ELSEVIER

The STATIS method: Characterization of conformational states of flexible molecules from molecular dynamics simulations in solution

R. Coquet, L. Troxler, and G. Wipff

Laboratoire MSM, URA 422 CNRS, Institut de Chimie, 67000 Strasbourg, France

STATIS, a data analysis method used when data can be expressed as matrices, seems particularly well suited to characterize the internal molecular motions and conformational states extracted from MD trajectories. We first outline this method and the "adapted STATIS" method. Applications are presented for 18-crown-6 (simulated for 2 nsec in acetonitrile solution) and for the $(L30)_2Cu^+$ catenate (stimulated for 150 psec in chloroform). STATIS should be valuable for the classification of molecular conformations and simplified visualization of MD trajectories. © *1996 by Elsevier Science Inc.*

Keywords: conformational analysis, STATIS, molecular dynamics, classification, 18-crown-6, catenate

INTRODUCTION

When simulated in solution by molecular dynamics ("MD") techniques,[1,2] flexible molecules undergo motions corresponding to overall translation, rotation, and internal motions ("vibrations" plus conformational changes). During the gas-phase simulations, as the translation of the center of mass and the rotation of the molecule are generally removed directly, the saved trajectory corresponds more or less to internal motions. Simulations in solution are generally performed in a solvent box where the solute diffuses, rotates, vibrates, and undergoes conformational changes. In the following, we focus on the important question of conformational states of the solutes produced by such simulations: the search for a small number of representative classes of conformers and the description of the time-dependent exchange between one class and the others. In some cases, for example the globular proteins, the conformational changes found at 300 K correspond to "microstates," i.e., to fluctuations of relatively small amplitudes around an average structure.[3–5] More often, however, the conformers are drastically different; this would be the case on the denaturation of proteins, or for small flexible solutes [e.g., peptides,[6] macro(poly)cyclic molecules in supramolecular chemistry[7–9]] on timescales ranging from 100 psec to a few nanoseconds.[10] As the temperature is increased, the conformational space is sampled more efficiency.[11]

To characterize the conformers, a statistical analysis can be performed on the Cartesian coordinates of the solute after removal of its global translational and rotational motions.[4,12–16] Such an approach has severe problems with very flexible molecules whose conformation and overall shape change during the dynamics.[12,13] We propose a classification approach based on a statistical study of scalar products between interatomic vectors, which remain invariant on global translational and rotational motions. The scalar products correspond to a superposition of "noselike" and of "conformational-like" motions, which are first separated. This is achieved by a principal component analysis (PCA) based on the eigenvectors of the covariance matrix of scalar products. A hierarchical classification is then performed on the "most significant" motions. The method is briefly outlined here, and followed by two applications in supramolecular chemistry: 18-crown-6 simulated for 2 nsec in acetonitrile and the $(L30)_2Cu^+$ catenate simulated for 150 psec in chloroform.

Address reprint requests to: G. Wipff, Laboratoire MSM, URA 422 CNRS, Institut de Chimie, 4, rue B. Pascal, 67000 Strasbourg, France.

Received 14 May 1996; accepted 14 August 1996.

655 Avenue of the Americas, New York, NY 10010

0263-7855/96/$15.00
PII S0263-7855(96)00064-1

The STATIS method

STATIS[17–22] is a data analysis method that has been developed to extract the most meaningful information stored in a set of two-dimensional (2D) matrices, indexed by time. It is fully explained in Refs. 22 and 23. To our knowledge, it has not been applied so far to analyze molecular motions or trajectories. The general idea is to perform a statistical analysis of "intrastructures,"[23] defined by vectorial inner products, in order to filter the "noiselike" deformations and keep only the statistically most significant deformations. This is followed by a hierarchical classification of the inner products from which the most representative conformers can be extracted. In the following we briefly outline the procedure we have applied.

Let's note N_{atom}, the number of atoms of the solute, and N_{set}, the number of MD sets of coordinates to analyze. The atomic coordinates are stored at each time t in a matrix X_t of $(N_{atom} \times 3)$ dimension.

First, at each time t, from the matrix $\mathbf{X}_t$ one calculates the *intrastructure* matrix $\mathbf{W}_t(N_{atom} \times N_{atom})$, defined by the scalar product

$$\mathbf{W}_t = \mathbf{X}_t \cdot \mathbf{X}_t^* \tag{1}$$

where $\mathbf{X}_t^*$ is the transposed matrix of $\mathbf{X}_t$. We then perform a PCA-like analysis of the $\mathbf{W}_t$ matrices, based on the eigenvalues and eigenvectors of a covariance matrix **S.** This matrix, called the *interstructure* matrix in original STATIS,[22] measures the similarity between the intrastructures. Each $S_{tt'}$ element of **S** is defined by a scalar product between the symmetrical $\mathbf{D} \cdot \mathbf{W}_t$ and $\mathbf{D} \cdot \mathbf{W}_{t'}$ matrices:

