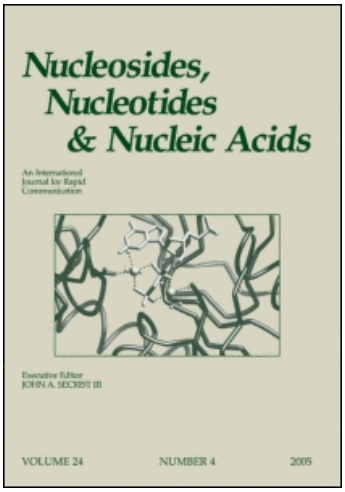

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Crystal and Solution Structure of 5'-O-(Guanosine-2'-O-Phosphonomethyl)Cytidine, an Isopolar Nonisosteric Phosphonate Analog Of GpC

J. Zachová[a]; I. Císařová[b]; M. Buděšíns[c]; R. Liboska[c]; Z. Točík[c]; I. Rosenberg[c]
[a] Institute of Physics, Charles University, Prague 2 [b] Department of Inorganic Chemistry, Charles University, Prague 2 [c] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague 6, Czech Republic

To cite this Article Zachová, J. , Císařová, I. , Buděšíns, M. , Liboska, R. , Točík, Z. and Rosenberg, I.(1999) 'Crystal and Solution Structure of 5'-O-(Guanosine-2'-O-Phosphonomethyl)Cytidine, an Isopolar Nonisosteric Phosphonate Analog Of GpC', Nucleosides, Nucleotides and Nucleic Acids, 18: 11, 2581 — 2599
To link to this Article: DOI: 10.1080/07328319908044628
URL: http://dx.doi.org/10.1080/07328319908044628

NUCLEOSIDES & NUCLEOTIDES, 18(11&12), 2581–2599 (1999)

CRYSTAL AND SOLUTION STRUCTURE OF 5'-O-(GUANOSINE-2'-O-PHOSPHONOMETHYL)CYTIDINE, AN ISOPOLAR NONISOSTERIC PHOSPHONATE ANALOG OF GpC.

J.Zachová,[a] I.Císařová,[b] M.Buděšínský,[c] R.Liboska,[c] Z.Točík[c] and I.Rosenberg[c]*

[a]*Institute of Physics, Charles University, Ke Karlovu 5, 12116 Prague 2*
[b]*Department of Inorganic Chemistry, Charles University, Hlavova 2030, 12840 Prague 2*
[c]*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo 2, 16610 Prague 6, Czech Republic*

ABSTRACT: X-Ray crystal structure of 5'-O-(guanosine-2'-O-phosphonomethyl)-cytidine (G-p_cC(2'-5')) was determined. The structural unit involves two molecules of G-p_cC(2'-5'), differing both in the conformation of the phosphonate internucleotide linkage and in the intramolecular base-plane distances (3.6 Å; 3.3 Å), and fourteen molecules of water. The crystal data of G-p_cC(2'-5') are discussed along with those obtained from NMR spectra.

Introduction

It has been widely recognized that any successful *in vivo* use of synthetic oligonucleotides as potential specific and selective therapeutics[1] depends on the stability of these compounds towards nucleases. Modification of the internucleotide linkage[2] of oligonucleotides, therefore, comes into consideration as the logical and efficient way how to modulate resistance against nuclease cleavage. On the other hand, however, the modifications of the internucleotide linkage are mostly connected with influencing the hybridization properties of such altered oligonucleotides because of the subsequent conformational changes of the sugar-phosphate backbone.[3]

Among the modified oligonucleotides, the dinucleoside monophosphate analogs represent the shortest possible oligonucleotide mimics possessing two principal attributes of their longer congeners, the base-stacking and base-pairing. Because of their

chemical availability, these "dimers" are suitable models for the study of the physicochemical properties acquired as a consequence of the modification of the internucleotide linkage and/or other changes in the molecule.[4-7]

Within the framework of our systematic investigation of phosphonate-based nucleotides and oligonucleotides of a high resistance towards nucleases, we have synthetized analogues of diribonucleoside monophosphates differing in the position of a bridging methylene group introduced into the sugar-phosphate backbone, featuring both the 2',5'- and 3',5'-isomers. These compounds have been investigated by Raman and 1D NMR spectroscopy revealing conformational similarity to the natural diribonucleoside monophosphates.[8,9] A base-stacking comparable with that of natural dinucleoside mono-phosphates was found and confirmed by a Raman spectroscopy study.[8] The existence of the strong intramolecular base-stacking interactions was unambiguously proved by measuring the stability of the complexes of polyU with a series of phosphonate ApA analogues . Despite the presence of a longer internucleotide linkage, we have found that the complexes have excellent thermal stability (as measured by T_m values).[10,11]

In order to examine in more detail the influence of the inserted methylene group in the modified internucleotide linkage on the sugar-phosphate backbone and the nucleo-base conformation and intramolecular base-stacking, we have studied these compounds by NMR techniques and also by X-ray diffraction analysis, as we have succeeded in the crystallization of the first and so far the only of the four possible regioisomers of GpC analog; no modified ApA dimer from this series has been obtained in crystalline form until now. Thus, we report here the X-ray crystal structure of 5'-O-(guanosine-2'-O-phosphonomethyl)cytidine [G-p_cC (2'-5')] together with a detailed study of the solution conformation. The data obtained are compared with those of the natural congener, guanylyl-2',5'-cytidine dihydrate (GpC (2'-5').[12]

Material and methods

The 5'-O-(guanosine-2'-O-phosphonomethyl)cytidine was synthetized as previously reported[13] Crystals were grown using the method of sitting drop, by slow evaporation of aqueous solution of G-p_cC(2'-5') in the presence of magnesium chloride (mole ratio G-p_cC(2'-5') : Mg^{2+} = 2 : 1; pH = 3.6). The presence of magnesium ions seemed to be necessary for successful crystallization, although they were not found in the crystal

TABLE 1. Crystallographic data for G-p$_c$C(2'-5').

Formula	$C_{40}H_{84}N_{16}O_{38}P_2$
Molecular weight	1459.17
Temperature	293(2) K
Space group	Monoclinic, $P2_1$ (No.4)
Cell dimensions	a = 12.4770(1) Å ; b = 13.050(2) Å; c = 20.381(2) Å ; β = 105.319(8)°
V	3200.5(6) Å^3
Z	2
D_{calc}	1.514 g cm^{-3}
μ	0.180 mm^{-1}
2 range	1.04 to 19.98 °
hkl ranges	−11→11, 0→12, 0→19
Diffractions collected / unique	3257 / 3155 [$R(int)^a$ = 0.036]
Diffraction observed; I>2Φ(I)	2293
Number of parameters	395
$GOF_{all}{}^a$	1.034
$R_{obs}(F)$; $R_{all}(F)^a$	0.0785; 0.1328
$wR_{obs}(F^2)$; $wR_{all}(F^2)^a$	0.2020; 0.2326
Δρ, e Å^{-3}	0.671 ; -0.409

^{a} $R(F) = \Sigma(||F_o|-|F_c||) / \Sigma|F_o|$, $wR(F^2)^b = [\Sigma(w(F_o^2 - F_c^2))^2 / \Sigma(w(F_o^2)^2]^2$, GOF = $[\Sigma(w(F_o^2 - F_c^2))^2 / (N_{diffrs} - N_{params})]^2$, $R(int) = [\Sigma|F_o^2 - F_o^2(\text{mean})| / [\Sigma F_o^2$

^{b} weighting scheme used: $w = [\sigma^2(F_o^2) + 0.1446\,P^2 + 9.42\,P]^1$; $P = [\max(F_o^2) + 2F_c^2] / 3$

structure. The obtained crystals were pentagonal, colorless plates with maximal dimension of 0.4 x 0.25 x 0.15 mm, and were measured at 293(2)K on a CAD4-MACHIII diffractometer with graphite-monochromated MoKα radiation (λ=0.71073A). The cell parameters were determined from 25 reflections in the 10–11° θ range. The intensities of reflections were measured by the θ–2θ scan and corrected on LP-factor; the absorption was neglected. Three standard reflections monitored every 1h, showed intensity variation of ±2%. The structure was solved by direct methods (SHELXS86)[14] and refined[15] by SHELXL97 using full-matrix least-square procedure based on F^2. The refinement was isotropic for all atoms except phosphorus. Hydrogen atoms were fixed in theoretical positions and given temperature factors of their bonding atoms multiplied by 1.5. The crystal and refinement parameters are summarized in Table 1.

