The Dimensions and Shapes of the Furanose Rings in Nucleic Acids

By STRUTHER ARNOTT and D. W. L. HUKINS

Department of Biological Sciences, Purdue University, Lafayette, Ind. 47907, U.S.A.

(Received 30 May 1972)

A survey was made of the geometry of furanose rings in β -nucleotides and β -nucleosides (as monomers related to nucleic acids) for which structures have been determined by X-ray crystallography. Mean values, and estimated standard deviations from them, were calculated for bond-lengths, bond-angles and conformation-angles. For parameters with values dependent on ring-puckering, separate calculations were made for each ring type. (The rings are puckered in one of three conformations: C-2- or C-3-*endo* or C-3-*exo*; C-2-*exo* has not been observed.) The results were used to compute standard furanose rings with C-2-*endo*, C-3-*endo* and C-3-*exo* conformations for use in nucleic acid molecular model-building. The survey also showed that the only other conformation-angle in nucleotides dependent on the furanose ring conformation corresponds to the relative orientation of the purine (or pyrimidine) base and the ring.

Molecules of biological polymers usually do not form crystalline arrays of sufficient extent and perfection for them to yield many X-ray-diffraction signals corresponding to periodicities less than 2.0 Å (0.2 nm). Polynucleotides are no exception, whether they are like DNA, elongated molecules in polycrystalline fibres, or like tRNA, folded molecules in single crystals. As a consequence, fine molecular structural details for polynucleotide chains cannot be obtained from X-ray-diffraction analyses unless the diffraction data from the polymer are supplemented by stereochemical information in the form of expected bond-lengths, bond-angles and sometimes even conformation-angles (rotations about bonds). X-ray analyses of fibrous polynucleotides by linked-atom least-squares methods (where the bond-lengths, bond-angles and sugar ring conformations are assumed to be fixed at the mean values observed in accurate, high-resolution X-ray-crystallographic studies of monomers) provide atomic positions with a precision of a few tenths of an Å and experimental values for chain conformation-angles in the polymers with estimated standard deviations of several degrees. The correctness of the results (as distinct from their precision) depends on the error introduced by assignment of fixed values to certain features of the polymers.

The furanose rings of nucleic acids are either β -D-ribose (ribose) or β -D-erythropentose (deoxyribose). In early molecular model-building studies of DNA the deoxyribose ring was assumed to be planar (Crick & Watson, 1954). Spencer (1959) predicted that the furanose rings of nucleic acids would be puckered. This prediction was confirmed by subsequent crystallographic studies of nucleotides and

nucleosides and prompted Sundaralingam & Jensen (1965*b*) to define four main conformations in which either C-2 or C-3 (as defined in Fig. 1) is displaced furthest from the plane through C-1, O-5 and C-4 either in the direction of C-5 (*endo*) or away from C-5 (*exo*). The number of X-ray-crystallographic analyses of appropriate, accurately determined nucleotide and nucleoside structures has doubled in the last 3 years (Table 1). We have made our survey so that the new stereochemical information from them can be exploited in the precise methods of analysing X-ray-diffraction data from fibrous polynucleotides (e.g. Arnott *et al.*, 1969) and in X-ray-crystallographic studies of tRNA (Hampel *et al.*, 1968; Blake *et al.*, 1970; Cramer *et al.*, 1970; Kim *et al.*, 1971; Sakurai *et al.*, 1971). In this survey we consider values of bond-lengths, bond-angles and conformation-angles. The mean values of these parameters have been reconciled by a linked-atom least-squares procedure to provide standard furanose rings for use (where appropriate) as fixed stereochemical features in nucleic acid models.

Survey of Furanose Ring Geometry

The structures considered are listed in Table 1. By the criteria of published agreement indices and estimated standard deviations, all have been determined very accurately. Cyclic nucleotides have been omitted since their sugar ring conformations are often distorted by the additional phosphate ester linkage (Watenpaugh *et al.*, 1968; Coulter, 1969; Coulter & Greaves, 1970; Saenger & Eckstein, 1970).

Histograms were prepared for each bond-length,

Table 1. *Sugar ring conformation described by conformation-angles and out-of-plane distances*

