Optimized halogen atomic radii for PBSA calculations using off-center point-charges

Andreia Fortuna[†,‡] and Paulo J. Costa[*,†]

[†]*BioISI – Biosystems & Integrative Sciences Institute, Faculty of Sciences, University of Lisboa, 1749-016, Lisboa, Portugal*

[‡]*Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, University of Lisbon, Av. Professor Gama Pinto, 1649-003 Lisbon, Portugal*

E-mail: pjcosta@ciencias.ulisboa.pt,pjcosta@fc.ul.pt

Phone: +351-21-7500196

Abstract

In force field methods, the usage of off-center point-charges, also called extra-points (EPs), is a common strategy to tackle the anisotropy of the electrostatic potential of covalently-bonded halogens (X), thus allowing the description of halogen bonds (XBs) at the molecular mechanics / molecular dynamics (MM/MD) level. Diverse EP implementations exist in the literature differing on the charge sets and/or the X–EP distances. Poisson–Boltzmann and surface area (PBSA) calculations can be used to obtain solvation free energies (ΔG_{solv}) of small molecules, often to compute binding free energies (ΔG_{bind}) at the MM-PBSA level. This method depends, among other parameters, on the empirical assignment of atomic radii (PB radii). Given the multiplicity of off-center point-charges models and the lack of specific PB radii for halogens compatible with such implementations, in this work we assessed the performance of PBSA calculations for the estimation of ΔG_{solv} values in water (ΔG_{hyd}), also conducting an optimization of the halogen PB radii (Cl, Br, and I) for each EP model. We not only expand the usage of EP models in the scope of the General AMBER Force Field (GAFF) but also provide the first optimized halogen PB radii in the context of the CHARMM General Force Field (CGenFF), thus contributing to improving the description of halogenated compounds in PBSA calculations.

keywords: hydration free energy, halogen radii, extra-point, PBSA calculations, halogenated ligands, halogen bonding

Introduction

The determination of binding free energies (ΔG_{bind}) is of extreme importance in computer-aided drug discovery[1,2] and, in order to obtain accurate results, the solvation free energy (ΔG_{solv}) must be properly estimated. Furthermore, ΔG_{solv} is by itself an important property in the study and characterization of compounds, therefore, the estimation of ligand hydration free energies, ΔG_{hyd}, is paramount as most biomolecular processes occur in water.[3,4] Free energy perturbation (FEP) and thermodynamic integration (TI) methods can be used to calculate both ΔG_{bind} and ΔG_{solv},[5–9] but at a cost of being computationally demanding. On the other hand, the molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) method has the advantage of requiring a much lower computational cost to obtain both quantities, hence, it can be applied to large libraries of compounds.[10,11] In this method, the hydration (solvation) free energies of the components of the system (e.g. protein–ligand) are estimated using the Poisson–Boltzmann (PB) equation and surface area (SA) calculations, which, among other terms, depends on atomic charges and a set of atomic radii, often called PB radii , which are used to define the solute (ϵ_{in}) and solvent (ϵ_{out}) regions (Figure 1).

Covalently-bonded halogens (R–X, with X = Cl, Br, I) possess anisotropic electronic features, showing a positive region in their electrostatic potential, called the σ–hole, which enables their interaction with a Lewis base (B).[12] This type of R–X$\cdots$B interaction, known as halogen bond (XB),[13] has found application in diverse fields of (bio)chemical sciences.[14,15]

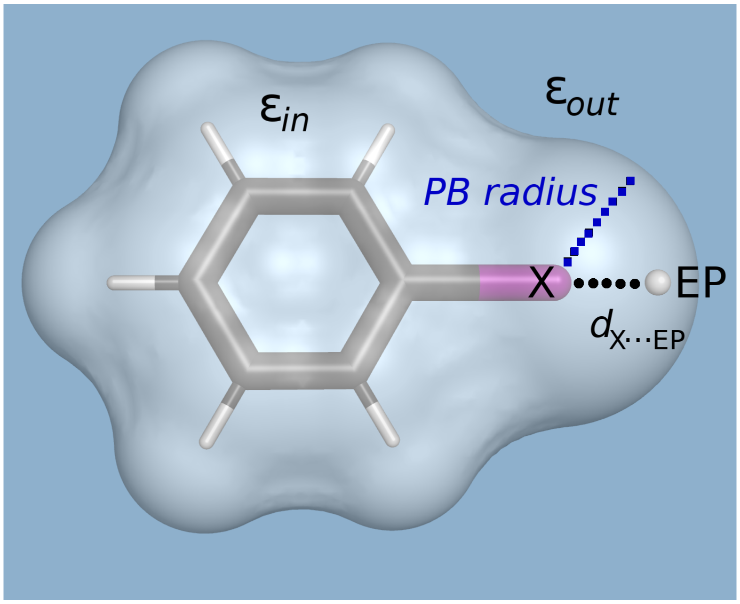

Figure 1: Schematic representation of a halogenated molecule in a PBSA calculation: an off-center point-charge (EP) is placed at a given distance from the halogen X ($d_{X\cdots EP}$) whereas the halogen PB radius along with those of the remainder atoms, is used to define the region of the solute (ϵ_{in}) and the solvent (ϵ_{out}).

Indeed, there are numerous examples were XBs play an important role in anion recognition,[16,17] catalysis,[18] and in protein-ligand systems,[19–21] this latter case showing their relevance in medicinal chemistry. In fact, halogenated compounds are strongly present in drug discovery,[20,22] mainly due to their propensity to improve drug-like properties such as membrane permeability and pharmacokinetic stability.[23,24] Therefore, computational methods aiming at describing halogenated ligands and their association with a biomolecular target must provide a proper description of the σ–hole which is not straightforwardly achieved by standard force fields (FFs) as these typically consider halogen atoms to carry a negative charge leading to unfavorable interactions with other negative atoms.[21] To overcome this problem, various strategies to emulate the σ–hole in empirical FFs emerged, namely, electric multipole expansions, aspherical interatomic potentials, and off-center point charges.[25] The latter methodology is quite simple and consists of placing a positive charge, often called extra point (EP), at a given distance from the halogen along the R–X bond axis to emulate the σ–hole (Figure 1). This strategy has the advantage of a rather low impact on the computational cost which is important for large-scale simulations. Thus, a variety of different implementations of the EP strategy were published[26–33] differing in the charge-fitting scheme or charges assigned, and on the X···EP distance ($d_{\mathrm{X\cdots EP}}$). These implementations are able to properly describe XBs in molecular dynamics simulations of protein–ligand systems.[21] In some studies, MD trajectories were used in MM-(GB/PB)SA (GB = Generalized Born) calculations to estimate ΔG_{bind}[26,34–37] and even though the presence of an EP usually leads to accurate results regarding protein–ligand geometries and sampling of XBs, its impact on the accuracy of ΔG_{solv} predictions when using such methods is yet to be properly evaluated.

Very recently,[38] we showed that a problematic issue arises when using PBSA calculations combined with a common off-center point-charge implementation for the General Amber Force Field (GAFF)[39] that uses a $d_{\mathrm{X\cdots EP}}$ value corresponding to the Lennard-Jones (LJ) parameter R_{min} of the halogen[26] along with restrained electrostatic potential (RESP) charges[40] fitted for all particles (here denoted ***EP1*** method). Indeed, since PB radii sets such as PARSE[41] or modified Bondi (mbondi)[42] do not include specific values for the heavier halogens, standard radii (r_{std}) are commonly taken from the literature[4,43] (Table S1). However, these values are smaller than R_{min}, thus placing the EP outside the halogen radius, i.e., in the solvent dielectric, leading to a poor estimation of ΔG_{solv}. By conducting an optimization we provided a new set of halogen PB radii (r_{opt}) that can be used with $d_{\mathrm{X\cdots EP}} = R_{\mathrm{min}}$ in the context of GAFF (***EP1***)[26] yielding mean absolute error (MAE) values between experimental and calculated ΔG_{solv} values lower than 1.5 kcal mol^{-1} as opposed to 4.2 kcal mol^{-1} when standard radii are used. This work also showed that even without EP addition, the halogen PB radii can be slightly optimized (MAE below 1.8 kcal mol^{-1} *vs* 2.0 kcal mol^{-1} without optimization). A grid of $d_{\mathrm{X\cdots EP}}$ values and halogen PB radii eventually showed the nonexistence of optimal parameters as a wide range of pairs yield low MAE values, also demonstrating that the optimization strategy can be ported into other FFs and/or off-center point-charge implementations.

Given the diverse EP strategies that can be used for performing MD simulations (and possibly MM-PBSA calculations), some of them also placing the EP outside the halogen radius (i.e. $d_{X\cdots EP} > r_{std}$) herein, we assess the performance of PBSA in the estimation of ΔG_{solv} in water (ΔG_{hyd}) using three different off-center point-charge implementations from literature in the context of GAFF[27,29] and CHARMM[30] force fields, also conducting an optimization of the PB radii for Cl, Br, and I for each EP model. With this work we aim at improving the description of halogenated compounds in structure-based virtual screening approaches by providing methods that accurately estimate ΔG_{solv} values which are paramount to understand protein–ligand binding events. Moreover, and since the addition of EPs is not limited to halogens, being also important to address anisotropic electronic features in other elements such sulfur (σ–hole),[44,45] oxygen or nitrogen atoms (lone pairs),[44,46] this work might also lay the ground for expanding the optimizations to those elements.

