Volume 11 Number 12 1983 Nucleic Acids Research

Heteronomous DNA

Struther Arnott*, R.Chandrasekaran, I.H.Hall and L.C.Puigjaner

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Received 8 February 1983; Revised and Accepted 27 April 1983

ABSTRACT

A fibrous form of poly d(A):poly d(T) has a heteronomous secondary structure which is the first to be confirmed for a polynucleotide duplex: although both chains are 10_1 helices, mutually hydrogen-bonded in the standard (Watson-Crick) fashion, each has a quite different conformation. One chain -- probably poly d(A) -- has C3'-endo-puckered furanose rings characteristic of the A family of polynucleotide secondary structures while the other -- probably poly d(T) -- has the C2'-endo-puckered rings of the B family. Since analogous heteronomous structures could be assumed by DNA-DNA or DNA-RNA duplexes containing more general base sequences the polymorphic range of polynucleotide double-helices may be even greater than we have come to suppose.

INTRODUCTION

Until now all successful models of fibrous polynucleotide duplexes containing Watson-Crick base-pairs have been analogous to the paradigmatic version (1) in having identical, anti-parallel chains. This is to say that they have diad axes perpendicular to their axes of screw symmetry. Diad axes lie midway between successive base-pairs in duplexes which have a dinucleotide as the repeated structural motif such as the left-handed allomorphs of (2 2 6_5) poly d(GC):poly d(GC) and (2 7_6) poly d(As^4T):poly d(As^4T) (2), and the wrinkled, right-handed allomorphs of (2 2 4_1) poly d(AT):poly d(AT) and (2 5_1) poly d(GC):poly d(GC) (3). Duplexes like classical A and B DNA, which have a mononucleotide as the repeated structural motif and more general base sequences, have additional diads in the mean planes of their base-pairs. The symmetry elements in these structures are, of course, only approximate since A:T, T:A, G:C and C:G base-pairs are only quasi-isomorphous.

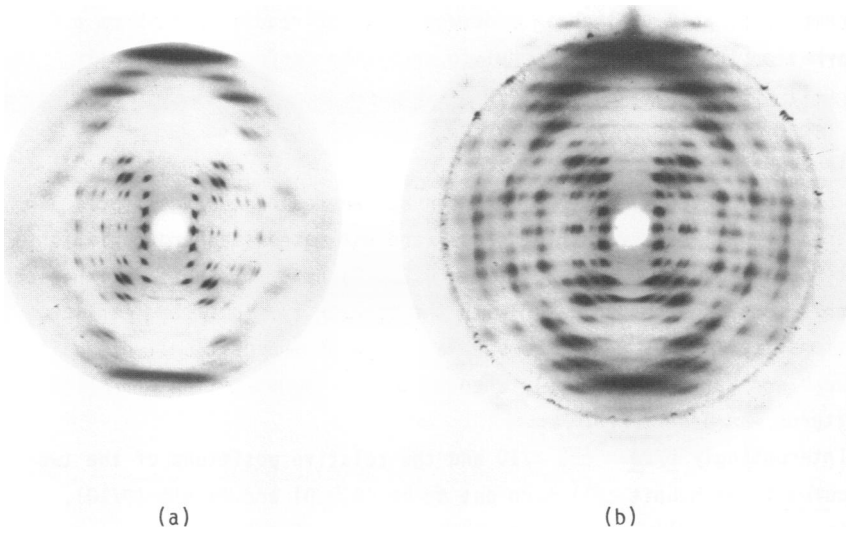
Clearly there is no absolute requirement that polynucleotide chains in multi-stranded complexes be conformationally identical. The triplex structures formed by poly(U):poly(A):poly(U) (4), by

poly d(T):poly d(A):poly d(T) (5), and by poly(I):poly(A):poly(I) (6), illustrate this clearly and one of the models (23) for poly r(A):poly d(T) with different furanose conformations in its two chains demonstrates this even more emphatically. Thus in duplexes where the complementary strands are rather different chemically we might expect that the traditional diadic or quasi-diadic symmetry would sometimes be abandoned. We have found this incontrovertibly to be the case in a polycrystalline, fibrous form of poly d(A):poly d(T). In this instance the information derived from the X-ray diffraction pattern (in the form of unit cell dimensions and the intensities of the Bragg reflexions) is sufficient not only to eliminate more conventional models but also to force the conversion of one of these unsatisfactory models into an acceptable model of a non-traditional kind during a least-squares refinement.

POLY D(A):POLY D(T)

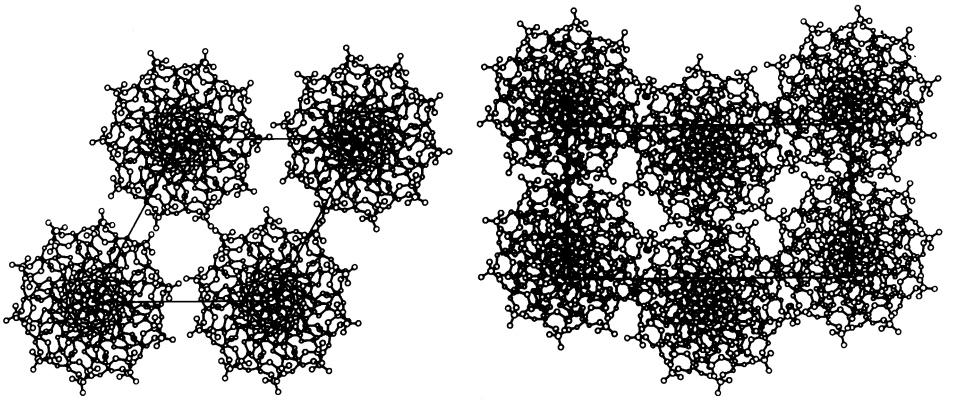
Fibers of poly d(A):poly d(T) were first analysed by Arnott and Selsing (5) who obtained X-ray patterns different from the classical A and B patterns of general sequence DNA. One pattern (α) (Fig. 1a), obtained above 85% relative humidity, indicated 10-fold helical molecules with a pitch (P = 3.29 nm) somewhat less than that (3.37 nm) of the common 10_1 B form observed with DNA from calf thymus, salmon sperm, etc. In the α form the poly d(A):poly d(T) molecules are packed in a screw-disordered, quasi-trigonal fashion with long axes 2.28 nm apart (Fig. 2a). A second (β) pattern (Fig. 1b), obtained below 77% r.h., indicated apparently isomorphous 10-fold helical molecules (P = 3.24 nm) packed in a fully crystalline fashion.

