

Derivation of a molecular mechanics force field for cholesterol*

Zoe Cournia, Andrea C. Vaiana, G. Matthias Ullmann, and
Jeremy C. Smith[‡]

*IWR – Computational Molecular Biophysics, University of Heidelberg,
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany*

Abstract: As a necessary step toward realistic cholesterol:biomembrane simulations, we have derived CHARMM molecular mechanics force-field parameters for cholesterol. For the parametrization we use an automated method that involves fitting the molecular mechanics potential to both vibrational frequencies and eigenvector projections derived from quantum chemical calculations. Results for another polycyclic molecule, rhodamine 6G, are also given. The usefulness of the method is thus demonstrated by the use of reference data from two molecules at different levels of theory. The frequency-matching plots for both cholesterol and rhodamine 6G show overall agreement between the CHARMM and quantum chemical normal modes, with frequency matching for both molecules within the error range found in previous benchmark studies.

INTRODUCTION

Cholesterol is a major component of the mammalian plasma cell membranes and represents about 50 % of the membrane lipid. It has been proposed that evolution has selected cholesterol because it optimizes the physical properties of membranes for biological function [1]. Cholesterol influences strongly the mechanical and thermodynamical properties of the membrane [2–7] and thus regulates membrane fluidity and permeability and adjusts the lateral mobility of membrane proteins [8–12].

Molecular dynamics (MD) simulation of biological membranes is a growing field of interest. As a first step toward atomic detail MD simulation of biological membranes, simulations of simple single-component lipid bilayers have been performed [13–19]. Extension to systems incorporating cholesterol have also been reported [20–27]. In the past few years, a growing number of simulation studies have addressed problems concerning the interactions between cholesterol and lipids and how cholesterol is organized in membranes [20–27]. However, the results are often contradictory. Although research over the past decades revealed that cholesterol plays an important role in biological membranes, many questions concerning the role of cholesterol, its organization, and interactions within biological membranes remain unanswered. Moreover, it is still unclear which particular characteristics of this molecule have led to its selection through evolution, even though its precursors (e.g., lanosterol) are structurally very similar.

Most empirical force fields used in common MD packages (such as CHARMM [28]), are equipped with parameter sets for modeling the basic building blocks of biomolecules, but often not for more exotic molecules such as cholesterol. Moreover, suitable experimental data (e.g., X-ray crystal structures, vibrational spectra, NMR measurements) against which one could refine new parameters are

*Lecture presented at the European Molecular Liquids Group (EMLG) Annual Meeting on the Physical Chemistry of Liquids: Novel Approaches to the Structure, Dynamics of Liquids: Experiments, Theories, and Simulation, Rhodes, Greece, 7–15 September 2002. Other presentations are published in this issue, pp. 1–261.

[‡]Corresponding author: Fax: +49-6221-54-8868; E-mail: biocomputing@iwr.uni-heidelberg.de

scarce. Consequently, it is often necessary to derive force fields using results from high-level quantum chemistry.

Here we present a new parameter set for cholesterol for the CHARMM force field. The parameters were obtained using an automated frequency-matching method [29]. The method involves careful choice of an initial parameter set, which is then fitted to match vibrational eigenvector and eigenvalue sets derived from quantum chemical calculations. This method is based on a harmonic approximation of the molecular potential energy surface and is thus well suited for modeling physical properties of rigid molecules, as is the case for cholesterol. The new results for cholesterol are presented here together with previously obtained results for another polycyclic molecule, the fluorophore rhodamine 6G. For both molecules, good agreement is obtained between the modes calculated with the new force field and with high-level quantum mechanics.

