Conformational chaos and biomolecular instability in aqueous solution

Vincenzo Villani,*[a] Antonio M. Tamburro[a] and Jose' M. Zaldivar Comenges[b]

[a] Universita' della Basilicata Dipartimento di Chimica, Via N. Sauro, 85, 85100 Potenza, Italy
[b] European Commission, Joint Research Center Institute for Systems, Informatics and Safety Systems Modelling and Assessment Unit, TP 250, 21020 Ispra (VA), Italy

The glycine-rich tropoelastin tetrapeptide Ac-Gly-Leu-Gly-Gly-NMe has been modeled in aqueous solution by means of molecular dynamics simulations and the conformational motions have been characterized using nonlinear dynamics theory. Large amplitude fluctuations of the peptide backbone and H-bond patterns are detected. The end-to-end vector R_{ee} undergoes anomalous diffusion with antipersistent fractional Brownian motion according to chaotic motions of molecules on fractal media. The vibrational picture of the intramolecular vectors shows a spatiotemporal self-similar disorder along the peptide chain on large scale observation demonstrating a high entropy state. The conformational chaos of the peptide is a consequence of the nonlinear effects of the attractive interactions between residues developed in aqueous solution due to water being a poor solvent. The viscous drag is high-lighted and is thought to be due to the percolation network of disordered H-bonded water molecules. The method of the reconstruction of the phase-space using the embedding theorem is presented and the invariant properties of the peptide are calculated. The existence of a low dimensional chaotic attractor according to dissipative systems has been demonstrated. The dynamical high entropy state of the peptide in solution is in agreement with the proposed mechanism of the transition-to-chaos for the elastin elasticity.

Introduction

What exactly is the dynamics of peptides in dilute aqueous solution? Of course, a simple answer is not possible: it will depend upon the *particular* primary structure and on the experimental conditions. Nevertheless, *universal* scale behaviors due to the nonlinear forces are to be expected.[1]

We are interested in aqueous solutions of essentially flexible peptides and we expect to observe complex dynamics. In this paper we describe a study of the tetrapeptide Ac-Gly-Leu-Gly-Gly-NMe, a typical sequence recurring in the tropoelastin chain. The glycine-rich sequences characterize the elastic performance of elastin,[2] owing to the high concentration of this residue (up to 33%). In fact, due to the absence of a side chain group, the glycyl residue is extremely flexible either from a *static* or from a *dynamic* point of view in that the conformational minima of the backbone are equally probable and the interconversion frequency is higher than those corresponding to the global molecular motions of the chain.[1]

In a series of previous theoretical papers[3–7] some of us have determined the relative stability of conformers and the dynamical behavior of that sequence, either as an isolated molecule or in aqueous solution, from the available experimental NMR and CD data.[8]

A number of accessible conformational states and a large chain-mobility characterize the peptide behavior. Nevertheless, a typical dynamical pattern of hydrogen bonds, characterized by the type II β-turn [Gly$_1$]C=O $\cdots$ HN[Gly$_4$], is observed.

An amazing nonlinear dynamic behavior with conformational *solitons* has been evidenced in the molecule *in vacuo* and a *chaotic* one for the molecule in aqueous solution. In the Villani–D'Alessio–Tamburro model, this has been related to the entropic mechanism of the elastin elasticity as a *chaos–soliton transition* from the relaxed to the stretched form in an aqueous medium.[5,6]

In this work, by means of molecular dynamics (MD) simulations in aqueous solution and original analysis of data, we have attempted to further define the conformational picture of the dissolved peptide in terms of diffusive behavior focusing on intramolecular structural parameters.

The advent of nonlinear time series analysis and the mathematical theorems associated with chaotic dynamics in the late eighties now make it possible, not only to qualify, but also to quantify the behavior of complex systems. The techniques, which allow complex dynamics to be represented in multi-dimensional phase space as a geometrical object, have had success in predicting chaotic behavior.[9]

The dynamic behavior of the system has been carefully defined in terms of mean square displacement, Hurst exponent, fractal and attractor dimensions and Lyapunov exponents. The dynamic Fourier surface method introduced in ref. 5 has been used successfully. We have tried to develop further this nonlinear time series analysis using delay coordinate embedding by applying recently developed numerical algorithms. In this way, we have investigated the complexity and instability of nonlinear conformational dynamics, showing the *spatiotemporal chaos* of the molecule which could be the basis of the role played by similar sequences in elastomeric proteins such as elastin.

Model and methods

AMBER 4.1 software[10] was used on a DEC Alpha Station 2100 under the UNIX 5.0 operating system. The data analysis was performed using homemade FORTRAN programs and the final plots using GNUPLOT[11] under WINDOWS NT on a Pentium II PC.

The molecular potential energy has been computed for the peptide using the Wiener *et al.* united-atoms force field model[12] (where carbon atoms with implicit inclusion of bonded hydrogen are used) and for water molecules the Jorgensen *et al.* TIP3P model[13] (which include only Lennard–Jones and electrostatic terms). This choice was suggested by the requirement of

DOI: 10.1039/b001105n

J. Chem. Soc., Perkin Trans. 2, 2000, 2177–2184 **2177**

optimizing the computer efforts and was encouraged both by a large amount of literature on aqueous solution simulations and on a heuristic basis by our previous MD simulated annealing of the same peptide in aqueous solution which successfully interpreted the most important experimental results.[7]

The MD simulation at constant temperature and pressure by means of the method of van Gunsteren and co-workers[14] was accomplished by coupling the system to external heat- and pressure-baths at reference values $T_o = 300$ K and $P_o = 1$ atm by the time constants $\tau_T = \tau_P = 0.2$ ps. These dynamics are not Hamiltonian (*i.e.* not at constant total energy) and the molecular model can be reduced to an ensemble of dissipative nonlinear Rayleigh oscillators.[5] Periodic boundary conditions were applied and the cutoff distance criterion of 9 Å for the nonbonded interactions were used.

The equations of motion were integrated in Cartesian coordinates *via* the Verlet *leap-frog* algorithm[15] with holonomic constraints of the bond distances at equilibrium values, using the SHAKE routine.[16] This results in rigid TIP3P water molecules being characterised by three bonds rather than two bonds and an angle.

A solution box of volume $32.2337 \times 30.4357 \times 27.3159$ Å³ and density 0.9733 g cm^{-3} with one tetrapeptide Ac-Gly-Leu-Gly-Gly-NMe molecule solvated by 852 waters is the starting point of our simulation. This is the state obtained by MD simulated annealing reported in ref. 7.

