Estimating the Water Solubilities of Crystalline Compounds from Their Chemical Structures Alone

James W. McFarland*

reckon.dat consulting, 217 Blood Street, Lyme, Connecticut 06371-3509

Alex Avdeef and Cynthia M. Berger

pION INC., 5 Constitution Way, Woburn, Massachusetts 01801

Oleg A. Raevsky

Institute of Physiologically Active Compounds, Chernogolovka, Moscow Region, 142432, Russia

Received January 5, 2001

Partial atomic charges are significant descriptors in predicting the water solubilities of crystalline organic compounds from their chemical structures. Lipophilicity remains the predominant factor. It was also found that quantitative estimates of hydrogen bond strengths (hydrogen bond factors) play important roles. These descriptors can be easily interpreted to guide chemists to the synthesis of compounds with increased or decreased water solubility. This work is based on a set of 22 compounds the aqueous solubilities of which were determined by a new potentiometric method, pSOL, and were confirmed, in part, by the traditional shake-flask method. A new software package, HYBOTPLUS, furnished the partial atomic charges and hydrogen bond factors.

INTRODUCTION

In general, the more water-soluble drugs are, the greater is the likelihood they will be absorbed by the oral route of administration. Of late, discovery lead compounds not only tend toward high in vitro potency and high lipophilicity but also tend toward low aqueous solubility and poor oral absorption.[1] Hence, there is a need to estimate the water solubilities of compounds before their syntheses to eliminate those that have a low probability of succeeding in later drug development stages.

In 1980, Yalkowsky and Valvani[2] successfully correlated the aqueous solubilities of nonelectrolytes with octanol−water partition coefficients and melting points. From theoretical considerations, they proposed the following relationship

$$\log S_w = - a \times \log P - b \times \mathrm{mp} + c \qquad (1)$$

where log S_w is the logarithm of the compound's molar solubility in water; where log P is the logarithm of its octanol−water partition coefficient; and where mp is the compound's melting point. Log P approximates the activity coefficient (γ) of the un-ionized solute in equilibrium with the un-ionized molecular species in the crystal. The mp is a proxy for the relative energy it takes to break the solute's crystal lattice; the higher the mp, the lower the solubility. While this has advanced our understanding of the forces involved in the solution process, the equation is not suited for accurately predicting the solubilities of compounds prior to their syntheses. Log P is fairly well estimated from the MEDCHEM or similar programs, but there are as yet no reliable means to predict melting point; the experimental value must be used.

More recently, methods have been proposed that can be used to predict the aqueous solubilities of compounds from their chemical structures alone.[3−15] We set forth here such an approach that includes partial atomic charge and quantitative estimates of hydrogen-bonding strengths as determined by the computer program HYBOTPLUS.[16] To verify this method, we used a set of 22 ionizable drugs or drug-like compounds whose solubilities were determined by a single laboratory using a new potentiometric method, pSOL.[17] These compounds consisted of drugs from several different therapeutic categories. Some were acids, some were bases, and one was a zwitterion, but none was a neutral compound.

METHODS

Solubility measurements for the compounds in Table 1 were performed by the pSOL Model 3 instrument.[18] Avdeef, Berger, and Brownell[17] give the details of the experimental method.

Hydrogen bond factors were determined by HYBOTPLUS software.[16] Raevsky et al.[19] and Raevsky[20] describe the derivation of hydrogen bond factors in greater detail. In essence, these factors are mathematical constructs that allow one to quantitatively estimate relative hydrogen bonding acceptor and donor strengths for a large variety of chemical atoms and groups. The electronic and steric environments of the groups are taken into consideration. The calculations are rapid and can be done on an IBM-compatible personal computer. To compute the hydrogen bond factors, this software must be run simultaneously with HYBOT,[21] a

* Corresponding author phone: (860)434-2271; fax: (860)434-5125; e-mail: reckon.dat@attglobal.net.

