Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 607–612
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Investigating the utility of momentum-space descriptors for predicting blood–brain barrier penetration

Jabir H.A. Al-Fahemi [a,1], David L. Cooper [a,*], Neil L. Allan [b]

[a] Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, UK
[b] School of Chemistry, University of Bristol, Cantocks Close, Bristol BS8 1TS, UK

Accepted 10 January 2007
Available online 14 January 2007

Abstract

We investigate the possible use of families of momentum-space descriptors and of trivial classical descriptors for the prediction of blood–brain barrier penetration, expressed as log *BB*. A 12-descriptor model based on entropy-like momentum-space quantities and on the numbers of atoms of each type has good statistical quality for a set of 42 structurally diverse molecules. We also consider the inclusion in our models of some of the other descriptors that have been used in earlier models for these molecules. The resulting models are not expected to be useful *as-is* for making genuine predictions for much larger test sets, but the various results do demonstrate the potential benefits of incorporating momentum-space descriptors into QSAR models for predicting log *BB*.

Keywords: Blood–brain barrier; Momentum-space; Molecular descriptors; QSAR

1. Introduction

An important consideration in the development of central nervous system active drugs is the degree of penetration of the blood–brain barrier. As such, it can be highly desirable to know values of log *BB*, in which *BB* is the ratio of the steady-state concentrations of the relevant compound in the brain and in the blood. Values of log *BB* tend to lie in the range from −2 to +1; compounds with log *BB* more negative than −1 are poorly distributed to the brain whereas those with positive values greater than 0.3 cross the blood–brain barrier fairly easily. Of course, the determination of values of log *BB* is just one rather small part of the drug discovery and design process, which also requires a range of absorption, distribution, metabolism and excretion (ADME) data and also toxicity data. Our emphasis here is on log *BB*.

The experimental determination of log *BB* values is not a simple task, involving as it does the direct measurements of the drug concentrations in the brain and in the blood of laboratory animals. Not surprisingly, there has been a great deal of effort to establish QSAR models based on experimentally determined and experiment-free molecular descriptors. The literature in this area is of course very extensive, but we restrict ourselves here to a brief discussion of previous work that is of particular relevance to the present study, in which we consider the possible use of momentum-space descriptors. It should be stressed from the outset that this is an exploratory study: we do not expect the particular QSAR models that we present here to be useful *as-is* for making genuine predictions for large test sets. Instead, the main aim is discover whether it could in principle be useful to include descriptors such as ours, alongside more traditional ones, in QSAR models for predicting log *BB* values.

Young et al. [1] considered QSAR models for central histamine H_2 receptors and found, for 20 molecules, that log *BB* does not correlate very well with either the octanol–water partition coefficient, log *P*, or the corresponding cyclohexane–water partition coefficient. Of greater utility was the difference between these two partition coefficients, Δ(log *P*), which is mainly a measure of the solute hydrogen-bond acidity. Their best correlation for the 20 molecules is characterized by a correlation coefficient (*R*) of 0.831, a standard deviation (S.D.) of 0.439 and a Fisher-*F* statistic (*F*) of 40. The same group of

* Corresponding author. Tel.: +44 794 3532; fax: +44 794 3588.
E-mail address: dlc@liv.ac.uk (D.L. Cooper).
[1] Current address: Umm Al-Qura University, Saudi Arabia.

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.01.002

20 molecules was also considered by Waterbeemd and Kansy [2] using as their descriptors the molecular volume and a quantity Λ_{alk} that is meant to characterize the total H-bonding capacity. The required values of Λ_{alk} were calculated from experimental log P values and from calculated molar volumes. The resulting correlation is characterized by $n = 20$, $R = 0.934$, S.D. = 0.290 and $F = 58$. Less satisfactory results were obtained when, instead of Λ_{alk}, they used values for the surface area of the hydrophilic part of the van der Waals surface.

Abraham et al. [3] expressed concern about the relatively small number of molecules (20) in the set examined by Young et al. [1] and subsequently by Waterbeemd and Kansy [2], and considered instead an extended set of $n = 65$ molecules. They used a regression model based on the McGowan volume (V_x) [4], an excess molar refraction (R_2), a solute dipolarity/polarizability parameter (π_2^H), and the solute hydrogen-bond acidity and basicity (α_2^H and β_2^H, respectively). For 57 molecules (8 were removed as outliers) they achieved $R = 0.952$, S.D. = 0.197 and $F = 99$. However, a close inspection of the descriptor values shows that a number of them are strongly correlated (R_2 with π_2^H, R_2 with β_2^H, R_2 with V_x, π_2^H with β_2^H, and α_2^H with V_x) [5]. Sun [6] has used these 57 molecules in tests of a large generic set of 'universal' atom-type descriptors; a three-component partial least-squares (PLS) model based on these descriptors gave $R^2 = 0.897$ and a root mean square error (RMSE) of 0.259.

