CHAMBER: Comprehensive Support for CHARMM Force Fields Within the AMBER Software

MICHAEL F. CROWLEY,[1] MARK J. WILLIAMSON,[2] ROSS C. WALKER[2]

[1]*National Renewable Energy Laboratory, 1617 Cole Blvd, Golden, CO 804013393*

[2]*San Diego Supercomputer Center, University of California San Diego, 9500 Gilman Drive, La Jolla, CA*

Received 5 May 2009; accepted 26 May 2009
Published online 30 July 2009 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.22372

ABSTRACT: The similarity of the AMBER force field's energy functional form with that of the CHARMM force field, gives the potential for direct translation of common bonding and nonbonding terms, along with their parameters, present in CHARMM topology and parameter files, with the intent of evaluation within the AMBER software; specifically the SANDER and PMEMD dynamics engines. To this extent, we have created a tool, CHAMBER, which can take a CHARMM protein structure file (PSF), coordinate file (COR) and associated forcefield files, and convert these to an AMBER topology file (prmtop) and associated coordinate file (inpcrd). CHAMBER opens a conversion route which enables the simulation of CHARMM parameterized models using AMBER's PMEMD engine; thus providing improved serial efficiency as well as parallel efficiency over large numbers of CPUs. Significant effort has been expended in ensuring a true representation of the CHARMM force field in AMBER providing energies and forces that are the same to the limits of machine precision. This software will be released in the upcoming version 1.3 of the free AMBERTools suite.[1] Int J Quantum Chem 109: 3767–3772, 2009

Key words: CHARMM; AMBER; converter; PSF; PRMTOP; SANDER; PMEMD

Correspondence to: R. C. Walker; e-mail: ross@rosswalker.co.uk
Contract grant sponsor: DoE SciDAC.
Contract grant number: DE-AC36–99G0–10337.
Contract grant sponsor: UC Lab award.
Contract grant number: 09-LR-06–117792.
Contract grant sponsor: NSF.
Contract grant number: TG-MCB090110.
Contract grant sponsor: San Diego Supercomputer Center Advanced User Support program.

[1]http://ambermd.org.

1. Introduction

AMBER [1, 2] and CHARMM [3–5] are two approaches to the parameterization of classical force fields that find extensive use in the modeling of biological systems. The high similarity in the functional form of the two potential energy functions used by these force fields, Eqs. (1) and (2), gives rise to the possible use of one force field within the other force field's dynamics engine. In the case of the CHARMM force field, its dynamics engine is also called CHARMM [6], whereas for AMBER, the main two dynamic engines available are SANDER and PMEMD, with the latter engine possessing a subset of the functionality of SANDER but having been heavily optimized for parallel performance.

$$V_{\mathrm{AMBER}} = \sum_{\mathrm{bonds}} k_r(r - r_{\mathrm{eq}})^2 + \sum_{\mathrm{angles}} k_\theta(\theta - \theta_{\mathrm{eq}})^2 + \sum_{\mathrm{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \sum_{i<j}\left[\frac{q_i q_j}{\epsilon R_{ij}}\right] \quad (1)$$

$$V_{\mathrm{CHARMM}} = \sum_{\mathrm{bonds}} k_b(b - b_0)^2 + \sum_{\mathrm{angles}} k_\theta(\theta - \theta_0)^2 + \sum_{\mathrm{dihedrals}} k_\phi[1 + \cos(n\phi - \delta)] + \sum_{\mathrm{Urey\text{-}Bradley}} k_u(u - u_0)^2 + \sum_{\mathrm{impropers}} k_\omega(\omega - \omega_0)^2 + \sum_{\phi,\psi} V_{\mathrm{CMAP}} + \sum_{\mathrm{nonbonded}} \epsilon\left[\left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{\epsilon r_{ij}} \quad (2)$$

For the implementation of the CHARMM force field within an AMBER dynamics engine, parameters that are of the same energy term can be directly translated. However, there are differences in the two functional forms, with CHARMM having three additional bonded terms. Additionally, the force fields scale 1–4 interactions in different manners.

The first additional term is CHARMM's two body Urey–Bradley term, which extends over all 1–3 bonds, the second is a four body quadratic improper term and the final additional term is a cross term, CMAP [7, 8], which is a function of two sequential carbon back bone dihedrals. This term originates from differences observed between classically calculated two-dimensional ϕ/ψ peptide free energy surfaces using the CHARMM22 force field and that of experiment. CMAP is a numerical energy correction which essentially transforms the 2D ϕ/ψ classical energy map to match that of a QM calculated map.

