Peptide-Induced Parallel DNA Duplexes for Oligopyrimidines. Stereospecificity in Complexation for Oligo(L-lysine) and Oligo(L-ornithine)

Marcel H. P. van Genderen,* Martin P. Hilbers, Leo H. Koole, and Henk M. Buck

Department of Organic Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Received March 6, 1990; Revised Manuscript Received May 9, 1990

ABSTRACT: It is shown that the cationic oligopeptides octadeca(L-lysine) (Lys_{18}) and octadeca(L-ornithine) (Orn_{18}) can induce a parallel duplex for the natural DNA oligomer dT_{10} with thymine–thymine base pairs. Complexation of the ammonium groups in the peptide side chains with the DNA phosphates leads to diminished electrostatic phosphate–phosphate repulsions, which allows this T-T base pair formation. From combined NOESY ^1H NMR and molecular mechanics studies, it follows that the parallel duplex is right-handed, with the peptide located in the groove of the duplex. For the natural DNA oligomers dC_{10}, $d(C_6T_6)$, and $d(T_6C_2T_2)$, only Lys_{18} is able to induce the formation of parallel duplexes with C-C and T-T base pairs. It is shown that, for Orn_{18}, a complexation must occur with one of the nonbonded oxygen atoms in the phosphate groups (O_R) in such a way that unfavorable steric interactions are present with the C-C base pairs, which have a larger propellor twist angle than T-T base pairs. An analogy is presented between peptide complexation with the phosphates and the neutralization of the phosphate groups by methylation, which is known to lead to parallel duplexes with T-T base pairs (for both the S_P and R_P configurations) and C-C base pairs (only for the S_P configuration).

Parallel DNA duplexes, in which the 5′ → 3′ vectors in the backbones run in the same direction, have been reported for natural oligodeoxynucleotides under extreme conditions (e.g., a highly acidic medium; Sarma et al., 1986), or for modified oligodeoxynucleotides. For instance, recently it was reported that, in the stem of hairpin structures with an artificial 3′–3′ or 5′–5′ linkage, a parallel arrangement of the strands can be forced (Germann et al., 1989). Also, the duplexes between oligodeoxynucleotides with α-oriented bases and natural DNA are found to be parallel (Morvan et al., 1987; Thuong et al., 1987). Spontaneous formation of parallel duplexes has been found for phosphate-methylated pyrimidine oligomers (Koole et al., 1986, 1987; Quaedflieg et al., 1990a). Methylation of the phosphate groups leads to slim, right-handed duplexes with pyrimidine–pyrimidine base pairs, due to the elimination of interstrand phosphate–phosphate charge repulsions. For phosphate-methylated thymidine oligomers (dT_n, n = 2–8) with thymine–thymine (T-T) base pairs, the configuration of the chiral phosphorus atom in the methyl phosphotriester group does not influence the duplex stability (Koole et al., 1987), even though for the S_P configuration the methyl points into the solvent while for the R_P configuration the methyl group is located in the groove of the helix (van Genderen et al., 1987a). However, for the phosphate-methylated nucleotides d(CpC) and d(TpC), miniduplexes with cytosine–cytosine (C-C) base pairs (and in the latter case also a T-T base pair) are formed exclusively for the S_P configuration (Koole et al., 1988; Quaedflieg et al., 1990a). Molecular mechanics calculations have shown that the C-C base pair has a considerably larger propellor twist angle between the bases compared to the T-T base pair, leading to unfavorable steric interactions with the phosphate methyl group in the R_P configuration only. Similar observations were recently made with the phosphate-methylated and 2′-O-methylated RNA dinucleotide r(CpU), which forms a parallel miniduplex only for the S_P configuration (Quaedflieg et al., 1990b).

Since it is of interest whether parallel duplexes with pyrimidine–pyrimidine base pairs also occur for unmodified DNA, we have previously studied complexation of the natural oligomers dT_8 and dT_{10} with cationic oligopeptides, such as oligo(L-lysine) and oligo(L-ornithine), which possess terminal ammonium groups on side chains of four and three methylene groups, respectively (van Genderen et al., 1987b, 1988). Cationic oligopeptides are known to bind electrostatically to the phosphate groups in antiparallel DNA duplexes and raise their stability by shielding of the phosphate charges (Tsuboi, 1967; von Hippel & McGhee, 1972). We were indeed able to demonstrate that both natural dT_8 and dT_{10} form duplex structures when either oligo(L-lysine) or oligo(L-ornithine) is added. In the present paper, we use high-resolution NMR[1] techniques to establish the parallel nature of these duplexes. Furthermore, peptide-induced parallel duplexes with T-T *and* C-C base pairs are characterized with NMR and UV spectroscopic studies and molecular mechanics studies. It is found that only oligo(L-lysine) is able to induce duplex formation when cytosine bases are present in the oligomer. This is explained by a difference in ammonium–phosphate complexation for oligo(L-lysine) and oligo(L-ornithine), which leads to steric interactions for the combination oligo(L-ornithine)/C-C base pair similar to those found for the R_P configuration in the phosphate-methylated compounds.

Materials and Methods

Oligonucleotides. The oligonucleotides were synthesized on a 10-μmol scale on an Applied Biosystems 381A DNA synthesizer, using the standard β-cyanoethyl phosphoramidite protocol. After removal from the column and deprotection in ammonia, the compounds were purified by alcohol precipitation at −20 °C. For the short oligomers we use [decamers dT_{10}, dC_{10}, and $d(T_6C_2T_2)$ and dodecamer $d(C_6T_6)$], the following approach proved to be the most efficient: to 80 μL

[1] Abbreviations: AMBER, assisted model building and energy refinement; EDTA, ethylenediaminetetraacetic acid; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TE, Tris-HCl/EDTA; UV, ultraviolet.

0006-2960/90/0429-7838$02.50/0

of DNA solution, 20 μL of a 3 M sodium acetate/acetic acid buffer solution (pH 5.6) and 900 μL of 96% ethanol were added. Overnight precipitation, centrifugation at 14000 rpm for 15 min, and drying afforded DNA pellets in ca. 50% yield, based on UV absorption at 260 nm (assuming a single-strand concentration of 33 μg/mL to give an A_{260} of 1; Maniatis et al., 1982).

Proteins. Oligo(L-lysine) (average molecular weight 3700, degree of polymerization 18) and oligo(L-ornithine) (average molecular weight 6200, degree of polymerization 32) were purchased as the hydrobromide salts from Sigma Chemical Co. Peptides with an average length of 18 residues were needed, since a duplex of the DNA decamers contains 18 phosphate groups to be shielded. The oligo(L-lysine) was used as received, and a stock solution was made by weighing. The oligo(L-ornithine) mixture was fractionated on a Sephadex G25 superfine gel filtration column, with an 0.1 M HCl solution as eluent. Fractions of 2 mL were collected at a flow rate of 17 mL/h. The average oligomer length n in the fractions could be calculated with an accuracy of 0.1 from the ^{1}H NMR spectrum, by comparing the intensities of the separately visible H_α resonances of the 2 end residues and of the $n-2$ residues in the interior of the peptide. Fractions with an oligomer length of 16.5–19.6 were combined, resulting in an average oligomer length of 18 ± 1.5. On the basis of the UV absorption of a standard solution of the oligo(L-ornithine) mixture at 209 nm, it was calculated that the concentration of ornithine residues equals $(4.28 \times 10^{-4})A_{209}$ M. By use of this relation, a stock solution was prepared of an oligo(L-ornithine) fraction with the correct length.

