

# Solution structure of RNA duplexes containing alternating CG base pairs: NMR study of r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> under low salt conditions

Mariusz Popena, Ewa Biala, Jan Milecki<sup>1</sup> and Ryszard W. Adamiak\*

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland and <sup>1</sup>Faculty of Chemistry, Adam Mickiewicz University, Poznan, Poland

Received July 14, 1997; Revised and Accepted October 6, 1997

PDB coordinates 1PBM, 1PBL

## ABSTRACT

Structures of r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> have been determined by NMR spectroscopy under low salt conditions. All protons and phosphorus nuclei resonances have been assigned. Signals of H5'/5'' have been assigned stereospecifically. All <sup>3</sup>J<sub>H,H</sub> and <sup>3</sup>J<sub>P,H</sub> coupling constants have been measured. The structures were determined and refined using an iterative relaxation matrix procedure (IRMA) and the restrained MD simulation. Both duplexes form half-turn, right-handed helices with several conformational features which deviate significantly from a canonical A-RNA structure. Duplexes are characterised as having C3'-endo sugar pucker, very low base-pair rise and high helical twist and inclination angles. Helices are overwound with <10 bp per turn. There is limited inter-strand guanine stacking for CG steps. Within CG steps of both duplexes, the planes of the inter-strand cytosines are not parallel while guanines are almost parallel. For the GC steps this pattern is reversed. The 2'-O-methyl groups are spatially close to the 5'-hydrogens of neighbouring residues from the 3'-side and are directed towards the minor groove of 2'-O-Me(CGCGCG)<sub>2</sub> forming a hydrophobic layer. Solution structures of both duplexes are similar; the effect of 2'-O-methylation on the parent RNA structure is small. This suggests that intrinsic properties imposed by alternating CG base pairs govern the overall conformation of both duplexes.

## INTRODUCTION

Understanding the structure of functionally important RNA domains is essential for learning the principles of RNA folding, its catalytic properties and the specificity of RNA-protein interactions (1). In recent years, advances in heteronuclear and multi-dimensional NMR spectroscopy (2–4), enzymatic and chemical oligoribonucleotide synthesis (5,6) and both uniform (4,7) and regioselective (8) isotopic RNA labelling have resulted in the determination of RNA structures of considerable complexity.

In view of this, it is perhaps surprising that structures of short RNA duplexes containing alternating CG base pairs, a motif often found in RNA stems, has until now escaped full description. In 1984, a double helical A-type structure was proposed for the r(CGCGCG)<sub>2</sub> studied by high resolution NMR under low salt conditions (9,10). However, the <sup>1</sup>H NMR spectrum was only partly resolved at that time and the structure of this duplex has not been established. This was not only due to a typical severe overlap of the ribose signals for RNA (2,3,11) but also to a very pronounced effect of CG motif repetition in the structure. Our long term interest in the structure of RNA duplexes containing alternating CG base pairs was prompted by the finding of a left-handed double helix named Z-RNA (12,13). Slow kinetics of A- to Z-RNA transition in 6 M sodium perchlorate was advantageous in determining the NMR structure of the spectrally distinctive Z-form of r(CGCGCG)<sub>2</sub> (14). Recently, it was found that helicity reversal of the r(CG)<sub>n</sub> duplexes might also be promoted by high pressure (15). This was true only for r(CGCGCG)<sub>2</sub> but not for the very stable (*T*<sub>m</sub> = 76°C) 2'-O-Me(CGCGCG)<sub>2</sub> (16). Taken together, these observations suggested unique structural features of right-handed RNA helices containing alternating CG base pairs and prompted us to prepare modified (CG)<sub>n</sub> duplexes (17) and to study their structures in detail.

In this work we present NMR structures of two self-complementary RNA hexamers, r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, under low salt conditions. The length of the duplexes is the minimum which reveals the properties of CG and GC steps in RNA helices. This would not be possible for a shorter duplex, due to the effect of 'fraying ends'. On the other hand, a very pronounced effect of alternation in the structure of longer (CG)<sub>n</sub> duplexes would have made an accurate NMR analysis of unlabelled molecules very difficult. Both structures form right-handed, overwound mini helices with a number of structural features deviating from canonical A-RNA. Since this is the first report of the solution structure of a 2'-O-methylated RNA duplex, it allowed us to see the effect of 2'-O-methylation on RNA. This issue is relevant to the understanding of the properties of nuclease resistant RNA and was discussed earlier for the DNA/2'-O-Me(RNA) heteroduplex in solution (18) and the self-complementary A-DNA duplex containing single 2'-O-methyladenosine in the crystal (19). The effect of 2'-O-methylation on the parent

\*To whom correspondence should be addressed. Tel: +48 61 8 528503; Fax: +48 61 8 520532; Email: adamiakr@ibch.poznan.pl

$r(\text{CGCGCG})_2$  is not very pronounced—both duplex structures are governed by the intrinsic properties of the alternating CG base pairs.

The X-ray crystal structure of  $2'-O\text{-Me}(\text{CGCGCG})_2$  at 1.30 Å resolution, the highest to be reported in the RNA field, is presented in the accompanying paper.

## MATERIALS AND METHODS

### Nuclear magnetic resonance spectroscopy

**Oligoribonucleotide samples.** The hexamer  $r(\text{CGCGCG})$  was synthesised by the phosphotriester method in solution (12,13).  $2'-O\text{-Me}(\text{CGCGCG})$  was prepared by automated solid-phase synthesis using phosphoramidites as described (13). Oligoribonucleotides were speed vacuum dried from  $\text{D}_2\text{O}$  (99.8%) three times, dissolved in  $\text{D}_2\text{O}$  (99.98%) buffer containing 150 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA, and dried and dissolved in  $\text{D}_2\text{O}$  (99.98%, 0.6 ml). Final concentrations were 2 and 3 mM for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$  respectively. Exchangeable proton signals were measured from the samples in  $\text{H}_2\text{O}/\text{D}_2\text{O}$  9:1 v/v solution at pH 6.5.

**NMR spectra and data analysis.** Homonuclear 2D  $^1\text{H}$  NMR spectra of duplexes in  $\text{D}_2\text{O}$  and  $\text{H}_2\text{O}/\text{D}_2\text{O}$  were recorded using the phase-sensitive method (20) on a Varian Unity<sup>+</sup> 500 MHz spectrometer. Spectra were processed with VNMR (Varian) and FELIX/NMRchitect (Molecular Simulations) software on SUN and SGI Iris Indigo<sup>2</sup> workstations. Proton chemical shifts were referenced to tetramethyl ammonium chloride as internal standard (3.18 p.p.m.). 2D NOESY 500 MHz spectra of  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$  in  $\text{D}_2\text{O}$  were recorded at 30°C. Standard pulse sequence (21) was applied at three different mixing times  $\tau_m = 80, 150$  and 300 ms and spectra were acquired with 1K complex data points in the  $t_2$  and 1K real points in the  $t_1$  dimension, with spectral width set to 3.7 kHz. A relaxation delay of 2.5 s between transients was used. Phosphorus decoupling was turned on during mixing and acquisition periods. The residual HOD signal was removed from FIDs during data processing using the linear prediction method (22). FIDs were zero-filled to 1K complex points in the  $t_1$  dimension and phase shifted sine-bell in both directions were applied as digital filtration functions. After the first Fourier transformation, a base line correction was applied using a fifth order polynomial function. In all spectra,  $t_1$  ridge noise was attenuated by multiplying the first row by one-half (23) prior to the second Fourier transformation. Quantitative NOEs were determined by the integration of 2D NOESY spectra ( $\tau_m = 80$  and 150 ms) cross-peaks optimised with Lorentzian line shape function using the FELIX software (MSI). 2D NOESY experiments in  $\text{H}_2\text{O}/\text{D}_2\text{O}$  were acquired at 0 and 10°C; mixing time 200 ms, 4K data points in  $t_2$  and 360 points in  $t_1$ . The WATERGATE method (24) was used to suppress the water line. In addition, 1D proton spectra at 300 MHz were acquired using jump and return pulse sequence (25) at temperatures ranging from -3 to +95°C. DQF-COSY experiments (26) were measured with and without  $^{31}\text{P}$  decoupling at narrow spectral width (1.5 KHz) in both dimensions covering all the resonance frequencies of sugar protons. Spectra were collected with 980  $t_1$  FIDs of 2048 complex points in  $t_2$  from 48 scans each. Low power preirradiation of the residual HDO peak during the 2 s relaxation delay was applied. A squared  $\pi/8$  shifted sine-bell was applied as a function of digital

