How accurate are continuum solvation models for drug-like molecules?

Kongsted, Jacob; Söderhjelm, Pär; Ryde, Ulf

Published in:
Journal of Computer-Aided Molecular Design

DOI:
10.1007/s10822-009-9271-6

2009

Document Version:
Peer reviewed version (aka post-print)

How accurate are continuum solvation models for drug-like molecules?

Jacob Kongsted [1], Pär Söderhjelm [2], & Ulf Ryde [2]*

[1] Department of Physics and Chemistry, University of Southern Denmark,

Campusvej 55, DK-5230, Odense M, Denmark

[2] Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124,

SE-221 00 Lund, Sweden

Correspondence to Ulf Ryde, E-mail: Ulf.Ryde@teokem.lu.se,

Tel.: +46 – 46 2224502, Fax: +46 – 46 2224543

2017-03-18

Abstract

We have estimated the hydration free energy for 20 neutral drug-like molecules, as well as for three series of 6–11 inhibitors to avidin, factor Xa, and galectin-3 with four different continuum solvent approaches (the polarised continuum method (PCM) the Langevin dipole method, the finite-difference solution of the Poisson equation, and the generalised Born method), and several variants of each, giving in total 24 different methods. All four types of methods have been thoroughly calibrated for a number of experimentally known small organic molecules with a mean absolute deviation (MAD) of 1–6 kJ/mol for neutral molecules and 4–30 kJ/mol for ions. However, for the drug-like molecules, the accuracy seems to be appreciably worse. The reason for this is that drug-like molecules are more polar than small organic molecules and that the uncertainty of the methods is proportional to size of the solvation energy. Therefore, the accuracy of continuum solvation methods should be discussed in relative, rather than absolute, terms. In fact, the mean unsigned relative deviations of the best solvation methods, 0.09 for neutral and 0.05 for ionic molecules, correspond to 2–20 kJ/mol absolute error for the drug-like molecules in this investigation, or 2–3000 in terms of binding constants. Fortunately, the accuracy of all methods can be improved if only relative energies within a series of inhibitors are considered, especially if all of them have the same net charge. Then, all except two methods give MADs of 2–5 kJ/mol (corresponding to an uncertainty of a factor of 2–7 in the binding constant). Interestingly, the generalised Born methods typically give better results than the Poison–Boltzmann methods.

Key Words: solvation energy, free energy of hydration, continuum methods, PCM, Langevin dipoles, FDPB, generalised Born.

Introduction

The great majority of chemical reactions on earth take place in aqueous solution. Therefore, it is of highest interest to know how this solvent tunes the reaction rates. Likewise, in biochemistry, it is crucial to know how different reactions proceed in water solution or inside an enzyme. This is one of the main goals of theoretical modelling. Naturally, solvent effects can be estimated by including explicit water molecules in the calculations. However, solvation effects are long-range, meaning that a very large number of water molecules need to be considered, making the calculations very expensive, especially if quantum mechanical (QM) methods are used. Moreover, the number of possible configurations of the system increases rapidly with the number of molecules, which makes it hard to obtain accurate and converged estimates of dynamic and entropic effects. Therefore, there has been a great interest in developing methods in which the solvent is treated as a structureless dielectric continuum [1,2,3,4].

There are several types of such continuum solvent methods. One is the polarised continuum method (PCM), which is an extension of the simple Born and Onsager [5,6] models for general charge distributions and molecular shapes [7,8,9]. In this method, the solute is assigned a dielectric constant (typically $\varepsilon = 1$) and the solvent another dielectric constant). The interface of the two volumes is described by a surface formed by small area elements (tesserae) and the reaction field from the solvent is then described as a charge on each area element. The PCM method has especially been popular in combination with QM calculations and many variants are now available [7,8,10], with the integral-equation formalism (IEFPCM) [11,12] and the COSMO models [13] currently being most widely used. However there are also implementations that can be used at the molecular mechanics (MM) level and for whole proteins [14,15].

In the protein-modelling community, methods based on the solution to the Poisson–Boltzmann (PB) equation have been more popular [16,17,18,19,20,21,22]. Again, the solute and solvent are assigned different dielectric constants, but the solute is described by a set of atomic charges. The method can also be used in QM calculations or with more complicated charge distributions [23,24,25,26,27].

The generalised Born (GB) methods can either be seen as an approximation to the PB method or as a generalisation of the Coulomb equation [1,3,4,28,29]. It approximates the solvation energy by a pair-potential between each charge, employing a specific screening function between each pair, which depends on the geometry of the solute. It has received much interest lately, both from the MM [30,31,32,33] and QM communities [34,35,36].

A conceptually more different method is the Langevin-dipole (LD) method, extensively used by Warshel and coworkers [37,38,39], but also by other groups [40], both at the QM and MM levels. In this method, the solvent is described by an array of explicit dipoles (and possibly polarisabilities), which is affected by the electrostatic field of the solute and all the other dipoles according to the Langevin equation [41]. Thus, in variance to the other three methods, this method is semi-discrete, with explicit solvent molecules that are described in a simplified way and that are fixed in space on a grid. Long-range solvation effects are estimated by the Born or Onsager models outside a certain radius. Moreover, the LD method models saturation effect of the solvent, which may be important for multiply-charged ions [42]. Unfortunately, the LD method has been shown to diverge if realistic values of the water dipoles (and polarisabilities) are used [37], so reduced values have to be used, compensated by other means. Several other methods have also been suggested [1,2,8,9,42], but they have not been equally widely used.

All methods described up to now only give the electrostatic (polar) part of the solvation energy. However, there are other contributions to the solvation energy, viz. the cost of forming a cavity in the solvent, the unfavourable exchange repulsion between the solvent and the solute, and the favourable dispersion energy between the solute and the solvent [1,8,9]. There are many ways to take these effects into account ranging from a single term that is

proportional to the solvent-accessible surface area (SASA) of the solute [43] to more sophisticated expressions for all three terms [44,45,46]. Traditionally, PB and GB methods have used only a SASA term, the QM PCM methods have been combined by the three-term expression, and the LD method has used an intermediate two-term expression [38,39], but this has started to change [47,48,49,50,51,52].

All types of continuum methods include several parameters. In particular, they all need a set of atomic radii to define the extent of the solute, both for the electrostatic and non-polar terms (the self-consistent isodensity PCM method [53] is a prominent exception to this, but it shows convergence problems and typically gives worse results than other PCM methods). The results of the solvation methods strongly depend on these radii and therefore a thorough calibration is needed. All types of continuum solvation methods have been calibrated in this way, typically using a set of experimental hydration free energies for both neutral and ionic small molecules [20,38,39,42,47,48,54], although some have been calibrated towards explicit solvent simulations instead [48,55]. If the parametrisation is thorough, mean absolute deviations (MADs) of 1–2 kJ/mol for neutral molecules have been reported [10,36,51,54,56], i.e. slightly larger than the typical experimental uncertainty (~0.8 kJ/mol) [57,58,59]. For ionic molecules, the results are appreciably worse, with MADs of ~20 kJ/mol, but the experimental uncertainty is also of a similar size [10,35,39,58].

Unfortunately, experimental solvation data is missing for most larger molecules, in particular for most drug-like molecules. Therefore, continuum solvation methods often play a prominent role in theoretical estimates of ligand-binding affinities. For example, one of the terms in the MM/PBSA and similar approaches [60,61] is the difference in solvation free energy between the bound and unbound ligand, estimated by the PB or GB methods, and combined with a non-polar SASA term.

A natural question is whether the accuracy of these estimates is similar to what is found for the calibration set of molecules. In fact, this is necessary, because ligand-binding methods typically aims at an accuracy of a few kJ/mol (6 kJ/mol corresponds to one order of magnitude in binding affinity). As a first approximation, it can be assumed that this term is dominated by the solvation energy of the free ligand (at least for a hidden binding site), because the solvation energy of the protein will mainly cancel between the free and complexed protein, especially as the geometry is typically not changed [60]. Even if we do not know the experimental solvation energy of the ligands, we can obtain an estimate of the accuracy of the various continuum solvation methods by calculating the free energy of hydration with several well-calibrated methods. If all methods give similar results, we can be quite confident that the results are correct. Otherwise, larger errors can be expected for drug-like molecules.

In this paper, we compare the predictions of the four major types of continuum solvation methods (PCM, LD, PB, and GB) and several variants of each for the hydration free energy of 20 neutral drug-like molecules and three series of inhibitors to avidin (7 biotin analogues), factor Xa (11 amidinobenzyl-indole-carboxamide inhibitors) and galectin-3 (6 lactose derivatives). We show that the absolute errors for the drug-like molecules are much larger than for small organic molecules. However, the results are much improved if only relative energies within each inhibitor series are considered.

Methods

Molecules

Five sets of molecules were used. First, we used a set of 40 small organic molecules for which the hydration free energies are known experimentally (20 neutral, 10 anions, and 10 cations). The molecules were primarily selected to include models of the peptide backbone and all amino-acid side-chains. This set was then enhanced with two neutral molecules with

fluorine and chlorine, because these elements are common in drug-like molecules, and the neutral molecules were complemented with ionic analogues, for which experimental data is available. This test set of molecules is listed in Table S1 (in the supporting information) together with experimental data [20,36,38,62,63]. Experimental data from other sources [36,39,42,47,52,54,57,64,65] for the neutral molecules differ by up to 0.55 kJ/mol, which reflects the experimental uncertainty. On the other hand, the values for the ions differ by up to 36 kJ/mol (or even 74 kJ/mol for an outlier for the $HCOO^-$ ion [47]), which reflects both the experimental uncertainty and the problem of fixing the scale of absolute solvation energies. We have used the recent values of Truhlar, et al. [63] (for imidazoleH^+, which is not included in that study, we used the value by Florian and Warshel [38], but corrected to use the same value of the absolute solvation energy of the proton). The prime use of this set was to check that we use the methods properly (all details of the calculations are not always presented and software tend to be updated), but also to compare variants of each type of method.

