

Margit Möllhoff · Ulrich Sternberg

Molecular mechanics with fluctuating atomic charges – a new force field with a semi-empirical charge calculation

Received: 14 August 2000 / Accepted: 24 October 2000 / Published online: 24 April 2001
© Springer-Verlag 2001

Abstract A new force field has been designed to implement the calculation of Coulomb interactions with fluctuating atomic charges. The charges are calculated by use of a semi-empirical quantum chemical method – bond polarization theory (BPT). The BPT method establishes a direct proportionality between molecular properties, for instance atomic charges or chemical shifts, and bond polarization energies. These energies are calculated from bond orbitals that are constructed for every bond of the force field. Thus the charges depend on the three-dimensional geometry of the molecular system, and it is possible to include all mutual polarizations in the term for electrostatic interaction. The primary goal of this new force field is better description of the intermolecular interactions of molecular systems. No special term within the force field is applied for the description of hydrogen bonds. The inclusion of the polarization effect over the whole system is one of the most important advantages of the method in respect of force fields that divide the molecular system into molecular mechanics and quantum chemical regions.

The force field was tested by being used to describe the structure and interaction energies of several small molecular systems (26 hydrogen-bonded dimers) from a web-based *ab initio* data collection by Halgren. The results show an overall RMS deviation of 2.5 kcal mol⁻¹ for the interaction energies, 0.06 Å for the hydrogen bond distances (X...Z) and 20.1° for the X-H...Z angles. This is comparable with most existing force fields. The results were obtained with the original parametrization of Halgren for the van der Waals interactions without any fine tuning of the interaction parameters.

Additional interaction energies and structures of selected DNA/RNA base pairs (see Figure) were studied. The geometries of hydrogen bonds, in particular, are reproduced satisfactorily – after geometry optimization the distances differ on average by 0.06 Å and in the angles by 6° from the *ab initio* Hartree–Fock results including correlation.

Keywords Molecular mechanics · Force field parametrization · Fluctuating atomic charges · Fast semi-empirical charge calculation · Intermolecular interactions

Introduction

Description of the intermolecular interactions with the help of potential functions enables the description of how reactions proceed in cells or how pharmaceuticals interact with proteins in a living body. Detailed information is needed to enable understanding of biochemical processes, but we can only use very simple models to obtain an idea of how the compounds interact. Force fields and molecular dynamics (MD) simulations give us the possibility of simulating interactions of large molecular systems with relatively simple potential energy functions.

If a new component, e.g. the BPT charge calculation, is introduced into a molecular mechanics force field, the classical part should be as simple as possible with a minimum number of parameters. Otherwise it seems impossible to fix the problems introduced by the quantum chemical part. We therefore developed our method as a universal force field with a small number of parameters. The parametrization of the different parts should be as independent as possible.

Many problems in force field calculations arise from the calculation of Coulomb interactions with fixed charges neglecting all mutual polarizations. Most of the commonly used force fields work with fixed charges CHARMM [1], AMBER [2], GROMOS [3], OPLS [4],

M. Möllhoff
Beckman Institute UIUC, 405 North Mathews Avenue, Urbana,
IL 61801, USA

U. Sternberg (✉)
Friedrich-Schiller-Universität Jena,
Institut für Optik und Quantenelektronik, Max-Wien-Platz 1,
07743 Jena, Germany
e-mail: sternberg@ioq.physik.uni-jena.de
Tel.: +49-3641-947283, Fax: +49-3641-947282

and MMFF [5], or adjustable dipoles MM3 (all elements) [6]. These charges are parameters adjusted to give reasonable interaction energies in structures not too far from equilibrium geometry. Atomic charges are often exaggerated to account for the energy contributions from polarization, because in classical force fields electric polarization is excluded. Moreover, special terms are sometimes included for the proper description of hydrogen bonds. However, the current trend seems to be that force fields are moving away from specialized potentials for the description of hydrogen bonds, and instead using the already existing van der Waals and Coulomb terms.

Recently a polarizable force field fitted to *ab initio* data was published by Friesner et al. [7] and also applied to selected amino acids [8]. Using the fluctuating charge (FQ) model of Berne et al. [9] the method is based on atomic charges from the electronegativity equalization method [10, 11]. A great advantage of this method, compared with models based on fixed charges is that the charges depend on the topology of the molecule – the two dimensional structure. Other force fields have been published which apply the FQ model to focus on very special tasks – description of hydrocarbons [12], organic molecules [13], or water [14].

The COSMOS force field presented in this paper includes the semi-empirical charge calculation based on bond polarization theory (BPT) [15, 16]. The most important advantage of BPT is that the atomic charges depend on the three-dimensional structure of the molecule. The algorithm is fast enough for molecular mechanics calculations, because only one matrix inversion is necessary (the matrix has the dimension of the number of atoms) to calculate the charges. Most computing time is consumed by the integrals needed for the matrix elements.

In the first section a short description of the potential functions is given. In the second some applications of the force field are presented – internal geometries of a broad variety of small molecules, interactions and donor–acceptor behavior of hydrogen-bonded dimer structures, and the interaction energies and geometries of selected DNA/RNA base pairs.

The COSMOS force field

The COSMOS (COmputer Simulation of MOlecular Structures) force field was designed to fulfil the following requirements:

1. The number of empirical parameters should be as limited as possible.
2. The parameters of the different energy contributions should be independent; in particular the intramolecular terms should be independent of the intermolecular terms.
3. The force field should be universally useful for most common compounds.

The second point, in particular, should be emphasized, because the semi-empirically calculated atomic charges

fluctuate with changing molecular structure. A correct calculation of Coulomb interactions with fluctuating charges can be accomplished only if the electrostatic energy is independent of the parametrization of other energy contributions. Only independent parametrization can guarantee that parts of the force fields can be re-parametrized without changing the rest of the force field.

Special attention was paid to the intermolecular interactions. An improved description of polar effects yields more reliable hydrogen-bond energies. Special terms for the description of hydrogen-bond interactions are not necessary.

Bond energy

The bond-stretching energy is associated with deformation of a bond from its equilibrium length r^0 . For the description of the bond energy we follow some ideas of Smith [17] applied in his PIMM force field. The bond energy between two atoms i and j is calculated using the well known potential:

$$E = \sum_{i>j}^{N_{bo}} k_{ij}(r_{ij})(r_{ij} - r_{ij}^0)^2 \quad (1)$$

where r_{ij} refers to the actual distance between atoms i and j , and $k_{ij}(r_{ij})$ is the force constant. The force constant is calculated depending on the actual bond distance r_{ij} as described by an inverse power series [18]:

$$k(r_{ij}) = \frac{C_2}{r_{ij}^2} + \frac{C_4}{r_{ij}^4} + \frac{C_6}{r_{ij}^6} \quad (2)$$

The parameters C_2 , C_4 , and C_6 and the equilibrium distances r_{ij}^0 for saturated bonds are taken from the literature [17] (see also Ref. [19]). This non-harmonic potential function is used for single bonds between the atoms C, N, O, S, Cl, and F. A fixed value of k_{Hj} (1007 kcal mol⁻¹ Å⁻¹) was introduced for the bonds to hydrogen.

Although we followed PIMM [17] in defining potential and parameters for the force constants and the equilibrium distances for saturated bonds, another approach was used to describe the equilibrium distances r_{ij}^0 for unsaturated bonds and missing parameters of Ref. [17]. The parametrization of the equilibrium distances for unsaturated bonds is more complicated because the equilibrium distances depend on the valence of the bond. The valence of conjugated double bonds may range from 1 to 2. This valence depends on the distribution of π -electrons within the system of conjugated π -bonds and also on the planarity of this π -system.

One possible means of solving this problem is parametrization of the many types of conjugated π -bonds. This option is used in many force fields but many parameters are needed and it is a complicated task to include “all the special cases”, that can occur. Thus, non-planar π -systems, like biphenyl, are “problem molecules” for some force fields.

