Conformationally "Concerted" Changes in Nucleotide Structures. A New Description Using Circular Correlation and Regression Analyses

Kunihiro Kitamura,* Akio Wakahara, Hiroshi Mizuno, Yasumasa Baba, and Ken-ichi Tomita

Contribution from the Faculty of Pharmaceutical Sciences, Osaka University, 133-1 Yamada-kami, Suita, Osaka, Japan 565, and The Institute of Statistical Mathematics, 4-6-7 Minamiazabu, Minato-ku, Tokyo, Japan 106. Received August 28, 1980

Abstract: The introduction of the concept of circular variate into the conformational analysis of nucleosides and nucleotides has allowed mathematical and statistical treatments of their angular data. Circular correlation analysis provides a survey of the quantitative relationships between torsion angles. The subsequent regression analysis allows for a prediction of the torsional pathway varied inherent in the conformational changes. The analyses using 127 known crystal structures of nucleosides and nucleotides indicate that there are high correlations in $\phi'-\chi$, $\chi-\phi$, $\chi-\tau_0$, $\tau_3-\psi'$ and $\psi'-\phi'$; i.e., each pair of torsions has a linear relation within small deviation along the circular regression line. The simultaneously changing processes of any paired torsions in a nucleotide are regulated by a total of the linear relationships of the paired torsions. These findings strongly suggest that the conformational changes of nucleosides and nucleotides are concerted.

The conformations of nucleotides are far more complex than those of peptides, because the sugar–phosphate backbone conformation is prescribed by four single-bond rotations (ϕ, ψ, ψ', and ϕ') in addition to five sugar torsions (τ_0, τ_1, τ_2, τ_3, and τ_4) and the glycosyl torsion (χ) as shown in Figure 1. Changes in nucleotide conformation are therefore likely to be caused by the effects of solvents, metal cations, or substituents. In fact, X-ray crystallographic and NMR studies of a number of nucleosides and nucleotides have reported a variety of conformations. Through these previous investigations, it was pointed out[1-4] that there are certain interrelations between χ and sugar pucker and between χ and backbone torsions. It is expected that in many cases the conformational changes among nucleotide structures obey a certain conformational pathway, and thus the analysis of nucleotide conformations has been of great concern. Torsion angles are the principal variates which characterize various molecular conformations. Accordingly histogram,[5] conformation circle,[1,2] and conformation map[1] have served as proper displays for the visualization of torsion angles in the preferred or allowed regions but have had little ability to inform us of the quantitative relationship between torsion angles. Although quantitative measurements of torsion angle relationships are useful for a prediction of conformational change, no such techniques have been suggested so far. It is clear that the difficulty mainly arises from the inherent character of angle, which makes the usual numerical treatments invalid if the angular distribution range is beyond 180°. Here it should be emphasized that the torsion angle is not the usual Euclidean-type variate because of having both directionality and periodicity. This fact prompted us to introduce the concept of the circular variate. Our statistical examination using crystal structure data provides an approach to the goal to parameterizing the relationships between torsions and also between conformers. Indeed, a novel method based on the linear type relation of circular variates led to (1) the calculation of circular correlation coefficients between torsion angles, (2) the estimation of the circular regression lines between torsions, and (3) the numerical representation of conformational similarities between nucleotides. The first and the second are of particular use in describing the extent of correlation and the mode of variation between torsion angles. The third, giving a useful method for the classification of various conformers, will be described elsewhere. Our study demonstrates that of any two different torsion angles within a nucleotide unit, certain torsion pairs show virtually linear correlations. This means that the movement of one torsion is linearly coupled with that of

the other. The conformational regularity deduced from this analysis thus has considerable use in understanding simultaneous movements between torsions within the nucleotide unit, allowing the prediction of conformational changes as a basis of the combination of torsion angles.

Methods

Ordinary correlation and regression analyses are of particular use in estimating a bivariate relationship in n observation pairs. However, variates such as torsion angles are regarded as observations on a circle and therefore have both directionality and periodicity. Hence, the statistics used for the ordinary data are generally inappropriate as descriptive measures of circular data because of following reasons. As shown in Figure 2, let the observations P_k and P_l be located on a unit circle and their values be 30° and 330°, respectively, as defined by a counterclockwise rotation from the x axis. In an angular domain [0°,360°], the arithmetic mean (180°) does not agree with the mean (0°) obtained intuitively. In another angular domain [−180°,180°], however, both mean values equal 0°. Therefore, the arithmetic mean obviously is dependent upon the selection of the defined angular domain or the choice of the zero direction. This dependence is actually seen in the variance and the correlation coefficient of circular data. Accordingly statistics must be constructed so as to be independent of both angular domains and the choice of zero directions. In this section, the concepts of mean direction and circular variance are introduced as the basic statistics of circular data. Then, our definition of the circular correlation coefficient is given together with the regression model for circular data (circular regression line).

Mean Direction and Circular Variance. Let the location of kth observation, P_k ($k = 1, ..., n$), on the circumference of a unit circle be represented by the unit vector $e^{i\theta_k}$. Then the mean direction $\bar{\theta}$ of n observations can be defined by eq 1. The right side of this

$$\frac{1}{n}\sum_{k=1}^{n} e^{i\theta_k} = \bar{\omega}e^{i\bar{\theta}} \tag{1}$$

equation shows the mean vector, that is, the centroid of the set

---

* To whom correspondence should be addressed at Osaka University.

(1) Sundaralingam, M. *Biopolymers* **1969**, 7, 821–860.
(2) Arnott, S.; Hukins, D. W. L. *Nature (London)* **1969**, 224, 886–888.
(3) Sarma, R. H.; Lee, C.-H.; Evans, F. E.; Yathindra, N.; Sundaralingam, M. *J. Am. Chem. Soc.* **1974**, *96*, 7337–7348.
(4) Berman, H. M.; Neidle, S.; Stodola, R. K. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 828–832.
(5) Arnott, S.; Hukins, D. W. L. *Biochem. J.* **1972**, *130*, 453–465.