For a set of 57 molecules, with log *BB* data taken from Young et al. [1] and from Abraham et al. [3], Lombardo et al. [7] used as a single descriptor the computed solvation free energy in water (ΔG_w^0). After the removal of two outliers, they found $n = 55$, $R = 0.82$, S.D. = 0.41 and $F = 108$. The reliable computation of ΔG_w^0 is of course relatively expensive. For the same group of 57 molecules, Clark [5] explored the use instead of the polar surface area and of calculated log P values. After removing two compounds, his best correlation was characterized by $n = 55$, $R = 0.887$, S.D. = 0.354 and $F = 96$.

Luco [8] modeled log *BB* for 61 structurally diverse compounds, mostly taken from Young et al. [1] and Abraham et al. [3], but with the addition of a small number of acidic compounds. Starting with as many as 25 molecular descriptors, including several topological and constitutional descriptors, he obtained a three-component PLS model characterized by $n = 58$, $R = 0.922$, S.D. = 0.318 and $F = 102$ (three molecules were excluded as outliers). Feher et al. [9] examined the same group of 61 molecules, using as descriptors just the calculated log P values, the number of hydrogen-bond acceptors in an aqueous medium and the polar surface area, obtaining $n = 61$, $R^2 = 0.730$, RMSE = 0.424, and $F = 51$. They found for the full set of 61 molecules that their results are only slightly inferior to those of Luco [8].

Given that it has been shown that certain momentum-space quantities can be useful as molecular descriptors in QSAR (and QSPR) studies [10–12], it seems very worthwhile to investigate their possible utility for the challenging problem of predicting blood–brain barrier penetration. We examine here the log *BB* data for a set of 45 structurally diverse molecules. As in some of our previous work, we consider three families of descriptors, as well as hybrid models based on combinations of these.

2. Methodology

We selected molecules for this study as follows. Our starting point was the training set of Clark [5], which consists of 27 named compounds and a group of larger molecules, which he labeled 1–30. Except for N_2, Luco's training set [8] includes all of these named molecules (and with the same log *BB* values). For our own study, we selected the 27 named compounds considered by Clark [5] (as listed in his Table 1) and augmented these with four acidic compounds considered by Luco [8] (specifically molecules listed in his Table 1 as **57**–**60**). We also returned to the work of Abraham et al. [3], from which many of these subsequent sets have in part been derived, and noticed in their Table 8 that there are three small/symmetrical molecules (specifically CS_2, NO and SF_6) that have not been included by Clark [5] or Luco [8]. We chose to include the data for those three molecules in the present study.

Our merged set of 34 molecules (27 named molecules from Clark [5], four acidic molecules from Luco [8], and three small/symmetrical molecules from the original work of Abraham et al. [3]) includes only three systems that have log *BB* values which are more negative than −0.31. In order to remedy this obvious deficiency, we selected 11 species from the first 20 of the numbered molecules in Clark's training set (as listed in his Table 1) [5]. The selection process was fairly arbitrary, except that we took care to include a range of structural features and to pick a number of molecules that have significantly negative log *BB* values. We note that all 11 of these molecules were also included in Luco's training set [8], and with the same log *BB* values. The final set of molecules selected for the present study consists of 45 structurally diverse species (see Table 1 and Scheme 1).

For each of our selected molecules, we performed AM1 geometry optimizations and then calculated the momentum-space (p-space) total electron density, $\rho(\mathbf{p})$, from the Fourier transform of the resulting wave function. One of the families of p-space descriptors that we consider here ('α') is based on the moments of momentum, as defined in the following equation:

$$\langle p^n \rangle = \int p^n \rho(\mathbf{p})\, d\mathbf{p} \tag{1}$$

Typically, we consider n values of −2, 0 and 2. In practice, we calculate $\langle p^{-2} \rangle$ and, instead of $\langle p^0 \rangle$ and $\langle p^2 \rangle$, we use the total number of electrons treated in the AM1 calculation (N) and the magnitude of the total energy, $|E|$. We also include in 'α' the molecular weight. This particular family of descriptors has previously proved useful in QSAR/QSPR models for various quantities, including gas-chromatography retention times, gas–hexadecane partition coefficients and tadpole narcosis concentrations [12].

