Abstract

Chemistry-driven *in silico* methods were used to construct a methodology to elucidate ionic liquids (ILs) that simulate various biological barriers, using a predictive model known as the Abraham solvation parameter model. A PCA (principal component analysis) was performed on experimental data gathered from previous literature containing log P Abraham coefficient correlations of 673 ILs and 29 biological systems. The resulting scatter plot indicated which ILs had the most potential to successfully replace these systems during testing. The liquid, 1-(2-hydroxyethyl)-3-methylimidazolium tris(pentafluoroethyl)trifluorophosphate ($[EtOHMIm]^+[FAP]^-$) was found to simulate several partitioning systems in the human body. We formulated linear equations using log P values of $[EtOHMIm]^+[FAP]^-$ and barriers including the blood-brain, blood-liver, and blood-fat. It was found that the trio could be effectively mimicked by $[EtOHMIm]^+[FAP]^-$. The procedure was rerun for aquatic organisms to produce IL mimics for *Daphnia pulex*, *Carassius auratus*, and *Lepomis macrochirus*, proving our mechanism's versatility. Statistical analysis, employing test and training methods, produced standard deviation (SD), absolute average error (AAE) and average error (AE) values . Overall, this study resulted in computational data that could be further utilized to reduce the millions of dollars allocated to experimental biological testing, particularly drug testing, conducted on animals and humans.

A Novel Methodology to Mimic Biological Properties using Ionic Liquids: An Extensive *in silico* Study

Introduction

Animal Rights and Drug Discovery Process

According to PETA, People for the Ethical Treatment of Animals, at least 100 billion animals are killed during laboratory testing in the U.S. alone [1]. A vast majority of these deaths are unnecessary as more than 92% of medicines found successful during animal trials ultimately fail during clinical testing [2]. This failure is most likely due to the fact that an animal's response is rarely ever precisely compatible to that of a human. Animal testing thereby results in needless animal tormenting, and a low percentage of successful drugs.

Efficiency of the drug-discovery process itself is lacking, as emphasized by the 12.5 years and $1.2 billion to $1.8 billion needed to produce each new drug [3]. The goal of 100 successful drug trials required by the Food and Drug Administration is seldom reached [4]. New drugs in the market are becoming increasingly complex in terms of molecular design, making it burdensome to measure absorption and distribution, which are two important ADMET properties (absorption, distribution, metabolism, excretion, and toxicity) that need to be evaluated [6]. Before testing begins, a drug must pass through numerous phases, including disease identification and drug molecule discovery. Toxicity and solubility data is utilized at the preclinical phase to verify the feasibility of the proposed medication. These stages accumulate to a combined cost of about $267 million dollars and take approximately 3-4 years to complete, a dilemma many researchers and companies are endeavoring to resolve [5].

To combat the high finances, animal abuse, and low rate of success present in clinical testing, we take an interdisciplinary approach between *in silico* and chemical methods. The Abraham model

complements inventions that improve the drug discovery process, such as chemosynthetic livers and microfabricated reactors, while reducing unnecessary steps [7]. The model is noted for being successful in anticipating the solute transfer of over 3000 calculated compounds, predicting bioavailability, and aiding in the identification unsuccessful drugs [6]. It is also able to determine a drug's solute distribution in the bloodstream without using impractical experimental methods [8]. This knowledge is critical in establishing how effective the drug is in eliciting the desired response. Therefore, our group employed the Abraham model's resourceful application in our study to distinguish chemical similarities between ionic liquids and the previously mentioned barriers.

Ionic Liquids:

Ionic liquid (IL) solvents have several physical and chemical properties that differ from conventional solvents, making them a highly significant subject of exploratory study. Made entirely of pure electrolytes, they are liquids at room temperature or lower, do not emit volatile organic compounds due to their absence of measurable vapor pressure, dissolve other compounds easily, and are stable at high temperatures, the latter three properties of which are major complications with organic solvents [9]. Although ILs are slightly more costly than organic solvents, the ecological benefits contributed are worthwhile as environmental rehabilitation can sometimes cost millions of dollars [10]. It is favorable to use them in place of other substitutes due to the fact that ionic liquids are "easy to handle" and their "...properties can be adjusted to suit the requirements of a particular process" (Bhaskar 2012) [11]. ILs are acclaimed as "green solvents" and "designer solvents" across many works of scientific literature. Due to all the reasons listed above, we chose them as a standard material for this project [9].

The Abraham Solvation Parameter Model:

We recognized that although several other contemporary models used to describe solute partitioning exist, including the quantitative structure-property relationships (QSAR) and quantum mechanical (QM) computer models, they fail to incorporate important intermolecular properties, and thus, lack essential information that facilitates solute-solvent interactions. [8]. Consequently, these models narrow their focus to only one chemical property, such as hydrogen bonding. Thus, we turned to the Abraham solvation parameter model, which analyzes division of solutes from a liquid to condensed phase (log P) or a gaseous to condensed phase (log K). The Abraham model excels in flexibility because its usage of several descriptors (E, S, A, B, V) in every derived correlation allows for the identification of more chemical properties. This aids in the discovery of relationships between solutes in different phases, ranging from organic and molecular solvents, ionic liquids, biological systems, and ecological barriers [8,12].

