ELSEVIER

Theoretical hydrogen bonding parameters for drug design

Emanuela Gancia, John G. Montana, and David T. Manallack

Celltech Chiroscience Ltd., Granta Park, Great Abington, Cambridge CB1 6GS, UK

Hydrogen bonding interactions play a major role in many chemical and biological processes. This article describes the development of a method for the quantitative estimation of the hydrogen-bonding donor strengths of OH/NH moieties and of the hydrogen bonding acceptor strengths of O/N atoms in different chemical structures. The method is based on the correlation of experimentally observed hydrogen-bonding strengths with quantum-mechanical derived properties, calculated on the acceptor atom (for hydrogen-bond acceptors) and on the heavy atom attached to the donor hydrogen (for hydrogen-bond donors). The properties giving the best correlation with the experimental hydrogen bonding scales were electrophilic superdelocalizability and self-atom polarizability. The best equations found have been implemented in a Web-based tool for hydrogen-bond strength prediction. © 2001 by Elsevier Science Inc.

Keywords: hydrogen bonding, quantum-mechanical properties, AM1, PM3, Web-based program, drug design, QSAR

INTRODUCTION

The importance of hydrogen bonding in drug design is well recognized; hydrogen-bonding capabilities deeply influence the transport and ADME (Adsorption, Distribution, Metabolism and Excretion) properties of a molecule as well as its specific interaction with biological receptors. Many QSAR studies have been reported in which hydrogen-bonding interactions play a key role in modeling a particular target activity.[1] The hydrogen-bonding parameters used in QSAR can be either experimentally determined or derived from calculated properties. Beside the obvious accuracy advantage in the use of experimental values, some disadvantages may be recognized. The scales are strongly dependent on the experimental conditions so that values can not be transferred easily from one scale to another. The process of searching and importing tabulated experimental data for a large set of molecules can be tedious, time consuming, and prone to errors. Finally, experimental values may be missing for particular structures. All of these problems can be overcome easily by theoretical methods; once a model for calculating a target property is determined, the procedure can be extended to any other molecule (within the confidence interval) and fully automated.

Numerous attempts have been made to replace experimental hydrogen-bonding values with theoretical descriptors. These attempts range from the simple count of the number of donors/acceptors in a molecule[1,2] to the application of additive or multiplicative empirical methods[3–5] and finally to the use of quantum-mechanical (QM) calculated properties. Dearden et al. have recently reviewed these studies.[6]

Focusing only on QM approaches, donor ability has been modeled by using:

- the energy of the lowest unoccupied molecular orbital (LUMO) and the atomic charge on the hydrogen atom,[7,8]
- the superdelocalizability and self-atom polarizability of the heavy atom attached to the donor hydrogen.[9]

Whereas, acceptor ability has been modeled by using:

- the energy of the highest occupied molecular orbital (HOMO) and the atomic charge on the most negatively charged atom in the molecule,[7,8]
- the molecular electrostatic potential (MEP).[10]

To the best of our knowledge, atomic properties such as the superdelocalizability and self-atom polarizability have not yet been applied to hydrogen-bond acceptors.

Despite what has been studied in the past, the correct prediction of hydrogen-bond capabilities is still an open problem. We have applied two different semi-empirical methods (AM1 and PM3) in calculating an extensive set of molecular/atomic properties and we have correlated them to experimental acceptor/donor scales. Our main objectives were to explore whether we could identify new and meaningful correlations between QM properties and hydrogen-bond ability and to extend to hydrogen-bond acceptors the work that has been performed previously on hydrogen-bond donors by employing superdelocalizability and self-atom polarizability properties. A further objective was to implement the results in an easy-to-use appli-

Corresponding author: Emanuela Gancia, Celltech Chiroscience Ltd., Granta Park, Great Abington, Cambridge, CB1 6GS, UK.

E-mail address: emanuela.gancia@celltechgroup.com

Journal of Molecular Graphics and Modelling 19, 349–362, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(00)00084-X

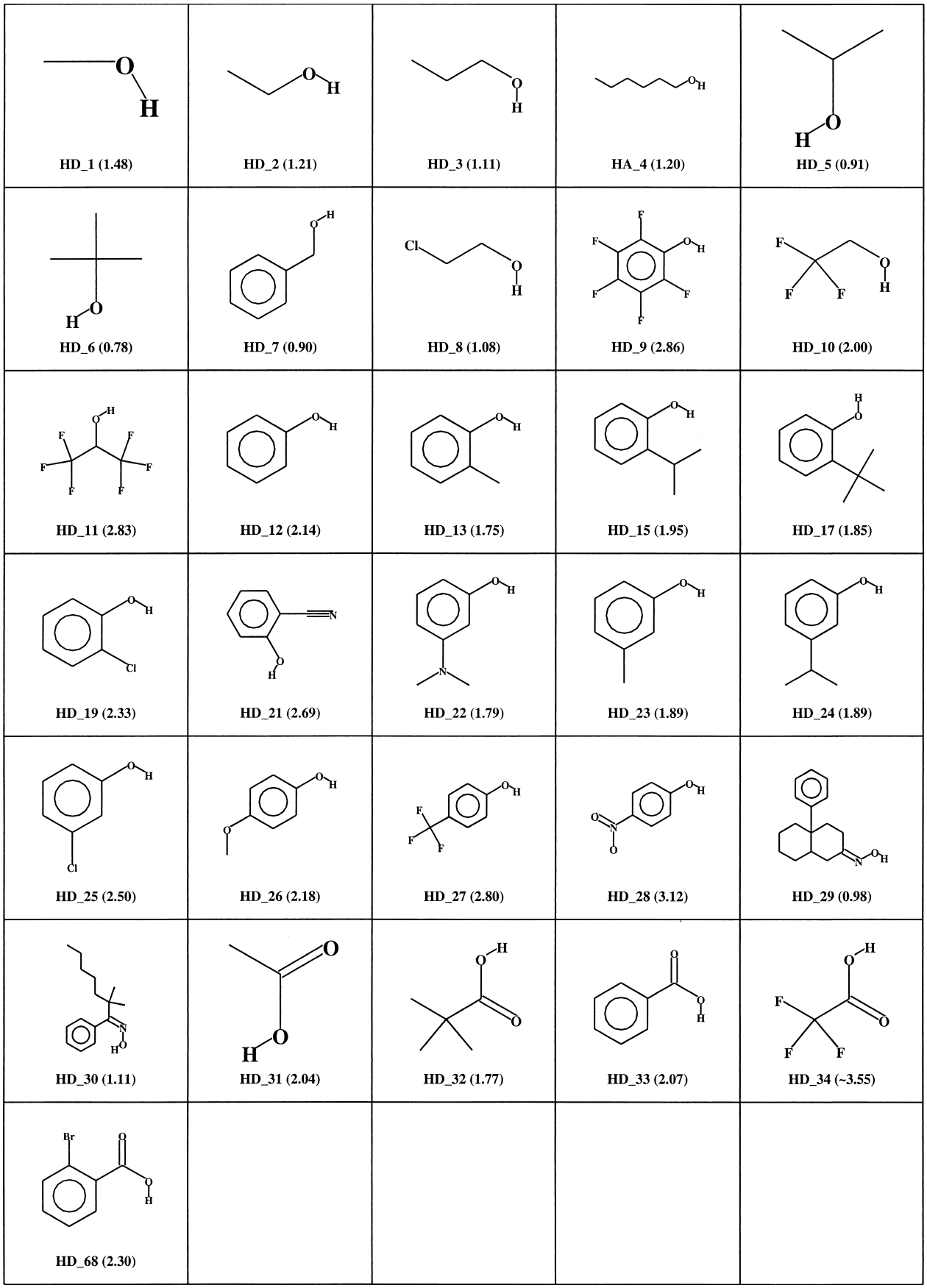


Figure 1. Thirty-one OH-type proton donors. The experimental logK_α values[11] are reported in parentheses. The values for compounds HD__9 and HD__68 were provided by Jeff Morris, AstraZeneca (personal communication).