The second family of p-space descriptors ('β') consists of entropy-like quantities, as defined in Eq. (2), in which

Table 1
Observed and predicted values of log *BB*

Molecule	Observed	12-Descriptor model	12-Descriptor model and ΔG^0_w
2,2-Dimethyl butane	1.04	0.91	0.94
3-Methyl pentane	1.01	0.84	0.88
2-Methyl pentane	0.97	0.82	0.87
Methyl cyclopentane	0.93	0.89	0.93
3-Methyl hexane	0.90	0.89	0.91
Heptane	0.81	0.83	0.85
Hexane	0.80	0.78	0.83
Pentane	0.76	0.70	0.77
Carbon disulfide	0.60	0.63	
Isoflurane	0.42	0.39	0.36
1,1,1-Trichloroethane	0.40	0.51	0.47
Benzene	0.37	0.47	0.47
Toluene	0.37	0.65	0.62
Sulfur hexafluoride	0.36	0.48	
Halothane	0.35	0.19	0.24
Trichloroethene	0.34	0.39	0.36
Teflurane	0.27	0.37	0.38
Enflurane	0.24	0.36	0.34
Fluroxene	0.13	−0.19	−0.05
1,1,1-Trifluoro-2 -chloroethane	0.08	0.03	0.16
Methane	0.04	−0.16	−0.07
Nitrous oxide	0.03	−0.11	
Nitrogen	0.03	0.02	−0.05
Diethyl ether	0.00	−0.08	0.07
Butanone	−0.08	0.02	−0.01
2-Propanol	−0.15	0.00	−0.09
Propanone	−0.15	−0.03	−0.10
Ethanol	−0.16	−0.07	−0.20
1-Propanol	−0.16	−0.05	−0.17
2-Methyl propanol	−0.17	0.05	−0.07
Valproic acid	−0.22	−0.06	
p-Acetamidophenol	−0.31	−0.45	
Acetyl salicylic acid	−0.50	−0.59	
Salicylic acid	−1.10	Outlier	
A	−0.04	−0.24	−0.21
B	−0.18	−0.09	−0.34
C	−1.15	Outlier	Outlier
D	−1.17	−0.88	−1.08
E	−0.66	−0.87	−0.68
F	−0.67	−0.61	−0.67
G	−1.42	Outlier	Outlier
H	−1.23	−1.17	−1.22
I	−0.82	−0.84	−0.63
J	−1.12	−1.07	−1.18
K	0.11	−0.07	−0.03

The different regression models are identified in the text, as are the sources of the experimental data.

$\sigma(\mathbf{p}) = \rho(\mathbf{p})/N$ is sometimes termed the 'shape function'.

$$S_n = -\int p^n \rho(\mathbf{p}) \ln \rho(\mathbf{p})\, \mathrm{d}\mathbf{p},$$

$$S'_n = -\int p^n \sigma(\mathbf{p}) \ln \sigma(\mathbf{p})\, \mathrm{d}\mathbf{p} \quad (2)$$

This family of descriptors (with n values of −2, 0 and 2) has previously proved useful in a study of log P [11]. The two sets of entropy-like descriptors are of course trivially related to one another, as is shown in the following equation:

$$S'_n = \frac{1}{N}(\langle p^n \rangle \ln N + S_n) \quad (3)$$

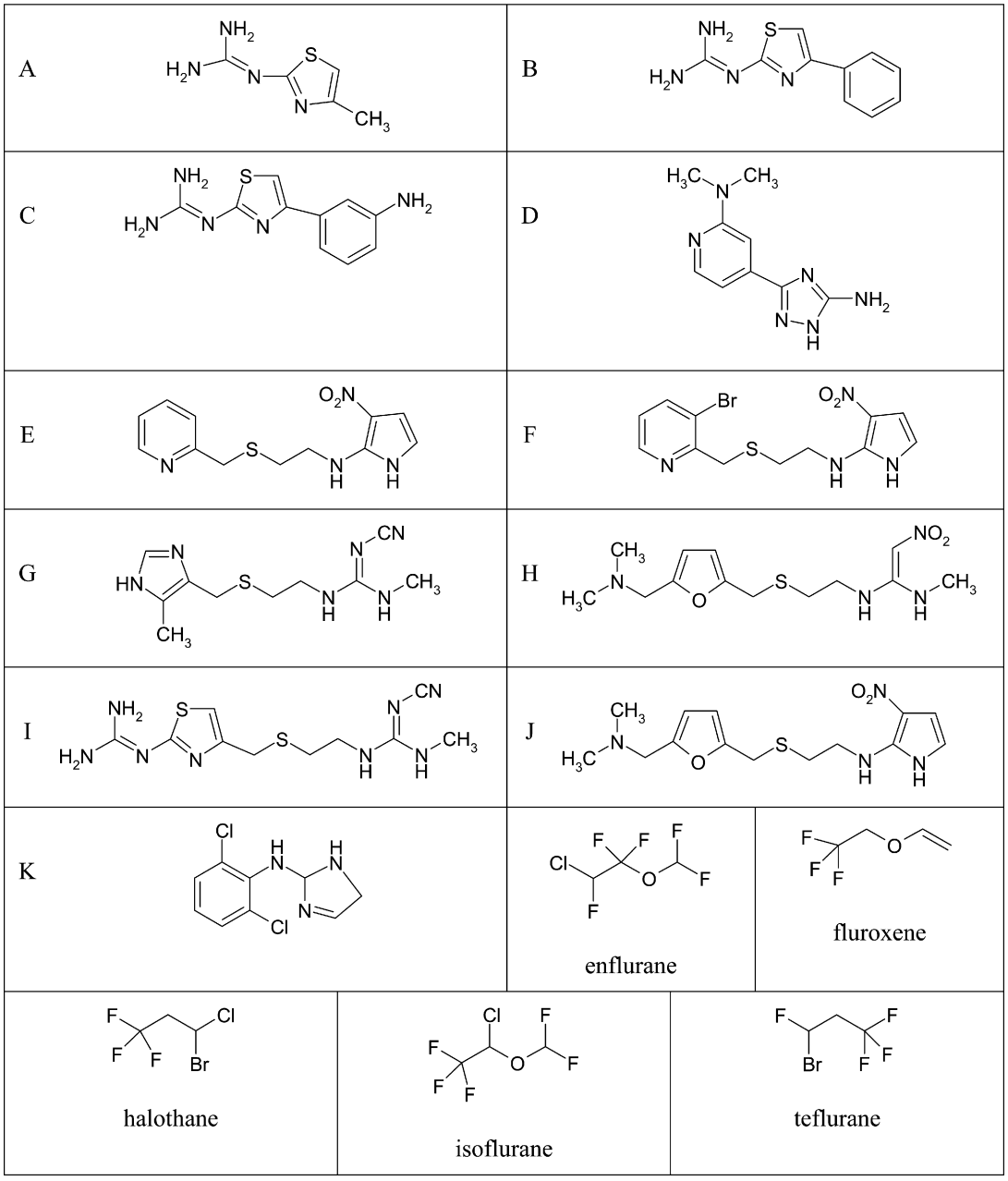
The third family of descriptors ('γ') consists of entirely trivial structural quantities, specifically the numbers of atoms of each type (n_X) and the number of bonds (n_{bond}), in the sense that this term is defined when calculating McGowan volumes [4]. We found in our previous work [12] that this family of descriptors can produce rather good results for a range of quantities, as characterized by (adjusted) R^2 and S.D. values, even if the somewhat increased number of descriptors leads to a poorer F statistic.

Multiple linear regression (MLR) models were constructed in the present work using SPSS, almost exclusively employing the 'simultaneous method', which SPSS calls the 'enter method' [13]. One simply specifies the set of descriptors that make up the chosen model, except that SPPS will reject descriptors that are too strongly correlated with the others. One of the most useful measures of the success of an MLR model is of course the adjusted R^2 value, which takes account of the number of variables and of the number of observations, and is far more informative than are raw correlation coefficients or coefficients of determination (whether quoted as R^2 or R). We do of course also quote values of the standard deviation and of the 'analysis of variance' or Fisher-F statistic.

As will be discussed later, two models were also constructed using instead the so-called 'stepwise method', in which SPSS considers each of the chosen descriptors in turn. If the descriptor contributes to the success of the model then it is retained, but all of the other included descriptors are then reassessed, to determine whether they are successfully contributing to the model and thus should be retained. This approach is recommended by statisticians not least because it tends to result in small sets of predictor variables. However, it is well known that this procedure does not always guarantee the best QSAR model [14].