Objective: We designed this project to develop predictive models that would accurately identify substitutes for different biological barriers using computational methods and ILs. It is our fundamental intent to generate a profitable and reliable way for scientists to ascertain, in early stages, which drugs will meet their desired needs with respect to the penetration of human partitioning barriers. Yet, our mechanism can also be implemented in many other subjects, as proven towards the end of this study.

Procedure

Abraham Model Methodology:

We applied the Abraham model for the subsequent procedures due to its versatility in fields of study including biology, environmental studies, and chemistry. The model uses two linear free-energy equations to describe solute partitioning between two partly miscible or immiscible phases.

Log P describes the transfer of solute between two miscible phases:

$$\log P = c + eE + sS + aA + bB + vV \quad (1)$$

Log K describes the transfer of solutes from two immiscible phases:

$$\log K = c + eE + sS + aA + bB + lL \quad (2)$$

Solute Descriptors		Solute Coefficients	
		c	Linear regression constant
E	Excess Molar Refraction in units of $cm^3/(10\ mol^{-1})$	**e**	Interaction of the process phase with solute through dispersion interactions
S	Dipolarity and Polarizability	**s**	Measure of phase's dipolarity and polarizability
A	Hydrogen bond acidity	**a**	Measure of solvent's acid properties
B	Hydrogen bond basicity	**b**	Measure of work needed to create solvent cavity dispersion forces of a gas
L	Logarithm of the solute gas phase dimensionless Ostwald partition coefficient into hexadecane at 298 K	**l**	Measure of work needed to create solvent cavity and dispersion forces of a gas.
V	McGowan volume in units of $cm^3/(100\ mol^{-1})$	**v**	Measure of work needed to create solvent cavity and dispersion forces

Both of the linear equations above consist of three characteristics: measured solute properties, calculated solute descriptors, and calculated equation coefficients. Knowing any two of the listed values provides for the calculation of the third.

We disregarded solute descriptor **L,** the Ostwald partition coefficient into hexadecane at 298 K used in log K (Equation 2). The barriers of interest in this study only embody water-to-solvent distribution, so we utilized log P, which is described by Equation 3:

$$\mathbf{P = [solute]_{condensed\ phase} / [solute]_{water}} \quad (3)$$

$$\text{Solute}_{(water)} \rightleftharpoons \text{Solute}_{(condensed\ phase)}$$

To recognize a trend that would allow us to create a model for solute distribution in biological systems, we analyzed and compared an exhaustive compilation of acquired equation coefficients, which included excess molar refraction, dipolarity/polarizability, hydrogen-bond acidity, hydrogen-basicity, and McGowan's volume. This resulted in a linear regression equation with the slope-intercept form:

$$\log P_{\text{biological system}} = m \log P_{\text{ionic liquid}} + b \qquad (4)$$

In the regression above, the partitioning of solutes across any boundary is dependent upon the solubilizing behavior of various tested solutes in the designated ionic liquid [8].

Research on Biological Systems and Barriers:

We accumulated biological barrier and IL Abraham model partitioning coefficients, calculated by multi-linear regression analysis, to incorporate into the Abraham model [13]. We also noted coefficients for barriers in aquatic organisms, although our initial focus was more heavily on human biological systems. This step was taken to broaden the scope of our study and potentially establish linear models for aquatic toxicities as well. Table 1 shows the linear regression constants (c) and Abraham coefficients (e, s, a, b, v) obtained:

Table 1. Abraham Model Log P Correlation Coefficients for Biological Systems[23,24]

Name	c	e	s	a	b	v
Fathead minnow(96)	0.996	0.418	-0.182	0.417	-3.574	3.377
Guppy(96)	0.811	0.782	-0.23	0.341	-3.05	3.25
Bluegill(96)	0.903	0.583	-0.127	1.238	-3.918	3.306
Goldfish(96)	0.922	-0.653	1.872	-0.329	-4.516	3.078
Golden orfe(96)	-0.137	0.931	0.379	0.951	-2.392	3.244
Medaka high-eyes(96)	-0.176	1.046	0.272	0.931	-2.178	3.155
Medaka high-eyes(48)	0.834	1.047	-0.38	0.806	-2.182	2.667
Daphnia Magna(24)	0.915	0.345	0.171	0.42	-3.395	3.521
Daphnia Magna(48)	0.841	0.528	-0.025	0.219	-3.703	3.591
Ceriodaphnia dubia	2.234	0.37	-0.04	-0.44	-3.28	2.76
Daphnia pulex	0.502	0.39	0.3	0.66	-3.59	3.57
Tetrahymena pyriformis	0.616	0.413	-0.048	0.348	-2.707	2.944

