

A new AMBER-compatible force field parameter set for alkanes

**Alexei M. Nikitin, Yury V. Milchevskiy &
Alexander P. Lyubartsev**

Journal of Molecular Modeling

Computational Chemistry - Life Science
- Advanced Materials - New Methods

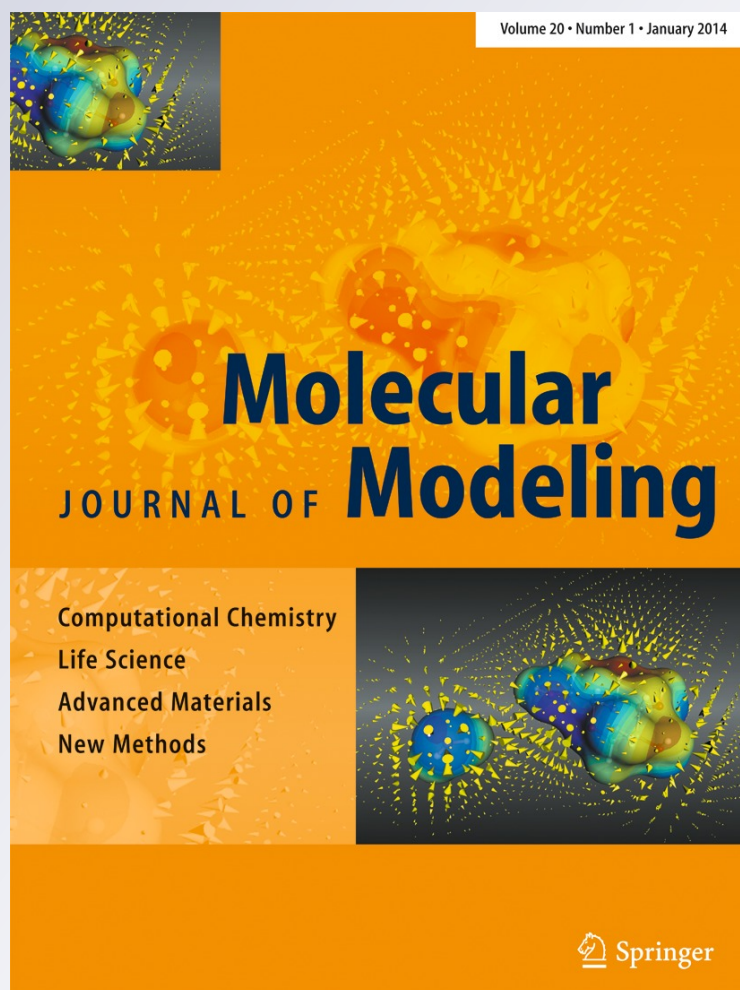
ISSN 1610-2940

Volume 20

Number 3

J Mol Model (2014) 20:1-10

DOI 10.1007/s00894-014-2143-6



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

A new AMBER-compatible force field parameter set for alkanes

Alexei M. Nikitin · Yury V. Milchevskiy ·
Alexander P. Lyubartsev

Received: 8 November 2013 / Accepted: 10 January 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract We present a new force field parameter set for simulating alkanes. Its functional form and parameters are chosen to make it directly compatible with the AMBER94/99/12 family of force fields implemented in the available software. The proposed parameterization enables universal description of both the conformational and thermodynamic properties of linear, branched, and cyclic alkanes. Such unification is achieved by using two essential principles: (1) reduction of the Lennard-Jones radius for all sp^3 carbons to 1.75 Å; (2) separate optimization of Lennard-Jones well depths for carbons with different degree of substitution. The new parameter set may prove to be optimal for description of alkyl residues in a broad range of biomolecules, from amino acids to lipids with their extended linear tails.

Keywords AMBER · Force field · Hydrocarbons · Molecular dynamics · OPLS

Introduction

In biomolecular computations, the most popular force fields have nearly unified and quite simple functional form. Classical examples are AMBER [1, 2], OPLS [3], and CHARMM [4]. Notwithstanding their success, much can (and should) be done to refine and extend these force fields, aiming at unified description of the broadest possible scope of substances. One

of the problem points is simulation of lipids, which have long aliphatic tails, whereas in the initial design of the above force fields, parameterization was done only for small hydrocarbons. Thus in OPLS-AA [5], nice results were reported in representing small hydrocarbons with just two types of C and H atoms. However, this does not seem to hold if the parameterization simultaneously includes long aliphatic chains. More exactly, the C and H parameters can be fitted so that the thermodynamic properties of simulated linear alkanes would coincide with the experimental ones, but this is achieved at the expense of worse description of branched and cyclic alkanes. This nuisance can be circumvented by using special parameters for linear alkanes. This is the way taken by CHARMM [6], which has a set of alkanes parameters designed specially for simulating the lipid hydrocarbon tails. These parameters cannot be directly used with force fields of the AMBER/OPLS type. Recent parameterizations of lipids (sLipids [7, 8], LIPID11 [9]) have been adapted to use with AMBER for membrane-associated molecules and polypeptides. They may be conventionally called «CHARMM-type parameterizations» for several reasons. First, as in CHARMM, they are unique for linear alkanes, as the corresponding C and H parameters differ from those used for amino acids. Second, the intramolecular parameters such as bond lengths and angles are inherited from CHARMM. For example, the term for the CCC angle is greater than tetrahedral; it approximately corresponds to the real angle in linear alkanes ($112^\circ \pm 1^\circ$ from gas electronography [10]; $112.4^\circ \pm 0.2^\circ$ from microwave spectra [11]; 112.14° from MP2/6-311+G(2d,p) calculations); i.e., CCC angles are supposed to not be strained. By contrast, AMBER assumes the tetrahedral value. Inasmuch as in alkanes the CCC angles are somewhat increased because of steric repulsion, in AMBER they always experience certain strain. The modification of AMBER proposed here envisions the use of a unified scheme with *tetrahedral* CCC angle terms.

A. M. Nikitin (✉) · Y. V. Milchevskiy
Engelhardt Institute of Molecular Biology Russian Academy of
Sciences, Moscow 119991, Russia
e-mail: agilemolecule@gmail.com

A. P. Lyubartsev
Division of Physical Chemistry, Arrhenius Laboratory, Stockholm
University, Stockholm 10691, Sweden

In this paper we want to show that it is possible to design a field that, on the one hand, would uniformly, on the AMBER principles, describe aliphatic C and H regardless of whether they are in amino acid side chains or in lipid tails; and on the other hand, would be suitable for describing aliphatic hydrocarbons as such, irrespective of biomolecules, that is in linear, branched, or cyclic hydrocarbons, including strained five- and eight-membered rings (here we do not consider three- or four-membered cycles).

