

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231394143>

Analytic Theory of Pseudorotation in Five-Membered Rings. Cyclopentane, Tetrahydrofuran, Ribose, and Deoxyribose

ARTICLE in THE JOURNAL OF PHYSICAL CHEMISTRY · APRIL 1995

Impact Factor: 2.78 · DOI: 10.1021/j100002a019

CITATIONS

27

READS

26

2 AUTHORS, INCLUDING:



Nobuhiro Go

Japan Atomic Energy Agency

238 PUBLICATIONS 10,099 CITATIONS

[SEE PROFILE](#)

Analytic Theory of Pseudorotation in Five-Membered Rings. Cyclopentane, Tetrahydrofuran, Ribose, and Deoxyribose

Masaki Tomimoto and Nobuhiro Gō*

Department of Chemistry, Faculty of Science, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Received: June 4, 1994; In Final Form: October 24, 1994[®]

An analytic theory of pseudorotation in five-membered-rings is developed in the form of a series of approximations for cyclopentane and extended to be applicable to tetrahydrofuran, ribose, and deoxyribose. We assume the fixed values for all bond lengths. In this treatment, five bond angles should satisfy one constraint for ring closure. We derived an exact expression of the constraint. All five dihedral angles are expressed explicitly as functions of the bond angles. Pseudorotation is defined as a low-energy path of an empirical potential energy function in the space of five bond angles satisfying the ring-closure constraint. The formula of the zeroth approximation coincides exactly with that obtained previously by Geise *et al.* The first and second approximations drastically improve the numerical accuracy. At the same time formula obtained in these approximations reveal the inherent mathematical structure of the phenomenon of pseudorotation, which can be used to develop very accurate theory of pseudorotation in tetrahydrofuran. The analytic functions derived for ribose and deoxyribose allow their phase angles to deviate from ideal values group theoretically required for cyclopentane. The derived pseudorotation path agrees with experimental results, and the inherent difference between ribose and deoxyribose is identified explicitly as coming from the absence of the oxygen atom in deoxyribose.

Introduction

Methods of computer simulation of conformational dynamics of macromolecules have become recently powerful enough to be practically important in both science and technology. There are two different systems of computation depending on the choice of independent variables, one in which atomic Cartesian coordinates are independent variables, and the other in which bond lengths and bond angles are kept fixed and only rotatable dihedral angles are treated as independent variables. The latter has an advantage in that it requires a smaller number (about one-eighth in case of proteins) of independent variables than the former. In fact, it proved to be very powerful as applied to proteins, especially in the form of normal-mode analysis.¹ It has a range of further important applications such as three-dimensional structure determination from experimental NMR data² and X-ray crystallographic refinement of protein dynamic structures.^{3,4}

It is therefore desirable to extend the method so that nucleic acids can be treated. The method applied for proteins is based on the assumption that a molecule has a topology of tree, i.e., does not have a ring in it. Therefore, it cannot be applied directly to nucleic acids, unless furanose rings are treated as rigid. However, furanose rings have important roles of making up the backbone of nucleic acid molecules together with phosphate groups, and of providing attachment points for bases. They take puckered conformations, because planar conformations of five-membered rings are considerably strained. It is known that a number of puckered states are possible. Changes among them would have significant influence on the overall molecular architecture. Hence, the study of its flexibility should be important for elucidation of the dynamic aspect of nucleic acid structures. For this purpose we need to have a precise analytic expression of all dihedral angles in the ring as a function of a variable of pseudorotation.

A prototype of puckering in five-membered rings can be observed in cyclopentane. This problem has a long history of theoretical and experimental studies since Kilpatrick *et al.* first introduced the concept of pseudorotation from analysis of the unusual thermodynamic properties.^{5,6} Many spectroscopic data have supported the concept, and empirical theories have been proposed that are good enough to explain the experimental data.^{7–16} However, they are specific to cyclopentane and cannot be extended easily for a furanose ring and its derivatives.

Cyclopentane has a path of an almost free pseudorotation.⁵ This is due to its highly symmetric structure. Tetrahydrofuran, $(\text{CH}_2)_4\text{O}$, with an oxygen atom providing a permanent electric dipole moment, is accessible to various spectroscopic experiments. It has been shown to undergo fairly rapid pseudorotational motions but not as freely as in cyclopentane.^{6,16} Microwave⁹ and far-infrared¹⁷ studies have shown presence of quadruplet energy minima along the pseudorotation path and energy barriers between them hindering free pseudorotation. A Monte Carlo simulation has been carried out¹⁸ to relate pseudorotational motion in tetrahydrofuran with its thermodynamics properties. The appearance of the energy barriers in the pseudorotation path should be a result of existence of an oxygen atom in the ring. The oxygen atom also destroys the very high symmetry in cyclopentane.

Since the furanose ring has some substituents, its pseudorotation is not free but hindered. It is known to have two notably preferred puckered states in one round of its pseudorotation, named C₃-endo and C₂-endo. These two stable forms characterize two distinctly different structural families, A-form and B-form, of nucleic acid structures, respectively.¹⁹ The former belongs to the N-range (north) and the latter to the S-range (south) in the traditional polar representation of sugar pucker conformational space.²⁰ Therefore, conformational change between the two states with some energy barrier is called N–S interconversion.^{21,22} For conformational change of nucleic acid from A- to B-forms there are two possible pathways: N–W–S (north–west–south) and N–E–S (north–east–south), since the two states exist on the closed path of pseudorotation.

* To whom all correspondence should be addressed. E-mail: go@qchem.kuchem.kyoto-u.ac.jp.

[®] Abstract published in *Advance ACS Abstracts*, December 15, 1994.

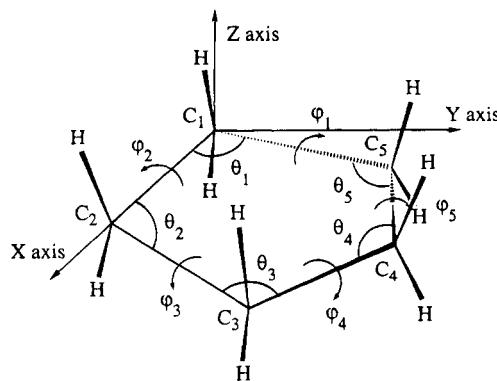


Figure 1. Definition of bond angles θ_j and dihedral angles φ_j and also of the coordinate system to express the Cartesian coordinates.

A number of studies^{5,10,23–25} have been done on mathematical description of the pseudorotation path. They describe a series of the puckered conformations on the pseudorotation path by two parameters: phase angle of pseudorotation P and amplitude of puckering τ_m . A plot of τ_m against the phase angle P is the above-mentioned polar representation. Following these works, a number of computational^{20,21,26–28} and experimental²⁹ studies have attempted to elucidate the energy profile along the pseudorotation path. All the results indicated that for the N–S interconversion the N–E–S pathway is favored over the N–W–S pathway. However, as to the value of energy barrier on the N–E–S pathway, they concluded two incompatible values. One is about 0.5 kcal/mol,²⁶ and the other is about 3–5 kcal/mol.^{20,21,27,29} The former value supports a relatively free pseudorotation, while the latter indicates a hindered pseudorotation.

What we try in this paper is to develop a logically translucent analytic treatment of pseudorotation in cyclopentane that can be easily extended to treat its derivatives. First, we develop a mathematical method to calculate a series of approximate but analytic functions to express the path of pseudorotation for cyclopentane. The highly symmetrical molecular structure of cyclopentane allows an approach based on group theory and results in a precise phase relation for variations of bond angles at five carbon atoms and also in the free rotation along the path of pseudorotation. Second, we extend the method for the pseudorotation in tetrahydrofuran. Replacement of a carbon atom by an oxygen atom in the five-membered-ring to obtain tetrahydrofuran lowers the symmetry, resulting in deviation from the precise phase relation and also in emergence of energy barrier along the path of pseudorotation. Third, we further extend the method developed for cyclopentane and tetrahydrofuran to be applicable to furanose rings, ribose, and deoxyribose and then study various properties of pseudorotation in them.

Cyclopentane

We first develop an analytic theory for the pseudorotation in cyclopentane as a prototype molecule of five-membered-rings.

Ring Closure Constraint. We treat all C–C bond lengths to have the same fixed value r . Bond angles θ_j and dihedral angle φ_j are defined as in Figure 1. Cartesian coordinates of carbon atoms are determined by placing the cyclopentane molecule as in Figure 1, i.e., C_1 at the origin, C_2 on the positive x axis, and C_3 on the first quadrangle of the x,y plane. Thus

$$C_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$C_2 = \begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}$$

$$C_3 = \begin{pmatrix} r[1 - \cos \theta_2] \\ r \sin \theta_2 \\ 0 \end{pmatrix} \quad (1)$$

Those of C_4 and C_5 are given in terms of further θ_3 and φ_3 , and θ_4 and φ_4 , respectively, by

$$C_4 =$$

$$\begin{pmatrix} r[1 - \cos \theta_2 + \cos \theta_2 \cos \theta_3 - \cos \varphi_3 \sin \theta_2 \sin \theta_3] \\ r[\sin \theta_2 - \cos \theta_3 \sin \theta_2 - \cos \varphi_3 \cos \theta_2 \sin \theta_3] \\ r[-\sin \varphi_3 \sin \theta_3] \end{pmatrix}$$

$$C_5 = \begin{pmatrix} C_x \\ C_y \\ C_z \end{pmatrix} \quad (2)$$

where

$$C_x = r[1 - \cos \theta_2 + \cos \theta_2 \cos \theta_3 - \cos \theta_2 \cos \theta_3 \cos \theta_4 - \cos \varphi_3 \sin \theta_2 \sin \theta_3 + \cos \varphi_3 \cos \theta_2 \sin \theta_2 \sin \theta_3 + \cos \varphi_3 \cos \varphi_4 \cos \theta_3 \sin \theta_2 \sin \theta_4 + \sin \varphi_3 \sin \varphi_4 \sin \theta_2 \sin \theta_4 + \cos \varphi_4 \cos \theta_2 \sin \theta_3 \sin \theta_4]$$

$$C_y = r[\sin \theta_2 - \cos \theta_3 \sin \theta_2 + \cos \theta_3 \cos \theta_4 \sin \theta_2 - \cos \varphi_3 \cos \theta_2 \sin \theta_3 + \cos \varphi_3 \cos \theta_2 \cos \theta_4 \sin \theta_3 + \cos \varphi_3 \cos \varphi_4 \cos \theta_2 \cos \theta_3 \sin \theta_4 + \cos \theta_2 \sin \varphi_3 \sin \varphi_4 \sin \theta_4 - \cos \varphi_4 \sin \theta_2 \sin \theta_4]$$

$$C_z = r[-\sin \varphi_3 \sin \theta_3 + \cos \theta_4 \sin \varphi_3 \sin \theta_3 + \cos \varphi_4 \cos \theta_3 \sin \varphi_3 \sin \theta_4 - \cos \varphi_3 \sin \varphi_4 \sin \theta_4] \quad (3)$$

Because we fix all bond lengths to have the same fixed value r , the distance between C_1 and C_5 must also be r . From this requirement and eqs 1 and 3, we have

$$0 = 3 - 2 \cos \theta_2 - 2 \cos \theta_3 + 2 \cos \theta_2 \cos \theta_3 - 2 \cos \theta_4 + 2 \cos \theta_3 \cos \theta_4 - 2 \cos \theta_2 \cos \theta_3 \cos \theta_4 - 2 \cos \varphi_3 \sin \theta_2 \sin \theta_3 + 2 \cos \varphi_3 \cos \theta_4 \sin \theta_2 \sin \theta_3 + 2 \cos \varphi_3 \cos \varphi_4 \cos \theta_3 \sin \theta_2 \sin \theta_4 + 2 \sin \varphi_3 \sin \varphi_4 \sin \theta_2 \sin \theta_4 - 2 \cos \varphi_4 \sin \theta_3 \sin \theta_4 + 2 \cos \varphi_4 \cos \theta_2 \sin \theta_3 \sin \theta_4 \quad (4)$$

It is easy to derive the following expression, which determines the value of a dihedral angle in terms of bond angles:

$$\cos \varphi_i =$$

$$\frac{1 - 2 \cos \theta_i + 2 \cos \theta_{i+2} - 2 \cos \theta_{i+4} + 2 \cos \theta_i \cos \theta_{i+4}}{2 \sin \theta_i \sin \theta_{i+4}} \quad (i = 1-5) \quad (5)$$

where suffix i in variable θ_i is to be understood to mean $i - 5$, when $i > 5$.