For a particular descriptor to be considered a useful predictor in the final MLR model, conventional guidelines are that its t value, *i.e.* the coefficient for that descriptor divided by its standard error, should lie outside the range −2 to +2. Such t values are also of use in determining the strongest and weakest descriptors. Statistical outliers in the models were identified in the present work as those molecules with absolute standardized residuals greater than 2, as listed in the table of Casewise Diagnostics in SPSS [13]. After eliminating outliers, and rerunning the regression analysis, there is no table of Casewise Diagnostics when the standardized residuals for all of the remaining systems lie inside the acceptable range.

One way of assessing the statistical significance of a given MLR correlation would be to fix the order of the molecules, and the values of the corresponding molecular descriptors, while randomizing the order of the log *BB* values. In general, a new model based on these 'wrong' data will have a poorer correlation coefficient. After repeating this randomization test many times, we could determine the probability that we can obtain a correlation coefficient greater than or equal to the



Scheme 1. Compounds A–K and some of the 'less familiar' molecules listed in Table 1.

'true' one. As is well known, the numerical value of the probability p_F that would be determined by following such a procedure corresponds to a standard statistical measure known as the significance of the F statistic. Models are typically considered statistically significant if p_F is less than 0.05, but we will insist here on achieving very much lower values.

3. Results

As measured only by adjusted R^2 and S.D., we found for the 45 molecules listed in Table 1 that the best single family of descriptors is 'γ', giving adjusted $R^2 = 0.748$ ($R^2 = 0.800$) and S.D. = 0.329 [14]. Taking account also of the F statistic, the best results are in fact for descriptor family 'α'. Looking at combinations of the families of descriptors, the most promising combinations are 'αβ' and 'βγ', with the first of these having both the best adjusted R^2 (0.804) and the best Fisher statistic ($F = 21$). We found that SPSS rejected attempts to retain S_{-2} in the 'αβ' model and also excluded all of $\langle p^{-2} \rangle$, $|E|$, molecular weight and n_C from the 'αβγ' model, which turned out to have much the same statistical quality as did the 'βγ' model [14].

In the light of our results for the 'αβγ' model, we returned to the 'βγ' families of descriptors. All of our attempts to include also $\langle p^{-2} \rangle$ and/or $|E|$ and/or N and/or ln N proved unsuccessful, because of strong linear dependence between various descriptors. Further investigations (including the construction of MLR models in which one or more descriptors is excluded) revealed that n_{bond}, S_0 and S'_0 contribute very little indeed to the quality

of the 'βγ' model, and so we chose to reject them. In this way, we arrived at our preferred 12-descriptor MLR model.

Our 12-descriptor model is based on $\{n_H, n_C, n_N, n_O, n_F, n_{Cl}, n_{Br}, n_S\}$ and $\{S_{-2}, S'_{-2}, S_2, S'_2\}$ and it is characterized by $n = 42$, $R^2 = 0.943$ (adjusted R^2 is 0.920), S.D. = 0.165 and $F = 40$. Based on their t values, all 12 descriptors satisfy conventional guidelines for statistical importance. The significance value of the F statistic is $p_F = 8.30 \times 10^{-15}$, indicating that the success of this model in reproducing the variation in the data is rather unlikely to be due to chance. Three of the original 45 molecules (salicylic acid and compounds C and G) have absolute standardized residuals greater than 2 and were thus treated as outliers, even though we have no explanation at the present time for their apparently aberrant behavior. The predictions of our 12-descriptor model for the log *BB* values of the remaining 42 molecules are listed in Table 1 and they are displayed in Fig. 1(a). Further numerical details for all of the models constructed in this study are available from ref. [14].

One of the most obvious criticisms of our 12-descriptor model is that the ratio of molecules to descriptors (often called the Topliss ratio [15]) is relatively small (3.5), whereas one normally aims for a value of at least 5. On the other hand, the model is characterized by a very low value of p_F.

