# NMR and molecular modelling studies of the DNA duplex formed in an oligothymidine/oligolysine complex

# NMR and molecular modelling studies of the DNA duplex formed in an oligothymidine/oligolysine complex

by Marcel H.P. van Genderen, Leo H. Koole and Henk M. Buck

*Department of Organic Chemistry, Eindhoven University of Technology, Eindhoven, the Netherlands*

Communicated at the meeting of November 30, 1987

SUMMARY

Oligothymidine strands form a parallel duplex based on thymine-thymine base pairs after methylation of the phosphate groups. A similar duplex can be obtained by shielding of the phosphate charges with the polycationic protein poly-L-lysine, as we have determined earlier with UV hyperchromicity experiments. Now we present $^{1}H$ and $^{31}P$ NMR studies on the model system $d(T_{10})/Lys_{18}$ that corroborate this model. Binding of the protein to the phosphate groups of the DNA was ascertained from an upfield shift in the $^{31}P$ spectrum. The presence of a DNA duplex structure was deduced from an imino chemical shift of 13.0 ppm, which indicates thymine-thymine base pairing. With variable temperature $^{31}P$ studies, a melting temperature of 30°C was found in water, agreeing with UV hyperchromicity experiments. Preliminary NOESY measurements show a right-handed helix structure for the DNA, and several DNA-protein NOEs have been established. The experimental data are in agreement with a triple-stranded DNA/protein helix with frayed ends. Based on this model, a molecular modelling study was performed to predict the protein conformation that accommodates a parallel DNA duplex.

## INTRODUCTION

The stability of DNA duplexes is enhanced considerably when the interstrand electrostatic repulsions between the phosphate groups are minimized. This can be achieved completely by methylation of the phosphate diester groups (Koole et al., 1987a). In addition, new structures are possible under conditions of phosphate shielding. Phosphate-methylated oligothymidine strands have been shown to associate in a parallel fashion (i.e. the $5' \rightarrow 3'$ vectors in the backbones run in the same direction) based on thymine-thymine (T-T) base pairing (Koole et al., 1987b). The shielding of the phosphate charges can also be achieved in

a different way, viz. by complexation with positive sites of a protein. For antiparallel DNA duplexes, it is known that addition of the polycationic protein poly-L-lysine increases the duplex stability, expressed as the duplex⇄coil transition temperature $T_m$ (Tsuboi, 1967). We have shown recently (van Genderen et al., 1987) that complexation of poly-L-lysine with natural $d(T_8)$ leads to the formation of a DNA duplex with a $T_m$ value of 30°C, as determined with UV hyperchromicity experiments. The most likely structure for this system is a parallel DNA duplex with T-T base pairs, around which a poly-L-lysine winds itself in one of the grooves (analogous to Feughelman et al., 1955). We now report $^{31}P$ and $^{1}H$ NMR spectroscopic investigations that corroborate this model, and present a molecular modelling study of the poly-L-lysine conformation in the complex with DNA.

## RESTULTS AND DISCUSSION

For the NMR studies, we chose a well-defined model system, consisting of the DNA $d(T_{10})$ and the protein octadecalysine ($Lys_{18}$). In the structural model, the lysine side chains alternatingly point up and down toward the phosphate groups on either side of the groove (see figure 1).

In the model system, each $Lys_{18}$ molecule can exactly associate with the eighteen phosphate groups in two base-paired $d(T_{10})$ strands, giving a stoichiometry of 2 : 1 (DNA : protein). In order to test whether the DNA strands from a duplex structure in the $d(T_{10})/Lys_{18}$ system, we measured the $^{1}H$ chemical shift of the imino protons in the thymine bases. The imino protons are involved in the hydrogen bonding of thymine, and therefore a change is seen from a chemical shift of 11.2 ppm for free thymine bases (Haasnoot et al., 1980) to 13.5 ppm upon T-T base pair formation (Koole et al., 1987b). At a sample temperature of 20°C, the imino resonances of the $d(T_{10})/Lys_{18}$ complex were found at 12.6 ppm, and at 10°C they are located at 13.0 ppm. This clearly indicates DNA duplex formation with T-T base pairing.