By using eq 5, we can eliminate dihedral angles from eq 4 to obtain

$$0 = 5 + 4S_i - 28S_i^2 - 16S_iS_{i+1} + 40S_iS_{i+2} + 16S_i^2S_{i+1} + \\ 16S_iS_{i+1}^2 - 48S_iS_{i+1}S_{i+3} + 16S_i^2S_{i+1}^2 - 32S_iS_{i+1}^2S_{i+2} + \\ 32S_iS_{i+1}S_{i+2}S_{i+3} \quad (6)$$

where

$$\begin{aligned} S_i &= s_1 + s_2 + s_3 + s_4 + s_5 \\ S_i^2 &= s_1^2 + s_2^2 + s_3^2 + s_4^2 + s_5^2 \\ S_iS_{i+1} &= s_1s_2 + s_2s_3 + s_3s_4 + s_4s_5 + s_5s_1 \\ S_iS_{i+2} &= s_1s_3 + s_2s_4 + s_3s_5 + s_4s_1 + s_5s_2 \\ S_i^2S_{i+1} &= s_1^2s_2 + s_2^2s_3 + s_3^2s_4 + s_4^2s_5 + s_5^2s_1 \\ S_iS_{i+1}^2 &= s_1s_2^2 + s_2s_3^2 + s_3s_4^2 + s_4s_5^2 + s_5s_1^2 \\ S_iS_{i+1}S_{i+3} &= s_1s_2s_4 + s_2s_3s_5 + s_3s_4s_1 + s_4s_5s_2 + s_5s_1s_3 \\ S_i^2S_{i+1}^2 &= s_1^2s_2^2 + s_2^2s_3^2 + s_3^2s_4^2 + s_4^2s_5^2 + s_5^2s_1^2 \\ S_iS_{i+1}^2S_{i+2} &= s_1s_2^2s_3 + s_2s_3^2s_4 + s_3s_4^2s_5 + s_4s_5^2s_1 + s_5s_1^2s_2 \\ S_iS_{i+1}S_{i+2}S_{i+3} &= s_1s_2s_3s_4 + s_2s_3s_4s_5 + s_3s_4s_5s_1 + \\ &\quad s_4s_5s_1s_2 + s_5s_1s_2s_3 \quad (7) \end{aligned}$$

with s_j meaning $\cos \theta_j$. Any five-membered ring has 3×5 degrees of freedom, six out of which are external (three translational and three rotational) degrees of freedom, leaving nine as internal degrees of freedom. Because we treat five bond lengths to be fixed in cyclopentane, only four remain as independent internal degrees of freedom. This agrees with what we have derived, i.e., we derived six relations, eqs 4 and 5, to be satisfied by 10 variables θ_j and φ_j ($j = 1-5$). Hereinafter we treat dihedral angles φ_j as functions of θ_i ($i = 1-5$) given by eq 5. Five bond angles satisfy one relation of eq 6, which we shall refer to as ring closure constraint. We will sometimes designate the right-hand side of eq 6 as $g(s_1, s_2, s_3, s_4, s_5)$ or simply g .

Energy Function. The actual structure of cyclopentane is determined not only by the ring closure constraint but also by its potential energy function. Various functions have been employed for theoretical studies of cyclopentane.^{7,30-32} They consist generally of five types of terms: (a) bond stretching strain, (b) bond angle bending strain, (c) dihedral angle strain, (d) nonbonded interactions, and (e) electrostatic interactions. Among them (a) is the hardest term, i.e., the energy changes most sharply for changes of positions of atoms. In this paper, we take the limit of very hard function for the bond stretching strain.²¹ In this limit we can treat bond lengths to be fixed as we did in the previous section. Among the remaining four types, (b) and (c) are harder than the others. They are expected to determine the essential structural properties of cyclopentane. In the following, concrete calculation is carried out for a specific case of the empirical energy function of the AMBER united atom type.³³ In this treatment terms of type (d) and (e) do not appear explicitly but are included implicitly in terms of type (b) and (c). Thus, the energy function E consists of the two types of terms (b) bond angle strain E_B and (c) dihedral angle strain E_D :

$$E = E_B + E_D \quad (8)$$

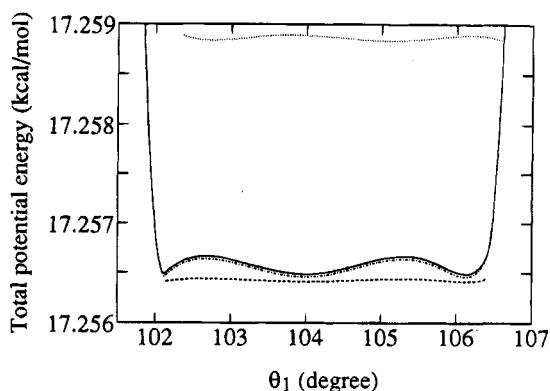


Figure 2. Total potential energy obtained by numerical minimization at each fixed value of θ_1 , solid line; by the zeroth-order approximation, dotted line; by the first-order approximation, broken line; by the second-order approximation, dash-dot line.

$$E_B = \sum_{i=1}^5 K(\theta_i - \theta_0)^2 \quad (9)$$

$$E_D = \sum_{i=1}^5 V[1 + \cos 3\varphi_i] \quad (10)$$

In the following, we will first carry out a numerical analysis and then develop analytic treatment of pseudorotation for a specific case of the numerical values taken from AMBER; $K = 63.0$ kcal/mol rad², $V = 2.0$ kcal, and $\theta_0 = 112.4^\circ$.³³ However, we will see that the analytic treatment is applicable not only for this specific case but also for a range of energy functions.

By definition bond angles θ_j take values in the range between 0 and π . Therefore $\sin \theta_j$ is positive and is equal to $(1 - \cos^2 \theta_j)^{1/2}$. Then, $\cos \varphi_j$ of eq 5, and further E_D of eq 10 (because $\cos 3\varphi = 4\cos^3\varphi - 3\cos\varphi$) are all given explicitly in terms of $s_j = \cos \theta_j$. Because E_B of eq 9 can also be regarded as a function of s_j , the total energy of eq 8 is a function of s_j , i.e., $E = E(s_1, s_2, s_3, s_4, s_5)$.

Numerical Treatment. Stable structures of cyclopentane should correspond to minima of $E = E(s_1, s_2, s_3, s_4, s_5)$ under the constraint of $g = 0$. Pseudorotation of cyclopentane should correspond to a one-dimensional low-energy path which goes through the above minima. Calculation of this low-energy path is carried out by minimizing E numerically under the constraint of $g = 0$ for a series of different given values of s_1 or θ_1 . For this calculation program IMSL is used. The obtained minimum energy and the values of other bond angles are shown in Figures 2 and 3, respectively. From these figures we see that, in the range of about 102.1° and 106.4° of θ_1 , there is a closed-loop low-energy path in the five dimensional space of $\theta_1-\theta_5$. The value of the potential energy along the path is very flat and, in fact, varies by no more than 0.0002 (kcal/mol). The values of dihedral angles along the path are calculated by eq 5 and are shown in Figure 4. Because the value of only $\cos \varphi_i$ is given in eq 5, the sign of $\sin \theta_i$ is undetermined, and therefore $-\varphi_i$ satisfies eq 5 if φ_i satisfies it. For this reason the graph of Figure 4 has a mirror symmetry with respect to the line of $\varphi_{i+1} = 0$. If both signs of five $\sin \varphi_i$ can be taken independently, there would be 2^5 different sets of φ_i ($i = 1-5$) for a given set of values of s_i ($i = 1-5$). However, we see in eq 4 that signs of neighboring $\sin \varphi_i$ are not independent. For this reason only two out of 2^5 different sets of φ_i correspond to the ring structure of cyclopentane. If one set is $(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5)$, then the other set is $(-\varphi_1, -\varphi_2, -\varphi_3, -\varphi_4, -\varphi_5)$. Corresponding two ring structures are mirror images of each other. Motion of the

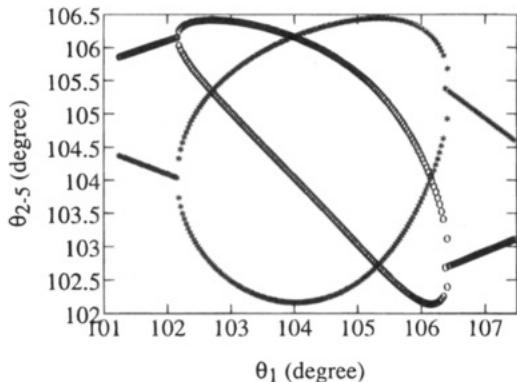


Figure 3. Values of other bond angles at the numerically obtained energy minimum point for each given value of θ_1 , circles, θ_2 and θ_5 ; stars, θ_3 and θ_4 .

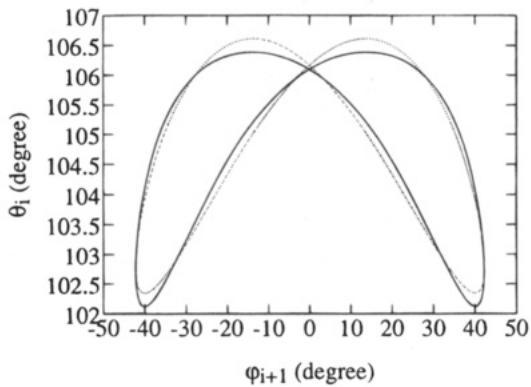


Figure 4. Relation between the value of dihedral angle φ_{i+1} and the bond angle θ_i , solid line, numerically obtained value at each fixed value of θ_1 ; thin dotted line, zeroth-order approximation.

molecule along the low-energy path is called pseudorotation. Outside this range of θ_1 , minimization of E gives a distorted structure with sharply increasing energy.

By this numerical study, we can see some of the characteristic feature of the pseudorotation in cyclopentane. In the following we develop analytic theory of the pseudorotation.

Analytic Treatment of the Zeroth-Order Approximation. We can point out that both g and E are functions of s_j ($j = 1-5$) which are invariant against the following transformation:

$$\begin{aligned} s &= (s_1, s_2, s_3, s_4, s_5) & \sigma_v s &= (s_1, s_5, s_4, s_3, s_2) \\ C_5 s &= (s_2, s_3, s_4, s_5, s_1) & C_5 \sigma_v s &= (s_5, s_4, s_3, s_2, s_1) \\ C_5^2 s &= (s_3, s_4, s_5, s_1, s_2) & C_5^2 \sigma_v s &= (s_4, s_3, s_2, s_1, s_5) \\ C_5^3 s &= (s_4, s_5, s_1, s_2, s_3) & C_5^3 \sigma_v s &= (s_3, s_2, s_1, s_5, s_4) \\ C_5^4 s &= (s_5, s_1, s_2, s_3, s_4) & C_5^4 \sigma_v s &= (s_2, s_1, s_5, s_4, s_3) \end{aligned} \quad (11)$$

where C_5 and σ_v are symbols in the group theory. Therefore, both g and E belong to the invariant representations of the C_{5v} point group. Because of this property it is expected that pseudorotational of cyclopentane can be better treated analytically in terms of the basis set of the irreducible representation of the point group C_{5v} than in terms of the original variables s_j ($j = 1-5$). The five-dimensional representation $\Gamma(s_1, s_2, s_3, s_4, s_5)$, of which s_j ($j = 1-5$) are the basis set, can be reduced into irreducible representations according to the character of C_{5v} point

TABLE 1: Character Table of the C_{5v} Point Group

	C_{5v}	E	$2C_5$	$2C_5^2$	$5\sigma_v$	
A_1	1	1		1	1	$x^2 + y^2, z^2$
A_2	1	1		1	-1	R_z
E_1	2	$2 \cos 72^\circ$	$2 \cos 144^\circ$	0	$(x, y); (R_x, R_y)$	(xz, yz)
E_2	2	$2 \cos 144^\circ$	$2 \cos 72^\circ$	0		$(x^2 - y^2, xy)$

TABLE 2: Basis Set of the Irreducible Presentations

$\Gamma(A_1)$	$u_1 = (1/5)^{1/2}[s_1 + s_2 + s_3 + s_4 + s_5]$
	$u_2 = (2/5)^{1/2}[s_2 \sin(2/5\pi) + s_3 \sin(4/5\pi) + s_4 \sin(6/5\pi) + s_5 \sin(8/5\pi)]$
$\Gamma(E_1)$	$u_3 = (2/5)^{1/2}[s_1 + s_2 \cos(2/5\pi) + s_3 \cos(4/5\pi) + s_4 \cos(6/5\pi) + s_5 \cos(8/5\pi)]$
	$u_4 = (2/5)^{1/2}[s_2 \sin(4/5\pi) + s_3 \sin(8/5\pi) + s_4 \sin(12/5\pi) + s_5 \sin(16/5\pi)]$
$\Gamma(E_2)$	$u_5 = (2/5)^{1/2}[s_1 + s_2 \cos(4/5\pi) + s_3 \cos(8/5\pi) + s_4 \cos(12/5\pi) + s_5 \cos(16/5\pi)]$

group in Table 1 as follows:

$$\Gamma(s_1, s_2, s_3, s_4, s_5) = \Gamma(A_1) + \Gamma(E_1) + \Gamma(E_2) \quad (12)$$

The basis set of the irreducible representation in terms of the original variables and its inverse relation are given in Tables 2 and 3, respectively. Representation of the group elements C_5 and σ_v in this basis set are given as follows:

$$C_5 u_1 = u_1$$

$$\sigma_v u_1 = u_1 \quad (13)$$

$$C_5 \begin{pmatrix} u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \cos(8/5\pi) & \sin(8/5\pi) \\ -\sin(8/5\pi) & \cos(8/5\pi) \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \end{pmatrix}$$

$$\sigma_v \begin{pmatrix} u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \end{pmatrix} \quad (14)$$

$$C_5 \begin{pmatrix} u_4 \\ u_5 \end{pmatrix} = \begin{pmatrix} \cos(16/5\pi) & \sin(16/5\pi) \\ -\sin(16/5\pi) & \cos(16/5\pi) \end{pmatrix} \begin{pmatrix} u_4 \\ u_5 \end{pmatrix}$$

$$\sigma_v \begin{pmatrix} u_4 \\ u_5 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_4 \\ u_5 \end{pmatrix} \quad (15)$$

Hereinafter we shall develop a theory taking u_j ($j=1-5$) as independent variables with one constraint $g = 0$. The explicit form of $g(u_1, u_2, u_3, u_4, u_5)$ can be derived from eq 6, but it is very lengthy to print. We shall later give an explicit form of an equivalent expression.

