

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/79452>

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

# The solution structure of the circular trinucleotide cr(GpGpGp) determined by NMR and molecular mechanics calculation

Margret M.W. Mooren<sup>+</sup>, Sybren S. Wijmenga, Gijs A. van der Marel<sup>1</sup>, Jacques H. van Boom<sup>1</sup> and Cornelis W. Hilbers\*

NSR Centre for Molecular Structure, Design, and Synthesis, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen and <sup>1</sup>Gorlaeus Laboratories, State University Leiden, PO Box 9502, 2300 RA Leiden, The Netherlands

Received February 25, 1994; Revised and Accepted May 31, 1994

## ABSTRACT

The 3'-5' circular trinucleotide cr(GpGpGp) was studied by means of 1D and 2D high resolution NMR techniques and molecular mechanics calculations. Analysis of the J-couplings, obtained from the <sup>1</sup>H and <sup>13</sup>C-NMR spectra, allowed the determination of the conformation of the sugar rings and of the 'circular' phosphate backbone. In the course of the investigations it was found that the Karplus-equation most recently parametrized for the CCOP J-coupling constants could not account for the measured J(C4'P) of 11.1 Hz and a new parametrization for both HCOP and CCOP coupling constants is therefore presented. Subsequent analysis of the coupling constants yielded 'fixed' values for the torsion angles  $\beta$  and  $\delta$  (with  $\beta = 178^\circ$  and  $\delta = 139^\circ$ ). The value of the latter angle corresponds to an S-type sugar conformation. The torsion angles  $\gamma$  and  $\epsilon$  are involved in a rapid equilibrium in which they are converted between the gauche(+) and trans and between the trans and gauche(-) domain respectively. We show that the occurrence of  $\epsilon$  in the gauche(-) domain necessitates S-type sugar conformations. Given the aforementioned values for  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  the ring closure constraints for the ring, formed by the phosphate backbone can only be fulfilled if  $\alpha$  and  $\zeta$  adopt some special values. After energy minimization with the CHARMM force field only two combinations of  $\alpha$  and  $\zeta$  result in energetically favourable structures, i.e. the combination  $\alpha(t)/\zeta(g-)$  in case  $\gamma$  is in a gauche(+) and  $\epsilon$  is in a trans conformation, and the combination  $\alpha(t)/\zeta(g+)$  for the combination  $\gamma(t)/\epsilon(g-)$ . The results are discussed in relation to earlier findings obtained for cd(ApAp) and cr(GpGp), the latter molecule being a regulator of the synthesis of cellulose in *Acetobacter xylinum*.

## INTRODUCTION

Low molecular weight nucleosides and nucleotides such as cyclic molecules in which the nucleotide units are connected by 3'-5' linkages, play important roles in many cellular functions. For example, the circular ribo-dinucleotides cr(UpUp) and cr(ApUp) are effective inhibitors of the DNA dependent RNA-polymerase of the bacterium *Escherichia coli* during the initiation phase of transcription (1,2). A completely different activity is exhibited by the circular dinucleotide cr(GpGp), which is an activator of the enzyme cellulose synthase in the bacterium *Acetobacter xylinum* (3). Some structural analogs, like cr(IpIp), cd(GpGp) and two diastereomers of the monophosphothioates cr(Gp(S)Gp) were found to be relatively potent activators of the cellulose synthase as well. On the other hand, closely related molecules were found to be only marginally stimulatory, like cd((Gp)r(Gp)) and cr(GpGpGp) or entirely devoid of stimulatory activity, like cd(ApAp), cr(CpCp), cr(UpUp) and cr(GpCp) (4). The structural reason for this is not clear. Therefore we investigated the solution structure of the circular trinucleotide cr(GpGpGp). This molecule is schematized in Fig. 1. A second reason for studying cr(GpGpGp) is that such circular molecules contain structural elements which may enhance the understanding of folding in larger DNA and RNA molecules. To obtain the solution structure of cr(GpGpGp) we have employed high resolution NMR and molecular mechanics calculations. The most salient finding is that the ribose rings of the molecule adopt a pure S-conformation at the temperatures studied, contrary to what is usually found for these types of sugars.

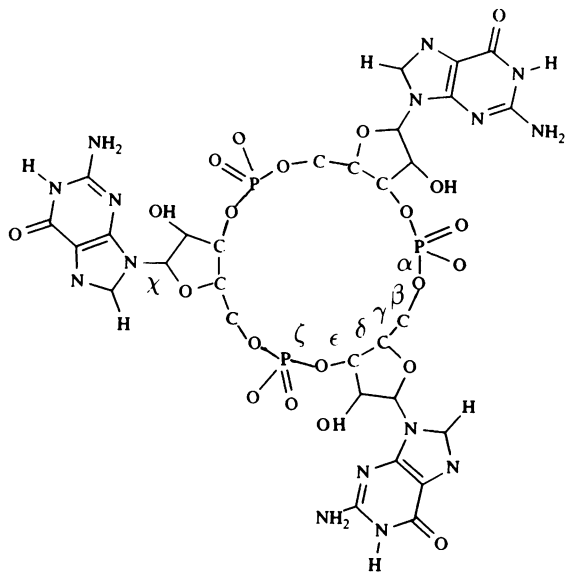
## MATERIALS AND METHODS

### Sample preparation

The circular trinucleotide cr(GpGpGp) was synthesized via an improved phosphotriester method (5). NMR samples were prepared by freeze drying the material two times from 99.8%

\*To whom correspondence should be addressed

<sup>+</sup>Present address: Unilever Research Laboratorium, Olivier van Noortlaan 120, 3133 AT Vlaardingen, The Netherlands



**Figure 1.** Schematic representation of the structure of the cyclic nucleotide, cr(GpGpGp). The ring system of the trinucleotide is closed by linking the 5'-end of the first guanosine to the 3'-end of the third. Definition of the torsion angles is indicated in one nucleotide and is according to the IUPAC/IUB convention for nucleic acids. For clearness, the sugar protons are omitted in this scheme.

D<sub>2</sub>O. Subsequently, 0.4 mg of the sodium form of freeze dried material was dissolved in 400 ml of a 25 mM phosphate buffer in 99.98% D<sub>2</sub>O, containing 1 mM sodium cacodylate, yielding a 1 mM sample of the circular trinucleotide. Sodium chloride was added until the total sodium concentration was 0.2 M.

### NMR spectroscopy

1D-<sup>1</sup>H-NMR spectra were recorded on a Bruker WM 200 spectrometer, equipped with an ASPECT 2000 computer, and on a Bruker AM 600 spectrometer, equipped with an ASPECT 3000 computer, at temperatures between 278K and 305K. 1D-<sup>13</sup>C spectra were recorded on a Bruker AM 400 spectrometer equipped with an ASPECT 3000 computer. In the latter experiment protons were broad-band decoupled during acquisition, using Waltz-decoupling. A phase sensitive 200 MHz NOESY spectrum (6) with a spectral width of 2000 Hz and with a mixing time of 300 ms was acquired with 2K points in the  $t_2$ - and 512 points, zero-filled to 1K points, in the  $t_1$ -direction. Prior to Fourier transformation the FIDs were multiplied with a squared cosine window function in both directions. An  $\omega_1$ -scaled double quantum filtered COSY spectrum ( $\omega_1$ -DQF-COSY) (7) was recorded on a Bruker AM 600 spectrometer with a spectral width of 6000 Hz and with 2K points in the  $t_2$ - and 512 points, zero-filled to 1K points, in the  $t_1$ -direction. The degree of  $\omega_1$ -scaling was equal to 0.5. Prior to Fourier transformation the FIDs were multiplied with a shifted sinebell window function in both directions.

