Published online 25 May 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 15* ***7573–7583***
doi:10.1093/nar/gks470

How does hydroxyl introduction influence the double helical structure: the stabilization of an altritol nucleic acid:ribonucleic acid duplex

Margriet Ovaere[1], Jiri Sponer[2,3], Judit E. Sponer[2,3], Piet Herdewijn[4] and Luc Van Meervelt[1,*]

[1]Department of Chemistry, Katholieke Universiteit Leuven, Biomolecular Architecture and BioMacS, Celestijnenlaan 200F, B-3001 Leuven, Belgium, [2]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-61265, Brno, Czech Republic, [3]CEITEC - Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic and [4]Laboratory of Medicinal Chemistry, Katholieke Universiteit Leuven, Rega Institute for Medical Research and BioMacS, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received February 28, 2012; Revised April 30, 2012; Accepted May 2, 2012

ABSTRACT

Altritol nucleic acids (ANAs) are a promising new tool in the development of artificial small interfering ribonucleic acids (siRNAs) for therapeutical applications. To mimic the siRNA:messenger RNA (mRNA) interactions, the crystal structure of the ANA:RNA construct a(CCGUAAUGCC-P):r(GGCAUUACGG) was determined to 1.96 Å resolution which revealed the hybrid to form an A-type helix. As this A-form is a major requirement in the RNAi process, this crystal structure confirms the potential of altritol-modified siRNAs. Moreover, in the ANA strands, a new type of intrastrand interactions was found between the O2′ hydroxyl group of one residue and the sugar ring O4′ atom of the next residue. These interactions were further investigated by quantum chemical methods. Besides hydration effects, these intrastrand hydrogen bonds may also contribute to the stability of ANA:RNA duplexes.

INTRODUCTION

In antisense technology, the antisense oligonucleotides have to hybridize strongly and selectively with their messenger ribonucleic acid (mRNA) complement. A variety of nucleic acid modifications have been synthesized for these purposes. Insertion of a methylene group between the ring oxygen atom and the anomeric carbon atom of the furanose ring of RNA gives altritol nucleic acid (ANA, Figure 1a). This chemical insertion has a profound effect on the physicochemistry and the biology of these nucleic acids. The nucleic acid becomes chemically and enzymatically more stable than RNA (1), while keeping very selective and strong hybridization properties following Watson-Crick rules (2). As a result of this, ANA has scored very well in a small interfering RNA (siRNA) screening assay (3).

When carrying out the same insertion in DNA, hexitol nucleic acid (HNA) is obtained. Likewise, HNA-modified RNA's show strong siRNA effects (4). The sugar rings of DNA and RNA (having a furanose sugar moiety) are more flexible than the sugar rings of HNA and ANA (having a reduced pyranose sugar moiety) and hybrids between ANA and DNA or RNA are more stable than hybrids between HNA and DNA or RNA, which points to the importance of the presence of the OH group for duplex stabilization.

Herein, we report on the structural and physiochemical reasons for the duplex stabilization effect of this OH group in an ANA:RNA hybrid, which is not observed in regular dsRNA's.

MATERIALS AND METHODS

Oligonucleotide synthesis

Synthesis and assembly of the ANA strand were performed according to Ovaere *et al.* (5). At the 3′-end of the ANA strand, an extra phosphate group was added because of solid support choice. The RNA sequence was purchased from Eurogentec. Hybridization was accomplished by titration and monitored by NMR.

Crystallization conditions

Crystals were formed after about 1 year by the vapour diffusion hanging drop method at 289 K using Crystal

*To whom correspondence should be addressed. Tel: +32 16 327609; Fax: +32 16 327990; Email: Luc.VanMeervelt@chem.kuleuven.be

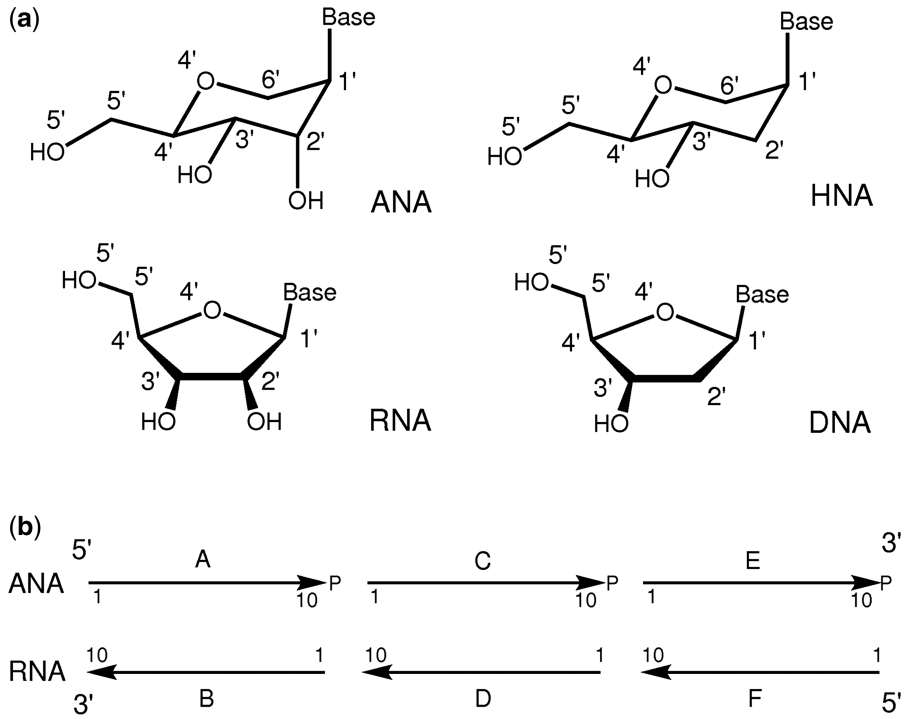

Figure 1. (**a**) Representation of an ANA and an HNA nucleoside in $^1C_4$ chair conformation together with the RNA and DNA nucleosides for comparison. The numbering scheme is assigned. (**b**) Numbering scheme for the three ANA:RNA duplexes in the asymmetric unit.

Screen II (Hampton Research) in a condition which includes 0.01 M nickel chloride hexahydrate, 1.0 M lithium sulphate monohydrate and 0.1 M Tris (tris(hydroxyl-methyl)aminomethane) pH 8.5 as buffer. For the crystallization drop, 1 µl of 0.5 M ANA:RNA decamer (duplex concentration) was combined with 1 µl screening condition and equilibrated against 500 µl crystallization screen. One rod-form crystal (dimensions 0.3 mm × 0.10 mm × 0.06 mm) was obtained and cryoprotected with 40% propylene glycol.

Data collection and processing

Diffraction data were recorded at the Swiss Light Source PXIII beamline (Paul Scherrer Institute, Villigen, Switzerland) on a MAR225 CCD detector (100 K, wavelength 1.000 Å, crystal to detector distance 200 mm). In total, 200 frames (φ increment 0.5°) covering a 100°-data range were collected to a resolution of 1.96 Å. Two crystals with similar unit cell could be found in the reflection data, and these data were integrated separately by MOSFLM (6). Nevertheless, the diffraction data originate mainly from one crystal and only these data could be used for structure solution. The data were scaled by Scala (7), and processing statistics is listed in Table 1. The crystal belongs to the orthorhombic crystal class 222 with $a = 26.07$, $b = 42.51$ and $c = 157.79$ Å. Systematic absences analysed by Pointless (7) suggested $P2_12_12_1$ as space group.