METHODS

Computational details

For cholesterol (Fig. 1), the quantum chemical calculations were performed with the NWChem 4.0.1 package [30]. A DFT (B3LYP) geometry optimization was performed of the isolated cholesterol molecule. To reduce computational time, an effective core potential (ECP) (SBKJC: Stevens–Basch–Krauss–Jasien–Cundari) [31] was used for the carbons and the oxygen. ECPs replace the core electrons with an effective potential, thus eliminating the need for the core basis functions, which usually require a large set of Gaussians to describe them. Normal mode analysis was then performed using the same level of theory. The vibrational frequencies resulting from the quantum calculations were then rescaled by a factor of 0.9614 to compensate for the use of the harmonic approximation to the potential energy surface [32]. For the calculation of the atomic partial charges, q_i , the CHELPG method [33] with the standard 6-31G* basis set within NWChem was used. This method employs a least-squares fitting procedure to determine the set of atomic partial charges that best reproduces the quantum mechanical electrostatic potential at selected grid points. The fitting is subject to the constraint that the sum of the charges should be equal to the net charge on the molecule. To ensure that the charges on symmetrically equivalent atoms are equal, additional constraints on the partial atomic charges were imposed during the fitting procedure. To do this, the molecule was grouped into subsets of atoms, which were constrained to have zero total charge. For instance, the methyl groups were restrained to zero charge with, in addition, all the hydrogens carrying identical charges.

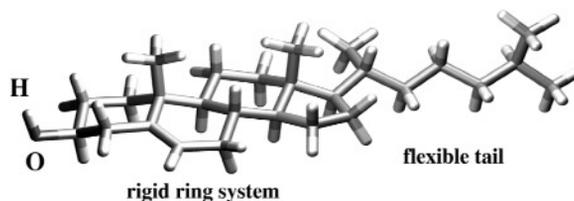


Fig. 1 Minimized structure of cholesterol (DFT SBJKC optimization).

For rhodamine 6G (R6G) (Fig. 2), the quantum chemical computations were performed with the GAUSSIAN-94 package using the standard 6-31G* basis set [34]. The restricted Hartree–Fock (RHF) level of theory was used for geometry optimizations and normal mode calculations. Frequencies resulting from the quantum calculations were rescaled by a factor 0.8929 to compensate for the neglect of electron correlation at the Hartree–Fock (HF) level and the harmonic approximation of the potential energy surface [35]. To derive the atomic partial charges, q_i , again the CHELPG method within

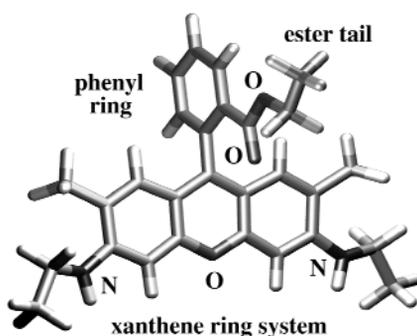


Fig. 2 Minimized structure of R6G (RHF 6-31G* optimization).

GAUSSIAN was used to fit the point-charge electrostatic potential to that obtained quantum mechanically [33].

Parameter refinement

All molecular mechanics calculations were performed using the CHARMM 27 package [28]. Except for the new parameters that are derived here, the existing CHARMM parameters were used [36]. The molecular mechanics minimizations performed were carried out using the steepest-descent algorithm for initial minimization and followed by Newton–Raphson minimization with a convergence criterion for the energy gradient of 10^{-6} kcal/mol/Å. No cutoff was applied to electrostatic interactions.

In CHARMM, the empirical potential energy function is given by eq. 1 [28]:

$$\begin{aligned}
 V(R) = & \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{ub} K_{ub}(s - s_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\chi(1 + \cos(n\chi - \chi_0)) + \\
 & + \sum_{\text{impropers}} K_\phi(\phi - \phi_0)^2 + \sum_{\text{nonbond}} \epsilon_{ij} \left[\left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^{12} + \left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{D r_{ij}}
 \end{aligned} \quad (1)$$

where K_b , K_{ub} , K_θ , K_χ , K_ϕ are the bond, Urey–Bradley, angle, dihedral and improper dihedral force constants, and b , s , θ , χ , and ϕ represent the bond length, Urey–Bradley 1–3 distance, bond angle, dihedral angle and improper torsion angle. The subscript zero, where present, is used to represent the corresponding equilibrium value. Nonbonded interaction between pairs of atoms (labeled i and j) at a relative distance r_{ij} are described by the Lennard–Jones 6–12 (LJ) and Coulomb interaction terms; R_{ij}^{\min} and ϵ_{ij} are, respectively, the distance between atoms i and j at which the LJ potential is minimum and the depth of the LJ potential well for the same pair of atoms. D is the effective dielectric constant and q_i is the partial atomic charge on atom i .