A second way is to solve the Hückel Hamiltonian as implemented within MMPI [17] or to perform an itera-

tive PPP calculation cycle as in MMP2 or MMP3 [20]. Initially this solution seems attractive and quite general. Problems show up if this procedure must work together with the rest of the force field:

1. a π -bond order for the planar π -system is needed for dihedral forces of conjugated bonds
2. in large molecules there can be several more or less non-overlapping π -systems, because of steric reasons (non-planarity of the π -system); and
3. when the whole molecule is treated as a single π -system, we arrive at very large Hamiltonian matrices.

We propose a treatment of conjugated π -bonds that makes a reasonable attempt to avoid the problems of the former procedures. The idea dates back to Pauling [21], who established a relationship between valence and bond distance. A new parameterization of Pauling's equation was obtained by O'Keeffe and Brese [22]. They proposed a different constant (0.37 instead of the Pauling value 0.35) and developed an empirical relationship between the shortening of an ideal single bond R_{ij} and the valence v_{ij} of a bond between the atoms i and j .

$$r_{ij}^0 = R_{ij} - 0.37 \ln v_{ij}. \quad (3)$$

The parameters for the calculation of the single bond distances R_{ij} are published for most elements of the periodic system [22]. The formula for R_{ij} takes into account bond-shortening arising as a result of electronegativity differences between the elements.

A procedure is necessary to estimate the valences of conjugated bonds. We identify the valence v_{ij} with the occupation number of the conjugated bond and suggest the following formula:

$$v_{ij}^\pi = \cos^2 \phi \sqrt{1/N_i^\pi} \sqrt{1/N_j^\pi}. \quad (4)$$

This formula describes the π -valence of a conjugated bond v_{ij}^π depending on the number of π -electrons forming a π -system with a central bond between the atoms i and j . The number of π -electrons N^π is obtained by counting the number of neighbors of the atom i with π -bonds (N_i^π) and the same procedure applies for atom j (N_j^π). This means that each π -bond of a neighbor of i (or j) contributes one electron to the π -system. If the bond partners of the atoms i (or j) supply lone pairs, a value $N^\pi=2$ is added for each lone pair, assuming that only one half of an electron participates in the π -system.

The weakening of the π -overlap depends on dihedral angle, ϕ , and has to be taken into account for non-planar π -systems. ϕ is the dihedral angle of the π -system with a central bond between i and j . The valence of an ideal conjugated planar π -bond is $v_{CC}^\pi=1/2$ ($N_i^\pi=N_j^\pi=2$ for a C-C bond in benzene). For the central C-C bond between two phenyl rings in biphenyl it is $N_i^\pi=N_j^\pi=3$. The torsion around the central C-C bond in biphenyl reduces v_{CC}^π by $\cos^2 \phi$, compared with the planar case.

The equilibrium distances for all kinds of unsaturated bonds (π -valence $\neq 0$) are calculated according to Eqs (3)

and (4). For single bonds between elements other than between the atoms C, N, O, S, Cl, F, and H it is also possible to obtain reasonable equilibrium bond distances by use of Eq. (3). The default value for all missing force constants is set proportional to the valence by multiplying v_{ij}^π by the default force constant for π -bonds (700 kcal mol⁻¹ Å⁻¹).

Angular energy

The bond angle term is associated with the deformation of an angle from its equilibrium value. Most force fields use harmonic potentials. In our approach we follow the valence bond (VALBOND) concept of Landis and Cleveland [23] for the calculation of the angular energy. The idea is based on the connection between the strength of hybrid orbitals and the bond angles at an atomic center. The angular energy of three connected atoms i - k - j is calculated from the deviation of the orbital strength $S_k(\alpha_{ij})$ of the hybrids from a maximum possible strength S_k^{\max} :

$$E = \sum_k^{cent.at.} \sum_i^{lig.} \sum_j^{lig.} k_k \{S_k^{\max} - S_k(\alpha_{ij})\} \quad (5)$$

where S_k^{\max} is calculated from:

$$S_k^{\max} = \sqrt{\frac{1}{1+m}} (1 + \sqrt{3m}) \quad (6)$$

where m refers to the hybridization of the central atom k (i.e. $m=3$ for a sp^3 hybrid). The preference of the ligands i and j bonded to an central atom k for hybridizational states is parameterized using weight factors k_k . Overlap integrals Δ_{ij} between the hybrids are needed to describe the deviation from the maximum possible orbital strength. The actual overlap integrals Δ_{ij} are calculated from the bond angles and the hybridizations in the following way:

$$\Delta_{ij} = \frac{1}{m} (1 + m \cos \alpha_{ij}). \quad (7)$$

The overlap integrals are used to obtain the generalized hybrid orbital strength $S_k(\alpha_{ij})$:

$$S_k(\alpha_{ij}) = S_k^{\max} \sqrt{1 - \frac{1 - \sqrt{1 - \Delta^2}}{2}}. \quad (8)$$

Thus a non-harmonic potential function is obtained which describes the dependence of the energy on the bond angles of the system with an appreciable reduction of the number of parameters, because only parameters for pairs of atoms and not, as usual, for atom triplets are needed. Using this energy function, a realistic crossing over angular energy barriers is described.

The parameterization works for all elements of the p-block as described in Ref. [23], which covers nearly the whole periodic table. For missing parameters (covalent bonds are excluded for actinides or lanthanides, elements of the second, the 8th and the first main group, with the exception of hydrogen, elements from 7th mean

group are not involved as central atoms k) the contribution to the angle energy will be zero.

Recently a paper appeared, written by Landis and Cleveland, which describes the application of the VALBOND to hypervalent molecules of the p-block [24]. An extension will be implemented in the COSMOS force field on the basis of the ideas of this work.

Dihedral energy

The term that describes the dihedral energy is associated with the tendency to prefer certain 1–4 conformations (e.g. *cis*, *trans*, *gauche*...). As usual, the periodic function:

$$E_{\sigma} = \sum_{k < l} \frac{1}{2} k_{kl} (1 \pm \cos n\tau) \quad (9)$$

was used to describe the dihedral energy. τ is the dihedral angle between the atoms k–i–j–l. The potential $(1 + \cos 3\tau)$ gives a typical threefold barrier in the case of σ -bonds, and for π -systems, $(1 - \cos 2\tau)$ yields twofold barriers. An additional energy term for B-systems takes into account the valence v_{ij} of the central π -bond in association with [17]:

$$E_{\pi} = k_{\pi}(\tau)(1 - \cos^2 \tau) \\ k_{\pi}(\tau) = k_{\pi} \{ v_{ij} + 0.15 v_{ij}^2 \cos^2 \tau + C_{ij}(1 - \cos^2 \tau) \}. \quad (10)$$

The force constant of the π -bond depends on the valence v_{ij} for the planar system k–i–j–l. The valence v_{ij} is calculated by use of Eq. (4), omitting the $\cos^2 \phi$ term. To obtain forces with a preference for the planar state of the π -system the constant is set proportional to the bond valence for the case $\phi=0$. C_{ij} are correction factors, depending on the atom type. Although two-parameter matrices are necessary (one for the force constant and one for the corrections), an appreciable reduction of the number of parameters was achieved, because only parameters for atom pairs are needed and not for atom quartets, as is usual. All parameters were taken from Ref. [17]. Parameters are available for all central π -bonds (i–j) between the atoms C, N, O and for the 1–4 interactions between the atoms H, C, N, O, S, F, and Cl. For missing parameters the contribution to the angular energy will be zero. The force constant k_{π} is obtained from the energy that is needed to break the ethene π -bond ($k_{\pi}=188.372/4$ kJ mol⁻¹) [25].

Bend energy

There are no proper dihedral angles that keep the π -system planar in π -bonds to terminal atoms, e.g. carbonyl oxygen atoms. An additional force must be introduced to readjust the system into planar conformations for improper dihedral angles. The potential applies only if the central atom is an sp^2 carbon or sp^2 nitrogen. The

energy contribution is proportional to deviation from planarity of a π -system:

$$E = 4k_{\pi} \sum_i^{sp^2} \sum_j^3 v_{ij}^{\pi} \cos^2 \beta_{\perp}. \quad (11)$$

There are three bond vectors a, b, and c from a central sp^2 hybridized atom i to the next neighbor atoms j, k, and l (ligands). The angle β_{\perp} is the angle defined by the vector normal to the plane spanned by b and c and the vector a of the π -bond. The resulting force is zero for a planar π -system, and $\neq 0$ for a non-planar sp^2 center. The force constant is set to the energy value needed to break a C–C π -bond. The valence v_{ij} is calculated by use of Eq. (4), omitting the $\cos^2 \phi$ term (by analogy with the dihedral angle the valence is taken for the planar case).

