Stereochemical, Structural, and Thermodynamic Origins of Stability Differences between Stereoisomeric Benzo[*a*]pyrene Diol Epoxide Deoxyadenosine Adducts in a DNA Mutational Hot Spot Sequence

Shixiang Yan,[†] Robert Shapiro,[†] Nicholas E. Geacintov,*[,†] and Suse Broyde*[,‡]

Contribution from the Department of Chemistry and Department of Biology, New York University, New York, New York 10003

Abstract: Benzo[*a*]pyrene (BP), a prototype polycyclic aromatic hydrocarbon (PAH), can be metabolically activated to the enantiomeric benzo[*a*]pyrene diol epoxides (BPDEs), (+)-(7*R*,8*S*,9*S*,10*R*)-7,8-dihydroxy-9,-10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene and the (−)-(7*S*,8*R*,9*R*,10*S*) enantiomer. These can react with adenine residues in DNA, to produce the stereoisomeric 10*S* (+)- and 10*R* (−)-*trans-anti*-[BP]-N^6-dA adducts. High-resolution NMR solution studies indicate that in DNA duplexes the 10*R* (−) adduct is intercalated on the 5′-side of the modified adenine, while the 10*S* (+) adduct is disordered, exhibits multiple adduct conformations, and is positioned on the 3′-side of the modified adenine. Duplexes containing the 10*S* (+) adduct positioned at A* within codon 61 of the human *N-ras* sequence CA*A are thermodynamically less stable and more easily excised by human DNA repair enzymes than those containing the 10*R* (−) adduct. However, the molecular origins of these differences are not understood and represent a fascinating opportunity for elucidating structure−function relationships. We have carried out a computational investigation to uncover the structural and thermodynamic origins of these effects in the 11-mer duplex sequence d(CGGACA*AGAAG)·d(CTTCTTGTCCG) by performing a 2-ns molecular dynamics simulation using NMR solution structures as the basis for the starting models. Then, we applied the MM-PBSA (molecular mechanics Poisson−Boltzmann surface area) method to compute free energy differences between the stereoisomeric adducts. The 10*R* (−) isomer is more stable by ∼13 kcal/mol, of which ∼10 kcal/mol is enthalpic, which agrees quite well with their observed differences in thermodynamic stability. The lower stability of the 10*S* (+) adduct is due to diminished stacking by the BP moiety in the intercalation pocket, more helix unwinding, and a diminished quality of Watson−Crick base pairing. The latter stems from conformational heterogeneity involving a *syn*−*anti* equilibrium of the glycosidic bond in the modified adenine residue. The lower stability and conformational heterogeneity of the 10*S* (+) adduct may play a role in its enhanced susceptibility to nucleotide excision repair.

Introduction

Chemical carcinogenesis by polycyclic aromatic hydrocarbons (PAHs) is of great interest because these substances are widespread in our environment. Intriguing stereochemical properties of their metabolic derivatives lead to distinct biophysical, biochemical, and biological properties in the products of their reaction with DNA, which likely play a key role in the carcinogenic potency of the PAH compounds. Benzo[*a*]pyrene (BP) is a prototype PAH which manifests these properties. It is present in tobacco smoke, automobile exhaust, and in foods, both as a product of pyrolysis of proteins and as a contaminant.[1−4] Metabolic activation can occur by a number of pathways, including the predominant diol epoxide route,[5] the one-electron oxidation path,[6] and through reactive *o*-quinones.[7] Among the reaction products through the major diol epoxide pathway are a pair of enantiomers, (+)-(7*R*,8*S*,9*S*,10*R*)-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene and its (−)-(7*S*,8*R*,9*R*,-10*S*) mirror image, known as (+)- and (−)-*anti*-BPDE[5] (Figure 1a).

Both (+)- and (−)-*anti*-BPDE can react with adenine[8−10] in DNA by *trans* epoxide opening, to form the 10*S* (+)- and 10*R* (−)-*trans-anti*-[BP]-N^6-dA adducts (Figure 1b); these are more predominant than the products of *cis* epoxide opening. The adenine adducts, while derived in lower yield than guanine adducts, are biologically important. They are mutagenic in cellular systems in which repair is apparently not a factor.[11] Also of interest is the finding that low, nontoxic doses of (+)-*anti*-BPDE produce a greater proportion of mutations at adenines than guanines,[12−14] when compared to the high dose of (+)-*anti*-BPDE employed in many experimental protocols. Mutations, if engendered in genes involved in cell-cycle control, such

* Authors to whom correspondence should be addressed. Telephone: 212-998-8407. Fax: 212-998-8421. Email: ng1@nyu.edu. Telephone: 212-998-8231. Fax: 212-995-4015. Email: broyde@nyu.edu.

[†] Department of Chemistry.

[‡] Department of Biology.

(1) Phillips, D. H. *Mutat. Res.* **1999**, *443*, 139−147.

(2) Grimmer, G. *Proc. 13th Int. Symp. Polynuclear Aromatic Hydrocarbons;* Gordon and Breach: Langhorne, PA, 1993.

(3) Harvey, R. G. *Polycyclic Aromatic Hydrocarbons;* Cambridge University Press: Cambridge, UK, 1991.

(4) Perrin, J. L.; Poirot, N.; Liska, P.; Hanras, C.; Theinpont, A.; Felix, G. *Proc. 13th Int. Symp. Polynuclear Aromatic Hydrocarbons;* Gordon and Breach: Langhorne, PA, 1993.

(5) Conney, A. H. *Cancer Res.* **1982**, *42*, 4875−4917.

(6) Cavalieri, E. L.; Rogan, E. G. *Xenobiotics* **1995**, *25*, 677−688.

(7) Burczynski, M. E.; Harvey, R. G.; Penning, T. M. *Biochemistry* **1998**, *37*, 6781−6790.

(8) Cheng, S. C.; Hilton, B. D.; Roman, J. M.; Dipple, A. *Chem. Res. Toxicol.* **1989**, *2*, 334−340.

(9) Meehan, T.; Straub, K. *Nature* **1979**, *277*, 410−412.

(10) Szeliga, J.; Dipple, A. *Chem. Res. Toxicol.* **1998**, *11*, 1−11.

(11) Page, J. E.; Zajc, B.; Oh-hara, T.; Lakshman, M. K.; Sayer, J. M.; Jerina, D. M.; Dipple, A. *Biochemistry* **1998**, *37*, 9127−9137.

10.1021/ja0043035 CCC: $20.00
Published on Web 06/30/2001

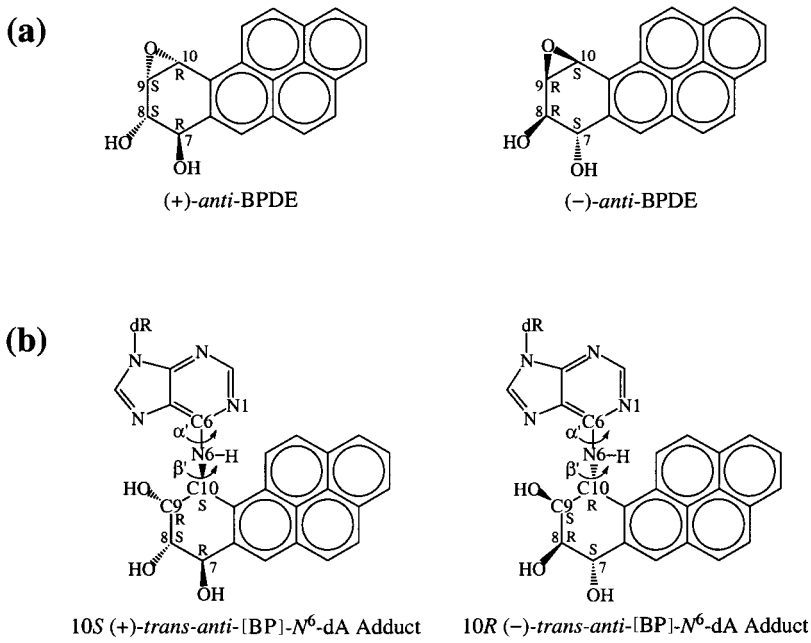

Figure 1. Structures of (a) (+)- and (−)-*anti*-BPDE and (b) 10*S* (+)- and 10*R* (−)-*trans-anti*-[BP]-N^6-dA adducts. Torsion angles α′ and β′ are defined as: α′, N1−C6−N6−C10(BP); β′, C6−N6−C10(BP)−C9(BP). (c) Human *N-ras* codon 61 sequence context. A*6 is the modified adenine.

as oncogenes or tumor suppressor genes,[15,16] are widely believed to be an initiating event in the multistage process of carcinogenesis.[17,18]

High-resolution NMR solution studies have revealed that both the 10*S* (+)- and 10*R* (−)-*trans-anti*-[BP]-N^6-dA adducts in duplex DNAs are intercalated into the DNA helix, without displacement of the modified adenine,[19−22] which remains Watson−Crick hydrogen bonded when opposed by a normal partner, thymine.[19] Conformational heterogeneity is manifested by the 10*S* (+) isomer in solution.[20−25] Moreover, the 10*S* (+) isomer is intercalated on the 3′-side of the modified adenine,[20,21] while the 10*R* (−) isomer is inserted, oppositely, on the 5′-side.[19,22] This opposite orientation phenomenon has now been observed in numerous (+)/(−) adduct pairs of activated PAHs,[19−22,26−40] irrespective of the specific conformation adopted, the specific base modified, or the specific PAH. The