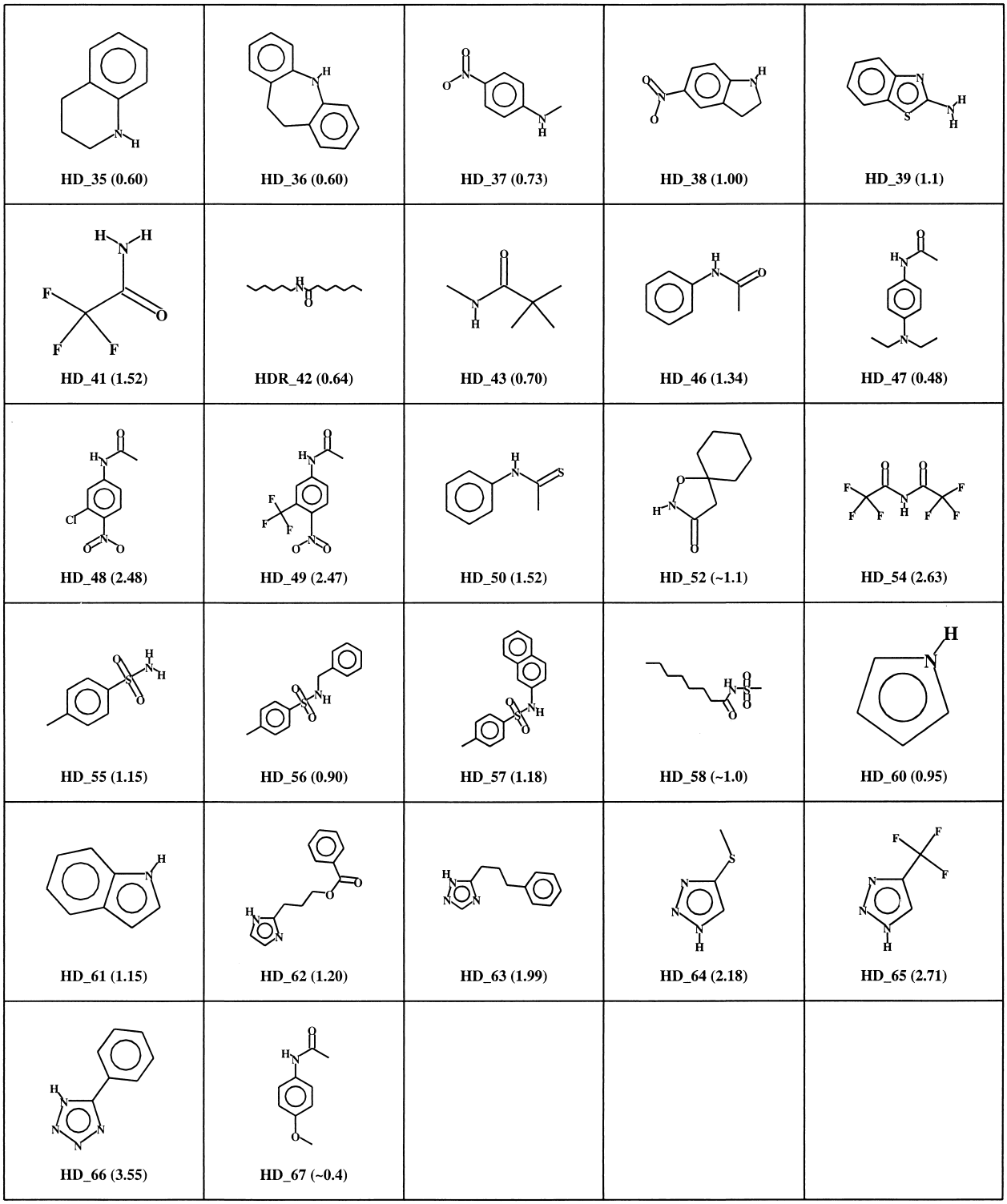


Figure 2. Twenty-seven NH-type proton donors. The experimental logK_α values [11] are reported in parentheses.

cation that could be directly used by chemists to predict hydrogen-bond strengths.

Many sources of experimental hydrogen bonding scales exist. As our interest was in predicting hydrogen-bonding strength in a drug design context, we found an ideal source of data in the measurements by Abraham et al. of the hydrogen-bonding equilibrium constants of a large variety of proton donors and acceptors, using 4-nitrophenol as the standard donor and N-methylpyrrolidinone as the standard acceptor.[11] 1,1,1-trichloroethane ($\epsilon = 7.53$) was used as solvent; due to its high dipolarity, it was considered to be a better model for real biological membranes than less polar solvents. Furthermore,

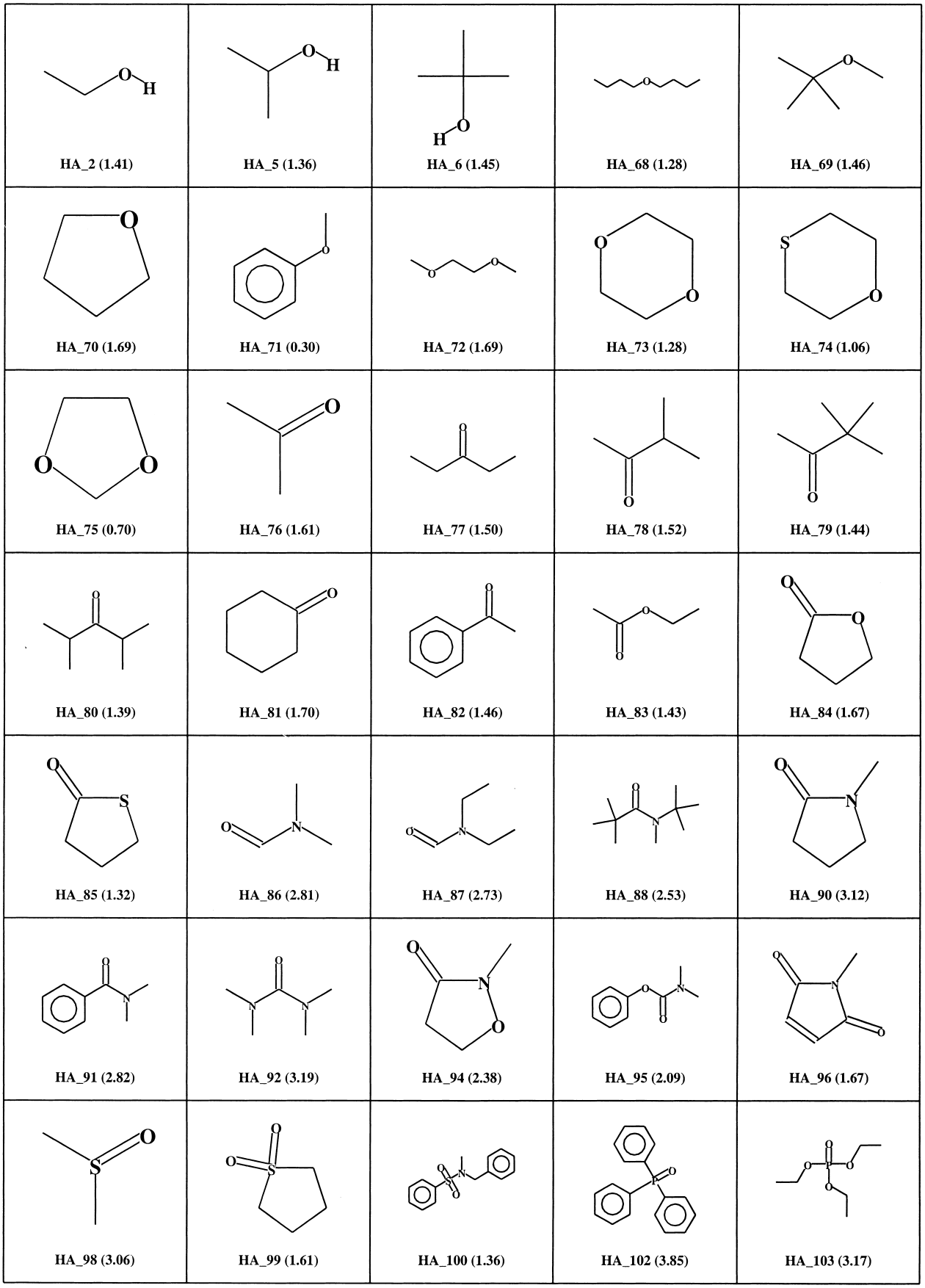


Figure 3. Thirty-five Oxygen containing proton acceptors. The experimental logK_β values [11] are reported in parentheses.

