# Building-Block Approach for Determining Low-Frequency Normal Modes of Macromolecules

**Florence Tama,**[1] **Florent Xavier Gadea,**[1] **Osni Marques,**[2] **and Yves-Henri Sanejouand**[1*]
[1]*Laboratoire de Physique Quantique, CNRS, IRSAMC, Université Paul-Sabatier, Toulouse Cedex, France*
[2]*National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory, Berkeley, California*

**ABSTRACT** Normal mode analysis of proteins of various sizes, ranging from 46 (crambin) up to 858 residues (dimeric citrate synthase) were performed, by using standard approaches, as well as a recently proposed method that rests on the hypothesis that low-frequency normal modes of proteins can be described as pure rigid-body motions of blocks of consecutive amino-acid residues. Such a hypothesis is strongly supported by our results, because we show that the latter method, named RTB, yields very accurate approximations for the low-frequency normal modes of all proteins considered. Moreover, the quality of the normal modes thus obtained depends very little on the way the polypeptidic chain is split into blocks. Noteworthy, with six amino-acids per block, the normal modes are almost as accurate as with a single amino-acid per block. In this case, for a protein of n residues and N atoms, the RTB method requires the diagonalization of an n × n matrix, whereas standard procedures require the diagonalization of a 3N × 3N matrix. Being a fast method, our approach can be useful for normal mode analyses of large systems, paving the way for further developments and applications in contexts for which the normal modes are needed frequently, as for example during molecular dynamics calculations. Proteins 2000;41:1–7. © 2000 Wiley-Liss, Inc.

Key words: collective motion; RTB method; rigid bodies; rotation-translation; B-factors; conformational change

## INTRODUCTION

Since the early 1980s, normal mode analysis has proved to be useful for studying collective motions of biological macromolecules.[1–3] Noteworthy, it has been shown that some of the lowest-frequency normal modes of several proteins are strongly correlated with the large amplitude conformational change of these proteins observed upon ligand binding.[4–10] Also, the corresponding normal mode coordinates have been used to study domain motions in proteins,[11–13] to analyze molecular dynamics trajectories through the quasi-harmonic approximation,[14] or to integrate the equations of atomic motion with large time-steps.[15,16]

Thus, it is now clear that normal coordinates form a well-suited coordinate reference system for studying the properties of the potential energy surface of proteins.

However, until recently, it was not feasible to perform a normal mode analysis for proteins of more than approximately 150 amino acids with an all-atom model, or to approximately 200 amino acids by using a united atom description in which hydrogens bound to carbon are not explicitly represented, that is, when 3N, the number of degrees of freedom of the system, is larger than 5,000–6,000,[17] the limiting phase being the numeric diagonalization of a matrix of size 3N × 3N. Recent progresses in the methodology have allowed this limitation to be overcome,[9,10] and it is now possible to perform such an analysis for proteins as large as the largest ones presently found in the Brookhaven protein data bank[18] (see below).

To date, two different approaches have been successfully used to compute low-frequency normal modes of large proteins at the atomic level. The first one makes use of a very efficient block Lanczos algorithm.[9] This algorithm, as implemented in the BLZPACK program package, for example, allowed the calculation of the 10 lowest-frequency normal modes of the closed form of dimeric citrate synthase (3N = 25,584), with 20 minutes of CPU time on a Cray C90.[9] However, in the present state of this package, such a calculation can be performed only when the matrix to be diagonalized is quite sparse, because a computer memory larger than 10 times the number of nonzero matrix elements may be required. With the second method, called DIMB,[10,19] implemented in the CHARMM package[20] since version 24, systems as large as the aspartate transcarbamylase (3N = 79,758) can be studied.[21] However, the corresponding process is quite slow. For instance, to compute the 197 lowest-frequency normal modes of phosphoglycerate kinase (3N = 11,529), 4.4 hours of CPU time on a Cray C98 are required.[22] In the present work, we show that it is possible to obtain very accurate approximations of the low-frequency normal modes of large systems, by using small amounts of CPU time on common kinds of

workstations. The method used is the first step of a method described in a previous work, in which only small test cases had been examined.[23] In this approach, hereafter called RTB (rotations-translations of blocks), the protein is divided into $n_b$ blocks, a block being made of one or of a few consecutive residues. Then, the lowest-frequency normal modes of the protein are obtained as a linear combination of the rotations and translations of these blocks.

From a practical point of view, a new code has been developed that avoids the storage of the whole Hessian matrix and paves the way for the determination of approximate normal modes of very large biomolecules.