To gain some insight into the nature of the new variables u_j ($j=1-5$), we study the low-energy path of E under the constraint of $g = 0$ in the subspace spanned by u_1 , u_2 , and u_3 , or, in other words, in the space defined by $u_4 = u_5 = 0$. In this subspace an explicit form of g is given by

$$\begin{aligned} 0 &= 16u_1^4 - 16\sqrt{5}u_1^3 - \\ &\quad 4u_1^2\{5 + 2(3 + \sqrt{5})(u_2^2 + u_3^2)\} + 4u_1\{5\sqrt{5} + \\ &\quad (35 + 13\sqrt{5})(u_2^2 + u_3^2)\} + 2(7 + 3\sqrt{5})(u_2^2 + u_3^2)^2 - \\ &\quad 10(17 + 7\sqrt{5})(u_2^2 + u_3^2) + 25 \end{aligned} \quad (16)$$

In this expression variables u_2 and u_3 appear only in the combination of $(u_2^2 + u_3^2)$. For this reason we make another change of variables, now from u_1 , u_2 , and u_3 to a , b , and ω , as follows:

$$u_1 = 5^{1/2}a$$

$$u_2 = (5/2)^{1/2}b \sin(2\omega)$$

TABLE 3: Relation Inverse to Table 2

$\Gamma(A_1)$	$\Gamma(E_1)$	$\Gamma(E_2)$
$s_1 = (1/5)^{1/2}[u_1]$	$+ \sqrt{2}u_3$	$+ \sqrt{2}u_5]$
$s_2 = (1/5)^{1/2}[u_1]$	$+ \sqrt{2}u_2 \sin(2/5\pi) + \sqrt{2}u_3 \cos(2/5\pi)$	$+ \sqrt{2}u_4 \sin(4/5\pi) + \sqrt{2}u_5 \cos(4/5\pi)]$
$s_3 = (1/5)^{1/2}[u_1]$	$+ \sqrt{2}u_2 \sin(4/5\pi) + \sqrt{2}u_3 \cos(4/5\pi)$	$+ \sqrt{2}u_4 \sin(8/5\pi) + \sqrt{2}u_5 \cos(8/5\pi)]$
$s_4 = (1/5)^{1/2}[u_1]$	$+ \sqrt{2}u_2 \sin(6/5\pi) + \sqrt{2}u_3 \cos(6/5\pi)$	$+ \sqrt{2}u_4 \sin(12/5\pi) + \sqrt{2}u_5 \cos(12/5\pi)]$
$s_5 = (1/5)^{1/2}[u_1]$	$+ \sqrt{2}u_2 \sin(8/5\pi) + \sqrt{2}u_3 \cos(8/5\pi)$	$+ \sqrt{2}u_4 \sin(16/5\pi) + \sqrt{2}u_5 \cos(16/5\pi)]$

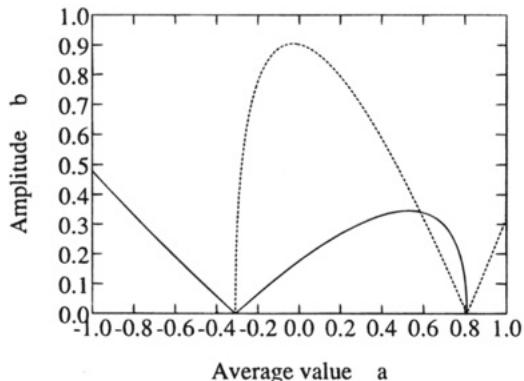


Figure 5. Amplitude b plotted as function of the average value a (eq 19), solid line, and amplitude b' or b'' against a (eq 28), broken line.

$$u_3 = (5/2)^{1/2}b \cos(2\omega) \quad (17)$$

The coefficient in front of a and b are chosen to make the following expression, eq 20, simple. The coefficient 2 in front of ω is chosen for the reason, which will be stated later. Because of eqs 13 and 14, new variables a , b , and ω change for the group operation C_5 as $(a, b, \omega) \Rightarrow (a, b, \omega + (n + 4/5)\pi)$, and for σ_v as $(a, b, \omega) \Rightarrow (a, b, \omega + m\pi)$, where n and m are arbitrary integers.

In terms of the variables introduced above g is now given by

$$\begin{aligned} 0 = 32a^4 - 32a^3 - 8a^2\{1 + (3 + \sqrt{5})b^2\} + \\ 4a\{2 + (13 + 7\sqrt{5})b^2\} + (7 + 3\sqrt{5})b^4 - \\ 2(17 + 7\sqrt{5})b^2 + 2 \end{aligned} \quad (18)$$

The constraint $g = 0$ can be solved explicitly for b as a function of a as

$$\begin{aligned} b = [(3 - \sqrt{5})/4] \sqrt{4a - (1 + \sqrt{5})} \times \\ \sqrt{2(3 + \sqrt{5})a - (9 + 5\sqrt{5}) + 2\sqrt{2}\sqrt{(25 + 11\sqrt{5}) - 2(15 + 7\sqrt{5})a}} \\ (19) \end{aligned}$$

Original variables are now expressed from Table 3 as

$$s_j = a + b \cos[2\{\omega + 4/5(j-1)\pi\}] \quad (j=1-5) \quad (20)$$

Here b is a function of a given by eq 19. This function is plotted in Figure 5. The value of b vanishes at $a = -0.309016 (= \cos 108^\circ)$. This value of a corresponds to the case of exactly planar pentagon.

By substituting eqs 19 and 20 in the expression of $E = E(s_1, s_2, s_3, s_4, s_5)$, we can express E as a function of a and ω . We seek a low-energy path of approximate pseudorotation in this two-dimensional space of a and ω . Cross section of $E(a, \omega)$ at $\omega = 0$ is given in Figure 6. A minimum exists at $a = -0.249$. The minimum value $E = 17.2588$ (kcal/mol) is very close to the numerically obtained values, ~ 17.2566 (kcal/mol), along the low-energy path of pseudorotation (Figure 2). Contour map of $E(a, \omega)$ in the area near $a = -0.249$ is shown in Figure 7. Value of a along the low-energy path in this map depends very

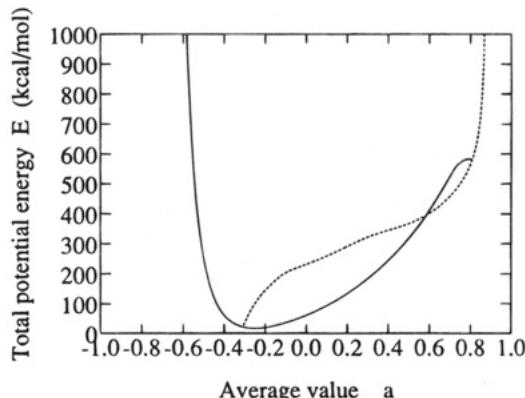


Figure 6. Cross section of the total potential energy at $\omega = 0$ plotted as a function of a , solid line, and cross section at $\omega' = 0$ or $\omega'' = 0$ against a , broken line.

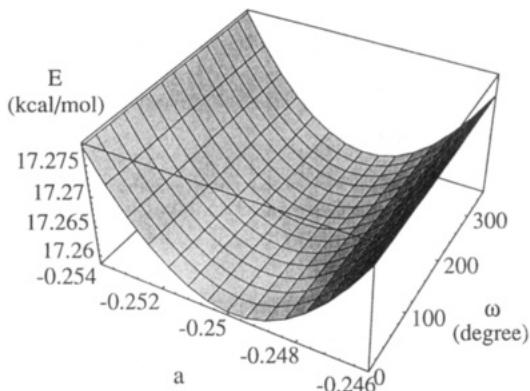


Figure 7. Contour map of the total potential energy in the area near $a = -0.249$.

little on ω . Therefore $a = \text{constant}$ is a good approximation of the low-energy path. This constant value of a is determined by requiring that the average value of E along the approximate path be minimum. This calculation is done by numerically minimizing the following quantity to obtain $a = -0.249858 \times 10^{-1}$:

$$\bar{E} = (1/2\pi) \int_0^{2\pi} E(a, \omega) d\omega \quad (21)$$

The value of E along this approximate low-energy path is shown in Figure 8 as the value of E in the zeroth-order approximation. For the above value of a , we have $b = 3.60652 \times 10^{-2}$ from eq 19. Values of dihedral angles for this set of s_i ($i=1-5$) are calculated by eqs 4 and 5 and are plotted in Figure 9 and also in Figure 4. The obtained graph in Figure 4 agrees quite well with the numerically obtained one. This indicates that the low-energy path of the zeroth-order approximation given by the analytic expression of eq 20 with $(a, b) = (-0.249858, 0.0360652)$ is already a good approximation of the pseudorotation in cyclopentane.

The expression of eq 20 generates a closed loop in the five dimensional space of $(s_1, s_2, s_3, s_4, s_5)$. The change of ω from 0 to 2π corresponds to two tours along this closed loop. When

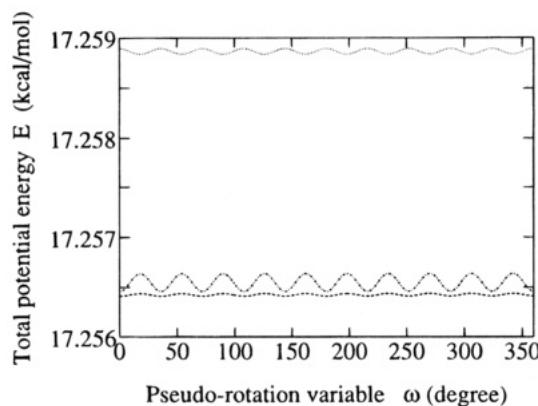


Figure 8. Total potential energy along the low-energy path of pseudorotation plotted against the pseudorotation variable ω , dotted line, zeroth-order approximation; broken line, first-order approximation; dash-dot line, second-order approximation.

the same loop is expressed in the 10-dimensional space of $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5; \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5)$, the change of ω from 0 to 2π corresponds to one tour of the closed loop in this 10-dimensional space. One tour along this closed loop in the 10-dimensional space corresponds to one complete pseudorotation. In this sense it is appropriate to call ω pseudorotation variable. When the closed loop in the 10-dimensional space is projected onto the five-dimensional space of $(s_1, s_2, s_3, s_4, s_5)$, it generates a twice degenerated closed loop. This is the reason why the coefficient two appears in front of ω of eq 20. The choice of the variable change of eq 17 was made to achieve such an expression. From Figure 9 we see that the group operations cause the following transformation of point in the 10-dimensional space:

$$\begin{aligned} C_5(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5; \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5) = \\ (\theta_2, \theta_3, \theta_4, \theta_5, \theta_1; \pm \varphi_2, \pm \varphi_3, \pm \varphi_4, \pm \varphi_5, \pm \varphi_1) \\ \sigma_v(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5; \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5) = \\ (\theta_1, \theta_5, \theta_4, \theta_3, \theta_2; \pm \varphi_2, \pm \varphi_1, \pm \varphi_5, \pm \varphi_4, \pm \varphi_3) \quad (22) \end{aligned}$$

Note that the above group operations were defined originally in the space of $(s_1, s_2, s_3, s_4, s_5)$. More usually in chemistry group operations are defined in terms of three-dimensional geometrical objects. There, σ_v is defined as the mirror image operation with respect to a plane perpendicular to the symmetry axis of C_5 operations. Then, it should correspond to $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5; \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5) \rightarrow (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5; -\varphi_1, -\varphi_2, -\varphi_3, -\varphi_4, -\varphi_5)$, which in turn corresponds to $\omega \rightarrow \omega + \pi$. The latter half of the loop of pseudorotation in the 10-dimensional space corresponds to the mirror image of the former half.

The curves of dihedral angles in Figure 9 appear to be very close to simple sinusoidal curves. In fact they can be well approximated by the following simple expression as shown in Figure 10:

$$\varphi_j = f \sin[\omega + \frac{4}{5}(j-1)\pi - \frac{2}{5}\pi] \quad (j=1-5) \quad (23)$$

where the amplitude f is determined analytically by substituting the values of s_i for $\omega = 1/10\pi$, the value which gives the maximum value of φ_2 , into eq 5 and is given by

$$\cos f = f_n/f_d$$

$$\begin{aligned} f_n = 4 - 8a + 8a^2 - 4(3 + \sqrt{5})b + 4(1 + \sqrt{5})ab + \\ (3 + \sqrt{5})b^2 \end{aligned}$$

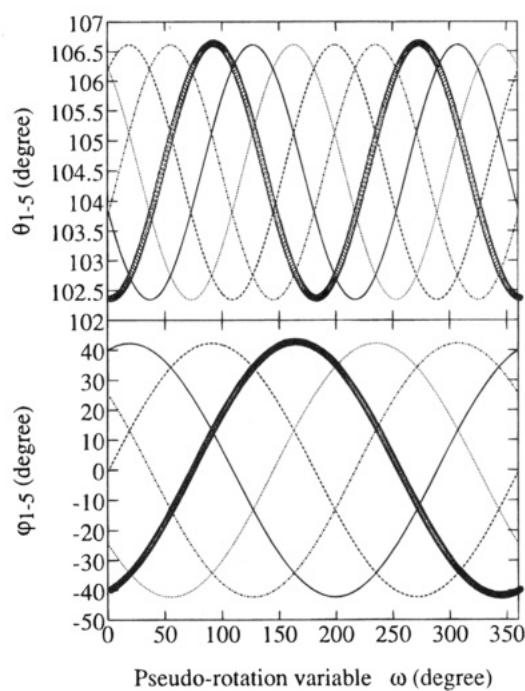


Figure 9. Values of bond angles (upper) and dihedral angles (lower) against the pseudorotation variable ω in the zeroth-order approximation. There is another set of solutions for dihedral angles, in which signs of all dihedral angles are opposite, circles for θ_1 and φ_1 ; solid line for θ_2 and φ_2 ; dotted line for θ_3 and φ_3 ; dashed line for θ_4 and φ_4 ; dash-dot line for θ_5 and φ_5 .

