Conformational non-linear dynamical behavior of the peptide Boc-Gly-Leu-Gly-Gly-NMe

Vincenzo Villani,* Luciano D'Alessio and Antonio M. Tamburro

Dipartimento di Chimica, Università degli Studi della Basilicata, Via Nazario Sauro 85, 85100 Potenza, Italy

The conformational motion of a peptide has been investigated by means of molecular dynamics simulations and has been characterized in the time and the frequency domains. The tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe has been considered as a model for some repetitive amino acid sequences of glycine-rich regions of elastin, which could play a key role in the entropic elasticity mechanism. The classical tools of dynamical analysis (time series, spectral density, time correlation functions, delay maps) have been used and new methods, introducing the idea of travelling trajectory packets and trajectories in the frequency space, have been developed. *In vacuo*, non-ergodic, essentially quasiperiodic motion has been revealed and solitons have been observed according to the non-linear dynamical behavior of systems with small anharmonic perturbation. This particularly concerns the dynamics of the end-to-end distance, a global variable very sensitive to the whole conformational flexibility of the model molecule. The dynamical picture has been joined with the experimentally observed conformational disorder and with the amplitude instability of KAM theory for non-linear systems: a transition is hypothesized from quasiperiodic (solitons) to ergodic motion (Hamiltonian chaos) for the peptide in solution and the entropic mechanism of elastin elasticity is interpreted from a new and unitary dynamical point of view.

Introduction

In recent years the study of complex dynamical systems in many different fields has received growing interest, and a unifying approach has been found to be useful for the description of the universal behavior of time evolutions.[1,2]

From the dynamical point of view a molecule is a complex system because of the highly non-linear bonded and non-bonded interaction forces. In the exact analytical solution of the linear approximate problem, the molecular motion, corresponding to small amplitude oscillations, is given by the superposition of the vibrational normal modes. Nevertheless, if the molecule is a flexible system, which can undergo large conformational changes, the motion cannot be described, even approximately, by the linear theory. In these cases the molecular dynamics approach provides an approximate numerical solution to the exact classical equation of motion and self-organized dynamical behavior, such as chaos or solitons that have no equivalent in the linearized problem, can emerge. This has important consequences in equilibrium statistical mechanics where an equal *a priori* probability for sampling any point on the constant energy surface of an isolated system is assumed: the so-called equiprobability principle or ergodic hypothesis. In Maxwell's words: 'The only necessary argument for a direct proof of the problem of the thermodynamic equilibrium is that the system, if left alone in its present motion state, will pass, sooner or later, through all the phases that are in agreement with the energy equation'.[3] Moreover, Boltzmann, searching for a molecular mechanism that assured at thermal equilibrium the validity of the Maxwell's distribution of the velocities, had to assume final molecular chaos from the outset. Nowadays, the theory of dynamical systems allows us to understand how stochastic behavior can take place in the deterministic non-linear description of reality.[4–7]

In molecular dynamics simulations, the equations of motion are non-integratable and the computed trajectories can be ergodic[8] and chaotic,[9] assuring the evolution of the system from a non-equilibrium to an equilibrium state.[10] Nevertheless, if the anharmonic perturbations are small, non-ergodic essentially quasiperiodic motion can occur and solitons can be observed.

In this work we have simulated the time evolution of an isolated peptide coupled to an external heat bath and characterized its dynamical behavior using the trajectories of structural parameters under the conceptual framework that 'the time series talk by themselves'.[11]

The typical tools of dynamical system analysis, such as time correlation functions, spectral densities and delay maps, have been used and new techniques have been developed by introducing the concepts of a traveling trajectory packet and the trajectory in the frequency domain, with the aim of studying the motion of an isolated tetrapeptide.

The amino acid sequence -Gly-X-Gly-Gly- (X = Val, Leu, Ala) is a repetitive block of elastin and it is believed to play an important role in the entropic protein elasticity.[12] Accordingly, some peptides corresponding to this sequence have been synthesized and studied from an experimental point of view.[13–15] Furthermore, the model tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe has been extensively investigated by using computational techniques.[16–19]

In a previous work[16] an exhaustive conformational analysis was carried out and the minimum energy structures (conformers) were identified and characterized. The most stable conformer is typically identified by a type II β-turn involving the hydrogen bond $[Gly_1]CO \cdots HN[Gly_4]$. In particular, from the anti-correlated behavior of a pair of torsion angles inside the β-turn the existence of a free librational motion of the peptide unit inside the turn was hypothesized, and later observed.[18] Moreover, large -Gly-Gly- chain motions were identified and interpreted as fluctuations occurring between the folded global minimum of the tetrapeptide and the transition state of its conversion to an extended conformation.[18] All these peptide motions could well be a source of molecular entropy.

In the present work a study of the dynamics of the most stable conformer, which best fits the experimental evidence (Fig. 1) has been carried out with the purpose of describing the nature of the molecular motion around this stationary point, and to identify the molecular origin of the possible

contributions to the entropy involved in the elasticity mechanism.

Theoretical model

AMBER 4.0 program[20] was used on a VAX 6510 under the VMS 5.5 operating system. The molecular potential energy has been computed by the Wiener *et al.* all atoms force field[21] where a partition of the energy is assumed, as a sum of stretching, bending, torsional, non-bonded and H-bonded terms.

The theoretical study of dynamics for large biological molecules involves the calculation and analysis of trajectories: the atomic positions and velocities as time functions are given by numerical solution of the classical equations of motion.[22–25] In molecular dynamics (MD) simulations the internal forces are calculated from the gradient of the potential energy, whereas the external ones are determined by the interaction with an external heat bath maintained at constant temperature according to Berendsen's method.[26]

The equations of motion were integrated in cartesian coordinates *via* the Verlet leap-frog algorithm[27] with holonomic constraints of the bond distances at equilibrium values, using the SHAKE routine.[28] In our MD simulation the integration time step was 2 fs, and coordinates and energies were stored every 20 steps. A time period of 200 ps was simulated.

We observe that Berendsen's method is not Hamiltonian: the total energy of the system, at thermal equilibrium with a constant temperature heat bath, can fluctuate. The corresponding dynamical equation includes a non-Newtonian isothermal acceleration term designed to keep the instantaneous kinetic energy constant. For a single oscillator this approach reduces to a Rayleigh non-conservative, non-linear differential equation,[29,30] shown in eqn. (1), which could sustain a self-

$$\ddot{x} - \eta(1 - \dot{x}^2)\dot{x} + F(x) = 0 \qquad (1)$$

excited oscillation,[31,32] η being a damping factor and $F(x)$ a non-linear force. Then, from this point of view, the molecular system can be considered as an ensemble of Rayleigh non-linearly coupled oscillators.

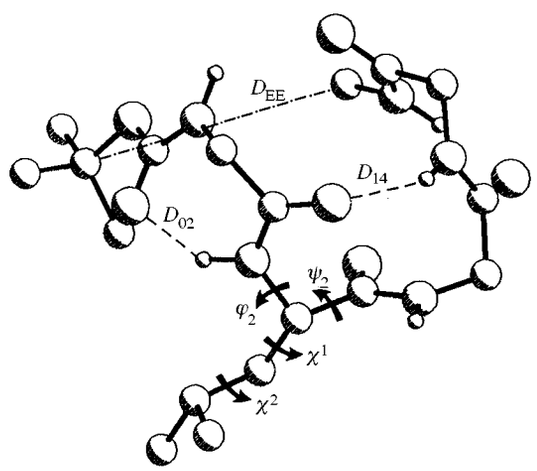

Fig. 1 Lowest energy conformer of the tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe. Hydrogen bonds D_{ij} and end-to-end distance D_{EE} are shown as dashed lines. The investigated torsion angles φ_2, ψ_2, χ^1, χ^2 are indicated.

Initial conditions and equilibrium approach

The starting point of our simulation is the final state (*i.e.* the set of atomic positions and velocities) which lies inside the global minimum well (not corresponding to the minimum point), obtained by MD simulated annealing at 300 K, as reported in a previous work:[17] in this way the initial state should be close enough to the target equilibrium temperature.

In Fig. 2 the temperature and the total, potential and kinetic energies as functions of time during the whole simulation are reported. The relaxation of the system to the equilibrium state, with invariant averages and amplitude fluctuations, is apparent. Moreover, the energy dissipation is due to the damping of the kinetic energy fluctuations which corresponds to the decrease of the potential energy toward the global minimum point.