## METHODS
### RTB Versus Standard Approaches

In standard approaches, the normal modes of a system are obtained through the diagonalization of the Hessian matrix $\mathbf{H}$, that is, the $3N \times 3N$ matrix of the second-derivatives of the potential energy with respect to the mass-weighted coordinates, where N is the number of atoms of the system.[24,25] In the RTB approach, $\mathbf{H}$ is first expressed in a basis defined by the rotations and translations of $n_b$ blocks, $\mathbf{H_b}$, the projected Hessian, being given by:

$$H_b = P^t H P$$

where $\mathbf{P}$ is an orthogonal $3N \times 6n_b$ matrix built with the vectors associated with the local rotations and translations of each block.[23] Approximate low-frequency normal modes of the system, thus, are obtained by diagonalizing $\mathbf{H_b}$, a matrix of size $6n_b \times 6n_b$, the corresponding (3N) atomic displacements being obtained as:

$$A_p = PA_b$$

where $\mathbf{A_b}$ is the matrix of the eigenvectors of $\mathbf{H_b}$.

As previously emphasized, the final objective is the treatment of huge macromolecules. This is achieved by a direct approach, implemented in a newly developed code, which borrows some ideas from the adiabatic-diabatic transformation needed for the vibronic treatment of small molecules[26] and proceeds in three steps. In the first step, the blocks of residues are defined and the six rotation-translation modes of each block, $U_i$, are determined and stored. These $6n_b$ vectors form the new basis of small dimension that corresponds to the projector $\mathbf{P}$. In the second step, the Hessian matrix is expressed in this new basis, separately for each coupling or diagonal block, $H_{ij}$

$$H_{ij}^b = U_i^t H_{ij} U_j$$

The set of $n_b^2$ $H_{ij}^b$ block-matrices forms the matrix $\mathbf{H_b}$. Thus, only the small dimension vectors $U_i$ and the small $6n_b \times 6n_b$ $\mathbf{H_b}$ matrix have to be stored. In the last step, $\mathbf{H_b}$ is diagonalized with standard methods. Hereafter, this was done either with BLZPACK or with the DIAGQ routine available in the VIBRAN module of the CHARMM package.

### Comparison of Two Sets of Normal Modes

A normal mode, $\vec{a}_j$, obtained with a standard approach, can be expressed as a linear combination of the $6n_b$ approximate normal modes $\vec{a}_k$ obtained with the RTB approach. Then, a value close to one for, $P_j$ with:

$$P_j = \sum_{k=1}^{6n_b} c_{jk}^2 \tag{1}$$

where $c_{jk} = \vec{a}_j . \vec{a}_k^p$, means that the j-th normal mode, $\vec{a}_j$, can be perfectly described in the subspace considered within the frame of the RTB approach.

### Normal Mode Analysis Versus Experimental Data
#### *Atomic fluctuations*

The atomic fluctuations are obtained from a normal mode analysis as[25]:

$$\langle x_i^2 \rangle = \frac{k_B T}{m_i} \sum_{j=1}^{n_v} \frac{a_{ij}^2}{\omega_j^2} \tag{2}$$

where $x_i$ is the atomic coordinate $i$, $m_i$, its mass, $n_v$, the number of modes considered, $v_j = \omega_j/2\pi$ the frequency of normal mode j, and $a_{ij}$ the corresponding coordinate displacement. For most proteins, accurate values of $\langle x_i^2 \rangle$ can be obtained with their 30 lowest-frequency normal modes ($n_v = 30$). Indeed, it has been shown that low-frequency normal modes of proteins, with frequencies under $30 \text{ cm}^{-1}$, are responsible for most of their atomic displacements.[5,27,28]

#### *Overlap*

To quantify how a given normal mode, $\vec{a}_j$, compares with an experimentally known conformational change, $\Delta \vec{r} = \vec{r}_o - \vec{r}_c$, the overlap between the two corresponding vectors can be calculated as[9,29]:

$$|\vec{a}_j \cdot \Delta \vec{r}| = \frac{|\sum a_{ij}(r_i^0 - r_i^c)|}{\sqrt{\sum a_{ij}^2 \sum (r_i^0 - r_i^c)^2}} \tag{3}$$

where $r_i^o$ and $r_i^c$ are, respectively, the atomic coordinate $i$ in conformations "o" and "c" of the protein, after both were superimposed. An overlap of one would mean that both kinds of collective atomic displacements occur along exactly the same direction of the configurational space.