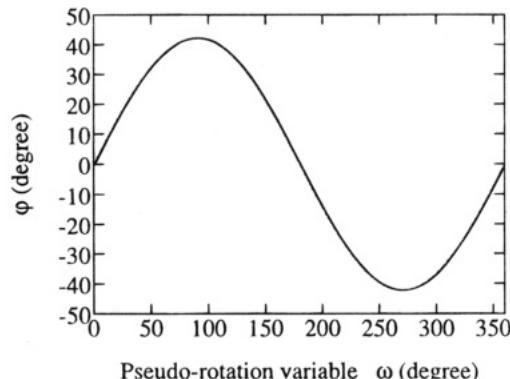


Figure 10. Comparison of the curves in Figure 9; solid line, and that of eq 23; dotted line. These two curves are in fact so similar that they are almost indistinguishable.

$$\begin{aligned} f_d = [2\{32 - 64a^2 + 32a^4 - 32(1 + \sqrt{5})ab + \\ 32(1 + \sqrt{5})a^3b - 8(3 + \sqrt{5})b^2 + 24(3 + \sqrt{5})a^2b^2 + \\ 16(2 + \sqrt{5})ab^3 + (7 + \sqrt{5})b^4\}]^{1/2} \quad (24) \end{aligned}$$

Comparison of eqs 20 and 23 with corresponding eqs in the paper by Geise *et al.*,¹⁰ reveals that the zeroth-order approximation in our paper coincides essentially with their result. Their phase angle P is related to our pseudorotation variable ω by $P = \omega - 2/5\pi$. In our analytic treatment, coefficients b in eq 20 and f in eq 23 are given as analytic functions of a , while in the empirical treatment of Geise *et al.*, they were independent parameters that were determined only empirically.

So far we have studied the low-energy path of E in the subspace spanned by u_1 , u_2 , and u_3 . The low-energy path in this subspace was found to be a good approximation to the true low-energy path of pseudorotation. We now have to examine the effects of neglect of variables u_4 and u_5 . To do this, let us find low-energy path in the subspace spanned by u_1 , u_4 , and u_5 as we did in the subspace spanned by u_1 , u_2 , and u_3 . In the

subspace spanned by u_1 , u_4 , and u_5 , an explicit form of g is given by

$$\begin{aligned} 0 = & 16u_1^4 - 16\sqrt{5}u_1^3 - \\ & 4u_1^2\{5 + 2(3 - \sqrt{5})(u_4^2 + u_5^2)\} + 4u_1\{5\sqrt{5} - \\ & (35 - 13\sqrt{5})(u_4^2 + u_5^2)\} + 2(7 - 3\sqrt{5})(u_4^2 + u_5^2)^2 - \\ & 10(17 - 7\sqrt{5})(u_4^2 + u_5^2) + 25 \quad (25) \end{aligned}$$

In this expression variables u_4 and u_5 appear only in the combination of $(u_4^2 + u_5^2)$. For this reason we make another change of variables as follows:

$$\begin{aligned} u_1 &= 5^{1/2}a \\ u_4 &= (5/2)^{1/2}b' \sin(4\omega') \\ u_5 &= (5/2)^{1/2}b' \cos(4\omega') \quad (26) \end{aligned}$$

or

$$\begin{aligned} u_1 &= 5^{1/2}a \\ u_4 &= -(5/2)^{1/2}b'' \sin(6\omega'') \\ u_5 &= (5/2)^{1/2}b'' \cos(6\omega'') \quad (26') \end{aligned}$$

Because of eqs 13 and 15, the new variables a , b' , and ω' change for the group operation C_5 as $(a, b', \omega') \rightarrow (a, b', \omega' + ((n/2) + (4/5))\pi)$, and for σ_v as $(a, b', \omega') \rightarrow (a, b', \omega' + (m/2)\pi)$, and the new variables a , b'' and ω'' as $(a, b'', \omega'') \rightarrow (a, b'', \omega'' + ((n/3) + (4/5))\pi)$ for C_5 , and $(a, b'', \omega'') \rightarrow (a, b'', -\omega'' + (m/3)\pi)$ for σ_v , where n and m are arbitrary integers. Choices of the variable change of eq 26 and 26' are made so that they change similarly as (a, b, ω) for the C_{5v} group operation. Necessity of this choice will become clear later.

In terms of the variables introduced above, g is now given (by writing b' and b'' as b) as

$$0 = 32a^4 - 32a^3 - 8a^2\{1 - (3 - \sqrt{5})b^2\} + 4a\{2 + (13 - 7\sqrt{5})b^2\} + (7 - 3\sqrt{5})b^4 - 2(17 - 7\sqrt{5})b^2 + 2 \quad (27)$$

The constraint $g = 0$ can be solved explicitly for b' or b'' as a function of a as

$$\begin{aligned} b' \text{ and } b'' = & [(3 + \sqrt{5})/4] \sqrt{4a - (1 - \sqrt{5})} \times \\ & \sqrt{2(3 - \sqrt{5})a - (9 - 5\sqrt{5}) - 2\sqrt{2}\sqrt{(25 - 11\sqrt{5})} - 2(15 - 7\sqrt{5})a} \quad (28) \end{aligned}$$

Original variables are now expressed from Table 3 as

$$s_j = a + b' \cos[4\{\omega' + 4/5(j - 1)\pi\}] \quad (j = 1-5) \quad (29)$$

or

$$s_j = a + b'' \cos[6\{\omega'' + 4/5(j - 1)\pi\}] \quad (j = 1-5) \quad (29')$$

Here b' and b'' are functions of a given by eq 28. This function is plotted in Figure 5. The value of b' (or b'') vanishes at $a = 0.809016$ ($= \cos 36^\circ$). This value of a corresponds to the case of exactly planar pentacle.

By substituting eqs 28 and 29 (or 29') in the expression of energy, we express it as a function of a and ω' (or ω''). This function again depends less on ω' (or ω'') but strongly on a . Cross section of it at $\omega' = 0$ (or $\omega'' = 0$) is given also Figure 6. We see that there is no stable minimum on this energy surface. The potential energy function in this subspace forces the molecule to assume a structure in the geometrical limit. This

figure suggests that when we extend the zeroth-order approximation to include effects from the space of u_4 and u_5 , the latter tends to make the value of a more negative.

Analytic Treatment of the First- and Second-Order Approximation. Now we proceed to find the low-energy path in the full five-dimensional space of u_1 , u_2 , u_3 , u_4 , and u_5 . Also in this five-dimensional space the variable changes of eqs 17 and 26 (or 26') are expected to be useful. Therefore, we make the following change of variables from u_1 , u_2 , u_3 , u_4 , and u_5 to a , b , c , d , and ω :

$$\begin{aligned} u_1 &= 5^{1/2}a \\ u_2 &= (5/2)^{1/2}b \sin 2\omega \\ u_3 &= (5/2)^{1/2}b \cos 2\omega \\ u_4 &= (5/2)^{1/2}(c \sin 4\omega - d \sin 6\omega) \\ u_5 &= (5/2)^{1/2}(c \cos 4\omega + d \cos 6\omega) \quad (30) \end{aligned}$$

In terms of these new variables, the original variables are given by

$$\begin{aligned} s_j = & a + b \cos[2\{\omega + 4/5(j - 1)\pi\}] + \\ & c \cos[4\{\omega + 4/5(j - 1)\pi\}] + d \cos[6\{\omega + 4/5(j - 1)\pi\}] \quad (31) \end{aligned}$$

The constraint $g = 0$ is now given exactly as

$$0 = K(a,b,c,d) + L(a,b,c,d) \cos(10\omega) + M(a,b,c,d) \cos(20\omega) \quad (32)$$

where

$$\begin{aligned} K(a,b,c,d) = & 10 + 40a - 40a^2 - 160a^3 + 160a^4 - \\ & 10(17 + 7\sqrt{5})b^2 - 10(17 - 7\sqrt{5})(c^2 + d^2) + \\ & 20(13 + 7\sqrt{5})ab^2 + 20(13 - 7\sqrt{5})a(c^2 + d^2) - \\ & 40(3 - \sqrt{5})a^2b^2 - 40(3 + \sqrt{5})a^2(c^2 + d^2) + \\ & 5(7 + 3\sqrt{5})b^4 + 140b^2(c^2 + d^2) + 5(7 - 3\sqrt{5})(c^4 + d^4) - \\ & 40(3 - \sqrt{5})b^2c + 80(1 + \sqrt{5})ab^2c - 40(3 + \sqrt{5})b^3d - \\ & 80(3 - \sqrt{5})bcd + 160(1 - \sqrt{5})abcd + \\ & 120(3 + \sqrt{5})bc^2d + 20(7 - \sqrt{5})c^2d^2 \end{aligned}$$

$$\begin{aligned} L(a,b,c,d) = & -40(3 + \sqrt{5})b^2(bc + cd) - \\ & 40(3 - \sqrt{5})b(c^3 + c^2 + d^2) + 80(1 - \sqrt{5})ab(c^2 + d^2) + \\ & 80(1 + \sqrt{5})ab^2d + 280b^2cd + 20(7 - 3\sqrt{5})cd(c^2 + d^2) - \\ & 120(3 - \sqrt{5})bcd^2 - 20(17 - 7\sqrt{5})cd + \\ & 40(13 - 7\sqrt{5})acd - 80(3 - \sqrt{5})a^2cd \end{aligned}$$

$$M(a,b,c,d) = d^2\{10(7 - 3\sqrt{5})c^2 - 40(3 - \sqrt{5})bd\} \quad (33)$$

We also have to express the total potential energy function E in terms of the new variables. Assuming that we expand the function E into the Fourier series of ω , let us study types of terms appearing in it. First, because E is originally a function of s_j ($j = 1-5$) and, in s_j , ω appears always in combination of 2ω , E must also be a function of 2ω . Second, because E is invariant for operations in the group C_{5v} , it must depend on ω in the form of 5ω . For these two reasons, we conclude that E contains Fourier terms that are multiples of 10ω .

Now our problem is to find a set of values of a , b , c , and d which gives a minimum of $E(a,b,c,d,10\omega)$ under the constraint of $g(a,b,c,d,10\omega) = 0$ at each given value of ω . Values of a , b , c , and d giving a minimum should be a function of 10ω .

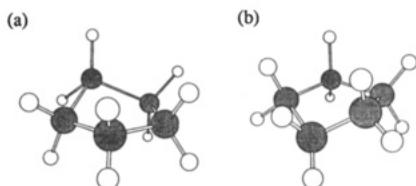


Figure 11. (a) C_s and (b) C_2 conformations along the path of pseudorotation.

TABLE 4: Values of the Parameters in Eq 31 in a Series of Approximations

	0th order	1st order	2nd order
<i>a</i>	-2.4986×10^{-1}	-2.4980×10^{-1}	-2.4980×10^{-1}
<i>b</i>	3.6065×10^{-2}	3.6059×10^{-2}	3.6059×10^{-2}
<i>c</i>		3.7289×10^{-3}	3.7321×10^{-3}
<i>d</i>			1.3800×10^{-4}

When we expand these quantities into Fourier series of 10ω and retain only constant terms, they should satisfy the constraint of E defined by eq 21.

The values of *a*, *b*, *c*, and *d* can be determined, as we did in the zeroth-order approximation, by minimizing the average of E defined by eq 21. By inspecting the expressions of K , L , and M , we realize that the zeroth-order approximation corresponds to the assumption of $c = d = 0$. Therefore, we carry out the calculation of the first-order approximation, in which only $d = 0$ is assumed in the above minimization. When even this assumption is not made, it is the calculation of the second-order approximation. The results of calculations of the first- and second-order approximations are shown in Table 4 and Figures 2 and 8. Deviations from the numerical result is within 0.0003 (kcal/mol) and 0.00003 (kcal/mol), respectively. The small deviation may not be entirely due to the approximate nature of the result, because the numerical result is a minimum value for a given fixed value of θ_1 , while the theoretical value is a minimum value for a given fixed value of the pseudorotation variable ω . These two minima can be slightly different. Values of dihedral angles are again calculated by eqs 4 and 5. When we plot the same curve as in Figure 4, the curves for the first and second approximations agree almost perfectly with the curve for the exact numerical result.

At this point we review how a conformation on the path of pseudorotation is transformed into another by the group operation of C_5 and σ_v . From eqs 13–15, it follows that, for group operations C_5 and σ_v , the variables *a*–*d* are invariant and the variable ω changes as $\omega \rightarrow \omega + (m + (4/5))\pi$ and as $\omega \rightarrow -\omega + n\pi$, respectively, where *m* and *n* are arbitrary integers. Thus, the group operations are expressed by these changes of the value of the pseudorotation variable ω . The minimum energy conformation corresponding to $\omega = 0$ is the C_s conformation and the maximum energy conformation corresponding to $\omega = \pi/10$ is the C_2 conformation; both are shown in Figure 11. Even though these two are minimum and maximum energy conformations, the energy difference between them is very small, as has long been recognized.

