Molecular and Thermodynamic Properties of $d(A^+-G)_{10}$, a Single-Stranded Nucleic Acid Helix without Paired or Stacked Bases[†,‡]

Nina G. Dolinnaya, Emory H. Braswell,[§] John A. Fossella, Horst Klump,[∥] and Jacques R. Fresco*

Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544-1014

Received April 1, 1993; Revised Manuscript Received July 8, 1993[●]

ABSTRACT: Previously (Dolinnaya & Fresco, 1992), on the basis of an analysis of UV absorption and CD properties as a function of temperature and pH, the secondary structure of the deoxyoligonucleotide $d(A^+-G)_{10}$ was hypothesized to be helical and intramolecular in origin, being stabilized not by stacked bases or hydrogen-bonded base pairs but instead by ionic bonds between positively charged adenine residues and distal negatively charged phosphates. Several other properties are now shown to be consistent with this unusual type of structure. The molecular weight determined for $d(A^+-G)_{10}$ by sedimentation equilibrium is that of the single strand, and consistent with this, there is no molecular weight change on helix disruption. Formation of the $d(A^+-G)_{10}$ helix is accompanied by cooperative uptake of nine protons, corresponding to nine adenine residues that can form ionic bonds with all the available distal phosphates, i.e., the $n+1$ or the $n+2$ phosphates. The thermodynamic parameters of this helical structure obtained from both van't Hoff analysis of the melting of the structure and calorimetric measurements are in keeping with the ionic properties of the proposed structure. So are the dependence of its stability on pH and ionic strength, and also on oligomer length when compared with the behavior of $d(A^+-G)_6$. The possible role of this type of secondary structure in protein recognition of the single-stranded homopurine element of H-DNA is evaluated.

The deoxyoligonucleotide $d(A-G)_{10}$ shows an unusual dichotomy in its optical properties in solution of modest ionic strength and acidic pH in that it displays intense near-UV circular dichroism but only slight UV hypochromicity (Dolinnaya & Fresco, 1992). The conformation displaying these properties exhibits a significantly elevated apparent pK_a ($_{app}pK_a$) for protonation of its adenine residues, the extent of elevation being inversely proportional to the ionic strength. Moreover, the structure is destabilized upon protonation of the dG residues, and its stability is independent of oligomer concentration.

This unusual mix of properties has been interpreted in terms of a single-stranded helical secondary structure stabilized not by hydrogen-bonded base pairs or base stacks but by Coulombic interactions between positively charged dA residues and distal negatively charged backbone phosphates, while the dG residues are situated so as not to overlap their dA nearest neighbors (Dolinnaya & Fresco, 1992). In this report, we describe in more quantitative terms a variety of properties whose parameters are shown to be fully in keeping with this general scheme, and suggest a potential biological role for this new type of nucleic acid secondary structure.

EXPERIMENTAL PROCEDURES

Oligodeoxyribonucleotides. $d(A-G)_{10}$ and $d(A-G)_6$ were synthesized and purified to homogeneity as previously described (Dolinnaya & Fresco, 1992).

Sedimentation Equilibrium Measurements. Solutions of $d(A^+-G)_{10}$ in 0.01 M sodium acetate, pH 4.0, varying from 1 to 0.01 mg/mL were centrifuged overnight in a Beckman XLA analytical ultracentrifuge at 40 000 or 52 000 rpm at 4 °C. External loading cells (Anservin et al., 1970) with three pairs of solvent–solute channels allowing solution columns of ~2.5 mm were used. The solution channels were loaded with 4 or 10 μL (for 4- and 12-mm path-length cells, respectively) of a high-density fluorocarbon oil (no. FC43, Minnesota Mining and Manufacturing Co.) in order to make the bottoms of the solutions visible by raising them. The solvent was used as the reference in the solvent compartment. Appropriate combinations of cell optical paths of 4 and 12 mm and wavelengths of light of 300 and 280 nm were used to measure by absorption the oligonucleotide concentration gradients established at equilibrium. Sedimentation equilibrium was achieved well within the overnight centrifugation time. One oligomer solution (0.01 mg/mL) was used to half-fill a 12-mm-path-length double-sector cell, resulting in a ~5-mm solution column. This solution was centrifuged at 48 000 rpm at 4 °C for ~24 h (during which equilibrium was attained), and the gradient was detected by absorption at 255 nm. Two oligomer solutions, 1 and 0.03 mg/mL in 4- and 12-mm-path-length cells, respectively, with a solution height of 2.5 mm were centrifuged overnight at 52 000 rpm at 40 °C. The resulting oligomer gradients were detected by absorption at 300 and 280 nm, respectively.

Data were analyzed using a nonlinear least-squares program. In order to perform a global fit using all the data taken at different speeds and cell path lengths, it is necessary to use speed factors that are proportional to the square of the running speeds and concentration factors that are equal to the concentration in mg/mL that can produce an absorbance of 1 for the particular cell and wavelength (Johnson et al., 1981).

Electrometric Titration. $d(A-G)_{10}$ (3.16×10^{-4} M residues) in 0.01 M NaCl was scrubbed with filtered N_2 equilibrated with H_2O to remove dissolved CO_2. Titration was then

[†] This work was supported by a grant to J.R.F. from the National Institutes of Health (GM42936). Support for the ultracentrifugation experiments was provided by a grant to E.H.B. from the National Science Foundation (DIR8717034).

[‡] This is paper no. 20 in the series entitled *Polynucleotides*, of which the last is Fossella et al. (1993).

* To whom correspondence should be addressed.

[§] National Facility for Analytical Ultracentrifugation, University of Connecticut, Storrs, CT 06269-3140.

[∥] Present address: Department of Biochemistry, University of Cape Town, South Africa.

[●] Abstract published in *Advance ACS Abstracts*, September 1, 1993.

0006-2960/93/0432-10263$04.00/0

performed at 23 °C with standardized 0.01 N HCl, pH being monitored with a Radiometer pH 26 meter. Proton uptake was calculated as the difference between H^+ added and H^+ measured from the pH. A solvent blank was similarly titrated, and its very slight "proton uptake", evident below pH 3.5, was used to correct the uptake by the oligomer solution.

CD Spectroscopy. This was performed as previously described (Dolinnaya & Fresco, 1992).