Structure solution and refinement

As the crystal structure of the ANA:RNA decamer was already partly solved in another unit cell but suffering from lattice translocation defects (unpublished results), a decamer of this crystal structure was used as molecular replacement model without the extra phosphate group and without the last base pair at the ANA 3′-side. Molecular replacement by Phaser (8) resulted in a helix oriented almost parallel to the *c*-axis. As the Matthews coefficient (2.22 Å^3/Dalton) suggested three decamers in the asymmetric unit, two more duplexes were searched for. The three decamers are stacked on each other in a head-to-tail manner.

Structure refinement was performed by Refmac (9) using the restraints based on the crystal structure of the altritol adenosine building block (Supplementary Data) for the ANA sugar rings and standard dictionary restraints for the ANA bases and the RNA strand. The $F_o - F_c$ and $2F_o - F_c$ electron density maps were carefully studied for any inconsistency, and bond distances, angles and chiral volumes were monitored. The 'Find Waters' option of the program Coot was used to localize 239 water molecules (10). Disordered solvent regions are modelled according to Babinet's principle (11). Finally, a R_1 value of 22.44% was obtained (R_{free} value: 23.42%). Refinement statistics is listed in Table 1. Final coordinates and structure factor amplitudes have been deposited with the Protein Data Bank (3OK2) and Nucleic Acid Data Bank (NA0770).

Quantum chemical calculations

Geometry optimizations were performed at B3LYP/6-311++G(2d,2p) level of theory. Initial positions of the C, O and P atoms were taken from the X-ray geometry. Selection of the starting geometries is described in the Results section. Starting H-atom positions were generated

Table 1. Data collection and refinement statistics for the crystal structure of the ANA:RNA decamer

Space group	$P2_12_12_1$
Resolution range (Å)	39.45–1.96 (2.07–1.96)
Measured reflections	50 925 (7258)
Unique reflections	11 922 (1775)
Completeness (%)	90.3 (93.7)
R_{merge} (%)	4.8 (70.2)
Multiplicity	4.3 (4.1)
Mean $I/\sigma(I)$	9.3 (2.1)
Average mosaicity (°)	0.70
Number of nucleic acid atoms	1302
Number of water molecules (treated as O)	239
R_1 value (%)	22.44
Mean B value of oligonucleotides (Å^2)	53.58
Mean B value of water molecules (Å^2)	58.06
RMS deviation from restraint target values	
Bond lengths (Å)	0.017
Bond angles (°)	2.508
Planes (Å)	0.019

Values in parentheses are for the outermost shell.

arbitrarily and later fully refined in the course of the geometry optimizations. The model systems carried a total charge of −1 and comprised two altritol moieties as well as a phosphate group linker. For simplicity, the hydroxymethyl group of the altritol unit phosphorylated at 3′ was substituted with a methyl group. All torsion angles along the chain O4′-C4′-C5′-O5′-P-O3′-C3′-C2′ (Supplementary Scheme S1) were fixed at the crystal value in the course of the geometry optimizations. The rest of the geometrical parameters were optimized. In this way, the interaction is derived in a manner consistent with the experimental structure (an overlay of the experimental and optimized geometries is presented in Figure 4c). The Gaussian09 computer code, Revision A.1 (Gaussian, Wallingford, CT, USA) was used for geometry optimizations. Electronic densities derived from the B3LYP/6-311++G(2d,2p) wave function were analyzed in the frame of Bader's Atoms in Molecule approximation (12) with the AIM2000 code (13,14). Interaction energies were evaluated at several theoretical levels using the B3LYP/6-311++G(2d,2p) optimized geometries (details of the interaction energy computations are provided in the Supplementary Data). Relative energies in the presence of bulk water were obtained from single point calculations using the B3LYP/6-311++G(2d,2p) optimized geometries and the polarizable conductor continuum solvent model (15,16) (C-PCM) in the standard parameterization supported by the Gaussian09 program: the average surface of a tesserae was 0.4 Å^2 and the minimum radius of the added spheres used to create the solvent excluded surface was 0.2 Å. The United Atom (UA0) topological model and a scaling factor of 1.1 were used to define the atomic radii.

RESULTS

Overall helical structure

The crystal structure of the ANA:RNA hybrid decamer a(CCGUAAUGCC-P):r(GGCAUUACGG) reveals a right-handed, anti-parallel double helix with three helices in the asymmetric unit (Figure 2, for numbering scheme see Figure 1b). The electron density map reveals a discontinuous sugar-phosphate backbone throughout helical columns with gaps between subsequent decamers. Both strands are held together by standard Watson-Crick base pairing. In the centre of the duplex, an open space can be seen which classifies the ANA:RNA hybrid as an A-like helix. The three double helices are very similar in the central part and at the 3′-ends (Supplementary Figure S1). At the 5′-end of both the ANA and RNA strands, some small deviations can be observed, possibly caused by the absence of phosphate groups at these ends.

Overall, the backbone torsion angles are similar within each strand type and typical for A-type helices (Supplementary Table S1). For the ANA strands, the standard deviations on the torsion angles are relatively small, whereas for the RNA strands, the standard deviations for the α- and γ-angles are quite large. The γ-angle values for the ANA strands show a remarkable pattern. All three C1-residues, which lack a phosphate group, have deviating values for the γ-angle compared with the other residues. This can be explained by steric repulsion between the O5′-atom of the C1-sugar moiety and the extra 3′-end phosphate group of the neighbouring ANA chain. All bases in the ANA:RNA hybrid have an antiorientation around the glycosidic bond with χ values being lower for the ANA residues than for the RNA residues.

The helical parameters calculated by the 3DNA code (17), including comparison with A- and B-type helices (18), are summarized in Supplementary Tables S2–S4. When considering the standard deviations, the values for all helical parameters lie within the same ranges for the three different duplexes. The large X-displacement (−5.4 Å, displacement of a base pair along its short axis, perpendicular to the helical axis), inclination (13.4°, the angle between the long axis of a base pair and a plane perpendicular to the helical axis), slide (−2.1°, displacement between two successive base pairs along their long axis) and roll (6.9°, the angle for rotation of a base pair around its long axis, with respect to its neighbouring base pair in the duplex) classify the three duplexes as belonging to the A-family double helices. A rather high X-displacement and slide are observed for all duplexes compared with the values for A-type DNA. As exactly 12 base pairs make one full helical turn and the twist angle between two successive base pairs is about 30°, the structure of the hybrid is more particularly organized in an A′-form helix (19,20).

Supplementary Tables S3 and S4 also list base pair step parameters for transitions between subsequent helices. The twist angle for these transitions are rather low, especially between the CD and EF duplexes where the twist is only 19.5°. For the X-displacement on the other hand, very large values are observed for the transitions of one decamer to another, with the largest X-displacement for the CD–EF transition (−9.3 Å). The other helical parameters for the transition steps do not show large deviations from the values found within the duplexes.

For the ANA residues, the altritol sugar ring puckering modes are assigned to the $^1C_4$ chair conformation.

