Published in final edited form as:
Theor Chem Acc. 2012 March ; 131(3): 1161–. doi:10.1007/s00214-012-1161-7.

Optimization of the explicit polarization (X-Pol) potential using a hybrid density functional

Jaebeom Han, **Donald G. Truhlar**, and **Jiali Gao**
Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455-0431, USA

Abstract

The explicit polarization (X-Pol) method is a self-consistent fragment-based electronic structure theory in which molecular orbitals are block-localized within fragments of a cluster, macromolecule, or condensed-phase system. To account for short-range exchange repulsion and long-range dispersion interactions, we have incorporated a pairwise, empirical potential, in the form of Lennard-Jones terms, into the X-Pol effective Hamiltonian. In the present study, the X-Pol potential is constructed using the B3LYP hybrid density functional with the 6-31G(d) basis set to treat interacting fragments, and the Lennard-Jones parameters have been optimized on a dataset consisting of 105 bimolecular complexes. It is shown that the X-Pol potential can be optimized to provide a good description of hydrogen bonding interactions; the root mean square deviation of the computed binding energies from full (i.e., nonfragmental) CCSD(T)/aug-cc-pVDZ results is 0.8 kcal/mol, and the calculated hydrogen bond distances have an average deviation of about 0.1 Å from those obtained by full B3LYP/aug-cc-pVDZ optimizations.

Keywords

Explicit polarization; X-Pol; Quantum force field

1 Introduction

The basic approach used in the most popular current parametrizations of molecular mechanics (MM) was established in the 1960s by Lifson, and this approach continues to play an essential role in providing force fields for dynamical simulations of macromolecular systems such as proteins and nucleic acids [1] as well as other nanomaterials. Despite its success, which was promoted by careful and laborious parameterization by many research groups over the past half century, there are also a number of well-known shortcomings, including redundancy of energy terms and parameters, the widespread use of harmonic approximations for bond stretching and angle bending, and the difficulty of treating electronic polarization and charge transfer (for a recent special issue on polarizable force fields, see Jorgensen [2]). Furthermore, molecular mechanics is not designed to treat chemical reactions and photochemical processes [3]. With continuing advances in computer architecture, it is natural to ask what type of force fields will be used for biomolecular and materials simulations in the future. To this end, we have introduced the explicit polarization (X-Pol) potential [4–8], which is an electronic structure method based on block localization of molecular orbitals [4, 5, 9]. The X-Pol method differs from the effective fragment potential (EFP) [10, 11] and SIBFA (sum of interactions between fragments computed ab

gao@jialigao.org.

initio) [12] potentials in that the latter models are derived by fitting results to ab initio results in terms of a multipole expansion of the electrostatics along with other energy terms. In the X-Pol method, a macromolecular system is partitioned into constituent blocks, also called fragments, each of which can be, for example, an individual solvent or solute molecule, an enzyme cofactor, a ligand or molecular fragment, or a peptide unit of a protein. The internal energies of the fragments are determined by an explicitly quantum mechanical method, and interfragment interactions are approximated in a way akin to a combined quantum mechanical and molecular mechanical (QM/MM) [3, 13–15] method. However, the electrostatic field in which each individual fragment is embedded is obtained from the corresponding instantaneous wave functions of all other fragments in the system, and the mutual electronic polarization among fragments is included self-consistently [4–8].

X-Pol can also be used as an electronic structure method such that any quantum chemical model, e.g., Hartree–Fock (HF) theory (or semiempirical models of HF), second-order Møller-Plesset perturbation theory (MP2), coupled cluster theory, or density functional theory (DFT), may be adopted to represent the individual fragment blocks. In this regard, one can treat all fragments by using the same method, or by mixing different electronic structure methods for different fragments (for example, MP2 for one fragment and DFT for all other fragments). Because a large system is partitioned into fragments, the X-Pol method can be made to scale well for fast calculations, and therefore, it can be used to establish a framework for the development of a next-generation force field [4] that goes beyond the conventional molecular mechanics by explicitly including a quantum mechanical treatment of electronic polarization and possibly charge transfer effects (which can be included, for example, by a recently proposed method [16] involving ensemble DFT). When X-Pol is used as a force field, we introduce a set of empirical terms to account for the missing exchange repulsion [17] and dispersion-like attractive, noncovalent interactions. Because these terms are empirical, they can increase the accuracy and, at the same time, reduce computational costs by using parameterization to compensate for errors introduced by using a low or modest level of electronic structure theory [8]. In the present study, we illustrate this by showing how we can use a modestly accurate density functional with a small basis set to treat the individual fragments in the X-Pol method. In particular, we employ the hybrid B3LYP model and a fairly small 6-31G(d) basis set, and we show that X-Pol with this choice can be parameterized to model hydrogen bonding interactions in good agreement with the results from full CCSD(T) calculations. Here, we emphasize that our goal is not to reproduce the geometries and energies at the B3LYP/6-31G(d) level that is used to represent the X-Pol fragment, but rather to obtain agreement with the higher-level CCSD(T) results by optimization of the parameters introduced in the X-Pol quantum force field [5, 8].

There are many other fragment-based molecular orbital methods [18]. For example, Zhang et al. [19, 20] developed a molecular fractionation with conjugated caps (MFCC) approach to treat proteins and protein–ligand interactions. In this method, the individual fragments are capped with a structure representative of the local structure of the original system, and the total energy is obtained by subtracting the energies of the common fragments used in the “caps”. The method provides a good means to evaluate interfragment interactions and a straightforward procedure to incorporate the local electronic structure into a fragment-based molecular orbital approach [21, 22]. Another way of separating the total energy into fragmental contributions is the general interaction energy expansion approach described by Stoll and Preuss [23]. The key to achieve fast convergence in this method, in contrast to early schemes [24], is to optimize the monomer, dimer and many-body fragmental molecular orbitals in the presence of all other fragments, rather than using isolated gas-phase fragment terms. There are a number of applications of this strategy, including the fragment molecular orbital (FMO) method [25, 26] and the electrostatically embedded many-body (EE-MB) expansion method [27–30]. The SCF procedure used in the FMO model is

identical to that developed in the X-Pol method [25, 26], whereas two-body and three-body exchange and charge transfer effects are included in the FMO2 and FMO3 implementations [23].

In Sect. 2, we briefly review the theoretical background of the X-Pol potential, and in Sect. 3, we present the computational details. In Sect. 4, we describe the optimization of parameters and compare the computed hydrogen bonding energies and geometries obtained from the X-Pol method with higher-level results. Finally, Sect. 5 summarizes the main findings from this work and presents concluding remarks.

2 Theoretical background

The X-Pol method has been described in detail elsewhere [4, 7, 8]. For completeness, we briefly describe the key aspects and approximations made in the X-Pol potential and the empirical parameters introduced to correct for these approximations. We note that the X-Pol method was developed based on block localization of the molecular wave function of the system, which includes a hierarchy of approximations [8]. There are numerous other methods based on localized molecular orbitals or molecular fragments. A recent review, which appeared online after the submission of this manuscript, contains an account of these methodologies [18].

First, we partition a macromolecular system into structural blocks, also called fragments. The molecular wave function, Φ, is approximated as a Hartree product of antisymmetric wave functions of the individual fragments, $\{\Psi^A\}$:

$$\Phi=\prod_{A=1}^{N}\Psi^A \tag{1}$$

where N is the number of fragments in the system, and Ψ^A is a Slater determinant of occupied molecular orbitals (MOs) that are constructed using an atomic orbital basis located on the atoms of fragment A. Thus, these MOs are block-localized by construction. In the present work, density functional theory is used to represent the molecular fragments, and the block-localized molecular orbitals (BLMO) are block-localized Kohn–Sham (BLKS) orbitals, in terms of which the electron density $\rho^A(\mathbf{r})$ of fragment A is given by

$$\rho^A(\mathbf{r})=2\sum_i|\phi_i^A|^2 \tag{2}$$

where ϕ_i^A is the ith doubly occupied Kohn–Sham orbital of fragment A. In the present work, the molecular fragments are closed-shell molecules.