Methods

Experimental Data

From FreeSolv[47](version 0.51[48]), a database of experimental and simulation data of hydration free energies, 142 halogenated molecules were taken, comprising 107 chlorinated, 23 brominated, and 12 iodinated compounds which were analysed independently. As before,[38] compounds containing multiple halogens from different elements in their structure (other than fluorine), namely, ID 4506634 (2-bromo-2-chloro-1,1,1-trifluoroethane) and ID 7768165 (1-bromo-2-chloroethane) were excluded from our analysis. Regarding fluorine-containing molecules, only compounds comprising also heavier halogens in their structure were considered.

Off-center point-charge models

We used three different off-center point-charge implementations taken from the literature termed ***EP2***,[27] ***EP3***,[29] and ***EP4***.[30] For all of them, only one EP is added along the C–

X covalent bond axis with the C–X$\cdots$EP angle being fixed at 180.0°. Their main features, namely, $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ and atomic charges assigned to the EP (q_{EP}) are summarized in Table 1. The values for ***EP1*** model, studied in an earlier publication,[38] are also shown for comparison.

Table 1: General information ($\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ in Å and EP charges, q_{EP}) for the different methods used in this work.

	GAFF						CHARMM	
	EP1		***EP2***		***EP3***		***EP4***	
Halogen	$\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$	q_{EP}	$\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$	q_{EP}	$\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$	q_{EP}	$\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$	q_{EP}
Cl	1.948	RESP fitted	Best fit to QM ESP[a]		1.00	0.1	1.64	0.05
Br	2.020				1.30	0.2	1.89	0.05
I	2.150				1.60	0.3	2.20	0.05

[a] The list of $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values is given in Tables S2-S4.

The first off-center point-charge model studied in this work was proposed by Sironi and co-workers[27] (Table 1, ***EP2***) in the context of AMBER/GAFF.[39] In this approach, atomic charges are fitted for all atoms (including the EP) using a RESP fitting procedure, and $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values are determined individually for each molecule by minimizing the error of the fit to the reference quantum-mechanical (QM) electrostatic potential (ESP). Thus, using previously optimized geometries at the B3LYP/6-311G(d,p) level of theory[49–54] with Gaussian 09, Revision A.2,[55] the reference ESPs were generated at the HF/6-31G(d)[56–58] level of theory for all elements, with the exception of iodine, for which the 6-311G(d)[54] basis set was used. The remainder parameters are described in an earlier publication.[38] Atomic partial charges were obtained by applying the restrained ESP (RESP)[40] procedure using the antechamber[59] module as implemented in AmberTools17.[60] During this RESP fitting procedure, an EP was placed along the C–X covalent bond by varying $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ from 1 Å to 4 Å with 0.01 Å increments. According to the ***EP2*** model strategy, the selected $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values were the ones minimizing the relative root-mean-square (RRMS) error of the fit. When $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ was larger than R_{min}, the value of R_{min} and respective atomic charges were attributed instead. The final $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values obtained for each compound are presented in Tables S2-S4.

Notice that in ***EP2***, and owing to the variable nature of the $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values, the EP is often placed in the water dielectric when using standard halogen PB radii (Table S1), i.e. $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}} > r_{\mathrm{std}}$ (see also the Off-center point-charge model ***EP2*** section in Results and Discussion).

The second method tested, ***EP3***, also called "no fit" explicit σ–hole (ESH), was proposed by Hobza and co-workers[29] in the context of AMBER/GAFF.[39] In this work a Br–EP distance of 1.5 Å and an EP charge of $0.20e$ were suggested as universal parameters. These values were reviewed in a subsequent publication,[61] also providing $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ and q_{EP} values for iodine and chlorine. In this case, the same RESP fitting procedure as described before was performed without EP addition. Then, the EP was added at the corresponding $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ value and a charge (q_{EP}) was attributed (Table 1, ***EP3***). The value of q_{EP} is then subtracted from the halogen charge, while the other atomic partial charges were not modified. Notice than once charges are obtained for a given set of molecules, no additional calculations are necessary to assign EP charges, thus saving computational time which could be useful for high-throughput calculations.[61] In this EP model, the $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values are always smaller than the corresponding r_{std} and therefore, the EP is correctly placed within the solute dielectric.

The last method explored in this work, named ***EP4***, is based on the parameterization of XBs in the CHARMM General Force Field (CGenFF).[30] In this case, and similarly to ***EP3***, the EP distances and charges for each halogen are fixed (Table 1, ***EP4***). To the best of our knowledge, no specific halogen PB radii exist to be employed along with CGenFF parameters. In this work, we used the halogen R_{min} values as r_{std} placeholders for the subsequent optimization (see Table S5 and Discussion below). Hence, using the standard values, $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}} < r_{\mathrm{std}}$ which correctly places the EP in the solute. The platform Ligand Reader & Modeler[62] from CHARMM-GUI,[63] or in some specific cases for which CHARMM-GUI failed to create a topology file, the CGenFF server (https://cgenff.umaryland.edu)[64] were used to generate CHARMM topologies containing atomic charges. Unfortunately for compounds ID 2996632 (chloroform), ID 6359135 (carbon-tetrachloride) and ID 8311321 (chloro-difluoro-methane), CHARMM topolgies could not be obtained, thus decreasing the number of chlorinated molecules analyzed to 104. A direct conversion of the CHARMM topology to an AMBER compatible one (e.g. using CHAMBER[65]), which could be read by the `pbsa` program of AmberTools 17,[60] was not performed since colinear lonepairs are not supported. Therefore the CHARMM topologies were converted to a free-format pqr file using in-house scripts.

PBSA Calculations

We followed a similar procedure as the one described in reference 21. Succinctly, PBSA calculations were performed using the `pbsa` program provided in AmberTools 17,[60] with a single conformation per compound. These type of data sets are most frequently constituted by compounds with few rotatable bonds,[4] thus averaging over multiple conformations has only a marginal impact on the results.[66] Indeed, in our previous work,[38] no significant improvement is achieved after systematic conformational search for selected compounds. The solvation free energy (ΔG_{solv}) was computed from the contribution of a polar component (ΔG_{polar}) and a nonpolar component ($\Delta G_{\text{nonpolar}}$):

$$\Delta G_{\text{solv}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \tag{1}$$

The ΔG_{polar} accounts for the electrostatic contribution to the total solvation free energy, obtained by a finite difference solution of the linearized Poisson-Boltzmann (PB) equation. On the other hand, the $\Delta G_{\text{nonpolar}}$ returns the contribution of solute-solvent hydrophobic interactions and cavitation. A solvent probe radius of 1.4 Å was used to define the solvent excluded surface and the solute and solvent dielectric constants were set to 1 and 80, respectively, thus setting $\Delta G_{\text{solv}} = \Delta G_{\text{hyd}}$. The remainder parameters can be found in reference 38.

In the context of AMBER/GAFF (methods ***EP2***, ***EP3***) we used three different calculation setups, termed ***pb1***, ***pb2***, and ***pb3***, representing different flavors usually employed for MM-PBSA calculations with this force field (Table S6). Briefly, for ***pb1*** and ***pb2*** setups, the PARSE[41] and mbondi radii,[42] respectively, are used for both ΔG_{polar} and $\Delta G_{\text{nonpolar}}$ calculations. For ***pb3***, ΔG_{polar} is calculated using mbondi radii whereas for $\Delta G_{\text{nonpolar}}$, attractive ($\Delta G_{\text{dispersion}}$) and repulsive interactions are treated separately.[10,67] This is done by using a surface-based integration method for the $\Delta G_{\text{dispersion}}$ term while the molecular volume enclosed by SASA (SAV) is estimated for the calculation of the cavity repulsive term using the R_{min} values as atomic radii. Atomic radii were assigned by default by `leap` in AmberTools 17[60] for each set (PARSE and mbondi), with minor corrections and the values are provided in Table S1.

In the context of CHARMM/CGenFF, we used a calculation setup, termed ***pb4*** (Table S6), similar to the one described in reference 68 which applied CHARMM36 (C36) and CGenFF to reproduce the experimental solvation free energies of a set of 70 molecules yielding an RMS of 2.5 kcal mol^{-1}. Similarly to this work, herein both ΔG_{polar} and $\Delta G_{\text{nonpolar}}$ were calculated using the set of radii from Nina et al.[69] and Banavali et al.[70] (see Table S5) and setting $\gamma = 0.005$ kcal mol^{-1} Å^{-2} while the constant term, β was set to zero.

For all tested off-center point-charge models and PBSA setups, the PB radii of the halogen atoms (Cl, Br, I) were varied between 1.00 and 4.00 Å at incremental values of 0.01 Å while keeping those of the remaining elements constant. A value of 0.00 Å was assigned to both the EP PB radius and its R_{min} value (a detailed explanation on how to run PBSA calculations with an EP was given in an earlier publication).[38] It is important to mention that in setups ***pb1***, ***pb2***, and ***pb4*** the variation of the halogen PB radius affects both ΔG_{polar} and $\Delta G_{\text{nonpolar}}$ terms. For ***pb3***, the halogen radius variation only affects ΔG_{polar}, while the contribution of $\Delta G_{\text{nonpolar}}$ is constant as it only depends on the R_{min} value.