$$S_{tt'} = \text{Trace}[\mathbf{D} \cdot \mathbf{W}_t \cdot \mathbf{D} \cdot \mathbf{W}_{t'}] \tag{2}$$

where **D** is a user-defined diagonal matrix of "weights" (for instance, $D_{ii} = 0$ on hydrogen atoms i and $D_{jj} = 1$ on nonhydrogen atoms, to perform the statistics on the latter only). The **S** matrix is of $N_{set} \times N_{set}$ dimension. The user can weight some of these sets via a diagonal matrix Δ (e.g., to skip the equilibration phase for the statistical analysis, while retaining the coordinates for further studies). Diagonalization of the $\mathbf{S}\Delta$ matrix leads to N_{set} normalized eigenvectors $u_i(t)$ $(i = 1, \ldots, N_{set})$ corresponding to N_{set} eigenvalues λ_i. The full basis of u_i vectors provides the same information as the interstructure matrix **S.** Most of the u_i values correspond to eigenvalues λ_i, which are close to zero. As explained in Ref. 22, the largest eigenvalues correspond to the largest dissimilarity (defined by the **S** matrix), i.e., to the largest molecular deformations. Selection of the p largest eigenvalues provides p eigenvectors u_i $(i = 1, \ldots, p)$, which correspond to the most significant motions $(p \ll N_{set})$.

At a given time t, a $\mathbf{W}_t$ matrix corresponds to a reduced list $u_1(t), u_2(t), \ldots, u_i(t), \ldots, u_p(t)$. We therefore perform a hierarchical classification on the $u_i(t)$ vectors, which is equivalent to the classification of the $\mathbf{W}_t$ matrices, and therefore of the 3D structures. This allows the comparison of structures at different times t. They were grouped into n_c typical classes following the technique described in Ref. 24, using SAS software.[25] In each class C_i of cardinal Card(C_i) the average matris is calculated as

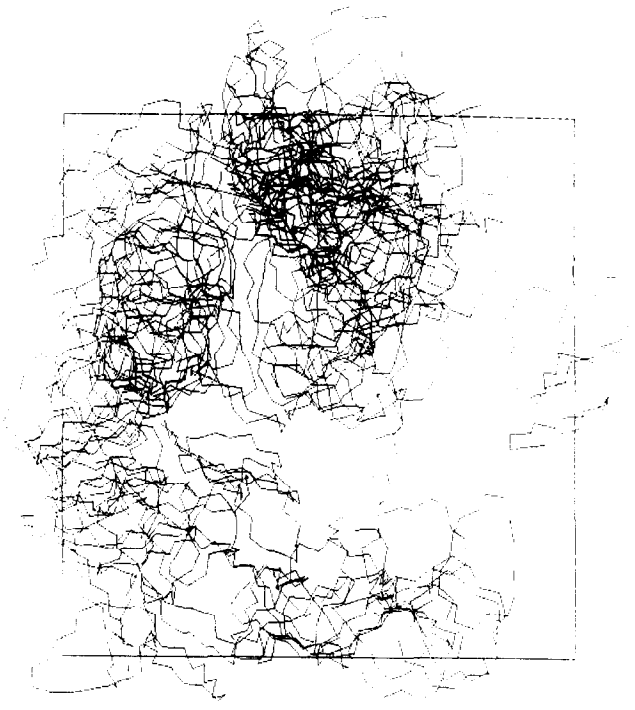

Figure 1. 18C6 in acetonitrile. Cumulated views after 2 nsec of MD simulation (solvent not shown).

$$\langle \mathbf{W} \rangle_{C_i} = \frac{1}{\text{Card}(C_i)} \sum_{W_t \in C_i} \mathbf{W}_t \tag{3}$$

Diagonalization of this matrix is used to calculate a 3D structure. Let X, Y, and Z be the three eigenvectors of $\langle \mathbf{W} \rangle_{C_i}$ corresponding to the three largest eigenvalues. These vectors indeed define, respectively, the Cartesian coordinates (x_i, y_i, z_i) of each atom $(i = 1, \ldots, N_{atom})$, i.e., correspond to an "average structure." The exchange from one class to another as a function of time can be represented as a two-dimensional plot, which simply depicts the conformational changes.

