Published in final edited form as:

Bioorg Med Chem. 2016 October 15; 24(20): 4812–4825. doi:10.1016/j.bmc.2016.06.034.

Parametrization of Halogen Bonds in the CHARMM General Force Field: Improved treatment of ligand-protein interactions

Ignacio Soteras Gutiérrez¹, Fang-Yu Lin^{†,1}, Kenno Vanommeslaeghe^{†,‡}, Justin A. Lemkul[†], Kira A. Armacost[§], Charles L. Brooks III[§], and Alexander D. MacKerell Jr[†]

Alexander D. MacKerell: alex@outerbanks.umaryland.edu

[†]Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, U.S

Life Sciences Department. Barcelona Super Computing Center. Jordi Girona 29, Barcelona 08034, Spain

[#]Department of Analytical Chemistry and Pharmaceutical Technology (FABI), Center for Pharmaceutical Research (CePhaR), Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium

[§]Department of Chemistry and Biophysics, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, U.S

Abstract

A halogen bond is a highly directional, non-covalent interaction between a halogen atom and another electronegative atom. It arises due to the formation of a small region of positive electrostatic potential opposite the covalent bond to the halogen, called the “sigma hole.” Empirical force fields in which the electrostatic interactions are represented by atom-centered point charges cannot capture this effect because halogen atoms usually carry a negative charge and therefore interact unfavorably with other electronegative atoms. A strategy to overcome this problem is to attach a positively charged virtual particle to the halogen. In this work, we extend the additive CHARMM General Force Field (CGenFF) to include such interactions in model systems of phenyl-X, with X being Cl, Br or I including di- and trihalogenated species. The charges, Lennard-Jones parameters, and halogen-virtual particle distances were optimized to reproduce the orientation dependence of quantum mechanical interaction energies with water, acetone, and N-methylacetamide as well as experimental pure liquid properties and relative hydration free energies with respect to benzene. The resulting parameters were validated in molecular dynamics simulations on small-molecule crystals and on solvated protein-ligand complexes containing halogenated compounds. The inclusion of positive virtual sites leads to better agreement across experimental observables, including preservation of ligand binding poses as a direct result of the improved representation of halogen bonding.

Correspondence to: Alexander D. MacKerell, Jr, alex@outerbanks.umaryland.edu.

¹These authors contributed equally to the study.

Keywords

sigma hole; halogen bond; CHARMM; halogen; drug design; CGenFF

1. Introduction

Halogen atoms are added to compounds during drug design to optimize binding affinities,^{1–3} to improve oral absorption and/or to enhance blood-brain barrier (BBB) permeation.⁴ A number of studies have shown that one of the reasons for the success of this strategy is the formation of non-bonded interactions referred to as “halogen-bonds.”⁵ A halogen bond is a highly directional, non-covalent interaction between a halogen atom and another electronegative atom. Typical halogen bond strengths vary between 0.5 and 10.5 kcal·mol^{–1}.⁶ The origin of this interaction is the formation of a small region of positive charge opposite the covalent bond to the halogen called the “sigma hole.”⁷ Based on energy decomposition analysis it has been suggested that this interaction is not only driven by electrostatics, but dispersion also makes a significant contribution.^{8,9}

Currently, molecular simulations of protein-ligand complexes are typically performed using force fields such as AMBER,¹⁰ OPLS,^{11,12} GROMOS,¹³ and CHARMM.¹⁴ As these Class I additive force fields represent electrostatic interactions using fixed atom-centered point charges, they typically do not adequately treat halogen bonds. Specifically, the electronegative halogen atoms typically bear negative charges, giving rise to repulsive interactions with other electronegative atoms that cannot be counterbalanced by other terms in the force field. A strategy to overcome this problem is to attach a charged virtual particle to the halogen,^{15–17} which will be referred to as a “sigma hole particle” in this remainder of the study.

The goal of this study is to extend the CHARMM additive force field to provide more accurate treatment of halogen bonds involving halogenated aromatic rings. The sigma hole particle strategy was chosen for its simplicity and successful predictive capabilities in other force fields,^{16–18} though other strategies exist.¹⁹ The charges, Lennard-Jones (LJ) parameters, and halogen-virtual particle distances were tuned to reproduce the orientation dependence of quantum mechanical (QM) interaction energies with water, acetone, and N-methylacetamide as well as experimental pure liquid properties, and relative hydration free energies with respect to benzene. Once parameters for the mono-halogenated models systems were obtained, we extended the parametrization to include phenyl rings with up to three halogen atoms in all possible configurations (Figures 1–3). The resulting parameters were validated in a set of small-molecule crystals containing halogenated compounds and by carrying out molecular dynamics simulations of solvated protein-ligand complexes in which halogen bonding contributes to ligand binding, including the proteins: human thrombin, *S. aureus* enoyl-ACP reductase saFabI, *P. falciparum* dihydrofolate reductase C59R/S108N double mutant and *A. orientalis* TDP vancosaminyltransferase GtfD. The inclusion of the sigma hole particle leads to better agreement across experimental observables, including preservation of ligand binding poses as a direct result of the improved representation of halogen bonding.

2. Computational Methods

2.1. Quantum mechanical calculations

QM geometry optimizations were performed with Gaussian03.²⁰ For chlorinated species, the gas phase optimizations were performed at the MP2 level of theory with the aug-cc-pVDZ basis set.²¹ For brominated and iodinated species, the optimizations were performed at the MP2 level of theory with aug-cc-pVDZ-(PP) basis set,^{22,23} in which the inner shells are represented by pseudopotentials that account for relativistic effects (and have a computational performance advantage). Single-point interaction energies for chlorobenzene with water, acetone, and NMA were obtained at the RIMP2 level of theory with the cc-pVQZ basis set²¹ using Q-Chem, version 4.0.1.²⁴ For bromobenzene and iodobenzene, single-point interaction energies with water and acetone were calculated with the sdb-cc-pVQZ basis set²⁵ at the MP2 level of theory with NWChem, version 6.0.²⁶ All interaction energies were corrected for basis set superposition error (BSSE) using the counterpoise method.²⁷ Interaction energies were calculated with the halogenated species monomer gas phase optimized geometries, water kept rigid in the TIP3P geometry and with acetone and NMA maintained in their respective gas phase MP2/aug-cc-pVDZ-optimized geometries.

The initial parameters of chlorobenzene, bromobenzene and iodobenzene were obtained from the additive CHARMM General Force Field (CGenFF).¹⁴ The sigma hole was represented by a massless, virtual particle attached to the halogen along the C-X bond axis (X = Cl, Br, or I) at a fixed distance, which bears a small positive charge. Optimization involved only the charge and the LJ terms (well depth, ϵ , and $R_{\min}/2$) for halogen atoms, the charge of the carbon atom to which the halogen is bound (C_{ipso}), and the distance between the halogen and the sigma hole particle. The optimization initially targeted intermolecular interactions of the halogenated compounds with water and acetone obtained by QM calculations. Two types of interaction energy profiles were computed. In the first interaction energy profile, referred as “radial scan” (Figure 4A), the oxygen atom of water and the O=C moiety of acetone were located along the C-X bond axis of the halogenated compounds. In these orientations, the water atoms and the non-hydrogen atoms of acetone were located in the plane of the aromatic ring. The interaction energy profile was calculated by varying the distance between the halogen and the oxygen atoms from 2.0 to 5.0 Å in intervals of 0.1 Å. In the second interaction energy profiles, referred as “angular scan” (Figure 4B), the C-X...O angle was varied from 90° to 270° in 5° intervals, keeping the X...O distance (d) at several selected values around the minimum found for the “radial” profiles. To optimize the atom pair-specific LJ parameters (CHARMM^{28–30} keyword NBFIX) between chlorine and the amide carbonyl oxygen, radial and angular scans were performed for chlorobenzene-NMA interactions using the same method, in which the oxygen of NMA was positioned along the C-X bond axis of chlorobenzene in the radial scan and the C-Cl...O angle was sampled in the angular scan.