(12) Wei, S. J.; Chang, R. L.; Wong, C. Q.; Bhachech, N.; Cui, X. X.; Henning, E.; Yagi, H.; Sayer, J. M.; Jerina, D. M.; Preston, B. D. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11227−11230.
(13) Wei, S. J.; Chang, R. L.; Bhachech, N.; Cui, X. X.; Merkler, K. A.; Wong, C. Q.; Hennig, E.; Yagi, H.; Jerina, D. M.; Conney, A. H. *Cancer Res.* **1993**, *53*, 3294−3301.
(14) Wei, S. J.; Chang, R. L.; Wong, C. Q.; Cui, X. X.; Dandamudi, N.; Lu, Y. P.; Merkler, K. A.; Sayer, J. M.; Conney, A. H.; Jerina, D. M. *Int. J. Oncol.* **1999**, *14*, 509−513.
(15) Bos, J. L. *Cancer Res.* **1989**, *49*, 4682−4689.
(16) Coles, C.; Condie, A.; Chetty, U.; Steel, C. M.; Evans, H. J.; Prosser, J. *Cancer Res.* **1992**, *52*, 5291−5298.
(17) Weinberg, R. A. *Sci. Am.* **1996**, *275*, 62−70.
(18) Garner, R. C. *Mutat. Res.* **1998**, *402*, 67−75.
(19) Zegar, I. S.; Kim, S. J.; Johansen, T. N.; Horton, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 6212−6224.
(20) Schwartz, J. L.; Rice, J. S.; Luxon, B. A.; Sayer, J. M.; Xie, G.; Yeh, H. J. C.; Liu, X.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1997**, *36*, 11069−11076.
(21) Yeh, H. J.; Sayer, J. M.; Liu, X.; Altieri, A. S.; Byrd, R. A.; Lakshman, M. K.; Yagi, H.; Schurter, E. J.; Gorenstein, D. G.; Jerina, D. M. *Biochemistry* **1995**, *34*, 13570−13581.
(22) Schurter, E. J.; Yeh, H. J.; Sayer, J. M.; Lakshman, M. K.; Yagi, H.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1995**, *34*, 1364−1375.
(23) Li, Z.; Kim, H. Y.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1999**, *38*, 16045−16057.
(24) Volk, D. E.; Rice, J. S.; Luxon, B. A.; Yeh, H. J. C.; Liang, C.; Xie, G.; Sayer, J. M.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **2000**, *39*, 14040−14053.
(25) Zegar, I. S.; Chary, P.; Jabil, R. J.; Tamura, P. J.; Johansen, T. N.; Lloyd, R. S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1998**, *37*, 16516−16528.
(26) Cosman, M.; de los Santos, C.; Fiala, R.; Hingerty, B. E.; Singh, S. B.; Ibanez, V.; Margulis, L. A.; Live, D.; Geacintov, N. E.; Broyde, S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1914−1918.
(27) Fountain, M. A.; Krugh, T. R. *Biochemistry* **1995**, *34*, 3152−3161.
(28) de los Santos, C.; Cosman, M.; Hingerty, B. E.; Ibanez, V.; Margulis, L. A.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1992**, *31*, 5245−5252.
(29) Cosman, M.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1994**, *33*, 11507−11517.
(30) Cosman, M.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1994**, *33*, 11518−11527.
(31) Cosman, M.; de los Santos, C.; Fiala, R.; Hingerty, B. E.; Ibanez, V.; Luna, E.; Harvey, R.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1993**, *32*, 4145−4155.
(32) Cosman, M.; Hingerty, B. E.; Luneva, N.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1996**, *35*, 9850−9863.
(33) Feng, B.; Gorin, A.; Kolbanovskiy, A.; Hingerty, B. E.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 13780−13790.
(34) Schurter, E. J.; Sayer, J. M.; Oh-hara, T.; Yeh, H. J. C.; Yagi, H.; Luxon, B. A.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1995**, *34*, 9009−9020.
(35) Cosman, M.; Fiala, R.; Hingerty, B. E.; Laryea, A.; Lee, H.; Harvey, R. G.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1993**, *32*, 12488−12497.

origin of this phenomenon has been revealed from computational studies on the nucleoside level[41−44] to stem from "primary steric hindrance"[45] between the modified base and the BP moiety when a (+) isomer is rotated into the conformational domain favored by the (−) isomer, and vice versa, and this is rooted in the mirror image nature of the benzylic ring in the two stereoisomers.

Differential interactions with key enzymes involved in DNA replication and transcription and lesion repair in adducts containing oppositely oriented bulky PAH derivatives could well play a significant role in their differing biological properties. In the case of the 10*S* (+)- and 10*R* (−)-*trans*-*anti*-[BP]-N^6-dA adducts, it has been observed that the 10*S* (+) isomer is more susceptible to repair by the human nucleotide excision repair (NER) system than the 10*R* (−) isomer, in the codon 61 CAA sequence of the human *N-ras* proto-oncogene,[46] with modification at the central A (Figure 1c). This codon is frequently mutated in human tumors, and the activation of the *N-ras* proto-oncogene to its oncogenic derivative through point mutations at codon 61 is understood to cause damage to the cellular machinery that regulates proliferation.[15] In solution, the 10*R* (−) isomer has an ∼8 °C higher melting temperature than the 10*S* (+) isomer in the same human *N-ras* codon 61 sequence context.[47]

In an effort to elucidate the structural, dynamic, and thermodynamic differences between these two adducts in this human *N-ras* codon 61 sequence context, which could explain their differences in thermal stability, conformational heterogeneity and repair susceptibility, we have carried out a computational investigation. This involved a 2-ns molecular dynamics simulation to create an ensemble of structures for each adduct, using NMR solution structures[19,20] as the basis for the starting models. From these we deduce the structural differences between the stereoisomeric adducts that explain the experimentally observed differences in melting, and that can rationalize the differences in repair susceptibilities. We applied the MM-PBSA (molecular mechanics Poisson−Boltzmann surface area) method[48−51] to compute free energy differences between the stereoisomeric adducts. The 10*R* (−) isomer is more stable by ∼13 kcal/mol, of which ∼10 kcal/mol is enthalpic. This agrees quite well with the observed melting differences. Our structural observations elucidate the origins of the lesser stability of the 10*S* (+) adduct: diminished stacking by the BP moiety in the intercalation pocket, more helix unwinding, and diminished quality of Watson−Crick base pairing stemming from conformational heterogeneity involving a *syn*−*anti* equilibrium of the glycosidic bond in the modified adenine residue. In turn, these differences may account for the observed greater susceptibility to repair of this adduct.[46]

Methods

Starting Structures. The NMR solution structure,[19] obtained from the Nucleic Acid Database,[52] was the starting structure for the DNA duplex undecamer in the human *N-ras* codon 61 sequence context (Figure 1c) containing a 10*R* (−)-*trans*-*anti*-[BP]-N^6-dA adduct with a normal partner dT. While there was no NMR structure available for the 10*S* (+)-*trans*-*anti*-[BP]-N^6-dA adduct in the same sequence, an experimental NMR solution structure[20] did exist in another sequence which is a reasonable starting model for our work. This NMR structure was a 9-mer, in the d(GGTCA*CGAG)·d(CTCGGGACC) sequence, and contained a mismatched dG opposite the lesion site. Coordinates were kindly provided by Dr. Jane Sayer. We remodeled it to the human *N-ras* codon 61 sequence context as follows: First, we truncated two terminal nucleotides at each end of the original nonamer duplex containing the 10*S* (+)-*trans*-*anti*-[BP]-N^6-dA adduct to avoid end effects; then, we appended three nucleotides corresponding to the human *N-ras* codon 61 sequence context at each end of the truncated DNA duplex. Finally, the rest of the DNA duplex was remodeled to the human *N-ras* codon 61 sequence. The starting structure for the unmodified DNA duplex d(CGGACAAGAAG)·d(CTTCTTGTCCG) in the human *N-ras* codon 61 sequence context was an energy-minimized B-form DNA computed with DUPLEX[53] from a B-DNA fiber diffraction model.[54] Insight II 97.0 from Molecular Simulations, Inc., a subsidiary of Pharmacopeia, Inc., was used for all of the modeling.

Force Field. To obtain partial charges for the 10*S* (+)- and 10*R* (−)-*trans*-*anti*-[BP]-N^6-dA nucleosides, we excised them from the NMR duplex DNA structures,[19−21] minimized, and used Hartree−Fock calculations with 6-31G* basis set to calculate the electrostatic potential, using Gaussian 94.[55] The charge was then fitted to each atomic center with restrained electrostatic potential fitting (RESP).[56] For the 10*S* (+) isomer two structures were used,[20,21] and final charges were averaged. These partial charges were then normalized to maintain a charge of −1 on the modified nucleotide, using our previously described protocol.[57] Bond angle parameters added to the force field for the two adducts were assigned by analogy to chemically similar atom types already available in the *parm98* parameter set.[58] Table S1 (Supporting

(36) Cosman, M.; Laryea, A.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 1295−1307.
(37) Zegar, I. S.; Setayesh, F. R.; DeCorte, B. L.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 4334−4348.
(38) Feng, B.; Zhou, L.; Passarelli, M.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1995**, *34*, 14021−14036.
(39) Feng, B.; Voehler, M.; Zhou, L.; Passarelli, M.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 7316−7329.
(40) Lin, C. H.; Huang, X.; Kolbanovskii, A.; Hingerty, B. E.; Amin, S.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *J. Mol. Biol.* **2001**, *306*, 1059−1080.
(41) Xie, X. M.; Geacintov, N. E.; Broyde, S. *Biochemistry* **1999**, *38*, 2956−2968.
(42) Xie, X. M.; Geacintov, N. E.; Broyde, S. *Chem. Res. Toxicol.* **1999**, *12*, 597−609.
(43) Tan, J.; Geacintov, N. E.; Broyde, S. *J. Am. Chem. Soc.* **2000**, *122*, 3021−3032.
(44) Tan, J.; Geacintov, N. E.; Broyde, S. *Chem. Res. Toxicol.* **2000**, *13*, 811−822.
(45) Geacintov, N. E.; Cosman, M.; Hingerty, B. E.; Amin, S.; Broyde, S.; Patel, D. J. *Chem. Res. Toxicol.* **1997**, *10*, 111−146.
(46) Buterin, T.; Hess, M. T.; Luneva, N.; Geacintov, N. E.; Amin, S.; Kroth, H.; Seidel, A.; Naegeli, H. *Cancer Res.* **2000**, *60*, 1849−1856.
(47) Krzeminski, J.; Ni, J.; Zhuang, P.; Luneva, N.; Amin, S.; Geacintov, N. E. *Polycyclic Aromat. Compd.* **1999**, *17*, 1−10.
(48) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(49) Jayaram, B.; Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Am. Chem. Soc.* **1998**, *120*, 10629−10633.
(50) Cheatham, T. E.; Srinivasan, J.; Case, D. A.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 265−280.
(51) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(52) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751−759.
(53) Hingerty, B. E.; Figueroa, S.; Hayden, T. L.; Broyde, S. *Biopolymers* **1989**, *28*, 1195−1222.
(54) Arnott, S.; Smith, P. J. C.; Chandrasekaran, R. *Handbook of Biochemistry and Molecular Biology*, 3rd ed.; CRC Press: Cleveland, OH, 1976.
(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P.; Johnsone, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Reploge, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision A.1; Gaussian, Inc.: Pittsburgh, PA, 1995.
(56) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(57) Wu, X.; Shapiro, R.; Broyde, S. *Chem. Res. Toxicol.* **1999**, *12*, 895−905.
(58) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.

Information) gives the partial charges and atom types, and Table S2 (Supporting Information) gives the added parameters.

Molecular dynamics (MD) simulations were carried out using the AMBER 5.0 package[59] with the Cornell et al. force field[60] and the *parm98* parameter set.[58] Helicoidal parameters[61] and DNA backbone torsional angles were computed with Dials and Windows.[62]

Molecular Dynamics Protocol. The particle mesh Ewald (PME) method[63,64] was used to treat long-range electrostatic interactions with a cubic B-spline interpolation and a 10^{-5} tolerance for the direct space sum cutoff. A 12-Å cutoff was applied to the nonbonded Lennard-Jones interactions. The SHAKE algorithm[65] was applied to constrain all bonds involving hydrogen atoms with a tolerance of 0.0005 Å, and a 1-fs time step was used in the dynamics simulation. The translational and rotational motion of the center of mass was removed about every 300 ps. In all, 20 Na^+ ions were added to the system for neutralization using the *Leap* module in AMBER 5.0. The system was then solvated with a rectangular box of TIP3P waters[66] which extended ~10 Å from the DNA atoms in each direction. This yielded a periodic box size of ~ 55 Å × 69 Å × 55 Å for the 10*S* (+) adduct, ~ 49 Å × 49 Å × 69 Å for the 10*R* (−) adduct and ~ 52 Å × 50 Å × 62 Å for the unmodified control. In all, 4710 water molecules were added for the 10*S* (+) adduct, 3640 for the 10*R* (−) adduct, and 3662 for the unmodified control. All systems followed the same minimization and equilibration protocols. First, the water molecules and counterions were minimized for 1500 steps of steepest descent, followed by 50 ps dynamics with the DNA fixed to allow the solvent to relax. The whole system was then minimized for 1000 additional steps of steepest descent, followed by 3 ps dynamics with 25 kcal/mol restraints on the DNA, which further allowed the waters to relax. Then the system was minimized for 5 rounds of 600 steps of steepest descent with the restraints on the DNA reduced by 5 kcal/mol each round, from 20 to 0 kcal/mol. Finally, the whole system was heated from 10 to 300 K over 40 ps using the Berendsen coupling algorithm[67] with a coupling parameter of 0.2 ps. Production simulation was then continued at atmospheric pressure with a 0.2 ps coupling parameter and 300 K for 2 ns.