10.1021/ci0102822 CCC: $20.00
Published on Web 06/08/2001

Table 1. Hydrogen Bond Factors, Other Physical Properties, and the Aqueous Solubilities of 24 Drug-like Compounds[a]

compound	α	$\Sigma(\lvert q^-_{min}\rvert + q^+_{max})$	ΣC_d	ΣC_{ad}	$\Sigma C_{ad}/\alpha$	log *P*	CLOGP	$\log(1/S_o)_{obs} \pm$ SD	$\log(1/S_o)$ predicted[b]	residual
atenolol	29.007	0.472	−7.844	15.299	0.5274	0.22	−0.109	1.30 ± 0.12	1.72	−0.42
diclofenac	29.525	0.531	−3.746	7.487	0.2536	4.51	4.711	5.59 ± 0.08	5.59	0.00
famotidine	32.736	0.669	−11.966	19.970	0.6100	−0.56	−0.556	2.48 ± 0.08	2.73	−0.25
flurbiprofen	25.936	0.537	−2.820	5.708	0.2201	3.99	3.754	4.36 ± 0.10	4.37	−0.01
furosemide	29.708	0.601	−8.172	13.967	0.4701	2.56	1.869	4.75 ± 0.10	4.18	0.57
hydrochlorothiazide	25.530	0.633	−7.299	14.610	0.5723	−0.07	−0.399	2.63 ± 0.05	2.48	0.15
ibuprofen	23.707	0.536	−2.820	5.037	0.2125	4.13	3.679	3.62 ± 0.02	4.23	−0.61
ketoprofen	28.016	0.537	−2.820	6.996	0.2497	3.16	2.761	3.33 ± 0.05	3.46	−0.13
labetolol	36.832	0.502	−9.950	17.272	0.4689	1.33	2.500	3.45 ± 0.11	4.36	−0.91
naproxen	25.746	0.537	−2.820	6.146	0.2387	2.81	2.816	3.89 ± 0.01	3.45	0.44
phenytoin	28.087	0.520	−3.692	9.106	0.3242	2.24	2.085	4.13 ± 0.11	3.10	1.03
propranolol	30.976	0.411	−3.684	8.800	0.2841	3.48	2.193	3.62 ± 0.11	2.43	1.19
benzoic acid	12.697	0.474	−2.890	4.265	0.3359	1.96	1.885	1.59 ± 0.02	2.72	−1.13
cimetidine	27.416	0.453	−5.834	11.152	0.4068	0.48	0.351	1.43 ± 0.08	1.38	0.05
diltiazem	44.054	0.617	0.000	8.593	0.1950	2.89	2.947	2.95 ± 0.08	3.73	−0.78
enalapril	39.188	0.623	−5.351	13.559	0.3460	0.16	0.667	1.36 ± 0.08	2.20	−0.84
metoprolol	29.848	0.445	−3.684	9.701	0.3250	1.96	1.346	1.20 ± 0.08	1.92	−0.72
nadolol	33.381	0.414	−7.048	14.827	0.4442	0.85	0.329	1.57 ± 0.08	1.39	0.18
propoxyphene	40.223	0.666	0.000	4.559	0.1133	4.37	5.210	5.01 ± 0.12	5.94	−0.93
quinine	37.702	0.435	−1.682	8.434	0.2237	3.50	2.485	2.82 ± 0.02	2.49	0.33
terfenadine	57.295	0.446	−3.549	9.822	0.1714	5.72	6.093	6.69 ± 0.17	6.17	0.52
trovafloxacin	38.000	1.193	−7.500	17.427	0.4586	0.13	−0.822	4.53 ± 0.06	4.12	0.41

[a] Data from pION Inc. and HYBOTPLUS. [b] From eq 4.