The parametrization was extended to compounds with up to three halogens in all possible relative positions, namely 1,2-, 1,3-, 1,4-, 1,2,3-, 1,2,4-, and 1,3,5-halogenated benzenes. The sigma hole particle was retained for all halogens, and its charge and distance to the halogen were transferred from the monohalogenated compounds. The remaining partial

atomic charges in the molecules were then adjusted until realistic dipole and quadrupole moments (with respect to their respective MP2/6-31G(d)³¹ values for chloro- and bromo-substituted benzenes, and MP2/6-311G* for iodo-substituted benzenes) were achieved. Charges of CH groups not adjacent to any C-X (X = Cl, Br, I) group were set to the same value as in benzene, -0.115 and +0.115 for carbon and hydrogen atoms, respectively.

2.2. Pure liquid simulations

LJ parameters of chlorine, bromine, and iodine were next optimized by targeting experimental pure liquid properties and relative hydration free energies with respect to benzene. Pure liquid properties included the enthalpy of vaporization and the molecular volume, and were computed by simulating the pure liquids using CHARMM²⁸⁻³⁰ following the procedure previously reported by our group.¹⁴ Briefly, the strategy was to construct a cubic unit cell by placing a copy of the molecule of interest on each grid point of an evenly spaced $6 \times 6 \times 6$ lattice at the experimental liquid density. For each system, an 11-ns NPT simulation was performed with the last 10 ns used for analysis, for which the per-molecule energy and the molecular volume in the condensed phase were obtained. The corresponding gas-phase energy was obtained by performing a 50-ps equilibration followed by a 50-ps production simulation on each of the molecules from the pure liquid systems (ie. 216 individual gas-phase simulations), as required for calculation of the heat of vaporization. All properties were calculated based on the average of three independent sets of simulations using different random number seeds to assign the velocities.

2.3. Free energy calculations

Relative hydration free energies with respect to benzene (G_{hydr}) of chlorobenzene, bromobenzene, iodobenzene, 1,2-dichlorobenzene, 1,3-dichlorobenzene, 1,4-dichlorobenzene, 1,2,3-trichlorobenzene, 1,2,4-trichlorobenzene, 1,3,5-trichlorobenzene, and 1,4-dibromobenzene were calculated via the free energy perturbation method. These free energy differences can be compared to experimental G_{hydr} using the thermodynamic cycle depicted in Figure 5. G_{hydr} is defined as $G_{benz \rightarrow X} = G_X - G_{benz}$, which can be computed as the difference: $G_{solv-benz \rightarrow X} - G_{gas-benz \rightarrow X}$, wherein X is the halogenated benzene. In the last expression, G_X and G_{benz} correspond to the absolute solvation free energies of X and benzene, respectively, whereas $G_{solv-benz \rightarrow X}$ and $G_{gas-benz \rightarrow X}$ correspond to the free energy difference of transforming benzene into X in solution and in the gas phase, respectively. G_{hydr} can also be obtained via a reverse process by transforming benzene into X in solution and in gas phase through calculating the difference: $-(G_{solv-benz \rightarrow X} - G_{gas-benz \rightarrow X})$. The G_{hydr} for halogenated benzenes were calculated by CHARMM²⁸⁻³⁰ and GROMACS (version 5.1),³² in both directions using the approaches described in the following paragraphs.

2.4. Relative free energy differences using CHARMM

The electrostatic contribution to the relative hydration free energies of the halogenated benzenes were calculated using thermodynamic integration (TI) in the PERT module of the 39b1 version of CHARMM.²⁸⁻³⁰ The benzene systems were solvated in a $35 \text{ \AA}^3$ box of CHARMM-modified TIP3P³⁴⁻³⁶ water consisting of 1416 water molecules at 298 K. The solvated structures were first minimized for 1000 steps of Adopted Basis Newton-Rhapon

(ABNR) minimization. Further, each lambda window was subjected to 100 ps of equilibration and a 1-ns production stage. Long-range electrostatics were treated with the particle mesh Ewald (PME) method,^{37,38} with $\kappa = 0.32 \text{ \AA}^{-1}$, an interpolation order of 6, and FFT grid spacing of approximately $1.0 \text{ \AA}$. The LJ interactions were switched smoothly to zero from 10 to $12 \text{ \AA}$ (VSWITCH in CHARMM nomenclature). The λ schedule was comprised of 21 λ windows, spaced evenly by 0.05. van der Waals (vdW) soft-core potentials were used to overcome potential endpoint instabilities.^{39,40} Five simulations were carried out for each system, and the LJ contribution to the relative free energy values were estimated using the Weighted Histogram Analysis Method (WHAM)⁴¹ by iterating until a tolerance of $0.001 \text{ kcal mol}^{-1}$ was reached. To account for truncation of the LJ interactions, a long-range correction (LRC)⁴² was included in the relative free energy calculation. The LRC was calculated by taking the difference in the van der Waals solvent-solute interaction energies using two cutoff schemes. Simply, the energies of the 10 final snapshots from each simulation were evaluated using LJ cutoffs of 12 and $50 \text{ \AA}$. The difference in LJ energy from these two schemes was taken as the LRC.

2.5. Relative free energy differences using GROMACS

Halogenated molecule parameters were converted to GROMACS format using an in-house Python script. To calculate G_{hydr} , each molecule was transformed in vacuo and in water, and G_{hydr} was evaluated by computing the free energy difference for each transformation using Bennett acceptance ratio (BAR) method.⁴³ Thirteen lambda states were used to describe the transformation from a halogenated molecule ($\lambda = 0$) to a benzene molecule ($\lambda = 1$). The LJ parameters were transformed in the first seven windows while holding charges in the $\lambda = 0$ state. While the LJ parameters were transformed, the sigma hole particle was gradually shifted inward towards the halogen atom to avoid Coulombic singularities arising from water molecules overlapping with the virtual particle if it was not shielded by the chlorine LJ radius during the transformation. Then, with the van der Waals parameters in the $\lambda = 1$ state, charges were transformed over the last six λ states. Soft-core potentials were used for transformation of the LJ terms, while Coulombic terms were interpolated linearly. For each value of λ during transformation in water, energy minimization, equilibration, and production runs were conducted. Energy minimization was carried out using the steepest descent method with a force tolerance of $10.0 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ and a maximum step size of $0.1 \text{ \AA}$. Equilibration was carried out for 50 ps in the NVT ensemble. Production was carried out for 1 ns in the NPT ensemble. The pressure was maintained at 1 bar using the Parrinello–Rahman barostat^{44,45}. In the equilibration and production runs, a leap-frog stochastic dynamics integrator⁴⁶ was used with a time step of 1 fs. Langevin dynamics method was used to control the temperature at 298 K with an inverse friction coefficient of 1.0 ps^{-1} . LJ interactions were force switched to zero over $10\text{--}12 \text{ \AA}$ (VFSWITCH in CHARMM nomenclature), and electrostatic interactions were treated with the smooth particle mesh Ewald (PME) method,^{37,38} using a real-space cutoff of $12 \text{ \AA}$. The LINCS^{47,48} algorithm was used to constrain all covalent bonds involving hydrogen atoms. For transformations in vacuo, production runs were conducted using infinite cutoffs after energy minimization of the compounds.

2.6. Crystal simulations

Crystal simulations on small molecules obtained from the Cambridge Structural Database (CSD)⁴⁹ were performed to determine which of several LJ parameter sets would be applied to chlorine. Twenty crystal structures containing chlorobenzene that involved chlorine-carbonyl oxygen interactions were selected in which the carbonyl was not in NMA for 10 crystals and 10 that involved NMA. The distance between chlorine and the oxygens are in the range of 3.0 to 3.4 Å in the crystal structures. Parameters for the non-halogenated molecules in the crystals were obtained from CGenFF.¹⁴ Crystal simulations were performed on the full unit cells using periodic boundary conditions corresponding to the length and angle parameters of the respective crystals. The Langevin thermostat was used, and each simulation was performed for 11.1 ns under the NPT ensemble. The reference temperature was set to that corresponding to the temperature at which the crystals were obtained. The pressure was maintained at 1 atm, allowing anisotropic relaxation of the unit cell. The last 10 ns were used for analyzing the unit cell vector lengths and the distance between the chlorine and oxygen atoms. The calculated unit cell dimensions and the distances are reported based on the average of three independent simulations.