Tetrahydrofuran

We extend the analytic theory of pseudorotation in cyclopentane developed in the preceding section to be applicable to tetrahydrofuran.

Ring Closure Constraint. Also in tetrahydrofuran we treat all bond lengths to be fixed. In this case we have to assume

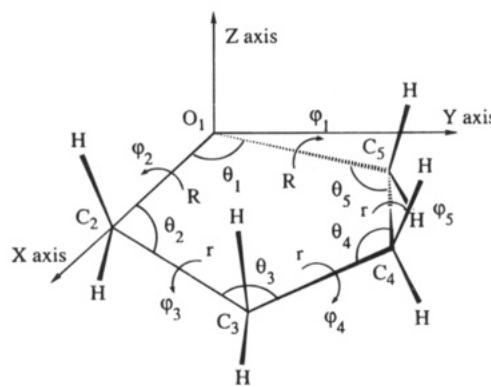


Figure 12. Definition of bond length R and r , bond angles θ_j , and dihedral angles φ_j , and also of the coordinate system to express the Cartesian coordinates.

two different bond lengths, R for $(CH_2)-O$ bonds, and r for $(CH_2)-(CH_2)$ bonds (Figure 12). Then eq 6, the ring closure constraint, becomes the following expression:

$$\begin{aligned}
 0 = & 1 - 16Q^2 + 36Q^4 - 16Q^6 - 4Q^2(1 + 6Q^2 - 8Q^4)s_1 + \\
 & 4Q(1 - 4Q^2 + 4Q^4)(s_2 + s_5) - 4(1 - 10Q^2 + 8Q^4)(s_3 + \\
 & s_4) - 4Q^4(3 + 4Q^2)s_1^2 + 4Q^2(1 - 12Q^2 + 4Q^4)(s_2^2 + s_5^2) + \\
 & 4(1 - 8Q^2)(s_3^2 + s_4^2) + 16Q^3(1 - 2Q^2)(s_1s_2 + s_5s_1) - \\
 & 16Q(s_2s_3 + s_4s_5) + 16(1 - 4Q^2 + 2Q^4)s_3s_4 + \\
 & 8Q^2(1 + 4Q^2)(s_1s_3 + s_4s_1) - 8Q(1 - 6Q^2)(s_2s_4 + s_3s_5) + \\
 & 8Q^2(1 + 8Q^2 - 4Q^4)s_5s_2 + 16Q^5(s_1^2s_2 + s_5s_1^2) - 16Q^2(1 - \\
 & 2Q^2)(s_2^2s_3 + s_4s_5^2) - 16(1 - 2Q^2)(s_3^2s_4 + s_3s_4^2) + \\
 & 16Q(s_4^2s_5 + s_2s_3^2) + 16Q^4(3 - 2Q^2)(s_5^2s_1 + s_1s_2^2) - \\
 & 48Q^3(s_1s_2s_4 + s_5s_1s_3) - 16Q^2(1 + 2Q^2)(s_2s_3s_5 + s_3s_4s_1 + \\
 & s_4s_5s_2) + 32Q(1 - Q^2)(s_2s_3s_4 + s_3s_4s_5) - 64Q^4(1 - \\
 & Q^2)s_5s_1s_2 + 16Q^6(s_1^2s_2^2 + s_5^2s_1^2) + 16Q^2(s_2^2s_3^2 + s_4^2s_5^2) + \\
 & 16s_3^2s_4^2 - 32Q^4(s_1s_2^2s_3 + s_4s_5^2s_1) - 32Q(s_2s_3^2s_4 + \\
 & s_3s_4^2s_5) + 32Q^6s_5s_1^2s_2 + 32Q^3(s_1s_2s_3s_4 + s_3s_4s_5s_1) + \\
 & 32Q^2s_2s_3s_1s_5 + 32Q^4(s_4s_5s_1s_2 + s_5s_1s_2s_3) \quad (35)
 \end{aligned}$$

where s_j and Q means $\cos \theta_j$ and R/r , respectively. The right-hand side of eq 35 will be written as $g(s_1, s_2, s_3, s_4, s_5)$.

Energy Function. As in the case of cyclopentane, pseudorotation in tetrahydrofuran should correspond to a low-energy path in the five-dimensional subspace of $(s_1, s_2, s_3, s_4, s_5)$ satisfying the constraint of eq 35. Many studies have previously been done by using empirical potential energy functions.^{7–9} Also we will develop an analytic theory of pseudorotation in tetrahydrofuran first by carrying out a numerical analysis for a particular choice of the following empirical energy function of the AMBER united atom type.³³ It consists of two types of terms: bond-angle bending strain E_B and dihedral angle strain E_D :

$$E = E_B + E_D \quad (36)$$

$$\begin{aligned}
 E_B = & K^{COC}(\theta_1 - \theta_0^{COC})^2 + \sum_{i=2,5} K^{CCO}(\theta_i - \theta_0^{CCO})^2 + \\
 & \sum_{i=3,4} K^{CCC}(\theta_i - \theta_0^{CCC})^2 \quad (37)
 \end{aligned}$$

$$E_D = \sum_{i=1}^2 V_1^{CO}[1 + \cos(3\varphi_i)] + \sum_{i=1}^2 V_2^{CO}[1 + \cos(2\varphi_i)] + \sum_{i=3}^5 V^{CC}[1 + \cos(3\varphi_i)] \quad (38)$$

The first term of E_B indicates the bond angle bending strain of $(\text{CH}_2)-\text{O}-(\text{CH}_2)$; the second term and the third term indicate that of $(\text{CH}_2)-(\text{CH}_2)-\text{O}$ and $(\text{CH}_2)-(\text{CH}_2)-(\text{CH}_2)$, respectively. The angle θ_0 is an equilibrium bond angle value for each case. The first and second terms of E_D indicate the dihedral angle strain of $(\text{CH}_2)-\text{O}-(\text{CH}_2)-(\text{CH}_2)$. While the former term represents a general torsional strain, the latter term expresses a special one for ensuring the correct gauche tendency. The third term of E_D indicates the dihedral angle strain of $(\text{CH}_2)-(\text{CH}_2)-(\text{CH}_2)-(\text{CH}_2)$. According to AMBER united atom parameters, $R = 1.425 \text{ \AA}$, $r = 1.526 \text{ \AA}$, $K^{COC} = 100.0 \text{ kcal/mol rad}^2$, $K^{CCO} = 80.0 \text{ kcal/mol rad}^2$, $K^{CCC} = 63.0 \text{ kcal/mol rad}^2$, $\theta_0^{COC} = 111.8^\circ$, $\theta_0^{CCO} = 109.5^\circ$, $\theta_0^{CCC} = 112.4^\circ$, $V_1^{CO} = 1.45 \text{ kcal}$, $V_2^{CO} = 0.10 \text{ kcal}$, and $V^{CC} = 2.0 \text{ kcal}$.³³ The term of special torsional strain has been introduced in AMBER as a perturbation to the first general term to reproduce the experimental value of the gauche-trans energy difference in methyl-ethyl ether. As we did for cyclopentane in the preceding section, the above energy function is regarded as a function of five bond angles, i.e., as $E = E(s_1, s_2, s_3, s_4, s_5)$. Reflecting the symmetry of the chemical structure of tetrahydrofuran, both $g(s_1, s_2, s_3, s_4, s_5)$ and $E(s_1, s_2, s_3, s_4, s_5)$ are invariant against the following variable transformation:

$$s = (s_1, s_2, s_3, s_4, s_5) \quad \sigma_v s = (s_1, s_5, s_4, s_3, s_2) \quad (39)$$

Numerical Treatment. Pseudorotation in tetrahydrofuran should also correspond to a low energy path of $E(s_1, s_2, s_3, s_4, s_5)$ under the constraint of $g(s_1, s_2, s_3, s_4, s_5) = 0$. This low-energy path is calculated by minimizing E for a series of θ_1 . The results are shown in Figure 13. From these figures we see that, in the range of about 107.0° and 110.6° of θ_1 , there is a closed-loop low-energy path in the five-dimensional space of $\theta_1, \theta_2, \theta_3, \theta_4$, and θ_5 . The value of potential energy along the path varies about 0.035 kcal/mol . Reflecting the invariance of E and g against the transformation of eq 39, θ_2 and θ_5 , and θ_3 and θ_4 behave similarly as we see in Figure 13b.

To see the effect of the special torsional strain, the second term of E_D , on the pseudorotation in tetrahydrofuran, we have done the numerical calculation of the low-energy path also for the case of an energy function in which it is assumed to be absent. Also in this case, a similar closed-loop low-energy path was obtained. However, it was found that the lowest energy structure is not on this closed-loop path, but on a branch extending from the loop to the right (larger value of θ_1) in Figure 13b. Therefore, the energy function without the special torsional strain does not correspond well to tetrahydrofuran.

Analytic Treatment. Now we will look for analytic descriptions of the pseudorotation in tetrahydrofuran. Because the essential feature of the phenomenon should be the same as in the case of cyclopentane, analytic description should be developed by taking an analogy to the result for it. Thus, the zeroth-order approximation should have the following form:

$$s_1 = a_1 + b_1 \cos(2\omega)$$

$$s_2 = a_2 + b_2 \cos[2\{\omega + (4/5)\pi\}]$$

$$s_3 = a_3 + b_3 \cos[2\{\omega + (8/5)\pi\}]$$

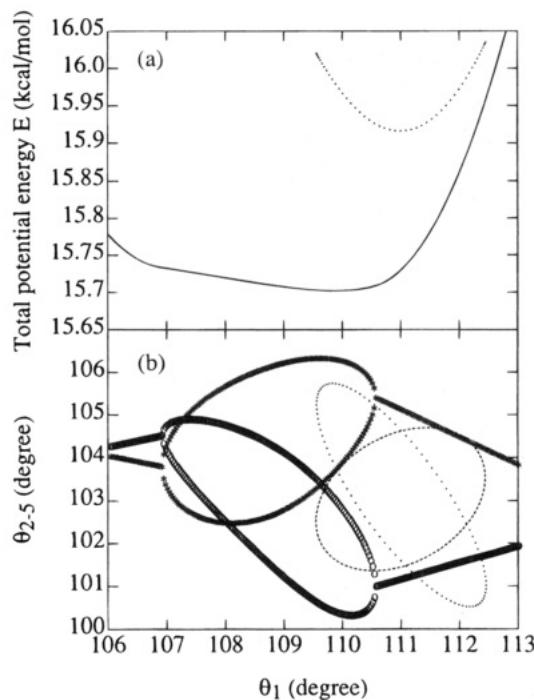


Figure 13. (a) Total potential energy obtained by numerical minimization at each fixed value of θ_1 , solid line; the zeroth-order approximation, dotted line. (b) The values of other bond angles at the numerically obtained energy minimum point for each given value of θ_1 , circles, θ_2 and θ_5 ; stars, θ_3 and θ_4 . Those for the zeroth-order approximation are given by broken lines, θ_2 and θ_5 ; by dotted lines, θ_3 and θ_4 .

$$s_4 = a_3 + b_3 \cos[2\{\omega + (12/5)\pi\}]$$

$$s_5 = a_2 + b_2 \cos[2\{\omega + (16/5)\pi\}] \quad (40)$$

This expression is taken from eq 20, but values of parameters a_i and b_i are assumed to be different to reflect the lower symmetry in tetrahydrofuran than in cyclopentane. The values for s_2 and s_5 and for s_3 and s_4 should be the same because of the symmetry of eq 39. As in the case of cyclopentane, the mirror image symmetry operation corresponds to the change of $\omega \rightarrow \omega + \pi$, and the symmetry operation of eq 39 to $\omega \rightarrow -\omega$.

In terms of the new variables, the ring-closure constraint $g = 0$ now has the following form:

$$0 = K(a_1, a_2, a_3, b_1, b_2, b_3) + \sum_{n=1}^4 L_{2n}(a_1, a_2, a_3, b_1, b_2, b_3) \cos(2n\omega) \quad (41)$$

Expressions of K and L_{2n} are very lengthy to print. We have obtained explicit expressions by using the computer program Mathematica. If we impose

$$K = 0, \quad L_2 = 0, \quad L_4 = 0, \quad L_6 = 0, \quad L_8 = 0 \quad (42)$$

then the constraint of eq 41 is always satisfied for any value of ω . Constant values of parameters a_i and b_i ($i = 1-3$) are then determined, as we did for cyclopentane in the preceding section, so as to minimize the average value of the potential energy E along the path described by variation of ω from 0 to 2π , i.e.,

$$\bar{E} = (1/2\pi) \int_0^{2\pi} E(a_1, a_2, a_3, b_1, b_2, b_3, \omega) d\omega \quad (43)$$

under the constraint of eq 42. The values of coefficients thus determined are given in Table 5.