Before refinement, an initial set of parameters was determined. The Van der Waals constants ϵ_{ij} and R_{ij}^{\min} depend mostly on atomic properties and are relatively insensitive to changes in the molecular environment. These were directly transferred from original CHARMM values and were not modified during refinement. Equilibrium values for bonds b_0 , angles θ_0 , and dihedrals χ_0 that were not existing in the original CHARMM force-field parameter file [36] were derived from the calculated quantum chemical structure. An initial guess, based on analogy to similar existing CHARMM parameters and on chemical intuition, was made for all other missing parameters. The second term in eq. 1 (the so-called Urey–Bradley term [36]) is not present in most other force fields and within the CHARMM model its use is limited to a few special cases. Here K_{ub} was set to zero wherever possible.

The initial parameter set was used for minimization and calculation of normal modes (eigenvalues and eigenvectors) with CHARMM. The normal modes obtained were then directly compared with

the normal modes calculated with the quantum chemistry methods, which are considered to be the reference, using the automated frequency-matching method [29]. Parameters were thus refined iteratively to fit the results of the quantum chemical normal mode calculations.

One problem of parametrization methods that fit to vibrational frequencies is identifying a calculated mode with the corresponding reference mode [37]. The fitting method used here optimizes frequency matching by using a penalty function that takes into account both frequencies and all the corresponding eigenvectors. In the ideal case of a perfect molecular mechanics model, the vibrational properties of the molecule, as calculated by molecular mechanics, should perfectly match those resulting from the quantum ab initio calculation. For perfect matching, not only must the frequencies coincide, but also each eigenvector, as calculated using the molecular mechanics (MM) force field, should be orthonormal to all but one (its corresponding eigenvector) of the eigenvectors calculated using quantum chemical methods.

We define the eigenvectors calculated from the MM force field as $\overline{\chi}_C^i$ and the eigenvectors calculated from quantum chemistry as $\overline{\chi}_Q^j$, where the superscripts i and j represent the normal mode number. An efficient way to check simultaneously for both orthonormality and frequency matching is to project each of the CHARMM eigenvectors $\overline{\chi}_C^i$ onto the reference set of eigenvectors $\overline{\chi}_Q^j$, and to find the frequency v_j^{\max} corresponding to the highest projection ($j: \overline{\chi}_C^i \cdot \overline{\chi}_Q^j = \max$). Plotting this frequency against the corresponding reference frequency, v_i , in the ideal case mentioned above would give a one-to-one relationship: $v_i = v_j^{\max}$ and $\overline{\chi}_C^i \cdot \overline{\chi}_Q^j = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Points that deviate from the ideal plot may indicate exchanged or mismatched frequencies. Automated refinement methods are mostly based on minimizing a penalty function, usually a weighted sum-of-square deviations from a set of reference values [38]. The optimization method used here is based on minimization of the weighted sum-of-squares, Y^2 of the deviations from the ideal situation, as follows:

$$Y^2 = \sum_{3N-6} \omega_i^2 (v_i - v_j^{\max})^2$$

$$\omega_i = \frac{1}{\max_j (\overline{\chi}_C^i \cdot \overline{\chi}_Q^j)}$$
(2)

where N is the number of atoms in the molecule and there are $3N-6$ independent vibrational frequencies. The weights ω_i are chosen to be the inverse of the highest eigenvector projection. This choice has the effect of biasing the penalty function, even in the case of a good frequency assignment, such that minimization of Y^2 leads to an improved eigenvector projection distribution. Refinement of parameter sets involves exploring a high-dimensional space in search of an optimal set. Consequently, as for any multidimensional search method, in parameter optimization there is always a risk of the search becoming trapped in a high local minimum. To reduce this risk, it is necessary to generate a physically reasonable set of initial parameters [37].