Distances d_2 and d_3 of C-2 and C-3 respectively from the plane C-1-O-5-C-4 are positive for an <i>endo</i> displacement ($1 \text{ \AA} = 0.1 \text{ nm}$).									
Structure	Reference	Conformation	τ_0 ($^\circ$)	τ_1 ($^\circ$)	τ_2 ($^\circ$)	τ_3 ($^\circ$)	τ_4 ($^\circ$)	d_2 ($\text{\AA}$)	d_3 ($\text{\AA}$)
Adenosine residue in uridylyl-(3',5')-adenosine phosphate 1	Seeman <i>et al.</i> (1971) and Sussman <i>et al.</i> (1972)	C-3- <i>endo</i>	4.5	-24.3	34.1	-32.4	17.3	-0.01	0.53
Uridine residue in uridylyl-(3',5')-adenosine phosphate 1		C-3- <i>endo</i>	10.7	-29.0	35.0	-30.0	12.8	-0.03	0.55
Adenosine residue in uridylyl-(3',5')-adenosine phosphate 2		C-3- <i>endo</i>	6.8	-23.6	31.1	-27.7	13.0	-0.02	0.48
Uridine residue in uridylyl-(3',5')-adenosine phosphate 2		C-3- <i>endo</i>	4.9	-26.2	35.3	-34.2	18.2	-0.02	0.56
Inosine	Munns & Tollin (1970)	C-3- <i>endo</i>	7.80	-30.62	40.56	-37.13	18.59	-0.20	0.47
4-Thiouridine	Saenger & Scheit (1970)	C-3- <i>endo</i>	-11.27	-14.93	33.11	-40.58	33.32	0.28	0.82
2'-Deoxycytidine hydrochloride	Subramanian & Hunt (1970)	C-3- <i>endo</i>	9.81	-29.51	36.89	-31.85	14.22	-0.24	0.36
Iodouridine I	Rahman & Wilson (1970)	C-3- <i>endo</i>	6.33	-25.45	35.25	-32.25	16.69	-0.16	0.42
Guanosine 5'-phosphate trihydrate	Murayama <i>et al.</i> (1969)	C-3- <i>endo</i>	6.47	-24.94	33.70	-31.42	15.33	-0.17	0.39
Uridine in β -adenosyl-2'- β -uridine 5'-phosphoric acid	Shefter <i>et al.</i> (1969)	C-3- <i>endo</i>	6.92	-28.53	38.20	-35.54	18.19	-0.17	0.46
5-Methyluridine	Hunt & Subramanian (1969)	C-3- <i>endo</i>	2.34	-25.72	38.49	-38.20	22.52	-0.06	0.56
Adenosine 3'-phosphate	Sundaralingam (1966)	C-3- <i>endo</i>	12.99	-19.73	33.89	-36.65	24.82	0.08	0.62
Adenosine (+5-bromouridine)	Haschemeyer & Sobell (1965a)	C-3- <i>endo</i>	5.00	-28.69	38.44	-36.85	20.70	-0.13	0.52
5-Bromouridine (+adenosine)	Haschemeyer & Sobell (1965a)	C-3- <i>endo</i>	-3.82	-21.29	38.23	-41.11	26.21	0.10	0.66
Adenosine 5'-phosphate	Kraut & Jensen (1963)	C-3- <i>endo</i>	4.76	-29.81	42.28	-40.06	22.85	-0.12	0.57
Calcium thymidylate	Trueblood <i>et al.</i> (1961)	C-3- <i>endo</i>	-3.87	-17.55	31.23	-34.81	24.99	0.10	0.60
Puromycin	Sundaralingam & Arora (1969)	C-3- <i>endo</i>	10.4	-31.2	39.9	-34.0	14.9	—	—
Cytidine	Furberg <i>et al.</i> (1965)	C-3- <i>endo</i>	-6.17	-27.99	37.40	-34.74	17.96	-0.16	0.46
Dihydrothymidine	Konnert <i>et al.</i> (1970)	—	-29.41	15.45	3.04	-20.37	31.24	0.71	0.76

3'-O-Acetyladenosine	Rao & Sundaralingam (1970)	C-2-endo	-26.55	37.13	-33.22	18.83	4.83	0.66	0.12
8-Bromoguanosine	Tavale & Sobell (1970)	C-2-endo	-26.69	36.92	-32.57	17.98	5.01	0.65	0.13
8-Bromoadenosine	Tavale & Sobell (1970)	C-2-endo	-23.38	37.66	-37.35	25.33	-1.79	0.59	-0.05
Monosodium inosine 5'-phosphate	Rao & Sundaralingam (1969)	C-2-endo	-21.00	37.20	-38.51	26.75	-3.61	0.53	-0.10
Adenosine in β -adenosyl-2'- β -uridine 5'-phosphoric acid	Shefter <i>et al.</i> (1969)	C-2-endo	-30.49	43.51	-37.84	20.20	6.11	0.74	0.16
Inosine dihydrate 2A	Munns <i>et al.</i> (1970)	C-2-endo	-20.92	31.79	-34.58	21.98	-1.11	0.54	-0.03
Inosine dihydrate 2B	Munns <i>et al.</i> (1970)	C-2-endo	-32.05	46.39	-34.23	17.44	10.75	0.84	0.27
Inosine dihydrate A	Thewalt <i>et al.</i> (1970)	C-2-endo	-22.55	36.98	-36.66	24.43	-1.27	0.56	-0.03
Inosine dihydrate B	Thewalt <i>et al.</i> (1970)	C-2-endo	-31.16	41.21	-34.78	17.88	7.98	0.78	0.20
Guanosine A	Thewalt <i>et al.</i> (1970)	C-2-endo	-22.05	34.74	-33.39	21.59	0.17	0.55	0.00
Guanosine B	Thewalt <i>et al.</i> (1970)	C-2-endo	-38.52	44.24	-32.41	11.08	16.98	0.92	0.43
Iodouridine II	Rahman & Wilson (1970)	C-2-endo	-16.87	36.45	-41.52	32.12	-9.58	0.43	-0.25
5-Bromouridine (+dimethyl sulphoxide)	Iball <i>et al.</i> (1968)	C-2-endo	-19.34	36.92	-40.13	29.00	-6.39	0.47	-0.17
Barium uridine phosphate	Shefter & Trueblood (1965)	C-2-endo	-15.82	31.48	-33.01	23.81	5.18	0.41	-0.14
Cytidylic acid b (monoclinic)	Bugg & Marsh (1967)	C-2-endo	-19.81	36.60	-38.79	28.11	-5.64	0.49	-0.15
Cytidylic acid b (orthorhombic)	Sundaralingam & Jensen (1965a)	C-2-endo	-15.58	33.38	-37.50	29.15	-9.15	0.39	-0.23
5-Iodo-2'-deoxycytidine	Camerman & Trotter (1965)	C-2-endo	-26.16	37.25	-30.71	17.99	5.12	0.68	0.14
5-Fluoro-2'-deoxy- β -uridine	Harris & MacIntyre (1964)	C-2-endo	-19.45	37.11	-40.31	30.11	-6.80	0.49	-0.17
5-Bromodeoxycytidine + deoxy-guanosine	Haschemeyer & Sobell (1965b)	C-2-endo	-17.61	30.65	-30.94	20.22	-2.14	0.41	-0.06
Deoxyadenosine monohydrate	Watson <i>et al.</i> (1965)	C-3-exo	-28.12	36.72	-31.98	15.80	6.93	0.68	0.18
Thymidine	Young <i>et al.</i> (1969)	C-3-exo	-2.42	23.93	-34.36	33.66	-19.87	0.06	-0.50
Deoxycytidine 5'-phosphate monohydrate	Viswamitra <i>et al.</i> (1971)	C-3-exo	-7.04	27.83	-36.93	33.20	-16.68	0.18	-0.42
			8.69	-12.08	-26.98	32.58	-26.18	-0.22	-0.64