In our MD simulations the integration time step δt was 2 fs, and energies and coordinates were stored every $\Delta t = 0.08$ ps. The time period of 1 ns was simulated, producing time series of 12500 points, and required about 85 CPU hours. The analysis is focused on the tetrapeptide dynamics and the trajectory of the end-to-end vector $\boldsymbol{R}_{ee}(t)$, a fundamental parameter characterizing macromolecule behavior,[1] has been considered. The vector $\boldsymbol{R}_{ee}$ is defined by means of the end carbon atoms of acetyl and *N*-methyl-amide groups.

Mean square displacement and Hurst exponent

The mean square displacement $\langle \boldsymbol{R}^2(\tau) \rangle$ is defined as the *time-dependent difference correlation function* [eqn. (1)], at correlation times which are delayed by an amount τ. The time average denotes the averaging over different *time origins t_0*. In this way intramolecular diffusive processes were monitored.

$$\langle \boldsymbol{R}^2(\tau) \rangle = \langle (\boldsymbol{R}(t_0) - \boldsymbol{R}(t_0 + \tau))^2 \rangle_{t_0} \qquad (1)$$

The *scaling law* of the mean square displacement of the diffusing variable as a function of time is given by eqn. (2)

$$\langle \boldsymbol{R}^2(\tau) \rangle \sim \tau^{2H} \qquad (2)$$

according to the diffusion law, where H is the critical *exponent of Hurst*.[18] We obtained the diffusion exponent H from the slope of the corresponding bilogarithm scale plot [eqn. (3)].

$$\ln \langle \boldsymbol{R}^2(\tau) \rangle \sim 2H \ln \tau \qquad (3)$$

When $H = 1/2$ ordinary Brownian motion occurs and Einstein–Fick's law is followed, while if $H \neq 1/2$ *fractionary Brownian motion (fBm)* takes place.[19] Both cases of enhanced and reduced diffusion rate are possible. For a random walk on a fractal object in the presence of fixed obstacles, the *antipersistent fBm* with $H < 0.5$ occurs. In fact, the irregularities existing at any lengh scales are responsible for the diffusion lag.[20]

Fractal dimension

How long is a dynamic path? Similar to the classical Mandelbrot[19] results, the length of a trajectory in its space as a function of the time resolution step P was computed to give the fractal dimension D of the path, which is a measure of its *jaggedness*.

The approach is in line with Perrin's historical observations on the Brownian motion of a fluctuating particle.[21] He described a *quite irregular* motion observing convoluted trajectories at every timescale: every straight segment at a certain scale is substituted by a longer than ever polygonal contour upon increasing the time resolution. In this way the lengths of the trajectories $\boldsymbol{R}_{ee}(t)$ have been measured (using Euclidean metrics) and eqn. (4) is derived from the bilogarithm

$$\ln (L(n)) \sim d \ln (n) \qquad (4)$$

plot (where $n = P/\Delta t$ is the resolution factor per measurement interval). The corresponding *fractal dimension $D = 1 - d$* was evaluated.

Dynamic Fourier spectra

This method, introduced for the first time in the study of the Boc-Gly-Leu-Gly-Gly-NMe tetrapeptide *in vacuo*, is a very powerful tool to characterize the nonlinear dynamic behavior.[5] In this way, the non-stationary vibrational picture was apparent and the characteristic excitation-sharing among few modes, a *fingerprint* of non-ergodic soliton motions, was observed.

Briefly, a family of delayed and bounded time series called traveling trajectory pockets (TTP) is generated. In particular, TTPs of 500 ps length starting at $t_n = nP$ (n is an integer number and $P = 50$ ps) from the initial point of the MD simulation, have been considered. This family of TTP was used to calculate the corresponding Fourier spectra. Accordingly, a family of power spectra $I_n(\omega, t_n)$ called *dynamic Fourier spectrum* (DFS) was generated and is represented as cross-section curves.

Reconstructing phase-space

Phase-space reconstruction is the first step in nonlinear time series analysis. Let us consider a system of d ordinary differential equations [eqn. (5)], where $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_d(t)]$ in R^d and $\boldsymbol{F} = [F_1, F_2, \ldots, F_d]$.

$$\mathrm{d}\boldsymbol{x}(t)/\mathrm{d}t = \boldsymbol{F}(\boldsymbol{x}(t)) \qquad (5)$$

A time series is a list of numbers which are assumed to be measurements of a quantity observable over time, which, in the absence of noise, is related to the dynamic system by eqn. (6).

$$s(t) = h(\boldsymbol{x}(t)) \qquad (6)$$

The system on which the observable is being measured is evolving with time. The phase-space reconstruction problem is that of recreating states when the only information available is contained in a time series, *i.e. how to go from scalar or univariate observations to the multivariable phase space which is required to study the system?* Typically, $\boldsymbol{F}$ and h are both unknown, so we cannot hope to reconstruct states in their original form. However, we may be able to reconstruct a state space that is equivalent to the original in the sense that differential properties are preserved. The work by Takens[22] has shown that if the dynamics is on a d-dimensional Euclidean space, an embedding of the system can be obtained with a $2d + 1$-dimensional reconstructed state space using derivatives or delay coordinates. The basic idea of this reconstruction is that if one has observed an orbit projected onto a single axis $s(t)$, then the orbit, which is presumed to come from an autonomous set of differential equations, may, by virtue of the projection, overlap with itself in the variables $s(t)$. There is no overlap of the orbit with itself in the true set of state variables by the uniqueness theorems about the solution of autonomous differential equations. If we can unfold the orbit by providing independent coordinates for a multidimensional space made out of the observations, then we

can undo the overlaps coming from the projection and recover orbits which are not ambiguous.

In this work delay coordinates have been used. Delay coordinates, $\{s(t), s(t - T), s(t - 2T), \ldots, s(n + (d_E - 1)T\}$ are easy to work with and can be effective for very high dimensional cases where it may not be practical to calculate the required number of derivatives. Most of the research on the state space reconstruction problem has focused the problems of choosing the *time delay T* and the *embedding dimension* d_E for delay coordinates.

Finding the time delay

The first step in phase-space reconstructon is to choose an optimum delay parameter T. Different prescriptions have appeared in the literature for choosing T but they are all empirical in nature. The most useful technique was suggested by Fraser and Swinney.[23] They propose using the first minimum of the *Average Mutual Information function $I(T)$*, as a kind of nonlinear correlation function to determine when the values of $s(n)$ and $s(n + T)$ are independent enough of each other to be useful as coordinates in a time delay vector but not so independent as to have no connection with each other at all.