2.7. Simulations of protein-ligand complexes

The final parameters were validated by molecular dynamics simulation of protein-ligand complexes that involve halogenated ligands. Starting coordinates for all protein-ligand complexes were taken from the Protein Databank: human thrombin (PDB 2ZFF),⁵⁰ *S. aureus* enoyl-ACP reductase saFabI (PDB 4ALI),⁵¹ *P. falciparum* dihydrofolate reductase (DHFR) C59R/S108N double mutant (PDB 1J3J),⁵² human JNK3 MAP kinase (PDB 4Z9L),⁵³ and *A. orientalis* TDP vancosaminyltransferase GtfD (PDB 1RRV).⁵⁴ Missing atoms or residues were added to the protein structures using the internal coordinate builder in CHARMM^{28–30}. To create the congeneric series of thrombin inhibitors, the crystallized active site water molecule was removed, consistent with displacement upon binding of a halogen-substituted inhibitor.⁵⁵ Coordinates for each halogen (Cl, Br, I) were subsequently added to the *meta* position of the ring (Figure S2). For all complexes, the ligands were initially processed in CHARMM^{28–30} to add missing hydrogens, and initial ligand topologies were generated using CGenFF.¹⁴ These topologies were subsequently modified to utilize the newly developed halogen nonbonded parameters and lone pairs. Parameters for the proteins and any bound cofactors (NADPH in saFabI, NADP in DHFR, and thymidine diphosphate in GtfD) were taken from the CHARMM36 force field.^{56,57}

Each protein-ligand complex was centered in a cubic box with a minimum solute-box distance of 10 Å, which was subsequently filled with CHARMM-modified TIP3P water^{34–36} and ~100 mM NaCl (including neutralizing counterions). All MD simulations were carried out with GROMACS,³² version 5.1. Bonds to hydrogen atoms were constrained with LINCS,^{47,48} allowing a 2-fs integration time step. Short-range van der Waals forces were force switched to zero from 10 - 12 Å. Electrostatic interactions were computed with the smooth particle mesh Ewald method^{37,38} with a real-space cutoff of 12 Å. Buffered neighbor lists were maintained using the Verlet cutoff scheme in GROMACS. Each system was energy minimized using the steepest descent method, after which equilibration was carried out for 100 ps under the NVT ensemble using a Langevin thermostat⁴⁶ with a friction

coefficient of 1 ps^{-1} to maintain the temperature at 298 K. Restraints were applied to non-hydrogen atoms in the protein during equilibration and were released at the outset of production simulations, which were conducted using the NPT ensemble. Pressure was maintained at 1 bar using the Parrinello-Rahman barostat^{44,45} with a relaxation time of 1 ps. Production simulations were carried out for 10 ns. All analyses were carried out using facilities within the GROMACS package.

3. Results and Discussion

As halogen bonds are highly orientation dependent, parameter optimization initially focused on the accurate treatment of interactions of the monohalogenated molecules with water and acetone as a function of relative orientation and on gas phase dipole moments, as that was anticipated to provide better G_{hydr} values. The parameters to be determined were halogen-sigma hole particle distances, charges on carbons, halogens and the sigma hole particle, as well as the LJ parameters for the halogens. The sigma hole particle bears a small positive charge in order to capture the positive region on the outer edge of the halogen. Once the initial LJ parameters and the halogen-sigma hole particle distances for the halogens were obtained, these values were applied to polyhalogenated molecules (Figures 1-3) to optimize charges by fitting to QM dipole moments.

As chlorinated benzenes are common in drug design⁵⁸ and more experimental data are available for these compounds relative to the brominated and iodinated species, LJ parametrization of chlorine was emphasized, targeting the reproduction of pure liquid properties and G_{hydr} . Parametrization initially focused on the pure liquid properties, yielding set LJ1. However, application of LJ1 for calculation of G_{hydr} yielded poor agreement with experiment. Accordingly, a second set of LJ parameters was optimized to reproduce the experimental G_{hydr} values as well as the QM radial and angular scans with water and acetone. This effort yielded the LJ2 parameter set. Subsequent QM calculations were undertaken on chlorobenzene-NMA interactions to assure that the developed parameters yielded acceptable interactions with the peptide bond carbonyl group. Doing so led to the introduction of an additional NBFIX term for the amide carbonyl oxygen, which is included in the LJ2 parameter set. Selection of the final parameter set was based on small molecule crystal and protein-ligand simulations. From these calculations, the LJ2 set was selected as the default parameters for chlorobenzenes. However, given the poor agreement with pure solvent heats of vaporization for di- and trichlorinated benzenes (see below), it is suggested that the LJ1 set be used when investigating non-aqueous systems involving chlorobenzenes. Unless noted, the results below are based on the LJ2 set of parameters for the halogenated benzenes.

3.1. Partial charges on the halogenated benzenes

Partial charges for the monohalogenated benzenes, namely chlorobenzene (CHLB), bromobenzene (BROB), and iodobenzene (IODB), are shown in Figures 1-3. Charges assigned to halogen atoms (q_X), the sigma hole virtual particle (q_{lp}), the carbon atom to which the halogen is bound ($q_{\text{C}_{\text{ipso}}}$), the distance between the halogen and the sigma hole particle ($r(\text{lp-X})$), and LJ parameters for the halogens are summarized and compared to the

CGenFF parameters in Table 1. The sigma hole particle attached to the halogen atom has no LJ terms assigned to it and always bears the same charge, $0.05 e$. Attempts to modify this value led to undesirable effects. Specifically, when the charge magnitude was made smaller, the favorable halogen bond with water was lost, which is consistent with a previous study performed using the AMBER force field.¹⁶ Conversely, when the charge was increased, interaction energies with water became too favorable as did the hydration free energy. With the charge on the sigma hole particle fixed, the polarization of the C-X bond decreases with halogen size. We note that maintaining the charge while varying the lp-X distance yielded dipole moments in good agreement with QM and experimental values, as discussed below.

The charges for the polyhalogenated compounds are shown in Figures 1-3 alongside the monohalogenated cases for comparison. As shown in the Figures, the addition of more halogens in the ring lowered the charges on the halogen atoms, as is expected considering the inductive effect. Thus, within the chlorinated series, the charge on chlorine drops from $-0.21 e$ in chlorobenzene to $-0.16 e$ in 1,2-dichlorobenzene (12DCB). This decrease in the charge is not so pronounced in 1,3-dichlorobenzene (13DCB), in which the charges on the chlorine atoms were $-0.19 e$. Conversely, for 1,4-dichlorobenzene (14DCB), the charges on the chlorines are the same as for 12DCB, i.e. $-0.16 e$. Similar trends are observed for the trichlorobenzenes. Specifically, the charges of chlorines in compound 1,2,3-trichlorobenzene (123TCB) were also smaller than the ones derived for CHLB. Furthermore, in this case the charge on the “inner” chlorine atom along the C-X bond axis is smaller than on the adjacent “outer” chlorines. For 1,2,4-trichlorobenzene, the charges on the chlorine atoms are similar to the ones obtained for 12DCB and 123TCB, although there are some differences corresponding to the lower symmetry of this molecule. Finally, the charges found for the chlorines in 1,3,5-trichlorobenzene 135TCB ($-0.17 e$) were also smaller than the ones found for CHLB ($-0.21 e$), and similar to the charges found for the chlorines in 123TCB. Similar trends were observed for the brominated as well as the mono- and diiodinated compounds. This led us to extrapolate the same trends to the triiodinated cases, for which little or no target data is available.

3.2. Dipole moments

Non-polarizable condensed phase force fields generally require model compounds to have dipole moments that are substantially larger than in the gas phase for polar neutral species in order to reproduce condensed phase data.¹⁴ However, the dipole moments of the parametrized gas phase halogenated benzenes are closer to their gas phase experimental or QM target values, as shown in Table 2. This can be explained by observing that a real sigma hole is the result of an anisotropic distribution of negative charge on the halogen atom, X. Specifically, the phenomena leads to a toroidal concentration of negative charge with a centroid that is located along the C-X bond,⁶⁰ giving X a net atomic dipole moment along this bond towards the carbon atom. The same electronic structure also results in a strong net atomic quadrupole moment that counteracts the aforementioned dipole moment in the small region along the same axis, resulting in a localized positive electrostatic potential that is the sigma hole. While our model specifically seeks to capture the latter effect with a positive particle, this results in a reversal of the aforementioned atomic dipole moment, partially counteracting the existing overpolarized dipole moment associated with the C-X bond and

ultimately bringing the molecular dipole moment in closer agreement with the gas-phase experimental values.

