Volume 8 Number 18 1980 Nucleic Acids Research

Polypurine DNAs and RNAs form secondary structures which may be tetra-stranded

Jeremy S.Lee, David H.Evans and A.Richard Morgan

Department of Biochemistry, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada

Received 9 May 1980

ABSTRACT

Polypurine DNAs and RNAs containing at least 33% guanine form a stable secondary structure at neutral pH and moderate ionic strengths. The t_m's of the polymers increase with increasing guanine content. To eliminate possible structures three novel polymers, $d(Gn^2A)_n$, $d(Gm^6A)_n$ and $d(IA)_n$*, as well as the random copolymer $r(G.A)_n$ were studied. Both $d(Gn^2A)_n$ and $d(IA)_n$ can form a secondary structure whereas $d(Gm^6A)_n$ and $r(G,A)_n$ cannot. Model building suggested two possible structures, one a duplex and the other a tetra-stranded polymer. The latter is considered to be the more likely, since previous X-ray diffraction studies have shown that rG_n and rI_n are tetra-stranded. Circular dichroism spectra are also consistent with such an interpretation.

INTRODUCTION

During our studies on the dismutation of $(pyrimidine)_n \cdot (purine)_n$ DNAs at low pH (1,2), we consistently observed a small hyperchromic change in the melting profile at t_m's nearly 60° below that of the triplex (see Figure 3a). Initially this small change was thought to be due to duplex fragments which had not been completely converted to triplexes. However, further investigation showed that the transition was due solely to the polypurine which was released during the dismutation:

$$2\ d(TC)_n \cdot d(GA)_n \rightarrow d(TC)_n \cdot d(GA)_n \cdot d(\overset{+}{C}T)_n + d(GA)_n$$

Consequently, it seemed probable that polypurine DNAs were forming some sort of self-structure. The isolation of polypurine DNAs and the synthesis of analogous RNA polymers has enabled us to study this phenomenon in more detail. The synthesis and properties of three novel polymers, $d(TC)_n \cdot d(IA)_n$, $d(TC)_n \cdot d(Gm^6A)_n$ and $d(TC)_n \cdot d(Gn^2A)_n$* will be described together with a new method for the production of polypurine DNAs using exonuclease III digestion.

MATERIALS AND METHODS

Ethidium bromide was a product of Sigma Chemical Co. and formaldehyde

was obtained as a 40% v/v aqueous solution from Fisher Chemicals. Nucleoside triphosphates including dITP were purchased from P-L Biochemicals. In addition dm^6ATP was synthesized essentially according to the procedure of Engel and von Hippel (3) and dn^2ATP was a gift of Dr. H. van de Sande.

RNA polymerase from E. coli was prepared according to the procedure of Chamberlin and Berg (4). The method of Jovin et al. (5) was used for the preparation of DNA polymerase I and exonuclease III except that the phosphocellulose chromatography (step 6) was omitted. The exonuclease III was not significantly contaminated by endonuclease activity as judged from its inability to convert ccc PM2 to oc PM2 DNA (6).

DNAs: Double-stranded repeating sequence DNAs were routinely prepared as described in detail by Morgan et al. (7). Three analogues of $d(TC)_n \cdot d(GA)_n$, that is $d(TC)_n \cdot d(IA)_n$, $d(TC)_n \cdot d(Gm^6A)_n$ and $d(TC)_n \cdot d(Gn^2A)_n$ were synthesized using $d(TC)_n \cdot d(GA)_n$ as template (at 0.1 A_{260} except for $d(TC)_n \cdot d(IA)_n$ where the template was at 0.5 A_{260}) in the presence of the appropriate purine nucleoside triphosphates. All incubations were performed at 37° except for $d(TC)_n \cdot d(Gn^2A)_n$ which was at 45°. The net fold syntheses were 8, 50 and 26 for $d(TC)_n \cdot d(IA)_n$, $d(TC)_n \cdot d(Gm^6A)_n$ and $d(TC)_n \cdot d(Gn^2A)_n$ respectively. However since degradation of the template by the exonucleolytic functions of DNA polymerase I always accompanies synthesis, then, even in the case of $d(TC)_n \cdot d(IA)_n$, the polymers will not be contaminated to any significant extent (<5%) by $d(TC)_n \cdot d(GA)_n$. This was born out by the sharpness of the thermal transitions (see below) where the majority of the hyperchromic change occurred within 2°. The synthetic reactions were routinely followed by using the ethidium fluorescence assay (6) (where the synthesis of $d(TC)_n \cdot d(IA)_n$ has already been described). Since after a heating and cooling step, only $d(AT)_n$ gives rise to fluorescence, then the reaction can be stopped before contamination by this polymer becomes a problem. By this criteria the samples of $d(TC)_n \cdot d(IA)_n$ and $d(TC)_n \cdot d(Gn^2A)_n$ used in this work contained $d(AT)_n$ to the extent of 2 and 3% respectively. The presence of $d(AT)_n$ could not be detected in any of the other polymers. t_m's and buoyant densities for the analogues of $d(TC)_n \cdot d(GA)_n$ are presented in Table I. As expected $d(TC)_n \cdot d(IA)_n$ has a lower t_m and $d(TC)_n \cdot d(Gn^2A)_n$ has a higher t_m than that of $d(TC)_n \cdot d(GA)_n$. The work of Engel and Von Hippel (3) shows that the presence of 6-methyl adenine destabilizes duplexes and $d(TC)_n \cdot d(Gm^6A)_n$ also demonstrates this trend. $d(TC)_n$, $d(TTC)_n$ and $d(TCC)_n$ were prepared by depurination of the appropriate duplexes according to the procedure of Harwood and Wells (8).

