PROTEINS: Structure, Function, and Genetics 23:557–560 (1995)

Hinge-Bending Motion in Citrate Synthase Arising From Normal Mode Calculations

Osni Marques[1] and Yves-Henri Sanejouand[2]

[1]C.E.R.F.A.C.S., 42 avenue Gustave Coriolis, 31057 Toulouse Cédex, France and [2]Laboratoire de physique quantique, C.N.R.S., U.R.A. 505, I.R.S.A.M.C., Université Paul-Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France

ABSTRACT **A normal mode analysis of the closed form of *dimeric* citrate synthase has been performed. The largest-amplitude collective motion predicted by this method compares well with the crystallographically observed hinge-bending motion. Such a result supports those obtained previously in the case of hinge-bending motions of smaller systems, such as lysozyme or hexokinase. Taken together, all these results suggest that low-frequency normal modes may become useful for determining a first approximation of the conformational path between the closed and open forms of these proteins.**

Key words: collective motion, normal mode analysis, low-frequency motion, hinge-bending, closed form of citrate synthase

INTRODUCTION

In many proteins, large conformational transitions involve the relative movement of almost rigid structural elements. In citrate synthase, a two-domain protein, coenzyme A binding induces a 18° rotation of the small domain around an axis close to residue 274, which represents the hinge.[1–3] One consequence of this motion is the closure of the cleft between the two domains, in which the substrate binding site lies (see Fig. 1). As in other cases—hexokinase,[4] phage T4 lysozyme,[5] etc.—such a hinge-bending motion has been probed by X-Ray crystallography.

One of the best suited theoretical methods for studying collective motions in proteins is the normal mode analysis,[6] which leads to the expression of the dynamics in terms of a superposition of collective variables, namely the normal modes coordinates. Until recently, normal mode analysis had been applied mostly to proteins of small size (200 residues or less). Actually, proteins in which hinge-bending occurs and that have been studied so far with this method are hen-egg white lysozyme,[7] human lysozyme,[8,9] phage T4 lysozyme,[10] and hexokinase.[11] In the two former cases, the lowest-frequency normal mode (3.6 and 3.7 cm^{-1}, respectively) compares well with the hinge-bending motion, as it had been previously calculated after an initial guess of the hinge axis.[12] In the case of the rather large hexokinase molecule (around 450 residues), two low-frequency normal modes (7 and 10 cm^{-1}, respectively) of the open form have strong components along the crystallographically observed hinge-bending motion. Though interesting, these later results need to be confirmed since they were obtained with both an approximate method[11] and an approximate protein model (in particular, at the time this calculation was performed, the primary sequence of hexokinase was only known from X-ray crystallography studies; it was later found to be only 30% identical with the exact one). In the case of phage T4 lysozyme, no detailed analysis or comparison with experimental data[5] has yet been published.

We report here a normal mode analysis of the closed form of *dimeric* citrate synthase. In order to study such a large system (around 900 residues), a new program, BLZPACK, was written by one of us (O.M.).

MATERIAL AND METHOD

The normal mode analysis is based on the following idea.[6] In the vicinity of a stationary point, the potential energy of a system, V, can be approximated by

$$V = \frac{1}{2}\sum k_{ij}(r_i - r_i^s)(r_j - r_j^s) \qquad (1)$$

where k_{ij}s are the second derivatives of the potential energy with respect to coordinates r_i and r_j and where r_i^s and r_j^s are the i and j coordinates of the stationary structure. With approximation (1), the equations of motion of the N atoms of the system can be solved analytically, leading to the following solutions:

$$r_i(t) = r_i^s + \frac{1}{\sqrt{m_i}} \sum_{j=1}^{3N} a_{ij} q_j(t),\ i = 1,3N \qquad (2)$$

with:

$$q_j(t) = C_j cos(\omega_j t + \phi_j) \qquad (3)$$

Received March 13, 1995; revision accepted July 13, 1995.

Address reprint requests to Yves-Henri Sanejouand, Laboratoire de Physique quantique, C.N.R.S., U.R.A. 505, I.R.S.A.M.C., Université Paul-Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France.