Furthermore, we want the force field to describe simultaneously both the conformational properties of alkanes and their thermodynamic properties in liquid phase. The success of OPLS force field demonstrates the importance of correct thermodynamic description. Correct thermodynamic behavior is important for description of mixtures, inhomogeneous systems where partitioning of different species is determined by their free energies. However, it is commonly believed that good thermodynamic capabilities are difficult to combine with conformational ones. The matter is that in OPLS (and in most analogous fields) the radius of an aliphatic carbon is taken to be about 2 Å. This value is thought to be suitable for thermodynamics but indeed can hardly fit into a conformational description, because excessive steric repulsion arises for gauche-conformations of the chain. We want to show that reduction of the carbon radius, making possible various chain conformations, can be compatible with adequate description of the thermodynamic properties.

On top of that, we want the force field satisfying these requirements to be defined by possible minimum number of varying parameters.

We achieve these goals by doing two main modifications:

- (a) **Reducing the sp^3 carbon radius from 2 to 1.75 Å.** The 2-Å radius is obviously excessive and precludes correct description of some native conformations. A vivid example is the gauche-gauche conformation of linear alkanes. Actually, such a radius is excessive even for gauche-butane. In AMBER, this problem is partially sidestepped by halving the 1–4 interaction, while CHARMM introduces special (reduced) 1–4 parameters. Yet the problem arises again for gauche-gauche-pentane. If we do not wish to introduce special rules for 1–5 interaction as well, we need to reduce the C radius. Taking 1.75 Å, we can obtain gauche-butane and gauche-gauche-pentane in conformations close to that computed within MP2/6-311+G(2d,p) theory.
- (b) **Taking that the well depth ϵ for carbon differs depending on substitution.** At the same time, the radius of sp^3 carbon and the parameters of its hydrogen may be taken as independent of the environment.

Here we depart from the simplest parameterization where by all aliphatic carbons are ascribed the same parameters. Yet

this appears necessary if we wish to simultaneously and accurately reproduce the thermodynamic properties of normal, branched, and cyclic alkanes.

This work proceeds from the unpublished AMBER-i field implemented in the Ascalaph package [12]. AMBER-i is based on non-tetrahedral angles and thus does not fit very well into the AMBER approach. Here we develop a new set of force field parameters trying to follow the same concept as that adopted in AMBER family of force fields.

Methodology

Parameterization

The parameters were optimized iteratively. The first step was fitting the Lennard-Jones parameters to the experimental thermodynamic data (liquids density and heat of vaporization). Then the bonded and angle interaction parameters were fitted to quantum-chemical data. Last, the torsion potentials were optimized by fitting to quantum-chemical calculations. Then the cycle was repeated again to ensure convergency. The quantum-chemical computations were performed mainly (if not otherwise said) at MP2/6-311+G(2d,p) level, which is substantially higher than the 6-31G(d) routinely used in AMBER parameterization.

For the hydrogens of aliphatic compounds, we define one molecular mechanics type HC as in AMBER99 [2]. For carbons, we introduce four types: methylene carbon is denoted CT, also as in AMBER99; methyl carbon is CH3; trisubstituted, CH; and tetrasubstituted, CR4. The final result for the fitting procedure is given in Table 1.

The 1–4 nonbonded Lennard-Jones interactions are reduced by half and the electrostatic ones by factor 1/1.2 as it is accepted in AMBER-type fields.

The initial bond and angle parameters were taken from AMBER99 and modified to better fit the quantum-chemical calculations. See [Ethane](#) section for details. The modified bond parameters are listed in Table 2.

The angle parameters were optimized from ab-initio computations for propane, except for the equilibrium CCC term, which was taken to be tetrahedral. See details in [Ethane](#) and [Propane](#) sections. The optimized angle parameters are given in Table 3.

Table 1 Lennard-Jones Parameters

Type	R*	ϵ
HC	1.495	0.0124
CH3	1.75	0.184
CT	1.75	0.156
CH	1.75	0.10
CR4	1.75	0.035

R* : van der Waals radius (the equilibrium distance), Å. ϵ : van der Waals well depth, kcal mol⁻¹

Table 2 Bond parameters

Type1	Type2	k_{bond}	R(eq)
CH3	HC	340	1.093
CT	HC	340	1.096
CH	HC	340	1.097
[C]	[C]	240	1.526

k_{bond} is force constant in $\text{kcal}(\text{mol} \text{Å}^2)^{-1}$ and R(eq) is the equilibrium distance (in Å). [C]: denotes identical parameters for all carbon species

The torsion angle parameters were optimized for describing *n*-butane and *n*-pentane, see [Gauche-butane](#) and [Gauche-gauche-pentane](#) section. The resulting values for torsion potential parameters are gathered in Table 4.

Further through the text, we will refer to a set of force field parameters derived in this work as AMBER-ii.

Conformational computations

In this section we demonstrate reproducibility of ab-initio computed energies by the developed set of parameters for a number of short alkanes.

Ethane

Figure 1 displays the dependencies of energy on the C–C bond length in ethane. It is evident that AMBER99 works nicely with small deformations at $1\text{--}2 \text{ kcal mol}^{-1}$ but worse with larger ones. The MP2-computed energy–length dependence is clearly not quadratic. If we do not want to use more complicated functions, staying within the AMBER functional form, then we face a choice of describing well either small or large deformations. The classical AMBER takes the former path, while here we take the latter. The reason is that we want to describe a broad range of compounds including cyclic ones which might be under stress, so it is important to more accurately represent severely distorted bonds. In the literature, one can find aliphatic C–C bonds up to 1.72 Å [13]. The force constant in this work was chosen such that at 1.72 Å length the energy would match the one calculated in MP2/6-311+G(2d,p). Thus the force constant in AMBER-ii is only $240.0 \text{ kcal}(\text{mol} \text{Å}^2)^{-1}$, versus 310.0 in AMBER99. We adopt this approach also for other force field parameters: whenever

Table 3 Angle parameters

Type1	Type2	Type3	k_{ang}	Theta(eq)
H	[C]	H	33	107
[C]	[C]	H	52	110.7
[C]	[C]	[C]	30	109.5

