

Can We Accurately Describe the Structure of Adenine Tracts in B-DNA? Reference Quantum-Chemical Computations Reveal Overstabilization of Stacking by Molecular Mechanics

Pavel Banáš,^{*,†} Arnošt Mládek,[‡] Michal Otyepka,[†] Marie Zgarbová,[†] Petr Jurečka,[†] Daniel Svozil,^{‡,§} Filip Lankaš,^{||} and Jiří Šponer^{*,‡,⊥}

[†]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, tr. 17 listopadu 12, 771 46, Olomouc, Czech Republic

[‡]Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic

[§]Laboratory of Informatics and Chemistry, Faculty of Chemical Technology, Institute of Chemical Technology, Technicka 5, 166 28 Prague, Czech Republic

^{||}Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 6, 166 10 Prague, Czech Republic

[⊥]CEITEC - Central European Institute of Technology, Masaryk University, Campus Bohunice, 625 00 Brno, Czech Republic

S Supporting Information

ABSTRACT: Sequence-dependent local variations of helical parameters, structure, and flexibility are crucial for molecular recognition processes involving B-DNA. A-tracts, i.e., stretches of several consecutive adenines in one strand that are in phase with the DNA helical repeat, mediate significant DNA bending. During the past few decades, there have been intense efforts to understand the sequence dependence of helical parameters in DNA. Molecular dynamics (MD) simulations can provide valuable insights into the molecular mechanism behind the relationship between sequence and structure. However, although recent improvements in empirical force fields have helped to capture many sequence-dependent B-DNA properties, several problems remain, such as underestimation of the helical twist and suspected underestimation of the propeller twist in A-tracts. Here, we employ reference quantum mechanical (QM) calculations, explicit solvent MD, and bioinformatics to analyze the underestimation of propeller twisting of A-tracts in simulations. Although we did not identify a straightforward explanation, we discovered two imbalances in the empirical force fields. The first was overestimation of stacking interactions accompanied by underestimation of base-pairing energy, which we attribute to anisotropic polarizabilities that are not reflected by the isotropic force fields. This may lead to overstacking with potentially important consequences for MD simulations of nucleic acids. The second observed imbalance was steric clash between A(N1) and T(N3) nitrogens of AT base pairs in force-field descriptions, resulting in overestimation of the AT pair stretch in MD simulations. We also substantially extend the available set of benchmark estimated CCSD(T)/CBS data for B-DNA base stacking and provide a code that allows the generation of diverse base-stacking geometries suitable for QM computations with predefined intra- and interbase pair parameters.

INTRODUCTION

The DNA double helix plays a major role in biology by storing genetic information. The expression of genetic information is a complex process that is regulated by molecular interactions between DNA and other molecules, such as histones, transcription factors, and many others.¹ One of the key issues needed to understand the expression and regulation of genetic information at the DNA level is knowledge of how the DNA sequence is recognized and interpreted by all of the other molecules.

Reading of DNA does not just involve direct “digital” decoding of the primary base sequence. The very first atomic-resolution X-ray structures of B-DNA revealed that the B-DNA double helix is irregular, with sequence-dependent variations of the double helix structure and flexibility, which are crucial for molecular-recognition processes.^{2–4} This finding prompted major efforts to understand the sequence dependence of properties of B-DNA, as knowledge of the relationship and

predictability could have major impacts in biology, genetics, and biomedical sciences. Since the early research, base stacking has been the prime candidate for the B-DNA sequence dependence. The sequence dependence of B-DNA stems primarily from the existence of 10 independent base-pair (dinucleotide) steps and may be further modulated via non-nearest neighbor effects.^{2–22} A limited amount of success was initially achieved using qualitative steric considerations.^{7,8} Nevertheless, after years of research, key aspects of the sequence dependence of B-DNA structure, dynamics, and flexibility remain unresolved, due to their complexity and fragility.⁵ The phenomenon involves not only subtle (but biochemically very important) variations around the equilibrium B-DNA geometry but also sequence-dependent properties of deformed B-DNA in various molecular complexes and molecular recognition processes.⁶

Received: February 9, 2012

Published: May 17, 2012



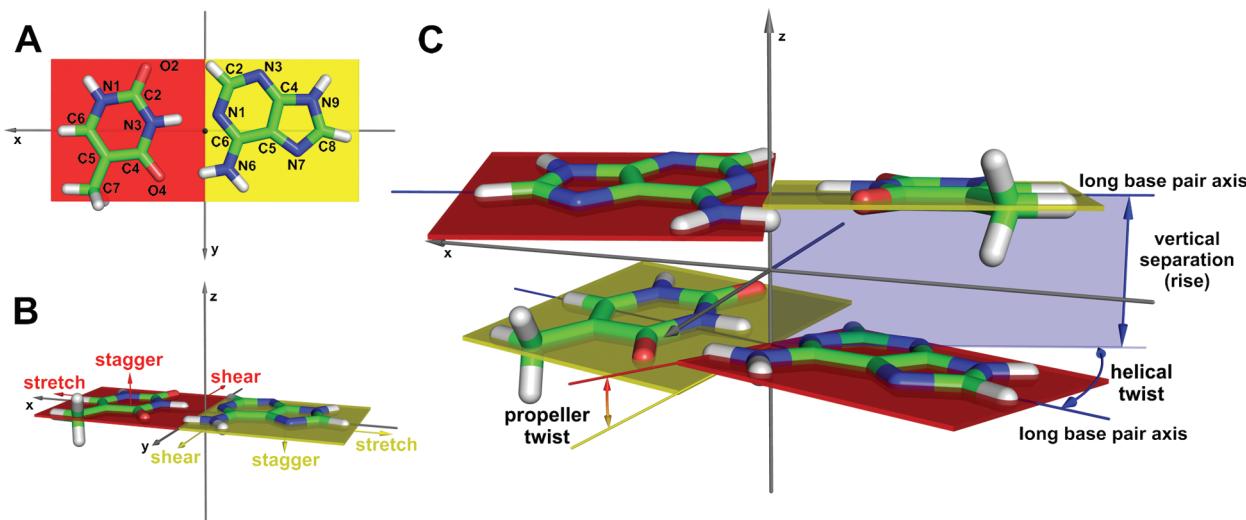


Figure 1. (A) The reference frame for an AT base pair with highlighted axes and atom numbering. (B) The scheme for base-pair translation parameters (shear, stretch, and stagger). (C) The idealized model geometry of an ApT base-pair step with highlighted propeller twist, helical twist, and vertical separation parameters.

There have been many attempts to provide theoretical predictions of the sequence dependence of B-DNA. The calculations so far have either exclusively considered base stacking or attempted to include the whole DNA structure. Base-stacking calculations have been typically carried out using molecular mechanics (MM) empirical potentials.^{9–11} Later on, some quantum-chemical (QM) calculations were attempted,^{13,14,16,21–28} although the QM technique has primarily been used to clarify the nature of base stacking and verify/parametrize other methods. The main obstacles (compared to MM) in QM studies of local conformational variability of DNA are the very limited sampling and inability to separate the electrostatic interactions, which are sensitive to solvent-screening effects and thus exaggerated by *in vacuo* computations. In the Supporting Information, we provide further details of the limitations of theoretical methods in studies of stacking and DNA local variations.^{13,29} Nevertheless, since the theoretical models should ideally consider a complete B-DNA molecule in its native environment and the sequence-dependence also includes thermal fluctuations, a molecular dynamics (MD) simulation with an explicit solvent is an ideal method for studying the sequence-dependence of B-DNA.^{21,22} To derive useful information, the simulations must be sufficiently long and the energy description sufficiently accurate. For example, a recent parmbsc0³⁰ version of the Cornell et al. force field³¹ eliminated spurious γ -trans backbone conformations gradually degrading the simulated B-DNA architectures.^{19,30–33} This development has enabled sufficiently long B-DNA simulations (see Supporting Information for details).

Among the DNA local variations, a major sequence-dependence effect is caused by phased adenine tracts, i.e., stretches of several (at least four) consecutive adenines in one strand without an intervening TpA step (typically 5'-A_n-3' or 5'-A_nT_m-3') that are regularly spaced every one full B-DNA turn.^{34–37} Such phased A-tracts induce a global bend in the DNA helix, with a magnitude of 17–21° per A₆ tract.³⁸ The A-tracts seem to be fairly rigid, with the consequence that they impair the formation of nucleosomes and thus regulate the nucleosome arrangement on the genome.³⁹ A-tracts also most likely contribute to DNA packing in bacteria lacking nucleosomes.⁴⁰ Valuable initial insights into A-tract properties

and DNA bending have been obtained in earlier simulations,^{34,35,41–44} although the conclusions might have been partially limited by the relatively short length of the simulations and force-field quality (see Supporting Information). The latest simulation study on A-tracts based on long parmbsc0 simulations, while achieving meaningful agreement with most experimental data on DNA bending, reported underestimation (in absolute values) of the propeller twist of the simulated A-tracts.³⁷

Propeller twist is an important conformational parameter of a DNA molecule. It describes the angle of counter-rotation of the two bases in a pair about the long base pair axis (Figure 1).⁴⁵ Note that since the sign of propeller twisting is negative, in this paper, all discussions of the magnitude of propeller twist consider its *absolute* value. A large absolute value of propeller twisting is an important signature feature of B-DNA A-tracts related to their specifically narrow minor groove.^{46,47} The propeller twist in the X-ray structures of sufficiently long A-tracts^{47,48} exceeds –20° and reaches ~–25° in some cases. In contrast, in our long MD simulation of a DNA oligomer containing a closely related tract, namely A₄T₄,³⁷ the mean propeller twist did not exceed –19° for any AT pair, and thus it seems likely that the A-tract propeller twisting is underestimated in contemporary atomistic simulations.³⁷ It should nevertheless be pointed out that the limited number and resolution of A-tract X-ray structures is responsible for some of the uncertainty in the interpretation of the experimental data (see Supporting Information and Discussion and Conclusions for more details).