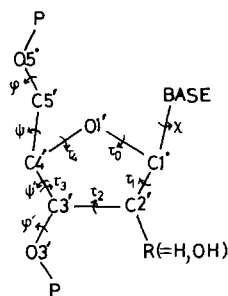**Figure 1.** A notation of ten torsion angles in nucleotide fragment. ψ' and τ_3 are torsion angles about the same bond C3'–C4', but ψ' is concerned with the backbone bond C5'–C4'–C3'–O3', while τ_3 is concerned with the sugar ring bond C2'–C3'–C4'–O1'.
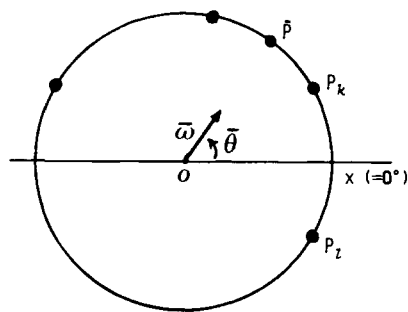

Figure 2. A schematic diagram of circular variates.

of n observations, and $\bar{\omega}$ is the mean resultant length. It is well-known that $\bar{\theta}$ and $\bar{\omega}$ are unique, and they are given by

$$\bar{\omega} = (\bar{\alpha}^2 + \bar{\beta}^2)^{1/2} \qquad \bar{\theta} = \tan^{-1}(\bar{\beta}/\bar{\alpha}) \qquad (2)$$

where

$$\bar{\alpha} = \frac{1}{n}\sum_{k=1}^{n}\cos\theta_k = \bar{\omega}\cos\bar{\theta} \qquad \bar{\beta} = \frac{1}{n}\sum_{k=1}^{n}\sin\theta_k = \bar{\omega}\sin\bar{\theta} \qquad (3)$$

This $\bar{\omega}$ has some suitable properties as a measure of the circular variance as follows: if the n observed directions are tightly bundled about the mean direction, then the length $\bar{\omega}$ will be close to unity. On the other hand, if the directions are broadly dispersed, $\bar{\omega}$ will be then close to zero. Hence, the circular variance V_c is expressed as

$$V_c = 1 - \bar{\omega} \qquad (0 \le V_c \le 1) \qquad (4)$$

Using (3), we get

$$\sum_{k=1}^{n}\cos(\theta_k - \bar{\theta}) = n\bar{\omega} \qquad (5)$$

$$\sum_{k=1}^{n}\sin(\theta_k - \bar{\theta}) = 0 \qquad (6)$$

Accordingly, from (4) and (5), the circular variance about $\bar{\theta}$ is given as

$$V_c(\theta) = 1 - \frac{1}{n}\sum_{k=1}^{n}\cos(\theta_k - \bar{\theta}) = \frac{1}{n}\sum_{k=1}^{n}[1 - \cos(\theta_k - \bar{\theta})] \qquad (7)$$

Here, it should be noted that this expression corresponds to the ordinary variance

$$\sum_{k=1}^{n}(x_k - \bar{x})^2/n$$

because the term of $1 - \cos(\theta_k - \bar{\theta})$ in (7) corresponds to half the square distance from P_k to $\bar{P}$ in Figure 2. And also (6) corresponds to the expression in the linear case

$$\sum_{k=1}^{n}(x_k - \bar{x}) = 0$$

i.e., the sum of deviations about the mean value equals zero.

Therefore $\bar{\theta}$ is analogous to the mean of an ordinary variate.

Circular Correlation Coefficient. When there is the set of n bicircular observations $(\theta_1, \phi_1), ..., (\theta_n, \phi_n)$, it is necessary to obtain a measure of correlation in order to evaluate to what degree the circular variates, θ and ϕ, are related one another.

Since the term of $\sin(\theta_k - \bar{\theta})$ in (6) is the deviation of θ_k from the mean direction, we have defined the circular covariance between two circular variates θ and ϕ by eq 8, where $\Theta_k = \theta_k - \bar{\theta}$

$$V_c(\theta, \phi) = \sum_{k=1}^{n}\sin\Theta_k\sin\Phi_k/n \qquad (8)$$

and $\Phi_k = \phi_k - \bar{\phi}$. If the relationship between Θ_k and Φ_k is $\Phi_k = b_c\Theta_k$, the circular covariance $V_c(\theta, \phi)$ has either positive or negative value depending on the sign of the gradient b_c. The circular correlation coefficient r_c is defined by eq 9. This coefficient is

$$r_c(\theta, \phi) =$$
$$\sum_{k=1}^{n}\sin\Theta_k\sin\Phi_k/(\sum_{k=1}^{n}\sin^2\Theta_k\sum_{k=1}^{n}\sin^2\Phi_k)^{1/2} \quad (|r_c| \le 1) \quad (9)$$

definitely independent of both angular domains and the choice of zero directions. If each circular variate of θ and ϕ has small variation about the mean direction of each, the right side of (9) is approximately equivalent to the ordinary correlation coefficient. Therefore, expression 9 can be regarded as a natural extension of the ordinary correlation coefficient.

Circular Regression Line. The $r_c(\theta, \phi)$ having a large absolute value suggests that there is a linear type relation between θ and ϕ. Let now an estimate of ϕ_k be represented as

$$\hat{\phi}_k = \hat{a} + \hat{b}_c\theta_k \qquad (k = 1, ..., n)$$

Then regression parameters $\hat{a}$ and $\hat{b}_c$ are estimated[6] by minimizing

$$V_c(\theta, \phi) = \frac{1}{n}\sum_{k=1}^{n}[1 - \cos(\phi_k - \hat{\phi}_k)] =$$

$$1 - \frac{1}{n}\sum_{k=1}^{n}\cos(\phi_k - \hat{a} - \hat{b}_c\theta_k)$$

because $1 - \cos(\phi_k - \hat{\phi}_k)$ is half the square distance between the points corresponding to directions ϕ_k and $\hat{\phi}_k$. This procedure is analogous to the method of a least-squares technique in the linear case. For actual calculations, an iteration method has been used.