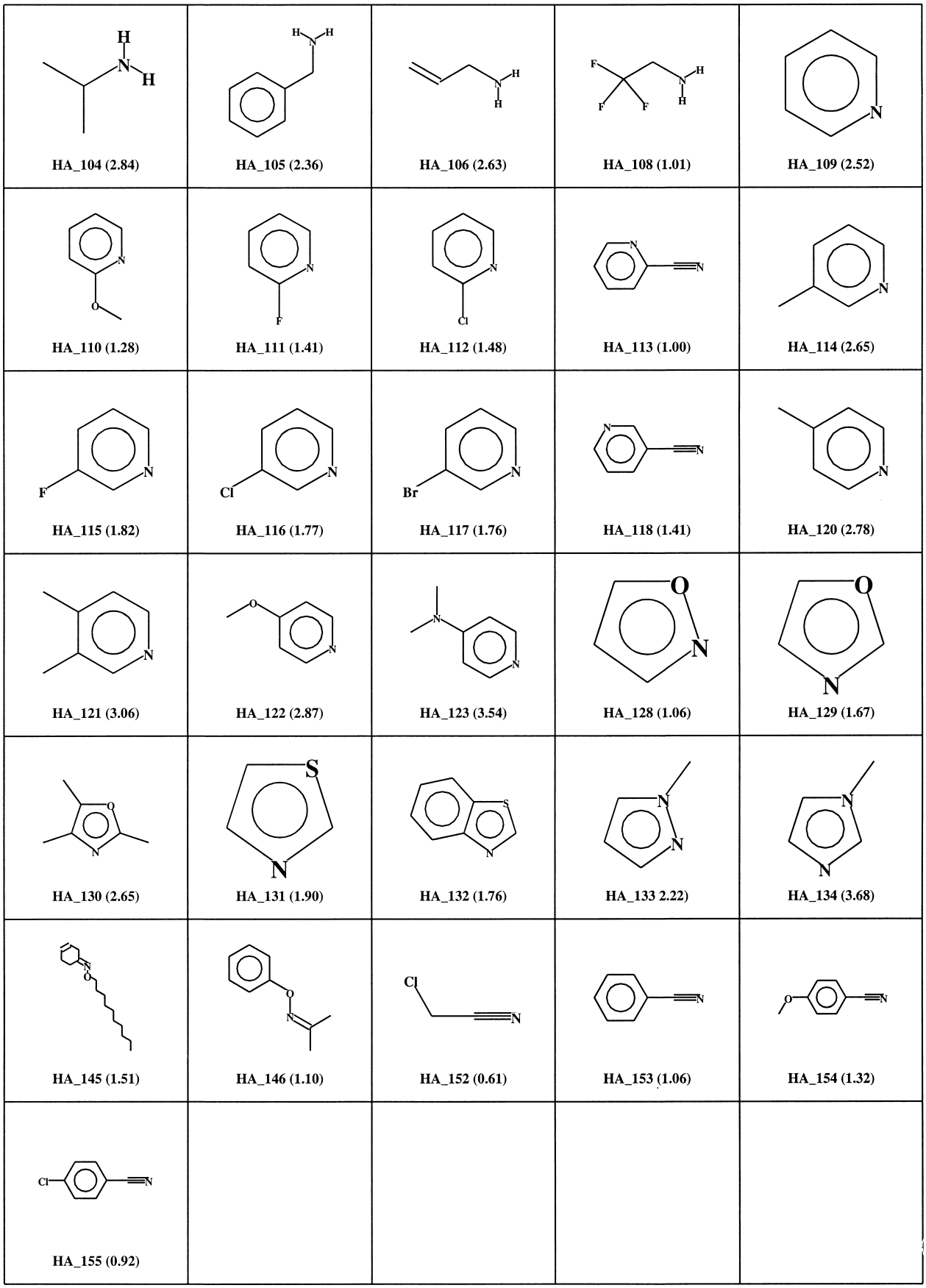


Figure 4. Thirty-one Nitrogen containing proton acceptors. The experimental logK_β values[11] are reported in parentheses.

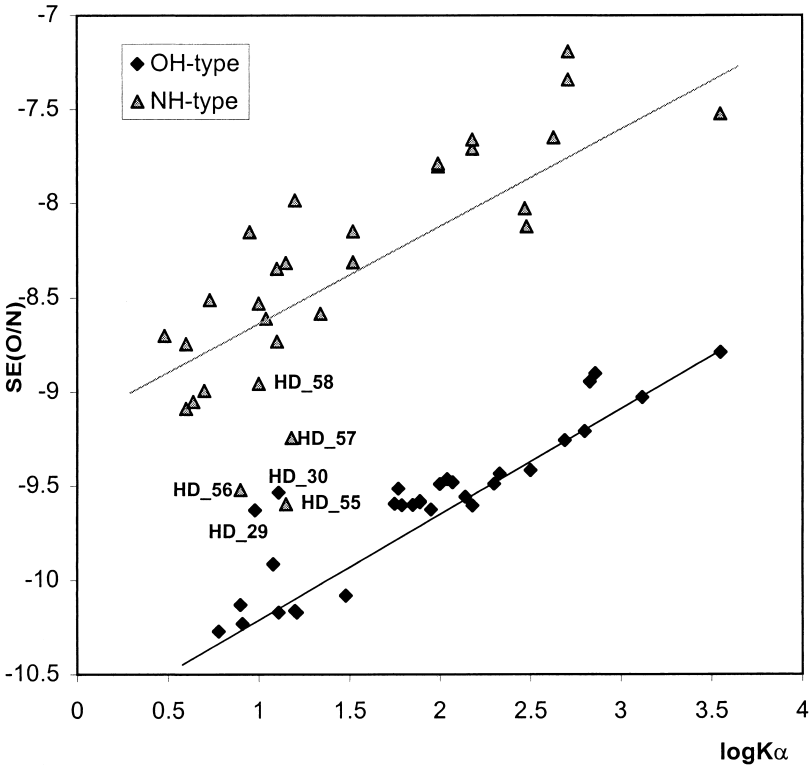


Figure 5. Scatter plot of $logK_\alpha$ versus SE of the electronegative atom bonded to the donor hydrogen for proton donors.

this data set offers good coverage of common donor/acceptor groups and is rich in heterocyclic compounds, thus making it an interesting set for medicinal chemists.

METHODS

Software

All the structures were generated as SMILES [12] strings and converted into 3D models using Corina.[13] The QM calculations were carried out with a modified version of Mopac 6.01 (Peter Bladon, Interchem Chemical Services, Glasgow), which is able to calculate additional atomic properties (keyword: PROPER).[14] Both the Sybyl[15] and Tsar[16] software packages were used for statistical analyses.

The modified version of Mopac 6.01 has been interfaced to Sybyl through an in-house SPL (Sybyl Programming Language) script. The script works on a Sybyl molecular database, creating the Mopac input file for each molecule in the databases and running Mopac. Once the QM calculation is completed for all the structures, a Sybyl spreadsheet is created and the relevant properties are imported. Subsequent QSPR (quantitative structure–property relationship) analyses are relatively straightforward.

Data set

The experimental hydrogen-bonding data have been taken from Abraham et al.[11] The values of $logK_\alpha$ and $logK_\beta$ were chosen as a measure of hydrogen bonding strength for hydrogen-bond donors and acceptors, respectively. All the structures used and the associated data are reported in Figures 1 and 2 (proton donors) and Figures 3 and 4 (proton acceptors). The numbering of the original paper has been kept for ease of reference. Some structures have been deleted from the original data set, in particular: (i) structures characterized by a potential ambiguity on the hydrogen-bond site (i.e. molecules containing more than one acceptor/donor of similar potency), (ii) structures where the acceptor site is neither nitrogen or oxygen; and (iii) structures where steric factors rather than the electronic distribution may heavily affect the logK (e.g., 2,6-disubsituted phenols). A total of 58 proton donors and 66 acceptors were analyzed.

Theoretical Descriptors

All the structures were fully optimized (EF routine), before any parameter calculation either using AM1[17] or PM3[18] Hamiltonian. In addition to the charge (Q), as defined in Mopac, some additional atomic properties[19] were computed from the eigenvectors $c_{\alpha j}$ and the eigenvalues λ_j, where α refers to the atomic orbital (i.e., s, p_x, p_y, and p_z) and j to the molecular orbital. Given a molecule with N molecular orbitals, whose levels from 1 to m are occupied, and an atom p with q atomic orbitals, these properties are defined as follows:

1. Electrophilic frontier electron density (FE)

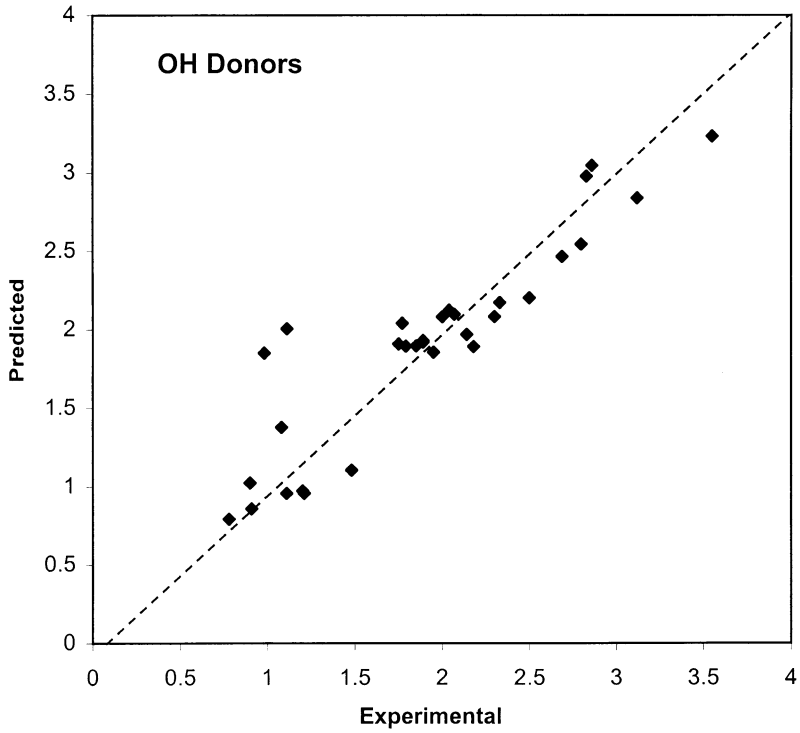
$$FE(p) = \sum_{a=1,\,q} c_{\alpha m}^2 \qquad (1)$$

is simply the sum of all the squared eigenvectors of p on the HOMO.