With $^{31}P$ NMR spectroscopy, it was established that this DNA duplex is associated with the $Lys_{18}$ via the phosphate groups. A solution of $d(T_{10})$ in water at room temperature showed a sharp resonance at 1.93 ppm in the $^{31}P$ NMR spectrum, indicating that all phosphate groups in the single-stranded system are nearly identical. Upon addition of an equivalent amount of $Lys_{18}$ to the NMR sample, a second resonance at 1.85 ppm appeared. After ca. 1h, the $^{31}P$ spectrum consisted of two peaks (see figure 2), a minor one at 1.93 ppm and a major one at 1.85 ppm.

The upfield resonance is broadened (5.5 Hz wide vs. 3.6 Hz for the downfield peak), which indicates a more rigid DNA structure, since a reduced rotational motion of the molecule results in a shortened $^{31}P$ transversal relaxation time (Gorenstein, 1981). Therefore, it can be concluded that the 1.85 ppm resonance is due to phosphate groups that are bound with a lysine residue. A similar upfield shift of the $^{31}P$ resonance has been observed by Davanloo and Crothers (1979) for the association of the duplex of $d(A_3GCT_3)$ with tetralysine, and might arise from changes in the electron density of the phosphorus atom as a

result of ionic interactions with the ammonium groups of $Lys_{18}$. Determination of the surface areas of the two $^{31}P$ signals with a curve-fitting program revealed that 24% of the phosphates is not associated with protein (downfield resonance). Addition of extra $Lys_{18}$ to the NMR sample did not change the $^{31}P$ spectrum, so the downfield signal can not be caused by remaining free $d(T_{10})$ strands, but free phosphates must be present in the $d(T_{10})/Lys_{18}$ system. Since eighteen phosphate groups are present in the complex, four of them (24%) are therefore not bound to protein. They most likely represent the relatively freely moving ends of the duplex, where association with protein is disturbed by a higher mobility of the phosphate groups. In each DNA strand, the two phosphates at each end remain uncomplexed, giving a total of four.

The $d(T_{10})/Lys_{18}$ complex was also studied with variable temperature $^{31}P$

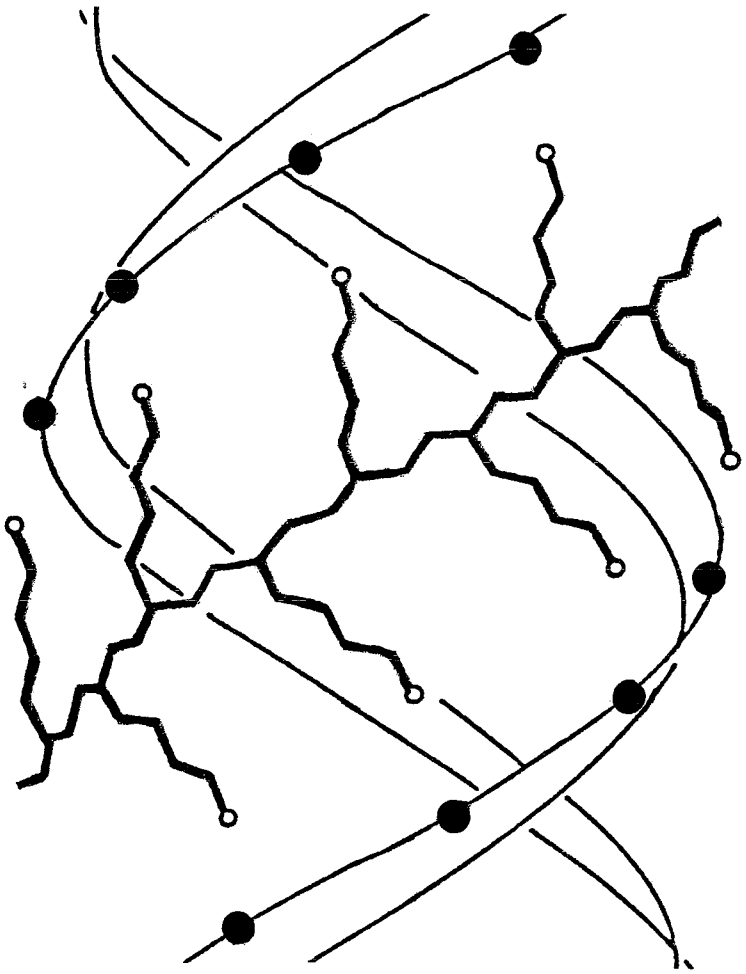

Fig. 1. Schematic structural model for the DNA/lysine complex. The protein molecule is drawn with heavy lines. Ammonium groups are indicated with open circles, and phosphate groups with filled circles.