## RESULTS AND DISCUSSION

Standard and RTB normal mode analyses of proteins of various sizes (46 up to 858 residues; see Table I) have been carried out, after a preliminary energy minimization performed with the CHARMM package,[20] by using extended atoms, the PARAM19 force field, a dielectric constant of 1, a shifting function for electrostatics and a switching function between 6.5 and 8.5Å for the Lennard-Jones interactions. The minimization process was stopped at a gradient root-mean-square (RMS) value of $10^{-6}$ kcal/(mole.Å); RMS deviations of the $C_\alpha$ atoms from the crystallographic structure are given in Table I. Then, $\mathbf{H}$ was computed with the VIBRAN module of CHARMM. Next,

**TABLE I. Preliminary Energy Minimization of the Proteins Considered in This Work: $C_\alpha$ Root Mean Square Deviation (Å) From the Crystallographic Structure**

| Protein | Pdb code | Number of residues | RMS deviation |
|---|---|---|---|
| Crambin | 1ccn | 46 | 1.4 |
| BPTI | 5pti | 58 | 1.1 |
| HIV-1 protease | 1hhp | 99 | 1.7 |
| Che Y protein | 3chy | 128 | 0.9 |
| CD4 | 3cd4 | 178 | 1.7 |
| Adenylate kinase | 1aky | 218 | 1.9 |
| TIM | 3tim | 250 | 1.6 |
| Triglyceride lipase | 3tgl | 265 | 1.3 |
| Tyrosine phosphatase | 1ypt | 278 | 1.2 |
| Alcool dehydrogenase | 8adh | 373 | 1.3 |
| Enolase | 3enl | 436 | 1.3 |
| Citrate synthase | 5csc | 858[a] | 1.9 |

[a]The dimeric form was considered.

following standard approaches, **H** was either diagonalized with BLZPACK, when the number of amino acids of the protein was less than 250 or by using the iterative diagonalization in a mixed basis (DIMB) method implemented in CHARMm 24 otherwise. In the latter case, the iterative process was stopped when the convergence criterion for the eigenvectors reached 0.05.[10]

Finally, the results obtained with such standard diagonalization methods were compared with those obtained with the RTB approach, by using 1, 2, 3, 5, or more, residues per block, that is, when $H$ is first projected into a subspace defined by the rotations and translations of a set of $n_b$ blocks, the projected Hessian, **$H_b$** being diagonalized with standard algorithms.

In Figure 1, $\nu_b$, the frequencies obtained as a result of the diagonalization of **$H_b$** are compared with $\nu_s$, the frequencies obtained with standard approaches. A linear relationship is observed, for frequencies up to at least 40 cm$^{-1}$, namely:

$$\nu_b = d_p \cdot \nu_s \qquad (4)$$

with $d_p = 1.7 \pm 0.1$. That $d_p$ is larger than 1 means that the protein potential energy surface is stiffer in the subspace considered within the framework of our approximation, that is, when each amino acid is assumed to behave like a rigid body. The surprising observation is that $d_p$ does not seem to depend on the protein size. This finding suggests that, in the case of the low-frequency normal modes of proteins, the slight deformations experienced by the amino acids during these motions are a consequence of local constraints, nearly identical on average. This is an appealing result, because it allows to use Eq. (4) to estimate rather accurately the exact frequency of a normal mode, once the approximate value is known. However, as shown in Figure 2 in the case of adenylate kinase, when **$H_b$** is calculated with five residues per block, the linear relationship observed in Figure 1 does not hold. This latter result is likely to be a consequence of the fact that $\Phi$, $\Psi$ dihedral motions inside each block are not represented,
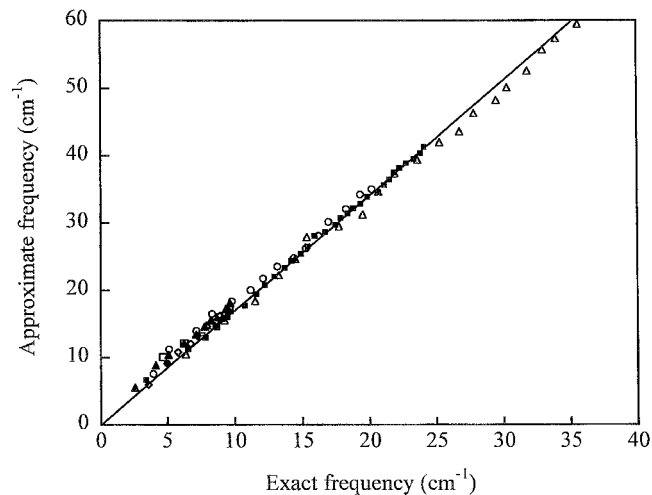


