Eur Biophys J (1999) 28: 369–379

ARTICLE

Atsushi Matsumoto · Masaki Tomimoto · Nobuhiro Go

Dynamical structure of transfer RNA studied by normal mode analysis

Received: 29 October 1998 / Revised version: 8 February 1999 / Accepted: 11 February 1999

Abstract The internal motion of yeast phenylalanine transfer RNA is studied by normal mode analysis in extended dihedral angle space, in which the flexibility of five-membered ribose rings is treated faithfully by introducing a variable for its pseudo-rotational motion. Analysis of global molecular motions reveals that the molecule is very soft. We show that this softness comes not from the property of the "material" comprising the molecule but from its slender shape. Analysis of thermal distance fluctuations reveals that this molecule can be regarded as consisting dynamically of three blocks. Thermal fluctuations of the mainchain dihedral angles show rigidity of the anticodon region. They also show flexibility of regions around non-stacking bases. Base-stacking interactions cause suppression of the correlated functions of mainchain dihedral angles beyond a ribose ring. We analyze the thermal fluctuation of parameters describing the positions of two adjacent bases. Fluctuations of relative translational parameters in the anticodon and acceptor stem regions are found to be larger than those in other stem regions. The relative translational motions cause the two stem regions to undergo global twisting and bending motions. We show that the role of pseudo-rotational motion of sugars is important in regions around bases which are involved in nonregular interactions.

Key words Normal mode analysis · Dihedral angle space · Pseudo-rotation · Interbase parameters

A. Matsumoto · M. Tomimoto[1] · N. Go (✉)
Department of Chemistry, Graduate School of Science,
Kyoto University,
Sakyo-ku, Kyoto 606-8502, Japan
e-mail: go@qchem.kuchem.kyoto-u.ac.jp

Present address:
[1] Central Pharmaceutical Research Institute of Japan Tobacco Inc.,
Takatsuki, Osaka 569-1125, Japan

Introduction

Transfer RNA molecules form a class of small globular polynucleotide chains about 75–90 nucleotides long. They act as vehicles for transferring amino acids from the free state inside the cell into the assembled chain of the protein. This vital function as an intermediary between the nucleic acid language of the genetic code and the amino acid language of the working cell has made transfer RNA a major subject of research in molecular biology.

In the process of protein synthesis, the tRNA molecules interact with a variety of other molecules, and in such interactions conformational changes of tRNA molecules are expected to play important roles. Therefore, it is important to study their internal motions.

Unlike the canonical double-helical DNA molecules, tRNA molecules consist roughly of two elements, double-helical stem regions and loop regions (Fig. 1). Some bases in the loop or terminal regions are free from stacking interactions. Tertiary interactions keep the L-shaped three-dimensional structure. In this paper, we study the dynamic features of the molecule both at the levels of the structural elements and the global molecular conformation.

Normal mode analysis is one of the methods now widely used for studying dynamic aspects of biopolymers (Go et al. 1983). In this method, molecular conformational dynamics is expressed as a superposition of mutually independent collective motions called normal modes. Underlying this picture of conformational dynamics is a basic assumption of the harmonicity of the energy surface, i.e., the energy surface is assumed to be given by a multidimensional parabola within the range of conformational fluctuations.

Normal mode calculations for proteins have been carried out frequently in dihedral angle space, because of the advantage of a considerable reduction of the number of independent variables associated with fixing bond lengths and bond angles. If we apply this treatment to nucleic acids, it leads to fixing the motions of sugar rings. However, it is well known that the five-membered sugar rings

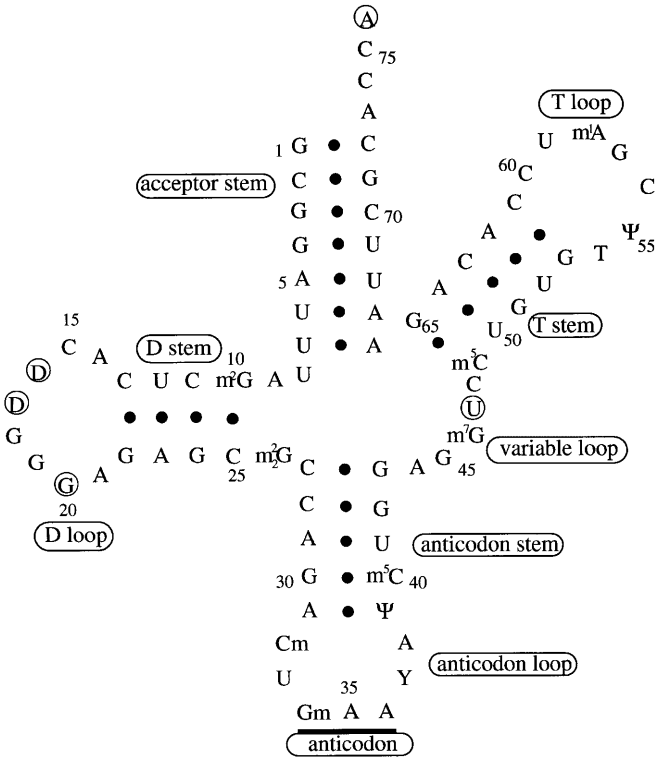

Fig. 1 Sequence and secondary structure of yeast phenylalanine transfer RNA shown in a clover-leaf diagram. The characters A, U, G, and C represent the usual nucleotides in RNA. Nucleotides with the character "m" are those modified by the addition of one or more methyl groups. Characters D, Y, Ψ, and T are dihydrouridine, wyosine, pseudouridine, and ribothymidine, respectively. *Circled characters* are those with nonstacking bases. Dots between two nucleotides mean that their bases are paired by hydrogen bonds. Nucleotide numbers are also shown

have very flexible puckering motions involving changes in bond angles so that their motions may have considerable influence on the entire molecular dynamics. Therefore, it is desirable that we deal with motions of sugar rings even in dihedral angle space. Recently, the algorithm developed in dihedral angle space was extended so as to deal with motions of sugar rings by introducing pseudo-rotation variables (Tomimoto and Go 1995). In this treatment, pseudo-rotational motion of a five-membered sugar ring is expressed as a function of a single pseudo-rotation variable, so that all bond angles and dihedral angles in a ring change as a function of this variable. This treatment of sugar rings also contributes to reduce the number of independent variables considerably. In this point, the algorithm has an advantage over others (Lavery et al. 1995; Mazur 1998) in which bond angles and dihedral angles in sugars are treated explicitly. Using the algorithm of Tomimoto and Go (1995), we have carried out normal mode analysis for yeast phenylalanine transfer RNA.