In MD simulations it is of paramount importance to solve the initialization problem, *i.e.* the discrimination of the initial non-equilibrium transient from the final thermal equilibrium state, which one is interested in. We propose a quantitative criterion based on the decay of total energy as a function of time and have found that the energy dissipation is closely exponential and can be fitted by eqn. (2) with $a = 46.0$ kcal mol^{-1},

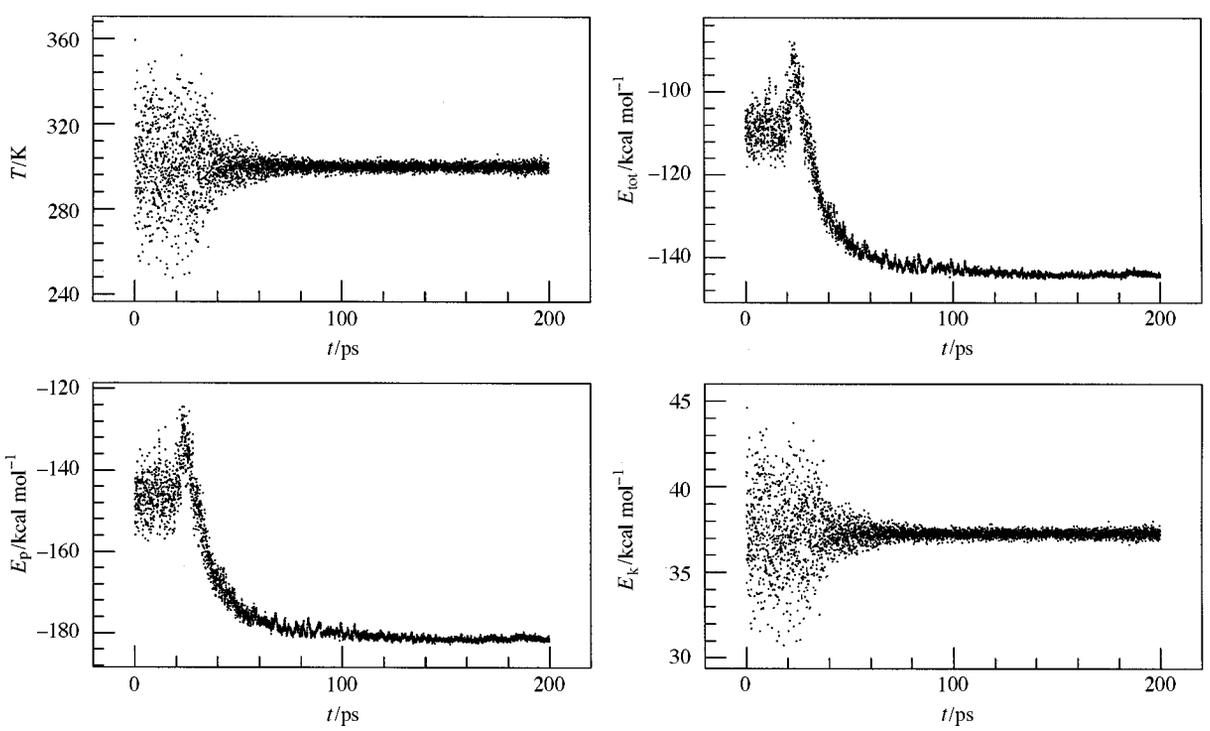

Fig. 2 History plots of temperature (T) and total (E_{tot}), potential (E_p) and kinetic (E_k) energies *vs.* time during 200 ps of MD simulation

$$E(t) = ae^{-bt} - c \quad (2)$$

$b = 0.074$ ps^{-1}, $c = 144.0$ kcal mol^{-1}. From this we assume the end of the transient stage as the time, about 100 ps, when the energy difference with respect to the asymptotic value is less than the fluctuation amplitude δE. Therefore we have studied the dynamics of the tetrapeptide molecule from this point.

Data analysis and numerical methods

The analysis has been focused on the study of the trajectories of the following structural parameters, both in the time and in the frequency domain, and phase space reconstruction from time series has been carried out. Local and global parameters have been studied. Four torsion angles φ_2, ψ_2, χ^1, χ^2 of the leucyl residue were taken into account, because they are the most characteristic of the motion of the tetrapeptide backbone. In addition, two hydrogen bond distances involving the seven- and ten-membered cycles, [Boc]C=O $\cdots$ HN[Leu$_2$] (D_{02}) γ-turn and [Gly$_1$]C=O $\cdots$ HN[Gly$_4$] (D_{14}) β-turns respectively, whose existence is supported by many theoretical and experimental evidences, were taken into account.

Finally a global parameter, namely the end-to-end distance (D_{EE}) between the tertiary carbon atom of Boc and the amidic carbon atom of NMe, has been considered for the study of collective motion, as it is an expression of the folding degree of the molecule.

All of these variables have been analyzed in terms of time normalized auto-correlation functions, see eqn. (3), where $\Delta x = x(t) - \langle x(t)\rangle$ and $\Delta^2 x = [x(t) - \langle x(t)\rangle]^2$.

$$C_x(\tau) = \langle \Delta x(t)\, \Delta x(t+\tau)\rangle / [\langle \Delta^2 x(t)\rangle \langle \Delta^2 x(t+\tau)\rangle]^{\frac{1}{2}} \quad (3)$$

In order to avoid the systematic decay of the statistical accuracy for the correlation function caused by the finite size of the time series,[33] we have chosen the shifted series as one-half of the series after the equilibration time, so that the number of the terms of the scalar product is constant (1250 points). Accordingly, we have considered 100 ps $\leqslant t \leqslant$ 150 ps and $0 \leqslant \tau \leqslant 50$ ps, with a time shifting step $\delta\tau = \delta t = 0.04$ ps. Since the correlation function is a generalization of the correlation coefficient, *i.e.* the cosine of the angle between two vectors whose components are the terms of the time series, we have normalized our results with respect to the square root of the product of variances of the stationary and shifted series. In a similar way we have studied the cross-correlation functions between all the structural parameters mentioned above [eqn. (4)].

$$C_{xy}(\tau) = \langle \Delta x(t)\, \Delta y(t+\tau)\rangle / [\langle \Delta^2 x(t)\rangle \langle \Delta^2 y(t+\tau)\rangle]^{\frac{1}{2}} \quad (4)$$

In addition the corresponding spectral densities are calculated from eqns. (5) and (6), where the frequency has been

$$F_{x(x,y)}(\omega) = \Sigma_\tau C_{x,(x,y)}(\tau) \cos(\omega\tau) - i\Sigma_\tau C_{x,(x,y)}(\tau) \sin(\omega\tau) \quad (5)$$

$$|F_{x(x,y)}(\omega)|^2 = [\Sigma_\tau C_{x(x,y)}(\tau) \cos(\omega\tau)]^2 + [\Sigma_\tau C_{x(x,y)}(\tau) \sin(\omega\tau)]^2 \quad (6)$$

varied up to the Nyquist critical frequency $\omega_{max} = 1/(2\delta\tau)$ with a step of $\delta\omega = 1/\Delta t$, imposed by the sampling and the simulation times. Finally, all structural parameters were studied in a reconstructed phase space,[34,35] through the plots of the two-dimensional delay maps, with the purpose of detecting the attractors. In this space underlying the coordinate axes are the time series and the shifted series is obtained from the original one by moving it to an integer multiple of a given time period. In our case, from the original time series $x(t)$ the translated series $y(t) = x(t + \tau)$ is constructed and a two-dimensional plot is obtained considering $x(t)$ and $y(t)$ as the coordinate of a point P(t). The examination of the path of P(t) gives some information about the attractor of the dynamical system or, more exactly, its projection in a subspace of lower dimension.

The shift parameter τ used in the present work (0.2 ps) was established through an appropriate screening and the selected value corresponds to the appearance of the most ordered structure.

The above analysis gave significant results (that is the appearance of a structured path) only for the global variable end-to-end distance. However, the local variables gave unstructured patterns, indicating a major complexity of motion.

Results

Time series and autocorrelation functions

In Fig. 3A–G the time series of the above mentioned key structural parameters, their autocorrelation functions and Fourier transforms are given. The history plots of dynamical variables are shown for the whole simulation ($0 \leqslant t \leqslant 200$ ps), whereas the autocorrelation functions and the power spectra are reported only for the equilibrium state ($t \geqslant 100$ ps).

The behavior of the time series shows a steady-state, which occurs after the transient stage defined above on the basis of the temperature and energy time evolution.

One observes that the transient is characterized by a number of different behaviors: (*a*) relatively large fluctuations around the equilibrium value of the global minimum conformer (Fig. 3A,B); (*b*) fast deviations from the equilibrium structure to extended short-lived conformations (Fig. 3E,F) as confirmed by the spike in the end-to-end distance trajectory (Fig. 3G); (*c*) a conformational transition from *trans* to *gauche*(+) of the torsion angle χ^2 (Fig. 3D); (*d*) finally, the formation of a medium-lived structure through a transition from *gauche*(−) to *trans* indicated by the jump in the χ^1 torsion angle (Fig. 3C). Points (*c*) and (*d*) agree with the expected mobility of the aliphatic chain.[36,37]

The long-time motion is analyzed in more detail in the autocorrelation functions and power spectra reported in Fig. 3. The correlation functions evidence the possible periodicity of the motion, whose frequency components appear as peaks in the corresponding power spectrum. According to the International System of units the frequencies are expressed in ps^{-1}. In order to facilitate the reading of plots note that 1 ps^{-1} corresponds to 33.356 409 5 cm^{-1}.

