

A conformational transition in the structure of a 2'-thiomethyl-modified DNA visualized at high resolution†

Pradeep S. Pallan,^a Thazha P. Prakash,^b Feng Li,^a Robert L. Eoff,^a Muthiah Manoharan^c and Martin Egli^{*a}

Received (in Cambridge, UK) 18th December 2008, Accepted 30th January 2009

First published as an Advance Article on the web 26th February 2009

DOI: 10.1039/b822781k

Crystal structures of A-form and B-form DNA duplexes containing 2'-S-methyl-uridines reveal that the modified residues adopt a RNA-like C3'-endo pucker, illustrating that the replacement of electronegative oxygen at the 2'-carbon of RNA by sulfur does not appear to fundamentally alter the conformational preference of the sugar in the oligonucleotide context and sterics trump stereoelectronics.

The conformation of nucleosides is subject to various stereoelectronic and steric effects.¹ In the case of the 2'-deoxy-ribose moiety the Southern and Northern regions in the pseudorotational cycle represent preferred ring pucker.^{2,3} Thus, conformations corresponding to the C2'-endo and C3'-endo phase angle (*P*) ranges adjacent to the South and North poles, respectively, give rise to the canonical B- and A-DNA duplex forms at the oligodeoxynucleotide level.⁴ By comparison, the presence of a hydroxyl group at the ribose 2'-position shifts the conformational equilibrium to the Northern half at the mononucleoside level and locks the RNA duplex in the A-form. Using molecular dynamics (MD) simulations, Ferguson and co-workers proposed that the sugar moieties in an oligo-2'-deoxy-2'-S-methyl-ribonucleotide (2'-SMe-RNA) paired to RNA preferentially adopt the C2'-endo conformation.⁵ This is in contrast to the 2'-deoxy-2'-O-methyl RNA analog (2'-OMe-RNA) paired to RNA, in that its sugars are limited to the C3'-endo range by the electronegative substituent.⁵ The latter observation is supported by structural data.^{6,7} Moreover, the conformational preorganization of the 2'-OMe-RNA strand for the RNA target is consistent with the increased thermodynamic stability of 2'-OMe-RNA–RNA duplexes compared to the corresponding 2'-SMe-RNA–RNA duplexes (ref. 5 and cited references). Although sulfur is less electronegative than oxygen, the switch from the North to the South conformation of the ribose in the oligonucleotide context as a result of replacing oxygen with sulfur at the 2'-position is surprising.

The conformational preorganization of the sugar furnished by electronegative and/or bulky 2'-substituents has been widely

used in oligonucleotides against RNA targets in antisense and RNAi applications.^{8–10} For example, the 2'-deoxy-2'-fluoro-ribo- (2'-F-RNA¹¹) and 2'-deoxy-2'-[(2-methoxy)ethyl]-ribonucleic acid modifications (2'-MOE-RNA¹²) exhibit a strong preference for the C3'-endo sugar conformation and their duplexes with RNA are restricted to the A-form. Conversely, the 2'-deoxy-2'-fluoroarabinonucleic acid modification (2'-FANA) shows a preference for the C2'-endo pucker at the level of the mono-nucleoside.¹³ But in duplexes, the 2'-FANA sugar has been observed to preferably adopt the O4'-endo (East; in B-form DNA or 2'-FANA–RNA hybrids)^{14–16} or even the C3'-endo or C4'-exo conformations (Northern; in A-form DNA).¹⁷

Individual nucleosides can adopt a range of conformations within duplexes without significantly affecting the overall geometry of the latter (*i.e.* O4'-endo in B-DNA).¹⁴ But neighboring residues whose *P* values differ more drastically (*i.e.* > 120°) can cause deformations of the duplex that include kinking and groove widening/narrowing.^{17,18} Therefore, it is evident that the conformational preferences of nucleosides affect the geometry and stability of oligonucleotide duplexes and, in turn, that the duplex geometry can have an effect on the conformation of individual nucleosides. However, in reality it is often not straightforward to predict the conformational properties of nucleosides, or to model the geometric features of duplexes containing a mix of residues that have deviating sugar pucker. In addition to the intrinsic sequence-dependent, conformational preferences of oligonucleotides, interactions with cations,^{19,20} water^{21,22} and proteins^{4,23,24} impinge on their geometry.