The second set of molecules was a diverse set of 20 neutral drug-like molecules, differing in size, shape, composition, functional groups, etc. (by neutral, we mean that the molecules were studied in the uncharged form and that we avoided obviously charged groups like carboxylate groups or aliphatic amines; however, we did no deeper investigation of the pK_a values of the molecules). For example, the number of atoms varies from 20 to 113. They are shown in Figure S1 and shortly described in Table S2.

The third set of molecules was seven biotin analogues, which has been used in several MM/PBSA studies of their binding to avidin [66,67,68]. They are shown in Figure S2. Btn1–Btn3 have a net negative charge (a carboxylate group), whereas the other four are neutral. They contain between 12 and 41 atoms. Btn4 was also included in set 2.

The fourth set is eleven amidinobenzyl-indole-carboxamide inhibitors of factor Xa (FXa) [69]. They are shown in Figure S3 and have been used in a recent MM/PBSA study [68]. They are larger than the biotin analogues, with 47–61 atoms. FXa39, FXa57, FXa63, FXa103, and FXa127 have a single positive charge, whereas the other six have a double positive charge.

The fifth set is lactose and five 3'-benzamido-N-acetyllactosamine galectin-3 inhibitors [70]. They are shown in Figure S4. The latter five only differs in the substituents of the benzene ring, being either H, F, or $–OCH_3$. They contain 45–71 atoms and are all neutral. Gal3 was also included in set 2.

In all calculations, both on the small molecules as well as the drug-like molecules, we restrict the calculation of the solvation energy to one molecular conformation. This might be a problem when comparing with experimental data for flexible molecules [71], but not for the drug-like molecules, for which no such comparison is made. The conformations used were either taken from crystal structures of proteins with the ligand of interest [69,70,72,73] or from the most stable conformation, according to a systematic conformational search at the MMFF94 level with the Spartan software [74].

PCM calculations

Five different sets of PCM calculations were performed. First, we tried to reproduce the original calibration of the UAHF radii (united atom for Hartree–Fock) [54] as closely as possible. Thus, the molecules were optimised at the Hartree–Fock (HF) level with the 6-31(+)G* basis set (i.e. 6-31G* for cations and neutral molecules, but 6-31+G* for the anions) [75]. Then, the solvation energy was calculated at the same level of theory using the DCPM method [7] and the ICOMP = 4 option [76]. The calculations employed the default UAHF radii [54] and parameters for water solvent (implying a dielectric constant of 78.39 and a solvent probe radius of 1.385 Å), as implemented in the Gaussian-98 software [77]. These calculations are called DPCM/HF below. Reported solvation energies are the sum of the electrostatic, cavitation, dispersion, and repulsion energy components.

Unfortunately, the PCM implementation in Gaussian-98 is quite unstable and the calculations failed for many of the larger molecules, viz. 061, cannabidiol, paclitaxel, progesterone, SB3, all factor Xa inhibitors, except FXa127, and all galectin-3 inhibitors, except lactose. Therefore, we performed similar calculations with the default solvation method in Gaussian-03 [78], IEFPCM (integral equation formalism PCM) [11,12]. All other methods, basis set, radii, etc. were the same as for DPCM. This is called IEFPCM/HF. However, a comparison showed that the radius of the C atom in CH_4 has been altered between the two versions of Gaussian, from 2.13 Å in Gaussian-98 to 2.04 Å in Gaussian-03. The presented results use these different radii, even if it gives a quite large error in Gaussian-03 (3.6 kJ/mol, compared to 0.6 kJ/mol, for IEFPCM/UAHF in Gaussian-03 with the two different radii).

For most of the molecules that failed at DPCM/HF level, the problem was related to the non-polar terms, whereas the electrostatic solvation energy could be calculated. To get DPCM/HF estimates also for these molecules, we complemented the electrostatic energies with the non-polar terms from the IEFPCM/HF calculations. Then, only paclitaxel, FXa39, FXa63, and the galectin-3 inhibitors failed with DPCM/HF due to convergence problems.

Next, we also tested the UAKS radii in Gaussian-03 (united atom for Kohn–Sham theory). Unfortunately, these radii are still not properly published, and the only information available is that they are optimised for the density-functional PBE0/6-31G* level [78]. Therefore, we used the hybrid PBE0 functional [79] together with the 6-31(+)G* basis set, using the IEFPCM method, the UAKS radii, and the same water solvent as for the other molecules. We tested to use both the same HF/6-31(+)G* geometries as for the other PCM calculations, or geometries optimised at the PBE0/6-31(+)G* level. These calculations are called IEFPCM/PBE0/HF and IEFPCM/PBE0/PBE0 in the following.

Finally, we also tested calculations at the B3LYP/6-31(+)G* level in the IEFPCM/UAKS calculations, using the same parameters as for PBE0, but only for the HF geometries. These calculations are called IEFPCM/B3LYP. A summary of all tested methods is given in Table 1.

All calculations with Gaussian-03 gave very poor results for the CH_3CONH^- ion (errors of 233–278 kJ/mol). This was caused by a poor radius of the N atom (1.04 Å). Therefore, we instead used the UAHF radius in Gaussian-98 (1.39 Å), which gave much more reasonable results (errors of 34–44 kJ/mol).

LD calculations

Next, we calculated hydration free energies with the LD method, as implemented in the ChemSol 2 software [39]. The geometry of all molecules was first optimised at the HF/6-31(+)G* level. Then, electrostatic potential (ESP) charges were calculated at the HF/6-31G* level with the Merz–Kollman scheme [80], as implemented in the Gaussian-03 software [18], but with a higher-than-default density of points, 10 shells with 17 points/unit area (giving ~2500 ESP points/atom; default is 6 shells with 1 point/unit area). These charges were used in the ChemSol software to calculate the solvation energies with the default radii and parameters of this software [16]. Reported energies are the sum of the Langevin dipole, entropy, van der Waals, and Born–Onsager energy components.

PB calculations

The PB calculations used the methods present in the Amber 9 software [81]. First, we used the default PB method, employing optimised Amber radii and separate cavitation and dispersion terms (called Tan2) [48]. Here, the cavitation energy is estimated from the SASA, whereas the dispersion term is obtained by a surface-based integration method, closely related to the one used in the PCM method. The calculations employed either charges obtained in the same way as the Amber 1994 (FF94) [82] or 2003 (FF03) [83] force fields, although the radii

are optimised for the former (R. Lou, personal information). In the former case, the electrostatic potential was obtained at the HF/6-31G* level in vacuum, whereas in the latter case, the B3LYP/cc-pVTZ method was used, with solvent effects treated with the IEFPCM and a dielectric constant of 4. The potential was calculated in points sampled by the Merz–Kollman scheme [80] (with the same higher-than-default point density as for the LD calculations) and the charges were estimated by the restrained ESP method (RESP) [84], using the antechamber module of Amber. Before the charge fitting, the molecules were optimised at the HF/6-31(+)G* level. We also tested the raw Merz–Kollman HF/6-31G* charges (employed in the LD calculations; called MK). Reported energies are the sum of the electrostatic, dispersion, and cavitation energy components. Default Amber parameters were used, indicating that

$$\Delta G_{\mathrm{cav}} = \gamma * \mathrm{SASA} + \beta \qquad (1)$$

with $\gamma = 0.182$ kJ/mol/Å^2 and $\beta = -4.52$ kJ/mol, and a 1.6 Å radius of probe molecule. However, the solvation grid was four times larger than the solute (FILLRATIO) and the grid spacing was 0.1 Å (SPACE), to ensure proper convergence of the solvation energies.

Second, we employed only a single non-polar (cavitation) term, using the same radii (Tan1), the modified Bondi radii (mbondi) also optimised with Amber [47], or the PARSE radii [20]. In this case, $\gamma = 0.0227$ kJ/mol/Å^2 and $\beta = 3.85$ kJ/mol was used in Eqn. 1, together with a probe radius of 1.4 Å. Otherwise, all parameters were the same as in the other calculations. These calculations were performed only for the FF94 charges. Below, the methods are specified by the charges and radii used, e.g. PB/FF94/mbondi (cf. Table 1).

We tested the convergence of the PB energies [85] by comparing the results obtained with the Amber pbsa module with those obtained by two other programs, Delphi II [86] and the adaptive PB solver (APBS) [87]. The Amber results agree with those of APBS within 2 kJ/mol for all molecules studied, although the Amber results are slightly more positive, with only one exception. The Delphi results are even more negative, and differing by up to 4 kJ/mol from both the Amber and APBS results. However, the Delphi results strongly depend on the parameters given to the program. For the default grid spacing (0.5 Å, instead of 0.1 Å), errors of up to 54 kJ/mol are observed.