From the plot analysis we observe the following patterns: (1) low frequency oscillations, about 1 ps^{-1} (Fig. 3E); (2) high frequency oscillations, about 7 ps^{-1} (Fig. 3C,D,F); (3) a wide frequency spread on the whole explored range (Fig. 3A,B,G). Case (1) is observed only for the hydrogen bond distance [Boc]C=O $\cdots$ HN[Leu$_2$] (D_{02}) involved in the formation of the γ-turn. Case (2) is typical of the torsion angles of the side-chain group and of the hydrogen bond distance D_{14}. A comparison of the frequency peak of D_{14} with that of D_{02} shows more stiffness in the H bond in the 10-membered ring, indicating a greater stability of the larger secondary structure with respect to the other. This confirms previous experimental evidence obtained by the study of NMR temperature coefficients of amide protons. In fact, the NH chemical shift showed a larger temperature dependence for the amide proton involved in the γ-turn with respect to that involved in the β-turn. This is an indication of a greater force constant for the latter bond. On the contrary, the situation is reversed when leucine is substituted by valine,[13] as we will subsequently explain. The vibrational frequency of the χ torsion angles is exactly the same, and this suggests a close coupling between them.

Finally the mixed behavior [case (3)] is observed for the backbone torsion angles and for the end-to-end distance, indicating a more complex motion of the chain. It is apparent that no integer ratios between the frequency values found in the spectrum exist. This situation may lead to complex motions in the cartesian coordinate space, similar to the open paths of the Lissajous figures.[38] In fact, if we consider the effective motion of the atoms around the equilibrium point,

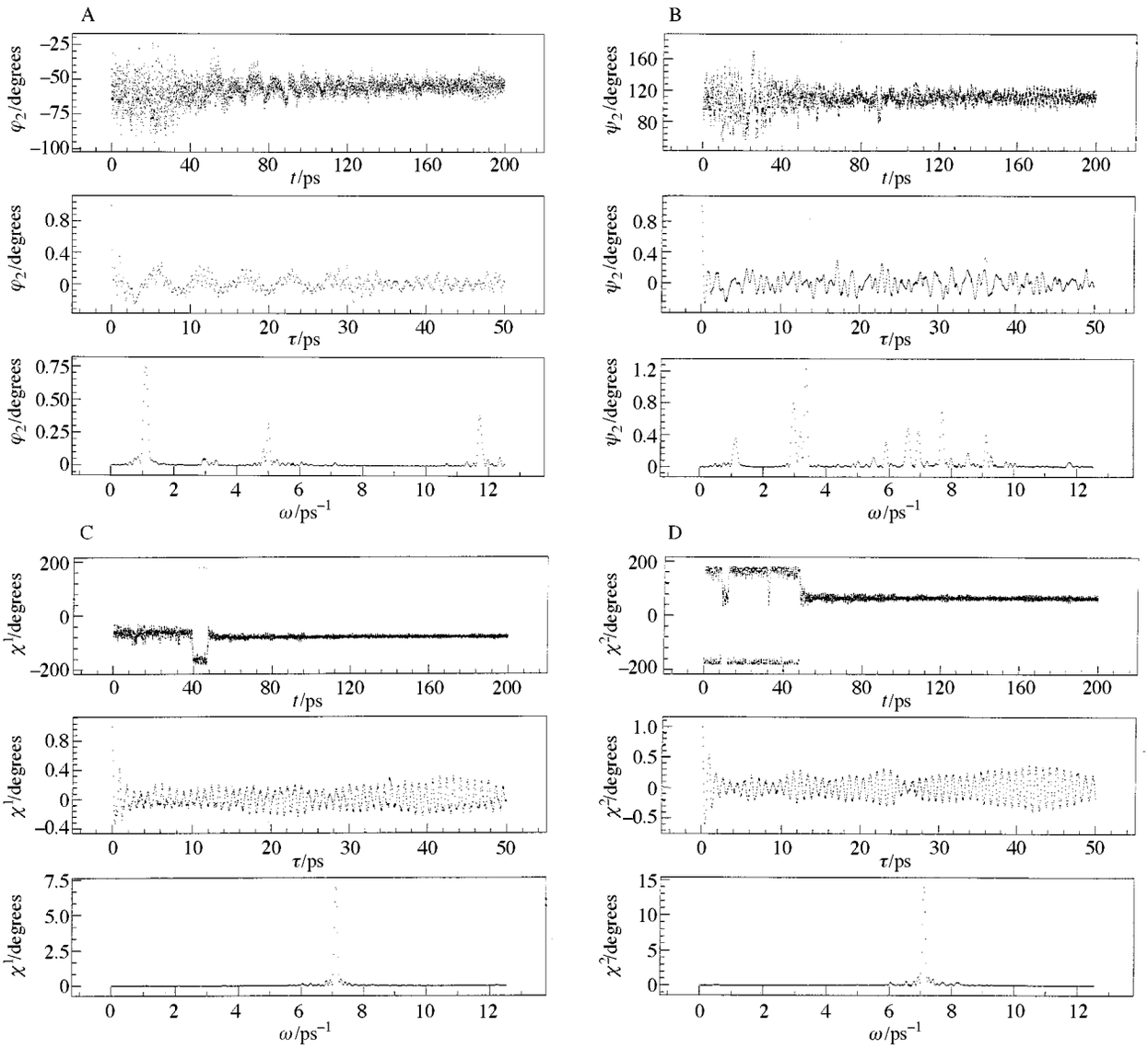

Fig. 3 Time series (top), time autocorrelation function (middle) and spectral density (bottom) of the following dynamical variables: torsional angles (A) φ_2, (B) ψ_2, (C) χ^1, (D) χ^2, (E) H-bond distances D_{02}, (F) D_{14} and end-to-end distance (G) D_{EE}

with irrational frequency ratios, as a superposition of a number of vibrational motions along various directions, then the trajectory will fill densely a subset of the phase space during the motion.

Time cross-correlation functions

In order to understand possible coupling between different motions inside the molecule, we have studied the cross-correlation functions between the dynamic behavior of the structural parameters and their Fourier transforms. The results (21 graphic outputs) indicate: (1) low frequency (about 1 ps^{-1}) correlations; (2) high frequency (about 7 ps^{-1}) correlations; (3) a superposition of low and high frequency correlations. The first type is observed in the following correlations: D_{02}/D_{14}, D_{02}/φ_2, D_{02}/χ^1, D_{02}/ψ_2, D_{14}/φ_2, D_{EE}/φ_2 and D_{EE}/D_{02}. A representative plot of the couple D_{02}/φ_2 is reported in Fig. 4 where a decrease in correlation as a function of time is also observed, leading to fully uncorrelated motions (dissipation of correlation). This indicates a tight low frequency correlation between the hydrogen bond in the seven-membered cycle and the torsion angle of leucine, which is in agreement with the low vibrational frequency of the [Boc]C=O···HN[Leu_2] hydrogen bond previously observed. The driving force of this motion is assigned to the practically free fluctuations of the high-inertial side chain of leucine, and this can be considered to be the cause of the weakness of the previously discussed γ-turn H-bond.

The second type of cross-correlation is observed in the following cases: D_{EE}/χ^2, D_{14}/χ^1, φ_2/χ^2, χ^2/ψ_2, χ^2/ψ_2, χ^1/χ^2 and D_{EE}/D_{14}. In Fig. 4 a sharp monofrequency plot (χ^1/χ^2) and a more representative multifrequency plot (D_{EE}/χ^2) are shown. This behavior seems to be typical of correlations involving the side-chain torsion angles χ^1 and χ^2, and may be explained by considering that the latter refers to displacement of few atomic masses. The observed frequency value is the same as that of the autocorrelations of those variables, showing the tight coupling between adjacent torsions in an aliphatic chain, probably due to the conservation of a local angular momentum. The high frequency cluster observed in the D_{EE}/χ^2 spectrum shows a non-trivial correlation between the global and a local parameter, indicating the high collective molecular motion due to significant coupling between degrees of freedom.

Finally, the third type of spectrum is characterized by a superposition of low and high frequency, and is observed for D_{EE}/ψ_2, D_{02}/χ^2, D_{14}/χ^2, D_{14}/ψ_2, φ_2/χ^1, φ_2/ψ_2, χ^1/ψ_2 and D_{EE}/χ^1. A typical plot is the D_{EE}/ψ_2 reported in Fig. 4. The observed behavior is related to the analogous behavior of the autocorrelations of the same variables and shows a strong coupling between them. The explanation can be based on the observation of the key role played by the Leu group as a central pivot of the molecular collective motion.