Fig. 1. Approximate frequencies calculated with the RTB approach, each block containing a single amino acid, as a function of the frequencies calculated with standard approaches. Proteins of various sizes were considered (see Table I): crambin (filled triangles), HIV-1 protease (open circles), adenylate kinase (open triangles), triglyceride lipase (open squares), alcohol dehydrogenase (crosses) and enolase (filled diamonds). For the sake of clarity, only some of the approximate values corresponding to exact frequencies calculated in the 10–20 cm$^{-1}$ range are shown.
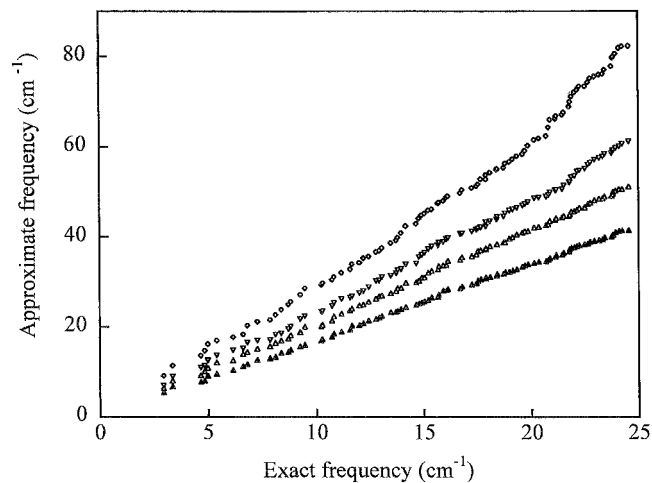


Fig. 2. Approximate frequencies calculated with the RTB approach, as a function of the frequencies calculated with standard approaches, in the case of adenylate kinase. The approximate frequencies were calculated by using one (filled triangles; bottom), two (open triangles), three (reverse triangles), or five (diamonds; top) amino acids per block.

although with $d_p = 3.0$ relation (4) still yields quite accurate estimates of $\nu_s$ for frequencies lower than 10 cm$^{-1}$, as can be verified when proteins of various sizes are considered (data not shown). Note also that $d_p$ seems to increase linearly, as a function of the number of residues per block. Indeed, $d_p$ is nearly equal to 1.7, 2.1, 2.4, and 3.0, when each block contains 1, 2, 3, or 5 residues, respectively.

In Figure 3, the RMS fluctuations of the $C_\alpha$ atoms of adenylate kinase are shown, as calculated with standard or RTB normal mode analysis, when the 30 lowest-
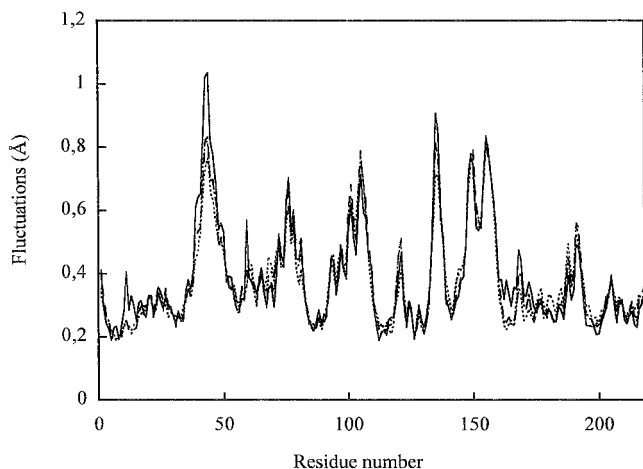
Fig. 3.   Root mean square fluctuations of the $C_\alpha$ atoms of adenylate kinase, calculated with standard (plain line) or RTB normal mode analysis, when the 30 lowest-frequency normal modes are taken into account, the RTB approach being used with one (broken line) or five (dotted line) amino acid(s) per block. The temperature was set to 300°K.