The underestimation of propeller twisting in A-tract B-DNA simulations may be due to numerous approximations, such as too steep Lennard-Jones r^{-12} repulsion^{49,50} and parametrization of amino groups that rather rigidly mimic purely planar sp^2 hybridization.^{17,51–54} The latter approximation can affect, e.g., bifurcated cross-strand A(N6)…T(O4) hydrogen bonds between consecutive base pairs in the major groove of ApA steps, stabilization of the close contacts of A(N6)…A(N6) amino groups^{54–56} in the major groove of the intervening ApT base-pair steps, and tolerance of large propeller twists within an AT pair, as flexible amino group hydrogens can more smoothly follow the T(O4) acceptor atom in the AT base pair itself.

Another factor could be an insufficiently accurate description of the network of structural water molecules in the minor groove mediating cross-strand contact between the O2 carbonyl of thymine and N3 nitrogen of adenine in consecutive base pairs. In addition, propeller twisting can be underestimated due to many other reasons not related to base–base interactions, such as the accuracy of the MM description of the glycosidic torsion and sugar pucker.⁵⁷

In the present study, we carried out a series of investigations to understand the underestimation of propeller twisting of A-tracts in simulations better. We first scanned the stacking energy between base pairs in ApT, TpA, and ApA base-pair steps and the base-pairing energy in the AT pair by varying the helical twist and propeller twist while using the high-quality SCS(MI)-MP2/cc-pV(D,T)Z QM method. Further, CBS(T) single-point energies were calculated for the most important geometries. CBS(T) stands for MP2/Complete basis set calculations corrected by the CCSD(T) method with a smaller basis set. The benchmark QM computations were then compared with two empirical potential models: (i) the conventional 6–12 Lennard-Jones form and (ii) exponential repulsion in combination with a damped dispersion term. We also analyzed helical parameters in explicit solvent molecular dynamics (MD) of A-tracts and monitored the stacking energy of base-pair steps obtained from MD using the RI-TPSS-D/TZVP QM method.¹⁶ Although we do not find any obvious reason why the force field seems to underestimate A-tract propeller twisting, we report several factors that are important for the performance of nucleic acids simulations. Namely, we demonstrate overstabilization of stacked structures in contemporary simulations that appears to stem from substantial anisotropy of polarizabilities not captured by the isotropic force fields.

METHODS

Preparation of Structures for Benchmark QM Computations. Structures of AT base pairs and ApA, ApT, and TpA base-pair steps (stacks of two consecutive AT base pairs) with different geometrical parameters, such as propeller twist, helical twist, vertical separation, shear, stagger, and stretch, were prepared with in-house software (<http://fch.upol.cz/en/software/>) using an approach similar to that in ref 10. Our program is able to generate base-pair-step geometries from given helical parameters and reference-frame geometries in a similar fashion to, e.g., the 3DNA software.^{58,59} In addition, our program calculates MM energies and performs MM relaxation of given helical parameters using either a standard 6–12 van der Waals term or damped 6-dispersion term complemented with exponential repulsion (see section on MM calculations). The base-pair and base-pair-step structures were derived from an AT base-pair reference structure optimized in the gas phase at the MP2/cc-pVTZ level with C_s symmetry (see Supporting Information for more details about the algorithm).

Explicit Solvent MD Simulations of A-Tracts and Preparation of Base-Pair Steps for Fast DFT-D Computations. To increase the sampling of stacking compared to idealized scans, we also used A-tract geometries obtained from previously published simulations of A_nT_n and T_nA_n tracts,³⁷ namely, the $d(GGCA_4T_4GCC)_2$ and $d(CCGT_4A_4CGG)_2$ systems referred to in the original study as AT_short and TA_short (see ref 37 for the simulation setup). To assess the influence of DNA dynamics on the stacking energies of ApT, TpA, and ApA base-pair steps, we utilized the previously

published protocol^{21,22} for selecting representative geometries from MD simulations. For each base-pair step, we generated in total 250 structures, each averaged over 400 ps, by using the PTraj module of AMBER 9.⁶⁰ Subsequently, the sugar-phosphate backbone was removed, and the force-field nucleobases were replaced by a reference geometry of monomers of the bases optimized separately at the RI-DFT-D/TPSS/TZVP level with C_s symmetry, i.e., assuming planarity of the amino groups. The sugars were replaced by hydrogens.

Additional Simulations. To extend the set of the simulations and enhance the statistics of the AT base-pair stretch and hydrogen-bonding distances within AT-base pairs, we performed new simulations of the A_3T_3 ($d(CGCA_3T_3GCG)_2$, starting structure PDB ID 1S2R) and A_6 (sequence $d(CCGGA_6CG)_2$, PDB ID 1D89) tracts. These simulations were carried out in AMBER 11⁶¹ using the GPU CUDA part of the code with the parmbsc0 force field,³⁰ net-neutralizing sodium ions employing Joung and Cheatham parameters for the sodium ion,⁶² and TIP3P and/or SPC/E water models.^{63,64} The systems were equilibrated using previously published protocols.^{57,65–68} See the Supporting Information for a comparison of mean propeller twists of AT base pairs obtained with these MD simulations with the corresponding experimental values.

QM Calculations. Interaction energies were calculated at three different levels of QM theory: RI-DFT-D, SCS(MI)-MP2, and CBS(T). Stacking and pairing energies calculated within the potential energy surface scans using the idealized geometries were calculated with the spin-component-scaled MP2 method optimized for molecular interaction (SCS(MI)-MP2)⁶⁹ with cc-pV(D,T)Z basis-set extrapolation according to Helgaker et al.⁷⁰ (see Supporting Information for more details). The stacking and pairing energies were corrected for basis set superposition error (BSSE) using the counterpoise procedure.⁷¹ Note that the base-pair-step stacking energy was calculated as a difference between the energy of the tetramer and the sum of dimer (base pair) energies. Thus, the SCS(MI)-MP2 stacking energies involve many-body terms. The SCS(MI)-MP2 calculations were performed using the Gaussian 03 program.⁷²

In order to validate the SCS(MI)-MP2 method, we evaluated a subset of the structures using the standard benchmark CBS(T) method, i.e., the MP2/CBS method corrected for higher-order correlation effects using CCSD(T) with a smaller basis set, namely aug-cc-pVDZ and cc-pVDZ basis sets for heavy atoms and hydrogens, respectively (see the Supporting Information for further details). The MP2/CBS energies were estimated according to Helgaker et al.^{70,73} with aug-cc-pVDZ and aug-cc-pVTZ basis sets.^{74,75} The density-fitting (DF) approximation⁷⁶ was employed. It should be noted that CBS(T) represents estimated CCSD(T) interaction energies extrapolated to CBS rather than exact CCSD(T)/CBS results. Both the MP2 and CCSD(T) calculations involved in the CBS(T) calculations were carried out with the MOLPRO 2006.1 software.⁷⁷ The pairwise interaction energies were corrected for BSSE using the counterpoise scheme.^{71,78}

On the basis of recent estimates,⁷⁹ we assumed that due to a compensation of errors (underestimation of the MP2/CBS attraction and Δ CCSD(T) repulsion with the basis sets specified above), our reference CBS(T) computations should be very close to the converged values and thus represent the most accurate energies for B-DNA stacking published so far. We recalculated the stacked UU and AT dimers (optimal gas-