Arnott and Selsing (5) assumed that the α and β forms indeed contained the same allomorph and that this was a minor variant of B DNA. They developed a provisional molecular model with 2 2 10_1 symmetry and fitted it to the Bragg diffraction data from the α form. At that time there was no pressing reason to consider alternative models. Since then, however, a more protean view of the structure of DNA has evolved. This prompted us to review the structure of poly d(A):poly d(T) using the richer data set from the β pattern. When we did this, we discovered that we could devise neither an A nor a B model with the traditional set of diad and screw symmetry elements, whether right-handed or left-handed, which could be accommodated in the unit cell. Only after we removed the constraint that the two chains had to be conformationally identical did we find a model that could be refined to a structure with a sterically unexceptional crystal arrangement and a good fit



(a) (b)

Figure 1: X-ray fiber diffraction patterns of poly d(A):poly d(T): (a) the quasi-hexagonal, screw-disordered, α form obtained at 92% r.h.; (b) the monoclinic, poly-crystalline, β form obtained at 71% r.h.



(a) (b)

Figure 2: The molecular packings of heteronomous duplexes of poly d(A):poly d(T): (a) in the α form, where the molecules have random orientations due to screw-disordering, four neighbouring molecules form a rhombus of side 2.28 nm and acute angle 60°; (b) in the more crystalline β form, quartets of molecules also have a rhombic arrangement but with side 1.87 nm and acute angle 72°.

with the observed intensities of diffraction. Moreover during the process of refinement the poly d(A) chain underwent a major readjustment from a B-like conformation in the starting model to an A-like conformation finally. The poly d(T) chain, however, remained B-like.

CRYSTAL PACKING

The X-ray pattern of the β form can be indexed on the basis of a rectangular unit cell with dimensions (and estimated standard deviations) $\underline{a}$ = 1.865(2) nm, $\underline{b}$ = 3.548(2) nm, $\underline{c}$ = 3.233(3) nm. This unit cell has one dimension doubled when compared with the version (5) published earlier. The additional reflexions which prompted a review of the unit cell dimensions are all weak and were detected only when intensive investigation of well-exposed β patterns was undertaken.

Interestingly $\underline{b}/2\underline{a} = \cos \pi/10$ and the relative positions of the two molecules in each unit cell turn out to be (0,0,0) and ($\underline{a} \sin(\pi/10)$, $\underline{b}/2$ (= $\underline{a} \cos(\pi/10)$), w) implying that quartets of molecules are arranged at the corners of a rhombus of side 1.87 nm and acute angle 2π/5 (Fig. 2b). Analogous arrangements are observed with other DNAs which have 10 or 5-fold screw axes (3).

REFINEMENT OF MODELS AND ARBITRATION BETWEEN COMPETITORS

The best molecular and crystal models of each form considered were obtained by the method of Linked Atom Least Squares (7,8). In this LALS method, the main molecular variables were the 2 × 6 conformation angles associated with the backbones of each dA and dT nucleotide, and the 2 glycosylic angles describing the orientation of each type of base with respect to its attached deoxyribose. Tilts and propeller twists of base-pairs were unrestrained variables. Additional molecular variables were the conformation angles of the furanose rings and their endocyclic bond angles which were allowed to vary in a stiffly elastic fashion. Packing variables were the positional and orientation parameters for each molecule. X-ray variables were a scale factor (K) for the observed structure amplitudes and an isotropic attenuation factor (B) for the calculated amplitudes.

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

$$\Omega = \Sigma\, \omega_{\underline{m}} \Delta F_{\underline{m}}^2 + \Sigma\, k_{\underline{i}} \Delta d_{\underline{i}}^2 + \Sigma\, \lambda_{\underline{h}} G_{\underline{h}}$$
$$= \underline{X} + \underline{C} + \underline{L} \quad , \text{ say.}$$

$\underline{X}$ involves the differences between the observed X-ray amplitudes and those calculated from the model structure. $\underline{C}$ involves both the close, non-bonded interatomic distances, d_j, which are driven towards some acceptably large value ${}_o\underline{d}_j$, and the conformation and bond angles which are elastically restrained. $\underline{L}$ involves constraints G_h 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, $\underline{X}$ or $\underline{C}$ as statistics in Hamilton's test (9,10). Details of this and other strategies which are now quite commonplace in the application of LALS to fiber diffraction analyses have been discussed many times before (e.g. 10,11) and will not be described further.

EMERGENCE OF A HETERONOMOUS MODEL

A contemporary version of the Arnott and Selsing B-like, 2 2 10_1 model with P = 3.29 nm is illustrated in Fig. 3d. A relaxed 10_1 version in which the poly dA and poly dT chains have similar, but not identical, conformations is shown in Fig. 3c. Unlike the 1974 models, the current ones include all the hydrogen atoms. These are not very important X-ray scatterers but contribute significantly to the steric properties of the models. There is no significant difference between these two models insofar as fitting the X-ray data from the semi-crystalline α form is concerned. Interestingly the root mean squared difference in the backbone conformation angles of the two chains in the relaxed model is only $\langle(\Delta\theta)^2\rangle^{1/2} = 3.2°$.

