Fractal analysis of protein potential energy landscapes

D. A. Lidar,[1,2,*] D. Thirumalai,[3] R. Elber,[2] and R. B. Gerber[2,4]
[1]*Department of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel*
[2]*Department of Physical Chemistry and The Fritz Haber Center for Molecular Dynamics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel*
[3]*Department of Chemistry and Biochemistry, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742*
[4]*Department of Chemistry, University of California—Irvine, Irvine, California 92697*
(Received 18 August 1998; revised manuscript received 11 November 1998)

The fractal properties of the total potential energy V as a function of time t are studied for a number of systems, including realistic models of proteins (pancreatic polypeptide, bovine pancreatic trypsine inhibitor, and myoglobin). The fractal dimension of $V(t)$, characterized by the exponent γ, is almost independent of temperature, and increases with time, more slowly the larger the protein. Perhaps the most striking observation of this study is the apparent universality of the fractal dimension, which depends only weakly on the type of molecular system. We explain this behavior by assuming that fractality is caused by a self-generated dynamical noise, a consequence of intermode coupling due to anharmonicity. Global topological features of the potential energy landscape are found to have little effect on the observed fractal behavior. [S1063-651X(99)11402-8]

PACS number(s): 87.15.−v, 36.20.−r, 61.43.Hv

I. INTRODUCTION

The energy landscape of proteins near the folding temperature is often described as complex [1,2]. That is, the landscape is expected to be rugged and composed of many minima and maxima, separated by barriers of varying heights, the result of frustration due to conflicting interactions between different segments of the protein. The complex nature of the underlying energy landscape has been demonstrated using ergodic measures [3], and by exact enumeration of minima and transition states of a tetrapeptide [4]. The goal of the present work is to investigate this energy landscape under a variety of conditions using tools from fractal analysis. The trajectories needed for analysis were generated using MOIL, a molecular dynamics (MD) program which allows realistic simulations of small proteins using semiempirical potentials [5]. Using MOIL, the total potential energy $V(t)$ was calculated as a function of time, for several proteins of choice [PPT (pancreatic polypeptide), BPTI (bovine pancreatic trypsine inhibitor) and myoglobin] at constant temperature. For comparison, we also simulated polyalanine and NaCl. The $V(t)$ curves so obtained under a variety of choices of parameters were then subjected to a fractal analysis. The main purpose was to correlate the resultant fractal dimensions with the topology of the energy landscape, and to understand the connection to the relevant physical parameters. Both MOIL and the method used to calculate the fractal dimension of the $V(t)$ curves are described in Sec. II.

There are several reasons for using a fractal approach in the analysis of the $V(t)$ curves. First, there is an intimate connection between ''roughness'' and fractality: much is known about the generation of rough surfaces using fractal algorithms and, conversely, on the analysis of rough surfaces in fractal terms [8,9]. In this sense the notion of the roughness of the energy landscape of proteins naturally calls for a fractal analysis, which can *quantify* the degree of roughness. Second, fractal dimensions often show up as ''universal exponents,'' and are related to the critical exponents of phase transition theory [10]. Conversely, there is a certain universality in protein dynamics, in the sense that most proteins exhibit a folding transition and multiexponential relaxation times. This suggests a common feature in their energy landscapes. While the rapid transition to folded states in small proteins is best rationalized by the existence of a dominant native basin of attraction or a folding funnel [2,11], the relaxation phenomenon might well be related to the roughness of the energy landscape, and hence to its fractal dimension. A ''universal'' fractal dimension would be a strong characterization of what it is that is common to different energy landscapes. Third, establishing that the energy landscape of a protein is indeed fractal would allow one to relate the subject of protein dynamics to a whole body of work on dynamics on fractals, e.g., diffusion on a fractal substrate [12].

The main problem in trying to characterize the roughness of a high-dimensional energy landscape, such as that exhibited by a protein, is the proper choice of a reaction coordinate. In this work we have chosen *time* for this purpose. The advantage is that time is a universal coordinate (unlike any of the spatial coordinates), and one can think of the protein potential energy dynamics in terms of this coordinate as a random walk taking place on a rough substrate. Indeed, the rate of sampling of conformation space in complex many-body systems has been shown [3] to be a diffusive process in the energy space. The universality of the time coordinate, however, can also be a drawback, in that it is not very sensitive to the global topology of the energy landscape, but

*Present address: Department of Chemistry, The University of California, Berkeley, CA 94720.

rather to the local structure of minima. Indeed, as will be presented in detail in Sec. III, our results appear to contain information mostly on the role of *anharmonicity* in determining the fractal features of the $V(t)$ curves. As such, the results reported here reflect more on the general properties of dynamics in anharmonic potentials than on protein dynamics per se. These and other issues will be discussed and analyzed in detail in Sec. IV. Conclusions and a summary are presented in Sec. V. Finally, to establish a simple reference model, an analytical solution for the fractal dimension of the $V(t)$ curve in the Rouse model is presented in the Appendixes.

II. METHODS

A. MOIL simulations

The MOIL program has been described extensively in the literature [5]. In a nutshell, it takes as input the x-ray structure of a given protein from the Brookhaven Protein Data Base, minimizes this structure, and runs molecular dynamics. MOIL makes it possible to simulate the energetics, dynamics, and thermodynamics of systems that vary in size, up to several thousands of atoms. It accounts for a number of different energy contributions: van der Waals and electrostatic (noncovalent), bonds, angles, torsions, and improper torsions (covalent):

$$V=\sum V_{\mathrm{vdW}}+\sum V_{\mathrm{electrostatic}}+\sum V_{\mathrm{bond}}+\sum V_{\mathrm{angle}} + \sum V_{\mathrm{torsion}}+\sum V_{\mathrm{improper}}. \tag{1}$$

The sums are over all interacting pairs of particles, and the components are given as follows:

$$V_{\mathrm{vdW}}=\frac{a_i a_j}{r_{ij}^{12}}-\frac{b_i b_j}{r_{ij}^{6}}, \tag{2}$$

where a_i and b_i are constants that depend on the type of atom i, and r_{ij} is the distance between atoms i and j;

$$V_{\mathrm{electrostatic}}=\frac{q_i q_j}{\epsilon r_{ij}}, \tag{3}$$

where q_i is the charge on atom i and ϵ is the dielectric constant; and

$$V_{\mathrm{bond}}=k_{ij}(r_{ij}-\langle r_{ij}\rangle)^2, \tag{4}$$

where k_{ij} is the bond force constant and $\langle x\rangle$ denotes the equilibrium value of coordinate x;

$$V_{\mathrm{angle}}=\kappa_{ijk}(\theta_{ijk}-\langle\theta_{ijk}\rangle)^2, \tag{5}$$

where κ_{ijk} is the angle force constant, and θ_{ijk} is the angle between two sequential bonds, i.e., bonds connecting atoms i,j and j,k; and

$$V_{\mathrm{torsion}}=\sum_{m=1}\alpha_{m,ijkl}\cos^m(\phi_{ijkl}). \tag{6}$$

Here i, j, k, and l are four sequentially connected atoms; ϕ_{ijkl} is the angle between the planes defined by atoms i, j, and k and j, k, and l; $\alpha_{m,ijkl}$ is a constant which depends on the types of atoms i, j, k, and l and on the periodicity of the rotation m (see Ref. [5] for details). Finally,

$$V_{\mathrm{improper}}=\mu_{ijkl}(\Phi_{ijkl}-\langle\Phi_{ijkl}\rangle)^2. \tag{7}$$

Here j, k, and l are all connected to atom i, and Φ_{ijkl} is the angle between the planes defined by atoms i, j, and k and j, k, and l.

In the present work, we considered only the time dependence of the *total* potential energy V due to all these terms. The time series for a number of systems were generated using numerical simulations employing either the canonical or microcanonical ensembles. As is customary, all simulations were preceded by a short initial annealing (heating) period. In the ''equilibration mode'' (used mainly for testing), the kinetic energies of all particles were rescaled in order to preserve the constant temperature (canonical ensemble). In the more important ''no-equilibration mode'' (default unless mentioned otherwise), no such rescaling was attempted, and the system was allowed to reach thermal equilibrium at the specified temperature (microcanonical ensemble).