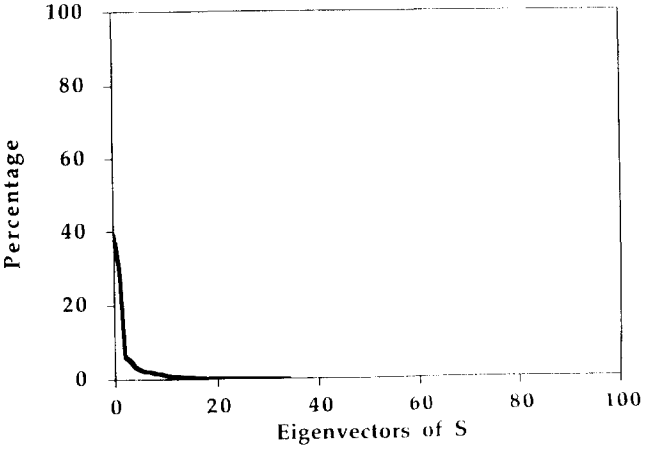

Figure 2. 18C6 in acetonitrile. Relative contribution of the first 100 μ_i eigenvectors to the total (20 000) eigenvalues of the **S** *matrix.*

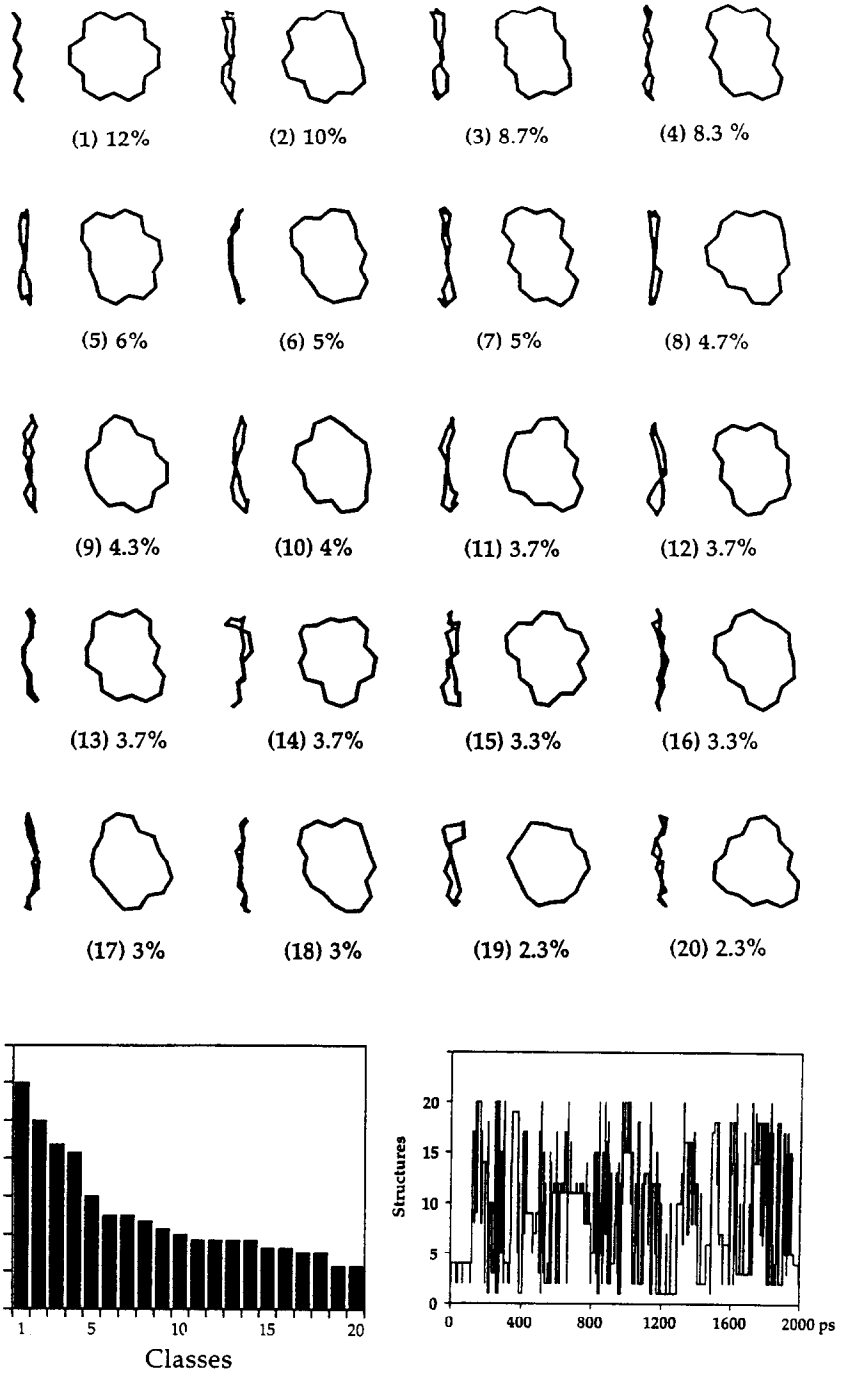

Figure 3. 18C6 in acetonitrile. Analysis of the conformational states on the basis of 20 classes. Top: *Average structure for each class and relative population (%); orthogonal views.* Bottom left: *Relative populations of the 20 classes.* Bottom right: *Time-dependent exchange between the 20 classes.*