Free Energy Analysis. Snapshots for calculating free energies were taken from the MD trajectories of the 10*S* (+) (**S**) and 10*R* (−) (**R**) adducts with water molecules and counterions removed. To assess the equilibration of the trajectories, we monitored the root-mean-square deviations (rmsd's) of the MD structures and the fluctuations in total energy, volume, pressure, and temperature. Then 150 snapshots were selected at 10-ps intervals from each equilibrated trajectory.

The free energies (G_{tot}) were estimated from the molecular mechanical (MM) energies (E_{MM}), the solvation free energies ($G_{solvation}$), and the vibrational, rotational, and translational entropies for the DNA, following previously developed methods:[48−51]

$$\Delta G_{tot} = \Delta E_{MM} + \Delta G_{solvation} - T\Delta S \quad (1)$$

where T is the temperature, and S is the entropy. The MM energies (E_{MM}) were calculated from internal energies (E_{int}) stemming from deviations of the bonds (E_{bonds}), angles (E_{angles}), and dihedral angles ($E_{dihedrals}$) from their equilibrium values, van der Waals interaction energies (E_{vdW}), and electrostatic energies ($E_{electrostatic}$), using the *Anal* module in the AMBER 5.0 package:[59]

$$E_{MM} = E_{int} + E_{vdW} + E_{electrostatic} \quad (2)$$

where:

$$E_{int} = E_{bonds} + E_{angles} + E_{dihedrals} \quad (3)$$

The same force field[60] and parameters used in the MD simulations was applied to calculate the MM energies (E_{MM}) but with no cutoff for nonbonded interactions.

The solvation free energies ($G_{solvation}$) were estimated from the electrostatic solvation energies (G_{PB}) and the nonpolar solvation energies ($G_{nonpolar}$):

$$G_{solvation} = G_{PB} + G_{nonpolar} \quad (4)$$

The electrostatic contribution to the solvation free energies (G_{PB}) for each snapshot was evaluated using the finite difference Poisson−Boltzmann (FDPB) method,[68,69] as implemented in the DelPhi program.[70,71] The FDPB method approximates the electrostatic solvation energy as the reaction field energy of taking a solute from a vacuum dielectric medium ($\epsilon = 1$) to an aqueous dielectric medium ($\epsilon = 80$).[48,70,71] A physiological 0.15 M ionic strength[72] was employed to calculate the effect of salt on the free energies. DNA atomic charges were taken from the Cornell et al. force field,[60] and those for the adducts are given in Table S1, Supporting Information. These are consistent with the charges used in the molecular mechanical energy calculations.[48] The atomic radii were taken from the PARSE parameter set,[73] and the dielectric boundary is defined by using a probe radius of 1.4 Å.[48,74] A grid spacing of 2.0 grids/Å, in which the longest linear dimension of the solute occupied 80% of the lattice, was used to determine the size of the cubic lattice, and the boundary potentials were set to the sum of the Debye−Hückel values.[75] A total of 300 linear iterations followed by 1000 nonlinear iterations was performed for each snapshot to reach convergence, as judged by the rmsd between successive iterated potential maps being less than 10^{-5}.

The nonpolar contribution to the solvation free energies ($G_{nonpolar}$) was estimated as:

$$G_{nonpolar} = \gamma SASA + b \quad (5)$$

where $\gamma = 0.00542$ kcal/Å^2, $b = 0.92$ kcal/mol,[73] and *SASA* is the solvent-accessible surface area which was estimated by using Sanner's algorithm implemented in the MSMS software.[76] A solvent probe radius of 1.4 Å and PARSE atomic radii values[73] were used.

Solute entropic contributions were approximated with normal mode calculations[77] by using the *Nmode* module in the AMBER 5.0 package.[59] Normal mode calculations of solute entropy are only rough estimates and computationally expensive,[50] and consequently it is usual to employ only one structure to compute this quantity.[48,78,79] We employed the following protocol: First, 10 structures at 200-ps intervals were selected from each trajectory; then, using a distance-dependent dielectric function

(59) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5.0*; University of California: San Francisco, CA, 1997.

(60) 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*, 5179−5197.

(61) Dickerson, R. E. *J. Biomol. Struct. Dyn.* **1989**, *6*, 627−634.

(62) Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669−699.

(63) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(64) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pederson, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(65) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(66) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Imprey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(67) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(68) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.

(69) Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1990**, *94*, 7684−7692.

(70) Nicholls, A.; Sharp, K. A.; Honig, B. *DelPhi*; Department of Biochemistry and Molecular Biophysics, Columbia University: New York, 1990.

(71) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.

(72) Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell*, 3rd ed.; Garland Publishing: New York, 1994.

(73) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.

(74) Massova, I.; Kollman, P. A. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.

(75) Gilson, M. K.; Honig, B. H. *Nature* **1987**, *330*, 84−86.

(76) Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305−320.

(77) Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 671−682.

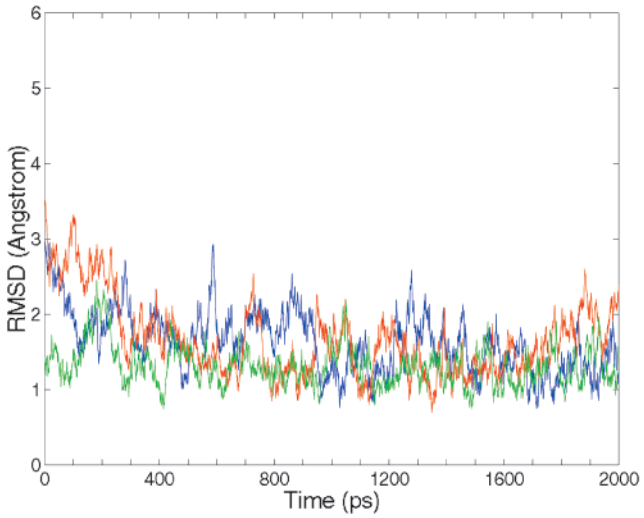

Figure 2. Root-mean-square deviations (rmsd's) for the 10*S* (+) adduct (**S**, red), 10*R* (−) adduct (**R**, blue), and the unmodified control duplex (**C**, green) over the 2-ns production MD simulation. The rmsd's were calculated relative to the average structures over 500−2000 ps for the 10*S* (+) (**S**) and 10*R* (−) (**R**) adducts, and 200−2000 ps for the unmodified control duplex (**C**).

($\epsilon = 4r$, where r is the interatomic distance in Å) to mimic solvent, steepest descent and conjugate gradient minimizations, followed by Newton−Raphson minimizations, were carried out with no cutoff for nonbonded interactions until the root-mean-square of the elements in the gradient vector was less than 10^{-4} kcal/(mol·Å) for each structure. Finally, we chose the one minimized structure with the smallest rmsd compared to the MD average structure for each adduct, to calculate the translational, rotational, and vibrational entropies at 300 K.

Quality of Watson−Crick Hydrogen Bonding Analysis. We employ a hydrogen bond quality index,[80] I_H, to quantitatively assess the deviation from ideal Watson−Crick hydrogen bonding distances and angles:

$$I_H = \sum_{\substack{\text{D−H···A} \\ \text{triplets}}} [(d_{DA} - d^0_{DA})^2 + (1 + \cos\gamma)^2] \qquad (6)$$

where d_{DA} is the instantaneous donor−acceptor distance, d^0_{DA} is an ideal donor−acceptor distance[81] (N6 (A) to O4 (T) is 2.95 Å, N1 (A) to N3 (T) is 2.82 Å) and γ is the instantaneous D−H···A bond angle with ideal value of 180°. The summation is over the two Watson−Crick hydrogen bonds in an A-T base pair. I_H adopts a value of 0 when ideal Watson−Crick hydrogen bonding is maintained.

Results

Simulation Stability. The rmsd's of the 10*S* (+) adduct (**S**), 10*R* (−) adduct (**R**) and the unmodified control (**C**) calculated relative to the starting structures are shown in Figure S1, Supporting Information. During the first ∼500 ps of production simulation, structures **S** and **R** exhibit continued equilibration, while structure **C** becomes equilibrated by ∼200 ps. We then calculated an average structure for the 10*S* (+) adduct (**S**) and 10*R* (−) adduct (**R**) over 500−2000 ps, and for the unmodified control (**C**) over 200−2000 ps. The rmsd's of structures **S**, **R**, and **C** were calculated relative to these average structures, as shown in Figure 2. The mean rmsd relative to the average structures is 1.5 ± 0.3 Å for the 10*S* (+) adduct (**S**), 1.5 ± 0.4 Å for the 10*R* (−) adduct (**R**) over 500−2000 ps, and 1.3 ± 0.2 Å for the unmodified control **C** over 200−2000 ps. These results show stable trajectories over the time frames of 500−2000 ps for structures **S** and **R**, and of 200−2000 ps for structure **C**.

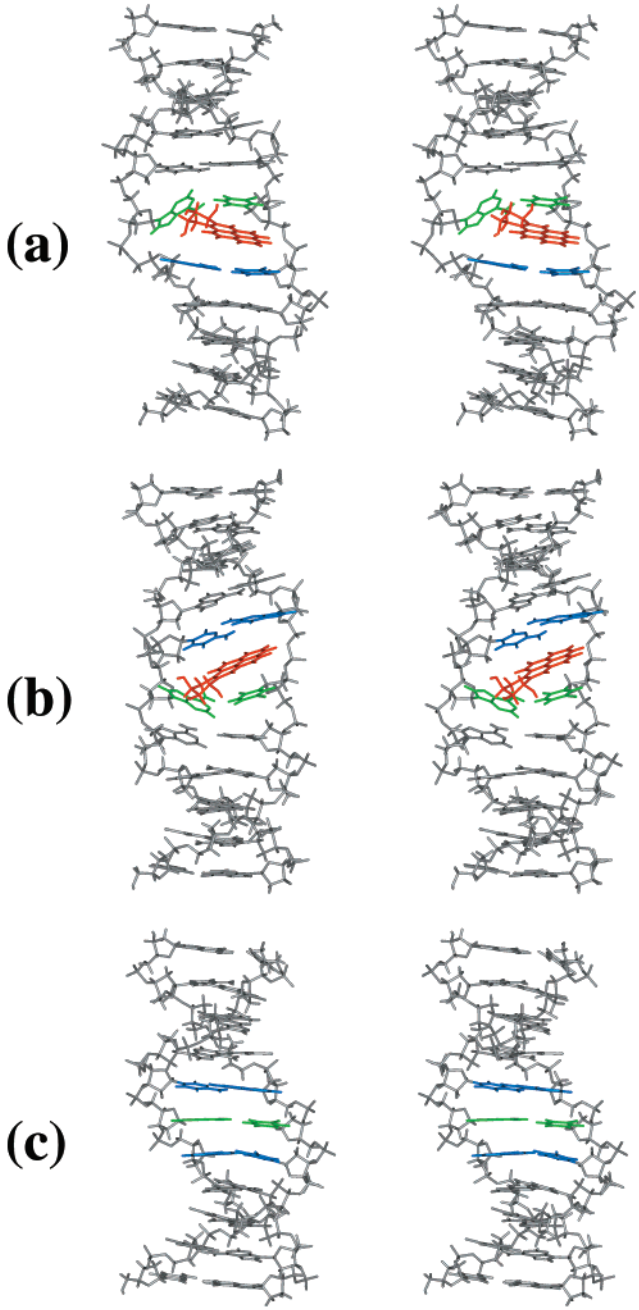

Figure 3. Stereoviews of the MD average structures. (a) 10*S* (+) adduct (**S**) over 500−2000 ps with BP in red, A*6-T17 in green, and A7-T16 in blue. (b) 10*R* (−) adduct (**R**) over 500−2000 ps with BP in red, C5-G18 in blue, and A*6-T17 in green. (c) Unmodified control duplex (**C**) over 200−2000 ps with C5-G18, A7-T16 in blue, and A6-T17 in green. The backbone, sugar atoms, and other residues are in gray. All structures are aligned to have 5′-OH at the top right. All stereo images are constructed for viewing with a stereoviewer.