The exponents characterizing the fractal nature of $V(t)$ reported below derive from averages over groups of 25 runs differering only in the initial random velocity distribution. A number of tests were performed to ensure that our results are not artifacts of the particular integration schemes used. We checked that using more than 25 runs did not significantly affect our results statistically. Between groups of 25 but for fixed molecule type, the runs differed, e.g., in temperature, addition of a water solvation shell, treatment of the terminus hydrogens, bond freezing, and deletions of residues. Most runs started at a temperature of $T_i=1$ K and heated in 1 K per time-step increments to the final temperature T_f, at which the rest of the simulation was run. The total duration of the simulations, once T_f was reached, was usually 10 or 25 psec, corresponding to 20 000 or 50 000 time steps, respectively, of 0.5 fsec each. Exceptions to these conditions include some shorter and longer runs, with some variation in time-step size. They will be discussed in detail below.

B. Fractal analysis

The $V(t)$ curves were subject to a standard fractal analysis, with the purpose of calculating their fractal dimension

$$\gamma=2-\alpha, \tag{8}$$

where α is the Hölder exponent, defined in Sec. II B.1.

1. General theory

Given any set of points $E=\{(x,y)\}$, one can estimate their fractal dimension by a number of (in principle) equivalent procedures. A popular one is the so-called ''Minkowski sausage,'' [13,14], where one draws a circle of radius R centered at each point in E, and calculates the area $A(R)$ of the union (''sausage'') of all such circles. The fractal dimension is then found as

$$\gamma(E)=\lim_{R\to 0}\left[2-\frac{\ln[A(r)]}{\ln(R)}\right]$$

$$\approx \text{slope}\{\log(1/R),\log[A(R)/R^2]\}, \tag{9}$$

where the second (approximate) equality follows by assuming that the first equality holds for all R and multiplying it throughout by $\log(1/R)$. This assumption is equivalent to a linear relationship between $\log(1/R)$ and $\log[A(R)/R^2]$, so that γ is found as the slope of the corresponding log-log plot. If such a linear relationship is not found, it is taken in practice as an indication that E is not a fractal set. Conversely, if at least one order of magnitude of linearity is observed, the convention is to consider E a fractal set in that range [15]. It should be noted that, rigorously speaking, the curve of a *function* is generally not a self-similar fractal, but rather a *self-affine* one. The latter refers to those cases where the axes have different units, so that the scaling might be different in the two directions. Indeed, a (deterministic) self-affine curve satisfies the scaling relation

$$h(t)=b^{-\alpha}h(b\,t), \tag{10}$$

which holds only on average for random processes. The ''Hölder,'' or self-affinity exponent α, is $2-\gamma$ [Eq. (8)] if the curve is subjected to the Minkowski analysis of Eq. (9) [16,17]. An example is the Weierstrass-Mandelbrot function [13] $h(t)=\Sigma_{n=-\infty}^{\infty}b^{-n\alpha}(1-\cos b^n t)$, which satisfies Eq. (10) and whose fractal dimension is $2-\alpha$. An equivalent characterization of statistically self-affine fractals is through the correlation function, which satisfies

$$C_h(t)=\langle[h(t+t_0)-h(t_0)]^2\rangle\sim t^{2\alpha}. \tag{11}$$

This relation is expected to hold for durations shorter than a typical correlation time t_c, found from

$$t_c=\frac{\int dt\, t\Gamma(t)}{\int dt\,\Gamma(t)}, \tag{12}$$

where $\Gamma(t)=\langle h(t_0)h(t+t_0)\rangle-\langle h(t_0)^2\rangle$.

2. Analysis of time series

In the present case the (self-affine) set under consideration is: $E=\{(t,V(t))\}_t$, and the fractal analysis we applied is a numerically well-behaved version of Eq. (9): the ''ϵ-variation method.'' We briefly describe the method in Appendix B; for a full description see Ref. [18]. We applied this analysis in order to extract the self-affinity exponent (which we refer to simply as the ''fractal dimension'' and denote by γ) of our $V(t)$ curves.

III. RESULTS

In this section we will present the findings from our numerical simulations. A discussion will be postponed to Sec. IV. We will start by presenting a sample of ''raw'' simulation results, then move on to a detailed presentation of each of the three proteins, compare them to each other, perform a comparison with several test cases, estimate the effect of time scales of different processes, and finally study the role of anharmonicity vs quasiharmonicity. All our MD simulations started out from the folded state (as determined by x ray), which we subsequently minimized using the Powell algorithm [5]. The protein was then heated from 1 K in increments of 1 K at every time step, with a random initial velocity distribution, until the desired final temperature was reached. Fractal analysis was performed on data from the constant temperature period only. Dynamics on all three proteins was run for 25 psec, with a time step of 0.5 fsec (with a few exceptions, to be detailed below). Fractal analysis was performed on the full 25-psec trajectories and on the first 10 psec of each such trajectory. This process was repeated for a wide selection of temperatures in the range 1–800 K, with a focus on the room temperature regime.

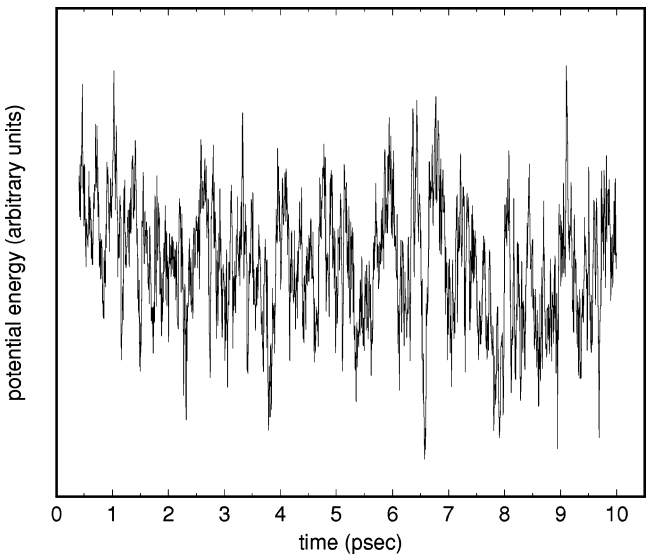

FIG. 1. Total potential energy as a function of time for a single run of BPTI at 310 K. Annealing from 1 K occurs during the first 0.4 psec, which are not shown. The highly irregular fluctuations are reminiscent of fractional Brownian noise.

A. ''Raw'' results

A graph of a typical $V(t)$ curve is displayed in Fig. 1, for BPTI at 310 K. We only present the time series after the temperature is stabilized, i.e., the first 0.4 psec during which annealing from 1 K occurs are omitted. The noisy data are strikingly similar to ''fractional Brownian motion'' [13], and clearly suggest the use of fractal-type analysis. This is done in Fig. 2, where the corresponding ϵ-variation analysis is shown. A straight line can be fitted with confidence to the ϵ-variation data over some 1.5 orders of magnitude (decades). The linear regression coefficient in this case is $\mathcal{R}^2=0.998$, where 1 indicates a perfect straight line. The fractal dimension (slope in Fig. 2) is $\gamma=1.762\pm0.004$, where the uncertainty is due to the linear regression. All log-log plots look strikingly similar to the one in Fig. 2, and have a similar range of linearity. The $\mathcal{R}^2$ coefficient rarely went below 0.99. Figure 3 presents all 25 fractal dimensions γ obtained from the runs for BPTI under the above-mentioned conditions. The dashed line is the average fractal dimension, which in this case yields $\gamma=1.75\pm0.02$. The uncertainty in

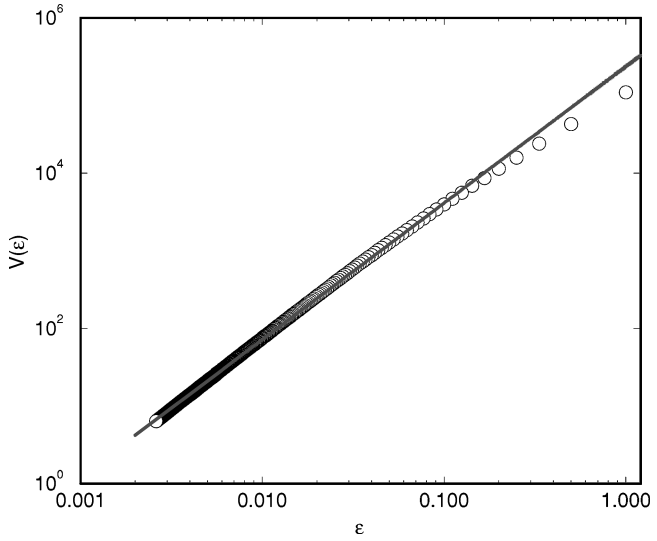

FIG. 2. ϵ-variation [18] result for the curve shown in Fig. 1. A straight line can be observed for about 1.5 decades. The bending at high ϵ values is due to finite size effects.

this number reflects both the individual linear regression errors on each data point, as well the standard deviation about the average. The procedure illustrated in Figs. 1–3 was repeated throughout this work. Thus every γ data point in the figures to follow is the result of an average over 25 fractal dimensions obtained from independent runs differing only in the initial velocity distribution.