The lateral separation of molecular centers is quite large (2.28 nm) in the α form but only 1.87 nm in the crystalline β form. In this form both the 2 2 10_1 and 10_1 B models lead to extremely short non-bonded interatomic distances between neighbouring molecules. (Distances are considered overshort when they are more than 0.04 nm less than the sum of the van der Waals radii of the two atoms involved. Hydrogen atoms are important.) This steric compression can be relieved somewhat but the resulting structures all have unacceptably high values (R > 0.4) of the crystallographic discrepancy index $R = \Sigma|\Delta\underline{F}_m|/\Sigma_o\underline{F}_m$. The progress of the least-squares refinements was oscillatory and convergence slow as would be expected for severely flawed models.

Persistent refinement of one 10_1 B model eventually produced a structure (Fig. 3a, Table 1) free of overshort interatomic distances within and between molecules, and possessed of a creditably low value of R (0.30) for the 113 measurable reflexions (Table 2). The crucial event leading to

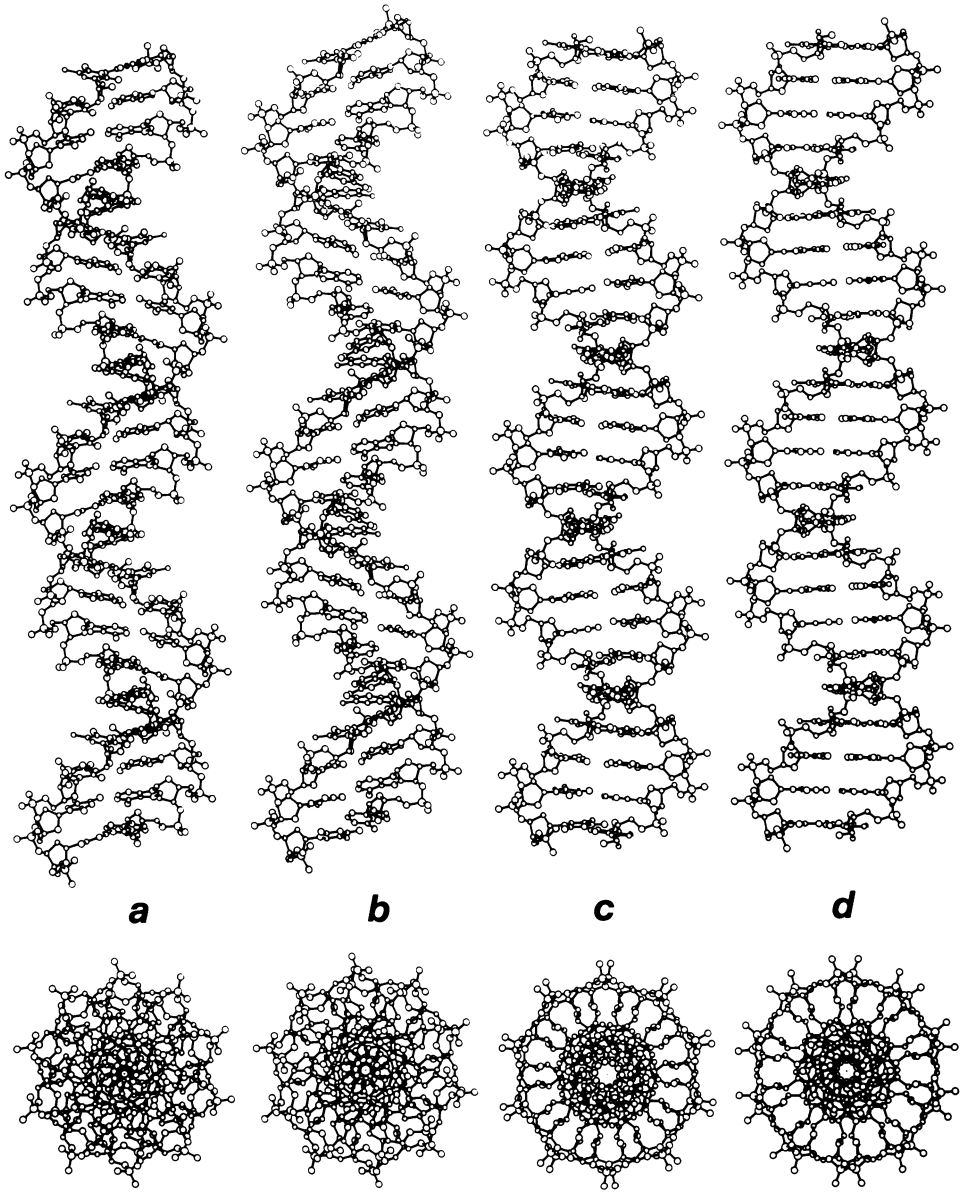


Figure 3: The morphology of the heteronomous and diadically symmetric duplexes of poly d(A):poly d(T). Views normal to the helix axes are shown at top and the corresponding helix axis projections are underneath. (a) and (b) show similar heteronomous models (each with A and B chains) developed to fit the β and α data respectively. (c) shows a mildly heteronomous model with similar but not identical B-like chains. (d) shows a highly symmetrical, 2 2 10_1 model.

Table 1: Cartesian and cylindrical polar coordinates of the atoms in the molecular asymmetric unit of the β form of poly d(A):poly d(T). Successive nucleotide pairs can be generated by adding multiples of 36° to φ and the same multiple of 0.323 nm to Z.