In all cases one may observe that an apparently very complex signal is always resoluble in a limited number of harmonic oscil-

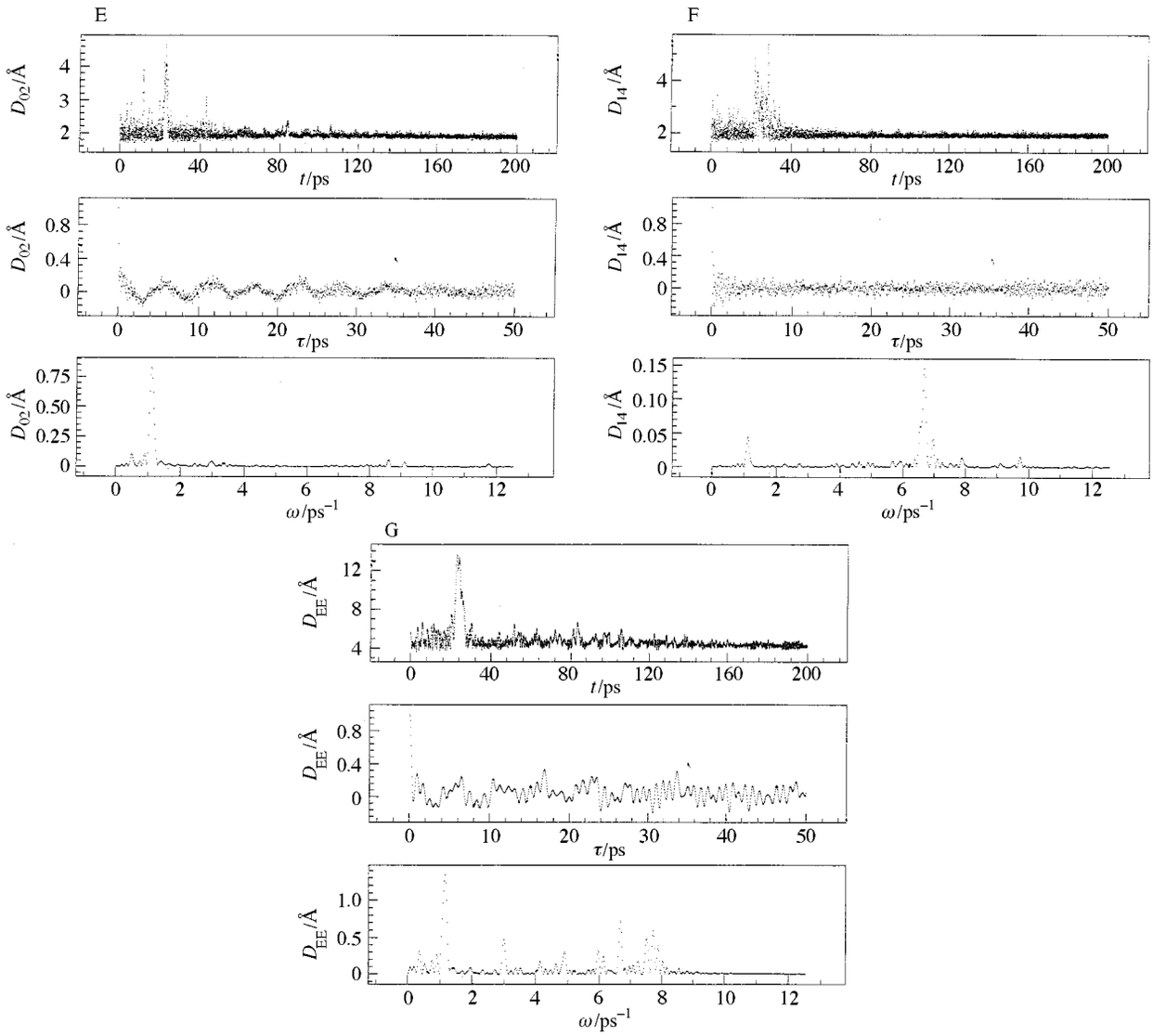

Fig. 3 (*continued*)

lations, indicating that the motion is essentially quasiperiodic. The above observation is in line with the expected behavior for a system moving around its stable equilibrium point.

Trajectory in the frequency domain

Fig. 5 shows the plot of the complex Fourier transform of the end-to-end distance trajectory as a function of the frequency. The trajectory in the frequency domain, that is the dual of the usual trajectory in the time domain, is thus defined. In our opinion, this kind of analysis can be useful for characterizing the motion of the system. In fact, it confirms the essentially quasiperiodic character of the collective molecular motion evidencing a disordered coil, although still well structured, similar to that detected by delay-maps analysis in the time domain. The plot of Fig. 5, whose shape is like a twisted elastic band,[39,40] shows that one can differentiate the essentially quasiperiodic motion seen for our system from a completely random motion and from a few frequency band period motion, where we have verified that a scattered coil or a structured loop are obtained, respectively, as a consequence of the phase correlations. In this way, and at variance with the usual spectral density plot, the whole dynamical information contained in the Fourier transform is retained and can be used as a sound basis for analysis.

Phase space reconstruction

Fig. 6 shows the phase–time trajectory of the end-to-end distance in the space $D_{EE}(t)$, $D_{EE}(t + \tau)$, t (delay time $\tau = 0.2$ ps) during the last 50 ps of the simulation. A structured motion in this fuzzy helix is apparent. In fact our observations strongly support the idea that the prevalent structure of the attractor is essentially a limit cycle described with roughly constant angular velocity. To test the above hypothesis we have studied the motion in the plane $D_{EE}(t)$, $D_{EE}(t + \tau)$ and we have calculated the time dependence of the polar angle from the approximate center of the orbits, and its time derivative. In the following, x is a short-hand notation for D_{EE}. The center C of the loops was calculated as the mean of the $x(t)$, $x(t + \tau)$ data [eqns. (7)

$$x_C = \langle x(t) \rangle \tag{7}$$

and(8)] then the coordinates were translated carrying the origin

$$y_C = \langle x(t + \tau) \rangle \tag{8}$$

in C and a new set of data $x'(t)$, $x'(t + \tau)$ was obtained. The polar angle in respect to the $x'(t)$ axis is calculated as a function of time from eqn. (9).

$$\theta = \tan^{-1} x'(t + \tau)/x'(t) \tag{9}$$

The corresponding discrete time derivative of θ, *i.e.* $\dot{\theta} = \Delta\theta/\delta t$, where $\Delta\theta$ is the angle described during the time interval δt, is obtained from Carnot's theorem [eqn. (10)], which can be satisfied by eqns. (11)–(13).

$$\Delta\theta = \cos^{-1}(r_1^2 + r_2^2 - s^2)/(2r_1r_2) \tag{10}$$

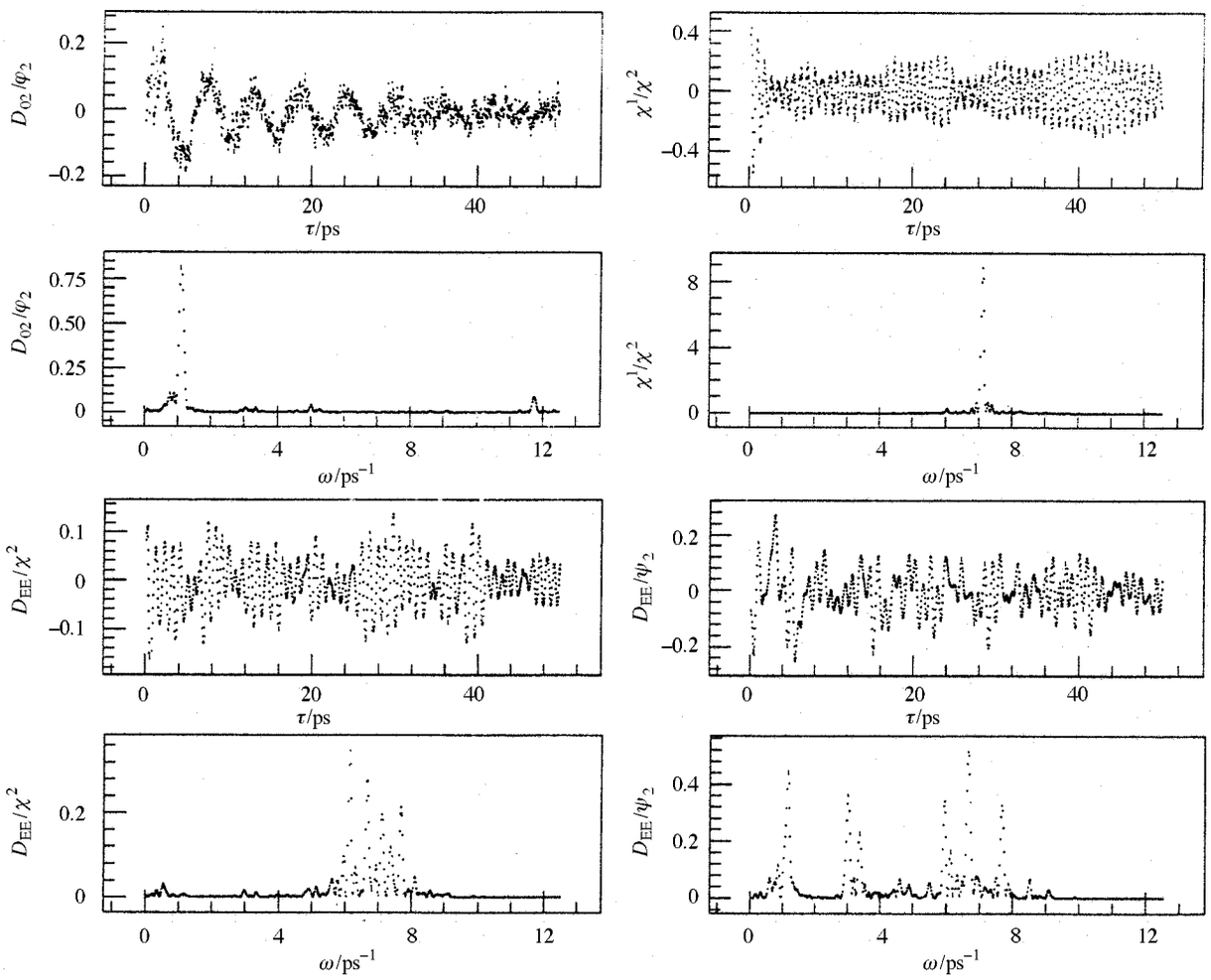

Fig. 4 Time cross-correlation functions and corresponding spectral densities for representative couples of dynamical variables: D_{02}/φ_2, χ^1/χ^2, D_{EE}/χ_2 and D_{EE}/ψ_2

