

Preserving the Integrity of Empirical Force Fields

Asuka A. Orr, Suliman Sharif, Junmei Wang, and Alexander D. MacKerell, Jr.*



Cite This: *J. Chem. Inf. Model.* 2022, 62, 3825–3831



Read Online

ACCESS |



Metrics & More

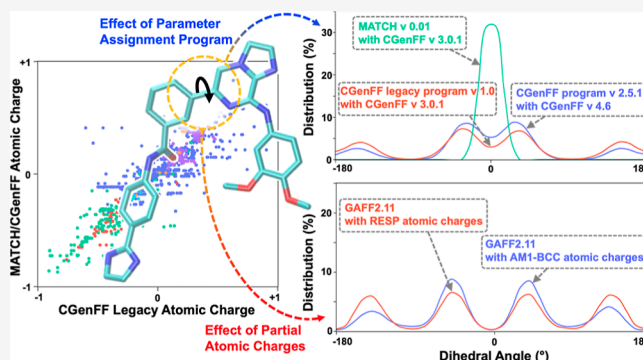


Article Recommendations



Supporting Information

ABSTRACT: Generalized force fields (FFs) act as extensions to biomolecular FFs to provide a wide coverage of organic molecules. However, their precise application to an arbitrary molecule presents a separate challenge. We show that MATCH assigns different atom types and bonded and nonbonded parameters than CGenFF, and the AM1-BCC charge model, commonly used with GAFF/GAFF2, does not exactly reproduce the performance of the RESP charge model. The results indicate the need for caution when employing FFs to ensure their integrity with respect to their implementation and validation.



Empirical force fields (FFs) represent the foundation of molecular modeling and simulation. Currently, a range of FFs are available, dominated by additive, pairwise all-atom models optimized primarily for macromolecular simulations, including CHARMM^{1–10} and AMBER,^{11–16} among others,^{17,18} that are implemented in different simulation programs.^{19–22} In the context of drug design as well as studies of a wide range of diverse systems such as nanomaterials and ionic liquids, additive FFs have been extended to small, organic compounds. These include the CHARMM General FF (CGenFF)^{23–26} in the context of CHARMM^{1–10} and General Amber FF (GAFF)^{27–31} and GAFF2³² in the context of AMBER.^{11–16}

While the small-molecule empirical FFs have seen wide use, users must be aware of the fidelity of their implementations from different sources and simulation programs. For example, a simple challenge is the use of different constants for Coulomb's law, leading to differences in electrostatic energies and forces. Other examples include using the correct units, joules (International System of Units) versus calories (centimeter–gram–seconds unit system), or R_{\min} versus σ values for the Lennard–Jones term. Such problems are typically solved through careful implementation of an FF in a simulation program, an ideally one-time task. However, a more insidious problem is the actual implementation of the FF itself, given the requirement for assigning atom types and parameters to a wide range of diverse chemical structures.

Recently, a study was published applying a high-throughput workflow to leverage the capabilities of high-performance supercomputers for free-energy perturbation calculations.³³ While the computational methods and workflows in the study were well described and validated, the definition and implementation of the FFs used raised important concerns

that went beyond that specific study.³³ Motivated by this³³ and other studies,^{34–36} three issues are addressed in this study. First is the integrity of CGenFF,^{23–26} where the program MATCH³⁷ rather than the CGenFF program^{38,39} was used for atom typing and parameter assignment.³³ As will be shown, the fidelity of MATCH³⁷ in producing CGenFF^{23–26} atom types and parameters is quite low, such that the “CGenFF” referenced in that study did not represent CGenFF^{23–26} of the MacKerell Laboratory and collaborating laboratories. What is presented as representing a widely used FF associated with a specific parameter assignment program is actually a very different FF than that implied based on the nomenclature used in the publication.³³ Second is the use of Austin Model 1-bond charge correction (AM1-BCC) model^{40,41} to assign atomic charges to assess the quality of GAFF2. GAFF2 parameters were developed using HF/6-31G* restrained electrostatic potential (RESP) charges;⁴² RESP charges have been shown to outperform AM1-BCC charges in hydration free-energy calculations with GAFF2.³² Thus, the reported GAFF2 is assessed based on a charge model that may not be able to achieve the original performance of GAFF2. Third, is the opposite issue, where two FFs, OpenFF 1.2.0⁴³ and GAFF,^{27–31} are presented as independent models, while detailed analysis of the nonbonded parameters show them to be nearly identical. This may reflect the fact that SMIRNOFF99Frost,⁴⁴ the predecessor of OpenFF 1.2.0,⁴³

Received: May 16, 2022

Published: August 2, 2022



originated from parent FFs AMBER parm99⁴⁵ and parm@Frosst⁴⁶ and has been reported to be related to, “but not identical” to, GAFF.^{43,44} Accordingly, the goal of the present study is to provide clarification of the integrity of small-molecule FFs and to raise user awareness of the importance of such clarification when applying such FFs.

