Structure of a pleiomorphic form of poly d(AT):poly d(AT)

R.P.Millane, J.K.Walker, Struther Arnott*, R.Chandrasekaran and D.L.Birdsall
Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, and

---

R.L.Ratliff
Los Alamos Scientific Laboratory, University of California, Los Alamos, NM 87544, USA

Received 22 March 1984; Accepted 13 June 1984

ABSTRACT

A chemically simple polynucleotide duplex, poly d(AT):poly d(AT), has been trapped in a fibrous form with a complex helical secondary structure with a large (7.4 nm) axial repeat 24 nucleotides long. The motif which is repeated by the symmetry elements is a hexanucleotide in which two residues (both TpA) have the less common *gauche minus* conformation at C3'-O3' and consequently distinctive phosphate orientations. This reinforces earlier conclusions that PypPu nucleotides tend to have different shapes from PupPy nucleotides and that DNA surfaces may signal what base sequences lie beneath them. The morphological differences between this pleiomorphic DNA polymer and closely-related, but more symmetrical allomorphs are just as great as those observed in short DNA fragments in crystals.

INTRODUCTION

Fibrous polynucleotide duplexes are very polymorphic (1-3). This is observed both with general sequence DNAs and with synthetic DNAs such as poly d(GC):poly d(GC) and poly d(AC):poly d(GT) (with their A, B, and Z forms) (4), poly d(AT):poly d(AT) (with A, B, and D forms) and poly d(A):poly d(T) (with its heteronomous or H form) (5). The Z forms of poly d(GC):poly d(GC) and poly d(AC):poly d(GT) dramatically reflect their dinucleotide chemical structures in their secondary structures by having quite different furanose rings and backbone conformations in alternate nucleotides (6, 4). Recently we showed (7) that the B and D forms of some poly d(PuPy):poly d(PuPy) DNAs also reflect their chemical structure in their secondary structures but in a more subtle way. This prompted us to describe them as having "wrinkled" versions of the common conformations. In these wrinkled DNAs, a dinucleotide is the repeated structural motif rather than a mononucleotide as in standard or "smooth" DNA. One of the two nucleotides in the motif has essentially the same conformation as in smooth DNA but the alternate nucleotide (probably PypPu) has a different conformation about the (C3'-O3') bond and, consequently, a marked change in the orientation of its phosphate group.

Figure 1: X-ray diffraction patterns of D DNAs: (a) tetragonal, pleiomic poly d(AT):poly d(AT); (b) tetragonal, wrinkled poly d(IC):poly d(IC); and (c) pseudo-hexagonal, smooth poly d(AAT):poly d(ATT).

The wrinkles appear to be robust since they are found in B-like DNAs with different molecular symmetries, with different crystal packings and with GC, AT and IC sequences (7). These results suggest that such perturbations may be common in stretches of duplex DNA where purines and pyrimidines alternate.

Now we have obtained a novel diffraction pattern (Fig. 1a) from the lithium salt of poly d(AT):poly d(AT). The overall intensity distribution resembles those in D forms (Figs. 1b,c). However, the long axis repeat ($a = 7.4$ nm) is about 3 times that of the usual D forms ($a = 2.5$ nm) implying that the repeated structure motif contains 6 nucleotides with different conformations rather than 2. Such a large unit of structure approaches in complexity that of the self-complementary, B-like duodecanucleotide studied

in single crystals (8) and therefore provides the first opportunity to explore the extent to which features observed in crystallized DNA fragments may survive in polymers.

The new molecular structure has a conformational complexity greater than its chemical complexity. We have therefore chosen to describe this kind of secondary structure as pleiomic (9), a botanical term used to describe whorls with greater than the normal number of structures.

The structure of pleiomic poly d(AT):poly d(AT) is presented here and compared with refined versions of previously reported wrinkled (7) and smooth (10) forms of D DNA.

EXPERIMENTS

The polynucleotide was dissolved in 0.75 M LiCl pH 7.0 buffer and precipitated with ethanol. The precipitate was washed repeatedly with ethanol, and finally acetone, then dried in a vacuum desiccator. The dried material was redissolved in water and spun into a fiber at a relative humidity of 75%. The X-ray diffraction pattern (Fig. 1a) was obtained at a relative humidity of 56% using a pinhole camera and a flat film.

The optical density in the diffraction pattern was digitized using an Optronics Photoscan P-1000 two-dimensional, computer controlled, scanning microdensitometer using a 100 μ raster step. The observed diffraction amplitudes were subsequently extracted using a Digital Equipment PDP 11/40 computer and Comtal imaging system.

The reciprocal space coordinates of a number of the sharp diffraction spots were computed and used to calculate the unit cell parameters.

Using optical density samples between the Bragg reflections, a radially symmetric background function was defined. Parameters defining the mosaic spread and disorientation were estimated and used to calculate approximate spot boundaries which were centered on the reciprocal lattice points on the film. The optical density was corrected for Lorentz and polarization effects (after subtraction of the background) and summed inside each boundary to estimate the observed structure amplitudes. This procedure will be described in more detail elsewhere.

Inspection of Fig. 1a shows that there is considerable overlap between adjacent reflections particularly in the vicinity of the equator and at wider angles. This is due primarily to the large axial repeat which results in closely spaced layer lines and to the arcing of the diffraction spots due to disorientation. Where reflection boundaries overlapped significantly, they

were combined to form a single boundary and a composite structure factor determined.

Molecular and crystal structure modelling were conducted using the method of Linked Atom Least Squares (LALS) (11,12). The main molecular variables were the six conformation angles associated with the backbone of each nucleotide together with each glycosidic torsion angle. The angles were elastically bound to the mean values in whatever domain they adopted. In addition, in the final refinements, the conformation angles and endocyclic bond angles of the furanose rings were allowed to vary subject only to bond angle restraints and ring closure constraints. In this particular crystal structure, there are no packing variables but certain specific alternatives (discussed later) have to be considered.

