Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. A New Description Using the Concept of Pseudorotation[1]

C. Altona and M. Sundaralingam*

Contribution from the Crystallography Laboratory, Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706. *Received February 9, 1972*

Abstract: The amplitude (τ_m) and phase angle of pseudorotation (P) of the sugar ring in several β-purine and β-pyrimidine nucleosides and nucleotides were calculated from the known endocyclic torsion angles. A statistical classification of the number of compounds for which P falls in a given range shows that only two relatively narrow pseudorotational ranges are preferred by β sugars in the solid, each occupying less than 10% of the total pathway.

The furanose ring is centrally located in the sugar phosphate chains of nucleic acids. From X-ray studies it is known that this five-membered ring occurs in two ranges of conformations, usually referred to as C(2′)-endo and C(3′)-endo forms.[2] A changeover from one conformer of the furanose ring into the other produces major structural changes in the naturally occurring nucleic acids [*e.g.*, B-DNA[3] (C(2′)-endo) and RNA[4] (C(3′)-endo)] and has an important bearing on their biological functions. The conformational differences of the furanose ring in B-DNA and RNA, respectively, are probably a manifestation of the polymeric structure, since the component ribo- and deoxyribonucleosides and nucleotides may crystallize in either of the two conformations, the ribose derivatives in particular showing little preference for one or the other conformational range. Changes in hydration and metal ions can markedly affect the furanose ring conformation, for example B-DNA (95% humidity, lithium salt) (C(2′)-endo)[2] and A-DNA (75% relative humidity, sodium salt) (C(3′)-endo).[5]

Detailed information on the conformation of the furanose ring cannot be obtained from X-ray diffraction studies of fibrous nucleic acids and polynucleotides because of the serious limitation in the resolution of the X-ray data. The main source of information on the furanose conformation has come from single-crystal X-ray diffraction studies of nucleosides and nucleotides.[2] Previous descriptions of furanose ring conformations were based on calculated three- and four-atom planes and the torsion angles about the furanose ring bonds. The description based on the "best planes" is only approximate. Moreover, it does not convey the concept of continuum of conformations within a single potential energy well, and it is based upon symmetry properties of a ring system which by virtue of its substitution pattern is not expected to show symmetry. In this paper we utilize a description which is based on the concept of pseudorotation and which allows unequivocal determination of the exact conformation of each furanoid ring in terms of two paramteres, the "phase angle" of pseudorotation P and the degree of pucker, τ_m.

Amplitude and Phase Angle of Pseudorotation

The concept of pseudorotation was first introduced by Kilpatrick, *et al.*,[6] in their discussion of the "indefiniteness" of the cyclopentane conformation. In cyclopentane the angle of maximum puckering rotates without substantial change in potential energy. However, the presence of one or more substituents, endocyclic or exocyclic, will give rise to an *induced* potential energy barrier opposing "free" pseudorotation.[7] The consequences of "limited" pseudorotation were more recently explored by Altona, *et al.*[8-13] A quantitative description of puckering and conformation in terms of the maximum angle of torsion (θ_m) and of the "phase angle" of pseudorotation (Δ) was developed,[10] based on the interrelationship between the five torsion angles in a nonplanar five-membered ring. In fact, the phase angle locates the exact position on the pseudorotational pathway relative to some chosen "standard" conformation and can be calculated easily when two or more torsion angles are known from experiment. The conformations of ring D in steroids[10] and of several other five-membered ring systems for which X-ray data are available have been described by this method.[11,12] In the present article, we propose a slight modification and purport to show its usefulness for the accurate description of ribose and deoxyribose rings in nucleosides and nucleotides.

The relationship between the five torsion angles (numbered θ_0, θ_1, θ_2, θ_3, and θ_4) of each of the infinite number of conformations met on the pseudorotational

(1) Part of this work was carried out during 1968–1969, when both authors were cooperating at Case Western Reserve University, Cleveland, Ohio 44106.

(2) M. Sundaralingam, *Biopolymers*, 7, 821 (1969), and references therein.

(3) R. Langridge, D. A. Marvin, W. E. Seeds, H. R. Wilson, C. W. Harper, M. H. F. Wilkins, and L. D. Hamilton, *J. Mol. Biol.*, 2, 38 (1960).

(4) S. Arnott, M. H. F. Wilkins, W. Fuller, and R. Langridge, *ibid.*, 27, 535 (1967); S. Arnott, *Progr. Biophys. Mol. Biol.*, 21, 265 (1970).

(5) W. Fuller, M. H. F. Wilkins, H. R. Wilson, and L. D. Hamilton, *J. Mol. Biol.*, 12, 60 (1965).

(6) J. E. Kilpatrick, K. S. Pitzer, and R. Spitzer, *J. Amer. Chem. Soc.*, 69, 2483 (1947).

(7) K. S. Pitzer and W. F. Donath, *ibid.*, 81, 3213 (1959).

(8) C. Altona, H. R. Buys, and E. Havinga, *Recl. Trav. Chim. Pays-Bas*, 85, 973 (1966).

(9) H. J. Geise, C. Altona, and C. Romers, *Tetrahedron Lett.*, 1383 (1967).

(10) C. Altona, H. J. Geise, and C. Romers, *Tetrahedron*, 24, 13 (1968).

(11) C. Altona and A. P. M. van der Veek, *ibid.*, 24, 4377 (1968).

(12) C. Romers, C. Altona, H. R. Buys, and E. Havinga, *Top. Stereochem.*, 4, 39 (1969).

(13) C. Altona, "Conformational Analysis," G. Chiurdoglu, Ed., Academic Press, New York, N. Y., 1971, p 1. The minus sign in eq 1 in this paper should be replaced by a plus sign.

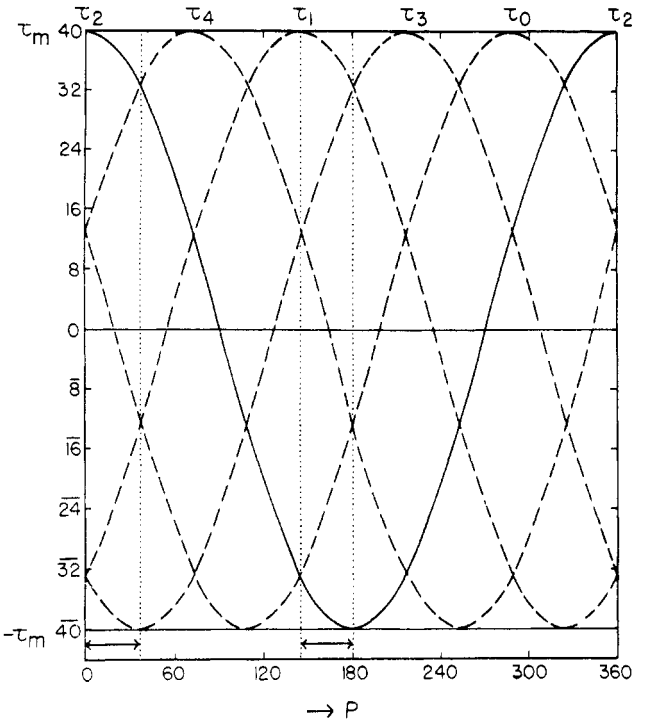

Figure 1. Calculated change of the five torsion angles of a puckered five-membered ring during one full pseudorotational cycle $0 \leqslant P \leqslant 360°$. Dotted vertical lines show the conformational limits usually encountered in single-crystal studies.

