A Phototautomerizable Model DNA Base Pair

Anthony K. Ogawa, Osama K. Abou-Zied, Vickie Tsui, Ralph Jimenez, David A. Case, and Floyd E. Romesberg*

Contribution from the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received May 22, 2000

Abstract: 2-(2′-Hydroxyphenyl)benzoxazole (HBO) undergoes rapid photoinduced proton transfer from the enol-imine to the keto-amine tautomer. When incorporated in duplex DNA opposite an abasic site, HBO appears to be a good mimic of a natural DNA base pair based on duplex stability, UV and CD spectroscopy, and molecular dynamics simulations. In the chosen sequence context, HBO exists exclusively as the *syn*-enol tautomer, which is consistent with an environment of the model base that is dominated by neighboring nucleobases and not by solvation. The *syn*-enol efficiently undergoes photoinduced proton transfer to the keto-tautomer, which should allow for the study of tautomerization within the duplex DNA.

Sequence-specific base pairing in duplex DNA is encoded by the specific hydrogen bonding (H-bonding) patterns of the adenine:thymine (dA:dT) and guanine:cytosine (dG:dC) base pairs. High-sequence specificity is critically dependent on the bases existing in their keto-amino tautomers. However, each base may be converted to its minor, but often relatively stable enol-immine tautomer by a double proton transfer (prototropic tautomerism, Figure 1a). These tautomeric base pairs are physiochemically distinct, possessing different hydrogen-bonding patterns and different dipole moments.[1] As a result, prototropic tautomerism may be involved in spontaneous mutagenesis, since the H-bonding patterns of the tautomers favor mispairing.[2] In addition, prototropic tautomerism may play an important role in the biological function of DNA[3] and may also be intimately related to the vibrational motions involved in duplex fluctuations.[4] Herein, we report a model DNA base pair that when incorporated into DNA is efficiently tautomerized by light absorption (phototautomerization). Thus, the model base pair should allow for the study of tautomerization in the biological environment of duplex DNA.

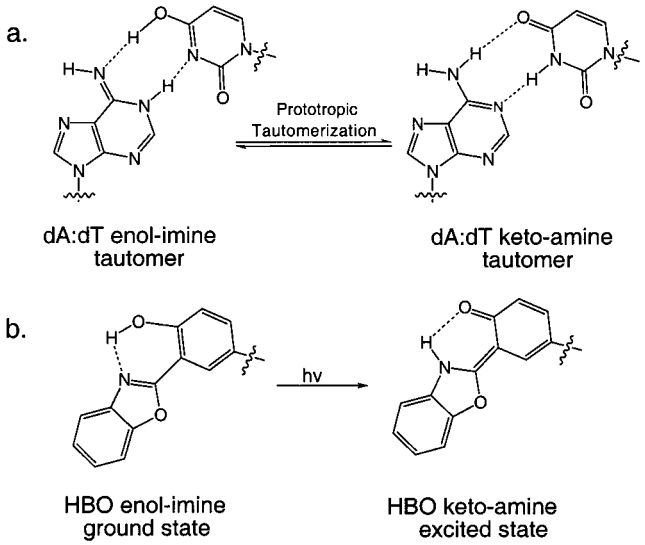

Figure 1. (a) Prototropic tautomerization of a natural base pair and (b) phototautomerization of model base pair 2-(2′-hydroxyphenyl)-benzoxazole (HBO).

Theoretical studies of base pair tautomerization have been reported,[5] but experimental investigations have been limited by an inability to initiate and follow proton transfer within a specific base pair in the DNA polymer. Experiments directed at examining tautomerization usually involve stable base analogues that are believed to favor rare tautomers, for example O^6-methyguanine or N^6-methoxyadenine.[2] These studies are not

* Author to whom correspondence should be addressed: Telephone (858) 784-7290, Fax (858) 784 7472, E-mail floyd@scripps.edu.