Fig. 1. Definition of a nucleotide unit with its atom numbering scheme and conformation-angles

In (a) N is the nitrogen atom of a purine or pyrimidine base. *cis* conformations have zero conformation-angles and for positive rotation the far end of a bond rotates clockwise relative to the near end; the zero angle for χ has (C-4-N-9) purine or (C-2-N-1) pyrimidine *cis* to C-1-C-2. [Standard base nomenclature is given in *J. Amer. Chem. Soc.* (1960) **82**, 5545]. In (b) are shown the ring conformation-angles that are defined by ring bonds alone (Sundaralingam, 1969). Views of furanose rings edge on to the plane C-1-O-5-C-4: (I) C-2-*endo*; (II) C-3-*endo*; (III) C-2-*exo*; (IV) C-3-*exo*.

bond-angle and conformation-angle (Figs. 2-6). For each parameter attention was paid to possible dependence of the values observed on the substitution at C-2 (ribose or deoxyribose), and to possible correlations among parameters as a result of ring-puckering. Where clearly bimodal distributions were detected separate histograms sometimes are provided for clarity.

The results of the survey are summarized in Table 2, where mean values of stereochemical features and estimated standard deviations from the means are provided. As in the survey of pyranose ring geometry (Arnott & Scott, 1972) we have not felt justified in differentially weighting different results except that in some instances (which we discuss below) extreme outlying values have been excluded from our calculations.

Ring bond-lengths appear to be independent of conformation and vary little from structure to structure (estimated standard deviations <0.002 nm). Exocyclic bond lengths are slightly more variable and some correlation with ring-puckering is observed. For example, a distinctive difference is observed for the mean lengths of C-1-N: 1.465 Å (0.1465 nm) for C-2-*endo* and 1.490 Å (0.1490 nm) for C-3-*endo*.

The ring bond-angles at O-5, C-2 and C-3 are independent of conformation. In any particular structure the angles at C-2 and C-3 may not be equal but there is no correlation with conformation as was formerly asserted by Sundaralingam & Jensen (1965b). Rather it appears that the angles at C-1 and C-4 are conformation-dependent. The former is smaller, 105.6° (1.3), for C-2-*endo* rings and larger, 107.6° (1.0), for C-3-*endo* (and *exo*). The trend is reversed for the angle at C-4, where the corresponding values are 105.5° (1.7) and 104.0° (1.8). The exocyclic bond-angles are somewhat more variable than the ring bond-angles. (Estimated standard deviations from the mean are twice as large.) Only those associated with the pendant oxygen atoms are conformation-dependent. The rule therefore appears to be: at C-2 and C-3 the magnitudes of bond-angles within the ring are not conformation-dependent whereas the exocyclic angles are; at C-1 and C-4 the reverse correlation is observed.

We now discuss the ring conformation-angles in more detail (Table 1). The newer structure determinations reveal the same kind of puckered-ring conformations with C-2 and C-3 the most out-of-plane atoms. The distances d_2 and d_3 by which these are displaced from the C-1-O-5-C-4 plane are given in Table 1. With one exception (dihydrothymidine with d_2 and d_3 both greater than 0.7 Å although only one of the ring conformation-angles, τ_2 , is required to be extremely unusual to achieve this), the pattern of displacement is that one atom is displaced about 0.5 Å (0.05 nm) and the other about 0.2 Å (0.02 nm). We are not able to predict whether the minor dis-

Fig. 2. Histograms of furanose ring bond-lengths

The ranges [0.001 nm (0.01 Å)] are for (a) C-1-C-2, (b) O-5-C-1, (c) C-2-C-3, (d) C-4-O-5, (e) C-3-C-4. ■, C-3-*endo*; □, C-2-*endo*; ▨, C-3-*exo*; ▩, dihydrothymidine. (This key to the histograms also applies to Figs. 3-6 inclusive.)

Fig. 3. Histograms of exocyclic bond-lengths

The ranges [0.001 nm (0.01 Å)] are for (a) C-4-C-5, (b) C-1-N (C-2-*endo*), (c) C-3-O-1, (d) C-1-N (C-3-*endo* and -*exo*), (e) C-2-O-6.

placement would be in the same or the opposite direction to the major displacement in any given structure; nevertheless a high value of the larger displacement seems usually to be accompanied by a smaller displacement of the same sense.

The most frequently observed ring conformations are C-2-*endo* and C-3-*endo*. There are three C-3-*exo* structures. (Interestingly, all are deoxyribose derivatives.) This conformation is similar to C-2-*endo* in four conformation-angles but like C-3-*endo* in τ_0 ; in deoxycytidine 5'-phosphate monohydrate

(Viswamitra *et al.*, 1971) τ_1 is also nearer the C-3-*endo* range. For β -monomers no example of a C-2-*exo* conformation has been observed, but examples do exist for α -monomers (the C-1 epimers) (Hodgkin *et al.*, 1962; Shoemaker *et al.*, 1964; Hawkinson *et al.*, 1970; Rohrer & Sundaralingam, 1970).