Support for these extra terms have required the development of extensible sections to AMBER's prmtop format to accommodate this new information. In concert with these prmtop additions, the appropriate modifications have been made within SANDER's and PMEMD's code to enable the calculation of the energy and derivatives corresponding to these new terms. Additionally, because of the modular nature of these additions, backwards compatibility within the prmtop file has been maintained. MPI support for the parallel evaluation of these three extra terms is also provided within SANDER and PMEMD.

Additional checks are carried out within the SANDER code at runtime when using a CHARMM force field; for instance the 1–4 nonbond electrostatic scaling factor (SCEE) is checked to be 1.0 and not 1.2. The same is true with the van der Waal's terms scaling factor (SCNB); this is checked to be 1.0 and not 2.0 as per the AMBER default; deviations from this will raise an error. The intention behind the approach of creating a CHARMM enabled prmtop file is that the use of this prmtop file should be transparent to the user. Once a CHARMM prmtop file is produced by CHAMBER the SANDER and PMEMD dynamics engines automatically detect the presence of CHARMM parameters in the prmtop file and automatically select the correct parameters and code paths.

CHAMBER and the corresponding modifications to SANDER and PMEMD have been implemented in a development version of AMBER and will be present in the release version of AMBER 11. Additionally, it is anticipated that patches to enable CHARMM support will soon be made available for AMBER 10.

1.1. SUPPORT FOR CHARMM IN OTHER MD PACKAGES

A number of packages claim to support the CHARMM force field. These include Gromacs [9], LAMMPS [10, 11], Desmond [12], and NAMD [13] to name a few. Although all of these packages claim to provide support it is not clear how extensive such support is and how comprehensively it has been validated. LAMMPS for example does not support CMAP, while Gromacs relies on a number of third party, poorly documented scripts, for conversion which do not appear to be actively maintained and could not be made to work as part of this project.

TABLE I

Timings to complete the Joint AMBER CHARMM (JAC) benchmark on a dual quad core Intel(R) Xeon(R) CPU 2.66 GHz desktop machine using 8 processors.

MD program	Time (s)	Throughput (ns/day)
PMEMD	32.83	2.63
SANDER	57.73	1.50
CHARMM	99.00	0.87

All programs were compiled with their default settings using the same version of the Intel Fortran compiler (v. 10.1.008) and linked against the same MPI implementation (MPICH2 v1.0.7). Version c35b1 of CHARMM was used and a development version of AMBER; these specific binaries are used for all evaluations in this article.

Desmond provides extensive support for the CHARMM force field including CMAP terms, however, the different approach used for dealing with floating point arithmetic along with alternative approaches for the treatment of periodic boundaries means that it does not necessarily offer a direct replacement for carrying out a simulation with CHARMM.

NAMD supports the CHARMM force field, if the PSF is written in X-PLOR format. Internal testing showed it had good correspondence to CHARMM's output across all energy terms, however, different electrostatic energies are seen since NAMD uses an inverse Coulomb constant value of 332.0636, whereas CHARMM historically uses 332.0716. This is a subtle point, but if one wants a faithful implementation of the CHARMM force field, then one needs to be using the same value as CHARMM.

2. Motivations

The motivations behind the creation of CHAMBER are numerous, however, the overriding motivation is one of being able to make use of AMBER's highly scalable and efficient PMEMD code to run CHARMM force field simulations within a validated framework that provides a faithful reproduction of the CHARMM force field to machine precision.

A trivial and nonexhaustive example of the respective parallel scaling of PMEMD, SANDER, and CHARMM, can be demonstrated with the Joint AMBER CHARMM (JAC) benchmark[2] as shown in Table I. The motivation lies within this trend; CHARMM is already falling behind performance wise on 8 processors. Since the additional CHARMM terms being incorporated into SANDER and PMEMD are valence in nature so the impact on the performance of the AMBER codes by calculating these additional terms is minimal.

[2]http://ambermd.org/jac1000.tar.gz.

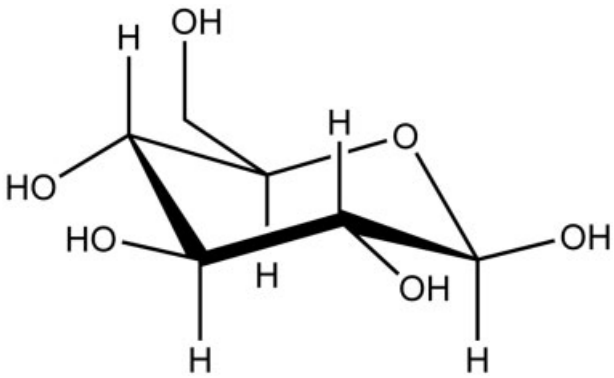

FIGURE 1. CHARMM generated monosaccharide, glucose.