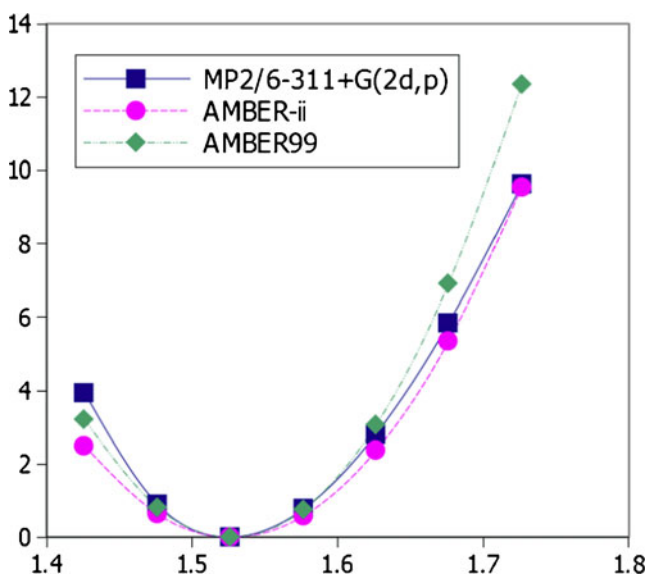
k_{ang} is angle force constant (in $\text{kcal}(\text{mol} \text{rad}^2)^{-1}$). [C]: denotes identical parameters for all carbon species and Theta(eq) is the equilibrium angle

Table 4 Torsional parameters

Type1	Type2	Type3	Type4	pathsNo	V(n)/2	Gamma	n
**	CH3	CH3	**	9	1.45	0	3
**	CH3	CT	**	9	1.45	0	3
**	CH3	CH	**	9	1.45	0	3
**	CH3	CR4	**	9	1.45	0	3
**	CT	CT	**	9	1.45	0	3
**	CT	CH	**	9	1.45	0	3
**	CT	CR4	**	9	1.45	0	3
**	CH	CH	**	9	1.45	0	3
**	CH	CR4	**	9	1.45	0	3
**	CR4	CR4	**	9	1.45	0	3
CH3	CT	CT	CH3	1	0.16	0	3
CT	CT	CT	CH3	1	0.16	0	3
CT	CT	CT	CT	1	0.16	0	3
CH3	CT	CT	CH3	1	0.09	180	2
CT	CT	CT	CH3	1	0.09	180	2
CT	CT	CT	CT	1	0.09	180	2
CH3	CT	CT	CH3	1	0.06	180	1
CT	CT	CT	CH3	1	0.06	180	1
CT	CT	CT	CT	1	0.06	180	1

PathsNo: Number of equivalent bond paths that the total V(n)/2 is divided into; this equals the product of the number of bonds to each of the middle two atoms. V(n)/2: Magnitude of torsion, kcal mol^{-1} . Gamma: Phase offset, deg. n: Torsion periodicity. Implicit definition of parameters ** is used if no explicit type is found for the given torsion.

we face a dilemma of correctly representing energy at equilibrium or at deformed geometry, the choice was always in favor of the latter. For this reason, AMBER-ii is a somewhat softer field. On the other side this means that it provides a worse description of vibration frequencies and consequently infrared spectra. Anyway, correct description of spectroscopic

**Fig. 1** Energy (kcal mol^{-1}) vs. C–C bond length (Å) in ethane

properties requires quantum treatment of vibrational motion, which is out of the scope of classical molecular mechanics.

Figure 2 displays the dependences of energy on the C-C-H angle in ethane. The parameters used in AMBER-ii are seen to nicely approximate the MP2/6-311+G(2d,p) computations.

The dihedral angle H-C-C-H parameters were taken from MP2/6-311+G(2d,p). The barrier for torsional rotation in the latter is 3.052 kcal mol⁻¹, which is nearly equal to that in AMBER-ii, 3.035 kcal mol⁻¹.

Propane

The C-C-C angle parameters were initially fitted to quantum-chemical data for propane. The angle was taken to be tetrahedral to correspond to the one used in AMBER-type fields, and the force constant was optimized to best approximate the conformational calculations with MP2/6-311+G(2d,p). In AMBER99 it is 40 kcal (mol rad²)⁻¹, and this provides very good description of quantum-chemical data. However, we had to reduce it to 30 kcal (mol rad²)⁻¹ in order to better describe the cyclopentane conformations and the gauche conformations of linear alkanes. Figure 3 shows how the energy depends on the C-C-C angle in propane with staggered and with eclipsed hydrogen.

Gauche-butane and gauche-gauche-pentane

Attempts to reproduce the gauche-butane energy have been made since the earliest works on conformational analysis. Regrettably, we still lack the exact experimental value, but the estimates clearly tend to go down. In 1970s this energy was estimated at ~1 kcal mol⁻¹ (e.g., 0.966±0.054 kcal mol⁻¹ [14]) was used in designing the MM2 force field [15]). In 2009,

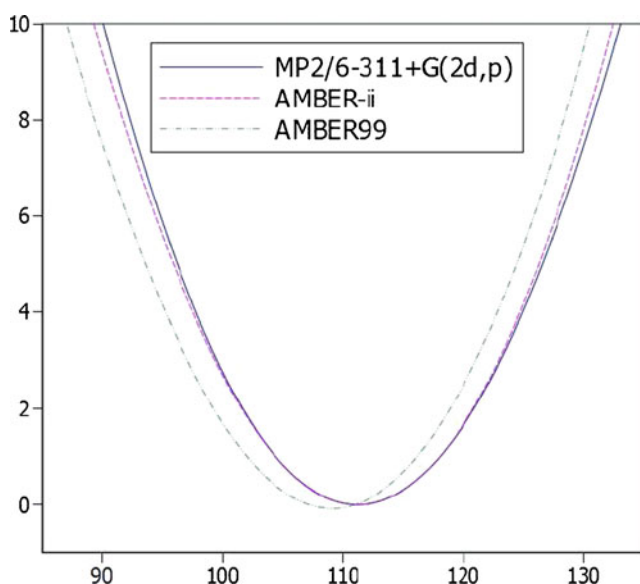


Fig. 2 Energy (kcal mol⁻¹) vs. C-C-H angle (deg) in ethane

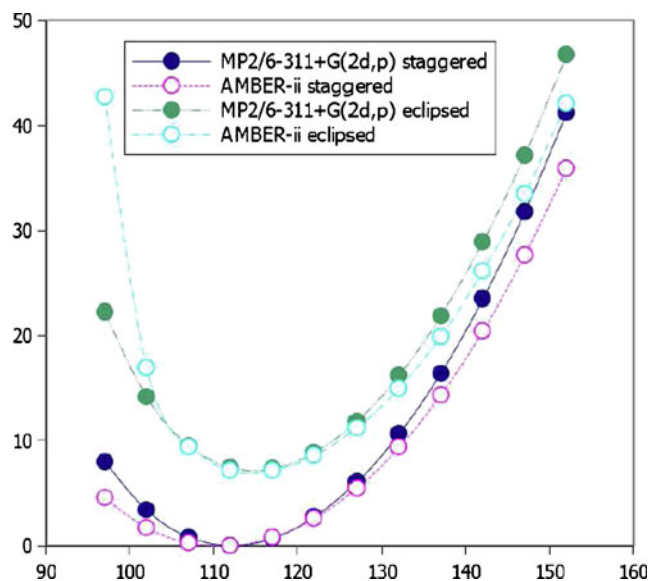


Fig. 3 Energy (kcal mol⁻¹) vs. the C-C-C angle (deg) in propane in staggered and in eclipsed conformations

a Raman spectroscopy study [16] gave 0.660±0.022 kcal mol⁻¹. Our quantum-chemical computations yield a range 0.56–0.62 kcal mol⁻¹.