filtration. The data were zero filled to give  $4\text{K} \times 4\text{K}$  and  $2\text{K} \times 8\text{K}$  matrixes with final digital resolution in the  $\omega_2$  direction of 0.8 and 0.4 Hz/point, respectively. Coupling constants  $J_{1',2'}$  were estimated from analysis of line shapes in the 1D spectra. All other vicinal homonuclear couplings were determined from cross-peak patterns of  $^{31}\text{P}$  decoupled DQF-COSY spectra and their subsequent simulation using SPHINX/LINSHA programs (27). The coupling constants  $J_{1',2'}$ ,  $J_{2',3'}$  and  $J_{3',4'}$  of both the ribose and  $2'-O\text{-methylribose}$  rings were interpreted in terms of pseudorotation phase angles ( $P$ ) and puckering amplitudes ( $\Phi$ ) with the aid of the program PSEUROT 6.2 (28). The program also gave the population of major N-type (C3'-endo) conformers. Heteronuclear couplings  $^3J_{\text{P,H}}$  were determined by deconvolution of 1D  $^{31}\text{P}$  spectra and cross-peaks analysis and simulation (27) of both coupled and decoupled DQF-COSY spectra.  $^{31}\text{P}$  NMR spectra with and without proton decoupling and 2D heteronuclear  $^{31}\text{P}\text{-}^1\text{H}$  chemical shift correlation experiments (COLOC) (29) were recorded at 121.4 MHz. Phosphorus chemical shifts were measured using triethylphosphate as an internal standard and referenced to  $\text{H}_3\text{PO}_4$ .

### Structure determination, refinement and analysis

Characteristic features of NMR spectra for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , e.g. very small coupling constants  $J_{1',2'}$ , strong inter-residue NOEs  $\text{H}2'\text{-H}6/\text{H}8$  (10) and values of imino protons chemical shifts allowed us to build all initial structures within the A-family of right-handed helices. Two sets of 16 initial structures, one for  $r(\text{CGCGCG})_2$  and one for  $2'-O\text{-Me}(\text{CGCGCG})_2$ , were built using the InsightII/Biopolymer software (MSI). For  $r(\text{CGCGCG})_2$ , the set consisted of one duplex structure with parameters as in the fibre model of A-RNA (30) and 15 duplexes generated by varying helical twist (giving 7, 9, 11, 13 and 15 bp/turn) and rise (2.5, 2.8 and 3.1 Å) parameters. For  $2'-O\text{-Me}(\text{CGCGCG})_2$ , the set contained one duplex structure with parameters as in the fibre model of A-RNA but built from  $2'-O\text{-methyl-cytidine}$  and  $-\text{guanosine}$  templates, for which atom charges were calculated using a semiempirical method (MOPAC/Insight II software). A further 15 structures within the set of  $2'-O\text{-methylated}$  duplexes were generated in a similar manner as for  $r(\text{CGCGCG})_2$ . All 32 initial structures were optimised by energy minimisation using the Amber force field (31) and conjugate gradient algorithm.

Distance restraints were obtained from quantitative analysis of NOE cross-peaks using an iterative relaxation matrix approach (32) (IRMA/NMRchitect, MSI). In addition, 18 constraints defining Watson-Crick base pairing ( $1.9 \pm 0.3$  Å) were applied. Three to five cycles of iteration, including 5 ps rMD, were necessary to reach convergence to 'distance refined' structures as inspected by analysis of  $R$ -factors described in IRMA/NMRchitect module. The IRMA not only refines the distances but also narrows the restraint bounds, so that the range of uncertainties for distance constraints varied from  $\pm 5$  to  $\pm 20\%$ . The structures from last IRMA cycles were further refined by introduction of torsion angle constraints and subsequent rMD simulation. The majority of torsion angle constraints were obtained from analysis of coupling constants  $^3J_{\text{H,H}}$  (33) and  $^3J_{\text{P,H}}$  (34). The range of uncertainties for dihedral constraints of  $\text{H}1'\text{-C}1'\text{-C}2'\text{-H}2'$ ,  $\text{H}2'\text{-C}2'\text{-C}3'\text{-H}3'$  and  $\text{H}3'\text{-C}3'\text{-C}4'\text{-H}4'$  angles was  $\pm 10^\circ$ ; for others  $\pm 20^\circ$ . Additionally, NOE-derived restraints for the *anti* glycosyl angle  $\chi$  were introduced. The backbone angles  $\alpha$  and  $\zeta$

constraints, derived from the  $^{31}\text{P}$  NMR chemical shifts analysis, were intentionally omitted due to their low accuracy (7). In the refinement of  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , the following restraints were applied (data for  $2'-O$ -methylated duplex in square brackets): 234 [120] distance constraints and 102 [102] dihedral angle constraints. The mean number of constraints per residue was 28 and 18.5 for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , respectively. Force constants of 20 kcal/mol/Å<sup>2</sup> and 30 kcal/mol/rad<sup>2</sup> were applied for distance and torsion angles restraints, respectively. *In vacuo*, 100 ps rMD simulations (300 K, with reduced phosphate charges and distance-dependent dielectric constant  $\epsilon = r_{ij}$ ) and energy minimisation were conducted using the Amber force field (31) implemented in the DISCOVER ver. 95.0 program (MSI) on Cray Y-MPEL and J916 computers. For each duplex, 16 rMD trajectories were analysed every 1 ps to explore the conformational space. For each trajectory, the last 30 frames were subjected to conformational analysis. For both duplexes, 16 \* 30 structures within conformational space satisfy the NMR data (average r.m.s.d. 0.8 Å for each rMD trajectory from last 30 ps). The coordinates of the last 30 frames were averaged giving two sets of 16 very similar structures as characterised by low average r.m.s.d.s of 0.17 and 0.24 Å for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , respectively. Therefore, for both duplexes, the closely similar 16 structures were averaged and subjected to unrestricted minimisation (100 iterations, conjugate gradient method) to deliver a single final, averaged duplex structure (as deposited in the PDB) for both,  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ . The criterion of good stereochemistry was fulfilled for both final structures. R.m.s. deviations from ideal values (Amber forcefield) were: 0.010 Å for bond lengths and 2.3° for bond angles. Application of experimentally obtained constraints to the rMD calculation conducted with B-DNA, as initial structure, resulted in rapid reversal to the A-type helix. The precision of the final structures were confronted with the experimental data with use of two factors. The first, based on differences between calculated and experimental NOE intensities (sixth-root  $R_{\text{factor}}$ ), given in the IRMA protocol,

of 0.020 and 0.010 for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , respectively. This factor is about twice higher when canonical A-RNA form of these duplexes is considered. Second factor, based on differences between all measured and theoretical  $^3J_{\text{H,H}}$  and  $^3J_{\text{P,H}}$  (named  $J_{\text{rms}}$  indicator), was 1.16 Hz for both duplexes. Theoretical coupling constants were obtained using the modified Karplus equation (33,34). Helical parameters for both final structures and conformational space were evaluated using the program CURVES 5.11 (35). All conformational parameters were subjected to statistical analysis.

## RESULTS AND DISCUSSION

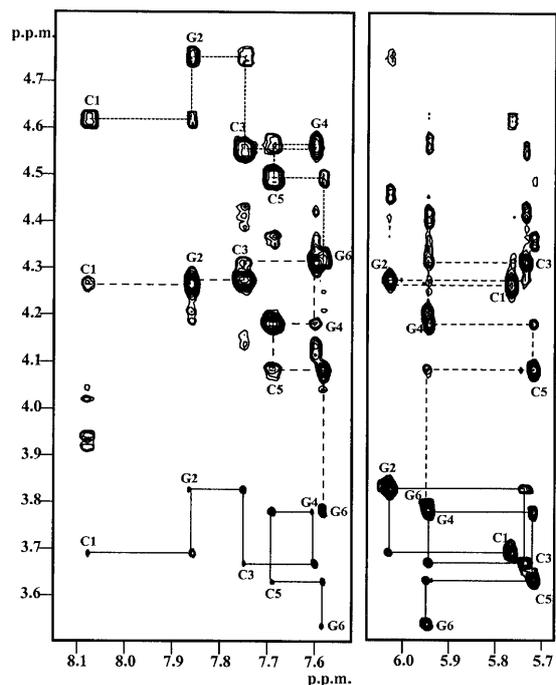
### NMR spectra analysis

*Assignments of non-exchangeable protons.*  $^1\text{H}$  NMR spectra of the  $2'-O\text{-Me}(\text{CGCGCG})_2$  were initially assigned as previously reported (36). The presence of  $2'-O$ -methyl groups in the duplex structure gave advantages (14,36) in the analysis of the parent  $r(\text{CGCGCG})_2$  spectra. Signals of H6/H8 and H1'/H5 protons were well separated from others of ribose and  $2'-O$ -methylribose residues in 1D  $^1\text{H}$  NMR spectra of both  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ . An analysis of 2D NOESY spectra in the H6/H8–H1'/H5 region was performed following the procedure for right-handed DNA/RNA double helices (10,11) based on probing intra- and inter-nucleotide connectivities between H1' and H6/H8. For  $r(\text{CGCGCG})_2$ , the results of H1'/H5 and H8/H6 assignment were the same as reported (9,10). The H1' resonances appear as singlets over a wide temperature range while H1' of G6 forms a doublet due to weak coupling with H2'. Small values of  $J_{1',2'}$  indicated the N-conformation for all sugar moieties in both duplexes. The chemical shifts of all aromatic protons (Table 1) and aromatic carbons (unpublished  $^{13}\text{C}$  NMR data) for  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$  were very similar. This points to structural similarity of the two duplexes in terms of stacking pattern. Due to the local influence of  $2'-O$ -methyl groups, the H1' of  $2'-O\text{-Me}(\text{CGCGCG})_2$  are deshielded  $\sim 0.15$  p.p.m., when compared to  $r(\text{CGCGCG})_2$ .