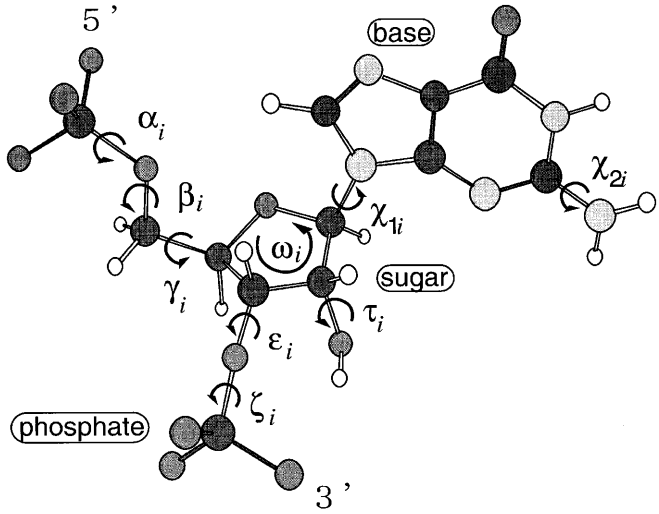

Fig. 2 Definition of independent variables in the *i*th nucleotide. Independent variables α_i, β_i, γ_i, ε_i, ζ_i, χ_{1i}, χ_{2i}, and τ_i are dihedral angles, and ω_i is a pseudo-rotation variable. The dihedral angle χ_{2i} does not exist in bases of uridines. In some modified nucleotides, more dihedral angles, for example χ_{3i}, χ_{4i}, … exist

Materials and Methods

Molecule studied

The yeast phenylalanine transfer RNA consists of 76 nucleotides. Its sequence is shown in Fig. 1. The 3′ and 5′ termini together form a Watson-Crick double stranded acceptor stem. There are 2510 atoms in this molecule. We take 610 rotatable dihedral angles and 76 pseudo-rotation variables as independent variables. Therefore, the total number of independent variables is 686 (see Fig. 2). We refer to this choice of independent variables as the calculation in extended dihedral angle space. This choice of the independent variables means that we are assuming fixed standard values for all bond lengths and bond angles, except for bond angles in ribose rings. It has the advantage of reducing the number of independent variables by a factor of about 11 as compared to a choice of atomic Cartesian coordinates. However, with this choice, we need to calculate a set of dihedral angles and pseudo-rotation variables that generates X-ray crystallographic conformation. This step is called regularization and is done before minimization of the conformational energy. The atomic coordinates of yeast $tRNA^{Phe}$ are taken from 4TRA in the Protein Data Bank.

Minimization of conformational energy

Starting from the regularized conformation, the conformational energy is minimized. The AMBER force field (Weiner et al. 1986) is used as an empirical energy function in this extended dihedral angle space. A sigmoidally distance-dependent function (Ramstein and Lavery 1988;

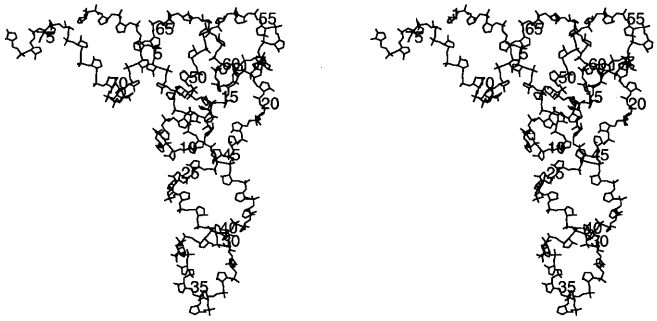

Fig. 3 Stereo-drawing of backbone atoms (heavy atoms of sugars and phosphates) at the minimum energy conformation, with nucleotide numbers

Daggett et al. 1991) is assumed for the dielectric constant ε in electrostatic interactions.

Mature tRNAs commonly contain many modified nucleotides (Kim 1976). Because we had no information about the electric partial charges for the modified nucleotides, we calculated these charges by the electrostatic potential fitting method (ESP fitting method) (Cox and Williams 1981; Singh and Kollman 1984) using the program MOPAC 6.0. The values of the calculated partial charges are available upon request to go@qchem.kuchem.kyoto-u.ac.jp.

The energy minimization is done by Newton's method, in the process of which both the gradient and the hessian (second derivative matrix) of the conformational energy function are used. This method has been a practical and in fact very efficient one for proteins, because of the algorithm for rapid calculation of the gradient and the hessian of the conformational energy function (Noguti and Go 1983b; Abe et al. 1984; Wako and Go 1987). This method has recently been extended so as to deal with nucleic acids with sugar puckering motions (Tomimoto et al. 1996). The minimized backbone structure is shown in Fig. 3. This molecule takes an L-shaped three-dimensional structure with one end formed by the accepter arm (nucleotides 1–7, 66–76) and the other by the anticodon loop region (nucleotides 32–38).

Normal mode analysis

Many normal mode calculations in dihedral angle space have been carried out for proteins and nucleic acids (Nishikawa and Go 1987; Ikura and Go 1993; Nakamura and Doi 1994). In this study, the calculation is carried out in extended dihedral angle space, but the mathematics remains mostly unchanged. Here, the basic mathematical formulation of the normal mode analysis is described only to the extent necessary to fix notations and to derive a few specific formulae used in this paper. Then a newly developed method is described to analyze the thermal fluctuations of the positions of a pair of adjacent bases.

In normal mode analysis, the conformational energy around a minimum point is approximated by a multidimensional parabola, (1/2) $\Delta q^t \boldsymbol{F}\, \Delta q$, where Δq is a vector whose ith component Δq_i is a small change of a value of the ith independent variable from that at the minimum, and $\boldsymbol{F}$ is the second derivative matrix of the conformational energy function at the minimum. The kinetic energy of the system is given by (1/2) $\Delta \dot{q}^t \boldsymbol{H}\, \Delta \dot{q}$, where $\boldsymbol{H}$ is the coefficient matrix, whose analytic expression is given in the literature (Noguti and Go 1983a). Note that external motions of the molecule are separated from internal motions in the above expression by the method of Eckart (1935).

Normal mode analysis is performed by simultaneously diagonalizing $\boldsymbol{H}$ to the identity matrix, and $\boldsymbol{F}$ to the eigenvalue matrix whose ith diagonal element gives the square of the angular frequency of the ith normal mode, yielding an eigenvector matrix $\boldsymbol{V}$.

Time averaged quantities

The correlation of the fluctuations of a pair of independent variables $\langle \Delta q_i\, \Delta q_j \rangle$ is given by

$$\langle \Delta q_i\, \Delta q_j \rangle = k_B T \sum_k v_{ik}\, v_{jk} / \omega_k^2 \tag{1}$$

where v_{ik} is the element of the eigenvector matrix $\boldsymbol{V}$, k_B the Boltzmannconstant, T the absolute temperature, and ω_k the angular frequency of kth normal mode. Equation (1) gives the mean-square thermal fluctuation of the ith independent variable $\langle \Delta q_i^2 \rangle$ when $j = i$.

Here, we define the correlation coefficient by

$$\langle \Delta q_i\, \Delta q_j \rangle / (\langle \Delta q_i^2 \rangle\, \langle \Delta q_j^2 \rangle)^{1/2} \tag{2}$$

The thermal displacement vector of the ath atom in the ith normal mode is defined as follows:

$$\boldsymbol{D}_{ai} = (kT)^{1/2} / \omega_i \sum_j \boldsymbol{k}_{aj}\, v_{ji} \tag{3}$$

Here, $\boldsymbol{k}_{aj}$ is given by $\partial \boldsymbol{r}_a / \partial q_j$, where $\boldsymbol{r}_a$ is the position vector of the ath atom. The method of calculation of $\boldsymbol{k}_{aj}$ is described in the literature (Noguti and Go 1983a).

