

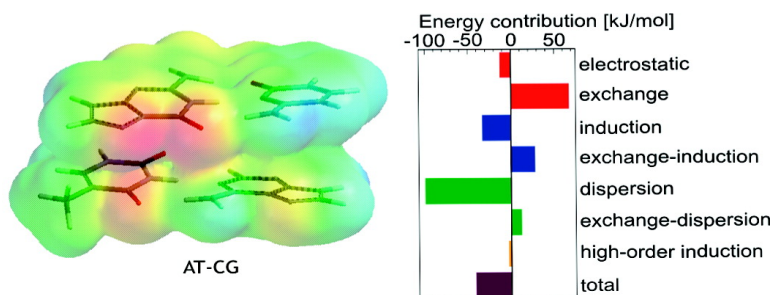
Communication

Stacking Energies for Average B-DNA Structures from the Combined Density Functional Theory and Symmetry-Adapted Perturbation Theory Approach

Annamaria Fiethen, Georg Jansen, Andreas Hesselmann, and Martin Schtz

J. Am. Chem. Soc., **2008**, 130 (6), 1802-1803 • DOI: 10.1021/ja076781m

Downloaded from <http://pubs.acs.org> on February 8, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 5 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



Stacking Energies for Average B-DNA Structures from the Combined Density Functional Theory and Symmetry-Adapted Perturbation Theory Approach

Annamaria Fiethen,[†] Georg Jansen,^{*,†} Andreas Hesselmann,[‡] and Martin Schütz[§]

Fachbereich Chemie, Universität Duisburg-Essen, Campus Essen, Universitätsstrasse, 5, D-45117 Essen, Germany,

Lehrstuhl für Theoretische Chemie, Universität Erlangen-Nürnberg, Egerlandstrasse 3, D-91058 Erlangen,

Germany, and Institut für Physikalische und Theoretische Chemie, Universität Regensburg,

Universitätsstrasse 31, D-93040 Regensburg, Germany

Received September 7, 2007; E-mail: georg.jansen@uni-due.de

The π -stacking energy between parallel layers of hydrogen-bonded base pairs is one of the important factors stabilizing the structure of DNA.¹ It is of considerable interest to achieve a detailed qualitative and quantitative understanding of the physicochemical nature of these interactions, for example, in the notions of intermolecular perturbation theory such as the first-order electrostatic, $E_{\text{el}}^{(1)}$, second-order induction, $E_{\text{ind}}^{(2)}$, and dispersion, $E_{\text{disp}}^{(2)}$, contributions to the interaction energy E_{int} . These terms have been defined in the early days of quantum mechanics and continue to form the basis of our understanding of the interactions between molecules. They also form the conceptual framework of the intermolecular part of most force-fields actually in use to simulate biopolymers such as DNA. For example, in present force-fields $E_{\text{el}}^{(1)}$ usually is approximated through Coulomb interactions between partial charges located at the atoms and $E_{\text{disp}}^{(2)}$ through atom–atom potentials decaying as $1/R^6$. In particular these two contributions also play an important role in the framework of quantum chemical calculations of stacking interactions: the former in attempts² to correlate E_{int} from supermolecular calculations with more or less elaborate approximations to $E_{\text{el}}^{(1)}$, and $E_{\text{disp}}^{(2)}$ in methods aiming at improving density functional theory (DFT) for stacking interactions via direct addition of damped atom–atom dispersion energies.³

In the present study stacking energies for all possible combinations of two base pairs arranged as in an averaged B-DNA structure were obtained directly from intermolecular perturbation theory. To this end the recently developed DFT-SAPT⁴ combination of symmetry-adapted intermolecular perturbation theory (SAPT)⁵ with a DFT description of the monomer properties entering SAPT has been employed. DFT-SAPT and the related SAPT(DFT)⁶ approach do not rely on any multipole approximation but rather employ the full charge densities and the frequency-dependent linear response densities of the interacting molecules to evaluate $E_{\text{el}}^{(1)}$, $E_{\text{ind}}^{(2)}$, and $E_{\text{disp}}^{(2)}$. Thus the important charge density penetration effects are fully taken into account—not only in the first-order² but also in the second-order contributions. Furthermore, the simultaneous exchange of electrons between the interacting molecules due to the antisymmetry requirement of the total wavefunction is taken care of in SAPT by adding the repulsive exchange corrections $E_{\text{exch}}^{(1)}$, $E_{\text{exch-ind}}^{(2)}$, and $E_{\text{exch-disp}}^{(2)}$ to the above terms. To these contributions the Hartree–Fock estimate of the third and higher order induction and exchange-induction contributions denoted as $\delta(\text{HF})$ is added. In a number of previous benchmark studies including the benzene-dimer,^{4f} the acetylene–benzene complex,⁷ and Watson–Crick (WC) and stacked structures of the guanine–cytosine (GC) and adenine–thymine (AT) dimers,⁸ it was shown that the resulting total interaction energies are in excellent agreement with the outcome of expensive coupled-cluster calculations including up to non-iterative triple excitations (CCSD(T)). CCSD(T) is accepted to yield the most accurate E_{int} for these systems to date. The present

contribution thus aims at providing accurate values for both the total interaction energies and all of its contributions.

In contrast to our previous work⁸ not only dimers but rather tetramers consisting of two stacked complementary dimers of the DNA bases were considered. While the dimer structures used previously came from ab initio gas-phase geometry optimizations⁹ and were not representative for their arrangement in DNA, the tetramer structures in this study correspond more closely to those they assume in B-DNA. They were generated with the 3DNA program¹⁰ using the complementary base pair ($\kappa = 0.5^\circ$, $\pi = -11.4^\circ$, $\sigma = 0.6^\circ$, $S_x = 0.00 \text{ \AA}$, $S_y = -0.15 \text{ \AA}$, $S_z = 0.09 \text{ \AA}$) and base-pair step parameters ($\tau = -0.1^\circ$, $\rho = 0.6^\circ$, $\omega = 36.0^\circ$, $D_x = -0.02 \text{ \AA}$, $D_y = 0.23 \text{ \AA}$, $D_z = 3.32 \text{ \AA}$) for B-DNA obtained as average values from high-resolution crystal structures.¹¹ Use of these parameters provides a well-defined point of reference for later studies investigating the influence of deviations from these average values. The hydrogen atom positions were determined in second-order Møller–Plesset (MP2) geometry optimizations of the complementary AT and CG dimers, employing the TZVPP basis set¹² and constraining the positions of the C, N, and O atoms to those generated by 3DNA. Since the current version of DFT-SAPT is restricted to dimer interaction energies, each complementary dimer was considered as a supermolecule, thus treating it actually as one of the monomers in DFT-SAPT. All DFT-SAPT and accompanying supermolecular calculations were done with Molpro,¹³ keeping core orbitals frozen. Density-fitting techniques as detailed in ref 4f were employed. The aug-cc-pVTZ basis set¹⁴ was used throughout, combined with the corresponding MP2 auxiliary basis set¹⁵ for density-fitting of DFT-SAPT dispersion contributions and supermolecular correlation energies, while density-fitting of all other SAPT contributions was done with the cc-pVQZ JK auxiliary basis set.¹⁶ Supermolecular E_{int} were counterpoise-corrected,¹⁷ which, by construction, is not necessary with DFT-SAPT. The exchange-correlation potential and kernel entering the DFT calculations were approximated with the LPBE0AC¹⁸/ALDA¹⁹ combination justified in ref 4f.

Figure 1 compares total E_{int} from DFT-SAPT to results from supermolecular electron correlation methods applicable to systems of this size, that is, MP2 and spin-component-scaled MP2²⁰ (SCS-MP2), also using the aug-cc-pVTZ basis set. MP2 is known to strongly overestimate the electron correlation contribution to π – π interacting systems, while SCS-MP2 yields E_{int} in much better agreement with CCSD(T).²⁰ The stacking energies from DFT-SAPT and the 3.5 times less time-consuming SCS-MP2 agree quite well, despite their entirely different theoretical background. Depending on the base pair step E_{int} varies between -30.7 (AT-TA) and -55.0 kJ/mol (CG-GC) with DFT-SAPT and between -27.7 and -49.5 kJ/mol with SCS-MP2. SCS-MP2 underestimates the magnitude of E_{int} with respect to DFT-SAPT by 0.6 (GC-AT) to 5.5 kJ/mol (CG-GC), on average by 2.5 kJ/mol. MP2, on the other hand, yields drastically larger magnitudes of E_{int} , that is, by 17.8 (CG-AT) to 21.7 kJ/mol (GC-CG), and on average by 20.1 kJ/mol.

[†] Universität Duisburg-Essen.