Rana temporaria	0.609	0.866	-0.347	-0.174	-2.808	3.054
Blood-brain *(in vitro)*	-0.057	0.017	-0.536	-0.323	-0.335	0.731
Brain/(blood,plasma)	-0.028	0.003	0.485	-0.117	0.408	0.703
Blood-brain (*in vivo*)	0.547	0.221	-0.604	-0.641	-0.681	0.635
Blood-muscle (*in vitro*)	-0.185	-0.209	-0.593	-0.081	-0.168	0.741
Blood-muscle (*in vivo*)	0.082	-0.059	0.01	-0.248	0.028	0.11
Blood-fat (*in vitro*)	0.474	0.016	-0.005	-1.577	-2.246	1.56
Blood-fat (*in vivo*)	0.077	0.249	-0.215	-0.902	-1.523	1.234
Blood-liver (*in vitro*)	-0.095	0	-0.366	-0.357	-0.18	0.73
Blood-liver (*in vivo*)	0.292	0	-0.296	-0.334	0.181	0.337
Blood-lung (*in vitro*)	-0.143	0	0	0	-0.383	0.308
Skin/aqueous solution	-5.426	-0.106	-0.473	-0.473	-3	2.296
Blood/Skin	-0.105	-0.117	0.034	0	-0.681	0.756
Water/Skin	0.523	0.101	-0.076	-0.022	-1.938	1.652
Skin permeation from water	-5.42	-0.102	-0.457	-0.324	-2.68	2.066
Blood-heart	0.132	-0.039	-0.394	-0.376	0.009	0.527
Blood-kidney	0.494	-0.067	-0.426	-0.367	0.232	0.41

Correlation Coefficients for Ionic Liquids:

To build a comprehensive database of ionic liquids, we compiled Abraham equation coefficients from several publications, including our previously published work. This data included correlations that had been found for individual cations/anions and specific ionic liquids that had been tested independently. For the cations/anions that were found separately, we used the following equation to produce every conceivable combination of cation/anion:

$$\text{Log P} = c_{anion} + c_{cation} + (e_{anion} + e_{cation})\mathbf{*E} + (s_{anion} + s_{cation})\mathbf{*S} + (a_{anion} + a_{cation})\mathbf{*A} + (b_{anion} + b_{cation})\mathbf{*B} + (v_{anion} + v_{cation})\mathbf{*V} \qquad [12]$$

The combinations composed a grand total of 673 log P correlation coefficient equations for ionic liquids that would later be integrated into a principal component analysis with biological coefficients. A full

compilation of the log P ionic liquid correlation coefficients is present in our database for this experiment, found at the successive link [14]

Principal Component Analysis:

After having collated data from a multitude of sources, our team used Principal Component Analysis, a mathematical algorithm, to compare equation coefficients and transform the five Abraham correlation coefficients and constants (c, e, s, a, b, v) into two linearly distinct variables, or two principal components [15,16]. We utilized the aforementioned steps to simplify the visualization of the relationships between the two categories of solvents.

We employed the statistical applications on IBM's Statistical Package for the Social Sciences (SPSS v20.0.0) to execute the PCA. After compiling all coefficient data into one Microsoft Excel spreadsheet, we imported and surveyed the data using the "Factor Analysis" function of SPSS. All Abraham solute coefficients were used as variables for testing and were set to "Univariate Variable" to carry out the description of a single variable in terms of both classes of solvents. Additionally, we needed Bartlett's test of sphericity and Kaiser-Mayer-Olkin index as indicators in order to recognize redundancy between the variance of different principal components [17]. We set "Factors to Extract," which represents the number of linearly uncorrelated variables, to 2 to accommodate the two-dimensionality that was required for our analysis. The first run consisted of sole comparison of the 673 ILs and the biological barriers present in the human body, as this was the initial focus of the experiment. The second run was identical, except additional Abraham correlation coefficients for several aquatic organisms were included as part of an additional investigation for further study.

Regression Analysis:

After our PCA, we further examined the ILs denoted as probable mimics on a linear graph alongside the biological solvents that were deemed to be their closest matches. To verify if a strong linear relationship existed between the two solvents, we assessed the coefficient of determination for the line of best fit by measuring its proximity to 1. Again, we searched previous literature to find log P correlation coefficients for organic solvents dissolved in both the biological system and the IL being assessed. Our group performed this procedure for each biological barrier and IL.We then graphed corresponding scatter plots, using the IL log P values and biological barrier log P values for organic solutes as x and y values, respectively. These scatter plots enabled us to quantify the line of best fit and coefficient of determination for each plot in the Microsoft Excel program. The ending result included 6 different equations with the form of Equation 2.

Utilizing SPSS Statistical Package to evaluate data:

To further ensure scientific validity, we performed a matrix-guided statistical analysis on the data from the results by dividing the data points into a test set and training set. We divided the data points as follows: (1) The IL mimic became the test set. (2) The training set consisted of a minimum of 30 data points including Abraham model organic solute descriptors and log P values for each respective biological system. We could then use the associated built in statistical package in SPSS to perform this multi-linear least squares analysis that produced standard deviation (SD), average absolute error (AAE), and average error (AE) values to be subsequently evaluated for accuracy and precision.

Results and Discussion:

PCA: Using SPSS software to make a careful numerical analysis of obtained data with a PCA between the human biological systems and the ionic liquids, our team produced the following scatter plot:

