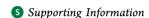
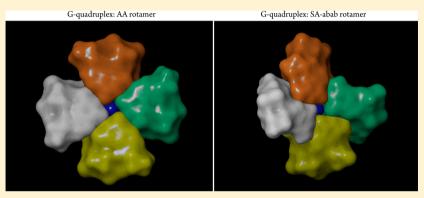


The Right Answer for the Right Electrostatics: Force Field Methods Are Able to Describe Relative Energies of DNA Guanine **Quadruplexes**

Jörg Grunenberg,*,† Giampaolo Barone,‡ and Angelo Spinello‡

[‡]Dipartimento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche, Università di Palermo, Palermo, Italy





ABSTRACT: Different force fields and approximate density functional theory were applied in order to study the rotamer space of the telomeric G-quadruplex DNA. While some force fields show an erratic behavior when it comes to the reproduction of the higher-order DNA conformer space, OPLS and MMFF implementations are able to reproduce the experimentally known energy order. The stabilizing effect of the AA (anti-anti) versus SA (syn-anti) conformer is analyzed applying mechanical bond strength descriptors (compliance constants). The fact that we observe the correct energy order using appropriate force fields is in contrast with results previously reported, which suggested the general inappropriateness of force fields for the description of Gquadruplex structures.

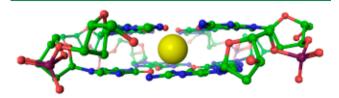
E ukaryotic chromosomes are terminated by hundreds or thousands of noncoding nucleotide repetitions (e.g., TTAGGG in humans or TTGGGG in ciliates, respectively) which play an important role in the protection of chromosomes from recombination and degradation.¹ In somatic cells, this telomeric DNA has a length of ~15 kb,2 displays a 3' termination (150-200 nucleotides), and is comprised by a single strand, which is progressively shortened during cellular replication until a critical length is reached. Normal cells thereupon enter a senescence state and stop replication: the programmed apoptosis. In tumor cells, however, the length of telomeres is kept constant due to the action of the telomerase enzyme. Consequently, the telomerase is overexpressed in 85-90% of human tumor cells,³ rescuing malignant cells from destruction by extending the protective caps on the ends of chromosomes. One of the most promising strategies for cancer therapy is therefore the inhibition of telomerase by small molecules, which are able to bind and stabilize the quadruplex structure, thus preventing the telomerase from binding to the chromosome.

In a variety of organisms, the telomeric repeats have been shown to form what is called G-quadruplex structures: certain G-rich oligonucleotide sequences⁴ that are formed by planar guanine quartets held together by Hoogsteen hydrogen bonds and $\pi - \pi$ stacking interactions. Metal ions further stabilize this quadruplex by a coordination of the O6 guanine atoms. Gquadruplexes form most readily in the single stranded 3' end of telomeric DNA;⁵ nevertheless, sequences able to fold into such a stable quadruplex conformation were found in many regions of the human genome, for example in oncogenes. In this case, the G-quadruplex has to be extruded from double stranded DNA, a process which is stabilized by small-molecule ligands such as the porphyrin TMPyP.6 Very recently, RNA guanine quadruplex structures were observed also in vivo in the cytoplasm of living human cells.⁷ Concerning the topology of G-DNA arrangements, the base of each nucleotide is attached via a glycosidic bond from the N9 nitrogen of the purine to the C1'-carbon of the deoxyribose sugar. The rotation about the glycosidic bond, the x-angle, defines two general conformational classes: the anti conformation with the base extended away from the sugar, and the syn conformation with the base essentially lying on top of the sugar ring (Figure 1). There is a

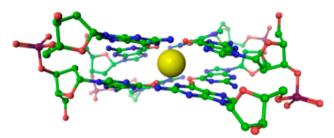
Received: April 16, 2014 Published: June 26, 2014

[†]Institute of Organic Chemistry, TU Braunschweig, Braunschweig, Germany

bulk of experimental evidence that an "anti-anti" stem orientation (AA) of the nucleobases is preferred against an "syn-anti" (SA) arrangement.⁴



AA conformer



SA-abab conformer

Figure 1. Two possible quartet G-DNA conformers: the 5'-anti-anti-3' (AA; above) and the 5'-syn-anti-3' (SA-abab; below) conformation, respectively. The channel cation K^+ is displayed as a yellow sphere.