Apart from its general importance for conformational analysis, a correct torsional energy is crucial for simulating long alkanes including hydrocarbon tails of lipids inasmuch as their properties strongly depend on the trans/gauche ratio. A liquid phase with a shortage of gauche conformers tends to excessive ordering: long alkanes begin to crystallize, and lipid bilayers fall into a gel phase [17]. Correct value of the gauche-energy is thus very important for description of liquid crystalline - gel transition in lipid bilayers.

Another aspect is that gauche conformations are important for the choice of the Lennard-Jones radius for carbon in alkanes. Most of the currently used force fields assume 1.9–2 Å, though appreciably smaller values are found in older fields, e.g., 1.85 Å [18]. However, analysis of gauche-conformations in alkanes clearly shows that the current values are overestimated. Thus, the distance between carbons 1 and 4 in gauche-butane is 3.11 Å in MP2/6-311+G(2d,p) and 3.10 Å in the MP2/sbsse-2 basis sets (Basis with a small superposition error. To be published elsewhere). Even assuming that bonded interactions make these atoms contact with substantial strain, the R* for carbon can hardly be expected to exceed 1.9 Å; most likely it would be 1.7–1.8 Å. As already mentioned, this problem is typically circumvented by scaling the Lennard-Jones parameters for 1–4 interactions (AMBER), or assigning special parameters for 1–4 interactions (Charmm). However, the problem arises again for carbons 1–5 in gauche-gauche-pentane, for which quantum mechanics predicts a distance of 3.68 Å, and here the possible strain is a much weaker argument. Taking this as the equilibrium distance, we have R*=1.84 Å, which may reasonably be regarded as the upper

bound. In all cases we have tested, smaller radii do not appreciably distort the geometry; the suitable range is 1.7–1.85 Å. In AMBER-ii, we take $R^*=1.75$ Å.

Tables 5 and 6 lists the 1–4 distances in gauche-butane and the gauche–trans energy difference. AMBER-ii is obviously best consistent with quantum-chemical data in both parameters.

Still more interesting is the situation in pentane. It has two torsional angles to admit the gauche form; if one dihedral is already gauche, the energy cost of the second transition is twice as low.

The parameters in AMBER-ii are chosen so as to be consistent with quantum mechanics. OPLS-AA overestimates the gauche-conformer energies and the second gauche transition costs as much as the first. Perhaps it is this feature that has not allowed adapting OPLS to lipid simulations. The force fields that successfully simulate lipid bilayers have lower gauche-conformation energy [7, 17]. It has been shown [19] that in OPLS-AA the C_4 to C_6 alkanes show a heat of vaporization very close to the experimental one, whereas in C_7 to C_{12} alkanes it is overestimated by 9–11 %. The matter may be that higher alkanes begin to crystallize during simulation for the lack of gauche forms, which in its turn is a result of setting the gauche energy too high. Indeed, if in AMBER-ii the torsional potential is altered so that the gauche–trans energy difference becomes ~ 1 kcal mol^{−1}, such a dodecane crystallizes completely in some 10 ns. A similar phenomenon can be observed in lipid membrane simulations [17]. At excessive gauche-conformer energy, the membrane passes into a more ordered structure, known as gel phase. It is to avoid such problems that we have paid special attention to alkane gauche conformations.

Charges

The question of partial atomic charges in alkanes still awaits a substantiated solution. Because of the symmetrical (tetrahedral) arrangement of substituents around each carbon, there are almost no dipole or quadrupole moments in alkanes. With the ESP [20] procedure, the charges on hydrogen at the aliphatic carbon may vary from +0.3e to negative values. As alkane molecules are rather flexible changing their conformations during a rather short time, it would be more appropriate to use conformationally averaged values for the charges (instead of charges of the optimized geometry) which turned out to be around zero [7]. There are several good force fields which choose this way.

There are several good force fields that simply ignore the partial charges. Such are the united atoms fields and some all-atoms fields such as TraPPE-EH [21, 22]. We also decided for zero charges, and found that this does not practically affect the thermodynamic properties. The exception is the ethyl group, where we ascribe $-0.012e$

Table 5 Gauche-butane parameters

	CH3 - CH3, ^a Å	dE, ^b kcal mol ^{−1}
MP2/6-311+G(2d,p)	3.114	0.558
AMBER-ii	3.110	0.531
AMBER-i	3.178	0.337
AMBER99	3.153	0.856
OPLS-AA		0.80 ¹⁹

^a Distance between carbons 1 and 4

^b Energy difference between gauche and trans-conformations

Data for OPLS-AA taken from [19]

to the methyl carbon and accordingly +0.012e to the methylene one; such charges enable reproduction of dipole moments in linear alkanes (MP2/aug-cc-pVDZ). A similar scheme is used in sLipids [7].

Thermodynamic calculations

As shown with OPLS, the van der Waals interaction parameters can be extracted from the experimental data on liquid densities and heats of vaporization. OPLS performs excellently on low-molecular alkanes, and hence on amino acids. In this case, the same parameters for all hydrogens and for all carbons do suffice, which is a great merit of the approach.

In our experience, it is also possible to describe linear alkanes with one-value parameters, e.g., upon increasing the Lennard-Jones well depth for hydrogen to $\epsilon=0.0375$ kcal mol^{−1} (unpublished). To our regret, this scheme accommodates neither branched alkanes, for which the CH_3 well depth has to be almost zeroed, nor methane, for which the ϵ must be increased. Since our goal is a unified description of the whole family of alkanes rather than just linear alkanes, we abandon the use of the same value of ϵ for all SP³ carbons. Upon fitting to the heats of vaporization, this parameter proves almost linearly dependent on the number of substituents (Fig. 4).