While the scaling factors for the dipole moments were close to 1.0 in most of the compounds, the scaling factors found for the quadrupole moments were in the range of 1.6-3.0. It should be emphasized that these values were obtained by trying to attain a compromise between predicted pure liquid properties and G_{hydr} . Keeping this in mind, the range of scaling factors for quadrupole moments was in line with the scaling factor of benzene (2.7); the large value reflects the fact that, in addition to the lack of polarizability in the MM model, it does not contain a representation of benzene's prominent π electron clouds, so its quadrupole moment needs to be further overestimated in order to implicitly capture associated effects, such as the cation- π interaction.⁶¹⁻⁶⁵ Conversely, the scaling factors for the dipole moments are markedly lower than the ones commonly encountered in additive condensed-phase force fields. As explained above, this is a likely consequence of the approximation of representing the sigma hole by a positive point charge.

3.3. Interaction energies

Computed interaction energy profiles of the halogenated benzenes with water, acetone, and NMA are presented in Figures 6 and 7, and the minima of the radial profiles are shown in Table 3. The interaction energies of chlorobenzene were calculated using LJ1 and LJ2 parameters as shown in Figure 6A-6B and Figure 7A-7B, respectively. Additionally, chlorobenzene and NMA interaction energies were calculated using LJ2 plus the NBFIX term for amide carbonyl oxygens as shown in Figure 6C and Figure 7C. Water interaction energies with chlorobenzene were reproduced well by using the LJ2 parameters that have a smaller radius and well depth. Notably, the LJ2 parameters also yielded better agreement for the acetone-water interactions, indicating their potential utility beyond interactions with water alone. In general, the interaction energies for water are in good agreement with the QM data while with acetone the MM energies are typically not as favorable as the QM data. However, the deviations found for acetone are relatively small and were considered acceptable.

In preliminary protein-ligand simulations, interactions of chlorinated aromatic rings with the peptide backbone were highly perturbed (not shown). Accordingly, we investigated the interaction of chlorobenzene with the carbonyl of NMA. Using the LJ2 parameters alone, the interactions with NMA are less favorable compared to QM results (Table 3). Addition of a NBFIX term between the chlorine and the amide carbonyl oxygen yields improved chlorobenzene-NMA interaction energies, which leads to better reproduction of interactions between chlorinated ligands and proteins, as shown below.

3.4. Pure liquid properties and G_{hydr}

Once the charges on the halogenated benzenes were obtained, the LJ parameters of chlorine, bromine, and iodine were adjusted to reproduce experimental pure liquid properties and G_{hydr} . The predicted values of enthalpy of vaporization (H_{vap}), molecular volume (V_m) and G_{hydr} using the final sets of the halogen parameters are summarized in Tables 4-7. The predicted values for G_{hydr} calculated using the GROMACS and CHARMM protocols

are summarized in Table 8. LJ parameters for brominated benzene and iodinated benzene yield good agreement with both the experimental pure liquid properties and G_{hydr} . However, as stated above, it was not possible to obtain a single LJ parameter set for chlorine suitable for reproduction of both pure liquid properties and G_{hydr} . For example, the pure liquid properties could be reproduced well by the LJ1 parameter set while the G_{hydr} values were generally too favorable among all the chlorinated benzenes (Table 4). Previous studies reported similar problems in reproducing closely both pure liquid properties and hydration free energies, leading to the use of atom pair-specific LJ parameters (i.e. NBFIX in CHARMM nomenclature).¹⁹ Given the importance of the force field reproducing the

G_{hydr} data, the LJ2 parameter set for chlorine was developed initially using the GROMACS free energy protocol. As can be seen in Table 5, the predicted G_{hydr} for chlorinated compounds are generally in good agreement with the experimental data. Use of the LJ2 parameter set for calculation of the pure liquid properties resulted in only qualitative agreement with pure liquid properties; V_m values were in satisfactory agreement but some differences greater than 10% occurred with the H_{vap} values. In the case of brominated and iodinated compounds, application of the LJ parameters based on the pure liquid properties yield satisfactory agreement with experimental data (Tables 6 and 7), with the deviation for G_{hydr} being 1 kcal·mol⁻¹ or less. The agreement is especially good for the two brominated benzene compounds for which experimental data are available, which showed deviations within 0.1 kcal·mol⁻¹.

To test the robustness of the calculated G_{hydr} values, we redid the calculations using a separate free energy perturbation protocol implemented in CHARMM (Table 8). The overall agreement between the two protocols is satisfactory, though the CHARMM G_{hydr} values are systematically more favorable. A potential contribution to this difference may be the effect of different switch smoothing functions applied on the LJ interaction. Potential switching was used in the CHARMM free energy calculations (VSWITCH in CHARMM nomenclature) versus the use of force switching in GROMACS (VFSWITCH). Calculation of the benzene or chlorinated benzene-water interaction energy differences using these two smoothing schemes over the range of 10 to 12 Å showed the use of VSWITCH to systematically yield more favorable LJ interaction energy difference (Table S3, supporting information). These systematic differences will contribute, in part, to the difference between the CHARMM and GROMACS G_{hydr} values, when considering that additional halogenation in the di- and trihalogenated species will increase this effect. Indeed, the correlation between the difference LJ interaction energy contribution differences and the difference in CHARMM and GROMACS G_{hydr} values is high ($R^2 = 0.69$), although quantitative analysis of the differences indicate that this contribution may not account for the total differences (Table S3 and Figure S1, supporting information). The remaining sources of the differences are difficult to ascertain given the known difficulty in converging G_{hydr} calculations.⁶⁶ Potential contributions may be the additional perturbation windows in the CHARMM protocol and/or the numerical LRC correction for the LJ term in CHARMM versus the analytical LJ LRC included in the GROMACS protocol. While elucidating the origin of this difference is important, it is beyond the scope of this work as we were able to reproduce the single-point energies and forces of the monomers and the solvated systems using the two programs using identical smoothing functions (not shown).

3.5. Crystal simulations of small molecules

Crystal simulations were performed using two sets of crystal structures obtained from the CSD to select the final LJ parameters for chlorine. One set included chlorinated benzenes interacting with a carbonyl, excluding amides, and the second set included NMA, as shown in Tables S1 and S2 in the Supporting Information, respectively; water was not present in any of the studied crystals. A summary of the results for the two sets of crystals is shown in Table 9. Both sets of parameters yielded significant improvements over CGenFF,¹⁴ especially with respect to the Cl...O interaction distances. For the crystals not containing NMA, the LJ2 set was significantly improved over LJ1, indicating the general utility of the parameters based on reproduction of the G_{hydr} data. In the crystals containing NMA, the average volume of the crystal was worse with LJ2 vs. LJ1, but the absolute unsigned error (AUE) and RMSD were similar and there was a significant improvement in the Cl...O interaction distances. The introduction of the NBFIX to the amide carbonyl O led to an additional improvement in the interaction distances. These results further indicate the improvements in the treatment of chlorine over CGenFF, as well as the ability of the new model to perform well in heterogeneous environments.

We note that the CGenFF parameters applied to the non-halogenated compounds in the crystals may affect the simulations. For example, in the CSD ID FUJWAJ crystal,⁴⁹ the amide carbonyl is part of a carbamate group with a large charge penalty, indicating that the charges on this compound may not be optimal. Such considerations complicate detailed interpretation of the results from the crystal simulations.

3.6. Protein-ligand simulations

The final test of the halogen parameters was a series of MD simulations of ligand-protein complexes to determine the ability of the force field to reproduce halogen-protein interactions. The results for five protein-chlorinated ligand simulation systems are summarized in Table 10, and results for the individual systems are shown in Figures S2-S6 and Tables S4-S10 of the Supporting Information. Simulations involving chlorinated ligands with the LJ2 set showed significant improvements over the original CGenFF parameters. The overall binding poses of the ligands in the five structures are improved according to ligand RMSD, as are the geometric parameters related to halogen bonding, specifically chlorine-oxygen interaction distances, $d(\text{Cl}\dots\text{O})$ and angles, $\phi(\text{Cl}\dots\text{O}=\text{C})$.