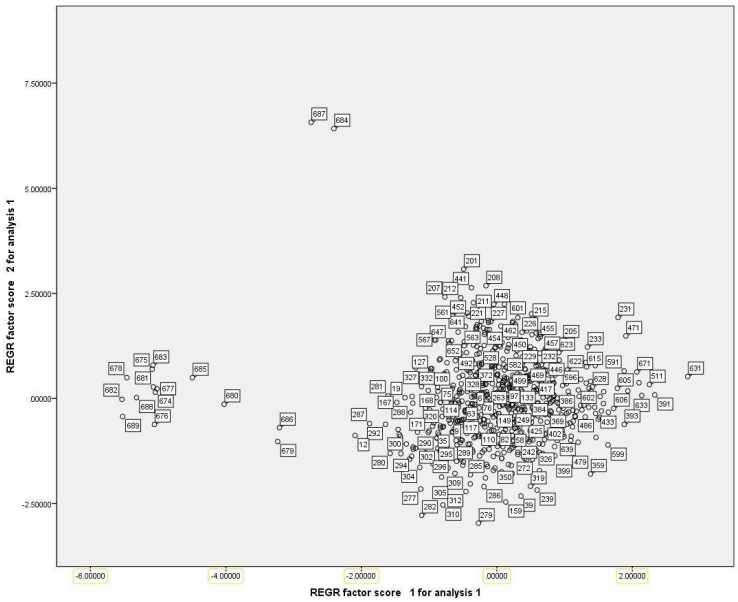

Figure 1. PCA of Human Biological Barriers and Ionic Liquids[13,14,18]

After expanding our focus to include more systems, such as *Daphnia Pulex*, we generated a PCA plot between all the biological systems/barriers and ionic liquids:

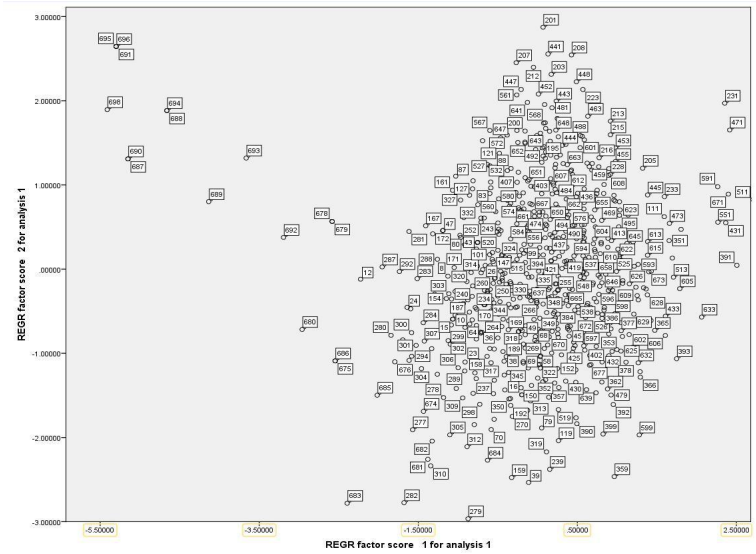

Figure 2. PCA of Ionic Liquids vs. All Biological Systems[13,14,18]

Our team utilized Figure 2 to optically decide which data points representing IL statistics were closest to those denoting biological solvents. The closer a pair of points, the more chemical properties shared between the two. The similarities established are portrayed in Table 5:

Table 5. Matched Data Points from Figure 2

Ionic	Ionic Name	Biological	Biological Name
70, 299	$[O4Am]^+$,b$[Tf2N]^-$, $[MeoeMPyrr]^+[FAP]^-$	684	Daphnia Pulex
310	$[O4Am]^+$,b$[FAP]^-$	681	Daphnia Magna(24)
276	$[MHIm]^+[FAP]^-$	676	Bluegill(96)
310, 277	$[O4Am]^+$,b$[FAP]^-$, $[MOIm]^+[FAP]^-$	682	Daphnia Magna(48)
73	$[OiQu]^+[Tf2N]^-$	677	Goldfish(96)
281	$[PM2Im]^+[FAP]^-$	678	Golden orfe(96)
12	$[EtOHMIm]^+[FAP]^-$	692	Blood-fat (in vitro)

Closer analysis of the graph from Figure 1 demonstrated that ionic liquid #12 ($[EtOHMIm]^+[FAP]^-$) was closest to the cluster of biological systems and would therefore most likely have a linear relationship with the following: 677 (blood-fat), 681 (blood-liver), and 683 (blood-brain).

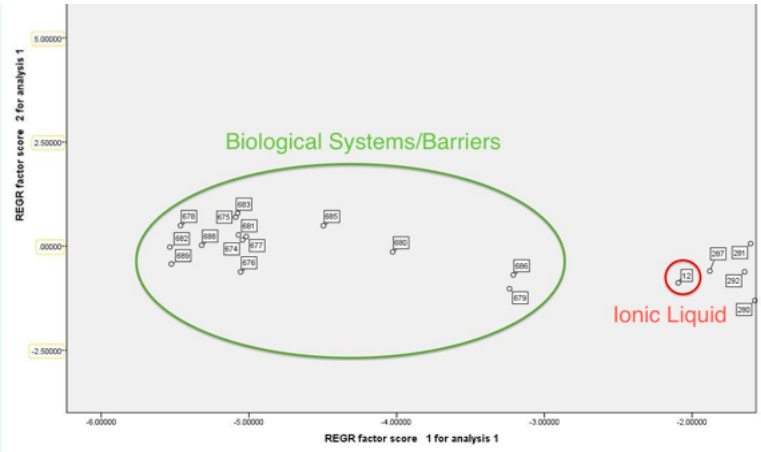

Figure 3. Closer Look at Figure 1 PCA graph

Linear Relationships:

The following are Abraham model log P correlation equations for each barrier, and linear relationships derived from plotting log P values of $[EtOHMIm]^+[FAP]^-$ against log P values of fat-blood partitioning, blood-liver partitioning, and blood-brain partitioning, respectively.