Data

The torsion angle data of 127 crystal structures of nucleosides, nucleotides, and their analogues were used for this analysis. Table I gives the abbreviations of 107 crystal structures which were obtained from Cambridge Crystallographic Database, December 1977. The remaining 20 structures are those of dinucleotides[7–12] which have not yet been filed but had already been published when this study was began. In all of 127 structures, the sugar moiety is either β-D-(deoxy)ribofuranose or β-D-arabinofuranose, while the base moieties are adenine (33 structures), guanine (19), xanthine (1), hypoxanthine (8), uracil (36), dihydrouracil (3), thymine (5), and cytosine (22). Each torsion angle of the structures was calculated. Angle notations[1] for the nucleotide unit are illustrated in Figure 1. The degree of correlation between any two torsion angles was evaluated for all of the nucleoside and

(6) Gould, A. L. *Biometrics* **1969**, *25*, 683–700.

(7) Seeman, N. C.; Rosenberg, J. M.; Suddath, F. L.; Kim, J. J. P.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 109–144.

(8) Hingerty, B.; Subramanian, E.; Stellman, S. D.; Sato, T.; Broyde, S. B.; Langridge, R. *Acta Crystallogr., Sect. B* **1976**, *B32*, 2998–3013.

(9) Tsai, C.-C.; Jain, S. C.; Sobell, H. M. *J. Mol. Biol.* **1977**, *114*, 301–315.

(10) Rosenberg, J. M.; Seeman, N. C.; Day, R. O.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 145–167.

(11) Jain, S. C.; Tsai, C.-C.; Sobell, H. M. *J. Mol. Biol.* **1977**, *114*, 317–331.

(12) Camerman, N.; Fawcett, J. K.; Camerman, A. *J. Mol. Biol.* **1976**, *107*, 601–621.