The values of energy and other bond angles along this approximate path of pseudorotation is shown in Figure 13. The

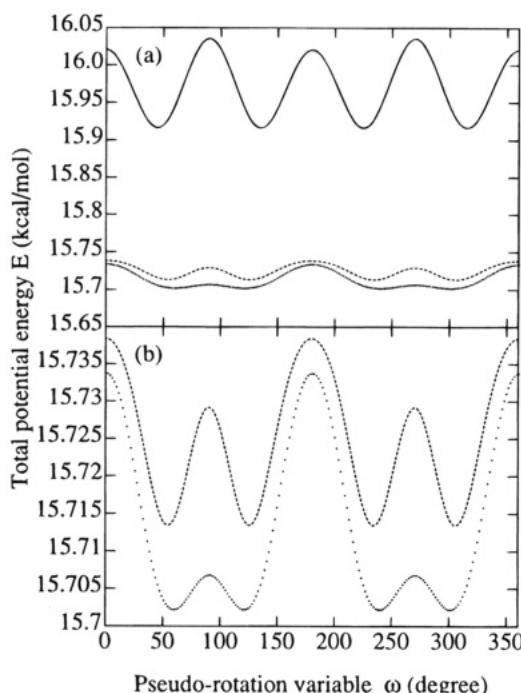


Figure 14. Values of the total potential energy plotted against the pseudorotation variable ω ; thin solid line in (a) is for the zeroth-order approximation; broken lines and dotted lines in both (a) and (b) are for the first- and the second-order approximations, respectively.

energy is also plotted against the pseudorotation variable ω in Figure 14. In the expressions of eq 40 there are six parameters, for which we impose five constraints of eq 42. Therefore, there is only one degree of freedom to be varied to minimize the quantity of eq 43. Considering this fact, it is impressive that zeroth-order approximation already gives fairly accurate description of the pseudorotation in tetrahydrofuran.

When we did the same analysis of the zeroth-order approximation for the case of the energy function without the special torsional strain, we obtained a curve of Figure 14 with only double minima and not with the quadruplet minima as in Figure 14. Experimental results indicate the presence of the quadruplet minima,^{9,17} thus indicating the necessity of the special torsional strain in the potential energy function for tetrahydrofuran.

Now let us consider higher order approximations. Comparison of eqs 20 and 31 suggests that after the zeroth-order approximation of eq 40, we should consider terms dependent on 4ω . However, we should remember that the phase terms in eq 40 are a consequence of the C_{5v} symmetry of cyclopentane. Because of the lack of such high symmetry in tetrahydrofuran, the phase terms assumed in eq 40 may not be exact. We should allow small deviations from the phases assumed there, or equivalently we should assume the following expression:

$$s_1 = a_1 + b_1 \cos(2\omega)$$

$$s_2 = a_2 + b_2 \cos\{2(\omega + \frac{4}{5}\pi + a_2\pi)\}$$

$$s_3 = a_3 + b_3 \cos\{2(\omega + \frac{8}{5}\pi + a_3\pi)\}$$

$$s_4 = a_3 + b_3 \cos\{2(\omega + \frac{12}{5}\pi - a_3\pi)\}$$

$$s_2 = a_2 + b_2 \cos\{2(\omega + \frac{16}{5}\pi - a_2\pi)\} \quad (44)$$

Here the symmetry of eq 39 is considered. We call this the first-order approximation. For this expression the ring closure

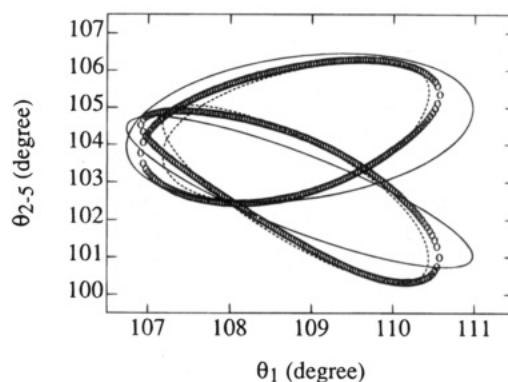


Figure 15. Values of other bond angles plotted against θ_1 ; the lower curves are for θ_2 and θ_5 , and the upper curves are for θ_3 and θ_4 ; circles, result of numerical minimization; solid line, the first-order approximation; broken line, the second-order approximation.

TABLE 5: Values of Parameters in Eqs 40, 44, and 46 for the Zeroth-, First-, and the Second-Order Approximations, Respectively

	0th order	1st order	2nd order
a_1	-3.5813×10^{-1}	-3.2299×10^{-1}	-3.2566×10^{-1}
b_1	2.3571×10^{-2}	3.5033×10^{-2}	2.6919×10^{-2}
c_1			3.4218×10^{-3}
a_2	-2.2533×10^{-1}	-2.4964×10^{-1}	-2.5093×10^{-1}
b_2	2.8427×10^{-2}	3.3811×10^{-2}	3.1668×10^{-2}
α_2		-5.1687×10^{-2}	5.4663×10^{-2}
c_2			3.5527×10^{-3}
β_2			4.9190×10^{-1}
a_3	-2.2697×10^{-1}	-2.2047×10^{-1}	-2.1719×10^{-1}
b_3	4.4294×10^{-2}	3.4312×10^{-2}	3.9858×10^{-2}
α_3		-5.8696×10^{-2}	2.0492×10^{-2}
c_3			-4.0258×10^{-3}
β_3			-6.4709×10^{-1}

constraint has the following form:

$$0 = K(a_1, a_2, a_3, b_1, b_2, b_3, \alpha_2, \alpha_3) + \sum_{n=1}^4 L_{2n}(a_1, a_2, a_3, b_1, b_2, b_3, \alpha_2, \alpha_3) \cos(2n\omega) \quad (45)$$

Again, if we impose eq 42, then the constraint of eq 45 is always satisfied for any value of ω . The values of parameters are determined so as to minimize the average value of the potential energy given by an expression corresponding to eq 43. The values of parameters thus obtained are given in Table 5. We see that deviations of phase angles α_2 and α_3 from ideal values are small. The value of energy is shown in Figures 14 and 15. The values of other bond angles are shown in Figure 15. Values of dihedral angles are shown in Figure 16.

The second-order approximation considers terms dependent on 4π . Thus we assume

$$s_1 = a_1 + b_1 \cos(2\omega) + c_1 \cos(4\omega)$$

$$s_2 = a_2 + b_2 \cos\{2(\omega + \frac{4}{5}\pi + \alpha_2\pi)\} + c_2 \cos\{4(\omega + \frac{4}{5}\pi + \beta_2\pi)\}$$

$$s_3 = a_3 + b_3 \cos\{2(\omega + \frac{8}{5}\pi + \alpha_3\pi)\} + c_3 \cos\{4(\omega + \frac{8}{5}\pi + \beta_3\pi)\}$$

$$s_4 = a_3 + b_3 \cos\{2(\omega + \frac{12}{5}\pi - \alpha_3\pi)\} + c_3 \cos\{4(\omega + \frac{12}{5}\pi - \beta_3\pi)\}$$

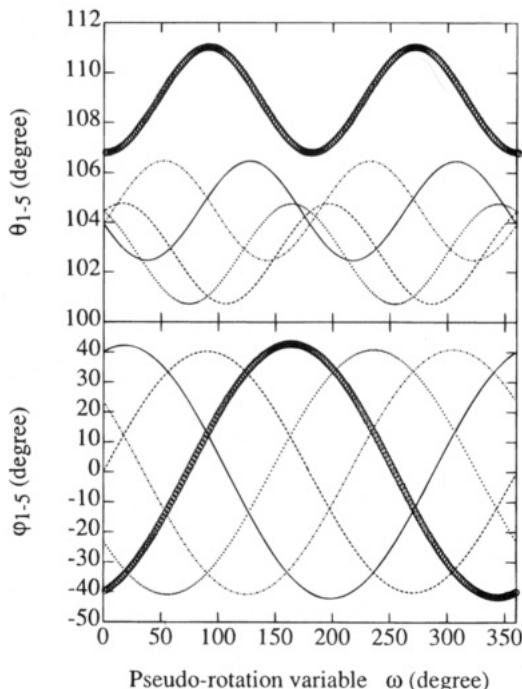


Figure 16. Values of bond angles (upper) and dihedral angles (lower) against the pseudorotation variable ω in the first approximation. There is another set of solutions for dihedral angles, in which signs of all dihedral angles are opposite, circles for θ_1 and φ_1 ; solid line for θ_2 and φ_2 ; dotted line for θ_3 and φ_3 ; dashed line for θ_4 and φ_4 ; dash-dot line for θ_5 and φ_5 .

$$s_5 = a_2 + b_2 \cos\{2(\omega + 16/5\pi - \alpha_2\pi)\} + c_2 \cos\{4(\omega + 16/5\pi - \beta_2\pi)\} \quad (46)$$

Here the symmetry of eq 39 is considered. For this expression the ring closure constraint has the following form:

$$0 = K(a_1, a_2, a_3, b_1, b_2, b_3, \alpha_2, \alpha_3, c_1, c_2, c_3, \beta_2, \beta_3) + \sum_{n=1}^8 L_{2n}(a_1, a_2, a_3, b_1, b_2, b_3, \alpha_2, \alpha_3, c_1, c_2, c_3, \beta_2, \beta_3) \cos(2n\omega) \quad (47)$$

Now, we impose

$$K = 0 \quad L_{2n} = 0 \quad (n = 1-8) \quad (48)$$

and minimize the average value of E . The values of parameters thus obtained are given in Table 5. The value of energy and the values of other bond angles are shown in Figures 14 and 15.

Ribose and Deoxyribose

For the furanose rings, we will follow essentially the method developed for tetrahydrofuran in the previous section. Numbering of atoms, etc., is given in Figure 17.

Analytic Treatment. The ring closure constraint is identical with eq 35. In the preceding section for tetrahydrofuran we have shown that the first-order analytic expression is accurate enough for most practical purposes. Therefore, for the furanose rings we assume the following first-order expression:

$$s_j = a_j + b_j \cos[2\{\omega + 4/5(j-1)\pi + \alpha_j\pi\}] \quad (j = 0-4) \quad (49)$$

Here $\alpha_0 = 0$, and a_j, b_j ($j = 0-4$) and α_j ($j = 1-4$) are parameters, whose values are to be determined below, and ω is the pseudorotation variable which varies from 0 to 2π . We should note here that our pseudorotation variable ω is related

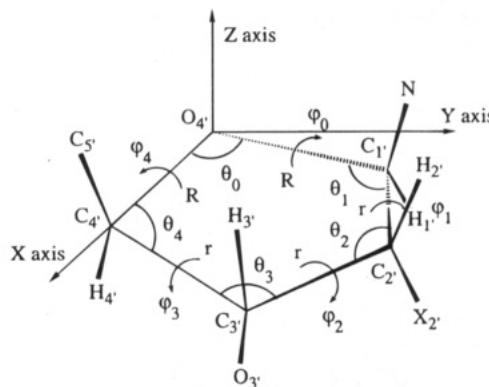


Figure 17. Schematic representation of the furanose ring which shows ring atoms, side-chain atoms, bond lengths R and r , valence bond angles θ_j , torsion angles φ_j , and coordinate system to express the Cartesian coordinates. Systems treated here are ribose (X_2' = oxygen) and deoxyribose (X_2' = hydrogen). Note that the numbering of atoms along the ring has been changed from those for cyclopentane and tetrahydrofuran, i.e., from anticlockwise to clockwise. Correspondence between two numbers, bond angles and dihedral angles have also been changed. These changes have been made to conform with the usual conventions.

to the phase angle of pseudorotation, P , by $\omega = P + 1/2\pi$.²³ The change from eq 44 in the preceding section for tetrahydrofuran to the above eq 49 reflects the lower molecular symmetry of furanose rings as compared with tetrahydrofuran. Eq 49 transforms the ring closure constraint into

$$g_1 + \sum_{k=1}^4 g_{2k} \sin(2k\omega) + \sum_{k=1}^4 g_{2k+1} \cos(2k\omega) = 0 \quad (50)$$

respectively, where g_j ($j = 1-9$) functions of the 14 parameters.