■ COVERAGE OF PARAMETER ASSIGNMENT PROGRAMS

Nine parameter sets were investigated (Table 1) on the 264 ligands reported in Gapsys et al.³³ The parameter set (first

Table 1. Overview of the Parameter Sets Investigated

parameter set	parameter assignment program	force field
CGenFF	CGenFF v 2.5.1	CGenFF v4.6
CGenFF legacy	CGenFF legacy v 1.0	CGenFF v3.0.1
MATCH/CGenFF	MATCH v 0.01	CGenFF v3.0.1
Gapsys/CGenFF ^a	MATCH v 0.01 ^b	CGenFF v3.0.1
GAFF2/RESP	Psi4, Antechamber, AmberTools21	GAFF2.11
GAFF2/AM1-BCC	Antechamber, AmberTools21	GAFF2.11
Gapsys/GAFF2 ^a	Antechamber, ACPYPE	GAFF2.11
GAFF/AM1-BCC	Antechamber, AmberTools21	GAFF1.81
OpenFF ^a	OpenFF Toolkit	OpenFF 1.2.0

^aParameters were extracted directly from Gapsys et al.³³ ^bAs shown in Figures S1 and S2, while the parameters were indicated to be generated using MATCH, there are differences between the MATCH charges calculated in the present study and those reported in the Gapsys et al. study.

column, Table 1) was used in conjunction with different parameter assignment programs (second column, Table 1) to assign atom types and parameters, yielding the corresponding generalized FF (third column, Table 1), to the investigated set of ligands. All parameter assignment programs, except for MATCH,³⁷ successfully generated parameters for all 264 ligands. Gapsys/CGenFF, Gapsys/GAFF2, and OpenFF parameters for all 264 ligands were already available in the database extracted from the Github repository provided by Gapsys et al.³³ MATCH successfully assigned parameters for 243 ligands and failed to assign parameters for 21 ligands. Thus, all reported data regarding MATCH/CGenFF exclude these 21 ligands.

■ DISCREPANCIES IN ASSIGNED CGENFF ATOM TYPES AND PARAMETERS

None of the studied ligands were assigned the same set of CGenFF or CGenFF legacy atom types and parameters in MATCH/CGenFF or Gapsys/CGenFF; all studied ligands were assigned at least one parameter that differed across the three programs. Despite Gapsys/CGenFF being based on MATCH and CGenFF v3.0.1,³³ the Gapsys/CGenFF parameters include lone pairs (LPs). Neither MATCH nor CGenFF v3.0.1 has LP functionality; thus, it appears that the parameters of ligands containing σ -holes were modified for Gapsys/CGenFF based on the procedure described by CGenFF developers applied to CGenFF v3.0.1 FF²⁶ and released in subsequent versions of CGenFF, including CGenFF v4.6 used in the present study. In addition, over 80% of the ligands contain atoms that are typed differently in MATCH/CGenFF and/or Gapsys/CGenFF (Table S1). Furthermore, nearly half of the studied ligands contain atoms that are typed differently between MATCH/CGenFF and Gapsys/CGenFF (Table S1).

Unsurprisingly, given the discrepancies in atom typing, discrepancies were observed in the bonded parameters (Table S2). Despite the bonded parameters of Gapsys/CGenFF being extracted from CGenFF v3.0.1 independent of MATCH,³³ discrepancies are also present between Gapsys/CGenFF and CGenFF legacy as well as CGenFF (Table S2). All ligands contain at least one differently assigned bonded parameter in Gapsys/CGenFF and MATCH/CGenFF. Notably, MATCH does not assign Urey-Bradley terms to angle parameters where appropriate. Thus, all angle parameters containing Urey-Bradley terms are described differently by MATCH/CGenFF regardless of atom type assignment.

Comparing the assigned partial atomic charges also revealed significant differences among the parameters (Figures 1A and S1 and S2). Compared to CGenFF and CGenFF legacy, over half of the atoms have different charges in MATCH/CGenFF and Gapsys/CGenFF, affecting all the studied ligands (Table S1). The differences in charges are often significant, with several charges differing by more than 0.5 e and including changes in the sign (Figure S2).