The variables were adjusted so as to minimize (in a least squares fashion) the expression

$$\begin{aligned}\Omega &= \sum_m \Delta F_m^2 + \sum_i k_i \Delta d_i^2 + \sum_n \lambda_n G_n \\ &= X + C + L.\end{aligned}\tag{1}$$

The term X involves the differences (ΔF_m) between the observed X-ray amplitudes and those calculated for the structure. C involves the close, non-bonded interatomic distances, d_i , which are driven towards some acceptably large values, ${}_0d_i$. The term L involves constraints, G_n , which become zero when residue connectivity and furanose ring closure have been achieved.

Competing molecular and crystal structures can be assessed for significant differences by using Ω , X or C as statistics in Hamilton's test (13, 14).

THE UNIT CELL AND ITS CONTENTS

The Bragg reflections in the diffraction pattern of pleiomic D DNA (Fig. 1a) can be indexed on the basis of a tetragonal unit cell (Fig. 2) with $a = b = 1.72$ nm, $c = 7.40$ nm. The a and b dimensions are very close to those of wrinkled D DNA and the c dimension is approximately three times that of wrinkled D DNA: for wrinkled poly d(IC):poly d(IC), $a = b = 1.71$ nm, $c = 2.45$ nm.

The diffraction intensity on the meridian is confined to layer lines with $\ell = 4n$, indicating a 4-fold helix and therefore a helical repeat 1.85 nm long. Assuming an average rise per residue within the range observed in other polynucleotide structures, there are 5, 6 or 7 nucleotides per helical

Figure 2: Crystal packing of pleiomic poly d(AT):poly d(AT).

repeat. Apart from the additional layer lines (due to the long 7.4 nm c -repeat), the overall intensity distribution on the diffraction pattern is so similar to those of D DNAs (Fig. 1) that the average rise per residue is undoubtedly approximately 0.31 nm which corresponds to a hexanucleotide helical repeat in each chain or 24 base pairs per c -repeat in the duplex. If the two chains of the duplex are identical then there are diad axes perpendicular to the screw axis, mid-way between every third stack of base pairs.

There is space for only one molecule in each unit cell. If the molecular screw and diad axes coincide with unit cell edges then the space group symmetry is $P224_3$ which also has diad axes at 45° to a . The crystallographic asymmetric unit is then the hexanucleotide. Since a particular molecular diad may lie either along a or at 45° to it and since the dinucleoside monophosphate spanning the diad may be either ApT or TpA, 4 possible crystal packing arrangements have to be evaluated.

Molecular structure refinement was conducted for each packing arrangement possible for pleiomic D DNA. This was accomplished from preliminary refinements using rigid C2'-*endo* furanose rings. In each nucleotide the conformations at P-O5', O5'-C5', C5'-C4', C3'-O3', O3'-P and C1'-N were varied. The results (Table 1) show that model I with the central diad of the hexanucleotide unit at $\mu = 0^\circ$ and spanned by an ApT sequence is best. Hamilton's test, which takes account of the number of observations and variables as well as the indices of agreement, applied to $(\Omega/N)^{1/2}$ shows that this model is superior to the next best at the 99.5% level of significance.

TABLE 1
Statistics of alternative models following refinement of
pleiomic poly d(AT):poly d(AT) with fixed furanose rings

Model	Sequence	$\mu(^{\circ})$	X	N_x	Ω	N	R
I	ApT	0	75	84	367	396	0.26
II	TpA	0	82	81	446	421	0.27
III	ApT	45	101	96	480	457	0.29
IV	TpA	45	115	90	421	402	0.32

μ is the angle between the α -axis and the diad. This diad lies within the ApT dinucleotide 4 (see Table 4).

N_x and N are the number of data contributing to X and Ω respectively defined in equation (1). N_x varies because reflections which are below the observational threshold but have larger calculated values are included in the refinement.

$R = \Sigma |\Delta F_m| / \Sigma O F_m$, where the $O F_m$ are the observed structure amplitudes, is the traditional X-ray discrepancy index.

Subsequently, an extensive refinement of model I with flexible furanose rings yielded the final crystal structure with the atomic coordinates listed in Table 2 and conformations in Table 4. The maximum and average deviations of the furanose conformation angles from the C2'-endo standard values (18) are 17° and 5° respectively, and for the endocyclic bond angles, 5° and 2° respectively. The maximum and average estimated standard deviations of the varied conformation angles (see (12) for details) are 8° and 4° respectively. The calculated X-ray structure amplitudes can be compared with the observed values in Table 3. Their discrepancy index ($R = 0.23$) has the kind of low value we now expect for well-determined structures of fibrous polynucleotides. Projections of the pleiomic structure normal to and along the helix axis are shown in Fig. 3 where they are compared with the wrinkled and smooth D DNA structures. A stereo view of one turn of the pleiomic structure is also shown in Fig. 4.

SMOOTH AND WRINKLED D DNA

To provide appropriate structures (Table 4, Fig. 3) for comparison with pleiomic D DNA, the models of smooth (10) and wrinkled (7) D DNA were re-refined with the furanose conformations not elastically tethered

TABLE 2

Cartesian (X,Y,Z) and cylindrical (R, ϕ ,Z) coordinates of atoms in a hexanucleotide of pleiometric poly d(AT):poly d(AT). The coordinates of the diadically related hexanucleotide in the antiparallel chain are given by (R,- ϕ ,Z). Successive hexanucleotides are given by (R, ϕ -90°, Z+1.85 nm).

