BIOCHEMISTRY
including biophysical chemistry & molecular biology

Rapid Report

pubs.acs.org/biochemistry

The Conformationally Constrained *N*-Methanocarba-dT Analogue Adopts an Unexpected C4′-*exo* Sugar Pucker in the Structure of a DNA Hairpin

Pradeep S. Pallan,[†] Victor E. Marquez,[‡] and Martin Egli*,[†]

[†]Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville, Tennessee 37232, United States

[‡]Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, Maryland 21702, United States

Supporting Information

ABSTRACT: Incorporation of a bicyclo[3.1.0]hexane scaffold into the nucleoside sugar was devised to lock the embedded cyclopentane ring in conformations that mimic the furanose *North* and *South* sugar puckers. To analyze the effects of *North*-methanocarba-2′-deoxythymidine (*N*-MCdT) on the B-form DNA, we crystallized d(CGCGAA[mcTmcT]CGCG) with two *N*-MCdTs. Instead of a duplex, the 12mer forms a tetraloop hairpin, whereby loop *N*-MCdTs adopt the C4′-*exo* pucker (*NE*; *P* = 50°). Thus, the bicyclic framework does not limit the pucker to the anticipated C2′-*exo* range (*NNW*; *P* = −18°).

Conformationally restricted or locked nucleoside analogues are being investigated in biochemistry, biotechnology, and medicinal chemistry and are of considerable interest in nucleoside-based drug discovery and, incorporated into oligonucleotides, for use as antisense agents, siRNAs, and aptamers.[1−8] There are multiple strategies for constraining the furanose sugar in the *North* conformation [i.e., C3′-*endo* (Figure 1)]. 2′-Carbohydrate modifications,[9,10] including locked nucleic acid (LNA),[11] and analogues with the 3′-oxygen replaced by a less electronegative moiety, i.e., nitrogen as in N3′ → P5′ phosphoramidate DNA,[12,13] all limit the sugar pucker to the northern range to various degrees. A large number of such analogues have been examined over the past 20 years because of their increased RNA affinity as a result of preorganizing the modified strand for the A-type conformation.[14,15] Although conformational preoganization is often assumed to be synonymous with an entropically favorable effect, the higher RNA affinity typically observed for analogues with *N*-sugars has been shown in some cases to be due to enthalpic gains.[16,17]

Compared with chemical modifications that constrain the furanose in the *N* range, those maintaining a *Southern* (*S*) pucker [i.e., C2′-*endo* (Figure 1)] are much less common. Remarkably, the bicyclo[3.1.0]hexane scaffold that contains a cyclopentane ring provides a surrogate of the five-membered furanose ring in nucleosides, whereby the cyclopentane is locked in an envelope conformation (Figure 1).[1] The location of the fused cyclopropane ring then determines whether the pseudoboat conformation adopted by bicyclo[3.1.0]hexane mimics *N* [C2′-*exo*, C4′-C6′-C7′ (Figure 1, red)] or *S* [C3′-*exo*, C1′-C6′-C7′ (Figure 1, blue)] puckers of furanose.

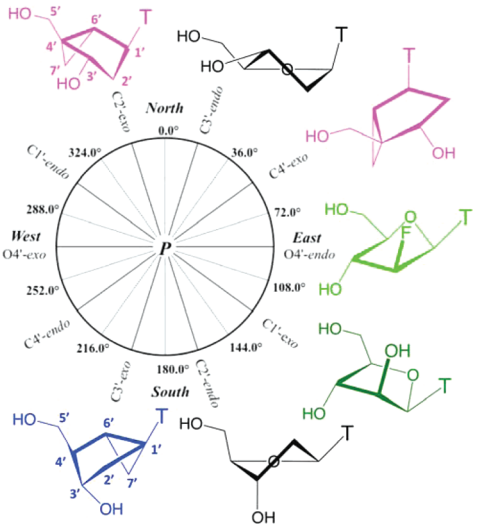

Figure 1. Pseudorotation cycle of the five-membered sugar ring in nucleosides with phase angles indicated in multiples of 36°.[18] Selected pucker modes are labeled and illustrated with representative examples of sugars adopting a particular conformation: C2′-*exo* *North*-bicyclo-[3.1.0]hexane (*N*-MCdT, magenta), C3′-*endo* 2′-deoxyribose (A-form duplex, similar for ribose, black), C4′-*exo* MCdT (this work, magenta), O4′-*endo* 2′-fluoro-2′-deoxyarabinose (FANA, light green), C1′-*exo* arabinose (ANA, green), C2′-*endo* 2′-deoxyribose (B-form duplex, black), and C3′-*exo* *South*-bicyclo[3.1.0]hexane (*S*-MCdT, blue).