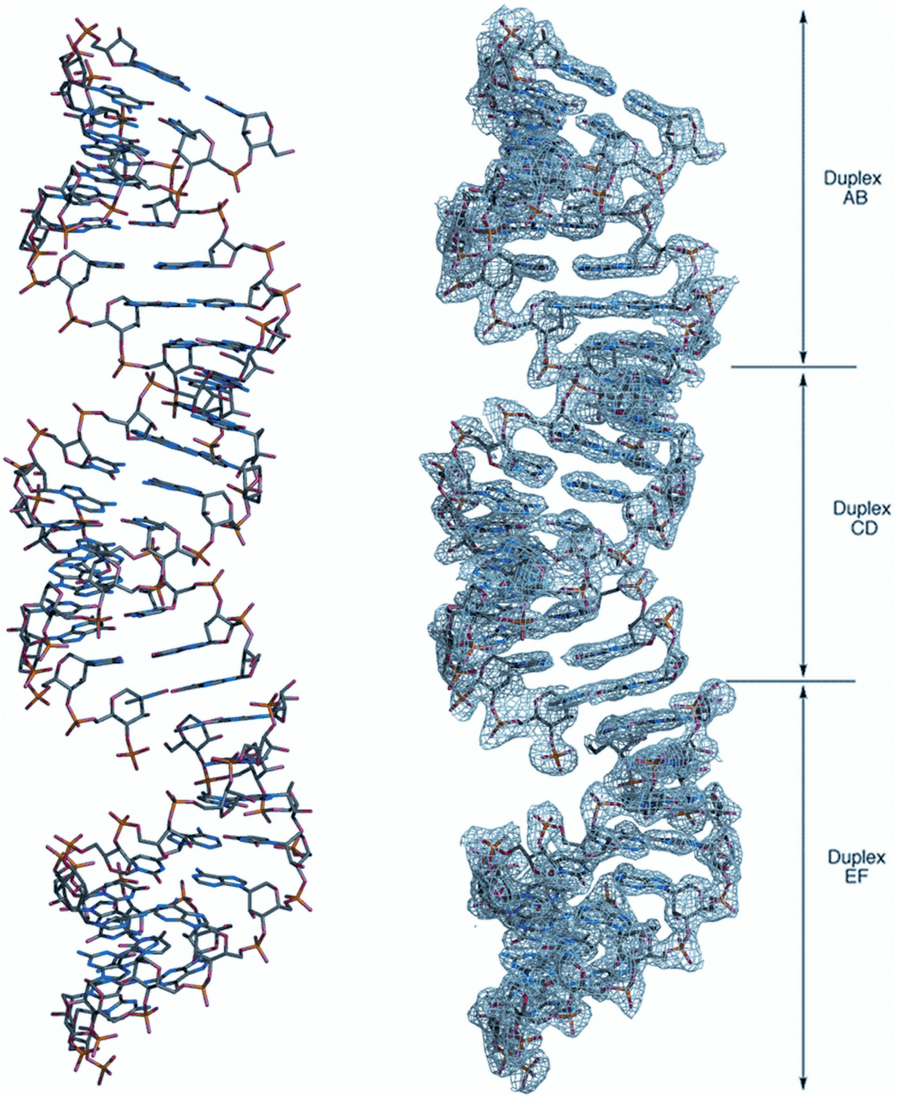

Figure 2. Side view of the asymmetric unit of the ANA:RNA decamer crystal. The right side shows the molecular structure in its contoured $2F_o-F_c$ electron density map (1.0 sigma level). The arrangement of the three different duplexes is assigned. Discontinuity of the helical column between duplexes CD and EF is clearly visible. Colour codes: nitrogen, blue; carbon, grey; oxygen, red and phosphorus, orange. Created with PyMOL (31).

The ribose sugar rings mainly adopt the C3′-*endo* conformation (Supplementary Table S5). This envelope puckering mode is observed for all sugar rings of strand D. One ribose ring of strand B and three sugar rings of strand F reside in the C2′-*exo* conformation, the envelope form with the C2′-atom displaced from the plane to the opposite side of the C5′-atom. The ribose ring at the 3′-end of strand F adopts the C1′-*endo* envelope conformation.

Within each strand, the sugar puckering is closely related to the inter-phosphate P_n-P_{n+1} distances. The ANA strands have an average inter-phosphate distance of 5.54 Å, whereas this distance is slightly higher for the RNA strands (5.91 Å). An outlying value of 6.9 Å is observed for the P_n-P_{n+1} distance between residues G9 and G10 of RNA strand F. This large distance originates from the deviating C1′-*endo* sugar puckering for the G10 sugar ring. For the transitions between subsequent ANA strands, the inter-phosphate distance between the 3′-end phosphate group and the G2 residue is on average 6.68 Å.

Intrastrand O2′−H···O4′ hydrogen bonds

The size of the six-membered altritol sugar ring compared with the smaller five-membered ribose can result in additional interactions between subsequent residues in ANA. The average distance between the O2′ atom of one residue and the sugar ring O4′ atom of the next residue is 3.12 Å (SD 0.24 Å), where it is 3.67 Å (SD 0.36 Å) in the RNA strand. To investigate the possibility of ANA backbone stabilization by O2′···O4′ interactions, the hydrogen atoms were added. Figure 3 visualizes positions of the H2′ atoms using the Protonate 3D algorithm employing the CHARMM force field (21), consistently with the prediction made using electronic structure QM computations