Choosing the embedding dimension

The time delay reconstruction of the system phase space provides a necessary number of coordinates to unfold the attractor called the *embedding dimension*,[24] d_E. This is a global dimension to unfold the dynamics which may be different from the real dimension.

Furthermore, this dimension depends on the time series measurement, and, hence, if we measure two different quantities from some system, there is no guarantee that the d_E from time delay reconstruction will be the same for each of them.

The usual method for choosing the minimum embedding dimension is to compute some invariant of the attractor. By increasing the embedding dimension used for the computations, one notes when the value of the invariant stops changing. Since these invariants are geometric properties of the attractor, they become independent of d for $d \geq d_E$, *i.e.* after the geometry is unfolded.

In this work we have used the Cao method[25] which is based on the idea of *False Nearest Neighbours* (FNN) developed by Kennel *et al.*[26] In this case, the condition of no self-intersection states that if the attractor is to be reconstructed successfully in R^d, then all the neighbour points in R^d should be also neighbours in R^{d+1}, where d is the embedding dimension. This method also provides a way to distinguish between deterministic and stochastic signals by plotting two functions $E1$ and $E2$ (see Cao[25] for a precise definition). When both quantities reach saturation we have found the embedding dimension. In case of noise $E1$ will never reach saturation and $E2$ will always remain 1 for any dimension.

Determining the dynamic dimension

Once one has determined the global number of dimensions required to unfold the attractor, there remains the problem of the number of dynamic degrees of freedom d_L, which are active in determining the evolution of the system as it moves around the attractor. To calculate this dynamic dimension we have used the method proposed by Kennel *et al.*[26] in which the percentage of local false nearest neighbours is evaluated. Using the same idea as the method of FNNs, they proposed a method to study the local structure of the phase-space to see if locally one requires fewer dimensions than d_E to capture the evolution of the orbits as they move on the attractor. Their approach was to work in a dimension, $d \geq d_E$, large enough to assure that the attractor has been unfolded. In this space, they studied for some
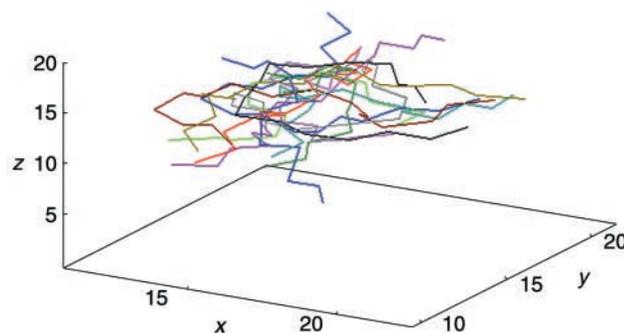

Fig. 1 Main chain conformation *snapshots* of the Ac-Gly-Leu-Gly-Gly-NMe peptide during 1 ns MD simulation in aqueous solution. A time sampling of 80 ps was used.

data point $y(k) = \{s(t), s(t - T), s(t - 2T), \ldots, s(n + (d - 1)T\}$, what subspace of dimension d_L one requires to make accurate local neighbourhood to neighbourhood maps of the data on the attractor. In fact, for a specified *number of neighbours N_B* of $y(k)$, they provided a local rule for calculating how these points evolve in one time step into the same N_B points near $y(k + 1)$. When the percentage of wrong predictions becomes independent of d_L and is also insensitive to the number of neigbours N_B, it is possible to say that the correct local dimension for the active degrees of freedom has been identified.

Global Lyapunov exponents

In the case of nonlinear systems one major feature that has emerged as a classifier are the *Lyapunov exponents*: they quantify how orbits on the attractor move apart (or together) under the evolution of dynamics.

Given a dynamic system in a d_L-dimensional phase-space it is possible to monitor the evolution of an infinitesimal d_L-sphere of initial conditions. This d-sphere will become a d_L-ellipsoid due to the locally deforming nature of the flow. The j-th one-dimensional Lyapunov exponent, λ_j, is then defined in terms of the length of the ellipsoidal principal axes $p_j(t)$ at time t (see Wolf *et al.*[27]) as eqn. (7).

$$\lambda_j = \lim_{t \to \infty} 1/t \log_2 p_j(t)/p_j(0), j = 1, \ldots, d_L \qquad (7)$$

The Lyapunov exponent monitors the behavior of two closely neighbouring points in a direction of the phase space as a function of time. If the points expand away from each other, the Lyapunov exponent will be positive, if they converge, the exponent becomes negative, if the two points stay the same distance apart, the exponent stays near zero. If base 2 is used, the exponents are measured in bits of information per time unit.

Results and discussion

Intramolecular trajectories

In order to clarify the motion of the peptide in solution, the conformations of the molecular backbone as a piecewise-line obtained by joining the bonded atoms are sketched in Fig. 1. The *snapshots* are sampled every 80 ps during the whole simulation and overlapped in tridimensional projection. The high flexibility of the peptide is apparent and the plot shows that the mobility is essentially intramolecular.

