Length-Dependent Formation of Parallel-Stranded DNA in Alternating AT Segments†

Markus W. Germann,*,‡ Bernd W. Kalisch,§ Richard T. Pon,§ and Johan H. van de Sande*,§
Department of Medical Biochemistry and Department of Biological Sciences, The University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4

Received April 11, 1990; Revised Manuscript Received July 9, 1990

ABSTRACT: Parallel-stranded DNA can be formed from alternating AT segments and is not restricted exclusively to homooligomeric AT sequences. DNA oligonucleotides 3′-d$(AT)_n$x$C_4(AT)_n$-3′ (where x indicates the location of the 5′–5′ phosphodiester linkage) form parallel-stranded hairpin structures at micromolar strand concentration for $n = 4$ or 5 but not for $n = 6, 7$. The spectral properties of the parallel-stranded structures are similar to those of the hairpin structures containing homooligomeric AT stems. However, parallel-stranded structures formed in alternating AT segments are significantly less stable than either their corresponding antiparallel control or the homooligomeric parallel AT hairpins as evidenced by their lower helix–coil transition enthalpy, melting temperature, and stability constant. This results in a remarkable polymorphism which is most pronounced for 3′-d$(AT)_5$x$C_4(AT)_5$-3′. This oligonucleotide can exist as a parallel-stranded hairpin, coil, or concatameric antiparallel structure(s), depending on temperature and strand concentration. These results suggest simple guidelines for the design of parallel-stranded DNA. In addition, we present a model for the assessment of the stability of parallel-stranded duplex structures formed from AT base pairs based on their sequence.

In 1986, Pattabiraman reported, based on molecular mechanics calculation, that d$(A)_6$·d$(T)_6$ might form a parallel-stranded right-handed double-helical structure with reverse Watson–Crick base pairing. We have provided experimental evidence that parallel-stranded structures are formed intramolecularly for homooligomeric AT sequences (van de Sande et al., 1988; Germann et al, 1989). Structural investigation on the hairpin 3′-d$(T)_8$xC_4A_8-3′ (where x indicates a 5′–5′ phosphodiester linkage) supported the predicted structural parameters for parallel-stranded DNA (Germann et al., 1989; Pattabiraman, 1986). Recently, Shchyolkina et al. (1989) reported that 3′-d$(AT)_5$-O-$(CH_2)_6$-O-$(AT)_5$-3′ could form an intramolecular structure that was assigned as parallel stranded at low strand concentration. A 1,6-hexanediol linker was also used to investigate the potential of two *Drosophila* sequences to form parallel-stranded DNA (Tchurikov et al., 1989).

Parallel-stranded DNA is, however, not restricted to hairpin constructs, containing either a 5′–5′ or 3′–3′ phosphodiester bond or a 1,6-hexanediol linkage, but can also be formed from natural oligonucleotides with the appropriate sequence homology (Germann et al., 1988; Ramsing & Jovin, 1988; Rippe et al., 1989). The sequence constraint in these systems was obtained either by using oligonucleotides containing blocks of A and T residues (Ramsing & Jovin, 1988) or by the introduction of an alternating AT segment (Germann et al., 1988). For both dimeric duplexes, the enthalpic stability is significantly lower than that expected for homooligomeric AT segments. This suggests that the alternating AT components present in both systems have a lower propensity to form parallel-stranded (PS)[1] DNA than do homooligomeric AT sequences. The potential relevance of PS DNA depends critically on the sequence requirement for its formation and on the relative stability of homoooligomeric and alternating AT tracts in parallel-stranded orientation. In order to compare the stability of alternating and homooligomeric AT sequences, we have employed an analogous experimental strategy as for the homooligomeric hairpin structures that were investigated previously (van de Sande et al., 1988; Germann et al., 1989).

Materials and Methods

Synthesis and Purification. The 5′-phosphoramidite derivatives were synthesized, using commercially available thymidine or *N*-benzoyl-2′-deoxyadenosine as starting materials (van de Sande et al., 1988). Oligodeoxynucleotides were synthesized and purified as described previously (Germann et al., 1989). The concentrations of the pure products were estimated from the extinction coefficients of the mononucleotides at 90 °C in 5 M $NaClO_4$ (Germann, 1989).

Gel Electrophoresis. Analysis of oligodeoxynucleotides was carried out on 15% polyacrylamide gels run under denaturing (8 M urea, 90 mM Tris–borate, and 5 mM EDTA, pH 8.3) and nondenaturing conditions (4 °C, 15 mM $MgCl_2$ and 90 mM Tris–borate, pH 8.3). Bands were visualized by UV shadowing (denaturing gels) or by staining with ethidium bromide (0.5 μg/mL, 50 mM NaCl at 4 °C for native gels).

Spectroscopy. Ultraviolet (UV) absorption spectra and melting curves were obtained on a Varian 2290 spectrophotometer equipped with temperature-regulated cuvette holders.

Circular dichroism (CD) measurements were recorded on a Jasco J-500C spectropolarimeter equipped with a temperature control unit.

Fluorescence measurements were obtained at 5 °C on a Perkin-Elmer 560-10 S fluorescence spectrophotometer. The excitation and emission wavelengths for Hoechst 33258 were λ_{ex} 355 nm, λ_{em} 480 nm.

Enzymatic Reactions. All enzymes were purchased from Pharmacia Canada. The oligonucleotide 5′-GAATTC-$(T)_3$GAATTC (referred to as R-1) was 5′-end-labeled with

† Supported by the Medical Research Council of Canada and the Alberta Heritage Foundation for Medical Research (AHFMR). M.W.G. is a holder of a postdoctoral fellowship from the AHFMR.

* Address correspondence to these authors.
‡ Department of Biological Sciences.
§ Department of Medical Biochemistry.

[1] Abbreviations: APS, antiparallel stranded; bpb, bromophenol blue; CD, circular dichroism; PS, parallel stranded; R-1, hairpin 5′-dGAATTCTTTGAATTC; UV, ultraviolet.

0006-2960/90/0429-9426$02.50/0

[γ-^{32}P]ATP (New England Nuclear) and T4 polynucleotide kinase (Chaconas & van de Sande, 1980). Ligation reactions of this hairpin to PS and APS oligonucleotides were carried out at 15 °C overnight under standard ligation conditions (66 mM Tris-HCl, pH 7.6, 5 mM dithiothreitol, and 1 mM ATP) and with 1 unit of T4 polynucleotide ligase (Maniatis et al., 1982). Each reaction (10 μL) contained 2 pmol of the labeled hairpin and a 10-fold excess of the (unphosphorylated) PS or APS oligonucleotides. After termination by several freeze–thaw cycles, 2-μL aliquots were analyzed by denaturing polyacrylamide gel electrophoresis.