The X-Pol total energy of the system can be written as follows [4, 8]:

$$E_{\text{tot}}=\sum_{A=1}^{N}\left(E^A+\frac{1}{2}E_{\text{int}}^A\right)+E_{\text{XD}} \tag{3}$$

where E^A is the energy of fragment A with the wave function Ψ^A, which can be calculated at any given theoretical level, including Hartree–Fock (HF), density functional theory (DFT), or post-HF theories such as Møller-Plesset perturbation theory or coupled cluster theory, E_{int}^A is the Coulomb interaction energy between fragment A and other fragments, and E_{XD} accounts for the exchange-repulsion (X) and dispersion-correlation (D) interactions between the fragments. It should be pointed out that the wave function Ψ^A in Eq. 3

corresponds to that of fragment A polarized by the remaining fragments in the system, and it differs from the wave function of an isolated fragment in the gas phase (Φ_0^A). The energy difference between the two states, Ψ^A and Φ_0^A, is the energy penalty paid for distorting the fragmental wave function due to many-body polarization [13].

The use of the Hartree-product wave function in Eq. 1 implies that the short-range exchange repulsion and long-range and medium-range dispersion and dispersion-like interactions (for brevity, will just call these dispersion in the rest of the article) as well as charge transfer among fragments are neglected [4]. The exchange repulsion and dispersion energies can be determined in various ways, for example by antisymmetrizing the block-localized (i.e., fragmental) orbitals in Eq. 1 [9, 17, 31, 32] or by perturbation methods such as symmetry-adapted perturbation theory (SAPT) [33–35]. However, these methods are not suitable for the construction of a fast quantum mechanical force field for large systems due to their high computational cost as compared to the method adopted here, which is discussed next.

Because the exchange repulsion is short-ranged and approximately pairwise additive [17] and the dispersion interactions can also be adequately modeled by pairwise potentials [4], such as those used in dispersion-corrected density functional theory (DFT-D) [36], we have used the Lennard-Jones potential to parametrically model the exchange-repulsion and dispersion interactions between each pair of fragments, A and B:

$$E_{\mathrm{ErD}}^{AB}=4\sum_{a}^{B}\sum_{b}^{B}\varepsilon_{ab}^{AB}\left[\left(\frac{\sigma_{ab}^{AB}}{R_{ab}}\right)^{12}-\left(\frac{\sigma_{ab}^{AB}}{R_{ab}}\right)^{6}\right] \tag{4}$$

where the A over the sum means that the sum is restricted to orbitals a on center A, and where ε_{ab}^{AB} and σ_{ab}^{AB} are parameters. These parameters are determined from atomic parameters by using standard combining rules: $\varepsilon_{ab}^{AB}=\sqrt{\varepsilon_a^A\varepsilon_b^B}$ and $\sigma_{ab}^{AB}=\sqrt{\sigma_a^A\sigma_b^B}$. The values of ε_a^A and σ_a^A depend on the atomic number of the atom and sometimes also on its hybridization. These parameters can be optimized for a particular electronic structure method used in the X-Pol potential, and the main objective of the present study is to illustrate the optimization of these parameters and the performance of the X-Pol potential with the B3LYP hybrid density functional and the modest 6-31G(d) basis set for calculating binding energies of bimolecular complexes. We will judge the accuracy by comparing to the results of higher-level CCSD(T) calculations.

For closed-shell fragments, the Kohn–Sham DFT energy of fragment A in the presence of the rest of the system is

$$E^A\left[\rho^A(\mathbf{r})\right]=\sum_i 2H_i^A+\sum_{i,j}2J_{ij}^A+E_{xc}^A\left[\rho^A(\mathbf{r})\right]+E_{\mathrm{nuc}}^A \tag{5}$$

where the superscript A labels the energies and densities for monomer fragment A, the indices i and j run through the doubly occupied, BLKS molecular orbitals of fragment A, H_i^A and J_{ij}^A are respectively the one-electron Hamiltonian integrals and the Coulomb integrals, $E_{\mathrm{xc}}^A\left[\rho^A(\mathbf{r})\right]$ is the exchange–correlation functional, and E_{nuc}^A is the nuclear repulsion energy.

The Coulomb interaction energy, E_{int}^A, in Eq. 3 is given by:

$$E_{\text{int}}^{A} = -2\sum_{i} <\psi_i^A | V^A(\mathbf{r}) | \psi_i^A > + \sum_{a}^{A} Z_a^A V^A\left(\mathbf{R}_a^A\right) \tag{6}$$

where Z_a^A is the nuclear charge of atom a of fragment A, and $V^A(\mathbf{r})$ is the total external electrostatic potential (ESP) due to all other fragments in the system. The external ESP is defined as follows:

$$V^A(\mathbf{r}) = -\sum_{B \neq A} \left(\int \frac{\rho^B(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \sum_{b}^{B} \frac{Z_b^B}{|\mathbf{r}-\mathbf{R}_b^B|} \right) \tag{7}$$

where Z_b^B is the nuclear charge of atom b of fragment B located at $\boldsymbol{R}_b^B$. The potential $V^A(\mathbf{r})$ could be determined analytically and used to compute the two-electron integrals in Eq. 6 (the terms in the first summation), but this is time-consuming and not a useful choice for fast calculations on macromolecular systems. Alternatively, the ESP in Eq. 7 can be treated by a distributed multipole expansion, and the simplest approximation is to retain only the distributed monopole terms, possibly with scaling to make up for this approximation. This yields the following expression.

$$V^A(\mathbf{r}) = \lambda \sum_{B \neq A} \sum_{b}^{B} \frac{q_b^B}{|\mathbf{r} - \mathbf{R}_b^B|} \tag{8}$$

where λ is a scaling parameter, and q_b^B is the partial atomic charge on atom b of fragment B. The partial atomic charges can be determined in various ways, for example by fitting electrostatic potentials or by using a charge population analysis method [5]. We make the latter choice [8] for the present calculations, using Mulliken population charges from the BLKS orbitals, and the single parameter λ is set to unity.

In calculating the total energy of the system by Eq. 3, we use the double self-consistent-field (DSCF) method [6, 7]. With an initial guess of the one-electron density matrix for each fragment, the electronic structure calculations for each fragment are performed in the presence of the Mulliken charges of all the other fragments until the change in the total electronic energy or density matrix reaches a predefined tolerance. Although the X-Pol theory has been formulated variationally to allow efficient calculations of energy gradients [7], here we use the older, nonvariational sequential optimization energy formulation.

3 Computational details

For the bimolecular complexes, in both the “high-level” reference calculations and the X-Pol calculations, partial geometry optimizations were performed in which the monomer geometries are held fixed at the corresponding level of theory. Thus, in each bimolecular complex, the hydrogen bond distance and angle between the donor and acceptor molecules, as illustrated in Figs. 1, 2, 3, 4 and 5, are optimized. (Only one angle is involved because we adopt a high symmetry for each hydrogen bond, as illustrated). In all cases, the monomer geometries that were optimized at the corresponding level of theory were held fixed.

The reference geometries were calculated by full (i.e., nonfragmental) B3LYP [37–39] calculations with the aug-cc-pVDZ basis set. The reference energies were obtained by full CCSD(T) single-point calculations with the aug-cc-pVDZ basis set at the geometries of the complexes optimized using B3LYP/aug-cc-pVDZ.

The X-Pol calculations were carried out using the B3LYP/6-31G(d) method as the quantum mechanical level with the geometries optimized by the same level of X-Pol calculation. The hybrid B3LYP functional was chosen in the present study because it is a popular model that has been used widely; one can certainly select a more accurate and recent functional, but the goal here is not to compare the quality and performance of different functionals. We sometimes use the notation XP@B3LYP/6-31G(d) to specify such an X-Pol calculation.