We analyze the observed conformational dynamics in terms of motions that the end-to-end vector undergoes. In Fig. 2 the tridimensional scatter-plot of vectorial trajectory $R_{ee}(t)$ is reported. $R_{ee}(t)$ appears anisotropic and inhomogeneous, with maxima of state density distributed disorderly within the self-similar point cloud. These fractal features suggest a chaotic intramolecular dynamics like the Brownian motions of molecules in disordered media.
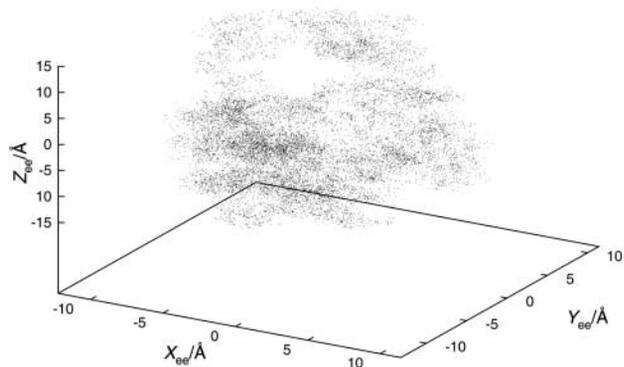
Fig. 2 Three-dimensional scatter-plot of the end-to-end trajectory $R_{ee}(t)$ during the MD simulation. A time sampling of 0.08 ps was used and the vectorial components are expressed in Å.
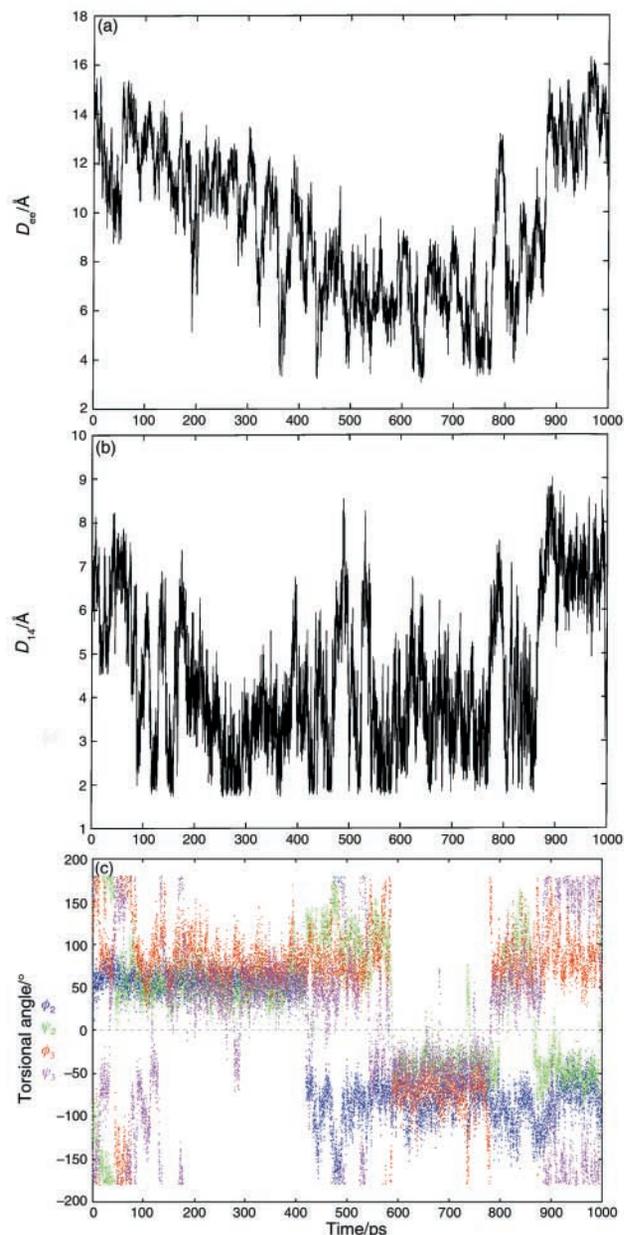

Fig. 3 Trajectories of (a) the end-to-end distance $D_{ee}(t)$ and of (b) the $[Gly_1]C=O\cdots HN[Gly_4]$ interaction distance $D_{14}(t)$ are plotted. Å and ps units are used. (c) Trajectories of φ_2, ψ_2, φ_3 and ψ_3 torsional angles are plotted. Torsions are in degrees and time in ps.

The end-to-end and $[Gly_1]C=O\cdots HN[Gly_4]$ distances D_{ee} and D_{14}, respectively, were considered. From the analysis of the $D_{ee}(t)$ and $D_{14}(t)$ trajectories of Fig. 3a and b it is possible to
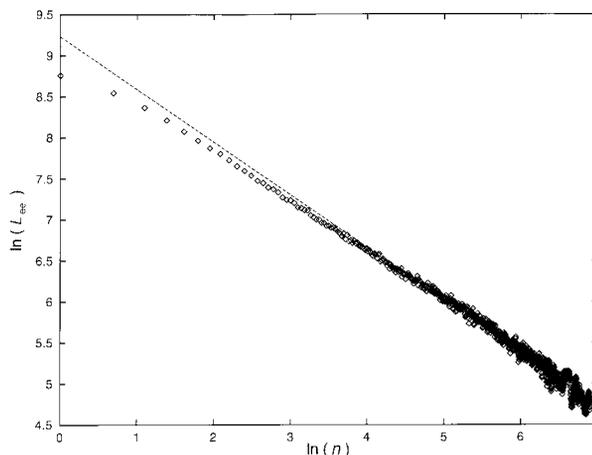
Fig. 4 Bilogarithm plot of the measured length for the trajectory $R_{ee}(t)$ as a function of the resolution time expressed as number of points per observation interval. The trajectory length is expressed in Å.

obtain information about the conformational states assumed by the peptide in aqueous solution and the transition patterns followed from one to another. The $D_{ee}(t)$ and $D_{14}(t)$ trajectories appear strictly correlated: the peptide is initially in an *unfolded* state, then changes to a *folded* one through a gradual transition characterized by very large fluctuations which involve the starting and final states. The large fluctuations are in agreement with the behavior of the systems at a phase transition[28] and confirm that the conformational transitions belong to the critical phenomena as described in ref. 4. Lastly, through a sudden *folded*-to-*unfolded* transition the final observed state is reached. Our dynamics seem to suggest *symmetry breaking* in the folding–unfolding process: the first being slow while the second was fast. This *self-organized criticality*[29] is in agreement with the cooperative helix–coil transition observed and modeled for proteins, *e.g.* in the denaturation of dispersed collagen in a dilute solution[30] or in recent simulations of water–plastocyanin systems.[31]

Lastly, from comparison of the torsion φ_2, ψ_2 and φ_3, ψ_3 trajectories reported in Fig. 3c, which identify the β-turn type $[Gly_1]C=O\cdots HN[Gly_4]$,[32] one can observe that the type II β-turn occurs in the time ranges of about 400 to 600 and 800 to 900 ps, according to both experimental and theoretical observations.[3–8]

Scaling law and fractal dimension

The $R_{ee}(t)$ trajectory representative of the intramolecular motions of the tetrapeptide in solution has been characterized quantitatively determining the critical exponent of the time scaling law and the corresponding fractal dimension.

Fig. 4 shows the bilogarithm scale plot of the length $L(n)$ of $R_{ee}(t)$ as a function of the resolution factor n. Then, the asymptotic straight trend has been fitted by the least-square regression line, whose slope d represents the critical exponent of the scaling law, from which the corresponding fractal dimension $D = 1 - d$ is obtained. The computed critical exponent, increasing the trajectory length, is $d = -0.64$. We have observed that the fractal dimension $D = 1.64$ is larger than that corresponding to ideal Brownian walks ($D = 1.5$) and is typical of *antipersistent fBm* related to critical self-organized phenomena.[29] This motion has a lower correlation time than the Brownian one, with greater entropy, and is a *fingerprint* of chaotic dynamics.