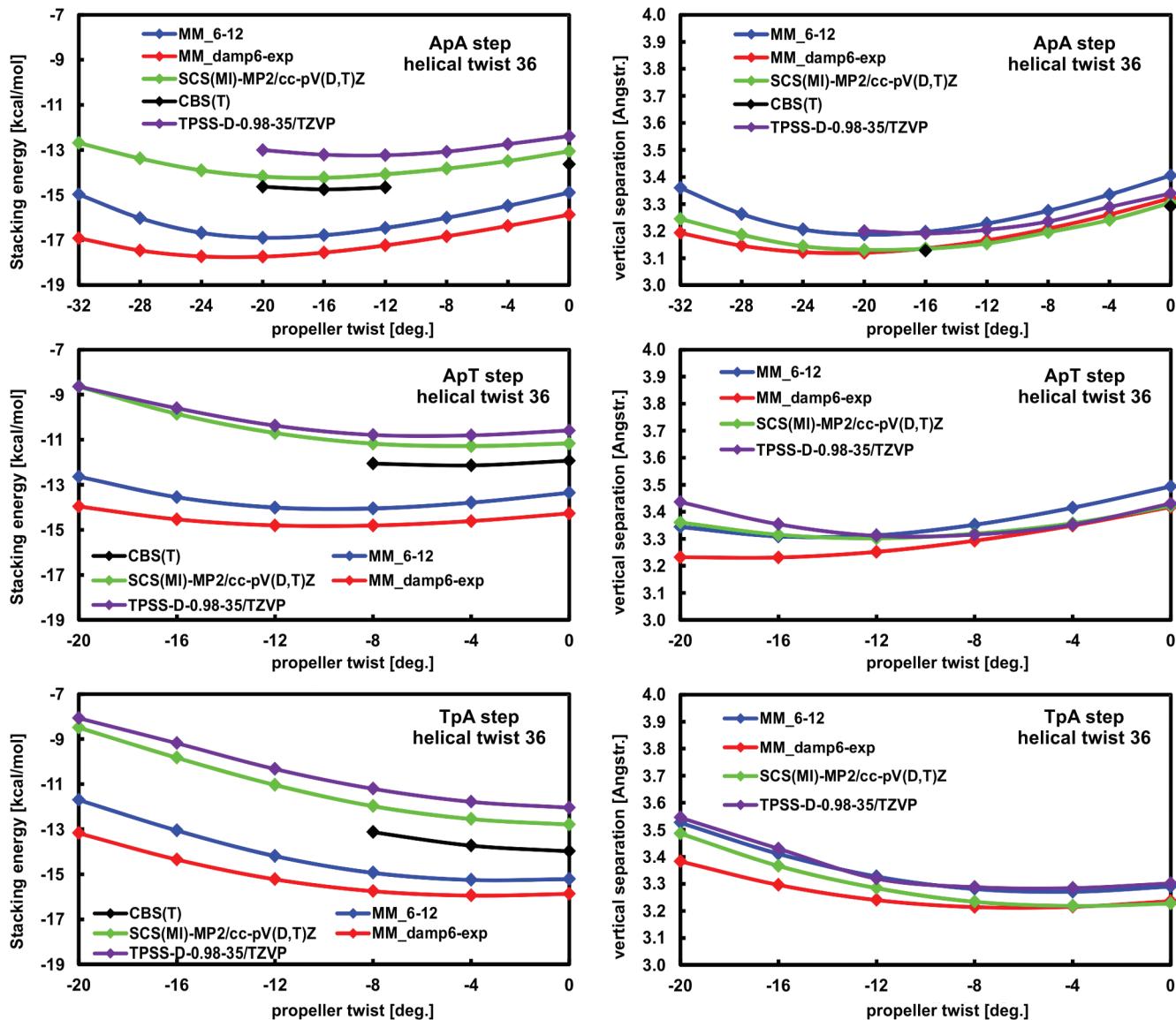


Figure 2. Plot of the stacking energy (left) and optimal vertical separation (right) as a function of propeller twist in ApA, ApT, and TpA base-pair steps. MM_6-12 (blue line) and MM_damp6-exp (red line) correspond to the MM energy (both electrostatic and van der Waals) using the standard 6–12 van der Waals term and exponential repulsion in combination with a damped dispersion term, respectively. SCS(MI)-MP2/cc-pV(D,T)Z (green line), TPSS-D-0.98-35/TZVP (violet line), and CBS(T) (black data) correspond to QM methods (see the Methods section). In this case, the vertical separation is equivalent to the vertical distance between the two base pair C6–C8 axes (see the Supporting Information).

phase structures) from the S22 database (structures 13 and 15)⁸⁰ using the CBS(T) level detailed above and obtained values of -9.68 and -11.63, which are almost identical to the most accurate results for these two geometries reported to date, i.e., -9.81 and -11.73.⁷⁹ Therefore, our CBS(T) computations appear to be more accurate than those originally used for the S22 and JSCH2005 (containing B-DNA data) data sets.^{80,81} For comparison, the respective UU and AT stack SCS(MI)-MP2 values were -9.42 and -12.00. Note, we use the CBS(T) abbreviation instead of CCSD(T)/CBS since we strictly avoid using the latter notation for data that are not true CCSD(T)/CBS extrapolations.

As the CBS(T) stacking energies had to be derived as a sum of four base/base stacks, the many-body terms are absent. However, the three base-pair steps studied in the present work show negligible nonadditivity of stacking.⁸¹ Thus, the difference in calculations does not affect the results.

To increase sampling of the mutual orientation of the bases within the base pair and the base pairs within the base-pair step, we used a set of structures generated on the basis of averaged snapshot geometries obtained by MD simulations (see above) and calculated the stacking and pairing energies using the efficient RI-TPSS/TZVP DFT method augmented by empirical dispersion according to Jurecka et al.⁸² The RI-DFT-D energies were not corrected for BSSE, as this method has been parametrized to include such a correction inherently.⁸³ The DFT-D stacking energies were calculated so that they included many-body terms. The DFT-D calculations were performed using the TurboMole 5.8⁸⁴ software package, in combination with a local code that computes the empirical dispersion correction (see <http://fch.upol.cz/en/software/>).⁸² For the DFT part of the calculation, the RI approximation,^{85,86} also known as density fitting, was used.

MM Calculations. Stacking and pairing energies calculated within the potential energy surface scans using the idealized geometries (geometries used in SCS(MI)-MP2 and CBS(T) calculations) were also calculated using the empirical potential to assess the performance of the force-field approximation with respect to the benchmark QM data. Within the force-field approximation, only electrostatic and van der Waals interactions are included.

In our calculations, we did not use HF-derived point charges that are typically included in force fields, such as the HF/6-31G(d) RESP charges applied in the Cornell et al. force field.³¹ HF charge distributions are intentionally overpolarized for simulations in water. Instead, we used MP2/cc-pVTZ ESP charges so that the MM Coulombic term was as close as possible to the electrostatic interactions in the reference QM calculations. Such charges are then consistent with the QM computations of interaction energies, as explained in detail in ref 13. Thus, the differences between QM and MM stacking and pairing energies can, as a first approximation, be interpreted as an inaccuracy of the van der Waals force-field term.

We used two different forms of the van der Waals force-field term: (i) the standard Lennard-Jones 6–12 potential (eq 1)

$$E_{6-12\text{VdW}} = \sqrt{\epsilon_i \epsilon_j} \left[\left(\frac{R_i + R_j}{r_{ij}} \right)^{12} - 2 \left(\frac{R_i + R_j}{r_{ij}} \right)^6 \right] \quad (1)$$

and (ii) an exponential repulsion term in combination with a damped dispersion term (eq 2)

$$E_{6-12\text{VdW}} = \sqrt{\epsilon_i \epsilon_j} \left[\frac{6}{\xi - 6} \exp \left(\xi \left(1 - \frac{r_{ij}}{R_i + R_j} \right) \right) - \frac{\xi}{\xi - 6} \frac{(R_i + R_j)^6}{r_{ij}^6 + \left(\frac{R_i + R_j}{\alpha} \right)^6} \right] \quad (2)$$

In both van der Waals forms, the R and ϵ parameters denote van der Waals radii and well depths, respectively, and were taken from the Cornell et al. force field.³¹ The formula for the exponential repulsion term was taken from the DREIDING force field,⁸⁷ where the scaling parameter $\xi = 12$ was used to retain the long-range dispersion interaction. It is possible to use our in-house program for calculating stacking energies and/or base-pairing energies of both base pairs and all 10 base-pair steps with an exponential force field (see <http://fch.upol.cz/en/software/>). In addition, we corrected the dispersion term with a damping parameter α , which is commonly used in soft-core potentials. We used α equal to 3, so that the dispersion term was only damped by approximately 0.1% in van der Waals minimum, which has a negligible effect on the presented interaction energies. However, such damping is crucial for stability in MD simulations with an exponential repulsion term.

RESULTS

Stacking Energy of Base Pairs in Base-Pair Steps. In the first set of calculations, we used idealized model geometries of ApA, ApT, and TpA base-pair steps. We performed a potential energy surface scan of helical- and propeller-twist parameters. The value of the vertical separation was optimized for each combination of helical and propeller twists by scanning

the vertical separation using 0.001 Å steps (for MM) or 0.1 Å steps together with a three-point quadratic interpolation scheme (for QM). The vertical separation was optimized in each method with the exception of the CBS(T) calculations, where the optimal vertical separation was explicitly evaluated under just two conditions (ApA step with a helical twist of 36° and a propeller twist of 0° or -16°, Figure 2), while the optimal value of vertical separation obtained by SCS(MI)-MP2/cc-pV(D,T)Z was used elsewhere. Figure 2 summarizes the stacking energies of all three studied base-pair steps for the most relevant helical twist angle of 36°. Complete data involving propeller twist scans at helical twist angles of 30°, 36°, and 42° are shown in Figures S1–S4 and Tables S1–S4 in the Supporting Information. The Supporting Information also includes all geometries of the structures for which the new reference CBS(T) B-DNA stacking energies were calculated. The other geometries can be easily generated using our in-house program and thus are not included in the Supporting Information (see the Methods section for details and <http://fch.upol.cz/en/software/> for a download of the program).

