# Restricting the Conformational Heterogeneity of RNA by Specific Incorporation of 8-Bromoguanosine

**ARTICLE** *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · APRIL 2003

Impact Factor: 12.11 · DOI: 10.1021/ja029176m · Source: PubMed

J|A|C|S
COMMUNICATIONS

Published on Web 02/11/2003

# Restricting the Conformational Heterogeneity of RNA by Specific Incorporation of 8-Bromoguanosine

David J. Proctor,[†] Elzbieta Kierzek,[‡] Ryszard Kierzek,*[,‡] and Philip C. Bevilacqua*[,†]

*Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland*

Received October 31, 2002; E-mail: pcb@chem.psu.edu; rkierzek@rose.man.poznan.pl

RNA typically folds in a hierarchical fashion, forming independently stable secondary structure before tertiary structure.[1] The function of many RNAs depends on a compact tertiary structure, as exemplified by a number of small ribozymes.[2] Unfortunately, secondary structure is prone to alternative pairings, or misfolds, which hinder the formation of native tertiary structure.[3] Because these interactions are strong, misfolds can lead to kinetic trapping, complicating mechanistic and structural studies of RNA. Considerable effort has been put into correcting misfolds to produce fast-folding RNAs. Since misfolding occurs most frequently at the secondary structural level, site-directed mutagenesis and antisense oligonucleotides have provided simple approaches to promote native folding,[3] as have nucleotide analogues.[4] Additionally, proteins have been shown to facilitate RNA folding in vitro and in vivo,[5] and variation in pH, temperature, metal ion, and RNA concentration can reduce RNA conformational heterogeneity.[3] One example of secondary structural misfolding is the dimerization of hairpins to give a duplex with a symmetric internal loop (Figure 1A). Dimerization is especially problematic at high RNA and salt concentrations, such as those required for NMR and X-ray crystallographic studies.[6] Even the unusually stable UUCG tetraloop hairpin can form a duplex during crystallization.[7]

Here, we describe limiting the conformational heterogeneity of RNA using the nucleotide analogue 8-bromoguanosine (8BrG). Structural studies on nucleosides and polymers have shown that 8BrG preferentially adopts the *syn* conformation, wherein the nucleobase is positioned over the ribose sugar.[8] This conformation, which is in contrast to the *anti* conformation typical of A-form RNA helices, arises because the steric bulk of bromine precludes its residence over the ribose ring. We demonstrate that 8BrG shifts a hairpin−duplex equilibrium toward the hairpin conformation primarily by destabilizing the duplex conformation (Figure 1).

The 8BrG analogue was introduced into loop position 4 of selected YNMG hairpin tetraloops (Figure 1A). The YNMG motif is comprised of 16 thermodynamically stable sequences that adopt structures similar to the UUCG tetraloop.[6] The YNMG motif was chosen as a model system since a *syn* guanosine occurs naturally at position 4 of the loop, and inspection of YNMG structures suggests substitution should not result in a significant steric clash. In addition, the YNMG motif gives rise to several diagnostic NMR spectroscopic features, including an unusually upfield-shifted imino proton resonance that is due to a bifurcated hydrogen bond between positions 1 (Y = C or U) and 4 (G) of the loop, and an unusually downfield-shifted $^{31}$P resonance caused by a sharp turn in the backbone at G9.[6,9]

8-bromoguanosine was synthesized as a phosphoramidite and incorporated into 12mer RNA oligonucleotides (Table 1) using standard chemistry.[10] Thermodynamic characterization (Table 1) revealed that substitution at position 4 of the UUCG loop, UUC**G**, had little effect upon stability, with $\Delta\Delta G^\circ_{37} = -0.08 \pm 0.21$ kcal mol$^{-1}$ and $\Delta T_M = 1.2$ °C relative to UUCG. Likewise, 8BrG-substitution at position 4 of the CGCG loop, CGC**G**, did not have a significant effect, with $\Delta\Delta G^\circ_{37} = -0.41 \pm 0.58$ kcal mol$^{-1}$ and $\Delta T_M = 1.5$ °C relative to CGCG.