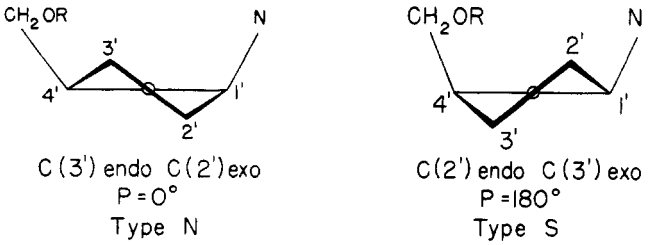

Figure 2. Diagrammatic projections of the furanose ring in two idealized twist conformations, looking toward the oxygen atom from the center of the C(2′)–C(3′) bond. Type N at $P = 0°$ represents the chosen standard, tpye S at $P = 180°$ represents its mirror image.

pathway of cyclopentane can be described by a simple cosine function,[10] eq 1

$$\theta_j = \theta_m \cos (P + j\delta) \qquad (1)$$

where $j = 0$–4, $\delta = 144°$ For $j = 0$, eq 1 reduces to

$$\theta_0 = \theta_m \cos P \qquad (2)$$

Equation 2 shows that θ_0 passes the values θ_m, 0, $-\theta_m$, 0, θ_m when P goes from 0 to 360°, *i.e.*, when the molecule has completed a full pseudorotational cycle and the original conformation is restored. It is important to remember that the maximum puckering travels *twice* around the ring during *one* full cycle; a pseudorotation over $P = 180°$ yields the mirror image of the original ring (all signs of θ_j inverted). In earlier work Δ ($= 2P$) was called the "phase angle" of pseudorotation, but for convenience of graphic representation we now propose to use P to represent the "phase angle."[14]

(14) Our P is equivalent to ϕ in Pitzer and Donath's equation.[7]

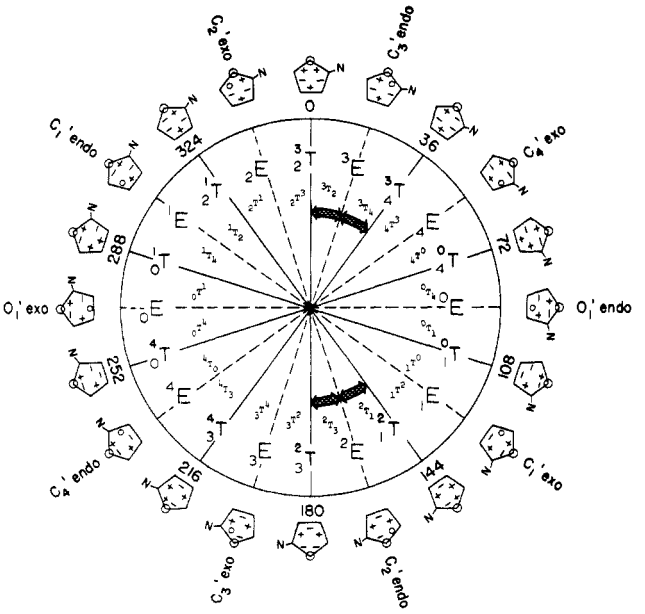

Figure 3. Pseudorotational pathway of the furanose ring. Each point on the circle represents a specific value of the phase angle of pseudorotation P. Heavy radial lines represent T conformations, dotted radials represent E forms; the corresponding signs of the ring torsion angles are also shown. Heavy arrows indicate the preferred pseudorotational regions. For details of abbreviated nomenclature, see ref 15b.

From eq 1 a useful formula (3) is easily derived[10]

$$\tan P = \frac{(\theta_2 + \theta_4) - (\theta_1 + \theta_3)}{2\theta_0(\sin 36 + \sin 72)} \qquad (3)$$

Now, considering a given sugar ring, characterized by five torsion angles about the ring bonds, as being frozen at a certain point along the total pseudorotation pathway available to this particular species, one can apply eq 3 to extract the phase angle P. This P represents the point lowest in energy, being determined by location and orientation of substituents and by packing forces in the crystal. Note that in cases where θ_0 is *negative* one should add 180° to the calculated value of P. Equation 2 yields the value for the second parameter θ_m.[15a] Figure 1 shows the successive change of each θ_j during one full pseudorotational cycle. The τ notation is used in this figure; see below.

Conformations of the Sugar Ring

We select a standard conformation ($P = 0°$) which is characterized by a maximally positive C(1′)–C(2′)–C(3′)–C(4′) torsion angle, *i.e.*, the symmetrical C(2′)-exo–C(3′)-endo ($^3T_2$)[15b] form, Figure 2. This definition of the $P = 0°$ point is most convenient because for the great majority of riboses and deoxyriboses P now falls in the ranges 0–36° and 144–180°. The

(15) (a) A Fortran program CONFOR is available upon request from one of the authors (C. A.). The program utilizes the permutational properties of eq 1–3, thus smoothing out the effect of experimental errors, with the added advantage of normalizing the results. Strictly speaking, eq 1–3 are approximate only for five-membered rings in which not all bond distances are equal. The effect of the shorter C–O bonds is small in our case, however, since comparison between input (experimental) torsion angles of furanoses and those computed from eq 1 using P_{av} and $\theta_{m,av}$ in the majority of cases shows agreement well within the limits of experimental error ($\leq \pm 0.6°$). (b) For abbreviated nomenclature see M. Sundaralingam, *J. Amer. Chem. Soc.*, **93**, 6644 (1971), and references therein.