A natural abundance <sup>1</sup>H-<sup>13</sup>C hetero correlation spectrum was recorded, at 305 K, on a Bruker AM 400 spectrometer, operating in the inverse mode (8), without <sup>13</sup>C-decoupling during data acquisition. The spectrum was acquired with a spectral width of 4000 Hz in the  $t_2$ - and 18000 Hz in the  $t_1$ -dimension. In the experiment 1536 transients were collected for each of the 114

FIDs which were sampled with 2K data points. The data were processed with zero filling to 512 data points in the  $t_1$ -dimension; Gaussian apodization was applied in the  $t_2$ -dimension to give resolution enhancement, and a  $\pi/2$  shifted squared sine bell window function was applied in the  $t_1$ -direction.

### Conformational analysis

The conformational analysis of the molecule was performed on the basis of J-coupling constants, which were derived from the 1D spectra. First, approximate J-couplings were deduced from the experimental spectra which subsequently, together with the chemical shifts, were used as input for the Bruker spectrum simulation program, PANIC. The parameters were adjusted in such a way that optimal agreement was obtained between the experimental and simulated spectrum. The pseudorotational analysis of the sugars was performed on the basis of the coupling constants available for the ribose ring, using a home-written computer program as described elsewhere (9). It uses a grid search routine to determine, from the available J-couplings, the parameters  $P_N$ ,  $P_S$ ,  $\phi_m^N$ ,  $\phi_m^S$  and  $x_S$  which describe the sugar conformation. In this way a region of allowed values for  $P_N$ ,  $P_S$ ,  $\phi_m^N$ ,  $\phi_m^S$  and  $x_S$  is found which satisfies the experimental J-couplings.

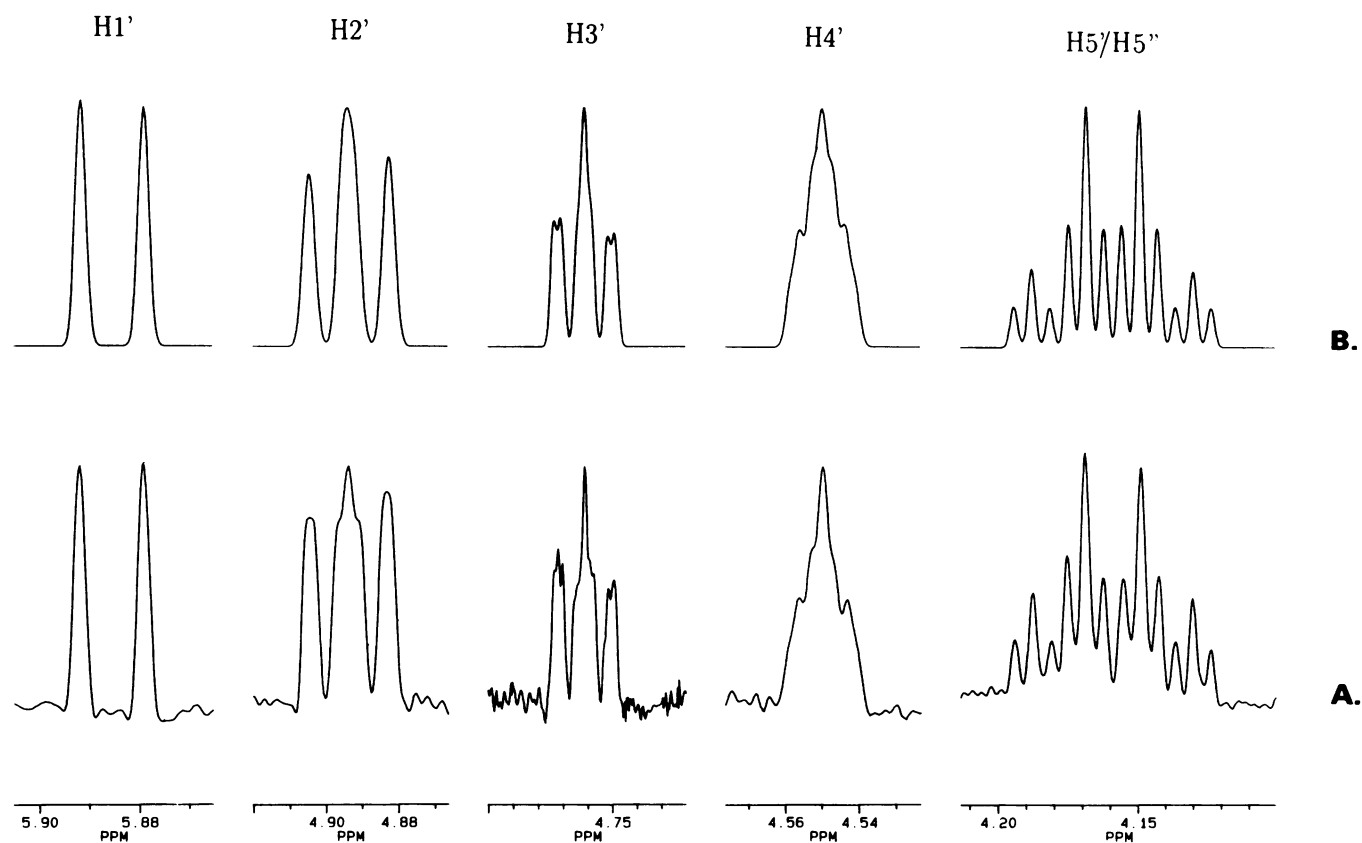
Derivation of torsion angles from hetero-coupling constants required reparametrization of the Karplus equations, relating the hetero-coupling constants between hydrogen and phosphorous,  $J(\text{HCOP})$ , and between carbon and phosphorous,  $J(\text{CCOP})$ , to the torsion angles  $\phi_{\text{HCOP}}$  and  $\phi_{\text{CCOP}}$ , respectively. A detailed description is given in Appendix 1.

Energy minimization of the circular trinucleotide was carried out with version 3.2.1 of the program Quanta (Polygen), which uses the CHARMM force field and runs on a Silicon Graphics Iris 4D-25G work station. A distance dependent dielectric constant was used to approximate the effect of solvent (10,11) on the molecule. The total charge of the molecule was distributed among the atoms using the Gasteiger method. Energy minimization, with an Adopted Basis Newton Raphson minimization method, was continued until the average energy gradient was equal to zero or until the differences between the energy values at each step during a cycle of minimization were less than 0.001 kJ/mole.