**Table 1.**  $^1\text{H}$  NMR chemical shifts  $\delta_{\text{H}}$  (p.p.m.) of  $r(\text{CGCGCG})_2$  (top) and  $2'-O\text{-Me}(\text{CGCGCG})_2$  (bottom) at 30°C in D<sub>2</sub>O

Residue	H6/H8	H5	H1'	H2'	H3'	H4'	H5'	H5''	OCH <sub>3</sub>	imino <sup>a</sup>	amino <sup>a</sup>
C1	8.03	6.00	5.53	4.57	4.58	4.32	4.02	3.93	–	–	8.31 7.00
G2	7.82	–	5.80	4.59	4.72	4.51	4.52	4.19	–	13.15	–
C3	7.72	5.31	5.55	4.55	4.56	4.45	4.57	4.15	–	–	8.49 6.77
G4	7.56	–	5.73	4.51	4.56	4.47	4.50	4.12	–	12.94	–
C5	7.58	5.25	5.50	4.34	4.46	4.39	4.53	4.07	–	–	8.40 6.85
G6	7.61	–	5.85	4.10	4.29	4.23	4.45	4.05	–	12.91	–
C1	8.08	6.05	5.76	4.27	4.61	4.27	4.03	3.93	3.69	–	8.34 7.02
G2	7.86	–	6.03	4.28	4.76	4.47	4.40	4.20	3.83	12.92	–
C3	7.75	5.33	5.73	4.32	4.56	4.43	4.40	4.15	3.67	–	8.68 6.79
G4	7.60	–	5.95	4.18	4.57	4.42	4.35	4.13	3.78	12.72	–
C5	7.69	5.27	5.73	4.08	4.50	4.36	4.37	4.08	3.63	–	8.59 6.85
G6	7.58	–	5.95	3.78	4.33	4.21	4.26	4.06	3.53	13.03	–

<sup>a</sup>Observed at 10°C on H<sub>2</sub>O/D<sub>2</sub>O 9:1 v/v.



**Figure 1.** Contour plots of the 2D NOESY spectrum (150 ms) of 2'-O-Me(CGCGCG)<sub>2</sub>. (Left) The 2'-O-methylribose to H6/H8 region with indicated pathways: OMe<sub>(i)</sub>-H6/H8<sub>(i)</sub>-OMe<sub>(i+1)</sub> (solid line); H2'<sub>(i)</sub>-H6/H8<sub>(i)</sub>-H2'<sub>(i+1)</sub> (dashed line); H3'<sub>(i)</sub>-H6/H8<sub>(i)</sub>-H3'<sub>(i+1)</sub> (dotted line). (Right) The 2'-O-methylribose to H1' region with indicated pathways: H1'<sub>(i)</sub>-OMe<sub>(i)</sub>-H1'<sub>(i+1)</sub> (solid line); H1'<sub>(i)</sub>-H2'<sub>(i)</sub>-H1'<sub>(i+1)</sub> (dashed line). Intra-residue cross-peaks of respective nucleotide residues are marked.

Difficulties arose for both duplexes in the assignment of resonances for H2', H3', H4', H5' and H5'' protons which lie close together in the 3.8–4.8 p.p.m. region. The H2' resonances of C1, G2, C3, G4 and C5 were assigned from NOESY spectra based on strong sequential inter-residue interactions between H6/H8<sub>(i)</sub> and H2'<sub>(i-1)</sub> characteristic of A-RNA helices (10). Previously, this approach allowed the assignment of H2' for r(CGCGCG)<sub>2</sub> (9). In addition, assignments of H2' signals were confirmed by analysis of cross-peak NOE intensities within the region H1'/H5-H2'/H3'/H4'/H5'/H5''. The NOEs between H1' and H2' were the strongest in this region. This is consistent with smaller H1'-H2' distances than H1'-H3' and H1'-H4' ones; assuming N-conformation for the sugars. The DQF-COSY spectrum in the latter region contained only one cross-peak resulting from a scalar coupling between H1'/H2' protons of G6.

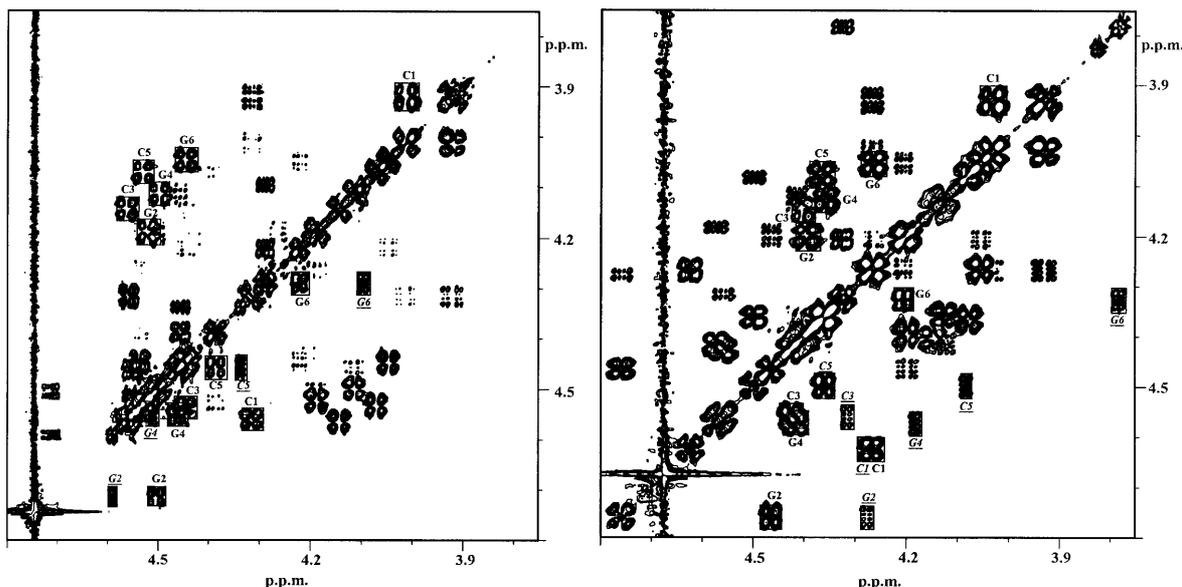
For r(CGCGCG)<sub>2</sub>, the 2D NOESY spectrum concerning H3' and H4' was very crowded making the search for both intra-nucleotide and sequential connectivities difficult and incomplete. Due to the effect of 2'-O-methylation, the H2' and H3' resonances of 2'-O-Me(CGCGCG)<sub>2</sub> were more dispersed (Table 1) allowing assignment of both intra- and inter-nucleotide connectivities (Fig. 1). Therefore, for both duplexes, the assignment of H3' and H4' signals was based on the analysis of H2'-H3' and H3'-H4' cross-peaks in the <sup>31</sup>P decoupled DQF-COSY spectra (Fig. 2).