ATOM	X(NM)	Y(NM)	Z(NM)	R(NM)	PHI(DEG)	ATOM	X(NM)	Y(NM)	Z(NM)	R(NM)	PHI(DEG)
PHOSPHATE						PHOSPHATE					
P	-0.4385	-0.7320	1.1276	0.8533	-120.92	P	0.1565	0.8256	0.3118	0.8403	78.27
O1	-0.3531	-0.7938	1.1980	0.9675	-124.87	O1	0.2597	0.9222	0.3553	0.9581	74.27
O2	-0.3681	-0.8182	1.0300	0.8972	-114.22	O2	0.0495	0.8797	0.2250	0.8811	86.78
O3	-0.3338	-0.8725	1.2345	0.7335	-118.29	O3	0.0314	0.7550	0.4337	0.7605	53.10
O5	-0.4860	-0.5972	1.0557	0.7700	-129.14	O5	0.2266	0.7013	0.2394	0.7371	72.09
SUGAR						SUGAR					
C1	-0.3219	-0.3251	0.9543	0.4575	-134.72	C1	0.2865	0.4299	0.0343	0.5166	56.32
C2	-0.4479	-0.3407	0.8639	0.5228	-142.74	C2	0.3414	0.5597	-0.0240	0.6556	58.52
C3	-0.5597	-0.3009	0.8563	0.6355	-151.74	C3	0.4615	0.5859	0.0588	0.7493	61.77
C4	-0.5055	-0.3533	1.0968	0.6202	-144.59	C4	0.4046	0.5415	0.2016	0.6759	53.23
C5	-0.5558	-0.4977	1.1330	0.7461	-138.15	C5	0.3518	0.6522	0.2907	0.7410	61.65
O4	-0.3603	-0.3509	1.0885	0.5030	-135.76	O4	0.3029	0.4406	0.1748	0.5347	55.50
H1	-0.2917	-0.2203	0.9555	0.3553	-142.93	H1	0.3516	0.3471	0.0062	0.4940	44.63
H2	-0.4591	-0.4456	0.8425	0.6336	-135.85	H2	0.2663	0.6336	-0.0116	0.6936	67.24
H3	-0.4440	-0.2635	0.7874	0.5194	-148.74	H3	0.3759	0.5427	0.1260	0.6502	55.29
H4	-0.6509	-0.3535	0.9330	0.7407	-151.49	H4	0.4820	0.6329	0.0636	0.8441	55.18
H5	-0.5262	-0.4906	1.1708	0.6011	-151.09	H5	0.4798	0.4840	0.2557	0.6815	45.05
H15	-0.6265	-0.5041	1.1129	0.8355	-142.74	H15	0.4242	0.7326	0.2942	0.9474	59.16
H25	-0.5390	-0.5158	1.2392	0.7461	-136.26	H25	0.3371	0.6136	0.3517	0.7901	61.22
ADENINE						THYMINE					
N1	0.1614	-0.3257	0.7797	0.3635	-63.64	N1	0.1468	0.3998	-0.0079	0.4259	69.84
C2	0.0789	-0.2275	0.8157	0.2408	-70.87	C2	0.1202	0.2714	-0.0490	0.2968	66.11
N3	-0.0447	-0.2348	0.8939	0.2391	-100.78	C3	0.2050	0.1829	0.0521	0.2754	41.89
C4	-0.0856	-0.3632	0.8356	0.3732	-103.27	C4	-0.0103	0.2472	0.0073	0.2474	92.38
C5	-0.0129	-0.4738	0.8334	0.4740	-91.56	C5	-0.1137	0.3395	-0.0879	0.3571	108.57
N6	0.1181	-0.4523	0.7874	0.4080	-75.79	C6	0.2267	0.3051	-0.1243	0.3801	126.36
N7	-0.0879	-0.5890	0.8332	0.5995	-98.49	C5	-0.0761	0.4705	-0.0430	0.4766	99.19
O8	-0.2017	-0.5455	0.8813	0.5812	-110.31	H6	-0.1827	0.5760	-0.0404	0.6043	107.60
N9	-0.2079	-0.4031	0.9079	0.4589	-116.94	H5	0.0501	0.4958	0.0052	0.5449	94.43
H2	0.1163	-0.1350	0.8080	0.1782	-49.27	H3	-0.0319	0.1544	-0.1177	0.1577	101.67
H16	0.2930	0.3313	0.7132	0.6057	-61.12	H6	0.0739	0.5889	0.0255	0.5935	82.85
H26	0.1638	-0.6472	0.7565	0.6691	-75.30	PHOSPHATE					
H8	-0.2780	-0.6052	0.9209	0.6660	-114.67	P	0.7200	0.5902	0.0347	0.9310	39.34
PHOSPHATE						O1	0.6067	0.7736	0.1717	0.9831	39.10
O1	-0.6658	-0.0620	0.8729	0.6708	-172.98	O2	0.7043	0.7144	-0.0443	1.0032	45.41
O2	-0.7574	0.0143	0.9370	0.7575	-178.32	O3	0.5804	0.5124	0.0406	0.7742	41.44
O3	-0.5837	-0.1619	0.9846	0.6058	-164.50	O5	0.8108	0.4863	-0.0463	0.9454	30.95
O5	-0.5466	-0.0026	0.7983	0.5466	-178.72	SUGAR					
SUGAR						C1	0.6662	0.2341	-0.2207	0.7250	18.84
C1	-0.2685	0.0643	0.6059	0.2761	166.54	C2	0.8262	0.2622	-0.2741	0.8668	17.61
C2	-0.3995	0.1140	0.5457	0.4155	164.07	C3	0.9183	0.2082	-0.1646	0.9416	12.77
C3	-0.4261	0.2355	0.5347	0.4860	151.07	C4	0.8363	0.2455	-0.0411	0.8715	16.36
C4	-0.3880	0.1755	0.7066	0.4523	125.24	C5	0.8723	0.3772	0.0247	0.9503	23.38
C5	-0.5028	0.1222	0.8334	0.5174	165.34	O4	0.6557	0.2405	-0.0793	0.7361	19.07
O4	-0.2951	0.0762	0.7463	0.2555	165.03	H1	0.5329	0.1302	-0.2405	0.6726	11.16
H1	-0.1886	0.1345	0.7821	0.5317	144.51	H2	0.8395	0.3700	-0.2329	0.9175	23.78
H12	-0.4775	0.0396	0.5615	0.4792	175.26	H4	0.8411	0.2049	-0.3557	0.8657	13.69
H22	-0.3624	0.1471	0.4433	0.4097	155.36	H5	0.8254	0.2654	0.1653	0.9453	14.71
H3	-0.5325	0.2291	0.6336	0.5262	154.06	H6	0.8426	0.1652	0.0223	0.8596	11.09