TABLE I.

DNA	t_m (°C)	Buoyant Density in Cs_2SO_4 (g/cm^3)
$d(TC)_n \cdot d(GA)_n$	65.5	1.430
$d(TC)_n \cdot d(IA)_n$	41.0	1.446
$d(TC)_n \cdot d(Gm^6A)_n$	53.5	1.410
$d(TC)_n \cdot d(Gn^2A)_n$	75.5	1.424

Buoyant densities at pH 8 in Cs_2SO_4 (ref. 2) and t_m's (see methods) measured in 10 mM Tris-HCl pH 8, 5mM NaCl and 0.1 mM EDTA for the analogues of $d(TC)_n \cdot d(GA)_n$.

Preparation of polypurine DNA's:Exonuclease III digestion: The two strands of a duplex DNA can frequently be separated by alkaline caesium chloride density centrifugation as described by Morgan et al. (7). However, if the polymer is not of very high molecular weight (i.e. $2\text{-}5 \times 10^5$ as is the case with the polymers used in this work) then some overlapping of the peaks will occur so that a further step is required to "clean up" the single-stranded polymers. Moreover, some polymers (e.g. $d(TC)_n$ and $d(GA)_n$) have very similar buoyant densities in alkaline caesium chloride so that this method is inappropriate. Consequently, we were fortunate to find that exonuclease III shows considerable specificity for digestion of the pyrimidine strand of many $(pyrimidine)_n \cdot (purine)_n$ DNAs. The production of the purine strand involves digestion followed by gel filtration.

Exonuclease III digestionswere performed in 50 mM KP_i, 10 mM $MgCl_2$ containing 4µg/ml exonuclease III and a duplex DNA at 1-4 A_{260}. The pH of the phosphate buffer and the temperature of incubation are extremely important in determining the extent of the specificity for degradation of the pyrimidine strand. Decreasing the pH and the temperature both cause a reduction in the rate of degradation but a large increase in the specificity. We have found that a pH of 6.4 and incubation at room temperature, in general, gives a good recovery of the purine strand as well as retaining a reasonable rate of degradation. A typical experiment is shown in Figure 1 for the preparation of $d(Gm^6A)_n$. The extent of degradation of the duplex can be followed by the fluorescence assay (6) and it is readily demonstrated by adding back excess $d(TC)_n$ or $d(GA)_n$ (to reform a duplex for fluorescence measurements) that it is the pyrimidine strand which is preferentially lost.

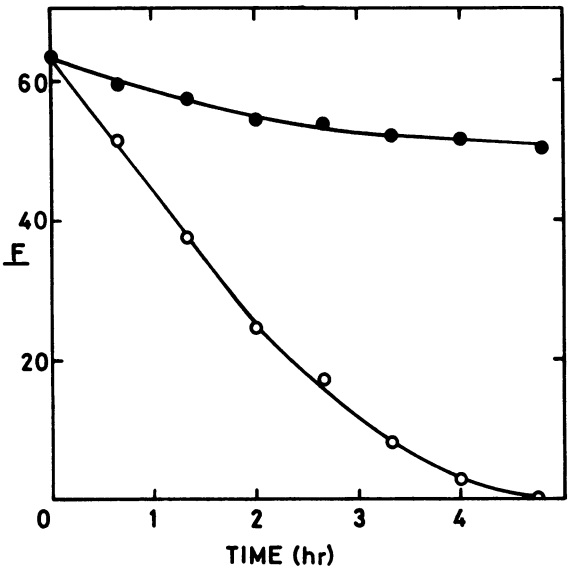

Figure 1. Exonuclease III digestion of $d(TC)_n \cdot d(Gm^6A)_n$ at pH 6.4 and 21° as described in Methods. At various times aliquots were taken from the reaction mixture and incubated in the absence (O) or presence (●) of excess $d(TC)_n$ at 21° in 0.2 M NaCl, 20 mM EDTA (to inactivate the nuclease) for ten minutes before measuring the duplex DNA by the pH 8 ethidium fluorescence assay (ref. 6). Since addition of excess $d(TC)_n$ returns the fluorescence to nearly its original level it can be deduced that it is the pyrimidine strand which is specifically degraded.

After the exonuclease III digestion EDTA was added to 20 mM and the reaction mixture was treated with 0.2 mg/ml of pronase (Calbiochem) at 37° for 2 hours. The $d(Gm^6A)_n$ was then isolated by gel filtration on a 0.5 M, 50-100 mesh agarose column using 1 mM NaOH as eluant to prevent non-specific losses. This method of purification is rapid and gives extremely good recovery of the DNA from the reaction mixture (90%). The DNA was stored frozen in a buffer containing 1 mM NaCl (from neutralization of the NaOH), 10 mM Tris-HCl and 0.1 mM EDTA in test-tubes pretreated with dimethyldichlorosilane, again to prevent non-specific losses.

