BIOCHEMISTRY
including biophysical chemistry & molecular biology

Rapid Report

pubs.acs.org/biochemistry

Guanine to Inosine Substitution Leads to Large Increases in the Population of a Transient G·C Hoogsteen Base Pair

Evgenia N. Nikolova,[†] Frederick Stull,[‡] and Hashim M. Al-Hashimi*[,§]

[†]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, United States

[‡]Program in Chemical Biology, The University of Michigan, Ann Arbor, Michigan 48109, United States

[§]Department of Biochemistry and Chemistry, Duke University School of Medicine, 307 Research Drive, Nanaline H. Duke Building, Durham, North Carolina 27710, United States

Supporting Information

ABSTRACT: We recently showed that Watson−Crick base pairs in canonical duplex DNA exist in dynamic equilibrium with G(*syn*)·C$^+$ and A(*syn*)·T Hoogsteen base pairs that have minute populations of ∼1%. Here, using nuclear magnetic resonance $R_{1\rho}$ relaxation dispersion, we show that substitution of guanine with the naturally occurring base inosine results in an ∼17-fold increase in the population of transient Hoogsteen base pairs, which can be rationalized by the loss of a Watson−Crick hydrogen bond. These results provide further support for transient Hoogsteen base pairs and demonstrate that their population can increase significantly upon damage or chemical modification of the base.

We recently provided evidence[1−4] based on nuclear magnetic resonance (NMR) $R_{1\rho}$ relaxation dispersion (RD)[5−7] and single-atom substitution experiments that G·C and A·T Watson−Crick (WC) base pairs (bps) transiently form Hoogsteen (HG) bps[8] with populations of ∼0.4 and 0.7% and lifetimes of 1.7 and 0.3 ms at pH ∼5.4 (Figure 1A). HG bps form through a 180° rotation of the purine base around the glycosidic bond from an *anti* to a *syn* conformation (Figure 1A). By modifying the structural and chemical presentation of DNA and modulating nucleobase accessibility, HG bps can play unique roles (reviewed in ref 9) in DNA−protein recognition,[10−13] DNA damage induction[14] and repair,[15,16] and replication.[17,18]

We provided evidence of transient HG bps based on analysis of $R_{1\rho}$ RD experiments, which quantify the degree of line broadening due to chemical exchange with a transient low-population species.[7] A two-state analysis of the RD data measured on a variety of G·C and A·T Watson−Crick bps in canonical duplex DNA was consistent with the existence of a transient species that features downfield-shifted purine base C8 and sugar C1′ chemical shifts. These chemical shifts are a hallmark of *syn* purine, where the base flips 180° about the glycosidic bond. Such a *syn* purine can be paired with pyrimidine via HG pairing. By using chemical modification, including N1-methyladenine (N1-Me-A) and N1-methylguanine (N1-Me-G), which are naturally occurring damaged forms of purine bases, we successfully trapped HG bps in duplex DNA and showed that they have the characteristic downfield-shifted carbon chemical shifts.[1] As further support for the assignment of HG bps for the observed transient state, we subsequently showed that the single-atom substitution, 7-deazaguanine and 7-deazaadenine, which specifically knocks out an HG H-bond between guanine N7 and cytosine N3 (N7···H−$^+$N3) without affecting WC H-bonds, destabilizes the transient state, such that it can no longer be detected by NMR RD.[2]

Here, we sought to shift the WC−HG equilibrium in the other direction, toward the HG state. In G·C bps, a transition toward HG bps preserves the O6···H−N4 H-bond present in WC pairing but leads to the loss of the N1−H···N3 and N2−H···O2 H-bonds, which are replaced with a single N7···H−$^+$N3 H-bond that further requires protonation of cytosine N3 (Figure 1). Therefore, one would expect that, in inosine, the loss of the exocyclic amino group at position C2 of guanine would lead to the loss of a WC H-bond without affecting HG

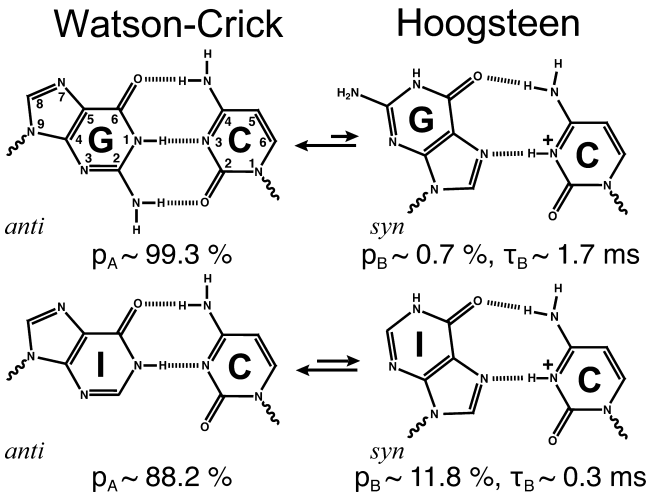

Figure 1. WC−HG equilibrium in canonical G·C and inosine-substituted I·C base pairs. Shown are the relative populations (p_A and p_B) and transient state lifetimes (τ_B) obtained from ^{13}C NMR RD at 26 °C and pH 5.4.