GB calculations

Finally, we also tested the four GB methods available in the Amber 9 software, viz. the Hawkins, Cramer and Truhlar pairwise GB^{HCT} method [34,88] with parameters described by Tsui and Case [89], the GB^{OBC} method by Onufriev, Bashford, and Case, either with the model I or model II parameters (called GB^{OBC1} and GB^{OBC2} below) [90], or the GBn model [91]. The four methods employ different atomic radii as is detailed in the Amber manual [81]: the modified Bondi radii [47] for GB^{HCT}, a second modified Bondi radii set (mbondi2) [90] for the two GB^{OBC} methods, and unmodified Bondi radii for GBn [92]. Default parameters were used for all methods.

Like for the PB methods, the electrostatic solvation energies from GB methods must be combined with an estimate of the non-polar solvation energy. The default SASA methods for GB in Amber employ a single surface-tension term (i.e. $\beta = 0$ in Eqn. 1), but this gave poor results for the values we tested ($\gamma = 0.0201$, 0.182, and 2.27 kJ/mol/Å^2 [81]). Therefore, we used instead the non-polar energies from the PB calculations, i.e. the dispersion and cavity function with optimised radii [48] (called Tan2), or Eqn. 1. with SASA calculated using the Tan, Bondi, mbondi, mbondi2, and PARSE radii (called Tan1, Bondi, mbondi, mbondi2, and Parse, as for the PB methods), preferably the same radii used in the GB method. The methods are specified by giving the GB method, the charges, and the radii used for the non-bonded term, e.g. GBn/FF94/Bondi. All 24 methods used are shortly described and summarised in

Table 1.

Result and Discussion

Test set

First, we used the various methods for our test set of 40 molecules with experimentally known free energies of hydration. A summary of the results of the 24 different methods is presented in Table 2, whereas the full data set is listed in Table S3 in the supplementary material. It can be seen that the DPCM/HF and IEFPCM/PBE0/PBE0 methods give the best results, with a mean absolute deviation (MAD) of only 1.3 kJ/mol for the 20 neutral molecules. This is similar to the results of the original UAHF calibration (i.e. DPCM/HF) [10], which involved 43 neutral molecules (but only 9 of those in our test set) and gave a MAD of 0.7 kJ/mol. This indicates that our calculations are correct and similar to the original calculations, although our results differ by up to 0.5 kJ/mol for the 9 overlapping compounds, showing that we cannot exactly reproduce the original results with Gaussian-98. The maximum error for the DPCM/HF method is 7.3 kJ/mol for trifluoroethanol, which is three times larger than the maximum error in the original investigation (2.3 kJ/mol). This is probably reasonable for molecules not included in the training set, especially as the training set only contained a single molecule with fluorine, CH_3F. The maximum error for IEFPCM/PBE0/PBE0 is 5.1 kJ/mol, for N-propylguanidine.

For the ions, the MAD is appreciably larger, 16 kJ/mol for DPCM/HF and 23 kJ/mol for IEFPCM/PBE0/PBE0. This is four times more than in the original investigation, 4 kJ/mol, but such a low uncertainty is highly suspicious, considering that the uncertainty in the experimental results is 8–29 kJ/mol [35,39,58,94]. In fact, the main reason for the large difference is that we have used different experimental data: The experimental estimates for the nine ions that are included in both investigations differ by up to 33 kJ/mol. On the other hand, the calculated solvation energies in the two investigations differ by only 0–4.9 kJ/mol, showing that we use the method properly.

The other three PCM variants give similar, but slightly worse results, as can be expected for methods not explicitly parametrised. For the neutral molecules, they give MADs ranging from 1.7 kJ/mol to 2.6 kJ/mol. Likewise, the maximum errors range from 6.3 to 8.8 kJ/mol. For the ions, the MADs are 19–25 kJ/mol.

For ChemSol (LD), the results are somewhat worse: the MAD is 3.5 kJ/mol for the neutral molecules, with a maximum error of 22 kJ/mol for N-propylguanidine. Again, this is similar to the original results [39], which showed a MAD of 3 kJ/mol and a maximum error of 13 kJ/mol for 40 neutral molecules. However, for the eight molecules that are included in both sets, there are differences of up to 4.3 kJ/mol (for water), which is quite unexpected, considering that we have used the original program with default parameters. It is possible that this is caused by details in the treatment, but it is more likely that some parameters of the method have changed after the calibration. In fact, results of the ChemSol method changed by up to 5.4 kJ/mol when going from the first [38] to the second version of the program [39].

For the ions, the differences are even larger, up to 31 kJ/mol for CH_3O^-. Therefore, we obtain a larger MAD (21 kJ/mol) than in the original investigation (11 kJ/mol, maximum error 29 kJ/mol). This is also partly owing to the use of different experimental data (they used a solvation free energy of the proton that is 19 kJ/mol different from the more recent value used by us [63]).

Next, we tested seven variants of the PB method, as implemented in the Amber 9 software. We used three different charge sets, viz. Amber 1994 and 2003 RESP charges (i.e. HF/6-31G* vacuum charges or B3LYP/cc-pVTZ continuum solvent charges), and the raw HF/6-31G* Merz–Kollman (MK) charges. Moreover, we used two different methods to calculate the non-polar solvation energy, viz. the two-term (dispersion and cavitation) method

of Tan et al. (Tan2) [48] and the one-term SASA method. For the former, we used only the optimised radii of Tan et al. [48], whereas for the latter method (and also for the electrostatic contribution), we tested three different sets of atomic radii, viz. those optimised by Tan et al. for the use of a two-term non-polar energy (Tan1) [48], the modified Bondi radii (mbondi), also optimised with Amber [47], and the PARSE radii [20]. In total, seven combinations were tested, as are summarised in Table 1. The results are given in Table 2.

It can be seen that the seven approaches give rather similar results for the test set, with MADs of 6–8 kJ/mol for the neutral molecules. This is similar to a previous investigation, employing 500 neutral molecules and modified Bondi radii with HF/6-31G* charges calculated with both RESP and MK, MAD = 6 kJ/mol [47]. Unfortunately, individual values are not listed in that investigation, so no direct comparison is possible. For the ions, a much larger variation is obtained, with MADs being 19–23 kJ/mol for the methods based on SASA and mbondi or Parse radii, but 57–58 kJ/mol for the methods based on the Tan radii (with either one or two terms). For their test set of 53 ions, Rizzo et al. obtained MADs of 27 kJ/mol [47].

For the neutral molecules, we obtain the lowest MAD and standard deviation for the FF03/Tan2 combination. However, it has systematically too positive results and gives large errors for the ionic compounds. The reason for this may be that the method was calibrated to reproduce calculations with explicit solvent, rather than experimental data [48]. The FF94/mbondi and MK/mbondi combinations, give slightly larger MADs for the neutral test set, but much lower MAD for the ionic compounds and no significant systematic error. Therefore, they are probably preferable for general use. They are the methods calibrated by Rizzo et al. [47].

Finally, we also tested the four GB methods available in the Amber 9 software, viz. the GB^{HCT}, GB^{OBC1}, GB^{OBC2}, and GBn methods [34,90,91]. As detailed in the Methods sections, they were combined with the same three charge models (FF94, FF03, and MK) and the two different methods for non-polar energies (the Tan2 cavity and dispersion energies [48] or the SASA estimate in Eqn. 1) and five different sets of radii for those energies (Tan1, Bondi, mbondi, mbondi2, or PARSE). In total, 11 different combinations were tested, as is described in Table 1. The results are given in Table 2.

It can be seen that the GB methods give similar results for the neutral test molecules, with MADs ranging from 3 to 6 kJ/mol. Thus, the GB methods actually perform better than PB methods for this test set, although they are often described as an approximation to PB and are calibrated towards PB data. The GB^{HCT}/mbondi combination with the FF94 and MK charges were also included in the test by Rizzo et al. and they obtained a similar result for their test set of 500 neutral molecules, with MADs of 6–7 kJ/mol [47]. However, at least for our test set, the results are strongly improved if the GBn method is used instead (the MAD is reduced to 4 kJ/mol, except with the Tan2 non-polar term). This is quite impressive, because the GBn method was never calibrated towards small-molecule solvation data. Instead, it was constructed to solve some known shortcomings in the GB^{OBC} methods and it was calibrated only towards PB and explicit-solvent calculations of proteins and amino-acid models. Apparently, the improved performance for small molecule is a by-product, indicating that the improvements were physically sound and that the non-polar SASA models are quite transferable.

It should be noted that another GB model, the surface GB model (SGB; not tested in this paper) [93] has previously been calibrated with both a standard SASA estimate and a 38-parameter SASA estimate (different values of γ and β in Eqn. 1 for 19 atom types), giving MADs of 3.7 and 1.3 kJ/mol for a set of ~200 neutral compounds, i.e. similar to our results for the first method and results approaching those of PCM with the heavily parametrised method [51]. Likewise, the SM8 method (not tested here), which is also based on GB calculations, gives MADs of 2.3–2.8 kJ/mol for 274 neutral molecules, using charges obtained at various QM levels [36].

For the ions, the GB methods give MADs of 25–30 kJ/mol, with only small differences between the four GB variants. Rizzo et al. obtained similar result (MADs of ~30 kJ/mol) [47], but with the heavily parametrised SGB+SASA method, a MAD of 8 kJ/mol was obtained for 32 carboxylate and ammonium ions [51] and SM8 gave MADs of 13–15 kJ/mol for 112 ions [36].