H4	-0.3334	0.2509	0.8295	0.4172	143.03	H15	0.8806	0.3898	0.0240	1.0552	21.68
H15	-0.5947	0.1941	0.8536	0.6161	161.63	H25	0.6375	0.3768	0.1280	0.9184	24.23
H25	-0.4653	0.1066	0.9580	0.4813	167.20	ADENINE					
THYMINE						N1	0.2058	0.2641	-0.4005	0.3348	52.08
N1	-0.2296	-0.0722	0.5605	0.2409	-162.55	C2	0.2778	0.1627	-0.3530	0.3220	30.35
C2	-0.0980	-0.0920	0.5271	0.1345	-136.81	N3	0.4015	0.1624	-0.3083	0.4331	22.02
O2	-0.0141	-0.0038	0.5333	0.0146	-164.92	C4	0.4552	0.2660	-0.3151	0.5376	32.14
N3	-0.0659	-0.2187	0.4825	0.2294	-106.63	C5	0.3941	0.3990	-0.3511	0.4508	45.36
C4	-0.1526	-0.3266	0.4748	0.3605	-115.04	C6	0.2617	0.3864	-0.4062	0.4667	55.89
O4	-0.1121	-0.4385	0.4363	0.4507	-104.40	H6	0.1897	0.4890	-0.4540	0.5245	68.79
C5	-0.0887	-0.2861	0.5120	0.4135	-134.28	H7	0.4803	0.5075	-0.3530	0.6388	46.58
H6	-0.3895	-0.4068	0.5033	0.5632	-133.76	C8	0.5891	0.4574	-0.3031	0.7458	37.63
H2	-0.3267	-0.1727	0.5529	0.3660	-151.84	N9	0.5814	0.3236	-0.2775	0.6834	29.10
N3	0.0295	-0.7264	0.4607	0.7382	-62.68	H2	0.2313	0.0742	-0.3511	0.2429	17.79
H6	-0.4176	-0.1538	0.5781	0.4450	-159.78	H16	0.0958	0.4740	-0.4847	0.4836	78.58
PHOSPHATE						H25	0.2300	0.5804	-0.4588	0.6243	68.38
P	-0.4104	0.4952	0.6482	0.6431	129.65	H8	0.6710	0.5119	-0.2851	0.8439	37.34
O1	-0.3059	0.5231	0.7381	0.6524	126.70	PHOSPHATE					
O2	-0.5501	0.5023	0.3269	0.7449	137.60	P	1.0297	0.0032	-0.2807	1.0297	0.18
O3	-0.3483	0.3523	0.6119	0.4954	134.67	O1	1.1653	-0.1121	-0.2283	1.1121	-5.75
O5	-0.3184	0.5927	0.5609	0.6728	118.24	O2	1.0894	-0.1142	-0.2142	1.1142	-5.75
SUGAR						O3	0.9433	0.0682	-0.1627	0.9488	4.14
C1	-0.0628	0.4924	0.3910	0.4964	97.27	O5	0.9148	-0.0488	-0.3791	0.9161	-3.05
C2	-0.0935	0.6312	0.3295	0.6390	98.96	SUGAR					
C3	-0.0476	0.7250	0.4386	0.7265	93.76	C1	0.6125	-0.0378	-0.5839	0.6137	-3.53
C4	-0.0809	0.6404	0.5615	0.6495	97.20	C2	0.7039	-0.0250	-0.6923	0.7409	-7.22
C5	-0.2150	0.6694	0.6266	0.7022	107.83	C3	0.7565	-0.2230	-0.5382	0.7887	-16.43
O4	-0.0669	0.5907	0.3266	0.5051	97.61	C4	0.7199	-0.1808	-0.4793	0.7727	-11.03
H1	-0.0419	0.4719	0.3684	0.4738	94.92	C5	0.6382	-0.1830	-0.4304	0.6519	-11.03
H12	-0.2091	0.6387	0.3237	0.6718	108.04	O4	0.5286	-0.0603	-0.4463	0.6415	-5.39
H22	-0.0451	0.6185	0.2369	0.6511	93.97	O6	0.5630	-0.1000	-0.5060	0.5355	-10.81
H3	-0.1083	0.8158	0.4357	0.8207	97.58	H12	0.0271	0.8204	-0.0171	0.8209	0.69
H4	-0.0017	0.6511	0.6356	0.6511	90.15	H22	0.7102	-0.1164	-0.7395	0.7197	-3.31
H15	-0.2375	0.7748	0.6190	0.8104	107.04	H3	0.0618	-0.2509	-0.5311	0.3876	-21.23
H25	-0.2108	0.6406	0.7319	0.6744	108.22	H4	0.5429	-0.0400	-0.5689	0.5429	-21.23
ADENINE						H15	0.9386	-0.2524	-0.3420	0.9334	-15.69
N1	-0.0688	0.0174	0.1953	0.0709	165.82	H25	0.7982	-0.1848	-0.2392	0.8119	-10.53
C2	0.0298	0.0949	0.2301	0.0995	72.56	THYMINE					
N3	0.0238	0.2167	0.2793	0.2180	83.72	N1	0.5815	0.1037	-0.6186	0.5907	10.11
C4	-0.1034	0.2616	0.2809	0.2913	57.77	C2	0.4529	0.1318	-0.6579	0.4717	16.22
C5	-0.2142	0.1941	0.2386	0.2030	137.82	O2	0.3658	0.0468	-0.6561	0.3688	7.29
O4	-0.1948	0.0645	0.1880	0.2052	181.67	C3	0.2632	0.0279	-0.6439	0.2632	31.67
H1	-0.2945	0.1429	0.2348	0.2467	47.77	C4	0.1957	0.3678	-0.6946	0.1939	35.34
N7	-0.3279	0.2719	0.2535	0.4260	140.33	H6	0.4846	0.4823	-0.7151	0.5037	44.86
O8	-0.2829	0.3823	0.3067	0.4736	126.51	H1	0.3095	0.3895	-0.6744	0.3732	26.77
H9	-0.1477	0.5848	0.2598	0.4739	126.51	H4	0.7560	0.4354	-0.7344	0.7274	29.94
C8	-0.2823	0.3823	0.3067	0.4736	126.51	C6	0.6781	0.2007	-0.6112	0.7071	16.49
H9	-0.1477	0.5848	0.2598	0.4739	126.51	H5	0.3347	0.7182	-0.5807	0.6815	16.49
H16	-0.0017	0.6511	0.6356	0.6511	90.15	H6	0.7704	0.1757	-0.5820	0.7902	12.85
H15	-0.2375	0.7748	0.6190	0.8104	107.04						
H25	-0.2108	0.6406	0.7319	0.6744	108.22						