For the automatic optimization of the force constants, a standard Monte Carlo (MC) scheme [39] was used to minimize Y^2 . Optimizations were performed separately on bond, angle, dihedral, and improper torsion constants. At each step i , all parameters were iteratively varied in the MC algorithm with a uniform distribution within a fixed range, Y_i^2 was evaluated, and, if $Y_i^2 < Y_{i-1}^2$, the new parameter set was used in the next step, $i + 1$. The fixed range was held within the limits of ± 100 kcal/mol/Å², ± 30 kcal/mol/rad², ± 1 kcal/mol, ± 10 kcal/mol/rad² for the bond, angle, dihedral, and improper force constants, respectively. To check for convergence of the penalty function, the MC optimizations were allowed to run for at least 50 more steps after the final value Y_i^2 was obtained, in which, in order to be accepted, the value of Y_i^2 had to remain constant.

When comparing results for different molecules, normalization of Y^2 can be rather tedious due to the different weights ω_i . For comparison purposes, then, after minimization of Y^2 the root-mean-square deviation, σ , from the reference case was calculated:

$$\sigma = \sqrt{\frac{\sum (v_i - v_j^{\max})^2}{3N-6}} \quad (3)$$

RESULTS

Parametrization of rhodamine R6G

The RHF 6-31G* ground-state structure for R6G is shown in Fig. 2. In CHARMM, atom type CA describes sp^2 carbons of aromatic rings. To ensure portability, this atom type could also be used for the case of the aromatic carbons in R6G. However, the use of CA-type carbons for the connection of two aromatic systems in a biphenyl-type configuration results in a structure that is highly biased toward a planar arrangement of the two rings. It was thus necessary to introduce a new atom type (CA1) for the connection between the phenyl ring and the xanthene moiety in R6G. The only changes to CA1 with respect to the standard CA type are the bond and torsion parameters (K_b , b_0 , K_χ , χ_0 , and n in eq. 1). The OS and NC2 types in CHARMM were used to describe the bridging sp^2 oxygen O6 and the nitrogen atoms in the partially charged amino groups in R6G. It was necessary to determine some new parameters for these atom types to account for the new bonding partners of these atoms in R6G. The resulting v_j^{\max} vs. v_i plot for R6G is shown in Fig. 3. The corresponding value of σ (51.5 cm^{-1}) is within the range of the previous benchmark studies [29]. The complete list of the parameters is available in ref. [29].