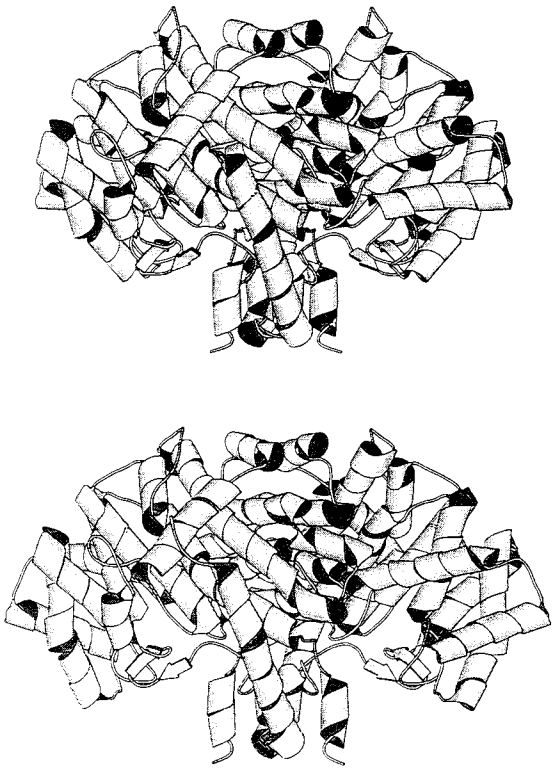

Fig. 1. Crystallographic structures of dimeric pig heart citrate synthase, represented with the Molscript program.[27] Top: closed form (file 2cts of Brookhaven Data Bank[28]). Bottom: open form (file 1cts).

which means that each atomic motion results from the superimposition of $3N$ independent sinusoidal contributions—the normal modes; in these equations, m_i is the atomic mass, C_j and ϕ_j are the amplitude and phasis, respectively, of normal mode j, which depend upon initial conditions, ω_j is the pulsation of normal mode j, obtained as the square root of the jth eigenvalue of the $3N \times 3N$ mass-weighted second derivatives of the potential energy matrix, and vector $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{3Nj})$ is the jth eigenvector of this matrix. Note that, at a given temperature, the lower the frequency of a normal mode, the larger its amplitude. Typically, for proteins, the normal modes whose frequencies lie under 30–100 cm^{-1} are found to be responsible for most of the amplitude of the atomic displacements.[13,14]

As a consequence of computer progress and improvement in the algorithms dedicated to the application of normal mode analysis to macromolecules,[10,15–17] large-size proteins are more and more commonly addressed, such as G-Actin[18] ($3N \approx 11{,}000$), phosphoglycerate kinase[19] ($3N \approx 12{,}000$), and hemoglobin[20] ($3N \approx 17{,}000$). In the present case of dimeric citrate synthase, the system we studied contains 8528 atoms ($3N = 25{,}584$), including coenzyme A and the substrate citrate (CH, CH_2, and CH_3 were treated as extended atoms).

In order to compute the 10 lowest-frequency normal modes of the closed form, besides the six zero-frequency modes corresponding to the overall translational and rotational degrees of freedom of the system, a Block Lanczos algorithm was used, with an inverted matrix formulation.[21,22] The basic idea behind the Lanczos algorithm is the generation of an appropriate basis of vectors for the target problem, so that the projection of this problem into such a basis leads to a smaller eigensystem. Note that the Block Lanczos algorithm is not an approximate method. It converges toward the exact solution of the eigenvalue problem. In order to determine the aforementioned modes with good precision, 120 vectors were required (20 steps, with 6 vectors per block). This computation was performed on one processor of a Cray C90, using 20 minutes of CPU time (it *could have* been performed on a workstation as well, with 128 Mb of memory and 800 Mb of disk space), while the preliminary energy-minimization and matrix computations were performed on an IBM RISC6000/320H workstation, with the CHARMm 21.3 program[23] and the same standard parameters and options used in a previous methodological study.[17] In particular, a cutoff at 7.5 Å was used in the calculation of nonbonded interactions, with a shifting function for electrostatics, and a switching function between 6.5 and 7.5 Å for van der Waals interactions.[23]