TABLE 3

(a) Observed and (b) calculated structure amplitudes for pleiometric poly d(AT):poly d(AT).

a		H	K	L=0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	0	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
1	0	N	(147)	181	497	230	582	(277)	(280)	(274)	820	818	(377)	(382)	(382)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)
1	1	500	*	108	302	(240)	301	930	218	428	480	802	801	820	848	(448)	(472)	(480)	(482)	(482)	(482)	(482)	(482)	(482)	(482)
2	0	(543)	(283)	804				422	483	(364)															
2	1		880	*				(328)	(343)	408	802	804	830	887	1078	1030	1270	1078	802	(887)	1102	1324	1228	(711)	
2	2	841	(363)	(279)	(384)	(412)	(430)	(448)	(487)																
3	0	(388)	(388)	(388)	(403)	(418)	(428)	(480)	(487)																
3	1	1277	*		817		788	703	1042	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027
3	2		1872	*			(483)	(500)	828	788	703	1018	985	885	884	884	884	884	884	884	884	884	884	884	884
4	0		1878	*			(522)	(549)	781	(864)	(802)	(822)	(844)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)
4	1	(528)	(821)	(528)	(580)	(587)	(582)	(601)	(621)	(640)	(662)	(684)	(708)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)
5	0		1431	*			(888)	(888)	(902)	(920)	(938)	(960)	(980)	(1002)	(1020)	(1042)	(1062)	(1082)	(1102)	(1122)	(1142)	(1162)	(1182)	(1202)	(1222)
5	1		2182	*			(927)	(950)	(975)	(1002)	(1028)	(1055)	(1082)	(1108)	(1135)	(1162)	(1188)	(1215)	(1242)	(1268)	(1295)	(1322)	(1348)	(1375)	(1402)
6	0	(778)	(688)	(674)	(681)	(708)	(728)	(744)	(764)	(784)	(798)	(808)													
6	1	(688)	(684)	(682)	(703)	(723)	(741)	(762)	(782)	(792)	(804)														
6	2	(780)	(740)	(744)	(768)	(778)																			

b		H	K	L=0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	0	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
1	0	N	(147)	181	497	230	582	(277)	(280)	(274)	820	818	(377)	(382)	(382)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)	(408)
1	1	500	*	108	302	(240)	301	930	218	428	480	802	801	820	848	(448)	(472)	(480)	(482)	(482)	(482)	(482)	(482)	(482)	(482)
2	0	(543)	(283)	804				422	483	(364)															
2	1		880	*				(328)	(343)	408	802	804	830	887	1078	1030	1270	1078	802	(887)	1102	1324	1228	(711)	
2	2	841	(363)	(279)	(384)	(412)	(430)	(448)	(487)																
3	0	(388)	(388)	(388)	(403)	(418)	(428)	(480)	(487)																
3	1	1277	*		817		788	703	1042	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027	1027
3	2		1872	*			(483)	(500)	828	788	703	1018	985	885	884	884	884	884	884	884	884	884	884	884	884
4	0		1878	*			(522)	(549)	781	(864)	(802)	(822)	(844)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)	(864)
4	1	(528)	(821)	(528)	(580)	(587)	(582)	(601)	(621)	(640)	(662)	(684)	(708)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)	(718)
5	0		1431	*			(888)	(888)	(902)	(920)	(938)	(960)	(980)	(1002)	(1020)	(1042)	(1062)	(1082)	(1102)	(1122)	(1142)	(1162)	(1182)	(1202)	(1222)
5	1		2182	*			(927)	(950)	(975)	(1002)	(1028)	(1055)	(1082)	(1108)	(1135)	(1162)	(1188)	(1215)	(1242)	(1268)	(1295)	(1322)	(1348)	(1375)	(1402)
6	0	(778)	(688)	(674)	(681)	(708)	(728)	(744)	(764)	(784)	(798)	(808)													
6	1	(688)	(684)	(682)	(703)	(723)	(741)	(762)	(782)	(792)	(804)														
6	2	(780)	(740)	(744)	(768)	(778)																			

Reflections below the threshold of observation are given in parentheses, and observed values are equal to the threshold and were included in the refinement only when the calculated value was greater than the threshold. M denotes a meridional reflexion, N a systematic absence and S a reflection which saturated the film. Because of arcing across the equator, the reflections marked * include intensities recorded above and below the equator. A constant isotropic attenuation factor $\exp(0.15(\sin^2\theta)/\lambda^2)$, where $\lambda = 0.15418$ nm, was applied to the calculated amplitudes.

Figure 3: Projections normal to and along the helix axes (denoted by the filled circles) of 24 base pairs of (a) smooth D DNA (poly d(AAT):poly d(ATT)); (b) pleiomic poly d(AT):poly d(AT); and (c) wrinkled poly d(IC):poly d(IC).

Figure 4: Stereo view of one turn of pleiomeric poly d(AT):poly d(AT).

to standard C2'-*endo* values as was the case with the earlier refinements (7). Although the backbone conformation angles in wrinkled D DNA differ from the earlier model by 9° on average, the assignment of base sequence is unaltered and the morphology is essentially the same. To represent wrinkled D DNA we chose poly d(IC):poly d(IC) since the set of diffraction data it provided is much better than that of wrinkled poly d(AT):poly d(AT) (7). The two molecules have essentially the same conformations although their packings are only pseudoisomorphous (7).