Van der Waals energy

Special attention is devoted to description of intermolecular interactions. The van der Waals term describes the repulsive and attractive energy contributions of non-bonded atoms. We used the potential functions of Halgren [26] for description of van der Waals interactions. For the intramolecular interactions a Exp-6 potential is used (Eq. 12):

$$E = \sum_{i < j} \epsilon_{ij} \left[18400 \exp \left(-12 \frac{R_{ij}}{R_{ij}^*} \right) - 2.25 \left(\frac{R_{ij}^*}{R_{ij}} \right)^6 \right] \quad (12)$$

and for the intermolecular interaction a 9–6 Lennard Jones potential (Eq. 13) is used:

$$E = \sum_{i < j} \epsilon_{ij} \left\{ 2 \left(\frac{R_{ij}^*}{R_{ij}} \right)^9 - 3 \left(\frac{R_{ij}^*}{R_{ij}} \right)^6 \right\} \quad (13)$$

where ϵ_{ij} is the well-depth and R_{ij}^* is the distance for the minimum energy separation. The parameters and the combination rules were taken from Halgren [26]. Some small modifications of Halgren's parameters were made in the COSMOS force field (to reduce the parameter set) – atom types are automatically assigned, and therefore atom types for special structures (i.e. carbon 4-rings refer to type 30 in Ref. [26]) are omitted. If the parameters are the same (e.g. for type 3, 4, 37, 41, 57, and 60 in Ref. [26]) only one type was used.

Some updates and extensions of the parameter set for the COSMOS force field are possible to take into account more recent work by Halgren on MMFF94 [5, 27] to improve the description of the van der Waals interaction.

Coulomb energy

The primary goal of this new force field is a good description of the intermolecular interactions of molecular systems by van der Waals and Coulomb potential functions. The calculation of Coulomb interactions with fluctuating atomic charges is possible by using a semi-empirical quantum chemical method [16] based on bond po-

larization theory (BPT) [15]. Within the framework of the BPT for every bond that was defined within the force field a two-center bond orbital is constructed. The polarity of the bonds is described by one free empirical parameter d_{ab} for every bond orbital. The wave function of our molecular system can now be written as a Slater determinant made up of bond orbitals. To include polarization we also introduce anti-bonds for every bond and add to the ground state Slater determinant configurations that contain bond to anti-bond excitations. The contributions of these excited configurations are calculated from a perturbational treatment. From this new wavefunction including polarization, it is possible to derive atomic charges as expectation values of projection operators. With some simplifications, and neglecting small contributions, we arrive at the following equation for the atomic charges q_a :

$$-q_a = \sum_b^{n_a} N_{ab}^\sigma d_{ab}^\sigma + \sum_b^{n_a} (N_{ab}^\sigma)^2 A_{ab}^\sigma \sum_{\chi \neq a,b} q_\chi \left(\left\langle \chi_b \left| \frac{1}{|\vec{R} - \vec{r}|} \right| \chi_b \right\rangle - \left\langle \chi_a \left| \frac{1}{|\vec{R} - \vec{r}|} \right| \chi_a \right\rangle \right) + \pi - \text{term}. \quad (14)$$

For every charge within the molecule we obtain an equation of the type Eq. (14) but these equations are coupled because the charge on atom a depends on the charges q_x of all other atoms. The first term in Eq. (14) represents the contribution of the n_a bonds of the first bond sphere. The second term accounts for the polarization of the first bond sphere by the charges q_x of the other atoms. Because only excitations from bonds into their own anti-bonds are taken into account, we have to calculate integrals over the hybrids χ_a and χ_b that form the bonds of atom a . If an atom forms π -bonds we have to add two terms of the same structure. To account for partial π -bonds each bond contribution is multiplied by a bond occupation number N_{ab} . These occupation numbers are estimated empirically by using the valence formula (Eq. 4).

For each bond two empirical parameters are needed – the polarity of the bond d_{ab} and the change of the charge with bond polarization A_{ab} . These parameters are calculated only once using atomic charges obtained from ab initio calculations on a set of calibration molecules. Inserting known charges q_a and q_x into Eq. (14) we obtain an over determined set of linear equations for the parameters d_{ab} and A_{ab} .

The BPT charges depend on the three-dimensional geometry of the molecular system. In this way, it is possible to include all mutual polarizations into the term for the electrostatic interaction using the atomic charges obtained from the BPT. Users of traditional QM-MM methods have to deal with the problem to decide which part should be included in the QM calculations and which part of the molecular system is treated by MM, and, further, how the two parts should be connected or divided. Usually the QM atoms can “feel” the polarization effects of the MM part, but not vice versa. This means the polarization effects can only act in one direction.

The BPT charge calculation is fast enough to treat the whole system (up to 6000 atoms), but on the other hand, options are provided to fix the charges of one part of the system and include only their polarizing influence. For large systems (more than 6000 atoms) this option can be used to reduce the time taken and the memory required by the matrix built up for charge calculation.

By introducing the BPT procedure into a force field, there is no boundary between QM and MM part. The inclusion of the polarization effect over the whole system is one of the most important advantages of the BPT method compared with QM-MM methods for treatment of systems of the same size.

Dealing with the calculation of large molecular systems, the number of non-bonded interactions dominates the calculation time. Small contributions of atoms to interaction energies at large distances can be neglected by using cutoff radii. Cutoff radii can also be introduced into the BPT integral calculations to reduce the number of polarization integrals, but there are problems associated with this simplification – the total charge is only conserved for relatively large cutoff radii (>12 Å). If two connected atoms “see” slightly different electrostatic surroundings, the ratio of the bond polarization energies for the connecting bond is different, leading to violation of charge conservation.

Because atomic charges derived from ab initio calculations depend on the basis set and on the method of population analysis, it is necessary to define atomic charges that represent the electrostatic energy with sufficient accuracy. In a previous paper [16] the parametrization of the BPT equations was performed by using atomic charges derived from ab initio STO-3G calculations and Mulliken population analysis. This type of charge is not well suited for calculation of intermolecular interaction energies. Charges fitted to ab initio calculations of electrostatic potentials (ESP charges) proved to give more realistic interaction energies, and ESP charges are therefore used in most modern force field parametrizations (see, e.g., Ref. [28]). In a recent paper [29] a good correlation ($R=0.992$) was found between BPT charges and ESP charges. The slope of this correlation was 2.05, and therefore all electrostatic energies (except ions) were scaled by $C_s=4.2$. The electrostatic interaction energy is calculated between all non-bonded neighbors (starting with 1–3 interactions) directly from Coulomb’s law, and no distance-dependent dielectric constants are used:

$$E = C_s \sum_{i < j} \frac{q_i q_j}{r_{ij}}. \quad (15)$$

Parameters for BPT charge calculations are available for the following bond types: C–C, C–N, C–O, C–H, C–F, C–Cl, N–H, O–H, Si–H, Si–C, Si–O, Si–Cl, and P–O. If there is no BPT parameter for the charge calculation available, the charges are calculated from a sum of contributions of Pauling electronegativity differences [30]:

$$q_{ab} = (0.16 + 0.035 |EN_a - EN_b|)(EN_b - EN_a). \quad (16)$$

There could also be problems connected with this approach, because the neighboring atoms are influenced by the differently calculated atomic charges. During extensive testing no problems or only minor influences on the interaction energies are registered.

No special term within the force field is included for description of hydrogen bonds. If a hydrogen bond is detected, the van der Waals' term is switched off, instead bond polarization supplies the stabilization energy of the hydrogen bridge. The detection of a hydrogen bond follows a test routine, which looks for a distance between an acceptor and donor of the hydrogen bond within a range of the bond distance plus 1.6 Å for the acceptor and donor.