## RESULTS

### Interpretation of NMR spectra and derivation of J-couplings and NOEs

In all spectra recorded for cr(GpGpGp) only a single set of resonances is observed for each set of the same type of protons. From this result we may conclude that the sugar phosphate backbone ring has an (averaged) three-fold axis of symmetry. Figure 2A shows the different multiplets observed in the <sup>1</sup>H-NMR spectrum of cr(GpGpGp), recorded at 298K. The spectral assignment was obtained from the connectivity pattern of a  $\omega_1$ -DQF-COSY spectrum, recorded at the same temperature (spectrum not shown). The available data did not allow us to make a stereospecific assignment for the H5' and H5'' protons (9) and these were tentatively assigned on the assumption that Remin and Shugar's rule (12) applies. The 1D spectra, measured between 278 and 305K, showed that the chemical shifts as well as the coupling constants are independent of temperature in this range.



**Figure 2.** Experimental (bottom, (A)) and simulated (top, (B)) multiplets of the sugar proton resonances of cr(GpGpGp) measured at 600 MHz. The experimental spectrum was resolution enhanced with Gaussian multiplication before Fourier transformation. All the given multiplets, except the H3' multiplet were recorded at 298K. The H3' multiplet is from the spectrum, recorded at 305K, because at room temperature the H3' proton resonates exactly at the position of the residual HDO peak. Because of the HDO shift the H3' multiplet becomes visible at higher temperatures. Although the J-coupling constants are the same at different temperatures, the spectrum at 298K was preferred for the other multiplets because of a better signal to noise ratio.

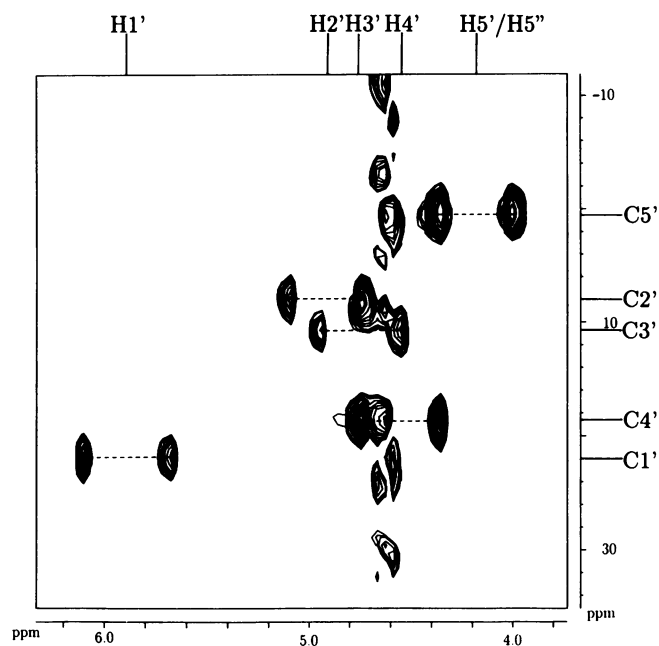
The  $^1\text{H}$ - $^1\text{H}$  and  $^1\text{H}$ - $^{31}\text{P}$  coupling constants necessary for a conformational analysis of the molecule, were derived from the  $^1\text{H}$ -spectra using an iterative simulation procedure (see Materials and Methods). Experimental and simulated multiplets, obtained for the spectrum recorded at 298K, are shown in Fig. 2. The resulting J-coupling constants have been collected in Table 1. Since we were not able to make stereospecific assignments for the H5' and H5'' protons it may be possible that the H5'P and H5''P coupling constants and/or the H4'H5' and H4'H5'' coupling constants have to be interchanged. The vicinal  $^{13}\text{C}$ - $^{31}\text{P}$  coupling constants, required for a full conformational analysis, were derived from a proton-decoupled  $^{13}\text{C}$  spectrum (not shown, results are listed in Table 1). The assignment of the resonances in the 1D- $^{13}\text{C}$  spectrum follows from the  $^1\text{H}$ - $^{13}\text{C}$  hetero-correlated spectrum shown in Fig. 3. A UV melting experiment, carried out for cr(GpGpGp) in a temperature range from 278 to 350K, did not show significant changes in the observed absorption, even at NMR concentrations, on increasing the temperature. This indicates that no dimers or aggregates are formed by this molecule in accordance with the results obtained from the NMR experiments mentioned above. (This is in contrast to the results obtained for the dinucleotide cd(ApAp) for which UV melting experiments showed that dimers are formed at NMR concentrations (13).) For the determination of the torsion angle  $\chi$  a NOESY spectrum was recorded. The cross peaks in the 200

**Table 1.** J-coupling constants (in Hz), obtained for cr(GpGpGp) after simulation of the one-dimensional  $^1\text{H}$ - and  $^{13}\text{C}$  spectra.

Coupling constant	value (Hz)
J(H1'H2')	7.5
J(H2'H3')	5.6
J(H2'P3)	0.8
J(H3'H4')	1.5
J(H3'P3)	7.3
J(H4'H5')	3.9
J(H4'H5'')	3.6
J(H4'P5)	1.5
J(H5'P5)	3.7
J(H5''P5)	4.0
J(C2'P3)	5.5
J(C3'P3)	5.6
J(C4'P3)	2.3
J(C4'P5)	11.1
J(C5'P5)	5.0

MHz NOESY spectrum have negative and the diagonal peaks have positive signs (data not shown) indicating rapid rotational motion of the molecule which precludes spin diffusion.

Examination of the NOE intensities shows that the intensity of the cross peak between the H8 and H1' protons is somewhat



**Figure 3.**  $^{13}\text{C}$ - $^1\text{H}$  hetero correlated spectrum of the trinucleotide at 305K. The assignment of the proton- and carbon resonances is indicated. In the f1-direction the C2' resonance position was arbitrarily set to 0 ppm.

larger ( $\sim 20\%$ ) than that of the cross peak between H8 and H2', indicating that torsion angle  $\chi$  is in the syn domain (vide infra).

### Structural analysis

The structural analysis carried out for cr(GpGpGp) relies almost entirely on the J-couplings deduced from the NMR experiments. The first thing that stands out is the large J-coupling between the H1' and H2' sugar protons (Table 1). Pseudorotational analysis shows that the sugar rings adopt a *pure S-type* conformation with a phase angle of pseudorotation, P, equal to  $154 \pm 5^\circ$  and a puckering amplitude,  $\phi_m$ , equal to  $35 \pm 1^\circ$  for the whole temperature range studied. Making use of the relationship (14)  $\delta = 120.6 + 1.1\phi_m \cos(P + 145.2)$  the exocyclic torsion angle  $\delta(\text{C5}'\text{-C4}'\text{-C3}'\text{-O3}')$  can be derived from these results leading to a value of  $139 \pm 5^\circ$ .

The torsion angle  $\gamma(\text{O5}'\text{-C5}'\text{-C4}'\text{-C3}')$  and the fraction of  $\gamma$  gauche(+) conformers was deduced from the H4'-H5' and the H4'-H5'' proton coupling constants (9,14). It is found that in cr(GpGpGp)  $\gamma$  occurs in the gauche(+) domain for 62% of the time irrespective whether two or three staggered conformations are considered. In these calculations the J-couplings for the pure gauche(+), gauche(-) and trans conformation were obtained by using the generalized EOS-Karplus equation (14). Considering the accuracy of the experimental data, the possible alternative assignment of the H5' and H5'' resonances (which requires that the values of the coupling constants  $J(\text{H4}'\text{H5}')$  and  $J(\text{H4}'\text{H5}'')$  are interchanged) leads to the same result for the fraction and the value of  $\gamma$ .