For r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> all resonances of prochiral H5' and H5'' protons were assigned. For C1, G2, G4 and

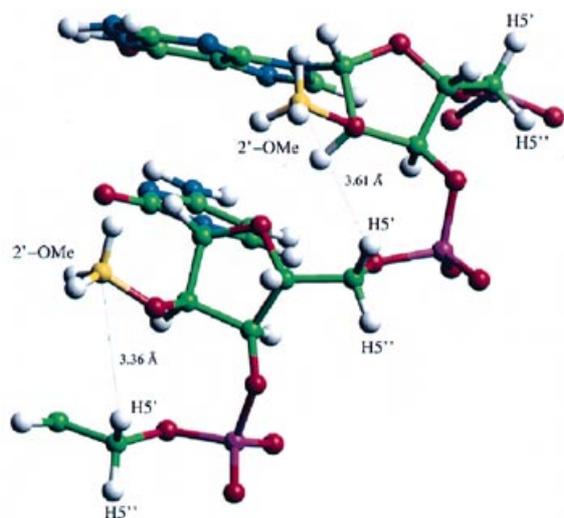
G6 residues, the DQF-COSY spectrum revealed correlation signals between H4'-H5' and H4'-H5''. No such correlations were observed for C3 and C5 due to overlap of the H4' and H5'' signals. Assignments of H5' and H5'' for the latter units were based on the analysis of weak correlation signals between H6/H8 and H5'' in the NOESY spectrum ( $\tau_m = 150, 300$  ms). Analysis of DQF-COSY spectra led to identification of six correlation signals for both duplexes reflecting geminal couplings H5'-H5''. The large effect of 2'-O-methylation on the H5' of G2, C3, G4, C5 and G6 residues, reflected by differences in the appropriate chemical shifts of both duplexes (Table 1), allowed stereospecific assignment of these protons. Due to the 2'-O-methylation the H5' signals are shifted up-field (~0.16 p.p.m.) whereas the chemical shifts of H5'' signals remain unchanged. This effect is caused by a close spatial contact between the 2'-O-methyl groups and the H5' protons of the 3'-end of the neighbouring residue of 2'-O-Me(CGCGCG)<sub>2</sub> (Fig. 3). Such a close contact is not seen for H5'' protons. For the C1 residue the stereospecific assignment of H5', H5'' was based on analysis of the NOESY cross-peaks of H3'-H5' and H3'-H5''. For 2'-O-Me(CGCGCG)<sub>2</sub>, 1D <sup>1</sup>H NMR spectra showed six well separated signals which were assigned from NOE sequential connectivities of 2'-O-methyls with H1' and with H6/H8 protons (Fig. 1). Sequential connectivities of 2'-O-methyls with H1' of 2'-O-methylated RNA strand have been also observed for the DNA/2'-O-Me(RNA) heteroduplex (18).

**Chemical shifts versus temperature profiles.** The already high thermal stability of r(CGCGCG)<sub>2</sub> (9,13) is further increased by the 2'-O-methylation. This is clearly seen from the analysis of chemical shifts versus temperature profiles. For both r(CGCGCG)<sub>2</sub> (5) and 2'-O-Me(CGCGCG)<sub>2</sub> duplexes, chemical shifts for H6/H8 of all but the C1 residues rise with temperature allowing estimation of  $T_m$  as 68 and 76°C, respectively. For 2'-O-Me(CGCGCG)<sub>2</sub>, all 2'-O-methyl groups, except G6, show the proton chemical shifts versus temperature profile as a plateau up to 45°C. In contrast to other proton signals, the chemical shifts for all 2'-O-methyl groups decrease from 45 up to 95°C. Most probably this is the effect of shielding of the 2'-O-methyl protons by the ring current of nucleobases (Fig. 3) upon duplex melting.

**Assignments of exchangeable protons.** Due to fast exchange processes with water, the signals of amino protons of guanines form very broad bands in the 7 p.p.m. region and are only visible at -3°C. These signals coalesce above 0°C. All signals of imino protons of guanines are sharp and well separated. Signals of amino protons of C1, C3 and C5 are much broader. They are shifted to 8.3–8.7 p.p.m., for N-H protons involved in hydrogen bond formation, and to 6.7–7.1 p.p.m. for those which are non-bonded (Table 1). Assignment of exchangeable imino protons of guanines and amino protons of cytosines for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> was accomplished by analysis of 2D NOESY spectra with  $\tau_m = 200$  ms at 0 and 10°C. In contrast to guanine imino proton signals, those for cytosine amino protons were seen at higher temperatures. With increase of temperature, three groups of hydrogen-bonded and no-bonded amino protons average to form three broad singlets observed up to the melting points of respective duplexes. Premelting of both duplexes was seen also in the broadening of all cytidine H5 and, to a lesser extent, H6 resonances. This shows that the exchange of cytosine amino protons with water is slow.



**Figure 2.** Ribose and 2'-*O*-methylribose proton regions of  $^{31}\text{P}$  decoupled DQF COSY spectra of:  $r(\text{CGCGCG})_2$  (left) and 2'-*O*-Me(CGCGCG) $_2$  (right); the H2'-H3' and H3'-H4' (underlined) correlation signals are marked right to the diagonal; H5'-H5'' are marked left of diagonal.



**Figure 3.** Structure of the single-stranded fragment of 2'-*O*-Me(CGCGCG) $_2$  showing a close spatial contact between 2'-*O*-methyl groups and H5' protons of the 3'-end neighbouring residues; an effect allowing for stereospecific assignment of H5' and H5''.

**Phosphorus nuclei assignments.**  $^{31}\text{P}$  NMR spectra of both duplexes show phosphorus nuclei signals in the narrow region of spectral widths  $<0.6$  p.p.m. This is typical for right-handed RNA duplexes (2,37). For  $r(\text{CGCGCG})_2$  and 2'-*O*-Me(CGCGCG) $_2$ , all five signals have been assigned and their  $^{31}\text{P}$  chemical shifts measured in the 5'  $\rightarrow$  3' direction are  $-0.72$ ,  $-1.12$ ,  $-0.78$ ,  $-1.26$ ,  $-0.83$  p.p.m. and  $-1.20$ ,  $-1.48$ ,  $-1.13$ ,  $-1.63$ ,  $-1.04$  p.p.m., respectively. In a previous report (9,10) on  $r(\text{CGCGCG})_2$ , the phosphorus signals were not assigned. Our assignment was based

on 2D heteronuclear  $^{31}\text{P}$ - $^1\text{H}$  chemical shift correlation (COLOC) spectra which clearly show strong P-H3' cross-peaks and weak signals corresponding to P-H5', P-H5'' and P-H4'. As expected (33), for both duplexes, chemical shifts for CpG and GpC steps are gathered in two groups, the first being offset to the lower field. Differences between  $^{31}\text{P}$  chemical shifts values are very small. This is in contrast to Z-RNA (12-14) for which much larger differences between CpG and GpC steps have been observed. All phosphorus signals of 2'-*O*-Me(CGCGCG) $_2$  were considerably shifted up-field compared to  $r(\text{CGCGCG})_2$ .

#### Structural features of the $r(\text{CGCGCG})_2$ and 2'-*O*-Me(CGCGCG) $_2$ under low salt conditions

**Conformation of ribose and 2'-*O*-methylribose rings.** The conformation of ribose and 2'-*O*-methylribose rings within both duplexes was studied applying a two-state model (38) and described in terms of N- and S-conformers. Analysis based on individual coupling constants  $J_{1',2'}$ ,  $J_{2',3'}$  and  $J_{3',4'}$  was done with the program PSEUROT 6.2 (28). Measured and calculated coupling constants and conformational parameters describing sugar pucker and fraction of N-conformers are presented in Table 2. As expected, the N-conformation dominates for all nucleoside units in both  $r(\text{CGCGCG})_2$  and 2'-*O*-Me(CGCGCG) $_2$ . For the terminal G6 in both duplexes, only a very small population of the S-conformer was detected. The C3'-endo sugar pucker is typical of A-RNA (39). Electronegative 2'-substitution including 2'-*O*-methylation leads to further C3'-endo pucker stabilisation (40,41). In the present study, most of the ribose rings of the parent  $r(\text{CGCGCG})_2$  are already stabilised in the C3'-endo conformation with elevated pucker amplitudes and hence the effect of 2'-*O*-methylation is much less pronounced than expected. This led to the overall similarity of the two duplexes. The puckering of ribose and 2'-*O*-methyl-ribose rings resulting from coupling constants analysis was confirmed

by the quantitative study of intra-sugar NOE connectivities between H1' and H2'/H3'/H4' protons. The integration of H1'–H2' cross-peaks was three to four times higher than H1'–H4' proving pure C3'-endo pucker for all residues. The rMD refinement protocol introduced no substantial changes to the experimentally derived sugar pucker parameters. Melting studies and above results show that end-effects, although measurable are very small if, especially in the case of the 2'-O-methylated analogue, not negligible.