**TABLE II. Correlation Between the $C_\alpha$ Root Mean Square Fluctuations, as Calculated According to Normal Mode Theory, When the Standard or When the RTB Approach Is Used, With One Residue per Block**

| Protein | Correlation |
|---|---|
| Crambin | 0.93 |
| BPTI | 0.93 |
| HIV-1 protease | 0.88 |
| Che Y protein | 0.85 |
| CD4 | 0.96 |
| Adenylate kinase | 0.93 |
| TIM | 0.88 |
| Triglyceride lipase | 0.84 |
| Tyrosine phosphatase | 0.95 |
| Alcool dehydrogenase | 0.86 |
| Enolase | 0.88 |

**TABLE III. Correlation Between the $C_\alpha$ Root Mean Square Fluctuations, as Calculated According to Normal Mode Theory, When the Standard or When the RTB Approach Is Used, With Different Partitionings**

| Protein | Number of residues per block | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 5 |
| HIV-1 protease | 0.88 | 0.90 | 0.85 | 0.74 |
| Adenylate kinase | 0.93 | 0.93 | 0.92 | 0.88 |
| Alcohol dehydrogenase | 0.86 | 0.83 | 0.82 | 0.79 |

**TABLE IV. Correlation Between the $C_\alpha$ Root Mean Square Fluctuations of Adenylate Kinase, as Calculated According to Normal Mode Theory, When the Standard or When the RTB Approach Is Used, With Different Partitionings**

| Number of blocks | Nb of residues per block | Correlation |
|---|---|---|
| 218 | 1 | 0.93 |
| 218[a] | 1[a] | 0.94 |
| 44 | 5 | 0.88 |
| 40[b] | 1–14[b] | 0.87 |
| 40[c] | 1–14[c] | 0.89 |
| 37 | 6 | 0.86 |

[a]The peptidic bond following each residue is included in each block.
[b]Each secondary structure element is put in a block.
[c]Blocks of same length, but randomly distributed along the polypeptidic chain.

frequency normal modes are considered (Eq. 2). Because in such calculations the weight of each normal mode is proportional to $1/v$, relation (4) was used to adjust the frequencies obtained with the RTB approach. As evidenced in Figure 3, the fluctuations calculated with standard and RTB normal mode analysis are highly correlated. In Table II, the corresponding correlation coefficient is given (0.93), as well as those obtained for several other proteins, when the RTB method is used with one residue per block. These coefficients have all quite high values, ranging between 0.84 and 0.96. As shown in Figure 3, for adenylate kinase, and in Table III, for two other proteins, high values are also obtained when the RTB method is used with blocks containing 2, 3, or 5 amino acids, although a slight decrease of these coefficients is observed in the latter case. Note that the way the polypeptide chain is split into blocks does not seem to matter much as far as such correlations are concerned. For instance, as shown in Table IV for adenylate kinase, results obtained are very similar, whether the peptidic bond is entirely included or not in the residue definition, as well as when the splitting is done according to the known distribution of secondary structure elements or not. In the former case, each secondary structure element is put in a block, each residue lying outside the secondary structure elements being put in its own block, whereas in the latter case, stretches of residues of the same length are put in the same number of blocks, but these stretches are randomly chosen.

These results suggest that the subspace considered within the framework of the RTB method captures, on average, most of the atomic displacements involved in the low-frequency normal mode motions of proteins. Figure 4 shows that, in the case of the HIV-1 protease, *each* low-frequency normal mode, as obtained with standard approaches, is also very well described in this subspace, at least for modes with frequencies under $20-40$ cm$^{-1}$. This result is in agreement with previous molecular dynamics and normal mode studies of myoglobin for which it was shown that both diffusional and vibrational motions with frequencies under 60 cm$^{-1}$ can be very well described by a rigid side-chain model.[30,31]

In Table V, the overlap between the conformational change of adenylate kinase upon substrate binding and each of its lowest frequency normal modes is shown (see eq. 3), the normal modes being obtained with either standard or RTB normal mode analysis. In both cases, the overlap of the mode the most " involved " in the conformational change[29] is underlined. When the RTB approach is used with blocks containing 1, 2, 3, or 5 residues, the value of this overlap ($0.53-0.60$) is found to be at least as large as when the standard approach is used (0.53), even though the rank of the corresponding mode is not the same. This
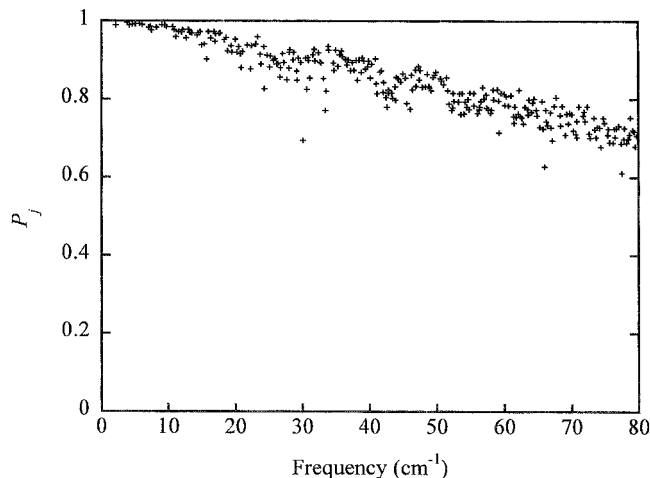
Fig. 4. Quality of the best possible description, within the frame of the RTB approach, of each low-frequency normal mode of the HIV-1 protease, as obtained with standard approaches. A value close to one for $P_j$ means that normal mode $j$ can be perfectly described with our approach [see Eq. (1)].