It is also possible to reconstruct the "filtered" 3D trajectory, based on the u_i vectors.[13]

Adapted STATIS method

The standard STATIS method described above may be computer time demanding if the number of initial structures (N_{set}) is large. Most of the computer time is devoted to the calculation of the **S** matrix. Significant acceleration can be obtained by modifying STATIS in the following way. After calculation of the N_{set} intrastructure matrices $\mathbf{W}_t$ ($t = 1, \ldots, N_{set}$) as above, one first calculates the average $\langle \mathbf{W} \rangle$ matrix over the *whole simulation* as

$$\langle \mathbf{W} \rangle = \frac{1}{N_{set}} \sum_{t=1}^{N_{set}} \mathbf{W}_t \qquad (4)$$

Diagonalization of $\langle \mathbf{W} \rangle$ provides N_{atom} "basis vectors" onto which the initial $\mathbf{W}_t$ matrices are projected, leading to $\mathbf{W}_t^{\pi}$ matrices. In practice, the dimensionality of this basis can be reduced from N_{atom} to q without important loss of information. The resulting $\mathbf{W}_t^{\pi}$ matrices are therefore of $N_{atom} \times q$, instead of $N_{atom} \times N_{atom}$ dimension ($q \ll N_{atom}$). This procedure can be easily understood in the case of small molecular deformations. As the $\mathbf{W}_t$ matrix is of rank 3 (i.e., has three nonzero eigenvalues, associated to the three moments of inertia of the molecule), the average $\langle \mathbf{W} \rangle$ matrix

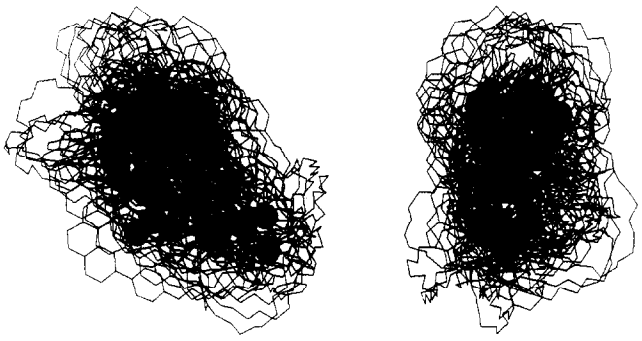

Figure 4. The $(L30)_2Cu^+$ catenate simulated for 150 psec in chloroform: cumulated views (solvent not shown).

also has three nonzero eigenvalues, while the remaining $(N_{atom} - 3)$ eigenvalues are close to zero.

The idea is thus to use the $\mathbf{W}_t^{\pi}$ instead of the $\mathbf{W}_t$ matrices, to compute an approximation $\mathbf{S}'$ of $\mathbf{S}$ [as defined by Eq. (2)] with

$$\mathbf{S}'_{tt'} = \text{Trace}[\mathbf{D} \cdot \mathbf{W}_t^{\pi} \cdot \mathbf{D} \cdot \mathbf{W}_t^{\pi}]. \quad (5)$$

The important time-saving feature, compared to the original STATIS method, results from the fact that the $\mathbf{S}'_{tt'}$ matrix elements are calculated about N_{atom}/q times faster than the $\mathbf{S}_{tt'}$ matrix elements (i.e., ratio of the sizes of $\mathbf{W}_t$ and $\mathbf{W}_t^{\pi}$ matrices). Storage requirements are also reduced by about N_{atom}/q.

The next steps are identical to those in the original STATIS method: diagonalization of the $\mathbf{S\Delta}$ matrix, selection of the p eigenvectors corresponding to the largest eigenvalues λ_i, and of the related conformations $u'_1(t)$, $u'_2(t)$, . . ., $u'_p(t)$.

We show, in the following, using two quite different examples, that $u'_1(t)$, $u'_2(t)$, . . ., $u'_p(t)$ obtained by modified STATIS are a good approximation of $u_1(t)$, $u_2(t)$, . . ., $u_p(t)$ obtained by the original STATIS method. They lead thus to a similar classification of conformations.

RESULTS

Conformation analysis of 18-crown-6 in acetonitrile solution

18-Crown-6 (18C6) is a small molecule (N_{atom} = 42). We analyze its trajectory calculated previously by Troxler and Wipff[26] by MD for 2 nsec in acetonitrile (N_{set} = 20000). As shown by the cumulated view (Figure 1), 18C6 undergoes significant translational and rotational motions in the solvent box.