GROUP	ATOM	X(NM)	Y(NM)	Z(NM)	R(NM)	PHI(DEG)
PHOSPHATE	P	-0.1484	0.7754	0.0675	0.7895	100.84
	O1	-0.2482	0.6836	0.0081	0.7273	109.95
	O2	-0.1428	0.9115	0.0095	0.9226	98.90
	O3	-0.1719	0.7853	0.2254	0.8039	102.34
	O5	-0.0033	0.7084	0.0612	0.7084	90.27
SUGAR	C1	0.2533	0.4571	0.1345	0.5226	61.01
	C2	0.3335	0.5143	0.0180	0.6130	57.04
	C3	0.2540	0.6389	-0.0204	0.6876	68.32
	C4	0.2299	0.6866	0.1229	0.7241	71.49
	C5	0.1078	0.7764	0.1225	0.7839	82.10
	O4	0.2015	0.5692	0.2044	0.6038	70.50
	H1	0.3203	0.4006	0.2009	0.5129	51.35
	H12	0.3316	0.4425	-0.0652	0.5529	53.15
	H22	0.4330	0.5441	0.0545	0.6953	51.49
	H3	0.1576	0.6155	-0.0679	0.6353	75.64
	H4	0.3151	0.7407	0.1664	0.8050	66.95
	H15	0.1301	0.8684	0.0664	0.8780	81.48
	H25	0.0817	0.8033	0.2260	0.8075	84.19
ADENINE	N1	0.1078	-0.0211	0.0156	0.1098	-11.09
	C2	0.2248	0.0307	0.0526	0.2268	7.77
	N3	0.2549	0.1553	0.0817	0.2985	31.35
	C4	0.1466	0.2349	0.0699	0.2769	58.03
	C5	0.0209	0.1965	0.0331	0.1976	83.94
	C6	0.0013	0.0605	0.0045	0.0605	88.79
	N6	-0.1168	0.0089	-0.0327	0.1171	-184.37
	N7	-0.0648	0.3057	0.0319	0.3125	101.96
	C8	0.0106	0.4051	0.0674	0.4052	88.49
	N9	0.1402	0.3699	0.0921	0.3956	69.24
	H2	0.3004	-0.0344	0.0592	0.3023	-6.54
	H16	-0.1248	-0.0889	-0.0519	0.1533	-144.54
	H26	-0.1968	0.0683	-0.0413	0.2083	-199.14
	H8	-0.0230	0.4989	0.0759	0.4994	92.64
PHOSPHATE	P	0.0595	-0.9691	-0.3494	0.9709	-86.49
	O1	0.1263	-1.0994	-0.3710	1.1066	-83.45
	O2	-0.0571	-0.9702	-0.2582	0.9719	-93.37
	O3	0.0168	-0.9066	-0.4903	0.9067	-88.94
	O5	0.1665	-0.8618	-0.2982	0.8777	-79.06
SUGAR	C1	0.3095	-0.5780	-0.1407	0.6556	-61.83
	C2	0.3204	-0.7193	-0.0845	0.7874	-65.93
	C3	0.4163	-0.7763	-0.1889	0.8808	-61.80
	C4	0.3614	-0.7196	-0.3199	0.8053	-63.33
	C5	0.2605	-0.8066	-0.3923	0.8476	-72.10
	O4	0.2996	-0.5935	-0.2814	0.6648	-63.22
	H1	0.3995	-0.5208	-0.1138	0.6564	-52.50
	H12	0.3679	-0.7158	0.0147	0.8048	-62.80
	H22	0.2223	-0.7687	-0.0902	0.8002	-73.87
	H3	0.4083	-0.8860	-0.1902	0.9755	-65.26
	H4	0.4449	-0.6962	-0.3876	0.8262	-57.42
	H15	0.2067	-0.7461	-0.4668	0.7742	-74.52
	H25	0.3130	-0.8884	-0.4440	0.9419	-70.59
THYMINE	N1	0.1884	-0.5044	-0.0947	0.5384	-69.52
	C2	0.2002	-0.3685	-0.0783	0.4194	-61.49
	O2	0.3039	-0.3080	-0.0995	0.4326	-45.39
	N3	0.0859	-0.3038	-0.0356	0.3157	-74.20
	C4	-0.0361	-0.3623	-0.0087	0.3640	-95.70
	O4	-0.1315	-0.2938	0.0290	0.3219	-114.11
	C5	-0.0387	-0.5052	-0.0289	0.5067	-94.38
	ME	-0.1678	-0.5768	-0.0019	0.6007	-106.22
	C6	0.0710	-0.5707	-0.0703	0.5751	-82.91
	H3	0.0923	-0.2048	-0.0229	0.2246	-65.74
	H6	0.0669	-0.6697	-0.0836	0.6730	-84.30

Table 2: Structure amplitudes for the β form of poly d(A):poly d(T): (A) observed and (B) calculated values. M denotes a meridional reflexion and N a systematic absence. Amplitudes in parentheses indicate reflexions too weak to be measured. In (A) the estimated threshold value is given. In (B) the calculated values are shown. Such reflexions were included in the refinement only when the calculated amplitudes had values greater than threshold. The spot marked * in (A) could not be densitometered due to severe overlap and high background and hence was not included in the refinement. The value of the attenuation parameter for the calculated amplitudes is B = -0.14 nm^2.