The results of the GBn method depends only slightly on the radii used for the SASA method. In fact, the Tan1 radii give the lowest MADs for both neutral molecules and ions, but the results with the Bondi radii are very similar (and they actually give slightly smaller standard deviations for the neutral molecules). Considering that Bondi radii are already used within the GBn method, we therefore suggest the consistent GBn/FF94/Bondi (or GBn/MK/Bondi) combination for general use. It actually gives a similar performance as the ChemSol method and is clearly better than all the PB methods for our test set.

It can also be noted that the GB methods are less sensitive to the details of the calculations than the PB methods: For example, the largest difference for the solvation energies of all the neutral molecules in the test set between the FF94 and FF03 charges is 15 kJ/mol for PB, but only 6 kJ/mol for GBn (the FF94 and MK charge sets differ by less than 10 kJ/mol for PB and 3 kJ/mol for GBn). Moreover, the average difference between GBn and the three other GB methods is only 6, 3, and 3 kJ/mol, i.e. less than for the PB methods with different sets of radii.

It can be noted that the mean *signed* deviation (MSD in Table 2) for the neutral molecules is close to zero for most methods. Only eight methods have a MSD that is significantly different from zero at the 95% level, viz. all PB methods, except those with the mbondi radii, and the IEFPCM/B3LYP, GB^{OBC1}/FF94/mbondi2, and GBn/FF94/Tan2 methods. This indicates that the other methods tested do not suffer from any significant systematic errors. For the ions, the standard deviations are so large that no significant systematic errors can be discerned.

Finally, we have in Table 2 also included the mean unsigned *relative* deviations (MURDs) of the various methods. It can be seen that they show similar trends as the MADs, ranging from 0.09 for DPCM/HF to 0.50 for GBn/FF94/Tan2 for the neutral molecules. Interestingly, they are appreciably smaller for the ions: Again DPCM/HF gives the best results 0.05, but all methods, except the PB methods with the Tan radii give MURDs of 0.09 or less. This is because all ions in our test set give quite similar solvation energies (in fact, assigning the same solvation energy, –285 kJ/mol, to all 20 ions gives a MURD of 0.08). Another problem with MURD is that it gives a high weight to neutral molecules with small solvation energies, e.g. toluene, CH_3SH, and $CH_2SCH_2CH_3$, which have solvation energies of only 4–6 kJ/mol.

An interesting question is whether the superior performance of the PCM methods reflects that PCM is inherently a better method (e.g. because it is directly based on self-consistent QM calculations) or if it is only caused by a more thorough parameterisation. It seems that the latter is the more likely answer. First, there are highly parameterised GB methods that give similar accuracy as PCM [36,56]. Moreover, if the polar part of the DPCM method is combined with any of the SASA estimates of the non-polar parts, much worse results are obtained (e.g. 5 kJ/mol MAD for the neutral test set with the non-polar Bondi SASA term). This shows that the polar and non-polar terms are not independent or transferable between the various methods.

Drug-like molecules

Next, we turn to the four sets of drug-like molecules, the main subject of this paper. Our aim is to investigate how accurate the solvation predictions are for such larger molecules with several functional groups. The raw results for these molecules are collected in Table S4.

First, we concentrate on the 20 neutral drug-like molecules, because these data are independent and can be compared to the test set of 20 small neutral molecules. From Table 3,

it can be seen that the 20 drug-like molecules show a much larger spread in the results of the 24 different continuum solvation methods than the 20 small molecules: The range for the various methods (largest minus smallest calculated solvation energy) vary between 34 kJ/mol for progesterone to 250 kJ/mol for paclitaxel (taxol), with an average of 80 kJ/mol. This is four times larger than for the small molecules, for which the range varied between 5 kJ/mol for propane to 61 kJ/mol for N-propylguanidine, with an average of 19 kJ/mol. Similar results apply if instead the standard deviation is used as a measure of the spread of the calculated results. The difference between the two sets of molecules is significant at a level of over 99.9%.

Of course it can be argued that this spread is caused mainly by the methods with the largest errors compared to experimental results. Therefore, we have made a similar analysis based only on the five PCM methods, which gave the lowest MADs compared to experiments for the small molecules. However, from Table 3, it can be seen that this does not change the results significantly (besides that the spread is appreciably smaller): The range for the drug-like molecules (4–26 kJ/mol, average 9 kJ/mol) is still three times larger than that for the small molecules (0.3–8 kJ/mol, average 3 kJ/mol).

For two of the drug-like molecules, the difference among the various methods is particularly large, viz. paclitaxel and SB3. In fact, the five PCM methods give positive solvation energies for these two molecules, whereas all the other methods give negative solvation energies. The difference can be traced entirely to the non-polar terms and may be connected to the use of pair-potentials for the dispersion and repulsion terms [94] or to the use of the van der Waals surface (which contains many small cavities inside large molecules), rather than the solvent-accessible surface, for the cavitation term. If these two molecules are removed from the analysis, the average range decrease to 70 kJ/mol, but this is still 4 times larger than for the small organic molecules.

The main reason for this larger variation for the drug-like molecules is that they are more polar than the small molecules. From Table 3, it can be seen that the small molecules have average solvation energies from –44 to + 9 kJ/mol (unsigned average 20 kJ/mol), whereas that of the drug-like molecules are from –25 to –162 kJ/mol, with an average of –73 kJ/mol. Thus, the drug-like molecules have on average almost four times larger solvation energies than the small molecules in our test sets, which is similar to the observed difference in the range and standard deviation. This indicates that the error in the continuum solvation methods are proportional to the solvation energy and this is confirmed in Figure 1, which shows how the standard deviation among the various methods (either all or only the five PCM methods) increases as the average solvation energy of the molecule increases, both for the small molecules and the drug-like molecules. The correlation coefficients (r^2) are 0.5–0.6 and the slopes are 0.16–0.18 for all methods and 0.05–0.06 for PCM. Thus, it seems clear that the *relative error* is a better measure of the accuracy of the continuum solvation methods than the absolute error, although most previous calibrations only report MADs [10,36,56]. Thus, the low MADs of neutral molecules, e.g. ~1 kJ/mol for DPCM/HF and IEFPCM/PBE0/PBE0 in Table 2 and 1–2 kJ/mol for the best calibrated continuum solvation methods [10,36,51,54,56], is not directly applicable to drug-like molecules. Instead, we expect a relative error of ~0.09–0.10 for these methods, which corresponds to 2–16 kJ/mol for the neutral drug-like molecules in our set.

However, it is also interesting to get more information about the accuracy of the individual methods. Of course, such a comparison is complicated by the fact that the experimental value for the solvation energy of these molecules is unknown. Still, we could use the average of all our theoretical estimates as a reasonable estimate of the solvation energy of the drug-like molecules. Statistically, the best choice is to use an average weighted by the inverse variance of each method for the neutral test set and also by the inverse number of each type of methods (5 PCM, 1 LD, 7 PB, and 11 GB) to avoid a bias from the number of variants of each method. This method-weighted average is also included in Tables 2 together

with the individual weights for each method. It can be seen that the method-weighted average gives MADs of 1.5 and 18 kJ/mol for the neutral molecules and ions compared to the experimental results, respectively. This shows that this type of weighting, which does not consider the signs of the errors, gives a proper estimate of the solvation energy. It should be pointed out that the selection of this average is somewhat arbitrary and this analysis is only performed to get results that are easier to interpret. However, the general results in Table 3 are independent of this selection of an average and gives qualitatively the same results.

The MADs of all the 24 methods compared to the method-weighted average are given in Table 4. If we concentrate on the MAD for neutral molecules (for which the data are independent), we can see that MAD is lowest for the PCM methods (8–17 kJ/mol; IEFPCM/HF best; DPCM/HF gave a lower MAD, but it failed for two of the larger molecules, for which the other PCM methods gave large deviations). Next come the PB/FF94/Tan1, GB^{OCB2}/FF94/mbondi2, and ChemSol methods with MADs of 18–21 kJ/mol. The other PB and GB methods had uniform MADs of 26–32 kJ/mol, except for the PB/FF94/Parse and GBn/FF94/Tan2 combinations, which gave higher MADs (47 and 51 kJ/mol, respectively).

However, the most interesting observation is that the MADs of all methods is larger for the neutral drug-like molecules than for the test set, by a factor 2 (PB/FF94/Tan1) to 9 (GBn/FF94/Tan2) with an average of 6 (it does not matter if we use the experimental or method-weighted average for the small molecules; the MAD differ by less than 1.1 kJ/mol or 27%). If paclitaxel and SB3 are removed from the analysis, the MADs are reduced to 7–13 kJ/mol for PCM, 14 kJ/mol for ChemSol, 14–38 kJ/mol for PB, and 13–45 kJ/mol for GB, which is still 2–7 times larger than for the small neutral molecules (average 4.6). This confirms the results obtained directly on the standard deviation of ranges (Table 3) and shows that the MADs obtained for small neutral molecules are not applicable for drug-like molecules.

Such large uncertainties for the drug-like molecules would actually be a disaster for any method that tries to estimate absolute ligand-binding affinities involving continuum-solvent methods: The MADs of the best PCM method of 10 kJ/mol corresponds to a factor of 50 in the binding constant and the MADs of the typical PB and GB methods (~29 kJ/mol), widely employed in the MM/PBSA and MM/GBSA approaches, correspond to an uncertainty in the binding constants of 100 000! This would make such approaches completely useless.