"Concerted" Changes in Nucleotide Structures

J. Am. Chem. Soc., Vol. 103, No. 13, 1981 3901

Table I. List of Abbreviation of Nucleosides and Nucleotides Used for the Analysis

```
ACADOS    S.T.RAO, J.AMER.CHEM.SOC., 92,4963(1970).
ADBURM    A.E.V.HASCHEMEYER, ACTA CRYST., 18,525(1965).
ADENOS10  T.F.LAI, ACTA CRYST.(B), 28,1982(1972).
ADENTP10  O.KENNARD, PROC.R.SOC.,A, 325,401(1971).
ADOSHC    K.SHIKATA, ACTA CRYST.(B), 29,31(1973).
ADPOSD    M.SUNDARALINGAM, ACTA CRYST., 21,495(1966).
ADPOSM    J.KRAUT, ACTA CRYST., 16,79(1963).
ADPOSM01  S.NEIDLE, ACTA CRYST.(B), 32,1850(1976).
ADURPO10  E.SHEFTER, ACTA CRYST.(B), 25,895(1969).
AGOPCD    K.AOKI, ACTA CRYST.(B), 32,1454(1976).
ANIMPH01  K.AOKI, BULL.CHEM.SOC.JPN., 48,1260(1975).
APAPAD10  D.SUCK, ACTA CRYST.(B), 32,1727(1976).
ARADEN10  G.BUNICK, ACTA CRYST.(B), 30,1651(1974).
ARATUR10  W.SAENGER, J.AMER.CHEM.SOC., 94,621(1972).
ARBCYT10  P.TOUGARD, ACTA CRYST.(B), 30,86(1974).
ARFCYT10  J.S.SHERFINSKI, ACTA CRYST.(B), 29,192(1973).
ARFUAD    A.K.CHWANG, ACTA CRYST.(B), 30,2273(1974).
ASTHYM10  W.SAENGER, ACTA CRYSY.(B), 27,2105(1971).
BAURIP    E.SHEFTER, ACTA CRYST., 18,1067(1965).
BEURID10  E.A.GREEN, ACTA CRYST.(B), 31,102(1975).
BRADOS    S.S.TAVALE, J.MOL.BIOL., 48,109(1970).
BRAXGU    S.C.JAIN, J.MOL.BIOL., 68,1(1972).
BRGUOS01  S.S.TAVALE, J.MOL.BIOL., 48,109(1970).
BROXUR10  J.IBALL, PROC.R.SOC.,A, 295,320(1966).
BRURID10  J.IBALL, PROC.R.SOC.,A, 295,320(1966).
BURDMS    J.IBALL, PROC.R.SOC.,A, 302,225(1968).
CATHYM    K.N.TRUEBLOOD, ACTA CRYST., 14,965(1961).
CDURID    D.SUCK, BIOCHIM.BIOPHYS.ACTA, 259,157(1972).
CLDOUR    D.W.YOUNG, ACTA CRYST.(B), 29,1259(1973).
CLPURB    H.STERNGLANZ, ACTA CRYST.(B), 31,2888(1975).
CLURID10  S.W.HAWKINSON, ACTA CRYST.(B), 27,34(1971).
CYTIAC    M.SUNDARALINGAM, J.MOL.BIOL., 13,914(1965).
CYTIAC01  C.E.BUGG, J.MOL.BIOL., 25,67(1967).
CYTIDI10  S.FURBERG, ACTA CRYST., 18,313(1965).
CYTIDN    J.J.GUY, ACTA CRYST.(B), 32,2909(1976).
DADPNH10  B.S.REDDY, ACTA CRYST.(B), 31,19(1975).
DGUBCY    A.E.V.HASCHEMEYER, ACTA CRYST., 19,125(1965).
DHTHUR10  B.KOJIC-PRODIC, ACTA CRYST.(B), 30,1550(1974).
DHURID01  D.SUCK, ACTA CRYST.(B), 28,596(1972).
DMGUAN10  T.BRENNAN, J.AMER.CHEM.SOC., 94,8548(1972).
DOCYPO    M.A.VISWAMITRA, J.AMER.CHEM.SOC., 93,4565(1971).
DOCYTC    E.SUBRAMANIAN, ACTA CRYST.(B), 26,303(1970).
DOURID    A.RAHMAN, ACTA CRYST.(B), 28,2260(1972).
DOXADM    D.G.WATSON, ACTA CRYST., 19,111(1965).
DTURID    G.H.-Y.LIN, ACTA CRYST.(B), 27,961(1971).
DXCYTD    D.W.YOUNG, ACTA CRYST.(B), 31,961(1975).
ERFIMP    A.H.-J.WANG, J.AMER.CHEM.SOC., 96,1205(1974).
ESMINM    N.NAGASHIMA, ACTA CRYST.(B), 30,1094(1974).
ETCYTC    A.H.-J.WANG, J.AMER.CHEM.SOC., 98,7401(1976).
FDOURD    D.R.HARRIS, BIOPHYS.J., 4,203(1964).
GUANPF    W.MURAYAMA, ACTA CRYST.(B), 25,2236(1969).
GUANSH10  U.THEWALT, ACTA CRYST.(B), 26,1089(1970).
GUOSBH    P.TOUGARD, ACTA CRYST.(B), 30,214(1974).
HDTURD10  B.KOJIC-PRODIC, ACTA CRYST.(B), 32,1090(1976).
HICYTM    B.KOJIC-PRODIC, ACTA CRYST.(B), 32,1103(1976).
HXURID    U.THEWALT, ACTA CRYST.(B), 29,1393(1973).
IDOXUR    N.CAMERMAN, ACTA CRYST., 18,203(1965).
IMPCOH    K.AOKI, BULL.CHEM.SOC.JPN., 48,1260(1975).
INOSIN10  A.R.I.MUNNS, ACTA CRYST.(B), 26,1101(1970).
INOSND01  A.R.I.MUNNS, ACTA CRYST.(B), 26,1114(1970).
INOSPH    N.NAGASHIMA, ACTA CRYST.(B), 30,320(1974).
IURIDN10  A.RAHMAN, ACTA CRYST.(B), 26,1765(1970).
MADENS10  P.PRUSINER, ACTA CRYST.(B), 32,161(1976).
MARAFC    G.I.BIRNBAUM, J.AMER.CHEM.SOC., 97,5904(1975).
MCYTMS10  E.SHEFTER, CRYST.STRUCT.COMM., 3,209(1974).
MEURID    D.J.HUNT, ACTA CRYST.(B), 25,2144(1969).
MEYRID    D.SUCK, J.AMER.CHEM.SOC., 94,6520(1972).
MRFPUR    T.TAKEDA, ACTA CRYST.(B), 31,1202(1975).
NAINPH10  S.T.RAO, J.AMER.CHEM.SOC., 91,1210(1969).
NEBULR    T.TAKEDA, ACTA CRYST.(B), 30,825(1974).
PALAUR10  H.M.BERMAN, BIOCHEMISTRY, 12,1809(1973).
SALCYS    C.TAMURA, CHEM.LETTERS, ,1221(1973).
SDGUNP    D.W.YOUNG, ACTA CRYST.(B), 30,2012(1974).
SURIDP    M.A.VISWAMITRA, ACTA CRYST.(B), 28,1108(1972).
TCYTDH    G.H.-Y.LIN, J.AMER.CHEM.SOC., 93,1235(1971).
TGUANS10  U.THEWALT, J.AMER.CHEM.SOC., 94,8892(1972).
THIRDN10  W.SAENGER, J.MOL.BIOL., 50,153(1970).
THOPAD10  O.KENNARD, J.CHEM.SOC.(B), ,1940(1971).
THPRIB    E.SHEFTER, J.PHARM.SCI., 57,1157(1968).
THPYUR    P.H.STOTHART, ACTA CRYST.(B), 29,2237(1973).
THYDIN    D.W.YOUNG, ACTA CRYST.(B), 25,1423(1969).
TRFBIM    P.PRUSINER, ACTA CRYST.(B), 29,2328(1973).
URARAF10  P.TOLLIN, ACTA CRYST.(B), 29,1641(1973).
URDTPE    W.SAENGER, EUR.J.BIOCHEM., 46,559(1974).
UROAME    K.MORIKAWA, ACTA CRYST.(B), 31,1004(1975).
URPOAD10  J.L.SUSSMAN, J.MOL.BIOL., 66,403(1972).
XANTOS    G.KOYAMA, ACTA CRYST.(B), 32,969(1976).
```


Figure 3. A τ_3-ψ' map for 127 nucleosides and nucleotides. The open circles are markedly distributed in two regions. Those at the upper right correspond to the C2' endo group and those at the lower left correspond to the C3' endo group.

Results and Discussion

In order to understand the mechanism of the conformational change in nucleotides, it is of interest to evaluate the quantitative relationships between torsion angles. As a result, the following link-movements have been observed: (1) the link-movement between the sugar ring torsion angle τ_3 and ψ'; (2) the link-movements between the sugar torsions and the glycosyl torsion angle χ; (3) the link-movements between χ and ϕ and between χ and ϕ'.

The Link-Movements between the Sugar Ring Torsion Angle τ_3 and ψ' and between τ_3 and ψ. The relationship between the sugar ring torsion angles and ψ' (or ψ) is relevant for understanding of the stereochemistry of sugar–phosphate linkage. The distribution of the torsion angle ψ in 127 structures shows a trimodal profile as follows: 35–75° range (96 structures), 165–195° range (24), and 280–295° range (6). These distribution ranges are similar to those reported by Sundaralingam.[1] Because of trimodal distribution, ψ does not correlate to any of the sugar torsions. This observation in crystalline state is in contrast to a solution study[3] which shows that there is interrelation between ψ and χ.

The distribution of ψ' is bimodal, ranging between 73 and 98° (63 structures) and between 132 and 160° (64), of which the former corresponds to the C3' endo pucker group and the latter to the C2' endo pucker group. Although statistically the two groups should be treated separately in order to calculate the circular correlation coefficient and the circular regression line, they apparently fit into a straight line on the τ_3-ψ' map (Figure 3). When both groups are merged, the correlation coefficient between τ_3 and ψ' is great ($r_c = 0.995$) enough to indicate a link-movement, the mode of which obeys the circular regression line, $\psi' = 121° + 1.11\tau_3$.