Blood-fat (in vivo), (Abraham *et al.,* 2012):

$$\text{Log } P_{\text{blood-fat}} = 0.077 + 0.249\mathbf{E} - 0.215\mathbf{S} - 0.902\mathbf{A} - 1.523\mathbf{B} + 1.234\mathbf{V}$$

Blood-liver (in vivo), (Abraham *et al.*, 2013):

$$\text{Log } P_{\text{blood-liver}} = 0.292 + 0.00\mathbf{E} - 0.296\mathbf{S} - 0.334\mathbf{A} + 0.181\mathbf{B} + 0.337\mathbf{V}$$

Blood-brain (in vivo), (Abraham *et al.*, 2012):

$$\text{Log P}_{\text{blood-brain}} = 0.547 + 0.221\mathbf{E} - 0.604\mathbf{S} - 0.641\mathbf{A} - 0.681\mathbf{B} + 0.635\mathbf{V}$$

Table 6. Log P Values for blood-fat and Ionic Liquid [14,19]

Organic Solutes	Log P $[EtOHMIm]^+[FAP]^-$	Log $P_{\text{blood-fat}}$ (*in vivo*)
1-Butanol	0.247	-0.13
1-Propanol	-0.18	-0.52
2-Pentanone	1.434	0.43
2-Propanol	-0.258	-0.67
Benzene	2.199	1.59
Ethanol	-0.573	-0.93
Ethyl acetate	1.363	0.28
Ethylbenzene	2.807	1.75
Heptane	3.246	2.01
Hexane	2.802	1.91
Methanol	-0.887	-1.08
Methyl acetate	0.997	0.02
Pentane	2.358	1.89
Toluene	2.467	1.78

NOTE: These are the x and y coordinates corresponding to the graphed scatter plots

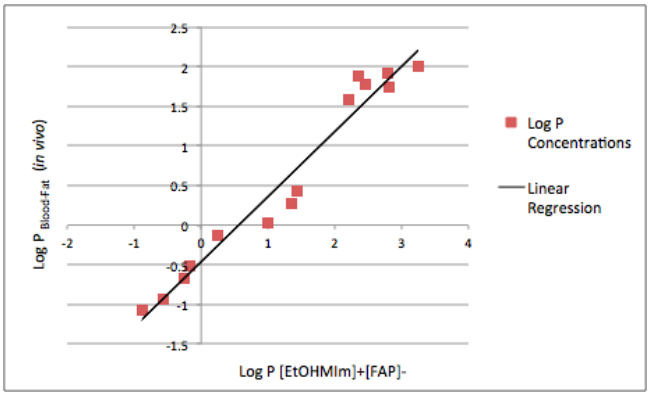

Figure 4. Log P Linear Relationship Between Blood-fat and Ionic Liquid

$y = 0.822x - 0.4632$
$R^2 = 0.96101$

Table 7. Log P Values for blood-liver and Ionic Liquid[14,20]

Organic Solutes	Log P $[EtOHMIm]^+[FAP]^-$	Log $P_{\text{blood-liver}}$ (*in vivo*)
Nonane	4.055	0.5
1-Propanol	-0.18	0.05

2-Propanol	-0.272	-0.03
1-Butanol	0.247	0.02
Acetone	0.826	0.02
Butanone	1.074	0.12
2-Pentanone	1.513	0.13
Ethyl acetate	1.363	0.13
Benzene	2.199	0.21
Ethylbenzene	2.807	0.31

NOTE: These are the x and y coordinates corresponding to the graphed scatter plots

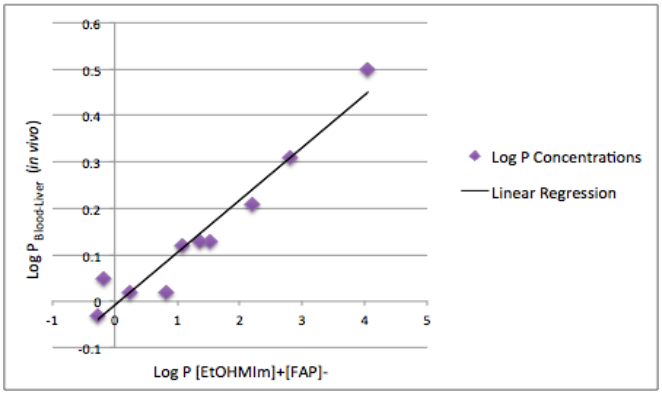

Figure 5. Log P Linear Relationship Between Blood-liver and Ionic Liquid

$y = 0.1132x - 0.0083$
$R^2 = 0.93341$

Table 8. Log P Values for blood-brain and Ionic Liquid [14,21]

Organic Solutes	Log P $[EtOHMIm]^+[FAP]^-$	Log $P_{blood\text{-}brain}$ (*in vivo*)
1-Decene	0.96	4.14
Acetone	-0.17	0.826
Benzene	0.2	2.199
Diethyl ether	-0.01	1.168
Ethyl acetate	0	1.363
Heptane	0.44	3.246
Methyl acetate	-0.13	0.997
Octane	0.69	3.71

Toluene	0.27	2.467

NOTE: These are the x and y coordinates corresponding to the graphed scatter plots