Table 6 Pentane parameters

	CH3 - CH3, ^a Å	dE, ^b kcal mol ^{−1}	
		gt–tt	gg–gt
MP2/6-311+G(2d,p)	3.685	0.545	0.251
AMBER-ii	3.707	0.561	0.268
AMBER-i	3.766	0.353	0.047
AMBER99	3.793	0.867	0.647
OPLS-AA		0.94 ¹⁹	0.95 ¹⁹

^a Distance between carbons 1 and 5

^b Energy difference between specified conformations: g gauche, t trans

Data for OPLS-AA taken from [19]

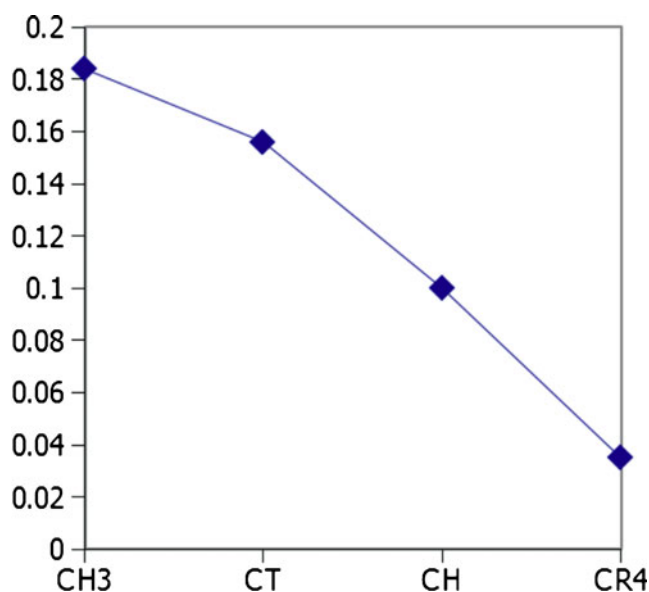


Fig. 4 Dependence of the Lennard-Jones well depth (kcal mol^{-1}) for carbon on the degree of substitution

Hydrogen parameters can be made the same, but to fit the experiment they have to be changed somewhat relative to AMBER. We take $R^* = 1.495 \text{ \AA}$; $\epsilon = 0.0124 \text{ kcal mol}^{-1}$.

The resulting scheme is consistent with the way of constructing AMBER94/99. Nonetheless, the question of hydrogen ϵ remains open, inasmuch as it can be varied quite broadly ($0.01\text{--}0.04 \text{ kcal mol}^{-1}$) and, compensating for the changes by adjusting the carbon ϵ , we can generate a series of fields with close properties.

As an additional test of our parameters, we have determined the solvation free energies of pure alkane liquids. Solvation free energy (or excess chemical potential) is determined as a Gibbs free energy change in transfer of a molecule from an (ideal) gas phase to the same volume of liquid phase. Experimentally, the solvation free energies of pure components are determined from the pressure of saturated vapor at equilibrium with the liquid. Experimentally consistent solvation free energies are important for correct description of phase equilibria and for partitioning (distribution) of substances across different phases. Simultaneous reproduction of the vaporization heat and solvation free energy means also that the entropy of the liquid phase is described correctly. We did not use experimental solvation free energies in derivation of AMBER-ii parameters; that is why agreement between simulated and experimental data would be an additional validation of our approach.

Computational details

Molecular dynamics simulations were run in Ascalaph and Abalone [12], accelerated with NVIDIA GeForce GTX 560 Ti and GeForce GTX 670. A four-level r-RESPA integrator [23] was used to integrate equations of motion. Bonded interactions were computed with a step of 0.25 fs, nonbonded ones with steps 0.5 f. at 3 Å cutoff; 1 f. at 6 Å cutoff; 2 fs up to the half-box length cutoff. In order to use a longer time step while integrating the bonded interactions, we assigned a mass of 10 to all hydrogens. Such substitution does not affect equilibrium structural and thermodynamics properties. Further, we used Andersen thermostat [24] and Berendsen barostat [25] with

Table 7 Thermodynamic properties of liquid alkanes computed in AMBER-ii and compared with experimental data

	ρ_{calc} , g/cm^3	ρ_{exp} , ^a g/cm^3	err, %	ΔH_{calc} , kcal mol^{-1}	ΔH_{exp} , ^a kcal mol^{-1}	err, %	$-\mu_{\text{ex, calc}}$	$-\mu_{\text{ex}}$, ^b exp
n-butane ^c	0.5725	0.5729	−0.07	5.1040	5.0349	1.37	2.41	2.73
n-pentane	0.6241	0.6214	0.44	6.3590	6.3162	0.68	2.77	3.18
n-octane	0.7048	0.6886	2.35	9.9860	9.9163	0.70	4.26	5.3
n-dodecane	0.7533	0.7452	1.08	14.7880	14.6479	0.96	6.74	7.29
cyclopentane	0.7233	0.7440	−2.78	6.4520	6.8181	−5.37	2.87	3.8
cyclohexane	0.7590	0.7739	−1.93	7.5720	7.8609	−3.68	3.31	4.41
isobutane	0.5514	0.5509	0.10	4.6350	4.5700	1.42	2.14	2.49
isopentane	0.6165	0.6142	0.37	5.9210	5.8800 ^d	0.70	2.53	3.22
neopentane ^c	0.6064	0.6012 ^e	0.87	5.3330	5.4381	−1.93	2.17	2.84

ρ : liquid density

ΔH : heat of vaporization, $\Delta H = E_{\text{inter}} + RT$

^a Experimental data from [36]

^b Calculated from data of [36]

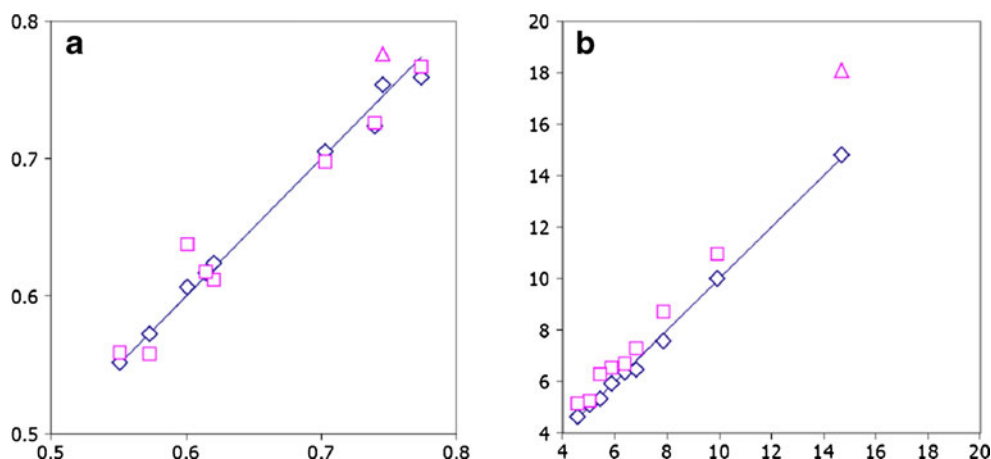
^c All at 298.15 K except neopentane (282.61 K). Butane regarded as overheated liquid