UV Measurements. All variable-temperature UV measurements were performed in 10-mm quartz cuvettes at 260 nm on a Perkin-Elmer 124 spectrophotometer. Samples of DNA/peptide complexes were prepared by diluting a DNA stock solution in the appropriate buffer to an A_{260} of ca. 0.2, after which such a quantity of peptide was added that the number of phosphates in the DNA equaled the number of terminal ammonium groups in the peptide. The samples were then incubated for 15 min at room temperature and after that for 15 min in ice. Usually, TE buffer (10 mM Tris-HCl/1 mM EDTA, pH 7.5) was used for these experiments. For the samples containing dC_{10} and $d(C_6T_6)$, however, it was necessary to adjust the pH of the buffer to 9, since otherwise the cytosine base is protonated at N_3 (Sarma et al., 1986; Adler et al., 1967; Guschlbauer, 1967). To obtain melting curves, sample temperatures were raised at 1 °C/min, and measurements were performed at 2 °C intervals. The UV hyperchromicity curves were computer fitted to obtain melting temperatures (T_m values). For this, we used the following formula for self-complementary strands (Marky & Breslauer, 1987). The fraction single-strand form (f) is determined by

$$f = 2/[1 + [1 + 8e^{(\Delta H^\circ/R)(1/T - 1/T_m)}]^{1/2}]$$

where ΔH° is the dissociation enthalpy of the DNA duplex. The entropy change is then determined from

$$\Delta S^\circ = \frac{\Delta H^\circ}{T_m} + R \ln C_T$$

where C_T is the total single strand concentration in the sample.

NMR Measurements. One-dimensional ^{1}H NMR measurements were performed at 200 and 600 MHz on Bruker AC 200 and AM 600 spectrometers, both interfaced with an Aspect 3000 computer. Usually, a 15 ppm sweep width was used with 16K data points, which were zero-filled to 32K points before Fourier transformation. Chemical shifts were referenced to the H_2O resonance or the residual HDO peak in D_2O, which were set at 4.68 ppm at room temperature and at 4.88 ppm at 4 °C. During the measurement, the water peak was suppressed by combined application of a semiselective observation pulse (using the 1331 pulse scheme) and digital shift accumulation (Hore, 1983; Roth et al., 1980). DNA/peptide complexes were dissolved either in D_2O or in an 85:15 (v/v) H_2O/D_2O mixture for measurements of the exchangeable protons, with a concentration of ca. 1 mM. All measurements were performed in 5-mm NMR tubes without addition of buffers or salts, since high salt concentrations disrupt the DNA/peptide interaction. Samples were thoroughly sparged with argon gas to remove dissolved oxygen and stored at 4 °C.

Two-dimensional ^{1}H NMR experiments were performed at 600 MHz on a Bruker AM 600 spectrometer interfaced with an Aspect 3000 computer. The phase-sensitive NOESY experiment was performed at 4 °C in D_2O with an f_2 size of 4K and 1024 experiments in the f_1 direction. Each experiment had 64 scans and a 90° pulse of 11.56 μs. A mixing time of 300 ms was used, and a sweep width of 6000 Hz. The residual HDO peak was suppressed by presaturation (Bodenhausen et al., 1984). The obtained file was weighed with a shifted squared-sine bell window in both directions before Fourier transformation. Zero-filling to 4K points was performed in the f_1 direction, and appropriate phase corrections were applied. Cross peaks in the resulting 16M file were integrated by volume to assess the strength of the NOE contact.

^{31}P NMR measurements were performed at 80.8 MHz on a Bruker AC 200 spectrometer interfaced with an Aspect 3000 computer. Usually, a sweep width of 150 ppm was used with 16K data points. Before Fourier transformation, zero-filling to 64K points was performed. Chemical shifts were referenced to 85% H_3PO_4 (0 ppm).

^{13}C NMR measurements were performed at 50.3 MHz on a Bruker AC 200 spectrometer interfaced with an Aspect 3000 computer. A size of 16K data points was used, and a sweep width of 12 kHz. Chemical shifts were related to tetramethylammonium bromide (56.4 ppm).

NMR Chemical Shift Calculations. The chemical shifts of the H_6 thymine protons and the C_α and C_β nuclei in the Lys_{18} peptide were calculated as a function of the orientation of the base pairs in a parallel duplex with the semiempirical method of Giessner-Prettre et al. (1970, 1976). This model accounts for the paramagnetic and diamagnetic shielding (Ribas-Prado & Giessner-Prettre, 1981) as well as ring current effects. The central base pair of a parallel duplex of dT_7 or dC_7 was used in the computations to avoid end effects. The H_6 chemical shifts were calculated for various values of $D_\parallel$, the distance from the middle of the base pair to the helix axis parallel along the line connecting the centers of the two bases. This geometric parameter is important for the description of base stacking in the parallel duplex. The chemical shifts of C_α and C_β in Lys_{18} are calculated as a function of the propellor twist angle in the T-T and C-C base pairs.

Model Building and Energy Refinement. Initially, the oligopeptide backbone conformation that fits the helical parameters of parallel DNA was established by calculating atomic coordinates in a growing peptide chain from standard bond lengths and angles, based on the X-ray structure of L-lysine hydrochloride, and variable torsion angles (van Genderen et al., 1988). The result of these calculations was used in model-building studies with the graphics program Chem-X, where the helical oligopeptide was fitted to a parallel DNA duplex by optimizing the conformation of the side chains.[2] Parallel DNA duplexes were constructed in Chem-X

Table I: Thermodynamic Parameters for the Melting Transitions of Peptide-Induced Duplexes

system	pH	T_m (°C)[a]	$\Delta H°$ [kJ/(mol bp)][a]	$\Delta S°$ [J/(mol bp·K)][a]
dT_{10}/Lys_{18}	7.5	21.0	41.8	131.3
	9.0	19.4	41.4	129.2
dT_{10}/Orn_{18}	7.5	21.3	43.9	136.7
	9.0	17.1	51.4	165.5
dC_{10}/Lys_{18}	9.0	25.3	31.8	96.1
dC_{10}/Orn_{18}	9.0			
$d(C_6T_6)/Lys_{18}$	9.0	20.1	45.6	146.3
$d(C_6T_6)/Orn_{18}$	9.0			
$d(T_6C_2T_2)/Lys_{18}$	7.5	19.7	35.1	108.7
$d(T_6C_2T_2)/Orn_{18}$	7.5			

Δh_{TT}, 37.6; Δh_{CC}, 35.1; Δh_{TC}, 40.1; Δh_{CT}, 50.2[b]