The mean values of the ring conformation-angles are given in Table 2. The conformation-angles are dramatically different from the values near $\pm 60^\circ$ that are observed in the six-membered pyranose sugar rings (Arnott & Scott, 1972). Some measure of the

Fig. 4. Internal bond-angles of the furanose ring

The histograms are in 0.5° intervals at (a) O-5, (b) C-1 (C-2-endo), (c) C-1 (excluding C-2-endo), (d) C-2, (e) C-3, (f) C-4 (C-2-endo), (g) C-4 (excluding C-2-endo).

Fig. 5. Histograms, in 1.0° intervals, of bond-angles involving exocyclic atoms

(a) O-5-C-4-C-5; (b) C-3-C-4-C-5; (c) C-4-C-3-O-1 (C-2-endo); (d) C-4-C-3-O-1 (excluding C-2-endo); (e) C-2-C-3-O-1 (C-2-endo + C-3-exo); (f) C-2-C-3-O-1 (excluding C-2-endo and C-3-exo); (g) C-3-C-2-O-6 (C-2-endo); (h) C-3-C-2-O-6 (C-3-endo); (i) C-1-C-2-O-6 (C-2-endo); (j) C-1-C-2-O-6 (C-3-endo); (k) N-C-1-C-2; (l) N-C-1-O-5.

Fig. 6. Histograms for furanose ring conformation-angles in 5° ranges

(a) τ_0 ; (b) τ_1 ; (c) τ_2 ; (d) τ_3 ; (e) τ_4 .

range of experimental values is provided by the estimated standard deviations from the means. The values of the deviations are comparable with those observed by Arnott & Scott (1972) for pyranose rings.

Standard Fixed Ring Conformations

It has been shown that all RNA double helices studied so far have a conformation similar to A-type DNA, in which the ribose ring conformation is C-3-*endo* (Arnott *et al.*, 1968). The main conformational difference between nucleic acids of the A-type and B-type and C-type DNA, is that the furanose ring conformation is altered from C-3-*endo* to C-2-*endo* (or to the rather similar C-3-*exo*) (Arnott, 1970). For most nucleic acid model-building purposes it will neither be convenient (because of the greatly increased number of parameters) nor justifiable (because of the inherent imprecision of the available data) to contemplate variability in the sugar ring conformations except the major transitions between the C-3-*endo* and C-2-*endo* (or C-3-*exo*) conformations. We have therefore prepared 'standard' rings for these three conformational variants. These have bond-lengths identical with the mean values of Table 2, and ring bond- and conformation-angles that deviate minimally, in a weighted, least-squares sense, from the appropriate mean values of Table 2.

To construct these standard rings we define a chain of eight atoms C-1', O-5', C-4', C-3, C-2, C-1, O-5, C-4 (as in Fig. 7), separated by the appropriate, fixed bond-lengths, to allow all ring bond- and conformation-angles to be explicitly described. (This requires three atoms and ϕ_0 to be included twice.) This description also emphasizes that the mean

values of bond-lengths and of ϕ and τ from our survey do not immediately lead to closed rings. To develop closed rings that have $\{\phi\}$ and $\{\tau\}$ deviating minimally from standard values, we resort to the constrained least-squares procedures used extensively by us in linked-atom refinements of helical structures (e.g. Arnott *et al.*, 1969). Here we vary $\{\phi\}$, $\{\tau\}$ and $\{\lambda\}$ to minimize

$$\sum k_p ({}_0D_p - D_p)^2 + \sum \lambda_n G_n$$

where $\{D_p\}$ are the bond- and conformation-angles, whose mean values from the survey are $\{{}_0D_p\}$ with estimated standard deviations $\{k_p^{-1}\}$, and $\{\lambda_n\}$ are (initially undetermined) Lagrange Multipliers (Dantzig, 1963) associated with the nine constraint relationships, $G_n = 0$, between co-ordinates of atoms in triangles C-4-O-5-C-1 and C-4'-O-5'-C-1', that have to be satisfied for closed rings. (Since ϕ_0 and ϕ_0' are variables but the bond lengths are not, seven of the nine G_n are sufficient.)

It should be noted that the ring-closing conditions decrease the number of degrees of conformational freedom by seven and that therefore there will, in general, be no solution without varying the bond-angles as well as the five conformation-angles. With both kinds of variables there are $11 - 7 = 4$ degrees of freedom. Although necessary for our present purposes, it should therefore be clear that variation in a continuous manner of each furanose ring conformation in a non-helical polynucleotide would be a complex and computationally expensive procedure.

In the standard rings (Table 3) no bond-angle is more than 0.6° different from the mean values of the survey; no conformation-angle is more than 2.3°

Table 2. *Summary of the results of the stereochemical survey of furanose rings*

Parameter values which are omitted are those that appear anomalous in the histograms (1 Å = 0.1 nm).