For energy calculations of nucleotides, ψ' is usually fixed at a value near 80° for the C3' endo group and near 140° for the C2' endo group. In fact, however, ψ' distributes over relatively broad ranges for both groups along the regression line as seen in Figure 3. This implies that the relationship between τ_3 and ψ' during the conformational changes from C3' endo to C2' endo via O1' endo and C1' exo obeys the circular regression line. This

nucleotide structures by using the eq 9. When the circular correlation coefficient was high, the circular regression line was also calculated to estimate in what manner the movements of two torsion angles are linked together. A pair of torsional movements may be called a "link-movement", when (1) the correlation coefficient is high enough to calculate the regression line and (2) the torsion angles are those in a closed system like the sugar ring, which are related by the defined equation.[13]
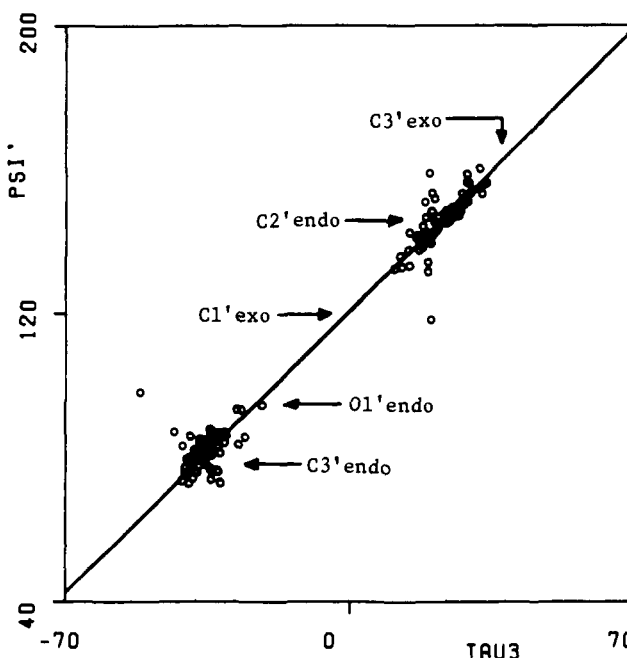
(13) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.

3902 *J. Am. Chem. Soc., Vol. 103, No. 13, 1981*
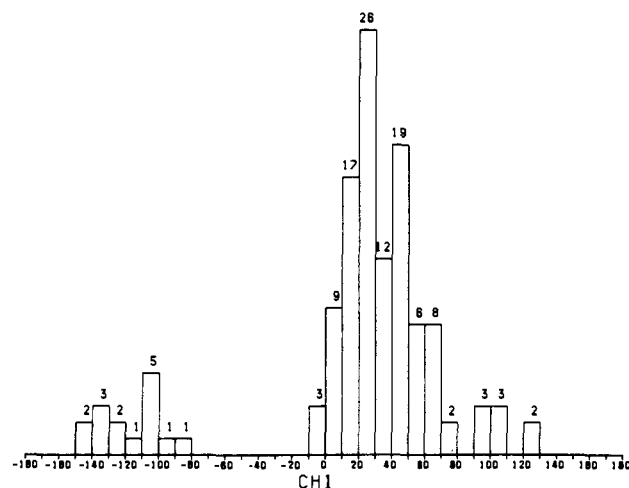
Kitamura et al.

Figure 4. A histogram of χ values at intervals of $10°$.

Table II. Circular Correlation Coefficients for χ-τ_0 and χ-τ_4

nucleoside group	N^a	circular correlation coeff	
		χ-τ_0	χ-τ_4
anti	112	-0.678	-0.303
pyrimidine	63	-0.768	-0.400
purine	49	-0.683	-0.380
adenine	26	-0.822	-0.309
guanine	16	-0.677	-0.787
syn	15	0.509	0.510

^{a} Numbers of the used compounds.

torsional pathway is similar to that deduced from the variation of the energy of the ribose ring with ψ',[14] where the energy is almost invariant in the region from $\psi' = 70°$ to $\psi' = 160°$ and that the sugar ring conformation varies continuously in this region. The circular regression line given here enables us to predict the ψ' value from the τ_3 value or vice versa.

The Link-Movements between the Sugar Ring Torsion Angles and χ. The distribution of the glycosyl torsion angle, χ, is bimodal in crystalline state (Figure 4), one group (112 structures) in the -6–$123°$ range (anti group) and the other (15) in the -149 to $-83°$ range (syn group). Each of the ring torsions τ_1, τ_2, and τ_3 is distributed bimodally, while τ_0 and τ_4 are distributed continuously. Therefore, τ_0, τ_4, and χ (anti group) were subjected to correlation analyses. As listed in Table II, a good correlation is found for χ-τ_0 but not for χ-τ_4. As an attempt to get a better correlation, the data were divided into two groups, a pyrimidine nucleoside group and a purine nucleoside group. For the pyrimidine nucleoside group, the correlation coefficient was considerably improved for χ-τ_0. For the purine nucleoside group, the correlation coefficients for both χ-τ_0 and χ-τ_4 were not improved enough to calculate circular regression lines. Therefore, further grouping into adenine nucleosides and guanine nucleosides was carried out, leading to better correlation of χ-τ_0 for the adenine nucleoside group and χ-τ_4 for the guanine nucleoside group. The results are summarized in Tables II and III and Figure 5. The regression line of χ-τ_0 (χ-τ_4) indicates that the conformational change from C3' endo to C2' endo is accompanied by the decrease of τ_0 (τ_4) and the increase of χ. On the contrary, the change from C2' endo to C3' exo is accompanied by the increase of τ_0 (τ_4) and the decrease of χ. Therefore, these regression lines can explain that the link-movement of χ-τ_0 (χ-τ_4) is regulated so as to prevent the H6 atom of pyrimidine nucleosides or the H8 atom of purine nucleosides from short contacts with the O5' atom and the hydrogen atoms attached to the C2', C3', or C5' atom.