Table I. The Calculated Values of the Phase Angle of Pseudorotation P and of the Amplitude of Pucker τ_m of the Sugar Ring for Compounds Having Type N Conformation[a]

Molecule no.	Compound	P, deg	τ_m[b]	τ_2 (θ_0)	τ_3 (θ_1)	τ_4 (θ_2)	τ_0 (θ_3)	τ_1 (θ_4)	χ_{CN}, deg	Side chain	Conventional notation	Ref
	β-Purines											
1	Uridylyl 3′-phosphate 5′-adenosine hemihydrate, adenosine residue molecule 1	1.9	31.3	30.7	−26.4	10.9	9.0	−24.9	48.7	52 gg	C(3′)-endo	*c*
2	Puromycin dihydrochloride pentahydrate	3.4	39.9	38.9	−34.0	14.9	10.4	−31.2	19.3	54 gg	C(3′)-endo	*d*
3	3′-Deoxy-3′-(dihydroxyphosphinylmethyl)adenosine (in complex with ethanol)	4.7	38.2	37.1	−32.9	15.2	9.0	−29.5	3.8		C(3′)-endo	*e*
4	Inosine	7.8	41.8	40.6	−37.1	18.6	7.8	−30.6	12.4	−169 gt	C(3′)-endo	*f*
5	Guanosine 5′-phosphate trihydrate[h]	8.0	34.8	33.8	−31.4	15.3	6.5	−25.0	12.4	46 gg	C(3′)-endo	*g, h*
6	Same as 1, molecule 2	10.8	36.8	35.1	−33.8	18.2	5.3	−25.4	37.5	57 gg	C(3′)-endo	*c*
7	Adenosine (in complex with 5-bromouridine)	11.2	40.6	38.5	−36.9	20.7	5.0	−28.7	12.4	41 gg	C(3′)-endo	*i*
8	Adenosine 5′-phosphate monohydrate	12.2	43.9	42.3	−40.0	22.8	4.8	−29.8	25.6	40 gg	C(3′)-endo	*j*
9	Adenosine 3′-phosphate dihydrate	23.3	38.4	34.6	−37.6	25.6	−3.4	−20.1	3.8	171 gt	C(3′)-endo	*k*
	β-Pyrimidines											
10	Deoxycytidine hydrochloride	3.3	37.7	36.9	−31.9	14.2	9.8	−29.5	−0.2	46 gg	C(3′)-endo	*l*
11	Same as 1, uridine residue, molecule 2	5.6	38.7	40.2	−32.5	15.3	8.2	−28.0	10.3	50 gg	C(3′)-endo	*c*
12	4-Thiouridine disulfide, residue 1	6.5	35.4	34.3	−31.0	15.3	7.5	−26.6	16.6	57 gg	C(3′)-endo	*m*
13	Adenylyl 2′-phosphate 5′-uridine tetrahydrate, uridine residue	8.7	39.5	38.1	−35.5	18.2	6.9	−28.5	4.9	57 gg	C(3′)-endo	*n*
14	5-Iodouridine, molecule 1	8.9	36.1	35.2	−32.3	16.7	6.3	−25.5	11.8	62 gg	C(3′)-endo	*o*
15	Cytidine	9.0	38.7	37.4	−34.7	18.0	6.2	−28.0	18.4	47 gg	C(3′)-endo	*p*
16	2-Thiocytidine dihydrate	11.1	38.7	37.1	−35.3	19.3	4.9	−26.8	20.3	55 gg	C(3′)-endo	*q*
17	Same as 11, molecule 1	11.4	38.6	36.9	−35.6	19.2	4.7	−27.7	20.0	54 gg	C(3′)-endo	*c*
18	Same as 12, residue 2	12.3	37.6	36.7	−34.1	19.2	3.4	−25.5	18.1	46 gg	C(3′)-endo	*m*
19	2,4-Dithiouridine monohydrate	12.5	38.8	37.2	−35.9	19.9	4.1	−25.9	19.5	41 gg	C(3′)-endo	*r*
20	5-Methyluridine hemihydrate	15.2	40.5	38.5	−38.2	22.5	2.3	−25.7	29.4	49 gg	(C3′)-endo	*s*
21	5-Bromouridine (in complex with adenosine)	22.8	41.5	38.2	−41.1	26.2	−3.8	−21.3	20.0	36 gg	C(3′)-endo	*i*
22	Calcium thymidine 5′-phosphate hexahydrate	25.2	35.4	31.2	−34.9	25.0	−3.9	−17.6	43.5	57 gg	C(3′)-endo	*t*
23	4-Thiouridine monohydrate	34.2	41.0	33.1	−40.6	33.3	−11.3	−14.3	−96.6	66 gg	C(3′)-endo	*u*
24	Dihydrothymidine	84.4	31.6	3.1	−20.4	31.2	−29.4	15.4	52.8	53 gg −169 gt	O(1′)-endo	*v*

[a] The ring torsion angles χ_{CN}[2] and the conformation of the C(5′)–O(5′) side chain are also shown. The ring torsion angles and χ_{CN} were calculated from the published atomic coordinates except where indicated otherwise. [b] τ values in degrees. [c] J. Rubin, T. Brennan, and M. Sundaralingam, *Science*, **174,** 1020 (1971); N. C. Seeman, J. L. Sussman, H. Berman, and S. H. Kim, *Nature (London), New Biol.*, **223,** 90 (1971). [d] M. Sundaralingam and S. K. Arora, *Proc. Nat. Acad. Sci., U. S.* **64,** 1021 (1969); M. Sundaralingam and S. K. Arora, *J. Mol. Biol.*, in press. [e] S. Hecht and M. Sundaralingam, *J. Amer. Chem. Soc.*, **94,** 4314 (1972). [f] A. R. I. Munns and P. Tollin, *Acta Crystallogr., Sect. B*, **26,** 1101 (1970). [g] W. Murayama, N. Nagashima, and Y. Shimizu, *ibid., Sect. B*, **25,** 2236 (1969). [h] The published coordinates of atom O(3′) seem to contain an error. [i] A. E. V. Haschemeyer and H. M. Sobell, *Acta Crystallogr.*, **18,** 525 (1965). [j] J. Kraut and L. H. Jensen, *ibid.*, **16,** 79 (1963). [k] M. Sundaralingam, *ibid.*, **21,** 495 (1966). [l] E. Subramanian and D. J. Hunt, *ibid., Sect. B*, **26,** 303 (1970). [m] E. Shefter and T. I. Kalman, *Biochem. Biophys. Res. Commun.*, **32,** 878 (1968), and private communication. [n] E. Shefter, M. Barlow, R. Sparks, and K. N. Trueblood, *Acta Crystallogr., Sect. B*, **25,** 895 (1969). [o] A. Rahman and H. R. Wilson, *ibid, Sect. B*, **26,** 1765 (1970). [p] S. Furberg, C. S. Peterson, and C. Romming, *ibid.*, **18,** 313 (1965). [q] G. H.-Y. Lin, M. Sundaralingam, and S. K. Arora, *J. Amer. Chem. Soc.*, **93,** 1235 (1971). [r] G. H.-Y. Lin and M. Sundaralingam, *Acta Crystallogr., Sect. B*, **27** 961 (1971). [s] D. J. Hunt and E. Subramanian, *ibid., Sect. B*, **25,** 2144 (1969). [t] K. N. Trueblood, P. Horn, and V. Lussati, *ibid.*, **14,** 965 (1661). [u] W. Saenger and K. H. Scheit, *J. Mol. Biol.*, **50,** 153 (1970). [v] J. Konnert, I. L. Karle, and J. Karle, *Acta Crystallogr., Sect. B*, **26,** 770 (1970).

previously suggested nomenclature[2] for the torsion angles about the sugar ring is as follows: τ_0 about O(1′)–C(1′), τ_1 about C(1′)–C(2′), τ_2 about C(2′)–C(3′), and so on in a clockwise manner. For purposes of using eq 3 in accordance with our selected standard conformation one should remember to convert as follows: $\theta_0 = \tau_2$, $\theta_1 = \tau_3$, $\theta_2 = \tau_4$, $\theta_3 = \tau_0$, $\theta_4 = \tau_1$.