Fortunately, the situation is not that bad in real applications, because only related drug candidates are normally considered, i.e. molecules with a common molecular framework, and only relative affinities are of interest. It is likely that a significant part of the errors related to each method cancel in such cases. This was tested by considering the remaining sets of ligands, consisting of three series of ligands to avidin, factor Xa, and galectin-3. In all three series, the general structure of the inhibitors is similar and only part of the structure show a variation.

The MADs for these three series of ligands together (compared to the method-weighted average) is rather similar to that for the 20 neutral drug-like molecules, 8–18 kJ/mol for PCM methods (IEFPCM/HF best), 20 kJ/mol for ChemSol, 13–35 kJ/mol for GB (GBn/FF94/Tan1 best), and 30–102 kJ/mol for PB (FF94/mbondi and MK/mbondi best). All PB methods with the Tan1 and Tan2 radii have large problems with the highly charged factor Xa inhibitors.

We therefore studied whether the predictions are more accurate if only similar compounds are considered, i.e. if we only try to predict the relative solvation energies for the avidin, factor Xa, and galectin-3 inhibitors, respectively. The results of such an approach is also included in Table 4 (column Target). It can be seen that the results for most of the methods improve by a factor of about three: The MADs are 3–6 kJ/mol for PCM, 16 kJ/mol for ChemSol, 7–36 kJ/mol for PB, and 6–11 kJ/mol for the GB methods.

Further improvements can be expected if we only consider molecules with the same net charge, i.e. if we further divide the biotin analogues into two sets (Btn1–3 with a negative

charge and the neutral Btn4–7) and the factor Xa inhibitors also into two sets with either +1 or +2 charge. The results of such a procedure are also shown in Table 4, and it can be seen that the MADs are now reduced to 2–4 kJ/mol for PCM, 3–6 kJ/mol for GB, 5 kJ/mol for ChemSol, and 5–9 kJ/mol for the PB methods. In fact, all the GB methods, except that with Tan2 radii can predict all the relative solvation energies within 3–4 kJ/mol and all PB methods except FF03/Tan2 within 4–5 kJ/mol. However, this still corresponds to a factor of 4–9 for the binding constants.

Of course, these results depend somewhat on the choice of the average used as a replacement for experimental data, especially the absolute MADs. However, as can be seen in Table S5 (which is based instead on the unweighted average that instead is biased towards the GB methods), once the relative solvation energies are considered, the results are much more stable. In particular, it can be seen that all methods except two can predict the relative solvation energies of the three inhibitor series with the same net charge within 2–5 kJ/mol, exactly as for the method-weighted average, showing that this is a general results.

Moreover, similar (but somewhat less informative) results can be obtained by studying the spread (standard deviation or range) of the various methods instead (as in Table 3). Table S6 shows that the spread for all methods is almost the same for the small molecules and the drug-like molecules, once only relative energies within series with the same net charge are considered for the latter. This shows that the results are general and do not depend of any comparison with a certain average value. However, it can be noted that for the PCM methods, results for the drug-like molecules are still 2–3 times worse than for the small molecules, indicating that the PCM methods probably perform worse for the drug-like molecules than for the small molecules.

Conclusions

In this paper we have tested how well the free energies of hydration of four series of drug-like molecules can be estimated by 24 variants of continuum-solvation methods of four widely used types (PCM, LD, PB, and GB). A continuum estimate of solvation energy is involved in most simplified physical methods to estimate ligand-binding affinities [95]. Typical examples are the widely used MM/PBSA and MM/GBSA methods [60], but also recent attempts to estimate ligand-binding affinities with quantum mechanical calculations of the whole ligand–protein complex use continuum solvation methods [15,96].

Our aim is to estimate the actual accuracy of available continuum-solvation models for drug-like molecules. Unfortunately, experimental data are not available for drug-like molecules. We have used two approaches to solve this problem. First, we have studied the spread of the results for the various methods. It turns out that both the range and the standard deviation of the methods is ~4 times larger for a set of 20 neutral drug-like molecules than for a test set of 20 small neutral molecules. The result is similar also if only the methods that give the best results for small organic molecules (PCM) are considered.

Second, we used the method-weighted average for our 24 different continuum solvation methods as an estimate of the experimental solvation energy. The weight is based on the variance of each method for the 20 neutral molecules. This approach also shows that all 24 methods actually have a higher MADs for the drug-like molecules than for the small organic molecules, by an average factor of 6.

The reason for this behaviour is that the drug-like molecules have larger (more negative) solvation energy than the small molecules, on average by a factor of almost four in our two test sets, and that the uncertainty in the solvation methods is proportional to the solvation energy (Figure 1). Thus, the accuracy of continuum solvation methods should be discussed in relative terms, rather than absolute. According to the data of the neutral small molecules, the PCM methods have relative errors of 0.09–0.17, ChemSol, GB, and PB methods have relative errors of 0.19, 0.25–0.50, and 0.34–0.47, respectively. For charged molecules, the relative

errors are somewhat smaller, 0.05–0.08, 0.06, 0.07–0.09, and 0.06–0.18, for PCM, ChemSol, GB, and PB, respectively. This means that for all the 38 neutral or singly-charged drug-like molecules studied in this paper, we would expect average errors of 2–20, 5–24, 7–41, and 9–55 kJ/mol for the best PCM, ChemSol, GB, and PB methods, respectively. Considering that 6 kJ/mol corresponds to an order of magnitude in the ligand-binding constant, this would be a catastrophe for any ligand-binding method involving such a continuum-solvation estimate.

Fortunately, the results are strongly improved if only relative energies within a series of similar drug-like molecules are studied, especially if only molecules of the same total charge are considered: The MADs are then reduced to 2–4, 3–4, 4–5, and 5 kJ/mol for all PCM, GB, PB, and LD methods, with only two exceptions. This is probably because there is a large degree of error cancellation when relative energies are considered, especially for binding energies, where we actually do not need to transfer the solutes to gas phase, but only move the ligand from water solution to the protein, i.e. within two condensed phases. It also shows that continuum-solvation methods are still useful if relative energies are considered and if errors of this size are acceptable, and also that the results are insensitive to the details and parameters of the methods. In fact, there seems to be little gain of using a quantum-mechanical method, such a PCM, to improve the continuum solvation energies for ligand-binding affinities.

These results explain why solvation energies for proteins differ so much between different continuum methods [97]: The electrostatic component of the solvation energy of a protein is dominated by the charged groups, which are of only four types (carboxylate, amine, guanidine, and imidazole). Therefore, the solvation energy will strongly depend on how the methods perform on these certain groups. As can be seen from Table S3, the various methods may differ by up to 115 kJ/mol for these groups, and this difference will be multiplied by the number of each group, giving very large differences in absolute terms. Moreover, the results explain why the MM/PB(GB)SA approaches sometimes give good, sometimes poor absolute binding affinities, depending on the target–ligand systems [60,66,67]. In this case, the solvation energy of the protein will mostly cancel and the results will mainly depend on the solvation energy of the isolated ligand. Therefore, the results will depend on how the method used performs for the absolute solvation energy of this certain ligand, which may differ significantly between different ligands, as can be seen in Table 3. However, once only relative energies between ligands of the same type are considered, most methods will give similar and quite reliable results.

Our calculations also give some indications concerning the accuracy of the various methods. Among our tested methods, the PCM variants clearly give the best results, with MURDs of 0.09 and 0.05 for the neutral and ionic test sets for the DPCM/HF method. The results are somewhat worse for the other PCM methods that are not equally thoroughly parametrised, but they are still better than all the other tested methods, at least for the small neutral molecules. However, there are indications that the PCM methods do not work equally well for drug-like molecules and that PCM fails completely for the two largest molecules.

Interestingly, our results also clearly show that the GB methods are more stable and more accurate than the widely used PB methods. This applies especially to the newer GBn method [91], which does not seem to have been calibrated towards small-molecule data before. Our results indicate that it is significantly more accurate than other GB variants in the Amber software, giving a MURD for neutral molecules of down to 0.25, when combined with a standard surface-area expression for the non-bonded term (but not with $\beta = 0$, as is default in Amber). This is clearly the method of choice among those available in the Amber software.

On the other hand, we do not see any clear preference for any set of radii to use for the non-polar part in the GB methods: Bondi, Parse, and Tan1 seem to give similar results (as do also different sets of ESP charges). Only the more advanced Tan2 method gives worse results, probably because it was calibrated with respect to explicit solvent calculations [48]. This most likely reflects that the non-polar part of the solvation energy is mostly quite small for our test molecules.

Finally, the ChemSol method gives slightly better results than the GB methods, with MURDs of 0.19 and 0.06 for the small neutral and ionic test sets.

In conclusion, we have from this investigation gained several results of interest for the prediction of ligand-binding affinities. In particular, we cannot expect to get more accurate results than 3 kJ/mol, even for relative energies, for any method that involves the estimation of the free energy of solvation by a continuum method. This restricts the use of such methods to ligand series, in which the binding constants differ by more than a factor of 3. Such a warning also applies to any attempt to estimate acid constants by continuum methods [98]. In fact, this problem is even harder, because the protonated and deprotonated species necessarily differ in charge. Therefore, at least one of the species will be charged, indicating that the accuracy will not be better than ~16 kJ/mol or 2.8 pK_a units.

Acknowledgements

This investigation has been supported by grants from the Swedish research council and by computer resources of Lunarc at Lund University. J. K. thanks the Villum Kann Rasmussen foundation for financial support. We thank Drs. J. Stålring, P. Rydberg, and L. Olsen for help with the selection and set-up of drug-like molecules.