We performed a STATIS classification on its 18 ring atoms, which define the conformational state. The calculation of the intrastructure $\mathbf{W}$ (42 × 42) and the interstructure S(20000 × 20000) matrices led, after diagonalization, to 20000 u_i eigenvectors and 20000 eigenvalues. The eigenvectors that correspond to the most significant conformational changes are characterized by the largest eigenvalues. Figure 2 represents the relative percentage of the first 100 of them and shows that the first three vectors $u_1(t)$, $u_2(t)$, and $u_3(t)$ contribute to 74% of the total eigenvalues. We therefore selected p = 3 to perform the classification of the 20000 $\mathbf{W}_t$ matrices.

The choice of number of classes n_c is not unique, but user defined. Several attempts were performed with n_c = 10, 20, and 30, respectively. We found that these three classifications led to similar information concerning the most populated classes. In the three cases, the six first classes contain more than 50% of the conformers. Furthermore, they correspond to similar 3D structures. We therefore present in Figure 3 the results for n_c = 20 only. We noticed that for n_c = 30, the last five classes correspond to structures not significantly different from one to another.

The most populated class correspond to the well-known D_{3d} form (12%), while the C_i form (8.3%) in class 4. In class 2 (10%) and class 3 (8.7%) the structures contain fragments of D_{3d} and C_i forms. The structures of Figure 3 can be compared with those of Figure 4 of Troxler and Wipff,[26] who characterized the conformers by the g^+/g^- t sequence of the 18 dihedral angles. According to Troxler and Wipff[26] D_{3d} is also the most populated form (14.3%), followed by the C_i (5.7%). There are interesting differences between the results of the two classification methods. In fact, according to the Troxler and Wipff criteria a change from 119 to 121° is depicted as a g^+ (60 ± 60°) to t (180 ± 60°) conformational change, which may introduce some artifacts. In the STATIS approach, the two structures would be classified in the same class. The populations of the 20 classes are depicted in Figure 3 (bottom left). The conformational flexibility of the crown is illustrated in Figure 3 (bottom right) by the time-dependent exchange between the 20 classes of conformers. As noticed previously, the lifetime of the D_{3d} form (class 1) is the longest. Another interesting feature concerns the sampling of the conformational space. Troxler and Wipff noticed that the number of conformers increased during the simulation without achieving convergence. Figure 3 (bottom right) clearly shows that 18C6 exchanges rapidly from one class to the other, i.e., is not trapped in any conformational region during the simulation.

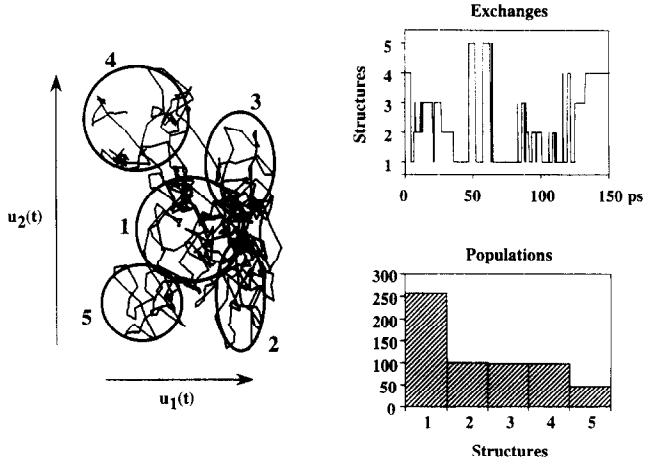

Figure 5. The $(L30)_2Cu^+$ catenate. Left: *$u_1(t)$ versus $u_2(t)$, as a 2D plot of the trajectory in its conformational space (see text).* Right top: *Time-dependent exchange between the five classes of conformations.* Right bottom: *Relative populations of the five classes.*

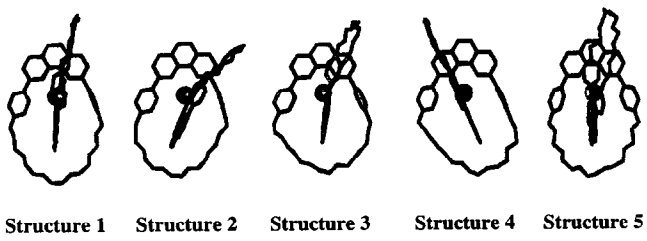

Figure 6. The $(L30)_2Cu^+$ catenate. The five average structures corresponding to classes 1 to 5.

Conformational states of $(L30)_2Cu^+$ catenate in chloroform solution

The $(L30)_2Cu^+$ catenate consists of two L30 interlocked chains wrapping around the Cu^+ cation via their phenanthroline moieties.[27] Each chain is composed of one phenanthroline moiety substituted by two phenyl residues linked by an $O(CH_2CH_2O)_5$ polyether chain. The simulation has been performed in chloroform at 300 K for 150 psec, starting from the X-ray structure.[27] Details of the simulation can be found in Ref. 13. In the simulation with the interlocked topology of the catenate, we were interested in the relative motion of one ring with respect to the other, which is a new feature compared to the 18C6 monocycle. Figure 4 presents cumulated views of this solute during the simulation in chloroform solution.