Received: September 24, 2014
Revised: October 19, 2014
Published: October 23, 2014

   dx.doi.org/10.1021/bi5011909 | *Biochemistry* 2014, 53, 7145−7147

pairing and, consequently, lead to an increase in the population of the transient state observed using NMR RD.

To examine the impact that replacement of guanine in the CA step with inosine would have on the WC−HG equilibrium, we prepared an A_6-DNA duplex (A_6-DNA^{I10}) (Figure 2A), in

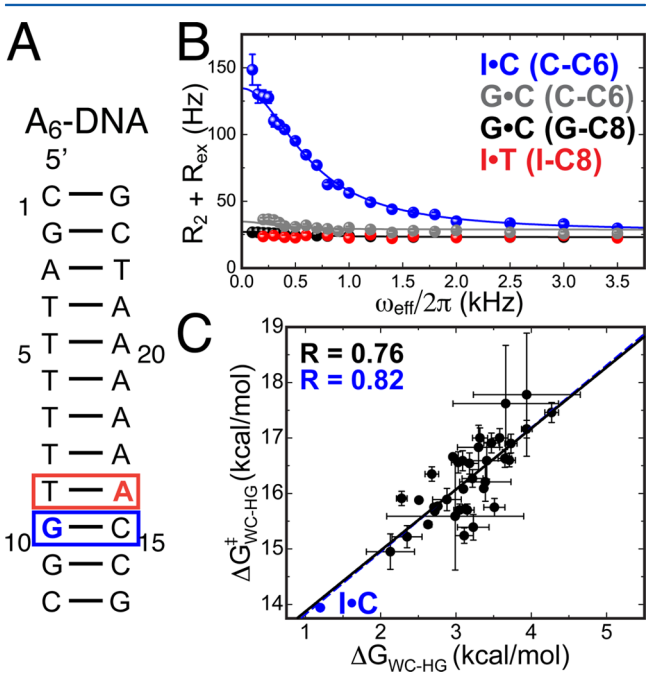

Figure 2. Increased level of $R_{1\rho}$ RD and HG bp formation at I·C. (A) A_6-DNA highlighting the A·T or G·C bp substituted with inosine at the purine base. (B) On-resonance ^{13}C $R_{1\rho}$ RD profiles comparing chemical exchange in G·C, I·C, and I·T bps (pH 5.2 for G·C C-C6 and pH 5.4 for all others). Solid lines represent best global fits to eq S1 of the Supporting Information. (C) Correlation between $\Delta G^{\ddagger}_{WC\text{-}HG}$ and $\Delta G_{WC\text{-}HG}$ for WC bps[4] (data for the I·C bp colored blue). Shown are the best-fit line and corresponding Pearson coefficient (R) with (---) or without (—) inclusion of the I·C bp.

which G10 was replaced with I10. We previously reported detailed NMR characterization of transient HG bps in A_6-DNA as well as variants that bear chemical modifications and mutations at the same position.[1−3] Because of difficulties in preparing isotopically $^{13}C/^{15}N$-labeled oligonucleotides with a modified inosine base that are required for obtaining high-quality off-resonance RD data, the strand complementary to the inosine-containing strand was ^{13}C- and ^{15}N-labeled. The two-dimensional $^{13}C-^{1}H$ heteronuclear single-quantum coherence NMR spectra of this selectively labeled strand were in very good agreement with spectra of A_6-DNA indicating formation of a WC-like I·C bp that does not impact neighboring bps (Figure S1 of the Supporting Information). Because I10 could not be isotopically enriched and the unlabeled strand concentration was not high enough (~0.5 mM), making it difficult to measure reliable $R_{1\rho}$ RD data at natural abundance, the complementary labeled C15 residue was used to probe chemical exchange at the I·C bp. In particular, formation of a G·C$^+$ HG bp is accompanied by a significant downfield shift of the C-C6 chemical shift due to protonation of C-N3, and this can be used to probe the WC−HG equilibrium.[3] Here, replacement of G10 with I10 resulted in a measurable downfield chemical shift change toward the transient HG state as well as significant line broadening at C15-C6, consistent with increased WC−HG exchange at the I·C bp.