The backbone torsion angle  $\beta(\text{P-O5}'\text{-C5}'\text{-C4}')$  can be derived with the aid of the C4'P coupling constant, which amounts to 11.1 Hz (Table 1). The most recently published parametrization of the Karplus equation for this coupling constant (15) does not account for such a large value. Therefore we derived improved Karplus relations of which the derivation is described in Appendix

**Table 2.** Eight possible conformers of cr(GpGpGp) with backbone torsion angles  $\beta, \gamma, \delta$  and  $\epsilon$  derived from the available experimental J-coupling constants and  $\alpha$  and  $\zeta$  from the ring closure constraints (see text) and their final values obtained after energy minimization without constraints using an all-atom approach. I, II and III denote the first, second and third guanine residue of the ring, respectively.

		$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	$\zeta$	$\chi$	P	$\phi_m$	E*
1. NMR Energy minimization	I	-13	178	53	139	-121	85	syn			$\geq 100$
	II	-10	-179	55	138	-122	85	-150			
	III	-10	-179	55	138	-122	85	-150			
2. NMR Energy minimization	I	167	178	53	139	-121	-55	syn			-109
	II	164	181	51	143	-130	-51	40	157	38	
	III	164	181	51	143	-130	-51	40			
3. NMR Energy minimization	I	-59	178	53	139	-59	61	syn			$\geq 100$
	II	-58	-173	62	139	-68	53	31			
	III	-71	162	66	150	-41	85	30			
4. NMR Energy minimization	I	133	178	53	139	-59	-77	syn			$\geq 100$
	II	139	178	57	139	-63	-74	62			
	III	131	179	51	138	-67	-77	62			
5. NMR Energy minimization	I	45	178	180	139	-121	-33	syn			-95
	II	6	175	179	140	-133	-45	171			
	III	62	156	176	97	-177	77	155			
6. NMR Energy minimization	I	-129	178	180	139	-121	111	syn			93
	II	-141	-161	173	150	-82	68	60			
	III	-141	-161	173	150	-82	68	60			
7. NMR Energy minimization	I	170	178	180	139	-59	81	syn			-97
	II	-169	177	-178	150	-75	83	40	169	38	
	III	-170	178	-177	150	-74	82	40			
8. NMR Energy minimization	I	39	178	180	139	-59	-65	syn			-80
	II	4	-175	179	150	-60	-49	29			
	III	6	-176	178	150	-60	-48	27			

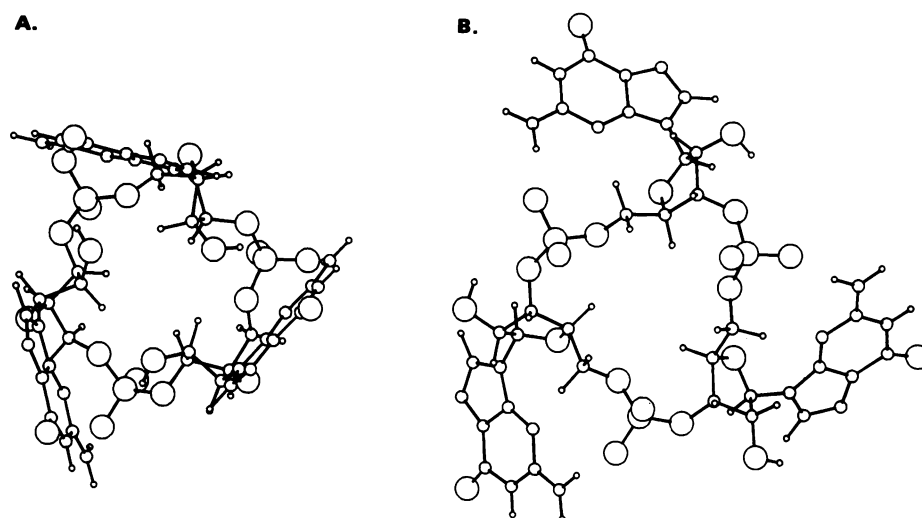
\*(kJ/mole)

1 and  $\beta$  was determined in the course of the parametrization of the CCOP-Karplus parameters (vide infra). It follows that, for the given  $J(\text{C4}'\text{P})$  coupling constant, two values, namely  $\beta = 164^\circ$  and  $\beta = 198^\circ$  satisfy the Karplus equation. The measured H5'P and H5''P coupling constants (Table 1) indicate that  $\beta$ , in addition to being in the trans domain, also adopts a  $\beta(\text{g}+)$  and/or a  $\beta(\text{g}-)$  conformation part of the time (Eqn. 5, Appendix 1). On the basis of Eqns. 6-9 (Appendix 1) we find that  $\beta$  is in the trans domain for 90-95% of the time with  $\beta(t) = 178^\circ$ , consistent with the value of  $J(\text{C4}'\text{P})$ , taking into account the accuracy of the data. An interchange in the assignment of the H5' and H5'' protons and a concomitant change of the values of  $J(\text{H5}'\text{P})$  and  $J(\text{H5}''\text{P})$  does not change this result.

The torsion angle  $\epsilon$  has only been observed to occur in the trans or gauche(-) domain (16);  $\epsilon(\text{g}+)$  is always forbidden because of steric hindrance and is therefore not considered in the subsequent analysis (see Discussion). The torsion angle  $\epsilon$  can be monitored by  $J(\text{C4}'\text{-P3})$ ,  $J(\text{C2}'\text{-P3})$  and  $J(\text{H3}'\text{-P3})$ . Making use of Eqns. 10-14 in Appendix 1 and assuming a possible error of 0.5 Hz in the experimental J-coupling constants, these two possibilities can be distinguished. We obtain a fraction  $\epsilon(t)$  of 0.7, and consequently the fraction  $\epsilon(\text{g}-)$  is 0.3, with  $\epsilon(t) = 240^\circ$  and  $\epsilon(\text{g}-) = 300^\circ$ .

The torsion angles  $\alpha(\text{O3}'\text{-P-O5}'\text{-C5}')$  and  $\zeta(\text{C3}'\text{-O3}'\text{-P-O5}')$  cannot be determined from NMR spectra, but can be derived mathematically (17), on the basis of ring closure constraints, using the values for  $\beta$  and  $\delta$  and the different combinations of torsion angles  $\gamma$  and  $\epsilon$ .

Since the  $^1\text{H}$  and  $^{13}\text{C}$  spectra of cr(GpGpGp) show only one set of resonances for each set of the same type of protons, the ring system has either a 'rigid' symmetrical conformation or it



**Figure 4.** Energy minimized structures obtained for cr(GpGpGp). **A.** Ball stick model of structure 2 from Table II. **B.** idem for structure 7 from Table II.

is averaged to a symmetrical conformation in time. We have made the analysis on the basis that the molecule is symmetric so that the three  $\alpha$ - and the three  $\zeta$  torsion angles each have the same value. This resulted in eight combinations of phosphate backbone torsion angles for which a closed symmetrical 18-ring could be formed (Table 2). Remarkably, of the four different, possible  $\gamma/\epsilon$  combinations always two and only two  $\alpha/\zeta$  combinations permit ring closure. For these eight structures, models were built to serve as starting conformations for energy minimization. The energy minimization was done in three steps. First, the torsion angles  $\chi$  were varied in steps of  $30^\circ$ , subject to the condition that within a molecule each of the three  $\chi$ -angles were kept at equal values and for each of these the energy of all of the eight aforementioned structures was minimized in 300 steps through variation of the backbone angles. In this way we obtained a minimum energy structure for a particular value of  $\chi$ , for each of the initial eight combinations of the phosphate backbone torsion angles. Secondly, the resulting eight energy minimized structures were subjected to a further minimization round until no changes in energy were detected between different steps. Finally, the resulting eight structures were energy minimized using an all-atom approach, without constraining any of the torsion angles. It turned out that all conformers remain approximately symmetric, except for one in which one of the sugar rings adopts an N-type conformation. The results, including the energy values of these structures, are given in Table 2. Ball and stick models of the lowest energy structures of cr(GpGpGp) are given in Fig. 4.