2. Nucleophilic frontier electron density (FN)

$$FN(p) = \sum_{\alpha=1,\,q} c_{\alpha(m+1)}^2 \qquad (2)$$

is simply the sum of all the squared eigenvectors of p on the LUMO.



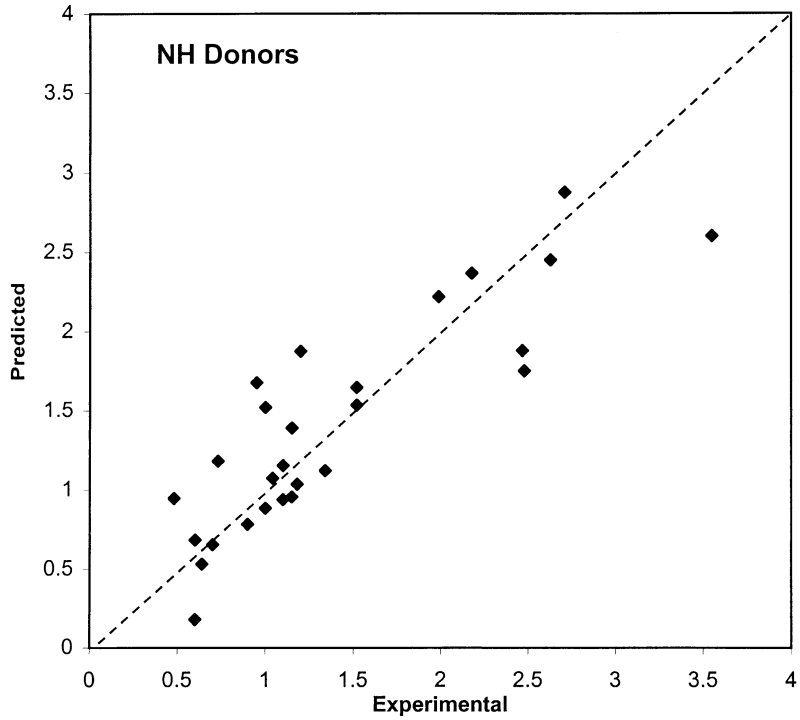


Figure 6. Predicted versus experimental logK_α for proton donors according to equations 1 (OH donors) and 4 (NH donors).

3. Electrophilic superdelocalizability (SE)

$$SE(p) = 2* \sum_{j=1,m} \sum_{\alpha=1,q} (c_{\alpha j}^2/\lambda_j) \qquad (3)$$

where the sum is over all the atomic orbitals of *p* and all the occupied molecular orbitals.

4. Nucleophilic superdelocalizability (SN)

$$SN(p) = 2* \sum_{j=m+1,N} \sum_{\alpha=1,q} (c_{\alpha j}^2/ - \lambda_j) \qquad (4)$$

where the sum is over all the atomic orbitals of *p* and all the unoccupied molecular orbitals.

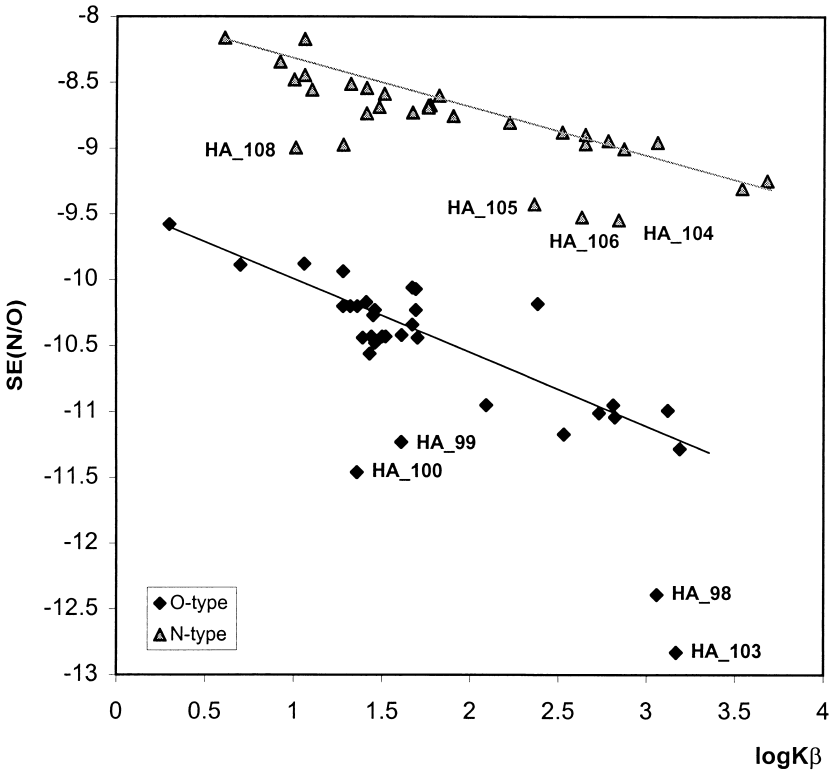


Figure 7. Scatter plot of $logK_\beta$ versus SE of the acceptor atom for proton acceptors. Compound HA__102 (SE=−14.72) is not reported in the plot.

5. Radical superdelocalizability (SR)

$$SR(p) = \sum_{j=1,m} \sum_{\alpha=1,q} (c_{\alpha j}^2/\lambda_j) + \sum_{j=m+1,N} \sum_{\alpha=1,q} (c_{\alpha j}^2/ - \lambda_j). \quad (5)$$

6. Atom self-polarizability (ALP)

$$ALP(p) = -4* \sum_{j=1,m} \sum_{\alpha=1,q} \sum_{k=m+1,N} (c_{\alpha j}^2 * c_{\alpha k}^2/\lambda_j - \lambda_k). \quad (6)$$

All the above properties have been calculated on the acceptor atom for proton acceptors and on the electronegative atom bonded to the donor hydrogen for proton donors. The following molecular properties were also taken into account:

- Energy of the highest occupied molecular orbital (E_{HOMO})
- Energy of the lowest unoccupied molecular orbital (E_{LUMO}).

RESULTS AND DISCUSSION

The two data sets of 58 proton donors and 66 proton acceptors have been analyzed with the following steps: (i) the correlation coefficients between experimental hydrogen bonding strengths and all the QM parameters were calculated; (ii) the scatter plots of experimental hydrogen-bonding strengths versus the most correlated properties were analyzed; and (iii) linear equations were derived by using multiple linear regression (MLR). The most significant findings are reported as follows.

Proton Donors

As the calculated atomic properties are strongly dependent on the atom type, we felt it would be reasonable to further divide the data set into two classes: 31 OH-type and 27 NH-type donors (Figure 1, Figure 2). The correlation between each electronic property and the experimental $logK_\alpha$ values was calculated for the entire data set as well as for each single class. When considering the whole data set, none of the calculated properties showed a high correlation to $logK_\alpha$. However, on separately analyzing the OH and NH donors, reasonable coorrelations were found. Considering the OH donor subset, the electrophilic superdelocalizability of the donor oxygen (SE_O) has a correlation coefficient of 0.91 (AM1) and 0.92 (PM3) with $logK_\alpha$. In the NH donor subset, SE_N is slightly worse, with r = 0.76 (AM1) and 0.68 (PM3). We found that, in some cases, AM1 and PM3 results were not consistent. For example (NH donors) SN_N calculated by AM1 has a correlation coefficient of −0.44 with $logK_\alpha$, whereas, when the same property is calculated by PM3, the correlation is −0.03. A closer inspection into each single value revealed that the SN values of a few compounds are completely outside the range covered by all of the other compounds. This is true for either AM1 or PM3, but affects different compounds in the two cases. We did not find any obvious explanation for these anomalous values, apart from the consideration that SN is based on the coefficients of all the unoccupied orbitals and refers to a "virtual" rather than to a real electronic density. The SN parameter was therefore not considered any further.