(1) (a) Bergman, E. D.; Weiler-Feilchenfeld *The Dipole Moments of Purines*; Bergman, E. D., Pullman, B., Eds.; The Israel Academy of Sciences and Humanities: Jerusalem, 1971; pp 21−28. (b) Lister, J. H. *The Purines. Supplement 1*; John Wiley and Sons: New York, 1996; Vol. 54. (c) Hurst, D. T. *An Introduction to the Chemistry and Biochemistry of Pyrimidines, Purines and Pteridines*; Page Bros.: Norwich, Great Britain, 1980. (d) Bugg, C. E. *Solid-state packing patterns of purine bases*; Bergman, E. D., Pullman, B., Eds.; The Israel Academy of Sciences and Humanities: Jerusalem, 1971; pp 178−204.

(2) (a) Nishio, H.; Ono, A.; Matsuda, A.; Ueda, T. *Nucleic Acids Res.* **1992**, *20*, 777−782. (b) Strazewski, P.; Tamm, C. *Angew. Chem., Int. Ed.* **1990**, *29*, 36−57. (c) Singer, B.; Chavez, F.; Goodman, M. F.; Essigmann, J. M.; Dosanjh, M. K. *Proc. Natl. Acad Sci. U.S.A.* **1989**, *86*, 8271−8274.

(3) (a) Chachisvilis, M.; Fiebig, T.; Douhal, A.; Zewail, A. H. *J. Phys. Chem. A* **1998**, *102*, 669−673. (b) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355. (c) Topiol, S.; Weinstein, H.; Osman, R. *J. Med. Chem.* **1984**, *27*, 1531. (d) Weinstein, H.; Chou, D.; Johnson, C. L.; Kang, S.; Green, J. P. *Mol. Pharmacol.* **1976**, *12*, 738. (e) Weinstein, H.; Mazurek, A. P.; Osman, R.; Topiol, S. *Mol. Pharmacol.* **1986**, *12*, 738. (f) Osman, R.; Weinstein, H.; Topiol, S.; Rubenstein, L. *Clin. Physiol. Biochem.* **1985**, *3*, 80. (g) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285. (h) Rhoads, R. E.; Hellman, G. M.; Remy, P.; Ebel, J. P. *Biochemistry* **1983**, *22*., 6048.

(4) Brauns, E. B.; Madaras, M. L.; Coleman, R. S.; Murphy, C. J.; Berg, M. A. *J. Am. Chem. Soc.* **1999**, *121*, 11644−11649.

(5) (a) Kwiatkowski, J. S.; Zeilinski, T. J.; Rein, R. *Adv. Quantum Chem.* **1986**, *18*, 85−130. (b) Sabio, M.; Topiol, S.; Lumma, W. C., Jr. *J. Phys. Chem.* **1990**, *94*, 1366−1372. (c) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811−6821. (d) Guallar, V.; Douhal, A.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. A* **1999**, *103*, 6251−6256. (e) Broo, A.; Holmén, A. *J. Phys. Chem. A* **1997**, *101*, 3589−3600. (f) Gorb, L.; Leszczynski, J. *J. Am. Chem. Soc.* **1998**, *120*, 5024−5032. (g) Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J. Mol. Struct. (THEOCHEM)* **1995**, *331*, 147−154. (h) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247−3276.

(6) (a) Chen, Y.; Gai, F.; Petrich, J. W. *J. Am. Chem. Soc.* **1993**, *115*, 10158−10166. (b) Douhal, A.; Guallar, V.; Moreno, M.; Lluch, J. M. *Chem. Phys. Lett.* **1996**, *256*, 370−376. (c) Takeuchi, S.; Tahara, T. *J. Phys. Chem. A* **1998**, *102*, 7740−7753, (d) Chou, P.-T.; Wei, C.-Y.; Wu, G.-R.; Chen, W.-S. *J. Am. Chem. Soc.* **1999**, *121*, 12186−12187.