In the following we develop an analytic theory of pseudorotation in ribose and deoxyribose by assuming a specific empirical energy function of AMBER potential³³ as we did for tetrahydrofuran. It should be noted, however, that our analytic formulation is quite general and is applicable to any reasonable empirical functions. In the treatment to be developed below, we neglect atoms attached to C_5' , O_3' , X_2' , and N (see Figure 17) and derive analytic expressions of pseudorotation in ribose and deoxyribose rings in which these atoms are neglected. In computer programs to treat nucleic acids, these neglected atoms and all interactions involving them should be restored. However, these restored interactions are assumed to have negligible effects on the pseudorotation path, so that the path determined without them can be used. For the conformational energy of the furanose molecules, the AMBER function consists of five types of terms: bond angle strain (E_B), torsion angle strain (E_T), gauche effect (E_G), electrostatic interactions (E_E) and nonbonded interactions (E_N). All 1-4 and 1-5 atom pairs, i.e., atom pairs with three bonds in between the four or more bonds in between, respectively, are considered in the electrostatic and nonbonded interactions. Those for 1-4 interaction terms are reduced by multiplying a factor of 0.5.³³ Since bond angles between side chains and main chains, torsion angles and inter-atomic distances can be represented by functions of the five bonds angles on the ring, the total potential energy function can be expressed as

$$E(s_1, s_2, s_3, s_4, s_5) = E_B + E_T + E_G + E_E + E_N \quad (51)$$

To be more precise, for a given set of bond angles or for a given $(s_1, s_2, s_3, s_4, s_5)$, two sets of dihedral angles $(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5)$ and $(-\varphi_1, -\varphi_2, -\varphi_3, -\varphi_4, -\varphi_5)$ are possible, as was discussed in detail in the case of cyclopentane. Conversion from one to the other set of the dihedral angles corresponds to a change of the variable of pseudorotation from ω to $\omega + \pi$. Among the

TABLE 6: Values of the 14 Parameters in Eq 49 for Three Cases of Energy Functions

	ribose			deoxyribose		
	case A ^a	case B ^b	case C ^c	case A ^a	case B ^b	case C ^c
a_0	-3.204×10^{-1}	-3.302×10^{-1}	-3.210×10^{-1}	-3.194×10^{-1}	-3.325×10^{-1}	-3.273×10^{-1}
b_0	3.281×10^{-2}	2.069×10^{-2}	1.938×10^{-2}	3.417×10^{-2}	2.364×10^{-2}	2.519×10^{-2}
a_4	-2.493×10^{-1}	-2.596×10^{-1}	-2.642×10^{-1}	-2.473×10^{-1}	-2.551×10^{-1}	-2.582×10^{-1}
b_4	3.508×10^{-2}	3.128×10^{-2}	3.459×10^{-2}	3.608×10^{-2}	2.891×10^{-2}	2.787×10^{-2}
a_4	1.130×10^{-2}	5.515×10^{-2}	5.235×10^{-2}	-2.915×10^{-2}	3.225×10^{-2}	3.225×10^{-2}
a_3	-2.209×10^{-1}	-2.159×10^{-1}	-2.079×10^{-1}	-2.277×10^{-1}	-2.323×10^{-1}	-2.330×10^{-1}
b_3	3.736×10^{-2}	4.162×10^{-2}	4.585×10^{-2}	3.578×10^{-2}	3.531×10^{-2}	3.336×10^{-2}
α_3	1.023×10^{-3}	1.685×10^{-2}	3.023×10^{-3}	-1.735×10^{-2}	1.130×10^{-2}	2.330×10^{-2}
a_2	-2.150×10^{-1}	-2.203×10^{-1}	-2.163×10^{-1}	-2.126×10^{-1}	-2.171×10^{-1}	-2.154×10^{-1}
b_2	3.541×10^{-2}	4.036×10^{-2}	4.343×10^{-2}	3.177×10^{-2}	3.357×10^{-2}	3.298×10^{-2}
α_2	-2.925×10^{-2}	5.140×10^{-2}	7.365×10^{-2}	-3.840×10^{-2}	-3.840×10^{-2}	-2.030×10^{-2}
a_1	-2.557×10^{-1}	-2.593×10^{-1}	-2.640×10^{-1}	-2.618×10^{-1}	-2.668×10^{-1}	-2.692×10^{-1}
b_1	3.213×10^{-2}	2.793×10^{-2}	2.766×10^{-2}	3.043×10^{-2}	2.676×10^{-2}	2.863×10^{-2}
α_1	-2.525×10^{-2}	-7.830×10^{-2}	-1.032×10^{-1}	-2.103×10^{-2}	5.035×10^{-2}	-3.265×10^{-2}

^a Total potential energy consists of E_B , E_T , and E_G . ^b E_B , E_T , E_G , E_E ($\epsilon = 4r$), and E_N . ^c E_B , E_T , E_G , E_E ($\epsilon = r$), and E_N .

five energy terms in the above eq 51, E_E and E_N , which are functions of interatomic distances, depend on a choice of one set of the dihedral angles put of the two possibilities. Such a choice is made in this paper that the phase angle of pseudorotation $P = 0$ corresponds to C₂-exo-C_{3'}-endo. Lack of mirror symmetry in ribose and deoxyribose rings necessitates such a choice and deprives from the path of pseudorotation an invariance for a change of $\omega \rightarrow \omega + \pi$. Equation 49 transforms the energy function of eq 51 into a function of the 14 parameters:

$$E(a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \omega) \quad (52)$$

The determination of values of these parameters is done by requiring that the average value of the conformational energy along the path of pseudorotation should be minimum (cf. eqs 21 and 43). The minimization of the energy function is to be done under the nine constraints, $g_j = 0$ ($j = 1-9$).

We apply the above procedure to both ribose and deoxyribose by assuming, following Weiner *et al.*,³³ that (case B) the dielectric constant ϵ in the electrostatic interaction is given by $4r$, where r is an interatomic distance in angstroms. To see the effects of nonbonded and electrostatic interactions in pseudorotation, we apply the above procedure also for a case where (case C) the dielectric constant is given by $\epsilon = r$ rather than by $\epsilon = 4r$ and for a case where (case A) the electrostatic and nonbonded interactions are assumed absent altogether.

Parameters. The sets of the 14 parameters in eq 49 determined for the above six cases are listed in Table 6. From these numerical values we see that deviations of phase angles, α_j , from the ideal values group theoretically required for cyclopentane (eq 31) are slightly larger than those found for tetrahydrofuran (Table 5) but still remain small.

At this point it may be appropriate to compare our analytic expression of path of pseudorotation, eq 49, with more commonly assumed ones.^{5,23-25} Differences are that we allow the deviations from ideal phase angles and j dependence of parameters a_j and b_j . The necessity and improvements attained by these differences have been clarified in the preceding sections for cyclopentane and tetrahydrofuran. Another important difference is that the values of the parameters in Table 6 are determined by requiring the pseudorotation path to be very accurately the low-energy path for the assumed empirical energy function of eq 51. This is in contrast to the treatment of, for example, by Olson,²⁷ in which an analytic but empirical expression of the pseudorotation path is assumed and an empirical energy function is used only to evaluate the value of energy on the assumed path. In such a treatment the pseudorotation path can be somewhat different from the actual low energy path of the assumed empirical energy function.

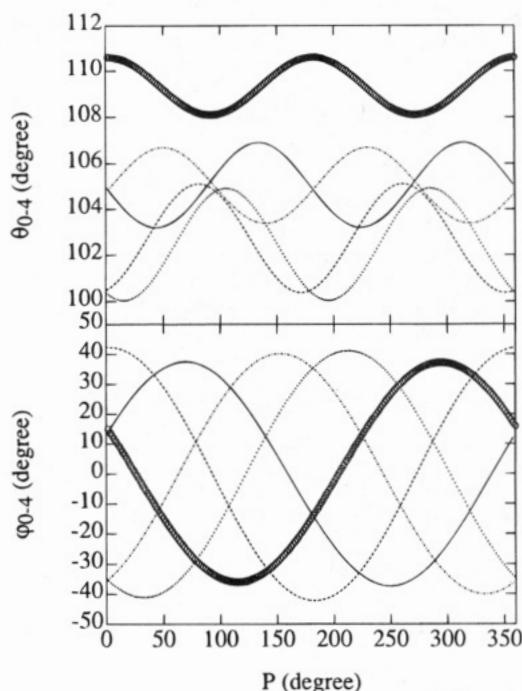


Figure 18. Values of bond angles (upper) and dihedral angles (lower) in ribose against the phase angle of pseudorotation P calculated for the energy function of case B, circles for θ_0 and ϕ_0 ; solid line for θ_4 and ϕ_4 ; dotted line for θ_3 and ϕ_3 ; dashed line for θ_2 and ϕ_2 ; dash-dot line for θ_1 and ϕ_1 .

Bond Angles and Dihedral Angles. Values of the bond angles and dihedral angles along the pseudorotation path in the energy function of case B are shown in Figure 18 for ribose and in Figure 19 for deoxyribose. Agreement with the experimental data summarized in Figures 5 and 6 of paper by Olson and Sussman²¹ is good. In Table 6 we see that the amplitudes of variation of θ_0 , which is the parameter b_1 , depends significantly on the choices of energy function both for ribose and deoxyribose. In units of degree, the amplitudes in cases A, B, and C are 1.9, 1.1, and 1.1 for ribose and 2.0, 1.2, and 1.2 for deoxyribose. Corresponding to these changes, amplitudes of variation of ϕ_0 and ϕ_4 also depend significantly on the choices of energy functions. In units of degree, those in cases A, B, and C are 41.9, 37.4, and 37.8 for ribose and 42.2, 37.4, and 37.6 for deoxyribose. These changes mean that out-of-plane motion of the oxygen atom is suppressed by the electrostatic and nonbonded interactions, making the furanose ring more planar. Because the furanose ring is a very crowded structure, the nonbonded interactions have an effect of preferring

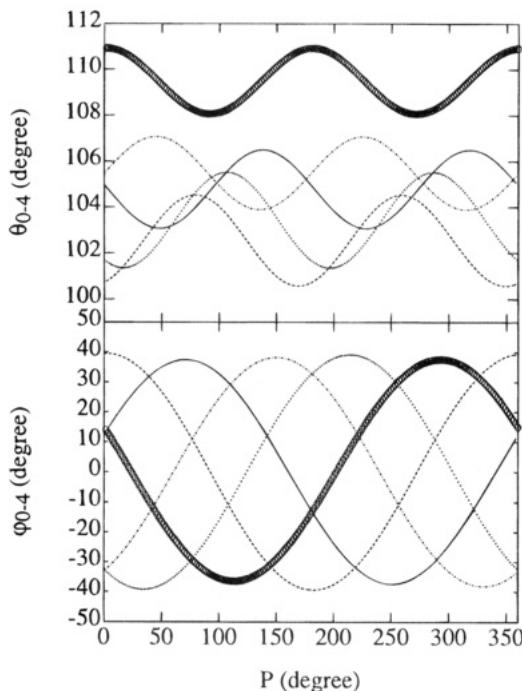


Figure 19. As in Figure 18, but for deoxyribose.

less crowded structures, i.e., more planar ones. Since all four carbon atoms but the oxygen atom in the rings have side chains, the preference for the planar structure becomes most pronounced at the oxygen atom.

Potential Energy. Values of potential energy along the path of pseudorotation for the three choices of the conformational energy function and their breakdown are shown for ribose in Figure 20 and for deoxyribose in Figure 21. In all figures, values of bond angle strain and torsional barriers depend relatively strongly on the phase angle of pseudorotation P . The former tends to stabilize O_4' -endo (E range in the polar representation; $P = \frac{1}{2}\pi$) and O_4' -exo (W range; $P = \frac{3}{2}\pi$) conformations. On the other hand, the latter tends to stabilize C_2' -exo- C_3' -endo (N range; $P = 0$) and C_2' -endo- C_3' -exo (S range; $P = \pi$) conformations. Furthermore, it is remarkable that their effects cancel each other almost completely. In Figures 20a and 21a, the total potential energy is almost identical with the gauche effect energy, indicating the importance of the gauche effect in the pseudorotation of furanose rings.²¹

The influence of the electrostatic and nonbonded interactions on the pseudorotation paths can be recognized by comparing curves in case A with those in cases B and C. As was already noted, the nonbonded interactions tend to make the ring planar. Associated with this tendency, we see that the phase angle dependence of the bond angle strain, the torsional barriers, and the gauche effect term increases, with the former two still canceling each other almost completely. In cases of B and C we see that the electrostatic and nonbonded interactions have some contributions to the total energy along the path of pseudorotation.

Ribose. Total potential energy have two stable states and two energy barriers on the pseudorotation path (Figure 20b,c). The stable states are in the N and S ranges, and N range (C_3' -endo) is more stable than S range (C_2' -endo) by 0.4 kcal/mol in Figure 20b and 1.5 kcal/mol in Figure 20c. The barriers are in the E and W ranges. The values of them are 1.9 and 2.8 kcal/mol, respectively, in Figure 20b, and 3.4 and 3.5 kcal/mol, respectively, in Figure 20c. Therefore, the N-E-S path is clearly more favored in Figure 20b than the N-W-S path in the N-S interconversion. As reviewed in the Introduction, two

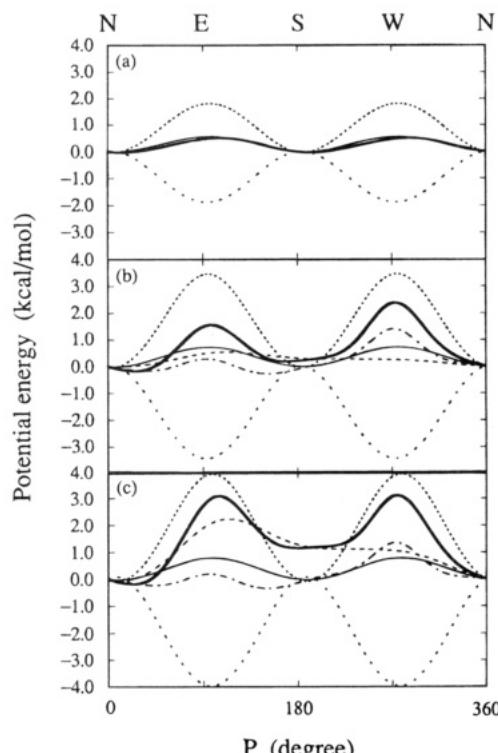


Figure 20. Values of potential energy along the path of pseudorotation and their breakdown for ribose: (a) for case A, (b) for case B, (c) for case C: In all panels, thick solid line, total potential energy; sparse dotted line, bond angle strain; dense dotted line, torsional barriers; thin solid line, gauche effect; dashed line, electrostatic interactions; dash-dot line, nonbonded interactions. The polar designation of the phase angle P is shown at the top.

puckered states of the furanose ring, one in the N and the other in the S range, with roughly the same stability have been observed experimentally for ribose with a favored barrier in the N-E-S range. These experimental results agree better with the result in Figure 20b, supporting the assumption of $\epsilon = 4r$ in the conformational energy function.