^d from [37]

^e NIST Chemistry WebBook [38]

μ_{ex} - solvation free energy. The experimental values were determined from the data on the pressure of saturated vapor (neglecting non-ideality of the vapor phase) according to $\mu_{\text{ex}} = RT \ln(P_{\text{vap}} V_m / RT)$ where V_m is molar volume of the liquid phase

Fig. 5 Correlation of the density, g/cm^3 (a) and vaporization heat, kcal mol^{-1} (b) between simulated (ordinate: AMBER99, purple squares, and AMBER-ii, blue rhombs) and experimental data (abscissa) for compounds listed in Table 7. Purple triangle shows result of AMBER99 for dodecane which falls into crystalline phase in simulations at normal conditions



relaxation time $\tau=2$ ps with pressure correction for potential cutoff. Thermodynamic properties were calculated with energy correction [26]. All simulations were carried out for 125 molecules in a cubic periodic box. Equilibration lasted 1 ns and the productive part was 4 ns. Hydrogen charges in AMBER fields were taken to be zero.

Solvation free energy calculations were carried out in the MDynaMix v.5.2.5 package [27]. The expanded ensemble method [28] was used according to which a soft-core potential reaction path [29] was introduced which decouples the interaction of one selected molecule with the rest of the system in a number of discrete steps (called also subensembles). During the simulation, the coupling parameter is periodically changed according to the Metropolis Monte Carlo rule, so that the system repeatedly change the subensemble from the case when the selected molecule fully interact with others to the last subensemble when it is completely decoupled. The resulting solvation free energy was calculated from equilibrated probabilities of visiting subensembles:

$$A_{\text{solv}} = -k_B T \ln \frac{p_M}{p_0} + \eta_M - \eta_0,$$

where p_M and p_0 are probabilities of visiting subensembles with fully interacting and completely decoupled interaction respectively and η_M and η_0 are the corresponding weighting factors determined within Wang-Landau optimization [30], see details in [29]. In the computations of this work, the number of subensembles was between 11 and 26 depending on the size of the molecule, which provided a statistical uncertainty of the computed free energy within $0.1 \text{ kcal mol}^{-1}$.

Quantum-chemical computations were performed with NWChem [31] and Firefly [32, 33] in the Ascalaph graphical shell. Conformational computations were run in the 6-311+G(2d,p) basis with MP2 electron correlation.

Comparisons of molecular conformations always included optimization of geometry.

Results and discussion

The thermodynamic properties of liquid alkanes are listed in Table 7.

The initial fitting of parameters was done with linear alkanes, so for them we have the best correspondence with the experiment. In parameterization of branched alkanes, we changed only the ϵ parameter for tri- and tetrasubstituted carbon. For cyclic alkanes, we got densities underestimated by 2 % and hence somewhat lower heats of vaporization. Thus far we have failed to identify the cause of this deficiency. In principle we could reduce the hydrogen radius: thereby the density of cyclohexane and cyclopentane would be underrated only by 1 %, but concomitantly we would overestimate by 1 % the densities of all linear alkanes. We decided against this.

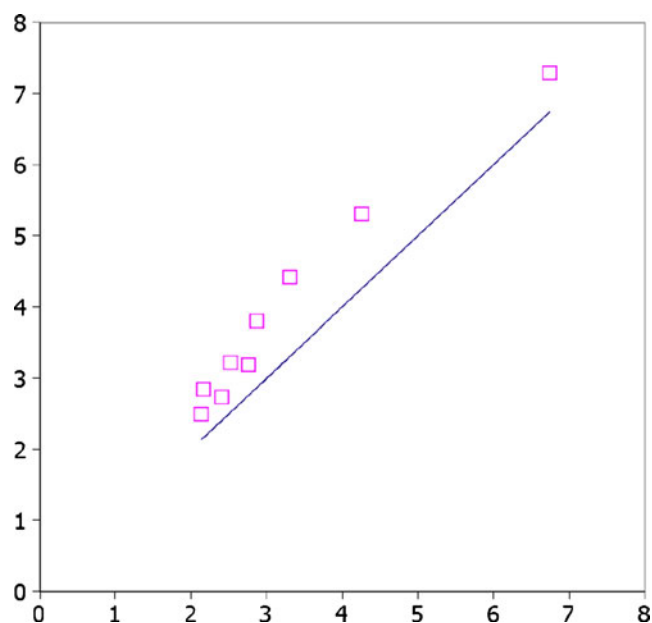


Fig. 6 Correlation of solvation free energy (kcal mol^{-1}) between experimental data (abscissa) and simulated in AMBER-ii (ordinate) for compounds listed in Table 7

Table 8 Comparison of conformation energy differences (kcal mol⁻¹) for some alkanes

	cyclo-C8H16		2,3-dimethylbutane	cyclohexane	methylcyclohexane	trans-1,2-dimethylcyclohexane	bicyclo[3.3.1]nonane
	crown I – boat-chair	crown II – boat-chair	gauche – trans	twist boat – chair	axial – equatorial	ax,ax – eq,eq	boat-chair – chair-chair
MP2/6-311+G(2d,p)	2.45	2.34	0.185	6.29	1.87	2.71	2.39
AMBER-ii	2.20	1.69	0.086	6.36	1.00	1.46	3.05
AMBER-i	0.89	<i>to I</i>	-0.293	7.07	0.43	<i>nd</i>	<i>nd</i>
AMBER99	1.58	1.51	-0.158	6.58	1.57	2.23	4.18
AMBER94	0.12	<i>to I</i>	-0.406	7.53	1.28	1.61	4.31

(*to I*) Crown II is unstable

(*nd*) Not determined because of the lack of required parameters in AMBER-i, which was parameterized only for alkanes found in the 20 natural amino acids

It should be noted that a similar underestimation (2.5 % for cyclohexane) was reported in OPLS-AA [5].

Figure 5 shows correlation of results computed using AMBER99 and AMBER-ii with experimental data. One can see that AMBER-ii provides clear overall improvement in reproduction of the experimental density and heat of vaporization in comparison with AMBER99. This improvement can be prescribed to the use of different epsilon parameter at carbon atoms with different degree of substitution contrary to AMBER99 which uses the same Lennard-Jones parameters for all tetrahedral carbons.