As this preliminary work consistently demonstrated the highest correlation between electrophilic superdelocalizability and $logK_\alpha$, we investigated the scatter plots of $SE_{O/N}$ versus $logK_\alpha$ (Figure 5). OH donors and NH donors are clearly separated in this plot. While the slopes of these lines appear to be similar, their intercepts differ. The few outliers belong to well-defined chemical classes: the only 2 oximes

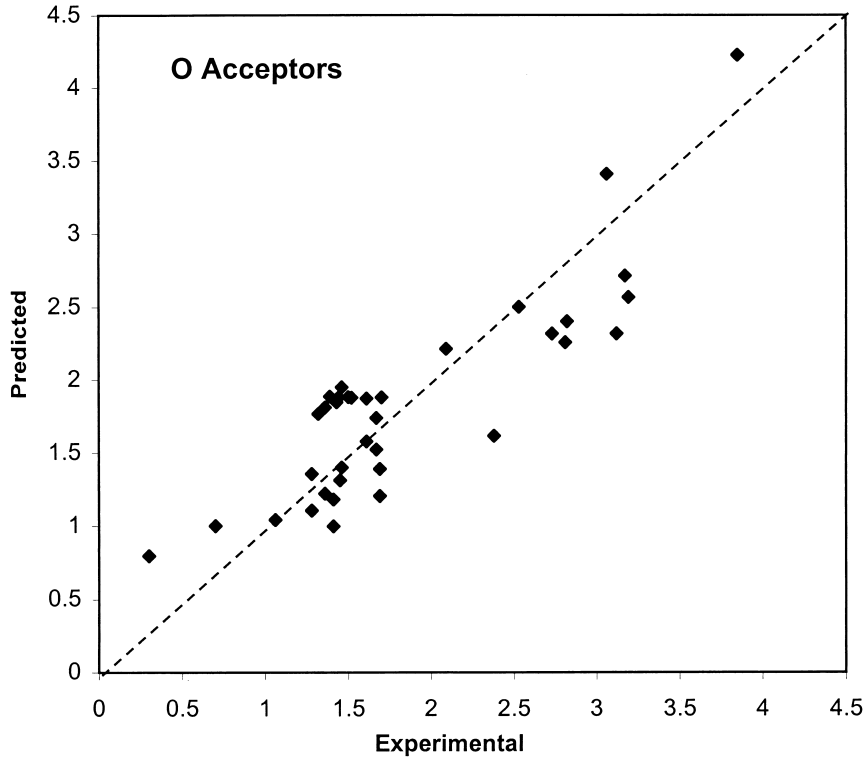
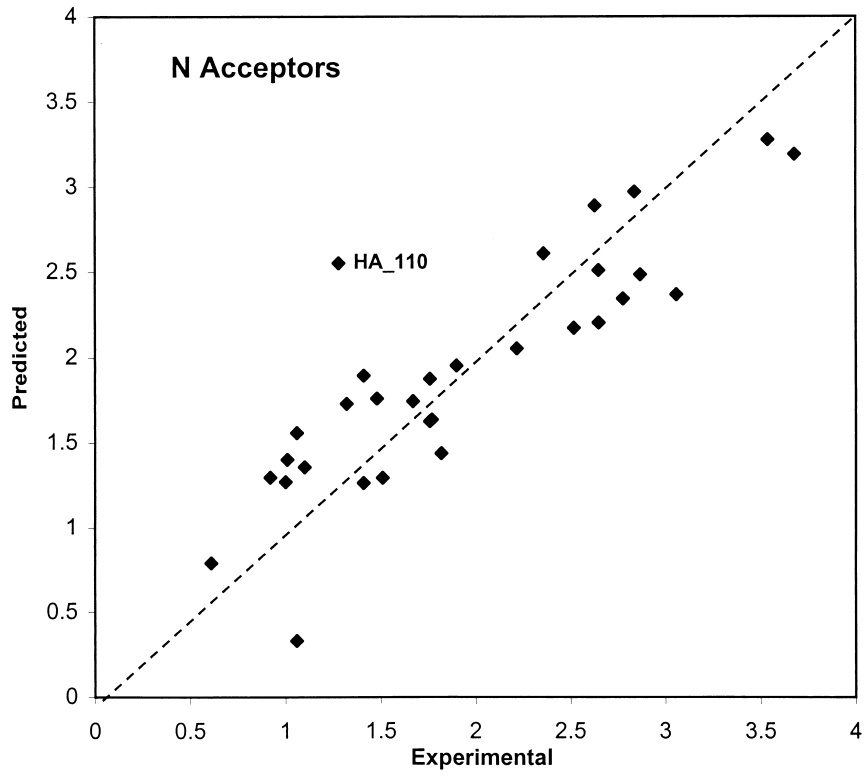


Figure 8. Predicted versus experimental logK_β for proton acceptors according to equations 7 (O acceptors) and 11 (N acceptors).



in the data set for OH donor class (HA__29 and HA__30, Figure 1) and the sulphonamides for NH donor class (HA__55−58, Figure 2).

MLR, as implemented in Tsar, was used to derive the following equations. Our objective was to obtain statistically significant models, while keeping the number of parameters as low as possible. The cross-validated results (r_{cv}^2) refer to the leave-one-out protocol.

- Best AM1 model for OH donors:
  $\log K_\alpha = 1.645 * SE_O + 17.689$ (1)
  $n = 31, r = 0.91, r^2 = 0.83, r_{cv}^2 = 0.81, F = 137, s = 0.306$
- Best PM3 model for OH donors:
  $\log K_\alpha = 1.713 * SE_O + 18.629$ (2)
  $n = 31, r = 0.92, r^2 = 0.85, r_{cv}^2 = 0.83, F = 92.45, s = 0.321$
- Best AM1 models for NH donors:
  $\log K_\alpha = 1.420 * SE_N + 13.276$ (3)
  $n = 23/27, r = 0.86, r^2 = 0.74, r_{cv}^2 = 0.69, F = 59.95, s = 0.437$
  $\log K_\alpha = 1.380 * SE_N + 0.535\ ALP_N + 22.475$ (4)
  $n = 27, r = 0.87, r^2 = 0.75, r_{cv}^2 = 0.68, F = 31, s = 0.434$
- Best PM3 model for NH donors:
  $\log K_\alpha = 0.939 * SE_N + 9.432$ (5)
  $n = 23/27, r = 0.78, r^2 = 0.60, r_{cv}^2 = 0.53, F = 32.62, s = 0.537$

The electrophilic superdelocalizability of oxygen (SE_O) is sufficient to model hydrogen bonding for the OH class, giving an r^2 of 0.83 (AM1) and 0.85 (PM3); this could be increased to 0.93 by removing the two oximes from the data set (data not shown). When analyzing the NH donors, the sulphonamides were detected as outliers and removed to obtain Equation 3 and Equation 5 (23 out of the possible 27 compounds were used). The introduction of a second parameter, the atom self-polarizability (Equation 4), gave a model with an r^2 of 0.75 (AM1) for all 27 compounds. For comparison, the r^2 against SE_N (AM1) alone for all the 27 compounds was 0.58 (equation not shown).

Previous studies have already reported good correlations between ALP and SE properties and hydrogen-bond donor ability.[9] Nevertheless, a more thorough comparison between our study and previous reports is difficult due to the different sources of experimental data used. Our study confirms that SE values are able to explain the hydrogen donor capability for a variety of drug-relevant groups. SE is calculated by summing up the squared coefficients of the occupied orbitals, and is therefore directly proportional to the electron density over the atom: the lower the electron density on the atom bonded to the hydrogen, the more easily the hydrogen will be donated. The scatter plots of predicted versus experimental $\log K_\alpha$, obtained by using Equation 1 and Equation 4 are reported in Figure 6.

Proton Acceptors

To analyze the 66 proton acceptors, we applied the same computational methods used for the donor set. The structures were divided into 35 O-type and 31 N-type acceptors (Figure 3 and Figure 4) and the correlation between various electronic parameters and $\log K_\beta$ were investigated.