10.1021/ja001778n CCC: $19.00
Published on Web 10/03/2000

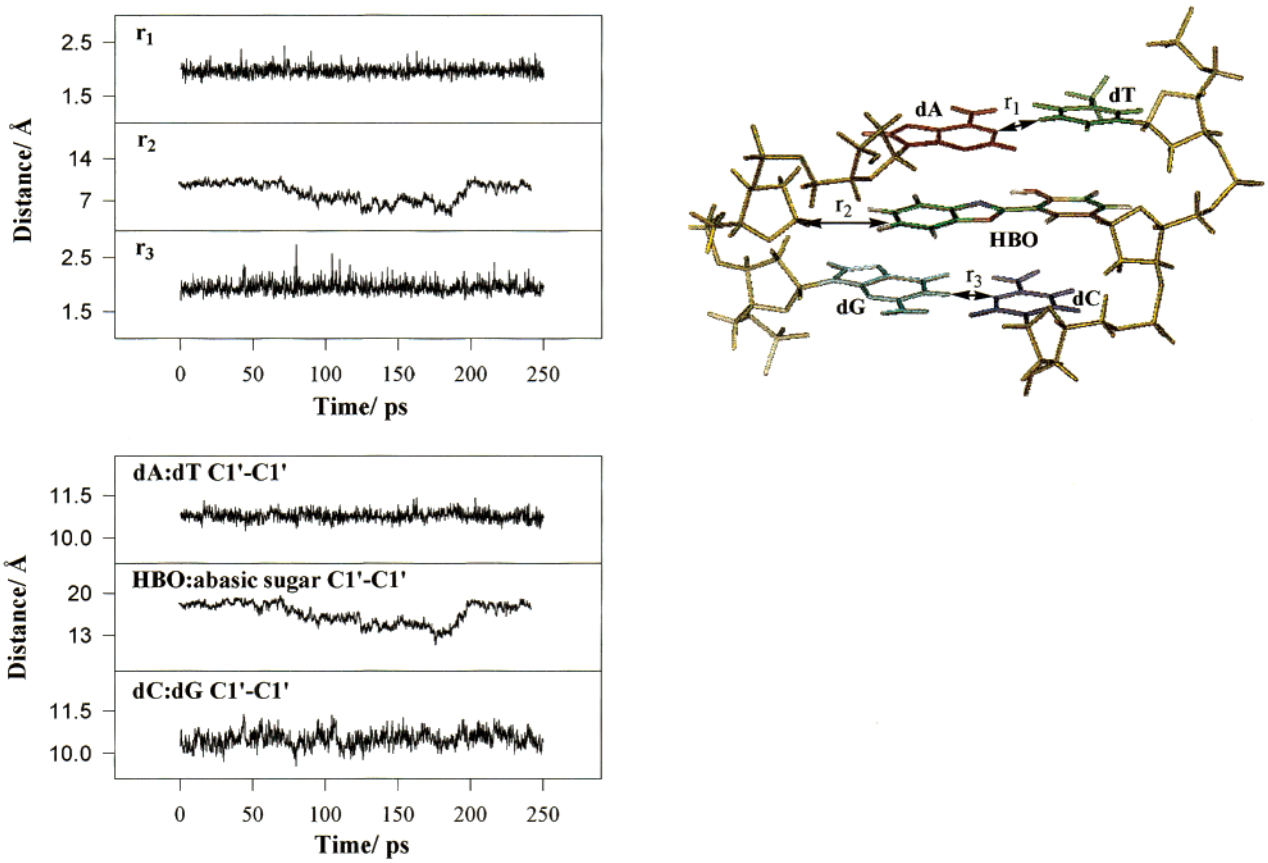

Figure 2. Distance fluctuations in duplex **1** containing HBO paired opposite an abasic site over the final 250 ps of simulation. Distances r_1, r_2, and r_3 (upper left panel) correspond to those labeled in the figure. The C1′−C1′ distance (lower left panel) corresponds to the distance between the C1′ sugar carbons at the model and flanking base pairs.

ideal, because the base analogues are known to cause structural perturbations to the DNA duplex. Tautomerization has also been examined with model nucleobases, free in solution, such as 7-azaindole.[6] These studies have shown that solvation may play a central role in facilitating the prototropic tautomerization of the free bases, either electrostatically or by direct proton exchange with the base H-bond donors and acceptors (water-facilitated prototropic tautomerization). However, base stacking in duplex DNA may result in an environment that is significantly different from that experienced by the base in solution. Overall, base pair tautomerization may be influenced by the water located in the major and minor grooves, as well as the details of the duplex environment, such as nucleotide sequence and thermal fluctuations.