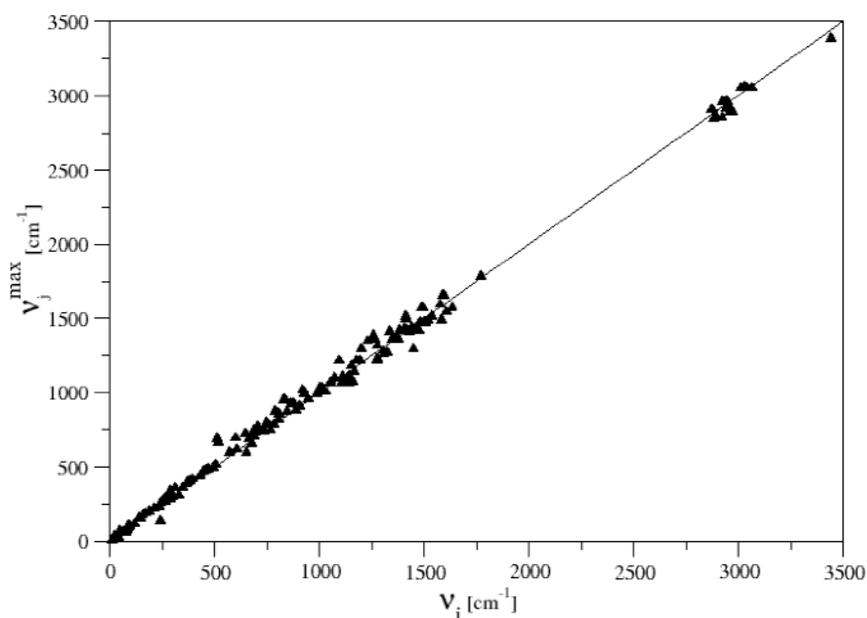


Fig. 3 Frequency-matching plot for R6G. Values of v_j^{\max} correspond to the maximum projection of the “classical” modes on the “quantum” reference modes v_i . The line is the ideal case where $v_j^{\max} = v_i$. Points refer to optimized parameters, $\sigma = 51.50 \text{ cm}^{-1}$.

Parametrization of cholesterol

The structure of cholesterol obtained from optimization with DFT (B3LYP) using the SBJKC basis set is shown in Fig. 1. It was not found necessary to define new atom types for CHARMM. The existing lipid atom types were used. For the sp^2 lipid atoms in cholesterol, atom type CEL1 was used. For the sp^3 atoms we used atom types CTL1, CTL2, CTL3 for saturated carbons with one/none, two, or three hydrogens respectively. The v_j^{\max} vs. v_i plot for cholesterol after parameter refinement is shown in Fig. 4. The corresponding value of σ (40.58 cm^{-1}) is within the range seen in previous benchmark studies [29]. Due to space limitations, the parameters cannot be listed here, but are available from the authors upon request.

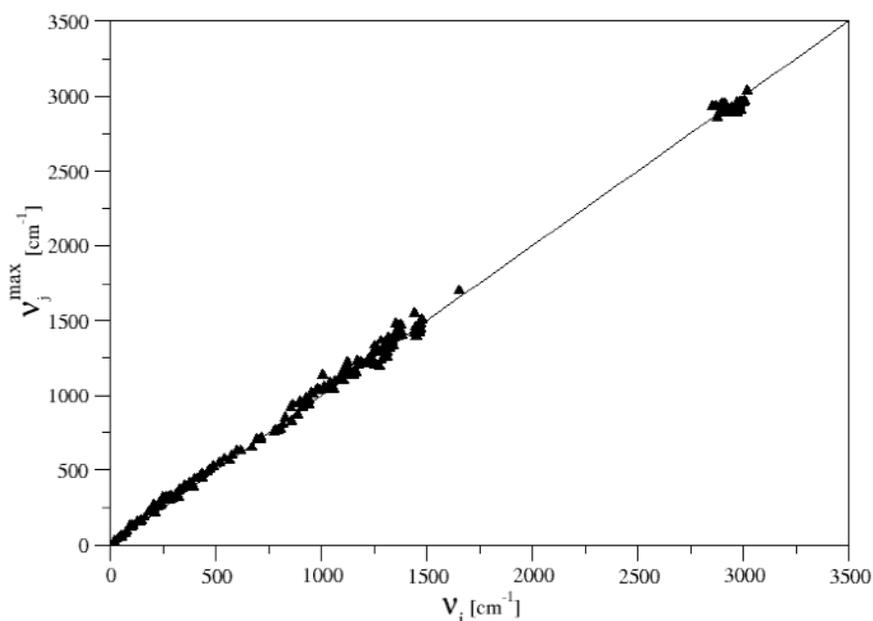


Fig. 4 Frequency-matching plot for cholesterol. Values of v_j^{\max} correspond to the maximum projection of the “classical” modes on the “quantum” reference modes v_i . The line is the ideal case where $v_j^{\max} = v_i$. Points refer to optimized parameters, $\sigma = 40.58 \text{ cm}^{-1}$.