According to the importance of this structure motive as a hot target in drug design, in past years many in silico studies appeared in the literature dealing with the structural and energetic properties of G-quadruplexes using both empirical force fields⁸ and approximate density functional theory.^{9,10} While molecular dynamic simulations of proteins and nucleic acids became standard in past decades and the development of sophisticated simulation methods was honored with the 2013 Chemistry Nobel Prize, a recent publication by Sponer and Grimme¹¹ nevertheless cast clouds concerning the reliability of MD simulations on noncanonical DNA structures. Their conclusion, based on the observed reversed stem stability comparing force field and DFT results, goes well beyond the known weak points of force fields (e.g., overstabilization of base stacking) stating that empirical force fields are not able to describe the G-quadruplex energy surface as a matter of principle, because of their pairwise-additive philosophy. If true, this verdict would have far reaching consequences. Due to Sponer's and Grimme's force field calculations, the SA-abab stem was described as the lowest minimum (-14 kcal/mol in favor of the AA stem) while approximate density functional theory puts the AA stem 5 kcal/mol below the SA-abab stem, in line with experimental evidence. Nevertheless, the robustness of this conclusion was not tested against different force field parameters or implementations. 12 In the present letter, we therefore compared several force fields (AMBER, OPLS-AA, MMFF) with DFT results in order to review the general capabilities of empirical force fields in describing the conformational space of G-quadruplex structures.

DFT calculations were performed with the Gaussian 09 suite of programs.¹⁸ The structures were fully optimized using the long-range corrected (LC) hybrid density functional wb97xd, which employs 100% Hartree–Fock (HF) exchange for long-range electron–electron interactions, while dispersion corrections are included via empirical atom–atom terms.¹³ A polarized medium sized all-electron Gaussian atomic orbitals

(AO) basis set (6-311G(d,p)) was used in the optimizations and compliance constants calculations. Continuum solvation effects were represented by the Polarizable Continuum Model (PCM) assuming the bulk dielectric constant of 78.35 for water. The force field calculations were carried applying the Macromodel 9.9 software (AMBER94, MMFF, and OPLS-2005) and with the Gaussian 09 set of programs (AMBER-bsc0). Since the description of biomolecular systems in the condensed phase relies first and foremost on an authoritative reproduction of the conformer space *in vacuo*, we, in a first step, conducted our simulations in the gas phase. Solvent effects have been included in a second step. A generalized-Born (GB) algorithm was applied.

Because we are dealing with different rotamers, it is possible to compare the different topologies by means of their computed steric force field energy directly. The starting conformers of our study—each containing tetrads of hydrogen-bonded guanine bases and different glycosidic torsions defining the nucleobase orientation relative to the sugar—were taken from the literature. All our computations started with the quantum chemically (gas phase) optimized geometries, noted "QMopt" in the original paper. Our own molecular geometries optimizations were conducted using the particular series of force fields mentioned below. Figure 2 demonstrates

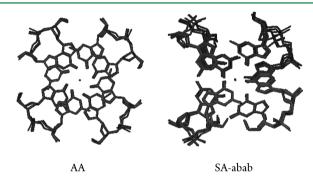


Figure 2. Superposition of the DFT, Hartree–Fock, OPLS, AMBER, and MMFF minimized AA (top) and SA-abab (bottom) rotamer of the model G-quadruplex.

that we are indeed comparing the same particular low energy conformer. The superposition of our optimized DFT and force field geometries shows only small dihedral and angle deviations. Comparing the energetics, we are therefore talking about the same energy wells, a prerequisite of any meaningful discussion in the following.

Table 1 compiles the energetic results of our optimizations: (1) In line with the results of Sponer and Grimme, the AMBER

Table 1. Calculated Relative Energies (kJ/mol) of the AA and the SA-abab Stems a

	vacuum	implicit solvent
WB97XD	+123.4	+78.7
Hartree-Fock	+117.3	+68.3
OPLS-2005 ^b	+47.4	+24.1
$MMFF^b$	+89.1	+10.9
AMBER_94 ^b	-72.3	-69.6
AMBER_bsc0 ^c	-70.9	-65.9
AMBER_bsc0 (Qeq) ^c	+127.1	+106.1

 $[^]a$ A positive sign is connected with a destabilized SA-abab rotamer. b Macromodel v99109. c Gaussian 09 A.1.

force field (both the AMBER94 and the AMBER-bsc0 implementation) sees the all-anti conformer AA higher in energy in comparison with the SA-abab (syn-anti) conformer. (2) Both the OPLS-2005 and the MMFF force field reproduce the experimental (and DFT) trend: a higher stability for the AA stem in comparison with the SA-abab stem. The overall picture is the same after the inclusion of implicit solvent effects. We would like to note that all three force field implementations (AMBER, OPLS-2005, and MMFF) are based on pairwise additive interactions. Nevertheless, different methodologies were used during their fitting process, respectively. The OPLS force field, for example, was trained using experimental properties of liquids, such as density and heat of vaporization, while the MMFF force field parameters were fitted against high level ab initio gas phase data. In the bottom line, the molecular training set for all three evaluated force fields (OPLS; AMBER and MMFF) did not include any of the noncanonical DNA structures described in this study.