Fig. 5. Relative displacements of the $C_\alpha$ atoms of one monomer of citrate synthase, as observed upon substrate binding (plain line), and as obtained with the RTB approach, by using two (broken line) or five (dotted line) residues per block, in the case of the third normal mode, that is, the most involved one. Observed and calculated $C_\alpha$ displacements were normalized so as to be of one unity for the largest one.

**TABLE V. Overlap Between the Conformational Change of Adenylate Kinase and Each of Its Lowest-Frequency Normal Modes, as Calculated With the Standard or With the RTB Method, With Different Partitionings[†]**

| | Standard method | | RTB method | | | |
| | | | Number of residues per block | | | |
| Mode number | Frequency $(cm^{-1})$ | Overlap | 1 | 2 | 3 | 5 |
|---|---|---|---|---|---|---|
| 1 | 3.0 | 0.32 | <u>0.58</u> | <u>0.60</u> | <u>0.53</u> | <u>0.60</u> |
| 2 | 3.4 | 0.48 | 0.37 | 0.41 | 0.48 | 0.44 |
| 3 | 4.7 | 0.25 | 0.21 | 0.14 | 0.12 | 0.27 |
| 4 | 4.8 | <u>0.53</u> | 0.23 | 0.26 | 0.29 | 0.02 |
| 5 | 5.0 | 0.17 | 0.37 | 0.23 | 0.26 | 0.12 |

[†]The overlap of the mode the most involved in the conformational change, that is, the one with the largest overlap value, is underlined.

**TABLE VI. Overlap Between the Conformational Change of Citrate Synthase and Its Most Involved Low-Frequency Normal Mode, as Calculated With the RTB Method, With Different Partitionings**

| Number of blocks | Nb of residues per block | Overlap (mode number) |
|---|---|---|
| 858 | 1 | 0.85 (3) |
| 429 | 2 | 0.85 (3) |
| 172 | 5 | 0.84 (3) |
| 149[a] | 1–27[a] | 0.75 (3) |
| 149[b] | 1–27[b] | 0.74 (3) |
| 143 | 6 | 0.82 (3) |

[a]Each secondary structure element is put in a block.
[b]Blocks of same length, but randomly distributed along the polypeptidic chain.

finding means that the RTB approach can slightly modify the ranking of the normal modes. Such a phenomenon was already noticed in our previous study of small test cases,[23] several avoided crossings being observed during the refinement process of the approximate modes obtained with the RTB method, even for a system as small as the decaalanine. However, this latter result is not a general one. For instance, in the case of the two N-terminal domains of CD4, when the three low-frequency normal modes calculated with the standard and with the RTB approach are compared, a one-to-one correspondence is observed, the overlap of the first (lowest-frequency) modes being 0.98, of the second ones, 0.99, and of the third ones, 0.98. In such a case, the RTB method is indeed expected to be very efficient, because the three lowest frequency normal modes of CD4 have been shown to be almost pure rigid-body motions of one domain with respect to another.[32,33]