Additional simulations of brominated and iodinated ligands bound to human thrombin were carried out (Tables S5-S6 in the Supporting Information). For the brominated inhibitor, the ligand RMSD is similar between CGenFF and the new parameters, as are the interaction distance and angle. For the iodinated inhibitor, the ligand binding pose was better maintained with the new parameters, with an RMSD of 1.34 Å versus 1.99 Å with CGenFF. In addition, the interaction distance (4.52 Å in the crystal structure) improved from 6.48 Å with CGenFF to 4.58 Å with the new parameters.

As with the small molecule crystals above, we note possible contributions from additional elements of the ligand topologies, unrelated to the halogens, on the obtained results. For example, desvancosaminy vancomycin is a large, flexible ligand built on a heterocyclic

backbone that binds to TDP-vancosaminyltransferase via a large binding pocket with extensive interactions with the protein (Figure S6). Suboptimal dihedral parameters may contribute to altered conformational sampling; though the error in any individual term is small, the combined effects will influence the simulation outcome and lead to systematic deviations in both CGenFF and LJ2 simulations. Given this fact, the improvements in the results (Table S10) can be attributed to the LJ2 halogen parameters developed in the present work. In the *P. falciparum* DHFR-pyrimethamine complex, the interaction between the pyrimethamine ring nitrogen (N1) is believed to be stabilized by hydrogen bonding to the side chain of Asp54, thus requiring the ring nitrogen to be protonated.⁵² However, the CGenFF parameters for this particular protonation state of the ligand had unacceptably high penalties for some of the assigned dihedrals, indicating the need for more general refinement within CGenFF, an effort that is beyond the scope of the present study. As an alternative, we treated Asp54 as protonated and the pyrimidine ring as neutral. Though not optimal, this approach led to stable ligand-protein interactions that allow a fair comparison between CGenFF and the new halogen parameters (Figure S4 and Table S8).

4. Conclusions

The work presented here details parametrization of chlorinated, brominated and iodinated benzenes by introducing a virtual particle to model the sigma hole on aromatic halogens that is responsible for halogen bonding. The electrostatic parameters were optimized to reproduce QM and experimental dipole and quadrupole moments along with QM water/acetone interactions. The initial LJ parameters were also adjusted to reproduce QM water/acetone interactions and subsequently optimized by targeting experimental heats of vaporization, molecular volumes, and relative hydration free energies. Bonded parameters were not optimized, as these parameters were taken directly from CGenFF, for which halogenated benzene parameters were already available. The LJ parameters for chlorine were further optimized by targeting QM NMA-chlorobenzene interactions and were validated by small molecule crystal simulations. The final parameters for chlorinated species include the parameters based on reproduction of free energies of hydration along with a specific NBFIX parameter for the carbonyl oxygen in amides. As this set of parameters leads to degradation of the pure liquid results for polyhalogenated benzenes (Table 5), for studies of such systems the use of the LJ1 parameter set is recommended. These parameters are listed in Table S11 of the supporting information. For brominated and iodinated species, the parameters based on the pure liquid properties were shown to be suitable for all the studied systems. Compared to CGenFF, the new parameters for halogenated molecules are better able to reproduce the targeted experimental and QM properties, including small molecule crystal structures as well as binding modes of ligand-protein complexes. Accordingly, the presented parameters are anticipated to be of utility for representing ligands involving halogenated aromatic rings, including those in drug-like molecules, in conjunction with CGenFF and the remainder of CHARMM36 additive force field.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health [grant numbers GM070855, GM072558, GM037554 and F32GM109632]. The University of Maryland Computer-Aided Drug Design Center and XSEDE are acknowledged for their generous allocations of computer time.

References

1. Buchini S, Buschiazio A, Withers SG. *Angew Chem Int Ed Engl*. 2008; 47:2700–3. [PubMed: 18300214]
2. Bundhun A, Ramasami P, Murray JS, Politzer P. *J Mol Model*. 2012; 19:2739–2746. [PubMed: 22968690]
3. Leite ACL, Moreira DR, de M, Cardoso MV, de O, Hernandez MZ, Alves Pereira VR, Silva RO, Kiperstok AC, Lima M, da S, Soares MBP. *ChemMedChem*. 2007; 2:1339–45. [PubMed: 17628867]
4. Gerebtzoff G, Li-Blatter X, Fischer H, Frentzel A, Seelig A. *Chembiochem Eur J Chem Biol*. 2004; 5:676–84.
5. Wilcken R, Liu X, Zimmermann MO, Rutherford TJ, Fersht AR, Joerger AC, Boeckler FM. *J Am Chem Soc*. 2012; 134:6810–6818. [PubMed: 22439615]
6. Politzer P, Murray JS, Clark T. *Phys Chem Chem Phys*. 2010; 12:7748–7757. [PubMed: 20571692]
7. Murray JS, Lane P, Clark T, Politzer P. *J Mol Model*. 2007; 13:1033–8. [PubMed: 17647029]
8. Riley KE, Hobza P. *J Chem Theory Comput*. 2008; 4:232–242. [PubMed: 26620655]
9. Riley KE, Murray JS, Fanfrlík J, ezá J, Solá RJ, Concha MC, Ramos FM, Politzer P. *J Mol Model*. 2012; 19:4651–4659. [PubMed: 22643966]
10. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. *J Am Chem Soc*. 1995; 117:5179–5197.
11. Jorgensen WL, Tirado-Rives J. *J Am Chem Soc*. 1988; 110:1657–1666. [PubMed: 27557051]
12. Jorgensen WL, Maxwell DS, Tirado-Rives J. *J Am Chem Soc*. 1996; 118:11225–11236.
13. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF. *J Phys Chem A*. 1999; 103:3596–3607.
14. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, MacKerell AD Jr. *J Comput Chem*. 2010; 31:671–690. [PubMed: 19575467]
15. Ibrahim MAA. *J Comput Chem*. 2011; 32:2564–2574. [PubMed: 21598284]
16. Ibrahim MAA. *J Mol Model*. 2012; 18:4625–38. [PubMed: 22643975]
17. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA. *J Chem Theory Comput*. 2016; 12:281–296. [PubMed: 26584231]
18. Jorgensen WL, Schyman P. *J Chem Theory Comput*. 2012; 8:3895–3901. [PubMed: 23329896]
19. Bereau T, Kramer C, Markus M. *J Chem Theory Comput*. 2013; 9:5450–5459. [PubMed: 26592281]
20. Frisch MJ, Trucks GW, Schlegel HB, S GE, Robb MA, Cheeseman JR, Montgomery JA Jr, V T, Kudin KN, Burant JC, Millam JM, Iyengar SS, T J, Barone V, Mennucci B, Cossi M, Scalmani G, R N, Petersson GA, Nakatsuji H, Hada M, Ehara M, T K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, K O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, C JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, S RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, O JW, Ayala PY, Morokuma K, Voth GA, Salvador P, D JJ, Zakrzewski VG, Dapprich S, Daniels AD, S MC, Farkas O, Malick DK, Rabuck AD, R K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, C S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, P P, Komaromi I, Martin RL, Fox DJ, Keith T, AL MA, Peng CY, Nanayakkara A, Challacombe M, G PMW, Johnson B, Chen W, Wong MW, Gonzalez C, JAP. *Gaussian03revD.01*. 2004
21. Woon DE, Dunning TH Jr. *J Chem Phys*. 1993; 98:1358–1371.
22. Peterson KA, Figgien D, Goll E, Stoll H, Dolg M. *J Chem Phys*. 2003; 119:11113–11123.