Thermal Denaturation Analysis. Thermally induced helix–coil transitions were monitored at 268 nm, using a temperature gradient of 0.5 °C/min. Absorbance readings were collected at 0.5 or 1.0 °C intervals and converted to $\Delta A/\Delta T$. The differentiated melting curve ($\Delta A/\Delta T$) was analyzed by using a concerted two-state model in which the absorbances of helix and coil forms vary linearly with the temperature

$$A = A_H(1 - \theta) + A_C\theta \quad (1)$$

The thermal behavior of the two states is described by α, the absorbance at 0 K, and the temperature coefficient β (i.e., $A_H = \alpha_H + \beta_H T$ and $A_C = \alpha_C + \beta_C T$), where the indexes H and C indicate absorbances and coefficients of the helix and coil form, respectively, and T is the temperature. The degree of transition (θ) is given by $K/(1 + K)$, where K is the equilibrium constant for the helix–coil transition. By use of the van't Hoff equation, θ and $d\theta/dT$ may be expressed as

$$\theta = \frac{1}{1 + \exp[(-\Delta H^\circ_{vH}/R)(1/T_M - 1/T)]} \quad (2a)$$

$$d\theta/dT = \theta(1 - \theta)\Delta H^\circ_{vH}/RT^2 \quad (2b)$$

where R is the gas constant (8.311 J mol^{-1} K^{-1}) and ΔH°_{vH} is the van't Hoff enthalpy for the helix–coil transition. The differentiated melting profile was fitted to eq 3 by using a

$$dA/dT = \beta_H + \theta(\beta C - \beta_H) + [\alpha_C - \alpha_H + (\beta_C - \beta_H)T]d\theta/dT \quad (3)$$

Simplex procedure (Caceci & Cacheris, 1984; Germann, 1989). The fitting of eq 3 to the differentiated melting profile requires only a five-parameter fit (e.g., $\Delta\alpha$, β_C, β_H, T_M, and ΔH°_{vH}) and was found to converge more rapidly than fitting the denaturation profiles with eq 1. Entropies are calculated from $\Delta S^\circ = \Delta H^\circ_{vH}/T_M$.

Helix Formation. Intramolecular helix formation was analyzed by using a simple all or none model in which the helix–coil equilibrium constant may be expressed as $K = \sigma s^N$ where σ is the nucleation parameter, σs and s are the equilibrium constant of forming the first base pair or adding one AT base pair on a preexisting one, respectively, and N is the number of base pairs formed (Cantor & Schimmel, 1984). For this model, the enthalpy and entropy of helix formation can be decomposed into contributions from helix initiation and helix growth [$\Delta H = \Delta H_\beta + (N - 1)\Delta H_s$ and $\Delta S = \Delta S_\beta + (N - 1)\Delta S_s$]. Thermodynamic data relating to chain initiation and growth are extracted from plots of enthalpy or entropy versus the number of nearest neighbors ($N - 1$). The equilibrium constant s is calculated from

$$\text{Ins} = -(\Delta H_s/RT) + \Delta S_s/R \quad (4)$$

Results and Discussion

Design and Structures of Alternating Parallel-Stranded Hairpin Oligonucleotides. We have exploited the ease of intramolecular hairpin formation for the demonstration of parallel-stranded alternating AT segments. The oligonucleotides 3′-d(AT)$_n$xC$_4$(AT)$_n$-3′ (with n = 4–7) and 5′-d-(AT)$_5$C$_4$(AT)$_5$-3′ have been synthesized. The parallel orientation of the hairpin stem sequences is obtained by a 5′–5′ phosphodiester linkage (marked by an x, Figure 1A). Similar to the homooligomeric parallel-stranded hairpin structures studied previously, the oligonucleotides investigated here can also form intermolecular structures (Figure 1B) that have the potential to form multimeric species (van de Sande et al., 1988; Germann et al., 1989). In addition, the alternating AT oligonucleotides containing the 5′–5′ phosphodiester bond also have the potential to form a dumbbell structure with antiparallel-strand alignment (Figure 1C). Both hairpin and dumbbell structures are formed intramolecularly, and consequently concentration independence of the physical properties (i.e., gel migration, UV spectra, and melting temperatures) does not distingiush between either structure. For the formation of the dumbbell, two unpaired hairpin loops are required. The hairpin loop size in Figure 1C was chosen to consist of four loop residues because the loops are closed by a weak AT base pair (Haasnoot et al., 1986). Smaller hairpin loops have also been reported; however, these are generally closed by a strong CG base pair (Orbons, 1987; Summers et al., 1985). The presence of two hairpin loops in the dumbbell structure only allows the formation of short double-helical segments and requires two nucleation events. Consequently, the resulting structure would be unstable for 3′-d(AT)$_n$xC$_4$-(AT)$_n$-3′ with $n \leq 5$; however, dumbbell structures will have to be considered for oligonucleotides with longer AT stem sequences.

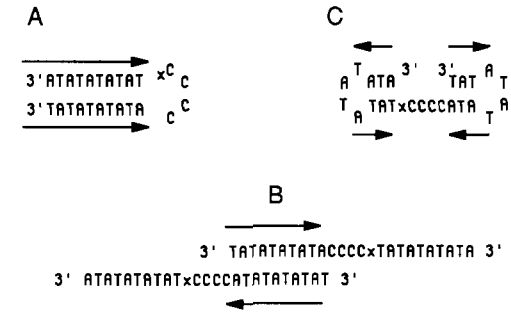

FIGURE 1: Structures for 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′. (A) Intramolecular parallel-stranded hairpin structure. (B) Intermolecular antiparallel dimer structure that has the potential to form concatamers. (C) Double hairpin structure with antiparallel strand alignment.