Stereoviews of the dynamics average structures from 500 to 2000 ps for the 10*S* (+) adduct (**S**) and the 10*R* (−) adduct (**R**), and from 200 to 2000 ps for the unmodified control (**C**) are shown in Figure 3. The carcinogen is intercalated between the modified A*6-T17 base pair and the 3′ neighboring A7-T16 base pair in the case of the 10*S* (+) adduct (**S**). In the 10*R*

(78) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330−14335.
(79) Reyes, C. M.; Kollman, P. A. *J. Mol. Biol.* **2000**, *297*, 1145−1158.
(80) Hingerty, B. E.; Figueroa, S.; Hayden, T. L.; Broyde, S. *Biopolymers* **1989**, *28*, 1195−1222.
(81) Saenger, W. *Principles of Nucleic Acid Structure;* Springer-Verlag: New York, 1984.

Table 1. Free Energy Analysis of 10*S* (+)- and 10*R* (−)-*trans*-*anti*-[BP]-N^6-dA DNA Adducts[a]

	10*S* (+)	10*R* (−)	10*S* (+) − 10*R* (−)
$\langle E_{electrostatic} \rangle$	364.3 (41.5)	346.0 (47.3)	18.3
$\langle E_{vdW} \rangle$	−181.6 (10.2)	−185.6 (10.2)	4.0
$\langle E_{int} \rangle$	1007.4 (17.8)	1003.2 (18.1)	4.2
$\langle E_{MM} \rangle$	1190.1 (41.9)	1163.7 (44.6)	26.4
$\langle G_{nonpolar} \rangle$	25.4 (0.2)	25.3 (0.2)	0.1
$\langle G_{PB} \rangle$	−5651.1 (40.1)	−5634.8 (44.0)	−16.3
$\langle G_{solvation} \rangle$	−5625.7 (40.0)	−5609.5 (44.0)	−16.2
$\langle G_{PB} + E_{electrostatic} \rangle$	−5286.8 (11.5)	−5288.8 (11.8)	2.0
$\langle E_{MM} + G_{PB} \rangle$	−4461.0 (15.6)	−4471.2 (15.7)	10.2
$-TS$	−591.6	−594.5	2.9
G_{tot}	−5027.2	−5040.3	13.1

[a] All energies are in kcal/mol. Values in parentheses are standard deviations.

(−) adduct (**R**), it is intercalated between the A*6-T17 and the C5-G18 base pairs, on the 5′-side of the modified adenine. Thus, both are similar to the parent duplex NMR structures[19,20] in this respect. Stereo snapshots along the trajectory for the 10*S* (+) adduct (**S**) are given in Figure S2, and for the 10*R* (−) adduct (**R**) in Figure S3, Supporting Information. Small adjustments throughout the duplexes account for the rmsd's from the starting structures. The BP moieties remain stably intercalated over the course of the simulations. This is demonstrated by the α′ and β′ torsion angles of the structures **S** and **R** over the 2-ns simulations, as shown in Figure S4, Supporting Information. The average α′ and β′ torsion angles are −153.4° ± 14.5° and −102.1° ± 17.5°, respectively, for the 10*S* (+) adduct; they are 152.2° ± 9.3° and 102.9° ± 9.3°, respectively, for the 10*R* (−) adduct, over 500−2000 ps. Our previous calculations for the 10*S* (+)- and 10*R* (−)-*trans*-*anti*-[BP]-N^6-dA adducts on the nucleoside level showed these α′, β′ domains to be among the low-energy wells for each adduct.[43]

The 10*R* (−) Adduct is More Stable Due to Enthalpic Contributions to the Free Energy. The calculated free energies for the 10*S* (+) (**S**) and 10*R* (−) (**R**) isomers are shown in Table 1 and Figure 4. The computed total free energies favor the 10*R* (−) adduct (**R**) over the 10*S* (+) adduct (**S**) by 13.1 kcal/mol. Analysis of the energy components shows that contributions from van der Waals and internal energies play a major role in the stabilization of the 10*R* (−) adduct (**R**).

The largest contribution to the free energy difference between the 10*S* (+) and 10*R* (−) adducts is the internal energy (E_{int}), stemming from bond, angle, and dihedral angle deviations from equilibrium values; this favors the 10*R* (−) adduct (**R**) by 4.2 kcal/mol. The van der Waals contribution to the free energy (E_{vdW}) favors the 10*R* (−) adduct (**R**) by 4.0 kcal/mol; this is the second largest contribution to the free energy difference between the two isomers. While the gas-phase electrostatic energy ($E_{electrostatic}$) favors the 10*R* (−) adduct (**R**) by 18.3 kcal/mol, the electrostatic solvation energy (G_{PB}) favors the 10*S* (+) adduct (**S**) by 16.3 kcal/mol at our 0.15 M physiological salt concentration.[72] As a result, the total electrostatic free energy difference favors the 10*R* (−) adduct (**R**) by only 2.0 (18.3−16.3) kcal/mol, because electrostatic solvation screens the gas-phase charge−charge interactions.[48] The big fluctuations observed in the gas-phase electrostatic energy are also greatly reduced in the total electrostatic free energy due to the screening effect of solvation.[48,82,83] The nonpolar term of the solvation free energy ($G_{nonpolar}$), the solvent-accessible surface area

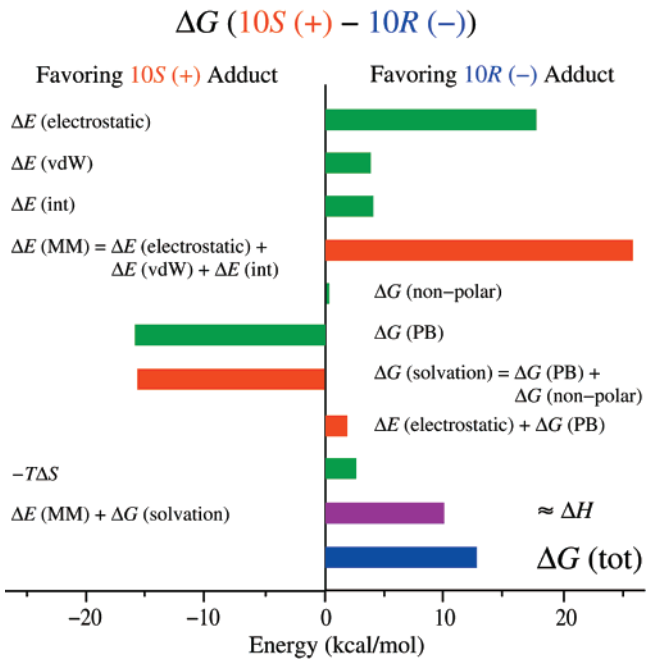

Figure 4. Free energy component analyses for the 10*S* (+) (**S**) and 10*R* (−) (**R**) adducts. The green bars are components of the free energy difference between the two adducts, and the red bars are the partial sums of these components. The magenta bar is the enthalpy difference between the two adducts, and the blue bar is the total free energy difference.

(SASA)-dependent term, favors the 10*R* (−) adduct (**R**) by only 0.1 kcal/mol, indicating hardly any difference in total solvent-accessible surface areas between the 10*S* (+) and 10*R* (−) adduct duplexes. As shown in Table 1, a free energy difference of 2.9 kcal/mol due to solute entropy favors the 10*R* (−) adduct (**R**). The molecular mechanical energies (E_{MM}) and the electrostatic solvation free energies (G_{PB}) are enthalpic contributions to the free energy, and these favor the 10*R* (−) adduct (**R**) by 10.2 kcal/mol.

Backbone and Helicoidal Parameter Analysis Shows the 10*S* (+) Adduct is More Distorted than the 10*R* (−) Adduct. The backbone torsional parameters for the 10*S* (+) adduct (**S**), the 10*R* (−) adduct (**R**) and the unmodified control duplex (**C**) are shown in Figure 5. For both the 10*S* (+) adduct (**S**) and the 10*R* (−) adduct (**R**), the greatest differences in the backbone parameters relative to the unmodified control (**C**) occur primarily at or near the intercalation pockets, namely the A*6-T17 and A7-T16 base pairs in the case of structure **S**, and the C5-G18 and A*6-T17 base pairs in the case of structure **R**. In addition, in the 10*S* (+) adduct (**S**), distortions in backbone torsions ϵ and ζ extend beyond the intercalation pocket. The 3′ intercalation of BP in the 10*S* (+) adduct (**S**) and the 5′ intercalation in the 10*R* (−) adduct (**R**) cause these distortions. Furthermore, we find greater deviations in backbone torsions in the 10*S* (+) adduct (**S**) compared to those in structure **C**. Comparing torsion angles in the intercalation pocket and neighboring residues on each side, namely C5, A*6, A7, G8, C15, T16, T17, and G18 for the 10*S* (+) adduct (**S**), and A4, C5, A*6, A7, T16, T17, G18, and T19 for the 10*R* (−) adduct (**R**), more distortions are seen (Figure 5) for the 10*S* (+) adduct (**S**), particularly in α, β, and γ. It is also observed that the χ angle at C1 in the unmodified control (**C**) has a *syn* conformation instead of the normal *anti* conformation, indicating end base-pair fraying, which has also been observed in a previous simulation.[84]

Helicoidal parameters are shown in Figure 6, and Table 2 gives average values of the helicoidal base pair and base pair

(82) Bashford, D.; Case, D. A.; Choi, C.; Gippert, G. P. *J. Am. Chem. Soc.* **1997**, *119*, 4964−4971.
(83) Demchuk, E.; Bashford, D.; Gippert, G. P.; Case, D. A. *J. Mol. Biol.* **1997**, *270*, 305−317.