To quantify transient HG bp formation, we collected on- and off-resonance ^{13}C RD dispersion on C15-C6 in A_6-DNA^{I10}. We observed substantially larger RD for C-C6 in the I·C bp than for C-C6 or G-C8 in the G·C bp under similar conditions (Figure 2B).[1] A two-state analysis of the RD data allowed for determination of all chemical exchange parameters, including the transient state chemical shift [$\Delta\omega_{AB(C6)} = \omega_{B(C6)} - \omega_{A(C6)}$], population ($p_B$), and lifetime ($\tau_B = 1/k_2$) (Figure S2 and Table 2 of the Supporting Information). The transient state chemical shift for C15-C6 is ~2.3 ppm downfield-shifted relative to the GS, consistent with formation of a HG bp, and is in excellent agreement with the downfield shift (~2.3 ppm) obtained when the G10·C15 bp in A_6-DNA is changed to a HG bp using guanine N1-methylation.[1]

The magnitude of the C15-C6 downfield chemical shift is similar to that previously seen in C-C6 (~2.2 ppm) or G-C8 (~3.7 ppm) in transient G·C HG bps[1] and cannot explain the RD observed in the I·C bp being much larger than that in the G·C bp. Rather, the larger RD arises primarily from a substantial 17-fold increase in the population of HG bps in I·C (p_B ~ 11.8%) as compared to G·C (p_B ~ 0.7%). This is in accord with our expectations based on the loss of WC H-bonds without significantly affecting HG H-bonds. In particular, the relative stability of the WC bp is decreased by ~1.8 kcal/mol in I·C compared to that in G·C. This is consistent with previous studies showing an ~2 kcal/mol loss of B-DNA stability from a single G·C to I·C substitution[19,20] and could be explained by the loss of a single N−H···O bond or replacement of the N−H···O bond with a weaker C−H···O bond.[21] The increase in the HG population is accompanied by an increase in chemical exchange rate k_{ex} ($k_{ex} = k_1 + k_2$ increases ~6-fold from ~600 s^{-1} in G·C to 3500 s^{-1} in I·C), which corresponds to an ~100-fold increase in the forward rate constant k_1 and a much smaller ~5-fold increase in the reverse rate constant k_2. This translates into a simultaneous decrease in the free energy difference between WC and HG states ($\Delta G_{WC\text{-}HG}$) and the free energy required to reach the transition state ($\Delta G^{\ddagger}_{WC\text{-}HG}$), which is in turn consistent with selective destabilization of WC due to the inosine substitution. We recently reported a similar trend based on ϕ value analysis of WC-to-HG transitions measured across diverse DNA sequence and position contexts, where k_1 increased in sequences with more abundant HG bps and k_2 exhibited a much weaker sequence dependence.[4] These results implied that the variations in the relative energetic stabilities of HG bps are dominated by variations in the stabilities of WC bps, as demonstrated by a direct correlation between $\Delta G_{WC\text{-}HG}$ and $\Delta G^{\ddagger}_{WC\text{-}HG}$.[4] Importantly, the observed correlation of sequence-dependent ΔG values for WC bps extrapolates to the ΔG values for the modified I·C bp (Figure 2C) and strongly supports our earlier proposal that WC bp stabilities, unlike HG bp stabilities, are highly sequence-dependent.

The 5-fold increase in k_2 could arise from destabilization of the HG bp, possibly because of the loss of H-bonding with solvent[11,22] or, alternatively, stabilization of the TS due to removal of the bulky amine group and more facile rotation of the purine base within the helix. While the structure of the TS remains unknown, most crystal structures of HG bps display H-bonds between G-NH_2 and the backbone phosphate or water molecules. The loss of the NH_2 group in inosine could contribute to the small (<1 kcal/mol) decrease in the stability of the transient HG I·C$^+$ versus G·C$^+$ state.

As a control, we conducted on-resonance ^{13}C $R_{1\rho}$ RD experiments at natural abundance on inosine C8 in an I·T bp,

dx.doi.org/10.1021/bi5011909 | Biochemistry 2014, 53, 7145−7147

which should not be capable of stably forming HG bps. This is because only one H-bond can form in the I·T HG bp without additional tautomerization and/or protonation. Indeed, we observed no evidence of chemical exchange at I-C8 (Figure 2B), supporting the idea that the observed transient state results from specific pairing of inosine with an appropriate HG pairing partner.

The results presented here provide additional evidence in support of transient HG bps in a canonical duplex and also show that HG bps can exist in much greater abundance in I·C bps. Inosine is a rare form of mutagenic damage that is recognized and excised by conserved repair enzymes such as DNA glycosylases and endonuclease V.[23] It is conceivable that the HG-type bps participate more broadly in damage recognition and repair, as could be the case for G·G, A·A, and A·C mispairs where the partially exposed HG face of the *syn*-purine is specifically recognized by the DNA mismatch repair enzyme MutS[15] or for the common oxidative damage 8-oxoguanine, which can form a stable HG-type bp with adenine.[24] HG-type bps involving inosine have already been implicated as the source of mutagenic substitutions, most frequently from A-T to G-C.[25] For example, *syn*-inosine has been observed to form an HG-type bp with protonated adenine at near-neutral pH,[26] which affords a possible mode of recognition for $I{\cdot}A^+$ mispairs. The observation of more abundant HG bps in I·C also raises the possibility that other forms of damage and chemical modifications may increase the abundance of these bps in genomes where they can potentially conduct unique biological functions.