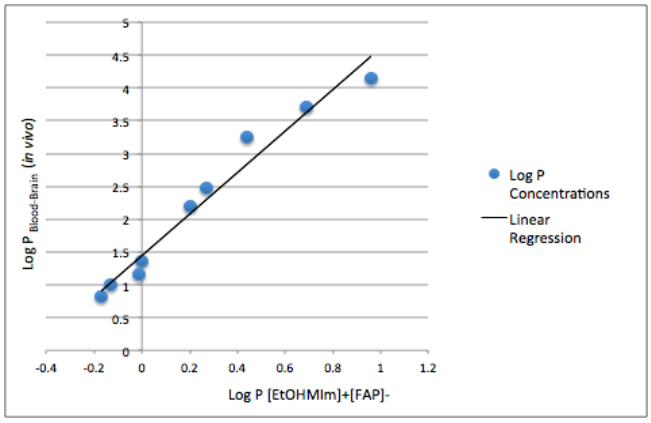

Figure 6. Log P Linear Relationship Between Blood-brain and Ionic Liquid

$y = 3.1641x + 1.4441$
$R^2 = 0.96676$

The three key barriers above pertain directly to current drug testing for pharmaceuticals targeting an assortment of organ systems. When attempting to predict the behavior of a drug's partitioning, for example, the blood-liver barrier, pharmaceutical companies will be capable of evaluating solubility of the compound in the ionic liquid system in order to acquire Abraham model solute descriptors (E,S,A,B,V). These descriptors would be utilized to formulate an Abraham model equation for that barrier and construct a log P value for the medication being studied in the correlated ionic liquid. Finally, an accurate log P linear relationship for the targeted biological system could be graphed with our derived equation:

Log BL (Blood-Liver) = mLog $P_{Compound}$+b , *where m and b are the values found for the linear relationship between blood-liver and ionic liquid, 0.1132 and -0.0083, respectively.*

This methodology is not restricted to the partitioning sites studied here and could be utilized for any plausible ionic liquid substitute of any barrier.

Further Linear Relationships:

Although the other organisms we tested did not portray the highest validity in terms of linear correlation and r-squared values, we still plotted their graphs:

Figure 7. Log LC and Log P Linear Relationship between *Daphnia pulex* and $[MeoeMPyrr]^+[FAP]^-$ [22]

y = 0.6336x + 1.7257
R^2 = 0.88934

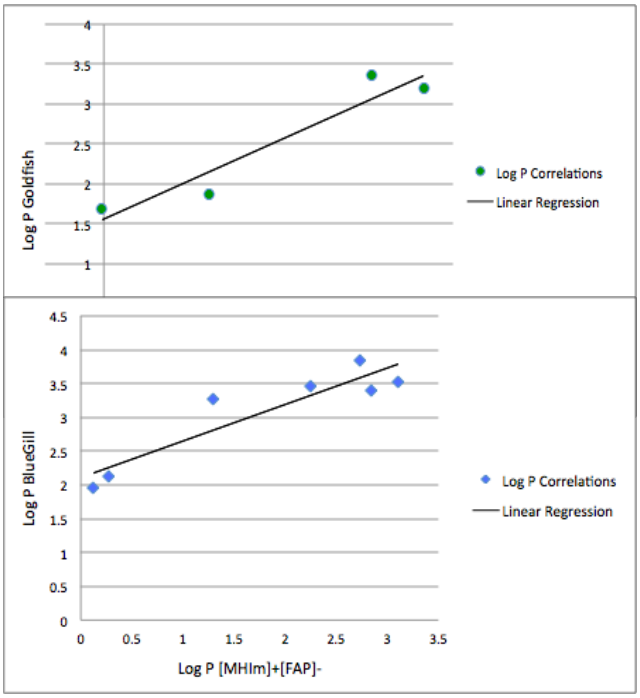

Figure 8. Log LC and Log P Linear Relationship between *Carassius auratus* and $[OiQu]^+[Tf_2N]^-$ [23]

y = 0.6551x + 1.5561
R^2 = 0.90748

Figure 9. Log LC and Log P Linear Relationship between *Lepomis macrochirus* and $[MHIm]^+[FAP]^-$ [24]
y = 0.5407x + 2.1093
R^2 = 0.84171

NOTE: Due to the enormity of this comprehensive study, it was not possible to include all data found and analyzed for the current project in this paper. Consequently, we procured a database congregating ionic liquid, organic solvent (w/d), and biological system Abraham correlation coefficients; log LC and log P values for the aquatic organisms analyzed above; additional PCA runs; and respective connections between biological systems/organic solvents and ionic liquids/organic solvents. Also, all tables/graphs found in this experiment are referenced. The data is found here: [18, 19, 20, 21, 22, 23, 24, 25]. The references are found here: [18].

SPSS Multi-linear Least Squares Analysis:

Analytical computation using Test and Training data sets yielded the following results:

Table 9. Test and Training Data Sets [25]

Biological System	Blood/Fat	Blood/Liver	Blood/Brain	*Lepomis macrochirus*	*Carassius auratus*

N (# of data points used)	49	50	34	59	53
SD (standard deviation)	.69671	0.83119	1.19536	.99603	0.62550
AAE (avg. absolute error)	.7847	1.6850	1.8329	1.5129	1.1835
AE (avg. error)	-.6066	-1.6514	-1.7806	1.5088	1.1835

Due to the fact that the living processes of an organism are dynamic and unpredictable, it is difficult to find a viable ionic liquid that mimics these biological characteristics. Our computational analysis indicated that the proposed mimics for each system contained statistical bias since AAE values were greater than SD values. Yet, note that *Carassius auratus* (goldfish) most strictly corresponded to its respective IL, determined from its lowest SD of 0.6255 log units. In sum, our results produced higher standard deviations and errors than what is normally accepted in similar studies regarding non-biological compounds. It is therefore necessary that we compile a larger data set of ionic liquid descriptors than found in this study, so that we may increase our chances of identifying a replacement IL with a low chance of statistical bias.