3. VALIDATION

In addition to preserving the high performance of the AMBER PMEMD software when running the CHARMM force field extensive work has been un-

TABLE II

Single point energy comparisons of a PSF and associated COR file evaluated in CHARMM c35bl and the corresponding CHAMBER produced prmtop and inpcrd evaluated in SANDER for the "Glucose" test case.

Energy	SANDER	CHARMM	Difference
BOND	1.2531	1.2531	0.0000
ANGLE	3.1015	3.1015	0.0000
DIHED	−24.8158	−24.8158	−0.0000
ELEC	83.8506	83.8506	−0.0000
VDW	3.1707	3.1707	−0.0000

The AMBER and CHARMM columns show the respective energy terms from the SANDER and CHARMM outputs. The CHARMM angle and Urey–Bradley term energies have been summed into one angle value, and the CHARMM dihedral, improper, and CMAP energies have been summed into one dihedral value. The normal and 1–4 contributions for AMBER's electrostatic and van der Waal energies have been combined respectively. Decompositions of all of these are shown in Table III. All energy values are in Kcal/Mol.

TABLE III
Energy decompositions for Table II.

Energy	Decompositions			
ANGLE	chm ang, ub	3.1015	0.0000	
DIHED	chm dhih, impr, cmap	−24.8158	0.0000	0.0000
ELEC	amb elec, ee14	−84.0824	167.9330	
VDW	amb vdw, 14vdw	−1.0273	4.1980	

All energy values are in Kcal/Mol.

dertaken to ensure that the CHARMM force field is faithfully reproduced in the AMBER software. This includes ensuring that results always match to machine precision.

The first stage of this validation is to ensure that the results for a given PSF (topology) and COR (coordinate) evaluated within the CHARMM engine, match that of the translated prmtop (topology) and inpcrd (coordinate) files evaluated within AMBER's respective engine. The validation protocol consists of the following sequence:

1. Construction of an arbitrary system within CHARMM.
2. Generation of the system's corresponding PSF and COR files.
3. Rereading of these two files and evaluation of the system's potential energy.
4. Conversion of the PSF and COR files, using the associated CHARMM parameter and topology files, with CHAMBER, to produce a prmtop and inpcrd output.
5. Evaluation of the system's potential energy within SANDER.
6. Comparison of CHARMM and SANDER's energy outputs; the differences being zero or close to, indicate that the CHARMM force field is being faithfully reproduced within SANDER.

The aforementioned protocol was applied to a series of test cases.

3.1. GLUCOSE

In the first test, the monosaccharide, glucose (see Fig. 1), was constructed and minimized within CHARMM. This initial test case is very simple consisting of just 24 atoms with no CHARMM specific extra energy terms being utilized. The CHARMM force field parameters used for this test were from the Carbohydrate Solution Force Field (CSFF/CHARMM) [14]. Table II shows the result of applying the validation protocol while Table III shows the explicit decompositions of these energy terms.

This test highlights the fact that any CHARMM like parameters can be used; it is not confined to the standard CHARMM [3, 4] parameter set.

3.2. ALA ALA ALA

In the second test, a tripeptide consisting of three Alanine residues was generated (see Fig. 2) within CHARMM. With parameters from the CHARMM22 force field with CMAP applied there are five improper terms with three unique types and one cross term with one cross term type within this 33 atom gas phase system. Table IV is the result of applying the validation protocol while Table V shows the explicit decompositions of these energy terms.

3.3. DHFR

A third, larger, test system represents a complete protein system and was built from an X-ray crystal structure of dihydrofolate reductase (DHFR). Tables VI and VII show the result of applying this process to the 2,489 atoms of a gas phase DHFR structure. There are 418 improper terms with 19 unique types and 157 cross terms with four unique types; overall a much

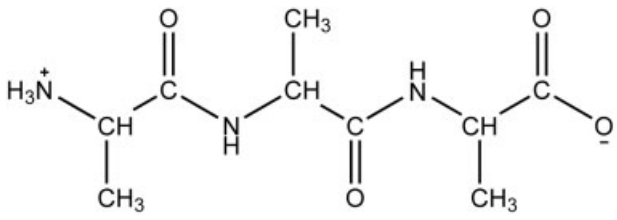

FIGURE 2. CHARMM generated tripeptide structure "ALA ALA ALA."

TABLE IV

Single point energy comparsions of a PSF and associated COR file evaluated in CHARMM c35b1 and the corresponding CHAMBER produced prmtop and inpcrd evaluated in SANDER for the "Ala Ala Ala" testcase.