The relationship between the usual endo–exo and *T–E* notations and the one presently utilized is illustrated in Figure 3. It is seen that symmetrical twist (*T*) conformations arise at *even multiples* of 18° of the phase angle P, symmetrical envelope (*E*) conformations arise at *odd multiples* of P. These symmetrical forms are *not* expected to represent the true minimum energy geometry of an asymmetrically substituted five-membered ring (except occasionally by accident) and indeed we find that in all but a few molecules P has some intermediate value. Figure 3 also shows the signs of the ring

Table II. The Calculated Values of the Phase Angle of Pseudorotation P and of the Amplitude of Pucker τ_m of the Sugar Ring for Compounds Having Type S Conformation[a]

Molecule no.	Compound	P, deg	$-\tau_m$[b]	τ_2 (θ_0)	τ_3 (θ_1)	τ_4 (θ_2)	τ_0 (θ_3)	τ_1 (θ_4)	χ_{CN}, deg	Side chain, deg	Conventional notation	Ref
	β-Purines											
25	Guanosine dihydrate, molecule 2	139.2	44.3	−32.4	11.1	17.0	−38.5	44.2	46.6	46 gg	C(1′)-exo	*c*
26	Formycin monohydrate	148.3	39.6	−33.1	15.7	9.7	−30.6	39.4	109.8	176 gt	C(2′)-endo	*d*
27	Tubercidin	149.3	43.8	−36.6	18.2	9.4	−33.5	43.8	73.1	−178 gt	C(2′)-endo	*e*
28	Inosine dihydrate, molecule 2	150.4	41.3	−34.8	17.9	8.0	−31.2	41.2	49 1	47 gg	C(2′)-endo	*c*
29	Same as 13, adenosine residue	153.4	43.3	−37.8	20.2	6.1	−30.5	43.5	54.6	45 gg	C(2′)-endo	*f*
30	8-Bromoguanosine	153.8	37.3	−32.6	18.0	5.0	−26.7	36.9	−130.1	54 gg	C(2′)-endo	*g*
31	3′-*O*-Acetyladenosine	154.5	37.7	−33.2	18.8	4.8	−26.5	37.1	−133.0	58 gg	C(2′)-endo	*i*
32	6-Thioinosine, molecule 2	159.6	40.5	−36.8	23.2	1.3	−25.8	39.3	−143.7	57 gg	C(2′)-endo	*j*
33	Same as 25, molecule 1	161.4	36.2	−33.4	21.6	0.2	−22.1	34.7	122.5	68 gg	C(2′)-endo	*c*
34	Same as 28, molecule 1	163.6	39.1	−36.7	24.4	−1.3	−22.6	37.0	120.0	64 gg	C(2′)-endo	*c*
35	Rubidium adenosine 5-diphosphate trihydrate	163.8	37.7	−34.4	24.0	−2.2	−22.4	36.0	38.0	57 gg	C(2′)-endo	*k*
36	8-Bromoadenosine	163.8	40.0	−37.3	25.3	−1.8	−23.4	37.7	−120.0	46 gg	C(2′)-endo	*h*
37	Same as 32, molecule 1	164.4	44.0	−41.4	27.7	−2.1	−24.4	42.0	−135.0	55 gg	C(2′)-endo	*j*
38	Deoxyguanosine (in complex with bromodeoxycytidine)	165.1	32.4	−31.0	20.2	−2.1	−17.6	30.6	−148.7	43 gg	C(2′)-endo	*l*
39	NaIMP·8H_2O	166.8	40.3	−38.5	26.8	−3.5	−21.2	37.3	40.9	59 gg	C(2′)-endo	*m*
40	N^2-Dimethylguanosine	173.7	36.9	−35.9	27.7	−7.9	−15.6	32.4	−103.9	−178 gt	C(2′)-endo	*n*
41	5′-Methylammonium-5′-deoxyadenosine	180.8	37.9	−36.9	31.4	−12.6	−11.5	30.9	−151.2	158 gt	C(3′)-exo	*o*
42	Deoxyadenosine monohydrate[q]	194.3	36.3	−34.4	33.7	−19.9	−2.4	23.9	10.9	−173 gt	C(3′)-exo	*p, q*
	β-Pyrimidines											
43	5-Bromo-2′-deoxyuridine	145.5	41.2	−33.1	14.6	11.6	−33.4	41.0	47.2	167 gt	C(2′)-endo	*r*
44	Bromodeoxycytidine (in complex with deoxyguanosine)	150.6	37.2	−32.0	15.8	6.9	−28.1	36.7	58.7	61 gg	C(2′)-endo	*l*
45	Dihydrouridine hemihydrate, molecule 2	151.3	33.1	−28.3	14.9	6.0	−24.7	32.8	57.1	−65 tg	C(2′)-endo	*s*
46	5-Iodo-2′-deoxyuridine	153.3	38.5	−33.1	18.7	5.5	−27.9	38.1	63.3	51 gg	C(2′)-endo	*t*
47	Same as 45, molecule 1	153.5	43.0	−37.6	20.4	6.1	−30.7	42.6	65.5	169 gt	C(2′)-endo	*s*
48	5-Bromouridine[u]	161.6	36.5	−34	22	0	−22	35	(51)	55 gg	C(2′)-endo	*r, u*
49	5-[1-2′-Deoxyuracilyl][q,w] ethyl sulfide, molecule 1						21.2		56.6	58 gg	C(2′)-endo	*q, v, w*
50	Same as 49, molecule 2[r]	163.8	37.9	−35.5	23.8	−1.5	−21.9	35.7	55.7	53 gg	C(2′)-endo	*q, v*
51	3′-*O*-Acetyl-4-thiothymidine	167.6	34.9	−33.2	23.7	−3.6	−18.0	32.2	54.0	51 gg	C(2′)-endo	*x*
52	5-Chlorouridine	168.8	35.4	−33.8	24.6	−4.5	−17.8	32.3	51.4	52 gg	C(2′)-endo	*y*
53	CMP (orthorhombic)	169.4	40.2	−38.7	28.1	−5.6	−19.8	36.6	41.8	44 gg	C(2′)-endo	*z*
54	BaUMP·7H_2O	170.2	34.2	−33.0	23.8	−5.2	−15.8	31.4	43.0	55 gg	(C2′)-endo	*aa*
55	5-Bromouridine–dimethylsulfoxide complex	170.5	41.1	−40.1	29.0	−6.4	−19.3	36.9	62.2	53 gg	C(2′)-endo	*bb*
56	5-Fluoro-2′-deoxyuridine	171.1	41.6	−40.3	30.1	−6.8	−19.4	37.1	59.0	−68 tg	C(2′)-endo	*cc*
57	Same as 53 (monoclinic)	171.7	38.7	−37.5	27.9	−7.0	−17.6	34.6	39.3	46 gg	C(2′)-endo	*dd*
58	Same as 14, molecule 2	174.8	42.1	−41.2	32.1	−9.6	−16.9	36.5	55.8	−65 tg	C(2′)-endo	*ee*
59	Thymidine	187.5	38.2	−36.9	33.2	−16.7	−7.0	27.8	39.1	173 gt	C(3′)-exo	*ff*
60	dCMP·H_2O	213.6	32.9	−26.9	32.6	−26.3	8.8	12.0	−5.9	57 gg	C(3′)-exo	*gg*