The mean absolute error (MAE), calculated as

$$\text{MAE} = n^{-1} \sum_{i=1}^{n} |\Delta G_{\text{solv}}(\text{calc})_i - \Delta G_{\text{solv}}(\exp)_i| \tag{2}$$

was used to evaluate the accuracy of the solvation free energy values. It consists of the unsigned difference between the computed (ΔG_{solv}(calc)) and experimental (ΔG_{solv}(exp)) values, averaged for each entire subset of compounds (n).

Results and Discussion

The solvation free energies (ΔG_{solv}) for a set of halogenated molecules taken from the FreeSolv database[47] were estimated using PBSA calculations with different calculation setups (***pb1***, ***pb2***, ***pb3***, and ***pb4***). Three off-center point-charge models (***EP2***, ***EP3***, and ***EP4***), meant to tackle the anisotropy of the halogens, were evaluated in the context of PBSA calculations aiming at the optimization of the halogen radii by comparing the calculated ΔG_{solv} values with the experimental ones. A comparison with the values obtained in a previous work using ***EP1***[38] is also performed. A discussion for each EP model is presented in the following sections.

Off-center point-charge model ***EP2***

As mentioned previously, model ***EP2*** is based on the work of Sironi and co-workers,[27] and considers the ability of the EP to reproduce the full anisotropy of the potential around the halogen by using the X$\cdots$EP distance that yield the best fit to the QM electrostatic potential. This means that, for each compound, a specific X$\cdots$EP distance is assigned along with atomic charges which are attributed to all atoms (including the EP) by the RESP procedure. We therefore applied this method to our library of halogenated compounds and the X$\cdots$EP distances ($\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$) that provide the best fit of the ESP against the QM potential are summarized in Tables S2-S4. Notice that for the subsequent PBSA calculations, when the obtained $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values are larger than R_{min}, $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ is set to the R_{min} value since placing the EP beyond this value might lead to instabilities during MD simulations.[25] The distribution of $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values for each dataset is shown in Figure 2. For iodinated molecules,

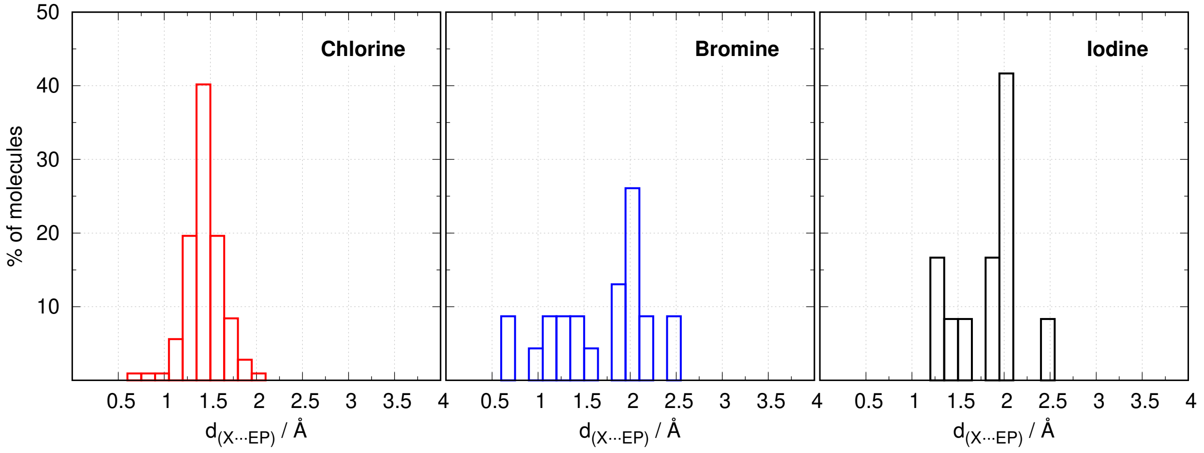

Figure 2: Distribution of $d_{X\cdots EP}$ values (Å) assigned using ***EP2***.

the assigned distances are in the 1.25–2.41 Å range with a maximum at ≈ 2.02 Å, slightly smaller than the R_{min} value for iodine (2.15 Å) and slightly larger than r_{std} (1.98 Å for both PARSE and mbondi sets, see Table S1) meaning that for several molecules, $d_{X\cdots EP} > r_{\text{std}}$. For the brominated dataset, the distances that led to the best fit to the QM are in a range of 0.72–2.40 Å, with a peak occurring at ≈ 2.02 Å, similar to the R_{min} value for bromine (2.02 Å) and larger than the standard PB radius (1.85 Å). The behavior of the chlorinated set is different as the $d_{X\cdots EP}$ values are evenly distributed around the maximum occurring at 1.42 Å (range 0.74-1.97 Å), a value much smaller than the R_{min} for chlorine (1.948 Å) and the standard PB radius (1.70 Å). These results show that placing the EP at R_{min} in iodinated and brominated compounds (as in ***EP1*** model) might not introduce larger errors on the fit of the ESP when compared with the reference potential, while for chlorinated molecules, shorter distances than R_{min} appear to reduce the mentioned error. Given the different size of the datasets, this issue must be further investigated, however, it is out of the scope of this manuscript.

In order to optimize the halogen PB radii, these were systematically varied for all compounds while using point charges derived for model ***EP2***, and the MAE values against experimental values were calculated (Figure 3). The plot reflects the non-constant nature

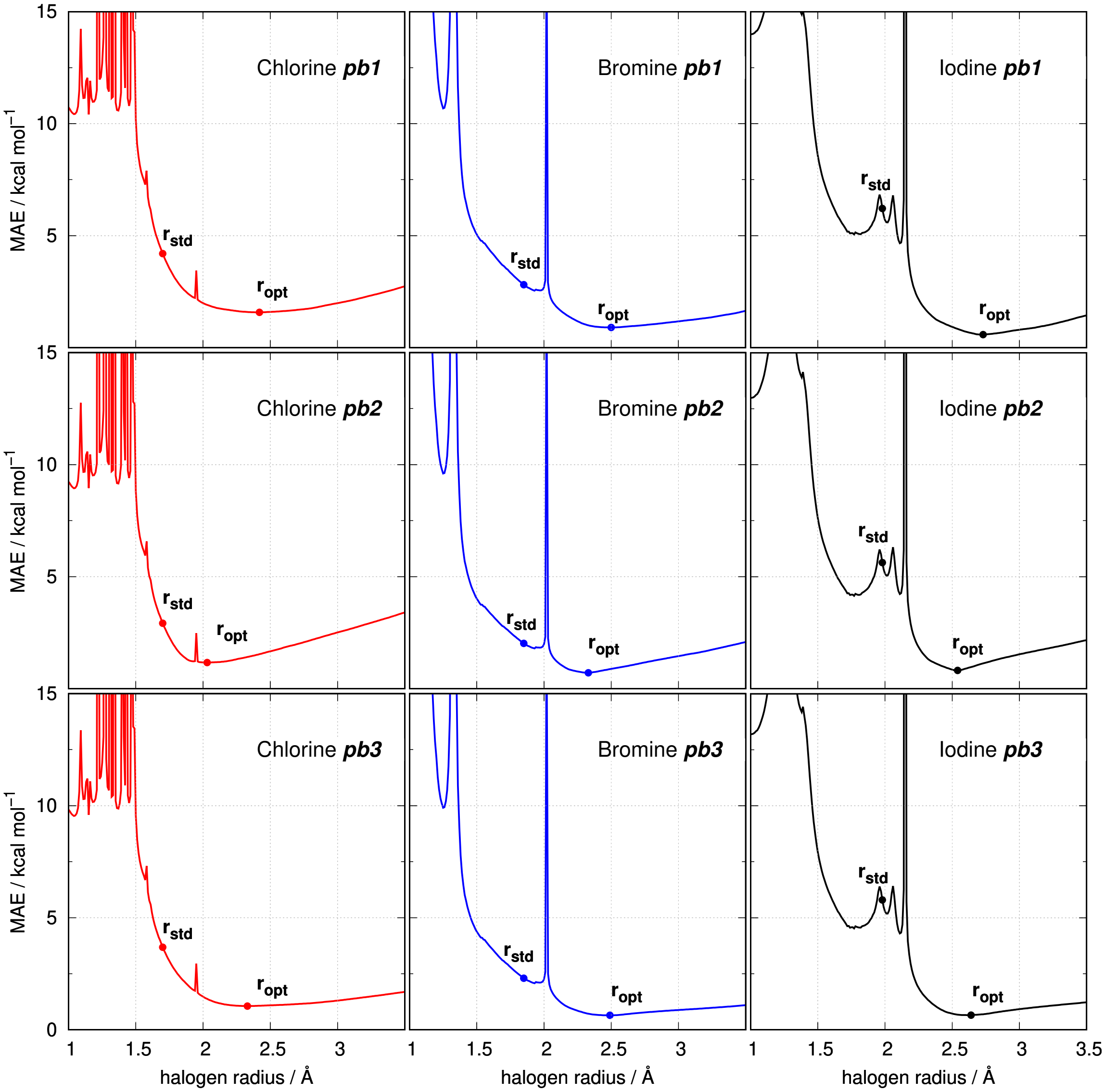

Figure 3: Variation of the MAE values with the halogen radius for each subset of halogenated compounds with an addition of an EP based in model ***EP2*** using PBSA setups ***pb1***, ***pb2***, and ***pb3***.

of the X$\cdots$EP distances in model ***EP2***. As we have shown before,[38] when the halogen PB radius is equal to the X$\cdots$EP distance, an asymptotic-like behavior in the error is observed. Therefore, several error peaks are observed whenever PB radii = $d_{X\cdots EP}$ for a given compound. These errors may result from instabilities associated with a point charge being placed at the interface between the two-dielectric media. In spite the occurrence of these peaks, all curves show a minimum corresponding to an optimal halogen PB radius (r_{opt}) that minimizes the MAE. The values are presented in Table 2 along with the MAEs obtained using r_{std} values. Using standard radii leads to considerably larger errors (up to c.a. 6 kcal mol^{-1} for ***pb1*** in the iodine dataset) than those obtained using the optimized values for which the larger MAE is c.a. 1.6 kcal mol^{-1} using ***pb1*** for chlorine, thus showing the importance of the halogen radii optimization in this context.