An identical method has also been used to isolate $d(Gn^2A)_n$, $d(GAA)_n$, $d(GGA)_n$ and $d(GA)_n$ on a preparative scale. Representative results for these and other $(pyrimidine)_n \cdot (purine)_n$ DNAs are shown in Table II. It should be pointed out that previous authors (9,10) have suggested that exonuclease III shows little specificity. The effects of pH and temperature are illustrated

TABLE II

DNA	% RECOVERY	pH	TEMPERATURE (°C)
$d(TC)_n \cdot d(Gm^6A)_n$	81% purine	6.4	21°
$d(TC)_n \cdot d(Gn^2A)_n$	88% purine	6.4	21°
$d(TC)_n \cdot d(GA)_n$	62% purine	6.4	21°
$d(TTC)_n \cdot d(GAA)_n$	60% purine	6.4	21°
$d(TCC)_n \cdot d(GGA)_n$	67% purine	6.4	21°
$d(TCC)_n \cdot d(GGA)_n$	30% purine	7.5	21°
$d(TCC)_n \cdot d(GGA)_n$	0% purine	7.5	37°
$d(TC)_n \cdot d(IA)_n$	50% pyrimidine	7.5	21°
$d(TC)_n \cdot d(IA)_n$	35% purine	6.4	21°

Effects of temperature and pH on the specificity of Exonuclease III digestion of various $d(pyrimidine)_n \cdot d(purine)_n$ polymers. Note that at pH 7.5 and 21° it is the pyrimidine strand of $d(TC)_n \cdot d(IA)_n$ which is preferentially recovered.

with respect to $d(TCC)_n \cdot d(GGA)_n$ and $d(TC)_n \cdot d(IA)_n$. For the latter polymer there is a dramatic changeover in specificity on lowering the pH. Even at pH 6.4 and 21° only 35% of the $d(IA)_n$ remained undigested and thus this is not a good method for preparing this polymer. However purine strands can effectively be "isolated" from $(pyrimidine)_n \cdot (purine)_n$ DNAs by lowering the pH to 5 so that they dismutate to a triplex together with a free purine strand (see Introduction and (2)). Since the triplex has a much higher t_m its presence does not interfere with t_m measurements of the purine self-structure. Thus this technique has been used to study $d(IA)_n$ which was not made preparatively by exonuclease III digestion.

RNAs: Polypurine RNAs were synthesized essentially according to the procedure of Morgan and Wells (11) using only ATP and GTP in the reaction mixture. After the synthesis stopped (about 15 hours with one fold net synthesis), the nucleic acid was isolated by gel filtration on an agarose 0.5 M column. The DNA was then removed by exhaustive digestion with DNAse I in the presence of 0.5 mM $CaCl_2$ and 4 mM $MgCl_2$. After addition of EDTA to 20 mM and incubation with 50μg/ml pronase at 37° for 12 hours the RNA was isolated by gel filtration on an agarose 0.5 M column. In all cases the resulting RNA contained no significant polypyrimidine DNA since it gave rise to no detectable fluorescence in the ethidium bromide fluorescence assay. However, ^3H-

labelling of the DNA purine strand used as template showed that the polypurine RNAs were contaminated between 2 and 5% with the analogous polypurine DNA. Random $r(G,A)_n$ was purchased from P.L. Biochemicals and contained 61% adenine and 39% guanine as base composition. Extinction coefficients for the purine RNAs and DNAs used in this work are shown in Table III.

Techniques: Circular dichroism spectra were recorded on a Cary 60 spectropolarimeter equipped with a model 6001 circular dichroism accessory with the cell thermostated at an appropriate temperature. Solutions were filtered through a Millipore filter or centrifuged on a bench centrifuge before use. Data reduction was according to Wells et al. (12) and spectra were computed manually with a wavelength spacing of 5 nm.

A Gilford model 2400 spectrophotomer equipped with an insulated cell compartment was used for the thermal-denaturation profiles. These were recorded automatically with the temperature being raised by a Haake circulating waterbath at approximately 0.5°/minute. Samples (300-600 µl) in the appropriate buffer were first purged with helium before overlaying with paraffin oil. All hyperchromicities were measured at 260 nm except for dismutated $d(TC)_n \cdot d(IA)_n$ which was at 250 nm.

RESULTS

Thermal-denaturation profiles for various polypurines under a variety of conditions are shown in Figure 2. The effect of ionic strength on the melting of $d(GGA)_n$ is investigated in Figure 2a. As expected, increasing

TABLE III. Molar Extinction Coefficients[a]

Polymer	ε_{260}	Polymer	ε_{260}
$d(GGA)_n$	9,200[b]	$r(GAA)_n$	10,200[d]
$d(GA)_n$	9,500[b]	$r(GA)_n$	8,800[d]
$d(GAA)_n$	7,500[b]	$r(GAA)_n$	7,500[d]
$d(Gm^6A)_n$	9,000[c]		
$d(Gn^2A)_n$	9,000[c]		

a. Extinction coefficients for polymers in 10 mM Tris pH 8, 0.1 mM EDTA at 21°.
b. Measured as described in ref. (2).
c. Assumed values for calculation of molar ellipticities.
d. From ref. (30).