Comparison of QM and MM Reveals Overstabilization of Stacking by MM. A comparison of QM stacking energy profiles with force-field computations revealed several interesting results.

- (i) The stacking energy obtained using MM with the standard 6–12 van der Waals term (i.e., gas-phase stacking calculated using the AMBER Cornell et al. force-field variant with MP2 point charges) was clearly overstabilized (cf. green vs blue energy curves in Figure 2 and Supporting Information Figures S1–S3). Since the overstabilization is seen for all steps and all values of helical twist, it appears to be a feature of the van der Waals term of the force field, as also suggested in ref 49. Switching to the exponential repulsion term stabilizes the stacking even more by an additional 0.6–2.1 kcal/mol (Figure 2). Although the comparison of MM with the true CBS(T) benchmark (black energy curves) reveals subtly smaller overstabilization, the effect remains significant. The results suggest that the force field is responsible for the overstabilization of intrinsic stacking in simulations. However, they do not necessarily imply equivalent overestimation of stacking free energies in aqueous environments, since these would be very dependent on the solvation energies. However, overstabilization of stacking has been suggested to occur in explicit solvent simulations of diagonal four-thymidine loops of d(G₄T₄G₄)₂ quadruplex DNA, forcing the loops into incorrect topologies.^{88,89}
- (ii) The exponential repulsion term improves the optimal value of vertical separation, which is overestimated by ~0.1 Å within the 6–12 van der Waals empirical description (Figure 2). This difference could potentially affect the results of simulations, although no data are presently available to assess if the effects are significant. The optimal vertical separation is improved by including the exponential repulsion term, particularly for low propeller twist values, whereas highly propeller-twisted structures of some base-pair steps, e.g., ApT, are well described by the standard 6–12 van der Waals potential (Figure 2). Therefore, neither form of the force field is able to fully capture the effect of propeller twist on vertical separation.

The vertical separation is related to a parameter known as rise in codes describing the double helix architecture (see the Methods section and Supporting Information for our definition of vertical separation). Numerical values of rise calculated by different codes may differ substantially, depending on its exact definition in a given code. However, we found that at least in the geometries studied here, the vertical separation obtained using our program matches the corresponding rise parameter calculated using the 3DNA code.^{58,59} Note that the optimization of vertical separation is crucial for a correct description of stacking, as any vertical compression or extension would be associated with a large van der Waals energy gradient. Performing computations with unrelaxed vertical separation would seriously compromise the resulting potential energy scans.^{10,11}

- (iii) Another interesting issue concerns whether a change of the force-field functional form can increase the propeller twisting, mainly for the ApA step, which is the most crucial one for A-tract architecture. Some differences were observed between the various descriptions, especially in the position of the energy minima. However, the profiles were relatively flat, and thus we suggest that such differences are likely to be insignificant (Figure 2 and Figures S1–S3, see also Table 1 for comparison with

Table 1. The Minimal and Maximal RI-DFT-D/TPSS/TZVP Stacking Energies (kcal/mol) of ApT, TpA, and ApA Base-Pair Steps among the Geometries Obtained from 400 ps Averaged Segments (250 points) of the 100 ns Long MD Simulations (see the Methods Section)

base-pair step	minimum energy	maximum energy
ApT	−11.74	−10.44
TpA	−12.69	−8.32
ApA	−12.34	−9.90

the range of thermal fluctuation of stacking energy). In other words, we suggest that the present data do not indicate that underestimation of the propeller twist in simulations is caused by a deficient description of base stacking by the force field.

- (iv) A potentially important QM vs MM difference can be seen by examining the dependence of the stacking energy on the helical twist angle. The reference QM method penalizes the helical twist of 30° in ApA and TpA base-pair-step geometries in comparison with the corresponding structures having a helical twist of 36° more strongly than both empirical potentials (see Figure S4). Thus, the empirical potentials used had a slightly reduced tendency to adopt higher values of the helical twist angle compared to QM data. Although this difference is rather small, it is significant at low propeller twists and might be amplified in a longer helix, as it should be multiplied by the number of base-pair steps. Since reduction of the helical twist leads to an increase in the stacking overlap, the observed difference between QM and MM helical twist energy dependence may stem from the general over-stabilization of stacking by MM.

DFT-D Computations on MD-Derived Structures. In order to evaluate fluctuations of stacking energy during thermal motions at finite temperature, we used MD simulations to generate 250 geometries of ApA, ApT, and TpA base-pair steps.

Subsequently, the stacking energies in these base-pair steps were calculated using the computationally efficient DFT-D/TPSS/TZVP method. Table 1 summarizes the minimum and maximum stacking energies for each base-pair step. The averaged, minimum, and maximum values of selected geometrical parameters through the MD trajectory are documented in Table 2. The geometries for the ApA step were taken from the ApA step between the second and third AT base pairs in the A_4T_4 simulation, i.e., ggcaAAatttgcc, highlighted by capitals.

Table 2. The Minimum, Maximum, and Mean Values of Helical Parameters of ApT, TpA, and ApA Base-Pair Steps among the Geometries Obtained from 400 ps Averaged Segments (250 points) of the 100 ns Long MD Simulations (See Methods Section) Calculated Using the 3DNA Code^a

base-pair step	average value	minimum value	maximum value
helical twist (deg)			
ApT	32	29	34
TpA	26	9	42
ApA	35	25	42
propeller twist (deg)			
ApT	−17/−17	−23/−23	−8/−9
TpA	−10/−10	−19/−18	−1/2
ApA	−17/−19	−24/−23	−6/−12
rise (Å)			
ApT	3.31	3.18	3.45
TpA	3.22	2.79	3.81
ApA	3.42	3.23	3.65

^aThe propeller twists of both individual AT base pairs are reported.

The results shown in Tables 1 and 2 demonstrate that the QM reference computations on idealized geometries more or less cover the range of main conformational parameters sampled by simulations, although this does not mean the geometries are identical, since the simulations also sample all of the remaining degrees of freedom (see the Supporting Information for a more detailed discussion of the structural and energy ranges observed in these calculations).

Comparison with Benchmark Energies. As noted above, for a subset of geometries, we calculated the benchmark CBS(T) (estimated CCSD(T)/CBS) energies. The SCS(MI)-MP2/cc-pV(D,T)Z energies agreed well with the reference CBS(T) data, although the stacking energy was slightly underestimated compared to CBS(T). The error in the SCS(MI)-MP2/cc-pV(D,T)Z method varied only negligibly with the propeller twist and was almost uniform even among different steps, with the best accuracy obtained in the case of the most important ApA step. For the ApA step, we also tested the accuracy of the predicted optimal vertical separation and found a close to perfect agreement between the CBS(T) and SCS(MI)-MP2/cc-pV(D,T)Z calculations. Taken together, the SCS(MI)-MP2/cc-pV(D,T)Z method provides a reasonable estimate of the stacking energy and is sufficiently accurate to provide benchmark data for the force fields, especially for the description of the shape of the potential energy surface. In addition, we found that the DFT-D method is sufficiently accurate to monitor fluctuations of stacking energies along MD trajectories, although it is slightly less accurate compared to SCS(MI)-MP2/cc-pV(D,T)Z (see the Supporting Information for detailed comparison of SCS(MI)-MP2 and DFT-D methods with reference CBS(T) data). It should be pointed

out that in drawing the conclusion about overestabilization of stacking by the force field, it is important to compare the stacking data for the different force fields (blue curve, Figure 2 left) with the true benchmark, i.e., the QM CBS(T) (black, Figure 2 left) data. The force-field overestimation of stacking is very visible for all studied base-pair steps. This overestimation, together with the QM vs MM data for base pairing (see below), indicates an imbalance in the force-field description. Assessment of other QM methods is a separate issue, which does not have any bearing on the force-field assessment.

Underestimation of Intrinsic Stability and Spurious Steric Clash in MM Description of AT Base Pair. The QM computations of base stacking did not reveal any clear reason why the force field would underestimate propeller twist, although the calculations clearly showed overestabilization of stacking by MM. Thus, we carried out reference computations on the AT base pair itself to determine whether its intrinsic properties may restrict the degree of propeller twist.

We used a set of idealized model geometries to scan the influence of the propeller twist on base-pairing energy within an AT base pair. In contrast to the stacking energy of base-pair steps, where we assumed that relaxation of the vertical separation was sufficient to provide meaningful data, the situation in the case of base pairing was much more complex. More parameters have to be taken into account to keep the scan within the low-energy region. For the sake of simplicity and to avoid problems with noncommutating rotation of bases within the AT base pair, we relaxed only spatial translation parameters (stretch, stagger, and shear) during the propeller twist scan (see Figure 1 for definitions of base-pair parameters). We found that the optimal shear varied between -0.006 and 0.003 \AA and thus virtually did not change with the propeller twist. We therefore fixed the shear at 0 \AA . Finally, only stagger and stretch parameters were relaxed for each propeller twist value. In the case of QM methods, we relaxed both stretch and stagger by scanning using 0.1 \AA steps combined with a quadratic interpolation scheme. In the case of empirical potentials, we relaxed the stretch and stagger values by scanning these parameters using 0.001 \AA steps. Opening and buckle degrees of freedom were not optimized in our study and were fixed at a value of 0° corresponding to the QM-optimized geometry of the AT base pair.