	Parameter	Mean value (Å)	Estimated standard deviation (Å)	Comments
Ring bond-lengths	C-4-O-5	1.457	0.020	Omitting 5-bromodeoxycytidine
	O-5-C-1	1.419	0.023	Omitting adenosine in a complex with 5-bromouridine and in β -adenosyl-2'- β -uridine 5'-phosphoric acid
	C-1-C-2	1.525	0.017	Omitting deoxyguanosine, 5-iodo-2'-deoxycytidine and inosine dihydrate 2B
	C-2-C-3	1.528	0.019	Omitting deoxyguanosine, 5-bromouridine (+adenosine) and inosine dihydrate 2A
	C-3-C-4	1.529	0.021	Omitting 5-iodo-2'-deoxycytidine
Pendant bond-lengths	C-4-C-5	1.516	0.022	Omitting iodouridine I and bromodeoxycytidine
	C-1-N	1.465	0.019	C-2- <i>endo</i> + dihydrothymidine
	C-2-O-6	1.490	0.018	C-3- <i>endo</i> + C-3- <i>exo</i>
	C-3-O-1	1.417	0.019	
Ring bond-angles	C-4-O-5-C-1	1.422	0.021	Omitting inosine dihydrate 2A
	O-5-C-1-C-2	109.7 (°)	1.0 (°)	Omitting inosine dihydrate (2A+2B), iodouridine 1 and guanosine dihydrate B
		105.6	1.3	C-2- <i>endo</i> omitting inosine dihydrate 2A
		107.6	1.0	C-3- <i>endo</i> + C-3- <i>exo</i>
		106.9	3.0	C-3- <i>endo</i> omitting 5-bromouridine (+adenosine) and 5'-methylammonium-5'-deoxyadenosine
	O-6-C-2-C-3	114.6	2.7	C-2- <i>endo</i> omitting inosine dihydrate 2B
		110.8	1.9	C-3- <i>endo</i>
	C-2-C-3-O-1	109.3	2.3	C-2- <i>endo</i> + C-3- <i>exo</i>
		113.3	2.5	C-3- <i>endo</i>
	O-1-C-3-C-4	108.9	2.7	C-2- <i>endo</i>
		111.9	2.4	C-3- <i>endo</i> + C-3- <i>exo</i> omitting 5-bromouridine (+adenosine)
	C-3-C-4-C-5	115.8	2.7	
	C-5-C-4-O-5	108.9	2.0	
	C-1-C-2-C-3	101.4	2.4	
	C-2-C-3-C-4	102.5	1.9	
	C-3-C-4-O-5	105.5	1.7	C-2- <i>endo</i> omitting inosine dihydrate 2B
		104.0	1.8	C-3- <i>endo</i> omitting 5-bromouridine (+adenosine)
		103.9	1.6	C-3- <i>exo</i>
Exocyclic bond-angles	O-5-C-1-N	108.0	1.8	
	N-C-1-C-2	113.7	2.7	
	C-1-C-2-O-6	112.5	2.4	C-2- <i>endo</i>

Table 2. *Continued*

	Parameter	Mean value (°)	Estimated standard deviation (°)	Comments
Ring conformation- angles	$\tau_0 = \tau[\text{C-4-O-5-C-1-C-2}]$	-23.7	6.1	C-2- <i>endo</i>
		4.1	5.3	C-3- <i>endo</i> omitting 4-thio- uridine
		-0.3	8.1	C-3- <i>exo</i>
	$\tau_1 = \tau[\text{O-5-C-1-C-2-C-3}]$	37.2	4.1	C-2- <i>endo</i>
		-25.4	4.6	C-3- <i>endo</i>
		25.9	2.8	C-3- <i>exo</i> omitting deoxycytidine 5'-phosphate monohydrate
	$\tau_2 = \tau[\text{C-1-C-2-C-3-C-4}]$	-35.5	3.3	C-2- <i>endo</i>
		36.2	3.1	C-3- <i>endo</i>
		-32.8	5.2	C-3- <i>exo</i>
	$\tau_3 = \tau[\text{C-2-C-3-C-4-O-5}]$	22.4	5.6	C-2- <i>endo</i>
		-35.0	3.7	C-3- <i>endo</i>
		33.1	0.5	C-3- <i>exo</i>
	$\tau_4 = \tau[\text{C-3-C-4-O-5-C-1}]$	-0.3	6.2	C-2- <i>endo</i> omitting guanosine dihydrate B
		19.6	5.4	C-3- <i>endo</i>
		-20.9	4.8	C-3- <i>exo</i>

Fig. 7. Eight-atom chain used to describe the furanose ring for building standard models

Atoms C-1, O-5 and C-4 are included twice, the second time as C-1', O-5' and C-4'. Bond-angles are given by ϕ_0 – ϕ_4 and ϕ_0 is included a second time as ϕ'_0 .

different (and most are much less). By the chance of it the standard C-2-*endo* ring has the 'envelope' conformation ($d_3 = 0.00\text{\AA}$, $d_2 = 0.59\text{\AA}$ or 0.059nm) because of the indiscriminate occurrence of the minor displacement (of C-3) on either side of the C-1–O-5–C-4 plane. The standard C-3-*endo* and C-3-*exo* rings have the minor displaced atom on the opposite side of the reference plane from the major. (Co-ordinates of these rings are given in Table 4.)

Furanose ring conformation-angles and other nucleotide chain angles

Fig. 8 shows histograms of the observed values of nucleotide conformation-angles obtained from the

structures listed in Table 1, together with a number of others for which only partial data are available. They represent an up-dated version of the preliminary survey of Arnott & Hukins (1969).

Of particular interest are the values of χ (defined in Fig. 1), which represent the orientation of the bases relative to the sugar rings. The majority of conformations are near 90° (the *anti* conformation) but seven are near -60° (the *syn* conformation). Of these seven, two are from monomers with C-3-*endo* furanose rings, although both have peculiar features: in 5'-methyl ammonium-5'-deoxyadenosine (Saenger, 1970) the conformation appears to be stabilized by a hydrogen bond between adenine N-3 and N⁺-H group bonded to the sugar C-5; 4-thiouridine has an extremely puckered sugar ring. The latter structure is the only *syn* monomer with a pyrimidine base.