In the thermally excited ith mode, the displacement vector of the ath atom fluctuates along positive and negative directions of vector $\boldsymbol{D}_{ai}$ with the root-mean-square amplitude of D_{ai}. Using this vector, the mean-square fluctuation of the ath atom is given as

$$\langle \Delta \boldsymbol{r}_a^2 \rangle = \sum_i \boldsymbol{D}_{ai}^2 \tag{4}$$

Interatomic distances also fluctuate around their mean values at the minimum energy conformation. The mean-square thermal interatomic distance between a pair of atoms, numbered a and b, is given as a sum of the contributions from individual modes by

$$\langle \Delta d_{ab}^2 \rangle = \sum_i [(\boldsymbol{D}_{ai} - \boldsymbol{D}_{bi}) \cdot (\boldsymbol{r}_a - \boldsymbol{r}_b) / |\boldsymbol{r}_a - \boldsymbol{r}_b|]^2 \tag{5}$$

In this study, the absolute temperature is set equal to 300 K.

Correlation of the directions of displacement vectors

As noted in the protein BPTI (Nishikawa and Go 1987), atomic motions in low-frequency normal modes are not

localized, and the directions of displacement vectors are essentially continuous, i.e., the directions of displacements of atoms close in space are near. In order to make a quantitative study of the extent of the spatial correlation of the direction of displacement vectors in each normal mode, the correlation function of the direction vectors of atomic displacements are defined and calculated as follows.

For a given normal mode, attention is paid to atoms whose displacement vectors are larger in magnitude than 1/20 of the largest displacement vector in the molecule. For such an atom (say, the *a*th atom), the direction vector of displacement $\boldsymbol{A}_i(\boldsymbol{r}_{a0})$ is defined by $\boldsymbol{D}_{ai}/|\boldsymbol{D}_{ai}|$, where $\boldsymbol{r}_{a0}$ is the position vector of the *a*th atom in the minimum energy conformation. This direction vector is normalized in the sense that $|\boldsymbol{A}_i(\boldsymbol{r}_{a0})| = 1$. Now the correlation function of the direction vector is defined by

$$C_i(d) = \frac{\sum_a \sum_b \mathbf{A}_i(\mathbf{r}_{a0}) \cdot \mathbf{A}_i(\mathbf{r}_{b0})\, \delta(d - |\mathbf{r}_{a0} - \mathbf{r}_{b0}|)}{\sum_a \sum_b \delta(d - |\mathbf{r}_{a0} - \mathbf{r}_{b0}|)} \qquad (6)$$

where δ is the Dirac's delta function, and the summation extends over all atoms for which direction vector $\boldsymbol{A}_i(\boldsymbol{r}_{a0})$ is defined. Therefore, the correlation function $C_i(d)$ is defined as an average of dot products between the direction vectors of displacement of a pair of atoms, which are separated in space by distance *d*.

Thermal fluctuation of the positions of adjacent bases

The positions of two rigid bodies in space are determined by 12 parameters. In our treatment in extended dihedral angle space, base core atoms constitute a rigid body. We pay our attention to these atoms. Then, the positions of adjacent bases are determined by 12 parameters. We can take 6 parameters for representing relative positions and orientations between the two bases, that is, shift, slide, rise, tilt, roll, and twist (Dickerson et al. 1989), and the remaining 6 parameters for representing absolute positions and orientations of the two bases as a whole. When the molecule fluctuates, these parameters also fluctuate. The thermal fluctuations of these parameters are calculated as follows.

The atomic fluctuations in a pair of bases Δx_i (i = 1, …, $3n$), where n is the number of atoms in the two bases, are related to the fluctuations of the 12 parameters Δf_k (k = 1, …, 12) by

$$\Delta x_i = \sum_k a_{ik}\, \Delta f_k \qquad (7)$$

where a_{ik} is an element of a constant $3n \times 12$ matrix A, whose kth column vector is given by a rate of change of the atomic positions for an infinitesimal displacement of the kth parameter. The definition of parameters Δf_k and the analytic forms of a_{ik} are given in the Appendix. Matrix A satisfies following condition:

$$\sum_i m_i\, a_{ij}\, a_{ik} = \delta_{jk}\, J_j \qquad (8)$$

where m_i is the mass of the ith atom, $\delta_{jk} = 1$ for $j = k$ and vanishes otherwise, and J_j is a constant which is given in the Appendix. It can be readily shown that the correlation of a pair of fluctuations of 12 parameters $\langle \Delta f_k\, \Delta f_l \rangle$ is given by

$$\langle \Delta f_k\, \Delta f_l \rangle = \sum_{ij} m_i\, m_j\, a_{ik}\, a_{jl}\, \langle \Delta x_i\, \Delta x_j \rangle / J_k\, J_l \qquad (9)$$

where $\langle \Delta x_i\, \Delta x_j \rangle$ is a correlation function of the displacements of coordinates of x_i and x_j. This equation gives mean-square thermal fluctuations of the parameters $\langle \Delta f_k^2 \rangle$ when $k = l$. For the calculation of fluctuations of relative translation, further treatment is needed and is shown in the Appendix.

Table 1 Values of Young's modulus for the seven lowest frequency normal modes with their frequencies

Mode number	Frequency (cm^{-1})	Young's modulus (10^8 dyn cm^{-2})
1	0.58	1.07
2	0.84	1.21
3	1.06	2.04
4	1.17	2.00
5	1.55	4.89
6	1.82	6.64
7	2.16	6.26

Results and discussion

Global molecular structure

Frequencies of the seven lowest frequency modes are given in Table 1 together with other quantities to be discussed later. The lowest frequency in this molecule with a molecular weight of 24 783 is 0.58 cm^{-1}. This value may be compared with the corresponding lowest frequency of 3.05 cm^{-1} in a globular protein, triosephosphate isomerase (TIM), which has a molecular weight of 26 757 (Kobayashi et al. 1997). This comparison suggests that transfer RNA is a molecule much softer than a protein with a similar molecular weight.

Figure 4 shows the displacement vectors of backbone atoms in this lowest frequency mode. Large bending motions are observed at both ends of the molecule, one at the acceptor arm, and the other at the anticodon arm. The directions of the two bending motions are parallel to the plane made by the L-shape of the molecule and opposite to each other. The motion of the acceptor arm seems to be caused by deformation of the region around U7 and A66, which exists at the junction of the T stem and the acceptor stem, and the motion of the anticodon arm seems to be caused by deformation of the region around C28 and G42. Nakamura and Doi (1994) analyzed motions of this molecule in low-frequency normal modes by assuming that they are hinge bending motions of two domains, one of which is made of the acceptor arm and the T arm, and the other made of the D arm and the anticodon arm. They identified the position and direction of the bending axis in each normal

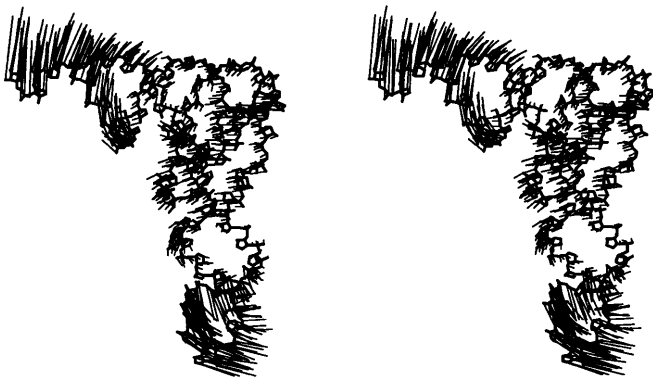

Fig. 4 Stereo-drawing of the displacement vectors of backbone atoms in the lowest frequency mode. Magnification factor of these vectors is 5