Results and discussion

Internal geometries of small molecules

We used a set of 30 test molecules from Halgren [31] as reference for our calculations. The data in Ref. [31] provide, among other information, dimer geometries and also HF/6-31G*-optimized geometries of 44 small molecules. A five digit code is given in parentheses behind the name of the molecule and refers to [31]:

FORMAMIDE (AM01A), TRANS-N-METHYL ACETAMIDE (AM04A), ACETATE ION (AN02A), PYRIDINE (AR02A), PYRROLE (AR06A), FURAN (AR07A), THIOPHENE (AR08A), IMIDAZOLE (AR09A), PYRIDINE N-OXIDE (AR21A), INDOLE (AR25A), CIS-ETHANOIC ACID (CA02A), CIS-METHYL FORMATE (CE01A), FORMALDEHYDE (CO01A), ACETONE (CO04A), FLUOROMETHANE (HL01A), FORMALDEHYDEIMINE (IM03A), N-METHYLAMINE CATION (NC02A), GUANIDINE CATION (NC06A), IMIDAZOLE CATION (NC12A), METHYLAMINE (NH01A), AMMONIA (NH10A), ANILINE, N-PUCKERED (NH14A), METHYLETHYLAMINE OXIDE, CNCC ANTI (NH20A), METHANOL (OH01A), PHENOL (OH08A), WATER (OH09A), DIMETHYL ETHER (OR06A), HYDROGEN SULFIDE (SR01A), DIMETHYL SULFIDE (SR03A) and METHYL DISULFIDE (SR06A).

In addition pyrazole was included for structural comparison. The geometry data of ab initio QCISD/6-31G** calculations were taken from Ref. [32]. We performed force field optimizations with COSMOS and MSI CHARMM (available over Quanta 97 [33]) for comparison. All calculated data were compared with proved experimental bond length and bond angles from three sources [34]. A complete table of the data (experimental, ab initio, COSMOS and CHARMM) is available on request from the authors. Linear correlations were estimated for calculated and experimental bond length and bond angle data; the linear correlations between experimental bond lengths and bond angles and values calculated by use of the COSMOS force field, for all 31 test molecules, are shown in Fig. 1.

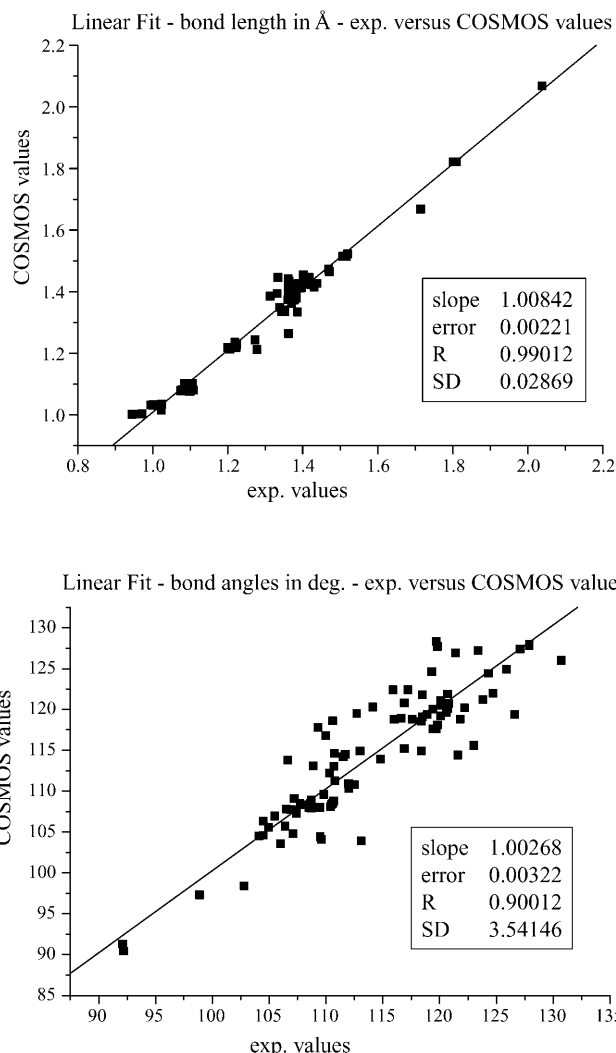


Fig. 1 Correlations of experimental bond lengths and the bond angles with values calculated by use of COSMOS for all 31 test molecules

A summary of the correlation results between experimental data and optimized geometries from different calculations (ab initio, CHARMM, and COSMOS) are given in Table 1. Results from linear correlation of bond length and bond angle data from force field calculations (COSMOS and CHARMM) with data from ab initio calculation (from Halgren [31]) are shown in Table 2. Tables 1 and 2 also list RMS values for the same data pairs. For visual comparison the linear fit through zero of bond length and bond angles from the COSMOS and ab initio calculations, for all 31 test molecules, are shown in Fig. 2.

As a result of this investigation, we can say that the COSMOS force field can satisfactorily reproduce bond lengths and bond angles. There is no doubt that ab initio calculations are able to reproduce the experimental values of small molecules at best (also with moderate basis sets). For our test set of 31 molecules we obtained RMS values of 0.01 Å for the bond lengths and 1.4° for the

Table 1 Summary of correlations between experimental bond lengths and bond angles and values calculated by use of different methods. B refers to the slope of the linear fit through zero, R is the correlation coefficient and SD is the standard deviation

Parameter	ab initio	COSMOS	MSI CHARMM
Number of data pairs	101	106	106
RMS (distances) in Å	0.010	0.029	0.015
B (distances)	0.99242	1.00842	0.99675
Error of B	8.64235E-4	0.00221	0.00116
R	0.99878	0.99012	0.99721
SD	0.01098	0.02869	0.01507
Number of data pairs	85	94	94
RMS (angles) in °	1.367	3.726	2.735
B (angles)	1.00188	1.00187	1.00337
Error of B	0.00131	0.00322	0.00248
R	0.98587	0.91276	0.94192
SD	1.37167	3.72521	2.74136

Table 2 Summary of correlations between ab initio calculated data and bond lengths and bond angles calculated by use of different force field methods

Parameter	COSMOS	MSI CHARMM
Number of data pairs	140	140
RMS (distances) in Å	0.336	0.205
B (distances)	1.01774	1.00432
Error of B	0.00228	0.00139
R	0.98586	0.99466
SD	0.03391	0.02069
Number of data pairs	121	121
RMS (angles) in °	3.103	2.217
B (angles)	0.99927	0.99906
Error of B	0.00247	0.00176
R	0.93499	0.95971
SD	3.10310	2.21555

bond angles. The force field calculations yield, as expected, larger deviations from the experimental values with 0.03 Å (COSMOS) and 0.02 Å (CHARMM) for the bond lengths and 3.7° (COSMOS) and 2.7° (CHARMM) for the bond angles, compared with ab initio values.

Fourteen structures were selected from Halgren's set [31] to calculate the overall RMS deviations (for all involved geometry parameters bond lengths, bond angles and dihedral angles) of the force field-optimized structures with regard to the ab initio-optimized structures and the experimental data. The selection was made with regard to the availability of proved experimental data from few sources as possible [34]. The following molecules were investigated FORMAMIDE, N-METHYL-ACETAMIDE TRANS, PYRIDINE, THIOPHENE, IMIDAZOLE, CIS-ETHANOIC ACID, CIS-METHYL FORMATE, METHYLAMINE, AMMONIA, ANILINE, METHANOL, WATER, DIMETHYL ETHER, HYDROGEN SULFIDE.