Using somewhat fewer descriptors from families 'α', 'β' and 'γ' for this particular set of molecules does of course tend to lead to poorer adjusted R^2 and S.D. values, but better Fisher statistics. Using the 'stepwise method' in SPSS for the reduced set of 42 molecules, we find that the preferred regression model is based on just three descriptors, namely n_N, n_O and $\langle p^{-2} \rangle$. This model is characterized by $R^2 = 0.822$ (adjusted R^2 is 0.808), S.D. = 0.257, $F = 58$ and $p_F = 2.68 \times 10^{-14}$. Using instead the 'stepwise method' for the full set of 45 molecules, the preferred regression model is again based on three descriptors but, this time, they are n_N, n_O and S_{-2}. The apparent importance of n_N and n_O seems likely to be a feature of this particular set of molecules, but we find it very encouraging that momentum-space descriptors such as $\langle p^{-2} \rangle$ or S_{-2} are included in these 3-descriptor models.

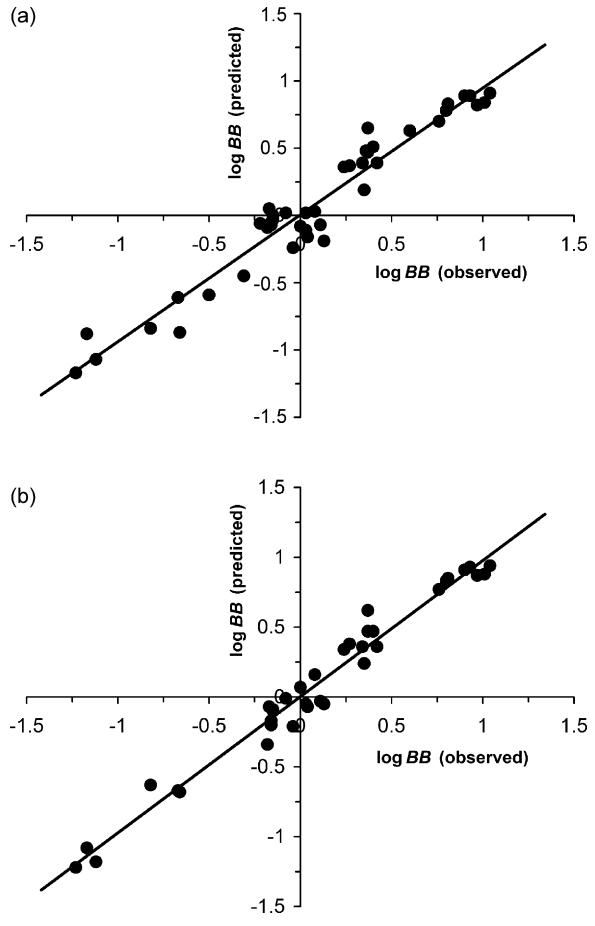


Fig. 1. Correlation between predicted and observed log *BB* values: (a) 12-descriptor model; (b) 12-descriptor model and ΔG^0_w.

We could in principle have carried out further statistical tests of our various sets of descriptors, such as the 'leave one out' and related procedures that are popular in QSAR studies. We could also have tried to use our models for test sets composed of molecules outside our chosen set of 45 molecules, but none of this seems at all likely to modify our key inference that some of our momentum-space descriptors do indeed appear to be useful, alongside simple feature counts, for log *BB*.

On the other hand, it does seem worthwhile to investigate also the performance of our 12-descriptor model when it is augmented with some of the quantities that have been utilized in previous studies of these molecules. Returning to the full set of 45 molecules, we find that values of ΔG^0_w [7], polar surface area [5] and calculated log P [5] are available for 38 of them. Attempting to use our preferred set of 12-descriptors for this set of 38 molecules (see also Table 1), we find that molecules C and G have standardized residuals of −2.76 and −2.53, respectively, and so they have again been treated as outliers. For the remaining $n = 36$ molecules, we find $R^2 = 0.955$ (adjusted $R^2 = 0.932$), S.D. = 0.159, $F = 41$ and $p_F = 1.22 \times 10^{-12}$. Subsequently including ΔG^0_w as an additional descriptor does indeed lead to improved statistics for this set of molecules: $n = 36$, $R^2 = 0.973$ (adjusted R^2 is 0.956), S.D. = 0.128, $F = 60$ and $p_F = 4.35 \times 10^{-14}$. The resulting predictions of log *BB* are listed in Table 1 and they are displayed in Fig. 1(b). Given the costs associated with reliable calculations of ΔG^0_w, we were very pleased to find that using instead the calculated values of log P gives results that are statistically only slightly inferior. The same was not true when using instead the polar surface area [14].