[a] Determined by computer fitting of the melting curves (see Materials and Methods). [b] Dimer enthalpy (kJ/mol).

as described before (van Genderen et al., 1987a). Two right-handed antiparallel DNA strands were generated and then reoriented to a parallel situation with preservation of the conformational characteristics of each strand. The DNA/peptide structures obtained in this way were then used as input geometries for energy minimization in the molecular mechanics program AMBER (Weiner et al., 1984). The coordinates of the input structures were energy-refined until the gradient was smaller than 0.4 kJ/mol·Å, which on the average took 4000 steps. All calculations were performed in the united atom approach, with a dielectric constant $\epsilon = r_{ij}$. This choice of dielectric constant has been found to simulate the long-range shielding of electrostatic interactions in aqueous solution in the most realistic way (Weiner et al., 1984). While the coordinates of the lysine residue are directly available from the AMBER database, the ornithine residue had to be defined with the PREP module of AMBER. The coordinates of ornithine were derived from those of lysine by removing one methylene group from the side chain.

Results and Discussion

Duplex Formation. For each of the oligomers dT_{10}, dC_{10}, $d(C_6T_6)$, and $d(T_6C_2T_2)$, the thermal denaturation behavior was studied with UV spectroscopy after complexation with octadeca(L-lysine) (Lys_{18}) and octadeca(L-ornithine) (Orn_{18}). As is shown in Table I, virtually identical duplex $\rightleftarrows$ coil transitions were seen with both oligopeptides for dT_{10}, whereas the cytosine-containing systems formed a duplex exclusively with Lys_{18}. It should be noted that the dC_{10} and $d(C_6T_6)$ oligomers were measured at pH 9 to prevent protonation of the cytosine base. As can be seen for dT_{10}, the change in pH from 7.5 to 9 hardly affects the stability of duplexes with T-T base pairs. By varying the ratio of peptide:DNA, it was established that the largest UV hyperchromicity effect occurs when one peptide molecule is present for two DNA strands.[3]

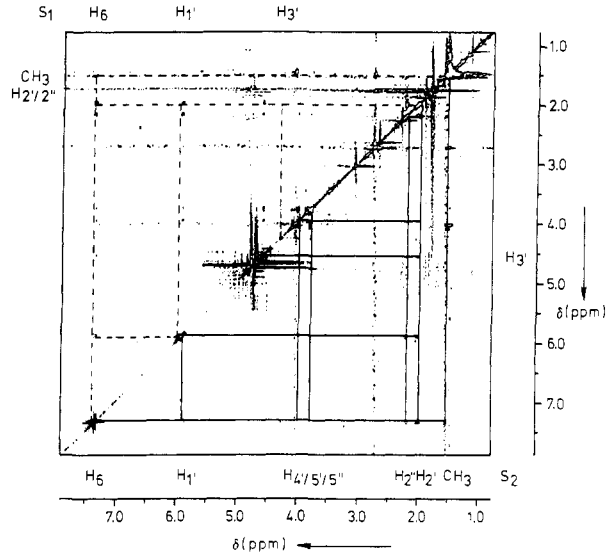

FIGURE 1: Assignment of proton resonances for the weak NOE contacts (---) and the strong contacts (—) in the dT_{10}/Orn_{18} complex in D_2O at 4 °C.

This indicates a structure with an oligopeptide strand in *one* of the grooves of the DNA duplex. The duplexes have comparable values for the transition enthalpy and entropy per base pair. The values of $\Delta H°$ can be analyzed in terms of the model of Breslauer et al. (1986), which states that the total enthalpy change can be obtained as an algebraic sum of contributions from individual dinucleotide units ($\Delta H° = \sum_k \Delta h_k$). Since in our systems four different dinucleotide units are present (i.e., TT, CC, TC, and CT), four Δh values can be abstracted from the experimental thermodynamic data of Lys_{18}-induced parallel duplexes (see Table I). The dimer contributions are all in the range 35–50 kJ/mol, which correlates well with thermodynamic data found for antiparallel base pairs. The present data allow the prediction of the $\Delta H°$ value of peptide-induced parallel duplexes for each pyrimidine base sequence.

The presence of DNA duplexes was ascertained independently by ^1H and ^{31}P NMR measurements on the dT_{10}/Lys_{18} complex. For thymine bases in single strands, the chemical shift of the imino protons is around 11 ppm in aqueous solution (Haasnoot et al., 1979), whereas for phosphate-methylated parallel duplexes with T-T base pairs, where the imino protons are involved in hydrogen bonding, chemical shifts of 13.4 ppm have been observed (Koole et al., 1987). At 10 °C, i.e., below the T_m value of 21 °C, the imino chemical shift in the dT_{10}/Lys_{18} complex is found at 13.0 ppm, which clearly indicates the formation of T-T base pairs. In the ^{31}P NMR spectrum, an aqueous solution of dT_{10} at room temperature shows a single resonance at 1.93 ppm, which shifts upfield to 1.85 ppm after addition of Lys_{18}. Such upfield shifts are known to occur in lysine–phosphate complexation (Davanloo & Crothers, 1979). The melting transition of the dT_{10}/Lys_{18} complex was followed by raising the sample temperature. The resonance at 1.93 ppm again appears around 30 °C and increases in intensity until it is equal to the 1.85 ppm signal. This indicates melting of the DNA duplex in such a way that the oligopeptide strand remains associated with one of the DNA strands. The high T_m value with respect to the UV measurements (21 °C) is due to the higher concentration in the

[2] The initial peptide conformation that forms a helix with a diameter and rise per base pair corresponding to parallel DNA (18 and 3.4 Å, respectively) was obtained by repeating a two amino acid fragment. The backbone torsion angles for the residue with the side chain pointed to S_1 are $\{\phi, \psi, \omega\} = \{173°, 160°, -170°\}$, and for the residue pointed to S_2 they are $\{-173°, -160°, -170°\}$ (van Genderen et al., 1988; cf. Figure 4C). For the carbon–carbon bonds in the side chains, it was found that all-trans torsion angles ($\{\chi^1, \chi^2, \chi^3\} = \{180°, 180°, 180°\}$) led to a good ammonium–phosphate complexation for half the strands (pointing toward S_2). For the side chains pointed toward S_1, it was necessary to use $\{\chi^1, \chi^2, \chi^3\} = \{180°, 60°, 60°\}$.

[3] The hyperchromicity effect as a percentage of the absorbance at the T_m value runs as follows for increasing peptide:DNA strand ratio, e.g., the dT_{10}/Orn_{18} complex: 0.25, 8.4%; 0.3, 6.0%; 0.4, 12.1%; 0.5, 14.6%; 0.8, 14.9%; 1.3, 15.2%.