A

N	K	L=0	1	2	3	4	5	6	7	8	9	10	11
0	0	N	N	N	N	N	N	N	N	N	N	N	N
0	1	N	(98)	(79)	(81)	(108)	(118)	(130)	(142)	(154)	(166)		
1	0	462											
0	2	488	888										
1	1	900	888	883	1191	880	220	248	282	808	883		
1	2	213	(140)	373	888	483	277	(225)	287				
0	3	N	(182)	(184)	(190)	(198)	(217)	(236)	(258)	831	842	2884	*
1	3	441	(179)	412	731	818	404	(282)	(283)				
2	0	(187)											
2	1 0	888	251	800	822	883	728	832	482				
2	2 1	383	(222)	(232)	473	487	(283)	838	(328)	2280	(371)	1148	1178
2	3	(240)	(243)										
0	5	N	(251)	(280)	488	(281)	388	1218	1030	1128	807	(427)	(488)
1	5	286	(288)										
2	4	(278)	(278)	(288)	(291)	(318)	833	1181	844	(403)	(428)	(458)	(484)
3	0 1	488	(282)	(300)	(313)	428	888	880	772				
0	8 3	(303)	874	878	(288)	(342)	(381)	(382)	(408)				
2	5 1	887	711	882	488	838	834						
3	3	(324)	(327)	(338)	(347)	(382)	(381)	(402)	838	1228		782	(838)
3	4 0	808	802	873									
2	6	(380)	(382)	(370)	(381)								
1	7	828	817	827	(388)	(408)	(428)	838	484	(488)			(883)
3	5 4	487	(384)	888	(412)	(427)	477	818	388	(818)			(808)
4	2 2	347	884	1240	887	478	838	(484)	378	(834)			
4	3 1	(428)	807	1123	(448)	(481)	714	(800)	(824)	(880)			738
4	4	(444)	818	801	(484)	(478)	(487)	(818)	(848)	(888)			(884)
2	8	(480)	(481)	(488)	(478)	(484)	(812)	(833)	(887)	(883)			
0	9 3	(474)	(478)	(483)	(484)	(808)	(828)	(847)	(871)	(888)			
1	8	838	878	481	488	(812)	(831)	(883)	(877)	(884)			
5	0 1	(802)	(804)	(810)	(821)	(836)	(883)	(874)	(888)	(828)			
4	6 2	(811)											
2	8 3	(818)	843	1088	824	(851)	888	(880)	(814)				
5	3 0	824	781	881	(848)	888	(880)	(801)	488	1188			
1	10	(840)	(842)	(848)		(873)		(812)	(838)	(884)			
5	4 4		(882)	(888)		(883)		(822)	(848)				
3	8 2		884	784		(808)			883	874			
0	11 4		(887)	(802)		(827)							
1	11 5		880					(881)	(708)				
6	0 1		(824)										

B

N	K	L=0	1	2	3	4	5	6	7	8	9	10	11
0	0	(2530)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(881)	(0)
0	1	(0)	(234)	(278)	(34)	(8)	(1)	(2)	(33)	(302)	(284)		
1	0	548											
0	2	827	838										
1	1	734	487	828	908	188	178	308	221	1247	1488		
1	2	124	(271)	828	818	208	303	(280)	231				
0	3	(0)	(133)	(443)	(242)	(100)	(88)	(24)	(131)	848	848	2120	(1882)
1	3	280	(118)	388	478	338	282	(814)	(888)				
2	0	(88)											
2	1 0	888	287	828	317	448	717	747	788				
2	2 1	388	(288)	(388)	383	288	(888)	882	(812)	1710	(883)	1881	1122
2	3	(338)	(348)										
0	5	(0)	(128)	(337)	300	(338)	270	781	880	724	284	(843)	(388)
1	5	47	(314)										
2	4	(188)	(288)	(128)	(481)	(402)	288	808	488	(710)	(830)	(888)	(882)
3	0 1	300	(408)	(188)	(344)	404	337	707	433				
0	8 3	(402)	488	141	(872)	(348)	(408)	(281)	(888)				
2	5 1	487	888	184	311	481	888						
3	3	(298)	(822)	(228)	(228)	(283)	(483)	(482)	832	1888		882	(818)
3	4 0	718	433	382									
2	6	(102)	(248)	(483)	(277)								
1	7	488	488	288	(188)	(782)	(388)	381	211	(823)			(388)
3	5 4	278	(888)	848	(488)	(218)	818	433	231	(387)			(381)
4	2 2	210	838	1201	888	438	888	(820)	242	(401)			
4	3 1	(332)	702	1011	(704)	(817)	873	(842)	(388)	(388)			828
4	4	(438)	837	488	(303)	(418)	(383)	(431)	(270)	(274)			(383)
2	8	(874)	(387)	(883)	(324)	(428)	(328)	(378)	(278)	(388)			
0	9 3	(888)	(880)	(834)	(887)	(883)	(801)	(838)	(808)	(321)			
1	8	488	888	388	188	(481)	(288)	(278)	(284)	(387)			
5	0 1	(388)	(887)	(383)	(128)	(874)	(188)	(178)	(218)	(443)			
4	6 2	(778)											
2	8 3	(487)	823	888	818	(718)	888	(884)	(834)				
5	3 0	82	838	847	(887)	847	(78)	(883)	128	1134			
1	10	(481)	(280)	(413)		(181)		(238)	(280)	(827)			
5	4 4		(488)	(883)		(343)		(277)	(431)				
3	8 2		473	747		(428)			888	421			
0	11 4		(427)	(818)		(338)							
1	11 5		404					(487)	(428)				
6	0 1		(388)										

this felicitious outcome was the change in the puckering of the sugar rings of the poly d(A) chain from C2'-endo to C3'-endo (Table 3).

TESTS OF OTHER MODELS

The most compelling argument for the credibility of the novel structure for poly d(A):poly d(T) is that it evolved spontaneously from a quite

Table 3: The values (and estimated standard deviations) of the molecular parameters for β poly d(A):poly d(T). Those of A and B DNA are given for comparison with dA and dT strands respectively.

Parameters	A-DNA	ApA	TpT	B DNA
Backbone Conformations (°)				
θ(C4'-C3'-O3'-P)	-145	-170(4)	171(4)	-141
θ(C3'-O3'-P-O5')	-78	-59(4)	-121(5)	-157
θ(O3'-P-O5'-C5')	-50	-60(5)	-41(5)	-33
θ(P-O5'-C5'-C4')	172	172(4)	-174(5)	138
θ(O5'-C5'-C4'-C3')	42	54(3)	43(4)	33
θ(C5'-C4'-C3'-O3')	79	81(1)	152(3)	142
Glycosylic Conformations (°)				
θ(C2'-C1'-N9-C4)	83	89(2)	-	141
θ(C2'-C1'-N1-C2)	83	-	147(3)	141
Furanose Conformations (°)				
θ(C4'-O4'-C1'-C2')	8	9(1)	-25(2)	-36
θ(O4'-C1'-C2'-C3')	-32	-35(2)	42(1)	46
θ(C1'-C2'-C3'-C4')	43	43(1)	-42(2)	-38
θ(C2'-C3'-C4'-O4')	-39	-40(1)	29(3)	19
θ(C3'-C4'-O4'-C1')	19	21(1)	-3(3)	10
Furanose Endocyclic Bond Angles (°)				
τ(C4'-O4'-C1')	110	107(2)	109(1)	107
τ(O4'-C1'-C2')	106	106(2)	106(2)	104
τ(C1'-C2'-C3')	100	103(2)	98(1)	99
τ(C2'-C3'-C4')	101	96(2)	103(1)	103
τ(C3'-C4'-O4')	103	108(2)	104(1)	106
Dihedral Angle between Planes in a Base-pair (°)				
Δγ	-6	-29(2)		-13