The correlation coefficients for acceptor groups were poorer, suggesting that finding a simple equation correlating hydrogen-bonding strengths and QM parameters is likely to be more difficult for acceptors than donors. This is not surprising if we consider that in the latter case much more chemical variation is possible (possible acceptor atoms include sp^2 and sp^3 oxygen, sp, sp^2 and sp^3 nitrogen). Once again, however, the best correlation coefficients were found to involve SE.

The inspection of the $\log K_\beta/SE_{O/N}$ scatter plot (Figure 7)

Table 1. Predicted and experimental hydrogen bonding scales for common functional groups[a]

	Donors ($\log K_\alpha$)		Acceptors ($\log K_\beta$)	
Group	Exp.	Calc.	Exp.	Calc.
Alk NH2	[b]	0	2.8	2.8
Ar NH2	0.6	0.3	1.0	1.9
Ar NH Ar	0.6	0.7	0.4	0.8
Alk NHO Alk	0.8	1.0		
Alk NH NO2	1.6	1.4		
Alk C=N OAlk			1.5	0.9
Alk C=N OAr			1.1	0.9
Amidine C=N			3.4	5.8
Guanidine C=N			4.2	7.2
Alk C#N			1.0	1.6
Ar C#N			1.2	1.5
Alk SCN			0.9	1.7
Cyanamide			2.0	2.9
Cyanoguanidine			2.9	5.1
Alk OH	1.2	1.1	1.4	1.1
Alk O Alk			1.5	1.2
Ar OH	2.1	2.0	0.6	0.7
Alk NOH	1.6	1.6		
Alk C=NOH	1.0	2.0		
Ar C=NOH	1.2	2.1		
Alk COOH	2.0	2.1		
Ar COOH	2.0	2.1		
Alk CHO			1.1	1.7
Ar CHO			1.2	1.7
Alk CO Alk			1.6	1.9
Ar CO Alk			1.4	1.9
Ar CO Ar			1.4	1.9
Alk COOAlk			1.4	1.8
Ar COO Alk			1.2	1.9
Lactone			1.7	2.0
Alk CONH Alk	0.7	0.7	3.0	2.5
Ar CONH Alk	0.7	0.7	2.8	2.5
Alk CONH Ar	1.3	1.1	2.5	2.1
Lactam			3.1	2.5
Cyclic imide	1.1	1.3	1.4	1.6
ALK NHCONH Alk	2.1	0.5	3.2	2.8
Alk OCONH Alk			2.4	2.2
Ar OCONH Alk			2.1	2.2
Alk OCONH Ar	1.1	1.1	2.0	2.2
ArNO2			0.7	1.6
Het N-Oxide			3.5	2.9
Alk SO Alk			3.0	3.4
Ar SO Alk			2.7	3.4
Ar SO Ar			2.4	3.3
Alk SO2 Alk			1.6	1.6
Ar SO2 Alk			1.4	1.7
Sulphonamide	1.2	1.2	1.4	1.6
Acylsulphonamide	1.0	0.9	1.0	1.7
Ar PO			3.9	2.8
AlkO3 PO			3.2	2.5
Alk CSNH Alk	1.0	1.2		
Ar CSNH Alk	1.0	1.2		
Alk CSNH Ar	1.5	1.5		
Alk NHCSNH Alk	2.1	0.9		

[a] Data from ref. 11, table 4.
[b] Negligible hydrogen bonding ability.

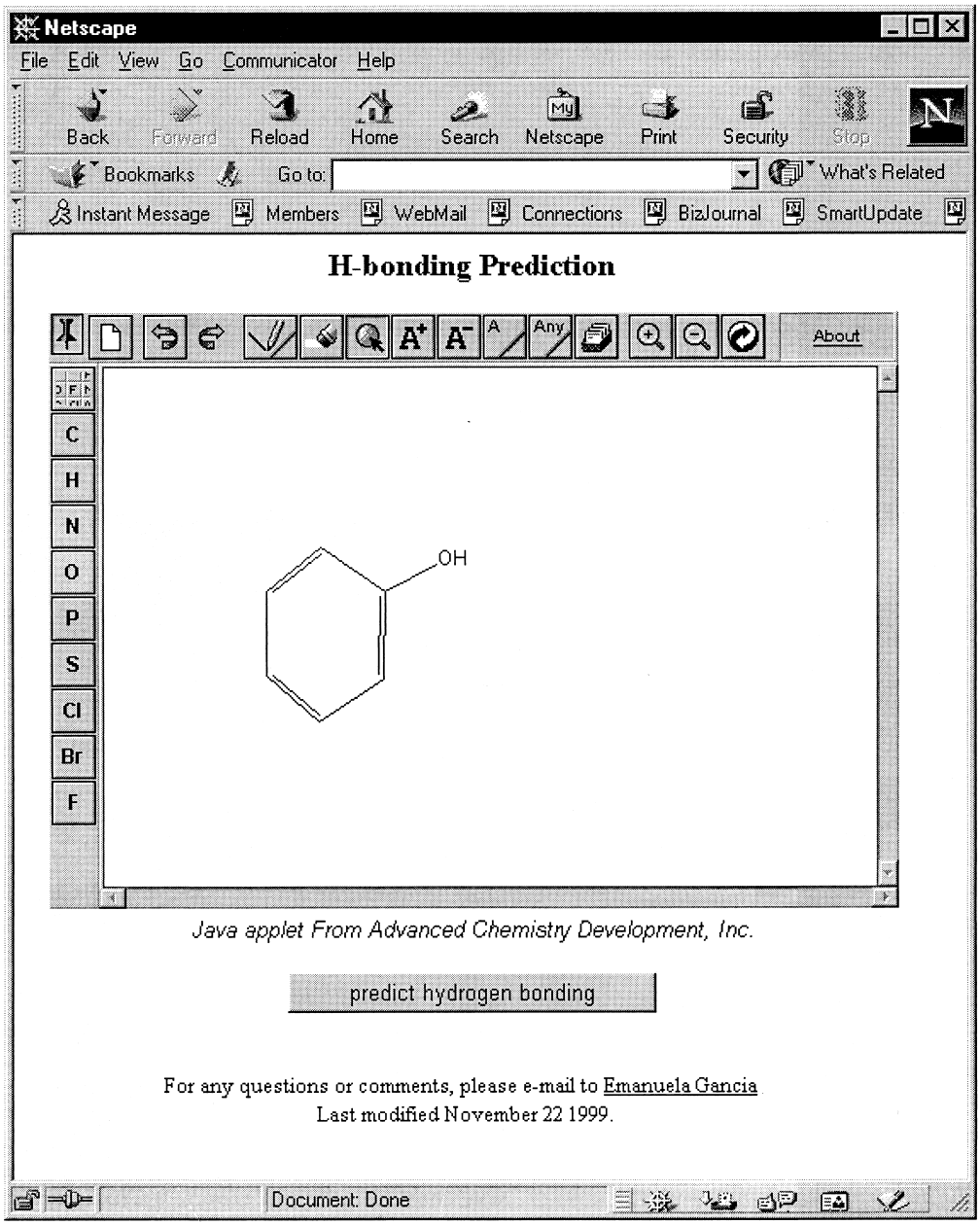


Figure 9. Input HTML page of the in-house Web-based hydrogen bonding calculator.

allowed for the identification of possible outliers in the O-type acceptor class: HA__98–103 (compounds containing SO and PO groups) and in the N-type acceptors: HA_104–108 (amines).

The best regression models found for both classes are reported here.

- Best models for O acceptors (AM1):
  $\log K_\beta = -1.430 * SE_O - 13.161$ (6)
  $n = 30/35$, $r = 0.87$, $r^2 = 0.76$, $r_{cv}^2 = 0.73$, $F = 91$, $s = 0.338$
  $\log K_\beta = -1.012 * SE_O - 0.302 * ALP_O - 12.973$ (7)
  $n = 35$, $r = 0.87$, $r^2 = 0.76$, $r_{cv}^2 = 0.69$, $F = 50$, $s = 0.404$
- Best models for O acceptors (PM3):
  $\log K_\beta = -1.430 * SE_O - 13.558$ (8)
  $n = 30/35$, $r = 0.82$, $r^2 = 0.67$, $r_{cv}^2 = 0.62$, $F = 56.06$, $s = 0.402$
  $\log K_\beta = -0.804 * SE_O - 0.386 * ALP_O - 159$ (9)
  $n = 35$, $r = 0.77$, $r^2 = 0.59$, $r_{cv}^2 = 0.49$, $F = 22.67$, $s = 0.527$
- Best models for N acceptors (AM1):
  $\log K_\beta = -2.642 * SE_N - 21.141$ (10)
  $n = 27/31$, $r = 0.89$, $r^2 = 0.80$, $r_{cv}^2 = 0.76$, $F = 100$, $s = 0.375$
  $\log K_\beta = -2.749 * SE_N - 0.590 * ALP_N - 33.976$ (11)
  $n = 31$, $r = 0.86$, $r^2 = 0.74$, $r_{cv}^2 = 0.67$, $F = 39.62$, $s = 0.432$
- Best models for N acceptors (PM3):
  $\log K_\beta = -2.497 * SE_N - 22.403$ (12)
  $n = 27/31$, $r = 0.83$, $r^2 = 0.69$, $r_{cv}^2 = 0.65$, $F = 55.90$, $s = 0.467$