In order to analyze the conformational properties of 2'-SMe-RNA and to establish the conformational boundaries of 2'-thiomethyl-ribonucleosides, we determined crystal structures of two DNA duplexes with incorporated 2'-SMe-rU residues. We chose two DNA oligomers to study the conformational properties of 2'-SMe-RNA, the decamer d(GCGTATACGC) and the dodecamer d(CGCGAATTCGCG) (the Dickerson–Drew dodecamer or DDD). The 10mer normally adopts an A-form conformation when at least one nucleoside is replaced by a 2'-modified residue²⁵ and the DDD constitutes a classic example of a B-form DNA.^{22,26} Here, we report crystal structures for duplexes of sequence GCGTAU*ACGC and CGCGAAU*U*CGCG containing 2'-SMe-rU (U*) residues at 1.04 and 1.25 Å resolution, respectively (see ESI†).‡ Examples of the quality of the final electron density are shown in Fig. 1 and final refinement parameters are listed in Table S1 (see ESI†).

The 10mer duplex adopts a regular A-form conformation with all residues including U*6 and U*16 (nucleosides are

^a Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232, USA.

E-mail: martin.egli@vanderbilt.edu; Fax: +1 615 322 7122

^b Department of Medicinal Chemistry, ISIS Pharmaceuticals Inc., Carlsbad, CA 92008, USA

^c Department of Drug Discovery, Alnylam Pharmaceuticals Inc., Cambridge, MA 02142, USA

† Electronic supplementary information (ESI) available: Oligonucleotide synthesis,²⁷ crystallization and data collection. Crystal data and refinement parameters (Table S1), illustrations of the 10- and 12-mer duplexes. See DOI: 10.1039/b822781k

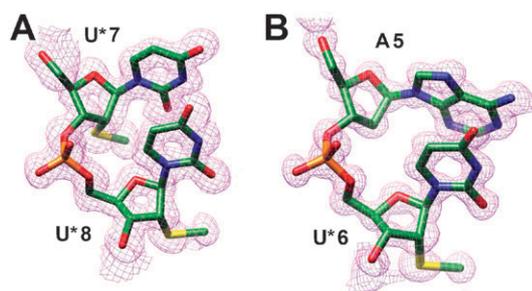


Fig. 1 Quality of the final structures. Fourier ($2F_o - F_c$) sum electron densities around the (A) U*7pU*8 step in the modified dodecamer and the (B) A5pU*6 step in the modified decamer.

numbered 1–10 in strand 1 and 11–20 in strand 2) exhibiting a C3'-endo pucker (Fig. S1, ESI†). Although it is possible that the 2'-SMe-ribofuranose is driven towards North as a result of its surroundings, this observation provides an indication that sulfur in place of oxygen (either in the form of 2'-OH or 2'-OMe) does not appear to fundamentally affect the conformational preferences of the sugar. The situation of U* embedded in a more or less canonical A-DNA duplex likely does not emulate the constraints put on a chimeric 2'-SMe-RNA/DNA or all-2'-SMe-RNA strands opposite RNA. However, it is noteworthy that 2'-arabino-T nucleosides displayed a C1'-exo (South-Eastern) pucker inside the same A-form duplex,¹⁷ indicating that the overall geometry of the 10mer does not necessarily limit the sugar conformation to North.

Unlike the A-form duplex that shows a regular overall geometry, the [CGCGAAU*U*CGCG]₂ duplex with U* residues replacing dT at positions 7 and 8 of the first and positions 18 and 19 of the second strand (nucleosides are numbered 1–12 in strand 1 and 13–24 in strand 2) undergoes a significant expansion in the central minor groove compared to the native 12mer (Fig. 2). At the ApU* step (underlined in the above sequence), the roll amounts to -20° (-1° ; values at ApT in the native DDD are in parentheses), the rise is 4.0 Å (3.34 Å), the twist is 41.5° (33°) and the slide is 0.44 Å (-0.30 Å). These distortions are a direct consequence of the C3'-endo sugar pucker of all four U* residues. In addition to the Northern sugar conformation of these residues, the deoxyriboses of 3'-adjacent C9 in strand 1 and C21 in strand 2 also adopt the C3'-endo pucker. Both cytidines display a C2'-endo pucker in the structure of the native DDD,²² indicating that the change is brought about by their vicinity to the U* residues. All adenosines display a C2'-endo pucker and the transition from A (South) to U* (North) is accompanied by a subtle change in the backbone conformations of U*7 and U*19. The β and γ torsion angles of both residues flip from the standard $ap/+sc$ to the $-ac/ap$ ranges.