B. Myoglobin

Figure 4 shows the results of our simulations of myoglobin. Every γ point was obtained as discussed above, so the error bars are due to both the averaging over the 25 data points and the linear regression error on each of these points. The dashed (solid) line connects the γ values resulting from the 25-psec (10-psec) runs. The inset shows the room temperature results. In the 10-psec case, γ appears to have a

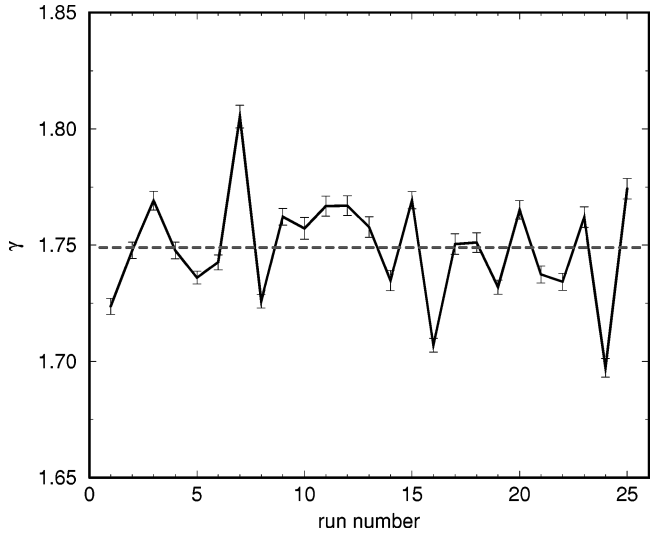

FIG. 3. All 25 fractal dimensions calculated for BPTI at 310 K. Individual error bars are due to linear regression. The dashed line is the average fractal dimension.

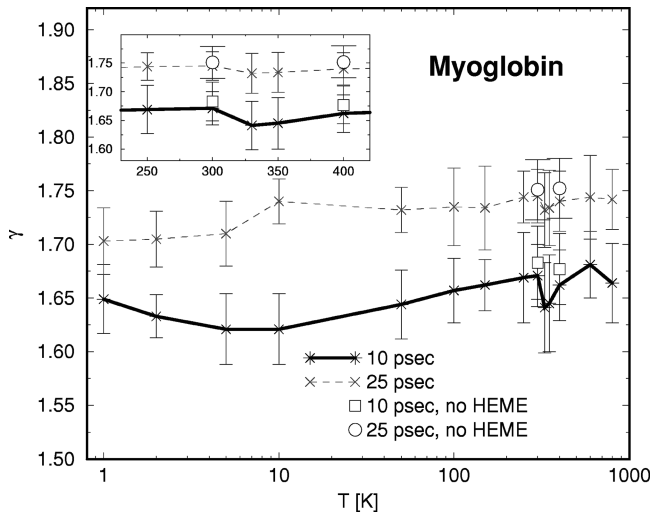

FIG. 4. Fractal dimension as a function of temperature for myoglobin. Lines connecting data points are guides to the eye only.

shallow minimum around 10 K, which is, however, absent in the full 25-psec simulations. The error bars are in fact too large to discuss confidently a trend in the curves. However, it is clear that *the fractal dimension increases with the simulation (or observation) time* τ: on average $\gamma \approx 1.67$ for 10 psec, and rises to $\gamma \approx 1.72$ for 25 psec. This behavior is seen in all our results ahead (see Figs. 5 and 6). As is clear from the inset, the fractal exponent γ is essentially constant around room temperature, apart from the rather sharp drop at 330 K.

The robustness of the results was tested by running the simulations at 300 and 400 K with the heme group deleted. The removal of the heme group makes myoglobin considerably less compact, causing it to unfold rather easily. Hence we expected a significant change in the underlying dynamics. However, Fig. 4 shows (squares and circles) that the influence of heme deletion is very minor, and results in a small increase in γ.

C. BPTI

The results of our simulations of BPTI are shown in Fig. 5. The overall behavior is very similar to that of myoglobin.

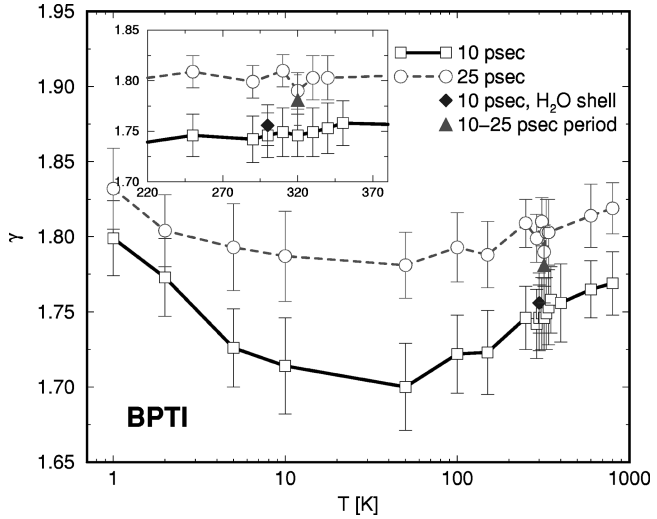

FIG. 5. Fractal dimension as a function of temperature for BPTI.

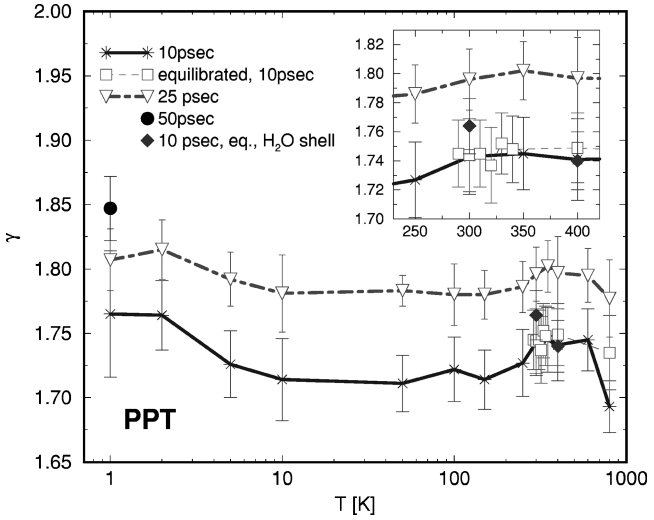

FIG. 6. Fractal dimension as a function of temperature for PPT.

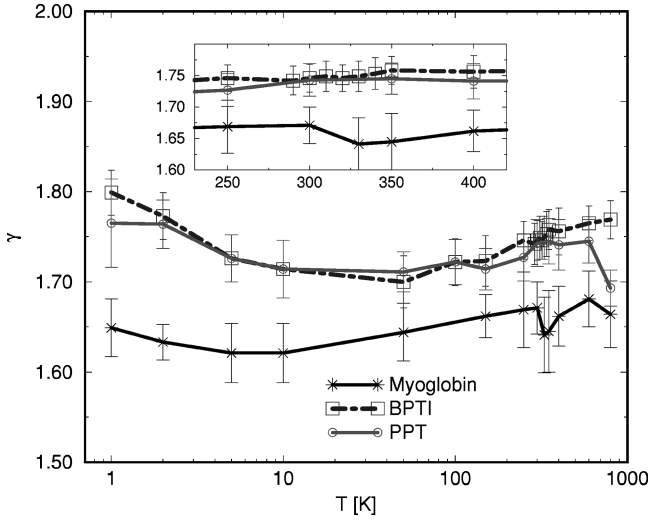

FIG. 7. Comparison among myoglobin, BPTI, and PPT for 10-psec simulations.

The differences are (1) overall the fractal dimensions are higher; and (2) the minimum is more pronounced (beyond statistical error), also visible in the 25-psec data, and has shifted to 50 K. The increase in γ with increasing observation time for the longer simulation is unmistakable for BPTI as well.

We also performed a 10-psec simulation including a water solvation shell (diamond). Clearly, this has no significant effect on the results.

As a consistency test, we calculated γ for the complementary 10–25 psec of the simulation (triangle). The result is a γ value in between the 10- and 25-psec values. Finally, the dip observed for myoglobin can be seen here as well as in the 25-psec simulation, at $T=320$ K.