Gel Electrophoretic Analysis of Alternating AT Oligonucleotides Containing a 5′–5′ Phosphodiester Linkage. The purity of the oligonucleotides was ascertained by denaturing gel electrophoresis (Figure 2A). All oligonucleotides appear as a single band and migrate according to their respective size. If the samples are run under native conditions (Figure 2B), the oligonucleotides 3′-d(AT)$_n$xC$_4$(AT)$_n$-3′ (n = 4, 5) migrate as one sharp band of similar mobility as the antiparallel hairpin 5′-d(AT)$_5$C$_4$(AT)$_5$-3′, demonstrating that intramolecular structures (hairpins) were formed. For the oligonucleotides with $n \geq 6$, bands corresponding to the formation of dimeric and trimeric antiparallel species are observed. While the trimer is observed for both n = 6 and n = 7, the dimeric species is only observed for n = 7. This demonstrates that neither intramolecular structure (hairpin, dumbbell) is formed for the oligonucleotides with $n > 5$ under the gel electrophoretic conditions. The dumbbell structure, which would be expected to be more stable for longer sequences, is, even under these favorable conditions, not observed, supporting our previous assignment of the parallel-stranded hairpin structure for n = 4, 5. This is also confirmed from the observation that if these samples are applied at a 10-fold higher strand concentration (30 μM) to the gel, only single bands with the same mobility

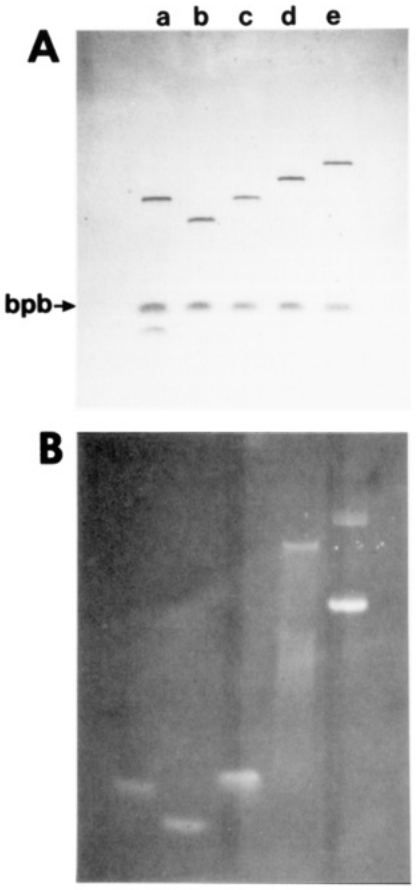

FIGURE 2: Gel electrophoretic analysis of alternating AT oligomers that have the potential to form parallel-stranded hairpin structures. (A) Denaturating 15% polyacrylamide gel (8 M urea, 90 mM Tris, and 5 mM EDTA, pH 8.3): lane a, 5′-d$(AT)_5C_4(AT)_5$-3′; lane b, 3′-d$(AT)_4xC_4(AT)_4$-3′; lane c, 3′-d$(AT)_5xC_4(AT)_5$-3′; lane d, 3′-d$(AT)_6xC_4(AT)_6$-3′; lane e, 3′-d$(AT)_7xC_4(AT)_7$-3′. Visualization by ultraviolet shadowing bpb, bromophenol blue (B) Nondenaturing 15% polyacrylamide gel (10 mM $MgCl_2$/90 mM Tris, pH 8.3) run at 4 °C. Approximately 0.01 OD_{260nm} of each oligonucleotide was loaded in 10 μL of loading buffer. Gels were stained with 0.5 μg/mL ethidium bromide/50 mM NaCl at 4 °C, and bands were visualized by fluorescence: lane a, 5′-d$(AT)_5C_4(AT)_5$-3′; lane b, 3′-d$(AT)_4xC_4(AT)_4$-3′; lane c, 3′-d$(AT)_5xC_4(AT)_5$-3′; lane d, 3′-d$(AT)_6xC_4(AT)_6$-3′; lane e, 3′-d$(AT)_7xC_4(AT)_7$-3′.

as in the low concentration gel are observed (data not shown). However, while for 3′-d$(AT)_4xC_4(AT)_4$-3′ the band appears sharp as in the low concentration gel, it is rather broad for 3′-d$(AT)_5xC_4(AT)_5$-3′. This indicates that for the latter oligonucleotide multimeric structures are formed in the loading buffer, which as they are diluted upon entering the gel convert to a hairpin. Thus, the gel electrophoretic data support PS hairpin formation for the oligonucleotides with 8 and 10 AT base pairs in the stem at a strand concentration <30 μM.

Optical Properties of Parallel Alternating AT Hairpins. The gel electrophoretic experiments demonstrate that the formation of intermolecular antiparallel multimeric structures is in competition with parallel-stranded hairpin formation. The oligonucleotides were therefore studied at different strand concentrations in order to establish the concentration range in which parallel-stranded alternating AT segment are stable. We and others have previously shown that the formation of parallel-stranded DNA structures is associated with characteristic spectral properties (Germann et al., 1988, 1989; Ramsing & Jovin, 1988; Ramsing et al., 1989). At 3.5–5.3 μM strand concentration, both 3′-d$(AT)_4xC_4(AT)_4$-3′ and 3′-d$(AT)_5xC_4(AT)_5$-3′ show blue-shifted absorption spectra at low temperatures while the longer sequences and the antiparallel control hairpin do not (data not shown). Blue-shifted spectra were also noted for the homooligomeric parallel-stranded hairpins (van de Sande et al., 1988). The most diagnostic plots for parallel-stranded structures that can be obtained from ultraviolet absorption data are the hyperchromicity profiles (Figure 3). For low strand concentrations, structures with both 8 and 10 AT stem pairs show plots that are consistent with PS DNA (Figure 3A,B) while the profile of 3′-d$(AT)_7xC_4(AT)_7$-3′ resembles that of the antiparallel control hairpin structure (Figure 3D). The hyperchromicity profile of 3′-d$(AT)_6xC_4(AT)_6$-3′ is also reminiscent of the antiparallel control, but the relatively low intensity at 245 nm indicates that a fraction of this oligonucleotide is also in a parallel-stranded conformation (Figure 3C). If the hyperchromicity profiles are recorded at a 10-fold higher strand concentration (44–64 μM), only 3′-d$(AT)_4xC_4(AT)_4$-3′ is still in a parallel-stranded conformation while 3′-d$(AT)_5xC_4(AT)_5$-3′ and 3′-d$(AT)_6xC_4(AT)_6$-3′ show hyperchromicity profiles that are characteristic for antiparallel DNA (Germann et al., 1989; Ramsing et al., 1989; Rippe et al., 1989). The inreased intensity in the hyperchromicity profiles for 3′-d$(AT)_7xC_4(AT)_7$-3′ is due to the increased stability of multimeric structures at the higher strand concentration. Thus, intramolecular PS DNA formation in the oligonucleotides studied here is strongly dependent on the chain length. In particular, the formation of alternating AT parallel-stranded segments is unfavorable for chain lengths ≥12 at micromolar strand concentration because of the stability of the competing antiparallel structures.