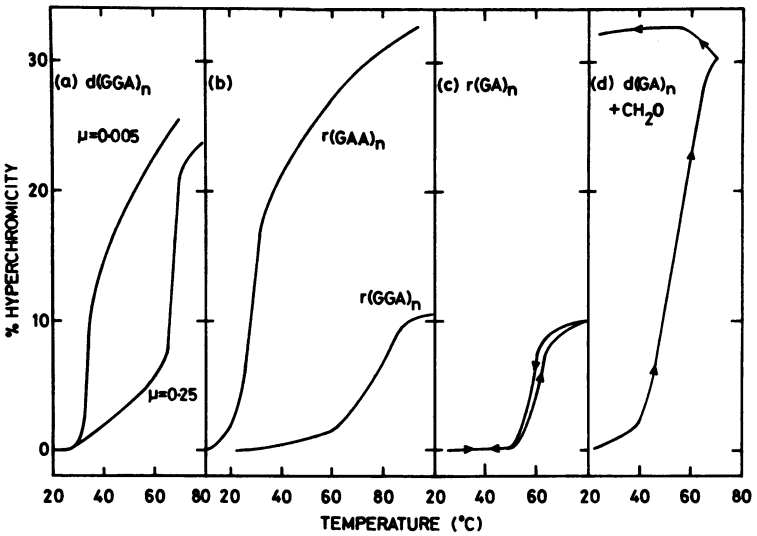

Figure 2. Thermal-denaturation profiles (a) $d(GGA)_n$ in 10 mM Tris-HCl pH 8, 0.1 mM EDTA (μ=0.005) or 10 mM Tris-HCl pH 8, 0.1 mM EDTA 0.25 M NaCl (μ=0.25). (b) $r(GAA)_n$, $r(GGA)_n$ and (c) $r(GA)_n$ in 10 mM Tris-HCl pH 8, 0.1 mM EDTA, 0.25 M NaCl. (d) $d(GA)_n$ in 10 mM Na acetate pH 5, 0.1 mM EDTA, 0.25 M NaCl and 1% formaldehyde (v/v).

the ionic strength increases the t_m but there is also a significant change in the shape of the curves. At low ionic strength the cooperative part of the transition is small and it is followed by a gradual increase in hyperchromicity. At high ionic strength the cooperative transition occurs at a higher t_m and is much larger but it is again followed by a gradual increase. Thus two types of transition are clearly visible which are interpreted as being caused by a cooperative dissolution of the polypurine self-structure followed by unstacking of the bases. As with other single-stranded polymers (13) this latter process would appear to be noncooperative and only slightly dependent on the ionic strength. Because of this slow unstacking of the bases the hyperchromicity change is very dependent on the temperature range over which it is measured and thus we have taken the t_m to be the mid-point of the cooperative transition as judged by eye. t_m measurements are thus thought to be accurate to ±2°. Another feature of the melting of $d(GGA)_n$, which can also be observed in some of the other melting profiles (see below), is the small increase in absorbance before the cooperative transition. This is thought to be due to length hetero-

geneity and imperfect matching up of the strands.

Figures 2b and c shows the transitions for $r(GAA)_n$, $r(GGA)_n$ and $r(GA)_n$ at high ionic strength. The hyperchromicity is much larger for $r(GAA)_n$ and the t_m's increase with increasing guanine content of the polymers. As with the melting of $d(GGA)_n$ they all show a cooperative transition followed by a slow increase in absorbance. Figure 2c also demonstrates that at high ionic strength the melting is rapidly reversible. We have found this to be true of all polymers tested as long as the ionic strength is greater than approximately 50 mM (data not shown). However in 10 mM Tris-HCl pH 8, 0.1 mM EDTA reformation of the self-structure is exceedingly slow unless the solution is frozen and thawed. (Note that $r(GAA)_n$ will not form a self-structure at this low ionic strength.)

The effect of formaldehyde on the melting of $d(GA)_n$ is shown in Figure 2d. In the absence of formaldehyde the transition is freely reversible as with $r(GA)_n$ but in its presence the reverse transition is abolished. Note that on reacting with formaldehyde there is an increase in absorbance at 260 nm (14) so that on cooling, any drop in absorbance due to restacking is expected to be masked. Since formaldehyde reacts primarily with amino groups (14) one can conclude that they are to some degree involved in the stabilization of the self-structure. Other experiments with $d(GA)_n$ have shown that the t_m is unaffected by changing the pH from 5 to 8. Presumably, therefore, the structure does not involve protonated bases. Moreover, in 0.25 M NaCl at an ethidium to phosphate ratio of 0.3, no increase in the t_m of $d(GA)_n$ could be detected, nor do any of the polypurine nucleic acids enhance the fluorescence of ethidium under the conditions of the pH 8 fluorescence assay (15). Thus it may be concluded that ethidium does not intercalate the polypurine self-structure.

The effect of modifying the bases as well as randomizing the base sequence is investigated in Figure 3. As mentioned in Methods $d(IA)_n$ was "isolated" by dismutating $d(TC)_n \cdot d(IA)_n$ at pH 5. The dismutation is extremely rapid (by comparison with - e.g. $d(TC)_n.d(GA)_n$) and no duplex DNA can be detected by the pH 5 fluorescence assay (2,15) even after only 5 minutes at pH 5 at room temperature. The melting profile of the dismutated material is shown in Figure 3a and two distinct transitions can be seen. This profile is very similar to that found for dismutated $d(TC)_n \cdot d(GA)_n$ at lower ionic strength (see 2). By comparison, the transition at 82° is due to the triplex $d(TC)_n \cdot d(IA)_n \cdot d(\overset{+}{C}T)_n$ and the transition at 48° is interpreted as being due to the self-structure of $d(IA)_n$. Both $d(GA)_n$ (t_m=58°) and