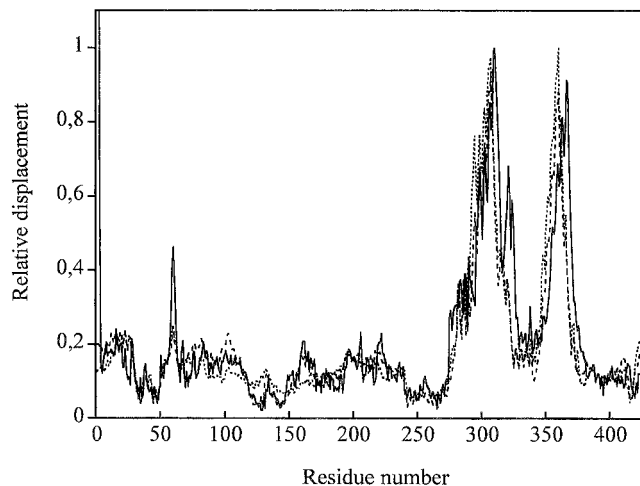In Figure 5, the conformational change of citrate synthase upon substrate binding is shown, as observed when crystal structures 5csc (open form) and 6csc (closed form) are superimposed.[34] In this case, the third lowest-frequency normal mode of the open form, as calculated with the RTB approach, is a collective motion looking strikingly similar to the conformational change, the overlap between these two motions ranging between 0.82 and 0.85, when 1, 2, 5, or 6 residues are gathered in each block (see Fig. 5; Table VI). When each secondary structure element is placed in its own block, each residue lying outside the secondary structures being placed in one block, the value of this overlap drops significantly, down to 0.75, but it is still a quite high value, larger than those obtained for adenylate kinase (see Table V), as well as larger than the value found in a previous normal mode study, namely 0.49, starting from a closed form.[9] Note that, here also, results of the same quality are obtained when stretches of residues of the same length than the secondary structure elements, but picked randomly along the sequence, are placed in their own block, each residue lying outside these stretches being placed in one block. From a practical point

**TABLE VII. CPU Time Required for the Calculation of 50 Normal Modes on a HP Workstation, Using RTB With One Residue per Block, or the Standard Method Available in CHARMM for Large Matrices, Namely, DIMB[†]**

| Protein | Standard method | | RTB method | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Matrix size | DIMB (min) | Projection (min) | Matrix size | DIAGQ* (min) | Total (min) |
| HIV-1 protease | 2,766 | 30 | 0.6 | 594 | 0.9 | 1.5 |
| Triglyceride lipase | 7,491 | 515 | 3.6 | 1,590 | 14 | 17.6 |

[†]Note that DIMB yield exact normal modes.
*The projected matrix was diagonalized with the DIAGQ routine found in the VIBRAN module of CHARMM.

of view, such results, although not in agreement with the intuition that motivated the initial developments of the RTB method,[23] are rather satisfactory because they suggest that the way the partitioning of the protein into blocks is performed has little qualitative consequences on the description of its low-frequency normal modes. However, the two proteins considered in details in the present study are made of only two structural domains. In further works, the case of multidomain proteins will be addressed.

## CONCLUSION

The RTB method rests on the hypothesis that the low-frequency normal modes of proteins, as obtained with standard normal mode analysis, can be described as pure rigid-body motions of blocks of consecutive amino-acid residues.[23] The results obtained during the present work strongly support this hypothesis, because we have shown that the RTB method yields very accurate approximations for the low-frequency normal modes of proteins of various sizes, various folds, etc.

Moreover, the RTB method is a very fast one. In Table VII, a comparison with DIMB is shown, for two proteins studied on a common HP workstation. In these cases, performed with one residue per block, the RTB approach is found to be 20–30 times faster than DIMB. Of course, it can be much faster than that. Indeed, when two residues are placed in each block, the $6n_b \times 6n_b$ matrix to be diagonalized has exactly the same size as in methods where only the $C_\alpha$ atoms are taken into account,[12,13] whereas, when six residues are placed in each block, it has the same size as in the fastest methods allowing for the calculation of B-factors, within the frame of harmonic analysis.[35] As a consequence, very large systems can be studied with all these methods, by using little CPU time and little computer memory. Furthermore, in the RTB method the potential energy function considered is the same as in standard normal mode analysis; other fast methods rest on highly simplified potential energy functions, in the line of the single-parameter one proposed by M. Tirion, in which all neighboring atoms are supposed to be linked together.[36] Also, the RTB method is the only quick method yielding all-atom normal modes in a straightforward way.

Thus, the RTB method should prove useful for performing normal mode analysis of large systems. Moreover, because it is a fast method, it can pave the way for the development and application of methods in which the normal modes and the corresponding normal mode coordinates are calculated many times.