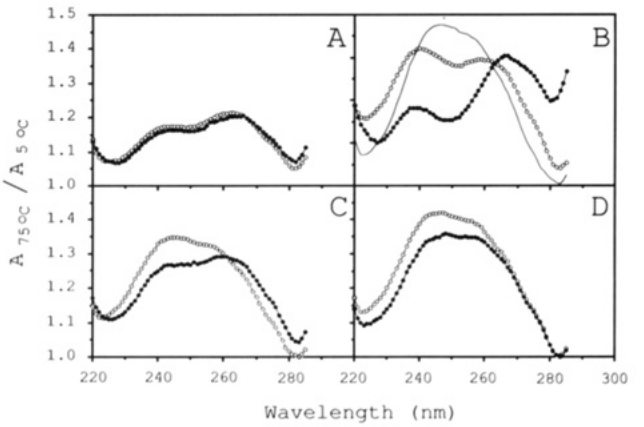

FIGURE 3: Wavelength-dependent hyperchromicity of alternating AT oligonucleotides in 0.8 M NaCl, 0.1 mM EDTA, and 10 mM sodium phosphate, pH 7.0. Absorbances of the oligonucleotides (~3.5 μM) at 70 °C were divided by the absorbance of the same samples at 6 °C. (A) 3′-d$(AT)_4xC_4(AT)_4$-3′: 5.3 μM (○) and 46 μM (●). (B) 3′-d$(AT)_5xC_4(AT)_5$-3′, 3.1 μM (○) and 55 μM (●); 5′-d$(AT)_5C_4(AT)_5$-3′, 9.1 μM (—). (C) 3′-d$(AT)_6xC_4(AT)_6$-3′: 3.7 μM (○) and 44 μM (●). (D) 3′-d$(AT)_7xC_4(AT)_7$-3′: 3.0 μM (○) and 53 μM (●).

The PS, APS, and denatured forms of 3′-d$(AT)_5xC_4(AT)_5$-3′ and the APS hairpin 5′-d$(AT)_5C_4(AT)_5$-3′ were also studied by circular dichroism (Figure 4). The most prominent feature in the CD spectra of the PS form (1.15×10^{-4} M residue, 4.8 μM strands) is the appearance of two maxima above 260 nm which is not observed for the APS control hairpin. In contrast, CD spectra of these oligonucleotides recorded under denaturing conditions are essentially the same as would be expected for denatured oligonucleotides containing similar sequences (Figure 4B). These experiments confirm that the PS structure is distinctly different from either APS or denatured forms since the CD spectrum of the PS form cannot be obtained as a linear combination of denatured and antiparallel structures. The circular dichroism of alternating

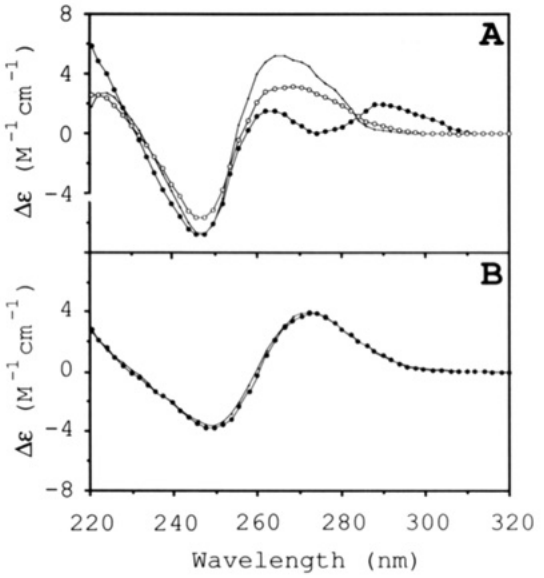

FIGURE 4: Circular dichroism of alternating AT oligonucleotides in 0.8 M NaCl, 0.1 mM EDTA, and 10 mM sodium phosphate pH 7.0. (A) Native conditions, 5 °C: 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′, 4.8 μM (●); 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′, 72 μM (○); 5′-d(AT)$_5$C$_4$(AT)$_5$-3′, 4.4 μM (+). (B) Denaturing conditions, 73 °C: 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′, 4.8 μM (●); 5′-d(AT)$_5$C$_4$(AT)$_5$-3′, 4.4 μM (+).

AT parallel-stranded sequences also differs from that observed previously for homooligomeric AT hairpins (van de Sande et al., 1988; Germann et al., 1989). In particular, we note the lower intensity and the appearance of a double peak above 260 nm.[2] Both alternating and homooligomeric AT parallel-stranded DNA show a positive band at 190 nm (data not shown) that is characteristic for right-handed helices (Germann et al., 1989).

We have utilized the spectral differences observed for parallel and antiparallel alternating AT segments to characterize the high-concentration form of 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′ (72 μM). In agreement with our assignment based on the hyperchromicity profiles, we find that the CD spectrum under these conditions is not compatible with parallel-stranded helix formation but resembles that of antiparallel-stranded DNA (Figure 4A).

Parallel-stranded DNA may also be identified from its weaker interaction and lower fluorescent quantum yield (63%), compared to antiparallel DNA, with the minor groove-specific drug Hoechst 33258 (Germann et al., 1989). The tendency of the alternating AT PS oligonucleotides to form intramolecular antiparallel structures, however, prevent us from recording complete binding isotherms for these oligonucleotides. We therefore measured the fluorescence of these oligonucleotides under saturating drug concentrations (1.6 μM Hoechst 33258, 0.3 μM oligonucleotide, 0.4 M NaCl, 0.1 mM EDTA, and 10 mM Tris, pH 8.3 at 5 °C). On the basis of both gel electrophoretic and spectroscopic evidence (UV, CD), 3′-d(AT)$_4$xC$_4$(AT)$_4$-3′ and 3′-d(AT)$_5$xC$_4$(AT)$_5$-3′ form a parallel-stranded hairpin at these strand concentrations. Both sequences show fluorescence of 65% and 63% of the APS control, respectively. These values compare well to the 63% fluorescence limit that was observed for the homooligomeric hairpin 3′-d(T)$_8$xC$_4$(A)$_8$-3′ and suggest structural similarities for both homooligomeric and alternating parallel-stranded AT sequences. In contrast, 3′-d(AT)$_6$xC$_4$(AT)$_6$-3′ and 3′-d-(AT)$_7$xC$_4$(AT)$_7$-3′ show a considerably higher fluorescence of 84% and 100%, respectively, that approaches the fluorescence of the antiparallel control hairpin. In addition, we observe that the fluorescence of these oligonucleotides also depends on the strand concentration (data not shown). This supports our previous conclusion that the oligonucleotides with $n \geq 6$ have a high tendency to form antiparallel structures. The lower fluorescence observed for 3′-d(AT)$_6$xC$_4$(AT)$_6$-3′ is consistent with the expected formation of a parallel-stranded hairpin for this oligonucleotide at lower strand concentrations.[3]