Table 2: MAE (kcal mol^{-1}) values obtained with standard (r_{std}) and optimized (r_{opt}) halogen radii (Å) for each subset of halogenated compounds with the addition of an EP using ***EP2*** model for PBSA setups ***pb1***, ***pb2***, and ***pb3***.

Halogen	r_{std}	***pb1***		***pb2***		***pb3***	
		MAE$_{\text{std}}$	MAE$_{\text{opt}}$(r_{opt})	MAE$_{\text{std}}$	MAE$_{\text{opt}}$(r_{opt})	MAE$_{\text{std}}$	MAE$_{\text{opt}}$(r_{opt})
Cl	1.70	4.201	1.579 (2.42)	2.923	1.172 (2.03)	3.686	1.059 (2.33)
Br	1.85	2.812	0.901 (2.50)	2.025	0.712 (2.33)	2.302	0.647 (2.49)
I	1.98	6.218	0.583 (2.73)	5.631	0.819 (2.54)	5.795	0.654 (2.64)

The setup that lead to lower MAE values for iodinated compounds was ***pb1***, with an MAE value of 0.583 kcal mol^{-1}, although ***pb3*** yields competitive results (0.654 kcal mol^{-1}). This latter setup achieved better results (0.647 kcal mol^{-1} and 1.059 kcal mol^{-1}) for brominated and chlorinated compounds, respectively. Moreover, for this latter dataset, larger MAEs were observed for all setups, which was also verified in the previous study,[38] and is related to the larger sample size (107 compounds).

The r_{opt} values follow the expected order of increasing atomic number (Cl < Br < I) in all the setups and are larger than the corresponding R_{min}, which is coherent with the results obtained in a previous work[38] for ***EP1***. Indeed, comparing the ***EP2*** results with those reported for ***EP1***, both the r_{opt} values and MAEs are quite similar.

The correlation between the calculated and experimental ΔG_{solv} values can be obtained using r_{opt} values (***EP2***, Figure 4). For all the setups, the correlations are high, although

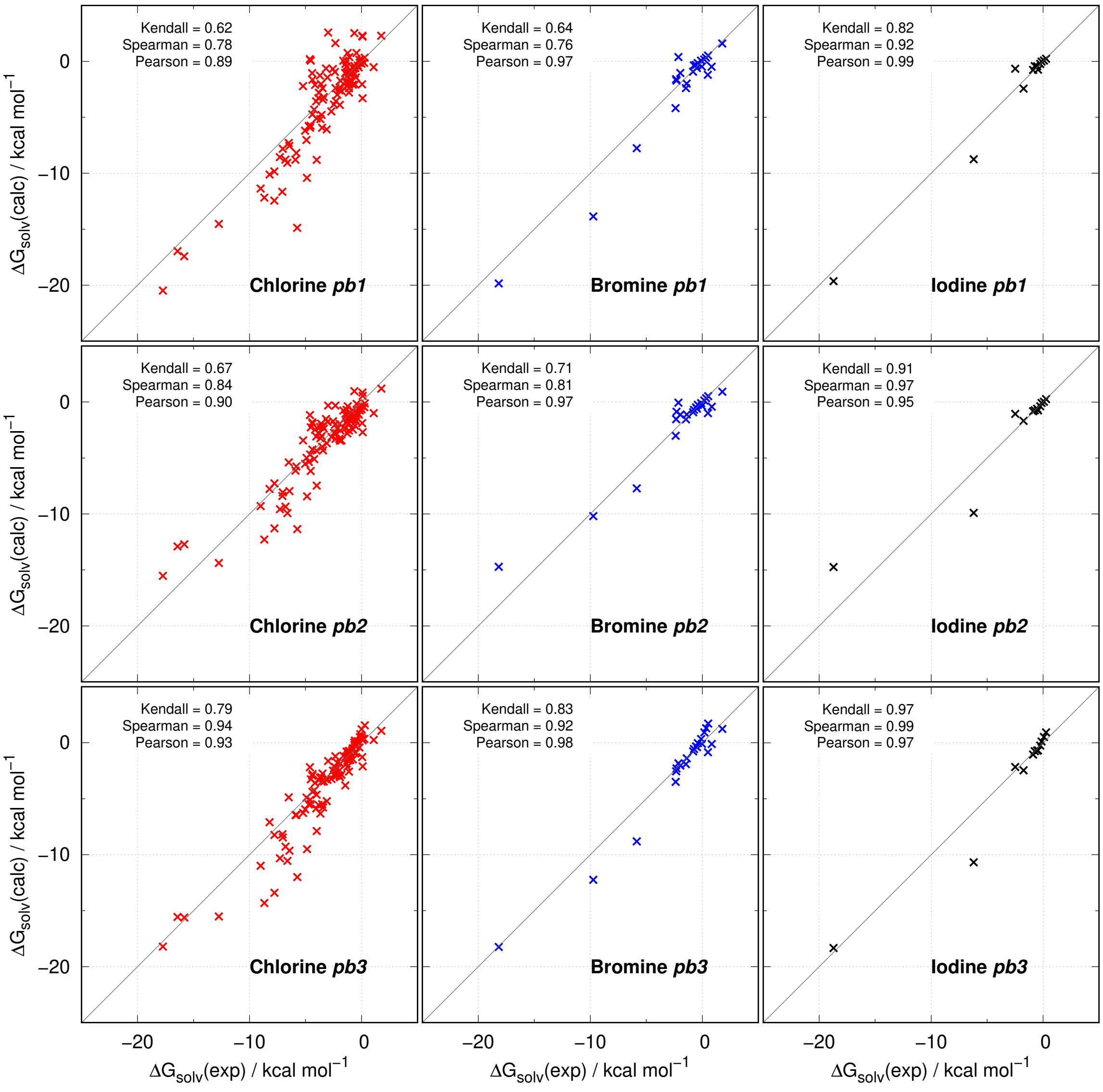

Figure 4: Correlation between experimental an calculated ΔG_{solv} values for each subset of halogenated compounds with with ***EP2*** model. The inset shows the resulting Kendall τ, Spearman rank correlation, and Pearson R correlation coefficients.

slightly lower for chlorinated compounds, which is consistent with the MAE values obtained. Additionally, the majority of the compounds possess an absolute deviation below 2 kcal mol^{-1} (Figure S1 and Figure S2), with higher deviations obtained for chlorinated compounds, as expected. These plots also show that the magnitude of the MAE is not concomitant with a systematic over/underestimation of ΔG_{solv} as the signed difference between calculated (ΔG_{solv}(calc)) and the experimental values (ΔG_{solv}(exp)) is reasonably distributed around 0.

Although the ability to predict ΔG_{solv} significantly varies with the PBSA method, some compounds are particularly difficult to tackle. For instance, in the iodinated set, 2-iodophenol presents the largest deviations (-4.48 to -2.55 kcal mol^{-1}). This systematic overestimation also occurs in a brominated analogue, 4-bromophenol (deviations up to -2.99 kcal mol^{-1}). As before,[38] the largest outlier is the chlorinated compound dialifor, the ΔG_{solv} values being highly overestimated especially when using the ***pb1*** setup (-9.14 kcal mol^{-1} deviation). Notice that the difficulties regarding the prediction of the hydration free energy of this compound was observed in other studies,[71,72] which can be associated with the presence of polar functional groups such as triophosphate, and also by a significant conformational floppiness.

Overall, using default radii leads to larger errors and therefore r_{opt} values should be used instead. When comparing the three tested PBSA setups, ***pb3*** setup is recommended under optimized conditions.

Off-center point-charge model ***EP3***

Contrarily to ***EP2***, in model ***EP3*** the values for the distances and charges of the EP are fixed (Table 1).[29] Thus, in the plots depicting the variation of the MAE against the PB radius of the halogen (Figure 5), error peaks are observed at 1.00 Å, 1.30 Å, and 1.60 Å, for chlorinated, brominated, and iodinate compounds, respectively. As explained before, these peaks occur when the PB radius value is equal to the X···EP distance ($\text{d}_{\text{X}\cdots\text{EP}}$) and the overall shape of the curves resemble those obtained with ***EP1*** where a fixed X···EP distance (equal to R_{min}) was also used,[38] although the peaks obtained herein are broader, specially for iodine.