Table 2. Descriptors Used

descriptor	definition
α	molecular polarizability
q^-_{min}	minimal value among negative partial atomic charges
q^+_{max}	maximal value among positive partial atomic charges
Σq^-	sum of negative partial atomic charges
Σq^+	sum of positive partial atomic charges
$\lvert q^-_{min}\rvert + q^+_{max}$	sum of the absolute values of q^-_{min} and q^+_{max}
C_{a_max}	maximal value of hydrogen bond acceptor factors (value always positive)
C_{d_min}	minimal value of H-bond donor factors (value always negative)
ΣC_a	sum of H-bond acceptor factors
ΣC_d	sum of H-bond donor factors
ΣC_{ad}	sum of the absolute values of H-bond acceptor and donor factors
$\Sigma q^-/\alpha$	sum of negative partial atomic charges/molecular polarizability
$\Sigma q^+/\alpha$	sum of positive partial atomic charges/molecular polarizability
$\Sigma C_a/\alpha$	sum of H-bond acceptor factors/molecular polarizability
$\Sigma C_d/\alpha$	sum of H-bond donor factors/molecular polarizability
$\Sigma C_{ad}/\alpha$	absolute sum of H-bond acceptor and donor factors/molecular polarizability
MW	molecular weight
CLOGP	calculated log *P* from the MEDCHEM program
CMR	calculated molecular refractivity from the MEDCHEM program

database-program also developed by Raevsky's group. Further, HYBOTPLUS calculates the molecular polarizabilities and partial atomic charges. Table 2 lists and defines these properties. In calculating these values for a compound with both a strong acidic and a strong basic group, the zwitterionic structure was used. Specifically, in the present work this means trovafloxacin.

CLOGP and CMR were determined using MEDCHEM software.[22] The definitions of these are also listed in Table 2.

For data analyses we used Statistica software[23] to perform multiple regression analyses. To compute Q^2, we used MODDE software.[24]

RESULTS

Avdeef et al.[17] published the aqueous solubilities of 12 ionizable drugs as determined by their new potentiometric method. These are reported as the first dozen compounds in Table 1. Since then, the solubilities of an additional 10 ionizable compounds (including a zwitterion) have been determined by the potentiometric method. These are reported as the last 10 compounds in Table 1. Because these compounds are ionizable, their solubilities depend on the pH of the aqueous solution. To avoid confusion about pH dependence, we define aqueous solubility, S_o, as the molarity of the un-ionized molecular species. In this regard S_o is analogous to *P*, the octanol−water partition coefficient; hence, the ionized species are implicitly factored out. Here we use *P* in the strict sense of the partitioning of the *un-ionized* species, rather than the *apparent P* where ionized species are taken into consideration.[25] The pSOL procedure involves knowledge of the dissociation constants; therefore, one can then compute the pH-solubility profiles of the various molecular species in solution.

Because the work of Yalkowsky and Valvani[2] demonstrated that the partition coefficient is a dominant factor in the aqueous solubility of a compound, we first examined the relationship between solubility and the measured and

Table 3. Correlation Matrix for the Independent and Dependent Variables Used in Eqs 4−6

	$\Sigma(\lvert q^-_{min}\rvert + q^+_{max})$	ΣC_d	$\Sigma C_{ad}/\alpha$	log P	CLOGP	log($1/S_o$)
$\Sigma(\lvert q^-_{min}\rvert + q^+_{max})$	1.000	−0.219	0.225	−0.316	−0.312	0.236
ΣC_d		1.000	−0.922	0.718	0.661	0.232
$\Sigma C_{ad}/\alpha$			1.000	−0.868	−0.847	−0.451
log P				1.000	0.951	0.699
CLOGP					1.000	0.702
log($1/S_o$)						1.000

calculated values for log P. Equations 2 and 3 show that both values correlate with log($1/S_o$) to about the same degree; hence, it seems the use of measured values can be avoided. In this work, the expression (±*n.nn*) in equations represents the standard error of the coefficient that precedes it.