Differential Scanning Calorimetry (DSC). Measurements were made using an adiabatic scanning microcalorimeter (DASM-4M Mashpriborintorg, Moscow) at the Biocalorimetry Center at Johns Hopkins University. In a typical run, a solvent blank containing 0.01 M sodium acetate, pH 4.0, and a solution containing 4.4×10^{-3} M residues of $d(A^+\text{-}G)_{10}$ were heated simultaneously after equilibration overnight at 1 °C. The solvent blank was used to correct the experimental DSC heat output plot for the oligomer, and the area under the resultant plot was used to determine the enthalpy change for the helix→coil transition between 0 and 70 °C by comparing it to an electrically generated calibration area (1.435 mcal).

RESULTS

Molecular Weight of $d(A^+\text{-}G)_{10}$. An essential feature of the model proposed for the $d(A^+\text{-}G)_{10}$ helix is that its secondary structure derives from intramolecular interactions. In that event, the molecular weight of the helix should be that of the single strand, and not some multiple thereof, so that the helix→coil transition for the structure should not be accompanied by a molecular weight change. To assess whether these expectations are met, sedimentation equilibrium measurements were made at pH 4.0 at 4 °C, where the helix is fully formed, and at 40 °C, where the helix is largely disrupted (Dolinnaya & Fresco 1992). At 4 °C, a 100-fold loading concentration range of monomer strands (1–0.01 mg/mL) was examined that spanned at equilibrium greater than a 1000-fold concentration range, while at 40 °C a 33-fold loading concentration range was examined.

To analyze the sedimentation gradient at equilibrium, one attempts to fit the observed oligomer distribution, assuming various models for monomer *strand* interaction. This was done with the aid of a computer program that produces three principal results: (a) the value of the *reduced molecular weight*, $M(1 - \phi'\rho)$, where ϕ' is the specific volume of the monomer strand at constant free energy (Eisenberg, 1990), and ρ is the density of the solvent; (b) the values of other constants related to the model (i.e., second virial coefficient, B, equilibrium constants for any monomer strand interaction, K_a, etc.); and (c) the value of the rms error. When a simple model, one for a single ideal species (without any assumption as to the molecular weight of that species), was fit to six sets of data (~1000 points) obtained at 4 °C, the rms error was small (0.011 absorbance unit), but the residuals displayed systematic error. On the other hand, a nonideal model of a single species with a second virial coefficient gave an even smaller rms error and, more important, an extremely small systematic error (see Figure 1 for a graph of the residuals). The values of $M(1 - \phi'\rho)$ and B obtained from this fit are $(3.01 \pm 0.04) \times 10^3$ g/mol and $(1.21 \pm 0.3) \times 10^{-5}$ L·mol/g^2, respectively.

A number of other models for equilibria between the single strand (i.e., the monomer strand) and specific multimeric complexes of that strand (e.g., dimer, trimer, etc.) were also tried (monomer–dimer, monomer–trimer, monomer–tetramer, monomer–hexamer, monomer–dimer–tetramer, and monomer–dimer–hexamer), but none yielded a fit as good as the single species nonideal model. Moreover, these association models all yielded extremely low values of ln K_a (e.g., −11.2 for monomer–dimer), indicating extremely small degrees of association, all the more so given that the oligomer concentration range examined (as a result of loading concentrations and centrifugation effects) was from ~1 μg/mL to over 1 mg/mL.

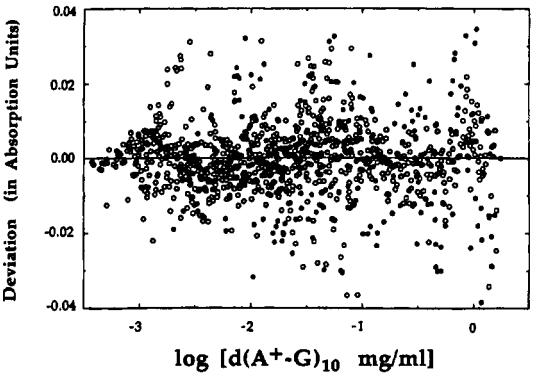

FIGURE 1: Residual plot showing the deviations for the sedimentation data at 4 (○) and 40 °C (●), using the nonideal model described in the text.

That data over such a range of loading concentrations, centrifugation speeds, detection wavelengths, and cell path lengths converge on a single model and value of $M(1 - \phi'\rho)$ eliminates concern for artifacts due to pressure effects or deviations from Beer's law. In the same manner, the best fit found for two sets of data (~200 points) at 40 °C was also the nonideal model. In this case, the values obtained for $M(1 - \phi'\rho)$ and B were $(3.01 \pm 0.23) \times 10^3$ g/mol and 3.03 ± 0.9 L·mol/g^2, respectively, with an rms error of 0.020 absorbance unit. Again, evidence of systematic error was insignificant (Figure 1).

Starting with the experimentally determined value of 3.01×10^3 g/mol for the *reduced molecular weight*, values for the density of the solvent of 1.0004 and 0.9957 at 4 and 40 °C, respectively, and a value of 6582.5 Da for the molecular mass of the single strand at pH 4.0, values for ϕ' can be calculated, assuming monomer, dimer, and higher aggregate molecular weights for the structure. Assumption of the monomer molecular weight yields values for ϕ' of 0.542 ± 0.007 and 0.54 ± 0.04 mL/g at 4 and 40 °C, respectively, in the same range as those for other nucleic acid structures, which for both DNA and RNA cases center about 0.5 mL/g (Eisenberg, 1990; Lindahl et al., 1965). In contrast, the assumption of molecular weights for dimer or higher aggregate models yields impossibly high values for nucleic acid-type molecules, e.g., 0.77 mL/g for a dimer. It is also noteworthy that the monomer model, which should be accompanied by no significant change in molecular weight on helix disruption at 40 °C (only the replacement of a few protons on dA residues by Na^+ ions associated with their phosphate partners; see below), gives essentially the same ϕ' values at the two temperatures. This could not be so if thermal denaturation of the helix involved an aggregate→monomer transition. Taken together, these results clearly show that the oligomer exists as a single strand at both 4 and 40 °C, albeit in different conformations.