For dodecane (marked in Fig. 5 by triangles) one can observe especially strong deviation of the AMBER99 results from the experimental data. This happens due to crystallization of dodecane at these conditions. We have already discussed that the temperature of transition to ordered phase in model alkanes depends strongly on the torsional potentials and the difference of the trans-gauche conformational energy. The torsional potential, obtained in this work using high-quality quantum-chemical calculations after specifying the LJ parameters, provides higher fraction of gauche-conformation in comparison with AMBER99, and results in dodecane remaining in the liquid phase under normal temperature 298K.

Figure 6 shows correlation of simulated and experimental solvation free energies. The solvation free energies are systematically underestimated though they follow the experimental trend. Since the vaporization energies (computed from the average potential energies) are in very good agreement, it is

the entropic contribution to the free energy of transfer from gas phase to liquid that is somewhat overestimated. Most likely that these deviations are related to the quantum features of the vibrational motion. Reproduction of solvation free energies is a very tough test of a molecular model, so the results obtained here can be regarded as satisfactory.

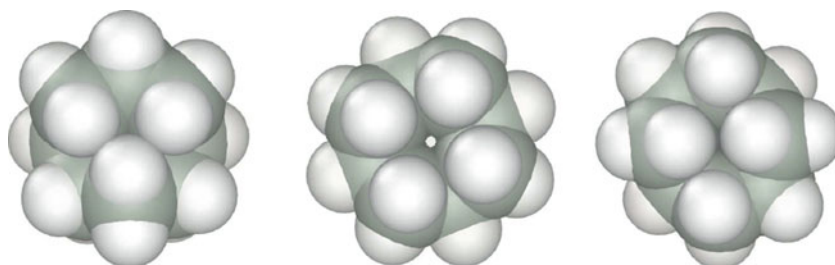
As was outlined above, the bonded interaction constants were fitted to quantum-chemical data for ethane, propane, *n*-butane, and *n*-pentane. In order to see to what extent these parameters are transferable, let us consider some molecules that often raise problems in conformational analysis [2]. Table 8 lists the energy differences for two characteristic conformers of cyclooctane, 2,3-dimethylbutane, cyclohexane, methylcyclohexane, trans-1,2-dimethylcyclohexane, and bicyclo[3.3.1]nonane. AMBER-ii and AMBER99 yield more correct values than AMBER-i or AMBER94 do. Such conformations as transitions between them are often present in many biological molecules, thus a correct (relative to high quality ab-initio calculations) description of their conformational behavior may be of importance for adequate modeling of biomolecular systems.

Strained molecules

Strained molecules offer a hard test for a force field. We therefore look closer at two special distinct cases, cyclooctane and Bicyclo[3.3.1]nonane.

Cyclooctane is an example of a compound that has large conformational freedom but no “relaxed” state. The global

Fig. 7 Selected cyclooctane conformations: boat-chair, symmetric crown, and distorted crown



minimum is the boat-chair (Fig. 7, left), the second best is probably a crown. The latter includes two close-energy varieties: symmetric with a four-fold axis (crown I in Table 8, middle in the figure) and distorted (crown II in Table 8, right in the figure). According to quantum-chemical calculations, crown I is a saddle point whereas crown II is a local minimum. In AMBER-i and AMBER94, the minimum appears to be crown I (because of crown II instability), while in AMBER-ii and AMBER99 it is the distorted crown II. All molecular mechanics models underestimate the energy difference relative to the boat-chair, but AMBER-ii yields the closest result (Table 8).

Bicyclo[3.3.1]nonane, as distinct from cyclooctane, is a conformationally rigid molecule, restricted to boat and chair configurations of each six-member ring. Its chair-chair conformation is strained because of the close approach of two methylene groups (Fig. 8). Therefore, its geometry is very sensitive to the van der Waals parameters for C and H as well as to bonded interaction parameters. Table 9 lists the calculated distances of the optimized geometry. In all cases the distances are overestimated relative to quantum-chemical calculations, but again the best result is obtained in AMBER-ii. The error in distance between carbons in AMBER-ii is twice as small than in the other fields (see also the energy differences in Table 8).

The example of Bicyclo[3.3.1]nonane once again confirms our major issue that the C radius in AMBER has been set too high. It might be possible that the H radius is also too high in all four fields. Note in this connection that in recent papers on extension of the AMBER force field to cyclic α,α dialkylated peptides [34, 35] the authors have decreased the carbon radius to 1.758 Å which is very close to the value of our work. This can be another confirmation that a smaller value of tetrahedral carbon radius [18] than the one used in the traditional force fields is realistic and can improve consistent description of molecular entities containing alkane fragments.

We have made additional test calculations adjusting the radii to describe the bicyclo[3.3.1]nonane geometry best,

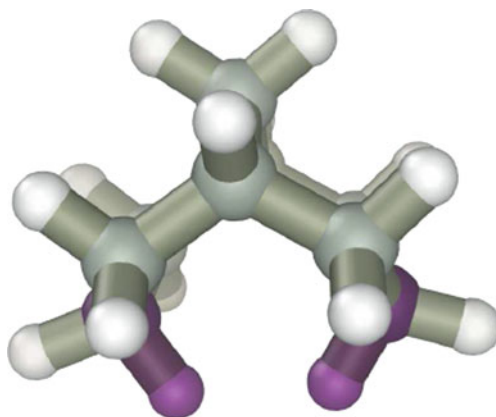


Fig. 8 Bicyclo[3.3.1]nonane in the chair-chair conformation. The highlighted carbons and hydrogens come into utmost contact, giving rise to considerable strain

Table 9 Distances between approaching atoms in chair-chair bicyclo[3.3.1]nonane

	H-H, Å	C-C, Å
MP2/6-311+G(2d,p)	1.9077	3.1189
AMBER-ii	2.0357	3.1633
AMBER99	2.0986	3.2180
AMBER94	2.0900	3.2159

and obtained $R^*_{\text{H}}=1.346$ Å and $R^*_{\text{C}}=1.85$ Å. These numbers show the spread of the possible range in which the parameters may vary. In principle it would be possible to find parameters which would provide even better precision for each specific case, but in expense of losing universality. Note also that optimizing parameters reproducing the geometry alone may induce more disagreements for the thermodynamic properties.