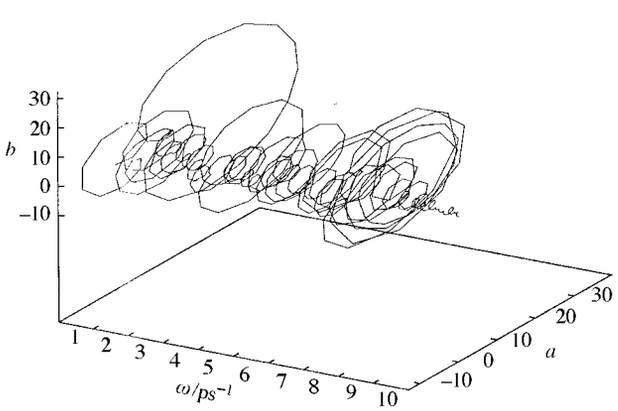

Fig. 5 Plot of the complex Fourier transform of the end-to-end distance D_{EE} as a function $F(\omega) = a + ib$ during the 100 ps of MD thermal equilibrium

$$r_1^2 = x'(t + \tau)^2 + x'(t)^2 \quad (11)$$

$$r_2^2 = x'(t + \delta t)^2 + x'(t + \delta t + \tau)^2 \quad (12)$$

$$s^2 = [x'(t + \delta t) - x'(t)]^2 + [x'(t + \delta t + \tau) - x'(t + \tau)]^2 \quad (13)$$

The results are shown in Fig. 7, where one can see that the polar angle varies from $\pi/2$ to $-\pi/2$ nearly periodically, and the angular velocity is stabilized around a constant value of 7.2 rad ps^{-1} (except for a few localized spikes) corresponding to a frequency of 1.15 ps^{-1}. This is in excellent agreement with the position of the low frequency peak in the power spectrum of

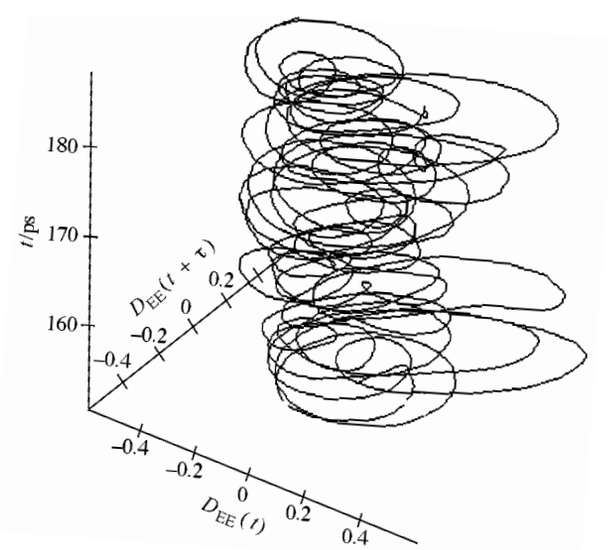

Fig. 6 Phase-time trajectory of the end-to-end distance in the space of $D_{EE}(t)$, $D_{EE}(t + \tau)$ and t for $\tau = 0.2$ ps and $150 < t < 200$ ps

the end-to-end distance correlation function. Therefore the observed coil is an expression of the low frequency motion of the global molecular parameter and may be considered to be a section of the attractor lying in the true phase space.

Note that attempts to obtain similar behavior utilizing local molecular parameters, including those whose time series seems to be rather ordered, failed completely and a random delay map was inevitably obtained. This is a clear demonstration that, at

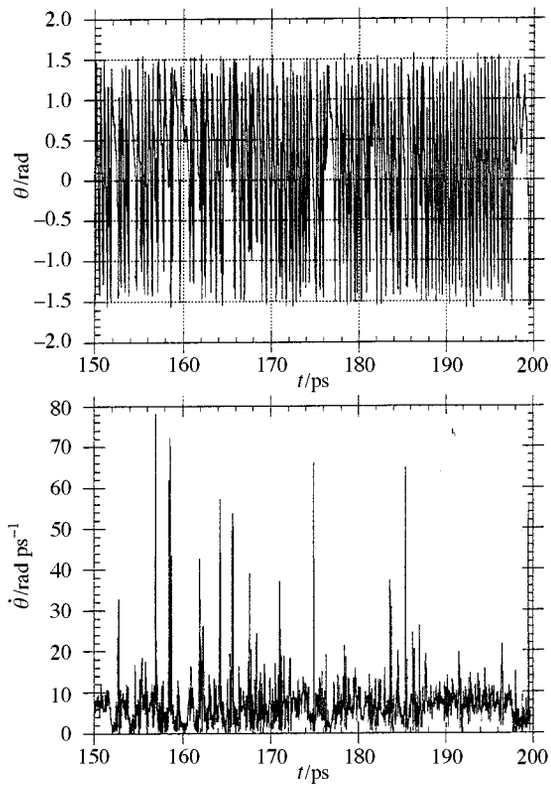

Fig. 7 Time evolutions of polar angle θ and time derivative $\dot{\theta}$ describing the angular position and velocity, respectively, for the projection in the plane $D_{EE}(t)$, $D_{EE}(t + \tau)$ of the trajectory reported in Fig. 6 (see text for detailed discussion)

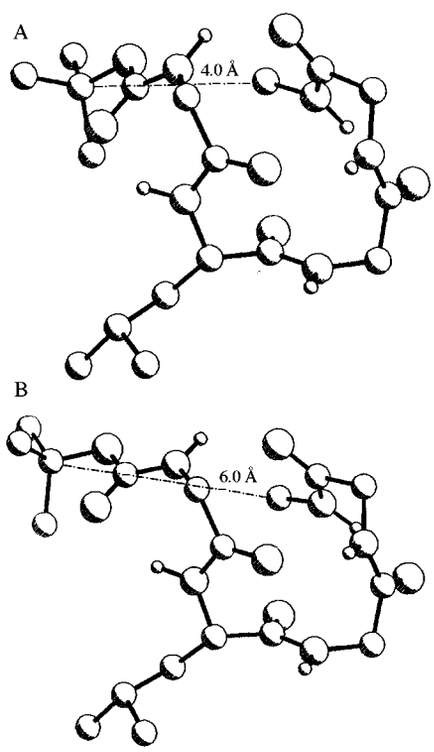

Fig. 8 Conformation snapshots corresponding to the maximum excursion of the end-to-end distance during the MD thermal equilibrium: D_{EE} = 4.0 (A) and 6.0 Å (B)

the considered temperature value, a low frequency ordered motion exists as the basis of the collective behavior of this coupled, non-linear molecular system.

Fig. 8 shows conformations corresponding to the maximum excursion of the D_{EE} distance, ranging from 4 to 6 Å. The corresponding shape of the molecular backbone suggests that

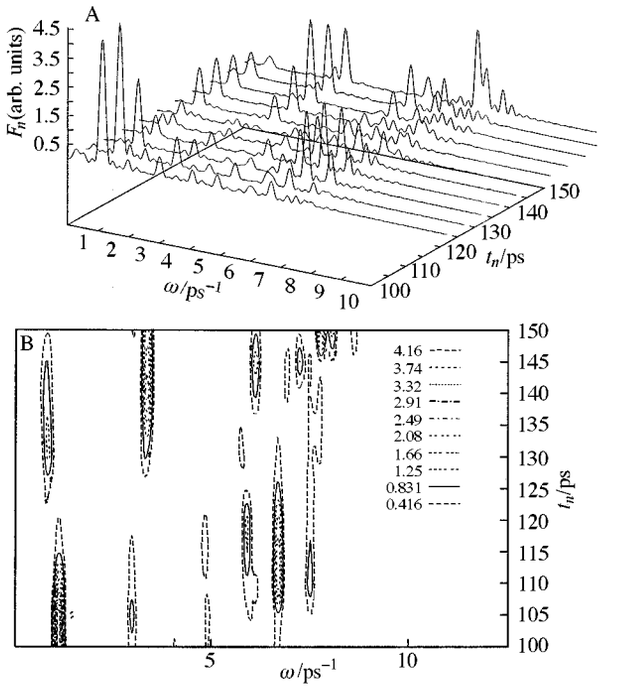

Fig. 9 Dynamical Fourier surface $F_n(\omega,t_n)$ of end-to-end distance D_{EE} as (A) a series of cross-sections and (B) an isointensity map during the MD thermal equilibrium

the low frequency vibration, largely prevalent in the collective molecular motion, corresponds to the fundamental mode of a U-shaped, diapason-like structure.