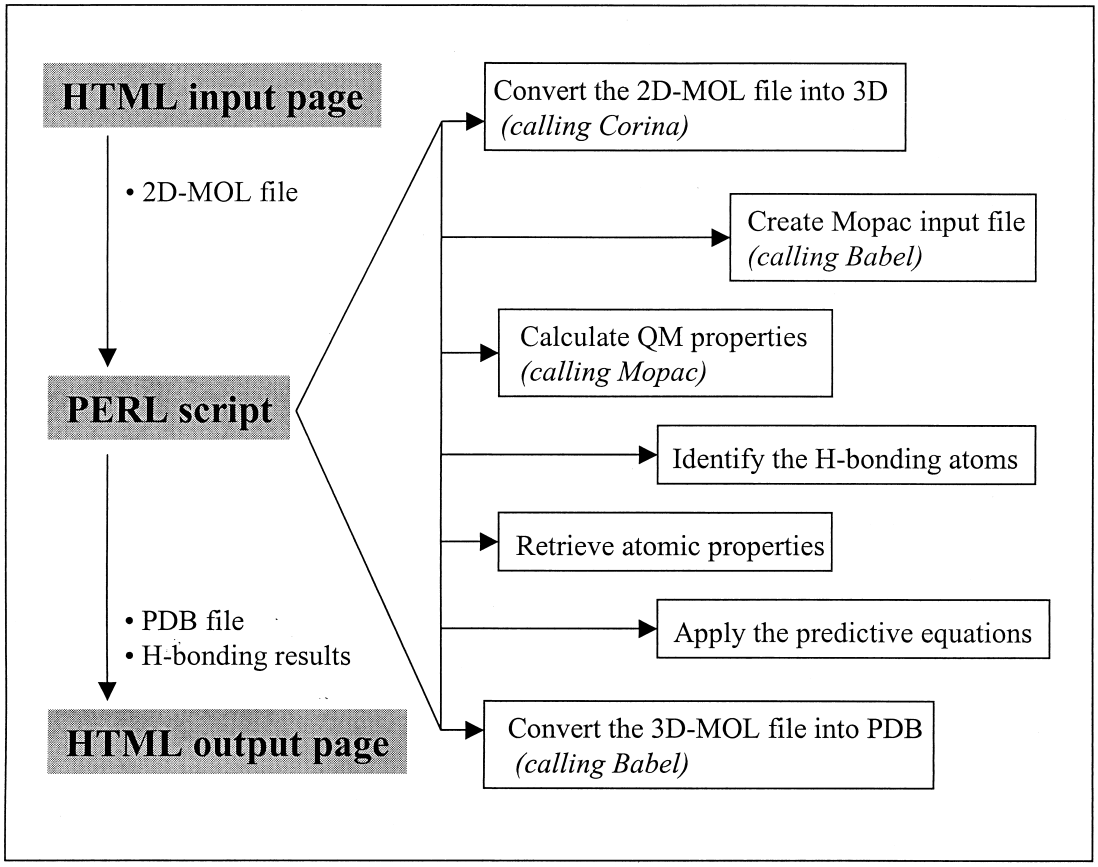


Figure 10. Flowchart of the in-house Web-based hydrogen bonding calculator.

$\log K_\beta = -2.202 * SE_N + 0.317 * ALP_N - 12.317$ (13)
$n = 31$, $r = 0.87$, $r^2 = 0.76$, $r_{cv}^2 = 0.72$, $F = 44.48$, $s = 0.414$

Following the removal of compounds containing SO or PO groups, we derived Equation 6 and Equation 8 (obtained using 30 out of the 35 compounds). As in the case of donors, the addition of ALP proved to be helpful in modeling hypervalent compounds (HA__98–103) and a 2-parameter equation (Equation 7) with an $r^2 = 0.76$ (AM1) could satisfactorily model all the 35 O acceptors. For comparison, the parameter SE_O (AM1) alone, applied to all 35 molecules, gave an r^2 of 0.62 (equation not shown). When considering the N acceptors, the amines (HA__104–108) were removed so that equations 10 and 12 were derived using 27 out of the original 31 compounds. The 2-parameter equations, Equation 11 and Equation 13, were obtained by using all 31 of the nitrogen compounds. For the entire training set of 31 compounds, the parameter SE_N alone (AM1) gave an r^2 of 0.61 (equation not shown), whereas the addition of ALP_N increased the r^2 to 0.74 (Equation 11).

The statistics of the equations for acceptors are not as good as those found for donors. However, they were considered of further use. Moreover, the use of consistent parameters (i.e., SE and ALP) between donors and acceptors was encouraging. AM1 and PM3 provided similar models, with slightly better results being shown for AM1 in modeling the proton acceptor capability. The scatter plots of predicted versus experimental $\log K_\beta$ (according to Equation 7 and Equation 11) are shown in Figure 8. The outlier HA__110 belongs to the 2-substituted pyrimidine class, whose anomalous behavior has been highlighted by Abraham et al. in the original article.[11]

PREDICTION OF THE HYDROGEN-BOND STRENGTHS OF COMMON FUNCTIONAL GROUPS

As all the reported equations contain only one/two parameters for 23–35 structures, the risks of chance correlation as well as of overfitting are minimized. Moreover, the selected parameters appear reasonable and can be interpreted in terms of the properties they represent. We further validated our study by applying equations 1, 4, 7, and 11 to predict the hydrogen-bonding values for common functional groups that are reported in Table 4 of Abraham et al.[11] This further analysis provided at the same time an estimate of the confidence level of our predictions for different chemical classes. The results are reported in Table 1. The agreement between calculated and experimental hydrogen strengths for hydrogen-bond donors (column 1 and column 2) is quite impressive: ureas and thioureas (AlkNHCONHAlk and AlkNHCSNHAlk) are the only classes poorly predicted, with an unsigned error of 1.6 and 1.2, respectively. When analyzing the hydrogen-bond acceptors (column 3 and column 4), the worst predictions involve mainly groups containing C=N, CN, SO, SO2 and PO functionalities.