Number of Protons Stabilizing $d(A^+\text{-}G)_{10}$. An electrometric titration from neutrality to pH 3.3 was performed at 23 °C to determine the number of ionic bonds that stabilize the structure. Figure 2 shows the course of the net uptake of protons by the oligomer (solid line) and for comparison the calculated uptake for an equimolar mixture of "unassociated" dA and dG residues, i.e., their 5 monophosphates (dashed line). It is apparent that proton uptake by the oligomer does

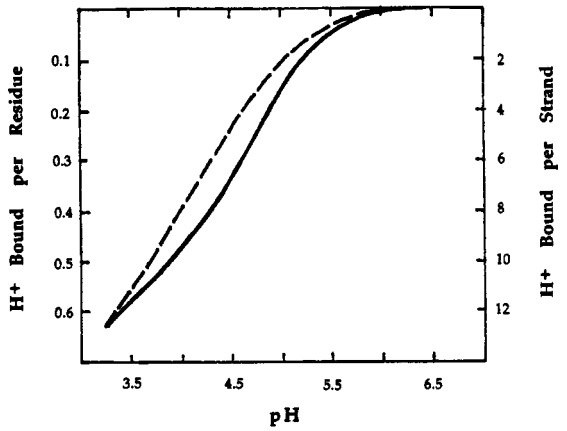

FIGURE 2: Titration curve of $d(A-G)_{10}$ at 23 °C (—), and for comparison a titration curve calculated for an equivalent mixture of noninteracting dA and dG residues (- - -).

not follow that simply expected from the pK_a values for the individual monomers. Rather, the uptake is more abrupt and sigmoidal, reflecting an upward $_{app}pK_a$ shift of 0.35 unit and the modest cooperativity associated with formation of a rather short helical structure with a limited number of linked interactions.

The titration curve in Figure 2 shows that 3 protons are bound per oligomer chain by pH 5, 6–7 by pH 4.5, and 9.2 by pH 4. As shown previously by UV difference spectral analysis (Dolinnaya & Fresco, 1992), the bulk of this binding is by dA residues. Moreover, since dG residues are more acidic (pK_a for the monomer being 2.9), their protonation barely begins below pH 4.5. In this connection, it should be noted that protonation of the dG residues leads to disruption of the helical structure formed maximally at pH 4.0, and therefore it must be somewhat suppressed. Near pH 3.3, where more than 12 protons are bound per oligomer chain, the calculated and experimental titration curves meet, presumably because all dA residues have been protonated, and protonation of dG residues follows essentially as if they were monomers.

After the fraction due to protonation of the dG residues (calculated using the monomer pK_as) from the total proton uptake at pH 4.0 is subtracted, it becomes apparent that >8.5 protons are bound by dA residues at 23 °C. When the electrometric titration curve is corrected for dG proton uptake, it correlates extremely well with the titration curve monitored by CD at room temperature under similar conditions (Figure 3). This coincidence is important because it shows that protonation is essential to the incorporation of each dA residue into helix. In addition, it permits the CD-monitored titration curve at 3 °C for the protonation of dA residues only, to be used in place of an electrometric one at that temperature to estimate the number of protons bound to the dA residues of $d(A-G)_{10}$ at any pH, using as the limits 0 protons bound at pH 6.0 and 9 protons bound at pH 4.0.[1] In fact, the latter is the theoretical maximum number of protons that can contribute to helix stability in the intramolecular structure of $d(A^+-G)_{10}$ if the ionic interactions between dA residues and backbone phosphates involve the $n+1$ or $n+2$ distal phosphates rather than the nearest-neighbor phosphates (see Discussion

[1] The substitution of the CD-monitored curve for a direct titration curve is valid here because only dA residues are being protonated. The chromophore of A is itself essentially insensitive to protonation, as evident from both the UV and the CD spectra of 5′-dAMP in both neutral and protonated forms. Hence, the CD changes accompanying protonation must be the consequence of placing the protonated dA residues in the helical conformation.

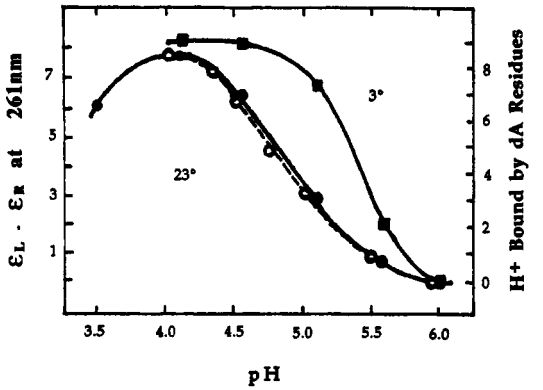

FIGURE 3: pH titration curves for $d(A-G)_{10}$. CD-monitored titrations were performed in 0.001 M sodium acetate and 0.009 M NaCl at 25 (●) and 3 °C (■). The CD curve at 3 °C was normalized to the axis showing proton uptake by dA residues (see text). The electrometric titration curve at 23 °C was performed in 0.01 M NaCl and corrected for the fraction of proton uptake due to dG residues to show the net uptake of protons by dA residues (○).

and Figure 14). Consequently, neither the first phosphate at the 5′ end nor the dA residue nearest the 3′ end would have partners with which to interact. This would leave one dA residue per oligomer to titrate below pH 4.0, with a pK_a similar to that of a monomer. As the normalized curve at 3 °C shows (Figure 3), the dA residues of $d(A-G)_{10}$ bind 3.2 protons at pH 5.5, 7.8 protons at pH 5.0, and 9 protons at pH 4.5 and 4.0.

Using the expression for the effective charge, Z, of a polyelectrolyte (Tanford, 1961),

$$Z^2 = 4Bm_sM^2$$

where B is the second virial coefficient, m_s is the molal concentration of the monovalent salt, and M is the molecular weight, one can use the values for B determined from the equilibrium sedimentation runs to estimate the effective charge on the molecule. At 4 and 40 °C, the calculation yields charges of 4.6 ± 0.4 and 7.2 ± 1 per strand, respectively. Of course, the effective charge and presumably the increase in effective charge are both expected to be less than the titratable charge and the change in the titratable charge. Nevertheless, the increase in effective charge at the higher temperature is consistent with the disruption of intramolecular ionic bonds that stabilize the helix at low temperature.