Test of conformational stability: 800 K MD simulation of cholesterol in vacuo

The conformation of cholesterol is essential for its in vivo function. Therefore, it is important that cholesterol preserves its stereochemistry during MD simulation [40]. In previous MD simulations [41], it was found that cholesterol inverted its stereochemistry, which is a simulation artifact. Therefore, to test this aspect of the new parameter set we performed a 2 ns MD simulation of a single cholesterol molecule in vacuo at 800 K. The tetracyclic ring system was found to be rigid and not to undergo conformational changes. The iso-octyl group (hydrocarbon tail) does undergo several conformational changes, as expected. During the simulation, the stereochemistry of all the seven asymmetrical centers was preserved.

CONCLUSIONS

In this paper, we present a parameter set for cholesterol for the CHARMM 27 force field. The method used here for force-field determination is particularly useful for deriving parameters for rigid molecules, for which the flexibility is determined principally by vibrations, as is the case for the cholesterol and R6G ring systems. Due to the computationally inexpensive nature of the penalty function proposed here, the method is also well suited for automation. In particular, the low-frequency modes are very well reproduced. These are of particular biological importance since they determine functional properties.

Deriving force-field parameters for cholesterol is an essential step toward reliable and realistic biomembrane simulations. Use of MD simulation should help gain insights into the dynamical effects of sterols in membranes and to derive biologically relevant structure–function relationships from a dynamical standpoint. Work is also in progress to derive new parameters for two other biologically important sterols, lanosterol and ergosterol.

Recent experimental studies of biological processes in membranes of living cells rely on single-molecule fluorescence measurements [42,43] in which fluorescently labeled lipid molecules are detected and traced. Although experimental methods have evolved very quickly, the important question of how these labels affect the dynamical properties of a membrane still remains open. Reliable force-field parametrization of fluorescent dyes such as R6G together with that of sterols will allow for detailed (atomic-level) simulations aimed at addressing these questions.

REFERENCES

1. K. Bloch. In *Biochemistry of Lipids and Membranes*, J. E. Vance and D. E. Vance (Eds.), pp.1–34, Benjamin/Cummings, Menlo Park (1985).
2. J. L. Thewalt and M. Bloom. *Biophys. J.* **63**, 1176 (1992).
3. L. Finegold (Ed.). *Cholesterol in Model Membranes*, CRC Press, Boca Raton (1993).
4. T. P. W. McMullen, R. A. N. Lewis, R. N. McElhaney. *Biophys. J.* **66**, 741 (1994).
5. T. P. W. McMullen and R. N. McElhaney. *Biochim. Biophys. Acta* **1234**, 90 (1995).
6. M. R. Vist and J. H. Davis. *Biochemistry* **29**, 451 (1990).
7. P. Slotte. *Biochim. Biophys. Acta* **1238**, 118 (1995).
8. P. L. Yeagle. *The Membranes of Cells*, Academic Press, San Diego (1993).
9. F. T. Presti. In *Membrane Fluidity in Biology*, Vol. 4, R. C. Aloia and J. M. Boggs (Eds.), pp. 97–146, Cellular Aspects, Academic Press, New York (1985).
10. R. Bittman. In *Subcellular Biochemistry “Cholesterol: Its Functions and Metabolism in Biology and Medicine”*, Vol. 28, R. Bittman (Ed.), Plenum Press, New York (1997).
11. T. H. Haines. *Prog. Lipid Res.* **40**, 299 (2001).
12. J. P. Incardona and S. Eaton. *Curr. Opin. Cell Biol.* **12**, 193 (2001).
13. L. R. Forrest and M. S. P. Sansom. *Curr. Opin. Struct. Biol.* **10**, 174 (2000).
14. D. J. Tobias. *Curr. Opin. Struct. Biol.* **11**, 253 (2001).
15. K. M. Merz. *Curr. Opin. Struct. Biol.* **7**, 511 (1997).
16. R. W. Pastor. *Curr. Opin. Struct. Biol.* **4**, 486 (1994).
17. D. J. Tobias, K. Tu, M. L. Klein. *Curr. Opin. Struct. Biol.* **2**, 15 (1997).
18. S. W. Chiu, M. M. Clark, E. Jakobsson, S. Subramaniam, H. L. Scott. *J. Comp. Chem.* **20**, 1153 (1999).
19. H. L. Scott. *Curr. Opin. Struct. Biol.* **12**, 495 (2002).
20. H. L. Scott. *Biophys. J.* **59**, 445 (1991).
21. O. Edholm and A. Nyberg. *Biophys. J.* **63**, 1081 (1992).
22. A. J. Robinson, W. G. Richards, P. J. Thomas, M. M. Hahn. *Biophys. J.* **68**, 164 (1995).
23. K. Tu, M. L. Klein, D. J. Tobias. *Biophys. J.* **75**, 2147 (1998).
24. A. M. Smordyrev and M. Berkowitz. *Biophys. J.* **77**, 2075 (1999).