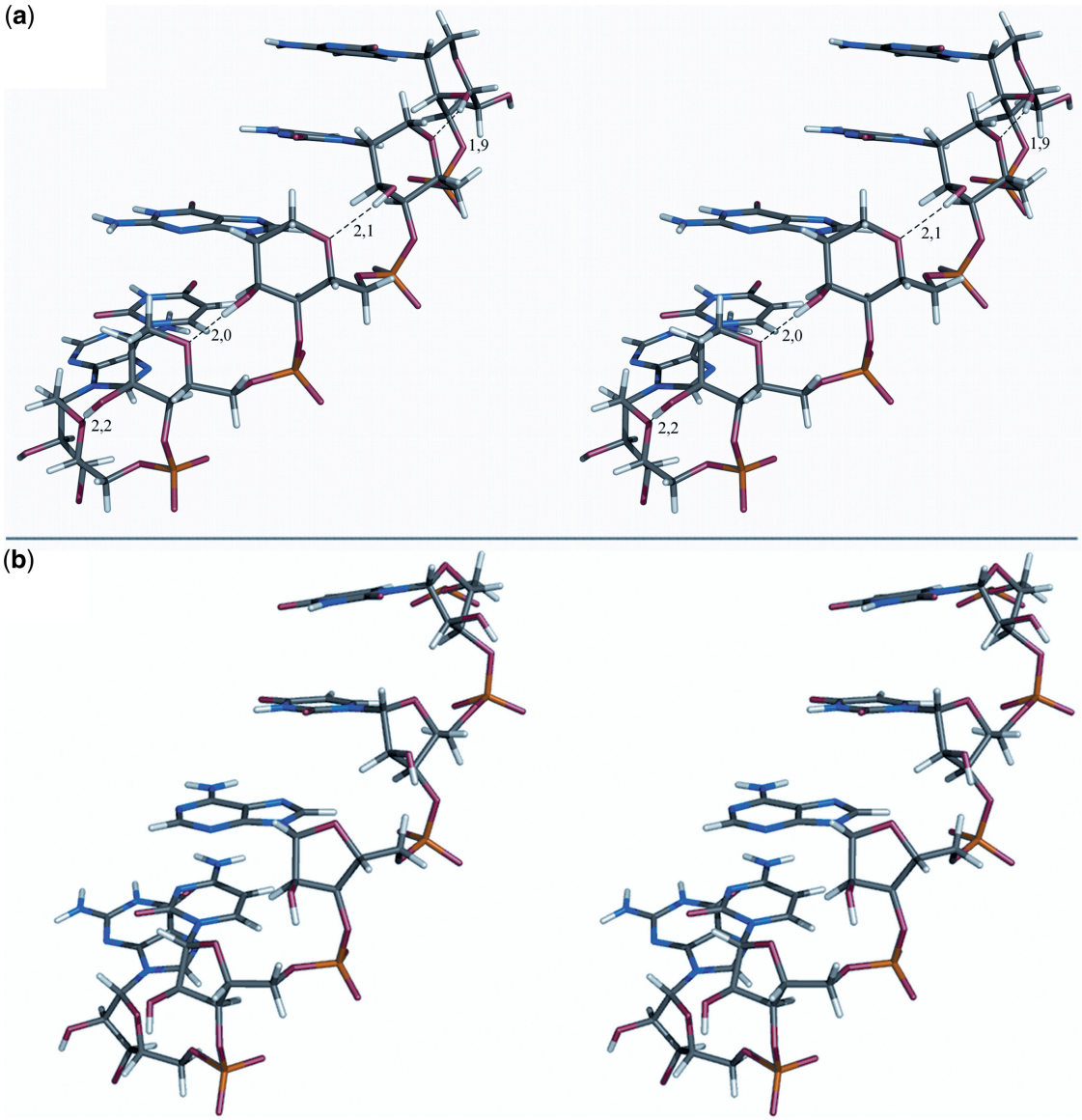

Figure 3. (**a**) Side view of a part of ANA strand E with the distances between the H2′ atom of one residue to the O4′ atom of the next assigned. (**b**) Side view of a part of RNA strand F. The H2′ atoms point away from the O4′ atom of the next residue. Colour codes according to atom type: carbon, grey; oxygen, red; nitrogen, blue and phosphorus, orange. Stereographic images created with PyMOL (31).

(see later). In the ANA strands, the H2′ hydrogen atoms mainly point towards the sugar ring O4′ oxygen atoms of the next residue, as was found for 20 of the 30 residues. For the RNA strands, on the other hand, the opposite pattern is observed. Only six H2′ hydrogen atoms out of 30 point towards the O4′-atom of the following residue. These findings lead to the assumption that interactions between the O2′ hydroxyl group of one residue with the O4′ oxygen atom of the next residue contribute to the stability of the ANA backbone.

On the basis of the interatomic O···O distances, in the crystal geometry of ANA, there are three potential acceptor sites for the H-bond donated by the O2′-hydroxyl group: either the O3′ or O5′-oxygens (average O···O distances are 2.78 Å with SD 0.14 Å and 3.84 Å with SD 0.42 Å, respectively) of the adjacent phosphate group or the neighbouring O4′-oxygen (average O···O distance 3.12 Å with SD 0.24 Å).

To assess the viability of these potential H-bonding interactions, we carried out high-quality quantum-chemical computations on small model systems (Figure 4). For a detailed description of the computational models, see the Interaction energy calculations part and Supplementary Scheme S1 in the Supporting Information. These model systems differed in the choice of the initial orientation of the O2′-hydroxyl group. In total, we have found two minima of practically identical stability (difference of the total electronic energies is 1.4 and 0.9 kcal/mol, in gas-phase and in bulk water, represented by the COSMO continuum solvent method, respectively), with the O2′-hydroxyl pointing towards O4′ and O3′, respectively. Optimization which was started from a structure

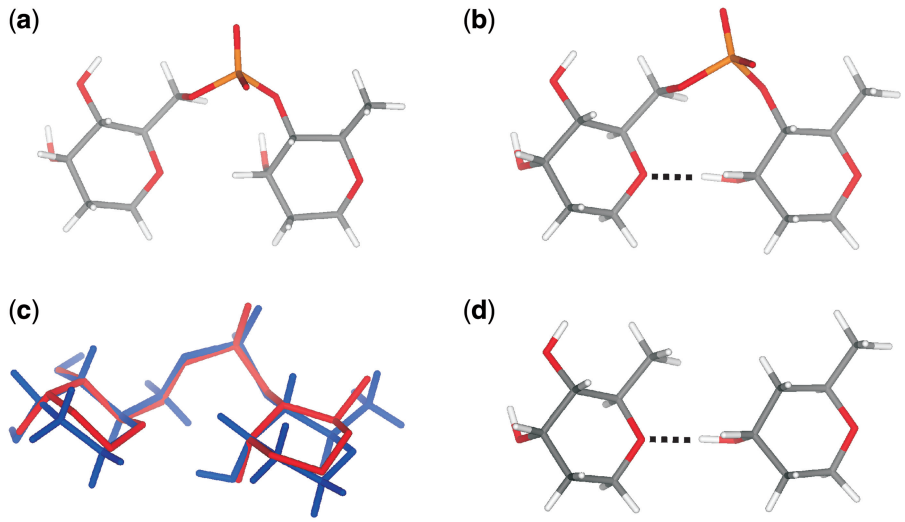

Figure 4. (**a** and **b**) Two optimized geometries obtained from quantum chemical calculations at the B3LYP/6-311++G(2d,2p) level. H-bonding contact confirmed by Bader-analysis is indicated with dotted lines. (**c**) Overlay of the B3LYP/6-311++G(2d,2p) optimized (blue) and crystal (red) geometries with H-bond between the O2′-hydroxyl and O4′. (**d**) Model used in interaction energy evaluation. For computational details, see the Supplementary Data.

with O2′-hydroxyl oriented towards O5′ converged to the structure with O2′-H•••O4′ H-bond.

We have carried out a Bader analysis of the electronic density for the geometry where the O2′-hydroxyl was oriented towards O4′ and indeed we have found a (3,−1) bond critical point (electron density is 0.0151, and the Laplacian of the electronic density is 0.0124), which unequivocally indicates the existence of an O2′-H•••O4′ H-bond. For comparison, we have evaluated the topological properties of the other optimized geometry, i.e. the one where the O2′-hydroxyl points nearly towards O3′. In this case, no (3,−1) bond critical point was found, demonstrating that no H-bond can be expected between O2′ and O3′, which is most likely because the O3′-H-O2′ angle (114.7°) is too small to form a stable H-bond. It has been shown that H-bonds between O or N donors (X) and acceptors (Y) require at least an X-H-Y bond angle of 120° in order to contribute to the stabilization to a noticeable extent (22). Therefore, in light of the Bader-analysis and the available structural data, viability of an O2′-H•••O3′ H-bond is much less likely than that of a concurrent O2′-H•••O4′ H-bond.