The corresponding *N*- and *S*-2′-deoxy-methanocarba-nucleosides (*N*- and *S*-MCdNs, respectively) or nucleotides have been tested for their antiviral activities and in SAR studies involving kinases, polymerases, and reverse transcriptases.[1,19−22] The effects of such nucleotides incorporated into DNA oligonucleotides[23] on the pairing stability and duplex conformation (i.e., bending) have also been reported.[24−26] Despite the extensive evaluation of methanocarba-nucleotides and -modified oligonucleotides, experimental, atomic-resolution models of *N*- or *S*-MCdNs in the oligonucleotide context and their effects on DNA conformation are presently lacking.

Received: February 16, 2012
Revised: March 9, 2012
Published: March 12, 2012

dx.doi.org/10.1021/bi300215k | *Biochemistry* 2012, 51, 2639−2641

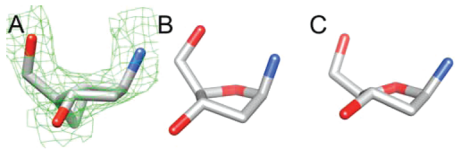

Figure 2. (A) Final Fourier ($2F_o - F_c$) sum electron density (1σ threshold) around the sugar of *N*-MCdT7 (C4′-*exo*; P = 50°). For comparison, 2′-deoxyriboses adopting (B) C2′-*exo* (P = −18°) and (C) C3′-*endo* (P = 18°) puckers are shown.

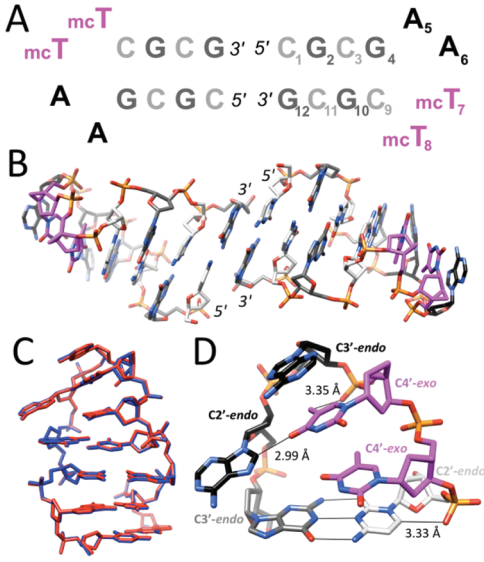

Figure 3. Hairpin structure of the *N*-MCdT-modified DDD. (A) DNA sequence, nucleotide numbering, and head-to-tail arrangement of left-handed $[d(CGCG)]_2$ stems. (B) Two independent hairpin molecules with residue coloring matching that in panel A. (C) Superimposition of the two independent hairpins reveals their nearly identical conformations. (D) Close-up of the AAmcTmcT loop.

To establish the preferred pucker of *N*-2′-deoxy-methanocarba nucleotides, we synthesized[23] and crystallized the Dickerson−Drew dodecamer (DDD), d(CGCGAAmcTmcTCGCG), with two *N*-MCdTs. The structure was determined by single-wavelength anomalous dispersion using two Ba^{2+} sites and refined to 1.8 Å resolution (see the Supporting Information for details).

An example of the quality of the final electron density is shown in Figure 2. The volume of the crystallographic unit cell was consistent with a single duplex per asymmetric unit. However, to our surprise, the experimental electron density revealed two independent single-stranded dodecamers, each folded into a hairpin with a left-handed $[d(CGCG)]_2$ stem and an AAmcTmcT loop (Figure 3). To our knowledge, this is the only stem−loop structure for the DDD reported to date.