NMR spectra of G-p$_c$C(2'-5') were measured on Varian UNITY-500 spectrometer (^{1}H at 500 MHz; ^{13}C at 125.7 MHz frequency) in D_2O (5 mg of sample in 0.6 ml

solution; pH = 3.6 adjusted by addition of phosphoric acid) at 20°C. 1D-Proton NMR spectra were taken at different temperatures in the range of 4–40°C. Chemical shifts and coupling constants of ribofuranose hydrogens, extracted from resolution enhanced spectra, were refined by spin simulation-iteration procedure using the program SPINI based on LAOCOON 3 (ref. 16). Proton double quantum filtered 2D-COSY spectrum[17,18] was recorded with a spectral width 5000 Hz, 512 time-increments and zerofilling to 2k x 2k data matrix before FT. Phase sensitive 2D-ROESY spectrum [19,20] was taken under similar conditions using mixing time 300 ms. A natural abundance ^{1}H-^{13}C heterocorrelated spectrum was recorded in the inverse mode (2D-HMQC pulse sequence[21,22]) with spectral width of 3000 Hz in t_2- and 24000 Hz in t_1-dimensions, 16 transients for each of 512 time-increments and zerofilling to 2k x 2k data matrix before FT. Carbon-13 chemical shifts and coupling constants, J(C,H) and J(C,P) were obtained by standard procedures.

The numbering of atoms in the ribofuranose ring and in the guanine and cytosine moiety as well as definition of torsion angles are based on the IUPAC/IUB recommendation.[23] The presence of an extra methylene group inserted into the internucleotide linkage leads to the introduction of additional torsion angle α' and to the modified definition of some other angles in G-p_cC(2'-5') (see Fig. 1). A slightly different labelling of atoms is used in the discussion of X-ray crystal structure in order to distinguish the atoms of the two independent molecules present in the unit cell (Fig. 2).

Results and discussion

X-Ray crystal structure. Perspective view of the molecular structure with atom labelling is shown in Fig. 2. Structural unit involves two symmetry-independent molecules of G-p_cC(2'-5') in the form of zwitterions protonated at nitrogen atoms N3 of cytosine and deprotonated at the phosphonate group, and 14 molecules of H_2O. The bond lengths and angles of both cytosine and guanosine are unexceptional and their values correspond to the data of GpC(2'-5').[24] Although no hydrogen atom was detected near the N3 of cytosines, the acidity of the crystallization solution (pH ~ 3.6), the geometry parameters (very short intermolecular distance N3...O3) and the angle C2-N3-C4 = 125° are in a good agreement with the published data on GpC(2'-5') in zwitterionic form[24] and other protonated structures.[25]

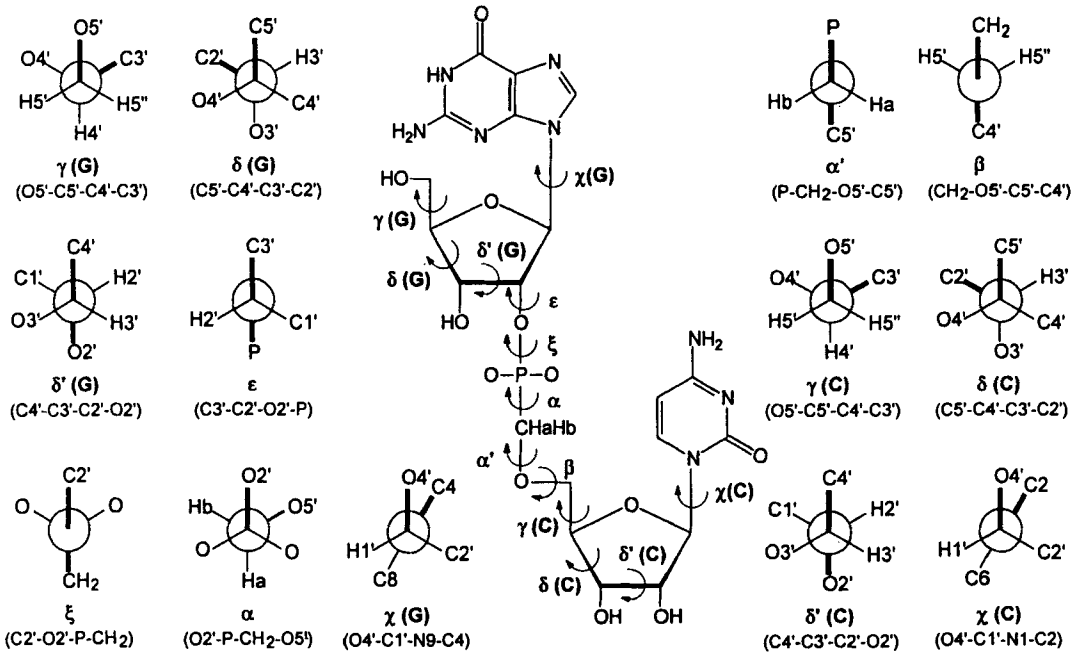

FIG. 1 Definition of torsion angles in G-p$_C$C(2′-5′).

The ribose puckering in both molecules of G-p$_c$C(2'-5') is very similar and close to that found[26] in GpC(2'-5'). Both guanosines are in *syn* (χ = 51° and 44°) and C2'-*endo* conformation while both cytidines exhibit *anti* (χ = −150° and −161°) and C3'-*endo* conformation. The possible influence of 2'-5' and 3'-5' linkage on the furanose puckering and *syn/anti* orientation of the base has been discussed.[12] No preferred conformation has been observed within the set of natural dinucleotide monophosphates with both types of linking (*e.g.*, the furanose puckering in CpA(3'-5') and CpA(2'-5') is C3'-*endo*/C2'-*endo*, in ApU(2'-5') it is C2'-*endo*/C3'-*endo* and in dCpG(3'-5') it is C2'-*endo*/C2'-*endo*). The known preference of the C2'-*endo* form to adopt *syn*-orientation[26] is confirmed also in the presented structure of G-p$_c$C(2'-5'). Conformation around C3'-C4' bond is also in agreement with Krishnan's and Seshadri's hypothesis[24] that the 2'-5' and 3'-5' purine-pyrimidine self-complementary structures differ essentially in the C3'-C4' bond con-formation (δ_1 and δ_2 values for the 2'-5' structures are < 0 while for the 3'-5' ones, they are > 0). We can conclude, therefore, that the conformation of nucleoside

