	1.81	1.25	1.79	0.27	1.19	1.81

D.6 Quantum-chemically Estimated (QCAP) Solvent Parameters for Primary through Quaternary Amine Cations

Table D-4. Quantum-chemically estimated Abraham p-LFER solvent parameters for quaternary amine cations. Parameters are for the Abraham equation of the form $\log(K_{\text{solvent-water}}) = eE_i + sS_i + aA_i + bB_i + vV_i + c$, where $K_{\text{solvent-water}}$ has units of (L water / kg solvent).

Solvent	c	e	s	a	b	v
1-Butanol	-2.21	-0.48	1.63	2511.67	-1.78	3.85
1-Decanol	-1.80	-0.37	1.11	2442.43	-2.48	3.78
1-Heptanol	-2.01	-0.43	1.42	2491.31	-2.08	3.82
1-Hexanol	-2.07	-0.45	1.48	2509.90	-2.00	3.86
1-Pentanol	-2.15	-0.47	1.57	2511.26	-1.86	3.86
1-Propanol	-2.29	-0.50	1.70	2538.15	-1.68	3.91
2-butanol	-2.15	-0.48	1.57	2441.86	-1.90	3.98
2-Methyl-1-propanol	-2.23	-0.48	1.65	2573.18	-1.78	3.98
2-Methyl-2-propanol	-2.05	-0.47	1.46	2440.99	-2.09	4.10
2-propanol	-2.24	-0.50	1.66	2468.11	-1.79	4.02
Acetone	-0.23	0.11	0.41	-43.46	-4.32	4.14
Acetonitrile	-0.09	0.17	0.51	-62.85	-4.24	3.85
benzene	1.24	1.04	-2.61	-904.44	-5.88	4.96
Benzonitrile	0.54	0.41	0.04	-819.01	-5.01	4.07
bromobenzene	1.73	0.74	-0.97	-796.23	-5.73	4.50
Butanone	0.02	0.19	0.17	-434.27	-4.68	4.19
ButylEthanoate	0.80	0.37	-0.86	-448.27	-5.12	4.18
Carbondisulfide	1.81	1.07	-2.41	-1097.26	-6.05	4.60
chlorobenzene	1.70	0.73	-0.88	-767.36	-5.72	4.54
Cyclohexane	1.35	1.22	-3.17	-1298.22	-6.09	4.60
Cyclohexanone	0.15	0.19	-0.03	-616.62	-4.71	3.74
Dibutylether	0.84	0.55	-1.76	-590.49	-5.30	4.22
Dichloromethane	0.77	0.43	0.12	315.05	-4.70	4.51
Diethylether	0.81	0.43	-1.00	-317.57	-5.19	4.57
DiMethylSulfoxide	-0.79	-0.12	0.26	-635.69	-4.10	3.09
Ethanol	-2.37	-0.52	1.77	2553.32	-1.58	3.92
Ethylbenzene	1.38	0.98	-2.47	-942.53	-5.86	4.68
EthylEthanoate	0.68	0.33	-0.61	-406.63	-5.05	4.21
Fluorobenzene	1.54	0.69	-0.83	-605.43	-5.64	4.76
Formamide	-4.16	-0.87	2.38	3077.42	-1.63	2.13

Heptane	1.12	1.25	-3.29	-1250.16	-6.06	4.81
iodobenzene	1.80	0.78	-1.25	-861.09	-5.79	4.43
Methanol	-2.82	-0.62	1.97	2800.20	-1.25	3.82
Methylcyclohexane	1.34	1.21	-3.14	-1260.99	-6.08	4.66
MethylEthanoate	0.63	0.30	-0.48	-443.27	-5.00	4.10
n-Decane	1.29	1.23	-3.21	-1290.52	-6.08	4.63
n-dodecane	1.34	1.22	-3.20	-1312.56	-6.09	4.57
n-hexadecane	1.42	1.21	-3.17	-1342.61	-6.10	4.47
n-hexane	1.04	1.26	-3.32	-1230.09	-6.05	4.88
Nitrobenzene	0.63	0.46	0.05	-928.27	-5.07	4.04
Nitromethane	0.11	0.22	0.34	-350.38	-4.38	3.51
n-Octane	1.19	1.24	-3.26	-1265.64	-6.07	4.74
n-Octanol	-1.95	-0.42	1.33	2485.11	-2.20	3.82
n-Pentane	0.92	1.27	-3.37	-1207.28	-6.03	4.99
n-undecane	1.32	1.24	-3.21	-1296.67	-6.10	4.66
p-Xylene	1.24	1.02	-2.66	-975.45	-5.86	4.74
tetrahydrofuran	0.57	0.28	-0.43	-469.26	-4.96	4.05
Toluene	1.34	1.01	-2.51	-915.19	-5.87	4.82
Tributylphosphate	-1.30	-0.33	0.18	387.02	-3.82	4.29

Appendix E
JOURNAL PERMISSIONS



RightsLink®

Home

Create
Account

Help

ACS Publications
Most Trusted. Most Cited. Most Read.**Title:** Modeling Nonlinear Adsorption to
Carbon with a Single Chemical
Parameter: A Lognormal
Langmuir Isotherm**Author:** Craig Warren Davis, Dominic M.
Di Toro**Publication:** Environmental Science &
Technology**Publisher:** American Chemical Society**Date:** Jul 1, 2015

Copyright © 2015, American Chemical Society

LOGIN

If you're a **copyright.com**
user, you can login to
RightsLink using your
copyright.com credentials.
Already a **RightsLink** user or
want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2016 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).
Comments? We would like to hear from you. E-mail us at customercare@copyright.com



RightsLink®

Home

Create Account

Help

ACS Publications
Most Trusted. Most Cited. Most Read.

Title:

Modeling Nonlinear Adsorption
with a Single Chemical
Parameter: Predicting Chemical
Median Langmuir Binding
Constants

Author:

Craig Warren Davis, Dominic M.
Di Toro

Publication:

Environmental Science &
Technology

Publisher:

American Chemical Society

Date:

Jul 1, 2015

Copyright © 2015, American Chemical Society

LOGIN

If you're a **copyright.com**
user, you can login to
RightsLink using your
copyright.com credentials.
Already a **RightsLink user** or
want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW