Nanosecond Molecular Dynamics of Zipper-like DNA Duplex Structures Containing Sheared G·A Mismatch Pairs

Nad'a Špačková,[†,‡,§] Imre Berger,[⊥] and Jiří Šponer*[,†,‡]

Contribution from the Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic, J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague, Czech Republic, Department of Physical Electronics, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic, and Institute for Molecular Biology and Biophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

Received February 7, 2000. Revised Manuscript Received May 18, 2000

Abstract: Molecular dynamics (MD) simulations are presented of an unusual DNA duplex structure with the sequence d(GCGAAAGC)$_2$ that adopts a central zipper motif of four unpaired and mutually intercalated adenines enveloped by sheared G·A mismatch base pairs and Watson−Crick G·C base pairs with B-form geometry at its end. On a nanosecond scale, the simulations show very stable trajectories and not only the Watson−Crick base pairs but also the central unpaired adenine zipper are revealed as predominantly rigid segments of the molecule. The sheared G·A mismatch base pairs in contrast are nonplanar and flexible, and bending of the structure can occur at the mismatch junctions. The pronounced flexibility of the sheared G·A mismatches is explained as a result of their intrinsic nonplanarity rather than being a consequence of any interactions with neighboring residues. The simulations clearly show that sheared G·A mismatches require extensive stacking with adjacent base pairs for their maintenance. Two stable local conformational substates of the d(GCGAAAGC)$_2$ zipper molecule are suggested by the simulations, involving cation-stabilized clustering of three negatively charged phosphate groups in the zipper region accompanied by adjustment of adenine stacking, sugar repuckering, and the presence of several highly ordered hydration sites with close to 100% occupancy and long-residing water molecules. Further, the capability of the zipper motif to incorporate guanine, cytosine, or thymine residues is tested. All simulations were carried out with the AMBER5 program with a force field created by Cornell et al. (Cornell, W. D.; et al. *J. Am. Chem. Soc.* **1995**, *117*, 5179) using the particle mesh Ewald (PME) technique for electrostatic interactions, with a total length reaching 30 ns. The overall results confirm an excellent performance of the PME MD technique and of the force field of Cornell et al. for unusual nucleic acid conformations.

Introduction

High-resolution analysis of DNA fragments provides a wealth of information at the atomic level about the multitude of conformations that the hereditary molecule can adopt. Since the discovery of the standard B geometry duplex form, it has become increasingly clear that DNA, far from being a rigid molecule, can adopt an astonishing variety of shapes. Besides the established B, A, and Z duplexes, DNA can form triplexes and also structures such as the G-DNA quadruplex formed by guanine-rich sequences,[1] the four-stranded intercalated cytosine-rich i-DNA motif,[2] and many more.[3] Based on single crystal analysis and NMR solution studies, it was recently discovered that base self-intercalation in DNA is not only confined to cytosine residues as in the i-DNA motif.[4] The crystal structure of the DNA duplex formed by d(GCGAAAGCT) showed, in an antiparallel double-stranded molecule, the presence of a central segment of four adenine residues that were not involved in base pairing and intercalated into each other such that a core stack of four consecutive adenine bases was formed.[4a] This central segment was enveloped by sheared G·A mismatch base pairs, upon which two Watson−Crick G·C base pairs were stacked in an essentially B-DNA form geometry. The two 3′ terminal thymine residues present in this duplex formed unpaired overhangs at either end of the molecule. The sequence d(GC-

* Address correspondence to this author at Institute of Physical Chemistry, Academy of Sciences of the Czech Republic. Fax: (420 2) 858 2307. E-mail: sponer@indy.jh-inst.cas.cz.

[†] Institute of Biophysics, Academy of Sciences of the Czech Republic.

[‡] Institute of Physical Chemistry, Academy of Sciences of the Czech Republic.

[§] Masaryk University.

[⊥] ETH-Hönggerberg.

(1) (a) Williamson, J. R.; Raghuram, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871−880. (b) Sundquist, W. I.; Klug, A. *Nature* **1989**, *342*, 825−829. (c) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. *Science* **1994**, *265*, 520. (d) Han, F. X.; Wheelhouse, R. T.; Hurley, L. H. *J. Am. Chem. Soc.* **1999**, *121*, 3561−3570. (e) Mergny, J. L.; Mailliet, P.; Lavelle, F.; Riou, J. F.; Laoui, A.; Helene, C. *Anticancer Drug Des.* **1999**, *14*, 327−339.

(2) (a) Gehring, K.; Leroy, J.-J.; Gueron, M. *Nature* **1993**, *363*, 561−565. (b) Nonin, S.; Leroy, J.-L. *J. Mol. Biol.* **1996**, *261*, 399−414. (c) Kang, C.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11636−11640.

(3) (a) Bouaziz, S.; Kettani, A.; Patel, D. J. *J. Mol. Biol.* **1998**, *282*, 637−652. (b) Lin, C. H.; Wang, W. M.; Jones, R. A.; Patel, D. J. *Chem., Biol.* **1998**, *5*, 555−572. (c) Kettani, A.; Bouaziz, S.; Skripkin, E.; Majumdar, A.; Wang, W. M.; Jones, R. A.; Patel, D. J. *Struct. Fold. Des.* **1999**, *7*, 803−815. (d) Chaput, J. C.; Switzer, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10614−10619.

(4) (a) Shepard, W.; Cruse, W. B. T.; Fourme, R.; de la Fortelle, E.; Prange, T. *Structure* **1998**, *6*, 849−861. (b) Chou, S.-H.; Zhu, L.; Reid, B. R. *J. Mol. Biol.* **1994**, *244*, 259−268. (c) Zhu, L.; Chou, S.-H.; Reid, B. R *J. Mol. Biol.* **1995**, *254*, 623−637. (d) Zhu, L.; Chou, S.-H.; Reid, B. R. *Nat. Struct. Biol.* **1995**, *2*, 1012−1017. (e) Chou, S.-H.; Zhu, L.; Reid, B. R. *J. Mol. Biol.* **1997**, *267*, 1055−1067.

10.1021/ja000460s CCC: $19.00
Published on Web 07/25/2000

GAAAGCT) is of interest as it incorporates the oligonucleotide sequence −GAAAGC− that is a consensus found at the replication origin of bacteriophages and parvoviruses.[5] In a variety of biochemical and biophysical studies, it was shown that oligonucleotides with this sequence have unusual properties such as high melting temperatures, nuclease resistance, and fast gel mobilities.[6] These findings lead to the proposal of intramolecular hairpin formation by these sequences, which was further corroborated by NMR studies.[5−7] Interestingly, the crystal structure of the nonanucleotide incorporating the consensus, d(GCGAAAGCT), showed not a hairpin but the unusual intercalated duplex structure, exemplifying the dynamic nature of this DNA. While NMR-based preliminary molecular dynamics calculations of the hairpin conformational state exist,[7a] there are no data available about the dynamics of the unusual duplex conformation involving the adenine zipper that was present in the crystal. This mode of zipper-like self-intercalation in a duplex was not only observed in crystals. NMR solution studies demonstrated that related pyrimidine−GNA−purine motifs can, besides unimolecular structures with a single base hairpin loop, also adopt bimolecular $(GNA)_2$ complexes with two unpaired intercalated bases in their center bracketed again by sheared G·A base pairs, with the hairpin−duplex equilibrium of PyGNAPu motifs being sequence-dependent.[4b−e]

The structure and dynamic behavior of nucleic acid molecules are primary targets of computational studies. In recent years qualitative methodological improvement has been achieved in this field.[8−10] Nanosecond-scale MD studies provide very valuable atomic-resolution information complementing the experimental studies.[11−15] We reported an extended set of nanosecond-scale MD simulations of unusual DNA assemblies: four-stranded intercalated cytosine-rich i-DNA[13a] and guanine quadruplex DNA molecules.[13b] In the present study we carry out a set of unconstrained MD simulations of intercalated zipper-like DNA duplex structures, the focus of our analysis being the role of base stacking, an assessment of the dynamical properties and stability of this structural motif, and the characterization of the properties of sheared G·A mismatches.

Methods

All calculations were carried out using the AMBER5[16] program with the Cornell et al.[17] force field. Where indicated, the simulations were carried out with a recent version of the force field.[18] The nucleic acid molecules investigated were surrounded by a periodic box of water molecules described by the TIP3P potential.[19] The periodic box was extended to a distance of 10 Å from any solute atom. The number of explicit water molecules included in the simulations varied from 1800 to 3100 depending on the solute molecule. The molecules were neutralized by Na^+ cations.[17,20] Cations were initially placed into the most negative locations using Coulombic potential terms with the LEAP module of AMBER5. Simulations were carried out using the Sander module of AMBER5 with SHAKE on the hydrogen atoms with a tolerance of 0.0005 Å and a 2 fs time step. A 9 Å cutoff was applied to Lennard-Jones interactions. Berendsen temperature coupling algorithm (with a time constant of 0.2 ps) was utilized. The nonbonded pair list was updated every 10 steps. Equilibration started by 1000 steps of minimization with the positions of the nucleic acid fixed. After this initial equilibration, all subsequent simulations were performed using the particle mesh Ewald method (PME).[8] The PME charge grid spacing was approximately 1.0 Å, and the charge grid was interpolated using a cubic B-spline with the direct sum tolerance of 10^{-6} at the 9 Å direct space cutoff. To speed up the fast Fourier transform in the calculation of the reciprocal sum, the size of the PME charge grid was chosen to be a product of powers of 2, 3, and 5. For dynamics runs after minimizations initial velocities were assigned from a Maxwellian distribution. Equilibration was continued by 50 ps of PME dynamics, with the position of the nucleic acid fixed. Subsequently, 1000 steps of minimization were carried out with 25 kcal/(mol Å^2) restraints placed on all solvent atoms, continued by 3 ps MD simulation using the same restraint. This equilibration was followed by five rounds of 1000-step minimization with solute restraints reduced by 5 kcal/(mol Å^2) in the course of each round. Then 20 ps of MD followed, with the system heated from 100 to 300 K over 2 ps. Equilibration was continued by several nanoseconds of production simulation. The center of mass velocity was removed during the production dynamics periodically at intervals of 10 ps.[21] No extra processing of the averaged structures obtained by the Carnal module was performed. Solvent and counterion distributions were monitored by binding atom positions from root-mean-square coordinate fit frames over all DNA atoms at 1 ps intervals into

(5) (a) Fiddes, J. C.; Barrell, B. G.; Godson, G. N. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1081−1085. (b) Sims, J.; Dressler, D. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 3094−3098.

(6) (a) Hirao, I.; Ishida, M.; Watanabe, K.; Miura, K. *Biochim. Biophys. Acta* **1990**, *1087*, 199−204. (b) Yoshizawa, S.; Ueda, T.; Ishido, Y.; Miura, K.; Watanabe, K.; Hirao, I. *Nucleic Acids Res.* **1994**, *22*, 2217−2221. (c) Hirao, I.; Naraoka, T.; Kanamori, S.; Nakamura, M.; Miura, K. *Biochem. Int*. **1988**, *16*, 157−162.

(7) (a) Tanikawa, J.; Nishimura, Y.; Hirao, I.; Miura, K. *Nucleic Acids Symp. Ser*. **1991**, *25*, 47−48. (b) Hirao, I.; Nishimura, Y.; Tagawa, Y.; Watanabe, K.; Miura, K. *Nucleic Acids Res*. **1992**, *20*, 3891−3896.

(8) (a) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys*. **1993**, *99*, 8345−8348. (b) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A *J. Am. Chem. Soc*. **1995**, *117*, 4193−4194.

(9) (a) Cieplak, P. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, UK, 1998; pp 1922−1930. (b) Follope, N.; MacKerrel, A. D. *J. Comput. Chem.* **2000**, *21*, 86−104. (c) Hobza, P.; Kabelác, M.; Šponer, J.; Mejzlik, P.; Vondrášek, J. *J. Comput. Chem.* **1997**, *18*, 1136−1150.

(10) (a) Hobza, P.; Šponer, J.*Chem. Rev.* **1999**, *99*, 3247−3276. (b) Šponer J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590−5596.

(11) (a) Young, M. A.; Jayram, B.; Beveridge, D. L. *J. Am. Chem. Soc*. **1997**, *119*, 59−69. (b) Feig, M.; Pettitt, B. M. *Biophys. J*. **1998**, *7*, *5* 134−149. (c) Young, M. A.; Beveridge, D. L. *J. Mol. Biol*. **1998**, *281*, 675−687. (d) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc*. **1998**, *120*, 9401−9409. (e) Cheatham, T. E., III; Srinivasan, J.; Case, D. A.; Kollman, P. A. *J. Biomol. Struct. Dyn*. **1998**, *16*, 265−280. (f) Simmerling, C.; Fox, T.; Kollman, P. A. *J. Am. Chem. Soc*. **1998**, *120*, 5771−5782. (g) Cheatham, T. E., III; Kollman, P. A. In *Structure, Motion, Interaction and Expression of Biological Macromolecules;* Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Schenectady, NY, 1998; pp 99−116. (h) Wigner, R. H.; Liedl K. R.; Ruedisser, S.; Pichler, A.; Hallbrucker, A.; Mayer, E. *J. Phys. Chem. B* **1998**, *44*, 8934−8940. (i) Feig, M.; Pettitt, B. M. *Biophys. J*. **1999**, *77*, 1769−1781. (j) Sprous, D. M.; Young, A.; Beveridge, D. L. *J. Mol. Biol*. **1999**, *285*, 1623−1632. (k) Langley, D. R. *J. Biomol. Struct. Dyn*. **1998**, *16*, 487−509. (l) Cubero, E.; Sherer, E. C.; Luque, F. J.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc*. **1999**, *121*, 8653−8654. (m) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J*. **1997**, *73*, 2313−2336. (n) Štefl, R.; Koča, J. *J. Am. Chem. Soc*. **2000**, *122*, 5025−5033. (o) Lankaš, F.; Šponer, J.; Hobza, P.; Langowski, J. *J. Mol. Biol*. **2000**, *299*, 697−712.

(12) (a) Weerasinghe, S.; Smith, P. E.; Pettitt, B. M. *Biochemistry* **1995**, *34*, 16269−16278. (b) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc*. **1997**, *119*, 1463−1469.

(13) (a) Špačková, N.; Berger, I.; Egli, M.; Šponer, J. *J. Am. Chem. Soc*. **1998**, *120*, 6147−6151. (b) Špačková, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc*. **1999**, *121*, 5519−5534.

(14) (a) Konerding, D. E.; Cheatham, T. E., III; Kollman, P. A.; James T. L. *J. Biomol. NMR* **1999**, *13*, 119−131.

(15) (b) Cheatham, T. E., III; Brooks, B. R. *Theor. Chem. Acc.* **1998**, *99*, 279−288.

(16) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Roos, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Singh, P.; Weiner, P.; Kollman, P. A. *AMBER*; University of California: San Francisco, CA.

(17) 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*, 5179−5197.

(18) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A *J. Biomol. Struct. Dyn.* **1999**, *16*, 849−866.

(19) Jorgensen, W. L.; Chandresekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926−937.

(20) Hardin, C. C.; Ross, W. S. *J. Am. Chem. Soc*. **1994**, *116*, 6080−6092.

(21) Harvey, S. C.; Tan, R.-K. Z.; Cheatham, T. E., III. *J. Comput. Chem*. **1998**, *19*, 726−740.

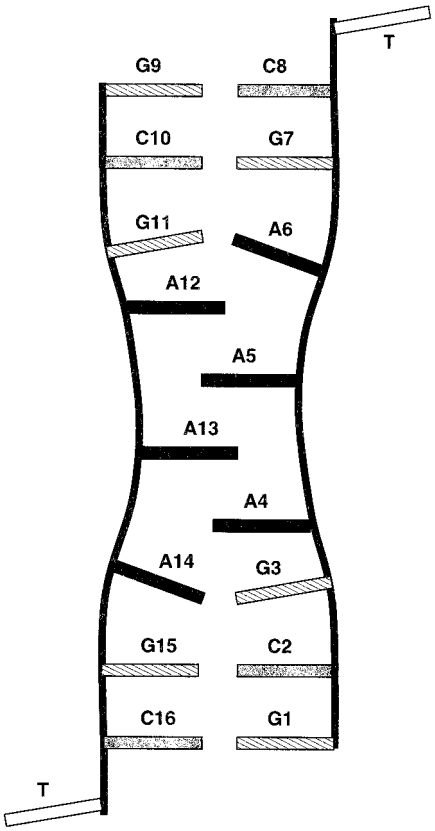

Figure 1. Schematic representation of the intercalated zipper-like DNA duplex, d(GCGAAAGCT)$_2$, adapted from ref 4a. The backbone is drawn in black, and bases are shown as bars. The numbering used throughout the text is indicated. The terminal thymine residues in the crystal structure (open bars) form the three-dimensional lattice and were omitted from the simulations. The central core of the two-fold symmetric molecule consists of four unpaired adenines arranged in a stacked zipper conformation, flanked by nonplanar G·A mismatch pairs and two further Watson−Crick G·C pairs at both ends.

(0.5 Å)^3 grids usually over whole trajectories,[22a] with the aid of the program UCSF MidasPlus.[22b] The calculations of stacking energies were carried out with the Cornell et al. force field and assuming averaged geometries of the simulated molecules. The sugar−phosphate backbone has been replaced by hydrogen atoms, and charges on these hydrogen atoms have been adjusted to achieve neutrality. A dielectric constant of 1 was utilized to evaluate intrinsic base stacking energies. More accurate data could be obtained by ab initio methods;[10] however, for our purpose the force field calculations appear to be sufficiently accurate.[9c]

Results

(A) Molecular Dynamics Simulation of the d(GCG-AAAGC)$_2$ Zipper-like DNA Duplex. Starting Geometry. Initial coordinates were taken from the d(GCGAAAGCT)$_2$ crystal structure solved at 2.1 Å resolution.[4a] The asymmetric unit of this crystal contains one DNA strand, and the antiparallel duplex is generated by a crystallographic dyad axis through the molecule center (Figure 1). The central part of the duplex consists of a zipper-like stem built by four consecutive unpaired and intercalated adenines that stack on top of each other via the base six-membered rings (the bases that are involved in forming this zipper motif are underlined in the following). The unpaired adenine zipper stem is enveloped both on top and on bottom by sheared G·A mismatch base pairs with N3(G)···N6(A) and N2(G)···N7(A) hydrogen bonds. The remaining guanine and cytosine bases of the sequence are arranged at both ends in two adjacent standard Watson−Crick G·C base pairs with essentially B-form geometry. In contrast, the G·A mismatch base pairs are significantly nonplanar. Only the guanine residues of the G·A mismatch base pairs stack on top of the outer adenine residues that are part of the central adenine zipper stem of the molecule, thus giving rise to a well-aligned oligo purine stack of six bases. It has been shown that an isolated (i.e., considering the base pair in the absence of any other contribution) G·A base pair is intrinsically propeller twisted, in contrast to standard GC and AT base pairs.[23] We nevertheless do not rule out that the stacking at the zipper−B-DNA junction may partly contribute to the nonplanarity of the mismatch pair.

The crystal structure was slightly modified for the purpose of the following simulations. The ^{5bromo}cytosines at positions 2 and 10 (Figure 1) were replaced by standard cytosines. The thymine residues at the 3′ termini were omitted. The crystal lattice is further stabilized by a cobalt hexaammine cation. Cobalt hexaammine cations are not expected to be essential for stabilizing the DNA duplex[4a] and were thus not considered in the simulations.

Trajectory. After standard equilibration of the structure, a 9 ns production simulation was carried out. The simulation produced a very stable trajectory with root-mean-square deviation (RMSd) values with respect to the crystal coordinates fluctuating around 2 Å along the entire trajectory (Figure 2a). Figure 2b shows a stereo overlay plot of the averaged theoretical structure and the crystal coordinates. Visual inspection shows a very good agreement between the two structures that is confirmed by an RMSd value of 1.7 Å between them. Figure 2c shows the RMSd of the simulated structure with respect to its averaged (1−5 ns) structure.

Flexibility of the Molecule. Thorough inspection of the conformational changes along the trajectory reveals only one larger alteration between the theoretical structure and the crystal coordinates, occurring in the region between adenine residue A12 and the adjacent G11·A6 mismatch base pair. Adenine A12 is slightly displaced from its original position stacked between G11 and A5 in a way that it approached adenine A6. The vertical interaction between the six-membered rings of guanine G11 and adenine A12 is disrupted, and the aromatic heterocycle of adenine A12 comes to stack on top of the guanine N2 and adenine N6 amino groups of the residues forming the G·A mismatch (Figure 3). This alteration propagates moderately, leading to local conformational variations in the zipper core architecture. Adenine A12 is slightly skewed from the zipper axis, and the two adjacent adenines A5 and A13 are tilted from their previously planar orientation with respect to each other by approximately 24°. This together with the displacement of adenine A12 further contributes to a bending of the adjacent base-paired segment with respect to the intercalated central zipper core of the molecule. This local deformation developed at 0.8 ns and persisted until the end of the simulation without any further conformational adaptation. Therefore, on our time scale, it appears to be a *static* deformation. This deformation was not observed at the other end of the duplex. Thus, despite the starting structure having 2-fold symmetry, the simulation provides a slightly asymmetrical averaged structure with an RMSd value between the two strands of 2.4 Å.

(22) (a) Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 4805−4825. (b) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *J. Mol. Graph.* **1988**, *6*, 13−27.

(23) Šponer, J.; Florián, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 827−833.

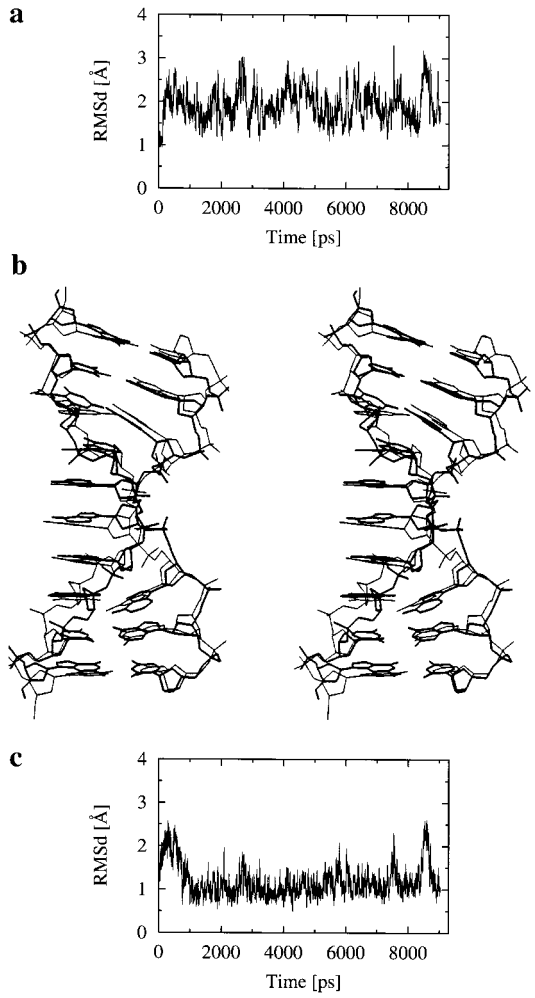

Figure 2. **(**a**)** RMS deviation along 9 ns MD trajectory of d(GCGAAAGCT)$_2$ between the theoretical structure and the crystal coordinates (unpaired residues involved in the zipper core are underlined). **(**b**)** Stereo overlay plot of the d(GCGAAAGC)$_2$ crystal (thin lines) and averaged theoretical (1−5ns) coordinates (thick lines). **(**c**)** RMS deviation along 9 ns MD trajectory of the simulated d(GCGAAAGC)$_2$ molecule calculated with respect to a theoretical structure averaged over nanoseconds 1−5 of the simulation.

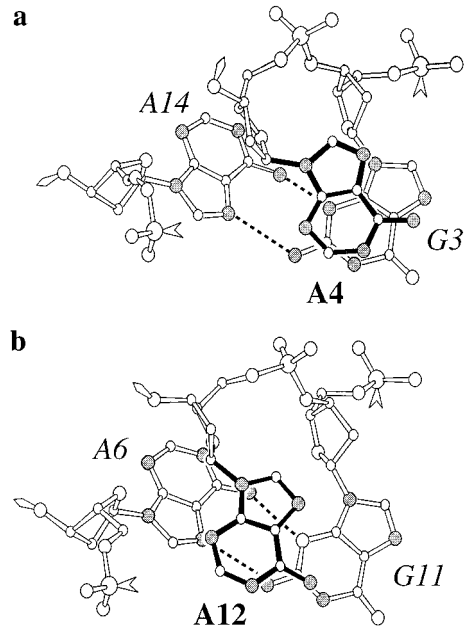

Figure 3. Ball-and-stick representation of the base stacking geometry at the junctions between unpaired zipper core and the base-paired segments in the averaged simulated structure of d(GCGAAAGC)$_2$. Nitrogen atoms are shaded, hydrogen bonds are drawn with dashed lines, and strand polarity is indicated by arrows. **(**a**)** Stacking between adenine A4 (solid black bonds) and the adjacent G3·A14 mismatch base pair (open bonds) showing overlay of the heterocycles of G3 and A4. **(**b**)** Stacking of the adenine residue A12 (solid black bonds) on the mismatch pair G11·A6 (open bonds) at the opposite end of the zipper core. The aromatic heterocycle of adenine A12 stacks on top of the guanine N2 and adenine N6 amino groups.

The molecule is composed of three basically rigid segments: the zipper-like core motif in the center and the two Watson−Crick base-paired double-stranded B-form parts at either end of this central segment. Bending and geometrical fluctuations occur at the junctions between these rigid segments of the molecule. These junctions coincide with the G·A mismatches. Figure 4 shows the significant temporary bending that occurred during the simulation. This can be assumed to represent a major statistical bending mode of this particular duplex.

Base Stacking. We calculated the base stacking interaction energies for all stacked steps along the structure (Table 1, see also Methods). The base stacking interactions in the individual steps of the crystal structure were in a range from −6.0 to −21.6 kcal/mol. The smallest base stacking energy value was calculated between the central two adenine residues (A5 and A13). This base step is characterized by the smallest overlap of the adjacent base heterocycles. Nevertheless, the base stacking energy of −6 kcal/mol still corresponds to the low-energy region of a stacked adenine dimer[10b] and represents approximately 40−70% of the stacking energy between consecutive Watson − Crick base pairs in regular DNA duplexes.[24] The largest base stacking energy value was found between the G·A mismatch and the adjacent regular Watson−Crick G·C base pair. The base stacking energy of approximately −22 kcal/mol is better compared to base stacking in standard B-DNA steps with canonical base pairs[24] and has a significant interstrand contribution of −8 kcal/mol.

The corresponding stacking energies in the averaged theoretical structure are very similar to the energies based on the crystal coordinates. An improvement was noted for the stacking of the central adenines where an energy of −9.7 kcal/mol was calculated for the theoretical structure, due to a more optimal overlay of the aromatic rings. A further difference in base-stacking energy was calculated between residues A12 and A5 with a reduced energy of only −5.9 kcal/mol as compared to ca. −8 kcal/mol found in the crystal geometry. This reflects the partial unstacking and bending between adenines A12 and A5 (see above).

Both crystal and simulated structures show sugar−base stacking[2g,3e,4a,e,24] between O4′ of the outer adenine residue of the zipper core and the adenine residue involved in the G·A mismatch pair. The interaction energy of such contacts is known to be about −3 to −4 kcal/mol.[24]

Backbone Geometry. The glycosidic torsion angles of the crystal structure are all anti with values in the high anti region for adenines involved in G·A mismatches and for guanines at the 5′ end. During the simulation substantial changes were not observed and the sugar moieties consistently remained in the anti conformation with glycosidic bond angle values between 207 and 274°. Most of the sugar rings in the crystal structure adopt C2′-endo or C1′-exo pucker (C1′-exo is a neighbor of

(24) Šponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. *Biophys. J.* **1997**, *73*, 76−87.

Figure 4. Conformational flexibility of the zipper-like DNA duplex d(GCGAAAGC)$_2$. (Left) The crystal structure in a side view looking at the stacked adenine zipper core. (Center) The theoretical molecule averaged over 1−5 ns (see Figure 2a). A slight bending originating at one G·A mismatch (on top of the zipper core in this representation) can be noted. (Right) Temporary bending occurring in the direction of the static bend, averaged over nanoseconds 8.5−8.6 of the trajectory.

Table 1. Net Base Stacking Energies (kcal/mol) Evaluated by Cornell et al. Force Field for the Crystal Coordinates and Averaged Theoretical Structure of d(GCGAAAGC)$_2$[a]

	crystal coordinates		theoretical structure	
stacked bases/ base pairs	total stacking	electrostatic term	total stacking	electrostatic term
G1C16/C2G15	−17.09	0.31	−17.37	0.42
C2G15/G3A14	−21.62	−6.10	−22.72	−5.64
G3A14/A4	−11.32	−1.85	−11.60	−1.94
A4/A13	−7.95	−0.66	−8.37	−1.01
A13/A5	−6.02	−0.05	−9.67	−1.65
A5/A12	−7.95	−0.66	−5.91	−0.22
A12/G11A6	−11.32	−1.85	−11.00	−2.69
G11A6/C10G7	−21.62	−6.10	−23.12	−5.30
C10G7/G9C8	−17.09	0.31	−17.51	0.26

[a] For numbering of residues see Figure 1.

C2′-endo in the pseudorotation circle), with the exception of the sugar rings of the adenine residues A4 and A12 at the ends of the zipper core, which were puckered C4′-exo. During the simulation, we observed repuckering of the sugars uniformly to C2′-endo. Only residue A4 at one end of the adenine zipper segment adopted C3′-endo for the whole trajectory, very close to the crystal value (C4′-exo). The corresponding adenine residue A12 at the opposite end of the zipper core also repuckered to C2′-endo after 0.8 ns and remained in this conformation until the end of simulation. The Cornell et al. force field is slightly biased toward a lower phase for C2′-endo pucker values while also underestimating the energy barrier for the C2′-endo−C3′-endo repuckering. This may explain certain differences between the X-ray and theoretical geometries.[18]

Interphosphate Distances and Phosphate Clustering. In the crystal, short interstrand distances were found in the central zipper core between phosphates P5 and P14 and the symmetry-related phosphates P6 and P13 (6.6 Å). The distance between P5 and P13 was 9.8 Å and between P6 and P14 8.7 Å. At the beginning of the simulation, the phosphate groups P6 and P13 approached each other more closely (6.0 Å) and remained spaced at this distance for 0.5 ns. At around 0.6 ns, the distance

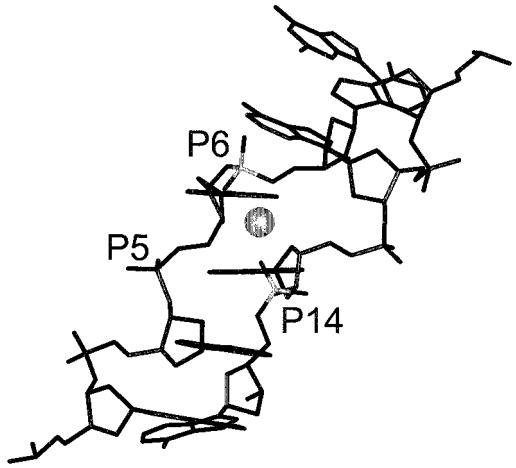

Figure 5. Clustering of phosphate groups observed during the simulation, stabilized by a cation (shown as a ball).

slowly increased to a value of 8 Å and remained stable at this value until the end of the trajectory. Instead, P6 contacted P14 (average distance during the trajectory of 6.4 Å) with a minimum distance of 5.8 Å in the interval from 2 to 4.5 ns. The closest interphosphate distance, however, involved phosphates P5 and P14 with an average distance of 6.0 Å. Thus, a tight clustering of P6, P14, and P5 was observed, stabilized by sodium cations that alleviate the electrostatic repulsion between these anionic groups. Sodium cations were in direct contact with at least one phosphate group in this region for around 55% of the simulation time, and in the remainder water molecules occupied the positions of the ions. A sodium cation was associated with phosphate group P14 and in addition with one phopshate oxygen of the P6 phosphate group within the period of 2.0−4.5 ns (Figure 5). Then, this sodium moved away from P6 and approached a phosphate oxygen of phosphate group P5 where it remained for an additional 0.5 ns, while being also associated with P14. At 3.8 ns the ion was found equidistant between the phosphate oxygens for 100 ps. The typical value of the Na^+

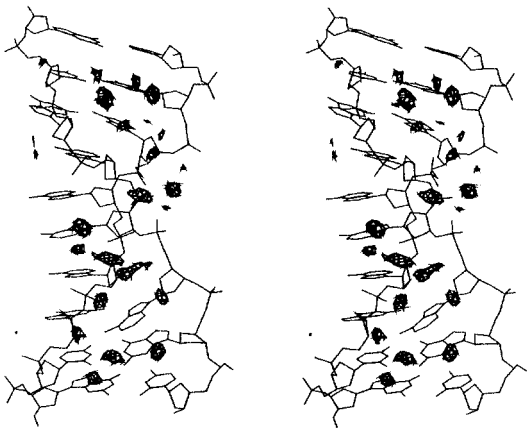

Figure 6. Stereo overlay plot showing the overall hydration of the simulated molecule. (Water density, contour level 100 hits per 0.5 Å^2, i.e., ca. 4 times bulk water density.)

phosphate oxygen distance was around 2.4 Å, indicating direct coordination. Further direct DNA−cation contacts were observed around 5 ns (near P6, residence time 100 ps) and around 7 ns (near P5, residence time 100 ps). While the sodium ion is localized between phosphate oxygens from P14 and P6, the space between P14 and P5 is filled with water molecules, and vice versa.

Intrastrand distances between phosphates in the averaged theoretical structure are within the range of 5.9−7.1 Å, in good agreement with the crystal coordinates.

Hydration of the Molecule. We have analyzed the overall hydration pattern in the vicinity of the DNA molecule (Figure 6). The structure is mainly hydrated in the grooves and near the phosphate groups. The highest density (∼7.2 times bulk water density) appears in the central region of the zipper groove.[25] As described above, water molecules and cations form bridges between the three clustered phosphates P5, P14, and P6. Additional ordered hydration sites close to that region were found with residence times typically around 0.05−0.3 ns.[11j,26] A distinct hydration site was found between oxygens O5′ and O4′ of adenine A5 and the N3 nitrogen of adenine A13. The occupancy of this hydration site is close to 100% along the trajectory. During the entire simulation only five water molecules exchanged position here, with very long individual residence times ranging from 0.5 to as much as 5.1 ns.

Cation−Base Interactions in the Zipper Core. Analysis of the trajectories revealed interactions of sodium cations with nitrogens N1 and N7 of the unpaired zipper adenines. For around 45% of the simulation time span at least one sodium cation was observed to be in direct contact with one of the zipper bases. The cations reside at N7 positions twice as frequently as in the vicinity of the N1 nitrogens. Residence times of these cations with respect to individual nitrogen atoms were between 0.1 and 0.2 ns. N1 and N7 nitrogen atoms in the zipper are alternately placed in one line along the *z*-axis, allowing continuous vertical movements of cations between adjacent adenines. An example was found in the interval between 1.1 and 2.2 ns. A sodium cation localized in the vicinity of the N1 nitrogen of adenine A5 moved after 0.1 ns to a position near the N7 atom of the adjacent base (A13) and after 0.7 ns back to the original position. Then, after 0.1 ns, it relocated to a position in the vicinity of the N7 nitrogen of adenine A12. Within the next 0.1 ns the cation returned again to the initial position for an additional 0.2 ns. We have ocassionally observed also association of more than one cation to the zipper motif. For example, at 4.3 ns, three cations were coordinated for 80 ps to the zipper bases.

Further Simulations of d(GCGAAAGC)$_2$. We have carried out two additional 3.0 ns simulations of the d(GCGAAAGC)$_2$ zipper-like duplex structure, with the aim to see whether local variations with tilted adenines and the phosphate cluster developed also in these simulations. One simulation was carried out with the same force field as the 9 ns simulation described above but using a slightly different box size and equilibration protocol. This simulation did not show any phosphate clustering or tilting of adenines, and the two strands remained symmetrical. The second simulation was then carried out using a recently released modified Cornell et al. force field, which has a different tuning of the sugar pucker parameters.[18] During this simulation, the local conformational variation (substate) developed again, including tilted adenines, repuckering, clustered phosphates, ion bridges, and long-residing water molecules. However, in contrast to the above-described 9 ns simulation, the bending here occurred at the opposite end of the molecule.

Finally, we extended the original 9 ns simulation described above by 1.1 ns at an elevated temperature of 400 K. In sharp contrast to our previous studies on four-stranded DNA conformations,[13] the elevated temperature destabilized this zipper-like DNA duplex structure. The RMSd value during the first 0.4 ns remained around 2 Å. Then, a rapid increase of the RMSd value was noted reaching 4.1 Å, caused by an out-of-plane rotation of the inner two stacked adenine residues of the zipper core. These two adenines remained stacked on top of each other while temporarily adopting a close to perpendicular position with respect to the other bases. After a subsequent 0.08 ns, this large geometrical deformation was entirely repaired. At ca. 0.65 ns a G·A mismatch base pair was disrupted. Its adenine looped away from the molecule and became stacked on the zipper core from the outside.

(B) MD Simulations of d(GCGGGAGC)$_2$, d(GCGCCAGC)$_2$, d(GCGAAAGC)$_2$, and d(GAAA)$_2$ Duplexes. We carried out a set of simulations with modifications in the DNA sequence of the zipper molecule in order to gather additional insight into the stability of this duplex DNA motif.

G-Zipper: d(GCGGGAGC)$_2$. We first replaced the four zipper adenine residues of the original molecule by guanine residues ("G-zipper"). For all other residues, the crystal coordinates were used as a starting model. Four nanoseconds of production simulation was carried out. The d(GCGGGAGC)$_2$ structure was stable with RMSd around 2 Å along the whole trajectory compared to the starting coordinates. To test the internal stability of the simulated structure, we calculated the RMSd along the trajectory with respect to a theoretical G-zipper structure avaraged from 1 to 4 ns. This RMSd value oscillated around 1.1 Å, indicating that the structure is internally as stable as the original molecule with an adenine zipper. In fact, the duplex molecules with either adenines or guanines in the zipper core adopted rather similar geometries (Figure 7). During the entire simulation we observed only one structural fluctuation. At around 2.5 ns the guanine residue G4 that is initially stacked on guanine G3 shifted toward and partially stacked on adenine A14. This temporary dislocation did not propagate to the rest of the molecule, and after 50 ps G4 moved back to the original position.

The backbone geometry of the theoretical d(GCGGGAGC)$_2$ structure is similar to the original molecule formed by d(GC-

(25) This type of analysis highlights the most ordered hydration sites. Lack of localized water density does not mean necessarily a lack of hydration, but there are no ordered hydration sites.[11i]

(26) Phan, A. T.; Leroy, J.-L.; Gueron, M. *J. Mol. Biol.* **1999**, *286*, 505−519.

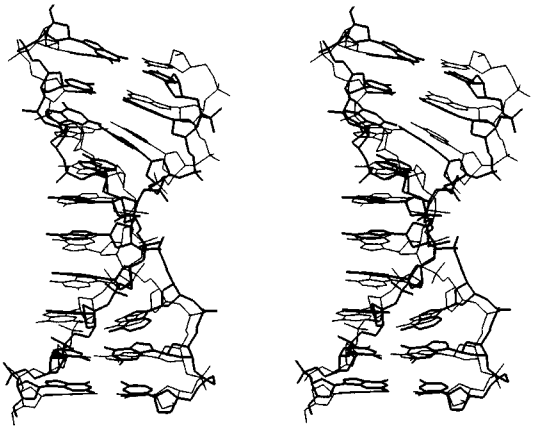

Figure 7. Stereo overlay plot of the simulated purine-rich zipper structures d(GCGAAAGC)$_2$ (thick lines) and d(GCGGGAGC)$_2$ (thin lines).

Table 2. Net Base Stacking Energies (kcal/mol) Evaluated by Cornell et al. Force Field for the Initial and Averaged MD Structures of G-Zipper d(GCGGGAGC)$_2$[a]

	starting coordinates		theoretical structure	
stacked bases/ base pairs	total stacking	electrostatic term	total stacking	electrostatic term
G1C16/C2G15	−17.09	0.31	−17.30	0.56
C2G15/G3A14	−21.62	−6.10	−22.89	−5.50
G3A14/G4	−7.52	2.33	−7.26	2.92
G4/G13	−8.59	0.06	−5.49	3.42
G13/G5	−13.10	−5.44	−9.65	−2.22
G5/G12	−8.59	0.06	−12.31	−4.08
G12/G11A6	−7.52	2.33	−7.33	3.05
G11A6/C10G7	−21.62	−6.10	−22.26	−5.29
C10G7/G9C8	−17.09	0.31	−17.81	−0.01

[a] Initial geometry was based on the crystal structure of d(GCGAAAGCT)$_2$ with the zipper adenines replaced by guanines (see text).

GAAAGC)$_2$. The glycosidic torsion angles are all anti except for the central zipper residues (G5, G13), which are in the syn conformation. The sugar rings are again puckered C2′-endo except for the outer zipper residues (G4, G12), which adopt C3′-endo sugar pucker. The hydration pattern of the G-zipper structure is similar to the pattern described above for the structure containing the adenine zipper core. The central core in the G-zipper structure shows a somewhat less structured hydration which can be partly explained by the different geometry of the sugar−phosphate backbone in the G-zipper, with close contacts between phosphate groups being absent here. The interstrand phosphate distances in the averaged theoretical structure were in a range between 7.0 (P6 to P13) and 10.0 Å (P6 to P14). The G-zipper shows a decreased propensity of direct cation−base contacts in the zipper region. For 88% of the simulation timespan, direct contacts between a cation and a zipper guanine were not observed. For the remainder, one Na^+ cation was coordinated to the zipper bases.

Table 2 shows stacking energies along the d(GCGAAAGC)$_2$ molecule for both the initial coordinates and the averaged (1−4 ns) theoretical structure. Guanine has a significantly more polar electrostatic potential compared to adenine. It results in a greater contribution of the electrostatic term and a larger geometry dependence of the stacking energy in guanine stacks.[10,24] The main difference between the stacking energies in the G-zipper and A-zipper starting structures occurs at the G·A mismatch stacked on the outer zipper base. In the molecule containing the adenine zipper core, the stacking is considerably better (−11.3 kcal/mol) than for the molecule with the guanine zipper core (−7.5 kcal/mol). This is due to an unfavorable electrostatic interaction between the two almost parallel guanines G3 and G4. On the other hand, the stacking of the central two bases in the G-zipper molecule yields −13.1 kcal/mol in the initial geometry, compared to −6 kcal/mol at the corresponding base step for the A-zipper crystal coordinates. The difference originates again in the electrostatic term.

Stacking energies for the averaged theoretical G-zipper molecule outside the zipper core are not significantly altered along the simulation trajectory. The stacking energy between the central two guanine residues in the zipper core deterioriated to −9.7 kcal/mol in the course of the simulation. A large stacking energy difference developed between the initially symmetry-related G4−G13 (−5.5 kcal/mol) and G5−G12 (−12.3 kcal/mol) zipper steps. The stacking of guanine G5 on guanine G12 shows a smaller overlay of the bases, but the electrostatic term is stabilizing, −4.1 kcal/mol. The geometry of the G4−G13 stack that exhibits a large overlay of the six-membered ring of one guanine stacked on the imidazole ring of the adjacent guanine and vice versa is less favorable, with a repulsive electrostatic term of +3.4 kcal/mol. The underlying geometrical alterations are quite small, and the RMSd value between the two strands in the averaged G-zipper structure is only 1 Å.

Unstable Pyrimidine Zipper: d(GCGCCAGC)$_2$ and d(GCGTTAGC)$_2$. We further carried out a 2.5 ns simulation of a hypothetical structure with cytosines in the zipper core, d(GCGCCAGC)$_2$. We observed a fast destabilization of the zipper core geometry, with central cytosines becoming unstacked and even temporarily rotating out of the stacking plane. The paired regions of the duplex molecule including the G·A mismatch in contrast remained stable. We have also carried out a 3.0 ns simulation of the d(GCGTTAGC)$_2$ molecule. This simulation produced a slightly more stable trajectory than d(GCGCCAGC)$_2$. During the first 2 ns, the inner zipper thymines became partially unstacked while the remainder of the molecule closely resembled the d(GCGAAAGC)$_2$ structure. At 2 ns the inner thymine residue T13 in the zipper core rolled somewhat, leading to a loss of stacking with the adjacent thymine T4. At the same time, partial stacking was established between the two inner thymines T13 and T5 with the exocyclic O2 keto oxygen of one thymine stacked on the aromatic ring of the other and vice versa. However, at 2.6 ns the zipper core was largely disintegrated.

Isolated Adenine Zipper: d(GAAA)$_2$. We have carried out a simulation (5.0 ns) of a truncated structure with the four Watson−Crick G·C base pairs removed from the molecule. Both G·A mismatches were disrupted within 0.6 ns. The guanine residues remained stacked on the adenine zipper core while the adenine residues previously involved in the mismatch base pairs stacked on top of the guanine residues. We have evidenced a structure characterized by extensive stacking of all eight bases, though at the end of the simulation the stacking of the two innermost adenines was disrupted, resulting in a structure consisting of two almost perpendicular four-base stacked segments.

Discussion and Conclusions

We have carried out altogether 30 ns of unrestrained MD simulations of DNA duplex structures containing a central segment of unpaired residues that are stacked on top of each other, forming an intercalated zipper core.

The simulations of the d(GCGAAAGC)$_2$ molecule containing a zipper motif of four unpaired and intercalated adenines yield stable trajectories, in very good agreement with the crystal structure. The zipper-like duplex molecule appears to be

considerably more rigid compared to DNA duplexes with regular Watson−Crick pairing geometry.[11] The d(GCGAAAGC)$_2$ molecule consists of three distinct segments: the central zipper core formed by the intercalated adenine residues and the Watson−Crick G·C base pairs at either end of the antiparallel duplex, with essentially B-form geometry. These rigid segments are separated by sheared G·A mismatch base pairs which are revealed as flexible junctions in the molecular structure.

It is well established that single isolated sheared G·A base pairs have never been found in bimolecular DNA duplexes and are rather associated with the reversal of the chain direction at the base of hairpin loops.[4e] Bimolecular sheared G·A base pairs always occur in a tandem arrangement, with or without an intervening intercalation motif.[4e] It has been suggested that the major factor influencing the stability of sheared G·A mismatch pairs is stacking between base pairs supplemented by sugar-base stacking.[4e] This view is fully supported by our results. The base stacking between the mismatch pair and adjacent G·C base pairs is considerably better than base stacking in any base pair step with two standard Watson−Crick pairs.[24] Further, a sugar-base stacking interaction exists between the zipper core and the mismatch base pair. The simulation of d(GAAA)$_2$, with the zipper core only enveloped by the mismatch pairs, provides further evidence for stacking being the major stabilizing contribution for G·A pairs: the sheared G·A base pairs swiftly desintegrate and associate with the rest of the structure via stacking.

The sheared G·A mismatch base pair, in contrast to G·C and A·T Watson−Crick base pairs, is intrinsically (i.e., being entirely isolated) nonplanar, and we suggest that this is an important feature to understand the structural role of G·A mismatch pairs in DNA molecules.[23] However, in RNA the intrinsic nonplanarity of the sheared G·A mismatch appears to be suppressed, possibly due to H-bonding between 2′-oxygens and adenine N6 amino groups.[27]

The MD simulations of the d(GCGAAAGC)$_2$ duplex molecule reproducibly show an interesting and, to our knowledge, hitherto unobserved development of a local conformational variation of a DNA molecule. The local variation is stabilized by a concerted interplay of repuckering of sugar moieties, rearrangement of base stacking, and crowding of specific phosphate groups accompanied by specific hydration interactions and sodium cation coordination. The simulated d(CGCAAAGC)$_2$ duplex structure shows, with respect to the crystal coordinates, an enlarged interphosphate distance between phosphate groups P6 and P13, associated with a C3′-endo to C2′-endo pucker transition of the sugar moiety of adenine A12, resulting in its partial unstacking (compare Figure 3b). Subsequently, this leads to a close clustering of phosphates P5, P14, and P6. This crowding of the negatively charged phosphate groups is stabilized through hydration by bridging water molecules with a significant occupancy of sodium cations (see Figure 5). This pattern is complemented by several closely spaced and highly ordered hydration sites. On the basis of this simulation we suggest that an interplay of different contributions acting simultaneously is behind many local conformational variations observed in nucleic acids, and these cannot be rationalized by considering just a single contribution. Our observations support a view that closely spaced phosphate groups can be energetically stabilized through mixed sodium−water bridges.

The observation of a hydration site with an average water residency time about 2 ns is rather unprecedented. Typical water residence times reported in MD simulations as well as in experiments are an order of magnitude shorter.[11i,26] However, this anomalously long residence time concerns a single hydration site clearly involved in a very specific local geometry, while other hydration sites in our study do not show anomalous residence times. To identify long-residence hydration sites, one should carefully monitor individual water molecules while the simulation should be well beyond the 1 ns scale. Only a fraction of MD studies reported so far satisfies these criteria, and we suggest that long-residence hydration sites may well be more frequently encountered in the future, with long simulations becoming routine.

The simulations suggest that the zipper core is partly stabilized by a partial penetration of monovalent ions into the primary hydration shell around the zipper adenines. This effect resembles the penetration of ions into the minor groove of ApT B-DNA steps first observed by Young and Beveridge[11a,m] and confirmed by others.[11n] Fractional occupation of sodium cations in the minor groove of B-DNA −AATT− sequences has also been reported in a crystallographic study,[28] though this particular result has not been unambiguously accepted by other crystallographic groups.[29,30] Penetration of ions into the DNA first hydration shell has also been evidenced by NMR.[31]

The 9 ns simulation of the initially symmetrical d(CGCAAAGC)$_2$ duplex resulted in a slightly nonsymmetrical molecule. The asymmetry that exists between the two termini in the simulated molecule and lack of any interconversion between the two substates mean that the structure is not fully equilibrated on the simulation time scale. It cannot be ruled out that the energetically favorable arrangement of the molecule is nonsymmetrical, i.e., when the above-described local variation develops at one end of the molecule only, as seen in the simulation. Then, in solution and on a sufficiently extended time scale, interconversion between two nonsymmetrical structures would still lead to a symmetrical averaged assembly. The packing of the molecule into the three-dimensional crystal lattice with a true dyad could prevent the experimental observation of this conformational variation. Similar restrictions imposed by crystal symmetry have been suggested for other local conformational variations in B-DNA.[32]

The zipper core architecture is characterized by attractive base stacking. Base stacking is one of the major sources of stability of this DNA molecule. All simulations presented clearly illustrate that adenines preferentially self-associate via stacking in an aqueous environment.[33]

The analogous d(GCGGGAGC)$_2$ zipper-like molecule with guanines is entirely stable on a nanosecond scale, in full agreement with NMR studies on d(−GGA−)$_2$ duplexes.[4e] However, the zipper-like duplex architecture is unstable when the purine bases in the zipper core are replaced by pyrimidines though the thymine residues have a more pronounced tendency to associate via stacking compared to cytosines. Thymines are considerably less polar than cytosines, rendering them more suitable for hydrophobic stacking interactions. NMR data show

(27) (a) Correll, C. C.; Freeborn, B.; Moore, P. B.; Steitz, T. A. *Cell* **1997**, *91*, 705−712. (b) Leontis, N. B.; Westhof, E. *Q. Rev. Biophys.* **1998**, *31*, 399−455.

(28) Shui, X.; McFail-Isom, L.; Hu, G. H.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.

(29) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1981**, *103*, 470−471.

(30) Chiu, T. K.; Kaczor-Grzeskowiak, M.; Dickerson, R. E. *J. Mol. Biol.* **1999**, *286*, 589−605.

(31) Hud, N. V.; Feigon, J. *J. Am. Chem. Soc.* **1997**, *119*, 5756−5757.

(32) Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709−714.

(33) (a) Leng, M.; Felsenfeld, G. *J. Mol. Biol.* **1966**, *15*, 455−476. (b) Solie, T. N.; Schellman, J. A. *J. Mol. Biol.* **1968**, *33*, 61−77. (c) Berger, I.; Cai, L.; Liqing, C.; Rich, A. *Biopolymers* **1998**, *44*, 257−267. (d) Cai, L.; Chen, L.; Raghavan, S.; Ratliff, R.; Moyzis, R.; Rich, A. *Nucleic Acids Res.* **1998**, *26*, 4696−705.

that (−GTA−) sequences form, to some extent, $(-G\underline{T}A-)_2$ duplexes containing *two-thymine* zippers.[4e] We believe this is not in disagreement with our observation of a rather unstable *four-thymine* self-intercalation. NMR studies show no duplex formation of d(−GCA−) sequences.[4e]

To understand the base stacking differences in the adenine versus guanine zipper molecules (see Tables 1 and 2), it is useful to consider the general difference of stacking properties of guanine and adenine.[10,24] Adenine has a low dipole moment. Stacking of adenines is dominated by isotropic and overlap-dependent dispersion attraction with only a weak electrostatic contribution. Guanine on the other hand is very polar and there is a strong and geometry-dependent electrostatic contribution in the G−G stacks. This explains why in the case of the $d(GCG\underline{GG}AGC)_2$ zipper a rather small geometrical asymmetry between G4−G13 and G5−G12 stacks gives rise to a large stacking energy difference and why the structure with smaller overlap of bases is considerably more stable. Let us give two additional examples of the importance of the electrostatic contribution to stacking of guanines. The stability of −GNNA− hairpin loops strongly depends on the polarity of the C·G base pair lying underneath the sheared G·A mismatch. The G−GGCA−C hairpin is much less stable than C−GGCA−G,[34a] due to an unfavorable electrostatic interaction between two almost parallel stacked guanines in the former sequence.[34b] Similarly, GpG stacking in a B-form duplex is characterized by significant electrostatic repulsion of both intrastrand homonucleotide stacks (G−G, C−C) of consecutive and almost parallel polar guanine and cytosine bases, making stacking in this step considerably weaker than in all other steps.[24] This could explain the unusual properties observed in G-tracts,[11o,35] such as high base pair opening rates[35a] and stacking geometry significantly shifted toward the A-form.[35b,c] *The stacking energy cannot be evaluated merely on the basis of the geometrical overlap of bases, as is common in experimental studies, since in this case the electrostatic portion of the stacking is neglected.* On the other hand, the solvent screening effects compensate for the net electrostatic contribution to the base stacking, and the effect of a polar solvent on stacking of polar bases such as guanine is considerably larger than for adenine.[36] The solvent screening effects can shift the stacking away from the presumably optimal arrangement.[36] This may explain why we observed rather an improvement of the net A−A stacking in the $d(GCG\underline{AA}AGC)_2$ MD simulation, while the net G−G base stacking appears to somewhat deteriorate (compared to the starting geometry) during the $d(GCG\underline{GG}AGC)_2$ zipper simulation.[37]

Let us briefly discuss the major limitations of our simulations, namely, the time scale of the simulations and the quality of the force field. The Cornell et al. force field provides a good description of base stacking and H-bonding of DNA bases.[9c] The original Cornell et al. force field possesses a somewhat imbalanced description of the sugar pucker, and a modified version of the force field was recently released.[18] We employed both variants of the Cornell et al. force field in our simulations and did not observe any significant force field dependence of the results. A further approximation concerns planarization of the amino groups of nucleobases by current force fields. The amino groups of isolated nucleic acid bases are intrinsically nonplanar and very flexible.[10a] This could allow for a formation of out-of-plane H-bonds and/or amino acceptor interactions[23,32,38] of unpaired amino groups in the zipper region which would not be captured by the simulation. Finally, the lack of a polarization term may influence the residence times and occupancies of the direct binding of cations to various chemical groups of the solute.[13b] Nevertheless, the sites for specific cation binding are determined by the electrostatic terms and should be predicted properly.

It is illustrative to compare the intercalated zipper-like DNA duplex with another intercalation motif, namely the i-DNA quadruplex. The two molecules have many similarities as well as striking differences and complementarities. The i-DNA stem is formed exclusively by cytosine intercalation and is characterized by a repulsive intrinsic base stacking due to a charge−charge repulsion.[13a,39] The zipper-like duplex intercalation shows considerably larger sequence variability, with a clear preference for purine bases over pyrimidines, and attractive intrinsic base stacking. In the i-DNA quadruplex molecule, two parallel duplexes held together by hemiprotonated $C{\cdot}C^+$ base pairs interdigitate into each other in a way that the $C{\cdot}C^+$ base pairs from one parallel duplex are intercalated between two $C{\cdot}C^+$ base pairs from the other parallel duplex, thus giving rise to a four-stranded assembly.[2] The architecture of the $d(GCG\underline{AA}AGC)_2$ zipper core actually would strongly resemble the geometry of half of the i-motif DNA. In fact, the i-motif qaudruplex can be considered as being built of an antiparallel "zipper core" of consecutively stacked cytosines that is hydrogen bonded by $C{\cdot}C^+$ self-pairs to an identical cytosine "zipper core". From our simulations, it is clear that the adenine zipper core, which is unpaired in contrast to four-stranded i-DNA, is a stable molecule on a nanosecond scale, provided that the zipper core is flanked by the Watson−Crick paired segments in addition to the sheared G·A pairs. We propose that the zipper-like architecture of consecutively stacked, mutually intercalated bases per se is a structural motif that may intrinsically require further energetical contributions to be stable. In the case of the cytosine-rich i-DNA motif, this additional stabilization energy is provided by the association of two zipper cores via hemiprotonated $C{\cdot}C^+$ base pairs. In the case of the adenine zipper, in contrast, additional stabilizing energy is provided by the paired segments enveloping the zipper core.

Acknowledgment. This study was supported by the grant A4040903 by IGA AS CR and by the grant 203/00/0633 by GA CR. I.B. is supported by a Liebig fellowship from the Fonds der Chemischen Industrie (FCI, Germany). All calculations were carried out in the Brno Supercomputer Center and Prague Supercomputer Center. We thank Dr. William Shepard for providing the X-ray coordinates of the $d(GCGAAACGT)_2$ zipper-like duplex structure.

JA000460S

(34) (a) Sandusky, P.; Wooten, E. W.; Kurochkin, A. V.; Kavanaugh, Th.; Mandecki, W.; Zuiderweg, E. *Nucleic Acids Res.* **1995**, *23*, 4717−4725. (b) Van Dongen, M. J. P.; Mooren, M. M. W.; Willems, E. F. A.; Van der Marel, G. A.; Van Boom, J. H.; Wijmenga, S. S.; Hilbers, C. W. *Nucleic Acids Res.* **1997**, *25*, 1537−1547.

(35) (a) Dornberger, U.; Leijon, M.; Fritzsche, H. *J. Biol. Chem.* **1999**, *274*, 6957−6962. (b) Ng, H.; Kopka, M. L.; Dickerson, R. E. *Proc. Natl. Acad. Sci U.S.A.* **1999**, *97*, 2035−2039. (c) Trantírek, L.; Štefl, R.; Vorličková, M.; Koča, J.; Sklenář, V.; Kypr, J. *J. Mol. Biol.* **2000**, *297*, 907−922.

(36) (a) Friedman, R. A.; Honig, B. *Biophys. J.* **1995**, *69*, 1528−1536. (b) Florián, J.; Šponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884−892.

(37) Note, however, that stacking of bases buried within DNA is not equivalent to stacking of two bases in water, and the solvent screening effects and expression of the electrostatic component of stacking depend on the particular DNA architecture.[13a]

(38) (a) Luisi, B.; Orozco, M.; Šponer, J.; Luque, F. J.; Shakked, Z. *J. Mol. Biol.* **1998**, *239*, 1123−1136. (b) Vlieghe, D.; Šponer, J.; Van Meervelt, L. *Biochemistry* **1995**, *34*, 16443−16451.

(39) Šponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 695−707.