Effects of C2′-Substitution on Arabinonucleic Acid Structure and Conformation

Divi Venkateswarlu and David M. Ferguson*

Department of Medicinal Chemistry and Minnesota Supercomputer Institute University of Minnesota, Minneapolis, Minnesota 55455

Received March 16, 1999

Sugar modifications play a central role in the development of potent antisense oligonucleotides.[1−3] Although a wide variety of C2′-ribose substitutions have been shown to increase antisense oligonucleotide (AO) affinity toward RNA, none have been shown to elicit RNase H activity which is critical to drug efficacy.[4] This failure can be traced to the conformational properties of the AO:RNA hybrid which are primarily determined by the preferred sugar puckers of the nucleotide. Most C2′-ribose substitutions favor C3′-endo puckering.[5] While this promotes preorganization of the antisense strand to a more stable A-form geometry, the resulting conformation of the AO:RNA hybrid is apparently not recognized by RNase H.[6] This enzyme is considered to be sensitive to the minor groove dimensions of the duplex, which assume an intermediate value between the canonical A- and B-form widths in the native DNA:RNA substrate.[7] The conformation of most ribose-derived AO:RNA hybrids is therefore too close to the ideal A-form geometry to activate RNase H cleavage of RNA.

Recently, arabino-based nucleotides have been shown to invoke RNAse H activity when hybridized with complementary RNA.[8] Although no X-ray crystallographic data is available for arabinonucleic acids (ANA) complexed with RNA, Egli and co-workers have determined the structure of an ANA:DNA complex.[9] Their work has shown that 2′-deoxy-2′-fluoroarabinose substitutions adopt an O4′-endo conformation in the duplex, producing stable B-form geometries when complexed with complementary DNA. In this paper, we report the results of a computational study of ANA:RNA hybrids containing C2′-arabino substitutions with two goals in mind. The first seeks to explain the basis for RNase H activity of ANA:RNA hybrids. It follows from previous work that this is most likely due to the preference of arabinonucleotides to adopt B-like conformations,[9,10] leading to ANA:RNA duplex structures that are geometrically similar to native DNA:RNA substrates. The second focuses on the effect of C2′-arabino substitution on conformation and stability. This is particularly important, given the vast array of potential modifications that could be incorporated at the C2′ position.[1−3]

The starting geometries for the ANA:RNA complexes were derived from experimental NMR structures of the hybrid sequence d(GCGTCAGG):r(CCUGACGC).[11] The deoxyribose units were modified appropriately to incorporate 2′-deoxy-2′-fluoroarabino (2′F), 2′-deoxy-2′-methoxyarabino (2′OMe), and standard arabino (2′OH) nucleotides. Ab initio calculations were initially performed on simple arabinonucleoside models to examine the puckering profiles of the three derivatives. The resulting energy profiles (which focused on the pseudorotation cycle from C3′-endo to C2′-endo) were further applied to parametrize the sugar fragments for subsequent molecular dynamics (MD) simulation. The general protocol as outlined by Cornell et al. was followed to parametrize the new residues for incorporation in the AMBER 4.1 force field database.[12,13] A listing of the partial atomic charges, valence and torsional parameters, and van der Waals parameters is reported in the Supporting Information. The three oligonucleotide complexes were placed in periodic boxes of TIP3P water and slowly equilibrated, following a procedure reported elsewhere.[14] Production runs were performed over 1 ns at constant temperature/pressure (300 K/1 atm) using 2-fs time steps. The SHAKE algorithm was applied to constrain all bonds to hydrogen, and coordinates were saved every 2 ps for further analysis.

The average physical properties of the ANA(2′OH):RNA and ANA(2′F):RNA hybrids were examined first, since these resemble complexes previously analyzed with X-ray crystallography as well as those used to investigate RNase H activity.[8,9] An analysis of the sugar puckers of the ANA(2′F) and ANA(2′OH) strands indicates that both adopt a more southern, or C2′-endo geometry, consistent with the ab initio model calculations. A close examination of the pucker values and structural details, however, shows that some substitution-dependent differences are evident in the two complexes. While the standard arabinose sugars favor southern pucker phase angles of ∼145−162° (C2′-endo), the values are slightly more northern and varied toward the O4′-endo domain (∼73−144°) in the fluoro-substituted strand. This difference is most likely due to the formation of an internal hydrogen bond between the C2′-hydroxyl group and the C5′-oxygen of the arabinose sugar that essentially "locks" the conformation in the C2′-endo form. This interaction is shown in Figure 1 and was found to be long-lived in the central six arabinose sugars over the course of the simulation (>95% populated). The lack of an analogous interaction, as well as the decreased size of the fluoro group, may also explain the increased variability noted in the phase angles of the ANA(2′F) strand. Although it had been reported that the fluorine makes steric contacts with flanking bases in the crystal structure of an ANA(2′F):DNA,[9] we see no evidence of that here. It is important to point out, however, that ANA:DNA structures adopt a B-form geometry,[9] which is inherently different than the A-like conformations reported for DNA:RNA hybrids.[11,15] In both ANA:RNA complexes, the sugars of the RNA complement strongly favor the northern or C3′-endo conformation (also shown in Figure 1). Overall, this produces a hybrid geometry that is structurally similar to that of the native DNA:RNA. The similarities were further verified by evaluating the helicoidal parameters of the duplexes. (A complete list of structural parameters is given in the Supporting Information.) A comparison of the minor groove widths (as estimated by interstrand phosphate distances) of the ANA(2′F):RNA, ANA(2′OH):RNA, and DNA:RNA structures shows all three share intermediate values (14−15 Å) when compared to ideal A- or B-form geometries. This may explain the RNase H activity of ANA:RNA hybrids reported