Summary of Results:

Our results prove that the IL $[EtOHMIm]^+[FAP]^-$ may conceivably be able to act as a substitute in current testing methods for the barriers involving blood-brain, blood-fat, and blood-liver. Supplementary studies performed concurrently with organisms point to ionic liquids [MeoeMPyrr]$^+$[FAP]-, $[OiQu]^+[Tf_2N]^-$, and $[MHIm]^+[FAP]^-$ as probable replacements for testing the lethal concentrations of compounds in species including *Daphnia pulex, Carassius auratus, and Lepomis macrochirus*, respectively.

The coefficients of determination found in our regression analyses are much higher than those reported in previous literature. When predicting permeability across the blood-brain barrier, we used

the QSAR computational method and ISIDA program to produce r-squared values of 0.781 and 0.872 [26]. Our r-squared values of .96101 for blood-fat, .93341 for blood-liver, and .96676 for blood-brain indicate a much more accurate finding than what has been accepted in previously conducted scientific research. However, SD and AAE values did indicate some statistical bias, owing to the enormous amount of compounds that have not yet been analyzed for log P in their respective biological barriers.

The compounds that were analyzed compose an integrative data network that will be of great use for researchers inevitably seeking to find parallels between ILs and biological systems. This accumulated database is continuously added to as more Abraham model coefficients are measured for distinctive ILs. In our previously co-authored publications, absorbency data from a Hewlett-Packard Diode Array spectrophotometer was used along with SPSS software to produce ion-specific equation coefficients for ionic liquid solvents such as the 4,5-dicyano-2-(trifluoromethyl)imidazolide anion ([BMIm]+[TDI]- and 3-MBPy]) [12]. These equation coefficients could later be put into a PCA to determine the corresponding biological system for that IL.

Conclusion and Future Work:

Pharmaceutical Drugs and Disease Treatment:

Drug companies are searching to develop medicines with low failure rates, and foreseeing solubility partitioning across biological barriers found in the human body is a key element of pharmacokinetics [8]. Knowledge regarding medication permeability across the three barriers for which we successfully identified IL counterparts is a prominent factor in several diseases, including Alzheimer's disease, multiple sclerosis, and brain tumors [27,28]. In diseases relating exclusively to the central nervous system, calculation of drug movement through the brain could be facilitated by our computed linear regression analysis for the blood-brain barrier [8]. Also, the blood-liver linear regression analysis

calculated by our *in silico* methods could enable better medications for diseases such as cirrhosis of the liver caused by alcoholism [29]. Lastly, our IL for the blood-fat barrier is promising for testing how certain fats pass through the barrier into the blood, potentially reducing the growing problem of obesity.

Producing Aquatic Toxicity Correlations:

By expanding our scope of investigation, we were able to produce unforeseen relationships between the log P partitioning of ionic liquids and the log LC (lethal concentration) of additional biological barriers. In-depth research into these barriers disclosed that our model could also be applied to aquatic toxicity and environmental rehabilitation. For example, *Carassius auratus*, also known as the common goldfish, is regarded as an invasive species in several parts of the U.S. [30]. Detecting a specific mimic for several different aquatic organisms potentially allows environmentalists to test toxic chemicals that are as specific as possible to goldfish. Invasive species could then be removed from waterways less expensively using our computed equations.

The Future of Ionic Liquids:

Ionic liquids solvents are quickly replacing traditional molecular solvents, as their unique properties make them easily adaptable to most chemical processes. However, the extent of log P relationships that currently exists on ionic liquids for the Abraham solvation model is narrowly limited to approximately 50 ILs, despite our PCA consisting of over 600 possible ILs. The PCA was able to find several IL counterparts to biological systems. Unfortunately, some of the ILs applied here have yet to be measured, because ILs are in a comparatively new domain. To procure more accurate and quantitative values for r-squared, SD, and AAE values in a broader study including aquatic organisms, it is paramount that we obtain more data on ILs and test a wider diversity of organic solutes dissolved in other ionic liquids and biological systems/barriers. On the other hand, our efficiently produced linear

equations and flexible methodology are the groundwork for future equations to be generated, all of which have the capacity to be extremely beneficial in an expansive array of fields, from environmental studies to pharmaceutical testing.

References

[1] Collins, F. People for the Ethical Treatment of Animals. http://www.peta.org/issues/animals-used-for-experimentation/animals-used-experimentation-factsheets/animal-experiments-overview (accessed July 1, 2014).

[2] Archibald, K. & Clotworthy, M. Comment on 'The ethics of animal research' by Festing & Wilkinson. *EMBO Reports* [Online] **2007**, *8*(9), 794-796.