Cooperativity of Proton Uptake. When compared with the titration curve for poly(A) (Fresco & Klemperer 1959), which forms a parallel-stranded duplex with protonated ($A^+ \cdot A^+$) base pairs, the oligomer titration curve, though clearly cooperative, is relatively broad even at 3 °C. There are two reasons for the much greater cooperativity of proton uptake by poly(A). One is the markedly greater length of the chains undergoing the conformational transition. The other relates to the fundamental difference between the two types of helical structure. Whereas poly($A^+ \cdot A^+$) contains stacked base pairs whose π electron interactions contribute substantially to helix stability, the $d(A^+-G)_{10}$ helix is without the benefit of much base stacking, depending instead primarily on intramolecular ionic bonds that reduce the electrostatic repulsion of adjacent phosphates along the chain and entropic contributions stemming from the liberation of sodium ions that formerly shielded those phosphates and from the relatively unconstrained dispositions of the unstacked dG residues. These differences in cooperativity between $d(A^+-G)_{10}$ and poly($A^+ \cdot A^+$) are manifest as differences in the slopes of Hill plots for their pH-induced conformational transitions (Figure 4), viz.,

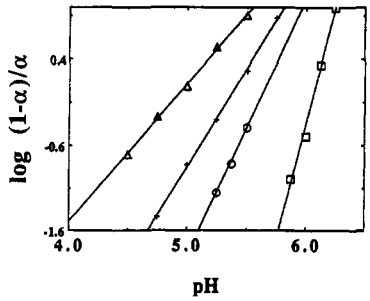

FIGURE 4: Hill plots: log $[(1 - \alpha/\alpha]$, where α is the fraction of proton binding sites that are occupied, vs pH for the coil→helix transitions of $d(A\text{-}G)_{10}$→$d(A^+\text{-}G)_{10}$ in 0.001 M sodium acetate + 0.009 M NaCl at 3 (△) and 25 (+) °C, and for poly(A)→poly(A^+·A^+) at 25 °C in 0.15 M NaCl (○) and in H_2O (□).

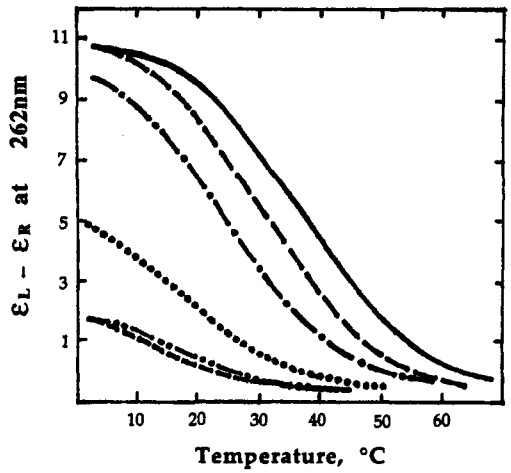

FIGURE 5: CD-monitored melting profiles for $d(A^+\text{-}G)_{10}$ in 0.001 M sodium acetate and 0.009 M NaCl at different pH values: 4.1 (—), 4.64 (– –), 5.1 (- • -), 5.57 (•••), 6.0 (••••), and 7.0 (- - -).

poly(A^+·A^+) shows the greatest slope in the absence of NaCl (5.4), but even in 0.15 M Na^+, it shows a much greater slope (3.0) than for $d(A^+\text{-}G)_{10}$ at 3 °C (2.3), which is greater still than at 25 °C (1.6). Thus, the cooperativity of proton uptake in these cases is directly related to the conformational transitions which the uptake makes possible.

Cooperativity in proton uptake is paralleled by evidence for cooperativity in CD-monitored melting profiles as a function of pH [Figure 5; Figure 2 in Dolinnaya and Fresco (1992)]. Thus, while the thermal stability of a high-ellipticity form markedly increases as the pH is lowered, the melting profile at pH 5.5 still shows a distinctly S-shaped character, even though the ellipticity at 3 °C is less than half the maximum possible. At 3 °C, an average of 3 protons per chain are bound (Figure 3). Were the uptake noncooperative or anticooperative, then all the chains would be one-third protonated. Since proton uptake is actually slightly cooperative, this value must reflect a situation wherein some oligomers are more than one third protonated and others less. In that case, the 3 protons taken up at pH 5.5 (Figure 3) probably represent the minimum number of Coulombic interactions required to nucleate the intramolecular helix. At the same time, it is apparent that additional protonation at lower pH helps to stabilize and extend the helical structure within the oligomer chain.

Dependence of Apparent pK_a on Ionic Strength. The $_{app}pK_a$ for the transition of $d(A\text{-}G)_{10}$ to the acid helix at 3 °C is raised 0.27 unit as the ionic strength is decreased from 0.01 to 0.001 M Na^+ (Figure 6). This is likely due to two effects, one being the lesser shielding of the negatively charged

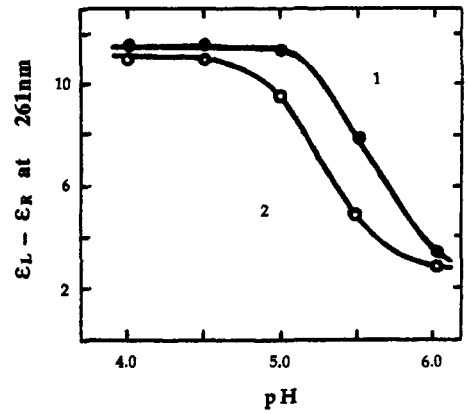

FIGURE 6: CD-monitored pH titration curves for $d(A\text{-}G)_{10}$ at 3 °C in 0.001 M sodium acetate (1) and in 0.001 M sodium acetate and 0.009 M NaCl (2). $c = 0.8 \times 10^{-4}$ M (per residue).

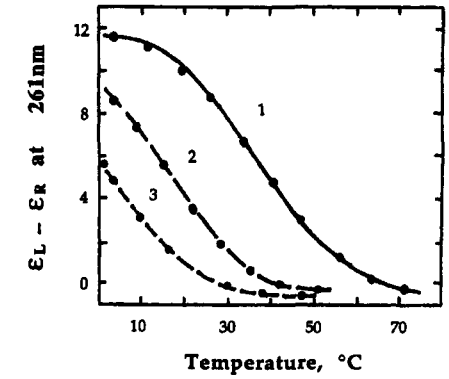

FIGURE 7: CD-monitored melting profiles for $d(A^+G)_{10}$ in 0.001 M sodium acetate, pH 4.0 (1), in 0.001 M sodium acetate, pH 5.2 (2), and in 0.001 M sodium acetate and 0.009 M NaCl, pH 5.2 (3). $c = 0.8 \times 10^{-4}$ M (per residue).