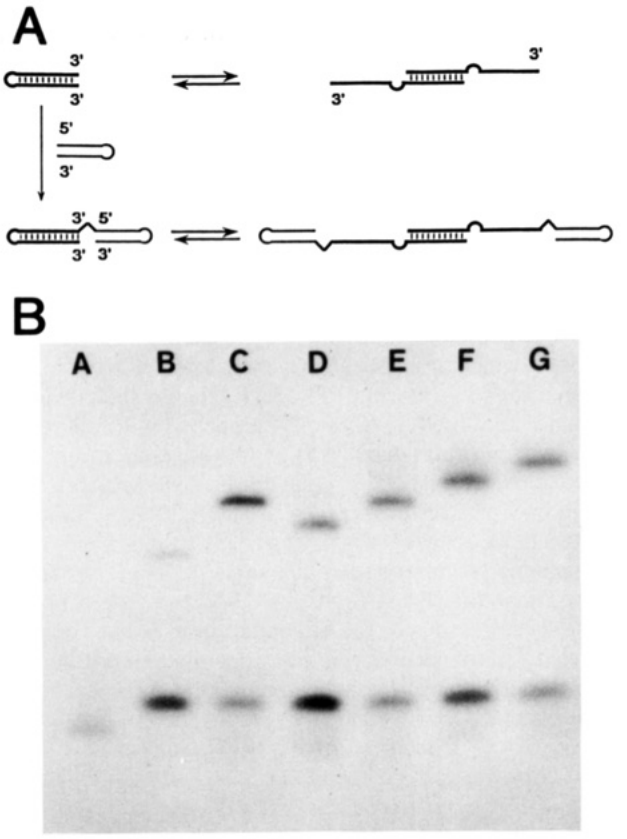

FIGURE 5: (A) Schematics of the ligation reaction of alternating AT oligonucleotides (in monomer–dimer equilibrium) with end-labeled R-1 hairpin. (B) Autoradiogram of a 15% denaturing polyacrylamide gel of the ligation reactions of 5′-end-labeled hairpin R-1 (0.2 μM) to unphosphorylated alternating AT oligonucleotides (2.0 μM): lane A, 5′-end-labeled hairpin R-1; lane B, ligation of R-1 to a 10-fold excess of unphosphorylated R-1; lane C, R-1 and 5′-d(AT)$_5$C$_4$(AT)$_5$-3′; lane D, R-1 and 3′-d(AT)$_4$xC$_4$(AT)$_4$-3′; lane E, R-1 and 3′-d-(AT)$_5$xC$_4$(AT)$_5$-3′; lane F, R-1 and 3′-d(AT)$_6$xC$_4$(AT)$_6$-3′; lane G, R-1 and 3′-d(AT)$_7$xC$_4$(AT)$_7$-3′.

Substrate Properties of Parallel Alternating AT Hairpins. Previous studies on homooligomeric parallel-stranded DNA have shown that both 3′–3′ and 5′–5′ hairpins ligate to a blunt-ended 5′–3′ oligonucleotide duplex in the presence of T4 polynucleotide ligase (van de Sande et al., 1988). This reaction proceeds at a lower rate than for a conventional ligation between antiparallel duplexes. The parallel-stranded hairpins studied here are incapable of forming antiparallel duplex structures that are substrates for T4 polynucleotide ligase (Figure 5A). To simplify the analysis of the products of ligation between parallel and antiparallel substrates by denaturing gel electrophoresis, we used a 5′-^{32}P-labeled blunt-ended hairpin (R-1 hairpin) to ligate to the 3′–3′ alternating A-T oligonucleotides. As demonstrated in Figure 5B, the R-1 hairpin ligated to a 10-fold excess (i.e., 2 μM) of the alter-

[2] These observations are consistent with the circular dichroism of a parallel-stranded intermolecular duplex structure containing both homooligomeric and alternating AT components (Germann et al., 1988).

[3] Note that Hoechst 33258 because of its high affinity to antiparallel-stranded DNA will stabilize antiparallel forms.

nating 3′-d$(AT)_n$x$C_4(AT)_n$-3′ (n = 4–7) oligonucleotides with good efficiencies (lanes D–G), and the products showed different mobilities than for the ligation product (30-mer) of R-1 to unphosphorylated R-1 (lane B). The fastest moving band that is visible in all ligation reactions corresponds to the AMP-activated ligation intermediate of R-1 which accumulates at low reaction temperatures (Harvey et al., 1971). The ligation products of the R-1 hairpin and 3′-d$(AT)_n$x$C_4(AT)_n$-3′ have the gel electrophoretic mobility of (19 + 4n)-mers that are the result of the addition of one R-1 hairpin to one 3′–3′ alternating AT oligonucleotide. Lane C shows the ligation of end-labeled R-1 and unphosphorylated 5′-d$(AT)_5C_4$-$(AT)_5$-3′, which produces a 39-mer that acts as a convenient size marker. The hairpin R-1 also ligated to the previously characterized 3′-d$(A)_{10}$xC_4T_{10}-3′ parallel-stranded hairpin as was expected (data not shown). Control reactions using 5′-end-labeled d$(A)_{10}$ or d$(T)_{10}$ single-stranded substrates and R-1 did not produce any ligation products, thereby ruling out the possibility of ligations involving single-stranded ends under our reaction conditions (data not shown). Thus, our results suggest that for 3′-d$(AT)_6$x$C_4(AT)_6$-3′ and 3′-d$(AT)_7$xC_4-$(AT)_7$-3′ a fraction of the oligonucleotide is in a parallel-stranded, blunt-ended hairpin conformation that can ligate to the R-1 hairpin while competing antiparallel forms containing single-stranded ends cannot (Figure 5A).