Measured values

$$\log(1/S_o) = 0.60(\pm 0.14)\log P + 1.92(\pm 0.39) \quad (2)$$

$n = 22$ $r = 0.70$ $r^2 = 0.49$ $s = 1.12$ $F_{1,20} = 19.150$ $p = 0.000292$

Calculated values

$$\log(1/S_o) = 0.57(\pm 0.13)\text{CLOGP} + 2.11(\pm 0.36) \quad (3)$$

$n = 22$ $r = 0.70$ $r^2 = 0.49$ $s = 1.11$ $F_{1,20} = 19.466$ $p = 0.000269$

The correlation coefficient between log P and CLOGP in the current data set is 0.95. While this is an expected result, it should be noted that on the basis of this correlation the calculated log P values for labetalol, enalapril, quinine, and propranol deviated respectively by −1.31, −0.85, 0.88, and 1.12 units from their measured values. These compounds and propoxyphene gave the least agreement between values of log($1/S_o$) predicted by eqs 2 and 3. Equation 3 estimated the measured solubilities of labetalol, quinine, propranolol, and propoxyphene more accurately than did eq 2. On the other hand, eq 2 estimated the solubility of enalapril more accurately than did eq 3.

While these equations are significant statistically, the r^2 values indicate that they are not good enough to make reliable predictions of a compound's solubility. We next explored whether additional calculable molecular properties could improve the correlations. HYBOTPLUS provides 10 such properties that can also be combined in various ways to expand the descriptor set. These are given and defined in Table 2. In addition, we also considered the molecular weight (MW) and the calculated molar refractivity (CMR).[21] The correlation matrix for the biological and physical properties in Table 1 is presented in Table 3. Tables 1 and 3 list only those descriptors that were found to be significant in the present work; the complete set of descriptors and the corresponding correlation matrix are available (see Supporting Information).

Forward-stepwise regression gave eq 4 as the best combination of descriptors where CLOGP was used as the estimator for lipophilicity.

$$\log(1/S_o) = 1.10(\pm 0.16)\text{CLOGP} + 4.95(\pm 0.99)(\lvert q^-_{min}\rvert + q^+_{max}) + 6.44(\pm 2.10)\Sigma C_{ad}/\alpha - 3.93(\pm 1.21) \quad (4)$$

$n = 22$ $R = 0.90$ $R^2 = 0.82$ $s = 0.70$ $Q^2 = 0.64$ $F_{3,18} = 26.946$ $p = 0.000001$

In this analysis, CLOGP was the first term entered and accounted for 49% of the variance in the dependent variable, log($1/S_o$). The next most important term, ($\lvert q^-_{min}\rvert + q^+_{max}$), accounted for 23% of the variance. Finally, $\Sigma C_{ad}/\alpha$ accounted for 10%. Hence, it appears that in addition to lipophilicity, partial atomic charge and hydrogen bonding play important roles in determining a compound's aqueous solubility. There were no statistical outliers among the 22 compounds. As a way of confirmation, backward-stepwise regression analysis gave the same result.[26]

Next we substituted log P for CLOGP. Forward-stepwise regression gave eq 5 as the best combination of the descriptors considered.

$$\log(1/S_o) = 1.10(\pm 0.11)\log P + 4.79(\pm 0.81)(\lvert q^-_{min}\rvert + q^+_{max}) - 0.28(\pm 0.06)\Sigma C_d - 3.21(\pm 0.71) \quad (5)$$

$n = 22$ $R = 0.94$ $R^2 = 0.88$ $s = 0.58$ $Q^2 = 0.83$ $F_{3,18} = 42.876$ $p < 10^{-6}$

In this analysis, log P was the first term entered and accounted for 49% of the variance in the dependent variable, log($1/S_o$). As with eq 4, the next most important term, ($\lvert q^-_{min}\rvert + q^+_{max}$), accounted for 23% of the variance. Finally, ΣC_d accounted for 16%. While eq 5 is a better correlation than eq 4, phenytoin was nevertheless an outlier. Log P corresponds in character to CLOGP in eq 4. The next most important term, ($\lvert q^-_{min}\rvert + q^+_{max}$), is the same as that in eq 4. The third most important descriptor in eq 5 is similar to the third most descriptor in eq 4 only in the sense that it reflects the involvement of H-bonding. It suggests that ΣC_d plays a more important role here than $\Sigma C_{ad}/\alpha$ does in eq 4.