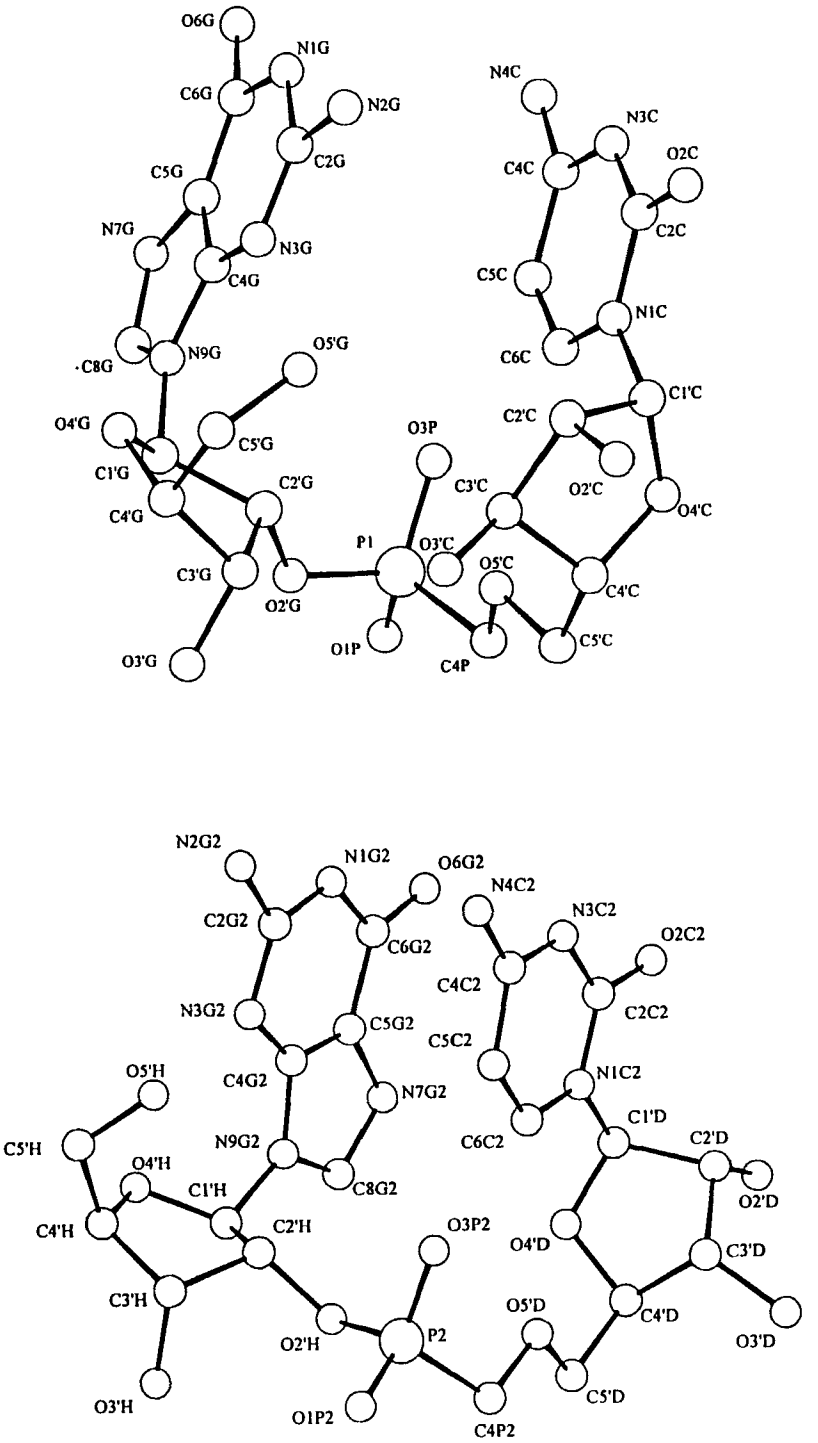

FIG. 2 Perspective view and numbering for "open" and "closed" structure of G-p$_C$C(2′-5′).

TABLE 2. Comparison of geometry parameters for G-p$_c$C(2′-5′) and GpC(2′-5′) in the crystal and in solution.

	Angle	Crystal (X-ray)			Solution (NMR)	
		G-p$_c$C(2′-5′)		GpC(2′-5′)	G-p$_c$C(2′-5′)	
		"open"	"closed"		N-type	S-type
Furanose (G)	Conf.	C2'-*endo*	C2'-*endo*	C2'-*endo*	C3'-*endo*	C2'-*endo*
	P	163	169	162	18 (0.44)	150 (0.56)
	ϕ_m	38.6	40.3	38	36	36
Furanose (C)	Conf.	C3'-*endo*	C3'-*endo*	C3'-*endo*	C3'-*endo*	C2'-*endo*
	P	9	11	19	0 (0.80)	150 (0.20)
	ϕ_m	39.7	40.2	40	38	36
O4'-C1'-N9-C4	χ(G)	51	44	54	*syn-* (37)	
O4'-C1'-N1-C2	χ(C)	-150	-161	-156	*anti-* (-176)	
O5'-C5'-C4'-C3'	γ(G)	53	55	54	γ^+(0.66), γ^t(0.24), γ^-(0.10)	
C5'-C4'-C3'-C2'	δ(G)	-96	-94	-96	-154 [a]	-105 [a]
C4'-C3'-C2'-O2'	δ'(G)	-152	-155	-154	-86 [b]	-151 [b]
C3'-C2'-O2'-P	ε	-117	-86	-97	-105	
C2'-O2'-P-C(P)	ξ	77	-165	-59 [c]	*d*	
O2'-P-C(P)-O5'	α	-63	70	-75 [c]	*d*	
P-C(P)-O5'-C5'	α'	174	-150	-	*trans-* (±160)	
C(P)-O5'-C5'-C4'	β	163	-177	-176 [c]	*trans-* (~180)	
O5'-C5'-C4'-C3'	γ(C)	51	58	54	γ^+(0.93), γ^t(0.02), γ^-(0.05)	
C5'-C4'-C3'-C2'	δ(C)	-157	-157	-159	-154 [a]	-96 [a]
C4'-C3'-C2'-O2'	δ'(C)	-77	-75	-76	-82 [b]	-155 [b]

[a] calculated using relation $\delta = \nu_3 - 120°$; [b] calculated using relation $\delta = \nu_2 - 120°$; [c] in GpC(2′-5′) these angles are defined differently: ξ_1 = C2'-O2'-P-O; α = O2'-P-O5'-C5'; β = P-O5'-C5'-C4'; [d] NMR data for estimation ξ_1 and α are not accessible.

moieties in G-p$_c$C(2'-5') corresponds to that published for natural diribonucleoside monophosphates.

Significant differences between the two structures of G-p$_c$C(2'-5') in the cell unit were found for the sugar-phosphonate backbone conformation (Table 2). The "open" structure adopts the gg^-t conformation with a long intramolecular base-plane distance (3.602 Å), an interplane angle of 11°, and a short distance between neighboring sugar carbon atoms (guanosine C2' ... cytosine C5' = 4.4 Å). The "closed" structure adopts *tgt* conformation with a short interplane distance (3.298 Å), a small angle between base-planes (5°), and a long distance between guanosine C2' and cytosine C5' (5.2 Å). For comparison, Saenger gives [26] an average intramolecular base-plane distance of ~ 3.4 Å.