Of course, even with such low p_F values, we should not take 13-descriptor models for relatively few molecules too seriously unless they can be shown also to be applicable to significantly larger set of molecules. Nevertheless, we believe that the statistical results for the various models that we have described here, ranging from 3 to 13 descriptors, have demonstrated that it could indeed be worthwhile to carry out further QSAR studies on much larger sets of molecules using a combination of simple feature counts, calculated log P values and momentum-space quantities of the types we have examined.

4. Conclusions

We have shown that a combination of entropy-like momentum-space descriptors and trivial classical descriptors (the numbers of atoms of each type) leads to a 12-descriptor

model, of good statistical quality, for the log *BB* values of 42 structurally diverse molecules. We have also found that it can be useful to combine our chosen set of descriptors with certain others that have been successfully used for these molecules, such as the computed solvation free energy in water or calculated octanol–water partition coefficients. Of course, we do not expect that the particular models described here will prove to be equally successful *as-is* for wide ranges of molecules outside the relatively small 'training sets', but we do believe that this exploratory study has demonstrated the potential benefits of incorporating momentum-space descriptors such as ours into QSAR models for predicting blood–brain barrier penetration (expressed as log *BB*), when used alongside more traditional descriptors such as calculated log *P* values and simple feature counts.

References

[1] R.C. Young, R.C. Mitchell, T.H. Brown, C.R. Ganellin, R. Griffiths, M. Jones, K.K. Rana, D. Saunders, I.R. Smith, N.E. Sore, T.J. Wilks, Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H_2 receptor histamine antagonists, J. Med. Chem. 31 (1988) 656–671.

[2] H. Waterbeemd, M. Kansy, Hydrogen-bonding capacity and brain penetration, Chimia 46 (1992) 299–303.

[3] M.H. Abraham, H.S. Chadha, R.C. Mitchell, Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain, J. Pharm. Sci. 83 (1994) 1257–1268.

[4] A. Mellors, J.C. McGowan, Uses of molecular volume in biochemical pharmacology, Biochem. Pharm. 34 (1985) 2413–2416.

[5] D.E. Clark, Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood–brain barrier penetration, J. Pharm. Sci. 88 (1999) 815–821.

[6] H. Sun, A universal molecular descriptor system for prediction of log *P*, log *S*, log *BB*, and absorption, J. Chem. Inf. Comput. Sci. 44 (2004) 748–757.

[7] F. Lombardo, J.F. Blake, W.J. Curatolo, Computation of brain–blood partitioning of organic solutes via free energy calculations, J. Med. Chem. 39 (1996) 4750–4755.

[8] J.M. Luco, Prediction of the brain–blood distribution of a large set of drugs from structurally derived descriptors using partial least squares (PLS) modeling, J. Chem. Inf. Comput. Sci. 39 (1999) 396–404.

[9] M. Feher, E. Sourial, J.M. Schmidt, A simple model for the prediction of blood–brain partitioning, Int. J. Pharm. 201 (2000) 239–247.

[10] (a) E.F. McCoy, M.J. Sykes, The estimation of molecular properties using momentum-space wavefunctions, Chem. Phys. Lett. 313 (1999) 707–712;
(b) M.J. Sykes, Estimation of Molecular Properties with Momentum-space Wavefunctions, PhD Thesis, The Flinders University of South Australia, 2000.
(c) E.F. McCoy, M.J. Sykes, Quantum-mechanical QSAR/QSPR descriptors from momentum-space wave functions, J. Chem. Inf. Comput. Sci. 43 (2003) 545–553.

[11] J.H. Al-Fahemi, D.L. Cooper, N.L. Allan, The use of momentum-space descriptors for predicting octanol–water partition coefficients, J. Mol. Struct. (THEOCHEM) 727 (2005) 57–61.

[12] J.H. Al-Fahemi, D.L. Cooper, N.L. Allan, The quantitative use of momentum-space descriptors, Chem. Phys. Lett. 416 (2005) 376–380.

[13] SPSS Base 10.0 Applications Guide, SPSS Inc., Chicago IL, USA, 1999

[14] J.H.A. Al-Fahemi, Momentum-space Descriptors for QSPR and QSAR Studies, PhD Thesis, Liverpool University, UK, 2006.

[15] J.G. Topliss, R.P. Edwards, Chance factors in studies of quantitative structure–activity relationships, J. Med. Chem. 22 (1979) 1238–1244.