**Glycosyl bond and backbone conformational analysis.** For r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, glycosyl bond torsion angles  $\chi$  were determined from analysis of NOEs. The low intensity of H1'–H6/H8 NOESY cross-peaks indicated an *anti* conformation for all nucleotide residues in both duplexes. The torsion angles  $\chi$  ranged from –154 to –177° and were quantitatively determined from intra-residue H6/H8–H3'/H2' NOEs. Spectral overlap of H2' and H3' signals for C1 and C3 residues in r(CGCGCG)<sub>2</sub> did not allow such a quantitative description. The backbone torsion angles  $\beta$ ,  $\gamma$  and  $\epsilon$  were determined from experimentally assessed coupling constants (33). Small homonuclear coupling constants  $J_{4',5'}$  ( $2.0 \pm 1$  Hz) and  $J_{4',5''}$  ( $2.5 \pm 1$  Hz) determined for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, indicate  $\gamma$  torsion angles ranging from 41 to 58°. Those are related to (+) *gauche* conformations. For the terminal ribo- and 2'-O-methyl C1 residues, coupling  $J_{4',5''}$  ( $3.7 \pm 1$  Hz) indicated the presence of a small fraction of the *trans* conformation. The heteronuclear coupling constants  $J_{PH5'}$ ,  $J_{PH5''}$ , with average value  $2.0 \pm 1$  Hz, were used to determine  $\beta$  torsion angles (178–183°). In addition, the presence of cross-peaks related to long-range couplings  $^4J_{PH4'}$  in COLOC spectra confirmed the (+) *gauche* for  $\gamma$  and *trans* for  $\beta$  torsion angles. Heteronuclear couplings  $J_{PH3'}$  ( $7.6$ – $8.8 \pm 1.0$  Hz) and ( $8.5$ – $9.6 \pm 1.0$  Hz) were measured for residues C1, G2, C3, G4 and C5 in r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, respectively. These data show that for both duplexes the *trans*

conformation is favoured with  $\epsilon$  values 201–209°. The  $\delta$  torsion angles were included in the evaluation of pseudorotation phase angles.

Experimentally determined torsion angles were compared with those resulting from the rMD refinement. For both duplexes, values for  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  and  $\chi$  describing final structures and those of conformational space are close to experimental values and to those of the A-type helices (Table 3). For the 2'-O-Me(CGCGCG)<sub>2</sub> there are, however, two exceptions concerning  $\alpha$  and  $\gamma$  torsion angles. The  $\gamma$  torsion angles are, on average, 20° higher than those of r(CGCGCG)<sub>2</sub> and of canonical A-RNA. This is compensated by a decrease of  $\alpha$  torsion angles. This effect is in accord with the anti-correlation principle concerning torsion angles  $\alpha$  and  $\gamma$  in the A-type helices (42). Most probably, higher  $\gamma$  angles are due to the influence of the 2'-O-methyl substituent which is spatially close to the H5' protons as indicated above. The  $\alpha$  torsion angles show lower values for GC than CG steps. These changes are also reflected by differences in the <sup>31</sup>P NMR chemical shifts; sensitive to conformational changes described by  $\alpha$  angles (2,4,37). The  $\zeta$  angles alternate more for CG than GC steps but to a lesser extent than for  $\alpha$ .

**Global structures and helical parameters.** The duplex structures of r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> (Fig. 4), are very similar overall with the root mean square deviation (r.m.s.d.) 1.0 Å. The similarity is even higher if only the four internal CG base pair core is superposed: r.m.s.d. 0.8 Å. This suggests that the intrinsic properties of alternating CG base pairs govern their overall conformation. Surprisingly, the effect of 2'-O-methylation on the parent RNA structure is relatively small. The r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> are A-type helices. The r.m.s.d. values referring for the initial structure of r(CGCGCG)<sub>2</sub>, built with parameters as in the fibre model of A-RNA (30), are 1.6 and 1.8 Å for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, respectively. This indicates that both structures deviate from canonical A-RNA.

**Table 2.** Experimental and calculated<sup>a</sup> constants  $J_{1',2'}$ ,  $J_{2',3'}$  and  $J_{3',4'}$ , sugar pucker parameters and percentage of N conformers for r(CGCGCG)<sub>2</sub> (top) and 2'-O-Me(CGCGCG)<sub>2</sub> (bottom)

Residue	Coupling constants $^3J_{H,H}$ (Hz)						Sugar pucker (°)		%N
	$J_{1',2'}$ exp.	calc.	$J_{2',3'}$ exp.	calc.	$J_{3',4'}$ exp.	calc.	P	$\Phi$	
C1	1.1 ± 0.2	1.1(0.2)	4.3 ± 1.0	4.6(0.6)	9.0 ± 1.0	9.0(0.8)	11(16)	41(6)	98(2)
G2	≤1.0	1.1(0.2)	4.3 ± 0.3	4.4(0.3)	9.3 ± 0.3	9.3(0.3)	13(9)	42(3)	98(2)
C3	≤1.0	1.1(0.2)	4.2 ± 1.0	4.4(0.4)	10.0 ± 1.0	9.5(0.5)	18(13)	44(5)	99(1)
G4	≤1.0	1.1(0.2)	4.3 ± 0.3	4.3(0.3)	9.6 ± 0.3	9.6(0.3)	19(8)	45(3)	98(2)
C5	≤1.0	1.1(0.2)	4.3 ± 0.3	4.3(0.3)	9.9 ± 0.3	9.8(0.2)	23(7)	46(3)	98(2)
G6	2.1 ± 0.2	2.1(0.2)	4.4 ± 0.3	4.4(0.3)	7.9 ± 0.3	7.9(0.3)	5(13)	42(3)	84(5)
C1	≤1.0	1.0(0.1)	4.3 ± 0.3	4.4(0.3)	9.1 ± 0.3	9.1(0.3)	5(12)	41(2)	99(1)
G2	≤1.0	1.0(0.1)	4.3 ± 0.3	4.3(0.3)	9.3 ± 0.3	9.2(0.3)	7(8)	42(3)	99(1)
C3	≤1.0	1.0(0.1)	3.8 ± 0.3	3.9(0.3)	9.9 ± 0.3	9.8(0.2)	17(8)	47(3)	99(1)
G4	≤1.0	0.9(0.1)	3.8 ± 0.3	3.9(0.3)	9.4 ± 0.3	9.7(0.3)	14(9)	47(2)	99(1)
C5	≤1.0	1.0(0.1)	4.2 ± 0.3	4.2(0.3)	9.8 ± 0.3	9.5(0.3)	12(8)	43(3)	99(1)
G6	1.7 ± 0.2	1.7(0.2)	4.9 ± 0.3	4.9(0.3)	7.9 ± 0.3	7.9(0.3)	3(12)	37(3)	90(4)

<sup>a</sup>With program PSEUROT 6.2 (28); standard deviations, from calculations using all permutations of possible coupling constants values, are given in parentheses.

**Table 3.** The  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\epsilon$  and  $\zeta$  backbone torsion angles<sup>a</sup> and  $\chi$  glycosidic angles for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> derived from the rMD trajectories<sup>b</sup>

Residue	$\alpha$		$\beta$		$\gamma$		$\epsilon$		$\zeta$		$\chi$	
	ribo	2'-O-Me	ribo	2'-O-Me	ribo	2'-O-Me	ribo	2'-O-Me	ribo	2'-O-Me	ribo	2'-O-Me
C1	–	–	–	–	51(7)	72(9)	195(4)	197(4)	296(5)	285(4)	218(5)	201(5)
G2	292(5)	277(9)	181(5)	180(4)	52(6)	70(7)	189(4)	203(4)	301(3)	291(5)	208(5)	202(4)
C3	281(4)	267(7)	189(4)	177(5)	49(4)	75(5)	193(4)	201(4)	290(4)	288(5)	208(5)	213(5)
G4	298(4)	262(8)	182(4)	179(5)	43(6)	79(6)	201(4)	208(4)	301(4)	294(4)	204(5)	207(5)
C5	266(7)	247(8)	175(4)	177(5)	81(5)	85(6)	200(4)	201(4)	300(4)	288(4)	204(5)	206(5)
G6	295(5)	274(8)	176(4)	173(4)	52(4)	72(6)	–	–	–	–	195(5)	191(5)
Mean	286	265	181	177	55	75	196	202	298	289	206	203
A-RNA <sup>c</sup>	294		186		49		202		294		202	
A-DNA <sup>c</sup>	285		208		45		178		313		206	

<sup>a</sup> $\alpha$  O5'  $\beta$  C5'  $\gamma$  C4'  $\delta$  C3'  $\epsilon$  O3'  $\zeta$  P.

<sup>b</sup>Standard deviations are given in parentheses.

<sup>c</sup>Refs 30,44.