The dominant conformation is *anti* but values in this range have a trimodal distribution. The mean values and estimated standard deviations from the mean are 81° (11), 114° (11) and 180° (8) for the three sub-divisions of the *anti* range. The first region has C-3-*endo* sugar conformations and the second two have C-2-*endo*. The three χ values near 180° are observed only for the purine bases hypoxanthine and guanine. The three C-3-*exo* monomers have χ values in the C-3-*endo* range.

The other conformation-angles of a nucleotide chain (except σ) do not show correlations with the furanose conformation. Their mean values in monomers and in polymers are shown in Table 5. The

Table 3. *Conformation-angles and ring bond-angles for the standard furanose rings*

Exocyclic bond-angles and bond-lengths are the same as the mean values of Table 2.

		C-2-endo	C-3-endo	C-3-exo
Conformation-angles (°)	τ_0	-23.5	3.2	-4.0
	τ_1	36.9	-25.6	24.7
	τ_2	-35.7	36.9	-34.8
	τ_3	22.9	-35.9	33.3
	τ_4	0.2	20.8	-18.6
Ring bond-angles (°)	C-4-O-5-C-1	109.8	109.8	110.0
	O-5-C-1-C-2	105.4	107.7	107.8
	C-1-C-2-C-3	101.3	101.6	102.0
	C-2-C-3-C-4	102.9	102.1	102.7
	C-3-C-4-O-5	106.0	104.0	104.4

Table 4. *Co-ordinates of standard furanose rings*

The co-ordinate origin is at O-5 with the x -axis along the extension of the C-1-O-5 bond and the y -axis coplanar with O-5-C-4 and C-1-O-5 within the obtuse angle formed by these two bonds ($1\text{\AA} = 0.1\text{ nm}$).

		x (Å)	y (Å)	z (Å)
C-3-endo	C-1	-1.419	0.000	0.000
	O-5	0.000	0.000	0.000
	C-5	1.761	1.454	0.795
	C-4	0.493	1.371	-0.031
	O-1	-0.608	3.553	0.051
	C-3	-0.688	2.199	0.476
	O-6	-2.018	1.730	-1.496
	C-2	-1.881	1.449	-0.114
	N	-1.879	-0.679	1.244
C-3-exo	C-1	-1.419	0.000	0.000
	O-5	0.000	0.000	0.000
	C-5	1.031	1.782	1.270
	C-4	0.500	1.366	-0.088
	O-1	-0.786	2.110	-2.030
	C-3	-0.688	2.172	-0.613
	C-2	-1.885	1.452	0.008
	N	-1.880	-0.783	1.181
C-2-endo	C-1	-1.419	0.000	0.000
	O-5	0.000	0.000	0.000
	C-5	1.440	1.549	1.175
	C-4	0.492	1.371	0.005
	O-1	-1.085	2.595	-1.329
	C-3	-0.749	2.264	0.012
	O-6	-3.153	1.708	0.254
	C-2	-1.825	1.346	0.591
	N	-1.871	-1.149	0.788

estimated standard deviations are typically two to three times greater than those observed for conformation-angles within the furanose ring.

Nucleic acid model-building

In creating and optimizing precise molecular models an important decision is which parameters should be assigned fixed values and which should be allowed to vary. This is dependent on the precision desired, and, ultimately, on the precision attainable with the available experimental data. Usually a precision of a few tenths of an Å in atomic positions would be desirable so that possibly important intra- and inter-molecular interactions, such as hydrogen bonds, would be identifiable. In such a case possible variation of bond-lengths (estimated standard deviation typically $0.02\text{\AA}$ or 0.002 nm) from the mean values observed in our surveys would certainly not be of concern and these parameters would be assigned fixed values. Assignment of fixed values to ring bond-angles, chain bond-angles, ring conformation-angles and chain conformation-angles might be made with decreasing confidence for each set of parameters in the order listed. Bond-angles are sufficiently invariant for us to conclude that it would rarely be necessary to contemplate treating them as variable parameters. In contrast, the observed variances of χ , ξ , θ , ψ , ϕ and ω make it imperative that these be treated as variables, the values of which should be determined, whenever possible, from experimental results from the system under examination. The situation of ring conformation-angles is somewhat ambiguous. Their smaller variances would justify assigning fixed values to them in most circumstances. Certainly this procedure avoids increasing the conformational variables of a nucleotide residue from 6 to (6+11) with the additional complications of imposing ring-closing constraints etc. Arnott & Scott (1972) have argued that variation of ring conformation-angles (with the inevitable complications) is inescapable for helical homopolysaccharides with pyranose rings. (Otherwise the only conformational variables would be two glycosidic bridge

Fig. 8. Other nucleotide conformation-angles from a survey of mono- and poly- β -nucleotide structures

Histograms above the base line are for polymers and those below it are for monomers: for polymers solid blocks represent structures whose agreement with X-ray-diffraction data has been optimized by a least-squares process, whereas the values in open blocks have not and are presumably less reliable. In (a) monomer ψ values are solid blocks and ϕ values are open blocks. For θ , ξ and ω , in (b), (c) and (d) respectively, all monomer values are solid blocks but for σ in (e) and χ in (f) the key given to Fig. 2 applies.

conformation-angles.) Their argument is not applicable to polynucleotides where there are six chain conformation-angles among which to distribute the errors arising from inaccurate assignment of fixed stereochemical features.