The Link-Movements between χ and ϕ and between χ and ϕ'. An evaluation of the link-movement between χ and ϕ (ϕ') on

(14) Levitt, M.; Warshel, A. *J. Am. Chem. Soc.* **1978**, *100*, 2607–2613.
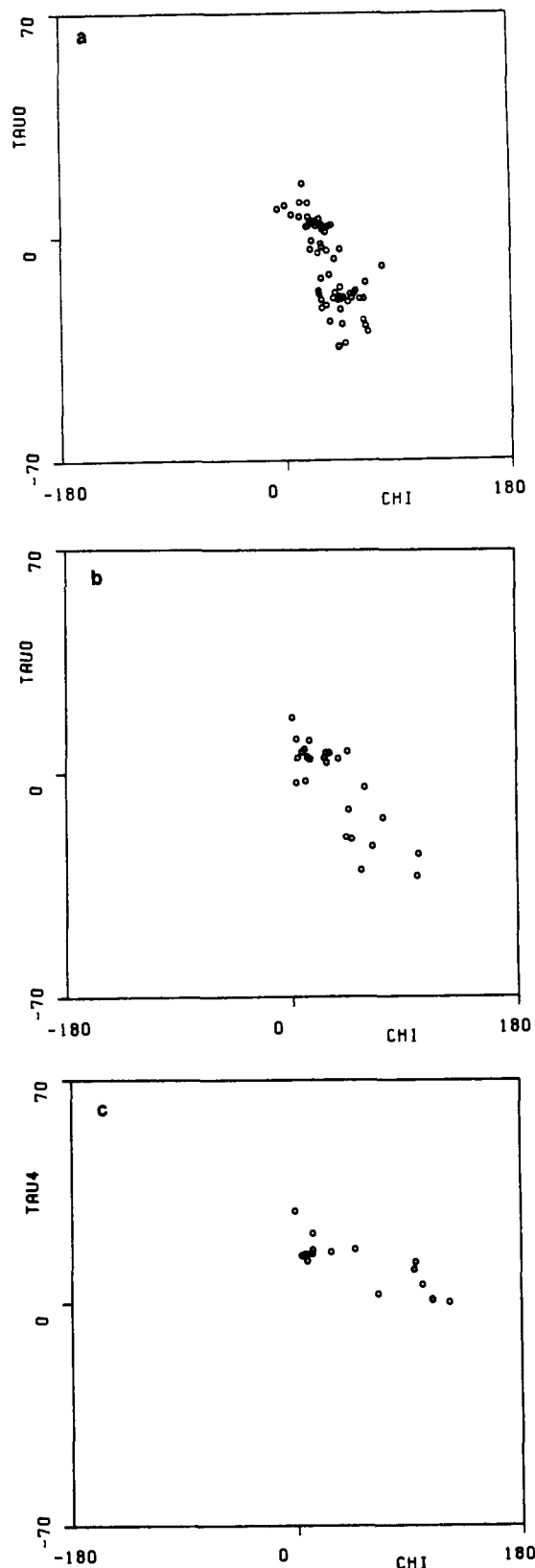

Figure 5. Scattering diagrams for χ and the sugar ring torsion angles: (a) a χ-τ_0 map for 63 pyrimidine nucleosides, (b) a χ-τ_0 map for 26 adenine nucleosides, and (c) a χ-τ_4 map for 16 guanine nucleosides.

5'(3')-nucleotides is of substantial importance to understand the mode of the backbone–base interaction. The data of 5'-nucleotides are required for the analysis of χ and ϕ. Eighteen of the analogues are of purine 5'-nucleotides and ten are of pyrimidine 5'-nucleotides. The χ values for all of the 5'-nucleotides fall in the $-6°$ to $+109°$ range and the ϕ values in the 154–237° range. All the

Table III. Circular Regression Lines^a between Torsions^b

$\tau_3 \leftrightarrow \psi'$ (All Structures, $N^c = 127, r_c{}^d = 0.995$)
$$\psi' = 121° + 1.11\tau_3 \qquad V_c{}^e = 0.002$$
$$\tau_3 = -107° + 0.89\psi' \qquad V_c = 0.002$$

$\chi \leftrightarrow \tau_0$ (Pyrimidine Nucleoside Group, $N = 63, r_c = -0.768$)
$$\tau_0 = 13° - 0.65\chi \qquad V_c = 0.012$$
$$\chi = 26° - 0.89\tau_0 \qquad V_c = 0.016$$

$\chi \leftrightarrow \tau_0$ (Adenine Nucleoside Group, $N = 26, r_c = -0.822$)
$$\tau_0 = 11° - 0.42\chi \qquad V_c = 0.010$$
$$\chi = 29° - 1.61\tau_0 \qquad V_c = 0.039$$

$\chi \leftrightarrow \tau_4$ (Guanine Nucleoside Group, $N = 16, r_c = -0.787$)
$$\tau_4 = 19° - 0.14\chi \qquad V_c = 0.003$$
$$\chi = 107° - 4.61\tau_4 \qquad V_c = 0.102$$

$\chi \leftrightarrow \phi$ (5'-Nucleotide Group, $N = 28, r_c = 0.803$)
$$\phi = 161° + 0.65\chi \qquad V_c = 0.029$$
$$\chi = -146° + 1.02\phi \qquad V_c = 0.043$$

$\phi' \leftrightarrow \chi$ (3'-Nucleotide Group, $N = 20, r_c = 0.785$)
$$\chi = -90° + 0.48\phi' \qquad V_c = 0.014$$
$$\phi' = 208° + 1.10\chi \qquad V_c = 0.033$$

$\psi' \leftrightarrow \phi'$ (3'-Nucleotide Group, $N = 20, r_c = 0.892$)
$$\phi' = 169° + 0.61\psi' \qquad V_c = 0.014$$
$$\psi' = -203° + 1.33\phi' \qquad V_c = 0.030$$

^{a} Reference 20.　^{b} Angular domain; $[-180°,180°]$ for τ_0, τ_3, τ_4, and χ and $[0°,360°]$ for ϕ, ψ', and ϕ'.　^{c} Numbers of the used crystal structures.　^{d} Circular correlation coefficients.　^{e} Circular variances.　^{f} Arrows indicate high correlations, and the broken lines indicate poor correlations between the paired torsion angles.

observed torsion angles (ψ) around the exocyclic C4'–C5' bond are in the gauche–gauche region.