RESULTS AND DISCUSSION

The relative displacements of the C^αs of monomer I of citrate synthase corresponding to the lowest-frequency (2.59 cm^{-1}) mode of the *closed* form are depicted in Figure 2c. These displacements happen to compare well with the C^α difference vectors between the open and closed forms (Fig. 2a), i.e., the experimentally observed hinge-bending motion. The comparison is even more convincing when the C^α difference vectors are computed between the open form and the energy-minimized closed one (Fig. 2b), i.e., the one we used to perform our study. In order to quantify the similarity between the hinge-bending motion and each of the calculated normal mode motions, the corresponding overlaps were computed, as

$$overlap(mode\ j) = \frac{|\sum a_{ij}(r_i^o - r_i^c)|}{[\sum a_{ij}^2 \sum (r_i^o - r_i^c)^2]^{1/2}} \quad (4)$$

where r_i^c, r_i^o are the atomic coordinates in the closed energy-minimized and in the open crystallographic structures, respectively, after both had been superimposed. An overlap of 1.00 would mean that the patterns of both kind of atomic displacements are perfectly similar. Though the normal modes were computed for the whole dimeric citrate synthase, the motions of each monomer were analyzed separately. In the case of the lowest-frequency mode, and when only atoms of monomer I are considered (Fig. 2c), the overlap is 0.49. Some of the other normal modes, like

those occurring at 4.00 and 4.64 cm^{-1}, are also found to contribute to the hinge-bending motion in monomer I (overlap of 0.39 and 0.21, respectively; see Table I), as well as one of the three translational motions (overlap of 0.49; data not shown). This latter result reflects the fact that the centers of mass of the two monomers are 3.2 Å closer in the closed form than in the open form of dimeric citrate synthase. For monomer II, the results are somewhat different, the 3.16 cm^{-1} normal mode being the closest to the hinge-bending motion (overlap of 0.35; data not shown). The difference between the results obtained for monomer I and II is primarily a consequence of the fact that we did not apply any constraints during the preliminary energy-minimization process; in particular, the symmetry between the structures of the two monomers, which was assumed by crystallographers, was not imposed on the system. Thus, in the energy-minimized dimeric structure we studied, monomers I and II have slightly different conformations (see Fig. 2b).

CONCLUSION

The present results of the normal mode analysis of dimeric citrate synthase support those obtained previously for lysozyme[7–9,12] and hexokinase.[11] In particular, the largest amplitude motion of the closed form of citrate synthase was found to compare well with the crystallographicaly observed hinge-bending motion. Note that this result was obtained without any explicit information on the open form of citrate synthase. In particular, coenzyme A, which is not present in the open structure, was included in our calculation (coenzyme A is known to induce the closure of citrate synthase[1,3]).

Taken together, all these results suggest that low-frequency modes may provide a useful approximation for the conformational path between the closed and open forms of citrate synthase. A method designed to obtain such a guess from low-frequency normal modes is presently being developed by L. Mouawad and D. Perahia for the transition between the T and R states of hemoglobin (L. Mouawad and D. Perahia, private communication). In their method, starting from the T state, the protein is first displaced in the subspace defined by a few low-frequency modes in such a way that it comes closer to the R state. Next, the energy is minimized, the normal modes are computed, and the protein is displaced