NMR spectroscopy. Upon raising the sample temperature, the DNA duplexes dissociate into single strands of $d(T_{10})$. Since each DNA duplex is stabilized by one protein molecule, only one of the strands can remain associated with the $Lys_{18}$. Therefore, the intensity of the downfield $^{31}P$ signal (free phosphates) must increase at higher temperatures. This melting behaviour was indeed

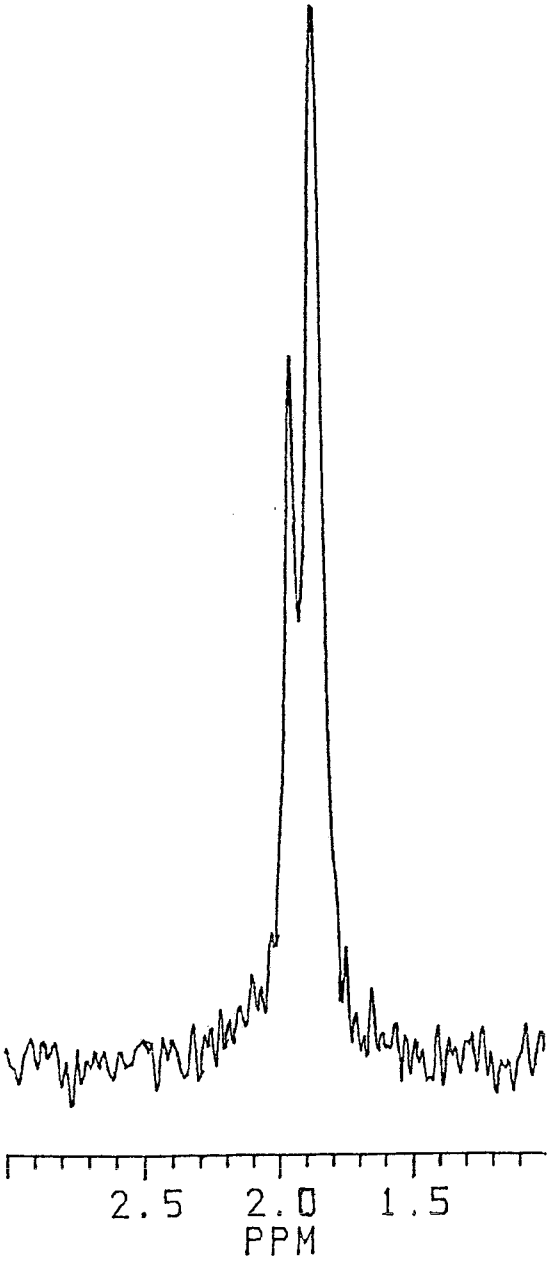


Fig. 2. $^{31}P$ NMR spectrum of the $d(T_{10})/Lys_{18}$ complex at room temperature.

observed (see figure 3), and a $T_m$ value of 30°C can be deduced in a salt-free medium.

It is of interest, that we also investigated the $d(T_{10})$/poly-L-lysine complex with UV hyperchromicity experiments. A $T_m$ value of 34°C was found in a low-salt (6 mM $MgCl_2$/4 mM Tris HCl) buffer solution, which is in good agreement with the present results. In figure 3, it can be seen that the intensity ratio of the resonances approaches unity at elevated temperatures, indicating that half of the $d(T_{10})$ strands remain complexed with $Lys_{18}$, which is in accordance with the stoichiometry of the DNA/protein system. The experiment was repeated with extra $Lys_{18}$ added. At room temperature, this gave no change in the $^{31}P$ spectrum (vide supra), but after dissociation of the DNA duplex the resulting free $d(T_{10})$ strands could now bind the extra protein. Indeed, a lower intensity ratio was observed at elevated temperatures (see figure 3), showing that less free DNA is present. These results support the present structural model (vide supra), consisting of a triple-stranded helix (two $d(T_{10})$ strands and one $Lys_{18}$ strand) with frayed ends. Further structural investigations have been performed on the $d(T_{10})$/$Lys_{18}$ system with two-dimensional nuclear Overhauser effect (NOE) measurements, which show the spatial proximity of protons within ca. 5 Å. Several NOE contacts have been identified, falling in two classes: NOEs within the DNA duplex, and DNA-protein NOEs. The presence of NOE cross-peaks between the base proton $H_6$ and the deoxyribose protons $H_{1'}$, $H_{2'}$, $H_{2''}$, and $H_{3'}$, give information concerning the conformation of the DNA duplex (see Table 1).