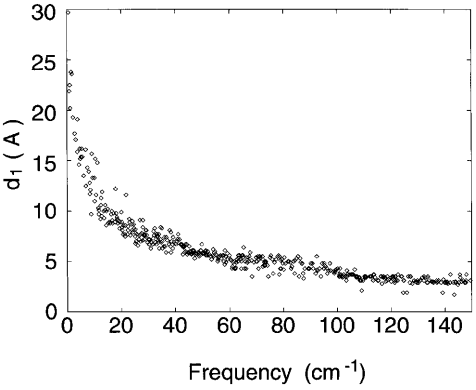

Fig. 6 Value of d_1, the smallest interatomic distance at which the correlation function $C_i(d)$ for the ith mode vanishes, plotted against frequency

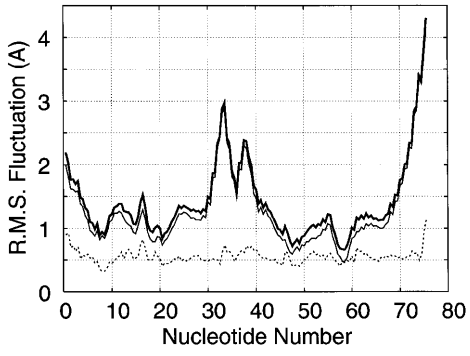

Fig. 5 Root-mean-square thermal fluctuation of the positions of C4′ and P atoms plotted by *thick solid line*, and square root of sums of contributions from modes with frequencies less than and larger than 2.5 cm^{-1} plotted by *thin solid line* and *broken line*, respectively

mode without prior knowledge of the axis and bend angles. They found the axis to be located around the elbow region of this molecule, which is made of the D loop and the T loop. Harvey and McCammon (1981) also used a two-rigid-body model for analyzing intramolecular flexibility in the molecule. However, as we show later, this molecule seems to consist of three rigid segments. Therefore, it seems appropriate that two hinge bending axes exist in this molecule instead of one.

Mean-square thermal atomic fluctuation is given by Eq. (4) and the root-mean-square fluctuation therefrom is plotted against nucleotide number for mainchain atoms (P and C4′) in Fig. 5. Equation (4) shows that the mean-square atomic fluctuation is given by a sum of the contributions from all normal modes. The contribution to the thermal atomic fluctuation from modes with a frequency below 2.5 cm^{-1} (7 modes) and that from all other modes (679 modes) are also plotted in Fig. 5. This figure shows that the pattern of atomic fluctuation specific to the molecular conformation is determined by a surprisingly small number of the lowest frequency modes. All other modes contribute to conformation-independent background fluctuations. Previous studies for proteins (Nishikawa and Go 1987; Ikura and Go 1993) have shown that the thermal atomic fluctuation is determined mainly by a relatively small number of low frequency normal modes. For instance, modes with a frequency less than 30 cm^{-1}, whose number is about the same as that of amino acid residues in a protein, account for more than 80% of the mean-square atomic fluctuation. However, in the present case of tRNA, a surprisingly smaller number of low frequency modes (seven modes) contribute to the same extent. As each normal mode contributes to the mean-square thermal fluctuation by a quantity given by $k_B T/\omega_i^2$, where ω_i is the angular frequency in the ith normal mode, the large contribution from the small number of low frequency modes is due to the fact that their frequencies are extremely low.

The appearance of the very low frequency modes in the transfer RNA molecule may be because either (1) the "material" comprising the molecule is very soft or (2) the molecule has a slender shape. To examine if (1) may be the reason, we try to estimate an effective Young's modulus for each normal mode. [The same analysis was done previously for proteins (Nishikawa and Go 1987; Ikura and Go 1993)]. For this purpose, each normal mode is regarded as a standing wave in a continuous body of nucleic acid. An average of the inter-nodal plane distances is estimated from the value d_1 at which the directional correlation function of Eq. (6) vanishes first. Young's modulus E is given roughly by $\rho\,(2ck\,d_1)^2$, where ρ is the mass density, c the light velocity, and k the frequency of the mode in wavenumbers. The estimated values of d_1 are plotted in Fig. 6 against the frequency of the mode. Estimated values of Young's modulus for the first seven modes are given in Table 1. The plot in Fig. 6 is very similar to that obtained previously for the small globular protein BPTI (Nishikawa and Go 1987), meaning that values d_1 for a given value k are essentially the same for proteins and RNAs. This in turn means that the "material" comprising proteins and RNA molecules has essentially the same mechanical property. Therefore the appearance of the very low frequency modes in transfer RNA should be due to the slenderness of its shape.

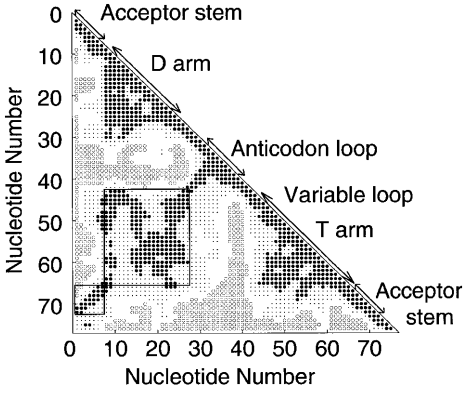

Fig. 7 Triangular map showing the thermal fluctuations of distances between a pair of nucleotides. The fluctuation of distance between two nucleotides is calculated as the average fluctuation between the mainchain atoms (O5′, C5′, C4′, C3′, O5′, P) of the two nucleotides. This quantity is normalized by dividing by the average fluctuation of all possible distances in the molecule. The normalized quantity S_{mn}, where m and n are nucleotide numbers, is classified into (a) $S_{mn} \geq 1.2$, (b) $1.2 > S_{mn} \geq 0.8$, (c) $0.8 > S_{mn} \geq 0.4$, and (d) $0.4 > S_{mn}$. Those in classes of (a), (c), and (d) are indicated in the map by symbols ○, "·" (dot), and ●. Those in class (b) are indicated by blanks

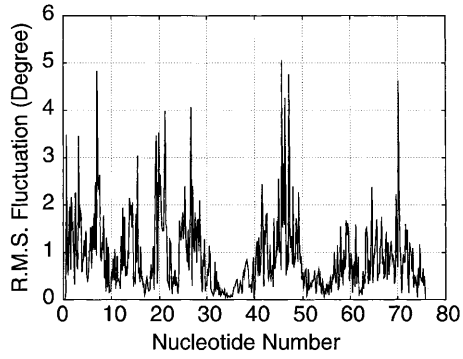

Fig. 8 Root-mean-square fluctuations of mainchain dihedral angles and pseudo-rotation variables due to the seven lowest frequency modes plotted against nucleotide number

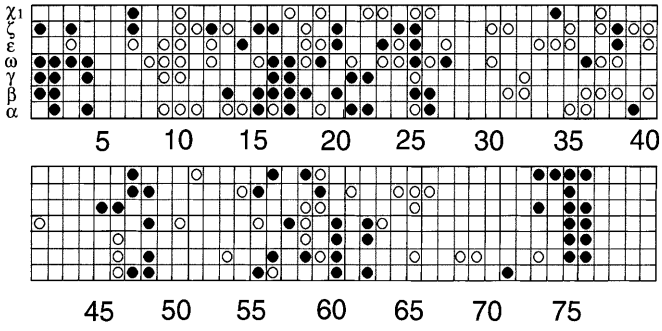

Fig. 9 Independent variables with relative thermal fluctuations higher than 1.2 or lower than 0.8 are plotted by ● and ○, respectively