**Table 4.** Selected helical parameters<sup>a</sup> for final structures of r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub> and those derived from the rMD trajectories<sup>b</sup>

Base pairs	x-Displacement (Å)				Inclination (°)				Propeller twist (°)			
	r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>		r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>		r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>	
	final	rMD	final	rMD	final	rMD	final	rMD	final	rMD	final	rMD
C1–G12	–3.9	–3.9(0.3)	–3.4	–3.3(0.4)	23	24(6)	23	21(5)	–33	–24(8)	–5	–2(11)
G2–C11	–4.0	–3.9(0.3)	–3.2	–3.1(0.4)	23	24(5)	26	24(5)	–23	–18(8)	–17	–11(8)
C3–G10	–4.0	–3.9(0.3)	–3.3	–3.2(0.3)	24	23(5)	28	25(5)	–29	–14(7)	–26	–19(8)
G4–C9	–4.0	–3.9(0.3)	–3.2	–3.3(0.3)	25	23(5)	30	28(5)	–29	–16(8)	–38	–28(8)
C5–G8	–4.0	–3.9(0.3)	–3.3	–3.2(0.3)	25	24(5)	24	20(5)	–26	–18(9)	–25	–16(8)
G6–C7	–3.9	–3.9(0.3)	–3.1	–3.0(0.4)	25	24(6)	20	17(5)	–33	–27(10)	–23	–25(8)

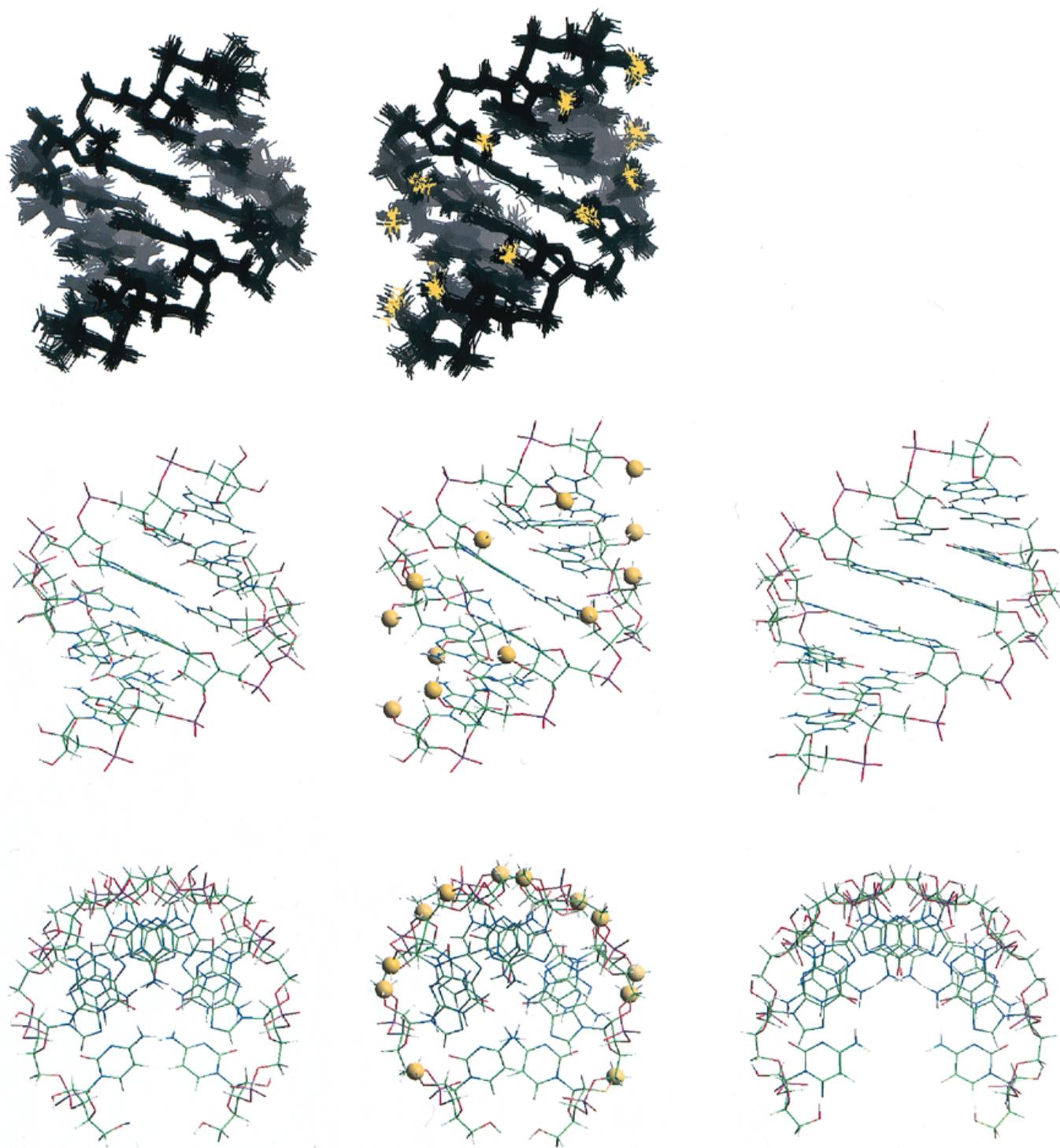
Base steps	Rise (Å)				Twist (°)				Roll (°)			
	r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>		r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>		r(CGCGCG) <sub>2</sub>		2'-O-Me(CGCGCG) <sub>2</sub>	
	final	rMD	final	rMD	final	rMD	final	rMD	final	rMD	final	rMD
C1–G2	2.1	2.1(0.4)	2.3	2.5(0.6)	37	37(3)	34	33(3)	2	6(4)	0	2(5)
G2–C3	2.6	2.5(0.3)	2.4	2.4(0.3)	37	36(4)	37	38(3)	–5	–5(4)	–3	–0(5)
C3–G4	2.0	2.3(0.3)	2.6	2.7(0.4)	39	44(4)	42	41(5)	2	1(5)	6	5(5)
G4–C5	2.5	2.6(0.3)	2.6	2.5(0.3)	38	36(3)	41	40(4)	–4	–7(6)	–6	–3(5)
C5–G6	2.1	2.1(0.4)	2.7	2.9(0.5)	38	34(4)	34	33(3)	3	8(4)	1	–1(4)

<sup>a</sup>Parameters were calculated with program CURVES (31).

<sup>b</sup>Standard deviations are given in parentheses.

To study the structure of duplexes in more detail, coordinates of the final structures and conformational space molecules were subjected to helical parameters calculation (35) and statistics (Table 4). In r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>, the Watson–Crick scheme of hydrogen bonds within CG base pairs is preserved, as expected. Typical for A-RNA the central hole viewed down the z-axis is relatively small (Fig. 4). This is consistent with the x-displacement parameters of –4 and –3.3 Å for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>. The r(CGCGCG)<sub>2</sub> is characterised by a very low helical rise which alternates from being as small as 2.0 Å for CG step and ~2.5 Å for the GC step. For 2'-O-Me(CGCGCG)<sub>2</sub> this parameter increases monotonically, reaching an average value 2.7 Å for the C5–G6 step, i.e., within the range typical for A-RNA. The helical twist angles are much higher than those in A-RNA and rise up to 42° in the centre of the 2'-O-Me(CGCGCG)<sub>2</sub> duplex. The duplexes form a little more than one half-turn of overwound helix. In both structures, a full-turn of

helix would contain only 9.6 bp. If terminal base pairs are excluded from parameterisation an average 9.4 and 9.0 bp per turn would be observed for r(CGCGCG)<sub>2</sub> and 2'-O-Me(CGCGCG)<sub>2</sub>. The inclination angles for r(CGCGCG)<sub>2</sub> show a high average value of 24°. They are even higher for 2'-O-Me(CGCGCG)<sub>2</sub>, reaching 30° in the core of the duplex. Analysis of the rMD trajectory shows that, due to base–base interactions, change in the inclination angle of a single base-pair induces concerted change of inclination in both neighbouring base-pairs (a 'domino effect') and two other parameters: helical twist and slide. An increase of the inclination and a helical twist led to an increase of the slide. The latter parameter also shows a tendency to alternate being smaller for CG steps. The propeller twist angles for both duplex base pairs are exceptionally high (Table 4) and vary so as to adjust the intra-strand stacking of guanines. In both cases, values of the roll angles are rather low and alternate. They are positive for CG steps, indicating their tendency to be open toward the minor groove, and negative