NUCLEOTIDE CONFORMATIONS

Smooth B-type (i.e., B, C or D) DNAs have a *trans* conformation at C3'-O3' and either a *trans* or *gauche minus* conformation at O3'-P. In wrinkled B-type DNAs the PupPy nucleotides have conformations like smooth DNAs, whereas the PypPu nucleotides have *gauche minus*, *trans* conformations at the C3'-O3', O3'-P bonds (7). Pleiomeric poly d(AT):poly d(AT) (Table 4) is similar to wrinkled DNA in having the C3'-O3', O3'-P conformations *trans*, *gauche minus* in all three ApT components of its hexanucleotide repeat and in having *gauche minus*, *trans* conformations in two of its TpA components.

TABLE 4
Conformation angles for pleiomic, wrinkled, and smooth D DNA.

D DNA	Pleiomic						Wrinkled		Smooth
Nucleotide†	1	2	3	4	5	6	1	2	
Parameters	TpA	ApT	TpA	ApT	TpA	ApT	CpI	IpC	NpN
Backbone Conformations (°)									
(C4'-C3'-O3'-P)	-152	-169	-96	172	-114	179	-109	-130	-163
(C3'-O3'-P-O5')	-149	-104	-160	-101	-160	-96	-152	-168	-127
(O3'-P-O5'-C5')	-64	-84	-103	-78	-83	-72	-81	-63	-66
(P-O5'-C5'-C4')	138	-178	124	-176	126	171	134	136	158
(O5'-C5'-C4'-C3')	78	74	85	74	81	68	65	60	68
(C5'-C4'-C3'-O3')	141	140	143	138	141	130	151	149	145
Glycosylic Conformations (°)*									
(C2'-C1'-N1-C2)	-	137	-	137	-	129	-	136	133
(C2'-C1'-N9-C4)	120	-	137	-	140	-	146	-	133

† The dinucleoside monophosphates and their base-paired mates are illustrated in Fig. 7.

* Glycosidic angles apply to the second base in the dinucleoside monophosphate.

However the TpA component which spans the molecular diad axis has the *trans*, *trans* conformation of smooth B DNAs. In other words, in pleiomic D DNA one in three TpA nucleotides has the conformations seen in smooth B DNAs rather than the special conformations observed for PypPu nucleotides in wrinkled B DNAs. In this sense pleiomic poly d(AT):poly d(AT) can be regarded as intermediate between the smooth and fully wrinkled forms.

PHOSPHATE ORIENTATIONS

The orientations of the phosphate groups are related to the conformations just discussed. The two angles that have been used traditionally to describe phosphate orientations are θ_p and ϕ_p , defined in Table 5. The values of these angles (Table 5) and the illustrations in Fig. 7 show that phosphate groups in nucleotides numbers 3 and 5 with C3'-O3', O3'-P *gauche minus*, *trans* are orientated quite differently from those in nucleotides where the two conformations are *trans*, *trans* or *trans*, *gauche minus*, which have similar orientations.

TABLE 5
Cylindrical polar radii (r_p) of phosphorous atoms and
phosphate orientations (θ_p and ϕ_p) in pleiomic,
wrinkled, and smooth D DNA.

D DNA	r_p (nm)	θ_p (°)	ϕ_p (°)
Pleiomic			
p 1	0.85	132	100
A 2	0.67	129	101
T 3	0.64	140	52
A 4	0.84	122	106
T 5	0.93	149	68
A 6	1.03	118	96
T			
Wrinkled			
p 1	0.81	141	60
I 2	0.83	148	85
p			
C			
Smooth			
p	0.85	131	100
N			

The θ_p are the angles between the helix axis and the O1...O2 lines. The ϕ_p are the angles between the helix axis and the bisectors of the O1-P-O2 angles.

MORPHOLOGY

Pleiomic poly d(AT):poly d(AT) is the first fibrous polymer to be analysed in which the complexity of the asymmetric unit is of the same order as in single crystal studies of DNA oligonucleotides (e.g. 8, 15). It is now possible to explore in a polymer the putative effects of sequence on the conformations and morphologies of B-type DNAs.

Of particular interest are the positions and orientations of the base pairs and the corrugations of the backbone. The orientations of the base pairs relative to the helix axis are described by the tilt and roll angles (Table 6). The relative roll angle measures the roll between successive base pairs. The variations of relative roll angles and turn angles between successive base pairs are also illustrated in Fig. 6.

Figure 5: (a) Illustration of parameters defining base pair position and orientation. The filled circle denotes the z axis (and the helix axis) which is directed towards the reader. In this figure, both d and s are positive. The distance A-N is equal to 0.107 nm. (b) Illustration of the effect of base pair shear. The lower base pair (with open bonds) is positioned relative to the helix axis (filled circle) as in smooth D DNA. The upper base pair (with solid bonds) corresponds to base pair 2 (and also 5) in pleiomic DNA which has the greatest shear.

The positions of base pairs can be described (Table 6, Fig. 5a) in terms of radial shift (d) and lateral shear (s) with respect to the helix axis. The base stacks (Fig. 7) illustrate that variations in these parameters (which are single-valued in smooth DNA) are exploited to maintain base-stacking in the presence of the excursions of the molecule about the helix axis. In order to identify the different phosphates and dinucleoside monophosphates, the phosphates within an asymmetric unit are labelled 1 to 6 in pleiomic DNA and 1 and 2 in wrinkled DNA (Tables 4, 5, 6 and Fig. 7). In pleiomic D DNA the base-pairs on either side of phosphate 1 are shifted out from the helix axis (and therefore from one another) compared with smooth D DNA and are, compensatingly, sheared (0.09 nm) towards one another to preserve base-stacking. Across phosphate 2 the stacking is enhanced by base-pair 2 having a reduced radial shift (compared with base-pair 1) and being sheared (0.22 nm) towards base-pair 1. Across phosphate 3 there is scarcely any shearing and a small radial shift in base pair 3 towards the helix axis is what is needed to provide good stacking. The stacking across phosphate 4 is similar to that in smooth D DNA, with the diadically-related base pairs scarcely being sheared and having radial shifts only slightly less (0.05 nm) than in the smooth form. Because of the diad axes there are only 4 unique base-pair stacks and therefore the arrangements of phosphates 5 and 6 are analogous to those of phosphates 3 and 2 respectively. It is interesting to note (Fig. 7) that despite the morphological variability of this molecule, the changes in positions of the base pairs in pleiomic (and also wrinkled)

TABLE 6

Positional and orientational parameters for the base pairs of pleiomic poly d(AT):poly d(AT), wrinkled poly d(IC):poly d(IC), and smooth D DNA.