The overall RMS value for the COSMOS-optimized structures was 0.084 and 0.053 for the CHARMM optimized molecules. If intramolecular geometries are required, COSMOS cannot compete with more elaborate programs, e.g. CHARMM. For all cases tested the RMS

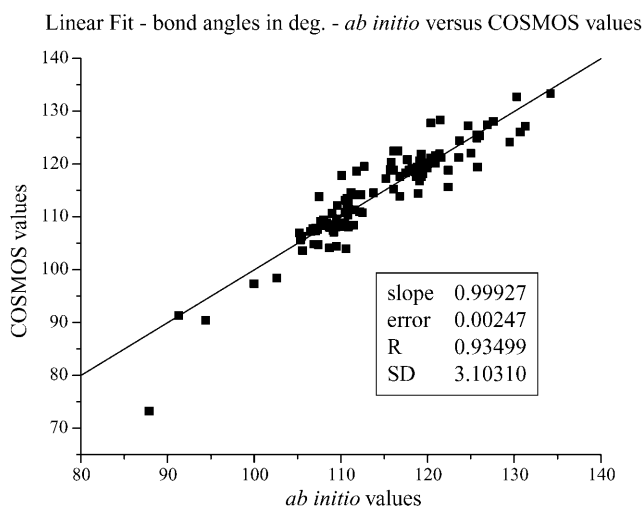
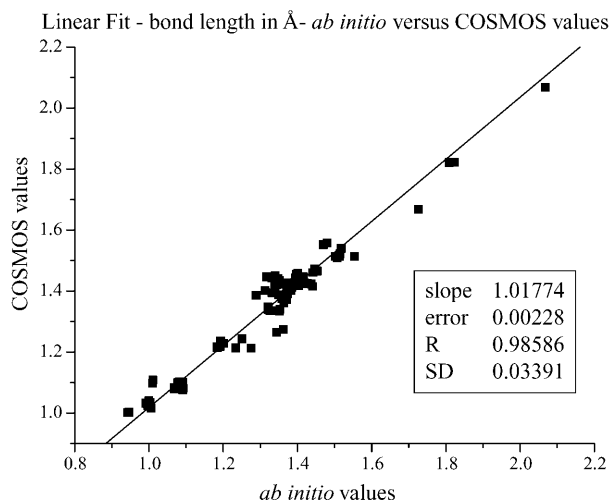


Fig. 2 Comparison of the linear fit through zero for bond lengths and bond angles calculated by use of COSMOS and by ab initio calculations for all 31 test molecules

deviations were larger for the COSMOS values than for the values calculated by CHARMM (force field compared with experimental data and force field compared with ab initio data). Force fields such as CHARMM contain many more parameters than COSMOS and many additional terms. It is, nevertheless, impressive that the simple valence formulas (Eqs 3 and 4) reproduce the bond length of π -systems in such a satisfactory way. This shows that the new concept of separated handling of potential terms with implementation of fully polarizable atomic charges is working. It is, therefore, easy to fix the sources of deviations and tune the parametrization when needed for COSMOS.

In addition to this investigation, we considered the challenge of Halgren [35] to reproduce the conformational energies of cyclohexanol and cyclopentanol (representative drawings are shown elsewhere [35]). In the COSMOS force field the hydrogen charges (q_{CH}) are between 0.063 and 0.064 for both molecules. The relative

Table 3 Interaction energies and hydrogen-bond geometries for pairs of small molecules (X-H...Y) from ab initio [31] (HF/6-31G* / unscaled=QM and scaled=SQM) and COSMOS calculations (FF)

Row no.	Dimer structure ^a	Dimer	Monomer 1	Monomer 2	Energy (kcal mol ⁻¹)				X...Z distance (Å)				H...Z distance (Å)				X-H...Z angle (°)	
					SQM	QM	FF		SQM	QM	FF		SQM	QM	FF		QM	FF
1	HOH...OH ₂	HB01A	OH09A	OH09A	-6.47	-5.62	-5.89		2.73	2.97	2.78	2.03	1.78			172.3	177.3	
2	HOH...OHCH ₃	HB02A	OH09A	OH01A	-6.10	-5.56	-7.12		2.72	2.95	2.81	2.01	1.82			169	169.9	
3	CH ₃ OH...OH ₂	HB02B	OH01A	OH09A	-6.15	-5.59	-7.13		2.72	2.96	2.75	2.01	1.75			174.9	179.5	
4	CH ₃ OH...OHCH ₃	HB03A	OH01A	OH01A	-6.09	-5.54	-8.79		2.71	2.95	2.82	2.00	1.83			174.6	170.0	
5	C ₆ H ₅ OH...OH ₂	HB04A	OH08A	OH09A	-8.10	-7.37	-7.07		2.67	2.90	2.69	1.95	1.69			174	174.4	
6	HOH...C ₆ H ₅ OH	HB04B	OH09A	OH08A	-5.18	-4.71	-6.89		2.79	3.03	3.03	2.12	2.04			160.8	169.3	
7	H ₃ CCOOH...HOOCCH ₃ cyclic	HB10A	CA02A	CA02A	-17.10	-15.60	-7.25		2.57	2.79	3.27	1.83	2.29			175.9	164.0	
8	HOH...Methyl formate (=O)	HB21A	OH09A	CE01A	-6.39	-5.81	-7.09		2.57	2.79	3.29	1.83	2.31			175.9	164.5	
9	HOH...Methyl formate (-O-)	HB21B	OH09A	CE01A	-3.82	-3.47	-6.85		2.76	3.00	2.91	2.06	1.93			167	164.3	
10	<i>t</i> -NMA...OH ₂	HB05A	AM04A	OH09A	-5.97	-5.43	-3.80		2.87	3.12	3.06	2.12	2.26			138.7	139.2	
11	HOH... <i>t</i> -NMA	HB05B	OH09A	AM04A	-8.03	-7.23	-5.42		2.70	2.94	3.31	2.00	2.32			178.3	120.5	
12	Formamide dimer cyclic	HB06A	AM01A	AM01A	-14.8	-13.44	-4.57		2.76	3.00	2.92	2.00	1.90			167.6	167.1	
13a	Formamide dimer parallel	HB06B	AM01A	AM01A	-8.12	-7.38	-3.78		2.76	3.00	2.93	2.00	1.90			171.5	171.7	
13b									2.80	3.04	3.04	2.09	2.16			156.7	140.5	
14	H2NH...HNH2 cyclic C2h	HB07A	NH10A	NH10A	-3.51	-3.19	-3.36		x	3.41	3.18	2.64	2.33			134.9	138.7	
15	H2NH...NH3 linear Cs	HB07B	NH10A	NH10A	-3.38	-3.07	-4.04		3.01	3.27	3.27	2.60	2.58			123.9	123.9	
16	Imidazole...OH2	HB12A	AR09A	OH09A	-7.00	-6.37	-1.82		3.15	3.43	3.42	2.42	2.39			179.0	179.7	
17	HOH...Imidazole	HB12B	OH09A	AR09A	-7.77	-7.06	-3.25		2.79	3.03	2.91	2.03	1.88			69.8	57.7	
18	HOH...NH3	HB08A	OH09A	NH10A	-7.22	-6.56	-6.95		2.73	2.97	2.96	2.10	1.98			178.2	176.3	
19	HOH...Pyridine	HB17A	OH09A	AR02A	-6.63	-6.03	-6.19		2.80	3.04	2.83	2.09	1.82			151.2	165.5	
20	HOH...Thiophene	HB19A	OH09A	AR08A	-2.66	-2.42	-5.52		2.77	3.01	2.93	2.12	1.96			176.4	179.6	
21	HOH...Pyridine <i>N</i> -oxide	HB20A	OH09A	AR21A	-10.69	-9.72	-6.22		3.59	3.90	4.01	3.50	3.68			155.8	161.6	
22a	Methylethylamine <i>N</i> -oxide...OHH	HB45A	NH20A	OH09A	-15.48	-14.10	-10.42		2.63	2.86	3.08	1.94	2.15			108.4	102.2	
22b									2.49	2.71	x	1.83	x			121.0	123.4	
23	HSH...OH ₂	HB31A	SR01A	OH09A	-2.93	-2.66	-1.86		x	2.87	2.94	1.83	2.15			161.0	153.8	
24	HOH...S(CH ₃) ₂	HB32A	OH09A	SR03A	-3.56	-3.24	-3.25		2.63	2.86	3.08	1.94	2.15			148.9	x	
25	HSH...SH ₂	HB37A	SR01A	SR01A	-0.96	-0.88	-2.73		2.49	2.71	x	1.83	x			124.2	162.7	
26	HOH...SH ₂	HB38A	OH09A	SR01A	-2.27	-2.06	-1.93		3.33	3.62	3.68	2.29	2.34			175.3	177.9	
									3.12	3.39	3.95	3.13	3.58			97.7	104.5	
									4.15	4.51	4.25	3.19	3.71			175.5	105.0	
									3.47	3.77	3.69	2.83	2.81			175.5	146.6	

^a For representative drawings see Ref. [27]

sequence of energies of four different conformers of cyclohexanol was reproduced with respect to the ab initio values from MP4SDQ/TZP calculations but the absolute values of the differences are overestimated. The relative position of the ab initio energies for cyclopentanol could not be reproduced by COSMOS. The force field always yields a lower energy for the equatorial conformers than for an axial bonded hydroxyl group, but this is in accordance with experimental values for cyclohexanol [36]. The cited value was measured for 4-*t*-butylcyclohexanol epimers equilibrated via Raney nickel catalysis.