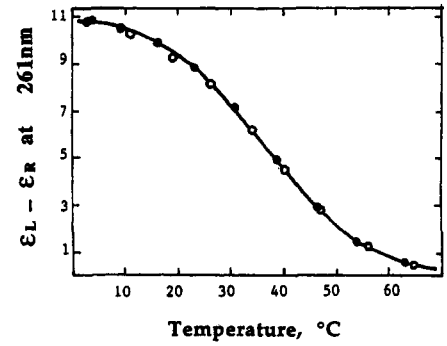

FIGURE 8: CD-monitored melting profiles of $d(A^+\text{-}G)_{10}$ in 0.001 M sodium acetate, pH 4.0 (●), and in 0.001 M sodium acetate and 0.009 M NaCl, pH 4.0 (○).

phosphate groups in the backbone, which makes for their stronger interaction with the positively charged dA residues, and the other being a rise in the pK_a of the acidic sites on the dA residues. It is not surprising, therefore, that increasing ionic strength significantly reduces the T_m of the self-structure formed by $d(A\text{-}G)_{10}$ in the region of pH close to the $_{app}pK_a$ for the transition. As can be seen from Figure 7, a 10-fold increase in Na^+ concentration near the $_{app}pK_a$ reduces T_m from 14.5 to ~0 °C. This difference in T_m correlates with the large difference in degree of protonation of dA residues in this pH region under these two ionic conditions. In contrast, at the pH of maximum stability of the acid structure (pH 4.0), the degree of protonation is not visibly affected by a 10-fold difference in Na^+ concentration (Figures 6 and 8). This must be the case because so high a proportion of protonated dA residues cannot be significantly altered by a

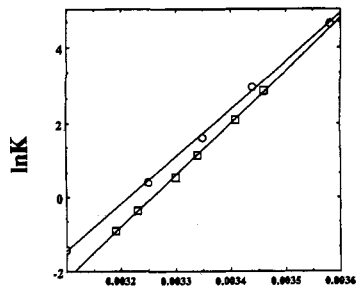

FIGURE 9: Plots of ln K vs $1/T$ obtained from CD-monitored thermal melting of d(A+-G)10 in 0.01 M sodium acetate, pH 4.0 (○), and from integration of the ΔC_p vs T plot in Figure 10 (□).

10-fold change in cation concentration.

In fact, at a pH where the structure is maximally stable, the absence of *positive* ionic strength dependence of T_m is especially informative. Were the structure multistranded in format, i.e., either a true double helix or a "hairpin" helix, the negatively charged phosphates that would have to be brought in proximity would lead to a positive dependence of stability on ionic strength (Elson et al., 1970), and this is clearly not observed.

Thermodynamics of d(A+-G)10 Formation. Given that the secondary structure of d(A+-G)10 is intramolecular in origin, that the melting process is reversible and monophasic, and that there are isoelliptic and isosbestic points in the CD and UV spectra between 3 and 50 °C(Dolinnaya & Fresco, 1992), an attempt was made to evaluate the energetics of the coil→helix transition,

$$\text{single-stranded coil} \xrightarrow{K} \text{single-stranded helix}$$

in terms of a two-state model, assuming that the cooperative length is the entire oligomer strand. Thus

$$K = \frac{C_o\alpha}{C_o(1-\alpha)} = \frac{\alpha}{1-\alpha} \quad (1)$$

where α is the fraction of strands that are helical and $(1-\alpha)$ is the fraction that are coiled. Therefore,

$$\alpha = \frac{(\Delta\epsilon(T) - \Delta\epsilon_c)}{(\Delta\epsilon_h - \Delta\epsilon_c)} \quad (2)$$

where $\Delta\epsilon(T)$ is $\epsilon_L - \epsilon_R(f_T)$ obtained from the CD melting profile and $\Delta\epsilon_h$ and $\Delta\epsilon_c$ are the CD parameters for d(A-G)10 in helix and coil forms, respectively. Combining eqs 1 and 2 gives

$$K = \frac{\Delta\epsilon(T) - \Delta\epsilon_c}{\Delta\epsilon_h - \Delta\epsilon(T)} \quad (3)$$

Since

$$-\ln K = \frac{\Delta H^\circ}{RT} - \frac{\Delta S^\circ}{R} \quad (4)$$

a plot of ln K vs $1/T$ should give a straight line. This result is seen in Figure 9, from whose slope and intercept are calculated values for the *formation* of d(A+-G)10 of ΔH = −22.8 ± 2.1 kcal/mol of oligomer, ΔS = −73.7 ± 6.2 eu/mol, and ΔG_{298} = −0.79 ± 0.1 kcal/mol.

As an alternative approach, the enthalpy change for the disruption of the acid helix was directly measured using differential scanning calorimetry (DSC). Figure 10 shows a DSC plot for the *melting* of d(A+-G)10 in 0.01 M sodium acetate, pH 4.0, from which one obtains values of ΔH = 11.7 kcal/mol, ΔS = 37.3 eu/mol, and ΔG_{298} = 0.596 kcal/mol of oligomer. For comparison, these values take on opposite sign, making it apparent that they are 50–75% of the van't Hoff values.

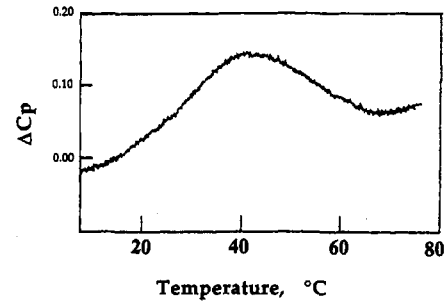

FIGURE 10: Excess heat capacity, ΔC_p, in kcal mol^{-1} deg^{-1}, for d(A+-G)10 in 0.01 M sodium acetate, pH 4.0, measured by differential scanning calorimetry.

This discrepancy may be understood in terms of the different observables in spectroscopic melting experiments and in differential scanning calorimetry. Both the van't Hoff enthalpy and the DSC ΔH value refer to the helix→coil transition over the same temperature range. However, whereas DSC provides a direct enthalpy measurement, the van't Hoff analysis depends on the physical properties to which CD is sensitive and on the assumption of a simple two-state process, i.e., all-helix strand→all-coil strand. In fact, the melting process must be much more complicated, viz., the broad transition, so that there are significantly populated intermediate states.