However, backward-stepwise regression analysis resulted in eq 6.

$$\log(1/S_o) = 1.31(\pm 0.16)\log P + 5.13(\pm 0.88)(\lvert q^-_{min}\rvert + q^+_{max}) + 8.24(\pm 2.00)\Sigma C_{ad}/\alpha - 5.34(\pm 1.20) \quad (6)$$

$n = 22$ $R = 0.93$ $R^2 = 0.86$ $s = 0.62$ $Q^2 = 0.83$ $F_{3,18} = 35.835$ $p < 10^{-6}$

When these descriptors were used in a forward stepwise regression, log P was the first term entered and accounted for 49% of the variance in the dependent variable, log($1/S_o$). The next most important term, ($\lvert q^-_{min}\rvert + q^+_{max}$), accounted for 23% of the variance. Finally, $\Sigma C_{ad}/\alpha$ accounted for 14%. Equation 6 has descriptors identical to those of eq 4. Further, the coefficients are within the 95% confidence limits of each other, and there are no outliers. The principal reason for this result was that ΣC_d was one of the terms that had to be removed to avoid ill conditioned matrices. It was correlated with three other terms at the level of $r > 0.90$, one of which was $\Sigma C_{ad}/\alpha$. On balance, we believe that eqs 4 and 6 best summarize the information in the data set. While eq 6 is somewhat inferior to eq 5, it has the advantage of predicting the measured solubilities with no outliers. Hence, through both forward and backward stepwise regression we arrive at the essentially the same result.

Referees of the original submission of this paper suggested that we also compare these results, which used solely calculated descriptors, to those that could be obtained by using the original Yalkowsky-Valvani relationship. However, we had to abandon this idea when we learned from a search of the literature that the mp's for the free bases of labetalol and enalapril and of the zwitterion of trovafloxacin have not been reported. Further, several significantly different mp's

were reported for each of diclofenac, flurbiprofen and nadolol, thus adding to the uncertainty of how to proceed. Already too many compounds would have to be omitted; no meaningful comparisons could be made under these circumstances.

DISCUSSION

Our studies show, that in addition to lipophilicity, partial atomic charges and hydrogen bond factors may play significant roles in determining a compound's aqueous solubility. It is plausible that in the crystalline state, intermolecular electrostatic interactions between partial atomic charges and intermolecular hydrogen bonding could contribute to crystal lattice stability. This in turn will relate a compound's water solubility. Together these descriptors could be proxies for mp.

We observed the important roles that the minimal negative partial atomic charge and the maximal positive partial atomic charge, ($|q^-_{min}| + q^+_{max}$), played. Katritzky et al.[3] analyzed the aqueous solubilities of 411 compounds of diverse structure that included entities that were gases, liquids, and solids under normal conditions. Interestingly, they found that one of the most important descriptors was Q_{min}, the most negative partial atomic charge. A like conclusion was reached by Mitchell and Jurs.[4] While there is some ambiguity here regarding whether the positive or negative partial atomic charge should be used, it seems clear that one or both of these terms point to the importance of intermolecular electrostatic interactions.