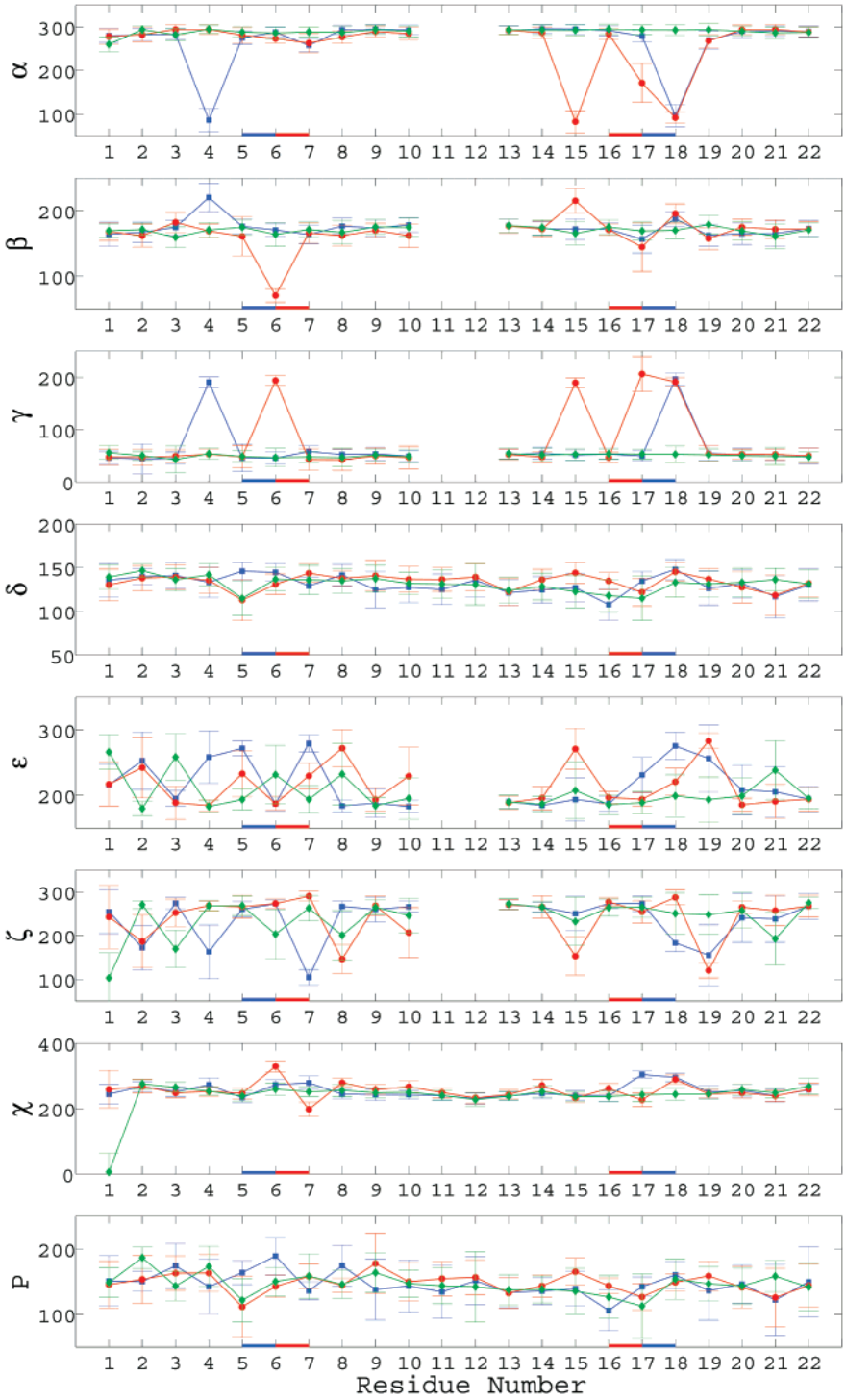

Figure 5. Average backbone torsional parameters for the structures of the 10*S* (+) adduct (**S**, red circles) over 500−2000 ps (3000 structures), 10*R* (−) adduct (**R**, blue squares) over 500−2000 ps (3000 structures), and the unmodified control duplex (**C**, green diamonds) over 200−2000 ps (3600 structures). The standard deviations are shown as error bars. The red and blue bars on the axes indicate the intercalation pocket of the 10*S* (+) adduct (**S**) (A*6-T17 and A7-T16), and the intercalation pocket of the 10*R* (−) adduct (**R**) (C5-G18 and A*6-T17), respectively. All values were calculated using Dials and Windows.[62] It should be noted that the residue numbers in Dials and Windows[62] differ from the IUPAC convention[81] as follows: For α, *β*, and *γ*, residue numbers 1−10 should be shifted +1, and for *ϵ* and *ζ*, residues 13−22 should be shifted −1 to accord with the IUPAC convention.[81]

step parameters in the vicinity of the lesion site, and for the corresponding residues in the unmodified structure **C**. Again, most of the deviations relative to the unmodified control (**C**) occur at the intercalation pockets. Notably, the buckle and propeller of the 10*S* (+) adduct (**S**) and the 10*R* (−) adduct (**R**) at the lesion site, A*6-T17, are in opposite directions in the two adducts, consistent with the 10*S* versus 10*R* stereochemistry and the associated 3′- versus 5′ intercalation in the 10*S* (+) (**S**) and 10*R* (−) (**R**) stereoisomeric adducts. Severe deviations in tilt are also observed at the A*6-T17 to A7-T16 base pair step in the 10*S* (+) adduct (**S**), and at the C5-G18 to A*6-T17 base pair step in the 10*R* (−) adduct (**R**), which are

(84) Spector, T. I.; Cheatham, T. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 7095−7104.

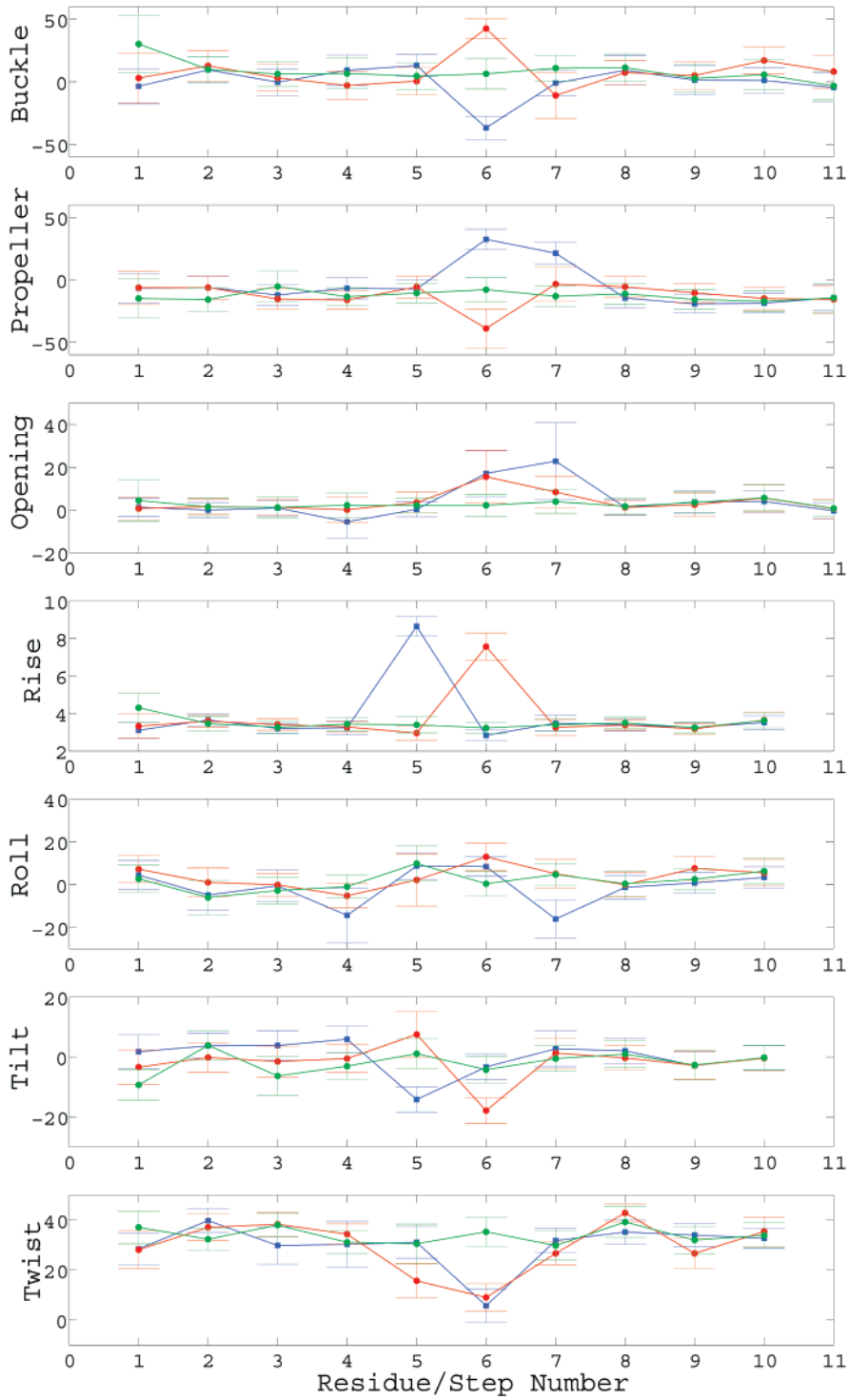

Figure 6. Average helicoidal parameters for the structures of the 10*S* (+) adduct (**S**, red circles) over 500−2000 ps (3000 structures), 10*R* (−) adduct (**R**, blue squares) over 500−2000 ps (3000 structures), and the unmodified control duplex (**C**, green diamonds) over 200−2000 ps (3600 structures). The standard deviations are shown as error bars. The numbering scheme for the nucleotide base pair steps is that the C1-G22 to G2-C21 is step 1, the G2-C21 to G3-C20 is step 2, ..., and so on. All values were calculated using Dials and Windows.[62]

the intercalation pockets for the two adducts, respectively. In addition, at the A7-T16 base pair in the 10*R* (−) adduct (**R**), substantial deviations in propeller and opening are observed. As shown in Figure 3, the modified adenine A*6 in the 10*R* (−) adduct (**R**) stacks well with its intrastrand adjacent base A7; however, the large deformations in the A7-T16 propeller and opening, and negative roll in the A7-T16 to G8-C15 base pair step in the 10*R* (−) adduct (**R**) appear to be the price for this enhanced stacking.