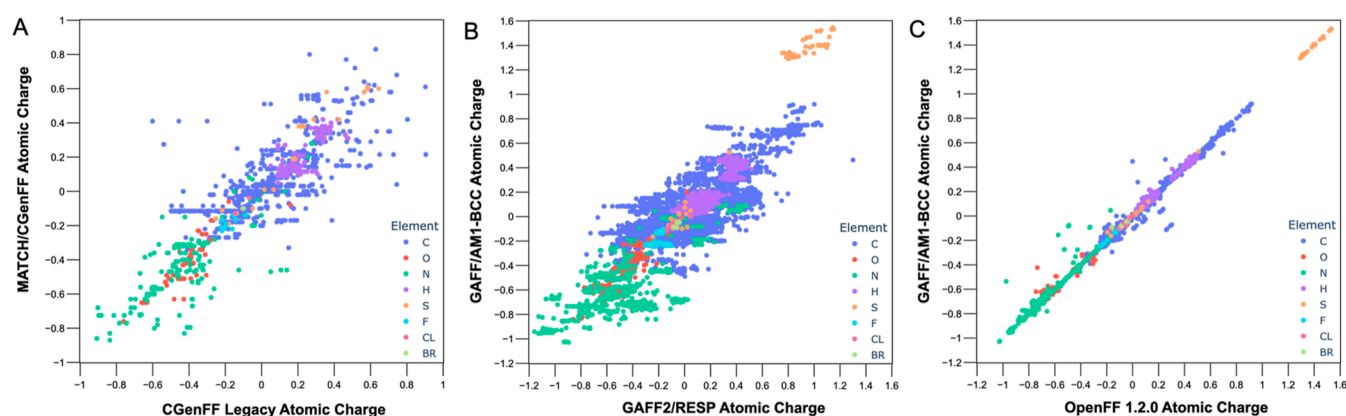


Figure 1. Correlation plots of atomic charges for (A) CGenFF legacy with MATCH/CGenFF, (B) GAFF2/RESP with GAFF2/AM1-BCC, and (C) OpenFF with GAFF2/AM1-BCC. Each datapoint corresponds to an atom.

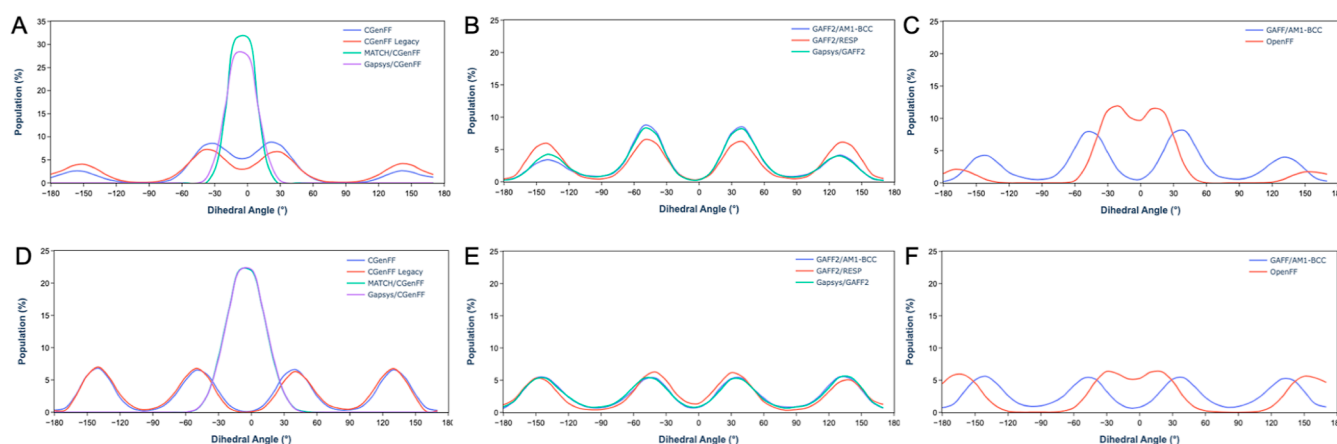


Figure 2. Dihedral angle distributions for the dihedral of atoms (A–C) 11-10-17-18 of lig_CHEMBL3265002 from the syk protein system of Gapsys et al.³³ and (D–F) 5-11-12-17 of lig_8e from the tnks2 protein system of Gapsys et al.³³

■ DIFFERENCES IN ASSIGNED GAFF2 ATOM TYPES AND PARAMETERS

Similar to Gapsys/CGenFF, Gapsys/GAFF2 also includes LPs.³³ Currently, GAFF2 does not have the LP functionality, thus it appears that the authors followed the published rules for off-site-charge generation for GAFF family of FFs,^{47,48} although details of the charge assignment approach were not presented.³³ Beyond the presence of added LPs, nearly 20% of the ligands contain discrepancies in atom typing by Gapsys/GAFF2 compared to GAFF2/AM1-BCC and GAFF2/RESP (Table S3). These discrepancies also affect the assigned bonded parameters (Table S4).