Compared to the structural changes observed in the central tetramer, the outer G-tract tetramers show helical geometries that are similar to those found in the native DDD (Fig. S2, ESI†). Thus, the strong roll observed at the ApU* step does not lead to a kink of the duplex as indicated by the relatively straight global helix axis (Fig. 2(B)). Moreover, there are no significant geometrical deviations between the two independent duplexes per crystallographic asymmetric unit. As a result of the wider minor groove there are no short contacts between 2'-SMe-substituents from opposite strands in the modified

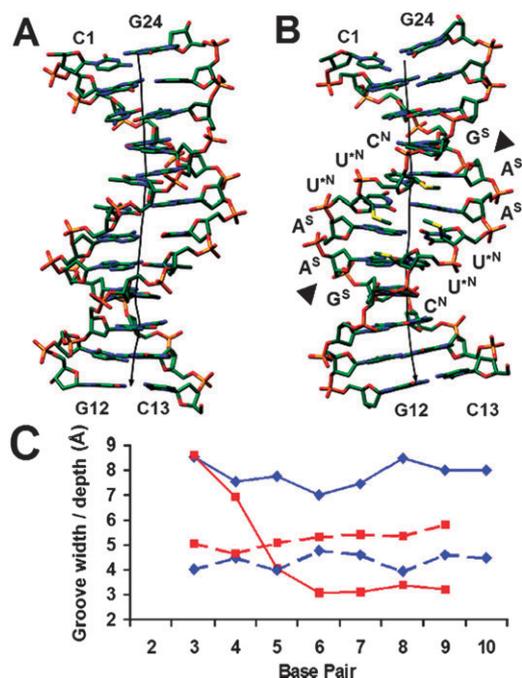


Fig. 2 (A) The native DDD and (B) the 2'-SMe-modified DDD viewed into the minor groove. Sulfur atoms in the modified duplex are highlighted in yellow, the helical axes are depicted as solid black lines, terminal and the six central residues are labeled (with the symbol in superscript next to the letter indicating the sugar pucker), and EcoRI cutting sites are marked by a triangle. (C) Minor groove width (solid line) and depth (dashed line) in the native²² (red) and modified (blue) DDDs.

DDD duplex. A close-up view into the central minor groove reveals intra- and inter-strand van der Waals interactions between 2'-SMe moieties and an interrupted minor groove hydration spine (Fig. 3).

Assuming that 2'-SMe-purines and -pyrimidines display similar conformational preferences, the finding that 2'-SMe-uridines adopt a C3'-endo (North) pucker inside both A- and B-form duplexes argues strongly against the earlier conclusion based on MD and *ab initio* calculations⁵ that the sugars of an oligo-2'-SMe-RNA strand paired to RNA, unlike oligo-2'-OMe-RNA,

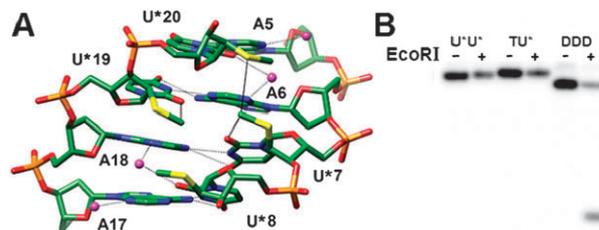


Fig. 3 (A) The U*-modified DDD viewed into the central minor groove. van der Waals interactions involving 2'-SMe-substituents and hydrogen bonds are indicated by solid and dashed lines, respectively, and water molecules are magenta spheres. Note the strong roll of uracils at the central base-pair step and the altered tip of the two inner compared with the two outer A-U* pairs. (B) PAGE assay of EcoRI cleavage experiments with the native DDD and modified dodecamers with either one (TU*) or both Ts (U*U*) replaced by 2'-SMe-U (from right to left). The - and + signs indicate absence and presence of the restriction enzyme, respectively.