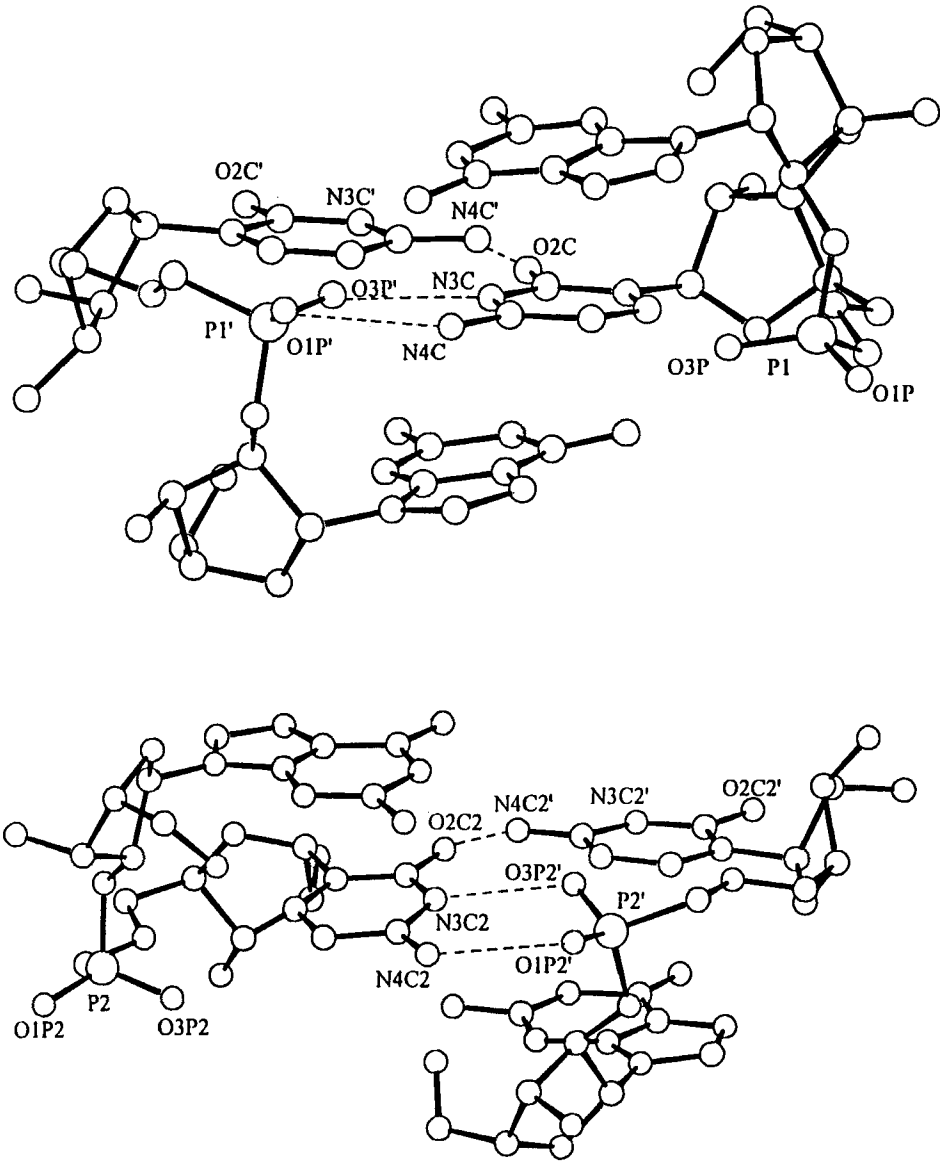

FIG. 3 Hydrogen bonds (dashed lines) connecting two "open" and two "closed" structures of G-p_CC(2′-5′).

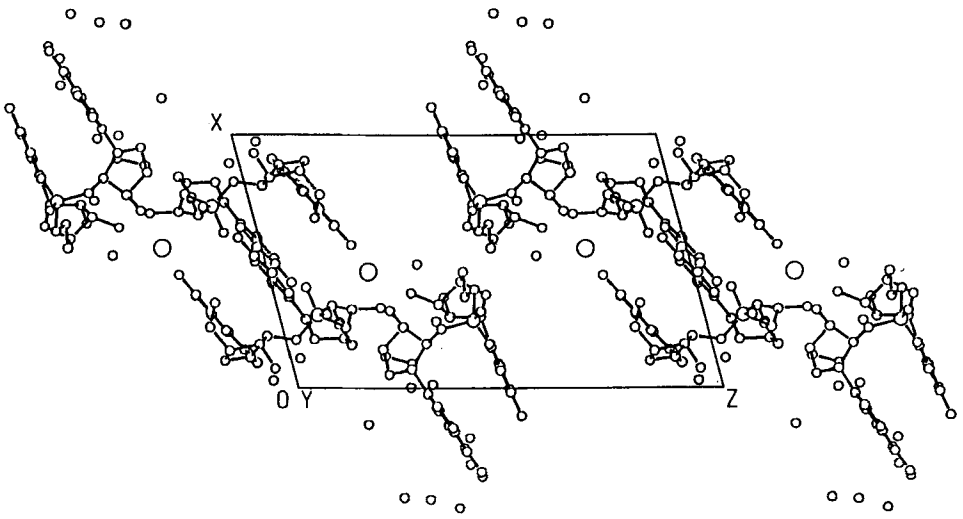

FIG. 4 Unit cell packing of G-p$_c$C(2′-5′) viewed along *b*-axis.

The considerable conformation difference between two structural forms of G-p$_c$C(2'-5') indicates conformational flexibility due to inserted methylene group. The geometry of phosphonate-sugar backbone in the "open" structure is in a good agreement with the g^-gt conformation (*i.e.* mirror symmetry) predicted for the simplest model of phosphonate internucleotide linkage (methyl methoxymethanephosphonate) by *ab initio* calculation.[27] The g^-gt conformation is also highly preferred for the duplex of d(T)$_{11}$ and d(A)$_{11}$ analog with alternating phosphonate phosphate linkage obtained by molecular dynamics simulation.[28]

Similarly as in GpC(2'-5') a pair of intermolecular hydrogen bonds, namely, cytosine N4 ... O1P (2.814 Å and 2.829 Å) and cytosine N3 ... O3P (2.728 Å and 2.678 Å), links the protonated cytosine to the negatively charged phosphonate (Fig. 3) and connects two conformationally identical molecules of G-p$_c$C(2'-5'). The third hydrogen bond is different and connects cytosine O2 and N4 (2.850 Å and 2.959 Å) of conformationally different molecules. Another difference against GpC(2'-5') is represented by the absence of intramolecular base stacking in the sense of their parallel orientation but the intermolecular base-base distance in G-p$_c$C(2'-5') is 3.43 Å. In contrast to GpC(3'-5'), the GpC(2'-5') and also G-p$_c$C(2'-5') do not create a miniature double helix connected by Watson-Crick base-pairing pattern (Fig. 4).

TABLE 3. Proton NMR data of G-p$_c$C(2'-5') (in D_2O at pH = 3.6; temp. 4– 40°C).

	Chemical shifts [ppm]						Coupling constants [Hz]				
Proton	4°	10°	20°	30°	40°	**Hi,Hj**	4°	10°	20°	30°	40°
Guanosine moiety											
H1'	5.984	5.984	5.982	5.987	5.986	**1',2'**	4.5	4.58	4.88	5.04	5.2
H2'	5.141	5.144	5.140	5.149	5.152	**2',3'**	5.49	5.42	5.42	5.42	5.4
H3'	4.721	4.703	4.666	4.648	4.628	**3',4'**	5.34	5.18	4.96	4.73	4.51
H4'	4.191	4.196	4.198	4.209	4.211	**4',5'**	2.6	2.6	2.75	2.82	2.9
H5'	3.887	3.884	3.874	3.878	3.874	**4',5''**	3.97	3.97	4.13	4.16	4.19
H5''	3.775	3.776	3.776	3.784	3.785	**5',5''**	-12.82	-12.8	-12.82	-12.82	-12.74
H8	7.989	7.987	8.002	8.000	7.992	**P,2'**	8.43	8.43	8.31	8.24	8.24
Cytidine moiety											
H1'	5.754	5.757	5.754	5.765	5.770	**1',2'**	2.15	2.25	2.40	2.64	2.94
H2'	4.162	4.156	4.144	4.146	4.144	**2',3'**	4.75	4.75	4.80	4.75	4.80
H3'	4.156	4.148	4.129	4.124	4.115	**3',4'**	7.35	7.25	7.10	6.90	6.70
H4'	4.103	4.100	4.091	4.096	4.091	**4',5'**	2.34	2.32	2.30	2.26	2.22
H5'	3.806	3.800	3.786	3.783	3.774	**4',5''**	1.90	1.95	2.00	2.10	2.20
H5''	3.661	3.657	3.644	3.645	3.640	**5',5''**	-11.40	-11.40	-11.40	-11.40	-11.40
P-CHa	3.813	3.806	3.788	3.780	3.764	**P,CHa**	8.1	8.1	~7.5	~7.5	7.47
P-CHb	3.67	3.664	3.652	3.649	3.642	**P,CHb**	10.9	10.9	10.83	10.76	10.6
H5	5.875	5.892	5.933	5.965	5.988	**Ha,Hb**	-13.05	-13.05	-13.2	-13.24	-13.28
H6	8.269	8.254	8.236	8.219	8.194	**5,6**	7.86	7.9	7.9	7.86	7.85

NMR study in aqueous solution

Structural assignment and extraction of NMR parameters. The assignment of guanine H8 and cytosine H5, H6 protons on the basis of chemical shifts and signal multiplicities is straightforward, similarly as the assignment of protons of "isolated" methylene group (P-CH_2). The ribose protons of both nucleotidyl parts of G-p$_c$C(2'-5') were assigned from 2D-COSY spectra. The additional spin-spin coupling of H1', H2' and H3' protons to phosphorus was used to identify a ribose of guanosine. Proton chemical shifts and coupling constants J(H,H) and J(H,P) were extracted from

TABLE 4. ^{13}C NMR data of G-p$_c$C(2'-5') (at 20°C; in D_2O at pH = 3.6)

Guanosine moiety				Cytidine moiety			
	δ(C)	J(C,P)	J(C,H) [a]		δ(C)	J(C,P)	J(C,H)
C-1'	88.50	C1',P=6.3	C1',H1'=168	**C-1'**	89.30		C1',H1'=173
C-2'	76.12	C2',P=5.4	C2',H2'=155.2	**C-2'**	74.67		C2',H2'=154.5
C-3'	69.57	C3',P=2.5	C3',H3'=153	**C-3'**	68.38		C3',H3'=148.7
C-4'	84.12		C4',H4'=148.8	**C-4'**	81.74		C4',H4'=148.8
C-5'	60.81		C5',H5'=143.4 C5',H5"=143.4	**C-5'**	70.32	C5',P=15.1	C5'H5'=143.4 C5',H5"=143.4
C-2	153.15			**PCHaHb**	66.64	CH_2,P=160.3	C,Ha=140.4 C,Hb=140.4
C-4	150.69		C4,H1'=7.3	**C-2**	156.96		C2,H1'< 2
C-5	116.59			**C-4**	165.43		C4,H6=10.1 C4,H5=3.1
C-6	158.50			**C-5**	95.39		C5,H5=177
C-8	138.68		C8,H8=212 C8,H1~4.5	**C-6**	141.23		C6,H6=185.5 C6,H5 ~3.0 C6,H1'~ 3.0

[a] The *J*(C4,H8) and *J*(C6,H8) were not observed due to the medium rate exchange of H8 proton.

resolution-enhanced 1D-proton spectra and refined by simulation-iteration analysis. Heteronuclear ^1H-^{13}C HMQC experiment was then used for structural assignment of all proton-bearing carbon atoms. The remaining quaternary carbons of guanine and cytosine were assigned on the basis of known chemical shift data of similar compounds and signal multiplicities observed in proton-coupled ^{13}C NMR spectra, from which we extracted also *J*(C,P) and some of *J*(C,H) couplings. Proton and carbon-13 NMR data are summarized in Tables 3 and 4.

Conformation of the ribofuranose ring. The conformation analysis of G-p$_c$C(2'-5') in aqueous solution is based mainly on the interpretation of vicinal coupling constants *J*(H,H), *J*(H,P), *J*(H,C) and *J*(P,C) using Karplus-type equations for corresponding interacting nuclei. Some additional arguments were derived from proton-proton contacts observed in 2D-ROESY spectrum. In solution, the furanose ring in ribonucleotides exists as an equilibrium mixture of conformers commonly denoted as N-type (roughly C3'-*endo*) and S-type (roughly C2'-*endo*).[29,30] The geometry of each of these conformers

is conveniently described by two geometrical parameters, namely, phase angle of pseudorotation P and puckering amplitude Φ_m. Using a model assuming two rapidly interconverting conformers (leading to population-weighed averaging of the observed NMR parameters), and a generalized Karplus equation,[31] one can calculate the pseudorotation parameters (phase angle P, puckering amplitude Φ_m, and molar fraction) of both conformers present[32,33] by means of the program PSEUROT[34].

Because only three couplings, *J*(H1',H2'), *J*(H2',H3') and *J*(H3',H4'), are available for a ribofuranose ring, a full pseudorotation analysis in terms of all five parameters demands the measurements of various sets of coupling constants at several temperatures (or pH values). Therefore we have measured the ^{1}H NMR spectra of G-p$_c$C(2'-5') at five different temperatures (in the range of 4 – 40°C) in aqueous solution at pH = 3.6 (identical with pH of the solution used for crystallization of the sample). While the ribose protons of guanosine moiety give the first order spectra, the ribose protons H-2', H-3' and H-4' of cytidine moiety form a strongly coupled spin system which requires simulation-iteration analysis. The refined values of chemical shifts and coupling constants are given in Table 2. All five sets of *J*(1',2'). *J*(2',3') and *J*(3',4') had been used for calculation of the "temperature independent" geometry parameters P_N, ϕ_N, P_S, ϕ_S and for molar ratio X_N which was taken as a temperature variable.[34] The pseudorotation parameters were calculated using a program[35] similar to the program PSEUROT. The puckering amplitudes (ϕ_N, ϕ_S) were optimized in the range of 30 to 45° in 2° steps, the pseudo-rotation phase angles for both conformers (P_N and P_S) were changed in the whole range of the possible values (–90 to +90° and +90 to 270°) in 3° steps, and the conformer population, expressed by molar fractions X_N and X_S, in the range of 0 to 1 in 0.02 steps. The best pseudorotation parameters P_N, ϕ_N, P_S, ϕ_S, common for all five sets of *J*-values, were evaluated as differences between the observed and calculated *J*-values, the decisive criterion being Σ_{rms} (final set of calculated pseudorotation parameters gives $\Sigma\Delta J$ = 0.55 Hz and Σ_{rms} = 0.22 Hz for guanosine and $\Sigma\Delta J$ = 0.43 Hz and Σ_{rms} = 0.17 Hz for cytidine ribose). For ribose ring of guanosine the two conformers, C3'-*endo* (P_N = 18°, ϕ_N = 36°) and C2'-*endo* (P_S = 150°, ϕ_S = 36°), were established to be equally populated at low temperature (4°C). With increasing temperature the population of C2'-*endo* conformation increases up to the mole ratio of 40:60 at 40°C. On the other hand, in cytidine moiety the C3'-*endo* form (P_N = 0°, ϕ_N =

38°) of ribose ring significantly prevails over the C2'-*endo* form (P_S=150°, ϕ_S=38°). Its population slightly decreases with increasing temperature from the mole ratio of 84:16 found at 4°C to 74:26 at 40°C. The highly preferred ribose conformation of cytidine in solution thus fits very well with the geometry in the crystal (see Table 2). Ribose ring of guanosine seems to manifest a much higher flexibility in solution.