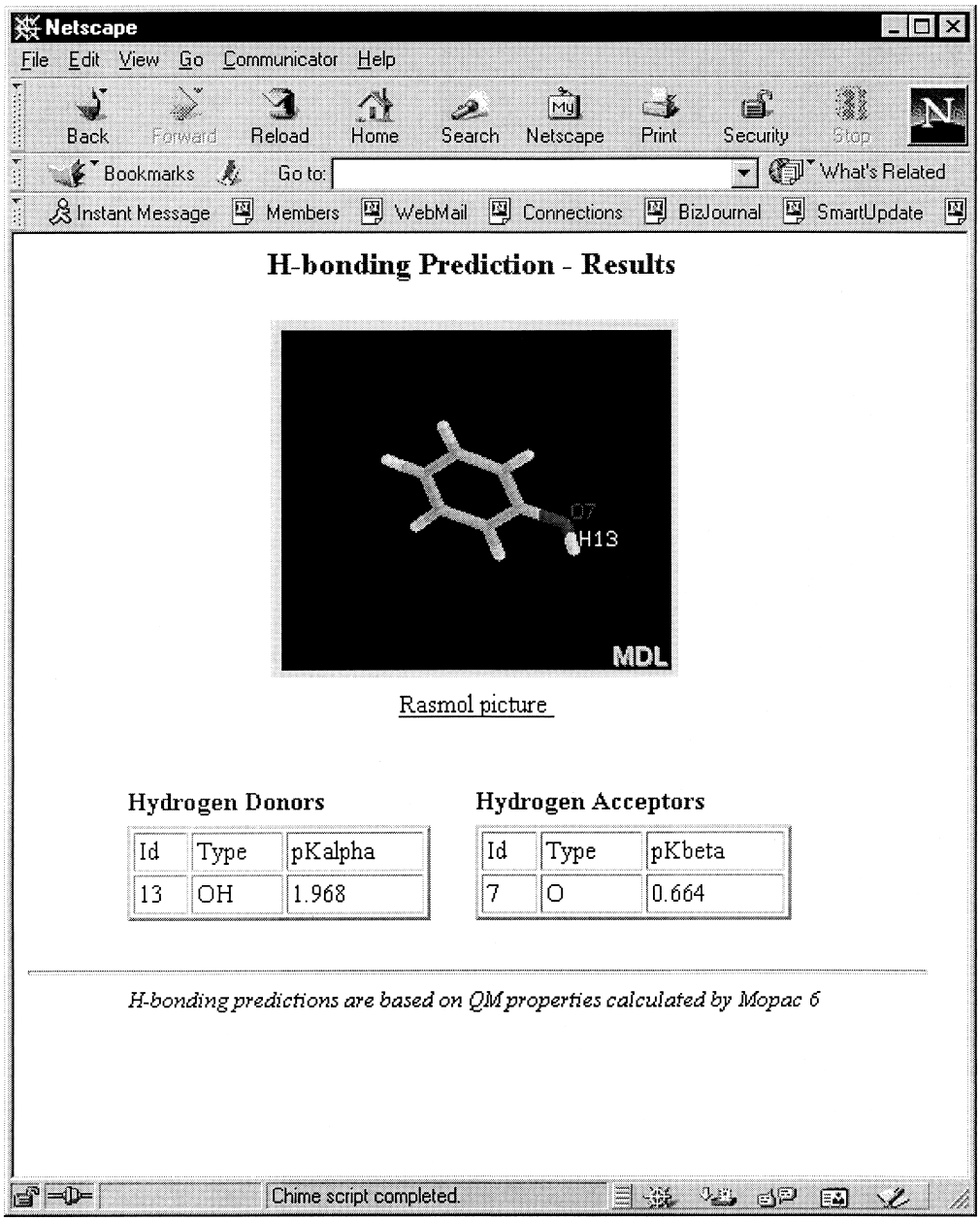


Figure 11. Output HTML page of the in-house Web-based hydrogen bonding calculator.

Development of a Web-Based Tool to Predict Hydrogen-Bonding Capabilities

The Web environment provides the means to create simple, interactive, easy-to-use tools for processing chemical information.[20,21] We therefore decided to implement our results via a Web-based calculator of hydrogen-bond strengths that could be readily accessible to the chemistry staff through the company Intranet site.

The PC user front-end of the tool is an HTML page containing a structure drawing applet, written in JAVA[22] (Figure 9). Chemical structures drawn into the applet can be exported as a MOL file. The file is then processed by an in-house cgi-bin PERL script, according to the scheme reported in Figure 10. The PERL script transfers the information contained in the MOL file to an Octane Silicon Graphics workstation, where all the calculations are performed. A 3D model of the molecule is created by Corina and then it is converted by Babel[23] into a DAT file, which is submitted to Mopac (with the keywords AM1, PROPER, and EF). At the end of the QM calculation, the atomic properties are extracted from the output file. Following this, all potential hydrogen-bonding atoms are identified and their strengths are calculated according to Equations 1, 4, 7, and 11.

Finally, the output results are converted in HTML format and sent back to the PC. An example of an output page is shown in Figure 11. The MDL Chime plug-in[24] is used to display the structure and view the atom numbering.

CONCLUSIONS

We have modeled the proton donor and acceptor capabilities of a large variety of functional groups by using QM parameters. In particular, we have highlighted a strong correlation between

hydrogen bonding strengths and the atomic superdelocalizability and self-polarizability of the atoms involved. While we have obtained satisfactory results by using calculated descriptors that are largely electrostatic in nature, a recent Compton-scattering experiment[25] supports a partially covalent picture of the hydrogen bond. These experimental results may encourage future modeling efforts to introduce parameters able to account for the covalent contribution of hydrogen-bonding interactions.

Despite some limitations (our method can predict only O/N acceptors and NH/OH donors, it cannot explain steric effects on hydrogen bonding, there are potential ambiguities in handling tautomers and conformational flexibility has not been addressed), we believe that the approach we have described can provide useful hydrogen-bonding strength estimates for drug design. To make the use of our models straightforward for chemists, we have implemented the hydrogen-bonding calculation in a Web-based tool.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Peter Bladon (Interchem Chemical Services, Glasgow) and Dr. Jeff Morris (AstraZeneca) for their helpful discussions. In addition, we would like to extend our special thanks to Dr. Bladon for access to his modified version of Mopac.

REFERENCES

1 Fujita, T., Nishioka, T., and Nakajima, M. Hydrogen-bonding parameters and its significance in quantitative structure–activity studies. *J. Med. Chem.* 1977, **20**, 1071–1081
2 Charton, M., and Charton, B.I. The structural dependence of amino acid hydrophobicity parameters. *J. Theoret. Biol.* 1982, **99**, 629–644
3 Seiler, P. Interconversion of lipophilicities from hydrocarbon/water into the octanol/water system. *Eur. J. Med. Chem.* 1974, **9**, 473–479
4 Moriguchi, I. Quantitative structure–activity studies 1. Parameters relating to hydrophobicity. *Chem. Pharm. Bull.* 1975, **23**, 247–257
5 Raevsky, O., Grigor'ev, V., and Mednikova, E. QSAR H-bonding descriptions. In: *Trends in QSAR and molecular modelling 92*, Wermuth, C.G., Ed., ESCOM, Leiden, 1993, pp. 116–119
6 Dearden, J.C., and Ghaufurian, T. Hydrogen bonding parameters for QSAR: comparison of indicator variables, hydrogen bond counts, molecular orbital and other parameters. *J. Chem. Inf. Comput. Sci.* 1999, **39**, 231–235
7 Wilson, L. Y., and Famini, G. R. Using theoretical descriptors in quantitative structure–activity relationships: some toxicological indices. *J. Med. Chem.* 1991, **34**, 1668–1674
8 Dearden, J.C., Cronin, M.T.D., and Wee, D. Investigation of calculated hydrogen bonding parameters for QSAR. In: *QSAR and molecular modelling: concepts, computational tools and biological applications*, Sanz, F., Giraldo, J., and Manaut, F., Eds., Proust Science Publishers, Barcelona, 1995, pp. 117–119
9 Dearden, J.C., Cronin, M.T.D., and Wee, D. Prediction of hydrogen bond donor ability using new quantum chemical parameters. *J. Pharm. Pharmacol.* 1997, **49 Suppl. 4**, 110
10 Murray, J.S., and Politzer, P. Correlation between the solvent hydrogen-bond-donating parameter α and the calculated molecular surface electrostatic potential. *J. Org. Chem.* 1991, **56**, 6715–6717
11 Abraham, M.H., Duce, P.P., Prior, D.V., Barratt, D.G., Morris, J.J., and Taylor, P.J. Hydrogen bonding. Part 9. Solute proton donor and proton acceptor scales for use in drug design. *J. Chem. Soc. Perkin Trans. II*, 1989, 1355–1375
12 Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 1988, **28**, 31–36
13 Sadowski, J., and Gasteiger, J. From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *Chem. Rev.* 1993, **93**, 2567–2581
14 Stewart, J. J. P. Mopac program package, quantum chemistry program exchange no. 455
15 Sybyl, Tripos Associates Inc., 1699 S. Hanley Rd., St. Louis, MO 63144
16 Tsar, version 3.21, Oxford Molecular Ltd, The Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, UK
17 Dewar, M.J.S., Zoebish, E.G., Healy, E.F., and Stewart, J.J.P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 1985, **107**, 3902–3909
18 Stewart, J.J.P. Optimization of parameters for semiempirical methods 1. Method. *J. Comput. Chem.* 1989, **10**, 209–220
19 Fukui, K., Yonezawa, T., and Nagata, C. Theory of substitution in conjugated molecules. *Bull. Chem. Soc. Japan* 1954, **27**, 423–427
20 Ertl, P., and Jacob, O. WWW-based chemical information system. *J. Mol. Struc.* (*Theochem)* 1997, **419**, 113–120
21 Ertl, P. World wide web-based system for the calculation of substituent parameters and substituent similarity searches. *J. Mol. Graphics Mod.* 1998, **16**, 11–13
22 ACD Structure Drawing Applet 3.1, freeware from Advanced Chemistry Development Inc., 133 Richmond Street West, Suite 605, Toronto, Ontario, Canada M5H 2L3, http://www.acdlabs.com/download/sda.html
23 Babel 1.6, P. Walters and M. Stahl, University of Arizona
24 Chime plug-in, distributed by MDL Information System Inc., San Leandro, CA, USA, http://www.mdli.com/support/chime/license.html
25 Isaacs, E.D., Shukla, A., Platzman, P.M., Hamann, D.R., Barbiellini, B., Tulk, C.A. Covalency of the hydrogen bond in ice: a direct X-ray measurement. *Phys. Rev. Lett.* 1999, **82**, 600–603