[‡] Universität Erlangen-Nürnberg.

[§] Universität Regensburg.

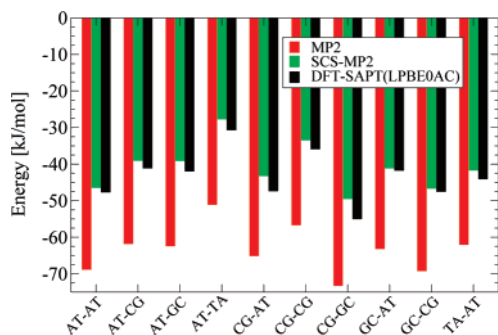


Figure 1. Comparison of total stacking energies for the ten unique steps of two DNA base pairs in average B-DNA from various methods.

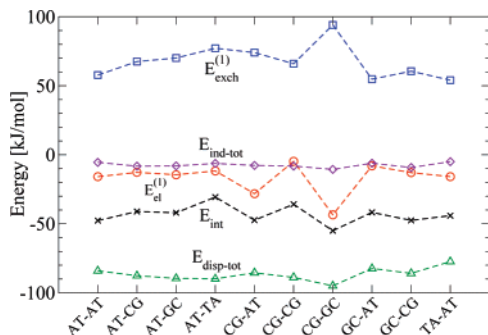


Figure 2. Stacking energy contributions for the ten unique steps of two DNA base pairs in average B-DNA from DFT-SAPT.

Figure 2 shows the DFT-SAPT contributions to E_{int} : the first-order energies $E_{\text{el}}^{(1)}$ and $E_{\text{exch}}^{(1)}$, the sum of all induction and exchange-induction energies $E_{\text{ind-tot}} = E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)} + \delta(\text{HF})$, and the second-order total dispersion energy $E_{\text{disp-tot}} = E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)}$ (cf. Supporting Information for complete data). By and large, the relative importance of the various contributions for all base-pair steps is similar to that found for π -stacked structures of the benzene dimer.^{4f} In detail, $E_{\text{disp-tot}}$ is the largest contribution, in which the repulsive $E_{\text{exch-disp}}^{(2)}$ cancels only 11 to 14% of $E_{\text{disp}}^{(2)}$. Depending on the base-pair step $E_{\text{disp-tot}}$ varies between -77.3 (TA-AT) and -94.9 kJ/mol (CG-GC). $E_{\text{ind-tot}}$ is the least important and most constant among all contributions. It lies between -5.0 (TA-AT) and -10.6 kJ/mol (CG-GC) and on average amounts to -7.5 kJ/mol. Note that $E_{\text{ind}}^{(2)}$ is much larger and varies much more strongly, that is, between -24.6 (TA-AT) and -52.7 kJ/mol (CG-GC). Yet, about 80 to 90% of this is compensated by the repulsive $E_{\text{exch-ind}}^{(2)}$. To the remainder the higher-order induction and exchange-induction estimate $\delta(\text{HF})$ adds between -1.7 (GC-AT) and -4.2 kJ/mol (CG-GC). The two first-order contributions, $E_{\text{el}}^{(1)}$ and $E_{\text{exch}}^{(1)}$, vary strongly with the base-pair step. While the electrostatic interaction energy is always negative, with values between -4.8 (CG-CG) and -43.5 kJ/mol (CG-GC), the repulsive first-order exchange energy is the decisive factor counterbalancing the dispersion contribution. It ranges from $+54.0$ (TA-AT) to $+94.0$ kJ/mol (CG-GC). The strikingly large contributions in the CG-GC step can be attributed to close contacts: three pairs of atoms have distances which are smaller by 0.2 Å than the sum of their standard van der Waals radii.²¹ In CG-AT there are two such pairs, in TA-AT one, and none in all other steps.