The 3′ intercalation of the bulky BP in the 10*S* (+) adduct (**S**) and the 5′ intercalation in the 10*R* (−) adduct (**R**) introduce a substantial increase in rise in the A*6-T17 to A7-T16 base pair step in the 10*S* (+) adduct (**S**) and the C5-G18 to A*6-T17 base pair step in the 10*R* (−) adduct (**R**) to accommodate the intercalated BP. A substantially decreased twist at the C5-G18 to A*6-T17 base pair step (15.7° ± 6.8°) and the A*6-T17 to A7-T16 base pair step (9.1° ± 5.1°) in the 10*S* (+) adduct (**S**), and at the A*6-T17 to A7-T16 base pair step (5.7°

Table 2. Average Helicoidal Parameters near Lesion Site[a]

	S	R	C	
buckle	42.7 (8.0)	−36.8 (9.2)	4.7 (10.9)	
propeller	−38.8 (15.5)	32.5 (8.1)	−7.9 (9.8)	
opening	15.6 (12.1)	17.1 (10.8)	2.2 (5.2)	
rise	7.6 (0.7)	8.6 (0.5)	3.4 (0.4)	3.2 (0.3)
roll	13.0 (6.5)	8.7 (6.3)	9.9 (8.2)	0.4 (5.7)
tilt	−17.9 (4.3)	−14.2 (4.2)	1.1 (5.0)	−4.2 (4.4)
twist	9.1 (5.1)	31.1 (6.4)	30.5 (7.9)	35.3 (5.8)

[a] Structures **S**, **R**, and **C** are the 10*S* (+) adduct, 10*R* (−) adduct, and the unmodified control duplex d(CGGACAAGAAG)·d(CTTCTTGTCCG), respectively. The base pair parameters buckle, propeller, and opening are for the A*6-T17 base pair of structures **S**, **R**, and **C**. The base pair step parameters rise, roll, tilt, and twist are for the A*6-T17 to A7-T16 base pair step of structure **S**, C5-G18 to A*6-T17 of structure **R**, C5-G18 to A6-T17 and A6-T17 to A7-T16 base pair steps of structure **C**. Values in parentheses are standard deviations.

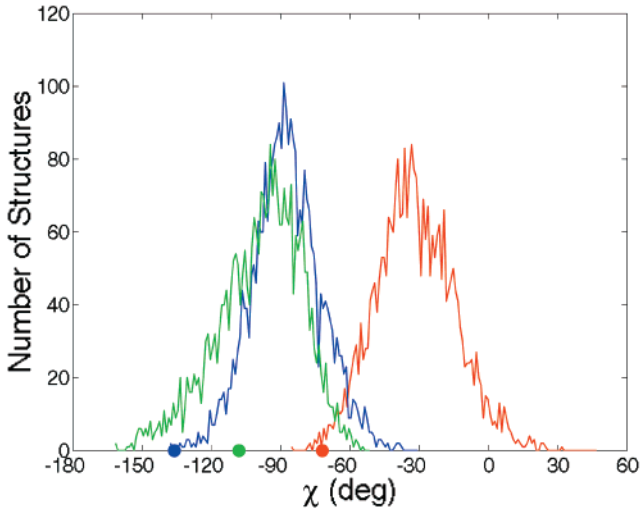

Figure 7. Distribution of the glycosidic torsion angle χ in modified adenine A*6 of the 10*S* (+) adduct (**S**, red), 10*R* (−) adduct (**R**, blue), and in A6 of the unmodified control (**C**, green) over 500−2000 ps. The red, blue, and green solid dots show the χ values of the starting structures of the 10*S* (+) adduct (**S**), 10*R* (−) adduct (**R**), and the unmodified control (**C**), respectively.

± 6.8°) in the 10*R* (−) adduct (**R**) is observed. Notably, the 10*S* (+) adduct (**S**) causes more unwinding than the 10*R* (−) adduct (**R**), relative to structure **C**.

***syn*−*anti* Equilibrium and More Distorted Watson−Crick Hydrogen Bonding in the 10*S* (+) Adduct.** Figure 5 reveals that the glycosidic torsion angle χ at the A*6 is displaced toward the *syn* domain in the 10*S* (+) adduct (**S**) compared to that in the 10*R* (−) one (**R**) and to A6 in the unmodified control structure (**C**). This is clearly shown in Figure 7, which displays the population of structures in the dynamics trajectories as a function of glycosidic torsion angle χ. Both *syn* and *anti* domains are encompassed by the χ range of the 10*S* (+) adduct (**S**), while only the *anti* region is found for the 10*R* (−) adduct (**R**) and the unmodified control (**C**). The range of χ values is −85.3° to 47.1°, with a mean value of −31.2° ± 17.1° in the 10*S* (+) adduct (**S**), and −137.7° to −30.6°, with a mean value of −87.0° ± 14.9° in the 10*R* (−) adduct (**R**). In the unmodified control (**C**) the range is −161.5° to −50.7°, with a mean value of −98.5° ± 18.3°. By comparison, the *syn* domain of DNA duplexes in the Z-DNA conformation is at χ ~60°−80°, while the *anti* domain of B-DNA has χ at ~(−120° to −100°).[81,85]

Figure 8 shows the two structures from the dynamics trajectory of the 10*S* (+) adduct (**S**) with the most extreme χ

(85) Berman, H. M. *Biopolymers* **1997**, *44*, 23−44.

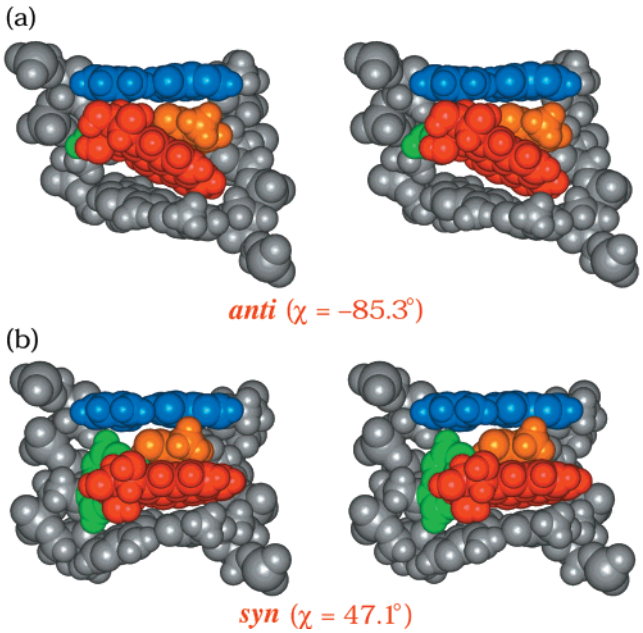

Figure 8. (a) *anti* and (b) *syn* structures from dynamics trajectory of the 10*S* (+) adduct (**S**) (see text).

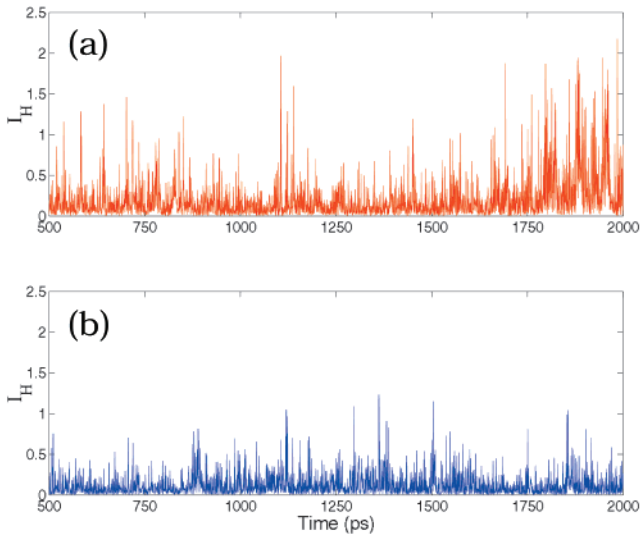

Figure 9. Time dependence of hydrogen-bond quality index, I_H, in (a) 10*S* (+) (**S**) and (b) 10*R* (−) (**R**) adducts over 500−2000 ps. I_H is 0 when Watson−Crick base pairing is ideal.

values: −85.3° (*anti*) and 47.1° (*syn*). These reveal the lack of normal Watson−Crick base pairing at the lesion site in the *syn* structure. Accordingly, we assessed the quality of Watson−Crick hydrogen bonding in this isomer and compared it with the 10*R* (−) adduct (**R**) using our hydrogen bond quality index I_H for each of the trajectories (see Methods). Figure 9 shows the time dependence of I_H for each isomer. Clearly, deviation from ideal Watson−Crick hydrogen bonding is much greater in the 10*S* (+) adduct (**S**). This is demonstrated by the summed values for this index; these are respectively 599.3 and 362.7 for the 3000 structures of the 10*S* (+) and 10*R* (−) adducts collected in the 500−2000 ps interval, with higher values for I_H corresponding to poorer hydrogen bonding quality.

Less Stacking of BP in the Intercalation Pocket of the 10*S* (+) Adduct than in the 10*R* (−) Adduct. Our results reveal that the overlap between the aromatic rings of BP and the adjacent base pairs is very different in the 10*S* (+) and 10*R* (−) isomeric adducts. As shown in Figure 10a, in the 10*S* (+)

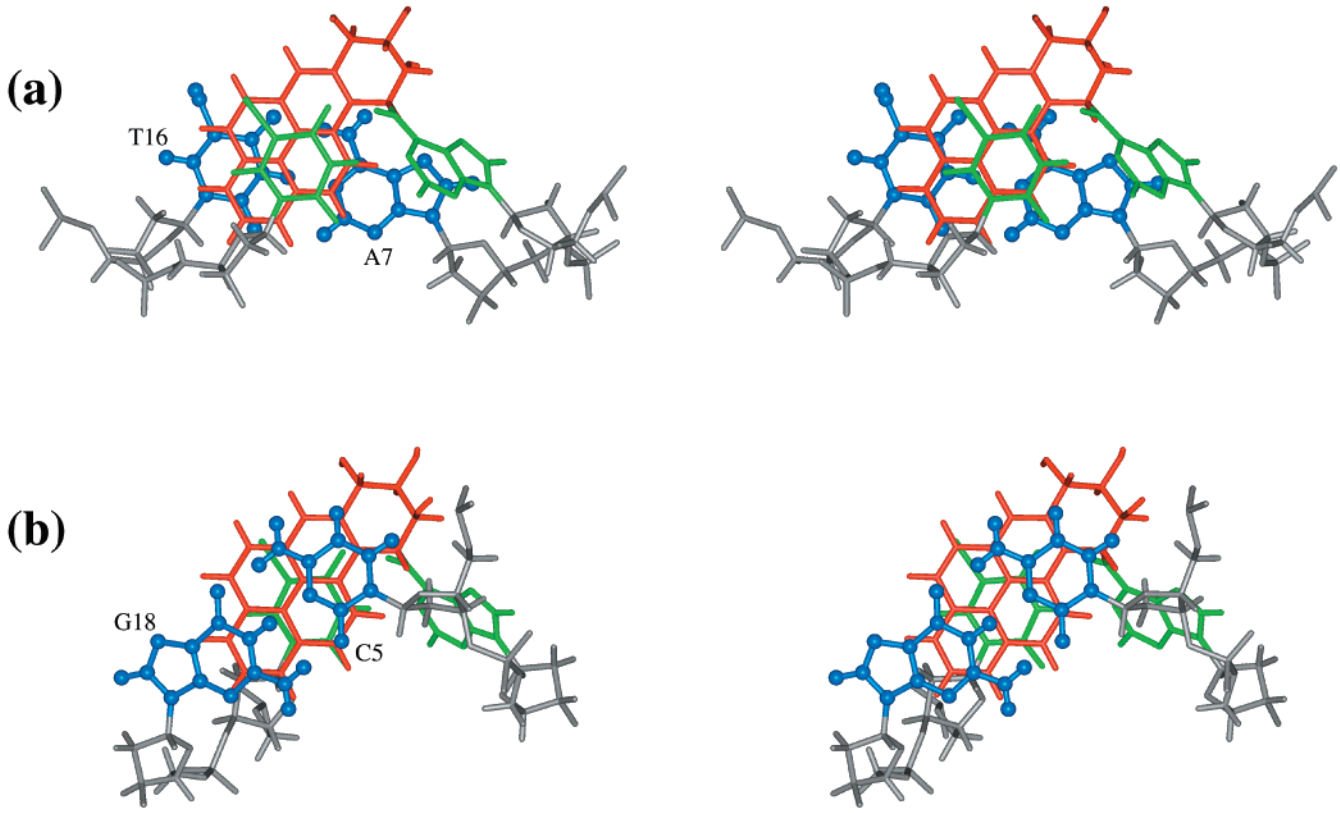

Figure 10. Stereoviews of the intercalation pockets of the 10*S* (+) (**S**, d(A*A)·d(TT)) and 10*R* (−) (**R**, d(CA*)·d(TG)) adducts. (a) 10*S* (+) adduct (**S**). BP is in red stick, A*6-T17 is in green stick, and A7-T16 is in blue ball-and-stick. (b) 10*R* (−) adduct (**R**). BP is in red stick, A*6-T17 is in green stick, and C5-G18 in blue ball-and-stick. All backbone and sugar atoms are in gray. The view is along the helix axis from 5′ (top) to 3′ (bottom) of the modified strand.