To see the energetic consequences of the O2′-H•••O4′ H-bonding interaction, we have carried out standard interaction energy computations at CBS(T) level (see Supporting Information for details) for the interaction of the two altritol units, by replacing the phosphate moiety with H atoms both at the 3′- and 5′-oxygens involved in the phosphodiester linkage (O-H distance for the added hydrogens was fixed at 0.96 Å). The CBS(T) abbreviation stands for complete basis set electron correlation calculations corrected for higher-order electron correlation effects. Such calculations are considered of chemical accuracy and represent established benchmarks (23). The computed interaction energy, −5.6 kcal/mol, shows visible stabilization contribution between the two altritol moieties and is comparable to the gas phase strength of a water dimer. However, the present system is almost exclusively stabilized by the dispersion energy (roughly the difference between the MP2 and HF terms, that amounts to −5.2 kcal/mol using the aug-cc-pVTZ basis set, for details see Supplementary Table S9). The interaction (although being intramolecular) appears to be even slightly stronger than that for the ribose-zipper type of interactions (computed interaction energy at CBS(T) level is ∼−4 kcal/mol), which are known to efficiently stabilize folded RNAs (24).

The calculations presented above were specifically based on the experimentally refined structure of chain C of the ANA:RNA duplex. As an example, we have taken the backbone geometry from the A6, A5 dinucleotide segment, with an experimental O2′•••O4′ distance of 3.15 Å, i.e. which is very close to the average of the experimentally determined O4′•••O2′ distances (3.12 Å). Note that despite constraining the key dihedral angles dictating the backbone conformation (see 'Materials and Methods' section), the O2′•••O4′ distance relaxed to 3.06 Å just due to optimization of valence angles and bond lengths. We then carried out a control computation of a second experimental structure (taken from strand C, dinucleotide step C2 and G3) with an experimental O2′•••O4′ distance of 3.57 Å. This geometry can be considered as the upper limit of the crystallographically observed O2′•••O4′ distances. Nonetheless, after geometry optimization (using the same optimization protocol as for the other geometry), the optimized O2′•••O4′ distance decreased to 3.11 Å, which is very close to the crystallographically found average value (3.12 Å). Due to the similarity of the optimized geometries obtained from the two different starting geometries, the CBS(T) interaction energy results were practically identical, see Supporting Information for further details.

The major groove widths, measured as the P_n-P_{n+2} distance minus 5.8 Å to account for the van der Waals radii of the phosphate groups, range from 9.5 Å to 10.8 Å with an average value of 10.1 Å. For the minor groove widths, the $C4'_n$–$C4'_{n+2}$ distances minus 3.4 Å (the van der Waals radii of two carbon atoms) were used as reference and have an average value of 7.9 Å with 7.3–9.0 Å range. The somewhat wider grooves, especially the major groove, are typical for the A′-form helix (20).

Hydrogen bond formation between the complementary bases in opposite strands holds the two strands of the double helix together. The distances of these hydrogen bonds lie well within standard hydrogen bond ranges (19).

Stacking interactions and crystal packing

The double helical structure of the ANA:RNA decamer is stabilized by intra- and inter-strand base stacking interactions. A representation of the stacking interactions between successive base pairs in duplex AB is shown in Supplementary Figure S2. Inter-strand stacking is observed for base pair steps 1, 2, 4, 7 and 9. Duplexes CD and EF follow the same pattern. Base pair stacking interactions for the transition between different duplexes are mainly intra-strand (Supplementary Table S6).

The three duplexes in the asymmetric unit stack on each other in a head-to-tail manner along the *c*-axis. By means of a 2-fold screw axis through the helix, six decamers stack on each other in the unit cell forming a helical column running through successive unit cells. A 2-fold screw axis perpendicular to the *c*-axis gives rise to the second 60 base pairs column in the unit cell. The columns are slightly curved and are therefore not exactly parallel to the *z*-direction (Supplementary Figure S3).

Between the ANA and RNA strands of neighbouring columns 16 interactions at hydrogen bond distances can be found (Supplementary Table S7). By these interactions, the side-by-side packing of the double helices is stabilized. Six times an RNA O2′-atom from one column is linked to two ANA oxygen atoms (O2′ and O3′ or O4′) of a neighbouring column. Also two hydrogen bonds linking an O2′-atom of one column to a phosphate group oxygen of another column are observed. A closer view of some of these interactions is shown in Figure 5.

Hydration

For the Matthews coefficient, a value of 2.22 $Å^3$/Dalton was obtained corresponding to a solvent content of 63%. In total, 239 water molecules were located in the crystal structure resulting in an average of 7.97 water molecules per base pair. The program Contact (25) was used to analyze the interactions between nucleic acid atoms and ordered water molecules in a distance range of 2.3–3.5 Å. Table 2 lists an overview of these contacts. Care has to be taken when considering these contacts as hydrogen bonds since no angle restrictions were used.

Hydration occurs mainly at the phosphate groups, with some more contacts for the ANA than for the RNA

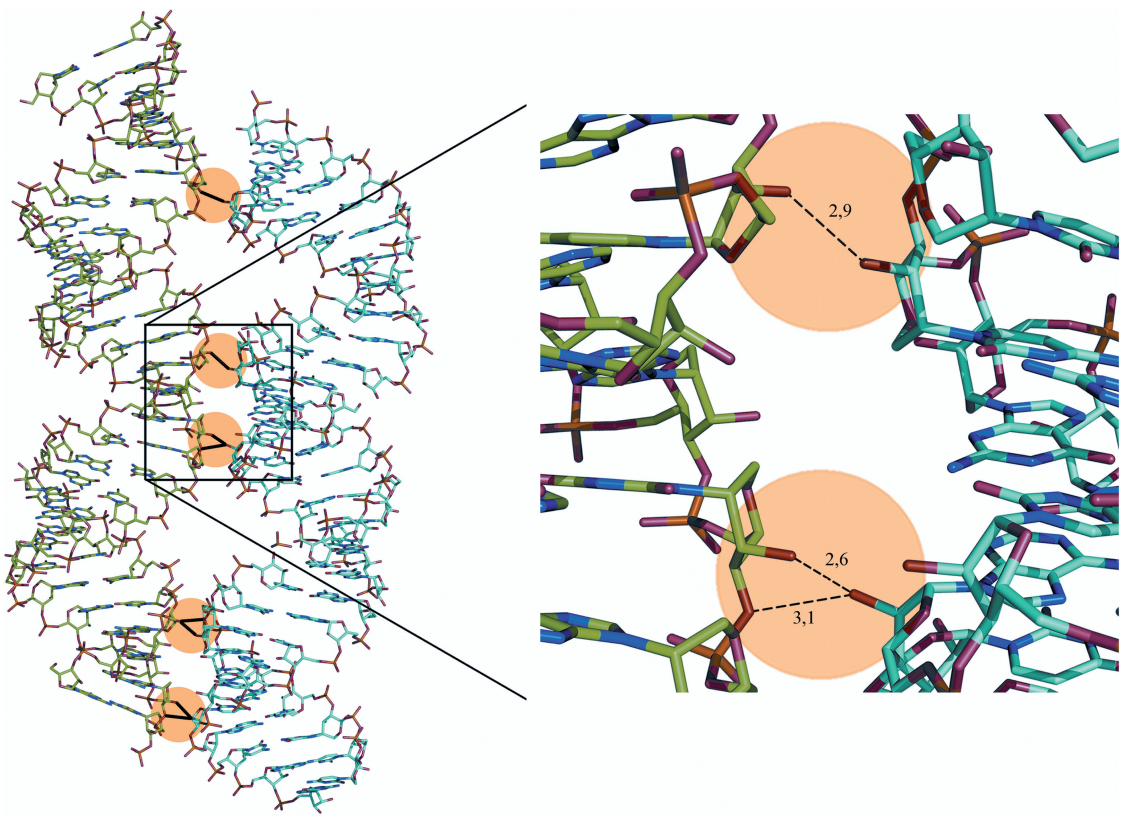

Figure 5. Interactions between neighbouring ANA:RNA duplexes. Created with PyMOL (31).