As becomes apparent from Figure 2 there are no quantitatively useful correlations between the total interaction energy and any of its contributions. Among all individual contributions $E_{\text{el}}^{(1)}$ still correlates best with E_{int} , but the Pearson correlation coefficient r_{xy} is only 0.756 and thus way too small to be quantitatively exploited. The present results thus confirm the previous criticism^{2c} of corresponding attempts.^{2a,b} Adding $E_{\text{exch}}^{(1)}$ with its partially com-

pensating trends to $E_{\text{el}}^{(1)}$ does not help: r_{xy} now drops to 0.687. There are, however, reasonable correlations between individual interaction energy components, the strongest of which is that between $E_{\text{ind}}^{(2)}$ and $E_{\text{exch-ind}}^{(2)}$ with $r_{xy} = -0.987$. As mentioned above, these two contributions cancel each other to a large extent, and their sum has about the same size as $\delta(\text{HF})$. The latter, however, does not correlate with $E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)}$ ($r_{xy} = 0.207$), thus preventing prediction of $E_{\text{ind-tot}}$ from any of its contributors. Reasonable correlations also exist between $E_{\text{disp}}^{(2)}$ and $E_{\text{exch-disp}}^{(2)}$ ($r_{xy} = -0.949$) and $E_{\text{exch}}^{(1)}$ and $E_{\text{disp}}^{(2)}$ ($r_{xy} = -0.920$). The most useful correlation for quantitative purposes, however, is that between $E_{\text{ind-tot}}$ and $E_{\text{disp-tot}}$: r_{xy} here is merely 0.776, but because of the relative smallness of $E_{\text{ind-tot}}$ a simple scaling of $E_{\text{disp-tot}}$ with 1.09 reproduces the sum $E_{\text{ind-tot}} + E_{\text{disp-tot}}$ within about ± 2 kJ/mol. This appears to be good enough to effectively account for the total induction contribution in a classical nonpolarizable force field for DNA simulations—at least as long as the surrounding environment (backbone, water, counterions) of the base pairs is neglected.

In summary, reliable modeling of DNA stacking requires accurate models of at least the first-order electrostatic and exchange and the second-order total dispersion contributions. The results presented here provide a well-defined basis for future developments of force-fields and quantum chemical models in this important area.

Supporting Information Available: Complete refs 11 and 13, further computational details, tables with energies and atomic coordinates, and figures of calculated structures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Žezáč, J.; Hobza, P. *Chem. Eur. J.* **2007**, *13*, 2983.
- (a) Joubert, L.; Popelier, P. L. A. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4353. (b) Hill, G.; Forde, G.; Hill, N.; Lester, W. A., Jr.; Sokalski, W. A.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *381*, 729. (c) Toczyłowski, R. R.; Cybulski, S. M. *J. Chem. Phys.* **2005**, *123*, 154312.
- Elstner, M.; Hobza, P.; Frauenheim, T.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149. Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515. Grimme, S. *J. Comp. Chem.* **2004**, *25*, 1463. Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comp. Chem.* **2007**, *28*, 555. Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397.
- (a) Jansen, G.; Hesselmann, A. *J. Phys. Chem. A* **2001**, *105*, 11156. (b) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464. (c) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319. (d) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778. (e) Hesselmann, A.; Jansen, G. *Phys. Chem. Chem. Phys.* **2003**, *5*, 5010. (f) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
- Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887. Szalewicz, K.; Patkowski, K.; Jeziorski, B. *Struct. Bond* **2005**, *116*, 43.
- Misquitta, A. J.; Szalewicz, K. *Chem. Phys. Lett.* **2002**, *357*, 301. Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2003**, *91*, 033201. Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005**, *123*, 214103.
- Tekin, A.; Jansen, G. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1680.
- Hesselmann, A.; Jansen, G.; Schütz, M. *J. Am. Chem. Soc.* **2006**, *128*, 11730.
- Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608.
- Lu, X.-J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108.
- Olson, W. K.; et al. *J. Mol. Biol.* **2001**, *313*, 229.
- Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.
- Werner, H.-J.; et al. *MOLPRO, a package of ab initio programs*, version 2006.3. <http://www.molpro.net> (accessed June 2007).
- Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.
- Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285.
- Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865. Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158. van Leeuwen, R.; Baerends, E. *J. Phys. Rev. A* **1994**, *49*, 2421. Grüning, M.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. *J. Chem. Phys.* **2001**, *114*, 652. Della Sala, F.; Görling, A. *J. Chem. Phys.* **2001**, *115*, 5718. Hesselmann, A.; Manby, F. *J. Chem. Phys.* **2005**, *123*, 164116.
- Gross, E. K. U.; Kohn, W. In *Adv. Quantum Chem.*; Trickey, S. B., Ed.; Academic: San Diego, 1990; Vol. 21, p 255.
- Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095. Antony, J.; Grimme, S. *J. Phys. Chem. A* **2007**, *111*, 4862.
- Bondi, A. *J. Phys. Chem. Phys.* **1964**, *68*, 441.

JA076781M