Dynamically rigid blocks

The fluctuation of the distance between a pair of atoms in the molecule is given by Eq. (5). The fluctuation of the distance between two nucleotides is given by averaging this atomic distance fluctuation and is plotted in Fig. 7. If the fluctuations among a certain number of nucleotides are small, the segment can be regarded as a dynamically rigid block. From the distribution of the symbols, several dynamically rigid blocks along the mainchain can be identified. They are the segments, consisting of 1–7 (acceptor stem), 8–27 (the D arm region), 31–41 (the anticodon loop), 44–65 (the variable loop and T arm region), and 66–71 (acceptor stem). Distance fluctuations in a contact region between segments 1–7 (acceptor stem) and 66–71 (acceptor stem) are also small, indicating that these two fragments, which are interacting with each other by base-pairing, behave as a single rigid unit. Similarly, fragments 8–27 (D arm region) and 44–65 (variable loop and T arm region) also behave as a single rigid unit. The tertiary interactions between the bases in these regions should be responsible for this behavior. From the distribution of the symbols, it is observed that the anticodon loop and the acceptor arm have large distance fluctuations to most of other segments. In consequence, this molecule can be regarded as consisting dynamically of three blocks.

Local structures

Rigid anticodon region

Figure 8 shows fluctuations of mainchain dihedral angles and pseudo-rotation variables due to the seven lowest frequency normal modes, which are responsible for the conformation-specific atomic fluctuations as shown in Fig. 5. Although the atomic fluctuations are large at the anticodon loop region 32–38, the fluctuations of the independent variables due to the seven lowest-frequency modes in this region are very small. Thus, in the very low frequency modes responsible for the conformation-specific atomic fluctuations, the anticodon loop region is fluctuating by a large amplitude but almost as a rigid body.

The thermal fluctuation of an independent variable is given by the square root of Eq. (1) with $i = j$. The amplitudes of the thermal fluctuations of the independent variables α_i, β_i, γ_i, ε_i, ζ_i, ω_i, and $\chi 1_i$, when averaged over i, are 9.5°, 7.6°, 6.9°, 6.3°, 6.7°, 7.9°, and 5.6°, respectively. Here, we divide each of the thermal fluctuations of the independent variables by its averaged value and call it the relative thermal fluctuation. The independent variables with high (larger than 1.2) and low (smaller than 0.8) relative thermal fluctuations are shown in Fig. 9.

In this figure, there are several regions with a small number of symbols, such as the regions made by nucleotides 5–6, 27–29, 41–45, 49–54 and 67–72. They are in the the stem regions, where double-stranded structure is seen. This shows that in the double-stranded regions the fluctuations of the dihedral angles and pseudo-rotation variables take average values. Although the D stem region also takes a double-stranded structure, there are many symbols in the

Table 2 Correlation coefficients between the fluctuations of independent variables averaged over i = 1–75[a]

5′	3′												
	$\chi 1_i$	ε_i	ζ_i	α_{i+1}	β_{i+1}	γ_{i+1}	ω_{i+1}	$\chi 1_{i+1}$	ε_{i+1}	ζ_{i+1}	α_{i+2}	β_{i+2}	γ_{i+2}
ω_i	0.029	–0.071	–0.160	–0.188	–0.018	0.038	0.138						
$\chi 1_i$		–0.221	0.073	–0.048	–0.035	0.071	0.008	0.132					
ε_i			**–0.539**	**0.352**	**–0.472**	**–0.149**	0.047	–0.062	0.012				
ζ_i				**–0.370**	**0.284**	**0.136**	0.072	–0.011	–0.020	–0.037			
α_{i+1}					**0.524**	**–0.603**	0.039	–0.014	0.019	–0.007	–0.011		
β_{i+1}						**–0.066**	–0.089	0.175	–0.135	0.032	0.010	–0.013	
γ_{i+1}							0.170	–0.264	0.107	–0.054	–0.034	0.033	–0.044

[a] Entries for pairs of mainchain dihedral angles existing between two adjacent sugars are shown by a bold face

Table 3 Correlation coefficients of the fluctuations of two dihedral angles (one of them α_i) separated by a sugar connected to a nonstacking base (D16, D17, G20 and U47). Correlation coefficients whose absolute value is larger than 10 times and 5 times of that of the average value (Table 2) are indicated by • and ⋆, respectively

	Dihedral angle				
i	ε_i	ζ_i	α_{i+1}	β_{i+1}	γ_{i+1}
16	⋆ 0.182	• –0.326	0.051	• –0.169	0.073
17	⋆ 0.120	⋆ 0.065	• 0.326	• –0.500	• –0.176
20	• 0.282	–0.005	⋆ 0.061	• –0.316	0.124
47	⋆ 0.182	• –0.192	• –0.173	• 0.593	• –0.349

figure. This will be due to the fact that some bases in this region are involved in tertiary interactions so that the interactions are not so simple as in the other double-stranded regions. The large fluctuation in nucleotide 1–4, which is also in the stem region, may be due to deformation of the structure after the energy minimization.

In the anticodon loop region, however, there are more symbols indicating smaller values of the fluctuations than in the double-stranded regions. Especially in the anticodon region (34th–36th nucleotides), successive dihedral angles, ε_{35}, ζ_{35}, α_{36}, and β_{36} take relative thermal fluctuations smaller than 0.8. This shows that the anticodon region is quite rigid. Probably this rigidity owes largely to interactions of the 33rd base with mainchain phosphate atoms in the anticodon region (Sundaralingam et al. 1976).

Nonstacking bases

Most of the bases in this molecule are involved in stacking interactions, but five out of the 76 bases are not. They are bases 16, 17, 20, 47, and 76 (see Fig. 1), and large fluctuations of the independent variables are observed around these bases (Fig. 9). In order to see the details of the fluctuations, correlation coefficients [Eq. (2)] are calculated and the values averaged over all the nucleotides, as shown in Table 2. This table shows that the correlations between two dihedral angles are generally large, when both of them exist between two adjacent sugars, that is, when both of them are dihedral angles out of ε_i, ζ_i, α_{i+1}, β_{i+1}, and γ_{i+1} (see Fig. 2). On the other hand, the correlations between two dihedral angles separated by a sugar are generally small.

The situation is different around nonstacking bases. Table 3 shows that many correlation coefficients between two dihedral angles separated by a sugar connected to a nonstacking base are large. This suggests that the large fluctuations of dihedral angles around nonstacking bases are due to the absence of stacking.

From the opposite point of view, this result suggests that the thermal fluctuations of the independent variables in regions where stacking is regular are restrained by the stacking.

Thermal fluctuation of the positions of adjacent bases

The thermal fluctuation of the positions of adjacent bases is given by the square-root of Eqs. (9) and (A9) when $k = l$, and is plotted against nucleotide number in Fig. 10. The fluctuations of the parameters of the common rotations are shown in Fig. 10a. The three parameters take large values at the anticodon arm and acceptor stem. This fact indicates that these regions show large bending and twisting motions, because the z axis in each base-fixed coordinate system is approximately parallel to the helix axis of each region, so that the large fluctuations of the common rotation parameters around the z axis in each region bring about a global twisting motion and those around the x and y axes, bring about global bending motions.