Traveling trajectory packets

In order to detect and characterize the time-dependent vibrational picture of the molecular system at the thermal equilibrium, we have developed a new method of analysis in which a family of delayed and bounded time series is generated, and their characteristic functions are constructed. In the analysis carried out by Wright,[41] characteristic functions of delayed series were used to monitor changes in dynamical systems. Moreover, the overlapping Fourier transforms are a well known technique in other areas, *e.g.* speech processing or signal processing.[42] By analogy with the wave packets of quantum mechanics the so defined time series may be called traveling trajectory packets (TTP). In particular, we have taken into account the end-to-end distance time series and the TTP of length Δt = 50 ps, starting at $t_n = nT$ (n is an integer number and T = 5 ps) from the equilibrium onset point (at 100 ps), have been considered. This family of packets can be represented as S_n (t, t_n) and it was used to calculate the correlation functions, normalized in each time interval, and the corresponding Fourier spectra. In this way we have constructed a family of correlation functions C_n (τ,t_n) and a family of power spectra F_n (ω,t_n) which can be called dynamical correlations and Fourier surfaces, respectively. In Fig. 9A and B the Fourier surface is reported as a series of cross-sections and an isointensity map, respectively. The plots of Fig. 9 clearly show a non-stationary vibrational state, *i.e.* an aperiodic motion-sharing among vibrational modes during the time evolution of the end-to-end distance as a consequence of the non-linear dynamics of the system. This is antithetic to the stationary behavior expected by superposition of normal modes of a linear problem. Moreover, at least for the considered timescale, we did not observe any recurrence time,[43] probably due to the high complexity of the investigated system. The vibrational behavior emerged from the previous overall correlation and Fourier analysis can be considered as the resultant of the superposition of all the vibrational non-stationary states developed during the equilibrium trajectory.

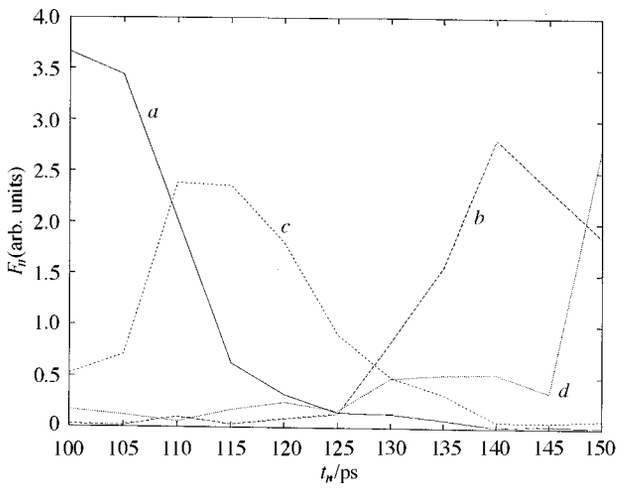

Fig. 10 Curves $F_n(t_n)$ corresponding to the frequency values giving the maximum intensity in Fig. 9: (*a*) $\omega = 1.2$, (*b*) $\omega = 6.72$, (*c*) $\omega = 3.36$ and (*d*) $\omega = 7.76$ ps^{-1}

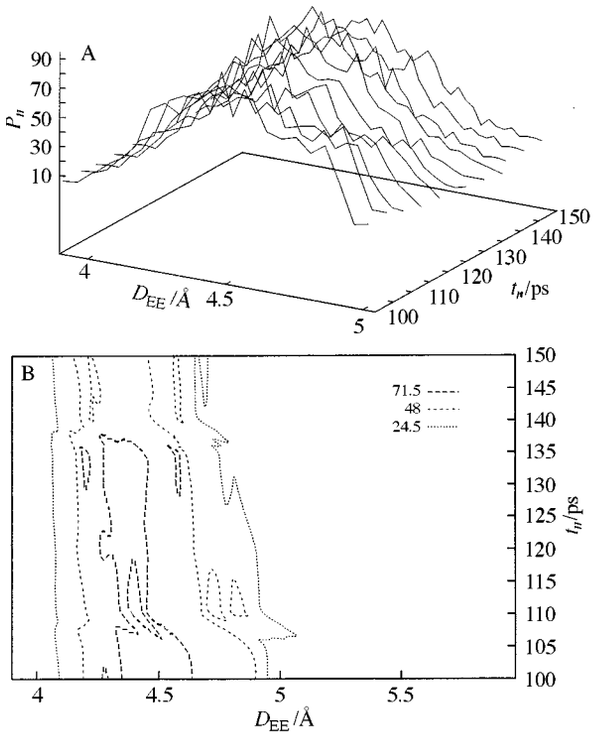

Fig. 11 Dynamical population surface $P_n(D_{EE}, t_n)$ of the end-to-end distance as (A) cross-sections and (B) an isodensity map during MD thermal equilibrium

During the time evolution a succession of largely prevalent vibrational modes is observed. The modes appear at increasing times, corresponding to $\omega = 1.2$, 6.72, 3.36 and 7.76 ps^{-1}. This time vibrational evolution is similar to the outcome of Fermi's problem[44] where the dynamics of a weakly non-linear oscillators lattice is interpreted in terms of solitons.[45] These are the analogues in non-linear systems of the normal modes of linear problems or the non-linear normal modes[46] and are found ubiquitously in many branches of science.[47] If the total energy is too low to be affected by the system's non-linearity, starting from a single vibration normal mode, neither the energy equipartition among all the molecular degrees of freedom nor an ergodic trajectory are observed, but instead an essentially quasiperiodic and non-ergodic behavior characterized by the presence of a soliton is observed. The activated soliton does not share energy with the others (in a similar way to the normal modes of a linear system) and is characterized by a recurrence time or period which grows in a polynomial way as a function of the number of interacting particles.[48]

The dynamics of a soliton are characterized by the energy transfer among few normal modes (the lowest frequency ones) in which the motion of the system from time to time is localized. The end-to-end distance dynamics shows just such a behavior, although, due to the system complexity and the time-scale used in our simulation, it is not possible to define any characteristic recurrence time.

Fig. 10 shows the curves $F_n(t_n)$, corresponding to the frequency values giving the maximum intensity of the most important observed vibrational modes. From the analysis of the TTP of Figs. 9 and 10 we observe that: (1) the modes evolve through a characteristic trend in the frequency–time plane. In particular, the initial motion is largely localized in the mode at 1.2 ps^{-1}. Then, this mode damps and motion and energy are shared and accumulated in the vibration at 6.72 ps^{-1}. After a short time interval during which no mode predominates (see the following), a vibration at $\omega = 3.36$ ps^{-1} is activated with half-frequency with respect to the preceding one. Then, the motion again is transferred to a non-correlated vibration at a higher frequency $\omega = 7.76$ ps^{-1}. This trend from low to high frequencies and *vice versa* shows a time drift toward the high frequencies. The turn-over of high and low frequency modes is observed also in Fermi's system. This characteristic motion sharing may be an expression of chaotic dynamics. A particular meaning could have the step with frequency halving: this could belong to a Feigenbaum's cascade of period-doublings.[49] Nevertheless from the inspection of the dynamical Fourier surface an increase in the number of vibrational modes is not observed as expected from the corresponding subharmonic cascade. (2) As previously mentioned, a time range exists around $t_n = 125$ ps where no prevalent vibrational mode is observed but a delocalization of the motion among the different degrees of freedom occurs, according to what is required at thermal equilibrium by the energy equipartition principle. We suggest that this may be a pattern of the universal coexistence of deeply different behaviors observed during the dynamical evolution of non-linear systems where order and chaos are closely joined. In this case quasiperiodic behavior as solitons and chaos as ergodic mixing motion restricted to around the examined energy minimum would be joined.

In order to characterize the TTP of the end-to-end distance from a statistical point of view the corresponding distribution histograms have been computed. As previously carried out with correlation functions and Fourier spectra, we have constructed a family of density distributions $P_n(D_{EE}, t_n)$ which can be called a dynamical population surface. In Fig. 11A and B this surface is shown as a series of cross-sections and as an isodensity map, respectively. It is apparent that during the time evolution of the end-to-end distance the statistical and vibrational evolution of the system show significant and complex behavior due to non-linear dynamics. The analysis shows that, apart from the various structures emerging in the plots, significant changes are observed only in the width of the distribution, due to the fluctuation of D_{EE}, while the mean value is practically time invariant.