D. PPT

Our PPT simulation results appear in Fig. 6. The general qualitative similarity to BPTI and myoglobin is noticeable, namely, the increase in overall fractal dimension with simulation time, and the shallow minimum at around 50 K. Several additional details are noteworthy: (1) We included a 50-psec-long simulation at 1 K (same 0.5-fsec time step), and the increase in γ persists (full circle). (2) We also performed simulations in the ''equilibration mode'' (see Sec. II A). These are indicated as squares, and can be seen most clearly in the inset. The results are virtually identical to the usual ''no-equilibration mode,'' demonstrating the agreement between the canonical and microcanonical ensembles for large (>100 atoms) systems. The only exception is the point at 800 K, where the ''no-equilibration'' value of γ drops rather sharply. (3) A water solvation shell (diamonds, also in the equilibration mode) again does not produce statistically significant results.

E. Comparison among myoglobin, BPTI, and PPT

Figures 7 and 8 compare the results for the three proteins in the 10- and 25-psec cases, respectively. These figures merely recapitulate the results from Figs. 1–6, and show no new data. The main points to notice are the striking similarity between PPT and BPTI, compared to the overall lower fractal dimension values for myoglobin. This holds in both the 10- and 25-psec cases.

F. PPT under a variety of conditions

We performed further tests of the robustness of our results by running 10-psec simulations of PPT at 300 K, under different conditions, as shown in Fig. 9. There are 11 data points in this figure, and the γ values are arranged in increasing order. Let us briefly describe the results, noting that case 7 is the ''standard'' one (i.e., the result shown in Fig. 6).

The case ''no α-helix'' refers to a simulation in which the entire PPT α helix was deleted. Under these conditions a significantly smaller γ value is obtained. On the other hand, data point 5 is the opposite case, where only the α helix remains. The corresponding γ is much closer to the reference value of data point 7. This shows *that the α helix is by and large responsible for the observed fractal dimension.*

Data point 2 corresponds to a freezing of all bond vibrations. The remaining motions are therefore proper and im-

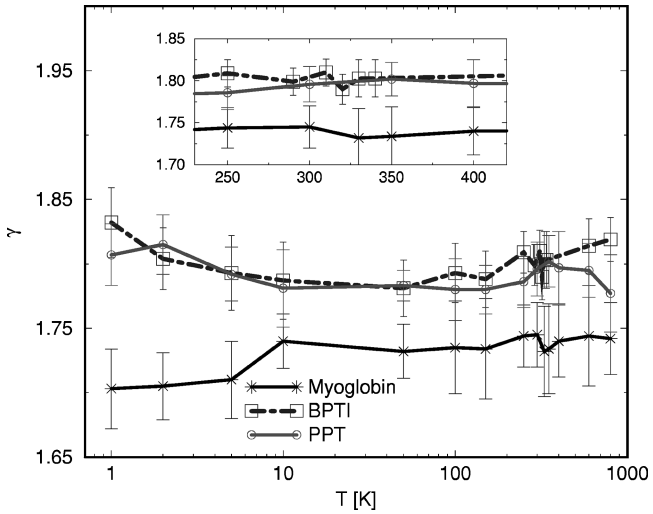

FIG. 8. Comparison among myoglobin, BPTI, and PPT for 25-psec simulations.

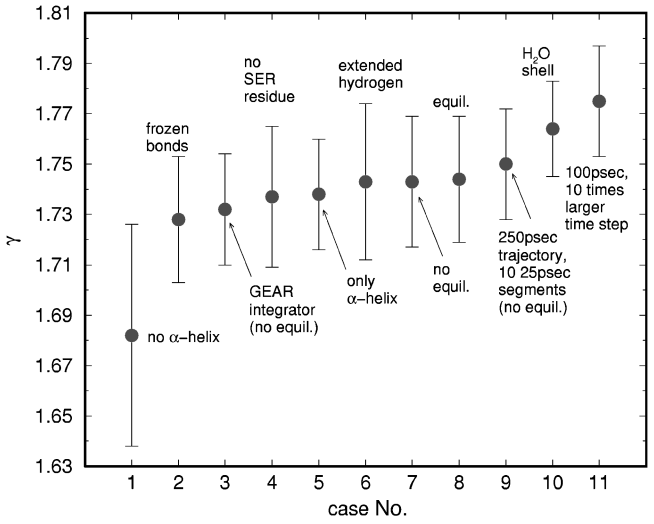

FIG. 9. A comparison of results for PPT at 300 K and 10 psec, under a variety of conditions.

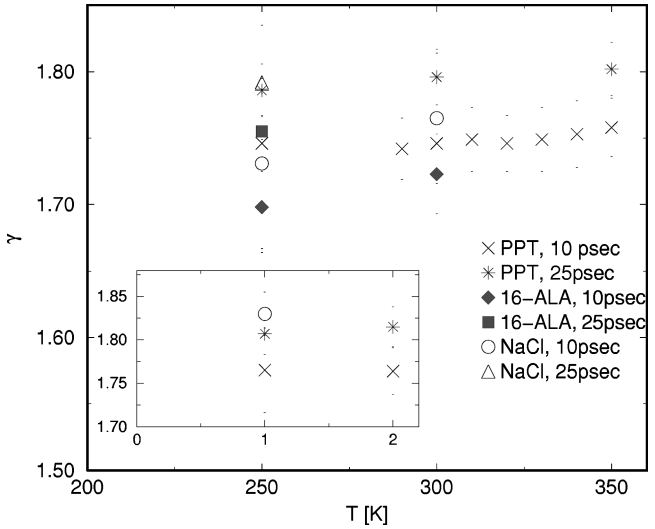

FIG. 10. A comparison between PPT, 16-alanine, and NaCl (40 atoms). Error bars have been omitted for clarity and are replaced by small dots.

proper rotations and center of mass translations. Since the deviation from the reference γ value is small, it appears that *the rotations are largely responsible for the observed fractal dimension*.

Data point 3 is the result when a different integrator is used, which appears to have only a minor effect. This test was performed to check whether the results are a consequence of numerical instabilities.

Data point 4 is another deletion experiment, this time of just the SER residue. This has a very small effect which, curiously enough, is slightly larger than leaving only the α helix, as indicated in case 5 (although due to the large error bars in fact this effect cannot be taken too seriously).

To calculate the γ corresponding to data point 6, the terminal hydrogens were treated as ''extended'': unlike in all other cases, the hydrogens were not treated as atoms but were ''absorbed'' into the carbons. This modifies the carbon mass and radius to appropriate effective values. The effect, however, is negligible.

Data points 8 and 10 relate to using the equilibration mode and adding a water solvation shell, respectively, which (as shown also in Fig. 6) do not have a significant effect either.

Data point 9 reports the result from a single 250-psec-long trajectory: the trajectory was divided up into ten segments of 25 psec each, and the γ's were averaged. Data point 11 is due to a 100-psec-long trajectory which, however, employed a time step of 5 fsec (ten times larger than the usual time step). In accordance with the observation of a fractal dimension which increases over time, the resulting average γ's of data points 9 and 11, are larger than the reference value.

G. PPT vs test systems: Polyalanine and NaCl

In order to estimate to what extent our results reflect properties unique to realistic proteins, we present calculations for two test systems: poly (16)-alanine and a 40-atom cluster of NaCl. Polyalanine folds into an α helix, so it has the crucial part of the protein topology (recall points 1 and 5 in Fig. 9), but the underlying energy landscape is not expected to be rugged. This is because all the residues are identical so there is no frustration due to conflicting interactions between different segments of the polymer. (Nevertheless each residue has internal conflicting interactions, giving rise to some ruggedness). The ground state of the nonpolymeric NaCl cluster is easily accessible [6]; i.e., starting from a disordered state, it finds its crystalline ground state well before the end of our 10-psec simulation. Thus it has neither the protein topology, nor its rugged energy landscape. The results of our simulations are shown in Fig. 10, which displays 10- and 25-psec runs of 16-alanine and NaCl at 250 K, as well as 10-psec runs at 300 K, and 10-psec runs for NaCl alone at 1 K. For clarity of presentation the error bars have been removed. However, it should be remarked that when the error bars are taken into account the results for a given temperature overlap. Indeed, it is striking that *the results for polyalanine and NaCl do not differ significantly from those for PPT*. In fact for 250 K, NaCl and PPT agree almost exactly, and the agreement for 10 psec at both 250 and 300 K is quite close. Only at 1 K does there seem to be a significant difference between NaCl and PPT. 16-alanine has lower γ values than PPT, but not very much so. Barring accidental agreement, these results clearly indicate that *neither the protein topology, nor the rugged energy landscape are essential in order to obtain fractal $V(t)$ curves, with fractal dimension close to that of realistic proteins*.