These observations confirm the previous hypotheses[5,6] of the change in the dynamic picture of solution with respect to the *in vacuo* molecule characterized by soliton vibrations.

Mean squares displacement and Hurst exponent

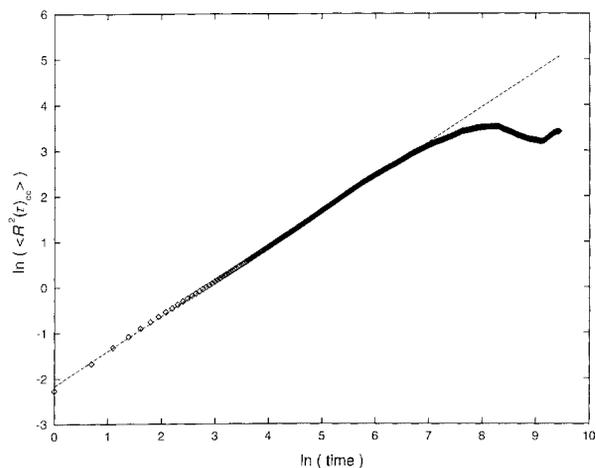Fig. 5 shows the bilogarithm plot of the time-dependent mean

Fig. 5 Bilogarithm plot of the mean square displacement for the end-to-end vector, $\langle \mathbf{R}^2(\tau)_{ee}\rangle$ *versus* the autocorrelation time τ. The square displacement is expressed in Å^2 and the time unit in 0.08 ps.

squares displacement $\langle \mathbf{R}^2(\tau)_{ee}\rangle$ *versus* the correlation time τ for the end-to-end vector $\mathbf{R}_{ee}$. In the meaningful time range, not affected by the finite length of the time series, the function has a linear trend and it is possible to evaluate the Hurst exponent $H = 0.38$. The observed value is lower than that corresponding to ideal Brownian walks ($H = 0.5$) pointing to a reduced diffusion rate and is consistent with the value of the observed fractal dimension of $\mathbf{R}(t)$.

The anomalous diffusion is in agreement with the *antipersistent fBm* of molecules in fractal media. From this point of view, we hypothesize that the diffusion lag is due to the viscous drag of the solution H-bond network. In other words, the water H-bonded molecules constitute a percolation cluster of obstacles against the conformational Brownian motion of the solute, and similarly to the findings obtained by means of simple lattice models, deviation from the ideal diffusion is observed.

We note that a correlation does exist between the scaling laws of the chain size as a function of the chain length N or the time period τ. The former is characterized by the *Flory critical exponent*[1] ν [eqn. (8)], the latter by the Hurst exponent. From

$$\mathbf{R}_{ee}(N) \sim N^{2\nu} \qquad (8)$$

this viewpoint, the exponents are similar and can be interpreted in the same way. Then, the observed H value is in agreement with $\nu = \frac{1}{3}$ as expected in *poor solvents* where the solute–solvent interactions are unfavored with respect to the intramolecular ones within the solute.

In conclusion, the dynamics of the elastin-related tetrapeptide in aqueous solution could be modeled by means of *antipersistent fBm* of chains on a fractal percolation medium of H-bonded water molecules.

DFS analysis

In Fig. 6 the DF spectrum of the end-to-end vector components is shown. Similar spectra also characterize the other *intramolecular vectors*,[1] as we have verified by taking into account a number of *ij* atom pairs of the backbone. In all cases the DF spectra testify a non-stationary vibrational motion localized in the low-frequency range (up to about 0.1 ps^{-1}). This behavior is typical of *fBm*, as we have verified. In these cases, in addition to the characteristic intensity damping toward the high-frequencies, non-stationary behaviors with time intensity fluctuations of excited modes are observed. This testifies to the variable and disordered spectral composition of the trajectory as a function of time.

Fig. 7a and b shows the DF spectra of the end-to-end and D_{01} (between C_{α_0} and C_{α_1} atoms) distances, respectively. The
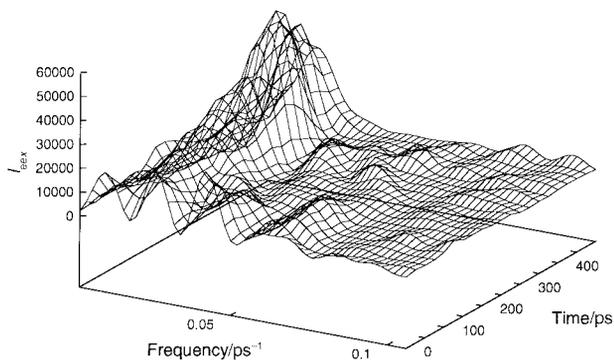

Fig. 6 DF spectrum of the *x*-component for the end-to-end trajectory $\mathbf{R}_{ee}(t)$ is reported. Time is in ps, frequency in ps^{-1} and TTPs of 500 ps have been used.
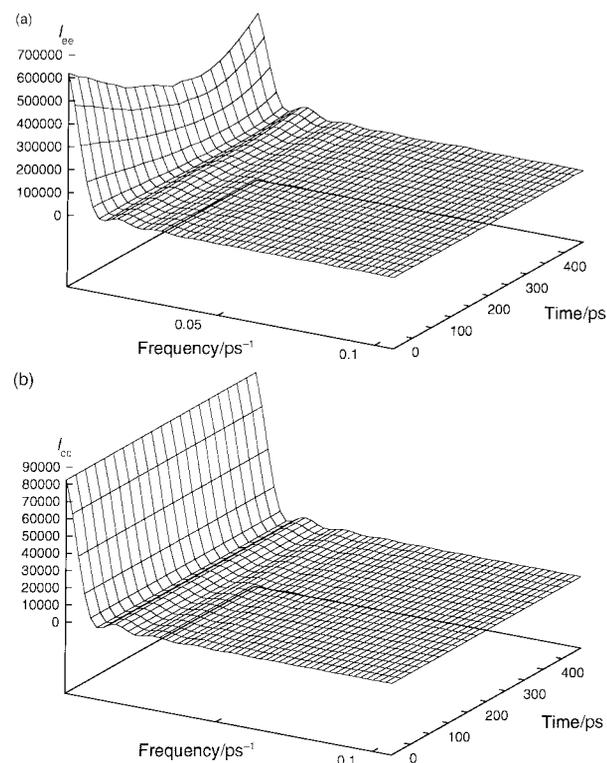