■ ASSOCIATED CONTENT

Supporting Information

Methods, Tables 1 and 2, and Figures S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: hashim.al.hashimi@duke.edu.

Funding

This study was supported by National Institutes of Health Grant GM089846 awarded to H.M.A.-H.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Nikolova, E. N., Kim, E., Wise, A. A., O'Brien, P. J., Andricioaei, I., and Al-Hashimi, H. M. (2011) *Nature 470*, 498−502.

(2) Nikolova, E. N., Gottardo, F. L., and Al-Hashimi, H. M. (2012) *J. Am. Chem. Soc. 134*, 3667−3670.

(3) Nikolova, E. N., Goh, G. B., Brooks, C. L., III, and Al-Hashimi, H. M. (2013) *J. Am. Chem. Soc. 135*, 6766−6769.

(4) Alvey, H. S., Gottardo, F. L., Nikolova, E. N., and Al-Hashimi, H. M. (2014) *Nat. Commun. 5*, 4786.

(5) Massi, F., Johnson, E., Wang, C., Rance, M., and Palmer, A. G., III (2004) *J. Am. Chem. Soc. 126*, 2247−2256.

(6) Korzhnev, D. M., Orekhov, V. Y., and Kay, L. E. (2005) *J. Am. Chem. Soc. 127*, 713−721.

(7) Palmer, A. G., III (2014) *J. Magn. Reson. 241*, 3−17.

(8) Hoogsteen, K. (1959) *Acta Crystallogr. 12*, 822−823.

(9) Nikolova, E. N., Zhou, H., Gottardo, F. L., Alvey, H. S., Kimsey, I. J., and Al-Hashimi, H. M. (2013) *Biopolymers 99*, 955−968.

(10) Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) *Cell 87*, 1295−1306.

(11) Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodadek, T., and Burley, S. K. (1999) *Genes Dev. 13*, 3217−3230.

(12) Aishima, J., Gitti, R. K., Noah, J. E., Gan, H. H., Schlick, T., and Wolberger, C. (2002) *Nucleic Acids Res. 30*, 5244−5252.

(13) Kitayner, M., Rozenberg, H., Rohs, R., Suad, O., Rabinovich, D., Honig, B., and Shakked, Z. (2010) *Nat. Struct Mol. Biol. 17*, 423−429.

(14) Bohnuud, T., Beglov, D., Ngan, C. H., Zerbe, B., Hall, D. R., Brenke, R., Vajda, S., Frank-Kamenetskii, M. D., and Kozakov, D. (2012) *Nucleic Acids Res. 40*, 7644−7652.

(15) Natrajan, G., Lamers, M. H., Enzlin, J. H., Winterwerp, H. H., Perrakis, A., and Sixma, T. K. (2003) *Nucleic Acids Res. 31*, 4814−4821.

(16) Yang, H., Zhan, Y., Fenn, D., Chi, L. M., and Lam, S. L. (2008) *FEBS Lett. 582*, 1629−1633.

(17) Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L., and Aggarwal, A. K. (2004) *Nature 430*, 377−380.

(18) Makarova, A. V., and Kulbachinskiy, A. V. (2012) *Biochemistry (Moscow) 77*, 547−561.

(19) Krepl, M., Otyepka, M., Banas, P., and Sponer, J. (2013) *J. Phys. Chem. B 117*, 1872−1879.

(20) Siegfried, N. A., Metzger, S. L., and Bevilacqua, P. C. (2007) *Biochemistry 46*, 172−181.

(21) Horowitz, S., and Trievel, R. C. (2012) *J. Biol. Chem. 287*, 41576−41582.

(22) Cuesta-Seijo, J. A., and Sheldrick, G. M. (2005) *Acta Crystallogr. D61*, 442−448.

(23) Kow, Y. W. (2002) *Free Radical Biol. Med. 33*, 886−893.

(24) Hsu, G. W., Ober, M., Carell, T., and Beese, L. S. (2004) *Nature 431*, 217−221.

(25) Nordmann, P. L., Makris, J. C., and Reznikoff, W. S. (1988) *Mol. Gen. Genet. 214*, 62−67.

(26) Leonard, G. A., Booth, E. D., Hunter, W. N., and Brown, T. (1992) *Nucleic Acids Res. 20*, 4753−4759.