23. Peterson KA, Shepler BC, Figgen D, Stoll H. *J Phys Chem A*. 2006; 110:13877–13883. [PubMed: 17181347]
24. Krylov AI, Gill PMW. *Wiley Interdiscip Rev Comput Mol Sci*. 2013; 3:317–326.
25. Martin JML, Sundermann A. *J Chem Phys*. 2001; 114:3408–3420.
26. Valiev M, Bylaska EJ, Govind N, Kowalski K, Straatsma TP, Van Dam HJJ, Wang D, Nieplocha J, Apra E, Windus TL, de Jong WA. *Comput Phys Commun*. 2010; 181:1477–1489.
27. Boys S, Bernardi F. *Mol Phys*. 1970; 19:553–566.
28. Brooks BR, Brooks CL III, MacKerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. *J Comput Chem*. 2009; 30:1545–614. [PubMed: 19444816]
29. Brooks BR, Bruccoleri RE, Olafson DJ, States DJ, Swaminathan S, Karplus M. *J Comput Chem*. 1983; 4:187–217.
30. MacKerell, AD., Jr; Brooks, CL., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M.; Schleyer, PvR; Allinger, NL.; Clark, T.; Gasteiger, J.; Kollman, PA.; Schaefer, HF., III; Schreiner, PR., editors. *Encyclopedia of Computational Chemistry*. Vol. 1. John Wiley & Sons; Chichester: 1998. p. 271–277.
31. Hariharan P, Pople J. *Theor Chim Acta*. 1973; 28:213.
32. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. *SoftwareX*. 2015; 1–2:19–25.
33. Kollman PA. *Chem Rev*. 1993; 93:2395–2417.
34. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J Chem Phys*. 1983; 79:926–935.
35. Neria E, Fischer S, Karplus M. *J Chem Phys*. 1996; 105:1902–1921.
36. Durell SR, Brooks BR, Ben-Naim A. *J Phys Chem*. 1994; 98:2198–2202.
37. Darden T, York D, Pedersen L. *J Chem Phys*. 1993; 98:10089–10092.
38. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. *J Chem Phys*. 1995; 103:8577–8593.
39. Zacharias M, Straatsma TP, McCammon JA. *J Chem Phys*. 1994; 100:9025–9031.
40. Steinbrecher T, Joung IS, Case DA. *J Comput Chem*. 2011; 32:3253–3263. [PubMed: 21953558]
41. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. *J Comput Chem*. 1992; 13:1011–1021.
42. Lagüe P, Pastor RW, Brooks BR. *J Phys Chem B*. 2004; 108:363–368.
43. Bennett CH. *J Comput Phys*. 1976; 22:245–268.
44. Parrinello M, Rahman A. *J Appl Phys*. 1981; 52:7182–7190.
45. Nosé S, Klein ML. *Mol Phys*. 1983; 50:1055–1076.
46. Goga N, Rzepiela AJ, de Vries AH, Marrink SJ, Berendsen HJC. *J Chem Theory Comput*. 2012; 8:3637–3649. [PubMed: 26593009]
47. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. *J Comput Chem*. 1997; 18:1463–1472.
48. Hess B. *J Chem Theory Comput*. 2008; 4:116–122. [PubMed: 26619985]
49. Allen FH. *Acta Crystallogr B*. 2002; 58:380–388. [PubMed: 12037359]
50. Baum B, Mohamed M, Zayed M, Gerlach C, Heine A, Hangauer D, Klebe G. *J Mol Biol*. 2009; 390:56–69. [PubMed: 19409395]
51. Schiebel J, Chang A, Lu H, Baxter MV, Tonge PJ, Kisker C. *Structure*. 2012; 20:802–813. [PubMed: 22579249]
52. Yuvaniyama J, Chitnumsub P, Kamchonwongpaisan S, Vanichtanankul J, Sirawaraporn W, Taylor P, Walkinshaw MD, Yuthavong Y. *Nat Struct Mol Biol*. 2003; 10:357–365.
53. Scapin G, Patel SB, Lisnock J, Becker JW, LoGrasso PV. *Chem Biol*. 2003; 10:705–712. [PubMed: 12954329]
54. Mulichak AM, Lu W, Losey HC, Walsh CT, Garavito RM. *Biochemistry (Mosc)*. 2004; 43:5170–5180.

55. Wang L, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, Romero DL, Masse C, Knight JL, Steinbrecher T, Beuming T, Damm W, Harder E, Sherman W, Brewer M, Wester R, Murcko M, Frye L, Farid R, Lin T, Mobley DL, Jorgensen WL, Berne BJ, Friesner RA, Abel R. *J Am Chem Soc.* 2015; 137:2695–2703. [PubMed: 25625324]
56. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD Jr. *J Chem Theory Comput.* 2012; 8:3257–3273. [PubMed: 23341755]
57. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M. *J Phys Chem B.* 1998; 102:3586–3616. [PubMed: 24889800]
58. Hernandez M, Cavalcanti SM, Moreira DR, de Azevedo Junior W, Leite AC. *Curr Drug Targets.* 2010; 11:303–314. [PubMed: 20210755]
59. Lide, DR. *CRC handbook of chemistry and physics* [electronic resource]. Boca Raton, Fla: CRC Press; Taylor & Francis; 2008. c2009
60. Novák M, Foroutan-Nejad C, Marek R. *Phys Chem Chem Phys.* 2015; 17:6440–6450. [PubMed: 25656525]
61. Blanco F, Kelly B, Alkorta I, Rozas I, Elguero J. *Chem Phys Lett.* 2011; 511:129–134.
62. Torrice MM, Bower KS, Lester H, Dougherty D. *Proc Natl Acad Sci U S A.* 2009; 106:11919–11924. [PubMed: 19581583]
63. Mahadevi AS, Sastry GN. *Chem Rev.* 2013; 113:2100–2138. [PubMed: 23145968]
64. Dougherty DA. *Acc Chem Res.* 2013; 46:885–893. [PubMed: 23214924]
65. Gallivan JP, Dougherty DA. *Proc Natl Acad Sci U S A.* 1999; 96:9459–9464. [PubMed: 10449714]
66. Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. *Curr Opin Struct Biol.* 2011; 21:150–160. [PubMed: 21349700]
67. [23rd March 2016] Pubchem 1,3,5-TRICHLOROBENZENE | C₆H₃Cl₃ - PubChem. Available at: https://pubchem.ncbi.nlm.nih.gov/compound/1_3_5-trichlorobenzene#section=Top
68. [23rd March 2016] Vapour pressures of crystalline 1,2,4,5-tetrachlorobenzene, and crystalline and liquid 1,3,5-trichlorobenzene and 1,2,4,5-tetramethylbenzene. Available at: <http://www.sciencedirect.com.proxy-hs.researchport.umd.edu/science/article/pii/S002196140090819X?np=y>
69. Chickos JS Jr, W EA. *J Phys Chem Ref Data.* 2003; 32:519–878.
70. [26th March 2016] 1,2,3-TRIBROMOBENZENE (CAS 608-21-9). Available at: http://www.chemicalbook.com/ProductChemicalPropertiesCB7223286_EN.htm
71. [25th March 2016] 1,2-Diiodobenzene (CAS 615-42-9) MSDS Melting Point Boiling Point Density Storage Transport. Available at: http://www.chemicalbook.com/ProductMSDSDetailCB9389618_EN.htm
72. [26th March 2016] 1,3-Diiodobenzene (CAS 626-00-6) - Chemical & Physical Properties by Cheméo. Available at: <https://www.chemeo.com/cid/64-958-8/1%2C3-Diiodobenzene#ref-joback>
73. [26th March 2016] 1-4-Diiodobenzene (CAS 624-38-4). Available at: <http://www.lookchem.com/1-4-Diiodobenzene/>
74. [26th March 2016] 1,2,3-triiodobenzene (CAS 608-29-7). Available at: <http://anchor.guidechem.com/dictionary/en/608-29-7.html>
75. Abraham MH, Andonian-Haftvan J, Whiting GS, Leo A, Taft RS. *J Chem Soc Perkin Trans 2.* 1994:1777–1791.
76. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA. *J Chem Theory Comput.* 2009; 5:350–358. [PubMed: 20150953]

Figure 1.

Final set of charges for the chlorinated benzenes. Charges on top correspond to chlorines or hydrogens, bottom charges to the carbon atoms to which they are attached. The atoms for which charges are not shown were kept as in benzene, +0.115 for hydrogens and -0.115 for carbons.

Figure 2.

Final set of charges for the brominated benzenes. Charges on top correspond to bromines or hydrogens, bottom charges to the carbon atoms to which they are attached. The atoms for which charges are not shown were kept as in benzene, +0.115 for hydrogens and -0.115 for carbons.

Figure 3.

Final set of charges for the iodinated benzenes. Charges on top correspond to iodines or hydrogens, bottom charges to the carbon atoms to which they are attached. The atoms for which charges are not shown were kept as in benzene, +0.115 for hydrogens and -0.115 for carbons.

Figure 4.

Relative orientations between the halogenated benzenes and water. The halogen is labeled as “X” and the carbon next to the halogen is labeled as “C”. (A) Radial scan, in which the water molecule is positioned directly along the C-X bond axis. (B) Angular scan, in which the water molecule is positioned at a distance “d” and changing the C-X...O angle θ .