(1) De Mesmaeker, A.; Haner, R.; Martin, P.; Moser, H. *Acc. Chem. Res.* **1995**, *28*, 366.
(2) Freier, S. M. and Altmann, K−H. *Nucleic Acids Res.* **1997**, *25*, 4429.
(3) Sanghvi, Y. S., Cook, P. D., Eds. *Carbohydrate Modifications in Antisense Research*; ACS Symposium Series No. 580; American Chemical Society: Washington, DC, 1994.
(4) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.
(5) (a) Lubini, P.; Zurcher, W.; Egli, M. *Chem. Biol.* **1994**, *1*, 39−45. (b) Martin, P. *Helv. Chim. Acta* **1995**, *78*, 486.
(6) Crooke, S. T.; Lemonidis, K. M.; Neilson, L.; Griffey, R.; Lesnik, E. A.; Monia, B. P. *Biochem. J.* **1995**, *312*, 599.
(7) Fedoroff, O. Y.; Salazar, M.; Reid, B. R. J. *Mol. Biol.* **1993**, *233*, 509.
(8) Damha, M. J.; Wilds, C. J.; Noronha, A.; Brukner, I.; Borkow, G.; Arion, D.; Parniak, M. A. *J. Am. Chem. Soc.* **1998**, *120*, 12976.
(9) Berger, I.; Tereshko, V.; Ikeda, H.; Marquez, V. E.; Egli, M. *Nucleic Acids Res.* **1998**, *10*, 2473.
(10) Ikeda, H.; Fernandez, R.; Wilk, A.; Barchi, J. J., Jr.; Marquez, V. E. *Nucleic Acids Res.* **1998**, *26*, 2237.
(11) Bachelin, M.; Hessler, G.; Kurz, G.; Hacia, J. G.; Dervan, P. B.; Kessler, H. *Nat. Struct. Biol.* **1998**, *5*, 271.
(12) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(13) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. R.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D. M.; Seibel, G. L.; Kollman, P. A. *Comput. Phys. Com.* **1995**, *91*, 1.
(14) Lind, K. E.; Mohan, V.; Manoharan, M.; Ferguson, D. M. *Nucleic Acids Res.* **1998**, *26*, 3694.
(15) Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 4805.

10.1021/ja990870k CCC: $18.00
Published on Web 05/29/1999

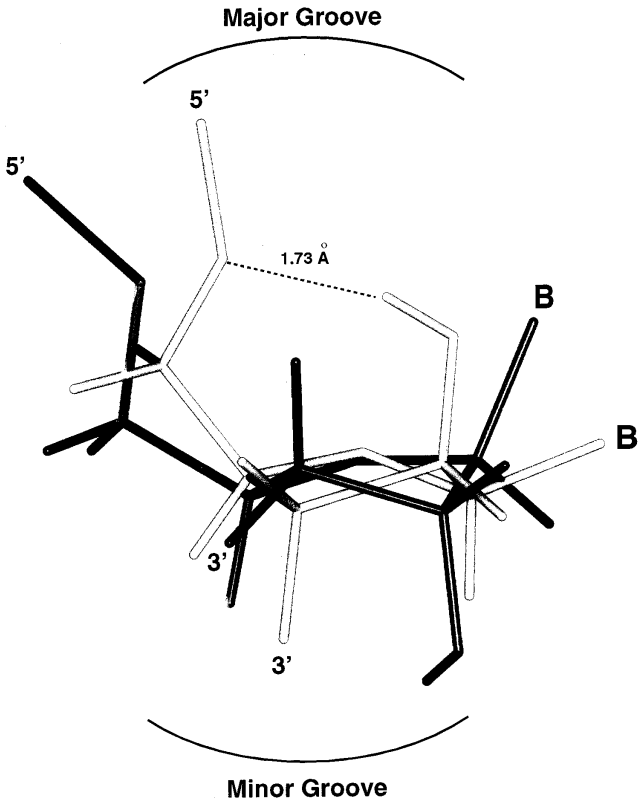

Figure 1. Superposition of ANA(2′OH) and RNA sugar conformations with major and minor groove directions given for perspective. The internal hydrogen bond interaction of the arabinose sugar is shown with a dashed line. The structures are taken from the central C:G complement of the ANA(2′OH):RNA complex.

in a previous paper.[8] It has also been suggested that the steric composition of the minor groove may be an important factor for RNase H recognition. For most ribose-based antisense nucleotides, C2′-substitutions project into the minor groove, potentially blocking recognition by RNase H.[14] The ANA C2′-substituents, however, adopt the opposite stereochemistry and occupy steric space in the major groove of the duplex. (See also Figure 1.) Although this suggests that most ANA C2′-modifications would project away from the minor groove, it is important to point out that this also depends on sugar pucker. Simple test simulations of the ANA:RNA duplexes show a C3′-endo pucker would dispose the C2′-substituent of the arabinose ring to the minor groove, dramatically changing the macromolecular conformation.