H. Role of observation time: Short simulations of PPT

In order to investigate systematically the role of the total duration of the simulation and the correspondingly participating physical processes, we ran simulations of 20 000 time steps of PPT at 300 K, with progressively smaller time steps. This is different from the bulk of our simulations, where a constant time step of 0.5 fsec was employed, so care should be exercised in the comparison. The main point is that by employing a smaller time step, previously inaccessible fast motions are now observable. The results are shown in Fig.

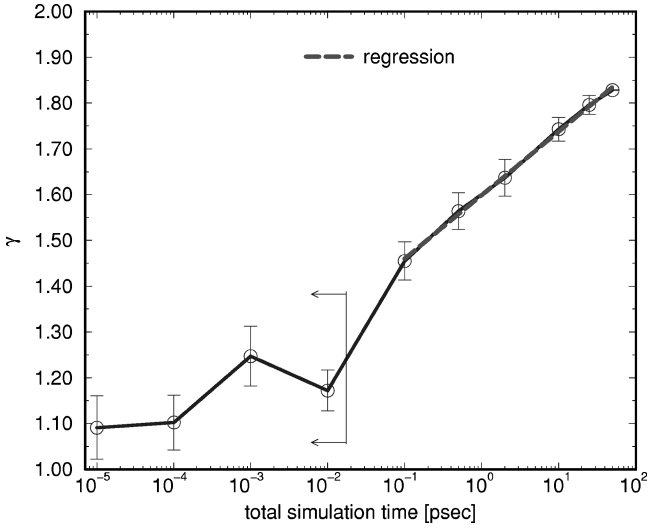

FIG. 11. 20 000 time step simulations of PPT at 300 K with varying total duration (x axis). Results to the left of the arrows are in the unphysical range.

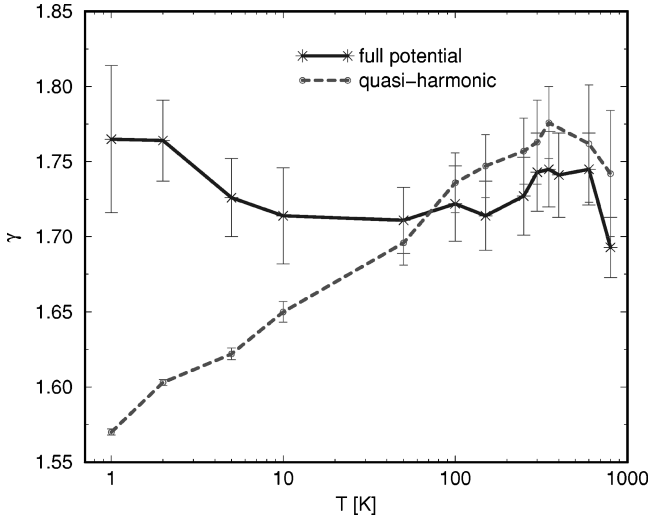

FIG. 12. 10-psec simulation results of PPT with the full potential (solid line), and just the quasiharmonic part (dashed line).

11. It is remarkable that the fractal dimension obeys an almost perfect logarithmic law over three orders of magnitudes,

$$\gamma(\tau) = 1.60 + 0.060 \ln(\tau), \tag{13}$$

for 0.1 psec$\leqslant\tau\leqslant$100 psec. The regression coefficient is $\mathcal{R}^2 = 0.999$. Figure 11 thus demonstrates decisively that γ is indeed strongly dependent on the total observation time [7].

Below $\tau = 100$ fsec (to the left of the arrows in Fig. 11) the behavior of γ appears rather erratic, and we believe this is due to numerical error. Indeed, the fastest motions in the protein are the OH or CH stretches, with a period of about 10 fsec. Thus $\tau = 10$ fsec would most likely not probe even a single period, and can confidently be regarded as the lower physical limit of our simulations. We turn next to an analysis and discussion of our results.

IV. DISCUSSION

The most striking observation presented in Sec. III is the rise of the fractal dimension with time. It suggests (at least) two likely hypotheses for its explanation: *As time evolves*, (1) there is increased access to local minima and ''fine structure'' of the potential energy landscape; and (2) low frequency vibrations and energy transfer between modes become increasingly activated.

The first of these attributes the growth in γ over time to the increased roughness of the sampled potential energy landscape. In other words, the protein ''visits'' more and more local minima (or basins of attraction) as it samples its conformational space. This process is consistent with an increase in fractal dimension, since the fractal potential energy landscape is not sampled all at once by the protein: instead the inner details are only progressively revealed. However, as shown by Czerminski and Elber [4], activated transitions between similar conformational structures confined to a small set of minima (separated by barriers on the order of $k_B T$) can occur at *room temperature* on a time scale of several picoseconds. A much larger time scale is required for such transitions at temperatures on the order of 1 K. The insensitivity of γ to temperature therefore appears to rule out the first hypothesis: at the low temperature end one would expect the protein to be confined to a single minimum within the time scale of our simulations, and the fractal dimension to be smaller as a consequence. Thus while increased access to local minima is consistent with the growth of γ with time, it is inconsistent with the lack of temperature dependence. We conclude that the fractal dimension measured in our simulations is not a good probe of the *global* potential energy landscape topology.

This brings us to the second hypothesis which, as will be argued shortly, is consistent with *both* the observed time and temperature dependencies. According to this hypothesis, the primary factor determining the fractal dimension of the potential energy curves is the self-generated *dynamical noise*, namely the very broad frequency spectrum appearing in the dynamics. This spectrum and the resulting self-generated noise is a consequence of the *anharmonicity* in the potential, as is demonstrated convincingly in Fig. 12. Shown in this figure, along with the full 10-psec simulation of PPT at 300 K, are the results of the same simulations with interaction terms $V_{\mathrm{vdW}}, V_{\mathrm{electrostatic}}$, and V_{torsion} deleted from the total potential energy of Eq. (1). Even with these terms absent in the potential energy we expect the system to exhibit quasiharmonic behavior only at sufficiently low temperatures. This is because anharmonicities are implicitly present in the potentials given in Eqs. (4) and (5). This becomes clear if the angle and stretch potentials are expressed in Cartesian coordinates. Thus at high enough temperatures the effect of the implicit anharmonicities should be evident. The results in Fig. 12 show these effects rather dramatically: there is a strong increase of γ as the temperature is increased. This shows that anharmonicity (even when included in describing internal degrees of freedom of the proteins) is of paramount importance in determining γ at high temperatures. It is interesting to note that the values of γ at low temperatures are greater than that for the Rouse model, which has an explicit quadratic Hamiltonian (in Cartesian coordinates) at all temperatures. This strongly implies that the value of $\gamma = 1.5$ may indeed be a lower bound for the systems considered here. At

high temperatures we find that the values of γ for PPT containing only Eqs. (4) and (5) are, within error bars, nearly the same as that obtained with the full potential. This also indicates that nonlinearities are the main origin of the apparent universal description of fractality of potential energy surfaces in many-body systems.

The mechanism we propose is thus as follows: $V(t)$ is a sum over many anharmonic terms, the number of which is proportional to the number of atoms in the protein. These anharmonic terms create *overtones*, which effectively fill the frequency spectrum and create a quasicontinuum. Now, on the one hand, as time evolves (at constant temperature) both very low and very high (previously inactive) frequency components become activated: the low frequencies are only sampled on long times, whereas the high frequencies correspond to highly energetic modes which must await energy transfer before equilibration sets in. On the other hand, as temperature increases (at constant τ) intermode coupling causing energy transfer from rigid modes to soft modes is facilitated, and more frequency components are activated. Thus in both cases one expects an increase in dynamical noise, with a corresponding increase in the fractal dimension of the signal. This accounts for the behavior seen in Figs. 4–6. Unfortunately we have not been able to study whether the fractal dimension converges on a limiting value upon full equilibration (at least as a function of temperature this does not seem to be the case—see Figs. 7 and 8), since this requires prohibitively long simulations [7].