STATIS was used to analyze the coordinates, saved every 0.25 psec (N_{set} = 600), considering all atoms (N_{atom} = 153 including the hydrogens). Diagonalization of the $\mathbf{S\Delta}$ matrix (600 × 600) leads to 600 eigenvectors $u_i(t)$, which describe fully the trajectory (t = 1, . . ., 600). As for 18C6, the three first (p = 3) represent the highest contribution to the eigenvalues (about 60%). A plot of two of them [e.g., $u_1(t)$ as a function of $u_2(t)$; see Figure 5, left] visualizes the conformational exchange as a function of time. Schematically, the points along these trajectories can be grouped in five main classes (Figure 5), which were therefore selected to classify the $\mathbf{W}_t$ matrices and the related conformers. The five "average structures" corresponding to the five average $\langle\mathbf{W}\rangle$ matrices are shown in Figure 6. Figure 6 clearly shows that the five structures correspond to the progressive "swinging" of one ring with respect to the other involving the phenanthroline and polyether chains. Figure 5 (right top and bottom) represent the time-dependent exchange between these five classes, and the relative population. Not surprisingly, class 1 is the most populated as it corresponds to an intermediate stage where one ring swings over the other i.e., exchanging from one class to another (Figure 5).

DISCUSSION AND CONCLUSION

We present a statistical approach to characterize the conformational states of flexible molecules obtained by MD