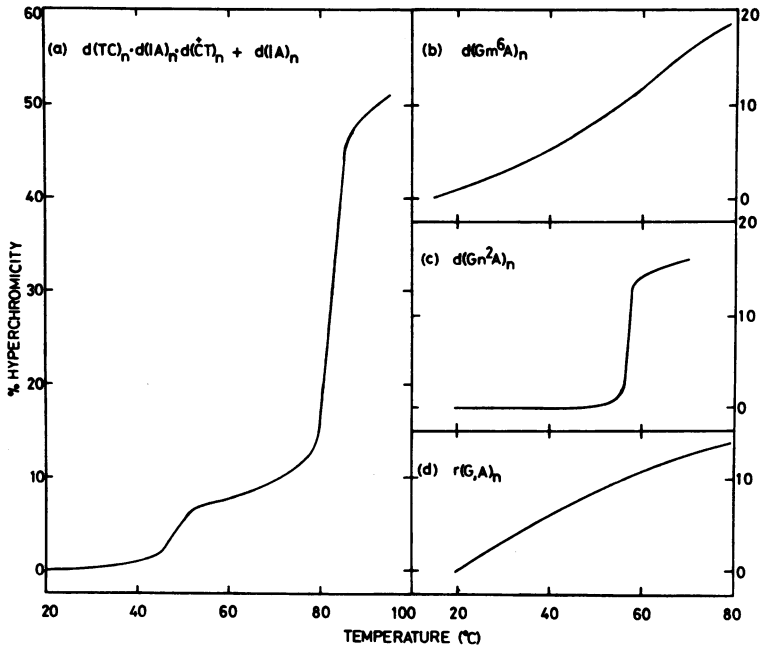

Figure 3. Thermal-denaturation profiles (a) $d(TC)_n \cdot d(IA)_n \cdot d(\overset{+}{C}T)_n$ + $d(IA)_n$ in 50 mM Na acetate pH 5, 0.1 mM EDTA, 0.25 M NaCl. (b) $d(Gm^6A)_n$ (c) $d(Gn^2A)_n$ and (d) $r(G,A)_n$ in 10 mM Tris-HCl pH 8, 0.1 mM EDTA and 0.25 M NaCl.

$d(GAA)_n$ (t_m=44°) give the same t_m whether the polymer is pure or "isolated" from dismutated duplex at pH 5. Thus the t_m for $d(IA)_n$ is considered to be reliable and shows that hypoxanthine can be substituted for guanine.

Figure 3b shows the melting profile for $d(Gm^6A)_n$. Clearly there is no cooperative transition in the temperature range 15-80° nor is one observed even in 1.2 M NaCl. The gradual increase in hyperchromicity is compatible with simple unstacking of the bases. $d(GA)_n$ under these conditions (0.25 M NaCl) has a t_m of 58° with a profile similar to that shown above for $d(GGA)_n$ (Figure 2a). Thus a methyl group on the 6-amino of adenine is sufficient to disrupt the self-structure. By comparison $d(Gn^2A)_n$ (Figure 3c) gives an exceedingly sharp transition with a t_m of 57°. Although the t_m is not significantly different from that of $d(GA)_n$ the transition is the sharpest observed for any of these polymers.

Finally, Figure 3d shows that randomizing the base-sequence abolishes

the cooperative transition (c.f. Figure 2c) although base unstacking still occurs. Thus this is further evidence that specific interactions between the bases are involved in formation of the polypurine self-structure.

The t_m data for all those polypurine nucleic acids which gave cooperative transitions is summarized in Figure 4. For completeness dI_n, rG_n and dG_n have been included from the work of previous authors (16,17,18) although under these conditions of high ionic strength (0.25 M NaCl) rG_n and dG_n would be expected to melt well above 100°. As mentioned above both the DNA and RNA polymers have increasing t_m's with increasing guanine content although the trend is much more marked with the RNA polymers. Another interesting feature is the t_m of $d(IA)_n$ which is very much higher than that of dI_n and falls between the t_m's of $d(GAA)_n$ and $d(GA)_n$. Finally it should be noted that rA_n and dA_n do not appear in Figure 4. dA_n has been extensively investigated by Riley et al. (13) and even in 5 M NaCl no cooperative transition was observed. Similarly, we have heated rA_n at high ionic strength and in the presence of $MgCl_2$ and spermidine. (Under certain con-

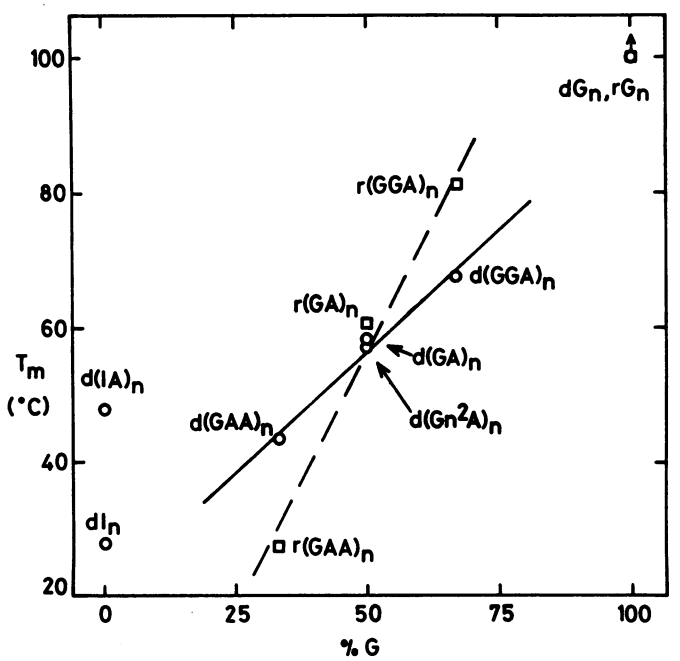

Figure 4. t_m as a function of % guanine content in 0.25 M NaCl. At this high ionic strength dG_n and rG_n would be expected to melt above 100°. The t_m for $d(IA)_n$ was measured at pH 5, all others were at neutral pH. See text for details.

ditions especially high polymer concentrations precipitation occurred. This might imply formation of an ordered structure.) Again no cooperative transition could be found though the absorbance did increase with temperature presumably due to unstacking of the bases. Thus the presence of guanine or hypoxanthine in the sequence would appear to be a necessary condition for the formation of self structure.