Our definition of the 12 parameters are based on the assumption that the adjacent bases are considered stacked on each other. Therefore, we cannot obtain meaningful values of parameters between two bases not stacked on each other, such as those involving nonstacking bases, those between the 33rd and 34th bases, and between the 55th and 56th bases. The values of ρ_z at nucleotide numbers 34, 47, and 56 are markedly smaller than the values of their respective neighbors. This may be due to the nonstacking nature of the respective adjacent base pairs.

Among the three relative rotation parameters, the fluctuation of twist is smallest and takes nearly the same value in all stem regions (Fig. 10b). The very large values at nucleotide numbers 17, 18, 34, 47, 48, and 76 may be again

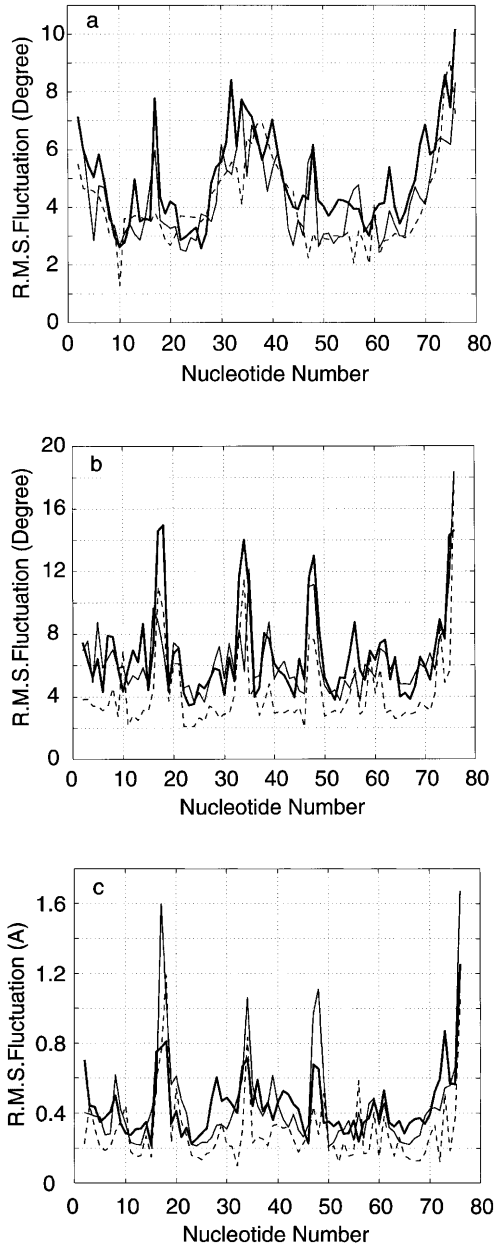

Fig. 10a–c Thermal fluctuations of parameters describing the positions and orientations of *i*th and $(i-1)$th bases plotted against nucleotide number. **a** Common rotational parameters: ρ_x by *thick solid line*, ρ_y by *thin solid line*, and ρ_z by *dashed line*. **b** Relative rotational parameters: tilt by *thick solid line*, roll by *thin solid line*, and twist by *dashed line*. **c** Relative translational parameters: shift by *thick solid line*, slide by *thin solid line*, and rise by *dashed line*

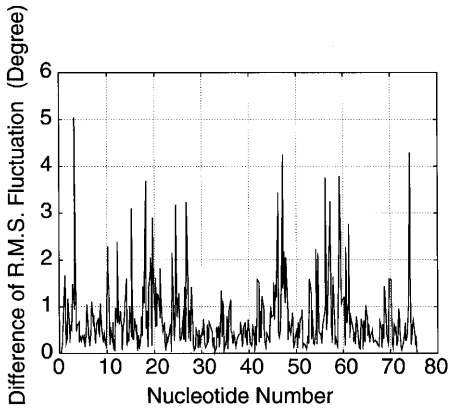

Fig. 11 Difference of root-mean-square fluctuations of mainchain dihedral angles calculated in dihedral angle space from those in extended dihedral angle space plotted against nucleotide number

due to the nonstacking nature. The other two parameters also take relatively constant values in all stem regions. However, the situation of the relative translational parameters is different (Fig. 10c). The values of these parameters are close in each stem region, but those in the anticodon stem region and acceptor stem region are generally larger than those in the remaining stem regions. As large bending and twisting motions are observed in the former regions, this suggests that the changes of these relative translational parameters are needed for these global conformational fluctuations.

Now, we pay attention to the anticodon region (34th–36th bases). The values of the parameters of relative translation between the 34th and 35th bases and those between the 35th and 36th bases are close to those in the anticodon stem region. Those of relative rotation between the 35th and 36th bases are close to those in the stem regions, but values of the fluctuation of tilt and roll between the 34th and 35th bases are larger.

This region is in a single-stranded region, and in a single-stranded region it is expected that a base can fluctuate more freely than in a double-stranded region owing to the lack of base-pair interaction. In fact, the values of the parameters for the 73rd–75th bases, which are in a single-stranded region and not base-paired with other bases by tertiary interaction, are fairly large. Therefore, the values of the parameters in the anticodon region close to those in stem regions suggest that the relative motions between adjacent bases in this region are restrained. This restraint may be due to the rigid mainchain structure of this region, shown earlier.

On the other hand, the large fluctuation of the tilt and roll between the 34th and 35th bases will be due to a large fluctuation of the dihedral angle χ_1 in the 34th nucleotide, whose base is at the end of the stack and interacts with another base by only one side so that it can fluctuate more freely than those sandwiched by two adjacent bases. Therefore, the fluctuation of tilt and roll between the 34th and 35th bases may not be influenced by the rigid mainchain structure.

Site dependence of the effect of internal motions of sugars

In order to access the effect of the pseudo-rotational motions of sugars, we have fixed the internal motions of sugars and then performed normal mode analysis. Normal

mode analysis by fixing the motions of sugars is equivalent to the normal mode analysis in dihedral angle space performed by Nakamura and Doi (1994). The difference of the mainchain thermal fluctuations of dihedral angles calculated in this space from those in extended dihedral angle space is shown in Fig. 11. They always become smaller when the motions of sugars are fixed, however, there is clear site dependence. The difference is large in the regions around nonstacking bases and bases involved in the tertiary interactions such as those in the D loop and T loop. On the other hand, they are generally small in the regions where regular stacking is observed, such as in stem regions. Therefore, we can conclude that the motions of sugars become more important in the irregular regions than in the regular double-stranded regions.

Summary

We have studied the internal motion of yeast phenylanine transfer RNA by normal mode analysis in extended dihedral angle space.

First we studied the global motions of the molecule. We have shown that this molecule is very soft from the fact that there are a few modes with very low frequencies. (Note that in normal mode analysis, fluctuation only around an energy minimum point is considered and transitions between distinct minima are neglected. This limitation may be underestimating the flexibility of the molecule. Thus, a more appropriate treatment may reveal a softer appearance of the molecule.) This soft appearance is shown to come from the slenderness of its shape and not from the property of the 'material' comprising the molecule. The thermal distance-fluctuation shows that there are three rigid segments in this molecule. They correspond roughly to the acceptor stem, the anticodon loop, and a remaining part which consists of the D arm, the variable loop, and the T arm.