The binding energy for a bimolecular complex, including the empirical Lennard-Jones terms to account for the exchange repulsion and dispersion contributions, is calculated by:

$$\Delta E_b(A\cdots B) = E(A) + E(B) - E_{\text{tot}}^{\text{X-Pol}}(A\cdots B) \tag{9}$$

where $E_{\text{tot}}^{\text{X-Pol}}(A\cdots B)$ is the X-Pol energy (Eq. 3) of the bimolecular complex in which each monomer, A or B, is treated as an individual fragment, and $E(A)$ and $E(B)$ are the B3LYP/6-31G(d) energies of the optimized monomer structures. All binding energies in the present article are zero-point-exclusive.

The Lennard-Jones parameters in Eq. 4 have been adjusted so that the XP@B3LYP/6-31G(d)-binding energies best reproduce the results calculated using CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ.

The full quantum mechanical calculations for all systems were performed using *Gaussian03* [40], whereas all X-Pol calculations were carried out using a local program that is coupled to a modified version of the GAMESS package [41].

4 Results and discussion

4.1 Optimization of the repulsion and dispersion interactions between fragments

We considered a total of 105 bimolecular complexes, each of which involves one water molecule and an organic or inorganic compound or ion; the organic compounds include ionic and neutral functional groups found in amino acids and nucleobases. Although experimental results for a number of hydrogen bonding complexes are available, we wish to examine a much larger dataset, and therefore, we used theoretical results as reference data, as explained in Sect. 3. The reference data for these complexes were used to optimize the Lennard-Jones parameters by an iterative procedure for the case where X-Pol fragments are treated by B3LYP/6-31G(d). In this process, we placed greater emphasis on the performance for binding energies than on hydrogen bond distances and angles. The resulting parameters for H, C, O, N, and S atoms and for F^-, Cl^-, and Na^+ ions are listed in Table 1.

Previously, we examined a small set of 14 bimolecular complexes and reported a set of parameters for several atoms [8], and the values listed in Table 1 for these atoms are very similar to those obtained in that work. In the present work, we introduced a new atom type for hydrogen attached to a sulfur atom (thiols and H_2S), and this atom type has a greater σ value than that used in other situations. Three atom types are assigned to oxygen, corresponding to an oxygen type in neutral functional groups and two types for anionic species. Previously, different Lennard-Jones parameters were used for sp^2 and sp^3 oxygen atoms [8], but a single oxygen type for both hybridizations is adequate here. The Lennard-Jones parameters for the carboxylate oxygen and neutral oxygen atoms are very similar; although it would be possible to use the same oxygen parameters in both cases, we kept the two atom types to increase flexibility. For nitrogen atoms, we distinguish atom types for neutral and protonated cases. For other elements, including carbon, a single set of parameters for each is sufficient for the present data set.

4.2 Energies and geometries of hydrogen bonded complexes

Figures 1, 2, 3, 4 and 5 depict the structural arrangements used in the present calculations, along with the optimized geometrical parameters from both the reference calculations and the X-Pol calculations. In many cases, more than one structure is considered for a given chemical species, corresponding to placing water molecules at different positions or in different orientations. Each structure is assigned a number for discussion purposes, and the computed binding energies are given in Tables 2, 3, 4, 5 and 6. The figures and tables are organized roughly according to functional groups.

Both the B3LYP/aug-cc-pVDZ and CCSD(T)/aug-cc-pVDZ-binding energies are given in the tables for comparison; however, the B3LYP-binding energies are inaccurate due to a poor treatment of dispersion contributions, and only the CCSD(T) values should be considered as reference values.

4.2.1 Small molecules and simple functional groups—Figure 1 and Table 2 give the results for water complexes with small molecules. Six sp^3 oxygen structures are included in our study, including water, two structures of methanol, and three structures of ethers (structures **1–6**).

The interaction energy for a water dimer (**1**) is calculated to be 5.7 kcal/mol by the X-Pol method, which yields a hydrogen bond length and bond angle of 1.93 Å and 135.5° (Fig. 1). These may be compared with the corresponding reference values of 5.2 kcal/mol, 1.91 Å, and 138.9°. The best estimate of the water dimer interaction energy is 5.0 kcal/mol using CCSD(T) with extrapolation to a complete basis set [42, 43].

For the methanol–water complexes (**2** and **3**), both XP@B3LYP/6-31G(d) and CCSD(T) calculations predict that methanol is a better hydrogen bond acceptor (structure 3) by 0.4–0.5 kcal/mol. As the number of alkyl groups on the oxygen increases in going from water to alcohol to ether, the calculated hydrogen bond strength is also enhanced, due to the electron donating effect of an alkyl group, to a final value of about 7 kcal/mol for the complex with tetrahydrofuran (**6**). The average unsigned errors in hydrogen bond lengths and angles are, respectively, 0.02 Å and 7° for the sp^3 oxygen-containing compounds.

The binding energies of methane (**7**) and benzene (**9**) with a single water molecule from the X-Pol optimizations are 0.3–0.4 kcal/mol greater than the CCSD(T) results, but the binding energy between ethane and water is 0.6 kcal/mol smaller (**8**).

We examined five complexes involving simple methyl amines (**10–14**). The primary and secondary amines are much better hydrogen bond acceptors than donors [44], both from the XP@B3LYP/6-31G(d) and CCSD(T)/aug-cc-pVDZ calculations. Although the X-Pol-binding energies for the donor complexes (**10** and **12**) are in good agreement with the reference data, the binding energies for the acceptor complexes are underestimated by 1.5–2.2 kcal/mol, and the deviation increases as the basicity of the amines increases with more methyl substitutions.

Figure 1 includes four sulfur compounds: hydrogen disulfide, methanethiol, dimethyl sulfide, and dimethyl disulfide (**15**–**21**). Similar to alkyl amine complexes, the binding energies for sulfur compounds are stronger when the sulfur atom acts as a hydrogen bond acceptor than a donor for H_2S and thiols; however, the difference is smaller than in the corresponding nitrogen compounds. These trends are correctly reproduced in the XP@B3LYP/6-31G(d) model in comparison with the reference data (Table 2). In fact, the X-Pol method performs very well, having an average unsigned error of less than 0.2 kcal/mol in binding energy. A somewhat less satisfactory finding is that an additional hydrogen

atom type for H_2S and thiols needs to be introduced, whereas the hydrogen bond distance for the acceptor complexes is significantly shorter in the X-Pol calculations than the values optimized using B3LYP/aug-cc-pVDZ.

4.2.2 Carbonyl-containing compounds—Figure 2 and Table 3 list results for carbonyl compounds, including aldehydes, ketones, carboxylic acids, esters, and amides. The X-Pol results for the aldehydes and acetone (**22–25**) are in reasonable agreement with the reference data, resulting in an average unsigned deviation in binding energy of 0.4 kcal/mol. It is especially encouraging that the X-Pol model correctly distinguishes the relative interaction energies between the two complexes of acetaldehyde with water, favoring the configuration with water oriented toward the methyl group. The optimized hydrogen bond geometries using XP@B3LYP/6-31G(d) are also in good agreement with the reference data.