Beyond atom typing and bonded parameters, a comparison of the atomic charges assigned in GAFF2/RESP, GAFF2/AM1-BCC, and Gapsys/GAFF2 shows that the AM1-BCC and RESP charge assignments are not equivalent (Figures 1B and S3A,B). While AM1-BCC provides an efficient method to generate high-quality atomic charges for small molecules, GAFF and GAFF2 were originally developed with HF/6-31G* RESP charges.^{27–32} Additionally, the use of GAFF2/AM1-BCC has been shown to result in systemic errors greater than 1.0 kcal/mol in solvation free-energy calculations for several specific functional groups,³² which may significantly limit the accuracy of computationally predicted protein–ligand binding free energies. A new AM1-BCC model (ABCG2)³² developed specifically for GAFF2 may be used to circumvent quantum mechanics calculations needed to derive RESP charges. This charge model paves the road for the development of the next generation of GAFF-GAFF3. How well ABCG2 performs in protein–ligand binding calculations is under investigation. Aside from the fact that AM1-BCC atomic charges with GAFF2 may introduce systemic errors in free-energy calculations,³² the atomic charges of Gapsys/GAFF2 are not equivalent to those of GAFF2/AM1-BCC for several atoms (Figure S3C).

■ SIMILARITIES IN GAFF AND OPENFF 1.2 PARAMETERS

Comparisons of GAFF and OpenFF showed striking similarities between the two FFs; this may be expected as OpenFF has been reported to share “common ancestry” with (yet be distinct from) GAFF.^{43,44} Unsurprisingly, due to the use of the AM1-BCC model to derive atomic charges in both GAFF/AM1-BCC and OpenFF, the assigned atomic charges

are nearly identical (Figure 1C). Interestingly, the assigned Lennard–Jones parameters σ and ϵ of GAFF/AM1-BCC and OpenFF are also similar (Figure S4A,B). Despite OpenFF assigning unique atom types to each individual atom, only a handful of σ and ϵ values are used, similar to GAFF/AM1-BCC (Figure S4A,B). There are larger apparent differences between the assigned σ and ϵ values of GAFF/AM1-BCC with GAFF2/AM1-BCC compared with OpenFF (Figure S4C,D).

In addition, the assigned bonded parameters by GAFF/AM1-BCC and OpenFF are often similar (Table S5). Approximately a sixth of the bond, angle, and dihedral force constants of OpenFF are within 10% difference of GAFF/AM1-BCC (Table S5), while all equilibrium bond parameters and the majority of equilibrium angles, dihedral periodicities, and dihedral phase parameters are within 10% difference. Moreover, nearly all the bond and angle parameters and over half of the dihedral parameters of OpenFF are within 35% difference of GAFF/AM1-BCC (Table S6). None of the improper parameters of OpenFF are similar to GAFF/AM1-BCC as the two FFs treat improper torsions differently.

■ STRUCTURAL AND ENERGETIC CONSEQUENCES OF ASSIGNED PARAMETERS

Molecular dynamics (MD) simulations of two representative ligands (lig_CHEMBL3265002 and lig_8e from Gapsys et al.³³), followed by dihedral angle distribution calculations, root-mean-square-deviation (RMSD)-based clustering, and interaction energy calculations show that the differences (and similarities) in assigned parameters directly affect the structural and energetic properties of the ligands.

CGenFF and CGenFF legacy allow more conformational flexibility in the simulated ligands compared to MATCH/CGenFF and Gapsys/CGenFF (Figures 2 and S6 and S7); MATCH/CGenFF and Gapsys/CGenFF frequently yield significantly different dihedral angle distributions from the simulations, with those distributions being more narrow (Figures 2 and S6 and S7). The atoms associated with these dihedral angles are assigned alternate atom types, have large discrepancies in charge, and/or are involved in the assignment of different bonded parameters (Figure S5, Tables S7 and S8). The rigidity imposed by MATCH/CGenFF and Gapsys/CGenFF disallows the ligands from adopting conformations allowed by CGenFF and CGenFF legacy (Figure S8). Furthermore, the average interaction energies of the ligands