Z-DNA is preferentially formed by alternating CG sequences at high salt concentrations, whereby G and C exhibit the typical C3′-*endo*/*syn* and C2′-*endo*/*anti* conformations, respectively.[27] At the transition between stems from adjacent hairpins oriented in a head-to-tail fashion, the virtual absence of a twist and a large slide between C:G pairs result in stacked guanines (Figure 3B). Both hairpins adopt virtually identical structures, and AAmcTmcT loops can be neatly overlaid (Figure 3C). The loop arrangement is stabilized by a partial stacking interaction between G4 and A5 and a continuous A6, mcT7, mcT8, and C9 stack (Figure 3D). Other stabilizing influences include C−H···O hydrogen bonds between C8H [A5] and O4 [T7] as well as C5M (methyl) and O2P (T7, intranucleotide). Most noteworthy in regard to our original goal of establishing the preferred conformation of the *N*-MCdT sugar is the C4′-*exo* conformation of both mcT7 and mcT8 (Figures 2A and 3D). The pseudorotation phase angles as calculated with PROSIT[28] are 50.1° (mcT7) and 52.5° (mcT8). The corresponding angles in the second hairpin are 48.5° and 51.5°, respectively. Therefore, *N*-methanocarba-nucleotides adopt a *Northeast* conformation in our structure (Figure 1).

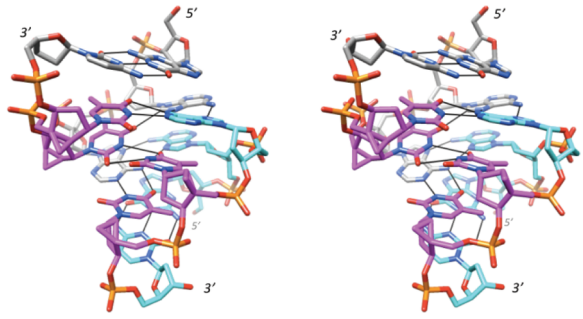

Figure 4. Cross-eyed stereoview of the interactions between hairpin loops in the crystal structure of the *N*-MCdT-modified DDD. Carbon atoms of residues G4, A5, A6, and C9 in hairpin 1 are colored cyan, and in hairpin 2, the corresponding atoms are colored gray. Hydrogen bonds are shown as thin lines, and water molecules were omitted.

In the crystal hairpins form infinite columns, stems are stacked at one end of each hairpin (Figure 3B), and at the other, loops are interacting under formation of three base pairs (Figure 4). Nucleotides mcT7 from joined hairpins form a sheared T:T# pair in the center (# marks a nucleotide from a second hairpin), stabilized by two N3−H···O2 hydrogen bonds. This pair is flanked by Watson−Crick type pairs between A6# and mcT8, and A5 engages in a Hoogsteen pair with the former via N6−H···O2 hydrogen bonds, thus resulting in base triples.

A search for hairpin structures that would allow a comparison with the DDD hairpin studied here did not yield cases of DNA AATT[29] or RNA AAUU[30] tetraloop sequences. Many years ago, Dickerson and co-workers determined the crystal structure of the DNA hexadecamer d(CGCGCGTTTTCGCGCG).[31] The oligonucleotide forms a hairpin with a left-handed $[d(CGCGCG)]_2$ stem and a dT_4 loop. However, the arrangement of thymidines in their structure is completely different from that of the AAmcTmcT loop observed here and thus does not allow a meaningful comparison or conclusions regarding the preferred sugar pucker of *N*-MCdT residues.

The DDD with two incorporated *N*-MCdT nucleotides crystallized as a hairpin, which is highly unusual. At the concentrations between 0.1 and 1 mM typically used for crystallization of an oligonucleotide, formation of a bimolecular species (i.e., the duplex) is the rule. Crystallization experiments with RNA hairpins featuring tetrameric stems and tetraloops invariably lead to duplex formation, with or without mismatch base pairs (see, for example, ref 32). Only unusually stable constructs with short stems, i.e., those capped with *trans*-stilbene,[33] crystallize as hairpins. Because the structure of the chemically modified DDD here represents the only case of such a dodecamer adopting a stem−loop conformation in the crystal, the question about the origins of this behavior arises.