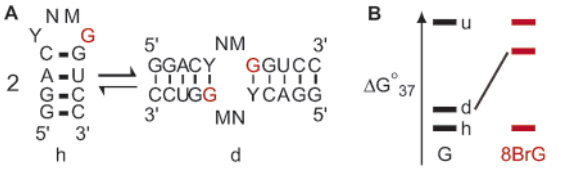


***Figure 1.*** (A) Equilibrium between hairpin and duplex conformations. A red G indicates 8BrG substitution; in the text this is indicated by a bold and underlined **G**. Y = C or U; N = A, C, G, or U; M = A or C. (B) Free energy diagram depicting the destabilization of the duplex conformation upon 8BrG substitution; h = hairpin; d = duplex; u = unfolded.

***Table 1.*** Thermodynamic Parameters for Hairpin Formation[10]

| sequence[a] | $\Delta G^\circ_{37}$ (kcal mol$^{-1}$)[b] | $T_M$ (°C)[c] |
|---|---|---|
| ggacUUCGgucc | −4.80 ± 0.13 | 71.7 |
| g**G**acUUCGgucc | −2.44 ± 0.12 | 57.2 |
| ggacUUC**G**gucc | −4.88 ± 0.17 | 72.9 |
| ggacCGCGgucc[d] | −3.60 ± 0.36 | 67.0 |
| ggacCGC**G**gucc | −4.01 ± 0.45 | 68.5 |

[a] The hairpin tetraloop is capitalized. [b] An extra significant figure is shown for $\Delta G^\circ_{37}$ to avoid round-off errors. [c] Maximum errors in $T_M$ are ∼1 °C. [d] Collected at 5−50 $\mu$M to favor the hairpin conformation.

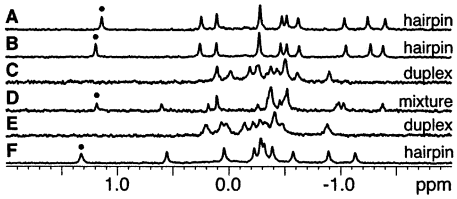


***Figure 2.*** $^1$H-decoupled $^{31}$P NMR spectra (202 MHz, 95% $H_2O$/5% $D_2O$ in 10 mM $NaH_2PO_4$/0.1 mM EDTA, pH 7.0, 45 °C) of (A) 0.4 mM UUCG, (B) 0.3 mM UUC**G**, (C) 0.4 mM CGCG, (D) 0.2 mM CGC**G**; and with 1M $Na^+$, (E) 0.2 mM CGAG, and (F) 0.5 mM CGA**G**.[13] The dominant conformation is given, and the downfield-shifted resonance diagnostic of the hairpin is indicated with a filled dot.

Structural characterization of unmodified UUCG by 1D $^1$H-decoupled $^{31}$P NMR spectroscopy revealed the expected 11 hairpin resonances dispersed over ∼2 ppm (Figure 2A). These include a resonance downfield-shifted to 1.15 ppm, which was previously assigned to G9P.[9] As expected on the basis of similar thermodynamic parameters (Table 1), the spectrum of 8BrG-substituted UUC**G** was nearly identical to UUCG (Figure 2B).

In contrast to UUCG, the $^{31}$P spectrum for CGCG comprised several densely packed resonances covering only ∼1 ppm (Figure 2C). One-dimensional $^1$H-decoupled $^{31}$P NMR spectra of A-form RNA helices have chemical shifts clustered near 0 ppm (referenced

[†] The Pennsylvania State University.
[‡] Polish Academy of Sciences.