Discussion

It is well known that the motion of a conservative system of multiple harmonic oscillators with small anharmonic perturbations is essentially non-ergodic and quasiperiodic: the motion in the phase-space is confined to a toroidal surface of smaller dimensions than that of the constant energy surface. The trajectory winds up the distorted torus densely and endlessly except for a negligible set of initial conditions which may lead to erratic motion on the whole energy surface (*i.e.* Arnold's diffusion).[50] The vibrational modes of the system are the solitons.

This dynamical behavior was discovered in the first molecular dynamics simulation performed by Fermi *et al.* in the early 1950s.[55] For such systems Kolmogorov stated and Arnol'd and Moser proved that the KAM theorem holds:[50] for a small non-linear perturbation the motion is essentially non-ergodic and quasiperiodic.

What happens when one violates the slightly non-integrable condition (small anharmonicity) of the KAM theorem? This appears to be a very difficult problem as Hamiltonian chaos (*i.e.* ergodicity and phase-space instability) may take place. KAM theory indicates the existence of an amplitude instability for conservative non-linear oscillators which allows a transition from essentially quasiperiodic motion to predominantly ergodic motion, when the extent of perturbation increases.[51] Zabusky and Deem[52] investigated this possibility in Fermi's problem and they demonstrated that for a large amplitude of motion the system exhibits widespread energy sharing among the vibrational modes. Nonetheless complete equipartition of energy was not achieved and the motion exhibited a correlation inconsistent with complete ergodicity at thermal equilibrium. It is unclear whether incomplete KAM instability or the constant high-order correlation of the type derived by Prigogine[53] was observed by Zabusky and Deem.[52]

Moreover, Saito and Hiraoka[54] found that in Fermi's problem the energy partition among modes takes place abruptly after an induction period starting from higher modes when the energy exceeds a threshold value. The Fermi's solitons seem to be due to the fact that the total energy, which depends on the initial mode, was not large with respect to the non-linearity of the system.

In molecular dynamics simulations the molecular chain is a system of multiple oscillators generally referred to as cartesian coordinates. The potential energy surface arises not only from simple harmonic terms (*e.g.* bond stretching, which is generally quadratic in both internal and cartesian coordinates), but also from terms that are anharmonic in cartesian coordinates (*e.g.* bond angle bending), even though they may be harmonic in internal coordinates, as well as terms that are anharmonic in both internal and cartesian coordinates (*e.g.* torsional and non-bonded strains).

For small amplitudes of conformational fluctuation around an equilibrium point at a sufficiently low temperature, the system behaves just like weakly non-linear coupled oscillators. Then, the observed essentially quasiperiodic, KAM motion of the tetrapeptide at room temperature is expected. Similarly, for large amplitude motions at higher temperature or presumably in solution,[55] we expect an ergodic and orbitally unstable chaotic motion to develop. It is interesting to note that the behaviors of our canonical dynamics are similar to micro-canonical dynamics, as expected from the statistical thermodynamic point of view.

The tetrapeptide exhibits essentially quasiperiodic dynamics of relatively low entropy at 300 K *in vacuo*. This is reminiscent of the mechanical brittle behavior of dry or poorly hydrated elastin which has been associated with a low entropy, enthalpically stable state.[16] Experimentally, strictly related peptides in aqueous solution show multiple conformational equilibria with substantial molecular flexibility.[13–15] This is a large internal rotational motion that, associated with the KAM amplitude instability, would ensure the motion in solution to be largely chaotic, mimicking the viscoelastic properties of the high-entropy, water-swelled protein.[12]

The KAM amplitude instability indicates that the extent of the non-linear perturbation determines to what extent the dynamical behavior will be chaotic and ergodic. Then, the essentially quasiperiodic and non-ergodic behavior of the system which satisfies the KAM theorem conditions and the Hamiltonian chaos (*e.g.* of a K-flow) for large amplitude motions represents opposite limit cases of a continuous range of intermediate behaviors. Similarly, the entropy (*e.g.* K-entropy that is essentially the sum of all positive Lyapunov exponents[50]), being a measure of the degree of motion instability of the system, will depend upon the extent of non-linear components.

The physico-chemical consequences of KAM theory and amplitude instability may be remarkable as far as macromolecular elasticity is concerned. In fact the picture we propose is that relaxed conformations are related to chaotic dynamical states, whereas the stretched conformations are connected to essentially quasiperiodic motions. Accordingly, the change in entropy on stretching, which is the driving force of the elasticity mechanism, is related to a variation of the K-entropy of the motion.

Let us briefly discuss the entropic mechanisms of elasticity of elastin. In Flory's theory of rubber elasticity[56] the relaxed state is a random chain network with random distribution of end-to-end chain lengths. This state corresponds to the maximum volume in the conformational space and then to the maximum entropy. When the system is deformed a decrease in conformational freedom occurs.

In our opinion, and as extensively proposed by one of us,[12] the entropic mechanism of elasticity of elastin is essentially of this kind. In the solvent entropy mechanism[57,58] the exposed hydrophobic side chains of elastin are surrounded on deformation by clathrate-like water of lower entropy than water around the relaxed structure. In this way the entropic elastomeric force is due to a decrease in the mobility of water molecules around the stretched protein. Lastly, in Urry *et al.*'s librational entropy mechanism of elasticity,[59] the relaxed polypeptide chain exhibits internal librations of large amplitude and low-frequency. The decrease in entropy on stretching is due to damping of internal motions of the chain giving torsional oscillations with a shift to higher frequencies. This mechanism is believed to occur within short peptide segments included within a regular structure named a β-spiral and constituted of recurring type II β-turns.

From a dynamical point of view, a common feature of the examined entropic mechanisms of elasticity is the maximum mobility state corresponding to the maximum entropy relaxed state. For the random coils of Flory's relaxed state, and also for the librational mechanism, the torsional skeletal vibrations can be treated as hindered rotators, with large amplitude rotational motions, displaying a largely non-linear dynamical behavior. In contrast, in the extended state internal chain motions decrease and small amplitude torsional oscillations take place. Similarly, in the solvent entropy mechanism, diffusive motions of bulk water in the relaxed maximum entropy state are substituted by vibrational motions of clathrate water in the extended state. Accordingly, it is possible to interpret by means of the dynamical systems theory the entropic mechanisms of elasticity in a global way and the entropy itself as a dynamical parameter.

A necessary condition for a macromolecular system to show elastomeric features is the highly chaotic dynamical behavior of the relaxed state due to the strong non-linearity of the protein and solvent large motions. The KAM amplitude instability is believed to ensure in such conditions the development of Hamiltonian chaos. Moreover, it is necessary that the molecular system, on deformation in the presence of external constraints, shows a dynamical essentially quasiperiodic behavior because of small amplitude vibrational motions developed in such conditions. This is not to say that in the extended state of the protein folded conformers are dominant, as are those assumed by our tetrapeptide *in vacuo*. Simply, whatever the conformation may be, it should exhibit quasiperiodic dynamics.

Preliminary results of MD simulations carried out on this peptide in dilute aqueous solution[60] show that all conformational motions increase because of the effect of the solvent on the dynamics of the formation/breaking of intramolecular H-bonds. Moreover, the time correlation functions, power spectra and Lyapunov exponents indicate that a chaotic

Brownian-like intramolecular motion occurs. In addition, we have verified that a similar conformational freedom also characterizes the constrained hydrated peptide, assumed to model a cross-linked chain in the relaxed state of the protein. In contrast, when the simulation is performed on a stretched constrained peptide, less freedom is observed and a very different dynamical behavior is obtained indicating a transition to low entropy, quasiperiodic soliton-like motion.[60]

Conclusions

This paper highlights the dynamic aspects of the molecular structure of the tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe, as a complementary approach to the static chemical description. We have analyzed a number of local molecular parameters (torsion angles, hydrogen bond distances) and a global variable (end-to-end distance) in order to gain insight into the conformational flexibility of the lowest energy conformer. Through analysis of the motion of the tetrapeptide we have observed that in contrast with the disordered behavior of local variables, the global variable shows a well structured pattern, characterized by a largely predominant low frequency motion. This suggests that conflicting micro-motions cancel one another and an ordered cooperative macro-motion results. Moreover, an analogous behavior was observed in a system of locally and globally coupled oscillators, where some macrovariables exhibit quasiperiodic or chaotic behavior, even if each element evolves independently.[61]

When the amplitude of the molecular motion is small (*e.g.* at sufficiently low temperature, or for a constrained state of stretching) the system can be considered to be a collection of weakly non-linearly coupled harmonic oscillators and the corresponding motion is essentially non-ergodic and quasi-periodic.[62,63] The vibrational motions of this system are the solitons, which localize the energy without equipartition: high frequency vibrational modes remain inaccessible and only low frequencies are significant.

This has been confirmed in a quite different molecular system;[64] poly(*p*-hydroxybenzoate), a semirigid liquid crystal polymer. The kinematic behavior of the latter, simulated by the same molecular dynamics approach, reveals a highly cooperative chain motion similar to that exhibited by our tetrapeptide. Although the two molecules are very different, their stiffness properties are analogous and originate in the first case from the covalent structure of the backbone chain, and in our more flexible tetrapeptide from the hydrogen bonded network which stabilizes the conformer.