Between χ and ϕ, a high circular correlation ($r_c = 0.803$) was found, and therefore the movements of χ and ϕ couple tightly together. The manner is such that the increase/decrease of ϕ is accompanied by the increase/decrease of χ, on the basis of the circular regression line, $\phi = 161° + 0.65\chi$. It should be noted that the movement of χ is linked to ϕ, although the two torsion angles are far apart from each other. This link must be spatial rather than transmissive, because ψ, located in between χ and ϕ, actually has no correlation with any torsion angles. Such a spatial link-movement probably arises from the inherent geometry that the glycosyl bond is nearly parallel to the C5'–O5' bond, if the conformations about the exocyclic C4'–C5' bond are gauche–gauche. An increase of χ at the 3'-end side of a dinucleotide short helix would result in a breakage of a parallel base stacking. To maintain the parallel base stacking, ϕ at the 3'-end side must increase concomitantly with χ. This mechanism is of particular importance for the formation of dinucleoside monophosphate–drug intercalation complexes: the concomitant changes of χ and ϕ keep the adjacent base planes parallel but lengthen the spacing between the adjacent bases. When the spacing doubles its normal value, an intercalative drug can be inserted into them. Indeed, the crystal structures of the dinucleoside monophosphate–drug intercalation complexes together with newly added ones[15-18] show that major conformational changes from RNA 11[19] are in the torsion angles

(15) Wang, A. H. J.; Nathans, J.; van der Marel, G.; van Boom, J. H.; Rich, A. *Nature (London)* **1978**, *276*, 471–474.

(16) Neidle, S.; Achari, A.; Taylor, G. L.; Berman, H. M.; Carrell, H. L.; Glusker, J. P.; Stallings, W. C. *Nature (London)* **1977**, *269*, 304–307.

(17) Wang, A. H. J.; Quigley, G. J.; Rich, A. *Nucleic Acids Res.* **1979**, *6*, 3879–3890.

(18) Shieh, H.-S.; Berman, H. M.; Dabrow, M.; Neidle, S. *Nucleic Acids Res.* **1980**, *8*, 85–97.

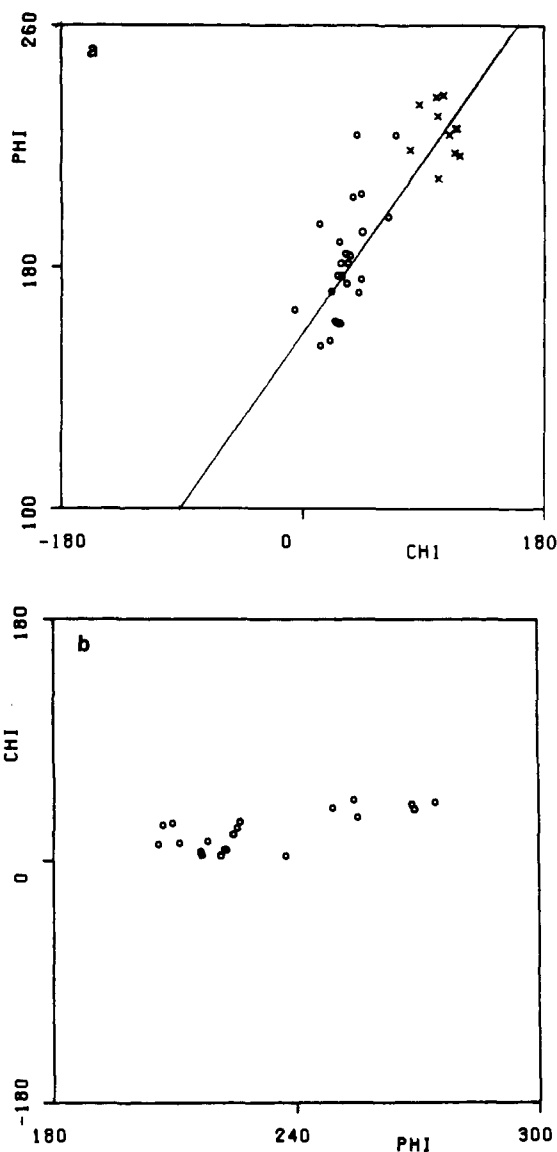(19) Arnott, S.; Dover, S. D.; Wonacott, A. J. *Acta Crystallogr., Sect. B* **1969**, *B25*, 2192–2206.

Figure 6. Scattering diagrams of (a) a χ–ϕ for 35 5'-nucleotides and (b) a ϕ'–χ for 20 3'-nucleotides. The (χ, ϕ) values for dinucleoside monophosphate–drug intercalation complexes are marked by $\times$. The regression line, $\phi = 161° + 0.61\chi$, estimated from 5'-nucleotides including additional data from intercalation complexes is shown by the solid line.

χ and ϕ at the 3'-end side. All of these (χ, ϕ) values are distributed around the upper right of the regression line, as marked by $\times$ in Figure 6a. In the course of the formation of the intercalation double helix from normal double helix, the concomitant changes of χ and ϕ would be governed by the circular regression line.

On the other hand, between ϕ' and χ, a high correlation ($r_c = 0.785$) was also found for 3'-nucleotide analogues consisting of eight purine and twelve pyrimidine nucleotides (Table III). The ϕ' values lie in the 206–274° range, while the χ values lie in the 4–45° range (Figure 6b). Both ϕ' and χ are the torsion angles which do not adjoin one another. There is the link-movement between χ and τ_0 (or τ_4) as listed in Table III, and the movement of τ_0 (or τ_4) is automatically linked to τ_3 because the sugar ring forms a closed system. Therefore the movement of ϕ' may be transmitted to χ via the sugar ring torsions. The mode of this

indirect link-movement can be estimated from the regression line, $\chi = -90° + 0.48\phi'$.