The behavior of the quasiharmonic case observed in Fig. 12 might seem curious in light of these arguments: after all, in the purely harmonic case there is no intermode coupling, nor any dynamical noise. However, we recall that due to the use of non-Cartesian coordinates [see Eqs. (4), (5), and (7)], a certain *residual degree of anharmonicity* remains with respect to the Cartesian components. At very low temperatures vibrations and rotations occur very close to their equilibrium values, and the anharmonic effects are barely noticeable. Only as the temperature rises do these vibrations and rotations start to deviate significantly from their equilibrium values and gain anharmonicity. This explains the rise of γ in the quasiharmonic case. In contrast, the other terms [Eqs. (2), (3), and (6)] make anharmonic contributions at all temperatures, which accounts for the relative constancy of γ in this case.

The different behavior of PPT and BPTI vs myoglobin can be attributed to the *size* of these proteins. Myoglobin has 156 residues, whereas PPT and BPTI only have 41 and 60 residues, respectively. Therefore, at any given temperature it takes longer to activate the overtones that give rise to the self-generated noise in myoglobin than in either PPT or BPTI, which explains the lower fractal dimension of myoglobin in Figs. 7 and 8. The proximity in size between PPT and BPTI, on the other hand, is the reason for their quantitatively similar behavior. A related argument makes clear why deletion of the heavy heme group from myoglobin yields a higher γ (Fig. 4): without heme the energy redistribution among the modes occurs faster.

Also the curious similarity between NaCl and PPT can be attributed to anharmonic effects: NaCl, with its purely Coulombic interactions, is of course strongly anharmonic, even more so than PPT. On the other hand, polyalanine is more harmonic, so as expected it has lower γ values.

It is thus seen that the ''anharmonicity'' hypothesis is capable of accounting for all of the salient features present in our results for the fractal dimension of the $V(t)$ curves. The role of multiple minima and ruggedness of the energy landscape seems to be more restricted: it is apparently not the global topology which is measured by γ. Nevertheless, it should be noted that *different* multiple minima will generally lead to a denser filling of the frequency spectrum or to the appearance of bands, which may overlap and thus induce additional intermode coupling.

An important remaining issue concerns the *value* of γ observed in our simulations: definitely greater than 1.5 (except perhaps for the $\tau=100$ fsec simulation, Fig. 11). *A priori*, it is not clear what the lower limit of γ for polymerlike systems subject to the usual dynamics (like Newtonian or Langevin) should be. A plausible lower limit of 1.5 is derived in the Appendixes for the Rouse model. The Rouse model considers a purely harmonic set of beads under the influence of a Gaussian random force. In our simulations (in the ''no-equilibration'' microcanonical mode) there is no randomness except for numerical truncation errors and ''randomness'' due to chaos caused by the anharmonic terms. However, we find $\gamma\approx 1.5$ as the result of our quasiharmonic simulations at very low temperature (Fig. 12). As argued above, this actually corresponds to the low-anharmonicity limit, so is in agreement with the result for the Rouse model, provided the truncation errors and/or the chaotic ''randomness'' can be modeled as a Gaussian random force acting on the protein (this is a speculation). Increasing the degree of anharmonicity can only increase the fractal dimension of the signal, by the mechanisms discussed before. These arguments then show that the Rouse model result could thus serve as a *lower bound* to the γ values found in our simulations.

V. SUMMARY AND CONCLUSIONS

In this work we have studied the fractal properties of the total potential energy V as a function of time t for several realistic proteins. We performed simulations at a wide range of temperatures and a number of time scales. Our main conclusion is that the fractal behavior seems to be caused by the self-generated dynamical noise due to intermode coupling and equilibration promoted by anharmonic effects. There are universal aspects to this behavior, as exemplified by the similarity between totally different systems such as PPT and NaCl. Thus it cannot serve to characterize individual proteins and their energy landscape, but is rather a property common to all sufficiently large and anharmonic systems. We believe, however, that probing the fractal aspects of protein potential energy landscapes is a useful approach to the characterization of the ruggedness of these landscapes. To achieve this, it would be very interesting to examine V as a function of appropriate *spatial* coordinates, rather than t. The question remains of whether there are other ways leading to useful geometrical characterizations of the energy landscape, e.g., identification of multiple minima and maxima by real and imaginary frequencies [3].

Note added: While this paper was in the final stages of preparation we became aware of an article by García *et al.*

[19]. The authors concluded that the time dependence of the mean-square displacement of crambin around the crystal structure is described by an effective fractal exponent $\gamma \approx 1.60$ (for t long). This result was interpreted as indicative of multi-basin dynamics which would be typical in a rugged energy landscape [3].

ACKNOWLEDGMENTS

We would like to thank Dr. Susan K. Gregurick and Dr. Adrian Roitberg for invaluable technical assistance. D.A.L. would like to thank the Institute for Physical Science and Technology at the University of Maryland, College Park, for its hospitality during a visit in which part of this work was carried out. We are grateful to Dr. Angel García for bringing Ref. [19] to our attention. This work was supported in part by a grant from the National Science Foundation (Grant No. NSF-CHE96-29845) to D.T.

APPENDIX A: AN ANALYTICALLY SOLVABLE EXAMPLE: THE ROUSE MODEL

It is useful to have a simple reference model in the background to compare the numerical results of Sec. III with. To this end, we consider the Rouse model (RM), which is standard in polymer theory. It describes a polymer as a set of N beads located at $(\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_N)$ connected along a chain, where each bead (n) undergoes Brownian motion under the influence of random force $\mathbf{f}_n(t)$. It should be remarked that this is essentially different from our simulations, in which no randomness (except for numerical truncation errors and chaotic ''randomness'' due to anharmonicity) is present. Nevertheless, it useful to study the RM in order to see analytically how a fractal $V(t)$ can come about.

The excluded volume and hydrodynamic interactions are disregarded in the RM. The dynamics is generally described by the Langevin equation [20]

$$\frac{\partial}{\partial t}\mathbf{R}_n(t)=\sum_m \mathbf{H}_{nm}\cdot\left(\mathbf{f}_m(t)-\frac{\partial U}{\partial \mathbf{R}_m}\right)+\frac{1}{2}k_BT\sum_m \frac{\partial}{\partial \mathbf{R}_m}\cdot\mathbf{H}_{nm}. \tag{A1}$$

In the RM the mobility tensor is assumed to be isotropic,

$$\mathbf{H}_{nm}=\frac{\mathbf{I}}{\zeta}\delta_{nm}, \tag{A2}$$

and the interaction potential is taken to be harmonic,

$$U=\frac{k}{2}\sum_{n=2}^{N}(\mathbf{R}_n-\mathbf{R}_{n-1})^2, \tag{A3}$$

so that Eq. (A1) reduces to

$$\zeta\frac{d\mathbf{R}_n}{dt}=-k(2\mathbf{R}_n-\mathbf{R}_{n+1}-\mathbf{R}_{n-1})+\mathbf{f}_n \qquad (n=2,3,\ldots,N-1), \tag{A4}$$

$$\zeta\frac{d\mathbf{R}_1}{dt}=-k(\mathbf{R}_1-\mathbf{R}_2)+\mathbf{f}_1, \tag{A5}$$

$$\zeta\frac{d\mathbf{R}_N}{dt}=-k(\mathbf{R}_N-\mathbf{R}_{N-1})+\mathbf{f}_N. \tag{A6}$$

The force and the potential define a time scale through

$$\tau\equiv\frac{\zeta}{k}. \tag{A7}$$

It is convenient to consider the last set of equations in the limit where n is taken as a continuous variable:

$$\zeta\frac{\partial\mathbf{R}_n}{\partial t}=k\frac{\partial^2\mathbf{R}_n}{\partial n^2}+\mathbf{f}_N, \qquad \frac{\partial\mathbf{R}_{n=0}}{\partial n}=\frac{\partial\mathbf{R}_{n=N}}{\partial n}=0. \tag{A8}$$

The random forces are assumed to be Gaussian, i.e., δ-function uncorrelated:

$$\langle\mathbf{f}_n(t)\rangle=0, \tag{A9}$$

$$\langle\mathbf{f}_n(t)\mathbf{f}_m(t')\rangle=2\zeta k_BT\delta(n-m)\delta(t-t').$$

In the continuum limit the potential energy becomes

$$U=\frac{k}{2}\int_0^N dn\left(\frac{\partial\mathbf{R}_n}{\partial n}\right)^2. \tag{A10}$$

The last three sets of equations define the continuum RM.