References

1 Cramer, C. J.; Truhlar, D. G. Chem. Rev. 1999, 99, 2161-2200.
2 Orozco, M.; Luque, F. J. Chem. Rev. 2000, 100, 4187-4225.
3 Simonson, T. Curr. Opin. Struct. Biol. 2001, 11, 243-252.
4 Feig, M.; Brooks, C. L. Curr. Opin. Struct. Biol. 2004, 14, 217-224.
5 Born, M. Zeitsch. Physik 1920, 1, 45.
6 Onsager, L. J. Am. Chem. Soc. 1936, 58, 1486.
7 Miertus, S; Scrocco, E.; Tomasi, J. Chem. Phys. 1981, 55, 117.
8 Tomasi, J.; Persico, M. Chem. Rev. 1994, 94, 2027-2094.
9 Tomasi, J.; Mennucci, B.; Cammi, R. Chem. Rev. 2005, 105, 2999-3093.
10 Barone, V.; Cossi, M. J. Phys. Chem. A 1998, 102, 1995.
11 Tomasi, J.; Mennucci, B.; Cances, E. J. Mol. Struct. (Theochem) 1999, 464, 211.
12 Chipman DM (2000) J Chem Phys 112:5558.
13 Klamt, A.; Schüürmann, G. J. Chem. Soc. Perkin Trans. 2, 1993, 799.
14 Vreven, T.; Mennucci, B.; da Silva, M. O.; Morokuma, K.; Tomasi, J. J. Chem. Phys. 2001, 115, 62-72.
15 Fedorov, D.G.; Kitaura, K.; Li, H.; Jensen, J. H.; Gordon, M. S. J. Comput. Chem. 2006, 27, 976-985.
16 Warwicker, J.; Watson, H. C. J. Mol. Biol. 1982, 157, 671.
17 Gilson, M.K.; Honig, B. Prot. Struct. Funct. Genet. 1988, 3, 32-52.
18 Sharp, K. A.; Honig, B. Annu. Rev. Biophys. Biophys. Chem. 1990, 19, 310.
19 Sharp, K. A.; Honig, B. Curr. Opin. Struct. Biol. 1995, 5, 323.
20 Sitkoff, D.; Sharp, K. A.; Honig, B. J. Phys. Chem. 1994, 98, 1978-1988.
21 Simonsson, T.; Brünger, A. T. J. Phys. Chem. 1994, 98, 4683.
22 Baker, N. A. Methods Enzymol. 2004, 383, 94.
23 Chen, Y; Noodleman, L.; Case, D. A.; Bashford, D. J. Phys. Chem. 1994, 98, 11875.
24 Baldridge, K.; Fine., R.; Hager, A. J. Comput. Chem. 1994, 15, 1217.
25 Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A.; Honig, B. J. Am. Chem. Soc. 1994, 116, 11875.
26 Gorgonea, V.; Merz, K. M. J. Phys. Chem. A. 1999, 103, 5171-5188.
27 Schnieders, M. J.; Baker N. A.; Ren P.; Ponder, J. W. J. Chem. Phys. 2007, 126, 124114.
28 Hoitjink, G. J.; de Boer, E.; Van der Meij, P. H.; Weijland, E. P. Recl. Trav. Pays-Bas 1956, 75, 487.
29 Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. J. Am. Chem. Soc. 1990, 112, 6127.
30 Onufriev, A.; Bashford, D.; Case, D. A. J. Phys. Chem. B 2000, 104, 3712-3720.
31 Bashford, D.; Case, D. A. Annu. Rev. Phys. Cehm. 2000, 51, 129-152.
32 Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. I. J. Comput. Chem. 2004, 25, 265-284.
33 Schnieders, M. J.; Ponder, J. W. J. Chem. Theory Comput. 2007, 3, 2083-2097.
34 Hawkins, G. D.; Cramer, C. J., Truhlar, D. G. Chem. Phys. Lett. 1995, 246, 122-129.
35 Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. J. Chem. Theory Comput. 2005, 1, 1133-1152.
36 Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. J. Chem. Theory Comput. 2007, 3, 2011-2033.
37 Warshel, A. J. Phys. Chem. 1979, 83, 1640.
38 Florián, J. Warshel, A. J. Phys. Chem. B 1997, 101, 5583.
39 Florián, J. Warshel, A. J. Phys. Chem. B 1999, 103, 10282-10288.
40 Hermida-Ramón, J. M.; Karlström, G.; Lindh, R. J. Phys. Chem. B 2002,106, 7115-7120.
41 Langevin, P. Annales de Chemie Physique, 1905, 8, 70.
42 Sandberg, L.; Casemyr, R.; Edholm, O. J. Phys. Chem. 2002, 106, 7889-7897.
43 Hermann, R. B. Theory of hydrophobic bonding. II. Correlation of hydrocarbon

solubility in water with solvent cavity surface area. J. Phys. Chem. 1972, 76, 2754-2759.
44 Cossi, M.; Tomasi, J.; Cammi, R., Int. Quantum. Chem. Quant. Chem. Symp. 1995, 29, 695.
45 Floris, F. M.; Tomasi, J. J. Comput. Chem. 1989, 10, 616.
46 Cossi, M.; Mennucci, B.; Cammi, R. J. Comput. Chem. 1996, 17, 57.
47 Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D.; J. Chem. Theory Comput. 2006, 2, 128-139.
48 Tan, C. ; Tan, Y.-H.; Luo, R. J. Phys. Chem. B 2007, 111, 12263-12274.
49 Maple, J. R.; Cao, Y.; Damm, W.; Halgren, T. A.; Kaminski, G. A.; Zhang, L. Y.; Friesner, R. A. J. Chem. Theory and Comput. 2005, 1, 694-715
50 Wagoner, J. A.; Baker, N. A. Proc. Natl. Acad. Sci. USA 2006, 103, 8331-8336.
51 Gallicchio, E.; Zhang, L. Y.; Levy, R. M. J. Comp. Chem. 2002, 23, 517-529.
52 Gallicchio, E.; Levy, R. M. J. Comp. Chem. 2004, 25, 479-499.
53 Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. J. Phys. Chem. 1996, 100, 16098.
54 Barone, V.; Cossi, M.; Tomasi, J. J. Chem. Phys. 1997, 107, 3210-3221.
55 Tan, C.; Yang, L.; Luo, R. J. Phys. Chem. B 2006, 110, 18686-18687.
56 Klamt A, Mennucci B, Tomasi J, Barone V, Curutchet C, Orozco M, Luque FJ (2009) Acc Chem Res, in press DOI 10.1021/ar800187p.
57 Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. J. Chem. Phys. 2003, 119, 5740-5761.
58 Li J, Zhu T, Hawkins GD, Winge P, Liotard DA, Cramer CJ, Truhlar DG (1999) Theor Chim Acta 103:9-63
59 Thompson JD, Cramer CJ, Truhlar DG (2004) 108:6532-6542.
60 Kollman, P. A.; Massova, I.; C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A . Case, T. E. Cheatham, Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc. Chem. Res. 2000, 33(12), 889-897.
61 Vorobjev, Y. N.; Hermans, J. Biophys. Chem. 1999, 78, 195-205.
62 Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. Biochemistry 1981, 20, 849-855.
63 Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. J. Phys. Chem. B 2006, 110, 16066-16081.
64 Alagona, G.; Ghio, C.; Nagy, P. I. Intern. J. Quant. Chem. 2004, 99, 161-178.
65 Takano, Y.; Houk, K. N. J. Chem. Theory Comput. 2005, 1, 70-77.
66 B. Kuhn, P. A. Kollman, Binding of a Diverse Set of Ligands to Avidin and streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models, J. Med. Chem. 2000, 43, 3786-3791.
67 A. Weis, K. Katebzadeh, P. Söderhjelm, I. Nilsson, U. Ryde (2006) "Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field", J. Med. Chem., 49, 6596-6606.
68 Kongsted, J.; Ryde, U., An improved method to predict the entropy term with the MM/PBSA approach, J. Comput.-Aided Mol. Design., 2008, ASAP; DOI: 10.1007/s10822-008-9238-z.
69 H. Matter, E. Defossa, U. Heinelt, P.-M. Blohm, D. Schneider, A. Müller, S. Herok, H. Schreuder, A. Liesum, V. Brachvogel, P. Lönze, A. Walser, F. Al-Obeidi, P. Wildgoose, J. Med. Chem. 45 (2002) 2749-2769.
70 Sörme, P.; Arnoux, P.; Kahl-Knutsson, B.; Leffler, H.; Rini, J. M.; Nilsson, U. J.; J. Am. Chem. Soc., 2005, 127, 1737-1743.
71 Mobley, D. L.; Chodera, J. D.; Dill, K. A. J. Phys. Chem. B 2008, 112, 938-946.
72 Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M.; J. Mol. Biol. 1993, 231, 698-710.