Circular dichroism(C.D.) spectra can also be used to demonstrate the formation of an ordered structure. As mentioned above the self-structure forms exceedingly slowly in 10 mM Tris-HCl pH 8, 0.1 mM EDTA but once formed in this buffer it is stable at room temperature (except for $r(GAA)_n$ whose t_m is very low in this buffer). The spectrum of the single-stranded form can then be obtained by heating above the t_m followed by cooling back down to room temperature. The spectra of the single-stranded and ordered forms of $d(GGA)_n$ obtained in this manner are shown in Figure 5a. The spectrum of

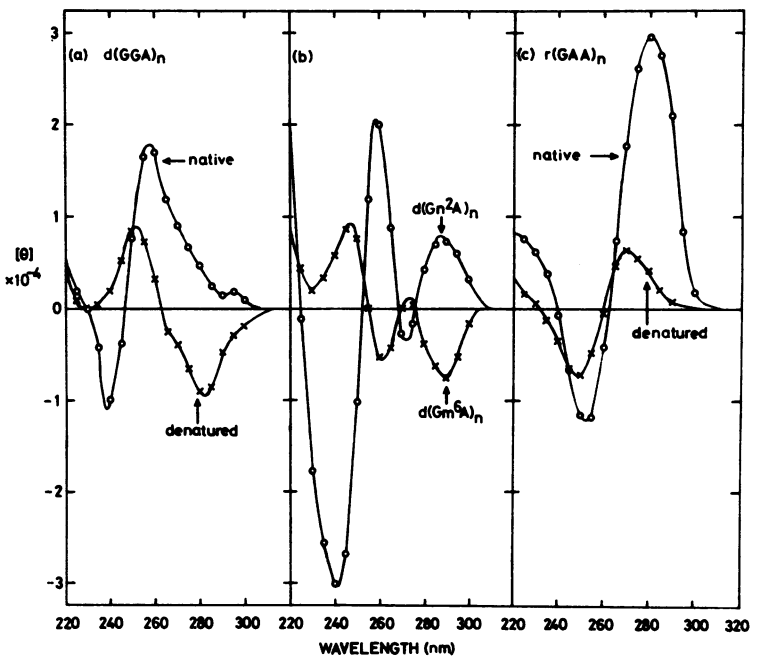

Figure 5. Circular dichroism spectra (a) $d(GGA)_n$ in 10 mM Tris-HCl pH 8, 0.1 mM EDTA, O=native at 20°, X=after heating to 50° and cooling to 27°. The low ionic strength prevents reannealing. (b) $d(Gn^2A)_n$ (O) and $d(Gm^6A)_n$ (X) in 10 mM Tris-HCl pH 8, 0.1 mM EDTA, 0.25 M NaCl. (c) $r(GAA)_n$ in 10 mM Tris-HCl pH 8, 0.1 mM EDTA, 0.25 M NaCl. O=native at 15° and X= denatured at 42°.

$d(GGA)_n$ before melting is very similar to that of $d(GA)_n$ which has previously been published (2). Moreover the spectra of dG_n and rG_n also show the same features (19,20) - that is they all contain positive bands centered around 260 nm and negative bands around 240 nm, suggesting that they share a common structure. On the other hand, after melting and cooling the spectrum of $d(GGA)_n$ changes dramatically with a negative band appearing at 280 nm and a positive band at 255 nm. The C.D. spectra of d(GpG), d(ApG) and d(GpA) all have negative bands above 260 nm (21) and thus this latter spectrum is consistent with the single-stranded form of $d(GGA)_n$.

The C.D. spectra of $d(Gn^2A)_n$ and $d(Gm^6A)_n$ recorded in 0.25 M NaCl are shown in Figure 5b. The spectrum of $d(Gn^2A)_n$ is similar to that of the ordered form of $d(GGA)_n$ (Figure 5a) except that the positive band (above 260 nm) is split into two components. On the other hand, the spectrum of $d(Gm^6A)_n$ is similar to that of the single-stranded form of $d(GGA)_n$ except that the negative band (above 260 nm) is split into two components. Thus as with the thermal-denaturation profiles the C.D. data is consistent with $d(Gn^2A)_n$ forming an ordered structure while $d(Gm^6A)_n$ remains single-stranded.

Finally the C.D. spectra of the ordered and single-stranded forms of $r(GAA)_n$ are shown in Figure 5c. Before heating the spectrum is similar to that of rG_n (19) except that the bands are shifted to longer wavelengths by about 20 nm. On heating to obtain the single-stranded form the positive and negative bands reduce in intensity and shift to shorter wavelengths by 20 and 5 nm respectively. This latter spectrum is almost identical to that of rA_n (22) which is single-stranded at this pH. Similar results have also been obtained for $r(GGA)_n$ and $r(GA)_n$ (data not shown).