2-(2′-Hydroxyphenyl)benzoxazole (HBO) is a member of a class of compounds that are known to undergo very rapid (subpicosecond) excited-state intramolecular proton transfer (ESIPT).[7,8] HBO may exist in two tautomeric forms, the keto and enol tautomers (Figure 1b), whose relative stabilities depend on the electronic state. In the electronic ground state, HBO exists as the enol-imine tautomer, while in the first excited singlet state the keto-amine tautomer is more stable. Therefore, light absorption triggers a very fast tautomerization by ESIPT from the hydroxyl group to the endocyclic nitrogen (Figure 1). The ring structure of HBO consists of a six-membered phenol ring covalently attached to the fused five−six membered heterocyclic benzoxazole. If positioned correctly in DNA, opposite an abasic site, the HBO ring structure resembles a Watson−Crick base pair. More specifically, the HBO ground-state enol is expected to resemble the enol-imine nucleobase tautomer, while the excited HBO keto-tautomer is expected to resemble the keto-amino nucleobase tautomer (Figure 1). In this manner, HBO should serve as a phototautomerizable model base pair.

The HBO phosphoramidite was synthesized as described in the Supporting Information. The phosphoramidite was incorporated into oligonucleotides 5′-CGTTTCXTTCTC and 5′-GAGAAXGAAACG at the positions labeled X with an ABI 392 DNA/RNA synthesizer. Each oligonucleotide was then hybridized with a complementary oligonucleotide containing an abasic site located at the position opposite HBO, resulting in duplexes **1** and **2**, respectively. To examine the stability of the model base pair the duplex melting temperature (T_m) of **1** and **2** was determined. Duplexes **1** and **2** melted with T_m values of 38 and 34 °C, respectively. The stabilities of both **1** and **2** compare favorably with that of the duplex containing a natural dA:dT pair at the same position, T_m = 39 °C. The higher stability of HBO in **1** may result from the increased overlap of the benzoxazole fragment with flanking pyrimidines (dA and dG) in this sequence context. However, in both sequence contexts, the duplex stability implies that HBO is paired stably opposite the abasic site, suggesting that HBO is stabilized by interstrand hydrophobic packing without significant distortions of the neighboring base pairs. These conclusions are further strengthened by circular dichroism studies (Supporting Informa-

(7) (a) Krishnamurthy, M.; Dogra, S. K. *J. Photochem.* **1986**, *32*, 235−242. (b) Arthen-Engeland, T.; Bultmann, T.; Ernsting, N. P.; Rodriguez, M. A.; Thiel, W. *Chem. Phys.* **1992**, *163*, 43−53. (c) Woolfe, G. J.; Melzig, M.; Schneider, S.; Dörr, F. *Chem. Phys.* **1983**, *77*, 213−221.

(8) (a) Kyrychenko, A.; Herbich, J.; Wu, F.; Thummel, R. P.; Waluk, J. *J. Am. Chem. Soc.* **2000**, *122*, 2818−2827. (b) Mosquera, M.; Penedo, J. C.; Ríos Rodríguez, M. C.; Rodríguez-Prieto, F. *J. Phys. Chem.* **1996**, *100*, 5398−5407. (c) Barbara, P. F.; Walsh, P. K.; Brus, L. E. *J. Phys. Chem.* **1989**, *93*, 29−34. (d) Smith, T. P.; Zaklika, K. A.; Thakur, K.; Walker, G. C.; Tominaga, K.; Barbara, P. F. *J. Phys. Chem.* **1991**, *95*, 10465−10475.

tion). CD spectra of duplex **1**, as well as the single stranded oligonucleotide, demonstrate right-handed helix formation with a structure similar to that of the A-form duplex.