10.1021/ja029176m CCC: $25.00 

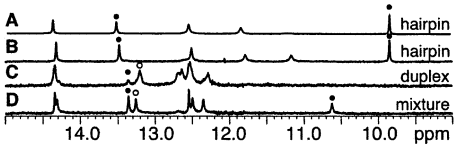


***Figure 3.*** Exchangeable $^1$H NMR spectra (600 MHz, 95% $H_2O$/5% $D_2O$ in 10 mM $NaH_2PO_4$/0.1 mM EDTA, pH 5.2, 1 °C) of (A) 1.1 mM UUCG, (B) 0.3 mM UUC**G**, (C) 0.2 mM CGCG, and (D) 0.2 mM CGC**G**. The dominant conformation is given, and selected resonances diagnostic of hairpin and duplex are indicated with filled and open dots, respectively.

to 85% $H_3PO_4$),[11] consistent with CGCG adopting the duplex conformation at NMR concentrations, as expected for this self-complementary sequence.[12] In contrast to CGCG, the $^{31}$P spectrum of CGC**G** was disperse, with a resonance at 1.2 ppm, consistent with a significant increase in the hairpin conformation upon 8BrG substitution (Figure 2D). In addition, in the presence of 1 M NaCl, known to favor the duplex conformation, a major shift in the equilibrium toward the hairpin conformation was observed for CGA**G** relative to CGAG (Figure 2, E and F).

One-dimensional $^1$H NMR spectroscopy of UUCG and UUC**G** revealed the expected upfield-shifted resonance for G9H1 at 9.8 ppm (Figure 3A, B).[6,9] In contrast, the spectrum of the CGCG tetraloop revealed primarily the duplex conformation, as illustrated by the relative intensities of hairpin and duplex resonances for G9H1 at 13.4 and 13.2 ppm, respectively,[14] and the absence of a resonance near 10 ppm (Figure 3C).[6] Strikingly, for CGC**G** the relative intensity of the resonances at 13.4 and 13.2 ppm inverts, with the hairpin conformation present at approximately twice the duplex concentration (Figure 3D). Moreover, the upfield-shifted resonance diagnostic of the tetraloop conformation is observed at 10.6 ppm (Figure 3D).[15] Overall, thermodynamic and NMR experiments on representative YNMG loops with 8BrG at loop position 4 are consistent with a *syn*-G, as expected from previous studies.[6]

Presumably the hairpin conformation of CGC**G** arises because of destabilization of the duplex. To estimate the extent of destabilization, 8BrG was substituted into the *stem* of a hairpin. The stability of g**G**acUUCGgucc was substantially decreased, with $\Delta\Delta G^\circ_{37} = +2.36 \pm 0.17$ kcal mol$^{-1}$ and $\Delta T_M = -14.5$ °C (Table 1). Since there are two potential C**G** base pairs in the duplex conformation of a CNMG sequence, the destabilization for the duplex is estimated at +4.7 kcal mol$^{-1}$. Thus, the shift in the equilibrium induced by 8BrG-substitution is considerable. These effects are depicted in the free energy diagram in Figure 1B.

We have demonstrated that through appropriate incorporation of the modified nucleoside 8BrG, the equilibrium for misfolding of a small RNA hairpin can be substantially altered. This effect appears to be due primarily to destabilization of the duplex state in which 8BrG cannot adopt the *anti* conformation. One application for this methodology is the study of salt effects on hairpins, which has been problematic due to dimerization. Also, more homogeneous secondary structure should favor the formation of native tertiary structure; thus, 8BrG-substituted YNMG hairpins could be engineered into large RNA molecules at positions where loop sequence is not critical in order to favor native structure. Since there are typically several such positions in large RNAs, this method should be generally useful for constraining conformation in catalytic, NMR, and X-ray crystallographic studies. In the latter case, the Br atom might also provide a useful heavy atom derivative.[16]

**Acknowledgment.** We thank Professor Doug Turner and his research group for use of their laboratory facilities. We also thank members of the Bevilacqua lab and Professors Chris Falzone, Barb Golden, and Juliette Lecomte for helpful comments. This work was supported by NSF Grant MCB-9984129; a Fellowship from the Alfred P. Sloan Foundation and a Camille Dreyfus Teacher-Scholar Award to P.C.B; and NIH Grant 1R03 TW1068 to R.K. and D. H. Turner.