Table 2. Number of contacts between ordered water molecules and the hydration sites of phosphate groups (OP1, OP2, O3′, O5′), sugar rings (O2′, O4′) and minor (O2, N2, N3) and major (N4, O6, N7, O4, N6) groove in the ANA:RNA hybrid crystal structure

Strands	Phosphate	Sugar	Minor	Major	Total
All	172	73	64	71	380
ANA	93	27	26	31	177
RNA	79	46	38	40	203
AB	65	24	26	18	133
CD	39	21	14	25	99
EF	68	28	24	28	148

Values for all strands, for ANA and RNA strands separately and for the three different duplexes in the structure are given.

phosphate groups. The OP2 atoms pointing into the major groove form more contacts to water molecules compared with the OP1 atoms (74 versus 52), a feature that has been observed in DNA and RNA A-family helices (26). Water molecules bridging successive phosphate groups are present in both the ANA and RNA strands (more for ANA than for RNA). For the O5′-oxygen atoms, more interactions with water molecules are found in ANA than in RNA strands (15 for ANA against 8 for RNA).

The lower hydration level for duplex CD can largely be ascribed to reduced hydration at the phosphate groups and the minor groove. More water molecules are located in the major groove than in the minor groove, which can be explained by the larger amount of possible donor/acceptor sites in the major groove (27).

The sugar moieties of the ANA strands are less hydrated compared with the RNA strands. A smaller O2′•••O4′ inter-residue distance is observed for the ANA strands (on average 3.12 Å for ANA versus 3.67 Å for RNA) together with more O2′•••O4′ contacts at hydrogen bond distances (2.5–3.5 Å) between successive ANA residues (28 for ANA and only 10 for RNA on a total of 30 distances). The lower hydration level of the ANA sugar moieties can therefore be caused by loss of potential hydration site due to formation of the above-discussed inter-residue interactions.

In Figure 6, the solvent accessible surface area (SASA) of the ANA:RNA decamers is shown. The difference between the larger six-membered altritol sugar rings and the ribose sugar rings is clearly visible. To obtain the polar and apolar contributions to the total SASA, these were calculated by the GETAREA program (28) (Supplementary Table S8). When comparing the total SASA of the ANA strands (without extra phosphate groups) with those of the RNA strands, no large differences are noticed. Due to the extra carbon atom in the altritol sugar ring, the apolar part is larger for ANA than for RNA strands. The polar contributions on the other hand are slightly smaller for ANA.

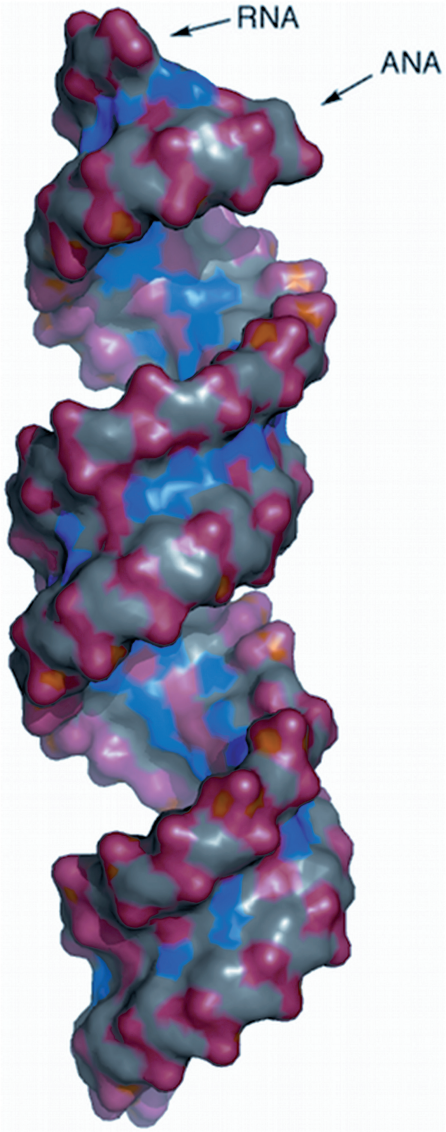

Figure 6. Side view of the solvent accessible surface area (SASA) of the three stacked ANA:RNA decamers in the asymmetric unit. The surface was calculated with a solvent probe radius of 1.4 Å. Colour codes correspond to the underlying atom type: C, grey; O, red; P, orange and N, blue. Created with PyMOL (30).

DISCUSSION

The altritol-modified nucleic acids were developed in search for oligonucleotides with an increased stability over HNAs. As a molecular-dynamics simulation revealed that solvation of the minor groove contributes to the difference in stability between a HNA:RNA and HNA:DNA duplex, an extra 2′-hydroxyl group in α-position which points towards the minor groove should increase the duplex stability (29). A study of the hybridization properties of ANA had already demonstrated that ANA:RNA duplexes are more stable than HNA:RNA duplexes (1). A crystal structure of a HNA:RNA decamer with a sequence analogous to that of the ANA:RNA decamer was solved by Maier *et al.* (30).

The HNA:RNA decamer has no extra phosphate group at the 3′-end of the HNA strand, the uracil bases in the ANA strand are replaced by thymine bases in the HNA