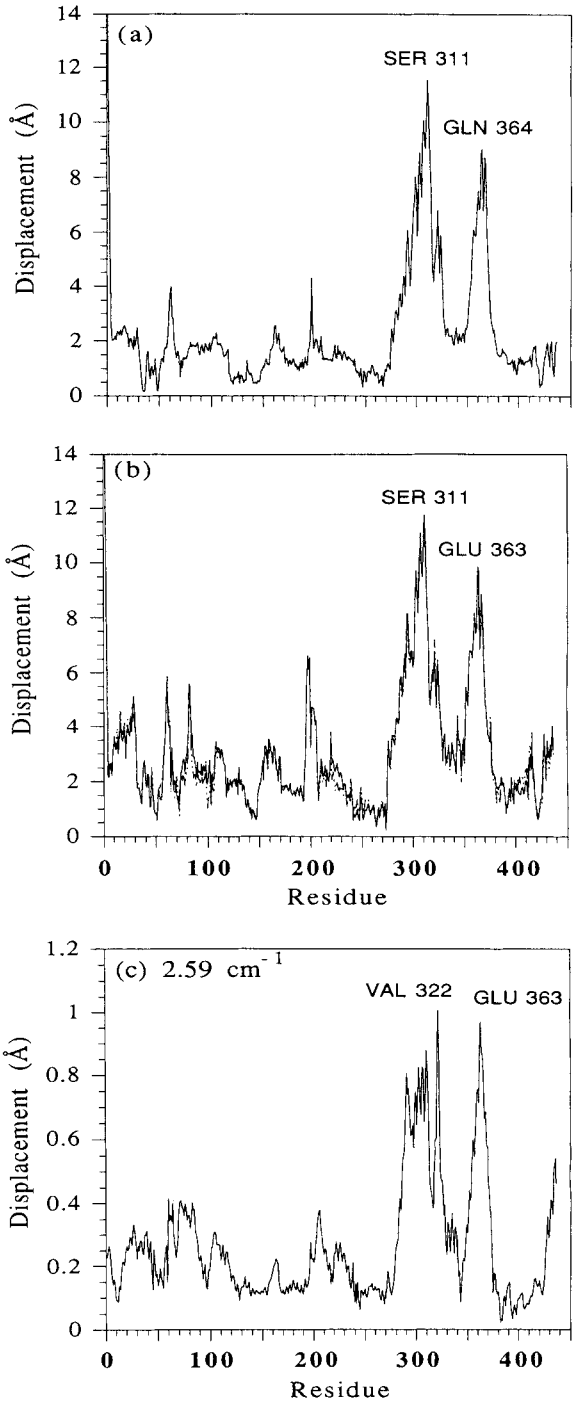

Fig. 2. Comparison of experimental and theoretical data. C^{α} displacements, as a function of residue number, (**a**) between the open and closed crystallographic structures of citrate synthase, after both were superimposed. The root-mean-square distance between the two structures is 3.0 Å (when only C^{α} atoms are considered). The largest secondary structures displacements are those of helices N (residues 274–291), O (residues 297–312), Q (residues 344–365), and R (residues 373–386); (**b**) between the open crystallographic and the closed energy-minimized structures, after both had been superimposed. Continuous line: monomer I. Dashed line: monomer II. Though the root-mean-square distance between the closed crystallographic and the closed energy-minimized structures is 1.7 Å, the overall characteristics of the experimental hinge-bending motion are found not to be altered by the minimization process. (**c**) 2.59 cm^{-1} normal mode. The relative atomic displacements from the closed energy-minimized structure of monomer I are computed according to Eq. (2), C_j being chosen in such a way that the potential energy cost is 3 kcal/mol. In order to have atomic displacements as large as in the experimentally observed hinge-bending motion, a potential energy of more than 400 kcal/mol would be required, according to the normal mode theory.

TABLE I. Comparison of Theoretical and Experimental Data*

Mode frequency (cm^{-1})	Overlap with hinge-bending
2.59	0.49
3.16	0.06
3.29	0.06
3.57	0.05
3.80	0.13
4.00	0.39
4.08	0.09
4.28	0.01
4.57	0.01
4.64	0.21

*Overlap between $\mathbf{a}_j$, the atomic displacements corresponding to each normal mode j, and $\Delta\mathbf{r}$, the difference vector between the closed energy-minimized and the open crystallographic structures, i.e., the hinge-bending motion in citrate synthase (see text).

again in this newly defined subspace. The process is then repeated until a point close enough to the R state is reached. In the citrate synthase case, it will be interesting to compare the path obtained with such a method with the two paths previously determined in our laboratory[24,25]: one with a modified version of the chain algorithm of Elber and Karplus,[26] and the other one with the 'directed dynamics' algorithm of Ech-Cherif El-Kettani and Durup.[24]

ACKNOWLEDGMENTS

We thank Roland Glowinsky, Mario Arioli, and Jean Durup, for fruitful discussions, and Georges Trinquier, for his critical reading of the manuscript. We also thank the referee for his helpful suggestions. BLZPACK, the program with the Block Lanczos algorithm we used, is available on the Mosaic Home Page of C.E.R.F.A.C.S. in Toulouse (hffp://www.cerfacs.fr).