One obvious reason for the wrong description of the conformer space could be the bogus account of dispersion effects. Nevertheless, both simple Hartree-Fock/6-31G(d) theory (without dispersion) and DFT (including dispersion, wb97xd) reproduce the correct experimental trend. Dispersion is obviously not the reason for the energy difference in our model systems. In fact, while dispersion is essential to reproduce the π - π stacking interactions in the canonical (double-helical) DNA conformation, the presence of the K⁺ channel cation seems to change the situation. A second reason for the erratic behavior of the AMBER force field could be the deficient description of the relevant hydrogen bonds. We therefore, in a second step, analyzed the relaxed force constants (inverse compliance constants) of the most important noncovalent interactions. The description in terms of compliance constants allows a direct comparison of individual covalent and noncovalent interactions for polyvalent supramolecular complexes. This is also true for soft interactions, albeit there still are irritations about this topic in the literature. ^{20,21} In past years, the accumulation of studies, which prove the utility of a deliberate application of compliance constants for strong and soft interactions, is nevertheless significant. 22-40

In the following, we are presenting the inverse of the computed compliance constants (relaxed force constants in N/cm) in order to facilitate the comparison with older literature values. Each G-DNA stem contains two quartets comprising four Hoogsteen hydrogen bonds of the NH···N and the NH···O types, as well as four dipole—ion interactions of the guanine···K⁺ type, respectively. Our own computed wb97xd relaxed force constants as a measure of the elasticity for the AA and the SA-abab stem are listed in Table 2. The analysis of the individual Hoogsteen hydrogen bonds confirms the recently

Table 2. Comparison of Noncovalent Interaction in the AA and the SA-abab Conformers of the Quartet G-DNA Computed at the DFT (WB97XD) and Force Field (AMBER) Levels of Theory in N/cm

vacuo	$NH\cdots N$	NH···O	K+O	mean
guanine dimer G ₂ (DFT)	0.11	0.24		0.17
AA (DFT)	0.34	0.34	0.19	0.29
SA-abab (DFT)	0.31	0.36	0.19	0.28
AA (AMBER)	0.41	0.29	0.19	0.29
SA-abab (AMBER)	0.30	0.31	0.20	0.27

reported H-bond cooperativity in G-tetrads. While for an isolated guanine dimer (G_2) the calculated relaxed DFT force constants of the hydrogen bonds point to quite soft interactions (NH···O, 0.24 N/cm; NH···N, 0.11 N/cm), the G-tetrads are characterized by strong NH···O (0.34 N/cm) and NH···N hydrogen bonds (0.31–0.36 N/cm). Comparing nevertheless the average (0.29 versus 0.28 N/cm) of all inter-residue hydrogen bonds as well as the eight contacts between the guanine oxygen and the channel cation K⁺—kept in place by eight soft interactions (0.19–0.20 N/cm)—there seems to be no difference between the AA and SA stems (in this case, the SA-abab rotamer) at all. This is true for both, the DFT and the AMBER description.