Simulations of a methoxy-substituted ANA:RNA complex have also been performed to investigate the potential effects of C2′−O−X derivatization on the macromolecular conformation. This site is of particular importance due to the wide array of C2′−O−X modifications that have been reported in the development of ribose-based antisense compounds.[1−3] The average structure taken from the last 500 ps of the simulation indicates the ANA(2′OMe):RNA complex adopts a conformation that is very similar to that of DNA:RNA. The pucker phase angles are comparable to those noted for the ANA(2′F) strand, lending further strength to the conclusion that an intramolecular hydrogen bond may lock the standard arabinose sugar in a more southern C2′-endo conformation. Given the previous report of potential steric contacts between the relatively small fluorine atom and flanking bases, we were particular concerned with the steric environment surrounding the methoxy group. Over the course of the simulation, nonbonded distances to the nearest neighbors of this group were therefore monitored for close contacts. Although we found no unfavorable steric contacts between flanking bases and the methyl hydrogens, some crowding of the methoxy oxygen was observed. As shown in Figure 2, the closest contacts occur

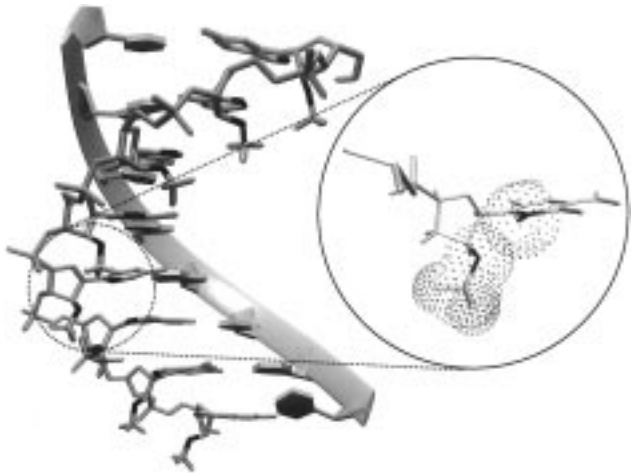

Figure 2. Structure of the ANA(2′OMe):RNA complex with the steric contacts of the methoxy group highlighted in black. The closeup view shows the relevant van der Waals surfaces.

between the oxygen and C6/C8 carbons of the flanking bases. While this interaction is slightly repulsive (with an average nonbonded distance of 3 Å), the contacts do not induce noticeable changes in the hybrid geometry as indicated by an analysis of helicoidal parameters of the duplex. (A listing of the close contact distances for all ANA:RNA complexes is given in the Supporting Information.)

Taken overall, the results indicate C2′-modified ANA:RNA complexes adopt conformations that are very similar to that of the native DNA:RNA substrate of RNase H. Subtle differences in the conformations, however, are noted depending on the arabino C2′-substituent. In particular, the sugar puckers of the ANA(2′OH) strand were found to adopt a more southern C2′-endo conformation when compared with the fluoro or methoxy derivatives. This may explain the observed decrease in stability and RNase H activity previously reported for ANA(2′OH):RNA complexes when compared with analogous ANA(2′F):RNA hybrids.[8] Our study indicates the ANA(2′OH) strand is more rigid due to the formation of an intramolecular hydrogen bond between the C2′-hydroxyl and the C5′-oxygen of the arabinose sugar. This most likely predisposes the ANA strand to a more B-like conformation, leading to less favorable ANA−RNA interactions in the A-like hybrid geometry. Perhaps the most provocative result of this study regards the potential use of C2′-modified ANA strands in the development of antisense compounds. While some steric contacts were noted between the C2′-methoxy oxygen and flanking bases, the simulations indicate that a vast array of C2′−O−X substitutions could be accommodated in the major groove of the ANA:RNA duplex. Given the past success of 2′-RNA substitutions in modulating the thermal stability of AO:RNA complexes,[1−3] further studies of C2′-ANA derivatives may prove highly fruitful in the development of antisense compounds. Of course, the design of potential ANA derivatives should also consider the steric environment about the C2′-oxygen as well as the relationship between sugar puckering and macromolecular conformation. Work is currently underway in our laboratory to investigate the effect of incorporating C2′−N−X, C2′−S−X, and larger alkoxy substituents on ANA:RNA structure and stability.

Supporting Information Available: Listing of AMBER force field parameters; tables of helicoidal parameters including C2′-contact distances; sugar pucker profiles; relative energies of arabinonucleoside conformations; coordinates for ANA(2′OH):RNA, ANA(2′F):RNA and ANA(2′OMe):RNA complexes in PDB format and details of figure generation (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA990870K