In this regard, K is related to temperature, T, for similar one-dimensional phase transitions such as are undergone by long-chain α-helical polypeptides, by

$$d(\ln K)/d(1/T) = -\Delta H/R\sigma^{1/2}$$

where σ is the nucleation parameter (Applequist & Doty, 1962). Hence, the sharpness of the transition is proportional to $\Delta H/\sigma^{1/2}$. For the thermally induced α-helix→coil transition of poly(benzyl glutamate), which proceeds with intermingled coil and helix regions along the chain, $\sigma = 2 \times 10^{-4}$ (Zimm et al., 1959). Notwithstanding that d(A+-G)10 is a relatively short chain, the value of the calorimetrically determined ΔH, 11.7 kcal/mol ÷ 20 residues = 0.59 kcal/mol of residue, would require that $\sigma = 6.7 \times 10^{-4}$ to agree with the van't Hoff ΔH. This value is remarkably close to the σ value for the α-helix of poly(benzyl glutamate). Such similarity of σ values for the two structurally related single-stranded helices suggests that the thermally induced coil→helix transition for d(A+-G)10 at pH 4.0 is cooperative, just as is the uptake of protons by d(A-G)10 when it undergoes the transition isothermally, as was deduced from the Hill plot slopes. Moreover, the σ value would suggest that intermediate states occur over the course of the thermal transition, just as they do for the pH-dependent isothermal transition.

In addition, as previously shown (Dolinnaya & Fresco, 1992), helix melting is accompanied by the stacking of dG residues with their nearest-neighbor dA residues, those dG residues being unstacked in the initial d(A+-G)10 helix. This stacking could contribute less to the signal in the CD-monitored melting process, as CD is generally more sensitive to the major loss of the chirality of the intact helix. However, such stacking must contribute to the enthalpy change measured by DSC as an exothermic process, diminishing the overall endothermic change.

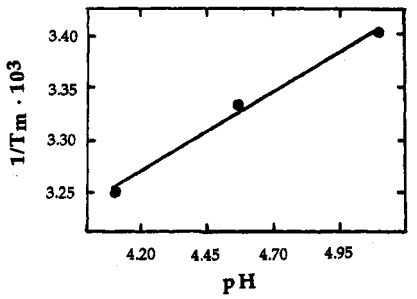

FIGURE 11: Dependence of T_m of d(A⁺-G)$_{10}$ on pH in 0.001 M sodium acetate and 0.009 M NaCl.

This circumstance was investigated by a comparison of the slopes of the plots in Figure 9 of ln K vs $1/T$ calculated from the CD melting profile (0.114) and of H obtained by integration of the ΔC_p curve vs T (0.119). The slopes of the two lines are nearly but not exactly the same, suggesting that nearly the same relative amounts of helix and coil are measured by the CD and DSC methods. While this is true, the discrepancy between the van't Hoff and DSC values may nevertheless be related to the base stacking which occurs in the course of single-strand helix melting (Dolinnaya & Fresco, 1992).

Since there are nine protonated dA residues at pH 4.0 that appear to form nine ionic bonds, the enthalpic cost of forming one ionic bond is −2.5 kcal/mol, assuming the van't Hoff ΔH, and −1.3 kcal/mol, assuming the DSC ΔH, the corresponding ΔG_{298} values being 0.088 and 0.066 kcal/mol, respectively.

The difference between the protons bound by coil and helix forms of d(A-G)$_{10}$ can be estimated from the dependence of T_m of the acid-induced structure on pH (between 4.1 and 5.1), thus (Record et al., 1978; Xodo et al., 1991):

$$dT_m/d(\log [H^+]) = -2.303\Delta n_{H^+} RT_m^2/\Delta H \qquad (5)$$

A plot of $1/T_m$ vs pH gives a straight line (Figure 11), whose slope, $-2.303R(\Delta n_{H^+})/\Delta H$) allows estimation of Δn_{H^+}, provided that the enthalpy change for the reaction is known. Using the van't Hoff enthalpy, one calculates that Δn_{H^+} is −0.74, whereas the DSC enthalpy gives $\Delta n_{H^+} = -0.4$. Assuming the pK_a values for dAMP and dGMP to represent those for the residues in the absence of formation of secondary structure, one calculates that ~7.9 protons would then be bound at pH 4.0 to d(A-G)$_{10}$,whereas the electrometric titration curve in Figure 1 shows that 9.2 protons are bound by d(A⁺-G)$_{10}$. Thus, values of Δn_{H^+} near 1 are reasonable.

Length Dependence of Stability of d(A⁺-G)$_n$. It was shown above that formation of three ionic bonds is required to nucleate the acid helix. Since, as noted, the dA residue nearest the 3′ terminus has no backbone partner with which to form an ionic bond, this would set the minimum length of an d(A-G)$_n$ oligomer capable of nucleating such a helix at d(A-G)$_4$ or d(A-G)$_5$. We therefore chose d(A-G)$_6$ for evaluating the length dependence, since this would allow one or two additional ionic bonds, i.e., a total of 4 or 5.

Whereas T_m of d(A⁺-G)$_{10}$ at pH 4 is 37 °C, it is only 10–12 °C for d(A⁺-G)$_6$ under identical conditions (Figure 12). Moreover, the $_{app}$pK_a for the transition to the acid structure at 0 °C is shifted from 5.3 for d(A-G)$_{10}$ to 4.5 for d(A-G)$_6$, which is much closer to the pK_a of 4 for dAMP. On the other hand, the $_{app}$pK_a for d(A-G)$_6$, like that for d(A-G)$_{10}$ (see Figure 6, is shifted up from 4.5 in 0.01 M Na⁺ to 4.8 in 0.001 M Na⁺ (not shown). Moreover, the CD spectrum at 0 °C for d(A⁺-G)$_6$ at pH 3.2, where it is maximally formed, is qualitatively identical to that of d(A⁺-G)$_{10}$ at that pH, and as evident from Figure 13, its intensity closely approaches that of d(A⁺-G)$_{10}$.

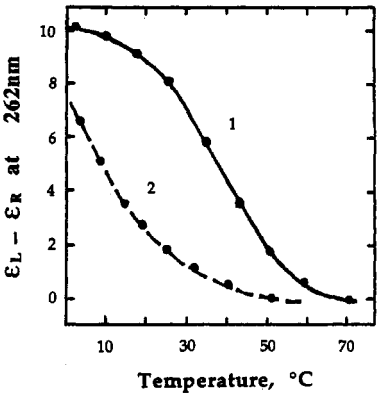

FIGURE 12: CD melting profiles for d(A⁺-G)$_{10}$ (1) and d(A⁺-G)$_6$ (2) in 0.01 M sodium acetate, pH 4.0. $c = 0.8 \times 10^{-4}$ M (per residue).