If both dispersion and hydrogen bonding are not responsible, what, in the end, could be the reason for the erratic behavior of the AMBER force field when it comes to the description of the conformer space? In order to describe the total energy, an accurate representation of the charge distribution throughout the whole molecule is essential for all atomic resolution modeling techniques and especially for the highly charged RNA/DNA systems. Since the guanine quadruplexes are noncanonical DNA structures, which were not included in the original force field training set, the standard assignment of the atomic charges might therefore be the source of small individual electrostatic errors, even if the atoms in question are not part of the hydrogen bond network. The accumulation of these small errors for the nearly 20 000 subtle, mutual nonbonded interactions might nevertheless lead to an erratic description of the total steric energy in force fields. A straightforward strategy to verify or falsify this hypothesis is therefore the calculation of adapted charges, which depend upon the environment of the starting conformation. Table 1 therefore also includes the result of a modified AMBER force field using quantum adapted charges applying Reppe's and Goddard's charge equilibrium (QEq) formalism. 42 The resulting adapted charges are based on (1) experimental atomic ionization potentials, (2) electron affinities, and (3) atomic radii. Using these experimental parameters, an atomic chemical potential is constructed leading to charges, which are in excellent agreement with experimental dipole moments. The AMBER-Qeq force field indeed reproduced the experimental (and DFT) trend: the AA rotamer is now calculated to be the most stable conformer using the AMBER force field. In comparison with the SA-abab stem, the stabilization is quite pronounced (127.1 kJ/mol) in a vacuum and shielded a little bit (106.1 kJ/mol) applying implicit solvent (see Table 1). Nevertheless, the AMBER-Qeq trend is in line with the experimental results: a higher stability for the AA-stem. A comparison of the relevant atomic charge before and after the Qeq formalism was applied reveals that the following atom types are involved, in decreasing order: the phosphorus atom, whose charge decreases by 0.6 units, and the N*, CA, and O2 atom types, whose charge values change by 0.4 units (see Table 3). Nevertheless, we refrain from any overinterpretation of the atomic charges, the moreso as any partial adaption of the force field parameters without adaption of all other nonbonded interaction or torsional parameters impedes their interpretation as physical observables. 43

In conclusion, our letter highlights again the prerequisite of any reliable force field simulation: an evaluation of the robustness against different force field parameters or implementations. Provided that the empirical potential function is chosen properly and flexible enough to adapt to the very

Table 3. Relevant AMBER_bsc0 Charges (a.u.) before and after the Qeq Formalism Was Applied^a

atom type	bsc0 ^b	bsc0 Qeq ^b	OPLS ^c
P	1.2	0.6	1.8
O2	-0.8	-0.4	-1.0
N*	-0.1	-0.4	-0.5
NB	-0.6	-0.4	-0.5
CA	0.7	0.3	0.5
N2	-0.9	-0.6	-0.8
NC	-0.7	-0.3	-0.5

 $^a\mathrm{The}$ OPLS charges are given for comparison. $^b\mathrm{Gaussian}$ 09 A.1. $^c\mathrm{Macromodel}$ v99109.

special electrostatics, force fields are indeed able to reproduce—and hence predict—the correct relative energetic ordering of different arrangements of guanine quadruplex (G-DNA) stems. Since the simulation of molecular recognition processes, where DNA guanine quadruplexes function as a receptor for small organic or inorganic molecules, relies especially on an authoritative and fast mapping of the system's conformational space, we believe that a combination of force field methods, controlled by DFT checkpoints will lead to a robust protocol for the development of effective and potent G-DNA binders.

■ ASSOCIATED CONTENT

Supporting Information

Geometries of the AA and SA-abab rotamers used in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: joerg.grunenberg@tu-bs.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Prof. Dr. M. Preller (MHH Hannover) for fruitful discussions.

■ REFERENCES

- (1) Hackett, J. A.; Feldser, D. M.; Greider, C. W. Cell 2001, 106, 275-286.
- (2) Cech, T. R. Angew. Chem., Int. Ed. 2000, 39, 34-43.
- (3) Shay, J. W.; Bacchetti, S. Eur. J. Cancer 1997, 33, 787-791.
- (4) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, 34, 5402–5415.
- (5) Todd, A. K.; Johnston, M.; Neidle, S. Nucleic Acids Res. 2005, 33, 2901–2907.
- (6) Huppert, J. L.; Balasubramanian, S. Nucleic Acids Res. 2007, 35, 406–413.
- (7) (a) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Nat. Chem. 2013, 5, 182–186. (b) Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. Nat. Chem. 2014, 6, 75–80.
- (8) Špačková, N.; Berger, I.; Sponer, J. J. Am. Chem. Soc. 1999, 121, 5519-5534.
- (9) van Mourik, T.; Dingley, A. J. Chem.—Eur. J. 2005, 11, 6064–6079.
- (10) Jissy, A. K.; Ashik, U. P. M.; Datta, A. J. Phys. Chem. C 2011, 115 (25), 12530–12546.
- (11) Šponer, J.; Mládek, A.; Špačková, N.; Cang, X.; Cheatham, T. E., III; Grimme, S. J. Am. Chem. Soc. 2013, 135, 9785–9796.