## DISCUSSION

To construct the 'NMR structures', listed in Table 2, the torsion angles  $\alpha$  and  $\zeta$  had to be calculated for the different possible combinations of the backbone torsion angles derived from the J-couplings. This was done algebraically (vide supra, 17). In these calculations the atoms are taken to be point masses. Thus, possible steric clashes leading to high, nonbonded energy terms are not accounted for. For example, in conformer 3 (Table 2), in the

18-ring the C4'-atoms give rise to severe steric hindrance with the C5' atoms of the neighbouring residue on the 3' side. Similar observations can be made for conformers 1, 4 and 6. All of these structures are therefore excluded from further consideration. The remaining conformers have structures which are energetically acceptable.

In the NMR spectra only one set of resonances is seen for each type of protons. This either means that cr(GpGpGp) has a threefold symmetrical structure or that there is rapid interconversion between different conformers, symmetrical or asymmetrical, giving rise to an averaged symmetrical structure. In fact rapid interconversion between different conformers does take place since we have found that  $\gamma$  and  $\epsilon$  and by inference also  $\alpha$  and  $\zeta$  do not occur in a single, staggered conformation. The 'NMR structures' (Table 2) were derived based on the assumption that cr(GpGpGp) is symmetrical and the question arises whether we should have considered asymmetric starting structures as well. From the energy minimizations, which started with symmetric structures, sometimes asymmetric conformers resulted with an acceptable low energy, e.g. conformer 5 (Table 2). From a comparison of its energy with that of the lowest energy conformer (2 in Table 2) and the occurrence of one of the sugars in an N-conformation which is at variance with the NMR data, we estimate that conformer 5 constitutes only a minor fraction and does not affect the NMR parameters. Thus, we tentatively conclude that asymmetrical structures are unfavourable. Because of its relatively high energy the contribution of conformer 8 may be disregarded as well. The two remaining conformers (2 and 7 in Table 2) correspond to the major and minor isomers, i.e. the lowest energy conformer is the major compound. Both conformers are symmetrical, their backbone angles are close to those of the NMR starting structure and the glycosidic torsion angle is in the syn domain. All torsion angles fall within the normal, energetically allowed regions (18), thus no unfavourable stress is induced in the course of the ring formation. In fact the backbone torsion angles of conformer 2 are not very different from those in B-DNA, except for  $\alpha$  which falls in the trans instead of the normal gauche(-) domain and  $\epsilon$  which has shifted

somewhat away from the trans domain. Apparently, this change is sufficient to achieve ring closure for the 18-membered ring.

In the minor component (structure 7),  $\gamma$  is in the trans and  $\epsilon$  in the gauche(-) domain and as a result  $\alpha$  is in the trans and  $\zeta$  in the gauche(+) domain. It can be seen in the ball and stick model of these conformers (Fig. 4) that the lowest energy conformer, 2, is more compact than conformer 7. In conformer 2 the guanine bases are in an axial position with respect to the 18-membered ring, while in 7 they are in a pseudo equatorial position. In both conformations the guanine amino groups are in a position to form a hydrogen bond with the phosphate group in the backbone, which is in correspondence with the NMR results (H8-H1' cross peaks were more intense than H8-H2' cross peaks) which indicates that the guanine bases are in a syn-orientation. This also followed from the molecular mechanics calculations discussed above, i.e.  $\chi(\text{syn})$  is 10 kJ/mole more favourable than  $\chi(\text{anti})$ . All this is in line with results seen in other systems (16).

The folding pattern found in structure 2, i.e. the combination of torsion angles between two adjacent sugar base moieties has been observed in a number of hairpins where sharp turns in the structure are induced by standard nucleoside geometry combined with non-helical torsion angles (9,19-24).

#### Comparison with earlier work on cd(ApAp)

Our earlier investigations on the circular dinucleotide cd(ApAp) demonstrated that at low temperature ( $\approx 5^\circ\text{C}$ ) the molecule occurs in a single conformation in which the deoxyribose rings adopt a pure N-type conformation (13). In a parallel X-ray diffraction study of this compound the same structure was obtained (25). The preferred state for deoxyribose sugars is, however, the S-type conformation and it turns out that in the lowest energy state of cd(ApAp) the N-pucker is imposed by ring closure constraints (vide infra). We were therefore surprised to find that in cr(GpGpGp), with the larger 18-ring system, the ribose sugar is forced into the S-conformation which is unfavourable for riboses. Moreover, within the range considered, the chemical shifts and J-couplings of cr(GpGpGp) are independent of temperature, leading to the result that the riboses stay S-puckered yet at the same time  $\alpha$ ,  $\gamma$ ,  $\epsilon$  and  $\zeta$  jump between different staggered conformations (vide supra). Thus we find a mixture of rapidly interconverting conformations for cr(GpGpGp), of which the fractions do not change in the temperature range studied, and a single conformation for cd(ApAp) which is sensitive to temperature changes, i.e. changes in chemical shifts and J-couplings are observed with increasing temperature. UV experiments, conducted to gain a better insight into these temperature dependent changes of the NMR parameters, showed that cd(ApAp) is involved in a novel type of stacking interaction in which, in the dimer formed by two cd(ApAp) molecules, the bases intercalate without being involved in basepairing. After disruption of the dimer structure the N-pucker may switch to a S-pucker but not necessarily so (13). For the circular cr(GpGpGp), NMR as well as UV melting experiments provide no indication for complex formation, neither for dimer, nor for higher order complexes.

Because the 12-membered ring of cd(ApAp) and the 18-membered ring of cr(GpGpGp) have to be closed and the values of all the torsion angles, except those of  $\alpha$  and  $\zeta$  are known, the possible combinations of values for  $\alpha$  and  $\zeta$  have been calculated with the aid of ring closure rules (17). From these calculations it followed for a symmetrical dinucleotide that for

S-type sugar conformations, perfect ring closure could only be achieved if torsion angle  $\epsilon$  is in a gauche(-) conformation. If  $\epsilon$  is in a trans conformation, which is the case for cd(ApAp) at low temperatures, the sugars have to be N-type to achieve ring closure in a symmetrical 12-ring.

The results are different for a symmetrical 18-membered ring. For an N- as well as an S-type sugar conformation, the ring can be closed for all values of torsion angles  $\gamma$  and  $\epsilon$ , although in the case of N-type sugars the values of torsion angles  $\alpha$  or  $\zeta$  in almost all cases are in an energetically unfavourable, eclipsed range. If S-type sugar conformations are present always two possible combinations of  $\alpha$  and  $\zeta$  values are found with both torsion angles in one of the staggered conformations.