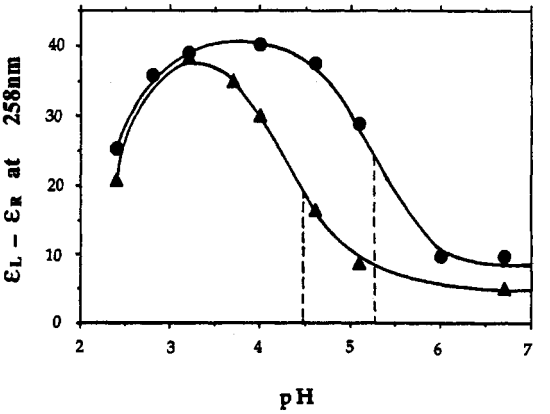

FIGURE 13: CD-monitored pH titration curves for d(A-G)$_{10}$ (●) and for d(A-G)$_6$ (▲) in 0.01 M sodium acetate at 0 °C. The dashed lines extrapolate to the $_{app}$pK_a values. $c = 1.13 \times 10^{-4}$ M (per residue).

Hence, their structures must be homologous, i.e., they must have the same general structure.

These results have important implications, particularly since they show that the ellipticity per residue is quite similar in the two fully formed homologs. If their helicity arose in both cases from hairpin helices, this could not be the case, since the ratio of unpaired residues in the turns of the two oligomers to total residues would necessarily be very different, i.e., 5 unpaired out of 12 vs 5 unpaired out of 20. This consideration is consistent with other arguments (Dolinnaya & Fresco, 1992) that rule out this type of intramolecular double-helical structure.

DISCUSSION

Molecular Properties. The data presented in this report allows us to substantially increase our confidence in the type of single-stranded secondary structure that we proposed earlier for d(A⁺-G)$_{10}$ in acidic solution (Dolinnaya & Fresco, 1992). Three new types of observations are particularly compelling in this regard. Whereas we had earlier deduced the single-strandedness of the helical form from the lack of dependence of T_m on oligomer concentration, we have now directly determined that the molecular weight of the helix is that of the oligomer itself and not of some aggregate, even over an enormous concentration range.

In addition, whereas we had earlier provided UV spectral data to show that the helical structure depends on protonation of the dA residues, we have now shown that nine protons are taken up cooperatively and that their uptake occurs in a pH range where only dA residues are likely to be protonated.

FIGURE 14: 5′→3′ sequence of $d(A^+-G)_{10}$ showing the possibility of forming up to nine bridges between dA^+ and the $n+1$ (- - -) or $n+2$ (—) phosphates. Only the possibilities from the 5′→3′ end are illustrated.

The observation that nine protons are taken up is structurally informative, for as may be seen in Figure 14, a maximum of nine of the ten dA residues in $d(A^+-G)_{10}$ can form ionic bonds with distal phosphates, and this only if the interactions are with the $n+1$ *or* the $n+2$ phosphates, counting from the 5′ end of the oligomer chain. Whereas nine ionic bonds are also possible (starting from the 3′-terminal dA residue) with their $n-1$ phosphates, only eight are possible if the helical twist brings these positively charged dA residues in proximity to the $n-2$ phosphates (not shown). Moreover, if the helical twist is such that ionic interaction is with the $n\pm3$ phosphates (counting from either end), then the number of possible ionic bonds is also reduced (not shown). Of course, while ionic bonds are energetically stabilizing, they are not too distance dependent or directionally sensitive and, hence, may by themselves not provide sufficient structural definition for the helix. In that case, additional helix rigidity can conceivably be mediated by the exocyclic amino hydrogens of the dA residues, with their potential for forming hydrogen bonds of relatively fixed distance and orientation toward the P=O moieties of the backbone. In this connection, we are currently completing a variety of spectroscopic, nuclease sensitivity, and molecular model building analyses of $d(A^+-G)_{10}$ that should provide much more structural detail.

The third critical observation in the present work has come from an analysis of the CD properties of $d(A^+-G)_{10}$ and $d(A^+-G)_6$ at the pH values where they are maximally formed. From their comparison, it becomes evident that these two oligomers form structurally homologous helices, which would not be the case if the two structures contained different *proportions* of dA residues that are not part of the helically twisted backbone, as would be expected if a hairpin-like structure were involved.

We have here also presented additional evidence for the importance of the ionic interactions to the maintenance of the helical structure by showing that increasing ionic strength destabilizes it. Finally, we have evaluated the thermodynamic parameters for the helical structure and shown their consistency with the Coulombic stabilization of the helix.

It is of interest to consider whether the type of single-stranded secondary structure found for $d(A^+-G)_n$ has wider applicability to homopurine sequences with more varied or randomized distributions of A and G residues. This type of structure clearly requires some minimal level of dA residues to provide sources of positive charge. Hence, a homo(dG) sequence cannot assume such a secondary structure. Moreover, homo(dA) sequences are likely to give rise to the acid poly(A^+·A^+) double helix unless the sequence is restrained at its ends by flanking sequences that link it to complementary double helices, as would be the case in H-DNA segments. This still allows a wide range of A,G-containing sequences which can meet the basic requirements for forming the type of single-stranded helical secondary structure stabilized by A^+-mediated ionic bonds with distal phosphates. In this connection, we see no reason why this type of helical structure could not be assumed by comparable RNA strand sequences just as we have shown that they are by DNA strand sequences.