Table 1. Observed NOE contacts in the $d(T_{10})$/$Lys_{18}$ complex, expressed relative to the $H_6$-$CH_3$ NOE in the thymine base.

| Intra-DNA | | DNA-protein | |
|---|---|---|---|
| NOE contact | relative intensity | NOE contact | relative intensity |
| $H_6$-$H_{1'}$ | 0.13 | $H_{5'}/_{5''}$-$H_\epsilon$ | 0.35 |
| $H_6$-$H_{2'}$ | 0.25 | $H_{5'}/_{5''}$-$H_\delta$ | 0.35 |
| $H_6$-$H_{2''}$ | 0.09 | $H_{4'}$-$H_\epsilon$ | 0.29 |
| $H_6$-$H_{3'}$ | 0.15 | | |

The dominance of the $H_6$-$H_{2'}$, NOE over the $H_6$-$H_{1'}$, NOE clearly shows an anti conformation of the thymine base. The relative intensities of the other NOEs are consistent with a $C_{2'}$,-endo sugar pucker (Sarma et al., 1986). Both conformational characteristics are a prerequisite for a right-handed DNA duplex structure. This is in good agreement with the fact that phosphate-methylated parallel structures also exhibit a right-handed duplex (Koole et al., 1987b). The DNA duplex is in close proximity with parts of the protein molecule, as can be seen in several DNA-protein NOEs (see Table 1). The observed contacts occur near the end of the lysine side chain ($H_\epsilon$ and $H_\delta$) and in the DNA backbone ($H_{4'}$ and $H_{5'}/_{5''}$), which confirms the association of the lysine ammonium groups with the DNA phosphates. The present NOE results