73 Perola, E.; Charifson, P. S.; J. Med. Chem. 2004, 47, 2499-2510.
74 PC Spartan Pro 1.0.6, 2001, Wavefunction, Inc. 18401 Von Karman Avenue, Irvine, CA 92612, USA.
75 Hariharan, P. C.; Pople, J. A. Theor. Chim. Acta 1973, 28, 213.
76 Mennucci, B.; Tomasi, J. J. Chem. Phys. 1997, 106, 5151.
77 Gaussian 98 (Revision A.9), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.
78 Gaussian 03, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.
79 Perdew, J. P.; Burke, K.; Ernzerhof, M. Phys. Rev. Lett. 78 (1997) 1396.
80 B. H. Besler, K. M. Merz, P. A. Kollman, *Atomic Charges Derived from semiempirical Methods*, J. Comput. Chem. 1990, 11, 431-439.
81 D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross, and P.A. Kollman (2006), AMBER 9, University of California, San Francisco.
82 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 1995, 117, 5179-5197.
83 Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee.; A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins, J. Comput. Chem. 2003, 24, 1999-2012.
84 C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Deriving Atomic Charges - the RESP Model, J. Phys. Chem. 1993, 97, 10269-10280.
85 M. Feig, A. Onufriev, M. S. Lee, W. Im, D. A. Case, C. L. Brooks, J. Comput. Chem. 2004, 25, 265-284.

86 W. Rocchia, E. Alexov, B. Honig. J. Phys. Chem. B 2001, 105:6507-6514.
87 Baker N.A.; Sept D.; Joseph S.; Holst M.J.; McCammon J.A. Proc. Natl. Acad. Sci. USA 98, 10037-10041 2001.
88 Hawkins, G. D.; Cramer, C. J., Truhlar, D. G. J Phys. Chem. 1996, 100, 19824-19839.
89 Tsui, V.; Case, D. A. Biopol. 2001, 56, 275-291.
90 Onufriev, A.; Bashford, D.; Case, D. A. Proteins, 2004, 55, 383-394.
91 Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. J. Chem. Theory Comput. 2007, 3, 156-169.
92 Bondi, A., J. Phys. Chem. 1964, 68, 441-451.
93 Ghosh, A.,; Rapp, C. S.; Friesner, R. A. J. Phys. Chem. B 1998, 102, 10983.
94 Curutchet C.; Orozco, M.; Luque, F. J.; Mennucci, B.; Tomasi, J. *J. Comp. Chem.* **2006**, *27*, 1769-1780.
95 Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.* **2002**, *41*, 2644-2676.
96 Raha, K.; Merz, K. M. Large-scale validation of a quantum mechanics based scoring function: Predicting the binding affinity and the binding mode of a diverse set of protein–ligand complexes. J. Med. Chem. 2005, 48, 4558-4575.
97 Gohlke, H.; Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras–Raf. J. Comput. Chem. 2004, 28, 238-250.
98 Jensen, J. H.; Li, H.; Robertson, A. D.; Molina, P. A.; J. Phys. Chem. B 2002, 106, 3486-3496.

Table 1. A tabular description of the various continuum solvation methods. The table lists the method for the geometry optimisation (HF/6-31(+)G* or PBE0/6-31(+)G*, the method for calculation of the polar solvation term, including the radii used, as well as for the PCM quantum mechanical level of the calculations and for the other methods the way the ESP radii were estimated (MK, FF94, or FF03; see the methods section), and finally the method (PCM, ChemSol, Tan2 [48], or SASA, i.e. by Eqn. (1)) and radii used for the non-polar solvation energies.

Short-name	Geometry	Polar solvation			Non-polar solvation	
	optimisation	Method	QM level or charges	Radii	Method	Radii
DPCM/HF	HF/6-31(+)G*	DPCM	HF/6-31(+)G*	UAHF	PCM	UAHF
IEFPCM/HF	HF/6-31(+)G*	IEFPCM	HF/6-31(+)G*	UAHF	PCM	UAHF
IEFPCM/PBE0/HF	HF/6-31(+)G*	IEFPCM	PBE0/6-31(+)G*	UAKS	PCM	UAKS
IEFPCM/PBE0/PBE0	PBE0/6-31(+)G*	IEFPCM	PBE0/6-31(+)G*	UAKS	PCM	UAKS
IEFPCM/B3LYP	HF/6-31(+)G*	IEFPCM	B3LYP/6-31(+)G*	UAKS	PCM	UAKS
ChemSol	HF/6-31(+)G*	LD	MK	ChemSol	ChemSol	ChemSol
PB/FF94/Tan2	HF/6-31(+)G*	PB	FF94	Tan	Tan2	Tan
PB/FF94/Tan1	HF/6-31(+)G*	PB	FF94	Tan	SASA	Tan
PB/FF94/mbondi	HF/6-31(+)G*	PB	FF94	mbondi	SASA	mbondi
PB/FF94/Parse	HF/6-31(+)G*	PB	FF94	Parse	SASA	Parse
PB/FF03/Tan2	HF/6-31(+)G*	PB	FF03	Tan	Tan2	Tan
PB/MK/Tan2	HF/6-31(+)G*	PB	MK	Tan	Tan2	Tan
PB/MK/mbondi	HF/6-31(+)G*	PB	MK	mbondi	SASA	mbondi
GB^{HCT}/FF94/mbondi	HF/6-31(+)G*	GB^{HCT}	FF94	mbondi	SASA	mbondi
GB^{HCT}/FF03/mbondi	HF/6-31(+)G*	GB^{HCT}	FF03	mbondi	SASA	mbondi
GB^{HCT}/MK/mbondi	HF/6-31(+)G*	GB^{HCT}	MK	mbondi	SASA	mbondi
GB^{OBC1}/FF94/mbondi2	HF/6-31(+)G*	GB^{OBC1}	FF94	mbondi2	SASA	mbondi2
GB^{OBC2}/FF94/mbondi2	HF/6-31(+)G*	GB^{OBC2}	FF94	mbondi2	SASA	mbondi2
GBn/FF94/Bondi	HF/6-31(+)G*	GBn	FF94	Bondi	SASA	Bondi
GBn/FF03/Bondi	HF/6-31(+)G*	GBn	FF03	Bondi	SASA	Bondi
GBn/MK/Bondi	HF/6-31(+)G*	GBn	MK	Bondi	SASA	Bondi
GBn/FF94/Tan2	HF/6-31(+)G*	GBn	FF94	Bondi	Tan2	Tan
GBn/FF94/Tan1	HF/6-31(+)G*	GBn	FF94	Bondi	SASA	Tan
GBn/FF94/Parse	HF/6-31(+)G*	GBn	FF94	Bondi	SASA	Parse

Table 2. Performance of the various methods on the test set, presented as the mean absolute deviation (MAD), maximum deviation (Max), mean signed deviation (MSD), standard deviation (Stdev), and mean unsigned relative deviation (MURD) from the experimental data calculated either for the 20 neutral molecules or the 20 molecules with a net charge. In addition, the weight used for the method-weighted average is given in the last column. The full data set, as well as the experimental data with references are shown in Table S1 in the supplementary material.

Method	Neutral					Charged					
	MAD	Max	MSD	Stdev	MURD	MAD	Max	MSD	Stdev	MURD	weight
DPCM/HF	1.3	7.3	0.0	2.3	0.086	16.3	47.3	11.6	11.6	0.052	0.143
IEFPCM/HF	1.7	7.9	0.0	2.8	0.117	18.7	55.5	17.2	14.4	0.059	0.097
IEFPCM/PBE0/HF	1.7	6.3	0.7	2.3	0.114	23.0	52.7	18.7	19.0	0.072	0.142
IEFPCM/PBE0/PBE0	1.3	5.1	0.0	1.9	0.104	22.9	55.1	20.3	17.8	0.072	0.209
IEFPCM/B3LYP	2.6	8.8	1.9*	2.8	0.166	24.6	54.6	20.3	19.2	0.076	0.094
ChemSol	3.5	21.5	1.3	6.0	0.194	20.6	52.5	-1.2	26.2	0.062	0.102
PB/FF94/Tan2	6.6	31.2	6.1*	7.2	0.394	56.7	111.6	23.8	60.6	0.173	0.010
PB/FF94/Tan1	8.4	37.8	8.2*	8.9	0.467	58.5	110.0	27.8	59.6	0.180	0.007
PB/FF94/mbondi	6.3	27.0	-2.9	8.8	0.369	18.9	47.8	-1.4	24.3	0.059	0.007
PB/FF94/Parse	7.7	18.6	-7.3*	6.3	0.476	23.3	61.8	-11.1	27.4	0.073	0.013
PB/FF03/Tan2	5.6	16.3	5.1*	5.0	0.372	56.6	135.0	19.4	63.5	0.173	0.021
PB/MK/Tan2	6.1	30.8	5.6*	7.0	0.358	56.5	111.6	23.5	60.6	0.173	0.011
PB/MK/mbondi	6.1	27.0	-2.8	8.6	0.339	19.0	47.8	-2.0	24.5	0.059	0.007
GB^{HCT}/FF94/mbondi	6.4	29.8	-3.5	9.1	0.381	27.3	93.2	-5.5	35.3	0.082	0.004
GB^{HCT}/FF03/mbondi	5.5	25.6	-2.9	7.9	0.334	28.2	84.9	-9.8	34.6	0.085	0.005
GB^{HCT}/MK/mbondi	6.1	29.8	-3.5	8.8	0.351	27.5	93.2	-6.1	35.4	0.083	0.004
GB^{OBC1}/FF94/mbondi2	4.8	17.3	-4.0*	5.0	0.378	29.8	64.6	-3.0	36.4	0.090	0.013
GB^{OBC2}/FF94/mbondi2	4.3	16.0	-2.1	5.4	0.317	30.4	68.5	2.5	37.1	0.092	0.011
GBn/FF94/Bondi	3.6	10.9	-1.9	4.2	0.296	25.0	85.2	-0.9	35.8	0.073	0.019
GBn/FF03/Bondi	3.9	11.5	-1.4	4.8	0.252	27.0	85.8	-5.3	37.3	0.080	0.014
GBn/MK/Bondi	3.6	10.4	-1.9	4.0	0.281	25.1	85.2	-1.4	36.1	0.074	0.021
GBn/FF94/Tan2	5.9	13.6	-3.4*	6.3	0.496	26.3	86.4	-4.5	36.9	0.078	0.009
GBn/FF94/Tan1	3.5	10.4	-1.3	4.3	0.278	24.8	85.1	-0.4	35.7	0.073	0.018
GBn/FF94/Parse	3.7	11.3	-2.1	4.3	0.306	25.1	84.9	-1.1	35.8	0.074	0.018
Method-weighted average	1.5	7.1	0.3	2.4	0.095	18.4	34.5	11.6	18.4	0.052	
Unweighted average	2.7	9.2	-0.5	3.9	0.187	23.8	51.3	5.2	26.9	0.074	

* The MSD is significantly (at the 95% level) different from zero.