DISCUSSION

The t_m data and C.D. spectra clearly demonstrate that all deoxyribose and ribose repeating-sequence purine polymers containing unmodified adenine and at least 33% guanine form a stable secondary structure. The fact that the t_m's increase approximately linearly with increasing guanine content (Figure 4) and that the C.D. spectra show similar features (Figure 5) would tend to support the view that these polymers share a common structure. Since polypurines with an unequal content of guanine and adenine (e.g. $d(GGA)_n$) form this structure then it is necessary to postulate hydrogen bonding only of guanine with guanine and adenine with adenine. Furthermore the structures form readily at neutral pH so that protonated bases are unlikely. We have been able to formulate only two types of isomorphous

structures which are consistent with this reasoning (Figure 6). The base-tetrad for guanine (and the analogous structure for hypoxanthine) have been postulated on the basis of X-ray fibre-diffraction studies of rG_n and rI_n (23,24). dG_n and dI_n are also thought to form the same tetra-stranded structure (tetraplex). However, it can be seen that adenine (or 2-amino adenine) can give rise to a completely isomorphous base-tetrad. This novel structure is sterically reasonable as judged from CPK models. A hydrogen-bonding scheme involving guanine and adenine base-pairs is also shown in Figure 6. These, too, are isomorphous and sterically reasonable.

Although it is not possible to unequivocally distinguish between these two possibilities, several factors suggest that the tetraplex is more probable. First, there is a strong precedent for the tetraplex model

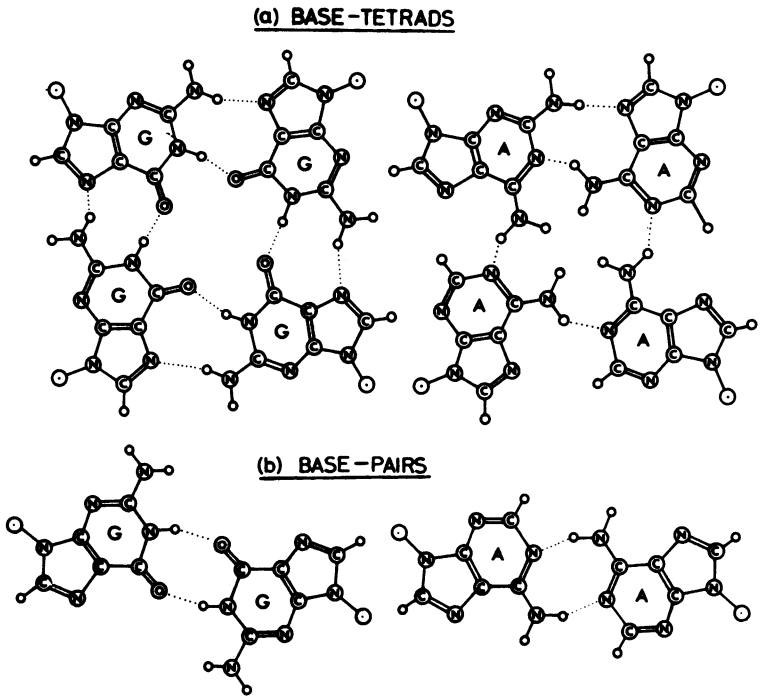

Figure 6. Possible isomorphous (a) base-tetrads and (b) base-pairs for polypurines. Note that hypoxanthine (which lacks a 2-amino group) can, replace guanine resulting in only one hydrogen bond per base for the tetrad structure. 2-amino adenine (see the top left hand base of the adenine tetrad) can replace adenine resulting in an additional hydrogen bound per base in the tetrad. Also 6-methylamino adenine can be accommodated in the adenine base-pair but not in the adenine base-tetrad.

since it is known to exist for both rG_n and rI_n (23,24). From model building studies we have not been able to identify any steric problems upon accommodation of the adenine tetrad into a tetraplex in the presence or absence of the guanine tetrad. Second, the C.D. spectra of the adenine-containing polymers share common features with those of rG_n and dG_n which are thought to be tetra-stranded (19, 20). However many other right-handed helices have a positive band at 260 nm and a negative band at 240 nm (B-DNA for example (25)). Thus the C.D. data is considered to be consistent with the tetraplex rather than definitive.

Finally the fact that $d(Gm^6A)_n$ is unable to form a self-structure provides strong evidence for the tetraplex since it is impossible to incorporate an extra four methyl groups in the middle of the proposed adenine tetrad. No such restriction exists for the purine duplex structure although the work of Engel and von Hippel (3) and the data of Table I would suggest some destabilization due to the presence of 6-methyl adenine. Thus by comparison with the t_m of $d(GA)_n$ that of $d(Gm^6A)_n$, if two-stranded, would be expected to be approximately 30° in 0.25 M NaCl. Yet no evidence for a transition could be found even in 1.2 M NaCl.

One objection which could be raised against the tetraplex model is the close promixity of four amino groups in the centre of the adenine tetrad. This is a rather unusual configuration but in the absence of detailed energy calculations there is no a priori reason for rejecting it. It is still possible to envisage a tetraplex model without adenine tetrads if the adenine residues are positioned extrahelically. For example, the structure of $d(GA)_n$ would consist of a core of stacked guanine tetrads surrounded by non-interacting adenine residues pointing away from the helix axis, (personal communication, Dr. D. Gray). Extrahelical bases have been proposed for the structure of several mismatched double helices (26). However, if in this case the adenine residues were extrahelical then one would expect both $d(GA)_n$ and $d(Gm^6A)_n$ to form a self-structure with equal facility. Thus the fact that $d(Gm^6A)_n$ remains single-stranded is good evidence against the formation of extrahelical structures.

It might be thought that the t_m data was also consistent with the tetraplex model since only in this case does the number of hydrogen bonds (and thus the t_m) decrease with the increase in adenine content of the polymers. However it is impossible to establish what the relative contributions are of hydrogen bonds and stacking interactions to the final stability of the polymer. Indeed in this regard if only hydrogen bonds are

considered for the tetraplex (see Figure 6), the t_m of $d(IA)_n$ is anomalously high while that of $d(Gn^2A)_n$ is anomalously low. Thus if these two polymers are forming tetraplexes it may be concluded that stacking interactions play a very important role in determining the final t_m. Another anomaly concerns the large difference in hyperchromicity (upon melting) between $r(GAA)_n$ and the other ribopolymers (Figure 2). One possible explanation is that $r(GGA)_n$ and $r(GA)_n$ remain largely stacked after the cooperative transition whereas $r(GAA)_n$ does not. Again this would emphasize the importance of stacking interactions.