#### Ribose sugars with S-type conformation

Ribose sugars with an S-conformation are relatively rare and occur in somewhat unusual structures (20,21,23,24,26,27) found in RNA hairpin loops or at intercalation sites where the spacing between adjacent bases is doubled. In the circular trinucleotide studied here a rapid equilibrium is observed between  $\epsilon(240^\circ)$  and  $\epsilon(300^\circ)$ , with  $\epsilon$  being in the gauche(-) domain for  $\approx 35\%$  of the time. This corresponds to the transformation from conformer 2 to 7 with the former being the energetically more favourable structure and present at the highest concentration. Surprisingly, the riboses remain S-puckered all the time, while their lowest energy state is the N-pucker (28). There may be different reasons for this behaviour. In the major conformer the syn-orientation of the guanine bases can be held responsible, i.e. the occurrence of N-type sugars is not allowed sterically for a purine nucleotide in which  $\gamma$  is in a gauche(+) conformation and  $\chi$  is in a syn orientation. Furthermore, for the resulting hydrogen bonding of the amino groups to the phosphate groups (in both conformers) an S-puckered sugar is preferred over an N-pucker. Such a conformation contributes a favourable energy term of about 10 kJ/mole (vide supra). When the sugar is in an S-type conformation an 'unhindered' transition from  $\epsilon(240^\circ)$  to  $\epsilon$  gauche(-) may take place leading to a favourable entropy contribution to the free energy. It is noted here that an  $\epsilon$  gauche(-) conformation can only occur in combination with an S-puckered sugar as has been suggested earlier by several authors (15,29,30) on the basis of available crystal structure and NMR data and follows unequivocally from a Ramachandran-type plot of P versus  $\epsilon$  (unpublished results). It is known that a water mediated hydrogen bond between the O2' atom and the 3'-phosphate may contribute to the stability of RNA. Model building shows that the trinucleotide can take advantage of this interaction but because of its cyclic nature the hydrogen bonds can only be formed for S-type sugars. Parallel studies on the circular molecule cd(ApApAp) (unpublished results) show that in this molecule, with deoxyribose instead of ribose sugars, an equilibrium exists between N- and S-type sugars. Because of the lack of the possibility to form the mentioned hydrogen bond, as is possible in RNA, the sugar pucker in this molecule is allowed to be flexible.

#### Comparison with the biologically active molecule cr(GpGp)

Circular r(GpGp) is a regulator of the biological synthesis of cellulose in the bacterium *Acetobacter xylinum*. It has been proposed that its biological function is associated with its ability to bind to the membrane-bound cellulose synthase (3). Relative to cr(GpGp), cr(GpGpGp) activates cellulose synthesis only a few percent. The propensity of the cyclic dinucleotides to form

self-intercalated dimers suggests that these molecules may also bind other aromatic groups and Liaw *et al.* (31) have proposed that the molecules which stimulate cellulose synthesis do this by forming an intercalative complex with aromatic side chains of cellulose synthase. In accordance with this hypothesis we find that the poor regulator cr(GpGpGp) does not form dimers in solution. However, the circular dinucleotide cd(ApAp) is a poor regulator, while it has been shown by us to form self-intercalated dimers in solution (13). In addition, other closely related cyclic dinucleotides, such as cr(ApAp), cr(CpCp), cr(UpUp) and cr(CpGp) are also devoid of stimulating activity. These data leads one to conclude that the possibility to dimerize is apparently necessary but not sufficient for stimulating activity.

The simple lock and key model is thus too restrictive to explain fully the observations and a study of the details of the interactions in the complex is required. Some points can nevertheless be made. In particular the nature of the bases seems to play an important role and not so much the nature of the backbone, since apart from cr(GpGp) all other active regulators are structural analogs namely cr(IpIp), cd(GpGp) and cr(Gp(S)Gp). A comparison of these cyclic nucleotides with the non-active ones suggests that the bases should have an imino group at position 1 and a carbonyl group at position 6. Uracyl has a set of donor and acceptor groups at comparable positions but cr(UpUp) is inactive suggesting that a carbonyl group is not allowed at position 2. These features affect stacking interactions and/or hydrogen bonding interactions with cellulose synthase.

## ACKNOWLEDGEMENTS

This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Research (NWO). Spectra were recorded at the Dutch National Hf-NMR Facility at Nijmegen (the Netherlands). We thank Dr M.C.M.Gribnau for the provision of the optimization routine and Ing. J.J.Joordens for technical assistance.

## REFERENCES

- Hsu, C.-Y.J. and Dennis, D. (1984) *Nucleic Acids Res.*, **10**, 5637–5647.
- Hsu, C.-Y.J., Dennis, D. and Jones, R.A. (1985) *Nucleosides Nucleotides*, **4**, 377–389.
- Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G.A., van Boom, J.H. and Benziman, M. (1987) *Nature*, **325**, 279–281.
- de Vroom, E. (1988) *Ph.D. Thesis* State University Leiden, the Netherlands.
- de Vroom, E., Broxterman, H.J.G., Sliedregt, L.A.J.M., van der Marel, G.A. and van Boom, J.H. (1988) *Nucl. Acids Res.*, **16**, 4607–4620.
- Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
- Chazin, W.J., Wüthrich, K., Hyberts, S., Rance, M., Denny, W.A. and Leupin, W. (1986) *J. Mol. Biol.*, **190**, 439–453.
- Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn. Res.*, **86**, 304–318.
- Blommers, M.J.J., van de Ven, F.J.M., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1991) *Eur. J. Biochem.*, **201**, 33–51.
- Gelin, B.R. and Karplus, M. (1975) *Proc. Natl. Acad. Sci.*, **72**, 2002–2006.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.
- Remin, M. and Shugar, D. (1972) *Biochem. Biophys. Res. Commun.*, **48**, 636–642.
- Blommers, M.J.J., Haasnoot, C.A.G., Walters, J.A.L.I., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1988) *Biochemistry*, **27**, 8361–8369.
- Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas*, **101**, 413–433.
- Lankhorst, P.P., Haasnoot, C.A.G., Erkelens, C. and Altona, C. (1984) *J. Biomolec. Struct. Dyn.*, **1**, 1387–1405.
- Saenger, W. (1984) *Principles of Nucleic Acids Structure*, Springer Verlag, New York.
- Hendrickson, J.B. (1961) *J. Am. Chem. Soc.*, **83**, 4537–4547.
- Pearlman, D.A. and Kim, S.-H. (1986) *J. Biomolec. Struct. Dyn.*, **4**, 69–98.
- Pieters, J.M.L., de Vroom, E., van der Marel, G.A., van Boom, J.H., Koning, T.M.G., Kaptein, R., and Altona, C. (1989) *Biochemistry*, **29**, 788–799.
- Cheong, C., Varani, G. and Tinoco, I. Jr. (1990) *Nature*, **346**, 680–682.
- Varani, G., Cheong, C. and Tinoco, I. Jr. (1991) *Biochemistry*, **30**, 3280–3289.
- Gupta, G., Sarma, M.H. and Sarma, R.H. (1987) *Biochemistry*, **26**, 7715–7723.
- Jack, A., Ladner, J.E., and Klug, A. (1976) *J. Mol. Biol.*, **108**, 619–649.
- Westhof, E., Dumas, P. and Moras, D. (1985) *J. Mol. Biol.*, **184**, 119–145.
- Frederick, C.A., Coll, M., van der Marel, G.A., van Boom, J.H. and Wang, A.H.-J. (1988) *Biochemistry*, **27**, 8350–8361.
- Stout, C.D., Mizuno, H., Rubin, J., Brennan, T., Rao, S.T. and Sundaralingam, M. (1976) *Nucl. Acids Res.*, **3**, 1111–1123.
- van den Hoogen, Y. (1988) *Ph.D. Thesis* State University Leiden, the Netherlands.
- Olson, W.K. (1982) *J. Am. Chem. Soc.*, **104**, 278–286.
- Dhingra, M.M. and Saran, A. (1982) *Biopolymers*, **21**, 859–872.
- Davies, D.B. and Sadikot, H. (1983) *Biopolymers*, **22**, 1843–1847.
- Liaw, Y.-C., Gao, Y.-G., Robinson, H., Sheldrick, G.M., Sliedregt, L.A.J. van der Marel, G.A., van Boom, J.H. and Wang, A.H.J. (1990) *FEBS Lett.*, **264**, 223–227.