25. M. Pasenkiewicz-Gierula, T. Rog, K. Kitamura, A. Kusumi. *Biophys. J.* **78**, 1376 (2000).
26. S. W. Chiu, E. Jakobsson, H. L. Scott. *Biophys. J.* **80**, 1104 (2001).
27. S. W. Chiu, E. Jakobsson, R. J. Mashl, H. L. Scott. *Biophys. J.* **83**, 1842 (2002).
28. B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, M. Karplus. *J. Comput. Chem.* **4**, 187 (1983).
29. A. C. Vaiana, A. Schulz, J. Wolfrum, M. Sauer, J. C. Smith. *J. Comput. Chem.* **24**, 632 (2003).
30. High Performance Computational Chemistry Group, *NWChem, A Computational Chemistry Package for Parallel Computers*, Version 4.0.1, Pacific Northwest National Laboratory, Richland, Washington, USA (2001).
31. W. J. Stevens, H. Basch, M. Krauss. *J. Chem. Phys.* **81**, 6026 (1984).
32. National Institute of Standards and Technology, USA, <<http://srdata.nist.gov/cccbdb>>.
33. C. N. Breneman and K. B. Wiberg. *J. Comput. Chem.* **11**, 361 (1990).
34. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople. *Gaussian 94, Revision D.4*, Gaussian, Inc., Pittsburgh (1995).
35. J. B. Foresman and M. J. Frisch. In *Exploring Chemistry with Electronic Structure Methods*, Pittsburgh (1993).
36. S. E. Feller, G. Gawrisch, A. D. MacKerell, Jr. *J. Am. Chem. Soc.* **124** (2), 318 (2002). S. E. Feller and A. D. MacKerell, Jr. *J. Phys. Chem. B* **104**, 7510 (2000); M. Schlenkrich, J. Brickmann, A. D. MacKerell, Jr., M. Karplus. In *Biological Membranes: A Molecular Perspective from Computation and Experiment*, K. M. Merz and B. Roux (Eds.), p. 31, Birkhauser, Boston (1996); D. Yin and A. D. MacKerell, Jr. *J. Comput. Chem.* **19**, 334 (1998); S. E. Feller, D. Yin, R. W. Pastor, A. D. MacKerell, Jr. *Biophys. J.* **73**, 2269 (1997).
37. J. Wang and P. A. Kollman. *J. Comput. Chem.* **22**, 1219 (2001).
38. P. O. Norrby and T. Liljefors. *J. Comput. Chem.* **19**, 1146 (1998).
39. F. Haeffner, T. Brinck, M. Haeberlein, C. Moberg. *Theochem.* **397**, 39 (1997).
40. C. M. Crowder, E. J. Westover, A. S. Kumar, R. E. Ostlund, Jr., D. F. Covey. *J. Biol. Chem.* **276**, 44369 (2001).
41. T. Rog and M. Pasenkiewicz-Gierula. *Biophys. J.* **81**, 2190 (2001).
42. G. J. Schutz, G. Kada, V. P. Pastushenko, H. Schindler. *EMBO J.* **19**, 892 (2000).
43. D. P. Dittrich, F. Malvezzi-Campeggi, M. Jahnz, P. Schwille. *Biol. Chem.* **382**, 491 (2001).