This contention is supported by our recent X-ray-diffraction experiments on B-type DNA (S. Arnott & D. W. L. Hukins, unpublished work). These experiments show that a model with standard C-3-*exo* furanose rings can be refined to give the best fit with the X-ray intensities and at the same time have no dubious stereochemical features. We judge that the best model with furanose rings fixed in the standard

C-2-*endo* conformation is less likely to be correct because many stereochemical features are only marginally acceptable. Nevertheless the fit with the diffraction data provided by this model is not significantly inferior to that of the C-3-*exo* model, showing that an equally good set of nucleotide atomic co-ordinates may be achieved even with an (albeit slightly) incorrect, fixed furanose conformation because of the number of degrees of structural freedom available. No significantly more precise set of atomic co-ordinates could have been obtained by continuous variation of the furanose ring conformation. Nor is it likely that a structure could be

Table 5. Mean values (and estimated standard deviations from the mean) for the conformation-angles (in degrees) of mono- and poly- β -nucleotides

Only those polymer models refined against X-ray-diffraction data by the method of linked-atom least-squares are included: underlined values of σ were fixed in these refinements.

	Monomers	Polymers
ϕ and ψ	64 (16), 179 (10), -60 (14)*	-69 (18)
θ	180 (15)	-166 (18)
ξ	52 (8), -179 (9), -67 (2)	50 (6)
σ	83 (6)†, 147 (6)‡, 156 (0)§	83¶, 146‡, <u>157**</u>
ω	-123 (24)	<u>-170 (22)</u>
χ	81 (11) , 114 (11)‡, 180 (8)‡, -67 (16)	80 (9)¶, 143**

* Omitting the anomalous ϕ value of β -adenosyl-2'- β -uridine 5'-phosphoric acid.

† C-3-*endo* omitting the anomalous value of guanosine 5'-phosphate trihydrate.

‡ C-2-*endo*.

§ C-3-*exo* omitting the anomalous value of deoxycytidine 5'-phosphate monohydrate.

|| C-3-*endo*+C-3-*exo*.

¶ C-3-*endo*.

** C-3-*exo*. Only one value (for B-type DNA) is available.

obtained that would be demonstrably more correct than that with the fixed, standard C-3-*exo* furanose ring.

Conclusions

Of the four main furanose ring conformations, C-2-*exo* is not observed for nucleic acid monomers and the C-3-*exo* conformation has been observed on only three occasions. Although the pucker of dihydrothymidine is unusual, only one of its ring conformation-angles is anomalous. Except for conformation-angles themselves, the only pucker-dependent parameters are the ring-bond-angles at C-1 and C-4, the exocyclic bond-angles at C-2 and C-3, and the length of the C-1-N bond. Presumably because of the limited number of structures available for their survey, Sundaralingam & Jensen (1965b) concluded (incorrectly) that ring bond-angles at C-2 and C-3 were pucker-dependent.

Apart from conformation-angles about furanose ring bonds, only χ (the orientation of the base relative to the ring) depends upon ring-pucker. The precise value of χ in the dominant *anti* range depends on ring conformation, as indicated by previous surveys (Arnott & Hukins, 1969; Sundaralingam, 1969; Wilson, 1970) and subsequent theoretical calculations (Lakshminarayanan & Sasisekharan, 1969, 1970; Berthod & Pullman, 1971; Wilson & Rahman, 1971; Kang, 1971). The C-2-*endo* furanose conformation is not necessary for the, less-common, *syn* χ values to be observed (cf. Sundaralingam, 1969).

Standard C-3-*endo*, C-2-*endo* and C-3-*exo* furanose rings derived in this work may be used in molecular model-building studies of nucleic acids. Fortunately, other conformation-angles are more variable than those of the furanose ring and the computational complexity of varying its conformation can often justifiably be avoided even in precise model-building studies. Of course, each of the three standard, fixed rings has to be tested as a candidate for a structural component in a particular nucleic acid molecule.

This research was supported by grants from the National Science Foundation (GB 18835), the National Institutes of Health (GM 17371) and the Eli Lilly Co.

References

- Arnott, S. (1970) *Progr. Biophys. Mol. Biol.* **21**, 256
- Arnott, S. & Hukins, D. W. L. (1969) *Nature (London)* **224**, 886
- Arnott, S. & Scott, W. E. (1972) *J. Chem. Soc. Perkin Trans. 2*, 324
- Arnott, S., Fuller, W., Hodgson, A. & Prutton, I. (1968) *Nature (London)* **220**, 561
- Arnott, S., Dover, S. D. & Wonacott, A. J. (1969) *Acta Crystallogr. Sect. B* **25**, 2192
- Berthod, H. & Pullman, B. (1971) *Biochim. Biophys. Acta* **232**, 595
- Blake, R. D., Fresco, J. R. & Langridge, R. (1970) *Nature (London)* **225**, 32
- Bugg, C. E. & Marsh, R. E. (1967) *J. Mol. Biol.* **25**, 67
- Camerman, N. & Trotter, J. (1965) *Acta Crystallogr.* **18**, 203
- Coulter, C. L. (1969) *Acta Crystallogr. Sect. B* **25**, 2055
- Coulter, C. L. & Greaves, M. L. (1970) *Science* **169**, 1098