REFERENCES

1. Remington, S., Weigand, G., Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. J. Mol. Biol. 158:111–152, 1982.
2. Huber, R., Bennett, W. S. Functional significance of flexibility in proteins. Biopolymers 22:261–279, 1983.
3. Wiegand, G., Remington, S. Citrate synthase: Structure, control, and mechanism. Annu. Rev. Biophys. Biophys. Chem. 15:97–117, 1986.
4. Bennett, W. S., Steitz, T. A. Structure of a complex between yeast hexokinase A and glucose. II: Detailed comparisons of conformation and active site configuration with the native hexokinase B monomer and dimer. J. Mol. Biol. 140:210–230, 1980.
5. Faber, H. R., Matthews, B. W. A mutant T4 lysozyme displays five different crystal conformations. Nature (London) 348:263–266, 1990.
6. Goldstein, H. "Classical Mechanics." Reading, MA: Addison-Wesley, 1950.
7. Brooks, B., Karplus, M. Normal modes for specific motions of macromolecules: Application to the hinge-bending mode of lysozyme. Proc. Natl. Acad. Sci. U.S.A. 82:4995–4999, 1985.
8. Gibrat, J. F., Go, N. Normal mode analysis of human lysozyme: Study of the relative motion of the two domains and characterization of the harmonic motion. Proteins 8:258–279, 1990.
9. Horiuchi, T., Go, N. Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes. Proteins 10:106–116, 1991.
10. Mouawad, L., Perahia, D. DIMB: Diagonalization in a mixed basis. A method to compute low-frequency normal modes for large macromolecules. Biopolymers 33:569–611, 1993.
11. Harrison, W. Variational calculation of the normal modes of a large macromolecule: Methods and some initial results. Biopolymers 23:2943–2949, 1984.
12. McCammon, J. A., Gelin, B. R., Karplus, M., Wolynes, P. G. The hinge-bending mode in lysozyme. Nature (London) 262:325–326, 1976.
13. Swaminathan, S., Ichiye, T., van Gunsteren, W. F., Karplus, M. Time dependance of atomic fluctuations in proteins: Analysis of local and collective motions in bovine pancreatic trypsin inhibitor. Biochemistry 21:5230–5241, 1982.
14. Levy, R. M., Perahia, D., Karplus, M. Molecular dynamics of an α-helical polypeptide: Temperature dependance and deviation from harmonic behavior. Proc. Natl. Acad. Sci. U.S.A. 79:1346–1350, 1982.
15. Hao, M.-H., Harvey, S. C. Analyzing the normal mode dynamics of macromolecules by the component synthesis method. Biopolymers 32:1393–1405, 1992.
16. Hao, M.-H., Scheraga, H. A. Analyzing the normal mode dynamics of macromolecules by the component synthesis method: Residue clustering and multiple-component approach. Biopolymers 34:321–335, 1994.
17. Durand, P., Trinquier, G., Sanejouand, Y.-H. A new approach for determining low-frequency normal modes in macromolecules. Biopolymers 34:759–771, 1994.
18. Tirion, M., ben-Avraham, D. Normal mode analysis of G-actin. J. Mol. Biol. 230:186–195, 1993.
19. Perahia, D. Application de la méthode des modes normaux à l'étude de la flexibilité des protéines. In "Images de la Recherche. Les systèmes moléculaires organisés." Paris, France: C.N.R.S., 1994:191–194.
20. 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, in press.
21. Parlett, B. N. "The Symmetric Eigenvalue Problem." Englewood Cliffs, NJ: Prentice-Hall, 1980.
22. Grimes, R. G., Lewis, J. G., Simon, H. D. "A shifted block Lanczos algorithm for solving sparse symmetric eigenvalue problems." Seattle, WA: Boeing Computer Services, 1991.
23. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comp. Chem. 4:187–217, 1983.
24. Ech-Cherif El-Kettani, M. A., Durup, J. Theoretical determination of conformational paths in citrate synthase. Biopolymers 32:561–574, 1992.
25. Ech-Cherif El-Kettani, M. A., Zakrzewska, K., Durup, J., Lavery, R. An analysis of the conformational paths of citrate synthase. Proteins 16:393–407, 1993.
26. Elber, R., Karplus, M. A method for determining reaction paths in large molecules: Application to myoglobin. Chem. Phys. Lett. 139:375–380, 1987.
27. Kraulis, P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. 24:946–950, 1991.
28. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.