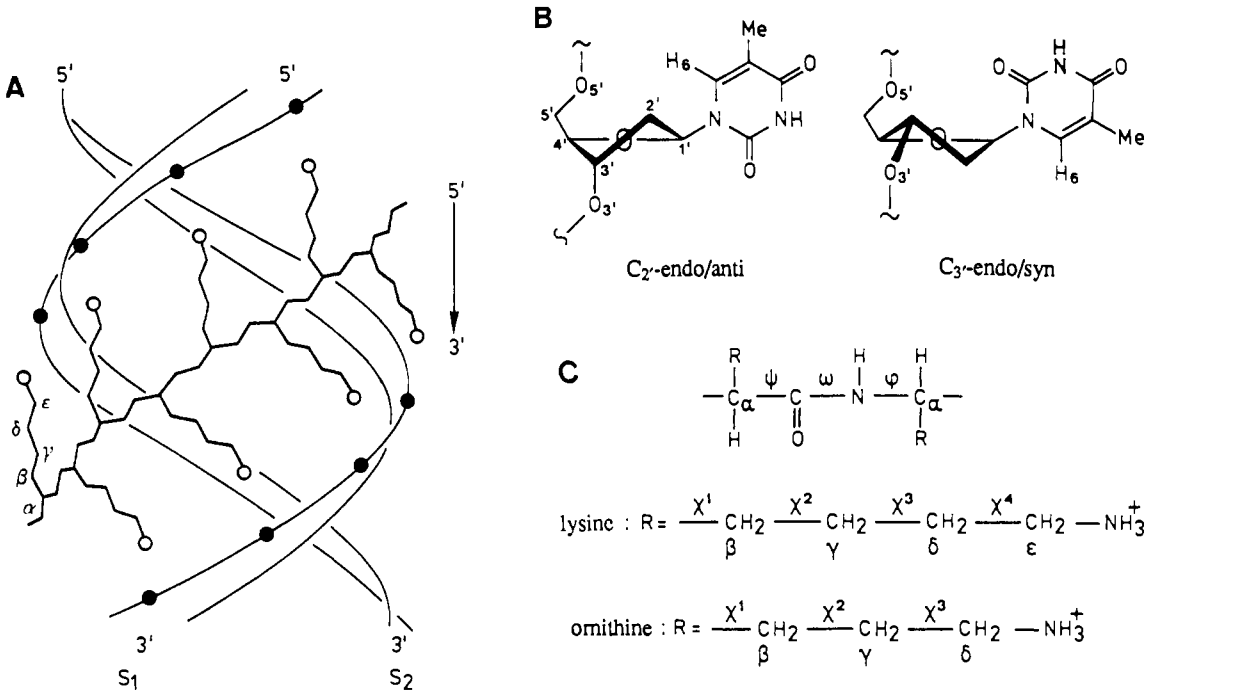

FIGURE 2: (A) Molecular model of a parallel duplex with an oligo(L-lysine) strand in one of the grooves. Phosphates (●) and ammonium groups (○) are indicated. (B) Possible conformations for sugar ring and base: $C_{2'}$-endo/anti (left) and $C_{3'}$-endo/syn (right). (C) Numbering scheme and definition of torsion angles in lysine and ornithine peptides.

NMR sample (1 mM vs 2 μM; Marky & Breslauer, 1987).

Two-Dimensional Proton NMR Studies. For a more detailed characterization of the peptide-induced parallel duplexes, we performed phase-sensitive NOESY measurements on the dT_{10}/Orn_{18} complex, in combination with AMBER molecular mechanics calculations. The measured NOE contacts can be related to short (<5 Å) interproton distances, and a distinction can then be made between various helical geometries of duplex DNA (e.g., right- or left-handed helices). Assignment of the proton resonances was obtained from the NOESY spectrum in a standard fashion (see Figure 1). For the deoxyribose rings, we followed the path from the $H_{1'}$ signal to the $H_{2'}$ and $H_{2''}$ resonances (assigned with the fact that the NOE contact $H_{1'}$–$H_{2'}$ is always stronger than $H_{1'}$–$H_{2''}$, regardless of sugar pucker; Kearns, 1987), which in turn led to $H_{3'}$. From there, $H_{4'}$ was identified and subsequently the $H_{5'}$ and $H_{5''}$ resonances, which were assigned according to the Remin and Shugar rule ($\delta_{5'} > \delta_{5''}$; Remin & Shugar, 1972). For the thymine bases, it was found that the H_6 protons had strong NOE contacts with the thymine methyl groups, which were used as a reference (interproton distance 2.5 Å). The location of the oligopeptide resonances was already known from one-dimensional measurements and was found to be almost unchanged.

As can be seen in Table II and Figure 1, virtually all proton resonances and NOE contacts are present in duplicate in the NOESY spectrum, with one partner weaker than the other. For the H_6 protons a double resonance was even seen in the one-dimensional spectrum, but here both peaks are of equal intensity. The upfield resonance was found to be associated with the strong NOE contacts. These findings suggest that the two DNA strands each give rise to a separate set of proton resonances, and also to different NOE intensities. In principle, the two DNA strands in a peptide-induced parallel duplex are indeed not identical, since they have a different orientation relative to the oligopeptide (see Figure 2A). Strand S_1 is complexed by a peptide residue pointing in the 5′ → 3′ direction, whereas strand S_2 is bound to a residue pointing in the 3′ → 5′ direction. It should be noted that a different location relative to the peptide is not possible for an antiparallel duplex, where the orientation of the two DNA strands is opposite to each other. In an antiparallel duplex, the NMR behavior of both DNA strands would therefore have been identical. It will be shown with molecular mechanics calculations that the two different ways of complexation between the DNA strands and the peptide for parallel duplexes results

Table II: Chemical Shifts (ppm) and Strength of NOE Contacts in the dT_{10}/Orn_{18} Complex

DNA protons	Chemical Shifts: weak NOEs	strong NOEs	peptide protons
H_6	7.32	7.28	H_α 3.99
$H_{1'}$	5.86	5.84	$H_{\beta/\gamma}$ 1.43
$H_{2'}$	2.05	1.97	H_δ 2.69
$H_{2''}$	2.05	2.15	
$H_{3'}$	4.23	4.52	
$H_{4'}$[a]		3.93	
$H_{5'}$[a]		3.77	
$H_{5''}$[a]		3.73	
CH_3(T)	1.51	1.49	

NOE Contacts DNA	weak NOEs	strong NOEs		weak NOEs	strong NOEs
H_6–CH_3(T)	2.34	6.63	$H_{1'}$–$H_{5''}$[b]		0.85
H_6–$H_{1'}$	0.62	3.07	$H_{2'}$–$H_{2''}$[a]		30.17
H_6–$H_{2'}$	2.15	7.03	$H_{2'}$–$H_{3'}$	1.26	4.40
H_6–$H_{2''}$	2.15	4.16	$H_{2'}$–$H_{4'}$[b]		0.68
H_6–$H_{3'}$	0.22	1.03	$H_{2''}$–$H_{3'}$	1.26	2.85
H_6–$H_{4'}$[b]		0.58	$H_{2''}$–$H_{4'}$[b]		1.23
H_6–$H_{5'}$[b]		0.37	$H_{3'}$–$H_{4'}$[b]		3.95
$H_{1'}$–CH_3(T)[b]		0.72	$H_{3'}$–$H_{5'}$[b]		2.43
$H_{1'}$–$H_{2'}$	0.97	5.16	$H_{3'}$–$H_{5''}$[b]		3.90
$H_{1'}$–$H_{2''}$	0.97	8.63	$H_{4'}$–$H_{5'}$[b]		5.63
$H_{1'}$–$H_{4'}$[b]		2.08	$H_{4'}$–$H_{5''}$[b]		8.15

NOE Contacts Peptide–DNA	with weak NOE protons	with strong NOE protons
CH_3(T)–H_α	6.75	
CH_3(T)–H_δ		1.64

[a] Separate resonances could not be observed. [b] Weak NOE contact was not detected.