Intermolecular interactions of small molecules

Correct reproduction of proton donor–acceptor behavior is essential for description of molecular interactions. This task is not easily solved by force fields, as shown by Halgren [35], and some force fields fail to reproduce the correct donor–acceptor behavior.

Molecular systems with high polarization effects were selected for calculation of intermolecular interactions. We started with 26 small hydrogen-bonded systems (out of a set of 66 homogeneous and heterogeneous dimers) also available in Ref. [31]. The resulting interaction energies and the geometry parameters of the hydrogen bonds from ab initio [31] and COSMOS calculations are shown in Table 3. Representative drawings are given elsewhere [5].

In his work [35] Halgren uses scaled interaction energies and distances (SQM) for comparison. In COSMOS force field calculations, fully polarizable charges are always used. The COSMOS results compare better with the non-scaled ab initio data (QM). Both sets of parameters (QM and SQM) are shown in Table 3 for comparison. The SQM data were taken from Ref. [35] and the QM values were calculated from Ref. [31]. The names of the dimer structures (column 2) follow Ref. [35] and the codes of the dimers and monomers (columns 3–5) are the same as in Ref. [31]. The interaction energies are calculated according to the scheme: $\text{interaction energy} = \text{energy}(\text{dimer}) - \{\text{energy}(\text{monomer 1}) + \text{energy}(\text{monomer 2})\}$.

Water often plays a key role in the parametrization of force fields. The interaction energies of the pairs water and methanol (rows 1–4) are similar to each other. The energies are in the same order as the ab initio values. Although the relative sequence of the four ab initio interaction energies is not correctly reproduced, the COSMOS-calculated donor–acceptor behavior (rows 2 and 3) is the same as that found experimentally for the interaction between water and methanol. MSI CHARMM and Amber failed here [35].

The stabilizing energies of the dimers are overestimated relative to results from ab initio calculations (QM). The value for the water dimer lies between the scaled ab initio data (SQM), which reproduces the experimental value in the condensed phase, and the non-scaled ab initio values (QM), which represent the interaction energy of two isolated molecule in a vacuum. The

COSMOS interaction energy of $-5.9 \text{ kcal mol}^{-1}$ is lower than the stabilizing energy of the successfully used water models TIP3P [37] and SPC [38], which is necessary to fit the heat of vaporization ($-6.5 \text{ kcal mol}^{-1}$) and other condensed-phase properties. However, the COSMOS value is in good agreement with the non-scaled ab initio value ($-5.6 \text{ kcal mol}^{-1}$), which represents the interaction of two isolated water molecules in the gas phase.

The O...O distances (rows 1–4) behave similarly to the energies – the COSMOS value again lies between the SQM and QM data. The COSMOS-calculated distance for the water dimer (2.78 \AA) is shorter than the comparable non-scaled ab initio value (2.97 \AA), but it is larger than the value that the above-cited water models use to reproduce the density and other properties of liquid water ($2.74\text{--}2.75 \text{ \AA}$).

For the water–methanol interaction pairs the H...O distances are approximately 0.2 \AA shorter than the QM values. The O–H...O angles deviate further (approx. 5°) from the QM data.

Calculation for the test dimers produced the correct donor–acceptor behavior for the water–methanol pair (rows 1–4). Correct donor–acceptor behavior also could be reproduced for phenol (rows 5 and 6), methyl acetate (rows 8 and 9), *N*-methylamide (rows 10 and 11), imidazole (16 and 17), ammonia, pyridine, and thiophene (rows 18–20) and their interaction with water, where other force fields failed (see Ref. [35]).

The interaction energies and geometries of the cyclic ethanoic acid and formamide dimers deviate substantially from the ab initio data. For ethanoic acid the distances (O...O and O...H) are much too long, and connected with this, the interaction energies are much too small. The donor–acceptor behavior is, nevertheless, reproduced correctly for formamide (rows 12 and 13). A re-parametrization of conjugated C=O bonds (Eq. 3) could possibly yield improved results.

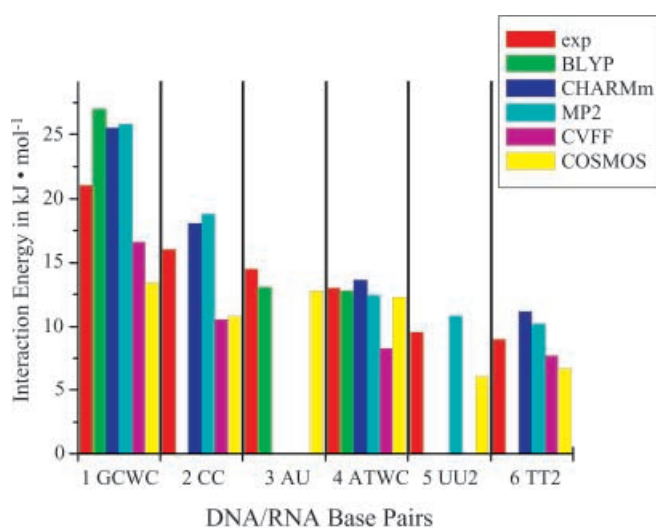
Differences among the interaction energies of the ammonia–water dimers (rows 14 and 15) are relatively small and the magnitudes are of the same order as ab initio results. Also the geometries obtained from the different methods are in good agreement. Unfortunately, COSMOS stabilizes the wrong dimer (linear more stable than cyclic) compared with the ab initio calculations, but the difference between the dimers is very small ($0.7 \text{ kcal mol}^{-1}$).

It seems very surprising to us that the interaction energies and geometries of sulfur-containing compounds also deviate only slightly from results from the ab initio calculations, because there are no parameters for the BPT charge calculation. In the BPT approach the atomic charge is a sum of contributions from all the bonds of that atom. The missing contributions, because of missing parameters, came from a simple method of calculating the atomic charges—charges derived from electronegativities. Deviations of the Coulomb energy and of the geometries should be expected.

The interaction energy of the thiophene–water dimer is overestimated relative to the ab initio values, but the

Table 4 Intermolecular interaction energies (kcal mol⁻¹) of DNA/RNA base pairs from different sources

Base pairs	Exp. [40]	ab initio BLYP [41]	CHARMm 23 [42]	ab initio MP2 [42]	CVFF [42]	COSMOS
GCWC	21	27	25.5	25.8	16.6	13.4
CC	16	x	18.1	18.8	10.5	10.8
AU ^a	14.5	13.1	x	x	x	12.8
ATWC	13	12.8	13.6	12.4	8.3	12.3
UU2 ^b	9.5	x	x	10.8	x	6.1
TT2	9	x	11.2	10.2	7.7	6.7

^a Geometry analogous to ATWC^b The ab initio values are, in this case, from Ref. [42]**Fig. 3** Comparison of experimental and calculated interaction energies for DNA/RNA base pairs. exp: I.K. Yanson et al. [40]; BLYP: D. Barsky, ab initio B3LYP/6-31++G** personal communication; CVFF, CHARMm: P. Hozba et al. [42]; MP2: M. Kratochvíl et al. (MP2/6-31G*) [42]; COSMOS: this work

interaction geometries calculated by different methods agree very well and the donor–acceptor interactions of ammonia and pyridine with water (rows 18–20) are correctly reproduced.