Even DDDs containing 2′-modified T or U residues at positions 7 and 8, i.e., 2′-SMe-U,[34] that display a C3′-*endo* pucker crystallized as duplexes with enlarged minor grooves

dx.doi.org/10.1021/bi300215k | *Biochemistry* 2012, 51, 2639−2641

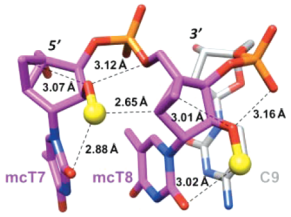

Figure 5. Shortest distances involving 2′-*O*-methyl groups (methyl carbons highlighted as yellow spheres) modeled onto the crystal structure of the *N*-MCdT-modified DDD hairpin. Only a portion of the loop, mcT7/8, and the adjacent stem, C9, is shown.

and minimal bending, but not as hairpins. Moreover, neither NMR nor CD experiments provided any indication of the presence of a hairpin in solution.[24,25] Addition of methoxy groups at C2′ of *N*-MCdT7 and -T8 results in several short distances between 2′-oxygen and methyl carbon and backbone and base atoms (Figure 5). Therefore, it is unlikely that a DDD with rU(T) or 2′-*O*Me-rU(T) nucleotides at positions 7 and 8 would adopt the hairpin conformation observed here for the *N*-MCdT-modified dodecamer. Interestingly, the chimeric DDD, 5′-fCfGfCfGaAaAaUaUfCfGfCfG-3′ [fN = FANA, aN = ANA (Figure 1)], was found to exist in an equilibrium between the duplex and a hairpin with an aAaAaUaU loop.[35]

The sugar puckers of *N*-MCdTs in the crystal structure differ by 70° in terms of phase angle from the value predicted for[1] and typically associated with this modified nucleotide (Figure 1).[24−26] At the level of the nucleoside, adoption of the C4′-*exo* pucker likely results in a higher potential energy compared with that of the C2′-*exo* pucker (see the Supporting Information for details). Although the structure leaves us without a model of the conformation of the *N*-MCdT sugar inside a DNA duplex, we have to conclude that the adoption of the hairpin by the modified DDD is linked to the presence of methanocarba residues. At the very least, we have identified an alternative, stable *NE* conformation for the bicyclic nucleoside analogue that is allegedly constrained to the *N* pucker region. Alternatively, the observed sugar pucker represents its actual preference in an oligonucleotide context, and the nucleoside should be termed *NE*-MCdT instead of *N*-MCdT.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Crystallization and structure analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The dodecamer is Protein Data Bank entry 4DKZ.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (615) 343-8070. Fax: (615) 322-7122. E-mail: martin.egli@vanderbilt.edu.