Figure 5. Thermodynamic cycle used to determine the hydration free energy relative to benzene, G_{hydr}

(A) Using LJ1 ($\epsilon = -0.31 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.94 \text{ \AA}$) **(B) Using LJ2 ($\epsilon = -0.23 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.86 \text{ \AA}$)** **(C) Chlorobenzene-NMA using LJ2**

Figure 6.

Radial interaction energy profiles between (A) chlorobenzene-water, (TIP3P)/acetone (ACO) using LJ1, (B) chlorobenzene-TIP3P/ACO using LJ2, (C) chlorobenzene-N-methylacetamide (NMA) using LJ2 without/with NBFIX term with the amide carbonyl oxygen, (D) bromobenzene- TIP3P/ACO, (E) iodobenzene- TIP3P/ACO. E_{tot} , E_{elec} , E_{vdW} corresponds to MM total, electrostatic, and van der Waals interaction energies. QM-TIP3P/ACO/NMA corresponds to QM calculated interactions with water, ACO, and NMA.

(A) Using LJ1 ($\epsilon = -0.31 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.94 \text{ \AA}$) (B) Using LJ2 ($\epsilon = -0.23 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.86 \text{ \AA}$)

(C) Chlorobenzene-NMA using LJ2

Figure 7.

Angular interaction energy profiles between chlorobenzene and water (TIP3P) or acetone (ACO) (A) using LJ1, (B) using LJ2, (C) chlorobenzene...N-methyl-acetamide (NMA) using LJ2 without or with the NBFIX term with the amide carbonyl, (D) bromobenzene...TIP3P or ACO, (E) iodobenzene...TIP3P or ACO at selected distances ($d(\text{Cl}\dots\text{O})$) where the QM interaction energies were at the minima in the radial interaction energy profiles. E_{tot} , E_{elec} , E_{vdW} corresponds to MM total, electrostatic, and van der Waals interaction energies, respectively. QM-TIP3P/ACO/NMA corresponds to QM calculated interactions with water, ACO, or NMA, respectively.

Table 1

Final (LJ2) parameters and previous values in CGenFF for monohalogenated benzenes.

Compound	$q_{C_{ipso}}$	q_X	q_{lp}	$r(lp-X)$ (Å)	e (kcal mol ⁻¹)	$R_{min}/2$ (Å)
Chlorobenzene ²	0.06	-0.21	0.05	1.64	-0.230	1.86
CGenFF	0.03	-0.13			-0.320	1.93
Bromobenzene	0.05	-0.18	0.05	1.89	-0.320	1.98
CGenFF	-0.08	-0.10			-0.420	2.07
Iodobenzene	-0.04	-0.17	0.05	2.20	-0.520	2.24
CGenFF	-0.08	-0.08			-0.550	2.19

¹ Lennard-Jones parameters are not placed on the sigma hole particle virtual sites

² The final parameter set includes an NBFIX term for chlorine and amide carbonyl oxygen ($e = -0.20$ kcal mol⁻¹, $R_{min} = 3.40$ Å).

Table 2

Dipole moments obtained from the model compounds compared to QM and experimental data. Dipoles were computed using the MP2/6-31G(d) and MP2/aug-cc-pVDZ model chemistry for chloro- and bromo-substituted benzenes, and MP2/6-311G(d) for iodo-substituted benzenes. All dipoles are given in units of D.

Halogen	Compound	MM Dipole	QM Dipole ^a	QM Dipole ^b	Exp. Dipole
Cl	CHLB	1.68	1.89	1.75	1.69 ± 0.03 ⁵⁹
	12DCB	2.50	2.74	2.50	2.50 ± 0.05 ⁵⁹
	13DCB	1.86	1.78	1.65	1.72 ± 0.09 ⁵⁹
	14DCB	0.00	0.00	0.00	
	123TCB	2.73	2.78	2.51	
	124TCB	1.78	1.40	1.26	
Br	135TCB	0.00	0.00	0.00	
	BROB	1.74	1.54	1.75	1.70 ± 0.03 ⁵⁹
	12DBB	2.87	2.64	2.35	
	13DBB	1.93	1.76	1.59	
	14DBB	0.00	0.00	0.00	
	123TBB	2.63	2.55	2.24	
I	124TBB	1.21	1.28	1.12	
	135TBB	0.00	0.00	0.00	
	IODB	1.69	1.63	1.63	1.70 ± 0.09 ⁵⁹
	12DIB	2.79	2.37	2.37	
	13DIB	1.95	1.65	1.65	
	14DIB	0.00	0.00	0.00	
	123TIB	2.15	2.11	2.11	
	124TIB	0.99	1.07	1.07	
	135TIB	0.00	0.00	0.00	

^aDipoles were computed using the MP2/6-31G(d) model chemistry.

^bDipoles were computed using the MP2/aug-cc-pVDZ model chemistry.

Dipoles were computed using the MP2/6-311G(d) model chemistry.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Interaction energies and distances obtained from our models for halogenated compounds interacting with water, acetone, or N-methylacetamide along the C-X bond axis of the halogenated compounds.

Compound	E_{int} (kcal mol ⁻¹)	Min. Distance O...X (Å)		
Water				
	CHARMM	QM	CHARMM	QM
Chlorobenzene ¹	-0.31	-0.46	3.4	3.2
Chlorobenzene ²	-0.39	-0.46	3.2	3.2
Bromobenzene	-1.20	-1.23	3.2	3.1
Iodobenzene	-1.30	-1.99	3.1	3.2
Acetone				
Chlorobenzene ¹	-0.30	-0.68	3.2	3.1
Chlorobenzene ²	-0.34	-0.68	3.1	3.1
Bromobenzene	-1.22	-1.51	3.0	3.1
Iodobenzene	-2.19	-2.34	3.2	3.1
N-methylacetamide				
Chlorobenzene ³	-0.44	-0.88	3.1	3.1
Chlorobenzene ⁴	-0.69	-0.88	3.0	3.1

¹ LJ1 parameters ($\epsilon = -0.31$ kcal mol⁻¹, $R_{min}/2 = 1.94$ Å)

² LJ2 parameters ($\epsilon = -0.23$ kcal mol⁻¹, $R_{min}/2 = 1.86$ Å)

³ LJ2 without NBFIX term for chlorine and amide carbonyl

⁴ LJ2 and NBFIX term for chlorine and amide carbonyl

Table 4

Bulk phase properties for the chlorinated benzenes using the LJ1 parameters ($\epsilon = -0.31 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.94 \text{ \AA}$). G_{hydr} values were calculated using the GROMACS protocol.

Mol.	V_m		H_{vap}		G_{hydr} ($\text{kcal}\cdot\text{mol}^{-1}$)	
	$\text{\AA}^3$	% Diff.	Temp (K)	$\text{kcal}\cdot\text{mol}^{-1}$	% Diff.	Temp (K)
CHLB	Exp. 169.00 ⁵⁹			9.79 ⁵⁹		
	Pred. 168.85 $\pm$ 0.11	-0.09	298	9.77 $\pm$ 0.01	-0.24	298
12DCB	Exp. 186.90 ⁵⁹			9.48 ⁵⁹		
	Pred. 188.09 $\pm$ 0.05	0.64	293	9.62 $\pm$ 0.03	1.51	453
13DCB	Exp. 189.40 ⁵⁹			9.24 ⁵⁹		
	Pred. 186.74 $\pm$ 0.14	-1.40	293	8.99 $\pm$ 0.04	-2.67	446
14DCB	Exp. 195.60 ⁵⁹			9.28 ⁵⁹		
	Pred. 192.13 $\pm$ 0.30	-1.77	328	9.10 $\pm$ 0.02	-1.90	447
123TCB	Exp. 207.32 ⁵⁹			11.79 ⁵⁹		
	Pred. 208.72 $\pm$ 0.07	0.67	298	11.24 $\pm$ 0.04	-4.64	423
124TCB	Exp. 206.52 ⁵⁹			11.84 ⁵⁹		
	Pred. 207.59 $\pm$ 0.03	0.52	298	11.24 $\pm$ 0.01	-5.12	406
135TCB	Exp. 206.90 ⁶⁷			12.01 ⁶⁸		
	Pred. 208.18 $\pm$ 0.02	0.62	293	12.56 $\pm$ 0.01	4.54	357

Bulk phase properties for the chlorinated benzenes using the LJ2 parameters ($\epsilon = -0.23 \text{ kcal mol}^{-1}$, $R_{\text{min}}/2 = 1.86 \text{ \AA}$). G_{hydr} values were calculated using the GROMACS protocol.