Table 2. Proposed Nomenclature for Generalized FFs, CGenFF and GAFF/GAFF2, Reported in Scientific Literature

parameter assignment program	force field	charge model	ideal nomenclature	abbreviated nomenclature
CGenFF v 2.5.1	CGenFFv4.6	N/A	CGenFF program v2.5.1/CGenFF v4.6	CGenFF ^a
CGenFF legacy v 1.0	CGenFFv3.0.1	N/A	CGenFF legacy program v1.0/CGenFFv 3.0.1	CGenFF legacy
MATCH v 0.01	CGenFFv3.0.1	N/A	MATCH/CGenFF v3.0.1	MATCH/CGenFF
Antechamber, AmberTools21	GAFF2.11	RESP	GAFF2.11/RESP	GAFF2/RESP
Antechamber, AmberTools21	GAFF2.11	AM1-BCC	GAFF2.11/AM1-BCC	GAFF2/AM1-BCC
Antechamber, AmberTools21	GAFF1.81	RESP	GAFF1.81/RESP	GAFF/RESP
Antechamber, AmberTools21	GAFF1.81	AM1-BCC	GAFF1.81/AM1-BCC	GAFF/AM1-BCC

^aThe abbreviation “CGenFF” should be reserved only for the latest versions of CGenFF program and FF at the time of the study.

with the surrounding solvent are more favorable using CGenFF and CGenFF legacy compared to MATCH/CGenFF or Gapsys/CGenFF (Figure S9).

Differences in the conformational behavior of the ligands assigned using GAFF2/RESP and GAFF2/AM1-BCC or Gapsys/GAFF2 were also observed. The dihedral angle distribution of the ligands are slightly different with GAFF2/RESP compared to GAFF2/AM1-BCC and Gapsys/GAFF2 despite the use of the same bonded parameters due to the differences in assigned atomic charges (Figures 2 and S6 and S7 and Table S9). The difference in the dihedral angle distribution with GAFF2/RESP is also reflected in the slightly larger number of distinct conformations adopted by the ligands with GAFF2/RESP (Figure S8). Additionally, despite the use of the same parameters, the assigned charges of lig_4-CHEMBL3265002 affect its water interaction energy, making it less favorable with GAFF2/RESP compared to GAFF2/AM1-BCC and Gapsys/GAFF2 (Figure S9). Variations in the conformations and interaction energies for lig_8e are relatively small, likely due to the small number of rotatable bonds (two rotatable bonds in lig_8e compared to nine in lig_4-CHEMBL3265002, Figure S5).

The overall conformational behavior of both ligands is similar with GAFF/AM1-BCC and OpenFF, despite the differences in their dihedral angle distributions. The dihedral angle distributions of OpenFF appear to loosely mimic those of GAFF/AM1-BCC, with a wider spread for several dihedral angles (Figures 2 and S6 and S7). Additionally, the ligands adopt nearly the same distinct conformations with OpenFF as with GAFF/AM1-BCC; only a few structures fall into mutually exclusive clusters (Figure S8). The average interaction energy of the ligands with OpenFF is similar to GAFF/AM1-BCC (Figure S9). Overall, it appears that OpenFF1.2.0 generally reproduces the behavior of the original GAFF through similar nonbonded and bonded parameters, despite being derived from refitting SMIRNOFF99Frost⁴⁴ parameters against several sets of high-level quantum mechanical data.⁴³

BEST PRACTICES FOR PROPOSING THE NOMENCLATURE OF SMALL-MOLECULE FFs

The present results show that the specific parameter assignment programs and charge models used have a notable effect on the generalized atom types and parameters assigned. This in turn affects the properties of the assigned generalized FF in molecular simulations. Thus, it is important to apply a descriptive naming scheme in the scientific literature to avoid misrepresenting FFs and misleading readers as to what atom types and parameters were used.

We propose the nomenclature presented in Table 2. It requires that the FF version, parameter assignment program version, and/or charge model used be explicitly stated prior to

using any abbreviation for the generalized FF being used. Further, researchers should reserve the use of the abbreviations “CGenFF” and “GAFF2” for the most current (at the time of their studies) versions of the FFs and their respective parameter assignment programs. Proper specification of the version number is particularly important as with any machine learning-based models, and the assigned parameters of newer versions of the CGenFF program may differ as the training sets and associated FF are updated. For CGenFF, the version of the CGenFF program and FF used is available in the header of the stream file output of the CGenFF program (Figure S10). For GAFF and GAFF2, the charge model is explicitly entered by the user in the Antechamber command line (Figure S11A), and the versions of GAFF and GAFF2 are available in the “gaff.dat” and “gaff2.dat” files, respectively, provided by AmberTools (Figure S11B). This information should be included in the description of the FF as shown in Table 2.