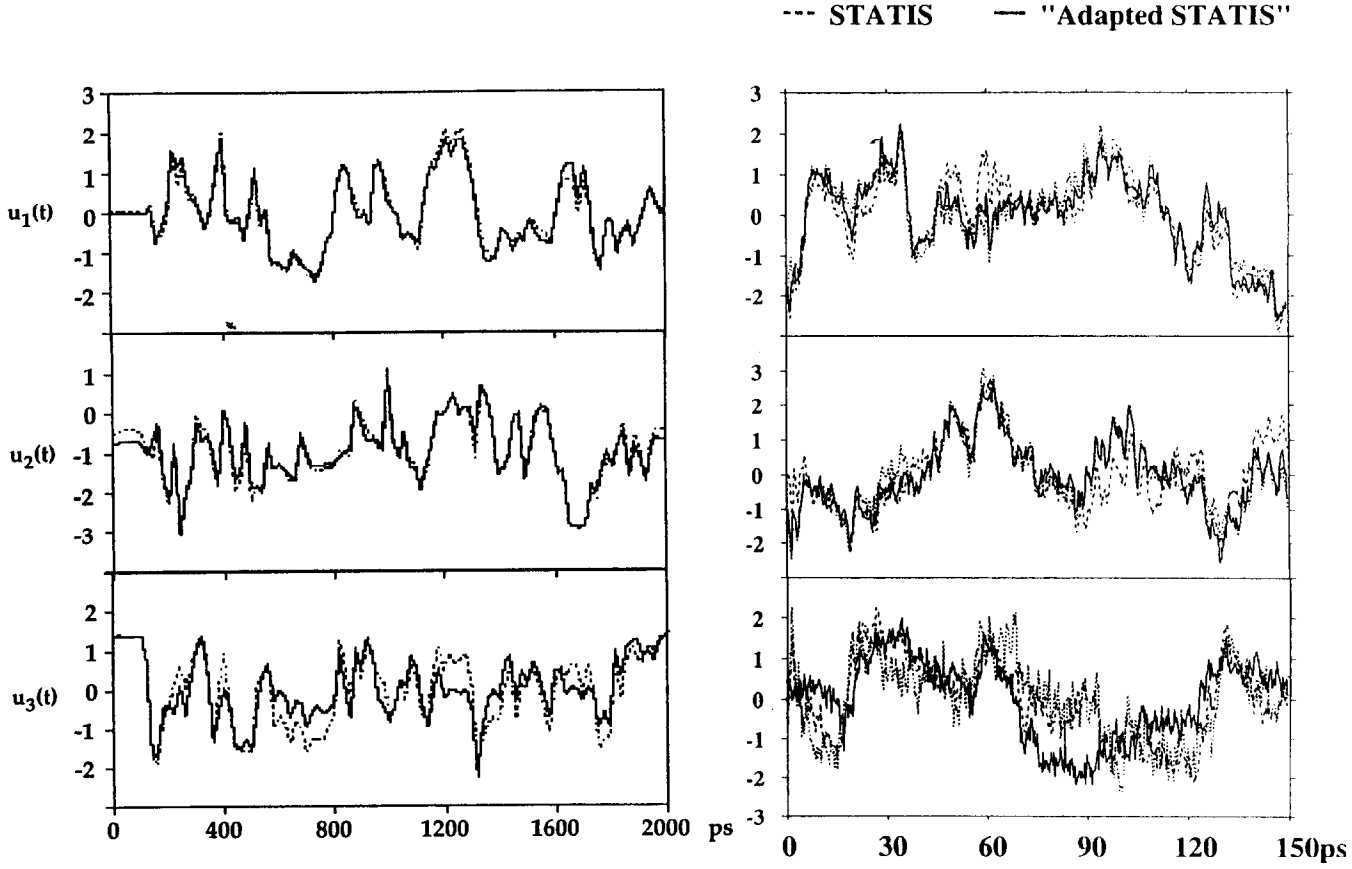

Figure 7. Conformational analysis of 18C6 in acetonitrile (left) and $(L30)_2Cu^+$ catenate in chloroform (right). Display of the $u_i(t)$ eigenvectors obtained by STATIS (solid line) and adapted STATIS (dashed line).

simulations. The trajectories of 18C6 and $(L30)_2Cu^+$ molecules in solution are quite different, as the former undergoes numerous conformational interconversions, and the second more limited "internal motions." An additional difficulty with 18C6 comes from the symmetry degeneracy of some of the conformers. We show that the STATIS approach, followed by a hierarchical classification, is quite robust, as it allows us to identify the most significant classes of conformer and their interconversions, without making a priori assumptions on the nature of the motions. First of all, the overall rotations and translations of the solute are implicitly subtracted from the initial trajectory. The internal motions are studied via the matrices of scalar products between atomic positions. The key feature of the method is the projection of the internal "motions" in a new space of low dimensionality p, based on the eigenvectors of a correlation (*interstructure*) matrix **S** associated to the largest eigenvalues. Only the p eigenvectors corresponding to the largest p eigenvalues correspond to the most significant internal motions.

The use of PCA to study molecular structures and reduce the dimensionality of a problem in chemistry is not new.[28] There have been several applications to analyze MD trajectories in a selected subspace. For instance, Amadei et al. showed that diagonalization of the covariance matrix of atomic displacements allows the separation of the "essential" motions of bovine pancreatic trypsin inhibitor (BPTI) from the local fluctuations.[4] Similar analysis of the MD trajectories have been reported for α helices[16] or calmodulin.[15] These molecules are, however, relatively compact, rigid, and asymmetrical, compared to crown ethers or macrocyclic molecules. As stated above and shown in Ref. 13, the choice of Cartesian coordinates may not be adequate to analyze highly deformable and symmetrical systems like 18C6, or protein like crambin, when it exchanges between very different regions of the conformational space.[29] We feel that performing the simplification and classification of conformational states on the inner vectorial products, instead of the Cartesian coordinates, may in such cases be an interesting alternative.

Adapted STATIS method

The original STATIS is highly computer demanding (35 CPU hours on an HP-735, and 65Mb of memory to analyze the above-described catenate). We therefore considered an alternative approach, adapted STATIS, which is much faster (less than 1 hr, and 30 Mb of memory for the catenate). We tested this new method on 18C6 in acetonitrile and on $(L30)_2Cu^+$ in chloroform solution. The key result is that the first eigenvectors $u_i(t)$ of the **S′** matrix (adapted STATIS) are very close to the first $u_i(t)$ eigenvectors calculated by STATIS. This is illustrated in Figure 7 for 18C6 and $(L30)_2Cu^+$. Since a "conformation" is defined by $u_1(t)$, $u_2(t)$, $u_3(t)$ or by $u'_1(t)$, $u'_2(t)$, $u'_3(t)$, both methods lead therefore to similar conformations, and to similar resulting classifications. The next eigenvectors are more different than the first ones, but they are also of less statistical importance. Thus, despite the use of an average $\langle \mathbf{W} \rangle$ matrix, the adapted STATIS method is performing well on the examples studied. It should be emphasized that the analysis of 18C6, a small molecule, is quite challenging, as the molecule undergoes complex motions like pseudorotation of the ring, and multiple conformational exchanges involving very symmetrical forms like the D_{3d} one, or asymmetrical ones.

These methods are not restricted to small molecules like the ones presented here. They should be particularly valuable in the study of large ones, e.g., biological macromolecules in motion. They should apply as well to structures that have been generated by methods other than MD (e.g., Monte Carlo, EMBED).

ACKNOWLEDGMENTS

The authors thank E. Engler for assistance and IDRIS for generous allocation of computer resources. R.C. thanks the French Ministry of Research for a research grant. G.W. thanks J. Blaudeau for linguistic assistance.