Table 3. The average, standard deviation, maximum, minimum, and range for all 24 methods or for the five PCM methods for the 20 drugs compared to the 20 neutral small molecules.

	All					PCM		
	Average	Stdev	Max	Min	Range	Average	Stdev	Range
061	-103.0	27.9	-55.9	-158.5	102.6	-60.3	3.2	8.7
caffeine	-55.2	12.0	-35.2	-83.1	47.9	-37.2	1.8	4.4
cannabidiol	-25.4	14.1	-0.4	-48.7	48.4	-5.0	3.1	8.2
TPD	-111.3	37.0	-69.7	-192.0	122.3	-83.1	3.6	9.7
efavirenz	-44.3	14.2	-19.7	-73.4	53.7	-22.1	2.1	5.2
ethoxyresorufin	-47.0	11.7	-29.3	-77.9	48.6	-34.5	3.0	7.0
Gal3[a]	-162.0	29.1	-117.5	-228.1	110.6	-129.5	11.6	25.8
MXA	-88.2	20.8	-50.8	-141.6	90.8	-72.6	4.4	11.5
omeprazole	-64.3	21.9	-29.6	-111.7	82.0	-32.0	1.8	3.8
paclitaxel[a]	-99.2	71.8	52.0	-197.8	249.9	46.3	6.7	15.5
paracetamol	-58.6	8.5	-36.9	-73.2	36.3	-51.7	2.3	6.0
phenolphtalein	-85.3	20.5	-45.6	-115.6	70.0	-53.4	5.9	15.4
phenylbutazone	-40.6	18.1	-5.4	-67.3	61.9	-9.2	2.5	6.1
progesterone	-41.0	9.3	-25.1	-59.4	34.2	-28.4	2.4	5.1
quercetin	-107.2	32.6	-64.3	-162.1	97.8	-71.9	7.1	18.2
R-thalidomide	-80.4	15.8	-56.3	-116.9	60.6	-58.7	2.2	5.6
SB3	-26.5	33.0	40.0	-60.2	100.2	36.1	2.6	6.9
S-warfarin	-70.7	21.5	-26.1	-97.8	71.7	-31.9	4.8	11.4
thiabendazole	-70.8	16.0	-44.7	-110.4	65.8	-60.0	2.3	5.8
Btn4	-86.5	9.5	-62.5	-110.3	47.8	-80.7	3.3	8.4
Average		22.3			80.1		3.8	9.4
H_2O	-34.3	11.4	-22.3	-56.2	33.8	-25.1	1.5	3.4
NH_3	-22.2	7.9	-8.0	-35.4	27.4	-17.8	0.7	1.7
CH_3OH	-21.6	6.4	-11.9	-34.7	22.8	-20.1	1.2	2.9
CH_3NH_2	-16.9	4.9	-5.6	-24.9	19.3	-19.4	0.7	1.7
CH_3CONH_2	-43.6	5.8	-33.3	-56.8	23.5	-38.5	1.4	3.4
CH_3SH	-6.5	2.2	-1.2	-10.3	9.1	-5.7	0.3	0.7
$CH_3C_6H_4OH$	-25.8	5.5	-12.6	-34.9	22.3	-24.9	1.9	4.8
CH_4	6.4	2.5	11.6	3.6	8.0	5.3	1.4	3.3
$CH_3CH_2CH_3$	8.9	1.7	12.7	7.3	5.4	8.2	0.1	0.3
isobutane	8.6	2.1	11.8	4.4	7.4	9.9	0.2	0.4
n-butane	9.3	1.7	13.0	7.4	5.6	8.8	0.1	0.3
CH_3CH_2OH	-18.4	6.1	-8.5	-29.1	20.6	-21.1	1.1	2.8
toluene	-5.8	3.0	0.6	-11.1	11.7	-3.8	1.0	2.5
$CH_3SCH_2CH_3$	-2.8	2.7	1.9	-7.6	9.5	-7.3	0.4	0.9
propionamid	-38.4	4.7	-27.5	-49.7	22.2	-37.3	1.4	3.3
3-methylindol	-25.2	5.5	-12.7	-38.2	25.5	-22.4	2.3	5.0
4-methyl-imidazole	-40.2	6.6	-25.0	-52.6	27.7	-40.0	2.2	5.3
N-propyl guanidine	-36.5	11.5	-7.9	-60.6	52.7	-40.8	3.0	8.2
CF_3CH_2OH	-26.4	5.5	-17.5	-35.6	18.1	-22.9	2.5	5.3
o-Cl-toluene	-6.1	3.2	-0.5	-14.8	14.3	-2.4	1.1	2.8
Average		5.0			19.3		1.2	2.9

[a] The DPCM/HF calculations failed for Gal3 and paclitaxel.

Table 4. MADs for the solvation free energies of the drug-like molecules for each method, calculated relative to the method-weighted average. The MADs were calculated either for the 20 neutral drug-like molecules, for all molecules in the three series (avidin, factor Xa, or galectin-3 inhibitors; called "All 3" in the Table), or individually for these three targets (avidin, FXa, and Gal). Moreover, the MADs were calculated for the relative solvation energies, either separately for the avidin, factor Xa, and galectin-3 inhibitors (Target), or separately for each total charge in these three series (Charge; in this case, lactose was omitted from the galectin-3 series).

Method	Neutral	3 Inhibitors				Target	Charge
	drugs	All 3	Avidin	FXa	Gal	All 3	All 3
DPCM/HF	7.1[a]	9.6[b]	4.4	13.4		3.4[b]	3.8[b]
IEFPCM/HF	9.6	7.7	3.0	11.7	5.9	3.2	2.6
IEFPCM/PBE0/HF	13.1	12.8	11.3	14.0	12.5	5.2	2.0
IEFPCM/PBE0/PBE0	10.7	14.5	6.8	23.8	6.4	5.1	3.4
IEFPCM/B3LYP	16.4	18.2	13.4	21.4	17.9	5.5	1.9
ChemSol	20.4	19.6	18.6	17.3	25.2	16.0	5.0
PB/FF94/Tan2	26.0	92.3	26.8	179.3	9.2	35.7	4.6
PB/FF94/Tan1	18.3	81.1	25.4	147.2	24.8	33.3	5.1
PB/FF94/mbondi	29.9	30.4	7.0	22.2	72.7	10.4	4.8
PB/FF94/Parse	47.1	54.8	19.7	63.3	80.1	7.5	5.4
PB/FF03/Tan2	26.6	101.5	37.0	191.9	11.0	35.2	9.2
PB/MK/Tan2	26.6	95.9	26.6	187.7	8.6	36.0	4.8
PB/MK/mbondi	32.0	30.3	7.4	20.5	74.8	10.3	4.5
GB^{HCT}/FF94/mbondi	27.7	29.7	10.6	33.1	45.9	11.1	4.3
GB^{HCT}/FF03/mbondi	27.4	22.8	6.6	23.9	39.5	10.5	4.5
GB^{HCT}/MK/mbondi	30.0	28.6	10.6	29.4	48.0	11.3	4.4
GB^{OBC1}/FF94/mbondi2	27.3	13.8	7.0	16.2	17.4	9.9	3.6
GB^{OBC2}/FF94/mbondi2	19.3	17.2	6.4	30.3	5.6	10.1	3.2
GBn/FF94/Bondi	28.6	13.3	5.1	19.3	11.7	6.2	3.7
GBn/FF03/Bondi	27.6	18.7	13.1	26.2	11.6	7.2	3.3
GBn/MK/Bondi	30.6	15.3	5.0	24.0	11.1	6.5	4.0
GBn/FF94/Tan2	50.6	35.1	17.4	50.6	27.4	8.0	5.6
GBn/FF94/Tan1	27.6	12.9	5.2	18.5	11.5	6.1	3.7
GBn/FF94/Parse	29.1	13.6	5.1	20.0	11.7	6.2	3.7

[a] Two of the calculations failed.
[b] Only for the biotin analogues and factor Xa inhibitors (two of which are missing).

Figure 1. The relation between the standard deviation of the various methods (either all or only the five PCM methods) and the average solvation energy (again either over all or over only the five PCM methods) for the 20 small molecules and the 20 drug-like molecules (paclitaxel and SB3 excluded with all methods).