[a] The ring torsion angles, χ_{CN}[2], and the conformation of the C(5′)–O(5′) side chain are also shown. The ring torsion angles and χ_{CN} were calculated from the published atomic coordinates except where indicated otherwise. [b] τ values are in degrees. [c] U. Thewalt, C. E. Bugg, and R. E. Marsh, *Acta Crystallogr., Sect. B*, **26,** 1089 (1970). [d] P. Prusiner, T. Brennan, and M. Sundaralingam, manuscript in preparation. [e] M. Sundaralingam and J. Abola, *Acta Crystallogr.*, in press. [f] See footnote *n*, Table I. [g] C. E. Bugg and U. Thewalt, *Biochem. Biophys. Res. Commun.*, **37,** 623 (1969); see also footnote *h*. [h] S. S. Tavale and H. M. Sobell, *J. Mol. Biol.*, **48,** 109 (1970). [i] S. T. Rao and M. Sundaralingam, *J. Amer. Chem. Soc.*, **92,** 4963 (1970). [j] E. Shefter, *J. Pharm. Sci.*, **57,** 1157 (1968). [k] F. M. Muller, private communication. [l] A. E. V. Haschemeyer and H. M. Sobell, *Acta Crystallogr.*, **19,** 125 (1965). [m] S. T. Rao and M. Sundaralingam, *J. Amer. Chem. Soc.*, **91,** 1210 (1969). [n] T. Brennan, C. Weeks, E. Shefter, S. T. Rao, and M. Sundaralingam, *ibid.*, in press. [o] W. Saenger, *ibid.*, **93,** 3035 (1971). [p] D. G. Watson, D. J. Sutor, and P. Tollin, *Acta Crystallogr.*, **19,** 111 (1969). G. H.-Y. Lin and M. Sundaralingam, to be published. [q] Coordinates used here represent the mirror image of the published structure. [r] J. Iball, C. H. Morgan, and H. R. Wilson, *Proc. Roy. Soc. (London), Ser. A*, **295,** 320 (1966). [s] M. Sundaralingam, S. T. Rao, and J. Abola, *Science*, **172,** 725 (1971). [t] N. Camerman and J. Trotter, *Acta Crystallogr.*, **18,** 203 (1965). [u] An error is evident in the published coordinates of C(1′). An attempt was made to calculate these coordinates from the known bond distances and angles; nevertheless, the angles τ_0, τ_1, τ_2, and τ_4 may be in error by a few degrees. [v] G. W. Frank, Ph.D. Thesis, University of New York, Buffalo, N. Y., 1968. [w] An error is evident in the published coordinates of C(3′). From τ_0 the approximate location of this residue on the P scale can be deduced. [x] W. Saenger and D. Suck, *Acta Crystallogr.*, in press, and private communication. [y] C. Coulter and S. W. Hawkinson, *Proc. Nat. Acad. Sci. U.S.*, **63,** 1359 (1969); S. W. Hawkinson and C. Coulter, *Acta Crystallogr.*, in press. [z] M. Sundaralingam and L. H. Jensen, *J. Mol. Biol.*, **13,** 914 (1965). [aa] M. Shefter and K. N. Trueblood, *Acta Crystallogr.*, **18,** 1067 (1965). [bb] J. Iball, C. H. Morgan, and H. R. Wilson, *Proc. Roy. Soc. (London), Ser. A*, **302,** 225 (1968). [cc] R. D. Harris and W. M. MacIntyre, *Biophys. J.*, **4,** 203 (1964). [dd] C. E. Bugg and R. E. Marsh, *J. Mol. Biol.*, **25,** 67 (1967). [ee] See footnote *o*, Table I. [ff] D. W. Young, P. Tollin, and H. R. Wilson, *Acta Crystallogr., Sect. B*, **25,** 1423 (1969); G. H.-Y. Lin and M. Sundaralingam, to be published. [gg] M. A. Viswamitra, B. S. Reddy, G. H.-Y. Lin, and M. Sundaralingam, *J. Amer. Chem. Soc.*, **93,** 4565 (1971).

Table III. Statistical Classification of the Number of β Nucleosides and β Nucleotides for Which the Pseudorotation Parameter P Falls in a Given Range[a]

	P range, deg	R purines	D purines	R pyrimidines	D pyrimidines
Type N	0–18	12		8	1 C_3'-endo
	18–36	1		2	1
	36–54				C_4'-exo
	54–72				
	72–90				1 O_1'-endo
Total type N		13	0	10	3
(P_{av})N		9.6°		10.8°[b]	*c*
Type S					
	90–108				O_1'-endo
	108–126				
	126–144	1			C_1'-exo
	144–162	9		3	3
	162–180	6	1	6	4 C_2'-endo
	180–198		2		1
	198–216				1 C_3'-exo
Total type S		16	3	9	9
(P_{av})S		157.7°	*c*	169.6°[d]	169.1°
Total N + S		29[e]	3	19	12

[a] The number of ribose (R) and deoxyribose (D) purines and pyrimidines is shown separately. [b] Molecule 23 was omitted from the calculation of P_{av}; its inclusion yields $P_{av} = 12.6°$. [c] The available data are considered insufficient to give a meaningful value of P_{av}. [d] Dihydrouridine (no. 45) was omitted from the calculation of P_{av} since the saturation of the double bond of the base seems to produce an effect on the pseudorotation of the ribose not present in fully aromatic bases. In fact, all "normal" β-pyrimidines of type S ribose conformations fall in the narrow range of $P = 161–175°$. [e] Including the compounds referred to in T. F. Lai and R. E. Marsh, private communication, and O. Kennard, N. W. Isaacs, W. D. S. Motherwell, J. C. Coppola, D. L. Wampler, A. C. Larson, and D. H. Watson, *Proc. Roy. Soc. London, Ser. A*, **325**, 401 (1971). The data on adenosine indicate that the conformation of the ribose ring in this compound (no. 61, type N) lies between those of molecules **3** and **4** of Table I: $P = 7.0°$, $\tau_m = 36.8°$. The conformation of the C(5′)–O(5′) side chain is gauche–trans.