REFERENCES

1 Allen, M.P. and Tildesley, D.J. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.
2 Karplus, M., Brooks, C.L., III, and Pettitt, B.M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics* (Prigogine, I. and Rice, S.A. Eds.). John Wiley & Sons, New York, 1988; Ichiye, T. and Karplus, M. *Proteins Struct. Funct. Genet.* 1987, **2,** 236–259
3 Ichiye, T. and Karplus, M. *Proteins Struct. Funct. Genet.* 1991, **11,** 205
4 Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. *Proteins Struct. Funct. Genet.* 1993, **17,** 412–425
5 Hayward, S., Kitao, A., Hirata, F., and Go, N. *J. Mol. Biol.* 1993, **234,** 1207–1217
6 Abagyan, R. and Argos, P. *J. Mol. Biol.* 1992, **225,** 519–532
7 Wipff, G. *J. Coord. Chem.* 1992, **27,** 7–37
8 Wipff, G. and Wurtz, J.-M. "Dynamic Views of Macrocyclic Receptors: Molecular Dynamics Simulations and Normal Modes Analysis." In *Transport through Membranes: Carriers, Channels and Pumps*, (Pullman, A., Ed.). Kluwer, Dordercht, 1988, p 1–26
9 Sun, Y. and Kollman, P.A. *J. Comput. Chem.* 1992, **13,** 33–40
10 Howard, A.E. and Kollman, P.A. *J. Med. Chem.* 1988, **31,** 1669–1675
11 Auffinger, P. and Wipff, G. *J. Comput. Chem.* 1991, **11,** 19–31
12 Coquet, R. and Wipff G. "Simplification and Characterization of MD Trajectories based on Principal Component Analysis, Multivariate Filtering and STATIS Method." In *XXVIe Journées de Statistique* (Dodge, Y., Ed.). Presses Universitaires de Neuchatel, Neuchatel, Switzerland, 1994, pp 113–118
13 Coquet, R. *Approaches Statistiques pour l'Analyse des Trajectoires de Dynamiques Moléculaires, Applications à des Molécules de la Chimie Supramoléculaire en Phase Gazeuse et en Solution*. Ph.D. Thesis. Université Louis Pasteur, Strasbourg, France, 199
14 Susnow, R., Schutt, C., and Rabitz, H. *J. Comput. Chem.* 1994, **15,** 963–980
15 Haiech, J., Koscielnaik, T., and Grassy, G. *J. Mol.*

Graphics 1995, **13,** 46–48; see also Broto, P., Moreau, G., and Vandycke, C. *Eur. J. Med. Chem.* 1984, **19,** 61
16 Basu, G., Kitao, A., Hirata, F., and Go, N. *J. Am. Chem. Soc.* 1994, **116,** 6307–6315
17 Coppi, R. "Analysis of Three-Way Data Matrices Based on Pairwise Relation Measures. In *COMPSTAT Proceedings in Computation Statistics* (De Antoni, Ed.) Physica-Verlag, Heidelberg Wein. 1986. pp 129–139
18 Carlier, A. *COMPSTAT* 1986, 140–145
19 Foucart, T. *Revue Statistique Appliquée* 1983, **XXXI,** 61–76
20 Lavit, C. Analyse Conjointe de Tableux Quantitatifs. International Center for Pure and Applied Mathematics, Proceedings of the Summer School, 1987
21 Lavit, C. *Statistiques Analyse Données* 1985, **10,** 103–116
22 Escoufier, Y. "Exploratory Data Analysis When Data Are Matrices." In *Recent Developments in Statistical Inference and Data Analysis* (Matusita, K., Ed.). North-Holland Pub. Co., 1980. pp 45–53
23 Lavit, C. *Analyse conjointe de Tableux Quantitatifs.* Masson, Paris, 1988
24 Saporta, G. *Probabilités, Analyse des Données et Statistique.* Ed. Technip., Paris, 1990
25 SAS Institute. *SAS Version 6,* 3rd ed. SAS Institute, Inc. P.O. Box 8000, Cary, North Carolina 27511, 1992
26 Troxler, L. and Wipff, G. *J. Am. Chem. Soc.* 1994, **116,** 1468–1480
27 Cesario, M., Dietrich-Buchecker, C., Sauvage, J.-P., Guilhem, J., and Pascard, C. *J. Chem. Soc. Chem. Commun.* 1985, 244
28 See for instance, Malinowski, E.R. and Howery, D.G. *Factor Analysis in Chemistry.* John Wiley & Sons, New York, 1980; Allen, F.H., Kennard, O., and Taylor, R. *Acc. Chem. Res.* 1983, **16,** 146; Murray-Rust, P. and Motherwell, W.D.S. *J. Am. Chem. Soc.* 1979, **101,** 4373 and references cited therein
29 Karplus, M. Private communication. G.W. thanks M. Karplus for this information.