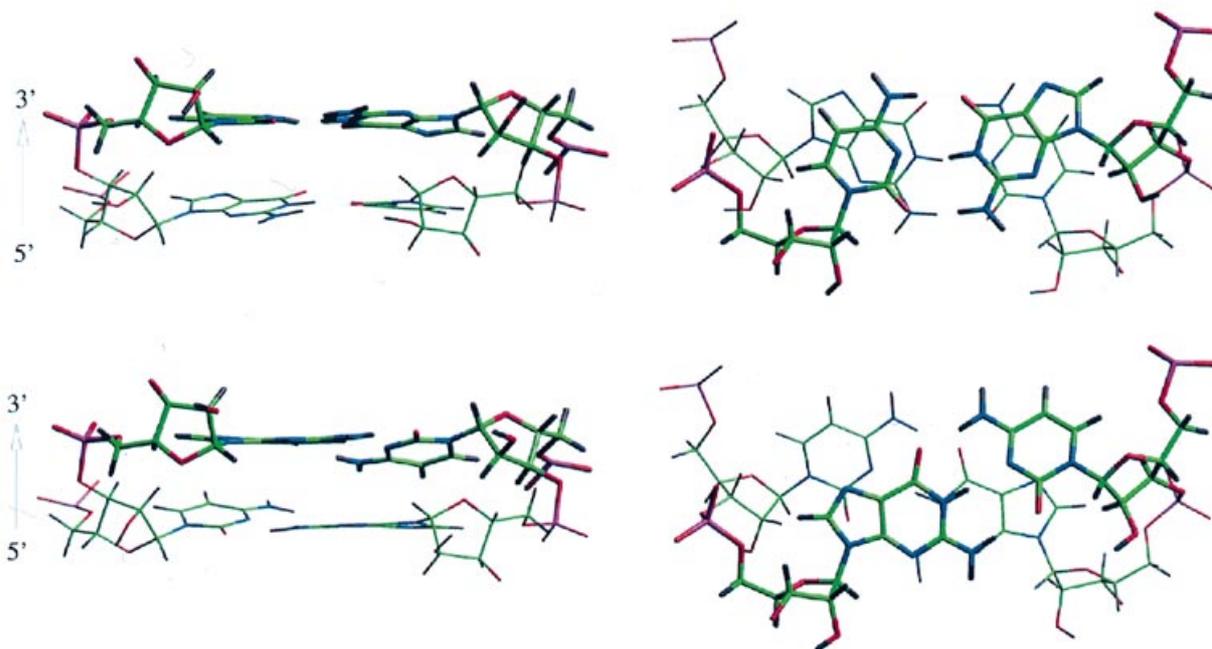


**Figure 4.** Top row: superposition of the last 30 frames representing one of 16 rMD trajectories for  $r(\text{CGCGCG})_2$  (left) and  $2'-O\text{-Me}(\text{CGCGCG})_2$  (right). Second and third rows: the overall views of the final structures of  $r(\text{CGCGCG})_2$  (left) and  $2'-O\text{-Me}(\text{CGCGCG})_2$  (middle) as viewed from the minor groove and along the helix axis. The  $2'-O\text{-Me}(\text{CGCGCG})_2$  carbon atoms of  $2'-O$ -methyl groups are marked (yellow, not in scale). The structure of  $r(\text{CGCGCG})_2$  built with parameters as in the fibre model of A-RNA is shown for comparison.

for GC steps, reflecting opening toward the major groove. A similar tendency was noted earlier for crystals of cytosine- and guanine-rich A-DNA duplexes (43).

For  $r(\text{CGCGCG})_2$  and  $2'-O\text{-Me}(\text{CGCGCG})_2$ , a striking, regular pattern of stacking has been found—different for CG and GC

steps (Fig. 5). Within CG steps of both duplexes, inter-strand cytosines are non parallel (with a high average dihedral angle  $40^\circ$  between their planes) while guanines are almost parallel. This is true even for terminal base pairs. For GC steps, guanines are non-parallel (on average  $25^\circ$ ) and cytosines show a tendency to



**Figure 5.** Stacking pattern within the CG and GC steps of  $r(\text{CGCGCG})_2$  as viewed from the minor groove and along the helix axis. For CG steps only limited inter-strand stacking of guanines is observed.

be parallel (on average  $14^\circ$ ). This conformational feature is even more pronounced for  $2'$ -*O*-Me(CGCGCG) $_2$  in the crystal (see the following paper). In CG steps there is only limited inter-strand stacking of guanines (Fig. 5), in contrast to a typical purine inter-strand stacking pattern observed in pyrimidine-purine steps of A-RNA helices (39).

Broad and shallow minor grooves are characteristic for both duplexes. The average minor groove widths based on distances between phosphorus atoms, when the sum of the van der Waals radii (5.8 Å) was subtracted, are 10.9 Å for  $r(\text{CGCGCG})_2$  and 11.5 Å for  $2'$ -*O*-Me(CGCGCG) $_2$ . This is in the range typical for A-RNA (39). Due to the short length of the duplexes, measurement of major groove width was not possible. Therefore, to overview the topography of major grooves the computer models of  $r(\text{CG})_{10}$  and  $2'$ -*O*-Me(CG) $_{10}$  helices were inspected. Models of two-turn helices were generated by a computer extension of four internal CG base pairs core from the NMR-derived coordinates. It appeared that the major grooves are exceptionally deep and narrow. We do not present detailed measurements here since the precision of such a model must be further evaluated.

Conformational features of RNA duplexes containing alternating CG base pairs have been studied on two duplexes: one native sequence  $r(\text{CGCGCG})_2$  and its  $2'$ -*O*-methylated analogue. It appeared that their low salt solution structures deviate significantly from a canonical A-RNA structure known from fibre diffraction data (30,44). The question arises: do their structural properties fall in the range typical for A-RNA helices or do they resemble other helices such as A'-RNA and A-DNA of the A-type genus? Values of parameters such as low rise and  $x$ -displacement and large helical twist and inclination ruled out similarities with poly(I)-poly(C) typical to the A'-RNA structure (30). Interestingly, structures of cytosine- and guanine-rich short, crystalline A-DNA duplexes are characterised by low inclination, 12 bp per turn and an average rise of 3.0 Å—parameters typical for A'-RNA (45,46). For  $2'$ -*O*-

Me(CGCGCG) $_2$ , conformational parameters such as helical twist, rise and inclination closely resemble that of the fibre model of calf-thymus A-DNA (44). In addition,  $\delta$  torsional angles, crucial for describing sugar pucker, with an average value of  $45^\circ$  for both duplexes were rather typical of A-DNA. With the limited number of cases evaluated, we are not yet in a position to claim to have found a sub-class in the A-RNA family. One might conclude that the canonical A-RNA form of both duplexes is within the envelope of converged structures (Tables 3 and 4) but only at low abundance. New RNA duplex structures must be solved to classify properties of this type of RNA duplexes.

*Effect of 2'-O-methylation.* Demand for both chemically- and enzyme-resistant hybridisation probes (antisense technologies) and  $2'$ -*O*-modified RNA to study mechanism of ribozymes action, resulted in a growing interest in  $2'$ -*O*-methylated oligoribonucleotides (47). The effect of the increasing  $T_m$  of the probe/target duplex structures was of additional benefit when using  $2'$ -*O*-methylated probes. However, only limited information is available on the structural effects of  $2'$ -*O*-methylation on RNA duplexes. The NMR study of DNA/ $2'$ -*O*-Me(RNA) hybrid duplex was reported earlier (18) but a refined model of the structure was not generated. The advantage of  $2'$ -*O*-methyl substituents when studying NMR spectra (18) was of importance to our work (36). In addition, the crystal structure of a self-complementary DNA duplex containing a single  $2'$ -*O*-methyladenosine residue has been reported (19). In the latter paper the authors attempted computer model building of a decamer duplex, made of all the four  $2'$ -*O*-methylribonucleoside units, based only on data from  $2'$ -*O*-methyladenosine. Our experiments, concerning model building of  $r(\text{CG})_{10}$  and  $2'$ -*O*-Me(CG) $_{10}$ , show that too much faith was given to that procedure. Both papers established a position for  $2'$ -*O*-methyl groups which point toward the minor groove of hybrid type duplexes.

The solution structure of 2'-O-Me(CGCGCG)<sub>2</sub> allows us for the first time to look at the effect of 2'-O-methylation on an RNA duplex. Since we have considered an RNA structure containing alternating CG base pairs of specific properties no general conclusions can be drawn concerning 2'-O-methylated RNA duplexes of random sequence. The effect of 2'-O-methylation on r(CGCGCG)<sub>2</sub> was clearly seen in the increase of T<sub>m</sub> from 68 to 76°C. Surprisingly, 2'-O-methylation of the parent r(CGCGCG)<sub>2</sub> has a smaller effect on the structure than expected. Further stabilisation of sugar C3'-endo puckering for end residues C1 and G6 was observed (Table 2). 2'-O-Methyl groups are spatially very close to the H5' of residues neighbouring from the 3'-side with an average distance 3.4 Å between carbon atoms of 2'-O-methyls and 5'-protons (Fig. 3). Their involvement in tuning the backbone conformation is manifested by increasing of γ angles up to 25° relative to r(CGCGCG)<sub>2</sub>. In addition, due to 2'-O-methylation, the x-displacement parameter is lowered and helical twist angles, inclinations and propeller twist are higher. The rise parameter is higher and there is no alternating of this parameter, in contrast to r(CGCGCG)<sub>2</sub>. The 2'-O-methyl groups are in close proximity to base pairs (Fig. 3) as indicated by NOEs between those groups and H6/H8 protons.