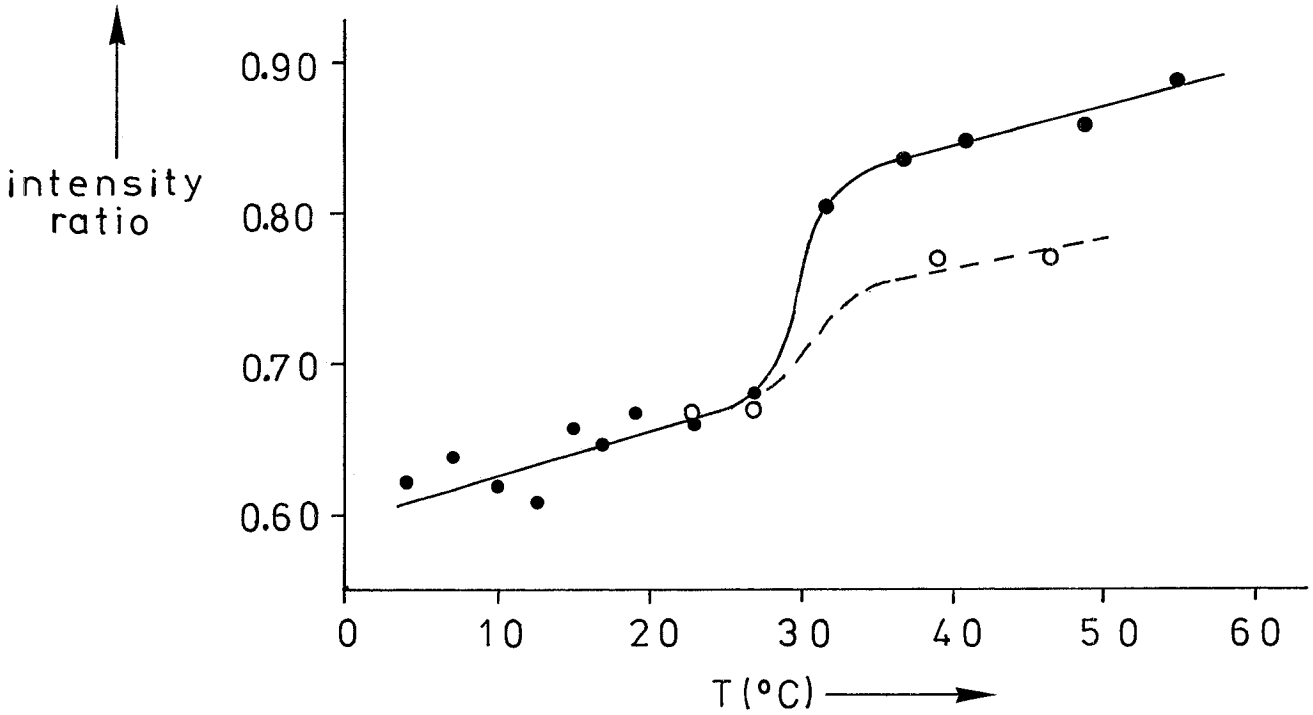


Fig. 3. Intensity ratio of the upfield to the downfield $^{31}P$ resonances in the $d(T_{10})/Lys_{18}$ complex as a function of the sample temperature, for a 2 : 1 stoichiometry (●) and with extra $Lys_{18}$ added (○).

Table 2. Conformation of an oligolysine chain that fits the parallel DNA duplex. Definitions of the torsion angles are given in the text.

| Backbone | | Side Chains | | | |
|---|---|---|---|---|---|
| $\phi_1$ | −160° | $\chi_1$ | 180° | $\chi_2$ | 180° |
| $\psi_1$ | 173° | $\chi_1'$ | 180° | $\chi_2'$ | 60° |
| $\omega$ | −170° | $\chi_1''$ | 180° | $\chi_2''$ | 60° |
| $\phi_2$ | 160° | $\chi_1'''$ | 180° | $\chi_2'''$ | 180° |
| $\psi_2$ | −173° | | | | |

point at a right-handed DNA duplex, complexed with a $Lys_{18}$ molecule of which the side chains are close to the DNA backbone in order to associate effectively with the phosphate groups.

Besides the experimental structural investigations, molecular modelling studies have been performed with the Chem-X program to predict the protein conformation that accommodates a DNA duplex. By fitting an oligolysine chain to a parallel DNA helix with T-T base pairs, it was determined that a repeating unit of two lysine amino acids must be present. The backbone conformation in this unit is given by the torsion angles $\phi$ (NH-$C_\alpha$), $\psi$ ($C_\alpha$-C(O)), and $\omega$ (C(O)-NH) (IUPAC-IUB, 1970). As can be seen in Table 2, a relatively small deviation from the all-trans situation is sufficient for the protein to assume a helical structure with 20 lysine residues per turn, a diameter of 20 Å, and a rise of 38 Å per turn.

For the side chains bearing the ammonium groups, a simple all-trans conformation is not enough to allow association with phosphate groups on both DNA strands. At least one of the chains in each repeating unit must adopt a bent structure, as can be seen in Table 2 ($\chi_{1/2}$ through $\chi_1'''_2$ represent the C-C torsion angles in both side chains going from the backbone toward the ammonium group). Although a more symmetrical orientation of the side chains may be possible, it is evident that the protein must adapt itself considerably to conform with the DNA helix. Undoubtedly, the DNA duplex itself will also change its conformation in the complexation with the protein. The present model will therefore serve as the starting point for AMBER molecular mechanics calculations, which will refine the structure of the DNA/protein complex. It is of interest to note, that we recently observed the formation of a parallel duplex stabilized by polyornithine, which differs from polylysine only in the length of the side chains (three vs. four carbon atoms). Polyornithine gives a more stable duplex structure, and therefore seems the fit the DNA better.

ACKNOWLEDGEMENT

Use of the services and facilities of the Dutch CAOS/CAMM Center, under grant numbers SON-11-20-700 and STW-NCH-44.0703, is gratefully acknowledged. We wish to thank Mr. M. Moody for performing the molecular modelling studies.