Funding

Supported by NIH Grant GM55237 (to M.E.). Use of the APS was supported by the U.S. DOE, Office of Science, Office of Basic Energy Sciences, Contract DE-AC02-06CH11357.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Marquez, V. E. (2008) in *Modified Nucleosides in Biochemistry, Biotechnology and Medicine* (Herdewijn, P., Ed.) pp 307−341, Wiley-VCH, Weinheim, Germany.
(2) Mathé, C., and Périgaud, C. (2008) *Eur. J. Org. Chem.*, 1489−1505.
(3) Crooke, S. T. (2007) in *Antisense Drug Technology: Principles, Strategies, and Applications*, 2nd ed., CRC Press, Boca Raton, FL.
(4) Bennett, C. F., and Swayze, E. E. (2010) *Annu. Rev. Pharmacol. Toxicol. 50*, 259−293.
(5) Rakesh, N., and Wengel, J. (2010) *Chem. Biodiversity 7*, 536−542.
(6) Herdewijn, P. (2010) *Chem. Biodiversity 7*, 1−59.
(7) Leumann, C. J. (2005) *Chimia 59*, 776−779.
(8) Trujillo, C. A., Nery, A. A., Alves, J. M., Martins, A. H., and Ulrich, H. (2007) *Clin. Ophthalmol. 1*, 393−402.
(9) Manoharan, M. (1999) *Biochim. Biophys. Acta 1489*, 117−130.
(10) Manoharan, M., Akinc, A., Pandey, R. K., Qin, J., Hadwiger, P., John, M., Mills, K., Charisse, K., Maier, M. A., Nechev, L., Greene, E. M., Pallan, P. S., Rozners, E., Rajeev, K. G., and Egli, M. (2011) *Angew. Chem., Int. Ed. 50*, 2284−2288.
(11) Wengel, J. (1999) *Acc. Chem. Res. 32*, 301−310.
(12) Tereshko, V., Gryaznov, S., and Egli, M. (1998) *J. Am. Chem. Soc. 120*, 269−283.
(13) Gryaznov, S. (2010) *Chem. Biodiversity 9*, 477−493.
(14) Prakash, T. P. (2011) *Chem. Biodiversity 8*, 1616−1641.
(15) Egli, M., and Pallan, P. S. (2010) *Chem. Biodiversity 7*, 60−89.
(16) Pallan, P. S., Greene, E., Jicman, P., Pandey, R., Manoharan, M., Rozners, E., and Egli, M. (2011) *Nucleic Acids Res. 39*, 3482−3495.
(17) Kaur, H., Wengel, J., and Maiti, S. (2008) *Biochemistry 47*, 1218−1227.
(18) Altona, C., and Sundaralingam, M. (1972) *J. Am. Chem. Soc. 94*, 8205−8212.
(19) Ezzitouni, A., and Marquez, V. E. (1997) *J. Chem. Soc., Perkin Trans. 1*, 1073−1078.
(20) Marquez, V. E., Ben-Kasus, T., Barchi, J. J. Jr., Green, K. M., Nicklaus, M. C., and Agbaria, R. (2004) *J. Am. Chem. Soc. 126*, 543−549.
(21) Boyer, P. L., Julias, J. G., Marquez, V. E., and Hughes, S. H. (2005) *J. Mol. Biol. 345*, 441−450.
(22) Eoff., R. L., McGrath, C. E., Maddukuri, L., Salamanca-Pinzón, S. G., Marquez, V. E., Marnett, L. J., Guengerich, F. P., and Egli, M. (2010) *Angew. Chem., Int. Ed. 49*, 7481−7485.
(23) Maier, M. A., Choi, Y., Gaus, H., Barchi, J. J. Jr., Marquez, V. E., and Manoharan, M. (2004) *Nucleic Acids Res. 32*, 3642−3650.
(24) Wu, Z., Maderia, M., Barchi, J. J. Jr., Marquez, V. E., and Bax, A. (2005) *Proc. Natl. Acad. Sci. U.S.A. 102*, 24−28.
(25) Maderia, M., Shenoy, S., Van, Q. N., Marquez, V. E., and Barchi, J. J. Jr. (2007) *Nucleic Acids Res. 35*, 1978−1991.
(26) Macias, A. T., Banavali, N. K., and MacKerell, A. D. Jr. (2007) *Biopolymers 85*, 438−449.
(27) Gessner, R. V., Quigley, G. J., and Egli, M. (1994) *J. Mol. Biol. 236*, 1154−1168.
(28) Sun, G., Voigt, J. H., Marquez, V. E., and Nicklaus, M. C. (2005) *Nucleosides, Nucleotides Nucleic Acids 24*, 1029−1032.
(29) Goodsell, D. S. (2010) *Nat. Educ. 3*, 39.
(30) Vanegas, P. L., Hudson, G. A., Davis, A. R., Kelly, S. C., Kirkpatrick, C. C., and Znosko, B. M. (2012) *Nucleic Acids Res. 40*, 439−444.
(31) Chattopadhyaya, R., Ikuta, S., Grzeskowiak, K., and Dickerson, R. E. (1988) *Nature 334*, 175−179.
(32) Holbrook, S. R., Cheong, C., Tinoco, I. Jr., and Kim, S. H. (1991) *Nature 353*, 579−581.
(33) Egli, M., Tereshko, V., Murshudov, G. N., Sanishvili, R., Liu, X., and Lewis, F. D. (2003) *J. Am. Chem. Soc. 125*, 10842−10849.
(34) Pallan, P. S., Prakash, T. P., Li, F., Eoff, R. L., Manoharan, M., and Egli, M. (2009) *Chem. Commun., 15*, 2017−2019.
(35) Martín-Pintado, N., Yahyaee Anzahaeeb, M., Campos, R., Damha, M. J., and González, C. (2011) 94th Canadian Chemistry Conference, June 5−9.