In Table 3, the MAE values obtained for both r_{std} and r_{opt} are summarized. For iodinated

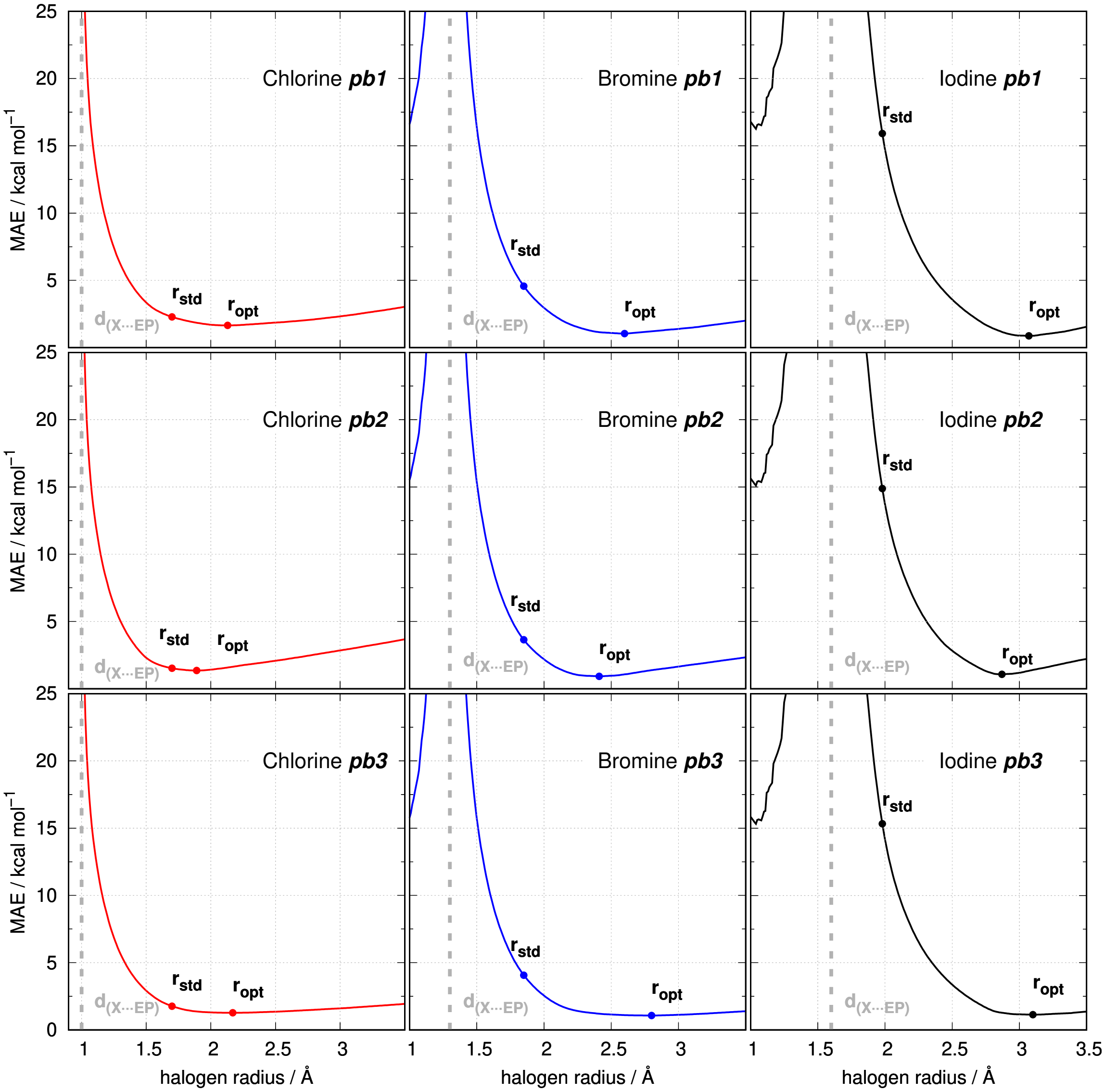

Figure 5: Variation of the MAE values with the halogen radius for each subset of halogenated compounds with an addition of an EP based in model ***EP3*** using PBSA setups ***pb1***, ***pb2***, and ***pb3***.

Table 3: MAE (kcal mol^{-1}) values obtained with standard (r_{std}) and optimized (r_{opt}) halogen radii (Å) for each subset of halogenated compounds with the addition of an EP using ***EP3*** model for PBSA setups ***pb1***, ***pb2***, and ***pb3***.

Halogen	r_{std}	***pb1***		***pb2***		***pb3***	
		MAE$_{\mathrm{std}}$	MAE$_{\mathrm{opt}}$(r_{opt})	MAE$_{\mathrm{std}}$	MAE$_{\mathrm{opt}}$(r_{opt})	MAE$_{\mathrm{std}}$	MAE$_{\mathrm{opt}}$(r_{opt})
Cl	1.70	2.281	1.648 (2.13)	1.525	1.352 (1.89)	1.766	1.273 (2.17)
Br	1.85	4.564	1.034 (2.60)	3.639	0.924 (2.41)	4.067	1.069 (2.80)
I	1.98	15.92	0.869 (3.07)	14.89	1.074 (2.87)	15.33	1.131 (3.10)

molecules, the errors obtained using r_{std} values were significantly large (> c.a. 15 kcal mol^{-1}) for all tested PBSA setups. This is caused by the fact that r_{std} is located very near the broad error peak. Thus, the standard PB radii for iodine (1.98 Å) along with and a X$\cdots$EP distance of 1.6 Å without charge refitting is not recommended in spite of the distances fulfill the criterion of placing the off-center point-charge inside the solute dielectric ($r_{\mathrm{opt}} > \mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$), again, highlighting the importance of the halogen PB radii optimization. Indeed, by using the optimized iodine PB radii, much lower MAE values (c.a. 1 kcal mol^{-1}) are obtained. For brominated molecules, the MAE values obtained using standard PB radii are also relatively large (c.a. 4 kcal mol^{-1}) whereas, noticeably, those obtained for the chlorinated database are not particularly high ($\approx$ 2 kcal mol^{-1}) since r_{std} is substantially far from the broad error peak located a 1.00 Å (Figure 5). Nonetheless, the MAE values decrease, specially in brominated molecules, when optimized halogen PB radii are used.

Again, for all the PBSA setups, the r_{opt} values follow the expected order of increasing atomic number (Cl < Br < I) and the setup that leads to lower MAE errors is dependent on the halogen set. For iodinated molecules, ***pb1*** performs better (MAE = 0.869 kcal mol^{-1}), while for brominated and chlorinated, ***pb2*** an ***pb3*** yield better results (0.924 kcal mol^{-1} and 1.273 kcal mol^{-1}, respectively). These are, however, slightly higher than those obtained for ***EP2*** in this work and for ***EP1*** in an earlier study,[38] although still acceptable, specially taking into account that in model ***EP3*** the charges are not specifically fitted with the EP. Nonetheless, a recommendation of the PBSA flavour that should be used with this off-center point charge model is not evident. By plotting the calculated and experimental ΔG_{solv} values under optimized conditions (Figure 5), the correlations are high, which was expected considering the low MAE values. However, ***pb3*** yields higher rank correlations for all halogens when compared with ***pb1*** and ***pb2***. It is also noteworthy to mention that our library of compounds qualitatively mimics the amount of each halogenated species present in the majority of databases, i.e. a huge presence of chlorinated molecules in relation to iodinated molecules. If we consider the weighted average of the MAEs for each setup ($\bar{x} = 1.489$ kcal mol^{-1} for ***pb1***; $\bar{x} = 1.264$ kcal mol^{-1} for ***pb2***; $\bar{x} = 1.230$ kcal mol^{-1} for ***pb3***), and the mentioned correlations (Figure 5) the latter one performs the best. As for ***EP2***, no obvious systematic over/underestimation of ΔG_{solv} is observed regardless of the used method (Figure S3 and Figure S4).

We also checked for systematic deviations on each setup/halogenated database. Again 2-iodophenol was the compound that yielded higher deviations (-5.60 to -3.58 kcal mol^{-1}), independently of the PBSA setup, while for brominated species, 4-bromophenol and bromacil are constantly ranked in the higher deviations (up to -2.94 kcal mol^{-1} and -5.13 kcal mol^{-1}, respectively). Once more in the chlorinated dataset, the calculated ΔG_{solv} values for dialifor were considerably more negative than the experimental ones resulting in a high deviation (up to -9.39 kcal mol^{-1} in ***pb1***).