The electrostatic interactions along give rise to one energy barrier existing in the E range. Detailed analysis indicates that an atom pair O_4' - O_2' is mainly responsible for this barrier. The nonbonded interactions alone give rise two energy barriers: smaller one in the E range and larger one in the W range. The major contribution to the former comes from an atom pair O_4' - O_2' , and that to the latter from atom pairs O_4' - O_2' and N - O_3' . Furthermore the nonbonded interactions stabilize C_3' -endo ($P = \frac{1}{10}\pi$) and C_2' -endo ($P = \frac{9}{10}\pi$) conformations. Therefore, as was already noted by Olson,²¹ the nonbonded interactions play a role together with the gauche effect to characterize the pseudorotational motion in ribose.

Examination of Figure 20b,c reveals that relative energy between the stable C_3' -endo and C_2' -endo conformations is determined mainly by the electrostatic interactions. As noted above, the contribution from the atom pair O_4' - O_2' dominates the electrostatic interactions in the N-E-S range. Thus, we see that the oxygen atom in ribose attached at C_2' , which is absent in deoxyribose, plays a crucial role in determining the basic character of pseudorotation in ribose.

Deoxyribose. Total potential energy has one stable state in the N range and one energy barrier in the W range on the pseudorotation path (Figure 21b,c). The value of the barrier is 1.3 kcal/mol in Figure 21b and 1.0 kcal/mol in Figure 21c. In the N-E-S range the total energy curves are relatively flat with variation of about 0.5 kcal/mol. Thus the N-E-S range

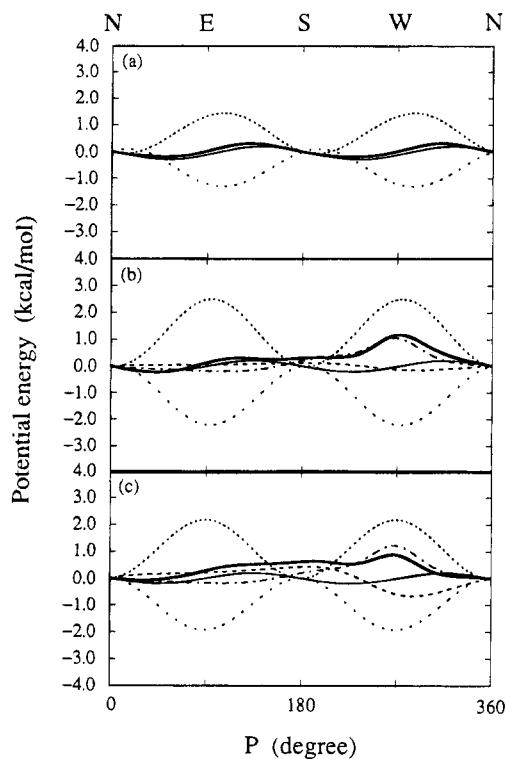


Figure 21. As in Figure 20, but for deoxyribose.

is more favored than the W range, which agrees with experimental results.²⁵

We see in Figures 21b,c that the electrostatic interactions favor the W range. The major contribution for this favorable interaction comes from an atom pair N–O_{3'}. Nonbonded interactions give rise to one energy barrier in the W range. Again the atom pair N–O_{3'} makes the major contribution for this barrier. The barrier in the E range, observed in ribose, is not observed in deoxyribose, because the O_{2'} atom, mainly responsible for the barrier in ribose, is absent in deoxyribose.

It is experimentally observed that deoxyribose has also two stable puckered states and that the state in the S range is slightly stabler than in the N range.^{22,23} Although our results shows existence of the two stable states on the pseudorotation path, the state in the S range is less stable than in the N range (Figure 21b,c). This may have come from the neglect of some side-chain atoms. As discussed previously, these neglected atoms should be restored in computer programs to treat nucleic acids.

Conclusions

We developed the analytic theory of pseudorotation in five-membered rings for cyclopentane in the form of series of approximations and extended it to be applicable to tetrahydrofuran, ribose, and deoxyribose.

For the pseudorotation in cyclopentane, eq 31 is a very good approximation, where constant values of the parameters $a-d$ are determined so as to minimize the quantity of eq 21 under the constraint of eq 34. The forms of eqs 31 and 34 are consequences of the ring structure and the symmetry of the chemical structure of the molecule, and are independent of the energetic details. Even though the analytical theory is developed in this paper by doing numerical calculations for one particular empirical energy function of the AMBER type, the formulation arrived at in the end is independent of an empirical energy function to be employed for E in eq 21. The numerical accuracy attained by the series of approximation is more than enough. The purpose of pursuing such a high accuracy in the case of

cyclopentane is to clarify the mathematical structure of the problem of pseudorotation. For practical applications of the present theory for cyclopentane the first-order approximation, i.e., the approximation of assuming $d = 0$ in eqs 31 and 34, would be enough for most purposes.

For the pseudorotation in tetrahydrofuran, a series of three approximate expressions, eq 40 (zeroth order), eq 44 (first order), and eq 46 (second order), are derived for the analytic representation and compared with the numerically obtained result. When we consider a term for the gauche effect (the second term in eq 38) in the empirical conformational energy function, even the zeroth-order approximation gives a fair agreement with the numerical result (Figure 13). In the zeroth approximation the phase relation holding exactly for cyclopentane is assumed. However, it should not hold exactly for tetrahydrofuran because of its lower structural symmetry. By allowing phase deviations, we obtain the first-order approximation. Agreement with the numerical result shows a drastic improvement (Figure 15). The first-order approximation consists of constant terms and terms dependent on 2ω . The second-order approximation obtained by allowing terms dependent on 4ω . The result is more accurate and almost identical with the numerical result (Figure 15). For most practical applications the first-order approximation would be accurate enough.

The energy profile along the pseudorotation path obtained in the first- and second-order approximations (Figure 14) agree with the experimental results.^{9,17} There are four energy minimum structures, C₃-exo ($\omega = 60^\circ$), C₄-endo ($\omega = 120^\circ$), C₃-endo ($\omega = 240^\circ$), and C₄-exo ($\omega = 300^\circ$). All of them, related to each other by symmetry operations, belong to the C₂ conformation. Between them, there are two large and two small energy barriers. The conformations of the former are O₁-exo ($\omega = 0^\circ$) and O₁-endo ($\omega = 180^\circ$), both belonging to the C₂ conformation, and those of the latter are C₃-exo-C₄-endo ($\omega = 90^\circ$) and C₃-endo-C₄-exo ($\omega = 270^\circ$), both belonging to C₅ conformation. It is known that furanose rings with even lower symmetry such as ribose and deoxyribose have two major stable states, C₂-endo and C₃-endo (in nomenclature common for ribose and deoxyribose, and corresponding to C₄-endo and C₃-endo in our tetrahydrofuran nomenclature, respectively), along the pseudorotation path. Therefore, we see that tetrahydrofuran already has the basic prototype character of the furanose rings.

For the pseudorotation in furanose rings of ribose and deoxyribose, we have obtained an analytic expression of eq 49 with the values of parameters listed as case B in Table 6. This expression gives an improvement over those currently used in literature in two aspects. The first is that it gives a more refined expression than the currently used ones. The latters correspond to assuming in eq 49 that (1) the values of parameters a_j and b_j do not depend on j and (2) deviations from ideal phase angles α_j are always zero. This assumption corresponds to our zeroth approximation in the treatment of cyclopentane. For cyclopentane with the high structural symmetry, this assumption was found to be reasonable. However, for tetrahydrofuran, a molecule with a lower symmetry, a further approximation with more parameters is found to be necessary. Equation 49 corresponds to the first-order approximation found reasonable for tetrahydrofuran. The points of refinement (1) and (2) mentioned above correspond to the loss of symmetry of the rings of ribose and of deoxyribose compared to cyclopentane. The second aspect of improvement of the expression of eq 49 is that the values of the parameters in Table 6 are determined by requiring the pseudorotation path to be very accurately the low-energy path for the assumed empirical energy function of eq

51. This is in contrast to the currently popular treatment in which an analytic but empirical expression of pseudorotation path is assumed and an empirical energy function is used only to evaluate the value of energy on the assumed path. In such a treatment the pseudorotation path can be somewhat different from the actual low-energy path of the assumed empirical energy function.

A detailed analysis is done of contributions of various energy terms to the total energy profile along the path of pseudorotation. Both for ribose and deoxyribose terms for bond-angle strain and torsional barriers have significant contributions but cancel each other almost completely. After cancellation of these two terms, the basic profile is determined by the term for the gauche effect. The electrostatic and nonbonded interactions add further contributions to modify the basic profile. The modification comes mainly from interactions between atom pairs O_{4'}—O_{2'} and N—O_{3'} in ribose and from interaction between an atom pair N—O_{3'} in deoxyribose. The nonbonded interaction between the atom pair N—O_{3'} gives rise to an energy barrier in the W range both in ribose and deoxyribose. The interaction between the atom pair O_{4'}—O_{2'}, which exists only in ribose, gives rise to energy barriers both in the E and W ranges in ribose. The absence of the oxygen atom in deoxyribose leads to a relatively flat energy profile in the N—E—S range of deoxyribose. This extra flexibility of the deoxyribose ring explains why DNA takes both B- and A-forms depending on environment, while RNA takes always A-form.

The analytic expressions of path of pseudorotation in ribose and deoxyribose rings will be used in a software we are now developing to treat nucleic acids in the dihedral angle space.

Acknowledgment. We are grateful to professor K. Harada for helpful discussions. This work has been supported by grants from Ministry of Education, Science and Culture, Japan.

References and Notes

- (1) Gō, N.; Noguti, T.; Nishikawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3696–3700.
- (2) Braun, W.; Gō, N. *J. Mol. Biol.* **1985**, *186*, 611–626.
- (3) Kidera, A.; Gō, N. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3718–3722.
- (4) Kidera, A.; Gō, N. *J. Mol. Biol.* **1992**, *225*, 457–475.
- (5) Kilpatrick, J. E.; Pitzer, K. S.; Spitzer, R. *J. Am. Chem. Soc.* **1947**, *69*, 2483–2488.
- (6) Lanne, J. *Vibrational Spectra and Structure*; D. Marcel: New York, 1972; Vol. 1, pp 25–50.
- (7) Pitzer, K. S.; Donath, W. E. *J. Am. Chem. Soc.* **1959**, *81*, 3123–3218.
- (8) Harris, D. O.; Engerholm, G. G.; Tolman, C. A.; Luntz, A. C.; Keller, R. A.; Kim, H.; Gwinn, W. D. *J. Chem. Phys.* **1969**, *50*, 2438–2445.
- (9) Engerholm, G. G.; Luntz, A. C.; Gwinn, W. D. *J. Chem. Phys.* **1969**, *50*, 2446–2457.
- (10) Geise, H. J.; Altona, C.; Romers, C. *Tetrahedron Lett.* **1967**, *15*, 1393–1386.
- (11) Carreira, L. A.; Jiang, G. J.; Person, W. B.; Willis Jr., J. N. *J. Chem. Phys.* **1972**, *56*, 1440–1443.
- (12) Miller, F. A.; Inskeep, R. G. *J. Chem. Phys.* **1959**, *18*,
- (13) Duing, J. R.; Wertz, D. W. *J. Chem. Phys.* **1968**, *49*, 2118–2121.
- (14) Altona, C.; Geise, H. T.; Romers, C. *Tetrahedron* **1968**, *24*, 12–32.
- (15) Ikeda, T.; Lord, R. C.; Malloy Jr., T. B.; Ueda, T. *J. Chem. Phys.* **1973**, *56*, 1434–1439.
- (16) Wurrey, C. J.; Durig, J. R.; Carreira, L. A. *Vibrational Spectra and Structure*; D. Marcel: New York, 1972; Vol. 5, pp 211–221.
- (17) Greenhouse, J. A.; Strauss, H. L. *J. Chem. Phys.* **1969**, *50*, 124–134.
- (18) Chandrasekhar, J.; Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5073–5079.
- (19) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (20) Westhof, E.; Sundaralingam, M. *J. Am. Chem. Soc.* **1980**, *102*, 1493–1500.
- (21) Olson, W. K.; Sussman, J. L. *J. Am. Chem. Soc.* **1982**, *104*, 270–278.
- (22) Herzyk, P.; Rabchenko, A. *J. Chem. Soc., Perkin Trans. 2* **1989**, 209–216.
- (23) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *93*, 8205–8212.
- (24) Cremer, D.; Pople, J. *J. Am. Chem. Soc.* **1974**, *97*, 1354–1367.
- (25) Rao, S. T.; Westhof, E.; Sundaralingam, M. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1981**, *37*, 421–425.
- (26) Levitt, M.; Warshel, A. *J. Am. Chem. Soc.* **1978**, *100*, 2607–2613.
- (27) Olson, W. K. *J. Am. Chem. Soc.* **1982**, *104*, 278–286.
- (28) Harvey, S. C.; Prabhakaran, M. *J. Am. Chem. Soc.* **1986**, *108*, 6128–6136.
- (29) Roder, O.; Ludemann, H. D.; Goldammer, E. V. *Eur. J. Biochem.* **1975**, *53*, 517–524.
- (30) Hendrickson, J. B. *J. Am. Chem. Soc.* **1961**, *83*, 4537–4547.
- (31) Hendrickson, J. B. *J. Am. Chem. Soc.* **1963**, *85*, 4059–4059.
- (32) Hoyland, J. R. *J. Chem. Phys.* **1969**, *50*, 2775–2775.
- (33) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

JP9413307