different structure during the course of an unbiased least-squares refinement designed only to abolish steric compression and at the same time to optimize the fit between the calculated and observed structure amplitudes. In addition, however, the final value of R from the heteronomous model is comparable with that obtained with other highly-hydrated, polycrystalline polynucleotide structures. By itself, of course, this index does not assure us that no other viable structural solutions are possible. Therefore we

generated and tested a variety of alternative models. Fortunately, the number of X-ray data is relatively large. (The field contains 113 measurable reflexions and 150 below threshold reflexions.) Also the unit cell is relatively compact. These two circumstances separately or jointly allow strong discrimination between different possibilities.

Firstly we simply exchanged the A and T bases. This had no effect on the X-ray index R since low resolution X-ray data cannot discriminate between A:T and T:A base-pairs. There were, however, short interstrand contacts between neighbouring, unpaired A and T bases which persisted even after refinement. For this reason we feel the structure is accurately presented with poly d(A) containing C3'-endo furanose rings and poly d(T) C2'-endo.

The two heteronomous molecules in each unit cell are related by a 2_1 axis parallel to b. An alternative model with a 2_1 axis parallel to a has a high R index (0.40) and very short intermolecular contacts.

Left-handed (10_9) heteronomous models are slim and do not give rise to unacceptable steric compression but in no packing arrangement was $R < 0.44$.

Molecular models with parallel or antiparallel chains hydrogen-bonded in the Hoogsteen (12) fashion which is a favorable alternative for A:T are also slim but in no instance produced $R \leqslant 0.42$.

Finally, as we mentioned before, mildly heteronomous models with two similar but unequal A-type chains or two similar but unequal B-type chains have large radii, unacceptable steric compression and large R indices ($\geqslant 0.42$).

Independent support for specific features of our exotic model for poly d(A):poly d(T) comes from Raman spectroscopy: Thomas and Peticolas (24) have shown that from samples of poly d(A):poly d(T) both C2'-endo and C3'-endo marker bands are produced simultaneously. At higher temperatures, but below the melting-point of the duplex, the C3'-endo band disappears.

DISCUSSION OF THE STRUCTURE

Conformations. Table 3 contains the values of the conformational variables in heteronomous poly d(A):poly d(T) and in A and B DNA. It is clear that the unlike chains in the new structure are each rather conventional members of their respective conformational genera.

Morphology. That the heteronomous model indeed has a distribution of atoms quite different from a diad-containing competitor is emphasised by the views (Fig. 3a,d) down the long axes of these alternative structures.

Viewed perpendicular to their long axes, however, the heteronomous 10_1

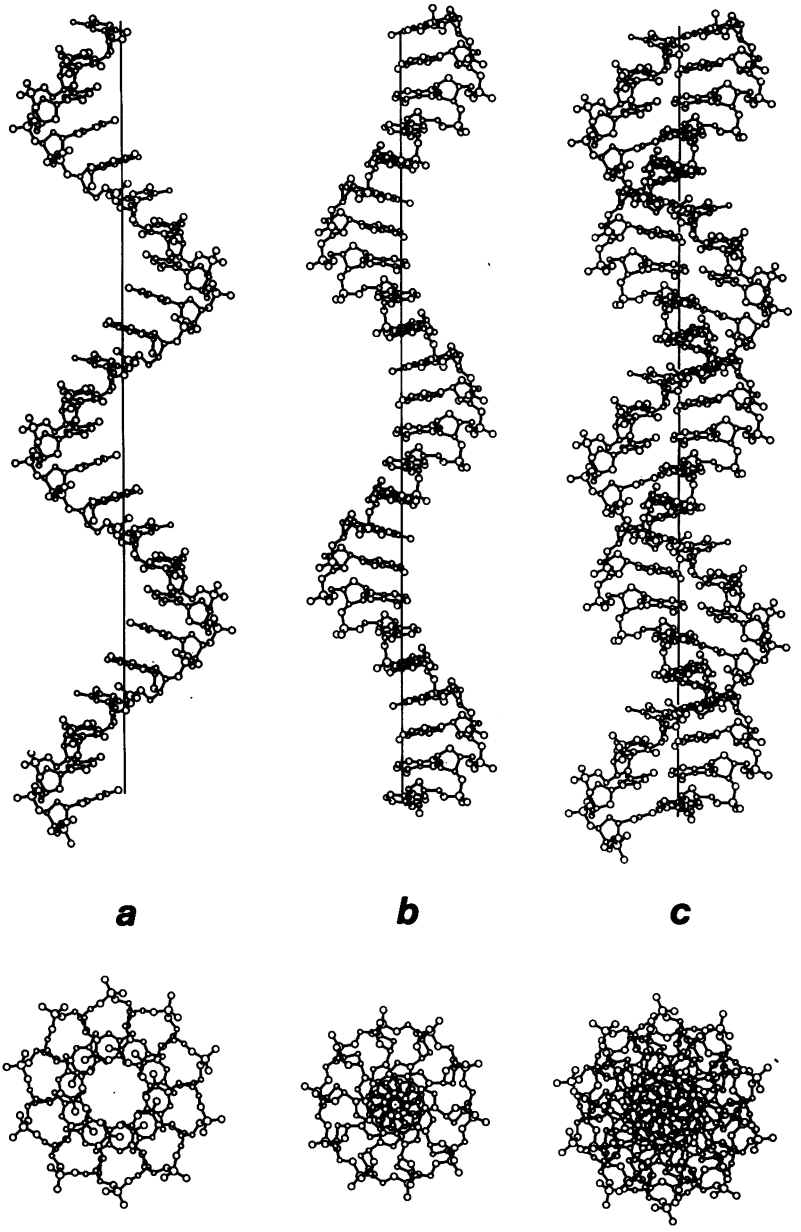


Figure 4: Views normal to (above) and down (below) the vertical helix axes of (a) the B-like poly d(T) strand which mates with (b) the A-like poly d(A) strand to produce (c) the heteronomous duplex.