Katritzky et al.[3] also noted that the fractional hydrogen donor surface area (FHDSA) was an important descriptor in correlating aqueous solubility with calculable physical properties. Abraham and Le[5] demonstrated the significance of quantitative estimates of hydrogen bond acidity and basicity in their amended solvation equation. Jorgensen and Duffy[6] also showed that hydrogen bonding is important. In the analysis of our data, $\sum C_{ad}/\alpha$ was a significant descriptor. In our view, this descriptor reflects that the density of hydrogen bonding sites is important as well as their number. These observations support the concept that hydrogen bonding is a key factor in aqueous solubility.

From the many QSPR studies on solubility, there is a sense emerging that there is a common set of calculable structural descriptors useful in predicting this important property. While each of the studies cited here defines these descriptors somewhat differently and often include descriptors not shared with the others, we believe that lipophilicity, partial atomic charges, and hydrogen bond factors, in one form or another, eventually will be the dominant ones.

This study shows that HYBOTPLUS descriptors in conjunction with CLOGP offers an approach to more accurate predictions of the water solubilities of crystalline organic compounds than does the use of CLOGP alone. This approach is rapid and computationally undemanding and has the advantage of giving results that are readily interpreted by chemists to provide direction in preparing compounds with altered aqueous solubilities.

ACKNOWLEDGMENT

We thank Ms. Melissa Strafford for valuable technical assistance in completing this work.

Supporting Information Available: The data considered in the analyses described in this article are available as a table. This includes the 15 summary variables from the HYBOT program as well as CLOGP, CMR (calculated molar refractivity), MW, and the experimental and calculated solubilities. In addition the correlation matrix for these data is also supplied. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.

(2) Yalkowsky, S. H.; Valvani, S. C. Solubilities and Partitioning I: Solubility of Nonelectrolytes in Water. *J. Pharm. Sci*. **1980**, *69*, 912−922.

(3) Katritzky, A. R.; Wang, Y, Sild, S.; Tamm, T. QSPR Studies on Vapor Pressure, Aqueous Solubility, and the Prediction of Water−Air Partition Coefficients. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 720−725.

(4) Mitchell, B. E.; Jurs, P. C. Prediction of Aqueous Solubiltiy of Organic Compounds from Molecular Structure. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 489−496.

(5) Abraham, M. H.; Le, J. The Correlation and Prediction of the Solubility of Compounds in Water Using an Amended Solvation Energy Relationship. *J. Pharm. Sci.* **1999**, *88*, 868−880.

(6) Jorgensen, W. L.; Duffy, E. M. Prediction of Drug Solubility from Monte Carlo Simulations. *Bioorg. Med. Chem. Lett*. **2000**, *10*, 1155−1158.

(7) Huuskonen, J.; Salo, M.; Taskinen, J. Aqueous Solubility Prediction of Drugs Based on Molecular Topology and Neural Network Modeling. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 450−456.

(8) A brief review of this subject: Katritzky, A. R.; Maran, U.; Lobanov, V. S.; Karelson, M. Structurally Diverse Quantitative Structure−Property Relationship Correlations of Technologically Relevant Physical Properties. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 1−18.

(9) Klopman, G.; Wang, S.; Balthasar, D. M. Estimation of Aqueous Solubility of Organic Molecules by the Group Contribution Approach. Application to the Study of Biodegradation. *J. Chem. Inf. Comput. Sci*. **1992**, *32*, 474−482.

(10) Pogliani, L. Modeling Purines and Pyrimidines with Linear Combination of Connectivity Indices-Molecular Connectivity "LCCI-MC" Method. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 1082−1091.

(11) Lee, P. H. Computational modeling of drug solubility. In *Book of Abstracts, American Chemical Society 220th National Meeting*; American Chemical Society: Washington, DC, August 20−24, 2000; COMP 40.

(12) Semus, S. F. Novel linear regression model for the prediction of aqueous solubility of drug molecules employing fundamental descriptors. In *Book of Abstracts, American Chemical Society 220th National Meeting*; American Chemical Society: Washington, DC, August 20−24, 2000; COMP 43.