Conclusion

The present study has proposed a method to formulate conformational movements using the circular correlation and regression analyses for the torsion angles of nucleosides and nucleotides. As a result, high correlations have been found in $\phi'-\chi$, $\chi-\phi$, $\chi-\tau_0$, $\tau_3-\psi'$, and $\psi'-\phi'$. The link-movements of these pairs can be visualized in a scheme (Table III). This diagram implies that movements of a torsion angle can be transmitted one after another, and eventually one conformation is changed to another one. In other words, if any torsion except ψ is given a value, an entire conformational model can be built by transmitted operations using regression lines listed in Table III. During the operations, the conformation about the exocyclic C4'-C5' bond may be fixed to be gauche-gauche, because this conformation is strongly favored in crystals of nucleotides.

In general, the molecular structures of nucleosides or nucleotides have been regarded as conformationally "rigid"[21] probably due to the small variances of torsion angles. However, within their conformationally allowed regions, it was clearly shown that, in many cases, changes in two torsion angles have a functional relation. The simultaneous changes of any paired torsions in a nucleotide are likely to be regularly governed by the combinations of their regression lines. Certain pairs of torsion angles which were known to be correlated so far[1-4] have now been parameterized. We here understand conformational changes could be explained quantitatively by concerted movements of the paired torsion angles.

(21) Sundaralingam, M. In "Proceedings of the 5th Jerusalem Symposium Quantum Chemistry and Biochemistry", Bergmann, E. D., Pullman, B., Eds.; Academic Press: New York, 1973; pp 417-455.

(22) Yamamoto, T.; Negishi, N.; Ushimaru, M.; Tozawa, Y.; Okabe, K.; Fujiwara, S. "TOOL-IR An On-Line Information Retrieval System at an Inter-University Computer Center", Proceedings of the 2nd USA-Japan Computer Conference, 1975; pp 159-165.

Mechanism of the Reduction and Oxidation Reaction of Cytochrome c at a Modified Gold Electrode

W. John Albery,[*1a] **Mark J. Eddowes,**[1b] **H. Allen O. Hill,**[*1b] **and A. Robert Hillman**[1a]

Contribution from the Inorganic Chemistry Laboratory, South Parks Road, Oxford, England, and the Department of Chemistry, Imperial College, London, England. Received November 18, 1980

Abstract: The reduction and oxidation of cytochrome c at a gold electrode modified with an adsorbed layer of 4,4'-bipyridyl has been investigated by using rotating-disk and ring-disk electrodes. The current voltage curves for both the oxidation and reduction reactions show that the system is nearly reversible, but the rotation speed dependences of the limiting currents in either direction indicate that there are additional potential independent rate-limiting processes before and after the electron transfer. From the dependence of the limiting currents on the concentration of reactant and product, we deduce that there is considerable adsorption of both reactant and product. This adsorption step appears to be essential for rapid electron transfer between the electrode and the protein and the adsorption and desorption rates are rapid, as expected from the near reversibility of the overall electrode process. The adsorption of both reactant and product was also measured by using a ring-disk electrode with modulation of the disk current. From these results a free-energy profile for the overall electrode reaction is deduced. This free-energy profile is symmetrical and the three transition states are of about equal energy at the standard electrode potential of cytochrome c. The relationship between the binding of cytochrome c to the modified electrode and its interaction with its physiological redox partners is discussed.

Rapid electron transfer between electrodes and metalloproteins in general has proved difficult to achieve. However, as we have shown,[2] cytochrome c will undergo a rapid and reversible electrode reaction, corresponding to oxidation and reduction of the heme iron, at a gold electrode in the presence of 4,4'-bipyridyl, which appears to act by adsorbing on the electrode surface, thus providing a suitable interface for interaction with the protein. Direct-current and ac cyclic voltammetry indicates[3] that the electrode process is quasi-reversible and is essentially diffusion controlled. Alternating-current impedance measurements[4,5] have provided a value

for the standard electrochemical rate constant, showing that the electrode reaction is indeed fast with $k_s = 1.5 \times 10^{-4}$ ms^{-1}. These studies have indicated[4,5] a specific binding interaction between cytochrome c and the modified electrode surface. Using rotating-disk and ring-disk techniques we have further investigated this proposed interaction and its effects on the kinetics of the observed electrode reaction.

Experimental Section

Rotating-disk voltammograms were recorded by point-by-point current measurement by using a 4-mm-diameter gold disk electrode in a 20-mL glass cell, incorporating a conventional three-electrode system, supplied by Oxford Electrodes Ltd. The gold working electrode was polished briefly before each experiment to remove surface contaminants by using a 1-μm alumina suspension in distilled water and then washed with distilled water. The reference electrode was a saturated calomel electrode (Radiometer, Type K 401) and the counter electrode was a platinum gauze, separated from the bulk solution by a fine glass frit. All potentials are reported with respect to the saturated calomel electrode. The potential and rotation speed of the electrode were controlled by a potentiostat and a rotating-disk rig, both supplied by Oxford Electrodes. The

(1) (a) Department of Chemistry, Imperial College; (b) Inorganic Chemistry Laboratory, University of Oxford.

(2) Eddowes, M. J.; Hill, H. A. O. J. Chem. Soc., Chem. Commun. 1977, 771-772.

(3) Eddowes, M. J.; Hill, H. A. O. J. Am. Chem. Soc. 1979, 101, 4461-4464.

(4) Eddowes, M. J.; Hill, H. A. O.; Uosaki, K. J. Am. Chem. Soc. 1979, 101, 7113-7114.

(5) Eddowes, M. J.; Hill, H. A. O.; Uosaki, K. Bioelectrochem. Bioenerg. 1980, 7, 527-537.