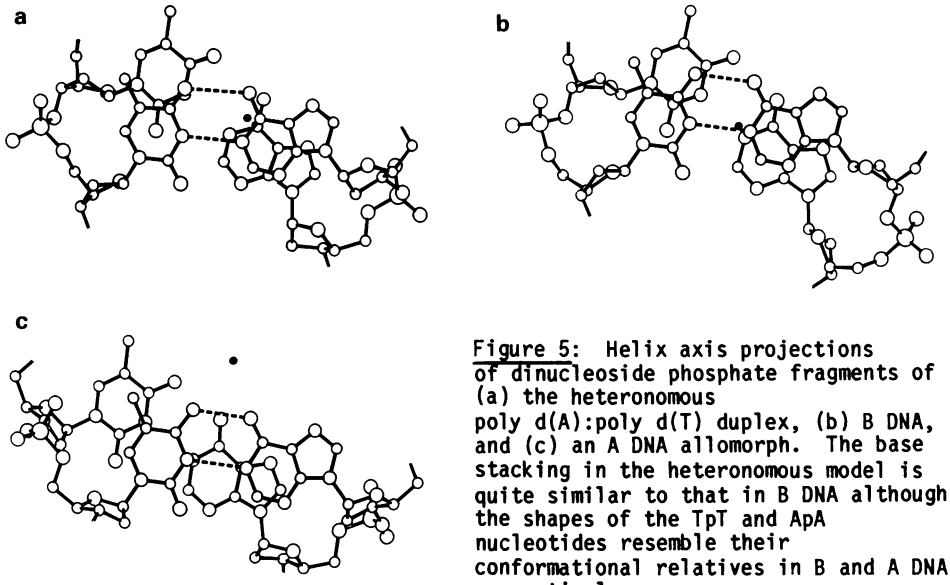


Figure 5: Helix axis projections of dinucleoside phosphate fragments of (a) the heteronomous poly d(A):poly d(T) duplex, (b) B DNA, and (c) an A DNA allomorph. The base stacking in the heteronomous model is quite similar to that in B DNA although the shapes of the TpT and ApA nucleotides resemble their conformational relatives in B and A DNA respectively.

duplexes (Fig. 3a,b) are not unlike the more symmetrical allomorphs. Thus the conceptual problem associated with a complementary duplex composed of one B-like chain (Fig. 4a) and one A-like chain (Fig. 4b) is difficult only for those who persist in thinking of reconciling chains from the classical A and B DNA structures which have quite different symmetries, pitches and base-tilts.

In Fig. 5 there are close-up views of dinucleoside phosphate fragments from the heteronomous poly d(A):poly d(T) duplex and corresponding sequences in B DNA and an A-like allomorph (5) with a similar axial translation per nucleotide (h = 0.33 nm). These show that the shape of d(ApA) is pretty much the same in poly d(A):poly d(T) as it is in A DNA. The same is true of d(TpT) in poly d(A):poly d(T) and in B DNA. What are rather different, of course, are the positions and arrangements of these residues with respect to the helix axes. The differences of radii are apparent in Fig. 5. Not apparent are the differences in orientation. In heteronomous duplexes the normals to the base planes do not need to make equal angles (γ) with the common helix axis and in poly d(A):poly d(T) γ_A = 16.5° is somewhat different from γ_T = 18.0°. In addition the dihedral angle between the planes of paired bases is quite high (29°) in keeping with the trend observed for A:T pairs in oligomers (13) and polymers (3): in oligomers the base-

reaches 27°; in poly d(AT):poly d(AT) it is 21° compared with 14° in poly d(GC):poly d(GC).

Base Stacks. The projections of the various stacked bases in Fig. 5 represent views down the helix axes. It is quite evident that the base-stacking in both chains of poly d(A):poly d(T) resembles that of the same sequences in B DNA but is quite different from what occurs in A-like DNAs.

α and β poly d(A):poly d(T). Our conclusion that β poly d(A):poly d(T) contains heteronomous duplexes may apply also to the α form. We have tested a variety of molecular models in the α system which is (as we implied earlier) less discriminating than β on two counts: its spacious unit cell can accommodate different kinds of duplex more or less equally well; it generates fewer Bragg X-ray reflexions. A B-like, 2 2 10_1 model (Fig. 3d), a relaxed version of this with similar but not identical 10_1 chains (Fig. 3c), and an emphatically heteronomous model with one B-like and one A-like chain (Fig. 3b) provide, after refinement, X-ray discrepancy indexes R = 0.43, 0.43, 0.33 respectively. The first value of R is larger than Arnott and Selsing obtained in 1974 because current modelling technology includes H atoms on the polynucleotide chains and insists on models with less steric compression. It is true also that steric compression within the heteronomous model is markedly less than is found even in the up-dated B-like models (<C> = 1.4 in the former but <C> = 2.7 in the latter). Although X-ray and steric consideration both favour heteronomous 10_1 over 2 2 10_1 poly d(A):poly d(T) molecular models in the more hydrated α system, as well as in the β system, the number of data in the former are not sufficient to make the conclusion statistically significant.

IMPLICATIONS

Certainly the persistently anomalous behaviour of poly d(A):poly d(T) in solution is compatible with its having usually an unconventional structure.