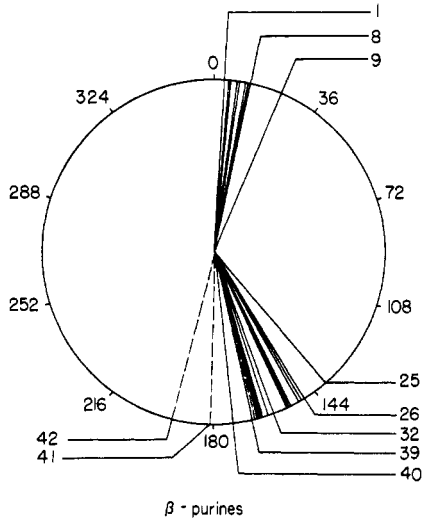

Figure 4. P values for β-purine glycosides. The preferred ring conformations are characterized by P in the ranges of 2–14° (type N) and 139–174° (type S). Only one β-purine riboside and the two β-purine deoxyribosides occur outside these ranges. Numbers shown refer to the numbers in the first column of Tables I and II.

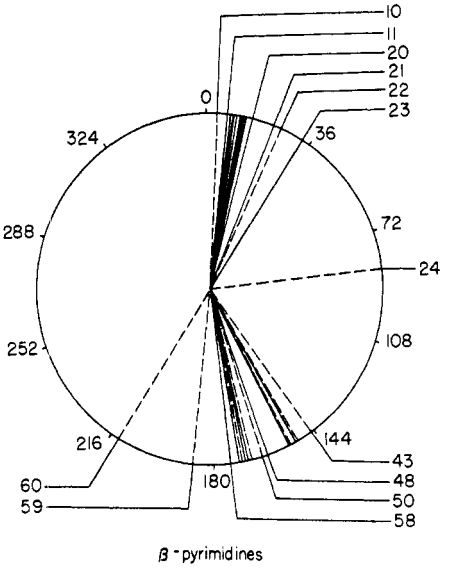

Figure 5. P values for β-pyrimidine glycosides. The preferred ring conformations are characterized by P in the ranges of 3–15° (type N) and 161–175° (type S). Dihydropyrimidines and several deoxyribosides occur outside these ranges. Numbers shown refer to the numbers in the first column of Tables I and II.

torsion angles for the ten T and ten E conformations. Given the signs and relative magnitudes of the τ values of a given sugar ring one can immediately deduce its approximate location on the pseudorotational circle without recourse to "best planes." Note that the upper or northern half of the circle ($P = 0 \pm 90°$) comprises all conformations that show a *positive* value of τ_2 (for purposes of quick identification denoted as type N) whereas the southern half ($P = 180 \pm 90°$) comprises all conformations with a *negative* value of τ_2, denoted as type S. Tables I and II show calculated values of P and of the amplitude of pucker τ_m for the 60 samples of β-purine and β-pyrimidine glycosides which are available to us at the time of writing, now arranged in meaningful order of progressive pseudorotation. The 3′,5′ cyclic phosphates and their analogs as well as the α nucleosides will be reported on elsewhere. The observed distribution of P values along the pseudorotational pathway of the purines and pyrimidines is visualized in Figures 4 and 5, respectively.

The τ_m values generally range from about 35 to 45°, yielding an average value of 39°. There seems to be little or no correlation between τ_m and P; in other words, crystal packing effects, hydrogen bond-

ing, and so on may cause a particular ring in a particular environment to flatten slightly (small τ_m) or to pucker (large τ_m) relative to the average puckering. A small difference (1–1.5°) is noted between τ_m of purine and of pyrimidine derivatives, purines being slightly more puckered than pyrimidines, regardless of type N or of type S preference: N-type purines, $\tau_{m,av}$ = 39.7°; S-type purines, $\tau_{m,av}$ = 39.5; N-type pyrimidines, $\tau_{m,av}$ = 38.4°; S-type pyrimidines, $\tau_{m,av}$ = 38.0°.

Of more interest from a conformational point of view are the insights gained by studying the behavior of the ribose and of the deoxyribose ring in terms of the pseudorototation parameter P. A statistical breakdown of the available data is presented in Table III. From Tables I–III we note the following.

(i) Two relatively narrow pseudorotational ranges are preferred by the β sugars in the solid, each occupying less than 10% of the total pseudorotational pathway. In the type N conformations P ranges from 3 to 34° (width 31°), but a definite preference for the lower part of the range is exhibited by both purines and pyrimidines. In the majority of cases P is found in the range 3–23°, with an average of about 10°. In the type S conformations a distinction between β-purines and β-pyrimidines is seen, although their preferred ranges definitely do overlap. The total allowed P range for the ribosides goes from 139 to about 175° (width 36°) but the purines seem to prefer the lower end of the scale (P_{av} = 158°), whereas the pyrimidines (omitting those in which the base is hydrogenated) are found preferentially at the higher end (P_{av} = 170°). The deoxyribosides show peculiar behavior. In type S purines the three examples known occur near to or beyond the upper limits displayed by the ribosides, up to P = 194° (deoxyadenosine, 42), although there seems to be no compelling reason for exclusive preference for this high P region. From the behavior of the deoxypyrimidines one would expect that, as more compounds are studied, deoxypurines with lower P values will be uncovered. The deoxypyrimidines do show a far wider range of P values. In type N conformations they occur on both ends of the P scale (but more data are needed to decide whether this is general behavior), but in type S the deoxyriboses show a far greater spread in P than the corresponding riboses: 145–215°, all values in between seem equally favored. This fact perhaps indicates a greater pseudorotational freedom of the ring in the 2′-deoxy compounds, *i.e.*, a much flatter potential energy well (in conformation S at least) as compared to the potential energy curve of the ribosides.[16]

(ii) The existence of these narrow ranges implies the existence of true minimum energy regions (conformers) separated by two pseudorotational barriers of unknown and not necessarily equal height located at approximately P = 90° and P = 270°, respectively.[17] The spread in P within each range is due to the flexibility of the five-membered ring; *i.e.*, for each compound *intra-* and/or *intermolecular* forces strike a balance, resulting in a species-specific geometry (conformation). Further X-ray investigations probably will extend the presently established ranges but the main picture now seems to be clear.

(iii) In the solid state the purine as well as the pyrimidine ribosides are distributed roughly 50:50 between the two conformational minima, which may be taken to indicate that energy differences between type N and type S conformers in the "free" state are rather small so that packing effects may well play an important role in deciding which of the two forms appears in the crystalline state. In solution *both* conformers are present in dynamic equilibrium in all cases so far investigated by nmr spectroscopy.[18]

(iv) The position, orientation, and nature of the base affect the sugar ring geometry. In α nucleosides[19] different ranges of P occur (340–0° and 187–224°, respectively). The strong correlations noted earlier[2] between the β-glycosidic torsion angle χ_{CN} and the sugar conformation hold rather well (Table IV).

Table IV. The Correlation between the β-Glycosidic Angle χ_{CN}[2] and the Sugar Conformation[a]

	χ_{CN}, deg
β-Purines, type N	−1 to 40
β-Purines, type S	38–73, 110–123, −149 to −103
β-Pyrimidines, type N	−1 to 44
β-Pyrimidines, type S	39–66

[a] One or two exceptions occur in each of the four groups listed.