- Cramer, F., von der Haar, F., Holmes, K. C., Saenger, W., Schlimme, E. & Schulz, G. E. (1970) *J. Mol. Biol.* **51**, 523
- Crick, F. H. C. & Watson, J. D. (1954) *Proc. Roy. Soc. Ser. A* **223**, 80
- Dantzig, G. B. (1963) *Linear Programming and Extensions*, pp. 140–144, Princeton University Press, Princeton
- Furberg, S., Peterson, C. S. & Romming, C. (1965) *Acta Crystallogr.* **18**, 313
- Hampel, A., Labanauskas, M., Connors, P. G., Kirkegaard, L., RajBhandary, U. L., Sigler, P. B. & Bock, R. M. (1968) *Science* **162**, 1384
- Harris, D. R. & MacIntyre, W. M. (1964) *Biophys. J.* **4**, 203
- Haschemeyer, A. E. V. & Sobell, H. M. (1965a) *Acta Crystallogr.* **18**, 525
- Haschemeyer, A. E. V. & Sobell, H. M. (1965b) *Acta Crystallogr.* **19**, 125
- Hawkinson, S. W., Coulter, C. L. & Greaves, M. L. (1970) *Proc. Roy. Soc. Ser. A* **318**, 143
- Hodgkin, D. C., Lindsey, J., MacKay, M. & Trueblood, K. N. (1962) *Proc. Roy. Soc. Ser. A* **226**, 475
- Hunt, D. J. & Subramanian, E. (1969) *Acta Crystallogr. Sect. B* **25**, 2144
- Iball, J., Morgan, C. H. & Wilson, H. R. (1968) *Proc. Roy. Soc. Ser. A* **302**, 225
- Kang, S. (1971) *J. Mol. Biol.* **58**, 297
- Kim, S.-H., Quigley, G., Suddath, F. L. & Rich, A. (1971) *Proc. Nat. Acad. Sci. U.S.* **68**, 841
- Konnert, J., Karle, I. L. & Karle, J. (1970) *Acta Crystallogr. Sect. B* **26**, 770
- Kraut, J. & Jensen, L. H. (1963) *Acta Crystallogr.* **16**, 79
- Lakshminarayanan, A. V. & Sasisekharan, V. (1969) *Biopolymers* **8**, 475
- Lakshminarayanan, A. V. & Sasisekharan, V. (1970) *Biochim. Biophys. Acta* **204**, 49
- Munns, A. R. I. & Tollin, P. (1970) *Acta Crystallogr. Sect. B* **26**, 1101
- Munns, A. R. I., Tollin, P., Wilson, H. R. & Young, D. W. (1970) *Acta Crystallogr. Sect. B* **26**, 1114
- Murayama, W., Nagashima, N. & Shimizu, Y. (1969) *Acta Crystallogr. Sect. B* **25**, 2236
- Rahman, A. & Wilson, H. R. (1970) *Acta Crystallogr. Sect. B* **26**, 1765
- Rao, S. T. & Sundaralingam, M. (1969) *J. Amer. Chem. Soc.* **91**, 1210
- Rao, S. T. & Sundaralingam, M. (1970) *J. Amer. Chem. Soc.* **92**, 4963
- Rohrer, D. C. & Sundaralingam, M. (1970) *J. Amer. Chem. Soc.* **92**, 4984
- Saenger, W. (1970) *FEBS Lett.* **10**, 81
- Saenger, W. & Eckstein, F. (1970) *J. Amer. Chem. Soc.* **92**, 4712
- Saenger, W. & Scheit, K. H. (1970) *J. Mol. Biol.* **50**, 153
- Sakurai, T., Rao, S. T., Rubin, J. & Sundaralingam, M. (1971) *Science* **172**, 1234
- Seeman, N. C., Sussman, J. L., Berman, H. M. & Kim, S. H. (1971) *Nature (London) New Biol.* **233**, 90
- Shefter, E. & Trueblood, K. N. (1965) *Acta Crystallogr.* **18**, 1067
- Shefter, E., Barlow, M., Sparks, R. A. & Trueblood, K. N. (1969) *Acta Crystallogr. Sect. B* **25**, 895
- Shoemaker, C. B., Cruickshank, D. W. J., Hodgkin, D. C., Kamper, M. J. & Pilling, D. (1964) *Proc. Roy. Soc. Ser. A* **278**, 1
- Spencer, M. (1959) *Acta Crystallogr.* **12**, 59
- Subramanian, E. & Hunt, D. J. (1970) *Acta Crystallogr. Sect. B* **26**, 303
- Sundaralingam, M. (1966) *Acta Crystallogr.* **21**, 495
- Sundaralingam, M. (1969) *Biopolymers* **7**, 821
- Sundaralingam, M. & Arora, S. K. (1969) *Proc. Nat. Acad. Sci. U.S.* **64**, 1021
- Sundaralingam, M. & Jensen, L. H. (1965a) *J. Mol. Biol.* **13**, 914
- Sundaralingam, M. & Jensen, L. H. (1965b) *J. Mol. Biol.* **13**, 930
- Sussman, J. L., Seeman, N. C., Kim, S. H. & Berman, H. M. (1972) *J. Mol. Biol.* **66**, 403
- Tavale, S. S. & Sobell, H. M. (1970) *J. Mol. Biol.* **48**, 109
- Thewalt, U. T., Bugg, C. E. & Marsh, R. E. (1970) *Acta Crystallogr. Sect. B* **26**, 1089
- Trueblood, K. N., Horn, P. & Luzzati, V. (1961) *Acta Crystallogr.* **14**, 965
- Viswamitra, M. A., Reddy, B. S., Lin, G. H.-Y. & Sundaralingam, M. (1971) *J. Amer. Chem. Soc.* **93**, 4565
- Watenpaugh, K., Dow, J., Jensen, L. H. & Furberg, S. (1968) *Science* **159**, 206
- Watson, D. G., Sutor, D. J. & Tollin, P. (1965) *Acta Crystallogr.* **19**, 111
- Wilson, H. R. (1970) *Nature (London)* **225**, 545
- Wilson, H. R. & Rahman, A. (1971) *J. Mol. Biol.* **56**, 129
- Young, D. W., Tollin, P. & Wilson, H. R. (1969) *Acta Crystallogr. Sect. B* **25**, 1423