CONCLUSIONS

Parameter assignment programs allow the application of generalized small-molecule FFs to diverse chemical structures. Here, we showed that (1) MATCH frequently assigns atom types and parameters that differ from CGenFF, affecting all ligands investigated in this study, (2) AM1-BCC charges are distinct from RESP charges, and thus GAFF2 performs differently based on the charge model used, and (3) OpenFF and GAFF, particularly their nonbonded parameters, are similar. These differences and similarities directly affect the conformational and energetic properties of ligands in MD simulations, as presented for two example compounds. Accordingly, researchers must carefully implement generalized FFs using the appropriate atom typing and parameter assignment programs when considering their research aims. For simulation studies using the CHARMM36 biomolecular FF in which the standard CGenFF parameters are desired, the CGenFF program with the latest CGenFF parameters (currently v4.6) should be used, while MATCH represents a viable option to assign CGenFF parameters, though the resulting FF differs significantly. For AMBER studies in which high accuracy is required, we recommend GAFF2/RESP over GAFF2/AM1-BCC, although AM1-BCC is a reasonable substitute in studies on a large number of molecules. Users should also inspect the format of their input ligand structure (e.g., correctly assigned bond orders in MOL2 files) to ensure the peak performance of the parameter assignment programs. The CGenFF program robustly assigns parameters with the use of MOL2 files converted from PDB by Open Babel⁴⁹ (as in this study), Molecular Operating Environment (MOE),⁵⁰ Schrödinger Maestro,⁵¹ and Marvin Sketch,⁵² resulting in the same topologies and parameters. However, variation with MATCH when using MOE generated MOL2 files were

observed as described in the Supporting Information (Impact of method for PDB to MOL2 conversion, Figure S12). Nonetheless, the presented results remain largely the same regardless of the source of the initial ligand files, which was further tested through use of SDF files, directly from Schindler et al.,⁵³ as input rather than PDB files, directly from Gapsys et al.³³ (Tables S12 and S13). In addition, users should be aware of differences in the standard⁵⁴ and CHARMM specific versions^{10,55} of the TIP3P water model and use the model recommended for the specific FF being used. Importantly, we encourage future publications to follow standard nomenclature, explicitly stating the versions of the FFs, parameter assignment program, and/or charge models used. Overall, the presented results highlight the importance of the selected parameter assignment program as well as the small-molecule FF to maintain the integrity of the FF used.

DATA AND SOFTWARE AVAILABILITY

All methodological details are provided in the Supporting Information. All simulation data are available upon request. The generated ligand topology and parameter files are available at https://github.com/mackereell-lab/small_mol_FF_toppar_JCIM.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c00615>.

Details of methods, correlation plots of non bonded parameters, distribution of atomic charge differences, atom numbers of simulated ligands, dihedral angle distributions of simulated ligands, population of RMSD-based clusters, average interaction energies of the simulated ligands with the surrounding solvent molecules, identification of CGenFF program and FF version through CGenFF program output, identification of GAFF version and charge model, tabulated data of discrepancies in atom types and partial atomic charges along with the number of ligands affected by discrepancies, tabulated data of percent bonded parameters with discrepancies and ligands affected by the discrepancies, tabulated data of similarities in assigned bonded parameters, and assigned atom types and atomic charges for simulated ligands (PDF)

AUTHOR INFORMATION

Corresponding Author

Alexander D. MacKerell, Jr. — Department of Pharmaceutical Sciences, University of Maryland, Baltimore, Maryland 21201, United States; orcid.org/0000-0001-8287-6804; Email: alex@outerbanks.umaryland.edu

Authors

Asuka A. Orr — Department of Pharmaceutical Sciences, University of Maryland, Baltimore, Maryland 21201, United States; orcid.org/0000-0003-4628-526X

Suliman Sharif — Department of Pharmaceutical Sciences, University of Maryland, Baltimore, Maryland 21201, United States

Junmei Wang — Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center,

University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States; orcid.org/0000-0002-9607-8229

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jcim.2c00615>

Notes

The authors declare the following competing financial interest(s): A.D.M. Jr. is co-founder and CSO of SilcsBio LLC.

ACKNOWLEDGMENTS

The authors acknowledge financial support from NIH GM131710.