1. Normal coordinates

We are interested in calculating the fractal scaling of the potential energy $U(t)$ defined in Eq. (A10). For this purpose it is convenient to transform to normal coordinates. It can be shown that in terms of the coordinates

$$\mathbf{X}_p\equiv\frac{1}{N}\int_0^N dn\cos(p\pi n/N)\mathbf{R}_n(t), \quad p=0,1,2,\ldots, \tag{A11}$$

the Langevin equation (A8) becomes

$$\zeta_p\frac{d\mathbf{X}_p}{dt}=-k_p\mathbf{X}_p+\mathbf{f}_p, \tag{A12}$$

where

$$\zeta_0=N\zeta, \tag{A13}$$

$$\zeta_p=2N\zeta \quad \text{for} \quad p=1,2,\ldots, \tag{A14}$$

$$k_p=\frac{2\pi^2kp^2}{N} \quad \text{for} \quad p=0,1,2,\ldots, \tag{A15}$$

and the forces satisfy

$$\langle \mathbf{f}_p(t)\rangle = 0, \tag{A16}$$

$$\langle \mathbf{f}_p(t)\mathbf{f}_q(t')\rangle = 2\zeta_p k_B T \delta pq\, \delta(t-t'). \tag{A17}$$

In this representation the random forces as well as the motions of the $\mathbf{X}_p$'s are independent, so that the motion of the polymer is decomposed into independent modes. Indeed, it is easily verified that the formal solution to Eq. (A12) is

$$\mathbf{X}_p(t) = \frac{1}{\zeta_p}\int_{-\infty}^{t} dt'\, e^{-(t-t')/\tau_p}\mathbf{f}_p(t'), \tag{A18}$$

where

$$\tau_p \equiv \frac{\tau_1}{p^2}, \quad \tau_1 \equiv \frac{\zeta_1}{k_1} = \frac{N^2}{\pi^2}\tau. \tag{A19}$$

Finally, the inverse transform of Eq. (A11) is

$$\mathbf{R}_n = \mathbf{X}_0 + 2\sum_{p=1}^{\infty} \mathbf{X}_p \cos(p\pi n/N). \tag{A20}$$

2. Scaling of the potential energy

We will now present a calculation of the fractal scaling of the potential energy in the RM, as defined in Eq. (11) [21]:

$$C_U(t) = \langle [U(t)-U(0)]^2\rangle. \tag{A21}$$

Inserting the expression for $\mathbf{R}_n$ into Eq. (A10) and expanding, yields

$$U = 2k\left(\frac{\pi}{N}\right)^2 \sum_{p,q=1}^{N} p\, q \mathbf{X}_p(t)\cdot\mathbf{X}_q(t) \times \int_0^N dn \sin(p\pi n/N)\sin(q\pi n/N). \tag{A22}$$

The integral evaluates to $(N/2)\delta_{pq}$, so that

$$U = k\frac{\pi^2}{N}\sum_{p=1}^{N} p^2 \mathbf{X}_p^2(t). \tag{A23}$$

Thus the fluctuation can be written as:

$$C_U(t) = \left(\frac{k\pi^2}{N}\right)^2 \left\langle \left[\sum_{p=1}^{N} p^2[\mathbf{X}_p^2(t)-\mathbf{X}_p^2(0)]\right]^2\right\rangle = \left(\frac{k\pi^2}{N}\right)^2 \sum_{p,q=1}^{N} p^2 q^2[\langle \mathbf{X}_p^2(t)\mathbf{X}_q^2(t)\rangle - 2\langle \mathbf{X}_p^2(t)\mathbf{X}_q^2(0)\rangle + \langle \mathbf{X}_p^2(0)\mathbf{X}_q^2(0)\rangle]. \tag{A24}$$

Inserting the formal solution for the normal coordinates [Eq. (A18)], the averages can be calculated from

$$\langle \mathbf{X}_p^2(t)\mathbf{X}_q^2(t')\rangle = \frac{1}{\zeta_p^2\zeta_q^2}\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t'} dt_3 \int_{-\infty}^{t'} dt_4 e^{-(2t-t_1-t_2)/\tau_p} e^{-(2t'-t_3-t_4)/\tau_q} \langle \mathbf{f}_p(t_1)\cdot\mathbf{f}_p(t_2)\, \mathbf{f}_q(t_3)\cdot\mathbf{f}_q(t_4)\rangle. \tag{A25}$$

To proceed, the averages over the forces must be evaluated. This can be done with the help of Wick's theorem [20]

$$\langle x_{n_1}x_{n_2}\cdots x_{n_{2p}}\rangle = \sum_{\text{all pairings}} \langle x_{m_1}x_{m_2}\rangle\cdots\langle x_{m_{2p-1}}x_{m_{2p}}\rangle, \tag{A26}$$

where the x_n's are Gaussian random variables. Employing the orthogonality relations from Eq. (A17) thus yields

$$\langle \mathbf{f}_p(t_1)\cdot\mathbf{f}_p(t_2)\mathbf{f}_q(t_3)\cdot\mathbf{f}_q(t_4)\rangle = (2k_BT)^2[\zeta_p\zeta_q\delta(t_1-t_2)\delta(t_3-t_4) + \zeta_p^2\delta_{pq}^2\delta(t_1-t_3)\delta(t_2-t_4) + \zeta_p^2\delta_{pq}^2\delta(t_1-t_4)\delta(t_2-t_3)]. \tag{A27}$$

In addition, we need the rule

$$\int_{-\infty}^{t} dt_2\, f(t-t_2)\delta(t_2-t_1) = f(t-t_1)\Theta(t-t_1), \tag{A28}$$

where $\Theta(x)$ is the Heaviside step function. Combining these results, the δ functions simplify the expression in Eq. (A25) as follows:

$$\langle \mathbf{X}_p^2(t)\mathbf{X}_q^2(t')\rangle = (2k_BT)^2\left[\frac{1}{\zeta_p\zeta_q}\int_{-\infty}^{t} dt_1\, e^{-(t-t_1)(1/\tau_p+1/\tau_1)}\int_{-\infty}^{t'} dt_3\, e^{-(t'-t_3)(1/\tau_p+1/\tau_1)} + 2\frac{\delta_{pq}}{\zeta_p^2}\left(\int_{-\infty}^{\min(t,t')} dt_1\, e^{-(t+t'-2t_1)/\tau_p}\right)^2\right]$$

$$= (2k_BT)^2\left[\left(\frac{\tau_p\tau_q}{\tau_p+\tau_q}\right)^2\frac{1}{\zeta_p\zeta_q} + 2\,\delta_{pq}\frac{\tau_p^2}{\zeta_p^2}e^{2[2\,\min(t,t')-t-t']/\tau_p}\right]. \tag{A29}$$

Inserting this into Eq. (A25) yields

$$C_U(t) = \left(\frac{k\pi^2}{N}\right)^2(2k_BT)^2\sum_{p,q=1}^{N} p^2\,q^2\left[\left(\frac{\tau_p\tau_q}{\tau_p+\tau_q}\right)^2\frac{1}{\zeta_p\zeta_q} + 4\frac{\tau_p^2}{\zeta_p^2}(1-e^{2t/\tau_p})\,\delta_{pq}\right]. \tag{A30}$$

Using the definitions of the various parameters appearing here, the final result becomes

$$C_U(t) = (k_BT)^2\left[\sum_{p,q=1}^{N}\frac{1}{p^2+q^2} + 4\sum_{p=1}^{N}(1-e^{-2\pi^2(t/\tau)(p^2/N^2)})\right]. \tag{A31}$$

The time dependence of $C_U(t)$ for large but finite N arises from the term $\Sigma_{p=1}^{N}\exp(-2\pi^2 tp^2/\tau N^2)$. To derive an expression for γ, it is useful to approximate the above sums by integrals. Defining $x=p/N,\ \ y=q/N$, we have

$$C_U(t) \approx (k_BT)^2\left[4N + \frac{1}{N^2}\int_{1/N}^{1} dx\int_{1/N}^{1} dy\,\frac{1}{x^2+y^2} - 4\int_{1/N}^{1} dx\, e^{-2\pi^2(t/\tau)x^2}\right]. \tag{A32}$$

The first integral is easily evaluated after a transformation to polar coordinates, and the second corresponds to the error function (in the second integral the error in taking $1/N=0$ is negligible)

$$C_U(t) \approx (k_BT)^2\left[4N + 2\pi\frac{\ln N}{N^2} - 2\sqrt{\pi}\frac{\mathrm{erf}\sqrt{\theta}}{\sqrt{\theta}}\right], \quad \theta \equiv 2\pi^2 t/\tau. \tag{A33}$$

In accordance with the preceding comments, we next take the $t \ll \tau$ limit. Then, using $\mathrm{erf}(x) \approx (2/\sqrt{\pi})[x - x^3/(3\times 1!) + x^5/(5\times 2!) - \cdots]$ to the first nontrivial (third) order, we can write

$$C_U(t\ll\tau) \approx (k_BT)^2\left[4N + 2\pi\frac{\ln N}{N^2} - 4\left(1-\frac{1}{3}\theta\right)\right]$$

$$\sim (k_BT)^2\frac{8\pi^2}{3}\frac{t}{\tau} \sim t^{2(2-\gamma)}, \tag{A34}$$

where in the last line the time-independent terms were ignored, and ultimately $\gamma = \frac{3}{2}$. Thus the Rouse model has a potential energy fluctuations which have a fractal dimension of $\frac{3}{2}$, just like a classical, fully correlated random walk.