For carboxylic acid systems, both the syn and anti conformations are considered (**26–33**). We note that the carboxylic acids are particularly good hydrogen bond donors, with computed binding energies of 7.6 and 8.4 kcal/mol for the syn (**28**) and anti (**33**) conformations of acetic acid using XP@B3LYP/6-31G(d), which may be compared with the reference values of 7.9 and 8.2 kcal/mol. The interaction energies on the carbonyl sites are of similar magnitude as those found in aldehyde and ketone complexes. The hydrogen bond accepting ability of the hydroxyl oxygen is relatively weak (3.3 kcal/mol) in the syn (**29**) configuration, while the structure in the anti conformer (**32**) enjoys a secondary hydrogen bonding interaction [45] to the carbonyl oxygen, increasing the XP@B3LYP-binding energy to 5.6 kcal/mol. The latter is 0.9 kcal/mol greater than the reference energy. The optimized geometrical parameters are also in excellent agreement between the two computational approaches. Analogously, both the syn and anti conformations for methyl formate are considered (**34–39**), and similar trends as the corresponding acids are found for these complexes. Overall, the average unsigned errors for all acid and ester complexes are just under 0.3 kcal/mol in binding energy and 0.1 Å in hydrogen bond distance.

The amides complexes are structures **40–49**. For the formamide-water complexes (**40–43**), the XP@B3LYP/6-31G(d) method predicts that the carbonyl group is a better hydrogen bond acceptor (**40** and **41**) than the amide group as a hydrogen bond donor (**42** and **43**) in agreement with the reference data, and the same trend is found in the *N*-methyl formamide and *N*-methyl acetamide complexes. However, for the donor complexes, XP@B3LYP/6-31G(d) yields a larger binding energy for **43** than complex **42**, due to the alignment of the carbonyl dipole in the direction of the N–H bond, but the opposite is found in the reference data. In the full QM calculation, there is apparently an overlap interaction between water and the carbonyl group [46], suggested by the smaller hydrogen bond angles (**42**). This is absent in the present X-Pol method [47]. In all amide complexes, the optimized hydrogen bond lengths using XP@B3LYP/6-31G(d) are in excellent agreement with those at the B3LYP/aug-cc-pVDZ level.

4.2.3 Heterocyclic compounds—We considered a number of heterocyclic compounds, which are displayed in Fig. 3, and the corresponding interaction energies are given in Table 4. For both imidazole and pyridine complexes with water, the XP@B3LYP/6-31G(d) method yields weaker binding energies by about 1 kcal/mol than the corresponding reference data. In these cases, the hydrogen bond lengths from the XP@B3LYP/6-31G(d) optimization are about 0.1 Å longer than the B3LYP/aug-cc-pVDZ results. However, good agreement is obtained between the X-Pol- and CCSD(T)-binding energies for the remaining heterocyclic compounds, even though we restricted the nitrogen atom type to just one for all neutral compounds.

Structures **53–72** represent hydrogen bonding interactions between a water molecule and the functionalities of nucleobases; cytosine and uracil are depicted in **59–64** and **69–72**, respectively, whereas only the six-member-ring portions of guanine (**53–58**) and adenine (**65–68**) are studied. In these complexes, each organic compound contains both hydrogen bond donor sites and acceptor sites, and the latter can be either an oxygen or a nitrogen atom. Thus, these species cover a large range of hydrogen bonding strengths. Although there are some variations, on average, the binding energies are about 1.3 kcal/mol (XP@B3LYP) and 1.4 kcal/mol (CCSD(T)) larger for structures that accept a hydrogen bond than for those that donate one to water. An exception is found for the two carbonyl groups in uracil, which have binding energies of just over 5 kcal/mol, without which the above difference would be even greater (2 kcal/mol). The binding energies for several complexes are particularly strong, with values greater than 9 kcal/mol (**57**, **60**, and **62**); this can be attributed to contributions from secondary hydrogen bonding interactions [45]. The X-Pol method correctly reproduces these features, in good agreement with the reference binding energies; however, structure **60** is predicted to have a stronger hydrogen bond than **62** from CCSD(T) calculations, but the opposite is obtained using XP@B3LYP. For the whole set of 23 heterocyclic complexes, the mean unsigned error in binding energy is 0.5 kcal/mol, and the differences in hydrogen bond distance are all under 0.1 Å.

4.2.4 Ions—Complexes involving anions and cations, ranging from simple monatomic ions to delocalized organic species, are shown in Figs. 4 and 5, and the corresponding binding energies are given in Tables 5 and 6. The binding energies for the monatomic F^- (**73**) and Cl^- (**74**) ions are fitted in exact agreement with the reference data, but the hydrogen bond distances are 0.19 and 0.14 Å longer in the X-Pol model than the reference values.

For the oxyanions, we consider hydroxide ion, (**75**) alkoxide ions (**76–78**), and conjugated species (**79–80**). In these complexes, there is strong electronic overlap between the two fragments, particularly for the smaller ions. Thus, the block localization of the fragment orbitals in the X-Pol method tends to introduce greater errors as reflected in the hydroxide–water complex (**75**), for which the X-Pol-binding energy is 4.5 kcal/mol greater than in the CCSD(T) calculations. The agreement for the larger and delocalized oxyanions is much improved, with an average error of 1.3 kcal/mol.

A total of twelve carboxylate–water plus allyl anion–water complexes are shown in Fig. 4. We found that a different set of Lennard-Jones parameters than those used for the alkoxide anions has to be adopted for the carboxylate anions, perhaps due to the more electron-delocalized nature of the carboxylate group. The binding energy ranges from 14.7 to 18.3 kcal/mol from CCSD(T) calculations. The agreement between these results and the XP@B3LYP ones is generally good with a mean unsigned deviation in binding energy of 0.3 kcal/mol. The XP@B3LYP/6-31G(d) method correctly predicts that the bifurcated forms of the complexes with water (**81**, **84** and **89**) are the most stable in each case [48], and the differences from the least stable complexes are from 2.3 to 2.7 kcal/mol. This agrees with CCSD(T) calculations except that the least stable complexes are reversed for structures **85** and **88** in the XP@B3LYP/6-31G(d) method.

The computed binding energies for cation–water complexes are given in Table 6. For the ammonium ions, the single-site hydrogen bonding complex (**94**, **96** and **98**) yields stronger interactions than the symmetric two-site structure (**95**, **97** and **99**) both from the XP/B3LYP and CCSD(T) models. However, the reference calculations predict that the latter complexes are 0.7–0.9 kcal/mol more stable than the predictions of the X-Pol method. In the alkyl ammonium series, binding energies decrease progressively as the number of methyl substituents increases, primarily due to charge delocalization. For imidazolium and pyridinium ions, the agreement between XP@B3LYP and CCSD(T) is also reasonable,

although the hydrogen bond distances from the X-Pol optimizations are 0.08 Å longer. Two structures are considered for the guanidium ion–water complex. In this case, the energy difference between the two structures is predicted to be smaller than that from CCSD(T) calculations, but the average of the two binding energies is consistent with the ab initio data. Finally, two carbocations are considered, both of which are found to be adequately modeled by the present X-Pol potential.

4.3 Overall assessment

The performance of the present XP@B3LYP/6-31G(d) method, based on comparisons to the results of CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ calculations, is shown in Fig. 6. Overall, the root mean square deviation (RMSD) in binding energy between the XP@B3LYP/6-31G(d) predictions and the CCSD(T)/aug-ccc-pVDZ//B3LYP/aug-cc-pVDZ reference data for all 105 bimolecular complexes, covering a range of binding energies of more than 20 kcal/mol, is 0.8 kcal/mol. For a similar set of bimolecular systems, combined QM/MM calculations using the AM1 [13, 49] and HF/3-21G [50] methods along with the three-point charge TIP3P model for water yielded RMSDs of 1.2 and 0.5 kcal/mol, respectively. In those studies, the optimization target was obtained from HF/6-31 + G(d) calculations. In other studies, Riccardi et al. [51] calculated the binding energies for a series of bimolecular complexes of water and organic compounds representing amino acid side chains using the SCC-DFTB/CHARMM potential and obtained an RMSD of 1.2 kcal/mol with respect to B3LYP/6-311 ++G(d,p)//B3LYP/6-31 +G(d) dataset. Freindorf et al. [52] obtained an RMSD of 1.5 kcal/mol for small organic molecule/water complexes using the B3LYP/6-31 +G(d)/AMBER potential. The present X-Pol potential, making use of B3LYP/6-31G(d) for each fragment, yields slightly better agreement with the target dataset than several combined QM/MM methods that employed the same strategy of optimizing the van der Waals parameters. Although the origin of the good performance of the X-Pol method needs to be more carefully investigated, a main difference from these QM/MM approaches is that the mutual electronic polarization effects between the two monomers in each complex are included in X-Pol, whereas fixed charge MM force fields are used in the QM/MM calculations.