The interaction energies of hydrogen sulfide pairs and dimethyl sulfide pairs with water (rows 23–26) are of the same order as the ab initio values. Because of the above mentioned problems, reproduction of the relative interaction energies was not expected. Deviations are larger for the hydrogen sulfide dimer (row 25) than for the other pairs. A possible explanation is that for this dimer all charges were derived from electronegativities, which influences the Coulomb energy, and the Coulomb energy is almost always the largest part of the interaction energy.

For amine-N-oxides (rows 21 and 22) two different species are possible – a zwitterion with a high dipole moment on the one hand and an uncharged molecule with a lower dipole moment on the other hand. The calculations of the dimers yield much better results for the molecules with the lower dipole moment and for the monomers with the lower dipole moment. The interaction energies are similar to results from ab initio calculations and the relative stabilization energies could be reproduced.

COSMOS reproduced the correct (experimental) donor–acceptor behavior in all the cases tested, with one exception – the interactions of hydrogen sulfide with partners. The error can be explained by the missing BPT charge parameterization of the sulfur bond types. The reproduction of the experimental donor–acceptor behavior is remarkable compared with results from other force fields, even if deviations from the ab initio data (≥ 2 kcal mol⁻¹) are larger for the COSMOS interaction energies.

Intermolecular interactions of DNA and RNA base pairs

The interaction energies of DNA/RNA base pairs are of interest for molecular genetics. We are aware that a large part of the interaction energy within the DNA and RNA is determined by the stacking of the base pairs. In this work we concentrated on the interaction energies of single DNA/RNA base pairs as a first step. DNA and RNA chains are highly charged molecular structures. It has been shown that the results of force field calculations can be highly dependent on the charge model (see, e.g., Ref. [39]).

Table 4 and Fig. 3 show the interaction energies for selected DNA/RNA base pairs from experiment and from different methods of calculation. The selection was made by the availability of proven experimental data from one source [40]. The following base pairs (shown in Fig. 4) were selected: guanine–cytosine Watson–Crick-paired (GCWC), cytosine–cytosine (CC), adenine–uracil (AU), adenine–thymine Watson–Crick-paired (ATWC), uracil–uracil (UU), and thymine–thymine (TT).

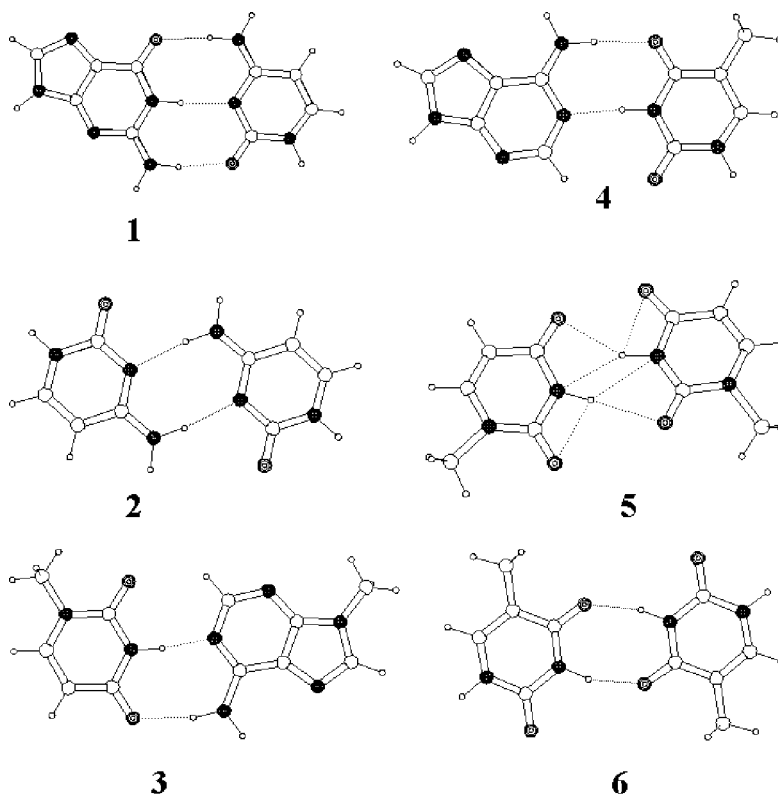
It should be noted that the experiments were performed with methylated bases. The methylations were made at those carbon atoms where the sugar component of the DNA/RNA chain is usually bonded. Methylated DNA/RNA bases were used in ab initio calculations by Barsky [41] (BLYP/6-31++G**) and in the COSMOS calculations. The ab initio MP2/6-31G*(0.25) with correction of the basis set superposition error (BSSE) and force field calculations (CHARMm and CVFF) from Šponer [42] were performed with non-methylated bases. The ab initio calculations of the base pairs from Šponer were made with fixed monomer geometries, and only the interaction parameters were determined by optimization.

Table 5 Comparison of COSMOS-optimized geometries of DNA/RNA base pairs with ab initio calculations and experimental data

Base pairs	Inter-acting atoms X–H...Z	X...Z distance (Å)					H...Z distance (Å)			X–H...Z angle (°)		
		Exp. [44]	Exp. [45]	ab initio [41]	ab initio [43]	COSMOS	ab initio [41]	ab initio [43]	COSMOS	ab initio [41]	ab initio [43]	COSMOS
GCWC	O6...H4–N4	2.91	2.98	2.79	2.92	3.15	1.75	1.91	2.13	179.5	177.0	166.5
	N1–H1...N3	2.95	2.95	2.94	3.04	3.12	1.90	2.03	2.09	178.1	176.1	172.4
	N2–H2...O2	2.86	3.03	2.92	3.02	3.00	1.90	2.02	1.97	179.7	178.1	177.6
CC	N4–H4...N3	x	x	x	2.92	2.89	x	1.89	1.85	x	173.6	176.5
	N3...H4–N4				2.92	2.96		1.89	1.93		173.6	173.9
AU ^a	N6–H6...O4	2.95	x	2.93	x	2.97	1.91	x	1.94	174.8	x	172.3
	N1...H3–N3	2.82		2.84		2.86	1.79		1.83	179.6		172.9
ATWC	N6–H6...O4	x	3.06	2.93	3.09	3.10	1.91	2.09	2.09	174.5	172.7	162.3
	N1...H3–N3		2.94	2.84	2.99	2.88	1.80	1.97	1.85	179.7	178.8	173.3
UU ^{2b}	N3–H3...O2	x	x	x	2.97	2.82	x	1.98	1.83	x	167.1	158.7
	O2...N3–H3				2.97	2.82		1.98	1.83		167.1	158.5
TT2	N3–H3...O2	x	x	x	2.98	2.75	x	1.99	1.76	x	167.4	157.8
	O2...N3–H3				2.98	2.76		1.99	1.78		167.4	158.2

^a Geometry analogous to ATWC^b The ab initio values are, in this case, from Ref. [42]**Fig. 4** DNA/RNA base pairs (for energies see Table 4)

1. guanine–cytosine Watson–Crick (GCWC),
2. cytosine–cytosine (CC),
3. adenine–uracil (AU),
4. adenine–thymine Watson–Crick (ATWC),
5. uracil–uracil (UU2),
6. thymine–thymine (TT2)



The ab initio calculations of Barsky were made without restrictions [41]. Since we found no proven experimental data relating to the CC, UU, and TT pairs, a problem appears – because the purine and pyrimidine bases have more than one donor and acceptor site, several variations of base-pairing are possible. For non-methylated bases there are even more possibilities. The only biologically relevant cases are those which are also possible if the molecule is methylated. There is only one possibility for CC. For TT and UU the biologically relevant pairs with

the lowest energy were chosen. The names of the base pairs are the same as in Refs [40] and [42]. Representative drawings are given elsewhere [42, 43].

COSMOS is able to reproduce the relative interaction energies satisfactorily, with the exception of the CC pair. The stabilization energy of this pair is much too small compared with the experimental and ab initio values. Within the COSMOS calculations the stabilization energy is largest for the GCWC pair, but this energy is too small compared with the ab initio values. As expected,

UU2 and TT2 have similar interaction energies. This is also reproduced by COSMOS, although the relative sequence of the stabilization energies is changed. In general, the COSMOS interaction energies are in the same range as the experimental values. Table 5 shows the length and angles of the hydrogen bonds from experiments (two different sources) and from different calculations.