D DNA	$\bar{d}$ (nm)	s (nm)	T(°)	γ (°)	θ_T (°)	θ_R (°)	$\Delta\theta_R$ (°)	τ (°)
Pleiomic								
p 1							-5.6	49.5
A	-0.39	-0.09	-14.5	20.1	-17.5	2.8		
p 2							5.6	36.5
T	-0.23	-0.22	-27.3	19.1	-15.8	-2.8		
p 3							-6.1	54.4
A	-0.09	-0.01	-21.4	23.3	-18.7	3.3		
p 4							6.6	38.7
T	-0.09	0.01	-21.4	20.0	-18.7	-3.3		
p 5							-6.1	54.4
A	-0.23	0.22	-27.3	22.7	-15.8	2.8		
p 6							5.6	36.5
T	-0.39	0.09	-14.5	18.0	-17.5	-2.8		
Wrinkled								
p 1							-5.7	46.5
I	-0.21	-0.02	-16.1	16.2	-13.4	6.8		
p 2							5.7	43.5
C	-0.21	0.02	-16.1	15.0	-13.4	1.1		
Smooth								
p							0.0	45.0
N	-0.14	0.00	-17.4	15.8	-13.2	0.0		

$\bar{d}$ and s are the radial shift and lateral shear respectively of the base pair as illustrated in Fig. 5a. Positive $\bar{d}$ implies that the helix axis is on the major groove side of the base pair roll axis. Positive shear corresponds to the base pair moving towards the chain directed (5' to 3') along negative z . T is the twist angle between two bases of a pair and γ is the angle between the normal to a single base and the helix axis. T is treated as a conformation angle so that positive T corresponds to the minor groove side of each base twisting in the direction of its sugar phosphate chain. θ_T and θ_R are the tilt and roll angles respectively of a base pair and are defined as the angles between the roll axis and the pseudo-diad respectively, and a plane normal to the helix axis. θ_T is positive if each base of the pair is tilted in the direction of its sugar phosphate chain. θ_R is positive if the minor groove side of the base pair is rolled in the direction of increasing z . $\Delta\theta_R$ is the relative roll angle between adjacent base pairs and is defined as θ_R for the base pair with the larger z -coordinate minus θ_R for the base pair with the lesser z -coordinate. With this convention, positive $\Delta\theta_R$ corresponds to the angle between the base pairs opening up towards the minor groove. τ is the turn angle between adjacent base pairs defined as the angle between the projections of the two base pair pseudo-diad axes onto a plane normal to the helix axis.

Figure 6: Variation of (a) relative roll angle and (b) turn angle between adjacent base pairs in one hexanucleotide of pleiomeric poly d(AT):poly d(AT) (—) and wrinkled poly d(IC):poly d(IC) (- - -). The locations of the diad axes in pleiomeric DNA are indicated.

DNA allow each base stack to be approximately similar to one of the two stacks in smooth D DNA.

In smooth D DNA the extra symmetry elements result in zero shear and a fixed radial shift.

In wrinkled D DNA there is modest shear and a radial shift comparable with the average value in pleiomeric D DNA.

The marked changes in positions of base pairs in pleiomeric poly d(AT):poly d(AT) relative to the smooth and even to the wrinkled version show no obvious correlation with the alternating purine-pyrimidine base sequence. By contrast the relative roll angles (Table 6 and Fig. 6a) oscillate from residue to residue and therefore may be correlated with sequence. These relative roll angles are positive for the ApT steps, negative for the TpA steps and they have the same sense as, and similar values to, the angles observed in wrinkled D DNA (positive for CpI and negative for IpC). Interestingly, these relative roll angles have a similar range of magnitudes ($-6^\circ < \Delta\theta_R < +7^\circ$) to those observed in a crystallized, B-like DNA fragment (17) but are opposite in sign with respect to PupPy and PypPu sequences.

Calladine (16) has proposed that rolling of adjacent base-pairs in the sense observed (17) relieves the steric compression resulting from propeller-twisting of base-pairs in B DNA. However, inspection of the base stacks (Fig. 7) for all the D DNAs shows that there is no systematic overlap between the purines. This difference between D and B DNA appears to be due to the helix axis passing through the minor groove in D DNA rather than

Figure 7: Base stacks (viewed along the helix axis, which is indicated by a filled circle) of (a) pleiometric poly d(AT):poly d(AT); (b) smooth D DNA; and (c) wrinkled poly d(IC):poly d(IC). The top base pair has solid bonds and its hydrogen bonds are shown by dashed lines. The numbers identify the dinucleosides monophosphates (Table 4).

through the centers of the base pairs as in B DNA. In pleiomer D DNA, any compression caused by base pair twisting would tend to be in the major groove for all base steps. Such compression could not be relieved by a systematic variation in roll angle (i.e. $\Delta\theta_R$ negative) since the average relative roll angle for a straight helix must be zero. In fact, in the B-like DNA fragment (8), the variation of relative roll angles may be due to bending of the helix, as much as to compression between base pairs. The relative roll angles observed in wrinkled and in pleiomer D DNA may be due to other interactions -- including attractive stacking.

The rotations between successive nucleotides in pleiomer D DNA (Table 6, Fig. 6b) follow the phosphate conformations (and, approximately, the base sequence); they are larger across the *gauche minus*, *trans* phosphates than across the *trans*, *gauche minus* and *trans*, *trans* phosphates. Inspection of Fig. 7 shows that the orientation of the *gauche minus*, *trans* NpN allows a larger turn angle between the base pairs than do the differently oriented *trans*, *gauche minus* and *trans*, *trans* NpN. The same applies to wrinkled poly d(IC):poly d(IC); a larger turn angle is observed for the *gauche minus*, *trans* than for the *trans*, *trans* NpN.

The gross morphologies of the different D DNAs appear much the same when viewed normal to their helix axes (Fig. 3), however when they are viewed along their helix axes the 3 structures are very distinct. In pleiomer D DNA the less symmetric distribution of nucleotides results in a structure of larger diameter and a more interdigitated packing of molecules in the crystal unit cell (Fig. 2).