The RMSD of the optimized hydrogen bond distance between the XP@B3LYP/6-31G(d) optimization and that of the B3LYP/aug-cc-pVDZ method is 0.13 Å (Fig. 7), which is slightly greater than a value of 0.08 Å for a much smaller set of 14 structures in a previous study [8]. This may also be compared with the AI-3/MM [50], SCC-DFTB/MM [51], and B3LYP/6-31 +G(d)/AMBER methods [52], which have RMSDs of 0.07 Å, 0.08 Å, and 0.11 Å relative to the respective datasets. The RMSD errors for hydrogen bond angles are 12° when comparing XP@B3LYP/6-31G(d) to B3LYP/aug-cc-pVDZ optimizations (Fig. 8). Although the errors in the optimized hydrogen bond angles are relatively large, the potential energy surfaces for these interactions are typically flat and they do not affect the binding energies significantly. Overall, the present XP@B3LYP/6-31G(d) model yields reasonable hydrogen bond geometries for a variety of organic functional groups interacting with a water molecule.

5 Concluding remarks

The X-Pol potential uses an electronic structure method to model the mutual polarization effects between interacting fragments. It employs an effective Hamiltonian to model energy contributions from interactions between fragments beyond the electrostatic ones that are included self-consistently in the quantum mechanical fragment calculations. These additional interactions are dominated by exchange repulsion and dispersion energies [8, 17], which are approximated by the Lennard-Jones model [4]. The parameters of the Lennard-Jones function can be optimized to reproduce experimental data or accurate results from

high-level calculations; in the present study, the CCSD(T)//aug-cc-pVDZ-binding energies and the optimized hydrogen bond distances and angles using B3LYP/aug-cc-pVDZ are chosen as the optimization target. Implicitly, the X-Pol results obtained with the optimized Lennard-Jones parameters for different atom types account for the energy component due to charge transfer in a way analogous to molecular mechanics force fields. If the specific description of charge transfer effects is important for a given problem, the fragment partitions need to be assigned in such a way that the electron donor and acceptor groups are both included in the same fragment. Alternatively, resonance charge delocalization effects can be modeled by the multiconfigurational, generalized X-Pol (GX-Pol) theory highlighted recently [47, 53] or by using ensemble DFT [16]. Here, however, we tested the simpler approach in which charge transfer is only implicit. The goal is to develop and test a computational approach involving a modest computational cost that can be applied for fast calculations on large systems.

The explicit polarization (X-Pol) method is based on block localization of molecular orbitals within each fragment. The fragments can be assigned, for example, as individual molecules such as solvent molecules or as or amino acid residues in a protein. If desired, important portions of the system can be treated as single large fragments, for example, one may take an entire active site of an enzyme as a single fragment. A key feature of the X-Pol method is that the block-localized orbitals of each fragment are optimized in the presence of the instantaneous electric field due to all other fragments, and the mutual polarization among all fragments is determined using self-consistent field methods [6, 7]. When the electronic integrals between different fragments are approximated by an external potential expansion, the computational costs can be greatly reduced [4, 8]. Furthermore, the fragment block-localization scheme naturally leads to linear scaling in electronic structural calculations on large systems. The X-Pol approach provides a theoretical framework for developing next-generation force fields for macromolecular simulations using an explicitly quantum mechanical electronic structure theory [54].

The present calculations employ the B3LYP hybrid density functional, and the Lennard-Jones parameters are optimized to higher-level reference data for a dataset containing 105 bimolecular, trimolecular, tetramolecular, pentamolecular, and heptamolecular complexes between one or more water molecules and an organic or inorganic compound or ion, representing the functional groups of amino acids and nucleobases. We found that the average deviation between the binding energies calculated by the XP@B3LYP/6-31G(d) and the CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods is about 0.8 kcal/mol, whereas the deviation in hydrogen bond distance is about 0.1 Å. It will be interesting to further test this kind of model through condensed-phase simulations, including the computation of liquid properties and solvation free energies.

Acknowledgments

We thank Dr. Yen-lin Lin for assistance. This work has been supported by the National Institutes of Health (RC1-GM091445 and GM46736) and the National Science Foundation (CHE09-56776 and CHE09–57162).

References

1. Levitt M. Nat Struct Biol. 2001; 8:392. [PubMed: 11323711]
2. Jorgensen WL. J Chem Theory Comput. 2007; 3:1877.
3. Gao J. Acc Chem Res. 1996; 29:298.
4. Gao J. J Phys Chem B. 1997; 101:657.
5. Gao J. J Chem Phys. 1998; 109:2346.
6. Xie W, Gao J. J Chem Theory Comput. 2007; 3:1890. [PubMed: 18985172]

7. Xie W, Song L, Truhlar DG, Gao J. J Chem Phys. 2008; 128 234108/1.
8. Song L, Han J, Lin YL, Xie W, Gao J. J Phys Chem A. 2009; 113:11656. [PubMed: 19618944]
9. Mo Y, Gao J, Peyerimhoff SD. J Chem Phys. 2000; 112:5530.
10. Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Krauss M, Garmer D, Basch H, Cohen D. J Chem Phys. 1996; 105:1968.
11. Gordon MS, Slipchenko L, Li H, Jensen JH. Ann Rep Comput Chem. 2007; 3:177.
12. Gresh N, Cisneros GA, Darden TA, Piquemal J-P. J Chem Theory Comput. 2007; 3:1960. [PubMed: 18978934]
13. Gao J, Xia X. Science. 1992; 258:631. [PubMed: 1411573]
14. Lin H, Donald G Truhlar. Theor Chem Acc. 2007; 117:185.
15. Senn HM, Thiel W. Angew Chem Int Ed. 2009; 48:1198.
16. Isegawa M, Gao J, Truhlar DG. J Chem Phys. 2011; 135 (in press).
17. Cembran A, Bao P, Wang Y, Song L, Truhlar DG, Gao J. J Chem Theory Comput. 2010; 6:2469. [PubMed: 20730021]
18. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV. Chem Rev. 2012; 112:632. [PubMed: 21866983]
19. Zhang DW, Xiang Y, Zhang JZH. J Phys Chem B. 2003; 107:12039.
20. Xiang Y, Zhang DW, Zhang JZH. J Comput Chem. 2004; 25:1431. [PubMed: 15224387]
21. Duan LL, Mei Y, Zhang DW, Zhang QG, Zhang JZH. J Am Chem Soc. 2010; 132:11159. [PubMed: 20698682]
22. Tong Y, Mei Y, Li YL, Ji CG, Zhang JZH. J Am Chem Soc. 2010; 132:5137. [PubMed: 20302307]
23. Stoll H, Preuss H. Theor Chem Acc. 1977; 46:12.
24. Hankins D, Moskowitz JW, Stillinger FH. J Chem Phys. 1970; 53:4544.
25. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M. Chem Phys Lett. 1999; 313:701.
26. Fedorov DG, Kitaura K. J Phys Chem A. 2007; 111:6904. [PubMed: 17511437]
27. Truhlar DG, Dahlke EE. J Chem Theory Comput. 2007; 3:1342.
28. Truhlar DG, Dahlke EE, Leverentz HR. J Chem Theory Comput. 2008; 4:33.
29. Tempkin JOB, Leverentz HR, Wang B, Truhlar DG. J Phys Chem Lett. 2011; 2:2141.
30. Leverentz HR, Truhlar DG. J Chem Theory Comput. 2009; 5:1573.
31. Y, Mo; Peyerimhoff, SD. J Chem Phys. 1998; 109:1687.
32. Mo Y, Gao J. J Comput Chem. 2000; 21:1458.
33. Jacobson LD, Herbert JM. J Chem Phys. 2011; 134:094118. [PubMed: 21384961]
34. Jeziorski B, Moszynski R, Szalewicz K. Chem Rev (Washington, DC). 1994; 94:1887.
35. Misquitta AJ, Podeszwa R, Jeziorski B, Szalewicz K. J Chem Phys. 2005; 123:214103. [PubMed: 16356035]
36. Grimme S. J Comput Chem. 2006; 27:1787. [PubMed: 16955487]
37. Lee C, Yang W, Parr RG. Phys Rev B. 1988; 37:785.
38. Becke AD. J Chem Phys. 1993; 98:5648.
39. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. J Phys Chem. 1994; 98:11623.
40. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, J.; J, A.; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; 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, JE.; Hratchian, HP.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Ayala, PY.; Morokuma, K.; Voth, GA.; Salvador, P.; Dannenberg, JJ.; Zakrzewski, VG.; Dapprich, S.; Daniels, AD.; Strain, MC.; Farkas, O.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Ortiz, JV.; Cui, Q.; Baboul, AG.; Clifford, S.; Cioslowski, J.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, RL.; Fox, DJ.; Keith,