Some trends in the results for the bond geometries are analogous with the interaction energies. With the exception of the CC pair the interaction geometries are quite well reproduced by COSMOS. Again it should be noted that the distances for GCWC deviate more from the experimental and ab initio values. One possible reason for deviation of the interaction energies are the too large distances of the interaction geometries for the CC and GCWC pair. Deviations of 0.2 Å in the X...Z distances from the ab initio values also occur in the TT2 structure, where again disagreement between the relative sequences of stabilization energies was found.

Differences between intermolecular hydrogen bridge distances obtained from ab initio calculations (geometry optimization basis set) and the COSMOS force field are, on average, approximately 0.06 Å. The largest differences are estimated for the GCWC and TT2 pairs. Differences between the angles (X-H...Z) of the intermolecular hydrogen bridges are relatively large – approximately 6° on average. This problem seems to be connected with the description of the hydrogen bond (only with the Coulomb and van der Waals potential) because no term in the potential controls the dependence of the strength of the hydrogen bond on the X-H...Z angle. Some modifications in that direction could substantially improve calculation of the interaction energies. On the other hand, the ab initio results are not at the limit of accuracy, because correlational corrections have a large influence on intermolecular interactions. Uncorrelated HF calculations disregard van der Waals attractions and should, therefore, give slightly incorrect distances and interaction energies.

Conclusions

The COSMOS force field is a powerful tool for calculating non-bonded intra- and intermolecular interactions. The differences between calculated and experimental bond lengths are approximately 0.03 Å; for bond angles the difference is ca. 3.5°. The main advantage of the COSMOS force field lies in description of the intermolecular interactions of polar systems. Differences between calculated and experimental interatomic distances (X...Z) for a polar system (X-H...Z) are approximately 0.2 Å. The interaction energies reproduce the correct experimentally determined donor–acceptor behavior for most of the examples tested.

The core of the new force field is semi-empirical charge calculation using bond-polarization theory. The bond polarization theory provides a fast method of calcu-

lating the atomic charges. The parameters are included for C–C, C–N, C–O, and P–O bonds of sp^3 and sp^2 hybrids and for the C–H, C–F, C–Cl, N–H, O–H, Si–H, Si–C, Si–O, Si–Cl, Zn–O, and Zn–N bonds of sp^3 hybrids. The model provides a good approach for inclusion of polarization effects of charged centers in a molecular system, and the influence of a solution or a lattice on a molecular system, in the calculation of the Coulomb term. The size of the molecular system studied is limited solely by CPU time. A typical calculation for a crystal lattice of 2000–3000 atoms, including net atomic charges, took only 2 min on a PC [46].

Acknowledgements This work was supported by the Deutsche Forschungsgemeinschaft. We wish to thank T.A. Halgren for providing valuable data via the internet. Thanks also to D. Barsky, who provided the results of high quality ab initio calculations.

References

- Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy Minimization and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187; MacKerell, A. D., Jr.; Bashford, D.; Bellot, M.; Dunbrack, R. L., Jr.; Evaseck, J. D.; Field, M. J.; Fisher, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnik, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179; Weiner, P. K.; Kollman, P. A.; Nguyen, D.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
- van Gunsteren, W. F.; Berendsen, H. J. C. Gromacs Molecular Simulation (GROMOS) Library Manual, Biomos, Groningen 1987.
- Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1995; Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.
- Allinger, N. L.; Zhou, X.; Bergsma, J. *J. Mol. Struct. Theorchem* **1994**, *312*, 69.
- Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.
- Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem.* **1999**, *103*, 4730.
- Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.
- Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, *108*, 4315.
- Rappé, A. K.; Goddard, W. A., III. *J. Phys. Chem.* **1991**, 3358.
- van Duin, A. C. T.; Baas, J. M. A.; van de Graaf, B. *J. Chem. Soc. Faraday Trans.* **1994**, *90*, 2881.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.
- Sternberg, U. *Mol. Phys.* **1988**, *63*, 249.
- Sternberg, U.; Koch, F.-T.; Möllhoff, M. *J. Comput. Chem.* **1994**, *15*, 524.
- Smith, A. Ph.D. Thesis, Darmstadt, 1989.

18. Dewar, M. J. S.; de Llano, C. *J. Am. Chem. Soc.* **1969**, *91*, 789.
19. Kroeker, M. Ph.D. Thesis, Darmstadt, 1994; Smith, A. E.; Lindner, H. J. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 235.
20. Sprague, J. T.; Tai, J. C.; Yuh, Y.; Allinger, N. L. *J. Comput. Chem.* **1987**, *8*, 581.
21. Pauling, L. *The Nature of the Chemical Bond*, Cornell University Press: Ithaca, NY, 1960.
22. O'Keefe, M.; Brese, N. E. *J. Am. Chem. Soc.* **1991**, *113*, 3226.
23. Root, D. M.; Landis, C. R.; Cleveland, T. *J. Am. Chem. Soc.* **1993**, *115*, 4201.
24. Cleveland, T.; Landis, C. R. *J. Am. Chem. Soc.* **1996**, *118*, 6020.
25. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. *J. Phys. Chem.* **1990**, *94*, 8897.
26. Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827.
27. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.
28. Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8262.
29. Sternberg, U.; Koch, F.-Th.; Bräuer, M.; Kunert, M.; Anders, E. *J. Mol. Mod.* **2001**, DOI 10.1007/s008940100007
30. Hannay, N. B.; Smith, C. P. *J. Am. Chem. Soc.* **1964**, *68*, 171.
31. http://server.ccl.net/cca/data/ff_evaluation_suite/
32. Llamas-Saiz, A. L.; Foces-Foces, C.; Mo', O.; Ya'ñez, M.; Elguero, E.; Elguero, J. *J. Comput. Chem.* **1995**, *16*, 263.
33. QUANTA 97, Molecular Simulations, Inc.; 9685 Scranton Road, San Diego, CA 92121-2777
34. Harmony, M. D.; Laurie, V. W.; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Lafferty, W. J.; Maki, A. G. *J. Phys. Chem. Ref. Data* **8** **1979**, 691; *Landolt-Börnstein, New Series, GroupII / 7*, Hellwege, K.-H. and A.M., Eds.; Springer, Berlin, Heidelberg, 1976; *Landolt-Börnstein, New Series, GroupII / 15*, Hellwege, K.-H. and A.M., Eds.; Springer, Berlin, Heidelberg, 1987.
35. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730.
36. Eliel, E. L.; Gilbert, E. C. *J. Am. Chem. Soc.* **1969**, *91*, 5487-5495.
37. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.
38. Berendsen, H. C. J.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; p 331.
39. Feig, M.; Pettitt, B. M. *J. Phys. Chem. B* **1997**, *101*, 7361.
40. Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.
41. Barsky, D., Lawrence Livermore National Lab. personal communication ab initio B3LYP/6-31++G** calculations
42. Hozba, P.; Hubálek, F.; Kabeláč, M.; Mejzlík, P.; Šponer, J.; Vondrášek, J. *Chem. Phys. Lett.* **1996**, *257*, 31; Kratochvíl, M.; Engkvist, O.; Šponer, J.; Jungwirth, P.; Hozba, P. *J. Phys. Chem. A* **1998**, *102*, 6921.
43. Šponer, J.; Leszczynski, J.; Hozba, P. *J. Phys. Chem.* **1996**, *100*, 1965.
44. Saenger, W. *Principles of Nucleic Acid Structure*, Springer, NY, Berlin, Heidelberg, London, Paris, Tokyo, 1983; Seemann, N. C.; Rosenberg, J. M.; Suddath, F. L.; Kim, J. R.; Rich, A. *J. Mol. Biol* **1976**, *104*, 109.
45. Spencer, M. *Acta Cryst.* **1959**, *12*, 59.
46. The authors provide free force field software which includes all routines for chemical shift and atomic charge calculations. Further information can be obtained by email from sternberg@ioq.physik.uni-jena.de