REFERENCES

- (1) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (2) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. Charmm Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370.
- (3) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the Charmm All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
- (4) Raman, E. P.; Guvench, O.; MacKerell, A. D. Charmm Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses. *J. Phys. Chem. B* **2010**, *114*, 12981–12994.
- (5) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. Impact of 2'-Hydroxyl Sampling on the Conformational Properties of Rna: Update of the Charmm All-Atom Additive Force Field for Rna. *J. Comput. Chem.* **2011**, *32*, 1929–1943.
- (6) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D. Charmm Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate-Protein Modeling. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180.
- (7) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive Charmm All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Φ , Ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (8) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the Charmm Additive Force Field for DNA: Improved Treatment of the Bi/Bii Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.
- (9) Mallajosyula, S. S.; Guvench, O.; Hatcher, E.; MacKerell, A. D. Charmm Additive All-Atom Force Field for Phosphate and Sulfate Linked to Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 759–776.
- (10) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. Charmm36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71–73.
- (11) Bergonzo, C.; Cheatham, T. E. Improved Force Field Parameters Lead to a Better Description of Rna Structure. *J. Chem. Theory Comput.* **2015**, *11*, 3969–3972.
- (12) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (13) Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T. E.; Galindo-Murillo, R.; Jurečka, P. Refinement of the Sugar-Phosphate Backbone Torsion Beta for Amber Force Fields Improves the

Description of Z- and B-DNA. *J. Chem. Theory Comput.* **2015**, *11*, 5723–5736.

(14) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12*, 4114–4127.

(15) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migués, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. Ff19sb: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16*, 528–552.

(16) Dickson, C. J.; Walker, R. C.; Gould, I. R. Lipid21: Complex Lipid Membrane Simulations with Amber. *J. Chem. Theory Comput.* **2022**, *18*, 1726–1736.

(17) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The Gromos Force-Field Parameter Sets 53a5 and 53a6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(18) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. Opls3: A Force Field Providing Broad Coverage of Drug-Like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296.

(19) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S. a.; Karplus, M. Charmm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(20) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(21) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1*–2, 19–25.

(22) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on Cpu and Gpu Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130.

(23) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D., Jr. Charmm General Force Field: A Force Field for Drug-Like Molecules Compatible with the Charmm All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(24) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D., Jr. Extension of the Charmm General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

(25) Hegazy, L.; Richards, N. G. Optimized Cgenff Force-Field Parameters for Acylphosphate and N-Phosphonosulfonimidoyl Functional Groups. *J. Mol. Model.* **2013**, *19*, 5075–5087.

(26) Soteras Gutiérrez, I.; Lin, F. Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L., 3rd; MacKerell, A. D., Jr. Parametrization of Halogen Bonds in the Charmm General Force Field: Improved Treatment of Ligand-Protein Interactions. *Bioorg. Med. Chem.* **2016**, *24*, 4812–4825.

(27) 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.

(28) Ozpinar, G. A.; Peukert, W.; Clark, T. An Improved Generalized Amber Force Field (Gaff) for Urea. *J. Mol. Model.* **2010**, *16*, 1427–1440.

(29) Jämbek, J. P.; Lyubartsev, A. P. Update to the General Amber Force Field for Small Solutes with an Emphasis on Free Energies of Hydration. *J. Phys. Chem. B* **2014**, *118*, 3793–3804.

(30) Boyd, N. J.; Wilson, M. R. Optimization of the Gaff Force Field to Describe Liquid Crystal Molecules: The Path to a Dramatic Improvement in Transition Temperature Predictions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24851–24865.

(31) Torsello, M.; Pimenta, A. C.; Wolters, L. P.; Moreira, I. S.; Orian, L.; Polimeno, A. General Amber Force Field Parameters for Diphenyl Diselenides and Diphenyl Ditellurides. *J. Phys. Chem. A* **2016**, *120*, 4389–4400.

(32) He, X.; Man, V. H.; Yang, W.; Lee, T. S.; Wang, J. A Fast and High-Quality Charge Model for the Next Generation General Amber Force Field. *J. Chem. Phys.* **2020**, *153*, 114502.

(33) Gapsys, V.; Hahn, D. F.; Tresadern, G.; Mobley, D. L.; Rampp, M.; de Groot, B. L. Pre-Exascale Computing of Protein-Ligand Binding Free Energies with Open Source Software for Drug Design. *J. Chem. Inf. Model.* **2022**, *62*, 1172–1177.

(34) Lundborg, M.; Lindahl, E. Automatic Gromacs Topology Generation and Comparisons of Force Fields for Solvation Free Energy Calculations. *J. Phys. Chem. B* **2015**, *119*, 810–823.

(35) Elisée, E.; Gapsys, V.; Mele, N.; Chaput, L.; Selwa, E.; de Groot, B. L.; Iorga, B. I. Performance Evaluation of Molecular Docking and Free Energy Calculations Protocols Using the D3r Grand Challenge 4 Dataset. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 1031–1043.

(36) Gapsys, V.; Pérez-Benito, L.; Aldeghi, M.; Seeliger, D.; van Vlijmen, H.; Tresadern, G.; de Groot, B. L. Large Scale Relative Protein Ligand Binding Affinities Using Non-Equilibrium Alchemy. *Chem. Sci.* **2020**, *11*, 1140–1152.