## APPENDIX 1

### Reparametrization of the Karplus equations relating the HCOP and CCOP J-coupling constants to torsion angles

A prerequisite for the interpretation of  $^1\text{H}$ - $^{31}\text{P}$  and  $^{13}\text{C}$ - $^{31}\text{P}$  coupling constants in terms of torsion angles is the availability of reliable parameters in the applied Karplus equations. So far, several parametrizations have been used in the literature, for the following type of equations:

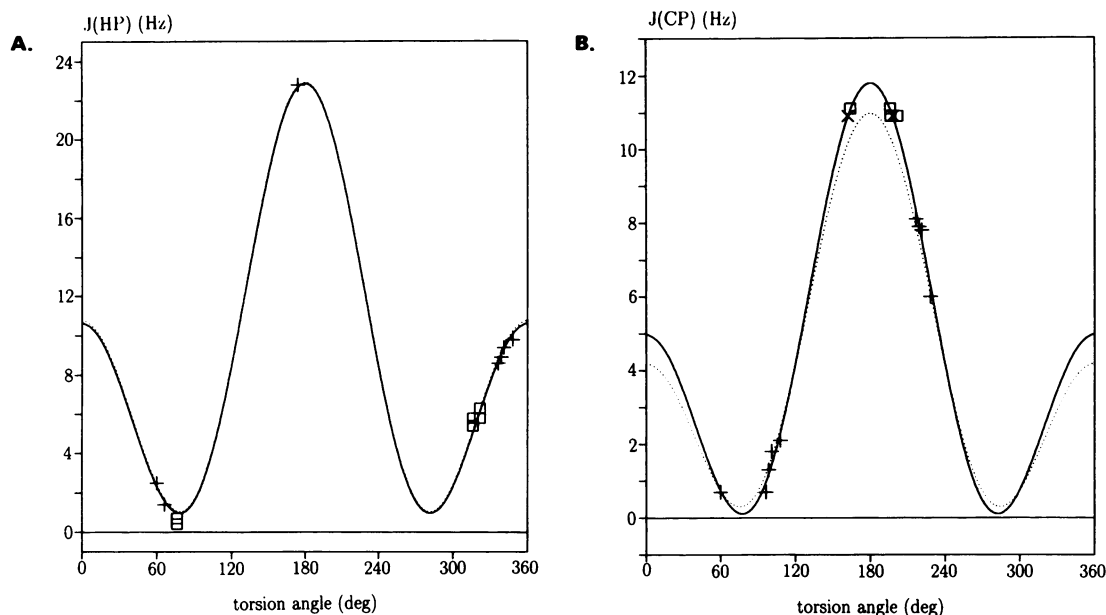
$$^3J_{\text{CCOP}} = A \cos^2(\phi) + B \cos(\phi) + C \quad (1)$$

$$^3J_{\text{HCOP}} = D \cos^2(\phi) + E \cos(\phi) + F \quad (2)$$

The latest values for A, B etc. were deduced by Lankhorst *et al.* (15). However, their parametrization cannot account for the carbon phosphorous coupling constant of 11.1 Hz measured for cr(GpGpGp). Neither can the observation, made earlier in our laboratory, for carbon-phosphorous coupling constants in the circular dinucleotide cd(ApAp), ( $J(\text{C4}'\text{P5})$  and  $J(\text{C4}'\text{P3})$ ), which amounted to 10.9 Hz, be accounted for (13). This is demonstrated in Figs. 5A and B where the J-couplings  $J(\text{HCOP})$ - and  $J(\text{CCOP})$  are plotted as a function of the corresponding torsion angles.

In order to derive a reliable Karplus parametrization the number of calibration points has to be as large as possible. As a starting point we used all but one of the data points used by Lankhorst *et al.* (15), i.e. we did not use the extrapolated value for the  $\beta(\text{trans})$  conformation,  $J(\text{C4}'\text{P}) = 10.9\text{ Hz}$ , which these authors assumed to correspond to a carbon-phosphorous torsion angle of  $180^\circ$ . We did not want to use this value, because small deviations may already influence the parametrization. The measured J-coupling constants and their corresponding unknown torsion angles as derived from (15) are given in Table 3. Of course, it is desirable to have at least one calibration point in the  $0$ – $60^\circ$  range. Unfortunately, a directly measured coupling constant in this range is, to our knowledge, not available and we had to rely on an indirectly estimated  $^{13}\text{C}$ - $^{31}\text{P}$  coupling constant of 0.7 Hz about a torsion angle of  $60^\circ$  (15). From the data obtained for the circular dinucleotide cd(ApAp) (13) we used the coupling constants  $J(\text{C4}'\text{P5})$  and  $J(\text{C4}'\text{P3})$  and their





**Figure 5.** A. HCOP Karplus function as derived by Lankhorst (.....) and as derived in this paper (—). The points indicated by (+) are the original data points used by Lankhorst (15). The data points indicated with a square are values as derived by NMR and X-ray diffraction studies (13,25) for cd(ApAp). B. CCOP Karplus function as derived by Lankhorst (.....) and the one derived in this paper (—). The experimental values used for the parametrization are indicated: the points given by (+) are the original points used by Lankhorst and the data points indicated with a square are the values derived for cd(ApAp) and for cr(GpGpGp). The exact value of the torsion angle  $\beta$ , corresponding to the C4'P5 coupling constant of cr(GpGpGp) was unknown before the reparametrization. The (x) denotes value of  $\beta$  earlier assigned to  $180^\circ$  by Lankhorst *et al.* (15).