To further understand the structure of duplex **1**, we conducted molecular dynamics simulations using AMBER 6.[9] The Cheatham et al.[10] force field was used, which is a slight modification of the Cornell et al. force field.[11] Atom charges for HBO were generated from electrostatic potential fitting using the Jaguar program,[12] at the 6-31G* level. The simulation was run from two starting structures corresponding to HBO in intercalated or extrahelical (rotated by 45° relative to helical axis) configurations. The DNA was solvated in a box of explicit water molecules with neutralizing sodium counterions. The DNA structures were first minimized, followed by water and sodium ion equilibration. A 2 ns constant volume production run was then executed at 300 K. No significant differences were found in the average structures which resulted from the two starting configurations. The trajectories were stable and predict that the DNA remains in a double helix structure with occasional fraying of base pairs at the duplex termini. Most importantly, the remaining base pairs, including those flanking HBO, were not fluctional and remained H-bonded throughout the simulation (Figure 2). HBO fluctuations were larger than those of the flanking base pairs (Figure 2), probably due to the absence of a covalent linkage between HBO and the abasic furan ring. Regardless, HBO remained intercalated throughout the simulation with the benzoxazole ring well stacked between the flanking bases of the opposite strand. There were two significant distortions in the DNA duplex resulting from the model base pair. The C1′−C1′ distance between nucleoside sugar rings of the HBO:abasic pair was greater than that for the natural pairs (Figure 2) leading to a slight (∼20%) widening of the minor groove as monitored by the closest phosphate distance. This is consistent with the structure apparent in the CD spectra, as discussed above. There was almost no change in the base pair rise; however, there was a slight (4 degree) decrease in helical twist at the model pair. Overall, the simulations predict that the structure of duplex DNA containing the model base pair is not significantly perturbed and implies that HBO paired opposite an abasic site is a good mimic of a natural base pair (Figure 3).

The photophysics of HBO is known to be complicated by solvent-dependent conformational isomerism.[7] The *syn*-enol (intramolecular O−H−N hydrogen bond), *anti*-enol (intramolecular O−H−O hydrogen bond), and in polar protic solvents the "strongly solvated enol" (intermolecular hydrogen bonding with solvent) are each observable in different solvents and have unique spectroscopic signatures. The more polar protic solvents favor the "strongly solvated enol" that is characterized by a long-wavelength absorption band (∼375 nm) with a minimally Stokes-shifted emission band (∼428 nm). In less polar solvents, the *syn*- and *anti*-enols dominate, and both exhibit characteristic absorption maxima at approximately 334 nm, while only the *syn*-enol shows a large Stokes-shift (emission maximum ∼500 nm) due to ESIPT.

The steady-state absorption and emission spectra of HBO were measured in hexane, DMSO, and pyridine, as well as in the DNA duplex **1**. Spectra of **1** (3 μM) were measured in 10 mM phosphate buffer (pH 7.0) and 100 mM NaCl. Our results for free HBO in each solvent agree well with previously reported data.[7] In double-stranded DNA, with the sequence context of duplex **1**, HBO exists exclusively as the *syn*-enol (Figure 4). This assignment is demonstrated by the 338 nm absorption maximum characteristic of the *syn*-enol in a hydrophobic environment, the absence of a red shifted absorption (∼375 nm) corresponding to a "strongly solvated" *anti*-enol, and the presence of a single, strongly Stokes-shifted fluorescence peak at 500 nm. The absence of a "strongly solvated" enol, even in aqueous solution, implies that the model base pair is sequestered in a very hydrophobic environment, and may be well shielded from the solvent. Alternatively, the electrostatic environment of DNA may sufficiently bias HBO in favor of the *syn*-enol. In either case, the environment of the model base pair appears to be dominated by interbase packing effects within the duplex and not by solvation effects, despite the availability of waters of solvation in both the major and minor grooves.[13]

In summary, HBO paired in duplex DNA opposite an abasic site appears to be a good mimic of a natural DNA base pair.