- (12) As trivial as it may seem, this is a shortcoming of many force field MD simulations. For a discussion, see: Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Biophys. J.* **2011**, *100*, L47–L49.
- (13) Chai, J.-D.; Head-Gordon, M. Phys. Chem. Chem. Phys. 2008, 10, 6615-6620.
- (14) MacroModel, version 9.9; Schrödinger, LLC, New York, 2012.
- (15) Cornell, W. D.; Cieplak, P.; Bayly, Ch. 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*, 5179–5197.
- (16) Halgren, T. A. J. Comput. Chem. 1996, 17, 490-519.
- (17) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. J. Comput. Chem. 2005, 26, 1752–1780.
- (18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, revision A.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (19) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **200**7, *92*, 3817–3829.
- (20) (a) Baker, J.; Pulay, P. J. Am. Chem. Soc. 2006, 128, 11324–11325. (b) For a recent example, presented in a particular disturbing tone, see: Baker, J. J. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2013, 4, 111–115.
- (21) For a recent discussion concerning possible misinterpretations, which are based the use of oversimplified force fields, see: Grunenberg, J.; Barone, G. RSC Adv. 2013, 3, 4757–4762.
- (22) Yurenko, Y. P.; Novotny, J.; Sklenar, V.; Marek, R. Phys. Chem. Chem. Phys. 2014, 16, 2072–2084.
- (23) von Frantzius, G.; Espinosa Ferao, A.; Streubel, R. Chem. Sci. 2013, 4, 4309-4322.
- (24) Markopoulos, G.; Grunenberg, J. Angew. Chem., Int. Ed. 2013, 52, 10648–10651.
- (25) Brovarets', O. h. O.; Hovorun, D. M. Phys. Chem. Chem. Phys. **2013**, 15, 20091–20104.
- (26) Henne, F. D.; Schnöckelborg, E.-M.; Feldmann, K.-O.; Grunenberg, J.; Wolf, R.; Weigand, J. J. Organometallics 2013, 32,
- (27) Brovarets', O. h. O.; Zhurakivsky, R. O.; Hovorun, D. M. J. Mol. Model. **2013**, 19, 4223–4237.
- (28) Shimasaki, T.; Nakayasu, K.; Shibata, M.; Yamaguchi, T. *J. Chem.* **2013**, 2013, 1–9.
- (29) Ponomareva, A. G.; Yurenko, Y. P.; Zhurakivsky, R. O.; Mourik, T. v.; Hovorun, D. M. J. Biomol. Struct. Dyn. **2014**, 32, 730–740.
- (30) Brovarets', O. h. O.; Yurenko, Y. P.; Hovorun, D. M. J. Biomol. Struct. Dyn. **2014**, 32, 993–1022.
- (31) Maekawa, M.; Daniliuc, C. G.; Jones, P. G.; Hohenberger, J.; Sutter, J.; Meyer, K.; Walter, M. D. Eur. J. Inorg. Chem. 2013, 2013, 4097–4104.
- (32) Brovarets', O. h. O.; Hovorun, D. M. J. Comput. Chem. 2013, 34, 2577–2590.
- (33) Danovich, D.; Shaik, S.; Rzepa, H. S.; Hoffmann, R. Angew. Chem., Int. Ed. 2013, 52, 5926-5928.
- (34) Shaik, S.; Danovich, D.; Wu, W.; Su, P.; Rzepa, H. S.; Hiberty, P. C. Nat. Chem. **2012**, *4*, 195–200.

- (35) Ponomareva, A. G.; Yurenko, Y. P.; Zhurakivsky, R. O.; van Mourik, T.; Hovorun, D. M. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6787–6795
- (36) Zou, W.; Kalescky, R.; Kraka, E.; Cremer, D. J. Chem. Phys. **2012**, 137, 084114.
- (37) Espinosa, A.; Streubel, R. Chem.—Eur. J. 2011, 17, 3166–3178.
- (38) Schreiner, P. R.; Reisenauer, H. P.; Romanski, J.; Mloston, G. Angew. Chem., Int. Ed. **2009**, 48, 8133–8136.
- (39) Madhav, M. V.; Manogaran, S. J. Chem. Phys. 2009, 131, 174112.
- (40) Pignedoli, C. A.; Curioni, A.; Andreoni, W. ChemPhysChem 2005, 6, 1795-1799.
- (41) Fonseca Guerra, C.; Zijlstra, H.; Paragi, G.; Bickelhaupt, F. M. Chem.—Eur. J. **2011**, *17*, 12612–12622.
- (42) Rappe, A. K.; Goddard, W. A. J. Phys. Chem. 1991, 95, 3358-3363.
- (43) As one of a multitude of parameters which are fixed either during the original force field parametrization or the charge equilibration process, they have to be considered as variational coefficients, optimized in order to minimize the force field's error function, no more, no less.