Like the stacking energy calculations, the SCS(MI)-MP2/cc-pV(D,T)Z and CBS(T) calculations provided very similar values for the interaction energies, as well as optimal stretch and stagger values. The much less expensive SCS(MI)-MP2/cc-pV(D,T)Z method was thus used as the benchmark for the force fields.

We obtained the following main results (Figure 3):

- In contrast to stacking energies, the base-pairing energy is significantly underestimated by both MM methods. This result is not unexpected, as the base-pairing energy should include some polarization energy, which is not included in the force-field approximations. However, at least in part, this can also be explained by the anisotropy of the dispersion interaction (see the Discussion section).
- While the QM method does not show any dependency of stagger on propeller twist, the propeller twisting can support large negative values of stagger under the empirical potential treatment. The negative stagger has the same effect as shifting the axis of propeller twist rotation toward the major groove, so that the heavy

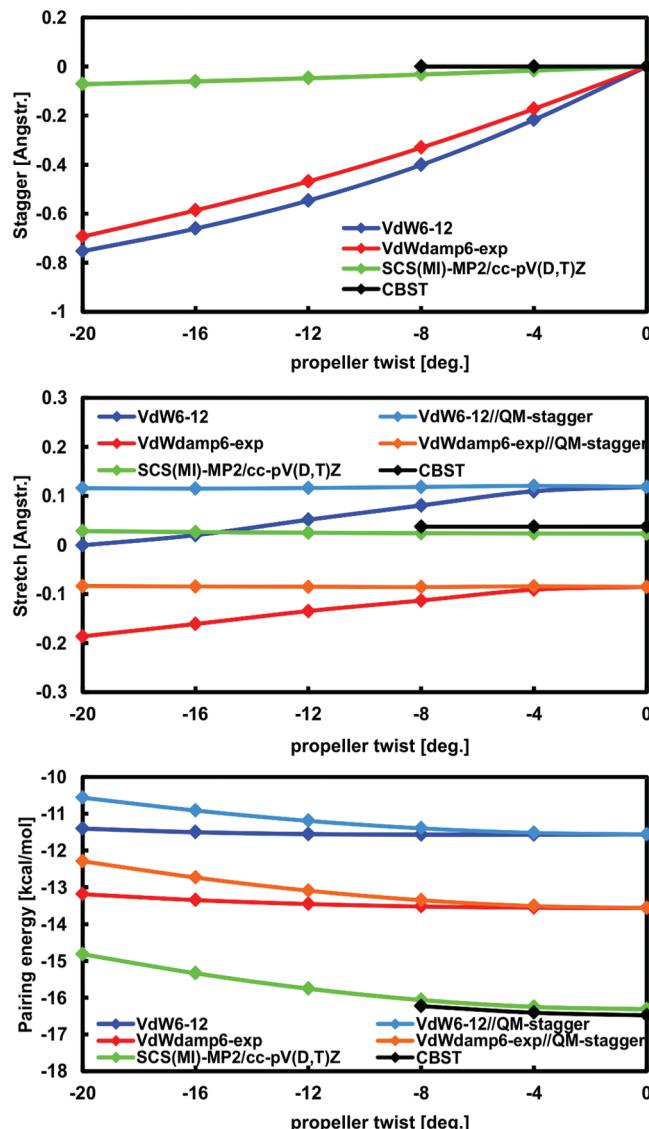


Figure 3. Plots of optimal stagger (top), optimal stretch (middle), and AT base-pairing energy as a function of propeller twist. VdW6-12 (blue line) and VdWdamp6-exp (red line) correspond to the empirical potential curves (both electrostatic and van der Waals) obtained by using the 6–12 Lennard-Jones potential and exponential repulsion with damped dispersion terms, respectively. SCS(MI)-MP2/cc-pV(D,T)Z (green line) and CBS(T) (black line) correspond to the QM methods. The curves labeled VdW6-12//QM-stagger (light blue line) and VdWdamp6-exp//QM-stagger (orange line) are the corresponding MM profiles obtained by minimization of the stretch for a given propeller twist while assuming the optimal stagger obtained from the SCS(MI)-MP2/cc-pV(D,T)Z QM scan. All calculations were performed with rigid monomers.

atoms of the A(N6)…T(O2) H-bond are juxtaposed, while those of the A(N1)…T(N3) H-bond are separated vertically (see Figures 1 and 4), which increases the A(N1)…T(N3) distance. This unexpected stagger is caused by excessive van der Waals force-field repulsion between the A(N1) and T(N3) atoms, which evidently is imbalanced in the force-field description. In order to address the effect of such repulsion in classical molecular dynamics simulations, we carried out a bioinformatics comparison of heavy atom distances of both AT base pair H-bonds through MD simulations and by examining the

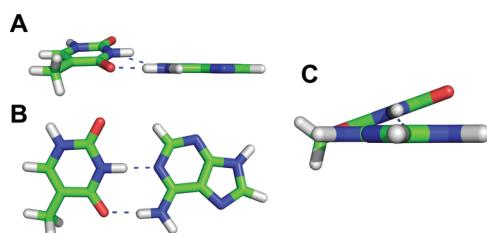


Figure 4. (A) Front, (B) top, and (C) side view of an AT base pair with high stagger as obtained by MM optimization due to an artificial A(N1)...T(N3) clash. The structure shown has a propeller twist of -20° , shear and a stretch equal to 0 Å and a stagger of 0.7 Å; note that the optimal stagger in MM calculations employing the 6–12 van der Waals term can be even slightly higher (see Table S5 in the Supporting Information).

database of DNA X-ray structures with resolutions below 2 Å. We indeed found apparently different distributions for heavy atom distances (see Figure S5 in Supporting Information). However, the significance of this comparison should not be overrated, as the differences might be caused not only by force-field imperfections but also by limited resolution of the X-ray data and the refinement protocols (see Figure S5 and further discussion in the Supporting Information). However, we consider that the differences are most likely caused by force-field imbalance.

In order to examine the empirical pairing-energy profiles for the geometries mimicking optimal QM conformations, the MM profiles were also calculated for the optimal stagger taken from the SCS(MI)-MP2/cc-pV(D,T)Z calculations, while only the stretch parameter was optimized using the force field (see Figure 3, light blue and orange curves). When constraining stagger at the QM optimal values (i.e., keeping it essentially constant and zero), the trends in the MM curves (both energies and optimal stretch) follow the SCS(MI)-MP2/cc-pV(D,T)Z curves, although they differ in absolute values. While the optimal stretch obtained using the standard 6–12 Lennard-Jones potential is overestimated compared to the SCS(MI)-MP2/cc-pV(D,T)Z values, the exponential repulsion potential generally underestimates the interbase distance. However, both force fields give rise to exaggerated A(N1)...T(N3) repulsion relative to the other molecular contacts, which causes the

propeller twist/stagger/stretch interdependence when optimized stagger values are used.

Relaxation of X–H Bonds and Amino Groups. Besides the base-pair parameters (stretch, stagger, shear, opening, and buckle), the base-pairing energy as a function of propeller twist depends on the relaxation of hydrogen atoms involved in hydrogen bonding. Relaxation of the hydrogens includes (i) X–H stretching and (ii) nonplanarity of the A(N6) amino group. As our primary interest in this study is the intermolecular potential, we performed most of the comparisons (specifically, all of the results given above) using QM calculations with rigid planar bases as obtained from full MP2/cc-pVTZ gradient optimization of an AT base pair with C_S symmetry (see the Methods section). However, we expected rather large differences between the force field and electronic structure description of hydrogen position relaxations. Thus, we optimized the hydrogen positions to examine whether MM treatment of hydrogen relaxation might cause an underestimation of propeller twisting in MD simulations.

Figure 5 compares the effect of relaxation of hydrogens on the subsequent SCS(MI)-MP2/cc-pV(D,T)Z and MM interaction energy (propeller twist dependence) computations. We performed two levels of hydrogen optimization: (i) We optimized only the X–H distances; such structures are for brevity labeled as “ C_S ” (mere X–H length optimization keeping C_S symmetry of base monomers; angular degrees of freedom of hydrogens within base plains were also not relaxed). (ii) We fully relaxed hydrogen atoms allowing nonplanarity of the adenine amino hydrogens; these structures are abbreviated as “ C_1 ”. Note that all heavy atoms were fixed, and thus the A(N6) nitrogen was kept within the base plane.