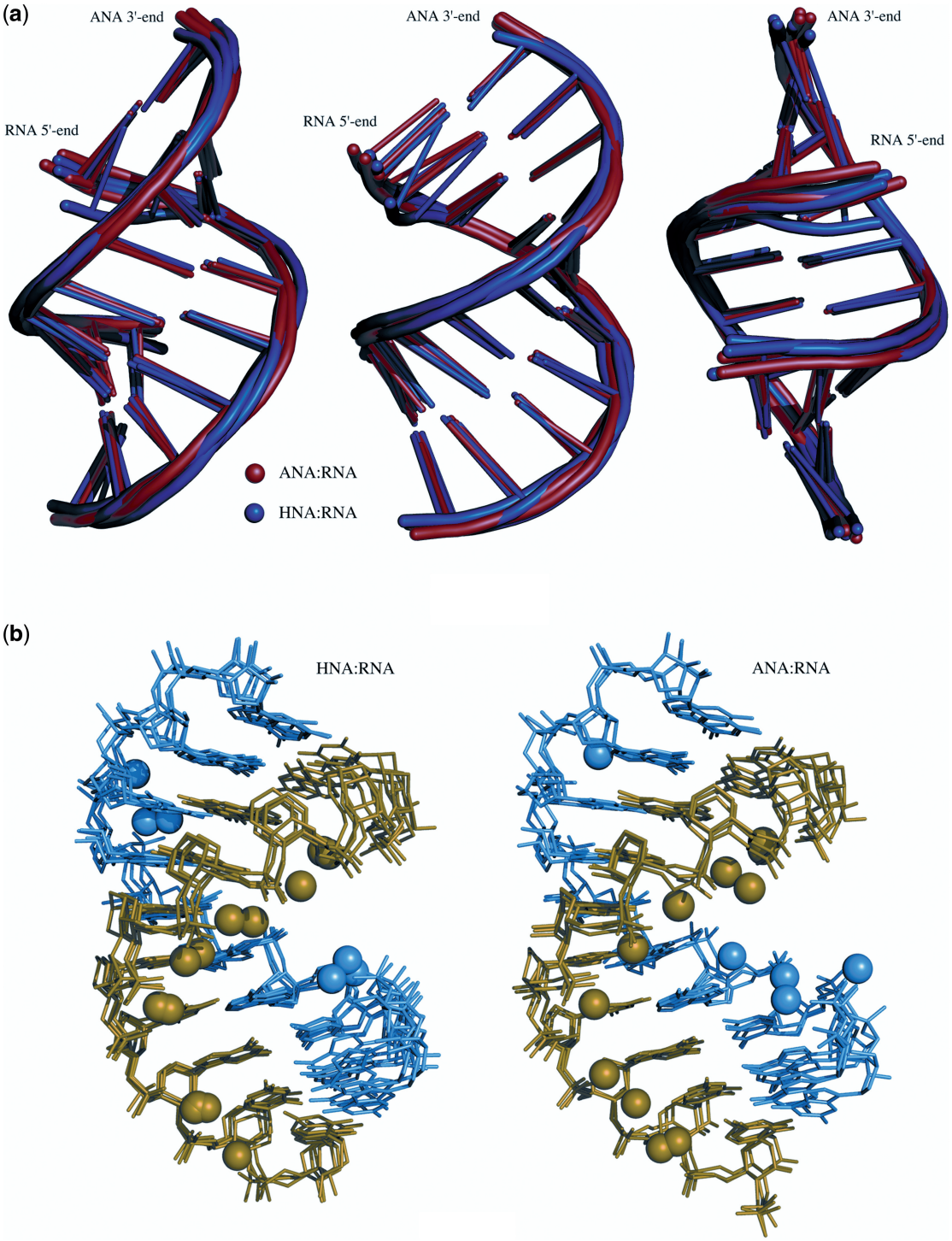

Figure 7. (**a**) Superposed decamers from the asymmetric unit of the ANA:RNA crystal structure (red) and the HNA:RNA crystal structure [blue, Maier *et al.* (30)]. Side view into the minor groove (left), into the major groove (right) and a side view in-between minor and major groove (centre). (**b**) Superposed decamers from the asymmetric unit of the HNA:RNA structure (left) and the ANA:RNA structure (right). The water molecules bridging between successive phosphate groups are presented as spheres. RNA strands and their bridging water molecules are coloured blue. HNA and ANA strands with their bridging water molecules are coloured green. Created with Superpose (25) and PyMOL (31).

strand, but in the RNA strand, the uracil bases are preserved. This crystal structure, with a resolution of 2.6 Å and solved in space group $P4_12_12$, has four duplexes in the asymmetric unit. In the crystal packing, pseudo-continuous helices containing eight decamers are formed. The HNA:RNA decamers fit well on the ANA:RNA decamers with an average positional RMS deviation of 0.95 Å, although some deviations can be noticed at the 5′-ends (Figure 7a). Similar hydration patterns are observed in the HNA:RNA and ANA:RNA

structures. The OP2 atoms of the modified strands are much more hydrated than those of the RNA strands. A higher hydration level of the O5′ atoms is also found for the modified strands compared with the RNA strands.

By the insertion of a methylene group in RNA and DNA, the inter-phosphate distances are shortened affecting the hydration pattern of the modified strands. The inter-phosphate distances are on average 5.54 Å in the ANA strands and 5.55 Å in the HNA strands of the ANA:RNA and HNA:RNA crystal structures, respectively. In the RNA strands of both crystal structures, these distances are on average nearly equal to 5.9 Å (5.91 Å and 5.85 Å in ANA:RNA and HNA:RNA, respectively), the expected value for A-type helices. By this shortened inter-phosphate distances in ANA and HNA, many more bridging water molecules between successive phosphate groups are observed, almost forming a water spine contributing to the duplex stability (Figure 7b). The bridging water molecules also interact with the O5′-atoms, which could explain their elevated hydration level.

For the ANA strands, the insertion of a methylene group introduces a new type of intrastrand interaction: because of the larger altritol sugar ring compared with the ribose ring, in the ANA strands the distance between the O2′-oxygen atom of one residue to the O4′-atom of the next is smaller compared with this distance in RNA, enabling O2′•••O4′ interresidue interactions. Force field computation indicated a possibility of O2′•••O4′ H-bonds in the ANA strand but not in the RNA strand. High-level quantum chemical calculations have shown that an O2′•••O4′ H-bond may contribute to the stabilization of the ANA backbone. It has rather remarkable intrinsic interaction energy ~5.5 kcal/mol with significant part of the stabilization coming from electron correlation (dispersion) contribution. Such stabilization is comparable with the energy of a good H-bond, whereas it is less electrostatic in nature, which may increase its relative importance in water environment. The findings reinforce the assumption that the higher stability of ANA:RNA over HNA:RNA is not only caused by extra hydration at the O2′-atoms but also is caused by inter-residue hydrogen bonding interactions between the O2′ atoms and O4′ sugar ring atoms of the next residue in the ANA strands, whereas HNA lacks the essential O2′-atom for these interactions.

AVAILABILITY

Crystallographic data (excluding structure factors) for the altritol adenosine building block reported in this article have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication no. CCDC-867845. Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: +44(0) 1223 336033 or email: deposit@ccdc.cam.ac.uk).

ACCESSION NUMBERS

CCDC-867845, Nucleic Acid Data Bank NA0770.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–9, Supplementary Figures 1–3, Supplementary Methods, Supplementary Scheme 1 and Supplementary References [17,18,25,28,31–45].

ACKNOWLEDGEMENTS

The authors thank the staff of the Swiss Light Source in Villigen (beamline X06SA) for help with the synchrotron experiments and Dr. Arnout Voet for help with CHARMM.

FUNDING

KU Leuven [to M.O.]; BioMacS, the KU Leuven Interfacultary Centre for Biomacromolecular Structure, KU Leuven (GOA Project to P.H., Impulse Project), the Hercules Foundation and FWO Flanders; project 'CEITEC - Central European Institute of Technology' [CZ.1.05/1.1.00/02.0068 to J.S. and J.E.S.]; European Regional Development Fund; Grant agency of the Czech Republic [P208/11/1822, P208/10/2302 and 203/09/1476 to J.S. and J.E.S.]. Funding for open access charge: KU Leuven.

Conflict of interest statement. None declared.

REFERENCES

1. Allart,B., Khan,K., Rosemeyer,H., Schepers,G., Hendrix,C., Rothenbacher,K., Seela,F., Van Aerschot,A. and Herdewijn,P. (1999) D-Altritol nucleic acids (ANA): hybridisation properties, stability, and initial structural analysis. *Chem. Eur. J.*, **5**, 2424–2431.
2. Abramov,M., Schepers,G., Van Aerschot,A., Van Hummelen,P. and Herdewijn,P. (2008) HNA and ANA high-affinity arrays for detections of DNA and RNA single-base mismatches. *Biosens. Bioelectron.*, **23**, 1728–1732.
3. Fisher,M., Abramov,M., Van Aerschot,A., Xu,D., Juliano,R. and Herdewijn,P. (2007) Inhibition of MDR1 expression with altritol-modified siRNAs. *Nucleic Acids Res.*, **35**, 1064–1074.
4. Fisher,M., Abramov,M., Van Aerschot,A., Rozenski,J., Dixit,V., Juliano,R.L. and Herdewijn,P. (2009) Biological effects of hexitol and altritol-modified siRNAs targeting B-Raf. *Eur. J. Pharmacol.*, **606**, 38–44.
5. Ovaere,M., Van Aerschot,A., Abramov,M., Herdewijn,P. and Van Meervelt,L. (2010) Crystallization and preliminary X-ray structure of the D-altritol oligonucleotide GTGTACAC. *Acta Cryst. F*, **66**, 460–462.
6. Leslie,A. (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint *CCP4+ESF-EAMCB Newslett. on Protein Crystallogr.*, **26.**
7. Evans,P.R. (2006) Scaling and assessment of X-ray data quality. *Acta Cryst. D*, **62**, 72–82.
8. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Cryst.*, **40**, 658–674.
9. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst. D*, **53**, 240–255.
10. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst. D*, **60**, 2126–2132.
11. Moews,P.C. and Kretsinger,R.H. (1975) Refinement of the structure of carp muscle calcium-binding parvalbumin by model building and difference Fourier analysis. *J. Mol. Biol.*, **91**, 201–225.