T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Gonzalez, C.; Pople, JA. Gaussian, Inc; Wallingford, CT:

41. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JS. J Comput Chem. 1993; 14:1347.
42. Bryantsev VS, Diallo MS, van Duin ACT, Goddard WAI. J Chem Theory Comput. 2009; 5:1016.
43. Zhang P, Fiedler L, Leverentz HR, Truhlar DG, Gao JL. J Chem Theory Comput. 2011; 7:857.
44. Gao J, Xia X, George TF. J Phys Chem. 1993; 97:9241.
45. Jorgensen WL, Pranata J. J Am Chem Soc. 1990; 112:2008.
46. Mo Y, Schleyer PvR WuW, M Lin, Q Zhang, J Gao. J Phys Chem A. 2003; 107:10011.
47. Mo YR, Bao P, Gao JL. Phys Chem Chem Phys. 2011; 13:6760. [PubMed: 21369567]
48. Gao J, Garner DS, Jorgensen WL. J Am Chem Soc. 1986; 108:4784.
49. Gao J. ACS Symp Ser. 1994; 569:8.
50. Freindorf M, Gao J. J Comput Chem. 1996; 17:386.
51. Riccardi D, Li G, Cui Q. J Phys Chem B. 2004; 108:6467. [PubMed: 18950136]
52. Freindorf M, Shao YH, Furlani TR, Kong J. J Comput Chem. 2005; 26:1270. [PubMed: 15965971]
53. Gao J, Cembran A, Mo Y. J Chem Theory Comput. 2010; 6:2402.
54. Xie W, Orozco M, Truhlar DG, Gao J. J Chem Theory Comput. 2009; 5:459. [PubMed: 20490369]

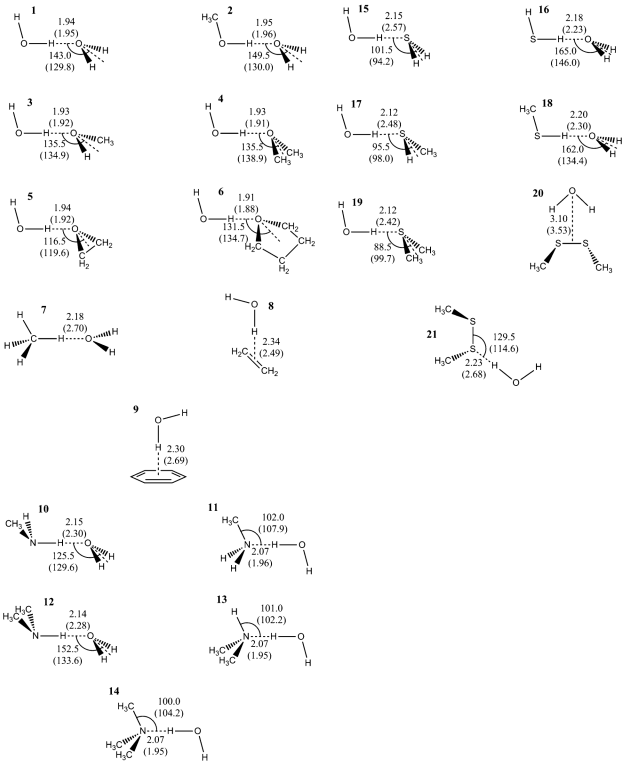

Fig. 1.

Schematic illustration of the bimolecular complexes between water and small molecules. Optimized hydrogen bond distances and angles from the X-Pol potential, XP@B3LYP/6-31G(d), are first given, followed by values in parentheses by B3LYP/aug-cc-pVDZ. Distances are given in angstroms and angles in degrees

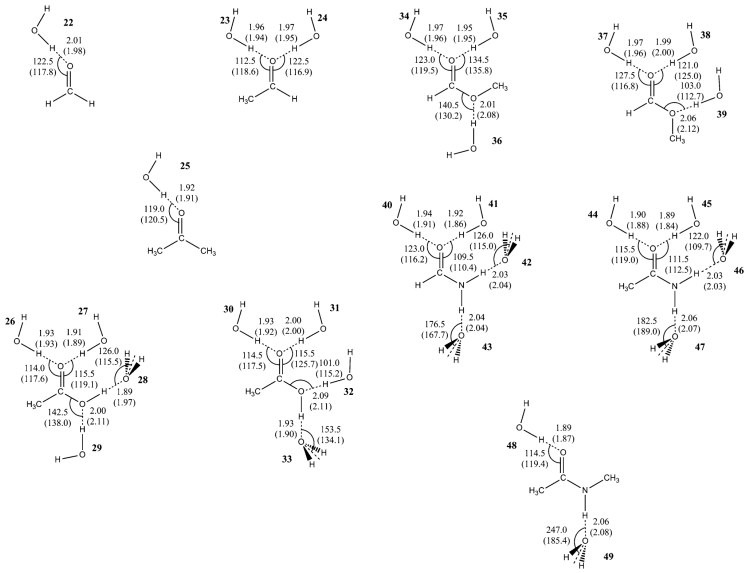

Fig. 2.

Bimolecular complexes depicting the interactions between water and a series of carbonyl-containing compounds. Optimized hydrogen bond distances and angles from the X-Pol potential, XP@B3LYP/6-31G(d), are given first, followed by values in *parentheses* by B3LYP/aug-cc-pVDZ. Distances are given in angstroms and angles in degrees

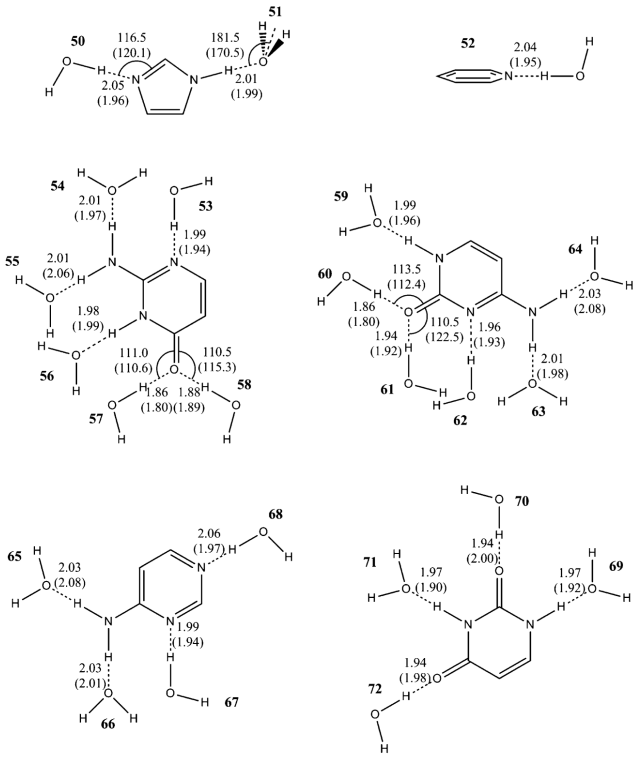

Fig. 3.