Thermodynamic Properties of Parallel-Stranded Alternating AT Segments. The observation that for 3′-d-$(AT)_n$x$C_4(AT)_n$-3′ with $n \geq 6$ antiparallel intermolecular structures were formed indicates that alternating AT segments are less favorable than homooligomeric tracts for PS DNA formation. This is also supported by the finding that 3′-d-$(AT)_5$x$C_4(AT)_5$-3′, when applied at higher strand concentrations, migrates as a broader band under nondenaturing gel conditions, which was not observed for 3′-d$(T)_{10}$xC_4A_{10}-3′ (van de Sande et al., 1988). In addition, the hyperchromicity profile for this oligonucleotide recorded at higher strand concentration (55 μM) indicated that antiparallel structures were formed. On the basis of the spectroscopic and gel electrophoretic experiments, it is evident that with exception of 3′-d-$(AT)_4$x$C_4(AT)_4$-3′ the structures formed are strongly dependent on the strand concentrations studied here. Consequently, the experimental conditions (i.e., strand concentration) for these oligonucleotides must be adjusted so that the PS form becomes dominant. The two oligonucleotides that form parallel-stranded hairpin structures at micromolar strand concentrations, 3′-d$(AT)_4$x$C_4(AT)_4$-3′ and 3′-d$(AT)_5$x$C_4(AT)_5$-3′, and the antiparallel control 5′-d$(AT)_5C_4(AT)_5$-3′ were further characterized by ultraviolet melting studies. For all three hairpins, the thermally induced helix–coil transition appeared monophasic and fully reversible and independent of the strand concentration used (0.4–2.0 OD/mL). The differentiated UV melting curves were fitted to a two-state model that incorporated temperature-dependent absorbances of helix and coil forms (Figure 6). From the data presented in Table I, it follows that the parallel-stranded 3′-d$(AT)_5$x$C_4(AT)_5$-3′ is both thermally and enthalpically (ΔH°_{vH} = 137.6 kJ mol^{-1}, T_M = 39.3 °C) significantly less stable than either the antiparallel control 5′-d$(AT)_5C_4(AT)_5$-3′ (ΔH°_{vH} = 244 kJ mol^{-1}, T_M = 57.2 °C) or the homooligomeric parallel-stranded hairpin 3′-d$(T)_{10}$xC_4A_{10}-3′ (ΔH°_{vH} = 205 kJ mol^{-1}, T_M = 40.6 °C at 0.1 M NaCl) (Germann et al., 1989). The helix–coil enthalpy obtained for the parallel-stranded hairpin 3′-d$(AT)_5$xC_4-$(AT)_5$-3′ is similar to 3′-d$(TA)_5$x$T_4(TA)_5$-3′, suggesting that the remarkably low enthalpy is a consequence of the alternating AT part of the hairpin and not the hairpin loop or the loop–stem interface (Table I). Using the data for the helix–coil transition for 3′-d$(AT)_4$x$C_4(AT)_4$-3′, we calculate that the enthalpy for disrupting an AT base pair in a parallel-stranded alternating AT helix of 16 kJ mol^{-1} is significantly lower than that for homooligomeric AT helixes (24 kJ mol^{-1}).

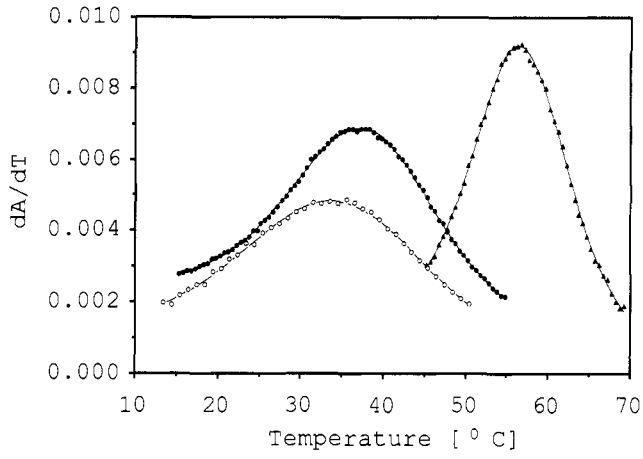

FIGURE 6: Differentiated ultraviolet melting curve of 3′-d-$(AT)_4$x$C_4(AT)_4$-3′ (○), 3′-d$(AT)_5$x$C_4(AT)_5$-3′ (●), and 5MHd-$(AT)_5C_4(AT)_5$-3′ (▲), recorded at 268 nm in 0.8 M NaCl, 0.1 mM EDTA, and 10 mM phosphate, pH 7.0. The total strand concentration of the samples was approximately 3.5 μM. The solid lines represent the best fit obtained from nonlinear least-squares curve-fitting to a two-state model (see Materials and Methods).

Table I: Thermodynamic Data for the Helix–Coil Transition of Alternating Parallel and Antiparallel AT Hairpin Structures[a]

sequence	T_M (°C)	ΔH°_{vH} (kJ mol^{-1})	ΔS° (kJ mol^{-1} K^{-1})
3′-d0AT)$_4$x$C_4(AT)_4$-3′	36.3 ± 0.4	106 ± 3	0.343 ± 0.01
3′-d$(AT)_5$x$C_4(AT)_5$-3′	39.3 ± 0.3	137 ± 4	0.441 ± 0.01
3′-d$(TA)_5$x$T_4(TA)_5$-3′	36.3 ± 0.4	132 ± 10	0.430 ± 0.03
3′-d$(T)_{10}$x$C_4(A)_{10}$-3′[b]	40.6 ± 0.2	205 ± 3	0.653 ± 0.01
5′-d$(AT)_5C_4(AT)_5$-3′	57.2 ± 0.4	244 ± 18	0.738 ± 0.05
5′-d$(TA)_5T_4(TA)_5$-3′	58.4 ± 0.2	229 ± 6	0.690 ± 0.02

[a] Melting temperatures were recorded in 0.8 M NaCl, 0.1 mM EDTA, and 10 mM phosphate, pH 7.0. Data are the averages of several independent measurements. T_M, melting temperature (θ = 0.5); ΔH°_{vH}, van't Hoff enthalpy; ΔS°, entropy. [b] The melting temperature for this sample was determined in 0.1 M NaCl (Germann et al., 1989).

A low helix–coil transition enthalpy (90 kJ mol^{-1}) was also observed for the intramolecularly formed parallel-stranded structure for 3′-d$(AT)_5$-O-$(CH_2)_6$-O-$(AT)_5$-3′ (Shchyolkina et al., 1989). This value is considerably lower than the enthalpy we obtain for 3′-d$(AT)_5$x$C_4(AT)_5$-3′ and may in part be due to the missing stacking interactions within and next to the C_4 hairpin loop. In addition, Shchyolkina et al. (1989) noted that intramolecular complex formation occurred only at very low strand concentration (~0.10 OD_{260nm}/mL), thereby limiting the accuracy of the data.

Estimation of the Chain Growth Enthalpy of Parallel-Stranded Structures Containing AT Base Pairs. To further study these apparent differences in stability of homooligomeric and alternating parallel-stranded AT tracts, we have analyzed the thermodynamic data for the intrmaolecular helix formation using a simple all or none model ($K = \sigma s^N$). According to this model, the alternating PS structures have a much lower chain growth parameter (3.2 at 0 °C)[4] than the homooligomeric parallel AT sequences (6.4 at 0 °C) (Table II). These data may be used to aid in the design and prediction of the chain growth enthalpy of putative parallel-stranded AT sequences

[4] The low chain growth parameter obtained indicates that more sophisticated models are indicated for an in-depth study of the helix formation in alternating PS hairpin structures.