Table III: Energy Contributions and Conformational Characteristics for the Energy-Refined Structure of the dT_{10}/Orn_{18} Complex

energy term[a] (kJ/mol)		DNA conformation[b]	strand S_1	strand S_2	protein conformation[c]	toward S_1	toward S_2
E_{tot}	−9825.3	α	−63.2°	−62.3°	ψ	161.4°	179.8°
E_{bond}	68.1	β	165.5°	162.4°	ω	178.9°	178.9°
E_{angle}	701.9	γ	57.4°	55.1°	ϕ	−169.2°	−153.6°
E_{dihed}	1379.9	δ	119.4°	117.4°[d]	χ^1	−151.3°	−170.8°
$E_{vdW,14}$	559.7	ϵ	178.3°	−170.3°	χ^2	46.0°	−158.9°
$E_{EEL,14}$	−8561.7	ζ	−86.3°	−71.0°	χ^3	52.4°	−173.1°
$E_{vdW,NB}$	−1450.0	χ	234.0°	237.0°			
$E_{EEL,NB}$	−2466.9	P	116.0°	161.0°[d]			
E_{Hbond}	−55.9	ν_{max}	0.37 Å	0.37 Å			
E_{int}	−327.5	hra	37.4°				
		twist	24.7°				

[a] Energy terms are, respectively: total energy, bond stretching energy, angle deformation energy, torsional energy, vicinal van der Waals energy, vicinal electrostatic energy, nonbonded van der Waals energy, nonbonded electrostatic energy, hydrogen-bond energy, and DNA interstrand energy. [b] Nomenclature according to IUPAC-IUB recommendations (1970) and Altona and Sundaralingam (1973). See also Figure 2B. hra and twist denote the helical repeat angle and the propellor twist angle, respectively. [c] Nomenclature according to IUPAC-IUB recommendations (1983). See also Figure 2C. [d] These values varied appreciably in the residues of the duplex.

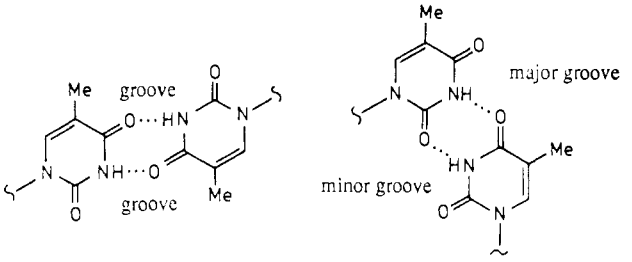

FIGURE 3: Possible T–T base pairs, with parallel (left) and antiparallel (right) backbone orientations. The location of the grooves in the duplex is indicated.

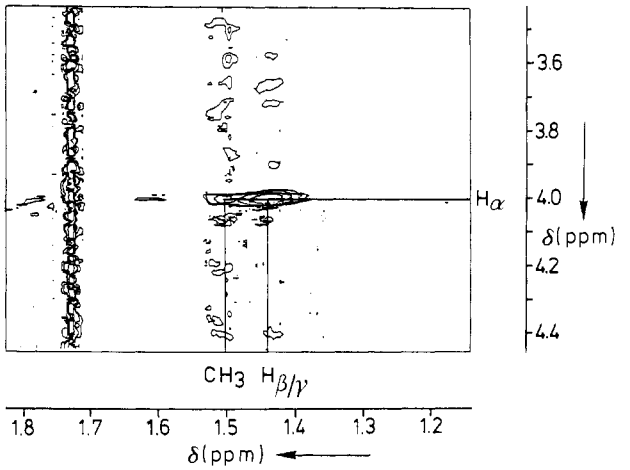

FIGURE 4: NOE contacts between H_α in the backbone of Orn_{18} and the base methyl group of thymine in the dT_{10}/Orn_{18} complex in D_2O at 4 °C.

in differences in conformation for the individual DNA strands (vide infra). On the basis of these DNA conformations, the strength and location of the NOE contacts will be rationalized. For now we will use the strong set of NOE contacts to deduce structural details of the dT_{10}/Orn_{18} complex. As is evident from Table II, the observable weak NOE contacts would give rise to the same conclusions.

Structural Details of the Peptide-Induced Duplex. The orientation of the thymine base can be deduced from the H_6–$H_{1'}$ NOE contact, which is very weak in the usual anti conformation (distance 3.7 Å; see Figure 2B) and strong in the syn conformation (distance 2.2 Å). The dT_{10}/Orn_{18} complex showed a relatively weak H_6–$H_{1'}$ contact compared to the reference contact H_6–CH_3(T) (2.5 Å), so it can be concluded that the T bases are in the anti conformation. Also, strong cross peaks are observed for H_6–$H_{2'/2''}$, which always occurs for DNA duplex structures where bases are in the anti orientation (Mirau & Kearns, 1984; Cohen, 1987).

On the basis of the anti conformation of the thymine bases, it can be established that the DNA backbones have a parallel orientation. In antiparallel duplexes, a major and a minor groove are present, and a cationic oligopeptide must be located in the minor groove to reach all phosphates (Tsuboi, 1967; Feughelman et al., 1955). In a symmetric parallel duplex, on the other hand, both grooves are identical, and both can accommodate an oligopeptide molecule. As can be seen in Figure 3, an antiparallel duplex would correspond with a large separation (13–17 Å) between the peptide backbone in the minor groove and the thymine methyl groups, whereas a short distance (5–7 Å) is always found in a parallel duplex. For dT_{10}/Orn_{18} a clear NOE contact, and hence a short distance, was observed for CH_3(T) and the H_α protons in the peptide backbone (see Figure 4), indicating that a parallel structure is present.[4]

The sugar conformation of the parallel duplex, which is in equilibrium between a $C_{3'}$-endo form and a $C_{2'}$-endo form (see Figure 2B), can also be assessed from the NOESY spectrum. Due to overlap of inter- and intrastrand cross peaks, no conclusion can be drawn from the H_6–$H_{2'}$ contacts, since the corresponding *intra*nucleotide distances are short (2.4 Å) for the $C_{2'}$-endo form, but the *inter*nucleotide distances are small (2.1 Å) for the $C_{3'}$-endo form. The intra- and internucleotide distances H_6–$H_{3'}$ are more suitable, since they are short (2.5–3 Å) in the $C_{3'}$-endo form, while they are much longer (>4 Å) in the $C_{2'}$-endo form. The very weak H_6–$H_{3'}$ contact in the dT_{10}/Orn_{18} complex therefore indicates a $C_{2'}$-endo sugar pucker. The observation of a relatively strong H_6–$H_{2''}$ cross