Bimolecular complexes depicting the interactions between water and heterocyclic compounds. Optimized hydrogen bond distances and angles from the X-Pol potential, XP@B3LYP/6-31G(d), are given first, followed by values in *parentheses* by B3LYP/aug-cc-pVDZ. Distances are given in angstroms and angles in degrees

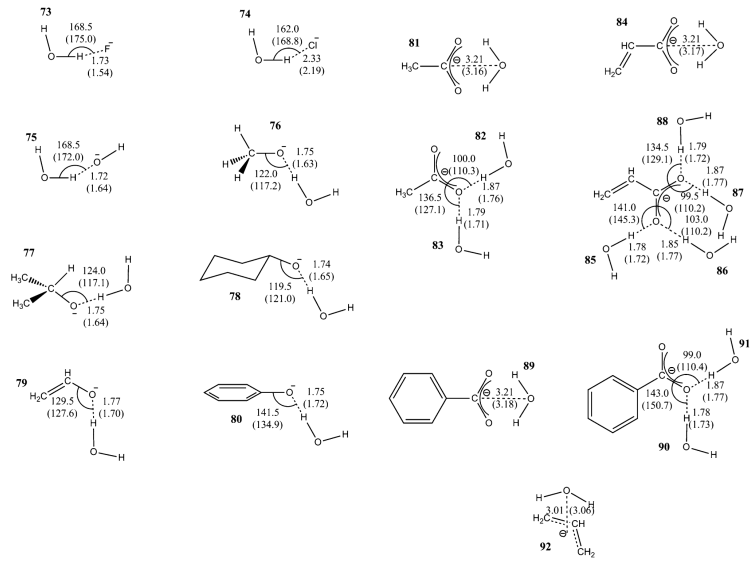

Fig. 4.

Bimolecular complexes depicting the interactions between water and anionic species. Optimized hydrogen bond distances and angles from the X-Pol potential, XP@B3LYP/6-31G(d), are given first, followed by values in parentheses by B3LYP/aug-cc-pVDZ. Distances are given in angstroms and angles in degrees

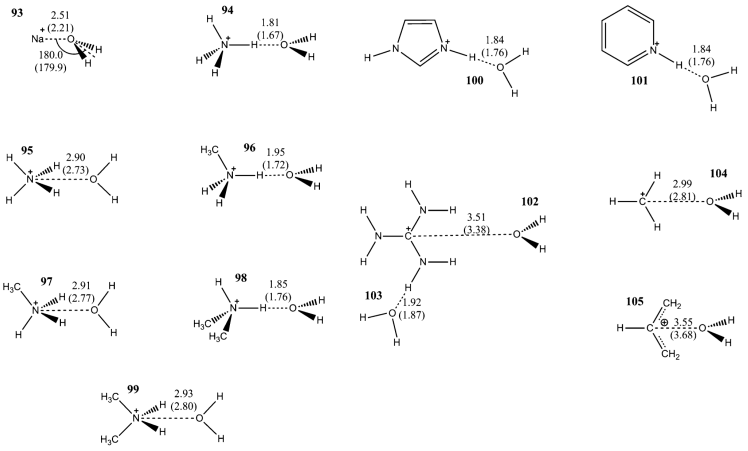

Fig. 5.

Bimolecular complexes depicting the interactions between water and cationic species. Optimized hydrogen bond distances and angles from the X-Pol potential, XP@B3LYP/6-31G(d), are given first, followed by values in parentheses by B3LYP/aug-cc-pVDZ. Distances are given in angstroms and angles in degrees

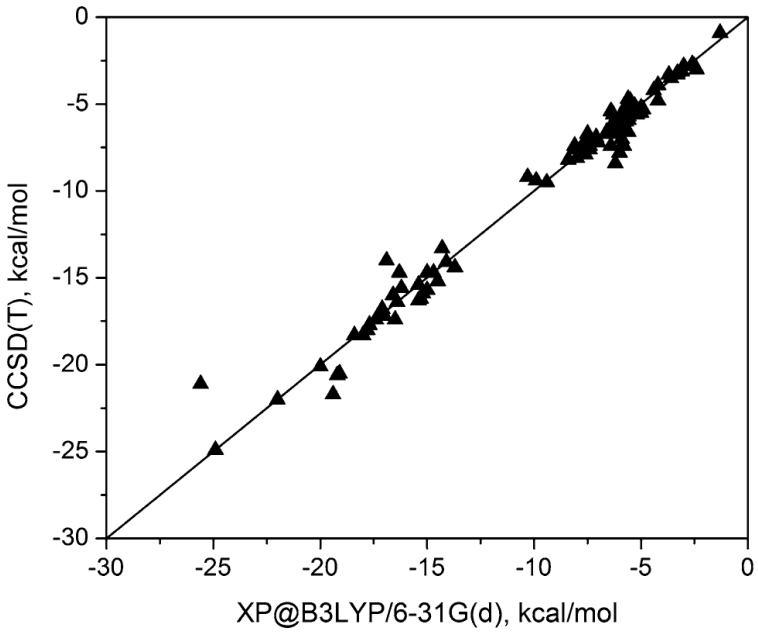

Fig. 6.

Comparison of the computed hydrogen bond interaction energies obtained using the XP@B3LYP/6-31G(d) and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods. The geometries used in the X-Pol calculations were obtained at the same level of theory, whereas those used in the coupled cluster energy evaluations were optimized with B3LYP/aug-cc-pVDZ

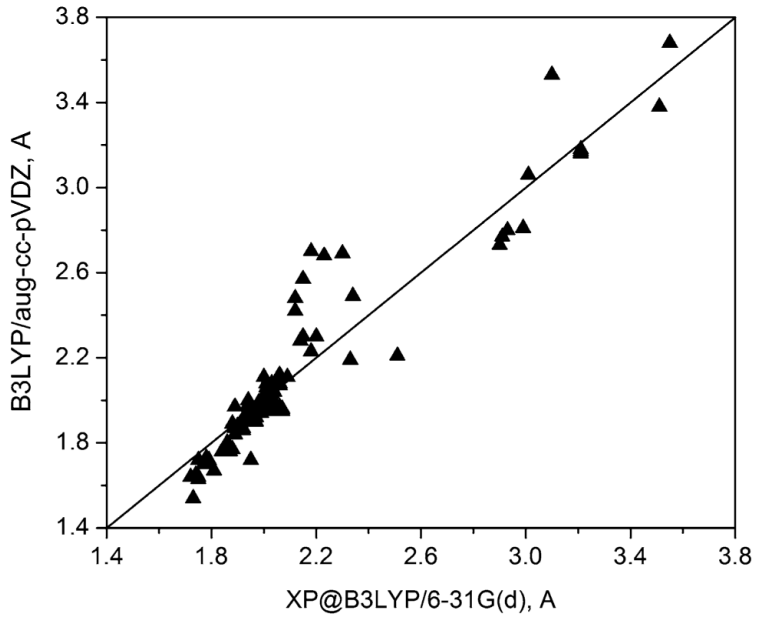

Fig. 7.
Comparison of the optimized hydrogen bond distances using the XP@B3LYP/6-31G(d) and B3LYP/aug-cc-pVDZ methods