(13) Cheng, A.; Merz, K. M. Prediciton of aqueous solubility of a diverse set of compounds using quantitative structure−property relationships. In *Book of Abstracts, American Chemical Society 220th National Meeting*; American Chemical Society: Washington, DC, August 20−24, 2000; COMP 137.

(14) McElroy, N. R.; Jurs, P. C. Aqueous solubility prediction of two classes of organic compounds from molecular structure. In *Book of Abstracts, American Chemical Society 220th National Meeting*; American Chemical Society: Washington, DC, August 20−24, 2000; COMP 150.

(15) Klamt, A.; Eckert, F. COSMO-RS: A Novel Way from Quantum Chemistry to Free Energy, Solubility, and General QSAR-Descriptors for Partitioning. In *Rational Approaches to Drug Design*; Holtje, H.-D., Sippl, W., Eds.; Proceedings of the 13th European Symposium on Quantitative Structure−Activity Relationships; Prous Science: Barcelona, 2001; pp 195−205.

(16) From the MolPro Project. HYBOTPLUS is no longer available; it has been replaced by a superior version, HYBOTPLUS-2000. This new version gives comparable results but also runs about 100 times faster and is based on a larger data set. Contact Professor O. A. Raevsky, Institute of Physiologically Active Compounds, Chernogolovka, Moscow Region, 142432, Russia; e-mail: raevsky@ipac.ac.ru or Dr. J. W. McFarland, reckon.dat consulting, 217 Blood Street, Lyme, CT 06371-3509; e-mail: reckon.dat@attglobal.net.

(17) Avdeef, A.; Berger, C. M.; Brownell, C. pH-Metric Solubility. 2: Correlation between the Acid−Base Titration and the Saturation Shake-Flack Solubility pH Methods. *Pharm. Res*. **2000**, *17*(1), 85−89.

(18) pSOL, pION INC., 5 Constitution Way, Woburn, MA 01801. E-mail: pion@pion-inc.com.
(19) Raevsky, O. A.; Grigor'ev, V. Y.; Kireev, D. B.; Zefirov, N. S. Complete Thermodynamic Description of H−Bonding in the Framework of Multiplicative Approach. *Quant. Struct.-Act. Relat*. **1992**, *11*, 49−63.
(20) Raevsky, O. A. Quantification of noncovalent interactions on the basis of the thermodynamic hydrogen bond parameters. *J. Phys. Org. Chem.* **1997**, *10*, 405−413.
(21) HYBOT and HYBOTPLUS, MolPro Project, Chernogolovka, Moscow Region, 142432 Russia. E-mail: raevsky@ipac.ac.ru; website: http://www.ibmh.msk.su/molpro/.
(22) MEDCHEM, version 3.54; Daylight Chemical Information Systems, Inc.: 18500 Von Karman Ave, Suite 450, Irvine, CA 92715.
(23) Statistica (Release 5, '97 Edition); Statsoft: 2300 East 14th Street, Tulsa, OK 74104.
(24) Modde 4.0, Graphical Software for Design of Experiments; Umetri AB, Box 7960, S-907 19 Umeå, Sweden, 1997.
(25) CLOGP refers to the logarithm of the partition coefficient of *un-ionized* species. The pSOL system of determining solubilities relies on this definition of log *P*. Yalkowsky and Valvani also intend this definition in their work. All too often, in the literature log *P* is used when in fact log *D* is what is meant. For a good review on this subject, see: Taylor, P. J. Hydrophobic Properties of Drugs. In *Comprehensive Medicinal Chemistry, Volume 4*. *Quantitative Drug Design*; Ramsden, C. A., Ed.; Pergamon Press: Oxford, 1990; pp 241−294.
(26) To avoid ill conditioned matrices, this required the removal of six descriptors that were highly correlated with at least one of those remaining: q^-_{max}, $\sum C_d$, $\sum q^+$, log *P* (when CLOGP was used) or CLOGP (when log *P* was used), CMR, and MW.

CI0102822