[4] For C–C base pairs, an antiparallel backbone orientation would allow only a single hydrogen bond, which leads to highly unstable duplexes (in surveys of all possible base pairs, antiparallel C–C is nowhere mentioned; Hobza & Sandorfy, 1987). This is confirmed by AMBER molecular mechanics calculations on the parallel and antiparallel duplexes of dC_6, which are good model systems for correlation with experimental stabilities (van Genderen et al., 1987a). Energy minimization resulted in a right-handed parallel duplex structure that is comparable to the one found for dT_6 (van Genderen et al., 1987a). The antiparallel dC_6 duplex forms a standard B-DNA structure. Although the conformations of the parallel and antiparallel systems are rather similar, the total energies (E_{tot}) and the interaction energies between the strands (E_{int}) differ widely. The antiparallel duplex of dC_6 has a higher E_{tot} compared to the parallel duplex (−1457.6 kJ/mol vs −1733.9 kJ/mol), and E_{int}, which represents the duplex stability, is markedly less negative (−115.8 kJ/mol vs −315.2 kJ/mol).

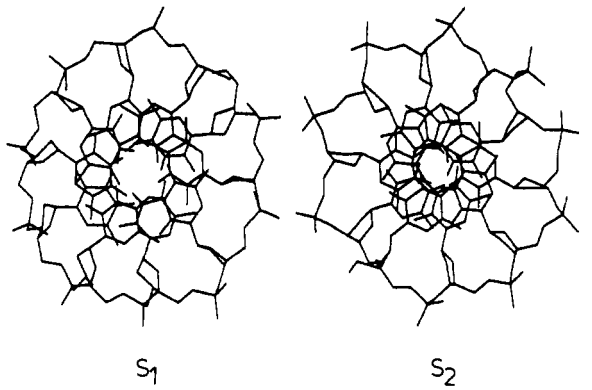

FIGURE 5: Top view of the strands S_1 and S_2 in the energy-refined structure of the dT_{10}/Orn_{18} complex.

peak is in agreement with this conclusion, since both intra- and internucleotide H_6–$H_{2''}$ distances are small for the $C_{2'}$-endo form. A very weak $H_{2''}$–$H_{4'}$ contact gives further support, since in the $C_{3'}$-endo form a strong cross peak (distance 2.3 Å) is expected, of the same intensity as the H_6–$CH_3(T)$ reference peak. We therefore conclude that the parallel duplex of dT_{10}/Orn_{18} is right-handed, with anti base orientation and $C_{2'}$-endo sugar puckers.

Molecular Mechanics Calculations. On the basis of the experimental data described above, we constructed a right-handed parallel duplex of two dT_{10} strands, with an Orn_{18} strand in one of the grooves (see Materials and Methods). This DNA/peptide complex was subjected to energy minimization with the AMBER force field. The energy terms and conformational parameters that resulted are given in Table III. The major energy interactions are the vicinal and nonbonded electrostatic attractions, which is to be expected for a complex that derives its stability from hydrogen bonding and ammonium–phosphate complexation. The interaction energy between the DNA strands (E_{int}) corresponds with the dissociation enthalpy determined in the UV hyperchromicity experiments. As can be seen, the theoretical value of 32.6 kJ/mol of base pairs is in fair agreement with the experimental value of 43.9 kJ/mol of base pairs (cf. Table I) when it is taken into account that no water molecules or counterions were used in the AMBER calculation.

After minimization, the peptide strand still adopts an alternating backbone conformation, although it has deviated from the input structure. The conformation of the DNA duplex stays in the parallel right-handed form, as was found in the NMR studies. The most striking feature is the difference in conformation for the two strands S_1 and S_2, especially with respect to the sugar rings. In Figure 5, it can be seen that a stronger base stacking (i.e., a larger overlap) is present for adjacent thymines in strand S_2, which implies a lower mobility for the residues in S_2 compared with those in S_1. This is consistent with the fact that in one of the strands the NOE contacts are more intense (vide supra), even when no variation in interproton distance is possible [e.g., for H_6–$CH_3(T)$], since the strength of NOE cross peaks is also determined by interproton cross relaxation rates, which are related to the correlation time of the interproton vector (Macura & Ernst, 1980; Clore & Gronenborn, 1984). We will therefore now identify the strong set of NOEs with strand S_2 and the weak set with S_1. In the AMBER calculations, the difference in stacking of the two DNA strands is due to a displacement of the T-T base pairs from the helix axis ($D_{\parallel}$ = 0.31 Å). As a result, the base methyl groups of S_1 point into the groove where Orn_{18} is located, which is in good agreement with the observation of a strong NOE between

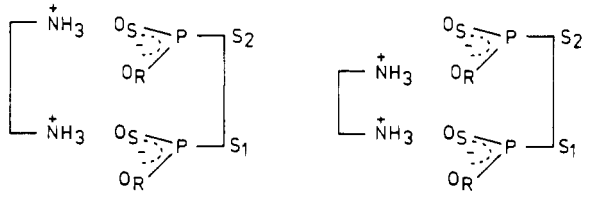

FIGURE 6: Schematic representation of the model-building studies on the mode of peptide complexation with the parallel duplex. Geometries for the Lys_{18} (left) and Orn_{18} (right) peptides are shown.

$CH_3(T)$ and the peptide H_α only for strand S_1 (see Table II). The value of this displacement can be related with the experimentally observed chemical shift difference for H_6 in the dT_{10}/Orn_{18} complex. The upfield H_6 resonance, connected with the strong NOEs, is due to strand S_2, in which base–base overlap is the strongest. This agrees well with the fact that H_6 is more shielded for this arrangement of bases (Patel, 1980). With chemical shift calculations (see Materials and Methods), it was found that the 0.04 ppm difference between the H_6 resonances of S_1 and S_2 can indeed be explained by an 0.3 Å value of $D_{\parallel}$.