Conclusions

In this work we have developed a new set of force field parameters for alkanes compatible with AMBER family of force fields. Parameterization was made on the basis of high-quality ab-initio computations together with experimental data for densities and vaporization heats of small alkanes, without explicit linking to biological molecules. That is why the new set of parameters can be used for modeling of hydrocarbons directly. Still, everything was done to keep compatibility with the biological context of AMBER and we hope that the developed parameters would be useful in modeling of alkane constituents of biomolecules such as lipids and peptides. Particularly, the new set of parameters provides better thermodynamic properties of liquid alkanes while keeping (and even improving) their conformational behavior. This was reached by introduction of two modifications to AMBER99. First, epsilon parameter of the Lennard-Jones potential for tetrahedral carbons was made dependent on the number of attached hydrogens. While this modification increases the number of atom types and parameters in the force field, it seems necessary for description of both conformational and thermodynamic properties in a consistent manner. Second, we substantially reduced the Lennard-Jones radius for sp^3 carbons. This modification seems to be necessary for adequate description of conformational properties of alkanes, and is manifested best on the example of the gauche-gauche conformation of pentane which cannot be reproduced with larger carbon radii. As a result, we have obtained a force field for simulations of linear, branched and cyclic alkanes which provides their conformational properties in agreement with high-quality quantum-chemical calculations and thermodynamics behavior in agreement with experimental data.

Acknowledgments This work has been carried out with the financial support of the Program of the Presidium of the Russian Academy of Sciences for Molecular and Cellular Biology and the Russian Foundation for Basic Research (Grant No. 11-04-02001), and Swedish Research Council (Vetenskapsrådet, Grant 621-2010-5005). The authors are grateful to Alexander V. Galkin for helpful comments and rendering the paper into English.

References

- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM Jr, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197
- Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074
- Jorgensen WL, Tirado-Rives J (1988) The OPLS force field for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666
- MacKerell AD, JrD B, Bellott M, Dunbrack RL, JrJD E, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kucsera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kucsera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
- Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
- Klauda JB, Venable RM, Alfredo Freitas J, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, JrRW P (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114:7830–7843
- Jämbeck JPM, Lyubartsev AP (2012) Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J Phys Chem B* 116:3164–3179
- Jämbeck JPM, Lyubartsev AP (2012) An extension and further validation of an all-atomistic force field for biological membranes. *J Chem Theory Comput* 8:2938–2948
- Skjevik ÅA, Madej BD, Walker RC, Teigen K (2012) LIPID11: a modular framework for lipid simulations using amber. *J Phys Chem B* 116:11124–11136
- Iijima T (1972) Molecular structure of propane. *Bull Chem Soc Jpn* 45:1291–1294
- Lide DR (1960) Microwave spectrum, structure, and dipole moment of propane. *J Chem Phys* 33:1514–1518
- Ascalaph molecular mechanics package <http://www.biomolecular-modeling.com/Products.html>
- Schreiner PR, Chernish LV, Gunchenko PA, Tikhonchuk EY, Hausmann H, Serafin M, Schlecht S, Dahl JEP, Carlson RMK, Fokin AA (2011) Overcoming lability of extremely long alkane carbon–carbon bonds through dispersion forces. *Nature* 477:308–311
- Verma AL, Murphy WF, Bernstein HJ (1974) Rotational isomerism: Aman spectra of n-butane, 2-methyl butane and 2,3-dimethyl butane. *J Chem Phys* 60:1540–1544
- Burkert U, Allinger NL (1982) *Molecular Mechanics* (ACS Monograph 177). American Chemical Society, Washington, D.C
- Balabin RM (2009) Enthalpy difference between conformations of normal alkanes: raman spectroscopy study of n-pentane and n-butane. *J Phys Chem A* 113:1012–1019
- Högberg C-J, Nikitin AM, Lyubartsev AP (2008) Modification of the CHARMM force field for DMPC lipid bilayer. *J Comput Chem* 29:2359–2369
- Dashevsky VG (1982) Conformational analysis of organic molecules (in Russian). Khimiya, Moscow
- Thomas LL, Christakis TJ, Jorgensen WL (2006) Conformation of alkanes in the gas phase and pure liquids. *J Phys Chem B* 110:21198–21204
- Cornell WD, Cieplak P, Bayly CI, Kollman PA (1993) Application of RESP charges to calculation conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc* 115:9620–9631
- Chen B, Siepmann JI (1999) Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of n-alkanes. *J Phys Chem B* 103:5370–5379
- Wick CD, Stubbs JM, Rai N, Siepmann JI (2005) Transferable potentials for phase equilibria. 7. United-atom description for nitrogen, amines, amides, nitriles, pyridine and pyrimidine. *J Phys Chem B* 109:18974–18982
- Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97:1990–2001
- Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72:2384–2393
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
- Allen MP, Tildesley DJ (1989) *Computer Simulation of Liquids*. Oxford University Press, Oxford
- Lyubartsev AP, Laaksonen A (2000) MDynaMix - a scalable portable parallel MD simulation package for arbitrary molecular mixtures. *Comput Phys Comm* 128:565–589
- Lyubartsev AP, Martsinovskii AA, Shevkunov SV, Vorontsov-Velyaminov PN (1992) New approach to monte Carlo calculation of the free energy: method of expanded ensembles. *J Chem Phys* 96:1776–1783
- Jämbeck JPM, Mocci F, Lyubartsev AP, Laaksonen A (2013) Partial atomic charges and their impact on the free energy of solvation. *J Comput Chem* 34:187–197
- Wang F, Landau DP (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett* 86:2050–2053
- Valiev M, Bylaska EJ, Govind N, Kowalski K, Straatsma TP, van Dam HJJ, Wang D, Nieplocha J, Apra E, Windus TL, de Jong WA (2010) NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput Phys Comm* 181:1477–1489
- Firefly QC package [a], which is partially based on the GAMESS (US) [b] source code. a. Granovsky AA, Firefly version 7.1.G, [www.http://classic.chem.msu.su/gran/firefly/index.html](http://classic.chem.msu.su/gran/firefly/index.html)
- Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) General atomic and molecular electronic structure system. *J Comput Chem* 14:1347–1363
- Grubisic S, Brancato G, Pedone A, Barone V (2012) Extension of the AMBER force field to cyclic α , α dialkylated peptides. *Phys Chem Chem Phys* 14:15308–15320
- Grubisic S, Brancato G, Barone V (2013) An improved AMBER force field for α , α -dialkylated peptides: intrinsic and solvent-induced conformational preferences of model systems. *Phys Chem Chem Phys* 15:17395–17407
- Mackay D, Shiu WY, Ma K-Ch, Lee SCh (2006) *Handbook of physical-chemical properties and environmental fate for organic chemicals*. Taylor & Francis, Boca Raton
- Jorgensen WL, Madura JD, Swenson CJ (1984) Optimized intermolecular potential functions for liquid hydrocarbons. 106:6638–6646
- NIST Chemistry WebBook <http://webbook.nist.gov/chemistry/form-ser.html>