(16) We wish to point out that the geometrical differences between, for example, the neighboring conformations C(1′)-exo, C(2′)-endo, and C(3′)-exo, are differences of degree, not of kind. It is incorrect to speak of interconversions between these conformations because they all have very similar potential energy wells (type S in this particular example), and "interconversion" implies the existence of a potential barrier which is at least higher than the lowest pseudorotational energy level. In fact, there is no reason whatsoever to postulate any potential barrier separating, *e.g.*, C(2′)-endo and C(3′)-exo forms of the compounds studied here. The conventional distinction between these and other conformations of the same type arose from the need to describe and compare the geometries of the sugar rings in various compounds but in the course of time this distinction seems to have acquired more physical meaning than warranted. In our view it is correct to speak of N ⇄ S interconversions (see points ii and iii) which occur in solution but not in the solid.

(17) In addition to the potential barriers near P = 90 and 270° (these barriers must be present, else one cannot explain the narrow ranges found), which arise when the molecule *pseudorotates* from one energy minimum to the other, a third barrier can be envisaged which corresponds to a planar sugar ring as intermediate in going from form N to S and *vice versa*. It has been argued[8] that for cyclopentane derivatives this "through the plane" mechanism of interconversion has much lower probability than the pseudorotational mechanism because of the high strain energies involved on forcing the molecule into a plane; similar arguments apply to the ribose ring. One is tempted to speculate that the pseudorotation pathways, especially the one along the P = 90° point (which is closest to N and S conformations of β glycosides), are the favored mode of N ⇄ S interconversion. The barrier height or heights have not been measured, but an upper limit may be placed on its value. In a recent application of the ultrasonic relaxation method to solutions of adenosine [L. M. Rhodes and P. R. Schimmel, *Biochemistry*, **10**, 4426 (1971)] a single relaxation curve was observed due to a process with an apparent activation energy of 6.2 kcal mol^{-1}. This process was shown to be the syn ⇄ anti base conformational equilibrium and was practically absent in pyrimidine nucleosides. Hence it may be argued that the pseudorotational barriers should be much smaller than 6.2 kcal mol^{-1}, else a double relaxation process would have been found. On the other hand the barriers cannot be much lower than 1.5–2 kcal mol^{-1} in order to have distinct conformations. In all, barrier heights in the range of 2–4 kcal mol^{-1} seem most probable.

(18) (a) F. E. Hruska and S. S. Danyluk, *J. Amer. Chem. Soc.*, **90**, 3266 (1968); (b) S. I. Chan and J. H. Nelson, *ibid.*, **91**, 168 (1969); (c) B. W. Bangerter and S. I. Chan, *ibid.*, **91**, 3910 (1969); (d) F. E. Hruska, A. A. Grey, and I. C. P. Smith, *ibid.*, **92**, 214 (1970); (e) B. J. Blackburn, A. A. Grey, I. C. P. Smith, and F. E. Hruska, *Can. J. Chem.*, **48**, 2866 (1970); (f) H. Dugas, B. J. Blackburn, R. K. Robins, R. Deslauriers, and I. C. P. Smith, *J. Amer. Chem. Soc.*, **93**, 3468 (1971); (g) F. E. Hruska, A. A. Smith, and J. G. Dalton, *ibid.*, **93**, 4334 (1971); (h) C. Altona and M. Sundaralingam, manuscript in preparation.

(19) C. Altona and M. Sundaralingam, manuscript in preparation.

(v) Table III shows a regrettable and surprising lack of information on the conformation of 2′-deoxyribose-purines. Both structures of this type known today show type S conformations and it is of great interest to know what P values would be preferred in type N forms. There seems to be no reason for exclusive preference of type S. Of course, in A-DNA the 2′-deoxyribose rings carrying purines occur as type N conformations, but no accurate geometrical details are known.

(vi) The calculated values of P and of τ_m for various proposed RNA and DNA models are shown in Table V. Some of the older models of the sugar moiety

Table V. Values of P and τ_m for Various Proposed RNA and DNA Models

		P	τ_m	Ref
Type N	RNA-11	4.5	42.0	*a*
	RNA-11	10.5	38.8	*b*
	A-DNA	11.6	42.1	*c*
	RNA-10	13.0	40.4	*c*
	Poly-A	18.4	39.9	*d*
	DNA	47.1	21.0	*e*
Type S	C-DNA	162.0	29.8	*f*
	B-DNA	174.7	17.3	*a*
	B-DNA	180.7	38.2	*b*

[a] Reference 3. [b] Reference 20. [c] Reference 5. [d] A. Rich, D. R. Davies, F. H. C. Crick, and J. D. Watson, *J. Mol. Biol.*, **3**, 71 (1961). [e] F. H. C. Crick and J. D. Watson, *Proc. Roy. Soc., London, Ser. A*, **223**, 80 (1954). [f] D. A. Marvin, M. Spencer, M. H. F. Wilkins, and L. D. Hamilton, *J. Mol. Biol.*, **3**, 547 (1961).

in these polymers are seen to be based on a highly flattened ring (small τ_m), whereas the more recent models[20] comply well with our conclusions concerning the preferred values of P and τ_m in riboses and deoxyriboses, respectively.

(vii) The conformation of the C(5′)–O(5′) side chain seems to be correlated with P. In the central parts of the allowed ranges (N or S, purines or pyrimidines) the gauche–gauche[2] conformation occurs exclusively. In type N and in type S β-purines, gauche–trans forms occur near the lower and upper limits of the allowed ranges. This is also true for type S β-pyrimidines, but in a few cases trans–gauche conformations occur instead of gauche–trans. However, in type N β-pyrimidines the trans conformations are conspicuously absent. The latter fact fits in with results of a recent pmr study[21] of a series of β-pyrimidines in which it was shown that with increasing proportion of C(2′)-endo (type S) conformer in the conformational equilibrium the proportion of gauche–trans (trans–gauche) rotamers also increases. The observed preference of deoxypyrimidines in the solid state for type S conformations (ratio 9:3, Table III) is also apparent in solution.[21]

Effect of Intermolecular Forces

It has often been a matter of debate whether or not conformational analyses based on studies of the crystalline state are relevant to the interpretation of physical measurement in the liquid state or in solution (*e.g.*, nuclear magnetic resonance, optical rotatory dispersion, circular dichroism). This problem is admittedly a complex one and seems to have given rise to various misunderstandings in the course of communication of results between X-ray crystallography and liquid state spectroscopy. Considering the importance of the problem we wish to summarize our points of view as follows. A given conformationally labile compound exists in solution as a dynamic equilibrium between two, three, or more distinct conformers, A ⇄ B ⇄ C, etc.[22] When this compound crystallizes[23] *usually* only one of these forms remains present in the solid state, say A* (case a). An asterisk signifies that the molecule exists in its solid state geometry; see below. Occasionally other phenomena are observed to occur: case b, a dynamic equilibrium still exists in the crystal, A* ⇄ B*, which gives rise to disorder as revealed by X-ray crystallography; case c, A* and B* exist as stable conformers side by side in the same unit cell, usually in a 1:1 ratio, the crystal gaining entropy of mixing in this way; case d, different crystalline states occur, each containing the same or a different pure conformer. Vibrational spectroscopy (infrared and Raman spectroscopy) is probably the most important method by which conformational behavior in both solid *and* solution can be studied and cases a or d on the one hand and b or c on the other are often readily distinguishable. However, we are unaware of vibrational studies of mononucleosides specifically aimed at elucidating similarities and differences in the spectra of solids and solutions.