Energy	SANDER	CHARMM	Difference
BOND	1.3460	1.3460	−0.0000
ANGLE	14.4829	14.4829	0.0000
DIHED	14.0846	14.0846	−0.0000
ELEC	8.4477	8.4477	−0.0000
VDW	0.7644	0.7645	−0.0001

Term amalgamations are as per Table II and decom-positions of all these are shown in Table V. All energy values are in Kcal/Mol.

TABLE VI

Single point energy comparisons of a PSF and associated COR file evaluated in CHARMM c35b1 and the corresponding CHAMBER produced prmtop and inpcrd evaluated in SANDER for the "DHFR" test case.

Energy	SANDER	CHARMM	Difference
BOND	145.1972	145.1972	−0.0000
ANGLE	465.2794	465.2794	−0.0000
DIHED	582.4940	582.4940	−0.0000
ELEC	−3515.0773	−3515.0771	−0.0002
VDW	−634.5064	−634.5064	0.0000

Term amalgamations are as per Table II and decompositions of all these are shown in Table VII. All energy values are in Kcal/Mol.

larger test case than the previous two. As can be seen from the tables, agreement between SANDER and CHARMM is very good.

It is important to note that for a faithful reproduction of a force field, correct energies alone are not sufficient: the associated analytical gradients must be correct as well. SANDER has an internal gradient checking method which calculates the gradient on a given atom numerically and then compares this to the analytical result. Five atoms within the "ALA ALA ALA" test case's peptide backbone (atom numbers 11, 13, 15, 21, 23) were selected for this analysis, with the criterion here being that these atoms were encompassed by the new energy terms; in this case these were covered by the one CMAP term and multiple improper terms. The result was the same analytical and numerical answers within a maximum RMS value of 0.264×10^{-7} Kcal/Mol/Å.

4. Conclusions

In conclusion, an open source tool has been produced that will convert CHARMM force field, psf and cor, pdb or restart files into prmtop and inpcrd files that can be read by the AMBER MD engines. Modified versions of AMBER's SANDER and PMEMD MD engines can use the information in such files to enable faithful CHARMM force field simulations that can reproduce CHARMM c35b1 simulations to machine precision. Both CHAMBER and the associated modified SANDER and PMEMD software have been extensively validated. The CHAMBER software will be made available as part of the next release of the AMBER tools package with support for CHARMM prmtop files being made available in the next version (v11) of the AMBER software. Additionally, it is anticipated that patches to enable such support in AMBER v10 will be made available shortly on the AMBER website.[3]

[3]http://ambermd.org.

TABLE V

Energy decompositions for Table IV.

Energy	Decompositions			
ANGLE	chm ang, ub	14.1159	0.3669	
DIHED	chm dhih, impr, cmap	14.2720	0.3357	−0.5220
ELEC	amb elec, ee14	−269.6955	278.1432	
VDW	amb vdw, 14vdw	−1.3299	2.0943	

All energy values are in Kcal/Mol.

TABLE VII

Energy decompositions for Table VI.

Energy	Decompositions			
ANGLE	chm ang, ub	433.6936	31.5858	
DIHED	chm dhih, impr, cmap	776.0660	22.4542	−216.0261
ELEC	amb elec, ee14	−10023.0276	6507.9503	
VDW	amb vdw, 14vdw	−1004.9337	370.4273	

All energy values are in Kcal/Mol.

References

1. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
2. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. AMBER 10. University of California: San Francisco, 2008.
3. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R.; Evanseck, J.; Field, M.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D.; Prodhom, B.; Reiher, W.; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
4. MacKerell, A. D., Jr.; Nilesh Banavali, N. F. Biopolymers 2000, 56, 257.
5. MacKerell, A. D, Jr.; Brooks, C. L., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and its Parameterization with an Overview of the Program. John Wiley & Sons: Chichester, 1998; p 271.
6. Brooks, B.; Bruccoleri, R.; Olafson, D.; States, D.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187.
7. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. J Am Chem Soc 2004, 126, 698.
8. Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. J Comput Chem 2004, 25, 1400.
9. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.; Berendsen, H. J Comput Chem 2005, 26, 1701.
10. Plimpton, S. J Comput Phys 1995, 117, 1.
11. Plimpton, S.; Pollock, R.; Stevens, M. PPSC SIAM, Minneapolis, Minnesota, 1997.
12. Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. ACM: New York, NY, 2006; p 84.
13. Phillips, J.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.; Kale, L.; Schulten, K. J Comput Chem 2005, 26, 1781.
14. Kuttel, M.; Brady, J. W.; Naidoo, K. J. J Comput Chem 2002, 23, 1236.