Altogether pleiomer D DNA has a more corrugated periphery. The shear of the base pairs produces a large variation in phosphate radii from the helix axis (Table 5) so that phosphates 2 and 3 are shifted towards the helix axis phosphates 5 and 6 are shifted away from the helix axis. Particularly noticeable is the one phosphate group (phosphate 6) per hexanucleotide which protrudes towards an adjacent molecule. Densely charged Li^+ cations may form intermolecular cation bridges involving this phosphate group. It is quite possible that this unusual poly d(AT):poly d(AT) structure originates through the formation of such cation bridges. The molecular distortions involved may not be possible with a period greater than 6 nucleotides.

DISCUSSION

Whatever the provenance of this pleiomer D DNA, its analysis permits us to determine what structural features emerge in DNA polymer molecules when

they are released from the strait-jacket of a highly symmetrical crystal. Comparison with the more symmetrical wrinkled D DNA shows us what features persist even when the environment is restrictive. Comparison with the B-like oligomer structure (17) may tell us the extent to which features observed in single crystals of DNA fragments may persist in polymers.

The most substantial differences between pleiomic and smooth DNA are in the positions of the base pairs with respect both to the helix axis and to adjacent base-pairs. These differences, several tenths of a nanometer in magnitude, are quite large in molecular terms. The orientations of the bases are different also: not only is there up to 3° of roll -- an excursion not permitted in smooth DNAs -- the base pairs are all tilted more and, with one exception, propeller-twisted more.

The variations in base positions result in phosphorus atoms being distant from the helix axis as little as 0.64 nm and as much as 1.03 nm (Table 5). Despite this, the orientations of the phosphate groups in the ApT nucleotides are remarkably like those in smooth D DNA. As in wrinkled B and D DNA it is the PypPu nucleotides that can have the altered phosphate orientations that are associated with the *trans* to *gauche minus* change in the C3'-O3' conformation. However, what the structure of this pleiomic D DNA reveals is that not every PypPu nucleotide is obliged to express this feature. As we have seen, one TpA nucleotide in three has a *trans* conformation at C3'-O3' and a phosphate orientation reminiscent of those in the ApT nucleotides and in every nucleotide in smooth D DNA.

We see in the wrinkled (polydinucleotide) B form of poly d(GC):poly d(GC) and in the wrinkled (polydinucleotide) D forms of poly d(IC):poly d(IC) and poly d(AT):poly d(AT) and in the pleiomic (polyhexanucleotide) D form of poly d(AT):poly d(AT) that the presence of an altered phosphate orientation reveals PypPu sequences. We now see that the absence of an altered phosphate orientation need not imply a PupPy sequence. On the other hand characteristic changes in relative roll angle (positive for PupPy and negative for PypPu) are found in both wrinkled and pleiomic DNA whether or not there are altered nucleotide conformations and phosphate orientations.

It is interesting also that the range of relative roll angles (13°) and the variation (18°) in the rotations between successive nucleotides are just as great in this high polymer structure as they are in oligomers (17) containing 4 to 12 nucleotides.

ACKNOWLEDGEMENTS

We thank Diane Neeb for typing the text, Bill Boyle for photography, and NIH and NCI for grants (GM17371 and CA30234).

*To whom correspondence should be sent

REFERENCES

1. Arnott, S. (1976) in *Organisation and Expression of Chromosomes*, Allfrey, V. G., Bautz, E.K.F., McCarthy, B. J., Schimke, R. T., and Tissieres, A., Eds., pp. 209-222, Dahlem Konferenzen, Berlin.
2. Leslie, A.G.W., Arnott, S., Chandrasekaran, R. and Ratliff, R. L. (1980) *J. Mol. Biol.* 143, 49-72.
3. Arnott, S. and Chandrasekaran, R. (1981) in *Biomolecular Stereodynamics*, Sarma, R. H. Ed., Vol. I (II), pp. 99-122, Academic Press, New York.
4. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A.G.W., and Ratliff, R. L. (1980) *Nature* 283, 743-745.
5. Arnott, S., Chandrasekaran, R., Hall, I. H., and Puigjaner, L. C. (1983) *Nucl. Acids Res.* 11, 4141-4155.
6. Wang, A.H.J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., vander Marel, G. and Rich A. (1979) *Nature* 282, 680-686.
7. Arnott, S., Chandrasekaran, R., Puigjaner, L. C., Walker, J. K., Hall, I. H., Birdsall, D. L. and Ratliff, R. L. (1983) *Nucl. Acids Res.* 11, 1457-1474.
8. Wing, R. M., Drew, H. R., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R. E. (1980) *Nature* 287, 755-758.
9. Arnott, S., Chandrasekaran, R., Banerjee, A. K., He, R.-G., Walker, J. K. (1983) *J. Biomol. Struct. Dynam.* 1, 437-452.
10. Selsing, E., Arnott, S. and Ratliff, R. L. (1975) *J. Mol. Biol.* 98, 243-248.
11. Arnott, S. and Wonacott, A. J. (1966) *Polymer* 7, 157-166.
12. Smith, P.J.C. and Arnott, S. (1978) *Acta Cryst.* A34, 3-11.
13. Hamilton, W. C. (1965) *Acta Cryst.* 18, 502-510.
14. Arnott, S. (1980) in *Fiber Diffraction Methods*, French, A. D. and Gardner, K. H. Eds., Vol 141, pp. 1-30, American Chemical Society, Washington, D.C.
15. Shakked, Z., Rabinovich, D., Cruse, W.B.T., Egert, E., Kennard, O., Sala, G., Salisbury, S. A. and Viswamitra, M. A. (1981) *Proc. Roy. Soc.* B213, 479-487.
16. Calladine, C. R. (1982) *J. Mol. Biol.* 161, 343-352.
17. Dickerson, R. E. (1983) *J. Mol. Biol.* 166, 419-441.
18. Arnott, S. and Hukins, D.W.L. (1972) *Biochem. J.* 130, 453-465.