prefer a C2'-endo (South) conformation. The 2'-substituents of U* residues in our crystal structures are pointing into the minor groove (Fig. 3(A)). Based on the conformational preferences predicted earlier, one would expect the substituents to lodge inside the major groove. The latter orientation would not interfere with RNase H, an endonuclease that cleaves the RNA portion of RNA–DNA hybrids,²⁷ binding the latter from the minor groove side. However, it was recently reported that 2'-SMe-RNA–RNA hybrids do not elicit RNase H,²⁷ consistent with our above structural observations [changes in the minor groove width of hybrid duplexes are also known to affect RNase H activity (ref. 33 and cited references)]. That the DDD duplex undergoes a drastic conformational change as a result of the substitution of T residues by U* is also demonstrated by assays with a restriction enzyme (Fig. 3(B)). EcoRI cleaves between G and A (underlined) in the recognition sequence 5'-GAATTC-3'/3'-CTTAAG-5' (contained in the DDD, Fig. 2(B)). Accordingly, the DDD with either one T or both Ts replaced by U* (labeled as TU* and U*U*, respectively) is not cleaved by EcoRI, an effect that must be largely due to conformational changes as a result of the 2'-SMe substituents and not the absence of the 5-methyl group in U compared with T.³⁷

Support by the National Institutes of Health is gratefully acknowledged (Grant R01 GM55237). We thank Dr. Z. Wawrzak for help with X-ray diffraction data collection. Vanderbilt University is a member institution of the LS-CAT at the Advanced Photon Source, Argonne, IL. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

Notes and references

‡ *Structure solution and refinement*: the structure of the 10mer was determined by the molecular replacement (MR) technique using an A-form DNA (PDB ID 411D) as the search model.^{28,29} All attempts to phase the 12mer data by the same approach failed, as did crystallization experiments for subsequent heavy atom phasing with 12mers containing either Br⁵C in place of C or, either one of the two, or both U* residues replaced by 2'-SeMe-U.^{30–32} The structure was ultimately determined by co-crystallizing the U*-modified DDD with *Bacillus halodurans* RNase H³³ and then phasing the structure of the complex by MR, using the coordinates of the enzyme alone as the search model. The coordinates of the 12mer obtained from the complex structure then served as the model to determine the crystal structure of the duplex alone by MR. Initial crystallographic and B-factor refinements of the 10- and 12mer structures were carried out with the program CNS,³⁴ setting aside 10% (5% for the 10mer) randomly chosen reflections for calculating the *R*-free. After the addition of water molecules and metal ions as well as incorporation of the U* residues and adaptation of the dictionary files, refinements were continued with the program SHELX,³⁵ using anisotropic temperature factors for all nucleic acid and solvent atoms. Final refinement parameters for the two duplex structures as well as the RNase H complex are listed in Table S1 (ESI†). All helical parameters were calculated with the program CURVES.³⁶

EcoRI cleavage assay: 500 nM DNA labeled with ³²P were mixed with 2 μL of 10X NEB EcoRI buffer (New England Biolabs), 2 μL of EcoRI (20K units mL⁻¹; New England Biolabs) and diluted to a reaction volume of 20 μL to obtain a final buffer concentration of 50 mM NaCl, 100 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, and 0.025% Triton X-100. The samples were incubated at 37 °C for 90 min, followed by PAGE.

Data deposition: final coordinates and structure factor files have been deposited in the Brookhaven Protein Data Bank (<http://www.rcsb.org>). The entry codes are 3EY1 (RNase H–dodecamer complex), 3EY2 (decamer) and 3EY3 (dodecamer).