Off-center point-charge model *EP4*

In order to test an off-center point-charge model out of the scope of AMBER/GAFF[39] such as ***EP2*** and ***EP3***, and given the similarity of the AMBER force field energy functional form with that of CHARMM,[65] we also performed an optimization of the halogen PB radii in the context of this latter force field. Indeed, recently, the CHARMM General Force Field (CGenFF) was updated with positively-charged virtual particles to describe halogen bonds[30]

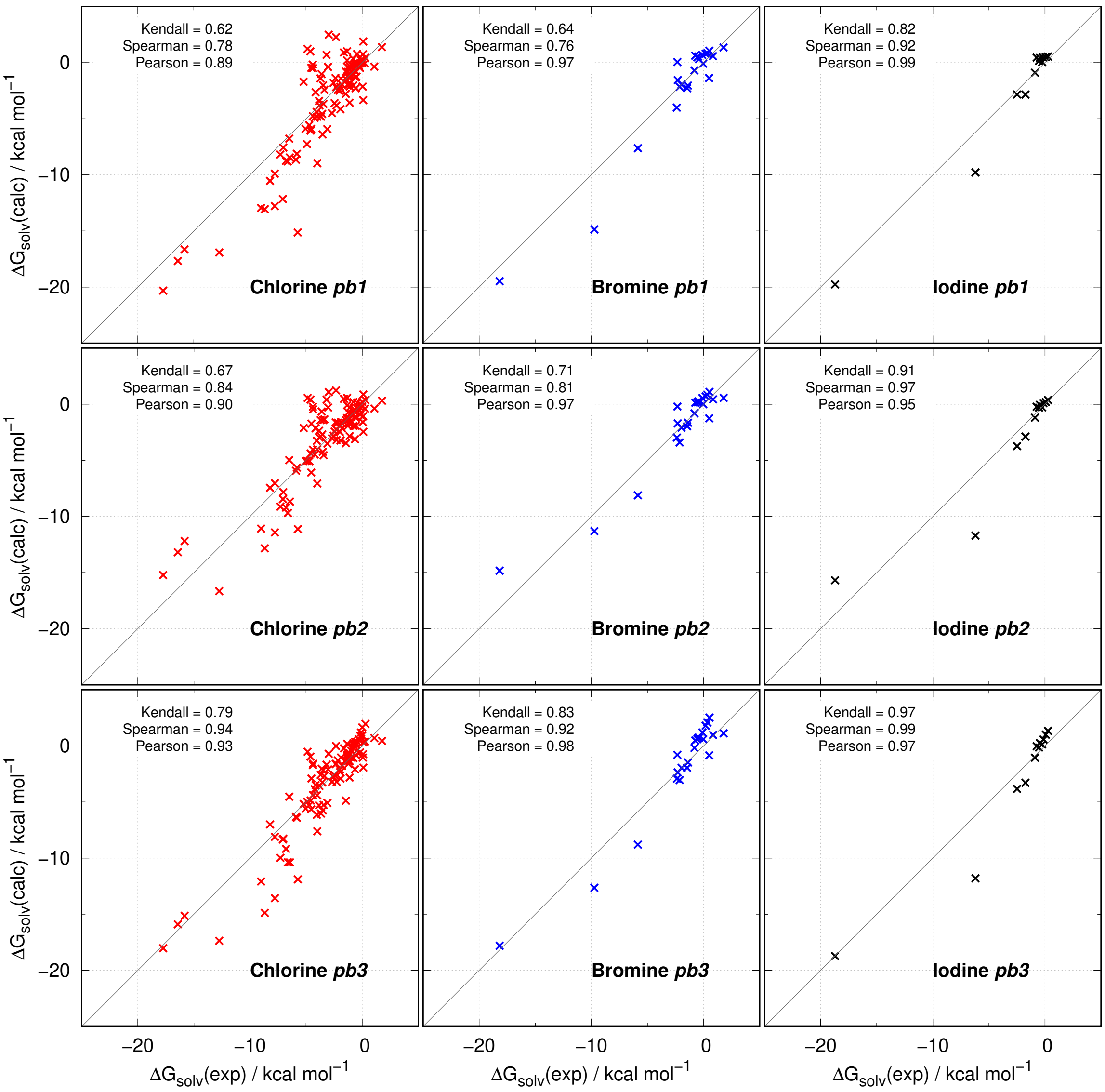

Figure 6: Correlation between experimental an calculated ΔG_{solv} values for each subset of halogenated compounds with with ***EP3*** model. The inset shows the resulting Kendall τ, Spearman rank correlation, and Pearson R correlation coefficients.

with fixed charges (q_{EP}) and distances ($d_{X\cdots EP}$) as summarized in Table 1. Although our set of halogenated molecules is diverse, it is important to notice that the optimization was based on interactions in model systems of halobenzenes and therefore, the strategy of porting the parameters to other molecules should be subject of further analysis. The variations of the MAEs with halogen radius using PBSA setup ***pb4*** is depicted in Figure 7. Not surprisingly,

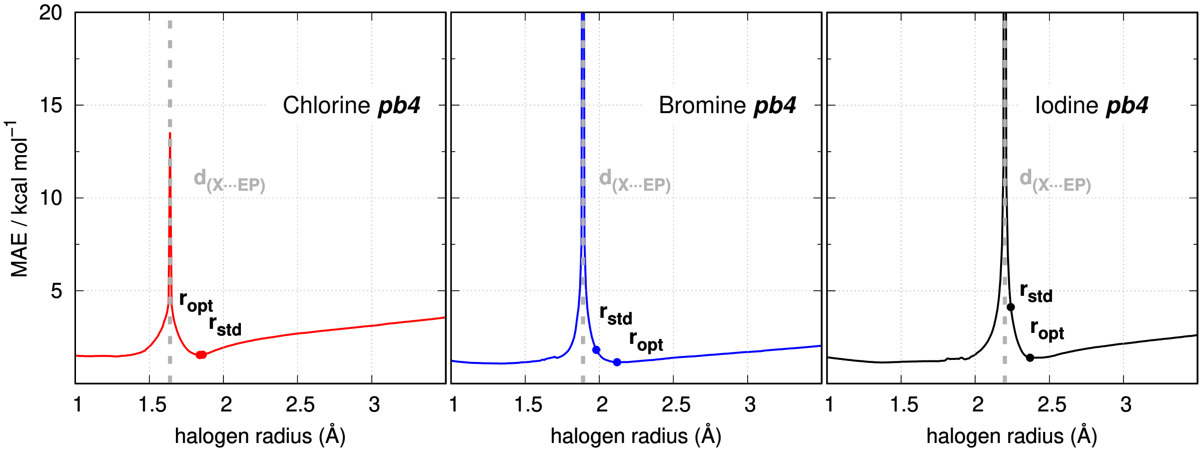

Figure 7: Variation of the MAE values with the halogen radius for each subset of halogenated compounds with an addition of an EP based in model ***EP4*** using PBSA setup ***pb4***.

the behavior of the curves is quite similar to the one observed for ***EP3*** (this work) and ***EP1*** (reference 38) where fixed $d_{X\cdots EP}$ values are used. When the PB radius = $d_{X\cdots EP}$, an error peak is observed while a minimum is observed at PB radius > $d_{X\cdots EP}$. Table 4 reports the MAEs for obtained with r_{std} along with those obtained with r_{opt}. Recall that while for the remainder elements CHARMM-compatible radii were used,[69,70] for the halogens their R_{min} value was used as r_{std}. Surprisingly, for chlorine, the standard radii (1.86 Å) is extremely similar to the optimized one (1.84 Å), yielding MAEs of ≈ 1.5 kcal mol^{-1}. For bromine, radii optimization leads to a slight improvement of the MAE values whereas for iodine, a large improvement of about 3 orders of magnitude is observed upon optimization owing to the fact that r_{std} (2.24 Å) is almost overlapping with $d_{X\cdots EP}$ (2.20 Å). A strong linear correlation between calculated and experimental values (Figure 8) is obtained for iodine and bromine

Table 4: MAE (kcal mol^{-1}) values obtained with standard (r_{std}) and optimized (r_{opt})halogen radii (Å) for each subset of halogenated compounds with the addition of an EP using ***EP4*** model for PBSA setup ***pb4***

Halogen	r_{std}	***pb4***	
		MAE$_{\mathrm{std}}$	MAE$_{\mathrm{opt}}$(r_{opt})
Cl	1.86	1.552	1.535 (1.84)
Br	1.98	1.816	1.148 (2.12)
I	2.24	4.120	1.392 (2.37)

with Pearson correlation coefficients > 0.9, tough a degradation of the monotonic character of the relationship is observed (lower Spearman rank correlation coefficients). For chlorine, the linear correlation (0.76) is the worse amongst tested PBSA setups and EP models and the same degradation of the monotonic character is observed.

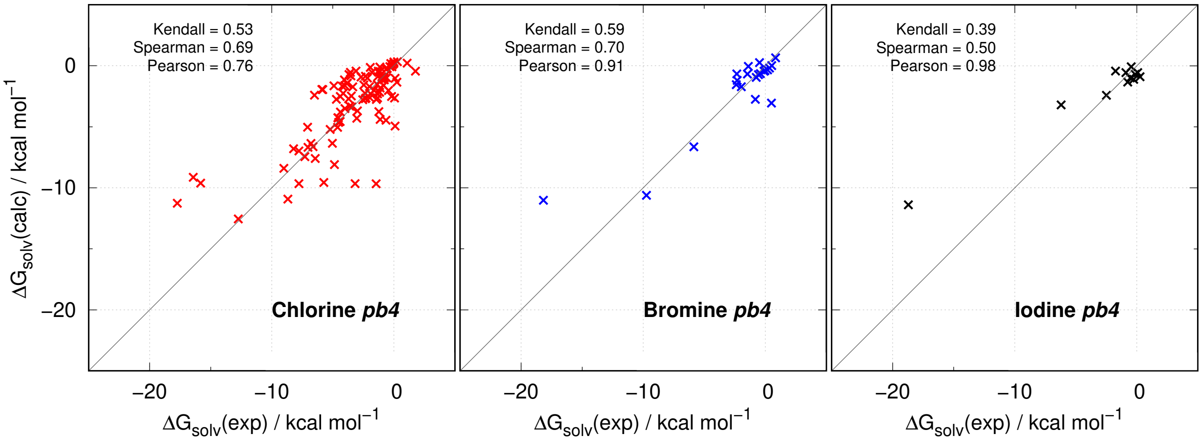

Figure 8: Correlation between experimental an calculated ΔG_{solv} values for each subset of halogenated compounds with with ***EP4*** model and ***pb4*** PBSA setup. The inset shows the resulting Kendall τ, Spearman rank correlation, and Pearson R correlation coefficients.