Table II: Thermodynamic Data for PS and APS Helix Growth and Nucleation in AT Hairpin Sequences[a]

	ΔH_s	ΔS_s	s(0 °C)	ΔH_β
PS $(AT)_n$	−16	−0.050	3.2	6
PS A_n-T_n[b]	−24	−0.073	6.4	14
APS A_n-T_n[b]	−29	−0.084	14.5	13

[a] The sequences were 3′-d$(AT)_4$x$C_4$$(AT)_4$-3′, 3′-d$(AT)_5xC_4$$(AT)_5$-3′, 3′-d$(T)_8xC_4$$(A)_8$-3′, 3′-d$(T)_{10}xC_4$$(A)_{10}$-3′, 5′-d$(T)_8C_4(A)_8$-3′, and 5′-d-$(T)_{10}C_4(A)_{10}$-3′. Data were derived from a plot of ΔH°_{vH} and ΔS° versus the number of nearest neighbors ($N - 1$). ΔH_s and ΔH_β, enthalpy of chain growth and nucleation, respectively; ΔS_s, entropy of chain growth. s is the equilibrium constant of forming one AT base pair on top of a preexisting base pair. Average values for nucleation of a PS helix in AT sequences: ΔH_β = 10 kJ mol^{-1}, ΔS_β = 0.005 kJ mol^{-1} K^{-1}. Note that these quantities refer to helix growth rather than helix dissociation. [b] Calculated from data given by Germann et al. (1989).

formed intra- or intermolecularly. The enthalpy is then calculated by using

$$\Delta H_{\text{growth calcd}} = n_1 \Delta H_s^{\text{altAT}} + n_2 \Delta H_s^{\text{homAT}} \quad (5)$$

where n_1 and n_2 indicate the number of alternating and homooligomeric AT nearest-neighbor interactions, respectively. The calculation of the helix–coil transition enthalpy requires the knowledge of the nucleation enthalpy in parallel-stranded sequences. For the hairpin sequences investigated here, an average of $\Delta H_\beta = -10$ kJ mol^{-1} is obtained. This value also contains the contribution of the hairpin loop and can therefore not be used a priori for the nucleation of dimeric helices. However, particularly for longer duplexes, the helix–coil transition enthalpy is primarily determined by the chain growth enthalpy and can thus be estimated by it. The predicted chain growth enthalpies for the parallel-stranded structure that are known to date are shown in Table III. As is evident, particularly in view of the simple model that was employed, $\Delta H_{\text{growth calcd}}$ approximates the experimental enthalpy in most cases. Investigation of additional parallel-stranded helixes in conjunction with more sophisticated models for the helix formation should enable us to predict the enthalpy of PS structures with increasing accuracy. It should be noted that we have not made any distinction between 3′-A-T-5′ and 3′-T-A-5′ base interactions since they are equivalent in a parallel-stranded structure but not in antiparallel DNA.

The conformational polymorphism of 3′-d$(AT)_5$x$C_4$$(AT)_5$-3′ was further investigated by recording hyperchromicity profiles at different temperatures for this sample (21.3 μM strand concentration). At low temperature, this sample exhibits a hyperchromicity that is characteristic for antiparallel DNA (Figure 7A,B). However, if the temperature is raised, the profile increasingly resembles that of parallel-stranded DNA, indicating that the sample converts to a parallel-stranded hairpin structure prior to melting. Our results show that the sizes of the regions of stability (e.g., temperature, strand concentration) for the parallel-stranded hairpins are strongly determined by the chain length of these oligonucleotides. For instance, for $n = 4$, PS hairpin conformation is preferred up to concentrations suitable for NMR experiments (0.5–1.0 mM), while for $n > 6$ the stability region for the antiparallel forms extends to micromolar strand concentrations.

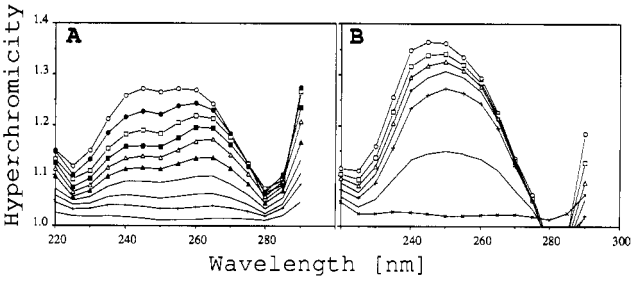

FIGURE 7: Wavelength-dependent hyperchromicity of (A) 3′-d-$(AT)_5$x$C_4$$(AT)_5$-3′ (21.3 μM and (B) 5′-d$(AT)_5C_4(AT)_5$-3′ (31 μM) in 0.8 M NaCl, 0.1 mM EDTA, and 10 mM sodium phosphate, pH 7.0: 5 (○), 10 (●), 15 (□), 20 (■), 25 (△), 30 (▲), 35 (△), 40 (♦), 45 (+), 55 (—), and 65 °C (×).

Conclusion

Gel electrophoretic and spectroscopic evidence shows that alternating AT sequences can form parallel-stranded segments. The formation of alternating PS segments is, however, less favorable than for homooligomeric sequences as evidenced by the lower chain growth parameter s. The consequence of a lower s is that for AT stretches longer than 10 base pairs competing conformations such as antiparallel-stranded concatamers are favored even at low strand concentrations. Thus, alternating AT segments are compatible with PS DNA formation, but longer alternating AT tracts (and to a lesser extent also homooligomeric AT tracts) prefer antiparallel strand orientation. This accounts for the observed polymorphic behavior of 3′-d$(AT)_5$x$C_4$$(AT)_5$-3′. The results presented in this paper enable us to present guidelines for the sequence selection in the design of parallel-stranded AT structures and allow the estimation of the chain growth enthalpy of parallel-stranded structures. For the design of parallel-stranded structures, both homooligmeric as well as alternating AT components can be used. However, as demonstrated, homooligomeric AT sequences adopt parallel-stranded conformation most easily, and their use should be maximized. At the same time, the possible formation of either intra- or intermolecular formed antiparallel structures must be minimized by the use of alternating AT components to destabilize these undesired antiparallel forms (Germann et al., 1988).