APPENDIX B: ϵ-VARIATION METHOD

Dubuc *et al.* [18] introduced a method particularly well suited for the evaluation of the self-affinity exponent. They demonstrated that their method has the most stable *local* exponent in comparison to a variety of other methods, such as box counting, Minkowski-Bouligand, and power spectrum. All methods are equivalent in the continuous domain, but significant differences arise when they are applied to digitized data. Systematic errors which plague other methods are eliminated in the ϵ-variation method.

If the graph of a function $f(x)$ is fractal, there exists st least one part of the domain $[0,1]$ of f on which f is nowhere, or almost nowhere differentiable. If $p(x,x')$ is the slope of the line connecting $(x,f(x))$ and $(x',f(x'))$, fractality is related to the following property: $\sup p(x,x') \to_{x'\to x} \infty$. It is the *rate of growth* of this convergence (as x' tends to x) that determines the fractal dimension. In order to measure this behavior, Dubuc *et al.* introduced the ''ϵ variation'' of f:

$$V(\epsilon,f) = \int_0^1 v(x,\epsilon)dx, \tag{B1}$$

where the ''ϵ oscillation'' $v(x,\epsilon)$ is

$$v(x,\epsilon) = \sup_{x'\in R_\epsilon(x)} f(x') - \inf_{x'\in R_\epsilon(x)} f(x'), \tag{B2}$$

and where $R_\epsilon(x) = \{s\in[0,1]: |x-s| < \epsilon\}$.

Since f is continuous, $V(\epsilon,f) \to_{\epsilon\to 0} 0$. It is the rate of growth of $V(\epsilon,f)$ that yields the fractal dimension, or (more precisely) the self-affinity exponent γ. The corresponding log-log plot for the calculation of γ is $[\log(1/\epsilon), \log(V(\epsilon,f)/\epsilon^2)]$, with the exponent given by the 2-slope.

Following is a MATHEMATICA program which implements an algorithm to calculate the ϵ variation. The algorithm closely follows the implementation suggested in Ref. [18].

```
     (*******************)
     (* Initialization: *)
     (*******************)

     <<Statistics`LinearRegression`;

k[i\_] := i ;
(      * Change this to a more spaced function if necessary;
       the restriction is k[i]-k[i-1]<k[i-1]. *)

data = Table[ {j dx, f[j]}, {j,Length[f]} ] ;
      (* f is the function whose self-affinity exponent is being
        calculated; dx is the spacing *)

datalength = Length[ data ] ;

Clear[u,uu,uuu,b,bb,bbb] ;
        (* The functions "u" and "b" are respectively
            upper and lower bounds to f. *)

     (**************************)
     (* variation calculations: *)
     (**************************)

     Table[ u[1,n] = Max[ Table[ data[[j,2]], {j,n-k[1],n+k[1]} ] ],
                                        {n,1+k[1],datalength-k[1]} ] ;
     u[1,0] = u[1,1] = Max[ Table[ data[[j,2]], {j,1,1+k[1]} ] ] ;
     u[1,datalength+1] = u[1,datalength] = Max[ Table[ data[[j,2]],
                                {j,datalength-k[1],datalength} ] ] ;
     Table[ b[1,n] = Min[ Table[ data[[j,2]], {j,n-k[1],n+k[1]} ] ],
                                        {n,1+k[1],datalength-k[1]} ] ;
     b[1,0] = b[1,1] = Min[ Table[ data[[j,2]], {j,1,1+k[1]} ] ] ;
     b[1,datalength+1] = b[1,datalength] = Min[ Table[ data[[j,2]],
                                {j,datalength-k[1],datalength} ] ] ;

     imax = Floor[ datalength/5 ] ; (* total no. of iterations *)
     R = Floor[ datalength/2] ;

     (* calculating the epsilon-variation: *)
     (* boundary conditions: *)
     For[ n=0,n<=R+1,n++,
        uu[n] = u[1,n] ;
        bb[n] = b[1,n] ;
     ] ;
     sum[1] = Sum[ u[1,n]-b[1,n], {n,R} ] ;
     For[ i=2,i<=imax,i++,  (* iterations of the algorithm *)
         delta = k[i]-k[i-1] ;
         For[ n=1,n<=R,n++,  (* running over the data values *)
            uuu[n] = Max[ Table[ uu[m], {m,n-delta,n+delta} ]//N ] ;
            bbb[n] = Min[ Table[ bb[m], {m,n-delta,n+delta} ]//N ] ;
         ] ;
         sum[i] = Sum[ uuu[n]-bbb[n], {n,R} ] ;
         For[ n=1,n<=R,n++,  (* running over the data values *)
            uu[n] = uuu[n] ;
            bb[n] = bbb[n] ;
         ]
     ] ;

     (* total variation: *)
     res = Table[ {1/k[i], 1/k[i]\^2 sum[i]}, {i,imax} ]//N ;

     (* regression: *)
     reg = Regress[ Log[10, res ],{1,x},x] ;

     Print[ OutputForm[ reg ] ] ; (* slope is gamma *)
```

[1] H. Frauenfelder and P. Wolynes, Phys. Today **47**(2), 58 (1994).
[2] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**, 1619 (1995).
[3] J. E. Straub, A. B. Rashkin, and D. Thirumalai, J. Am. Chem. Soc. **116**, 2049 (1994).
[4] R. Czerminski and R. Elber, J. Chem. Phys. **92**, 5580 (1990).
[5] R. Elber, A. Roitberg, C. Simmerling, R. Goldstein, H. Li, G. Verkhivker, C. Kaeser, J. Zhang, and A. Utilsky, Comput. Phys. Commun. **91**, 159 (1995).
[6] K. D. Ball, R. S. Berry, R. E. Kunz, Li Feng-Yin, A. Proykova, and D. J. Wales, Science **271**, 963 (1996).
[7] It is clear that the upper bound for $\gamma(\tau\rightarrow\infty)$ is 2, so Eq. (13) can be valid only for short τ.
[8] *Dynamics of Fractal Surfaces*, edited by F. Family and T. Vicsek (World Scientific, Singapore, 1991).
[9] J. C. Russ, *Fractal Surfaces* (Plenum, New York, 1994).
[10] *On Growth and Form*, Vol. 100 of *NATO Advanced Study Institute, Series E: Applied Sciences*, edited by H. E. Stanley and N. Ostrowsky (Martinus Nijhoff, Dordrecht, 1986).
[11] D. Thirumalai and S.A. Woodson, Acc. Chem. Res. **29**, 433 (1996).
[12] *The Fractal Approach to Heterogeneous Chemistry: Surfaces, Colloids, Polymers*, edited by D. Avnir (Wiley, Chichester, 1992).
[13] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
[14] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications* (Wiley, Chichester, 1990).
[15] P. Pfeifer and M. Obert, in *The Fractal Approach to Heterogeneous Chemistry: Surfaces, Colloids, Polymers* (Ref. [12]).
[16] A.-L. Barabási and H. E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995).
[17] D. A. Hamburger-Lidar, Phys. Rev. E **54**, 354 (1996).
[18] B. Dubuc, J.F. Quiniou, C. Roques-Carmes, C. Tricot, and S. W. Zucker, Phys. Rev. A **39**, 1500 (1989).
[19] A. E. García, R. Blumenfeld, G. Hummer, and J. A. Krumhansl, Physica D **107**, 225 (1997).
[20] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, New York, 1986), p. 91.
[21] It may be tempting to consider this fluctuation in the potential energy as related to the specific heat C_V. However, that is actually quite a different fluctuation: $k_B T^2 C_V = \langle (E - \langle E \rangle)^2 \rangle$, where E is the *total* energy, and the fluctuation is time independent, near equilibrium.