In the parent r(CGCGCG)<sub>2</sub>, the minor groove is broad enough to easily accommodate all ten 2'-O-methyl groups without much distortion. In 2'-O-Me(CGCGCG)<sub>2</sub>, the average distance between intra-strand 2'-O-methyl carbon atoms is 6.3 Å (r.m.s. 0.2 Å). This corresponds to the value 6.6 Å (r.m.s. 0.4 Å) reported for a computer model of the 2'-O-Me(RNA) decamer duplex (19). The distance between inter-strand 2'-O-methyl carbon atoms is 7.5 Å (r.m.s. 0.4 Å). This means that the distance between 2'-O-methyl groups across the minor groove is only 3.5 Å when subtracting 4 Å for the van der Waals radii for two methyl groups. The latter value is lower than that by the X-ray analysis of 2'-O-methylated A-DNA duplex (4.8 Å) and higher when estimated from model of 2'-O-Me(RNA) duplex (2.4 Å) (19). Closely spaced 2'-O-methyl groups resulted in the formation of a hydrophobic cushion within the minor groove of 2'-O-Me(CGCGCG)<sub>2</sub>; a feature confirmed by X-ray analysis of this duplex (see the accompanying paper).

## ACKNOWLEDGEMENTS

This work was supported by grants from the State Committee for Scientific Research, Republic of Poland (2P30300704 and 8T11F010008p08). Generous support of the Foundation for Polish Science and the Poznan Supercomputing and Networking Centre is acknowledged. We would like to thank Professors C.Altona, K.Wutrich and R.Lavery for kindly supplying us with programs PSEUROT 6.2, SPHINX/LINSHA and CURVES 5.11 respectively.

## REFERENCES

- Wyatt,J.R. and Tinoco,I.,Jr (1993) In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, New York, pp. 465–496.
- Varani,G. and Tinoco,I.,Jr (1991) *Q. Rev. Biophys.*, **24**, 479–532.
- Pardi,A. (1995) In James,T.L. (ed.), *Methods in Enzymology*. Academic Press, New York, Vol. **261**, pp. 350–383.
- Varani,G., Aboul-ela,F. and Allain,F.H.-T. (1996) *Prog. NMR Spectrosc.*, **29**, 51–127.
- Milligan,J.F. and Uhlenbeck,O.C. (1990) In Dahlber,J.E. and Abelson,J. (eds), *Methods in Enzymology*. Academic Press, New York, Vol. **180**, pp. 51–62.

- Usman,N., Ogilvie,K.K., Jiang,M.-J. and Cedergren,R.J. (1987) *J. Am. Chem. Soc.*, **109**, 7845–7854.
- Tolbert,T.J. and Williamson,J.R. (1996) *J. Am. Chem. Soc.*, **118**, 7929–7940.
- Glemarec,C., Kufel,J., Foldesi,A., Maltseva,T., Sandström,A., Kirsebom,L.A. and Chattopadhyaya,J. (1996) *Nucleic Acids Res.*, **24**, 2022–2035.
- Westerink,H.P., van der Marel,G.A., van Boom,J.H. and Haasnoot,C.A.G. (1984) *Nucleic Acids Res.*, **12**, 4323–4338.
- Haasnoot,C.A.G., Westerink,H.P., van der Marel,G.A. and van Boom,J.H. (1984) *J. Biomol. Struct. Dyn.*, **2**, 345–360.
- Wutrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York.
- Hall,K., Cruz,P., Tinoco,I.,Jr, Jovin,T.M. and Van de Sande,J.H. (1984) *Nature*, **311**, 584–586.
- Adamiak,R.W., Galat,A. and Skalski,B. (1985) *Biochim. Biophys. Acta*, **825**, 345–352.
- Davis,P.W., Adamiak,R.W. and Tinoco,I.,Jr (1990) *Biopolymers*, **29**, 109–122.
- Krzyzaniak,A., Barciszewski,J., Furste,J.P., Bald,R., Erdmann,V.A., Salanski,P. and Jurczak,J. (1994) *Int. J. Biol. Macromol.*, **16**, 159–162.
- Krzyzaniak,A., Salanski,P., Adamiak,R.W., Jurczak,J. and Barciszewski,J. (1996) In Hayashi,R. and Balny,C. (eds), *High Pressure Bioscience and Biotechnology*. Elsevier Science, Amsterdam, pp. 189–194.
- Biala,E., Milecki,J., Kowalewski,A., Popena,M., Antkowiak,W.Z. and Adamiak,R.W. (1993) *Acta Biochim. Pol.*, **40**, 1–7.
- Blommers,M.J.J., Pieleś,U. and de Mesmaeker,A. (1994) *Nucleic Acids Res.*, **22**, 4287–4194.
- Lubini,P., Zurcher,W. and Egli,M. (1994) *Chem. Biol.*, **1**, 39–45.
- States,D.J., Haberhorn,R.A. and Ruben,D.J. (1982) *J. Magn. Reson.* **48**, 286–292.
- Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4593.
- Barkhuisen,H., de Beer,R., Bovee,W.M.M.J. and van Ormondt,K. (1985) *J. Magn. Reson.* **61**, 465–481.
- Otting,G., Widmer,H., Wagner,G. and Wutrich,K. (1986) *J. Magn. Reson.* **66**, 187–193.
- Sklénar,V., Piotto,M., Leppik,R. and Saudek,V. (1993) *J. Magn. Reson. A*, **102**, 241–250.
- Plateau,P. and Gueron,M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
- Piantini,U., Sørensen,O.W. and Ernst,R.R. (1982) *J. Magn. Reson.*, **48**, 286–292.
- Widmer,H. and Wütrich,K. (1986) *J. Magn. Reson.* **70**, 270–279.
- de Leeuw,F.A.A.M. and Altona,C. (1983) *J. Comp. Chem.*, **4**, 428–437.
- Kessler,H.J., Griesinger,C., Zarbock,J. and Loosli,H.R. (1984) *J. Comp. Chem.*, **5**, 331–336.
- Arnott,S., Hukins,D.W.L., Dover,S.D., Fuller,W. and Hodgson,A.R. (1973) *J. Mol. Biol.*, **8**, 107–122.
- Weiner,S.J., Kollman,P.A., Nguyen,D.T. and Case,D.A. (1986) *J. Comp. Chem.*, **7**, 230–252.
- Boelens,R., Koning,T.M.G., van der Marel,G.A., van Boom,J.H. and Kaptein,R. (1989) *J. Magn. Reson.*, **82**, 290–308.
- Haasnoot,C.A.G., de Leeuw,F.A.A.M. and Altona,C. (1980) *Tetrahedron Lett.*, **21**, 2783–2792.
- Lankhorst,P.P., Haasnoot,C.A.G., Erkelens,C. and Altona,C. (1984) *J. Biomol. Struct. Dyn.*, **1**, 1387–1405.
- Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.
- Popena,M., Milecki,J., Biala,E. and Adamiak,R.W. (1995) *Nucleosides Nucleotides*, **14**, 983–984.
- Gorenstein,D.G. (1984) *Phosphorus-31 NMR: Principles and Applications*. Academic Press, New York.
- Altona,C. (1983) *Recl. Trav. Chim. Pays-Bas*, **101**, 413–433.
- Saenger,W. (1984) *Principles of Nucleic Acids Structure*. Springer, Berlin, Heidelberg.
- Uesugi,S., Miki,H., Ikehara,M., Iwahashi,H. and Kyogoku,Y. (1979) *Tetrahedron Lett.*, 4073–4076.
- Guschlbauer,W. and Jankowski,K. (1980) *Nucleic Acids Res.*, **8**, 1421–1433.
- Shakked,Z., Rabinovitch,D., Kennard,O., Cruse,W.B.T., Salisbury,S.A. and Viswamitra,M.A. (1983) *J. Mol. Biol.*, **166**, 183–201.
- Tippin,D.B. and Sundaralingam,M. (1997) *Biochemistry*, **36**, 536–543.
- Chandrasekaran,R. and Arnott,S. (1989) In Saenger,W. (ed.), *Landolt-Bornstein, New Series, Group VII*. Springer, Berlin, Vol. **1b**, pp 31–170.
- Kennard,O. and Hunter,W.W. (1989) In Saenger,W. (ed.), *Landolt-Bornstein, New Series, Group VII*. Springer, Berlin, Vol. **1a**, pp 255–360.
- Heinemann,U. (1991) *J. Biomol. Struct. Dyn.*, **8**, 801–811.
- Sproat,B.S. (1995) *J. Biotechnol.* **41**, 221–238.