adduct (**S**), there is substantial stacking between the aromatic moiety of BP and the partner T17 of the modified nucleotide A*6, but there is little overlap between the aromatic moiety of BP and the adjacent base pair A7-T16, so that the BP is stacked on only one face. On the other hand, in the 10*R* (−) adduct, as shown in Figure 10b, the BP aromatic moiety not only stacks with the partner T17 of modified A*6, but also stacks well with its adjacent base pair C5-G18, so that both faces of the BP are stacked. One aromatic ring of the BP is exposed in the case of the 10*S* (+) adduct (**S**), while all the rings are nearly fully stacked on both sides in the 10*R* (−) adduct (**R**). Thus, the 5′ intercalation in the 10*R* (−) adduct (**R**) imposes a much better overlap between the BP and the adjacent DNA base pair than the 3′ intercalation of BP in the 10*S* (+) adduct (**S**). The van der Waals interaction energies involving these residues reflect these differential stacking interactions. Specifically, we computed the van der Waals interaction energies between the BP and T17, A7, and T16 in the case of the 10*S* (+) adduct (**S**), and between the BP and T17, C5, and G18 in the case of the 10*R* (−) adduct (**R**). The results are shown in Table 3. The van der Waals interaction energy between the BP and T17 in the 10*S* (+) and 10*R* (−) adducts differs by only 0.1 kcal/mol, favoring the 10*S* (+) adduct (**S**). On the other hand, the van der Waals interaction energy between the BP and the adjacent base pair A7-T16 is −14.6 (−6.6 + (−8.0)) kcal/mol in the case of the 10*S* (+) adduct (**S**), but, for the 10*R* (−) adduct (**R**), the van der Waals interaction energy between the BP and the adjacent base pair C5-G18 is −18.5 (−10.2 + (−8.3)) kcal/mol. Therefore, in the intercalation pockets of the 10*S* (+) and 10*R* (−) adducts, the van der Waals interaction energy favors the 10*R* (−) adduct (**R**) by 3.7 kcal/mol.

Table 3. van der Waals Interaction Energies at the Intercalation Pocket[a]

	10*S* (+)	10*R* (−)	10*S* (+) − 10*R* (−)
BP-T17	−12.3	−12.2	−0.1
BP-A7/C5	−6.6	−10.2	3.4
BP-T16/G18	−8.0	−8.3	0.3
total	−26.9	−30.7	3.7

[a] BP-A7 and BP-C5 are for 10*S* (+) and 10*R* (−) adducts, respectively. BP-T16 and BP-G18 are for 10*S* (+) and 10*R* (−) adducts, respectively. All energies are in kcal/mol.

Discussion

Intercalation Causes Distortion near the Lesion Site in Both 10*S* (+) and 10*R* (−) Adducts Which is Modulated by the Stereochemistry at the Covalent Linkage Site. Intercalation of a planar aromatic moiety into a DNA double helix, in which all base pairs are maintained, requires the duplex to stretch and unwind to accommodate the inserted 3.4 Å thick aromatic ring system.[81,86,87] Our observed increase in rise from the canonical B-DNA value of 3.4 Å[88] to more than double this value (Table 2), is consistent with the required stretching; however, the rise is greater than that for equilibrium-bound intercalators due to the constraints of the covalent linkage. This constraint also causes the severe deviations in buckle, propeller, and opening at the modified A*6-T17 base pair, as well as tilt in the intercalation pocket. We find considerable unwinding in the case of each intercalated isomer; associated with the

(86) Pigram, W. J.; Fuller, W.; Davies, M. E. *J. Mol. Biol.* **1973**, *80*, 361−365.
(87) Ornstein, R. L.; Rein, R. *Biopolymers* **1979**, *18*, 1277−1291.
(88) Arnott, S.; Hukins, D. W. L. *J. Mol. Biol.* **1973**, *81*, 93−105.

stretching and unwinding are backbone torsional angle distortions. The intercalation-induced distortions are, however, modulated by the 10*S* versus 10*R* stereochemistry and the associated 3′- versus 5′ intercalation in the 10*S* (+) and 10*R* (−) stereoisomers. Specifically, our results have shown greater distortion in the 10*S* (+) adduct (**S**), in terms of unwinding, torsional angle deviations, and quality of Watson−Crick hydrogen bonding. In addition, we note opposite buckle and propeller of the A*6-T17 base pair in the stereoisomeric pair, stemming from the 10*S* versus 10*R* stereochemistry at the linkage site. The effects on unwinding and Watson−Crick base pairing are discussed below in detail.

Better BP-Base Stacking Causes Stronger van der Waals Interactions in the 10*R* (−) Adduct. Our results show that the BP is stacked on only one face in the 10*S* (+) adduct, while it is stacked on both faces in the 10*R* (−) adduct (Figure 10); this is consistent with our analysis of the van der Waals interaction energies in the 10*S* (+) and 10*R* (−) adducts (Table 1), which reveals that the total van der Waals interaction difference between the 10*S* (+) and 10*R* (−) adducts is 4.0 kcal/mol, favoring the 10*R* (−) adduct (**R**). This is the second largest contribution to the free energy difference between the two adducts. Moreover, the difference in the van der Waals interaction energy of the BP and the residues involved in the intercalation pocket in the 10*S* (+) and 10*R* (−) adducts is 3.7 kcal/mol, favoring the 10*R* (−) adduct (**R**) (Table 3), indicating that the total van der Waals interaction difference in the two adducts is due mainly to stacking differences in the intercalation pockets. The key stacking difference, accounting for the different van der Waals interaction energies, between the 10*S* (+) and 10*R* (−) adducts is that the BP in the 10*S* (+) adduct (**S**) is hardly stacked with its adjacent A7-T16 base pair, while the BP in the 10*R* (−) adduct (**R**) is well stacked with C5-G18, its adjacent base pair.

B-DNA Right-Handed Helical Twist is the Origin of Better Stacking of BP in the 10*R* (−) Adduct. As discussed above, the differing BP-base stacking interactions are key contributors to the differential stabilities of the two adducts in our free energy computations. The question arises: why is this the case? In Figure 10, the A*6-T17 base pair is oriented in a constant frame of reference in the 10*S* (+) and 10*R* (−) adducts. In the case of the 10*S* (+) adduct, intercalated on the 3′-side of the modified adenine A*6, stacking is with the A7-T16 base pair (Figure 10a); however, in the 10*R* (−) adduct, intercalated on the 5′-side of A*6, stacking is with C5-G18 (Figure 10b).

The stacking difference between the two adducts is imposed by the right-handed helical twist of the B-DNA. Although the intercalation of the bulky carcinogen substantially distorted the DNA helix in both the 10*S* (+) and 10*R* (−) adducts, the integrity of the B-DNA conformation is still maintained (Figure 3). As shown in Figure 10b, the BP is intercalated on the 5′-side of the modified A*6, in the 10*R* (−) adduct, and the long axis of the BP is positioned almost parallel to the long axis of the adjacent base pair C5-G18; therefore, the BP is well-stacked with the C5-G18. However, in the 10*S* (+) adduct, the BP is intercalated on the 3′-side of the modified A*6; due to the right-handed helical twist of B-DNA, the A7-T16 base pair, adjacent to the BP, is now twisted away from the optimal position for stacking (Figure 10a). As a result, the BP is not optimally overlapped with its adjacent base pair. In turn, the van der Waals interaction energy is much less favorable in the 10*S* (+) adduct than that in the 10*R* (−) adduct.

The 10*S* (+) Adduct is More Unwound than the 10*R* (−) Adduct. The right-handed helical twist not only imposes better stacking between the BP and the adjacent C5-G18 base pair in the 10*R* (−) adduct, it also introduces different steric effects in the two adducts. As shown in Figure 11a, the BP, intercalated on the 3′-side of the modified adenine A*6 in the 10*S* (+) adduct, has the benzylic ring of BP facing toward the 5′-side of A*6. On the other hand, in the 10*R* (−) adduct, the BP, intercalated on the 5′-side of the modified adenine, has the benzylic ring facing toward the 3′-side of A*6, as shown in Figure 11b. Due to the right-handed helical twist of B-DNA, C5-G18, the 5′-neighboring base pair to the modified A*6-T17 in the 10*S* (+) adduct, is twisted toward the benzylic ring of BP, facing the 5′-side of the modified adenine A*6. This causes steric crowding between the bulky BP benzylic ring and C5, which is destabilizing to the DNA helix. In the 10*R* (−) adduct, A7-T16, the base pair neighboring A*6-T17 on the 3′-side, is twisted away from the benzylic ring of BP, which faces the 3′-side of A*6. As a result, there is no steric crowding in the 10*R* (−) adduct. This phenomenon was discussed previously by Schwartz et al.,[20] and our MD simulations support these conclusions. We also note that the twist of the C5-G18 to A*6-T17 base pair step in the 10*S* (+) adduct is only ∼16° (Figure 6), which is substantially smaller than the average value in structure **C**, ∼31°. This local unwinding at the C5-G18 to A*6-T17 base pair step in the 10*S* (+) adduct is part of the accommodation to alleviate the steric crowding between the benzylic ring and the C5-G18 base pair imposed by the right-handed helical twist of B-DNA.[20] In turn, this local unwinding, additional to the unwinding associated with the intercalation of the BP, causes more structural distortion, namely more deviations in the backbone torsions α, β, and γ, in the 10*S* (+) adduct than in the 10*R* (−) adduct.