- C. Thibaudeau, P. Acharya and J. Chattopadhyaya, *Stereoelectronic Effects in Nucleosides and Nucleotides and Their Structural Implications*, Uppsala University Press, Uppsala, SE, 1999.
- C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205–8212.
- G. Sun, J. H. Voigt, I. V. Fillippov, V. E. Marquez and M. C. Nicklaus, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1752–1762.
- S. Neidle, *Principles of Nucleic Acid Structure*, Academic Press, London, UK, 2008.
- D. Venkateswarlu, K. E. Lind, V. Mohan, M. Manoharan and D. M. Ferguson, *Nucleic Acids Res.*, 1999, **27**, 2189–2195.
- P. Lubini, W. Zürcher and M. Egli, *Chem. Biol.*, 1994, **1**, 39–45.
- D. A. Adamiak, J. Milecki, M. Popena, R. W. Adamiak, Z. Dauter and W. R. Rypniewski, *Nucleic Acids Res.*, 1997, **25**, 4599–4607.
- M. Egli, *Antisense Nucleic Acid Drug Dev.*, 1998, **8**, 123–128.
- M. Manoharan, *Biochim. Biophys. Acta*, 1999, **1489**, 117–130.
- M. Manoharan, *Curr. Opin. Chem. Biol.*, 2004, **8**, 570–579.
- A. M. Kawasaki, M. D. Casper, S. M. Freier, E. A. Lesnik, M. C. Zounes, L. L. Cummins, C. Gonzalez and P. D. Cook, *J. Med. Chem.*, 1993, **36**, 831–841.
- M. Teplova, G. Minasov, V. Tereshko, G. Inamati, P. D. Cook, M. Manoharan and M. Egli, *Nat. Struct. Biol.*, 1999, **6**, 535–539.
- J. J. Barchi, Jr, L. S. Jeong, M. A. Siddiqui and V. E. Marquez, *J. Biochem. Biophys. Methods*, 1997, **34**, 11–29.
- I. Berger, V. Tereshko, H. Ikeda, V. E. Marquez and M. Egli, *Nucleic Acids Res.*, 1998, **26**, 2473–2480.
- G. Minasov, M. Teplova, P. Nielsen, J. Wengel and M. Egli, *Biochemistry*, 2000, **39**, 3525–3532.
- J.-F. Trempe, C. J. Wilds, A. Y. Denisov, R. T. Pon, M. J. Damha and K. Gehring, *J. Am. Chem. Soc.*, 2001, **123**, 4896–4903.
- F. Li, S. Sarkhel, C. J. Wilds, Z. Wawrzak, T. P. Prakash, M. Manoharan and M. Egli, *Biochemistry*, 2006, **45**, 4141–4152.
- Z. Wu, M. Maderia, J. J. Barchi, Jr, V. E. Marquez and A. Bax, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 24–28.
- M. Egli, *Chem. Biol.*, 2002, **9**, 277–286.
- M. Egli and V. Tereshko, Curvature and Deformation of Nucleic Acids: Recent Advances, New Paradigms, *ACS Symp. Ser.*, 2004, **884**, 87–109.
- M. Egli, S. Portmann and N. Usman, *Biochemistry*, 1996, **35**, 8489–8494.
- V. Tereshko, G. Minasov and M. Egli, *J. Am. Chem. Soc.*, 1999, **121**, 470–471.
- W. K. Olson, A. A. Gorin, X. J. Lu, L. M. Hock and V. B. Zhurkin, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 11163–11168.
- A. Sarai and H. Kono, *Annu. Rev. Biophys. Biomol. Struct.*, 2005, **34**, 379–398.
- M. Egli, N. Usman and A. Rich, *Biochemistry*, 1993, **32**, 3221–3237.
- R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Nature*, 1980, **287**, 755–758.
- W. F. Lima, J. G. Nichols, H. Wu, T. P. Prakash, M. T. Migawa, T. K. Wyrzykiewicz, B. Bhat and S. T. Crooke, *J. Biol. Chem.*, 2004, **279**, 36317–36326.
- A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022–1025.
- Collaborative Computational Project Number 4, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 760–763.
- Q. Du, N. Carasco, M. Teplova, C. J. Wilds, X. Kong, M. Egli and Z. Huang, *J. Am. Chem. Soc.*, 2002, **124**, 24–25.
- M. Teplova, C. J. Wilds, Z. Wawrzak, V. Tereshko, Q. Du, N. Carrasco, Z. Huang and M. Egli, *Biochimie*, 2002, **84**, 849–858.
- P. S. Pallan and M. Egli, *Nat. Protocols*, 2007, **2**, 647–651.
- P. S. Pallan and M. Egli, *Cell Cycle*, 2008, **7**, 2562–2569.
- A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, 1998, **54**, 905–921.
- G. M. Sheldrick and T. R. Schneider, *Methods Enzymol.*, 1997, **277**, 319–343.
- R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1989, **6**, 655–667.
- C. R. Aiken, L. W. McLaughlin and R. I. Gumport, *J. Biol. Chem.*, 1991, **266**, 19070–19078.