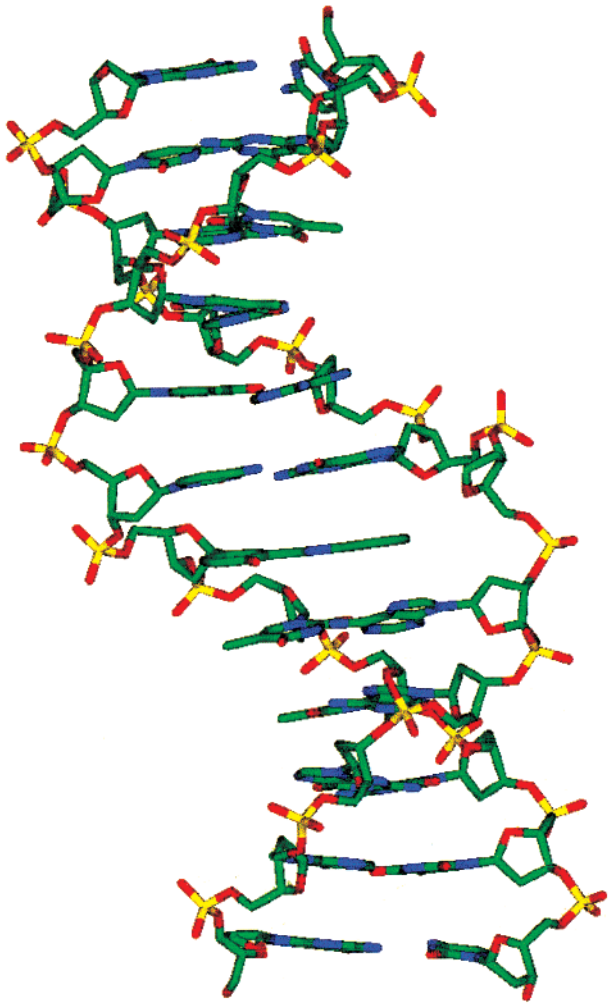

Figure 3. The average structure of duplex **1** over the entire 2 ns of simulation.

(9) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E. I.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER6*; University of California: San Francisco, 1999.

(10) Cheatham, T. E. I.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.

(11) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. J.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(12) *Jaguar 3.5*, Schrödinger, Inc., Portland, OR, 1988

(13) (a) Drew, H. R.; Travers, A. A. *Cell* **1988**, *37*. (b) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−56. (c) Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129−46. (d) Liepinsh, E.; Otting, G.; Wüthrich, K. *Nucleic Acids Res.* **1992**, *20*, 6549−53. (e) Kubinec, M. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 8739−8740.

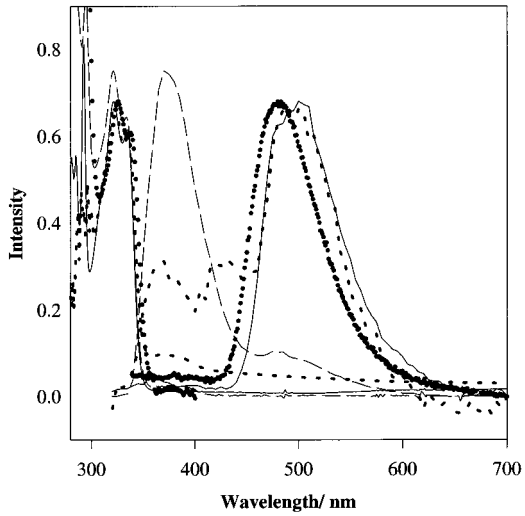

Figure 4. Absorption and fluorescence spectra of HBO in hexane (solid line), DMSO (dashed line), pyridine (dotted line), as well as duplex **1** (solid circles). The excitation wavelength was 334 nm.

The duplex DNA containing the model base pair is only marginally destabilized, and is significantly more stable than a mispair between bases. In addition, the CD spectra and the simulations show a duplex that is structurally and dynamically very similar to native DNA. Moreover, each tautomeric form of HBO has a characteristic spectral signature that accurately reports on the prototropic tautomerism of the model base pair. The steady state absorption and fluorescence data are consistent with an HBO environment that is very hydrophobic, implying that HBO efficiently interstrand packs and fills the space vacated by the abasic site. The steady state spectra also show that in duplex DNA, HBO exists exclusively as the *syn*-enol, and efficiently undergoes ESIPT, resulting in the keto-tautomer. Therefore, the model base pair should allow for a fast tautomerization within the duplex, free from complicating structural effects. Time-resolved experiments are currently underway to determine the kinetics of the proton transfer and the lifetime of the resultant HBO-keto tautomer, as well as any sequence dependencies.

Acknowledgment. Funding was provided by the La Jolla Interface in Science Fellowship (V.T.) and NIH grant RR.122555 (D.A.C.).

Supporting Information Available: Experimental procedures, characterizations, CD spectra, and simulation details and results for the entire duplex (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA001778N