Opposed to what was observed in the previous sections, using ***EP4*** along with PBSA setup ***pb4*** leads to an underestimation of the calculated values when the experimental ΔG_{solv} values are highly negative (Figure S5). Moreover, while for chlorine and bromine the difference between ΔG_{solv}(calc) and ΔG_{solv}(exp) values is reasonably distributed around 0, for iodine the distribution is tilted towards positive values (Figure S6) owing to a large underestimation of ΔG_{solv} for 2-iodophenol (experimental -6.20 kcal mol^{-1} *vs* calculated -3.19 kcal mol^{-1}) and 5-iodouracil (experimental -18.72 kcal mol^{-1} *vs* -11.39 kcal mol^{-1}). For brominated compounds the largest deviation is observed for 5-bromouracil (an underestimation of c.a. 7 kcal mol^{-1}) whereas 4-bromophenol and bromacil which were difficult cases using the previous EP models/PBSA setups are now nicely predicted (deviations of -0.79 kcal mol^{-1} and -0.89 kcal mol^{-1}, respectively). For chlorinated molecules, the difficult case of dialifor has only a moderate deviation (-3.82 kcal mol^{-1}) specially taking into account the deviations found in the other EP models while the largest outlier is now trichloronitromethane (experimental -1.45 kcal mol^{-1}, calculated -9.67 kcal mol^{-1}).

Conclusions

In classical force field methods, the σ–hole of halogen atoms is often emulated using an off-center point-charge (EP) placed along the R–X bond. The inclusion of this EP might introduce artifacts in the Poisson–Boltzman calculation if the EP is placed at the dielectric interface separating the solute (treated explicitly) and solvent media (treated implicitly), i.e., if the halogen PB radius is equal to the X–EP distance ($\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$).[38] Moreover, standard halogen PB radii (r_{std}) found in the literature can be smaller than $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$, thus yielding a non-realistic model where the EP is placed in the solvent dielectric. This occurs specially for models placing the off-center point-charge at a fixed distance corresponding to the Lennard-Jones parameter R_{min}[26] as suggested for AMBER/GAFF (herein termed ***EP1***), for which a set of compatible halogen PB radii were previously provided.[38]

Since other EP models with diverse $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values and associated charged sets exist, a need for parameters for such models is paramount. In the scope of AMBER/GAFF we selected two different models, one with variable $\mathrm{d}_{\mathrm{X}\cdots\mathrm{EP}}$ values, selected according to the best fit to a reference quantum-mechanical electrostatic potential (***EP2***)[27] and another (***EP3***) with a fixed $\mathrm{d_{X\cdots EP}}$ value and EP charge ($\mathrm{q_{EP}}$).[29] By taking a set of 142 halogenated compounds for which experimental ΔG_{solv} values are known, we conducted a PB radii optimization using three different PBSA calculation setups. Owing to the variable nature of $\mathrm{d_{X\cdots EP}}$ in ***EP2*** the variation of the mean absolute errors (MAEs) with the halogen radius shows several error peaks whenever $\mathrm{d_{X\cdots EP}}$ = PB radius for a given compound. Nonetheless, minima could be located and the optimized radii provide MAE values substantially lower than those obtained in standard conditions. Using the best performing PBSA setup which treats the attractive and repulsive interactions separately in the calculation of $\Delta G_{\mathrm{nonpolar}}$ (***pb3***), the largest error was c.a. 1 kcal mol^{-1} for chlorine, the method also providing excellent linear correlations between calculated and experimental values. The optimized radii and respective MAE values are extremly similar to those obtained for ***EP1***[38] which is explained by the fact that the distribution of $\mathrm{d_{X\cdots EP}}$ values in model ***EP2*** possesses maxima values very close to R_{min} (where the EP is placed in ***EP1***). For ***EP3*** model, a single but broad error peak occurs at $\mathrm{d_{X\cdots EP}}$ = PB radius, with ***pb3*** also being recommended. The optimized radii for this model are different from those obtained earlier with ***EP1*** and ***EP2*** in this work and the errors under optimized conditions are slightly larger but still acceptable (MAEs $<$ 1.3 kcal mol^{-1} using ***pb3***).

Aiming at expanding the range of force fields, we also provide the first optimized halogen PB radii that can be used in the context of the CHARMM General Force Field (CGenFF) which was recently updated to include positively-charged virtual particles to describe halogen bonds (***EP4***).[30] Since no standard PB radii were available for halogens, we initially took the R_{min} values as placeholders for r_{std}. Surprisingly, for chlorine, the R_{min} value is a good approximation for the PB radii as the optimized one is similar (1.86 Å and 1.84 Å, respectively), yielding similar MAE values (c.a. 1.5 kcal mol^{-1}), while for bromine and iodine, the r_{opt} are slightly larger than r_{std}, providing smaller MAE values (c.a. 1.1 kcal mol^{-1} and 1.4 kcal mol^{-1}, respectively).

A correct description of halogen bonds is extremely important in the context of computer-aided drug design[21] and the above EP models were developed to tackle this issue. However, the discussion of which model/force field performs better for such task is out of the scope of this manuscript but it could depend on the target system. After an appropriate choice of the off-center point-charge model to perform MM/MD simulations, we now provide the user with reliable parameters to perform PBSA calculations when halogenated compounds are involved which could prove very useful in the context of virtual screening.

Supporting Information Available

Supporting Figures and Tables (PDF)

Full list of compounds contained in each subset along with individual experimental and calculated ΔG_{solv} values for each EP model / PBSA setup under optimized conditions (CSV)

Acknowledgement

The authors thank Fundação para a Ciência e a Tecnologia (FCT), Portugal, for the doctoral grant SFRH/BD/146447/2019 (A.F.), and strategic projects UIDB/04046/2020 - UIDP/04046/2020 (BioISI) and UID/DTP/04138/2019 (iMed.ULisboa). This work was financed by FCT, Programa Operacional Regional de Lisboa (Lisboa 2020), Portugal 2020, FEDER/FN, and the European Union under project number 28455 (LISBOA-01-0145-FEDER-028455, PTDC/QUI-QFI/28455/2017).

References

(1) Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, *303*, 1813–1818.

(2) Mishra, S. K.; Koča, J. Assessing the performance of MM/PBSA, MM/GBSA, and QM–MM/GBSA approaches on protein/carbohydrate complexes: effect of implicit solvent models, QM methods, and entropic contributions. *J. Phys. Chem. B* **2018**, *122*, 8113–8121.

(3) Perlovich, G. L.; Bauer-Brandl, A. Solvation of drugs as a key for understanding partitioning and passive transport exemplified by NSAIDs. *Curr. Drug Deliv.* **2004**, *1*, 213–226.

(4) Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D. Estimation of absolute free energies of hydration using continuum methods: accuracy of partial charge models and optimization of nonpolar contributions. *J. Chem. Theory Comput.* **2006**, *2*, 128–139.

(5) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6*, 1509–1519.

(6) Christ, C. D.; Fox, T. Accuracy assessment and automation of free energy calculations for drug design. *J. Chem. Inf. Model* **2014**, *54*, 108–120.

(7) Cabeza de Vaca, I.; Qian, Y.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. Enhanced Monte Carlo methods for modeling proteins including computation of absolute free energies of binding. *J. Chem. Theory Comput.* **2018**, *14*, 3279–3288.

(8) Martins, S. A.; Sousa, S. F.; Ramos, M. J.; Fernandes, P. A. Prediction of solvation free energies with thermodynamic integration using the general AMBER force field. *J. Chem. Theory Comput.* **2014**, *10*, 3570–3577.

(9) Khavrutskii, I. V.; Wallqvist, A. Improved binding free energy predictions from single-reference thermodynamic integration augmented with hamiltonian replica exchange. *J. Chem. Theory Comput.* **2011**, *7*, 3001–3011.

(10) Wang, C.; Nguyen, P. H.; Pham, K.; Huynh, D.; Le, T.-B. N.; Wang, H.; Ren, P.; Luo, R. Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *J. Comput. Chem.* **2016**, *37*, 2436–2446.

(11) Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R. Recent Developments and Applications of the MMPBSA Method. *Frontiers in molecular biosciences* **2017**, *4*, 87.

(12) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen bonding: the σ-hole. *J. Mol. Model.* **2007**, *13*, 291–296.

(13) Desiraju, G. R.; Ho, P. S.; Kloo, L.; Legon, A. C.; Marquardt, R.; Metrangolo, P.; Politzer, P.; Resnati, G.; Rissanen, K. Definition of the halogen bond (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013**, *85*, 1711–1713.