**Supporting Information Available:** Protocols for the synthesis and characterization of 3′-*O*-phosphoramidite 5′-*O*-dimethoxytrityl-2′-*O*-tertbutyldimethylsilyl-2*N*-((dimethylamino)methylene)-8-bromoguanosine and RNA oligonucleotides (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

## References

(1) (a) Costa, M.; Michel, F. *EMBO J.* **1995**, *14*, 1276. (b) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol.* **1997**, *26*, 113. (c) Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271.
(2) (a) Scott, W. G.; Murray, J. B.; Arnold, J. R.; Stoddard, B. L.; Klug, A. *Science* **1996**, *274*, 2065. (b) Ferre-D'Amare, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567. (c) Rupert, P. B.; Massey, A. P.; Th Sigurdsson, S.; Ferre-D'Amare, A. R. *Science* **2002**, *298*, 1421.
(3) (a) Pan, J.; Woodson, S. A. *J. Mol. Biol.* **1998**, *280*, 597. (b) Treiber, D. K.; Williamson, J. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 309. (c) Chadalavada, D. M.; Knudsen, S. M.; Nakano, S.; Bevilacqua, P. C. *J. Mol. Biol.* **2000**, *301*, 349. (d) Chadalavada, D. M.; Senchak, S. E.; Bevilacqua, P. C. *J. Mol. Biol.* **2002**, *317*, 559.
(4) Olivas, W. M.; Maher, L. J., III. *Nucleic Acids Res.* **1995**, *23* 1936.
(5) For example (a) Tsuchihashi, Z.; Khosla, M.; Herschlag, D. *Science* **1993**, *262*, 99. (b) Mohr, S.; Stryker, J. M.; Lambowitz, A. M. *Cell* **2002**, *109*, 769. (c) Lorsch, J. R. *Cell* **2002**, *109*, 797.
(6) Proctor, D. J.; Schaak, J. E.; Bevilacqua, J. M.; Falzone, C. J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 12062.
(7) Holbrook, S. R.; Cheong, C.; Tinoco, I., Jr.; Kim, S. H. *Nature* **1991**, *353*, 579.
(8) (a) Tavale, S. S.; Sobell, H. M. *J. Mol. Biol.* **1970**, *48*, 109. (b) Michelson, A. M.; Monny, C.; Kapuler, A. M. *Biochim. Biophys. Acta* **1970**, *217*, 7. (c) Ikehara, M.; Uesugi, S.; Yoshida, K. *Biochemistry* **1972**, *11*, 830.
(9) Varani, G.; Cheong, C.; Tinoco, I., Jr. *Biochemistry* **1991**, *30*, 3280.
(10) See Supporting Information for experimental details.
(11) Gorenstein, D. In *$^{31}$P NMR, Principles and Applications*; Academic Press: New York, 1984.
(12) Duplex formation is also supported by thermodynamic data, where the $T_M$ of CGCG is concentration dependent between ≈50−200 $\mu$M strand concentration.
(13) Spectra were obtained at the maximum sample concentrations available. Slight differences in concentration occurred in some instances. For CGCG, the duplex conformation was also dominant at 0.2 mM, as seen in Figure 3C. CGCG was not chosen at 1 M $Na^+$ since it is perfectly complementary, and its shift in equilibrium is not in the detectable range for NMR.
(14) In general, for a given imino proton in a small RNA hairpin the resonance for the hairpin conformation is slightly downfield of that for the duplex conformation. See ref 6 for further examples.
(15) The precise value of the chemical shift has been found to be slightly larger when position 1 of the loop is a C. See ref 6 for details.
(16) Correll, C. C.; Freeborn, B.; Moore, P. B.; Steitz, T. A. *J. Biomol. Struct. Dyn.* **1997**, *15*, 165.

JA029176M