Stereospecificity of Peptide Complexation. The phosphodiester linkage in DNA contains two nonbonded oxygen atoms that are prochiral, i.e., a preferential complexation of a peptide ammonium group to one oxygen will result in a chiral ammonium–phosphate complex. Since the presence of an ammonium group relatively close (nitrogen–oxygen distance ca. 2.6 Å) to the phosphate will result in steric interactions with the DNA strands, a situation may arise where two complexation modes (i.e., selection of either oxygen) are comparable to the two configurations in methyl phosphotriester DNA. We therefore designate the prochiral oxygens as O_S and O_R, when attachment of a methyl group would lead to an S_P or R_P configuration, respectively. An analysis of the contacts between the ammonium groups of Orn_{18} and the phosphate oxygen atoms in the dT_{10} strands for the energy-minimized structure of dT_{10}/Orn_{18} reveals that the peptide selects the O_S atom (87%) in strand S_1, while O_R is preferred (65%) in strand S_2. This agrees well with model-building studies of the dT_{10}/Orn_{18} complex, where the length of the ornithine side chains (5.8 Å from C_α to terminal N) forces complexation with O_S in strand S_1 and with O_R in S_2 (see Figure 6). In the case of the dT_{10}/Lys_{18} complex, model studies show that the longer chains of lysine (6.8 Å) allow complexation with the most easily available O_S oxygen atoms in both DNA strands (see Figure 6). Therefore, complexation with the oligopeptides Lys_{18} and Orn_{18} may lead to a different duplex stability in the case of C-C base pairs (which in earlier studies on phosphate-methylated parallel duplexes were found to have a larger propellor twist angle compared to T-T base pairs), due to steric hindrance when complexation with the O_R oxygen occurs. This situation is then completely analogous to the two phosphorus configurations in phosphate-methylated systems, of which only S_P allows C-C base pair formation.

An experimental indication that the difference between peptide-induced parallel duplexes with T-T and C-C base pairs is due to the propellor twist angle follows from a comparison of the ^{13}C NMR spectra of the oligopeptide Lys_{18} in the systems dT_{10}/Lys_{18} and dC_{10}/Lys_{18}. A marked difference is observed for the shielding of C_α and C_β (see Table IV). For T-T base pairs C_α is strongly deshielded, while for C-C base pairs deshielding is more pronounced at C_β. Since the position of the peptide in the groove is similar for both systems, this difference in shielding may be due to a change in the base-pair conformation. In molecular models it can be seen that a large

Table IV: ^{13}C NMR Chemical Shifts (ppm) in the Lys_{18} Oligopeptide before and after Complexation with dT_{10} and dC_{10} at 4 °C and pH 9

atom	Lys_{18}	dT_{10}/Lys_{18}	dC_{10}/Lys_{18}	$\Delta\delta_{T-T}$	$\Delta\delta_{C-C}$
C_α	54.37	54.58	54.39	+0.21	+0.02
C_β	31.24	31.16	31.48	−0.08	+0.24
C_γ	22.85	22.82	22.85	−0.03	0.00
C_δ	27.18	27.16	27.14	−0.02	−0.04
C_ϵ	40.09	40.03	40.02	−0.06	−0.07

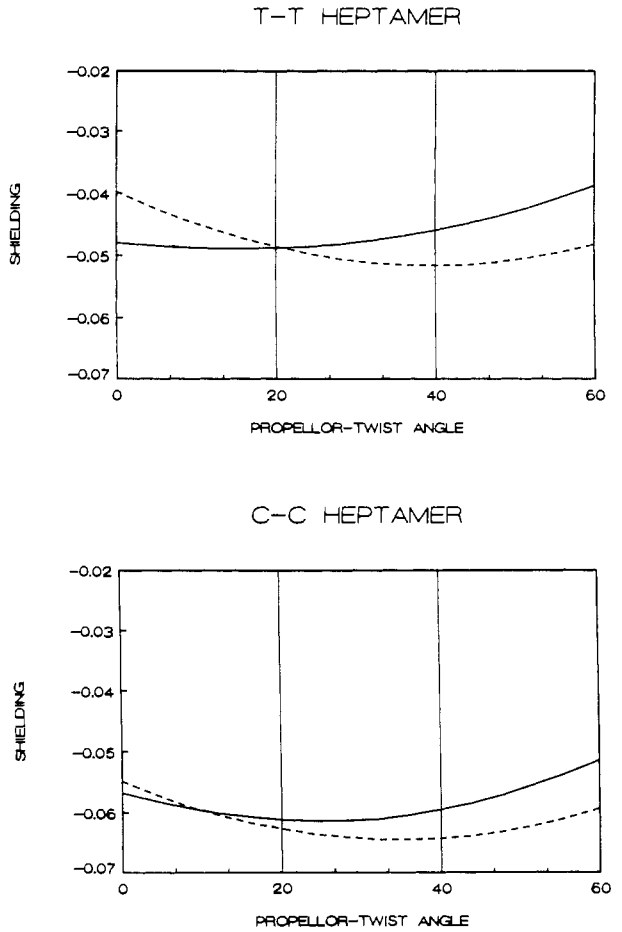

FIGURE 7: Calculated shielding effects of T-T base pairs (top) and C-C base pairs (bottom) on C_α (—) and C_β (---) in Lys_{18} as a function of the propellor twist angle in the base pairs.

propellor twist angle locates C_β in the plane of the ring current of the base, where a strong deshielding effect occurs, whereas a small twist angle leads to an in-plane location for C_α. This effect was studied with chemical shift calculations for C_α and C_β near a parallel duplex of dT_7 and dC_7 (see Materials and Methods). As can be seen in Figure 7, the smaller shielding of C_α compared to C_β as observed for T-T base pairs can only occur for small values of the propellor twist angle (<20°). In contrast, the larger shielding of C_α compared to C_β observed for C-C base pairs is present only for large values of the twist angle. Roughly the same difference in shielding between C_α and C_β is found for angles >20°. The interpretation of the ^{13}C NMR data is therefore in good agreement with the proposed difference between parallel duplexes with T-T and C-C base pairs.

Conclusions

On the basis of the present results, we conclude that natural DNA oligomers with pyrimidine base sequences can form right-handed parallel duplexes after complexation with the cationic peptide oligo(L-lysine). For thymine bases, a parallel duplex is formed also after complexation with oligo(L-ornithine) as peptide. However, the presence of C-C base pairs in the parallel duplex cannot be realized for oligo(L-ornithine), since complexation with the O_R phosphate oxygen in this case leads to steric hindrance with the C-C base pairs. This stereospecific behavior is completely analogous to the situation for phosphate-methylated parallel duplexes with T-T and C-C base pairs, where the latter systems exclusively occur for the S_P configuration. On the basis of the existence of peptide-induced parallel duplexes with only the two pyrimidine bases, it may be speculated that these structures have a relevance as primitive DNA duplexes. Only three components (thymidine, cytidine, and lysine) would have been sufficient to constitute a simple genetic material with a two-letter code.

Acknowledgments

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.

Registry No. Thymine, 65-71-4; cytosine, 71-30-7.

References

Adler, A., Grossman, L., & Fasman, G. D. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 423–430.

Altona, C., & Sundaralingam, M. (1972) *J. Am. Chem. Soc. 94*, 8205–8212.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

Breslauer, K. J., Frank, R. Blöcker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.

Clore, G. M., & Gronenborn, A. M. (1984) *FEBS Lett. 172*, 219–225.

Cohen, J. S. (1987) *Trends Biochem. Sci. 12*, 132–135.

Davanloo, P., & Crothers, D. M. (1979) *Biopolymers 18*, 2213–2231.

Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K., & Hamilton, L. D. (1955) *Nature (London) 175*, 834–838.

Germann, M. W., Vogel, H. J., Pon, R. T., & van de Sande, J. H. (1989) *Biochemistry 28*, 6220–6228.