Table 5

Mol.	V _m		H _{vap}		G _{hydr} (kcal·mol ⁻¹)	
	Å ³	% Diff.	Temp (K)	kcal·mol ⁻¹	% Diff.	Temp (K)
CHLBr	Exp. 169.00 ⁵⁹			9.79 ⁵⁹		
	Pred. 168.01 ± 0.23	-0.58	298	9.23 ± 0.06	-5.67	298
12DCBr	Exp. 186.90 ⁵⁹			9.48 ⁵⁹		
	Pred. 185.87 ± 0.21	-0.55	293	8.54 ± 0.05	-9.89	453
13DCBr	Exp. 189.40 ⁵⁹			9.24 ⁵⁹		
	Pred. 184.63 ± 0.31	-2.52	293	7.67 ± 0.03	-17.01	446
14DCBr	Exp. 195.60 ⁵⁹			9.28 ⁵⁹		
	Pred. 191.22 ± 0.31	-2.24	328	7.79 ± 0.09	-16.02	447
123TCBr	Exp. 207.32 ⁵⁹			11.79 ⁵⁹		
	Pred. 205.46 ± 0.33	-0.90	298	9.64 ± 0.03	-18.20	423
124TCBr	Exp. 206.52 ⁵⁹			11.84 ⁵⁹		
	Pred. 203.76 ± 0.18	-1.33	298	10.14 ± 0.02	-14.37	406
135TCBr	Exp. 206.90 ⁶⁷			12.01 ⁶⁸		
	Pred. 204.37 ± 0.18	-1.22	293	10.06 ± 0.03	-16.25	357

Table 6

Bulk phase properties for brominated benzenes. G_{hydr} values were calculated using the GROMACS protocol.

Mol.	V _m		H _{up}		G _{hydr} (kcal·mol ⁻¹)	
	Å ³	% Diff.	Temp (K)	kcal·mol ⁻¹	% Diff.	Temp (K)
BROB	Exp. 174.40 ⁵⁹			10.66 ⁵⁹		
	Pred. 172.42 ± 0.24	-1.14	293	10.26 ± 0.01	-3.75	298
						-0.69 ± 0.06
12DBB	Exp. 197.42 ⁵⁹			11.99 ⁶⁹		
	Pred. 196.86 ± 0.23	-0.28	293	10.18 ± 0.04	-15.06	403
						N/A
13DBB	Exp. 200.65 ⁵⁹			11.56 ⁶⁹		
	Pred. 195.61 ± 0.10	-2.51	293	9.36 ± 0.04	-19.04	432
						N/A
14DBB	Exp. 173.26 ⁵⁹			11.94 ⁶⁹		
	Pred. 192.01 ± 0.31	10.82	290	10.81 ± 0.02	-9.46	388
						-1.34 ± 0.48
123TBB	Exp. 196.67 ⁷⁰			N/A		
	Pred. 216.79 ± 0.20	10.23	293	---	---	---
						N/A
124TBB	Exp. N/A			N/A		
	Pred. ---	---	---	---	---	---
						N/A
135TBB	Exp. 198.24			N/A		
	Pred. 224.11 ± 0.20	13.05	295	---	---	---
						N/A

Table 7

Bulk phase properties for iodinated benzenes. G_{hydr} values were calculated using the GROMACS protocol.

Mol.	V_m		H_{vap}		G_{hydr} (kcal·mol ⁻¹)
	$\text{\AA}^3$	% Diff.	Temp (K)	% Diff.	
IODB	Exp. 185.04 ⁵⁹				-0.88
	Pred. 186.61 ± 0.22	0.85	293	-4.94	-1.16 ± 0.11
12DIB	Exp. 215.68 ⁵⁹				N/A
	Pred. 225.27 ± 0.24	4.44	293	-2.82	298 ^a
13DIB	Exp. 221.79 ⁵⁹				
	Pred. 226.64 ± 0.17	2.18	298	21.35	298 ^a
14DIB	Exp. 221.88 ⁷³				
	Pred. 225.19 ± 0.16	1.49	298 [/]	-0.68	417
123TIB	Exp. 258.50 ⁷⁴				N/A
	Pred. 265.39 ± 0.18	2.66	298 [/]	---	---
124TIB	Exp. N/A				N/A
	Pred. ---	---	---	---	---
135TIB	Exp. N/A				N/A
	Pred. ---	---	---	---	---

[/] Assumed to be 298K

Relative hydration free energies of halogenated molecules with respect to benzene (G_{hydr}) calculated using the GROMACS and CHARMM protocols.
All values are in kcal mol⁻¹.

Table 8

X	Exp, $G_X^{a,75,76}$	Exp, $G_{benz}^{a,75,76}$	Exp, G_{hydr}^b	G_{hydr} (GROMACS)	G_{hydr} (CHARMM)
CHLB	-1.12	-0.86	-0.26	-0.39 ± 0.05	-0.48 ± 0.07
12DCB	-1.36	-0.86	-0.50	-0.37 ± 0.32	-0.88 ± 0.05
13DCB	-0.98	-0.86	-0.12	-0.55 ± 0.22	-0.89 ± 0.08
14DCB	-1.01	-0.86	-0.15	0.09 ± 0.33	-0.20 ± 0.14
123TCB	-1.24	-0.86	-0.38	-0.17 ± 0.67	-1.28 ± 0.07
124TCB	-1.12	-0.86	-0.26	0.30 ± 0.64	-1.38 ± 0.10
135TCB	-0.78	-0.86	0.08	0.27 ± 0.87	-0.65 ± 0.09
BROB	-1.46	-0.86	-0.60	-0.69 ± 0.06	-1.11 ± 0.18
14DBB	-2.30	-0.86	-1.44	-1.34 ± 0.48	-2.85 ± 0.09
IODB	-1.74	-0.86	-0.88	-1.16 ± 0.11	-1.90 ± 0.05

^aThe experimental hydration free energy error is assumed to be 0.6 kcal.mol⁻¹, 75,76

^bThe relative hydration free energies error is assumed to be 0.85 kcal.mol⁻¹, which is the square root of sum of errors of 0.6 kcal.mol⁻¹, squared.

Table 9

Summary of the crystal simulation results for crystals containing chlorobenzene that involved chlorine-carbonyl oxygen interactions that do not include amides. Average differences, absolute unsigned error (AUE), and root-mean-square differences (RMSD) are reported based on ten crystals in each set.

Unit Cell Volume (Å ³)				Cl...O Distance (Å)			
	Avg. Diff.	AUE	RMSD	Avg. Diff.	AUE	RMSD	
Crystals not containing NMA							
CGenFF	29.64	49.74	85.56	0.39	0.39	0.42	
LJ1	23.81	40.98	71.06	0.30	0.30	0.35	
LJ2	7.03	33.83	63.11	0.20	0.21	0.29	
NMA-containing crystals							
CGenFF	-12.56	94.80	169.31	0.36	0.36	0.58	
LJ1	-12.49	100.68	176.86	0.31	0.32	0.57	
LJ2 ¹	-35.07	96.36	166.41	0.10	0.14	0.20	
LJ2 ²	-42.55	94.70	169.40	0.04	0.14	0.19	

¹Without the NBFIX to the amide carbonyl oxygen

²With the NBFIX

Table 10

Summary of structural properties in the protein-ligand simulations for chlorinated inhibitors. Ligand RMSD is computed for all ligand atoms after a least-squares fit of the protein backbone, using the crystal structure as a reference. Results of simulations for the five individual complexes are reported in Tables S4, S7-S10 of the Supporting Information.

	Ligand RMSD, Å		d(Cl...O), Å		ϕ (Cl...O=C), degrees		
	Average	Avg. Diff.	AUE	RMSD	Avg. Diff.	AUE	RMSD
CGenFF	1.80	1.04	1.04	1.71	-19.68	28.72	42.78
LJ21	1.33	0.32	0.40	0.47	-13.68	14.98	18.45

¹With NBFIX to the amide carbonyl oxygen.