Fig. 7 DF spectra of (a) the end-to-end distance, $I_{ee}(\omega,t)$ and of (b) D_{01} distance, $I_{CC}(\omega,t)$ are reported. Time is in ps, frequency in ps^{-1} and TTPs of 500 ps have been used.

spectra shown by any other peptide intramolecular distance are similar. In both cases time domains of 500 or 100 ps have been considered varying the length of the TTPs. In this way, it has been possible to characterize the space-time scale behavior of the intramolecular dynamics. At the considered scale lengths we observe self-similar Brownian motions, fingerprint of chaos dynamics. The observed scale invariance testifies that the peptide in solution exhibits time disordered motions along the peptide chain similarly to turbulent fluids. In contrast, the corresponding spectrum for the *in vacuo* peptide showed the typical soliton mode-sharing of essentially quasiperiodic motions. An analogous hypothesis on the vibrational state of stretched elastin has been proposed and is supported by the dynamics of a partially extended single DNA macromolecule[33] that can be described by linearly independent normal modes.

Attractor dimensions and Lyapunov exponents

To analyze the chaos quantitatively the Lyapunov exponents of the modulus and of the coordinate components for the end-to-end trajectory $\mathbf{R}_{ee}(t)$ were calculated. The time series data of
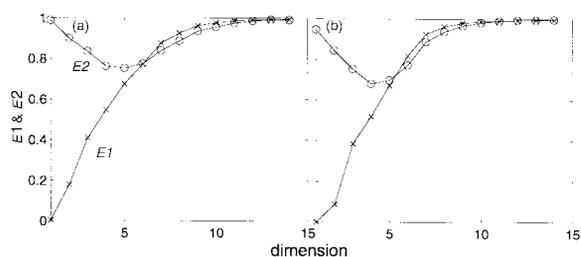
Fig. 8 Values of $E1$ and $E2$ functions (see text) for (a) $D_{ee}(t)$ and (b) x-coordinate of the end-to-end vector.
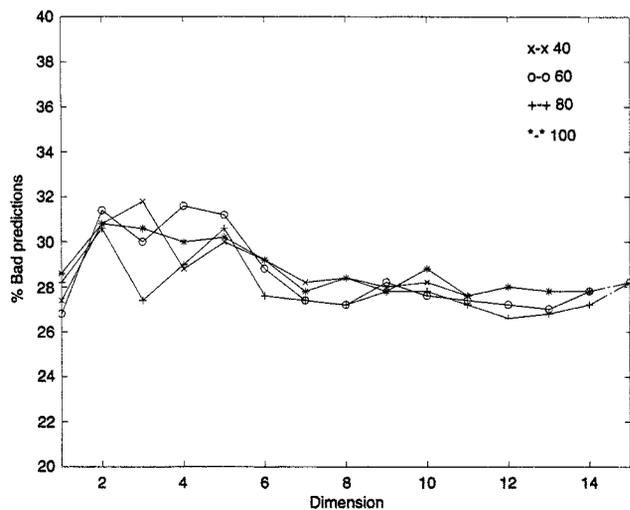

Fig. 9 Percentage of local false nearest neigbours as a function of the embedding dimension for the end-to-end trajectory $D_{ee}(t)$. $N_B = 40$, 60, 80 and 100 neigbours are considered. From this view $d_L \sim 9$ might be chosen. Note $d_E \sim 11$.
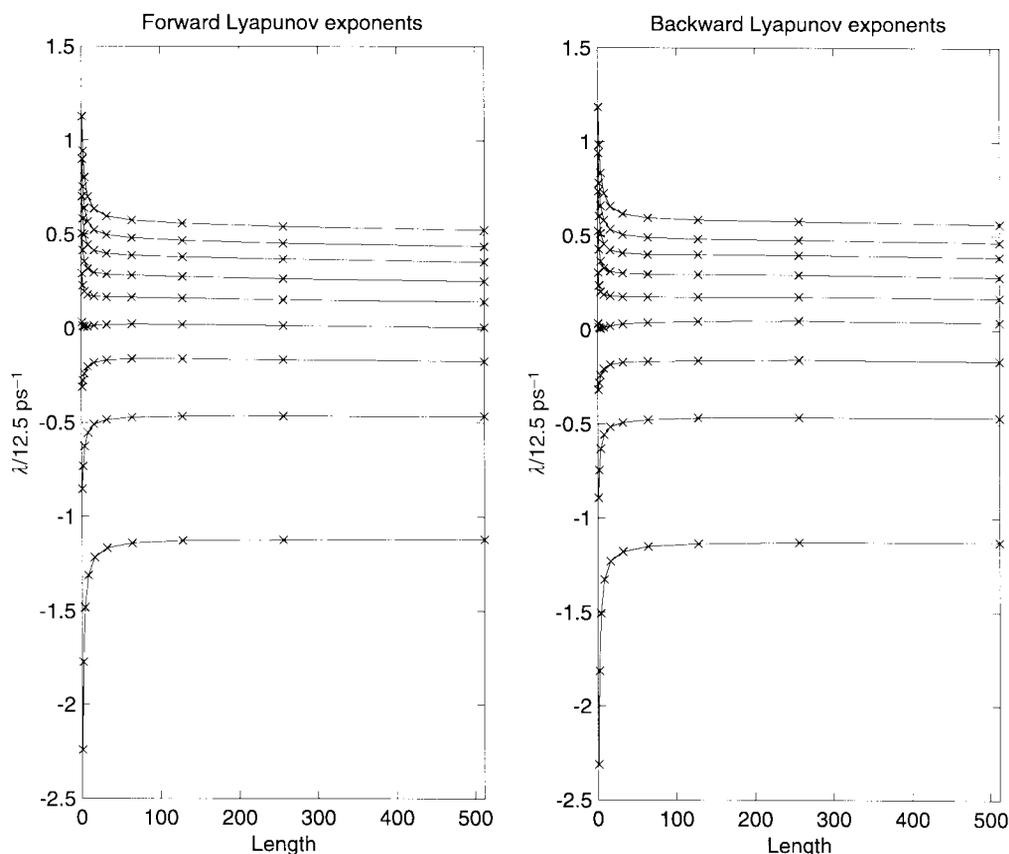
end-to-end distance was used to calculate the $I(T)$ function. The first minimum of the average mutual information function occurs at $T = 17$ (1.36 ps). Similar results, $T = 18$ (1.44 ps), were obtained using the data from the x-coordinate.