Torsion angle γ (O5'-C5'-C4'-C3'). The relative γ^+, γ^t, γ^- rotamer populations around C4'-C5' bond were calculated from vicinal coupling constants *J*(H4',H5') and *J*(H4',H5") using equations [1]–[3] with values 11.5 and 1.8 Hz for *J*(*trans*) and *J*(*gauche*).[36] According to Remin and Shugar[37] the H5' and H5" proton signals were assigned such that δ(H5') > δ(H5"), where H5' is gauche to H4' and O4'.

$$\gamma^+ = ((J(trans) + J(gauche) - (J(H4',H5') + J(H4',H5")) / (J(trans) - J(gauche) \quad [1]$$

$$\gamma^t = (J(H4',H5") - J(gauche) / (J(trans) - J(gauche)) \quad [2]$$

$$\gamma^- = (J(H4',H5') - J(gauche) / (J(trans) - J(gauche)) \quad [3]$$

The results of the calculation show, in general, preference for the γ^+ form which is almost the sole rotamer (> 90%) for the cytidine part and still the most populated one (~65%) in the more flexible guanosine part of G-p_cC(2'-5'). The population of the γ^+ form only very slightly decreases with increasing temperature . The γ^+ rotamers were also found in the crystal (angles γ(G) = 53°, 55° and γ(C) = 51°, 58°).

Torsion angle ε (C3'-C2'-O2'-P). The conformation about the C2'-O2' bond can be determined from vicinal coupling constants observed either in ^{1}H NMR spectra [*J*(H2'-C2'-O-P)] or ^{13}C NMR spectra [*J*(C1'-C2'-O-P)] and J(C3'-C2'-O-P) using Karplus type equations [4] and [5] derived by Lankhorst[38] and recently reparametrized.[39]

$$^3J(HCOP) = 15.3 \cos^2 \phi - 6.2 \cos \phi + 1.5 \quad [4]$$

The torsion angle ε has been only observed to occur in the *trans* or *gauche*(–) domain[26] since *gauche*(+) is always forbidden because of steric hindrance and it is not considered in the subsequent analysis. Using a single rotamer model the observed values of ^3J(C3',P) = 2.5 and ^3J(C1',P) = 6.3 Hz lead to an angle ε –109° while ^3J(H2',P) = 8.3 Hz indicates a lower value of ε (–84°). Both calculated ε values are quite close to those

found in the crystal (–118° and –86°). The alternative analysis based on the two-state model[39] can be applied to the calculation of ε(t), ε(g⁻) and population of rotamers (X(t), X(g⁻)) using Karplus-type equations [4] and [5] and relations [6–9]. The combination of three observed *J*-values gives a good agreement ($|J_{obs} - J_{calc}| < 0.5$ Hz) for the mole ratio X(t) : X(g⁻) = 66 : 34 with ε(t) = –120° and ε(g⁻) = –60°.

$$J(C3',P) = X(t).J(C3',P)(\varepsilon(t)) + X(g^-).J(C3',P)(\varepsilon(g^-)) \quad [6]$$

$$J(C1',P) = X(t).J(C1',P)(\varepsilon(t)-120°) + X(g^-).J(C1',P)(\varepsilon(g^-)-120°) \quad [7]$$

$$J(H2',P) = X(t).J(H2',P)(\varepsilon(t)+120°) + X(g^-).J(H2',P)(\varepsilon(g^-)+120°) \quad [8]$$

$$X(t) + X(g^-) = 1 \quad [9]$$

Torsion angles ξ (C2'-O2'-P-C(P)), α (O2'-P-C(P)-O5') and α' (P-C(P)-O5'-C5'). As follows from definition of torsion angles ξ and α (Fig. 1) there are no proper coupling constants which could be used for their estimation from NMR spectra. The conformation around the C(P)-O5' bond is coded in vicinal coupling constants *J*(P-C(P)-O5'-C5'). From the observed value *J*(P,C5') = 15.1 Hz and the Karplus-type equation [10] derived for phosphonates,[40] two possible values of the angle α' = ±160° could be estimated (compare with α' = 174° and –150° found in the crystal). Preferred *trans* orientation of phosphorus atom and carbon C5' is further supported by very small vicinal couplings observed in proton-coupled ^{13}C NMR spectrum between carbon C5'and PCH_2 protons (<2 Hz) which both adopt *gauche*-relation to C5' in that form.

$$^3J(C,P) = 7.86 \cos 2\phi - 1.76 \cos \phi + 7.35 \quad [10]$$

Torsion angle β (C(P)-O5'-C5'-C4'). The only vicinal couplings related to the conformation around O5'-C5' bond are two heteronuclear C-H couplings *J*(C(P)-O5'-C5'-H5') and *J*(C(P)-O5'-C5'-H5''). As determined from proton-coupled ^{13}C NMR spectra, their values are very small (< 2 Hz). Their comparison with torsion angle dependence of $^3J(C,H)$ derived for C-O-C-H arrays[41] indicates *gauche*-relation of both protons H5', H5'' to carbon atom C(P). Such an arrangement corresponds to preferred *trans* orientation of carbon atoms C(P) and C5' in solution which is the crystal.

Conformation of the guanine and cytosine bases about the glycosyl bond. *Syn-* and *anti*-orientations of guanine (or cytosine) differ in the distance between base protons H8

(or H6) and ribose protons H1'and H2'and can be distinguished, in principle, by NOE measurements.[42] In case of G-p$_C$C(2'-5'), we have observed a strong NOE contact of guanosine H8 proton to H1' and a very weak NOE to H2' clearly indicating a preferred *syn*-orientation of guanine. An opposite situation has been observed for cytosine where its H6 proton shows strong NOE contact to H2' and H3' while NOE to H1'is very weak which means a preferred *anti*-orientation of cytosine. This is in agreement with the orientation of bases found in the crystal.

The knowledge of vicinal coupling constants of proton H1' with carbon atoms C4, C8 of guanine and/or carbon atoms C2, C6 of cytosine should allow for determination of the torsion angle χ (at least for "rigid" nucleotide molecule) using Karplus-type curve adjusted for ^3J(H1'-C1'-N-C) [11] (refs 43,44) and approximate relations of angles: $\chi = \phi$(H1',C4 or C2) + 60° and $\chi = \phi$(H1',C8 or C6) – 120° for guanosine or cytidine. In case of G-p$_c$C(2'-5') we have observed J(H1',C4) = 7.3 Hz and J(H1',C8) ~ 4.5 Hz for guanosine and J(H1',C2) < 1.5 Hz and J(H1',C6) ~ 3.0 Hz for cytidine moiety. Due to the periodicity of Karplus relation these J-values lead to two possible χ angles of 37° and 83° for guanosine of which only the first one (χ = 37°) is typical for the *syn*-glycosidic bond and fits well the values χ = 51° and 44° found for G-p$_c$C(2'-5') in the crystal. The J-values of cytidine result in the alternative χ values of *ca* –176° and –64° but only the first one fits an NOE conclusion about the *anti*-glycosidic bond and agrees with crystal data (χ = -150° and –161°).

$$J(\text{H1'-C-N-C}) = 7.0 \cos^2 \phi - 1.5 \cos \phi \qquad [11]$$

The quantitative NOE data are commonly used in the determination of 3D-structures of biomolecules in solution.[45] Their application to the more flexible molecules is limited due to the existence of several conformations contributing to the observed NOEs.[46] Nevertheless, even in this case the qualitative or semiquantitative interpretation of the observed NOE contacts can provide some useful information about conformational behavior of the molecule under study. Its use for the determination of preferred *syn*- or *anti*-orientation of guanine and cytosine has been discussed above. A surprisingly high similarity of conformational features of G-p$_c$C(2'-5') in solution and in the crystal, as followed from the estimated torsion angles, prompted us to compare the

TABLE 5. The interproton distances for G-p$_c$C(2′-5′) found in the crystal and the NOE contacts observed in aqueous solution.