Biological Relevance. The sequences $d(A-G-G-A-G)_{28}$ and $d(A-G)_{21,27}$, as well as neighboring randomized A,G sequences, occur in the murine immunoglobulin switch region (Reaban & Griffin, 1990), in the origin of replication of the Chinese hamster *dhfr* gene (Brinton et al., 1991), and in some synthetic replication templates (Baran et al., 1991; Rao et al., 1988; Baran et al., 1987). There, they are involved respectively in the regulation of transcription (Reaban & Griffin 1990) and in the arrest or slowing of DNA replication (Brinton et al., 1991; Baran et al., 1987, 1991; Rao et al., 1988) through the formation of non-B-DNA structures that are stabilized by superhelical stress and acidic pH. Such structures include the acid-induced intramolecular H-DNA triplex which consists of T:AT and C^+:GC base triplets and a corresponding free homopurine strand. The acid-induced coil→helix transition observed here for $d(A-G)_{10}$ seems an ideal assignment for the free purine strand segment in H-DNA, which till now has been considered to have a nondescript random coil-type structure. As we have previously noted (Dolinnaya & Fresco, 1992), such a structural transition accompanying intramolecular triplex formation could conceivably provide stabilizing energy and also relieve superhelical stress. That the ${}_{app}pK_a$ for the $d(A-G)_n$ single-strand coil→helix transition would approach neutrality *in vivo* seems quite feasible for genomic H-DNA-forming regions, especially when Figure 13 is considered, which shows that ${}_{app}pK_a$ for the coil→helix transition rises with increasing length of the oligomer. The constraints imposed by flanking sequences would prohibit formation of competing multistranded homopurine secondary structures such as those observed *in vitro* at neutrality with moderate Na^+ and Mg^{2+} concentrations. Another driving force for the coil→helix transition could come from proteins which recognize and upon binding raise the ${}_{app}pK_a$ for the transition.

Recently, the cDNA sequence from the human *pur* gene was reported (Bergemann et al., 1992), as well as the sequence of the DNA target for the specific single-stranded DNA binding protein, the Pur factor, which the gene encodes. The Pur factor found in HeLa cell nuclei binds the purine-rich single strand of highly conserved sequences found upstream of the *c-myc* gene, near the center of a region implicated in initiation of chromosomal replication. Other regions containing the consensus d(G-A-G-G-G-A-G-A) sequence are found upstream of *int-2*, *lck*, and the genes for human β-globin and histone H-4 (Bergemann et al., 1992). The sequence deduced from the cDNA of the Pur α protein contains several features common to nucleic acid binding proteins, including glycine-rich regions, leucine and isoleucine repeats, PEST repeats, and several conserved basic amino acids. The most striking feature of the protein sequence is the presence of numerous aromatic and acidic amino acid residues which are regularly spaced and conserved. On the basis of the sequence, it has been suggested (Bergemann et al., 1992) that the aromatic residues could function in binding exposed purine residues in a single strand. In the case of the homopurine secondary structure reported here, the exposed dG residues seem eminently poised for such interaction. In addition, the acidic amino acid residues which are abundant and conserved especially in one segment of the protein seem exceptionally well disposed for the stabilization of an acid-induced coil→helix transition. This segment (class II repeat) shares RDYL, DD, and EE motifs in common with the acidic domain of the GAL4

protein (Mitchell & Tjian, 1989) and in addition contains several more acidic residues not present in the GAL4 acidic domain. In this connection, it is noteworthy that the Pur factor does not bind oligo[d(A)] or -[d(G)] strands and that the minimal sequence for binding is d(G-G-A-G-G-A). The fact that oligo[d(A)] and -[d(G)] cannot undergo the single-stranded acid-induced coil→helix transition that we have described here for $d(A\text{-}G)_{10}$ and that they are not bound by the Pur protein suggests that the Pur factor recognizes some sort of single-stranded helical secondary structure rather than a randomly coiled one. But it is not yet known whether or not $d(A\text{-}G)_{10}$ or other random homopurine sequences bind well. Given that the Pur factor sequence does not resemble any known DNA binding motif yet binds to single-stranded purine sequences that undergo a single-stranded secondary structural transition, it is tempting to suggest that this factor represents a novel protein motif that recognizes the nucleic acid secondary structure described here.

ACKNOWLEDGMENT

We are grateful to Jon Applequist for helpful suggestions and a critical reading of the manuscript.

REFERENCES

Anservin, A. T., Roark, D. E., & Yphantis, D. A. (1970) *Anal. Biochem. 34*, 237–261.

Applequist, J., & Doty, P. (1962) in *Polyamino Acids, Polypeptides, and Proteins*, (Stahmann, M. A., Ed.) pp 161–177, University of Wisconsin Press, Madison, WI.

Baran, N., Lapidot, A., & Manor, H. (1987) *Mol. Cell. Biol. 7*, 2636–2640.

Baran, N., Lapidot, A., & Manor, H. (1991) *Proc. Natl. Acad. Sci. U.S.A., 88*, 507–511.

Bergemann, A. D., & Johnson, E. M. (1992) *Mol. Cell. Biol. 12*, 1257–1265.

Bergemann, A. D., Ma, Z., & Johnson, E. M. (1992) *Mol. Cell. Biol. 12*, 5673–5681.

Brinton, B. T., Caddle, M. S., & Heintz, N. H. (1991) *J. Biol. Chem. 266*, 5153–5161.

Dolinnaya, N. G., & Fresco, J. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9242–9246.

Eisenberg, H. (1990) in *Landolt–Bornstein, Numerical Data and Functional Relationships in Science and Technology, New Series, Group VII Biophysics, Vol. 1, Nucleic Acids, Subvolume c*, pp 257–272, Springer-Verlag, New York and Berlin.

Elson, E. L., Scheffler, I. E., & Baldwin, R. L. (1970) *J. Mol. Biol. 54*, 401–415.

Fossella, J., Kim, Y. J., Shih, H. T., Richards, E. G., & Fresco, J. R. (1993) *Nucleic Acids Res.* (in press).

Fresco, J. R., & Klemperer, E. (1959) *Ann. N.Y. Acad. Sci. 81*, 730–741.

Johnson, M. L., Correla, J. J., Halvorson, H. R., & Yphantis, D. A. (1981) *Biophys. J. 36*, 575–588.

Lindahl, T., Henley, D. D., & Fresco, J. R. (1965) *J. Am. Chem. Soc. 87*, 4961–4963.

Mitchell, P. J., & Tjian, R. (1989) *Science 245*, 371–378.

Rao, B. S., Manor, H., & Martin, R. G. (1988) *Nucleic Acids Res. 16*, 8077–8094.

Reaban, M. E., & Griffin, J. A. (1990) *Nature 348*, 342–344.

Record, M. T., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103–134.

Tanford, C. (1961) *Physical Chemistry of Macromolecules*, p 229, John Wiley and Sons, Inc., New York.

Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G.A., & van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625–5631.

Zimm, B. H., Doty, P., & Iso, K. (1959) *Proc. Natl. Acad. Sci. U.S.A. 45*, 1601–1607.