Using a time lag of 1.36 ps, the functions $E1$ and $E2$ were calculated for the two time series. As can be seen in Fig. 8a and b, saturation is reached at $d_E \sim 11$ after which it remains approximately constant. This provides evidence that we are dealing with a low dimensional system. The strength of this conclusion is enhanced when similar results are obtained using both time series data.

In the case of end-to-end distance the percentage of bad predictions seen in Fig. 9 is independent of the number of neighbours N_B and of the local dimension at $d_L \sim 9$, telling us that this attractor may be adequately described by nine degrees of freedom. Similar results are obtained using the x-coordinate of the end-to-end vector. This means that models for simulating the dynamic behavior of this peptide should have local ninth-dimensional dynamics regardless of the dimensions of the overall space within which the model is embedded.

These results shed light on the amazing chance of performing simulations only along the few active dynamic degrees of freedom. By reducing drastically the dimension of the conformational phase-space for dispersed molecules in solution, it would be possible to perform long time force-field simulations, characterizing slow molecular motions as helix–coil transitions and protein folding in biopolymers or reptation in synthetic polymers, which at the present cannot be investigated and is unlikely to be for many years.[34] In this framework Amadei *et al.* have developed the *Essential Dynamics* method.[35–37] This approximation is linear and therefore can be improved. Nevertheless, the analysis outcome using chaos techniques confirms it.

Fig. 10 shows the nine computed local Lyapunov exponents forwards and backwards in time. As can be seen, *five Lyapunov exponents are positive*, one is close to zero and the others are

Fig. 10 Average local Lyapunov exponents λ for the end-to-end distance trajectory $D_{ee}(t)$ computed forward and backward in time using $d_L = 9$. The length is expressed in 2^{L-1} where L is the number of time steps away from the time of perturbation. λ is measured in 12.5 ps^{-1}. For large L, local exponents become global (they converge to stable values).

Table 1 Lyapunov exponents of Fig. 10 for $D_{ee}(t)$ and the x-coordinate of the end-to-end vector, computed forward and backward in time, are summarized. Also the total sums D_L are reported

λ_j	$D_{ee}(t)$		x-Coordinate	
	Forward	Backward	Forward	Backward
1	0.5244	0.5617	0.5268	0.5425
2	0.4376	0.4644	0.4356	0.4534
3	0.3556	0.3870	0.3487	0.3673
4	0.2519	0.2810	0.2564	0.2734
5	0.1425	0.1684	0.1378	0.1509
6	0.0082	0.0402	−0.0057	0.0141
7	−0.1734	−0.1656	−0.2039	−0.1762
8	−0.4646	−0.4669	−0.4887	−0.4426
9	−1.1203	−1.1277	−1.1059	−1.1359
Sum	−0.0380	0.1426	−0.0988	0.0470
D_L	8.96605		8.91064	

negative. The results are summarized in Table 1 which also gives a negative total sum of the Lyapunov exponents, as expected for a dissipative system.

Taking into account the fact that the calculation of Lyapunov exponents is very susceptible to contamination, the agreement between the different data sets is quite satisfactory. Furthermore, the computation of the Lyapunov exponents forwards and backwards in time gives approximately the same results which tell us that the dynamical dimension has been correctly calculated.

The existence of positive Lyapunov exponents demonstrates further and strictly the chaos of conformational dynamics. The chaotic behavior of the tetrapeptide in solution could play a role in the expression of the restoring force in the entropic mechanism of elastin elasticity. Our findings are in agreement with the experimental evidence of Gaspard *et al.*[38] for microscopic chaos in fluid systems obtained by the observation of Brownian motions of a colloidal particle suspended in water. Moreover, they are in the framework of the *chaotic hypothesis* of Krylov[39] (who referred to microscopic dynamic instabilities) and Gallavotti and Cohen[40] which assume that the properties of statistical mechanics can be predicted by treating the systems as chaotic.

Conclusions

In a previous paper[5] the nonlinear dynamic behavior of Boc-Gly-Leu-Gly-Gly-NMe *in vacuo* was described in terms of solitons. Moreover, on the basis of *amplitude instability theorem* of these vibrations for large oscillations,[41] the transition from quasiperiodicity to chaos was hypothesized in aqueous solution, where high intramolecular mobility is expected.

In this work the conformational flexibility and the transition to chaos is demonstrated by modeling the tetrapeptide Ac-Gly-Leu-Gly-Gly-NMe in aqueous solution. The simulated end-to-end trajectory $R_{ee}(t)$ has been analyzed by determining its fractal dimension ($D = 1.64$), the anomalous diffusion of its motion ($H = 0.38$), the vibrational time-dependent picture, attractor dimension ($d_L = 9$) and Lyapunov's exponents (five positive values).

The peptide backbone undergoes large fluctuations around folded and unfolded states as exemplified by the behavior of the β-turn [Gly$_1$]C=O$\cdots$HN[Gly$_4$]. The chance to develop these larger amplitude motions is a *necessary condition* for the observed chaotic behavior. The *antipersistent fBm* of R_{ee} is the fingerprint of correlations inside the randomness of the trajectory, typical of chaos. The observed behavior departs from the ideal Brownian motion because of the nonlinear interactions in solution, as the *excluded volume effects*[42] which reflect the interaction between residues and the water H-bond network. It is similar to the behavior of collapsed chains in

poor solvents[43] characterized by the Flory exponent $v = \frac{1}{3}$. The observed anomalous diffusion is coherent with the *fBm* in fractal media: the diffusion lag of the peptide could be due to the H-bonded water molecules that constitute a solution percolation network.

The end-to-end trajectory $R_{ee}(t)$ has been analyzed using delay coordinate embedding in an effort to understand the nonlinear dynamic behavior of the system. The analysis using standard time-delay embedding techniques indicates low dimensional chaotic dynamics. Even though the original system has a large number of degrees of freedom, the results seem to indicate that only a few degrees of freedom are active in the final attractor and that the dynamics of such a system could be described by a reduced number of differential equations.

The existence of a low dimensional chaotic attractor with a few dynamic degrees of freedom active in determining the evolution of the system, is typical of nonlinear dissipative systems[44] to which molecules in solutions belong, and is consistent with Langevin dynamics description. Similar results have been obtained by a group currently researching the folding of a small protein simulated as a Langevin system[45] (with internal, viscous and external random forces) and by Zhou and Wang on a polyalanine peptide.[46]

The vibrational picture of intramolecular vectors is space-time self-similar testifying to a disordered behavior of the peptide chain on large scale. Chaos is sufficient but not a necessary condition for Brownian motion.[47] In fact, the analysis of simplified models showed that the erratic motion of dispersed particles can be due largely to the randomness of the initial conditions of the fluid molecules.[48] However, chaos is defined mathematically in terms of positive Lyapunov exponents. In this way the authors showed that the conformational motion of the elastin peptide in aqueous solution indeed has chaotic features.