The full relaxation of hydrogens visibly increases tolerance of the propeller twisting at the QM level. On the other hand, at the force-field level, the tolerance of the propeller twisting is reasonably high even in the case of planar base geometries. Further, after full relaxation of hydrogens, the tolerance of the propeller twist in MM matches the corresponding QM value. In other words, when assuming full relaxation of hydrogens (which is relevant to MD simulations), the MM and QM have similar energy penalties for propeller twisting. When assuming fixed hydrogens, the QM procedure resists propeller twisting more than MM. This probably reflects the less realistic description of the directionality of H bonds in MM; i.e., QM

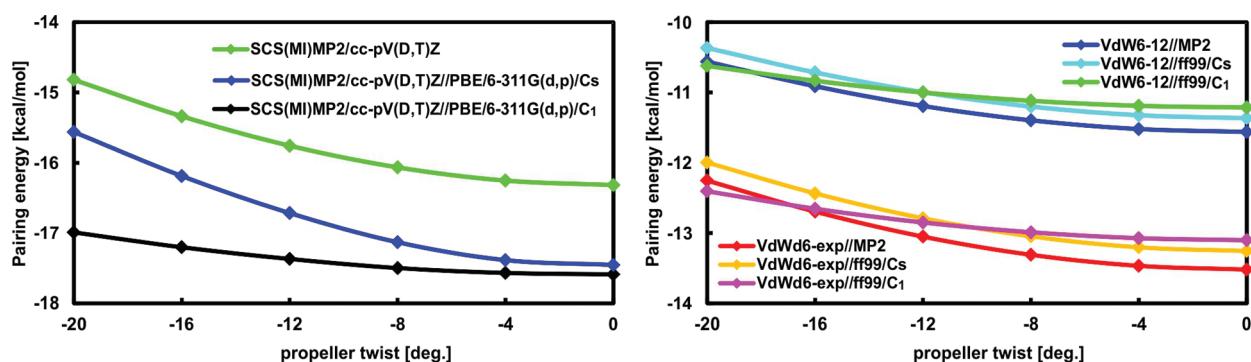


Figure 5. The QM (left) and MM (right) base-pairing energies between (i) rigid planar bases with an X–H bond length as in the reference frame optimized by MP2/cc-pVTZ (green line on the left and blue and red lines on the right), (ii) bases with DFT-relaxed X–H distances (“ C_S ” calculations, blue line on the left and yellow and cyan lines on the right), and (iii) bases with fully DFT-optimized positions of all hydrogens (“ C_1 ” calculations, black line on the left and green and magenta lines on the right) for an AT base pair. Note that the QM and MM graphs have different energy ranges (but the same scale) on the y axis.

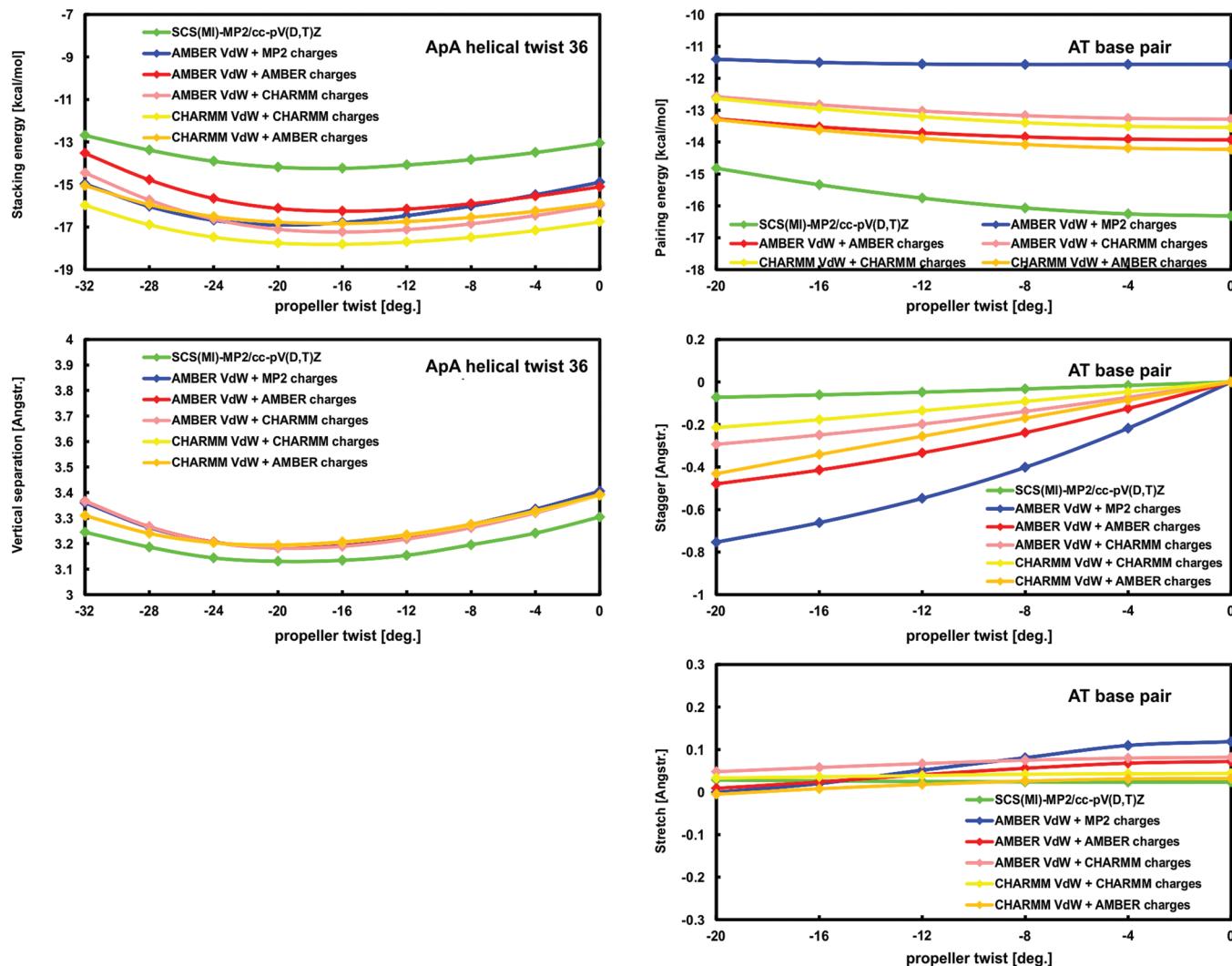


Figure 6. Stacking energies and corresponding optimal vertical separations in an ApA step, together with base-pairing energies and corresponding optimal stretches and stagers as a function of propeller twist in an AT base pair. The data are shown for all four combinations of the CHARMM and AMBER van der Waals parameters and the original force-field charges. For better comparison, the SCS(MI)-MP2 and present MM (AMBER van der Waals parameters and MP2 ESP charges) profiles are also shown.

is more sensitive to losing the perfect H-bond directionality upon propeller twisting.

It should be noted that our hydrogen relaxation procedure has some limitations that affect the absolute energies. This leads to the systematic energy differences between the individual curves shown in Figure 5. However, these differences are not expected to affect the energy changes upon propeller twisting (see the Supporting Information for further details). The basic conclusion from the calculations is that they do not indicate any underestimation of propeller twisting by MM description due to amino group parametrization.

■ DISCUSSION AND CONCLUSIONS

B-DNA molecules with stretches of at least four consecutive adenines in one strand (A-tracts) are intimately involved in DNA bending and regulation of nucleosome binding processes. The A-tracts are known to adopt a specific structure with high propeller twisting, which according to some X-ray structures exceeds -20° . If this experimental result were correct, the A-tract propeller twisting would be underestimated by contem-

porary atomistic simulations.³⁷ This may indicate potential inaccuracy in the force-field description of B-DNA.

This study initially aimed to explain the apparent underestimation of propeller twisting of A-tracts in simulations. Propeller twisting improves the intrastrand stacking overlap of bases, which is reduced by helical twisting. Propeller twisting is opposed by base pairing and by interstrand stacking.^{7,8} We assumed that the underestimation of propeller twisting in A-tracts by MD simulations can be caused by several approximations, such as a potentially inaccurate description of sliding of the stacked nucleobase plates along each other, which, for example, could be related to too steep repulsion of the r^{-12} Lennard-Jones terms. Many other reasons are also possible as mentioned in the Introduction and Supporting Information.

We have exploited the recent major progress in accurate QM computations of base stacking and carried out extensive investigations to evaluate several MM approximations that could potentially lead to underestimation of propeller twist in the simulations of A-tracts. We tested a standard force field as well as a modified force field with an exponential repulsion term and damped dispersion. In addition, we used molecular

dynamics extended by bioinformatics analyses to better understand the observed difference between the QM and MM data.

We found that empirical potentials do not seem to reduce propeller twist through the stacking energy term, at least not in a trivial way. We did not observe either a significant shift of energy minima in the propeller twist direction toward regions of small propeller twists (in absolute values) or any energetic penalty for highly propeller twisted structures in MM profiles compared to QM data. Thus, consideration of the pure stacking energy curves alone does not reveal any straightforward explanation for the potential underestimation of the propeller twist in simulations of B-DNA A-tracts. In addition, the 6-exp term in the force field does not appear to improve the simulation description in this respect.

Furthermore, our investigations of flexibility of the AT base pairs did not reveal any significant differences between the QM and MM stiffness of base pairs with respect to the propeller twist degree of freedom. We carried out two sets of computations, one involving consideration of the amino group flexibility and the other with planar nucleobases. When allowing the amino groups to be nonplanar, the intermolecular energetic penalty of introducing a propeller twist of -20° was -0.6 kcal/mol for both the QM and MM calculation. The energy penalty was even underestimated by the force field when comparing QM and MM computations with rigid nucleobases. In other words, the base pairing data also do not suggest any straightforward explanation for the underestimated propeller twist in A-tract MD simulations.