## REFERENCES

1. Go N, Noguti T, Nishikawa T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc Natl Acad Sci USA 1983;80:3696–3700.
2. Brooks BR, Karplus M. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc Natl Acad Sci USA 1983;80:6571–6575.
3. Levitt M, Sander C, Stern PS. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. J Mol Biol 1985;181:423–447.
4. Harrison W. Variational calculation of the normal modes of a large macromolecule: methods and some initial results. Biopolymers 1984;23:2943–2949.
5. Brooks B, Karplus M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. Proc Natl Acad Sci USA 1985;82:4995–4999.
6. Gibrat JF, Go N. Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. Proteins 1990;8:258–279.
7. Seno Y, Go N. Deoxymyoglobin studied by the conformational normal mode analysis: I. Dynamics of globin and the heme-globin interaction. J Mol Biol 1990;216:95–109.
8. Seno Y, Go N. Deoxymyoglobin studied by the conformational normal mode analysis: II. The conformational change upon oxygenation. J Mol Biol 1990;216:111–126.
9. Marques O, Sanejouand YH. Hinge-bending motion in citrate synthase arising from normal mode calculations. Proteins 1995;23: 557–560.
10. Perahia D, Mouawad L. Computation of low-frequency normal modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin. Comput Chem 1995;19:241–246.
11. Hayward S, Kitao A, Berendsen HJC. Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. Proteins 1997;27:425–437.
12. Hinsen K. Analysis of domain motions by approximate normal mode calculations. Proteins 1998;33:417–429.
13. Hinsen K, Thomas A, Field MJ. Analysis of domain motions in large proteins. Proteins 1999;34:369–382.
14. Hayward S, Go N. Collective variable description of native protein dynamics. Annu Rev Phys Chem 199546:223–250.
15. Sanejouand YH. E'tude thêorique des mouvements internes de grande amplitude de la décaalanine et du fragment C-terminal de la proteine ribosomale L7/L12. Ph.D. thesis, Orsay (France), 1990.
16. Elezgaray J, Sanejouand YH. Modeling large-scale dynamics of proteins. Biopolymers 1998;46:493–501.
17. Case D. Normal mode analysis of protein dynamics. Curr Opin Struct Biol 1994;4:285–290.
18. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
19. Mouawad L, Perahia D. DIMB: diagonalization in a mixed basis. a method to compute low-frequency normal modes for large macromolecules. Biopolymers 1993;33:569–611.
20. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan

S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comp Chem 1983;4: 187–217.

21. Thomas A, Field MJ, Mouawad L, Perahia D. Analysis of the low frequency normal modes of the T-state of aspartate transcarbamylase. J Mol Biol 1996;257:1070–1087.

22. Guilbert C, Pecorari F, Perahia D, Mouawad L. Low frequency motions in phosphoglycerate kinase. a normal mode analysis. Chem. Phys. 1995;204:327–336.

23. Durand P, Trinquier G, Sanejouand Y-H. A new approach for determining low-frequency normal modes in macromolecules. Biopolymers 1994;34:759–771.

24. Goldstein H. Classical mechanics. Reading, Mass: Addison Wesley; 1950.

25. Brooks BR, Janezic D, Karplus M. Harmonic analysis of large systems: I. Methodology. J Comp Chem 1995;16:1522–1542.

26. Gimperle F, Gadea FX. Beyond Born-Oppenheimer spectroscopic study of the C state of LiH. J Chem Phys 1999;110:11197–11205.

27. Swaminathan S, Ichiye T, Van Gunsteren WF, Karplus M. Time dependence of atomic fluctuations in proteins: analysis of local and collective motions in bovine pancreatic trypsin inhibitor. Biochemistry 1982;21:5230–5241.

28. Levy RM, Perahia D, Karplus M. Molecular dynamics of an α-helical polypeptide: temperature dependence and deviation from harmonic behavior. Proc Natl Acad Sci USA 1982;79:1346–1350.

29. Ma J, Karplus M. Ligand-induced conformational changes in ras p21: a normal mode and energy minimization analysis. J Mol Biol 1997;274:114–131.

30. Furois-Corbin S, Smith JC, Kneller GR. Picosecond timescale rigid-helix and side-chain motions in deoxymyoglobin. Proteins 1993;16:141–154.

31. Kneller GR, Smith JC. Liquid-like side-chain dynamics in myoglobin. J Mol Biol 1994;242:181–185.

32. Sanejouand YH. Normal-mode analysis suggests important flexibility between the two N-terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. Protein Eng 1996;9:671–677.

33. Sanejouand YH. Rôle du changement de conformation du CD4 lors de la fusion VIH/cellule. C R Acad Sci 1997;320:163–170.

34. Liao DI, Karpusas M, Remington SJ. Crystal structure of an open conformation of citrate synthase from chicken heart at 2.8-A resolution. Biochemistry 1991;30:6031–6036.

35. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold Des 1997;2:173–181.

36. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys Rev Lett 1996;77:1905–1908.