(37) Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L., 3rd Match: An Atom-Typing Toolset for Molecular Mechanics Force Fields. *J. Comput. Chem.* **2012**, *33*, 189–202.

(38) Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the Charmm General Force Field (Cgenff) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.

(39) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. 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.

(40) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. Am1-Bcc Model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132.

(41) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. Am1-Bcc Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

(42) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(43) Qiu, Y.; Smith, D. G. A.; Boothroyd, S.; Jang, H.; Hahn, D. F.; Wagner, J.; Bannan, C. C.; Gokey, T.; Lim, V. T.; Stern, C. D.; Rizzi, A.; Tjanaka, B.; Tresadern, G.; Lucas, X.; Shirts, M. R.; Gilson, M. K.; Chodera, J. D.; Bayly, C. I.; Mobley, D. L.; Wang, L.-P. Development and Benchmarking of Open Force Field V1.0.0—the Parsley Small-Molecule Force Field. *J. Chem. Theory Comput.* **2021**, *17*, 6262–6280.

(44) Mobley, D. L.; Bannan, C. C.; Rizzi, A.; Bayly, C. I.; Chodera, J. D.; Lim, V. T.; Lim, N. M.; Beauchamp, K. A.; Slochower, D. R.; Shirts, M. R.; Gilson, M. K.; Eastman, P. K. Escaping Atom Types in Force Fields Using Direct Chemical Perception. *J. Chem. Theory Comput.* **2018**, *14*, 6076–6092.

(45) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (Resp) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(46) Bayly, C.; McKay, D. An Informal Amber Small Molecule Force Field: Parm@Frosst. 2010. http://www.ccl.net/cca/data/parm_at_Frosst/.

(47) Ibrahim, M. A. Molecular Mechanical Study of Halogen Bonding in Drug Discovery. *J. Comput. Chem.* **2011**, *32*, 2564–2574.

- (48) 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.
- (49) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3*, 33.
- (50) Chemical Computing Group ULC. *Molecular Operating Environment (Moe)*, 2020.09, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2022.
- (51) Schrödinger, L. *Maestro*. Schrödinger Release 2022-2, New York, NY, 2021.
- (52) Chemaxon *Marvinsketch Version 17.1.30.0*, 2017.
- (53) Schindler, C. E. M.; Baumann, H.; Blum, A.; Böse, D.; Buchstaller, H.-P.; Burgdorf, L.; Cappel, D.; Chekler, E.; Czodrowski, P.; Dorsch, D.; Eguida, M. K. I.; Follows, B.; Fuchß, T.; Grädler, U.; Gunera, J.; Johnson, T.; Jorand Lebrun, C.; Karra, S.; Klein, M.; Knehans, T.; Koetzner, L.; Krier, M.; Leiendecker, M.; Leuthner, B.; Li, L.; Mochalkin, I.; Musil, D.; Neagu, C.; Rippmann, F.; Schiemann, K.; Schulz, R.; Steinbrecher, T.; Tanzer, E.-M.; Unzue Lopez, A.; Viacava Follis, A.; Wegener, A.; Kuhn, D. Large-Scale Assessment of Binding Free Energy Calculations in Active Drug Discovery Projects. *J. Chem. Inf. Model.* **2020**, *60*, 5457–5474.
- (54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (55) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98*, 2198–2202.

Recommended by ACS

ULYSSES: An Efficient and Easy to Use Semiempirical Library for C++

Filipe Menezes and Grzegorz M. Popowicz

AUGUST 05, 2022
JOURNAL OF CHEMICAL INFORMATION AND MODELING

READ 

Assessing the Interplay between Functional-Driven and Density-Driven Errors in DFT Models of Water

Etienne Palos, Francesco Paesani, *et al.*

MAY 04, 2022
JOURNAL OF CHEMICAL THEORY AND COMPUTATION

READ 

Observing Noncovalent Interactions in Experimental Electron Density for Macromolecular Systems: A Novel Perspective for Protein–Ligand Interaction Research

Kang Ding, Bo Huang, *et al.*

MARCH 29, 2022
JOURNAL OF CHEMICAL INFORMATION AND MODELING

READ 

Data-Driven Many-Body Potential Energy Functions for Generic Molecules: Linear Alkanes as a Proof-of-Concept Application

Ethan F. Bull-Vulpe, Francesco Paesani, *et al.*

SEPTEMBER 16, 2022
JOURNAL OF CHEMICAL THEORY AND COMPUTATION

READ 

Get More Suggestions >