Even if the results obtained for a short tetrapeptide cannot by any means be transferred to a complex molecule such as elastin, however, they are compatible with the transition-to-chaos mechanism of elasticity.[5,6] The interpretation of the observed chaotic dynamics in terms of elastin elasticity should be regarded as a hypothesis needing to be proven in further studies, using larger elastin substructures, control sequences and updated dynamics. It is reasonable to suggest that the physical behavior of elastin is a consequence of its nonlinear effects developed at the protein–water interface where glycine-rich regions, similar to the sequence of our peptide, are likely to be located and where chaotic dynamics is to be expected.

References

1 P. G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, 4th edn., 1979, 1993.
2 L. Debelle and A. M. Tamburro, *Int. J. Biochem. Cell Biol.*, 1999, **31**, 261.
3 V. Villani and A. M. Tamburro, *J. Chem. Soc., Perkin Trans. 2*, 1993, 1951.
4 V. Villani and A. M. Tamburro, *J. Biomol. Struct. Dyn.*, 1995, **12**, 1173.
5 V. Villani, L. D'Alessio and A. M. Tamburro, *J. Chem. Soc., Perkin Trans. 2*, 1997, 2375.
6 V. Villani, L. D'Alessio and A. M. Tamburro, in *Elastin and Elastic Tissue*, ed. A. M. Tamburro, Mario Armento Publ., Potenza, Italy, 1997, p. 31.
7 V. Villani and A. M. Tamburro, *J. Mol. Struct.* (*THEOCHEM*), 1998, **431**, 205.
8 A. M. Tamburro, V. Guantieri, L. Pandolfo and A. Scopa, *Biopolymers*, 1990, **29**, 855.
9 H. D. I. Abarbanel, *Analysis of observed chaotic data*, Springer, New York, 1996.
10 D. A. Pearlman, D. A. Case, J. C. Caldwell, W. S. Ross, T. E. Cheatham III, D. M. Ferguson, G. L. Seibel, U. Chandra Singh, P. Weiner and P. A. Kollman, AMBER 4.1, University of California, San Francisco, 1995.
11 T. Williams and C. Kelley, GNUPLOT 3.5, 1993.

12 S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comput. Chem.*, 1986, **7**, 230.
13 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein L., *J. Chem. Phys.*, 1983, **79**, 926.
14 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.
15 L. Verlet, *Phys. Rev.*, 1967, **98**, 159.
16 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327.
17 A. Bunde and S. Havlin, *Fractals and Disordered Systems*, Springer-Verlag, Berlin, 1992.
18 J. Feder, *Fractals*, Plenum Press, New York, 1988.
19 B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1982.
20 R. Metzler, U. G. Glockle and U. F. Nonnenmacher, *Physica A*, 1994, **211**, 13.
21 J. Perrin, *Les atomes*, Presses Universitaires de France, Paris, 1948.
22 F. Takens, in *Lecture Notes in Mathematics*, ed. A. Rand and L. S. Young, Springer-Verlag, Berlin, 1981, vol. 898, p. 366.
23 A. Fraser and H. Swinney, *Phys. Rev. A*, 1986, **33**, 1134.
24 P. Grassberger and I. Procaccia, *Phys. Rev. Lett.*, 1983, **50**, 346.
25 L. Cao, *Physica D*, 1997, **110**, 43.
26 M. B. Kennel, R. Brown and H. D. I. Abarbanel, *Phys. Rev. A*, 1992, **45**, 3403.
27 A. Wolf, J. B. Swift, H. R. Swinney and J. A. Vastan, *Physica D*, 1985, **16**, 285.
28 H. E. Stanley, *Phase Transitions and Critical Phenomena*, Oxford University Press, Oxford, 1971.
29 P. Bak, K. Tang and K. Wiesenfeld, *Phys. Rev. Lett.*, 1987, **59**, 381.
30 P. J. Flory and E. S. Weaver, *J. Am. Chem. Soc.*, 1960, **82**, 4518.
31 A. R. Bizzarri, C. Rocchi and S. Cannistraro, *Chem. Phys. Lett.*, 1996, **263**, 559.
32 G. Nemethy and H. A. Scheraga, *Biochem. Biophys. Res. Commun.*, 1980, **95**, 320.
33 S. R. Quarke, H. Babcock and S. Chu, *Nature*, 1997, **388**, 151.
34 H. S. Chan and K. A. Dill, *Physics Today*, 1993, **46**(2), 24.
35 A. Amadei, A. Linssen and J. C. Berendsen, *Proteins: Struct., Funct., Genet.*, 1993, **17**, 412.
36 A. Amadei, A. B. M. Linssen, B. L. de Groot, D. M. F. van Aalten and H. J. C. Berendsen, *J. Biomol. Struct. Dyn.*, 1996, **13**, 615.
37 D. M. F. van Aalten, B. L. de Groot, J. B. C. Findlay, H. J. C. Berendsen and A. Amadei, *J. Comput. Chem.*, 1997, **18**, 169.
38 P. Gaspard, M. E. Briggs, M. K. Francis, J. V. Sengers, R. W. Gammon, J. R. Dorfman and R. V. Calabrese, *Nature*, 1998, **394**, 865.
39 N. Krylov, *Nature*, 1944, **153**, 709.
40 G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.*, 1995, **74**, 2694.
41 G. H. Walker and J. Ford, *Phys. Rev.*, 1969, **188**, 416.
42 C. Domb, in *Advances in Chemical Physics vol. XV, Stochastic Processes in Chemical Physics*, ed. K. E. Shuler, John Wiley & Sons, New York, 1969, p. 229.
43 M. Nierlich, J. P. Cotton and B. Farnoux, *J. Chem. Phys.*, 1978, **69**, 1379.
44 H. G. Schuster, *Deterministic Chaos*, VCH, Weinheim, 1995.
45 J. M. Zaldivar and A. Abecasis, 5th SIAM Conference on Dynamical Systems, Snowbird, Utah USA, 25–27 May 1999.
46 H. Zhou and L. Wang, *J. Phys. Chem.*, 1996, **100**, 8101.
47 D. Durr and H. Spohn, *Nature*, 1998, **394**, 831.
48 H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer, Berlin, 1991.