In order to complete the investigation of the compositional requirements for the formation of parallel-stranded DNA, the potential existence of GC base pairs in a parallel-stranded helix must be addressed. We have therefore investigated the consequence of inserting a single CG base combination in a parallel-stranded duplex consisting of AT base pairs. The presence of a GC base combination in a parallel-stranded helix results in a destabilization of the parallel-stranded structure

Table III: Prediction of the Chain Growth Enthalpy of Several Intra- and Intermolecularly Formed Parallel-Stranded Helices Based on Chain Growth Parameters Derived from Parallel-Stranded Hairpin Structures (Table II)

	intramolecular PS DNA				intermolecular PS DNA			
	homo AT		alt AT					
					a	*b*	*c*	*d*
AT bp homo	7[e]	9[e]			7	15	17	17
AT bp alt			7	9	7	5	7	7
$\Delta H^\circ_{vH\,exp}$ (kJ mol^{-1})	154	205	106	137	309	353	459	507
$\sum \Delta H_{\text{growth calcd}}$ (kJ mol^{-1})	168	216	112	144	280	440	520	520

[a] 15 base pair dimeric PS DNA (Germann et al., 1988). [b] ps C2–C3. [c] ps D1–D2. [d] ps D3–D4 parallel-stranded structures (Ramsing & Jovin, 1988; Ramsing et al., 1989). [e] Germann et al. (1989).

to a similar extent as a CT base combination (submitted for publication).

References

Caceci, M. S., & Cacheris, W. P. (1984) *Byte 9* (5), 340–362.

Cantor, C. R., & Schimmel, P. R. (1980) in *Biophysical Chemistry*, W. Freeman and Co., San Francisco.

Chaconas, G., & van de Sande, J. H. (1980) *Methods Enzymol. 65*, 75–85.

Germann, M. W. (1989) Ph.D. Thesis, The University of Calgary.

Germann, M. W., Kalisch, B. W., & van de Sande, J. H. (1988) *Biochemistry 27*, 8302–8306.

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

Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N., & Kollman, P. A. (1986) *J. Biomol. Struct. Dyn. 3*, 843–857.

Harvey, C. L., Gabriel, T. F., Wilt, E. M., & Richardson, C. C. (1971) *J. Biol. Chem. 246*, 4523–4530.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Orbons, L. P. M. (1987) Ph.D. Thesis, University of Leiden.

Pattabiraman, N. (1986) *Biopolymers 25*, 1603–1606.

Ramsing, N. B., & Jovin, T. M. (1988) *Nucleic Acids Res. 14*, 6659–6676.

Ramsing, N. B., Rippe, K., & Jovin, T. M. (1989) *Biochemistry 28*, 9528–9535.

Rippe, K., & Jovin, T. M. (1989) *Biochemistry 28*, 9542–9549.

Rippe, K., Ramsing, N. B., & Jovin, T. M. (1989) *Biochemistry 28*, 9536–9541.

Shchyolkina, A. K., Lysov, Y. P., Il'ichova, I. A., Chernyi, A. A., Golova, Y. B., Ckernov, B. K., Gottikh, B. P., & Florentiev, V. L. (1989) *FEBS Lett. 244*, 39–42.

Summers, M. F., Byrd, R. A., Gallo, K. A., Samsom, C. J., Zon, G., & Egan, W. (1985) *Nucleic Acids Res. 13*, 6375–6386.

Tchurikov, N. A., Chernov, B. K., Golova, Y. B., & Nechipurenko, Y. D. (1989) *FEBS Lett. 257*, 415–418.

van de Sande, J. H., Ramsing, N. B., Germann, M. W., Elhorst, W., Kalisch, B. W., Kitzing, E. V., Pon, R. T., Clegg, R. C., & Jovin, T. M. (1988) *Science 241*, 551–557.

Organization, Structure, and Polymorphisms of the Human Profilaggrin Gene‡

Song-Qing Gan,§ O. Wesley McBride,∥ William W. Idler,§ Nedialka Markova,§ and Peter M. Steinert*,§

Dermatology Branch and Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received April 6, 1990; Revised Manuscript Received July 12, 1990

ABSTRACT: Profilaggrin is a major protein component of the keratohyalin granules of mammalian epidermis. It is initially expressed as a large polyprotein precursor and is subsequently proteolytically processed into individual functional filaggrin molecules. We have isolated genomic DNA and cDNA clones encoding the 5′- and 3′-ends of the human gene and mRNA. The data reveal the presence of likely "CAT" and "TATA" sequences, an intron in the 5′-untranslated region, and several potential regulatory sequences. While all repeats are of the same length (972 bp, 324 amino acids), sequences display considerable variation (10–15%) between repeats on the same clone and between different clones. Most variations are attributable to single-base changes, but many also involve changes in charge. Thus, human filaggrin consists of a heterogeneous population of molecules of different sizes, charges, and sequences. However, amino acid sequences encoding the amino and carboxyl termini are more conserved, as are the 5′ and 3′ DNA sequences flanking the coding portions of the gene. The presence of unique restriction enzyme sites in these conserved flanking sequences has enabled calculations on the size of the full-length gene and the numbers of repeats in it: depending on the source of genomic DNA, the gene contains 10, 11, or 12 filaggrin repeats that segregate in kindred families by normal Mendelian genetic mechanisms. This means that the human profilaggrin gene system is also polymorphic with respect to size due to simple allelic differences between different individuals. The amino- and carboxyl-terminal sequences of profilaggrin contain partial or truncated repeats with unusual un-filaggrin-like sequences on the termini. Such sequences are reminiscent of propeptides encountered in other structural protein systems. We suggest these sequences are required for the assembly of the accumulating protein into large keratohyalin granules among the keratin filaments in the granular cells and aid in later processing events.

Filaggrins represent an important class of intermediate filament-associated proteins (IFAPs) that function, at least in part, in the aggregation of keratin intermediate filaments into an organized "keratin pattern" during terminal stages of normal differentiation in mammalian epidermis (Dale et al., 1978, 1989; Steinert et al., 1981; Steinert & Roop, 1988). On the basis of data from both protein chemical studies (Harding & Scott, 1983; Resing et al., 1984, 1985) and more recent cloning experiments (Haydock & Dale, 1986; Rothnagel et al., 1987; Rothnagel & Steinert, 1990; McKinley-Grant et al., 1989), filaggrins are initially synthesized as large polyprotein precursors ("profilaggrins") consisting of many protein repeats

‡ The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02929.

* To whom correspondence should be addressed at the Laboratory of Skin Biology, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Building 10, Room 12N238, Bethesda, MD 20892.

§ Dermatology Branch.

∥ Laboratory of Biochemistry.