**Table 3.** In- and output of the reparametrization of the HCOP- and CCOP Karplus equations (15).

Coupling constant	Input		Output		
	$J_{\text{obs}}$ (Hz)	torsion angle ( $^\circ$ )	$J_{\text{calc}}$ (Hz)	torsion angle ( $^\circ$ )	$J_{\text{calc}} - J_{\text{obs}}$ (Hz)
C2'-P3	1.8 [17]	$\phi_1 - 120$	1.4	101	-0.4
C2'-P3	1.3 [17]	$\phi_2 - 120$	1.2	99	-0.1
C2'-P3	0.7 [17]	$\phi_3 - 120$	0.9	96	+0.2
C2'-P3	2.1 [17]	$\phi_4 - 120$	2.3	108	+0.2
C4'-P3	7.8 [17]	$\phi_1$	7.5	221	-0.3
C4'-P3	7.9 [17]	$\phi_2$	8.0	219	+0.1
C4'-P3	8.1 [17]	$\phi_3$	8.3	217	+0.2
C4'-P3	6.0 [17]	$\phi_4$	6.3	228	+0.3
C4'-P3	10.9 [14]	202	10.5	202 $\diamond$	-0.4
C-P	10.9 [17]	$\phi_5$	10.9	198/162	0.0
C-P ( $60^\circ$ )	0.7 [17]	60	0.7	60 *	0.0
C4'-P5	10.9 [14]	196	11.1	196 $\diamond$	+0.2
C4'-P5	11.1 [ $\Delta$ ]	$\phi_6$	11.1	196/164	0.0
H3'-P	9.4 [17]	$\phi_1 + 120$	9.4	341	0.0
H3'-P	8.9 [17]	$\phi_2 + 120$	9.1	339	+0.2
H3'-P	8.6 [17]	$\phi_3 + 120$	8.7	336	+0.1
H3'-P	9.8 [17]	$\phi_4 + 120$	10.1	348	+0.3
H3'-P	5.8 [14]	316	6.1	316 $\diamond$	+0.3
H3'-P	6.3 [14]	322	6.1	322 $\diamond$	-0.2
H5'-P	0.4 [14]	76	0.9	76 $\diamond$	+0.5
H5'-P	0.7 [14]	76	0.9	76 $\diamond$	+0.2
H5'-P	5.8 [14]	322	6.1	322 $\diamond$	+0.3
H5'-P	5.4 [14]	316	5.0	316 $\diamond$	-0.4
H-P ( $174^\circ$ )	22.8 [17]	174	22.7	174 *	-0.1
H-P ( $66^\circ$ )	1.4 [17]	66	1.6	66 *	+0.2
H-P ( $60^\circ$ )	2.5 [17]	60	2.3	60 *	-0.2

A comparison is made between the measured J-coupling constants,  $J_{\text{obs}}$ , and the J-coupling constants as calculated,  $J_{\text{calc}}$ , from the newly parametrized Karplus equations. The torsion angles marked by an asterisk (\*) were kept fixed during optimization and were first introduced in (15). The torsion angles marked by ( $\diamond$ ), which were determined for cd(ApAp) by NMR as well as X-ray studies (13,25), were also kept fixed during optimization.  $\Delta$  indicates the value of the torsion angle derived in this work for the title compound.

corresponding known torsion angles  $\beta$  and  $\epsilon$  as input for the reparametrization as well as a set of H-P coupling constants, namely  $J(\text{H3}'\text{P})$ ,  $J(\text{H5}'\text{P})$  and  $J(\text{H5}''\text{P})$  at two different temperatures. At this point ten carbon-phosphorous and seven

proton-phosphorous coupling constants of pure conformers were taken from Lankhorst *et al.* (15), whereby five of the torsion angle values corresponding to these coupling constants are unknown and incorporated as such in the calculations. Two carbon-phosphorous and six proton-phosphorous coupling constants are from Blommers *et al.* (13), whereby the values of the corresponding torsion angles are known from NMR and X-ray data (Table 3). From the work described here one carbon-phosphorous coupling constant has been derived; the value for the corresponding torsion angle  $\beta$  is unknown (Table 3). The  $\text{H5}'\text{P}$  and  $\text{H5}''\text{P}$  coupling constants derived for cr(GpGpGp) could not be used, because we were not able to make stereospecific assignments for the  $\text{H5}'$  and the  $\text{H5}''$  signals; neither could the  $\text{C2}'\text{P3}$  nor the  $\text{H3}'\text{P3}$  coupling constant of cr(GpGpGp) be employed for the reparametrization, because they do not correspond to a pure conformational state (*vide supra*). This results in a total of 26 calibration points (Table 3) on the basis of which two three-parameter Karplus equations and the unknown magnitudes of six torsion angles (five torsion angles from the compounds of (15) and torsion angle  $\beta$  from cr(GpGpGp)) could be derived. This has been achieved by a least squares optimization routine using a conjugate gradient minimization method. The resulting parameters are given in the following equations:

$${}^3J_{\text{CCOP}} = 8.0 \cos^2(\phi) - 3.4 \cos(\phi) + 0.5 \quad (3)$$

$${}^3J_{\text{HCOP}} = 15.3 \cos^2(\phi) - 6.2 \cos(\phi) + 1.5 \quad (4)$$

Plots of these equations, as a function of torsion angle  $\phi$ , are shown in Figs. 5A and B together with the 26 experimental calibration points. The fit is excellent as can be seen in Table 3. As mentioned above in the  $0-60^\circ$  region only one estimated calibration point for the carbon-phosphorous Karplus equation

was available and the reliability in this region is therefore less than that of the remaining part of the curve. We note that the difference between the previous proton-phosphorous Karplus equation and the new one is negligible. On the basis of the new parametrization the following equations are applicable to the J-couplings and fractions of  $\beta$  and  $\gamma$  conformations.

$$x_{\beta\text{trans}} = (25.5 - J_{\text{H5}'\text{P}} - J_{\text{H5}''\text{P}})/20.5 \quad (5)$$

$$J_{\text{H5}'\text{P}} = x_g + J_{\text{H5}'\text{P}}(60 - 120^\circ) + x_t J_{\text{H5}'\text{P}}(\beta - 120^\circ) + x_g J_{\text{H5}'\text{P}}(300 - 120^\circ) \quad (6)$$

$$J_{\text{H5}''\text{P}} = x_g + J_{\text{H5}''\text{P}}(60 + 120^\circ) + x_t J_{\text{H5}''\text{P}}(\beta + 120^\circ) + x_g J_{\text{H5}''\text{P}}(300 + 120^\circ) \quad (7)$$

$$J_{\text{C4}'\text{P5}} = x_g + J_{\text{C4}'\text{P5}}(60^\circ) + x_t J_{\text{C4}'\text{P5}}(\beta) + x_g J_{\text{C4}'\text{P5}}(300^\circ) \quad (8)$$

$$x_{\beta(g+)} + x_{\beta(t)} + x_{\beta(g-)} = 1 \quad (9)$$

$$J_{\text{C4}'\text{P3}} = x_t J_{\text{C4}'\text{P3}}(\epsilon^t) + x_g J_{\text{C4}'\text{P3}}(\epsilon^g) \quad (10)$$

$$J_{\text{C2}'\text{P3}} = x_t J_{\text{C2}'\text{P3}}(\epsilon^t + 120^\circ) + x_g J_{\text{C2}'\text{P3}}(\epsilon^g + 120^\circ) \quad (11)$$

$$J_{\text{H3}'\text{P3}} = x_t J_{\text{H3}'\text{P3}}(\epsilon^t - 120^\circ) + x_g J_{\text{H3}'\text{P3}}(\epsilon^g - 120^\circ) \quad (12)$$

$$x_{\epsilon(t)} + x_{\epsilon(g-)} = 1 \quad (13)$$