Although we could not explain the underestimation of propeller twisting in the simulations using the above-mentioned computations, it could be caused by other factors not investigated in this work, or by a combination of several subtle factors. However, we also cannot rule out the possibility that apparent underestimation of propeller twist in molecular dynamics compared to experimental structures might be, at least to some extent, caused by limited accuracy of the experimental structures. There are only two crystallographic studies of long A_6 A-tracts available in the literature. In the A_6 structure of DiGabriele and Steitz⁴⁸ (2.3 Å, 1D89), the reported values of propeller twist in A-tracts mostly exceed -20° and reach up to -28° in one case.⁴⁸ However, the propeller twist did not exceed -18.5° for the central of the three monomers resolved in the unit cell, which has the lowest B-factors and more localized electron density, and thus is expected to be a more highly resolved part of the structure. In the Supporting Information, we summarize average propeller twist values of A_3T_3 and A_6 tracts obtained by MD simulations compared with experimental X-ray data. Comparing molecular dynamics data of the A_6 tract with the central duplex of the structure of DiGabriele and Steitz,⁴⁸ the agreement with the experimental structure is found to be good (Figure S6). The other crystal structure of the A-tract (1D98) suggests that the propeller twisting in the A-tract reaches almost -26° , thus supporting large propeller twisting in sufficiently long A-tracts.⁴⁷ Therefore, with the available experimental data, the exact magnitude of the propeller twist in long A-tracts is somewhat uncertain.

Despite the fact that we could not unambiguously resolve the issue of magnitude of propeller twisting in A-tracts, the reference QM computations revealed two surprising substantial differences between the QM and MM description of base stacking and base pairing: (i) an imbalance in the description of stacking and pairing energy, whereby the force fields over-

estimate stacking and underestimate base pairing and (ii) significant steric clash between A(N1) and T(N3) atoms in AT base pairs. The first imbalance in particular may have far-reaching implications for simulation studies of nucleic acids.

The imbalance in stacking and base-pairing descriptions is apparent from Figure 2, which shows that the force-field stacking energies are overstabilized with respect to the reference QM values by about 25%. In contrast, in the case of the H-bonded AT pair (Figures 3 and 5), the force field underestimates the interaction by about 30%. Similar results were observed in other MM calculations, where we tested the original force-field charges and all four combinations of CHARMM and AMBER van der Waals parameters and charge sets (Figure 6). This could be due to the anisotropy of the dispersion (van der Waals) interaction, as demonstrated previously by Zgarbova et al.⁴⁹ Whereas the van der Waals force-field term is isotropic, the actual dispersion of nucleic acid bases is strongly anisotropic. Consequently, the actual dispersion interaction is relatively weakened in the stacking direction (perpendicular to the base plane) and strengthened in the in-plane directions (H-bonding) compared to the isotropic case. As this effect is not included in the force field, the force-field stacking interactions are overestimated and the corresponding H-bonding interactions are underestimated for nucleic acid bases if averaged isotropic van der Waals parameters are used. This is consistent with the above-described behavior, and it further generalizes our previous observations for uracil dimers.^{49,90} Work is in progress in our laboratories to investigate this issue further.

It should be emphasized that the magnitude of dispersion anisotropy is relatively large. For instance, in cytosine dimers, the C_6 component in the direction perpendicular to the ring plane is about 3-fold smaller than in ring plane.⁹¹ Fortunately, the observed errors in the force-field energy are much smaller due to fortuitous error cancellation.⁴⁹ It should be noted that while neglect of dispersion anisotropy in the force field is the most likely explanation for overestimation of the stacking energies, the other factors contribute to the underestimation of the H-bonding interactions as well, such as missing polarization or electrostatic overlap energy.⁴⁹ Note also that the errors stemming from the neglect of anisotropy cannot be improved by using the C_6 -exp force-field form. Although the exponential term improves (strengthens) the H-bonds by increasing the interaction energy, it also strengthens the stacking interaction, which is already too strong. Resolving this problem would require anisotropic treatment of the van der Waals forces.

It should be noted that the dispersion anisotropy and associated imbalance in the force field may have a dramatic impact on simulations. Let us assume that the anisotropy also affects the hydration of nucleobases since it includes H-bonding between bases and water molecules. As a net result, it could drive nucleic acids simulations into spurious overstabilization of conformations with stacked nucleobases. In fact, at least two such cases of overstabilization of stacking have been already reported.^{88,89,92} One example involved simulation of diagonal four-thymidine loops of d(G₄T₄G₄)₂ quadruplex DNA. Instead of the native loop topology, the simulations favored a thymine triad stacked on the outer guanine quartets of the quadruplex stem, which is in disagreement with a number of independent NMR and X-ray structures.^{88,89} Recently, likely significant overstabilization of stacking has also been suggested in studies of end-to-end stacking of B-DNA duplexes.⁹² Thus, overstacking may be a widespread problem in MM-based MD

simulations. Since a reduction of helical twist is associated with increased base-pair van der Waals overlap, the overestimation of stacking could perhaps even contribute to the known tendency of at least some force fields to underestimate the helical twist of B-DNA. This was observed as smaller differences in the ApA and TpA base-pair-step stacking energies between helical twists of 30° and 36° in MM calculation compared to the QM values (see Supporting Information Tables S1–S3 and Figure S4). Further, since propeller twisting arises as a response to helical twist, it could also, through reduction of helical twist, contribute to the underestimation of propeller twist. However, the latter two suggestions are rather speculative for the time being, as they deal with very subtle effects.

The second identified force-field imbalance is steric clash between A(N1) and T(N3) nitrogens forming one of two hydrogen bonds of the AT base pair. We found that this clash is still present even when exponential repulsion is used, which indicates that the van der Waals minimum of A(N1)…T(N3) and/or A(N1)…T(H3) contacts is shifted to too large a distance. This is supported by the comparison of MD data with the X-ray database. In particular, we found that the distances between the heavy atoms in both AT base-pair hydrogen bonds differed significantly between the MD data and X-ray database, although the comparison might be limited by the resolution and refinement protocol of the X-ray structures (see above and Supporting Information). Furthermore, according to AMBER MD simulations, the AT base-pair stretch was $0.0 \pm 0.1 \text{ \AA}$, whereas the typical X-ray value was -0.3 \AA , i.e., well outside the range of MD fluctuations. Note that this steric clash was more subtle when CHARMM van der Waals parameters were used, but it was still definitely present (Figure 6). On the other hand, it is not clear whether this localized A(N1)…T(N3) clash might influence some other parameters of DNA helical structure beside the above-mentioned base-pair stretch, and thus its impact on the simulation behavior may be rather limited.

As a result of the study, we extended the available set of benchmark estimated CCSD(T)/CBS data for base stacking. We suggest that these geometries, which are biochemically relevant, may be useful in benchmarking other computational methods. We also developed a code that allows generation of diverse base-stacking geometries suitable for QM computations with predefined intra- and interbase-pair parameters. Such geometries have considerable advantages over geometries taken directly from X-ray diffraction experiments, since the latter may be affected by data and refinement errors, which may bias accurate energy computations.⁹³

The present work demonstrates that the highest-accuracy QM computations can increasingly be directly applied for studying interesting aspects of nucleic acids structure and can complement the picture provided by atomistic simulations.

ASSOCIATED CONTENT

Supporting Information

Complete graphs of the stacking energies in ApA, ApT, and TpA base-pair steps as a function of propeller twist and helical twist (Figures S1–S4); tabulated stacking energies and optimal vertical separations (Tables S1–S4); pairing energy of AT base pair and optimal values of stretch and stagger (Tables S5–S7); averaged propeller twists of A-tract's AT base pairs obtained from MD (Figure S6); and more detailed discussions of DNA sequence dependence, a comparison of X-ray and MD data about A-tract's structure, possible reasons for underestimation

of the propeller twist in A-tract simulations, the algorithm of our software for generation of the structures, bioinformatics analysis, and details of QM calculations. Furthermore, all geometries, for which CBS(T) energies were calculated, are also provided. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pavel.banas@upol.cz, sponer@ncbr.muni.cz.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the project “CEITEC - Central European Institute of Technology” CZ.1.05/1.1.00/02.0068 (J.S.), “RCPTM – Centre of Advanced Technologies and Materials” and “RCPTM-TEAM” CZ.1.05/2.1.00/03.0058, CZ.1.07/2.3.00/20.0017 (M.O., P.B.) from the European Regional Development Fund, Ministry of Education Research grant MSM 6046137302 (D.S.), the J. E. Purkyne Fellowship and project no. Z40550506 from the Academy of Sciences of the Czech Republic (F.L.), and by grants 203/09/H046 (J.S., A.M.), 203/09/1476 (J.S., A.M.), P208/11/1822 (J.S., A.M., D.S.), P208/12/1878 (J.S., A.M.), P305/12/G03 (J.S., A.M.), P208/10/1742 (P.J.), and P301/11/P558 (P.B.) by the Grant Agency of the Czech Republic. We would like to thank the reviewers for very careful reading and many stimulating comments.