For example, it is possible to predict the ultraviolet circular dichroism spectrum of a DNA of complex sequence by making an appropriate linear combination of CD spectra of DNAs of repeated simple sequences that are fragments of the complex sequence (14). A prerequisite for success is that all the DNAs involved are isomorphous. Interestingly, the measured CD spectrum of poly d(A):poly d(T) (15) has been shown to be quite anomalous in this connection (16). The predicted spectrum which would form a self-consistent set with those of other DNAs is dramatically different from the measured version (16).

Another quite different experiment, involving incomplete scission of poly d(A):poly d(T) molecules laid down on mica or calcite surfaces, indicates that the pitch of this polymer is significantly different from that of typical DNAs (17).

Further, it has been shown (18) that nucleosomes will not form over a sufficiently (e.g. 80 nucleotide) long segment of poly d(A):poly d(T) in a recombinant DNA molecule. This implies not only that the segment has a structure substantially different from general sequence DNA but also that the special structure is sufficiently robust to resist the homogenising effect of flanking sequences. Interestingly, however, smaller (e.g. 20 nucleotide) segments can be accommodated in nucleosomes (18) although poly d(A):poly d(T) tracts even of this size appear to retain their idiosyncratic secondary structure (19,20). This may imply that such tracts contrive to lie between the sites of closest histone-DNA interaction (17,21) and could therefore play a role in phasing nucleosomes on DNA. Obviously any unconventional DNA secondary structure could produce the phenomena just discussed but it would be surprising if they did not result from heteronomous duplexes resembling (if not identical to) the ones we have been describing.

Monotonous $d(A)_n$:$d(T)_n$ tracts need not, of course, be the only sequences with these properties: clearly it is not unthinkable that any oligo d(Pu):oligo d(Py) segment, or any fragment that had approximately such a sequence, might serve equally well.

As mentioned above, Raman spectroscopy (24) has indicated that the precise manner in which poly d(A):poly d(T) may differ from orthodox DNAs may be its simultaneous accommodation of nucleotides with C2'-endo and with C3'-endo furanose rings.

Heteronomous duplexes are not confined to polynucleotides with only A:T base pairs. Patterns similar to those of poly d(A):poly d(T) have been observed also with poly d(I):poly d(C) and with poly d(AI):poly d(CT) (25).

DNA-RNA hybrid duplexes are also prime candidates to form heteronomous structures. Although many such hybrids are known to be essentially isomorphous with well-characterized RNA-RNA duplexes which do have diad axes, some have been thought to have heteronomous structures: poly d(I):poly (C) which has a 10-fold helix of pitch P = 3.1 nm (22); poly (A):poly d(T) which has a pitch P = 3.4 nm and is thought by some (23) to be a 10-fold helix but by us to be an 11-fold helix isomorphous with poly d(A):poly (U) (4) in which case the structure is unlikely to be heteronomous. There is a need,

therefore, for more intensive investigations of the polymorphic range of DNA-RNA hybrids to determine how commonly heteronomous structures occur.

ACKNOWLEDGEMENTS

We thank Carol Jacobson for typing the text, Bill Boyle for photography and the NIH for a grant (GM17371).

*To whom correspondence should be sent

REFERENCES

1. Watson, J.D. and Crick, F.H.C. (1953) Nature 171,737-738.
2. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A.G.W. and Ratliff, R. L. (1980) Nature 283,743-745.
3. 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.
4. Arnott, S. and Bond, P. J. (1973) Nature New Biology 244,99-101.
5. Arnott, S. and Selsing, E. (1974) J. Mol. Biol. 88,509-521.
6. Arnott, S. and Bond, P. J. (1973) Science 181,68-69.
7. Arnott, S. and Wonacott, A. J. (1966) Polymer 7,157-166.
8. Smith, P.J.C. and Arnott, S. (1978) Acta Cryst. A34,3-11.
9. Hamilton, W. C. (1965) Acta Cryst. 18,502-510.
10. 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.
11. Cael, J. J., Winter, W. T. and Arnott, S. (1978) J. Mol. Biol. 125,21-42.
12. Hoogsteen, K. (1963) Acta Cryst. 16,907-916.
13. Fratini, A. V., Kopka, M. L., Drew, H. R. and Dickerson, R. E. (1982) J. Biol. Chem. 257,14686-14707.
14. Gray, D. M. and Tinoco, I., Jr. (1970) Biopolymers 9,223-244.
15. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E. and Cantor, C. R. (1970) J. Mol. Biol. 54,465-497.
16. Arnott, S. (1975) Nucl. Acids Res. 2,1493-1502.
17. Klug, A., Rhodes, D., Smith, J., Finch, J. T. and Thomas, J. O. (1980) Nature 287,509-516.
18. Kunkel, G. R. and Martinson, H. G. (1981) Nucl. Acids Res. 9,6869-6888.
19. Strauss, F., Gaillard, C. and Prunell, A. (1981) Eur. J. Biochem. 118,215-222.
20. Peck, L. J. and Wang, J. C. (1981) Nature 292,375-378.
21. Finch, J. T., Brown, R. S., Rhodes, D., Richmond, T., Rushton, B., Lutter, L. C. and Klug, A. (1981) J. Mol. Biol. 145,757-769.
22. Arnott, S., Chandrasekaran, R., Hall, I. H., Puigjaner, L. C., Walker, J. K. and Wang, Manlin. (1982) in Symposia on Quantitative Biology: Structures of DNA (Proceedings of the Forty-Seventh Cold Spring Harbor Symposium on Quantitative Biology: Structures of DNA, Cold Spring Harbor Laboratory Publications), Vol XLVII, pp. 53-65.
23. Zimmerman, S. B. and Pheiffer, B. H. (1981) Proc. Natl. Acad. Sci. U.S.A. 78,78-82.
24. Thomas, G. A. and Peticolas, W. L. (1983) J. Am. Chem. Soc. 105,993-996.
25. Leslie, A.G.W., Arnott, S., Chandrasekaran, R. and Ratliff, R. L. (1980) J. Mol. Biol. 143,49-72.