Then we studied the conformational fluctuations of the local structures. The anticodon loop region is shown to fluctuate by a large amplitude in a few low frequency normal modes, which make major contributions to the thermal atomic fluctuation, but almost as a rigid body. The thermal fluctuations of the dihedral angles show that the mainchain structure of the anticodon region (34–36th nucleotide) is quite rigid. This rigidity is probably largely due to the interaction of the 33rd base with mainchain phosphate atoms in the anticodon region. On the other hand, the mainchain structures around nonstacking bases in general are shown to be flexible, from the fluctuations of the dihedral angles. The correlation coefficients of a pair of fluctuations of mainchain dihedral angles suggest that this flexibility comes from the absence of stacking interactions of bases.

We have calculated the thermal fluctuations of 12 parameters describing the positions and orientations of adjacent bases. The fluctuations of relative translational parameters in the anticodon and acceptor stem regions are found to be larger than those in other stem regions. The relative translational motions cause the two stem regions to undergo twisting and bending motions. The fluctuations of parameters in the anticodon region have values close to those in double-stranded regions even though the anticodon region is in a single-stranded region. This suggests that the motions of bases in the anticodon region are restrained by the rigid mainchain structure.

In our calculation, the internal motions of sugar rings are treated faithfully. In order to estimate the effect of their motions, we have done another normal mode calculation in which their motions are fixed. The mainchain dihedral angles have shown that the influence of the fixing is important in regions around bases which are involved in nonregular interactions.

Appendix

Here, we describe our definition of the 12 parameters for the description of motions between a pair of adjacent bases, and derive detailed expressions for a_{ij} and J_k in Eqs. (7) and (8). Then, we show the treatment for the calculation of the fluctuation of relative translation between two bases.

At the beginning, we define the local coordinate system fixed on each base. The origin is set at the position of the center of mass of the base and the $x\,y$ plane is set on the plane made by the heavy atoms on the base. Therefore, the position of the ith atom on the nth base $\boldsymbol{r}_i^{(n)}$ can be written as $\boldsymbol{r}_i^{(n)} = (x_i^{(n)}, y_i^{(n)}, 0)$. We determine the directions of the x and y axes so that the following condition is satisfied:

$$\sum_i m_i^{(n)} x_i^{(n)} y_i^{(n)} = 0 \tag{A1}$$

where $m_i^{(n)}$ is the mass of the ith atom on the nth base. Furthermore, the directions of the x and y axes are selected so that the C1′ atom, which is a sugar atom connected to the base directly, is in the second quadrant of the $x\,y$ plane. We take the positive direction of z axis along the 5′-to-3′ direction of the strand in the usual Watson-Crick base-paired double strand. The x, y, and z axes construct a right-handed system.

Next, we consider a pair of adjacent, nth and $(n+1)$th, bases. The equation for translation of the coordinate expressed in the jth base-fixed system $\boldsymbol{r}_i^{(j)}$ to that in the space-fixed system $\boldsymbol{s}_i^{(j)}$ is given as follows.

$$\boldsymbol{s}_i^{(j)} = S^{(j)}\, R\, \boldsymbol{r}_i^{(j)} + \boldsymbol{t} + \boldsymbol{t}^{(j)} \tag{A2}$$

where R and $\boldsymbol{t}$ are the rotation matrix and the translation vector common to both the nth and $(n+1)$th bases, and $S^{(j)}$ and $\boldsymbol{t}^{(j)}$ are those specific to each base. The infinitesimal change of $\Delta\boldsymbol{s}_i^{(j)}$ is given as

$$\Delta\boldsymbol{s}_i^{(j)} = S^{(j)} R\,(\Delta\boldsymbol{\rho}\times\boldsymbol{r}_i^{(j)}) + S^{(j)} R\,(\Delta\boldsymbol{\sigma}^{(j)}\times\boldsymbol{r}_i^{(j)}) + \Delta\boldsymbol{t} + \Delta\boldsymbol{t}^{(j)} \tag{A3}$$

where $\Delta\boldsymbol{\rho}$ and $\Delta\boldsymbol{\sigma}^{(j)}$ are rotation vectors, around which each base rotates to an infinitesimal extent, and they are common to both and specific to each of paired bases, respec-

tively. To derive Eq. (A3) from Eq. (A2), we used the relation $\Delta\boldsymbol{\sigma}_s^{(j)}\times(S^{(j)}\ R\ \boldsymbol{r}_i^{(j)}) = S^{(j)}\ R\ (\Delta\boldsymbol{\sigma}^{(j)}\times\boldsymbol{r}_i^{(j)})$ where $\Delta\boldsymbol{\sigma}^{(j)} \equiv (S^{(j)}\ R)^{-1}\ \Delta\boldsymbol{\sigma}_s^{(j)}$ is the rotation vector expressed in a base-fixed coordinate system, while $\Delta\boldsymbol{\sigma}_s^{(j)}$ is the same rotation vector expressed in a space-fixed coordinate system. In the text, we refer to $\Delta\boldsymbol{\rho}$ as the common rotation vector.

When we define the kth term in Eq. (A3) as $\Delta\boldsymbol{s}_{ik}^{(j)}$, the orthogonal conditions

$$\sum_{j=n,n+1}\sum_i m_i^{(j)}\,\Delta\boldsymbol{s}_{ik}^{(j)}\Delta\boldsymbol{s}_{il}^{(j)} = 0 \tag{A4}$$

where $k \neq l$, lead to the following conditions:

$$\begin{aligned} M^{(n)}\,\Delta\boldsymbol{t}^{(n)} + M^{(n+1)}\,\Delta\boldsymbol{t}^{(n+1)} &= 0 \\ I^{(n)}\,\Delta\boldsymbol{\sigma}^{(n)} + I^{(n+1)}\,\Delta\boldsymbol{\sigma}^{(n+1)} &= 0 \end{aligned} \tag{A5}$$

where $M^{(j)} = \sum_i m_i^{(j)}$ ($j = n, n+1$)and $I^{(j)}$ is a diagonal matrix whose kth diagonal element is $I_k^{(j)}$, which is an inertial moment and given as follows:

$$\begin{aligned} I_1^{(j)} &= \sum_i m_i^{(j)}\,y_i^{(j)2} \\ I_2^{(j)} &= \sum_i m_i^{(j)}\,x_i^{(j)2} \\ I_3^{(j)} &= \sum_i m_i^{(j)}\,(x_i^{(j)2} + y_i^{(j)2}) \end{aligned} \tag{A6}$$

Here, we define new vectors $\Delta\boldsymbol{\sigma}$ and $\Delta\boldsymbol{u}$ by $\Delta\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma}^{(n)} - \Delta\boldsymbol{\sigma}^{(n+1)}$ and $\Delta\boldsymbol{u} = \Delta\boldsymbol{t}^{(n)} - \Delta\boldsymbol{t}^{(n+1)}$, respectively. From this definition, it is natural to call these vectors as the relative rotation vector and the relative translation vector, respectively. From the conditions in Eq. (A5), $\Delta\boldsymbol{\sigma}^{(j)}$ and $\Delta\boldsymbol{t}^{(j)}$ are represented by these vectors as follows:

$$\begin{aligned} \Delta\sigma_k^{(n)} &= I_k^{(n+1)}\,\Delta\sigma_k/(I_k^{(n)} + I_k^{(n+1)}) \\ \Delta\sigma_k^{(n+1)} &= -I_k^{(n)}\,\Delta\sigma_k/(I_k^{(n)} + I_k^{(n+1)}) \\ \Delta\boldsymbol{t}^{(n)} &= M^{(n+1)}\,\Delta\boldsymbol{u}/(M^{(n)} + M^{(n+1)}) \\ \Delta\boldsymbol{t}^{(n+1)} &= -M^{(n)}\,\Delta\boldsymbol{u}/(M^{(n)} + M^{(n+1)}) \end{aligned} \tag{A7}$$

where $\Delta\sigma_k^{(n)}$ and $\Delta\sigma_k$ are the kth elements of vectors $\Delta\boldsymbol{\sigma}^{(n)}$ and $\Delta\boldsymbol{\sigma}$, respectively. In the text, we regard $\Delta\sigma_k$ for k = 1, 2, 3 as the change of tilt, roll, and twist, respectively.