12. Bader,R.F.W. (1990) *Atoms in Molecules: A Quantum Theory*. Oxford University Press, Oxford.
13. Biegler-König,F.W., Schönbohm,J. and Bayles,D. (2001) AIM2000. *J. Comput. Chem.*, **22**, 545–559.
14. Biegler-König,F.W. and Schönbohm,J. (2002) Update of the AIM2000-Program for atoms in molecules. *J. Comput. Chem.*, **23**, 1489–1494.
15. Cossi,M., Rega,N., Scalmani,G. and Barone,V. (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.*, **24**, 669–681.
16. Barone,V. and Cossi,M. (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A*, **102**, 1995–2001.
17. Lu,X.-J. and Olson,W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.*, **3**, 1213–1227.
18. Olson,W.K., Bansal,M., Burley,S.K., Dickerson,R.E., Gerstein,M., Harvey,S.C., Heinemann,U., Lu,X.J., Neidle,S., Shakked,Z. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.
19. Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer, New York.
20. Tanaka,Y., Fujii,S., Hiroaki,H., Sakata,T., Tanaka,T., Uesugi,S., Tomita,K. and Kyogoku,Y. (1999) A′-form RNA double helix in the single crystal structure of r(ugagcuucggcuc). *Nucleic Acids Res.*, **27**, 949–955.
21. Labute,P. (2007) Protonate 3D: assignment of macromolecular protonation state and geometry. http://www.chemcomp.com/journal/proton.htm (April 2012, date last accessed).
22. Wood,P.A., Allen,F.H. and Pidcock,E. (2009) Hydrogen-bond directionality at the donor H atom—analysis of interaction energies and database statistics. *CrystEngComm*, **11**, 1563–1571.
23. Sponer,J., Jurecka,P. and Hobza,P. (2004) Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *J. Am. Chem. Soc.*, **126**, 10142–10151.
24. Sponer,J.E., Réblova,K., Mokdad,A., Sychrovský,V., Leszczynski,J. and Sponer,J. (2007) Leading RNA tertiary interactions: Structures, energies, and water insertion of a-minor and p-interactions. A quantum chemical view. *J. Phys. Chem. B*, **111**, 9153–9164.
25. Bailey,S. (1994) The CCP4 suite—programs for protein crystallography. *Acta Cryst*, **D 50**, 760–763.
26. Egli,M., Portmann,S. and Usman,N. (1996) RNA hydration: a detailed look. *Biochemistry*, **35**, 8489–8494.
27. Wahl,M.C. and Sundaralingam,M. (1999) A-DNA duplexes in the crystal. In: Neidle,S. (ed.), *Oxford Handbook of Nucleic Acid Structure*. Oxford University Press, New York, pp. 133–135.
28. Fraczkiewicz,R. and Braun,W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.*, **19**, 319–333.
29. Van Aerschot,A., Verheggen,I., Hendrix,C. and Herdewijn,P. (1995) 1,5-anhydrohexitol nucleic acids, a new promising antisense construct. *Angew. Chem. Int. Ed.*, **34**, 1338–1339.
30. Maier,T., Przylas,I., Strater,N., Herdewijn,P. and Saenger,W. (2005) Reinforced HNA backbone hydration in the crystal structure of a decameric HNA/RNA hybrid. *J. Am. Chem. Soc.*, **127**, 2937–2943.
31. Delano,W.L. (2002) The Pymol molecular graphics system. http://www.pymol.org (April 2012, date last accessed).
32. Halkier,A., Helgaker,T., Jørgensen,P., Klopper,W., Koch,H., Olsen,J. and Wilson,A.K. (1998) Basis-set convergence in correlated calculations on Ne, N_2 and H_2O. *Chem. Phys. Lett.*, **286**, 243–252.
33. Halkier,A., Helgaker,T., Jørgensen,P., Klopper,W., Koch,H. and Olsen,J. (1999) Basis-set convergence of the energy in molecular Hartree-Fock calculations. *Chem. Phys. Lett.*, **302**, 437–446.
34. Eichkorn,K., Treutler,O., Oehm,H., Haeser,M. and Ahlrichs,R. (1995) Auxillary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.*, **242**, 652–660.
35. Weigend,F. and Haeser,M. (1997) RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.*, **97**, 331–340.
36. Weigend,F., Haeser,M., Patzelt,H. and Ahlrichs,R. (1998) RI-MP2: optimized auxillary basis sets and demonstration of efficiency. *Chem. Phys. Lett.*, **294**, 143–152.
37. Jurečka,P., Nachtigall,P. and Hobza,P. (2001) RI-MP2 calculations with extended basis sets—a promising tool for H-bonded and stacked DNA base pairs. *Phys. Chem. Chem. Phys.*, **3**, 4578–4582.
38. Boys,S.F. and Bernardi,F. (1970) The calculation of small molecular interactions by the difference of separate total energies. Some procedures with reduced errors. *Mol. Phys.*, **19**, 553–566.
39. Šponer,J., Jurečka,P., Marchan,I., Luque,F.J., Orozco,M. and Hobza,P. (2006) Nature of base stacking: reference quantum-chemical stacking energies in ten unique B-DNA base-pair steps. *Chem. Eur. J.*, **12**, 2854–2865.
40. Hobza,P. and Šponer,J. (1996) MP2 and CCSD(T) calculations on H-bonded and stacked formamide…formamide and formamidine…formamidine dimers. *Theochem-J. Mol. Struct.*, **388**, 115–120.
41. Jurečka,P. and Hobza,P. (2002) On the convergence of the (Delta E-CCSD(T)-Delta E-MP2) term for complexes with multiple H-bonds. *Chem. Phys. Lett.*, **365**, 89–94.
42. Werner,H.-J., Knowles,P.J., Lindh,R., Manby,F.R., Schütz,M. *et al.* (2008) MOLPRO, version 2008.1, a package of ab initio programs, http://www.molpro.net (April 2012, date last accessed).
43. Hampel,C., Peterson,K.A. and Werner,H.-J. (1992) A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods. *Chem. Phys. Lett.*, **190**, 1–12.
44. Deegan,M.J.O. and Knowles,P.J. (1994) Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories. *Chem. Phys. Lett.*, **227**, 321–326.
45. Jurecka,P., Sponer,J., Cerny,J. and Hobza,P. (2006) Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.*, **8**, 1985–1993.