Giessner-Prettre, C., & Pullman, B. (1970) *J. Theor. Biol. 27*, 87–95.

Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L., & Ts'o, P. O. P. (1976) *Biopolymers 15*, 2277–2286.

Guschlbauer, W. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1441–1448.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, F. J. M., van Boom, J. H., & Altona, C. (1979) *Nature (London) 281*, 235–236.

Hobza, P., & Sandorfy, C. (1987) *J. Am. Chem. Soc. 109*, 1302–1307.

Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.

IUPAC–IUB Commission on Biochemical Nomenclature 1969 (1970) *Biochemistry 9*, 3741–3749.

IUPAC–IUB Joint Commission on Biochemical Nomenclature (1983) *Eur. J. Biochem. 131*, 9–15.

Koole, L. H., van Genderen, M. H. P., Frankena, H., Kocken, H. J. M., Kanters, J. A., & Buck, H. M. (1986) *Proc. K. Ned. Akad. Wet., Ser. B 89*, 51–55.

Koole, L. H., van Genderen, M. H. P., & Buck, H. M. (1987) *J. Am. Chem. Soc. 109*, 3916–3921.

Koole, L. H., Broeders, N. L. H. L., van Genderen, M. H. P., Quaedflieg, P. J. L. M., van der Wal, S., & Buck, H. M. (1988) *Proc. K. Ned. Akad. Wet., Ser. B 91*, 245–249.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, p 486, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601-1620.
Mirau, P. A., & Kearns, D. K. (1984) *Biochemistry 23*, 5439-5446.
Morvan, F., Rayner, B., Imbach, J.-L., Lee, M., Hartley, J. A., Chang, D.-K., & Lown, J. W. (1987) *Nucleic Acids Res. 15*, 7027-7044.
Patel, D. (1980) in *Nucleic Acid Geometry and Dynamics* (Sarma, R. H., Ed.) Pergamon Press, New York.
Quaedflieg, P. J. L. M., Broeders, N. L. H. L., Koole, L. H., van Genderen, M. H. P., & Buck, H. M. (1990a) *J. Org. Chem. 55*, 122-127.
Quaedflieg, P. J. L. M., van der Heiden, A. P., Koole, L. H., van Genderen, M. H. P., Coenen, A. J. J. M., van der Wal, S., & Buck, H. M. (1990b) *Proc. K. Ned. Akad. Wet. 93*, 33-38.
Remin, M., & Shugar, D. (1972) *Biochem. Biophys. Res. Commun. 48*, 636-642.
Ribas-Prado, F., & Giessner-Prettre, C. (1981) *J. Mol. Struct.: THEOCHEM 76*, 81-92.
Roth, K., Kimber, B. J., & Feeney, J. (1980) *J. Magn. Reson. 41*, 302-309.
Sarma, M. H., Gupta, G., & Sarma, R. H. (1986) *FEBS Lett. 205*, 223-229.
Thuong, N. T., Asseline, U., Roig, V., Takasugi, M., & Hélène, C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5129-5133.
Tsuboi, M. (1967) in *Conformation of Biopolymers*, Vol. II, pp 689-702, Academic Press, New York.
van Genderen, M. H. P., Koole, L. H., Aagaard, O. M., van Lare, C. E. J., & Buck, H. M. (1987a) *Biopolymers 26*, 1447-1461.
van Genderen, M. H. P., Koole, L. H., & Buck, H. M. (1987b) *Proc. K. Ned. Akad. Wet., Ser. B 90*, 181-187.
van Genderen, M. H. P., Koole, L. H., & Buck, H. M. (1988) *Proc. K. Ned. Akad. Wet., Ser. B 91*, 171-178.
von Hippel, P. H., & McGhee, J. D. (1972) *Annu. Rev. Biochem. 41*, 231-300.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. K. (1984) *J. Am. Chem. Soc. 106*, 765-784.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, p 208, Wiley and Sons, New York.

Organization of the Human Protein S Genes[†,‡]

Dyann K. Schmidel, Alicia V. Tatro, Lisa G. Phelps, Jennifer A. Tomczak, and George L. Long*
Department of Biochemistry, University of Vermont, Burlington, Vermont 05405
Received January 30, 1990; Revised Manuscript Received May 2, 1990

ABSTRACT: Human genomic clones that span the entire protein S expressed gene (PSα) and the 3′ two-thirds of the protein S pseudogene (PSβ) have been isolated and characterized. The PSα gene is greater than 80 kilobases in length and contains 14 introns and 15 exons, as well as 6 repetitive "Alu" sequences. Exons I and XV contain 112 and 1139 bp 5′ and 3′ noncoding segments in addition to the amino and carboxyl termini, respectively. Exons I–VIII encode protein segments that are homologous to the vitamin K dependent clotting proteins and are bounded by introns whose position and type are identical with other members of this protein family. Exons IX–XV encode protein segments homologous to sex hormone binding globulin (SHBG) and are bounded by introns of identical type and position as in the SHBG gene. Genomic clones for the PSβ gene cover a distance of greater than 55 kilobases and contain segments corresponding to amino acids 46–635 of the mature protein and the 1.1-kb 3′ noncoding region of the cDNA. The presence of multiple base changes in the coding portions of this gene, resulting in termination codons and frame shifts, suggests that it is a pseudogene. Comparison of DNA sequences for the two genes reveals 97% identity for coding and 3′ noncoding, and 95.4% for intronic regions, suggesting divergence of the two genes is a relatively recent event.

Human protein S is a 69 000-Da vitamin K dependent plasma glycoprotein (Di Scipio & Davie, 1979) that acts as a cofactor for activated protein C in the coagulation cascade to inactivate factors Va and VIIIa (Walker, 1981; Suzuki et al., 1983; Solymoss et al., 1988). Protein S is synthesized in hepatocytes (Fair & Marlar, 1986), endothelial cells (Fair et al., 1986; Stern et al., 1986), and the megaloblastic cell line MEG-01 (Ogura et al., 1987). It is found circulating in the blood in equimolar amounts free and bound to the complement protein C4b binding protein (C4BP), in a 1:1 ratio (Dahlback, 1983). Protein S in the bound form is not available as a cofactor for APC (Dahlback, 1986).

Hereditary protein S deficiency has been reported by several groups and is often associated with symptoms found in protein C deficient individuals [for a review, see Engesser et al. (1987)], including familial thrombophilia. Recently, a molecular alteration in the protein S gene has been reported in a family exhibiting protein S deficiency (Ploos van Amstel et al., 1989). We have also recently described different alterations in the expressed gene from four independent families (Schmidel et al., 1989).

The cDNA for human protein S has been cloned and fully characterized (Lundwall et al., 1986; Hoskins et al., 1987; Ploos van Amstel et al., 1987). The translated precursor

[†] This work was supported in part by National Institute of Heart, Lung and Blood Grants R01 HL 38899 and C06 39745.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02917.
* To whom correspondence should be addressed.

0006-2960/90/0429-7845$02.50/0