By substituting Eq. (A7), the right-hand side of Eq. (A3) is described in terms of 12 variables (4 vectors×3 elements), and this gives the detailed expression of a_{ij} in Eq. (7). By using this expression, we get the following equation:

$$\begin{aligned} \sum_{j=n,n+1}\sum_i m_i^{(j)}\Delta\mathbf{s}_i^{(j)}\,\Delta\mathbf{s}_i^{(j)} &= (M^{(n)} + M^{(n+1)})\,(\Delta\mathbf{t})^2 \\ + \sum_k (I_k^{(n)} + I_k^{(n+1)})\,(\Delta\rho_k)^2 &+ \frac{M^{(n)}M^{(n+1)}}{M^{(n)} + M^{(n+1)}}\,(\Delta\mathbf{u})^2 \\ + \sum_k \frac{I_k^{(n)}I_k^{(n+1)}}{I_k^{(n)} + I_k^{(n+1)}}\,(\Delta\sigma_k)^2 & \end{aligned} \tag{A8}$$

where $\Delta\rho_k$ is the kth element of $\Delta\boldsymbol{\rho}$. By using Eqs. (7) and (8), the left-hand side of this equation is given by $\sum J_j\,f_j^2$. Therefore, this equation gives the detailed expression of J_k in Eq. (8).

The relative translation vector $\Delta\boldsymbol{u}$ in Eq. (A3) is expressed in the space-fixed coordinate system. In order to know the changes of translational parameters between two bases, that is, shift, slide, and rise, we have to project this vector onto the midway coordinate system between the nth and $(n+1)$th bases. When the rotation matrix of the midway coordinate system relative to the space-fixed coordinate system is given by T, the relative translation expressed in the midway coordinate system $\Delta\boldsymbol{w}$ is given by $\Delta\boldsymbol{w} = T^t\,\Delta\boldsymbol{u}$, so that the correlation function $\langle\Delta w_k\,\Delta w_l\rangle$ is given as follows:

$$\langle\Delta w_k\,\Delta w_l\rangle = \sum_{i,j} T_{ik}\,T_{jl}\,\langle\Delta u_i\,\Delta u_j\rangle \tag{A9}$$

where $\langle\Delta u_i\,\Delta u_j\rangle$ is given by Eq. (9). In the text, we regard $\langle\Delta w_k^2\rangle$ for k = 1, 2, 3 as the mean-square fluctuation of shift, slide, and rise, respectively.

Acknowledgements We would like to acknowledge helpful discussion with Wilma K. Olson. This work has been supported by grants-in-aid from the Ministry of Education, Japan, to N. G. A part of the computation has been done in the Computer Center of the Institute for Molecular Science.

References

Abe H, Braun W, Noguti T, Go N (1984) Rapid calculation of first and second derivatives of conformational energy with respect to dihedral angles for proteins. General recurrent equations. Comput Chem 8: 239–247

Cox SR, Williams DE (1981) Representation of the molecular electrostatic potential by a net atomic charge model. J Comput Chem 2: 304–323

Daggett V, Kollman PA, Kuntz ID (1991) Molecular dynamic simulations of small peptides: dependence on dielectric model and pH. Biopolymers 31: 285–304

Dickerson RE, Bansal M, Calladine CR, Diekmann S, Hunter WN, Kennard O, Kitzing E von, Lavery R, Nelson HCM, Olson WK, Saenger W, Shakked Z, Sklenar H, Soumpasis DM, Tung CS, Wang AHJ, Zhurkin VB (1989) Defintions and nomenclature of nucleic acids structure components. Nuclejic Acids Res 17: 1797–1803

Eckart C (1935) Some studies concerning rotating axes and polyatomic molecules. Phys Rev 47: 552–558

Go N, Noguti T, Nishikawa T (1983) Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc Natl Acad Sci USA 80: 3696–3700

Harvey SC, McCammon JA (1981) Intramolecular flexibility in phenylalanine transfer RNA. Nature 294: 286–287

Ikura T, Go N (1993) Normal mode analysis of mouse epidermal growth factor: characterization of the harmonic motion. Proteins 16: 423–436

Kim SH (1976) Three-dimensional structure of transfer RNA. Prog Nucleic Acid Res Mol Biol 17: 181–216

Kobayashi N, Yamato T, Go N (1997) Mechanical property of a TIM-barrel protein. Proteins 28: 109–116

Lavery R, Zakrzewska K, Sklenar H (1995) JUMNA (junction minimisation of nucleic acids). Comput Phys Commun 91: 135–158

Mazur AK (1998) Accurate DNA dynamics without accurate long-range electrostatics. J Am Chem Soc 120: 10928–10937

Nakamura S, Doi J (1994) Dynamics of transfer RNAs analyzed by normal mode calculation. Nucleic Acids Res 22: 514–521

Nishikawa T, Go N (1987) Normal mode of vibration in bovine pancreatic trypsin inhibitor and its mechanical property. Proteins 2: 308–329

Noguti T, Go N (1983a) Dynamics of native globular proteins in terms of dihedral angle. J Phys Soc Jpn 52: 3283–3288

Noguti T, Go N (1983b) A method of rapid calculation of a second derivative matrix of conformational energy for large molecules. J Phys Soc Jpn 52:3685–3690

Ramstein J, Lavery R (1988) Energetic coupling between DNA bending and base pair opening. Proc Natl Acad Sci USA 85:7231–7235

Singh UC, Kollman PA (1984) An approach to computing electrostatic charges for molecules. J Comput Chem 5:129–145

Sundaralingam M, Mizuno H, Stout CD, Rao ST, Liebman M, Yathindra N (1976) Mechanisms of chain folding in nucleic acids. The (ω',ω) plot and its correlation to nucleotide geometry in yeast $tRNA^{Phe}$. Nucleic Acids Res 3:2471–2484

Tomimoto M, Go N (1995) Analytic theory of pseudo-rotation in five-membered rings. Cyclopentane, tetrahydrofuran, ribose and deoxyribose. J Phys Chem 99:563–577

Tomimoto M, Wako H, Go N (1996) Conformational analysis of nucleic acid molecules with flexible furanose rings in dihedral angle space. J Comput Chem 17:910–917

Wako H, Go N (1987) Algorithm for rapid calculation of hessian of conformational energy function of proteins by supercomputer. J Comput Chem 8:625–635

Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) An all atom force field for simulations of proteins and nucleic acids. J Comput Chem 7:230–252