↓↓↓ Interproton distances in the crystal (normal – "open" structure; italics – "closed" structure) ↓↓↓																	
H	**6-C**	**8-G**	**1′-G**	**5-C**	**1′-C**	**2′-G**	**3′-G**	**4′-G**	**2′-C**	**3′-C**	**4′-C**	**5′-G**	**CHa**	**5′-C**	**5″-G**	**CHb**	**5″-C**
δ	**8.24**	**8.00**	**5.98**	**5.93**	**5.75**	**5.14**	**4.67**	**4.20**	**4.14**	**4.13**	**4.09**	**3.87**	**3.79**	**3.79**	**3.78**	**3.65**	**3.64**
5″-C	3.8	8.1	6.3	5.9	5.0	4.4	3.1	5.6	4.9	3.0	2.3	6.3	2.6	1.6	5.0	2.3	
3.64	*3.8*	*5.0*	*6.0*	*6.0*	*5.0*	*6.2*	*7.4*	*8.7*	*4.9*	*2.9*	*2.4*	*9.8*	*2.5*	*1.6*	*9.7*	*2.6*	
CHb	3.2	6.9	6.0	4.4	5.8	4.1	4.4	6.9	6.4	4.4	4.5	7.5	1.6	3.1	6.7		*
3.65	*3.9*	*5.7*	*5.6*	*5.0*	*6.5*	*4.9*	*5.7*	*7.6*	*6.1*	*3.9*	*4.7*	*8.6*	*1.6*	*3.2*	*8.1*		
5″-G	5.1	7.0	4.8	6.2	6.2	3.7	2.5	2.3	3.9	3.0	5.3	1.6	6.0	6.1			
3.78	*7.2*	*7.0*	*4.8*	*6.1*	*9.9*	*3.7*	*2.6*	*2.3*	*10.9*	*9.8*	*10.0*	*1.6*	*7.6*	*8.8*			
5′-C	3.3	8.5	7.1	5.5	4.5	4.8	4.3	6.9	5.2	3.5	2.4	7.3	2.5			< S >	
3.79	*3.6*	*3.6*	*4.7*	*5.7*	*4.6*	*5.4*	*6.7*	*7.6*	*5.2*	*3.6*	*2.4*	*8.8*	*2.3*				
CHa	3.9	6.5	5.2	5.2	6.5	3.8	3.5	5.9	6.5	4.3	4.5	6.9		*		< S >	
3.79	*3.8*	*4.6*	*4.5*	*5.3*	*6.2*	*4.5*	*5.2*	*6.7*	*6.5*	*4.4*	*4.4*	*8.0*					
5′-G	5.7	6.4	4.5	6.3	7.0	3.9	3.5	2.3	4.7	4.1	6.6				S		
3.87	*7.1*	*6.4*	*4.4*	*6.1*	*9.3*	*3.8*	*3.5*	*2.3*	*10.6*	*9.8*	*10.3*						
4′-C	3.8	9.6	7.8	6.2	3.2	5.4	4.4	6.7	3.6	2.8				S			S
4.09	*3.9*	*4.5*	*6.4*	*6.3*	*3.2*	*6.9*	*8.6*	*9.5*	*3.6*	*2.8*							
3′-C	2.4	7.2	5.5	4.4	3.6	3.0	2.5	4.5	2.3		S			< S >			S
4.13	*2.6*	*5.5*	*6.8*	*4.5*	*3.5*	*6.1*	*7.9*	*9.4*	*2.3*								
2′-C	3.7	8.9	7.4	5.3	2.6	4.8	4.6	6.0		*	S				W		
4.14	*3.8*	*6.2*	*7.9*	*5.3*	*2.6*	*7.3*	*9.4*	*10.6*									
4′-G	6.0	5.6	3.2	6.7	8.0	3.6	2.6					S			S		
4.20	*6.9*	*5.6*	*3.2*	*6.5*	*9.2*	*3.6*	*2.6*										
3′-G	3.8	5.8	3.7	5.2	6.0	2.3		S		M		W			W		W
4.67	*5.7*	*5.8*	*3.6*	*5.0*	*8.7*	*2.3*											
2′-G	2.9	4.3	2.9	3.2	6.0		S	W		M			M		M		W
5.14	*3.6*	*4.2*	*2.8*	*3.0*	*6.5*												
1′-C	3.6	10.0	8.8	5.3					< S >		S						
5.75	*3.5*	*4.2*	*6.4*	*5.3*													
5-C	2.2	5.5	5.5			S											
5.93	*2.5*	*5.1*	*5.2*														
1′-G	5.7	2.5				S	M	M									
5.98	*4.5*	*2.5*															
8-G	6.6		S			W											
8.00	*3.7*																
6-C				S	W	S			< S >				< M >			< M >	
8.24																	
↑↑↑ **Observed NOE contacts** [a] ↑↑↑																	

[a] The intensities of NOE cross-peaks are classified as: S = strong, M = medium, W = weak; Symbols < ... > indicate cross-peaks which can belong to more than one overlapping protons; * cross-peak could not be safely detected due to overlap of corresponding signals.

interproton distances found for two independent structures of G-p$_c$C(2′-5′) in the crystal and the observed NOE-contacts.

Detailed comparison of X-ray versus NOE data requires a stereochemical assignment of geminal protons at C5′ and PCH$_2$. Such stereochemical assignment is closely connected with conformation about C4′-C5′ and P-CH$_2$ bonds. In case of C4′-C5′ bond the preferred γ+ conformation could be determined from *J*(H4,H5′) and *J*(H4′,H5′) and

then protons H5' and H5' assigned from NOE on the basis of their different distances to proton H3'. On the other hand, there is no NMR evidence for conformation about P-CH_2 bond and these protons could not be stereochemically assigned. The interpretation of some NOE contacts is further complicated by overlap of protons H2'(C), H3'(C) at δ~4.14, PCHa, H5'(C), H5'(G) at δ~3.78 and PCHb, H5' at δ~3.65.

The results are shown in Table 5. There are in total 136 distances between 17 nonequivalent protons present in the molecule. Out of this number, 71 distances in "open" structure and 61 distances in "closed" structure are shorter than 5.0 Å and therefore, in principle, observable by NOE contacts.[45] The fourteen contacts between geminal and vicinal protons are "trivial" and all of them have been detected. The NOE contacts were also detected in all cases where interproton distance is <3.5 Å in both "open" and "closed" structures present in the crystal. The existence of a folded conformation in solution is supported by observed interresidue contacts between H2'(G) and both H5(C), H6(C), and between H3'(C) and both H2'(G), H3'(G). Some interproton distances in "open" and "closed" structure differ in such a way that NOE contact of given pair of protons should be observed only in one structure (d < 4 Å) while the distance is too large in the other (d >> 5 Å). For the "open" structure, such observable contacts are H2'(G)/H3'(C), H6(C)/H3'(G), H3'(G)/H3'(C), H3'(G)/H5'(C), H3'(G)/PCHa, H2'(C)/H5'(G) and H3'(C)/H5'(G) and for the "closed" structure, the respective contacts are H8(G)/H6(C) and H8(G)/H5'(C). Although not all of these contacts can be unequivocally checked in ROESY spectrum the absence of both detectable contacts for "closed" structure and the observation of the most of those characteristic for "open" structure suggest that "open" type conformation should be significantly populated in solution.