REFERENCES

- Zhurkin, V. B.; Olson, W. K. *Curr. Opin. Struct. Biol.* **2011**, 21, 348–57.
- Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, 287, 755–58.
- Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, 149, 761–86.
- Dickerson, R. E. *Sci. Am.* **1983**, 249, 94–111.
- Peters, J. P.; Maher, L. J. *Q. Rev. Biophys.* **2010**, 43, 23–63.
- Olson, W. K.; Gorin, A. A.; Lu, X. J.; Hock, L. M.; Zhurkin, V. B. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 11163–68.
- Calladine, C. R. *J. Mol. Biol.* **1982**, 161, 343–52.
- Dickerson, R. E. *J. Mol. Biol.* **1983**, 166, 419–41.
- Srinivasan, A. R.; Torres, R.; Clark, W.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1987**, 5, 459–96.
- Sponer, J.; Kypr, J. *J. Biomol. Struct. Dyn.* **1993**, 11, 277–92.
- Sponer, J.; Kypr, J. *J. Biomol. Struct. Dyn.* **1993**, 11, 27–41.
- Goodsell, D. S.; Kaczorzeskowiak, M.; Dickerson, R. E. *J. Mol. Biol.* **1994**, 239, 79–96.
- Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, 100, 5590–96.
- Sponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. *Biophys. J.* **1997**, 73, 76–87.
- Suzuki, M.; Amano, N.; Kakinuma, J.; Tateno, M. *J. Mol. Biol.* **1997**, 274, 421–35.
- Sponer, J.; Florian, J.; Ng, H. L.; Sponer, J. E.; Spackova, N. *Nucleic Acids Res.* **2000**, 28, 4893–902.
- Sponer, J.; Hobza, P. *Collect. Czech. Chem. Commun.* **2003**, 68, 2231–82.
- Lankas, F.; Sponer, J.; Langowski, J.; Cheatham, T. E. *Biophys. J.* **2003**, 85, 2872–83.
- Dixit, S. B.; Case, D. A.; Cheatham, T. E.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; Thayer, K. M.; Varnai, P.; Beveridge, D. L. *Biophys. J.* **2005**, 89, 3721–40.
- Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.

- Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J. *Nucleic Acids Res.* **2010**, *38*, 299–313.
- (21) Svozil, D.; Hobza, P.; Sponer, J. *J. Phys. Chem. B* **2010**, *114*, 2547–47.
- (22) Svozil, D.; Hobza, P.; Sponer, J. *J. Phys. Chem. B* **2010**, *114*, 1191–203.
- (23) Cysewski, P. *J. Mol. Model.* **2009**, *15*, 597–606.
- (24) Cysewski, P. *J. Mol. Model.* **2010**, *16*, 1721–29.
- (25) Czyzniowska, Z. *THEOCHEM* **2009**, *895*, 161–67.
- (26) Langner, K. M.; Sokalski, W. A.; Leszczynski, J. *J. Chem. Phys.* **2007**, *127*.
- (27) Hill, J. G.; Platts, J. A. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2785–91.
- (28) Cooper, V. R.; Thonhauser, T.; Langreth, D. C. *J. Chem. Phys.* **2008**, *128*.
- (29) Jurecka, P.; Sponer, J.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 5466–71.
- (30) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–29.
- (31) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–97.
- (32) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193–94.
- (33) Barone, F.; Lankas, F.; Spackova, N.; Sponer, J.; Karran, P.; Bignami, M.; Mazzei, F. *Biophys. Chem.* **2005**, *118*, 31–41.
- (34) Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1998**, *281*, 675–87.
- (35) Strahs, D.; Schlick, T. *J. Mol. Biol.* **2000**, *301*, 643–63.
- (36) Curuksu, J.; Zacharias, M.; Lavery, R.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 3766–73.
- (37) Lankas, F.; Spackova, N.; Moakher, M.; Enkhbayar, P.; Sponer, J. *Nucleic Acids Res.* **2010**, *38*, 3414–22.
- (38) Koo, H. S.; Drak, J.; Rice, J. A.; Crothers, D. M. *Biochemistry* **1990**, *29*, 4227–34.
- (39) Segal, E.; Widom, J. *Curr. Opin. Struct. Biol.* **2009**, *19*, 65–71.
- (40) Tolstorukov, M. Y.; Virnik, K. M.; Adhya, S.; Zhurkin, V. B. *Nucleic Acids Res.* **2005**, *33*, 3907–18.
- (41) McConnell, K. J.; Beveridge, D. L. *J. Mol. Biol.* **2001**, *314*, 23–40.
- (42) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1999**, *285*, 1623–32.
- (43) Lankas, F.; Cheatham, T. E.; Spackova, N.; Hobza, P.; Langowski, J.; Sponer, J. *Biophys. J.* **2002**, *82*, 2592–609.
- (44) Sherer, E. C.; Harris, S. A.; Soliva, R.; Orozco, H.; Laughton, C. A. *J. Am. Chem. Soc.* **1999**, *121*, 5981–91.
- (45) Dickerson, R. E. *Nucleic Acids Res.* **1989**, *17*, 1797–803.
- (46) Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8385–89.
- (47) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221–26.
- (48) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231*, 1024–39.
- (49) Zgarbova, M.; Otyepka, M.; Sponer, J.; Hobza, P.; Jurecka, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10476–93.
- (50) Morgado, C. A.; Jurecka, P.; Svozil, D.; Hobza, P.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 1524–44.
- (51) Sponer, J.; Hobza, P. *Int. J. Quantum Chem.* **1996**, *57*, 959–70.
- (52) Luisi, B.; Orozco, M.; Sponer, J.; Luque, F. J.; Shakked, Z. *J. Mol. Biol.* **1998**, *279*, 1123–36.
- (53) Vlieghe, D.; Sponer, J.; Van Meervelt, L. *Biochemistry* **1999**, *38*, 16443–51.
- (54) Sponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709–14.
- (55) Shatzky-Schwartz, M.; Arbuckle, N. D.; Eisenstein, M.; Rabinovich, D.; BareketSamish, A.; Haran, T. E.; Luisi, B. F.; Shakked, Z. *J. Mol. Biol.* **1997**, *267*, 595–623.
- (56) Sponer, J.; Kypr, J. *Int. J. Biol. Macromol.* **1994**, *16*, 3–6.
- (57) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E.; Otyepka, M.; Sponer, J. *J. Chem. Theory Comput.* **2010**, *6*, 3836–49.
- (58) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108–21.
- (59) Lu, X. J.; Olson, W. K. *Nature Protocols* **2008**, *3*, 1213–27.
- (60) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–88.
- (61) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B. P.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossavary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, X. W.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; University of California: San Francisco, CA, 2010.
- (62) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020–41.
- (63) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–35.
- (64) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–71.
- (65) Banas, P.; Rulisek, L.; Hanosova, V.; Svozil, D.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2008**, *112*, 11177–87.
- (66) Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114*, 8701–12.
- (67) Mlynky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114*, 6642–52.
- (68) Sklenovsky, P.; Florova, P.; Banas, P.; Reblova, K.; Lankas, F.; Otyepka, M.; Sponer, J. *J. Chem. Theory Comput.* **2011**, *7*, 2963–80.
- (69) Distasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073–83.
- (70) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639–46.
- (71) Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65–73.
- (72) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.
- (73) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437–46.
- (74) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–23.
- (75) Dunning, T. H. *J. Phys. Chem. A* **2000**, *104*, 9062–80.
- (76) Werner, H. J.; Manby, F. R.; Knowles, P. *J. J. Chem. Phys.* **2003**, *118*, 8149–60.
- (77) *MOLPRO*, version 2006.1; Cardiff University: Cardiff, U.K., 2006.
- (78) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–66.
- (79) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *135*.
- (80) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–93.
- (81) Sponer, J.; Jurecka, P.; Marchan, I.; Luque, F. J.; Orozco, M.; Hobza, P. *Chem.—Eur. J.* **2006**, *12*, 2854–65.
- (82) Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–69.
- (83) Cerny, J.; Jurecka, P.; Hobza, P.; Valdes, H. *J. Phys. Chem. A* **2007**, *111*, 1146–54.
- (84) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–69.

- (85) Eichkorn, K.; Treutler, O.; Ohm, H.; Haser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–89.
- (86) Whitten, J. L. *J. Chem. Phys.* **1973**, *58*, 4496–501.
- (87) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897–909.
- (88) Fadrna, E.; Spackova, N.; Stefl, R.; Koca, J.; Cheatham, T. E.; Sponer, J. *Biophys. J.* **2004**, *87*, 227–42.
- (89) Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E.; Kulinski, T.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 2514–30.
- (90) Kolar, M.; Berka, K.; Jurecka, P.; Hobza, P. *ChemPhysChem* **2010**, *11*, 2399–408.
- (91) Haley, T. P.; Graybill, E. R.; Cybulski, S. M. *J. Chem. Phys.* **2006**, *124*.
- (92) Maffeo, C.; Luan, B.; Aksimentiev, A. *Nucleic Acids Res.* **2012**, *40*, 3812–3821.
- (93) Sponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595–610.