[3] Daniel, G. W., Griffing, D., Socker, E., & Mayer, S. (2014, July 3). Right Drug, Right Patient: Streamlining Clinical Trials to Speed Drug Development. *The Brookings Institution*. Retrieved July 11, 2014, from http://www.brookings.edu/blogs/up-front/posts/2014/07/02-clinical-trials-pharmaceutical-drug-development

[4]Bernstein, M., & Cottingham, K. (2014, March 18). An end to animal testing for drug discovery?. *An end to animal testing for drug discovery?*. Retrieved July 1, 2014, from http://www.acs.org/content/acs/en/pressroom/newsreleases/2014/march/an-end-to-animal-testing-for-drug-discovery.html

[5] Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). Principles of early drug discovery. *British Journal of Pharmacology*,*162*(6), 1239-1249.

[6] Acree, W. E.; Grubbs, L. M.; Abraham, M. H. Prediction of Partition Coefficients and Permeability of Drug Molecules in Biological Systems with Abraham Model Solute Descriptors Derived from Measured Solubilities and Water-to-Organic Solvent Partition Coefficients. *UNT Digital Library* [Online] **2012**.

[7] Bernstein, M., & Cottingham, K. (2014, March 18). An end to animal testing for drug discovery?. *ACS: Chemistry For Life*. Retrieved July 11, 2014, from http://www.acs.org/content/acs/en/pressroom/newsreleases/2014/march/an-end-to-animal-testing-for-drug-discovery.html

[8]Stovall, D. M. Thermodynamics of the Abraham General Solvation Model: Solubility and Partition Aspects. Ph.D Dissertation, University of North Texas. Denton, Texas, 2006.

[9] Earle, M. J., & Seddon, K. R. (2000). Ionic liquids. Green solvents for the future. *Pure and Applied Chemistry* , *72*(7), 1391-1398.

[10] Park, S., & Kazlauskas, R. J. (2003). Biocatalysis in ionic liquids – advantages beyond green technology. *Current Opinion in Biotechnology*, *14*(4), 432-437.

[11]Bhaskar, K. (2012). First title: Ionic liquids-useful reaction green solvents for the future Second title: ionic liquids are the replacements for environmentally damaging solvents in a wide range of chemical processes.. *Journal Of Biomedical And Pharmaceutical Research, 1*(02). Retrieved from http://jbpr.in/index.php/jbpr/article/view/24

[12] Stephens, T.; Hart, Erin.; Kuprasertkul, N.; Mehta; S.; Wadawadigi, A.; Acree, W.; Abraham M. Abraham model correlations for describing solute transfer into ionic liquid solvents: calculation of ion-specific equation coefficients for the 4,5-dicyano-2-(trifluoromethyl)imidazolide anion. Physics and Chemistry of Liquids [Online] 2014, 56, DOI: 10.1080/00319104.2014.929949

[13] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Biological Systems Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc330328/. Accessed August 29, 2014.

[14] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Ionic Liquids Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307526/. Accessed August 6, 2014.

[15] Shlens, Jon; *A Tutorial On Principal Component Analysis Derivation, Discussion and Singular Value Decomposition;* Princeton University: Princeton, New Jersey, 2003

[16]Palmer, M. Principal Component Analysis. http://ordination.okstate.edu/PCA.htm. (Accessed July 2014)

[17]ERIC. Bartlett's sphericity test and the KMO index (Kaiser-Mayer-Olkin). http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/en_Tanagra_KMO_Bartlett.pdf. (Accessed June 2014)

[18] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *[Dataset Supplemental Material and References]*. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc330321/. Accessed August 29, 2014.

[19] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log P Blood Fat Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307527/. Accessed August 29, 2014.

[20] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log P Blood Liver Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307531/. Accessed August 29, 2014.

[21] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log P Blood Brain Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307529/. Accessed August 29, 2014.

[22] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log LC Daphnia pulex Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307528/. Accessed August 29, 2014.

[23] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log LC Goldfish Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307536/. Accessed August 29, 2014.

[24] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Log LC Bluegill Dataset*. UNT Digital Library.http://digital.library.unt.edu/ark:/67531/metadc307530/. Accessed August 29, 2014.

[25] Kupra, Nina; Mehta, Sumedha & Wadawadigi, Akash. *Test and Training Bio Systems Dataset*. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc307532/. Accessed September 4, 2014.

[26] Katritzky A. R.; Kuanar M.; Slavov S.; Dobochev D. A.; Fara D. C.; Kareleson M.; Acree W.; Solo'vev. V.P.; Varnek A.; Correlation of blood-brain penetration using structural descriptors. Bioorganic & Medicinal Chemistry[online] **2006**.DOI: 10.1016/j.bmc.2006.03.012

[27] Marques, F.; Sousa, J.C.; Sousa, N.; Palha, J.A.Blood Brain Barriers in aging and Alzheimer's disease. Molecular Neurodegeneration [Online] **2013**, *8*, 38. DOI: 10.1186/1750-1326-8-38.

[28]Rosenburg G.; Neurological diseases in relation to the blood-brain barrier. Journal of Cerebral Blood Flow & Metabolism[online] **2012**, 32, 7. doi:10.1038/jcbfm.2011.197

[29] *Canadian Liver Foundation*. http://www.liver.ca/liver-disease/types/cirrhosis/ (Accessed June 2014)

[30]Winter, T.J. Population size, growth, and control of exotic goldfish (*Crassius atratus*) in a small impoundment: implications for managing future invasions. *Undergraduate Research Opportunities Program* **[2005]**.