Acknowledgements

This work was supported by Ministry of Education of the Czech Republic (OK 169) and by grant No. A4055902/1999 of Grant Agency of the Academy of Sciences of the Czech Republic.

REFERENCES

1. Vlassov, V.V.; Vlassova, I.E.; Pautova, L.V. in *Progress in Nucleic Acid Research and Molecular Biology*, vol. 57 (Moldave, K.; ed.), p. 95. Academic Press, San Diego **1997**.

2. Hunziker, J.; Leumann, C. in *Modern Synthetic Methods 1995* (Ernst, B.; Leumann, C.; eds.), p. 333. VHCA, Basel **1995**.
3. Kool, E. T. *Chem. Rev.*, **1997**, *97*, 1473-1487.
4. Tazawa, I.; Tazawa, S.; Stempel, L.M.; Ts'o, P.O.P. *Biochemistry*, **1970**, *9*, 3499-3514.
5. Letsinger, R.L.; Bach, S.A.; Eadie, J.S. *Nucleic Acids Res.*, **1986**, *14*, 3487-3499.
6. Miller, P.S.; Yano, J.; Yano, E.; Carroll, C.; Jayaraman, K.; Ts'o, P.O.P. *Biochemistry*, **1979**, *18*, 5134-5143.
7. Muraoka, M.; Takahashi, S; Uesugi, S. *Nucleosides Nucleotides*, **1995**, *14*, 1503-1518.
8. Mojzeš, P.; Štěpánek, J.; Rosenberg, I.; Točík, Z.; Burian, M.; Pavelčíková, M.; Réfrégiers, M.; Videlot, H. *J. Mol. Struct.*, **1995**, *348*, 45-48.
9. Zajíček, J.; Rosenberg, I.; Liboska, R.; Točík, Z.; Fazakerley, G.V.; Guschlbauer, W.; Holý, A. *Collect. Czech. Chem. Commun.*, **1997**, *62*, 83-98.
10. Štěpánek, J.; Hanuš, J.; Endová, M.; Dole, J. Točík, Z.; Rosenberg, I., in *Spectroscopy of Biological Molecules: Modern Trends* (Carmona, P.; ed.), Kluwer Acad. Publ., Dordrecht **1997**, pp 239-240.
11. Pressová, M.; Endová, M.; Točík, Z.; Liboska, R.; Rosenberg, I. *Bioorg. Med. Chem. Lett.*, **1998**, *8*, 1225-1230.
12. Krishnan, R.; Seshadri, T.P.; Viswamitra, M.A. *Nucleic Acids Res.*, **1991**, *19*, 379-384.
13. Rosenberg, I.; Holý, A. *Collect. Czech. Chem. Commun.*, **1987**, *52*, 2572-2588.
14. Sheldrick G.M. *Acta Crystallogr. A.*, **1990**, *46*, 467-473.
15. Sheldrick G.M. SHELXL97. *Program for Crystal Structure Refinement from Diffraction Data.* University of Gottingen, Gottingen **1997**.
16. Castellano, S., Bothner-By, A.A. *J. Chem. Phys.*, **1964**, *41*, 3863-3869.
17. Bodenhausen, G.; Freeman, R.; Niedermeyer, R.; Turner, D.L. *J. Magn. Reson.*, **1977**, *26*, 133-164.
18. Bachmann, P.; Aue, W.P.; Muller, L.; Ernst, R.R. *J. Magn. Reson.*, **1977**, *28*, 29-39.
19. Bax, A.; Davis, D.G. *J. Magn. Reson.*, **1985**, *63*, 207-213.
20. Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R.R. *J. Am. Chem. Soc.*, **1987**, *109*, 607-609.
21. Muller, L. *J. Am. Chem. Soc.*, **1979**, *101*, 4481-4484.
22. Bax, A.; Griffey, R.H.; Hawkins, B.L. *J. Magn. Reson.*, **1983**, *55*, 301-315.
23. IUPAC/IUB JCBN Recommendations 1982, *Eur. J. Biochem.*, **1983**, *131*, 9-15.
24. Krishnan, R.; Seshadri, T.P. *Biopolymers*, **1994**, *34*, 1637-1646.
25. Taylor, R., Kennard, O. *J. Mol. Struct.*, **1982**, *78*, 1-25.
26. Saenger, W. in *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, **1984**, p. 56 and 132.
27. Štrajbl, M.; Florián, J. *J. Biomol. Struct. Dyn.*, **1996**, *13*, 687-694.
28. Barvík, I.; Bok, J.; Štěpánek, J. unpublished results.
29. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.*, **1972**, *94*, 8205-8212.
30. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.*, **1973**, *95*, 2333-2344.
31. Haasnoot, C.A.G.; de Leeuw, F.A.A.M.; de Leeuw, H.P.M.; Altona, C. *Org. Magn. Reson.*, **1981**, *15*, 43-52.
32. Haasnoot, C.A.G.; de Leeuw, F.A.A.M.; Altona, C. *Tetrahedron*, **1980**, *36*, 2783-2792.

33. de Leeuw, F.A.A.M.; Altona, C. *J. Chem. Soc., Perkin Trans. 2*, **1982**, 375-384.
34. de Leeuw, F.A.A.M.; Altona, C. *J. Comp. Chem.*, **1983**, *4*, 428-437.
35. Raich, I. *PhD Thesis,* Prague Institute of Technology, Prague **1994**.
36. Altona, C. *Recl. Trav. Chim. Pays-Bas*, **1982**, *101*, 413-433.
37. Remin, M.; Shugar D. *Biochem. Biophys. Res. Commun.*, **1972**, *48*, 636-642.
38. Lankhorst, P.P.; Haasnoot, C.A.G.; Erkelens, C.; Altona, C. *J. Biomol. Struct. Dyn.*, **1984**, *1*, 1387-1405.
39. Mooren, M.M.W.; Wijmenga, S.S.; van der Marel, G.A.; van Boom, J.H.; Hilbers, C.W. *Nucleic. Acids Res.*, **1994**, *22*, 2658-2666.
40. Thiem, J.; Meyer, B. *Org. Magn. Reson.*, **1978**, *11*, 50-51.
41. Hamer, G.K.; Balza P.; Cyr, N.; Perlin, A.S. *Can. J. Chem.*, **1978**, *56*, 3109-3116.
42. Wijmenga, S.S.; Mooren, M.M.W.; Hilbers, C.W. in *NMR of Macromolecules*, Oxford Univ. Press, Oxford, **1993**, p.217-288.
43. Lemieux, R.U.; Nagabushan, T.L.; Paul B. *Can. J. Chem.*, **1972**, *50*, 773-776.
44. Delbaere, L.T.J.; James, M.N.G.; Lemieux, R.U. *J. Am. Chem. Soc.*, **1973**, *95*, 7866-7868.
45. Wüthrich, K. in *NMR of Proteins and Nucleic Acids*, J.Wiley, New York, **1986**, p.177.
46. Lane, A.N. *Magn. Reson. Chem.*, **1996**, *34*, S3-S10.

Received : 3/15/99
Accepted : 6/1/99