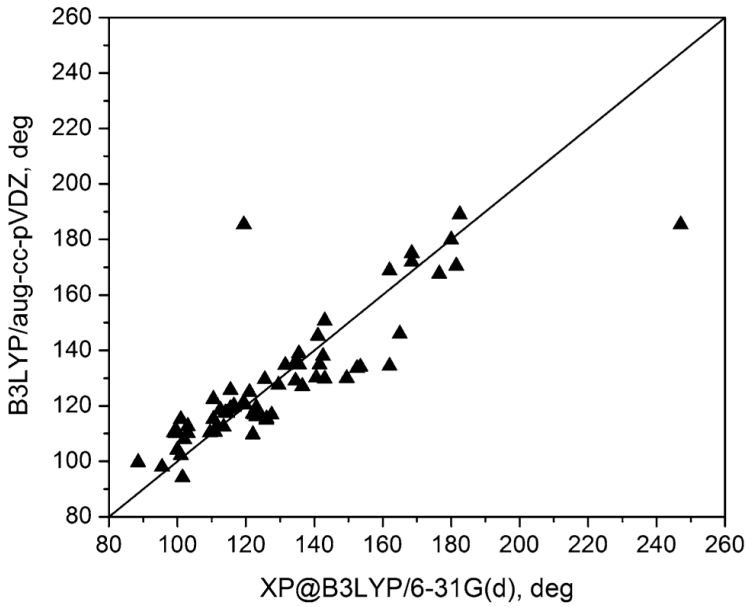

Fig. 8.
Comparison of the optimized hydrogen bond angles using the XP@B3LYP/6-31G(d) and B3LYP/aug-cc-pVDZ methods

Table 1

Optimized Lennard-Jones parameters in the X-Pol potential with B3LYP/6-31G(d)

Atom	σ (Å)	ε (kcal/mol)
H	1.31	0.04
H (−SH)	1.81	0.04
C	3.67	0.16
N (neutral)	3.60	0.20
N (cation)	3.47	0.20
O (both *sp*2 and *sp*3)	3.25	0.15
O^- (RO^-)	3.21	0.15
O^- (RCO_2^-)	3.24	0.15
S	3.11	0.56
Na^+	2.51	0.30
Cl^-	4.37	0.21
F^-	2.97	0.45

Table 2

Binding energies of bimolecular complexes between water and simple functional groups containing oxygen, nitrogen, and sulfur atoms computed using the XP@B3LYP/6-31G(d), B3LYP/aug-cc-pVDZ, and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods

Complex	XP@B3LYP/6-31G(d)	B3LYP/aug-cc-pVDZ	CCSD(T)/aug-cc-pVDZ
1	5.7	4.6	5.2
2	5.5	4.4	5.4
3	5.9	5.0	5.9
4	6.1	5.0	6.3
5	6.0	4.9	6.0
6	7.0	5.6	7.2
7	1.3	0.4	0.9
8	2.4	1.9	3.0
9	4.2	1.4	3.9
10	2.6	1.9	2.8
11	5.9	6.6	7.4
12	3.0	2.0	2.8
13	6.0	6.5	7.8
14	6.2	6.5	8.4
15	3.7	2.6	3.3
16	3.1	2.3	3.1
17	4.4	3.3	4.2
18	2.5	1.8	2.8
19	5.0	4.0	5.2
20	3.3	1.8	3.3
21	2.6	1.6	2.7

Energies are given in kilocalories per mole

Table 3

Binding energies of bimolecular complexes between water and carbonyl-containing compounds computed using the XP@B3LYP/6-31G(d), B3LYP/aug-cc-pVDZ, and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods

Complex	XP@B3LYP/6-31G(d)	B3LYP/aug-cc-pVDZ	CCSD(T)/aug-cc-pVDZ
22	4.2	4.2	4.8
23	5.8	5.2	6.4
24	5.0	4.9	5.5
25	6.6	5.7	6.6
26	6.1	5.2	6.2
27	6.0	5.1	5.9
28	7.6	7.9	7.9
29	3.3	2.1	3.2
30	6.4	5.5	6.6
31	5.4	4.6	5.4
32	5.6	3.5	4.7
33	8.4	6.6	8.2
34	4.9	4.6	5.3
35	6.2	5.1	6.5
36	3.6	2.2	3.5
37	5.2	4.9	5.6
38	5.2	4.5	5.3
39	5.5	3.5	4.8
40	6.0	5.8	6.4
41	7.4	6.7	7.6
42	5.7	5.0	6.0
43	5.9	4.7	5.6
44	7.5	6.5	7.5
45	8.0	7.1	8.1
46	5.6	4.7	5.9
47	5.6	4.3	5.6
48	7.9	6.8	7.9
49	5.6	4.1	5.7

Energies are given in kilocalories per mole

Table 4

Binding energies for bimolecular complexes between water and heterocyclic compounds computed with the XP@B3LYP/6-31G(d), B3LYP/aug-cc-pVDZ, and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods

Complex	XP@B3LYP/6-31G(d)	B3LYP/aug-cc-pVDZ	CCSD(T)/aug-cc-pVDZ
50	5.8	6.6	7.4
51	5.6	5.4	6.6
52	6.4	6.4	7.4
53	7.4	6.1	7.4
54	6.2	4.8	5.9
55	7.5	5.6	6.7
56	7.1	5.3	6.9
57	9.9	8.2	9.4
58	8.1	6.4	7.4
59	6.4	5.0	6.3
60	9.4	8.3	9.5
61	8.1	6.6	7.6
62	10.3	7.9	9.2
63	5.7	4.5	5.7
64	6.3	4.3	5.6
65	6.4	4.2	5.4
66	5.3	3.9	5.1
67	7.5	6.3	7.4
68	5.9	6.1	7.0
69	7.7	6.2	7.3
70	5.5	4.1	5.1
71	6.5	5.3	6.7
72	5.7	4.3	5.1

Energies are given in kilocalories per mole

Table 5

Binding energies for bimolecular complexes between water and anions computed using the XP@B3LYP/6-31G(d), B3LYP/aug-cc-pVDZ, and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods

Complex	XP@B3LYP/6-31G(d)	B3LYP/aug-cc-pVDZ	CCSD(T)/aug-cc-pVDZ
73	24.9	24.8	24.9
74	14.1	13.8	14.1
75	25.6	21.2	21.1
76	19.4	21.1	21.7
77	19.2	19.5	20.6
78	19.1	19.3	20.5
79	16.4	15.8	16.4
80	16.3	14.0	14.7
81	18.0	17.1	18.3
82	17.8	15.3	18.0
83	15.3	17.0	16.2
84	17.7	16.4	17.7
85	15.4	14.2	15.4
86	17.4	16.4	17.4
87	17.4	16.4	17.4
88	15.0	14.7	15.7
89	17.3	15.7	17.2
90	15.0	13.3	14.7
91	17.1	15.7	16.8
92	15.2	13.9	15.9

Energies are given in kilocalories per mole

Table 6

Binding energies for bimolecular complexes between water and cations computed using the XP@B3LYP/6-31G(d), B3LYP/aug-cc-pVDZ, and CCSD(T)/aug-cc-pVDZ//B3LYP/aug-cc-pVDZ methods. Energies are given in kilocalories per mole

Complex	XP@B3LYP/6-31G(d)	B3LYP/aug-cc-pVDZ	CCSD(T)/aug-cc-pVDZ
93	22.0	23.9	22.0
94	20.0	20.4	20.1
95	15.4	15.8	16.3
96	18.4	17.9	18.3
97	14.5	14.3	15.2
98	17.0	16.3	17.2
99	13.7	13.2	14.4
100	16.2	14.8	15.6
101	16.6	14.9	16.0
102	16.5	15.8	17.4
103	14.3	12.3	13.3
104	14.7	12.8	14.7
105	16.9	13.9	14.0