Long repeating polypurine sequences have been discovered in a variety of organisms; for example $d(TC)_{25}\cdot d(GA)_{25}$ in the spacer region of histone genes in the sea urchin (27); the sequence $d(GGA)_6G$ in polyoma virus (28) as well as $d(TTCTC)_n\cdot d(GAGAA)_n$ and $d(TTCTCTC)_n\cdot d(GAGAGAA)_n$ in Drosophila satellites (29). A biological role for the tetra-stranded structure proposed here seems possible for DNA polymers although there are plectonemic aspects to be considred. It also seems possible that this hydrogen-bonding scheme could be utilized in the tertiary folding of RNA for example. The discovery of rA_n tails on most eukaryotic mRNAs is intriguing in this regard since this polymer appears to be unable to form a self-structure even in the presence of Mg^{2+} or polyamines. Proteins, however, might be able to constrain rA_n to form a tetraplex and this could then be used for storage or transportation of mRNAs.

ACKNOWLEDGEMENTS

We would like to thank Dr. C. Kay and K. Oïkawa for performing the C.D. measurements and Dr. H. van de Sande and J. Lewis for their gift of dn^2ATP. This work was supported by the Medical Reaearch Council of Canada.

Abbreviations*

I=Inosine, m^6A=6-methylamino adenosine, n^2A=2-amino adenosine.

REFERENCES

1. Johnson, D.A. and Morgan, A. R. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 1637-1641.
2. Lee, J.S., Johnson, D.A. and Morgan, A.R. (1979) Nucleic Acids Res. 6, 3073-3091.
3. Engel, J.D. and von Hippel, P.H. (1978) J. Biol. Chem. 253, 927-934.
4. Chamberlin, M. and Berg, P. (1962) Proc. Natl. Acad. Sci. U.S.A. 48, 81-94.
5. Jovin, T.M., Englund, P.T. and Bertsch, L.L. (1969). J. Biol. Chem. 244, 2996-3008.

6. Morgan, A.R., Evans, D.H., Lee, J.S. and Pulleyblank, D.E. (1979) Nucleic Acids Res.7, 571-594.
7. Morgan, A. R., Coulter, M.B., Flintoff, W.F. and Paetkau, V.H. (1974) Biochemistry 13, 1596-1603.
8. Harwood, S.J. and Wells, R.D. (1970) J. Biol. Chem. 245, 5625-5634.
9. Richardson, C.C., Lehman, I.R. and Kornberg, A. (1964) J. Biol. Chem. 239, 251-258.
10. Donelson, J.E. and Wu, R. (1972) J. Biol. Chem. 247, 4661-4668.
11. Morgan, A.R. and Wells, R.D. (1968) J. Mol. Biol. 37, 63-80.
12. Wells, R.D., Larson, J.E., Grant, R.C., Shortle, B.E. and Cantor, C.R. (1970) J. Mol. Biol. 54, 465-497.
13. Riley, M., Maling, B. and Chamberlin, M.J. (1966) J. Mol. Biol. 20, 359-389.
14. Grossman, L. (1968) In Methods in Enzymology, Vol. XII, Nucleic Acids, Part B. 467-480. L. Grossman and K. Moldave eds. Academic Press, New York and London.
15. Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L. and Evans, D. H. (1979) Nucl. Acids. Res. 7, 547-569.
16. Chamberlin, M.J. and Patterson, D.L. (1965) J. Mol. Biol. 12, 410-428.
17. Miles, H.T. and Frazier, J. (1964) Biochim. Biophys. Acta. 79, 216-220.
18. Englander, J.J., Kallenback, N.R. and Englander, S.W. (1972) J. Mol. Biol. 63, 153-164.
19. Thiele, D. and Guschlbauer, W. (1971) Biopolymers 10, 143-157.
20. Marck, C. and Thiele, D. (1978) Nucleic Acids Res. 5, 1017-1028.
21. Cantor, C.R., Warshaw, M.M. and Shapiro, H. (1970) Biopolymers 9, 1059-1077.
22. Brahms, J., Michelson, A.M. and Van Holde, K.E. (1966) J. Mol. Biol. 15, 467-488.
23. Zimmerman, S.B., Cohen, G.H. and Davies, D.R. (1975) J. Mol. Biol. 92, 181-192.
24. Arnott, S., Chandrasekaran, R. and Marttila, C.M. (1974) Biochem. J. 141, 537-543.
25. Chan, A., Kilkuskie, R. and Hanlon, S. (1979) Biochemistry 18, 84-91.
26. Lomant, A.J. and Fresco, J.R. (1975) Prog. Nucleic Acid Res. and Mol. Biol. 15, 185-218.
27. Sures, I., Lowry, J. and Kedes, L.H. (1979) Cell 15, 1033-1044.
28. Soeda, E., Arrand, J.R., Smolar, N., Walsh, J.E. and Griffin, B.E. (1980) Nature 283, 445-449.
29. Fry, K. and Brutlag, D. (1979) J. Mol. Biol. 135, 581-593.
30. Murray, N.L. (1972) M. Sc. thesis, University of Alberta. "Three-stranded polynucleotide structures".