The soliton concept has been used to investigate collective phenomena in biophysical systems as the proton transport in hydrated proteins or non-linear energy localization in DNA thermal denaturation and transcription.[65]

When the molecular motion amplitude becomes large (*e.g.* at higher temperature or for the flexible hydrated molecule) the KAM amplitude instability gives rise to a Hamiltonian chaos. The transition between these two kinds of motions should be related to a change in a thermodynamic property such as entropy.

Speculatively, the elastin macromolecule, of which our tetrapeptide is a recurring sequence, could exhibit an analogous behavior. As a matter of fact, one may suggest that the elastin domains responsible for elasticity could be characterized by similar non-linear conformational dynamics, especially because sequences of the general type -Gly-X-Gly-Gly- are frequent there. In explicit terms, one may envisage that an essential contribution to the entropy of the relaxed state could come from internal chaotic motion. On stretching, the motion would became quasiperiodic, soliton-like and therefore a decrease of entropy would occur. Overall, a re-interpretation of classical elasticity theory[56] in terms of non-linear dynamics of the polypeptide chain(s) is proposed.

Acknowledgements

This work was supported by MURST and LAMI funds.

References

1 *Molecular-Dynamics Simulation of Statistical-Mechanical Systems*, ed. G. Ciccotti and W. G. Hoover, North-Holland, Amsterdam, 1986.
2 H. G. Schuster, *Deterministic Chaos*, VCH, Weinheim, 1995.
3 J. C. Maxwell, *Philosophical Transactions of the Royal Society of London*, 1867, CLVII, 49.
4 J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.*, 1985, **57**, 617.
5 D. Ruelle, *Chaotic evolution and strange attractors*, Cambridge University Press, Cambridge, 1989.
6 D. Ruelle, *Hasard et Chaos*, Editions Odile Jacob, Paris, 1991.
7 F. A. L. Anet, *J. Am. Chem. Soc.*, 1990, **112**, 7172.
8 J. E. Straub, A. B. Rashkin and D. Thirumalai, *J. Am. Chem. Soc.*, 1994, **116**, 2049.
9 H. Zhou and L. Wang, *J. Phys. Chem.*, 1996, **100**, 8101.
10 J. M. Haile, *Molecular dynamics simulation*, Wiley, New York, 1992.
11 D. R. Cox and P. A. W. Lewis, *The statistical analysis of series and events*, Chapman and Hall, London, 1966.
12 A. M. Tamburro, in *Elastin: Chemical and Biological aspects*, ed. A. M. Tamburro and J. M. Davidson, Congedo, Italy, 1990, p. 126.
13 A. M. Tamburro, V. Guantieri, L. Pandolfo and A. Scopa, *Biopolymers*, 1990, **29**, 855.
14 M. A. Castiglione Morelli, A. Scopa, A. M. Tamburro and V. Guantieri, *Int. J. Biol. Macromol.*, 1990, **12**, 363.
15 M. A. Castiglione Morelli, M. De Biase, A. De Stradis and A. M. Tamburro, *J. Biomol. Struct. Dyn.*, 1993, **11**, 181.
16 V. Villani and A. M. Tamburro, *J. Chem. Soc., Perkin Trans. 2*, 1993, 1951.
17 V. Villani and A. M. Tamburro, *J. Mol. Struct.* (*THEOCHEM*), 1994, **308**, 141.
18 V. Villani and A. M. Tamburro, *J. Biomol. Struct. Dyn.*, 1995, **12**, 1173.
19 V. Villani, *J. Mol. Struct.* (*THEOCHEM*), 1996, **363**, 43.
20 D. A. Pearlman, D. A. Case, J. C. Caldwell, G. L. Seibel, U. Chandra Singh, P. Weiner and P. A. Kollman, AMBER 4.0, University of California, San Francisco, 1991.
21 S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comput. Chem.*, 1986, **7**, 230.
22 J. A. McCammon, B. R. Gelin and M. Karplus, *Nature*, 1977, **267**, 585.
23 M. Levitt, *Nature*, 1981, **294**, 379.
24 W. F. van Gunsteren, H. J. C. Berendsen, J. Hermans, W. G. J. Hol and J. P. M. Postma, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 4315.
25 J. A. McCammon and S. C. Harvey, *Dynamics of protein and nucleic acids*, Cambridge University Press, Cambridge, 1988.
26 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
27 L. Verlet, *Phys. Rev.*, 1967, **159**, 98.
28 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327.
29 Lord Rayleigh, F.R.S., *Philos. Mag.*, 1883, **15**, 229.
30 W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 3.
31 J. M. T. Thompson and H. B. Stewart, *Nonlinear Dynamics and Chaos*, Wiley, Chichester, 1987.
32 J. K. Hale, *Oscillations in nonlinear systems*, Dover, New York, 1992.
33 P. J. Rossky and M. Karplus, *J. Am. Chem. Soc.*, 1979, **101**, 1913.
34 N. Packard, J. P. Crutchfield, J. D. Farmer and R. S. Shaw, *Phys. Rev. Lett.*, 1980, **45**, 712.
35 L. Colace and A. R. Giona, in *Chaos and fractals in chemical engineering*, ed. G. Biardi, M. Giona and A. R. Giona, World Scientific, Singapore, 1995.
36 R. M. Levy, M. Karplus and J. A. McCammon, *Chem. Phys. Lett.*, 1979, **65**, 4.
37 F. Lelj, P. Grimaldi and P. L. Cristinziano, *Biopolymers*, 1991, **31**, 6638.
38 H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, 1965.
39 F. C. Moon, *Chaotic and Fractal Dynamics*, Wiley, New York, 1992.
40 H. A. Davies and F. C. Moon, *Chaos*, 1993, **3**, 93.
41 J. Wright, *Chaos*, 1995, **5**, 356.
42 M. Schwartz and L. Shaw, *Signal Processing*, McGraw-Hill, New York, 1975.
43 M. Toda, *Theory of nonlinear lattices*, Springer-Verlag, Berlin, 1989.
44 E. Fermi, J. Pasta and S. Ulam, in *Los Alamos Rpt. LA-1940* (*1955*) *Collected Papers of Enrico Fermi*, University of Chicago Press, Chicago, 1965, vol. 2, p. 968.

45 N. J. Zabusky and M. D. Kruskal, *Phys. Rev. Lett.*, 1965, **15**, 240.
46 J. Waters and J. Ford, *J. Math. Phys.*, 1966, **7**, 399.
47 Y. R. Shen, *Science*, 1997, **276**, 1520; A. W. Snyder and D. J. Mitchell, *ibid*., 1997, **276**, 1538.
48 N. J. Zabusky and G. S. Deem, *J. Comput. Phys.*, 1968, **2**, 207.
49 M. J. Feigenbaum, *J. Stat. Phys.*, 1979, **21**, 669.
50 H. Bai-Lin, *Chaos*, World Scientific, Singapore, 1984.
51 G. H. Walker and J. Ford, *Phys. Rev.*, 1969, **188**, 416.
52 N. J. Zabusky and G. J. Deem, *J. Comput. Phys.*, 1967, **2**, 126.
53 I. Prigogine, *Non-equilibrium Statistical Mechanics*, Wiley-Interscience, New York, 1962.
54 N. Saito and H. Hiraoka, *J. Phys. Soc. Jpn.*, 1967, **23**, 167.
55 Z. R. Wasserman and F. R. Salemme, *Biopolymers*, 1990, **29**, 1613.
56 P. J. Flory, *Principles of Polymer Chemistry*, Cornell, Ithaca, NY, 1953.
57 W. R. Gray, L. B. Sandberg and J. A. Foster, *Nature*, 1973, **246**, 461.
58 J. M. Gosline, *Biopolymers*, 1978, **17**, 677.
59 D. W. Urry, C. V. Venkatachalam, M. M. Long and K. U. Prasad, in *Conformation in Biology, the Festschirft celebrating the sixtieth birthday of G. N. Ramachandran F.R.S.*, ed. R. Srinivasan and R. H. Sarna, Adenine Press, Guilderland, NY, 1982, p. 11.
60 V. Villani, L. D'Alessio and A. M. Tamburro, in *Elastin and elastic tissue*, ed. A. M. Tamburro, Procedings of an International Conference, Maratea (Italy), 17–20 October 1996, p. 31.
61 N. Nakagawa and Y. Kuramoto, *Physica D*, 1995, **80**, 307.
62 E. Thiele and D. J. Wilson, *J. Chem. Phys.*, 1961, **35**, 1256.
63 D. Bunker, *J. Chem. Phys.*, 1962, **37**, 393.
64 R. Fusco, L. Longo, L. Caccianotti, C. Aratari and G. Allegra, *Makromol. Chem., Theory Simul.*, 1993, **2**, 685.
65 *Nonlinear Excitations in Biomolecules*, ed. M. Peyrard, Les Editions de Physique, Springer Verlag, Paris, 1995.

Paper 7/02217D
Received 2*nd April* 1997
Accepted 8*th July* 1997