The energetics of the molecular surroundings in the solid state differ from those in solution and the conformer favored in the crystal may or may not be the predominant one in the liquid. In other words, packing forces are able to overcome a certain energetic disadvantage but just how much this disadvantage maximally can be in individual cases is a question yet to be answered. An example is formed by 4-thiouridine (molecule no. 23) which exists in the solid as a type N ribose with syn conformation of the base, whereas from the nmr coupling constants and proton chemical shifts it follows[18f,g] that the predominant conformer in solution is type S with anti conformation of the base. However, regarding the *total* available X-ray and pmr data[18h,19] one gains the strong impression [see also iii and vii above] that in general the conformational preferences in the solid find their parallel in solution as well.

The central question to be answered now is whether or not the assumption is justified that the geometries of A and A* in our example above are sufficiently similar in order to utilize information from the solid state to construct reliable models on which interpretation of liquid state measurements may be based with confidence. We propose to show that the wealth of available X-ray data on ribosides does provide at least some relevant arguments to the affirmative, which arguments are mainly based on the fortunate occur-

(20) S. Arnott, S. D. Dover, and J. A. Wonacott, *Acta Crystallogr., Sect. B*, **25**, 2192 (1969).

(21) F. E. Hruska, A. A. Smith, and J. G. Dalton, *J. Amer. Chem. Soc.*, **93**, 4334 (1971).

(22) We limit the discussion to cases where the free-energy differences are such that measurable amounts of the various conformers are present in the mixture, *e.g.*, $\Delta G^\circ \leq 2$–3 kcal mol^{-1}.

(23) Assuming that the conditions are chosen in such a way that the rate of interconversion between the conformers in solution is fast compared to the rate of crystallization.

Table VI. Difference in P between Eight Pairs of Structurally Identical (But Crystallographically Different) Ribose Residues[a]

Residue no. I	Residue no. II	Purine or pyrimidine derivative	$P_{II} - P_I$, deg	Conformation type	$\chi_{CN}{}^2{}_I$, deg	$\chi_{CN}{}^2{}_{II}$, deg
45	47	Pyrimidine	2	S[b]	57	66
53	57	Pyrimidine	2	S[c]	42	39
11	17	Pyrimidine	6	N	9	20
12	18	Pyrimidine	6	N	17	18
14	58	Pyrimidine		N + S[d]	12	56
1	6	Purine	7	N	48	38
32	37	Purine	5	S	−144	−135
28	34	Purine	13	S	49	120
25	33	Purine	22	S	47	123
4				N[e]	12	
62[f]	63[f]	Purine		N + S	69	39

[a] The compound numbers refer to those given in Tables I and II. [b] The members of the pair show different side chain C(5′)–O(5′) conformations, trans-gauche and gauche-trans, respectively. [c] Molecules 53 and 57 crystallize in different space groups. [d] Conformations of the side chain are gauche-gauche and trans-gauche, respectively. [e] The conformation of the side chain in residue 4 is gauche-trans, in residues 25 and 33 it is gauche-gauche. [f] The crystal structure of the hydrated sodium salt of adenosine triphosphate has recently been elucidated [O. Kennard, N. W. Isaacs, W. D. S. Motherwell, J. C. Coppola, D. L. Wampler, A. C. Larson, and D. G. Watson, *Proc. Roy. Soc. London, Ser. A*, **325**, 401 (1971)]. Two independent molecules (A and B, numbered 62 and 63) occur in the unit cell. We calculate the following values for P and τ_m from the published torsion angles: 62, $P = 13.4°$, $\tau_m = 41.7°$, type N; 63, $P = 155.6°$, $\tau_m = -43.5°$, type S. In both molecules the orientation of the base is anti and the orientation of the C(5′)–O(5′) side chain is gauche–gauche.

rence of cases where the asymmetric unit contains two molecules.

The material collected in the present paper contains eight pairs of structurally identical riboside residues, the members of which occupy crystallographically independent positions (seven pairs occur in the same unit cell, one pair, cytidine monophosphate, occurs in two different space groups). Nevertheless, the conformers found are identical (isotypical) and the geometries differ but slightly (Table VI). Two pairs are recorded (5-iodouridine and adenosine triphosphate) that behave differently and show a 1:1 distribution of N and S conformers in the solid, case c referred to above. Inosine deserves special mention because in the hydrated state it forms one of the seven pairs mentioned above (type S), whereas in the dehydrated crystal the molecule prefers a type N conformer. The isotypical pairs provide us with the opportunity to study the magnitude of the effect of the crystal environment on the pseudorotation. Notwithstanding the fact that the two members of each pair experience a *different intermolecular potential field*, these pairs usually show a surprisingly small (2–7°) difference in P. This difference is certainly small compared to the six times greater entire preferred ranges of P values, and occurs in three type N and in three type S cases. In two type S pairs, molecules no. 28,34 and 25,33 (Table VI), the differences are greater, 13 and 22°, respectively. However, the two pairs in question are at the same time the only ones in the series that also exhibit a different (syn *vs.* anti) conformation of the base and one may assume, therefore, that the different intramolecular forces between base and sugar play a certain role in these two cases.

Summarizing, the small effect of crystal surroundings on the geometry of isotypical ribose rings leads us to support a hypothesis that the spread in P observed for each paired (or nonpaired) N or S type conformer (which spread is again small compared to the total available pseudorotation pathway) is largely due to intramolecular forces peculiar to each molecular species (although one should keep in mind that special factors such as a favored intermolecular hydrogen bonding scheme possible in the crystal may sometimes play a decisive role). *It follows that going from solid to solution should not give rise to dramatic differences between A and A* geometries.* Hence, in our opinion, the combined information from the solid state data constitutes a reliable and accurate basis for the interpretation of physical measurements, much more so than mechanical models can provide. The existence of pairs of residues which show both type N and S conformers in no way refutes this argument but simply means that the free-energy differences between the crystal forms (and between the conformers as well) is relatively small.[18h]

Interestingly, no deoxyribosides have as yet been found that show the phenomenon of having two molecules in the asymmetric unit. In view of our conclusions concerning the relative shallowness of the deoxyribose potential energy curve it would not be surprising if pairs turned up that display a greater difference in P than is shown by the ribose ring.

Acknowledgment. This work was supported by Grant No. GM-17378 from the National Institutes of Health of the United States Public Health Service. We gratefully thank the various authors for communicating their results in advance of publication.