Total unwinding due to intercalation itself is almost the same in the 10*S* (+) and 10*R* (−) adducts, but manifests differently in the two stereoisomers. Specifically, the twist in the intercalation pocket, between C5-G18 and A*6-T17, in the 10*R* (−) adduct (**R**) case is ∼31°, quite similar to that of structure **C** (Table 2). However, for the A*6-T17 to A7-T16 intercalation pocket in the 10*S* (+) adduct (**S**), the twist is only ∼9°. In each case the local twist angle appears to be optimized for stacking. The 9° twist for the 10*S* (+) adduct (**S**) is very unwound, compared to the unmodified control (**C**), but appears to produce better stacking with the neighboring base pair A7-T16 than would the normal value of ∼35° seen in structure **C** (Table 2). In addition, in the 10*R* (−) adduct (**R**), there is marked local unwinding next to the intercalation pocket, at the A*6-T17 to A7-T16 base pair step, with a twist angle of ∼6°. This unwinding is likely due to the intercalation at the neighboring C5-G18 to A*6-T17 base pair step. Unwinding seems to have been displaced from the intercalation pocket in the 10*R* (−) adduct (**R**) to permit optimal stacking between the BP and the adjacent C5-G18 base pair.

The total unwinding near the lesion site in the 10*S* (+) adduct (**S**), due to intercalation itself and due to the BP steric effect, is ∼41°, compared to structure **C**, while the 10*R* (−) adduct (**R**) is unwound ∼29°, compared to structure **C** (Figure 6). Therefore, the 10*S* (+) adduct is more unwound than the 10*R* (−) adduct by ∼12°, compared in each case to structure **C**. This additional unwinding imposes greater distortion in the backbone torsions.

***syn*−*anti* Conformational Heterogeneity and a Diminished Quality of Watson−Crick Hydrogen Bonding in the 10*S* (+) Adduct.** NMR solution structures[20−25] have revealed that the 10*S* (+) adduct is conformationally heterogeneous in solution, to the extent that structural models were obtainable in the A*-G

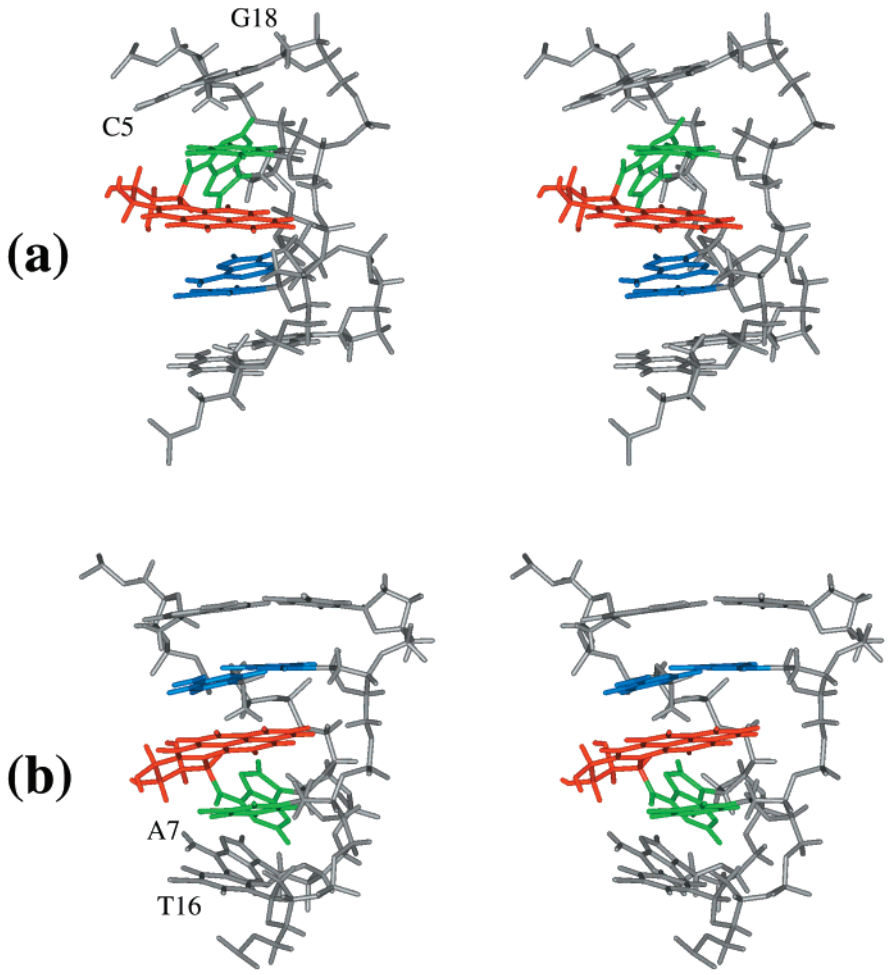

Figure 11. Stereoviews of the d(CA*AG)·d(CTTG) region of the 10*S* (+) adduct (**S**), and the d(ACA*A)·d(TTGT) region of the 10*R* (−) adduct (**R**). (a) 10*S* (+) adduct (**S**). BP is in red, A*6-T17 is in green, and A7-T16 is in blue. (b) 10*R* (−) adduct (**R**). BP is in red, A*6-T17 is in green, and C5-G18 in blue. All backbone, sugar atoms and other residues are in gray. The structures are aligned to have the 5′ residue of the modified strand at top left.

mismatch context, but not when the normal partner T was opposite the modified A. In the A*-G mismatch context a *syn*−*anti* equilibrium was observed,[20,21] and it has been suggested that there may be a similar equilibrium in the A*-T case.[23,24] Our dynamics results do indeed reveal such a *syn*−*anti* equilibrium for the 10*S* (+) isomer only, and this equilibrium induces greater deviation from ideal Watson−Crick base pairing in this isomer. The explanation for this *syn*−*anti* equilibrium in this isomer has been suggested by Volk et al.,[24] who noted that the steric crowding caused by the benzylic ring in this isomer is, in part, also alleviated by rotation of the glycosidic bond toward the *syn* region. This is clearly seen in Figure 8.

Computed Enthalpy Difference in the 10*S* (+) and 10*R* (−) Adducts Agrees Reasonably Well with Thermal Melting Data. The melting temperature for the DNA duplex d(CGG-ACA*AGAAG)·d(CTTCTTGTCCG) containing a 10*S* (+)-*trans-anti*-[BP]-N^6-dA adduct is 25 °C, while that of the modified DNA duplex containing a 10*R* (−)-*trans-anti*-[BP]-N^6-dA adduct in the same sequence context is 33 °C.[47] This is the same sequence employed in our computations. On the basis of the melting temperatures of the two isomeric DNA adducts and calorimetric studies of DNA melting,[89] we estimate the transition enthalpy difference between the two isomers to be 5.7 ± 1.9 kcal/mol. This is reasonably close to our computed enthalpy difference of 10.2 kcal/mol. However, the enthalpies in the calorimetric studies are melting transition enthalpies which refer to the double helix-to-coil transition, and they would not be expected to be the same as our computed enthalpies. The helix-to-coil melting transition measures the enthalpy difference between the double helix and single-stranded coils for each isomer,

$$D^*_+ \xrightarrow{\Delta H^+_{Tm}} S^*_+ + S$$

$$D^*_- \xrightarrow{\Delta H^-_{Tm}} S^*_- + S$$

where D^*_+, S^*_+, and ΔH^+_{Tm} are the modified double helix, modified single strand, and the melting transition enthalpy for the 10*S* (+) adduct; D^*_-, S^*_-, and ΔH^-_{Tm} are the corresponding quantities for the 10*R* (−) adduct, and *S* is the unmodified single strand. However, we computed the enthalpy difference between the 10*S* (+) and 10*R* (−) adducts as double helices.

$$\Delta H_{computed} = H_{D^*_+} - H_{D^*_-}$$

where $H_{D^*_+}$ and $H_{D^*_-}$ are the enthalpy of the 10*S* (+) adduct and that of the 10*R* (−) adduct, respectively. These cannot be

(89) Chalikian, T. V.; Völker, J.; Plum, G. E.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7853−7858.

identical because the two modified single-stranded coils would not be identical in the two isomers, although they may be quite similar. Consequently, the reasonable agreement is encouraging.

Conclusions

Free energy analysis reveals that the 10*R* (−)-*trans-anti*-[BP]-N^6-dA adduct is more stable than the 10*S* (+)-*trans-anti*-[BP]-N^6-dA isomer. This is consistent with melting temperature measurements for the two isomeric adducts in the same human *N-ras* codon 61 sequence context. The origin of the stability difference resides in better stacking of the intercalated BP moiety with adjacent base pairs in the 10*R* (−) adduct, together with steric differences involving the stereoisomeric benzylic rings.[20] Specifically, in the more stable isomer, there is much greater overlap between the BP aromatic rings and the neighboring base pair, with only one face of the BP stacked in the intercalation pocket in the 10*S* (+) adduct, while both faces are stacked in the 10*R* (−) adduct. The difference stems from the 3′- versus 5′ intercalation in the 10*S* (+) and 10*R* (−) isomers, respectively, and the associated different neighboring base pair position relative to the modified base pair, imposed by the B-DNA right-handed helical twist. While the DNA is unwound due to intercalation in both isomers, additional unwinding is imposed on the 10*S* (+) adduct to accommodate steric crowding by the benzylic ring of BP in this stereoisomer;[20] this, in turn, causes more distortion. Furthermore, we observe a *syn*−*anti* equilibrium in the dynamics trajectory of only the 10*S* (+) isomer, with attendant reduced quality of Watson−Crick base pairing, which is also due to benzylic ring crowding.[24] Together, the better stacking, lesser unwinding, and better quality of Watson−Crick base pairing account for the greater thermal stability of the 10*R* (−) isomer. The observed greater susceptibility to nucleotide excision repair in human cell extracts in the 10*S* (+) isomer[46] is in accord with its lesser stability, in agreement with recent evidence that DNA helix destabilization facilitates human nucleotide excision repair.[90−92] In turn, the greater stability and lower repair susceptibility of the 10*R* (−) isomer may contribute to genotoxicity.

Acknowledgment. *We dedicate this work to the memory of the late Peter Kollman.* This research is supported by NIH Grant CA-28038 and DOE Grant DE-FG02-90ER60931 to S.B., and NIH Grant CA-76660 to N.E.G. We thank Dr. Thomas Cheatham, III, for reading the manuscript and providing very helpful suggestions. Computations were carried out at the NSF San Diego Supercomputer Center, the NSF Advanced Computing Center for Engineering and Science at the University of Texas at Austin, and the DOE National Energy Research Scientific Computing Center. Our own SGI workstations, as well as those of the Academic Computing Services at New York University were used for visualization, modeling and data analysis.

Supporting Information Available: Table S1 gives partial charges, atom type assignments and topologies for the 10*S* (+)- and 10*R* (−)-*trans-anti*-[BP]-N^6-dA nucleotide adducts; Table S2 gives added force field parameters; Figure S1 shows the rmsd's relative to the starting structures for the 10*S* (+) adduct (**S**), 10*R* (−) adduct (**R**), and the unmodified control duplex (**C**) over the 2-ns production MD simulation; Figures S2 and S3 show stereo snapshots along the trajectory over the 2-ns production MD simulation for the 10*S* (+) and 10*R* (−) adducts, respectively; Figure S4 shows torsion angles α' and β' for the 10*S* (+) and 10*R* (−) adducts over the 2-ns production MD simulation (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0043035

(90) Hess, M. T.; Schwitter, U.; Petretta, M.; Giese, B.; Naegeli, H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6664−6669.

(91) Hess, M. T.; Naegeli, H.; Capobianco, M. *J. Biol. Chem.* **1998**, *273*, 27867−27872.

(92) Buschta-Hedayat, N.; Buterin, T.; Hess, M. T.; Missura, M.; Naegeli, H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6090−6095.