MATERIALS AND METHODS

The DNA $d(T_{10})$ was synthesized on a 15-$\mu$mol scale on an Applied Biosystems 380 A DNA synthesizer, and purified by alcohol precipitation. The octadecalysine was purchased as the hydrobromide salt from Sigma Chemical Co. Analysis with gel filtration through a Sephadex G-50 column showed the octadecamer to be indeed the major constituent. $^{31}P$ and $^{1}H$ NMR measurements were performed on a Bruker AC-200 spectrometer, interfaced with an ASPECT-3000 computer. Measurements in water refer to a $H_2O/D_2O$ mixture (85 : 15, v/v), with deuterium serving as the field-frequency lock. The strong $H_2O/HDO$ signal was suppressed with the CWEFT technique (Haasnoot and Hilbers, 1983). Chemical shifts of $^{31}P$ resonances were referenced to an 85% $H_3PO_4$ solution ($\delta$ 0 ppm), and were designated positive when downfield of the standard. Chemical shifts of $^{1}H$ resonances were referenced to the $H_2O/HDO$ peak ($\delta$ 4.68 ppm) before suppression. For the two-dimensional NOE spectra, a spectral width of 1350 Hz was used. Each of the 512 experiments in the $f_1$ direction was recorded with 64 transients, 2K data points, and a relaxation delay of 2 s. A mixing time of 200 ms was used, and spectra were recorded in the absolute magnitude mode. Free induction decays were weighed with a squared-sine function before Fourier transformation.

REFERENCES


Davanloo, P. and D.M. Crothers - Nuclear magnetic resonance investigation of lysine-oligopeptides and a complex with $d(pA)_3pGpC(pT)_3$. Biopolymers **18**, 2213-2231 (1979).

Feughelman, M., R. Langridge, W.E. Seeds, A.R. Stokes, H.R. Wilson, C.W. Hooper, M.H.F. Wilkins, R.K. Barclay and L.D. Hamilton - Molecular structure of deoxyribose nucleic acids and protein. Nature **175**, 834-838 (1955).

Genderen, M.H.P. van, L.H. Koole and H.M. Buck - DNA duplex formation of the $d(T_8)$/polylysine complex by shielding of the phosphate groups. Proc. Kon. Ned. Akad. van Wetensch. **B 90**, 181-187 (1987). Communicated by H.M. Buck at the meeting of June 22, 1987.

Gorenstein, D.G. - Nucleotide conformational analysis by $^{31}P$ nuclear magnetic resonance spectroscopy. Ann. Rev. Biophys. Bioeng. **10**, 355-386 (1981).

Haasnoot, C.A.G., J.H.J. den Hartog, J.F.M. de Rooij, J.H. van Boom and C. Altona - Loop structures in synthetic oligonucleotides. Nucleic Acids Res. **8**, 169-181 (1980).

Haasnoot, C.A.G. and C.W. Hilbers - Effective water resonance suppression in 1D- and 2D-FT-proton-NMR spectroscopy of biopolymers in aqueous solution. Biopolymers **22**, 1259-1266 (1983).

IUPAC-IUB Commission on Biochemical Nomenclature 1969 - Abbreviations and symbols for the description of the conformation of polypeptide chains. Biochemistry **9**, 3471-3479 (1970).

Koole, L.H., M.H.P. van Genderen, R.G. Reiniers and H.M. Buck - Enhanced stability of a Watson & Crick DNA duplex by methylation of the phosphate groups in one strand. Proc. Kon. Ned. Akad. van Wetensch. **B 90**, 41-46 (1987). Communicated by H.M. Buck at the meeting of December 15, 1986.

Koole, L.H., M.H.P. van Genderen and H.M. Buck - A parallel right-handed duplex of the hexamer d(TpTpTpTpTpT) with phosphate triester linkages. J. Am. Chem. Soc. **109**, 3916-3921 (1987).

Sarma, M.H., G. Gupta and R.H. Sarma - 500-MHz proton NMR study of poly(dG)·poly(dC) in solution using one-dimensional nuclear Overhauser effect. Biochemistry **25**, 3659-3665 (1986).

Tsuboi, M. - Helical complexes of poly-L-lysine and nucleic acids. In Conformation of Biopolymers, Academic Press, New York. Vol. II, pp. 689-702 (1967).