(14) Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. The halogen bond. *Chem. Rev.* **2016**, *116*, 2478–2601.

(15) Costa, P. J. The halogen bond: nature and applications. *Phys. Sci. Rev.* **2017**, *2*, 20170136.

(16) Brown, A.; Beer, P. D. Halogen bonding anion recognition. *Chem. Commun.* **2016**, *52*, 8645–8658.

(17) Nunes, R.; Costa, P. J. Ion-Pair Halogen Bonds in 2-Halo-Functionalized Imidazolium Chloride Receptors: Substituent and Solvent Effects. *Chem. Asian J.* **2017**, *12*, 586–594.

(18) Sutar, R. L.; Huber, S. M. Catalysis of organic reactions through halogen bonding. *ACS Catalysis* **2019**, *9*, 9622–9639.

(19) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16789–16794.

(20) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J. Med. Chem.* **2013**, *56*, 1363–1388.

(21) Costa, P. J.; Nunes, R.; Vila-Viçosa, D. Halogen bonding in halocarbon-protein complexes and computational tools for rational drug design. *Expert Opinion on Drug Discovery* **2019**, *14*, 805–820.

(22) Lu, Y.; Shi, T.; Wang, Y.; Yang, H.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W. Halogen bonding–a novel interaction for rational drug design? *J. Med. Chem.* **2009**, *52*, 2854–2862.

(23) Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A. Halogenation of drugs enhances membrane binding and permeation. *ChemBioChem* **2004**, *5*, 676–684.

(24) Ford, M. C.; Ho, P. S. Computational tools to model halogen bonds in medicinal chemistry. *J. Med. Chem.* **2016**, *59*, 1655–1670.

(25) Kolář, M. H.; Hobza, P. Computer modeling of halogen bonds and other σ-hole interactions. *Chem. Rev.* **2016**, *116*, 5155–5187.

(26) Ibrahim, M. A. Molecular mechanical study of halogen bonding in drug discovery. *J. Comput. Chem.* **2011**, *32*, 2564–2574.

(27) Rendine, S.; Pieraccini, S.; Forni, A.; Sironi, M. Halogen bonding in ligand–receptor systems in the framework of classical force fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19508–19516.

(28) Jorgensen, W. L.; Schyman, P. Treatment of halogen bonding in the OPLS-AA force field: application to potent anti-HIV agents. *J. Chem. Theory Comput.* **2012**, *8*, 3895–3901.

(29) Kolář, M.; Hobza, P. On extension of the current biomolecular empirical force field for the description of halogen bonds. *J. Chem. Theory Comput.* **2012**, *8*, 1325–1333.

(30) Gutiérrez, I. S.; Lin, F.-Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L.; MacKerell, A. D. Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand–protein interactions. *Bioorg. Med. Chem.* **2016**, *24*, 4812–4825.

(31) Franchini, D.; Dapiaggi, F.; Pieraccini, S.; Forni, A.; Sironi, M. Halogen bonding in the framework of classical force fields: The case of chlorine. *Chem. Phys. Lett.* **2018**, *712*, 89–94.

(32) Nunes, R.; Vila-Viçosa, D.; Machuqueiro, M.; Costa, P. J. Biomolecular Simulations of Halogen Bonds with a GROMOS Force Field. *J. Chem. Theory Comput.* **2018**, *14*, 5383–5392.

(33) Titov, O. I.; Shulga, D. A.; Palyulin, V. A. Quadrupole correction: from molecular electrostatic potential to free energies of halogen bonding. *J. Chem. Theory Comput.* **2018**, *15*, 1159–1167.

(34) Ibrahim, M. A. A. Molecular mechanical perspective on halogen bonding. *J. Mol. Model.* **2012**, *18*, 4625–4638.

(35) Luchi, A. M.; Angelina, E. L.; Andujar, S. A.; Enriz, R. D.; Peruchena, N. M. Halogen bonding in biological context: a computational study of D2 dopamine receptor. *J. Phys. Org. Chem.* **2016**, *29*, 645–655.

(36) Ibrahim, M. A. A.; Hasb, A. A. M.; Mekhemer, G. A. H. Role and nature of halogen bonding in inhibitor$\cdots$receptor complexes for drug discovery: casein kinase-2 (CK2) inhibition as a case study. *Theor. Chem. Acc.* **2018**, *137*, 38.

(37) Zhou, Y.; Wang, Y.; Li, P.; Huang, X.-P.; Qi, X.; Du, Y.; Huang, N. Exploring Halogen Bonds in 5-Hydroxytryptamine 2B Receptor–Ligand Interactions. *ACS Med. Chem. Lett.* **2018**, *9*, 1019–1024.

(38) Nunes, R.; Vila-Viçosa, D.; Costa, P. J. Tackling Halogenated Species with PBSA: Effect of Emulating the σ-hole. *J. Chem. Theory Comput.* **2019**, *15*, 4241–4251.

(39) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(40) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Chem. Phys.* **1993**, *97*, 10269–10280.

(41) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.

(42) Tsui, V.; Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **2000**, *56*, 275–291.

(43) Onufriev, A. V.; Aguilar, B. Accuracy of continuum electrostatic calculations based on three common dielectric boundary definitions. *J. Theor. Comput. Chem.* **2014**, *13*, 1440006.

(44) Horton, J.; Allen, A. E.; Dodda, L. S.; Cole, D. J. QUBEKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Theory Comput.* **2019**,

(45) Yan, X. C.; Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Description of Sulfur Charge Anisotropy in OPLS Force Fields: Model Development and Parameterization. *J. Phys. Chem. B* **2017**, *121*, 6626–6636.

(46) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Hardner, E. D. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**,

(47) Mobley, D. L.; Guthrie, J. P. FreeSolv: a database of experimental and calculated hydration free energies, with input files. *J. Comput. Aided Mol. Design* **2014**, *28*, 711–720.

(48) Mobley, David L. `http://www.escholarship.org/uc/item/6sd403pz`, 2013; Experimental and Calculated Small Molecule Hydration Free Energies. UC Irvine: Department of Pharmaceutical Sciences, UCI.

(49) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

(50) McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z= 11–18. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

(51) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(52) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(53) Curtiss, L. A.; McGrath, M. P.; Blaudeau, J.-P.; Davis, N. E.; Binning Jr, R. C.; Radom, L. Extension of Gaussian-2 theory to molecules containing third-row atoms Ga–Kr. *J. Chem, Phys.* **1995**, *103*, 6104–6113.

(54) Glukhovtsev, M. N.; Pross, A.; McGrath, M. P.; Radom, L. Extension of Gaussian-2

(G2) theory to bromine-and iodine-containing molecules: Use of effective core potentials. *J. Chem. Phys.* **1995**, *103*, 1878–1885.

(55) Frisch, M. J. et al. Gaussian 09 Revision A.2. Gaussian Inc. Wallingford CT 2009.

(56) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(57) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

(58) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* basis set for third-row atoms. *J. Comput. Chem.* **2001**, *22*, 976–984.

(59) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

(60) Case, D. A. et al. AMBER 2017. University of California, San Francisco, 2017.

(61) Kolář, M.; Hobza, P.; Bronowska, A. K. Plugging the explicit σ-holes in molecular docking. *Chem. Commun.* **2013**, *49*, 981–983.

(62) Kim, S.; Lee, J.; Jo, S.; Brooks III, C. L.; Lee, H. S.; Im, W. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *Journal of computational chemistry* **2017**, *38*, 1879–1886.

(63) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **2008**, *29*, 1859–1865.

(64) Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr, A. D. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.

(65) Crowley, M. F.; Williamson, M. J.; Walker, R. C. CHAMBER: Comprehensive support for CHARMM force fields within the AMBER software. *International Journal of Quantum Chemistry* **2009**, *109*, 3767–3772.

(66) Bordner, A.; Cavasotto, C.; Abagyan, R. Accurate transferable model for water, n-octanol, and n-hexadecane solvation free energies. *J. Phys. Chem. B* **2002**, *106*, 11009–11015.

(67) Tan, C.; Tan, Y.-H.; Luo, R. Implicit nonpolar solvent models. *J. Phys. Chem. B* **2007**, *111*, 12263–12274.

(68) Aleksandrov, A.; Lin, F.-Y.; Roux, B.; MacKerell Jr, A. D. Combining the polarizable Drude force field with a continuum electrostatic Poisson–Boltzmann implicit solvation model. *J. Comput. Chem.* **2018**, *39*, 1707–1719.

(69) Nina, M.; Beglov, D.; Roux, B. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.

(70) Banavali, N. K.; Roux, B. Atomic radii for continuum electrostatics calculations on nucleic acids. *J. Phys. Chem. B* **2002**, *106*, 11026–11035.

(71) Riquelme, M.; Lara, A.; Mobley, D. L.; Verstraelen, T.; Matamala, A. R.; Vöhringer-Martinez, E. Hydration Free Energies in the FreeSolv Database Calculated with Polarized Iterative Hirshfeld Charges. *J. Chem. Inf. Model.* **2018**, *58*, 1779–1797.

(72) Guthrie, J. P. A blind challenge for computational solvation free energies: introduction and overview. *J. Phys. Chem. B* **2009**, *113*, 4501–4507.

Graphical TOC Entry