OFFPRINT

Energy transfer in nonlinear network models of proteins

F. PIAZZA and Y.-H. SANEJOUAND

Please visit the new website
`www.epljournal.org`

TARGET YOUR RESEARCH WITH EPL

Sign up to receive the free EPL table of contents alert.

www.epljournal.org/alerts

Energy transfer in nonlinear network models of proteins

F. PIAZZA[1](a) and Y.-H. SANEJOUAND[2](b)

[1] *Ecole Polytechnique Fédérale de Lausanne, Laboratoire de Biophysique Statistique, ITP-SB - BSP-720, CH-1015 Lausanne, Switzerland*
[2] *Laboratoire U3B, UMR 6204 du CNRS, Faculté des Sciences et des Techniques - 2, rue de la Houssinière, 44322 Nantes Cedex 3, France, EU*

Abstract – We investigate how nonlinearity and topological disorder affect the energy relaxation of local kicks in coarse-grained network models of proteins. We find that nonlinearity promotes long-range, coherent transfer of substantial energy to specific functional sites, while depressing transfer to generic locations. In some cases, transfer is mediated by the self-localization of discrete breathers at distant locations from the kick, acting as efficient energy-accumulating centers.

It is now well established that the functional dynamics of proteins is deeply rooted in the peculiar topological arrangement of their native folds, as revealed by many experimental and computational studies [1,2]. In particular, the success of coarse-grained elastic network models (ENMs) in describing atomic fluctuations at room temperature have helped elucidate, at the harmonic level, the subtle interplay between structure and dynamics on one side and biological function on the other [3–10].

However, protein dynamics is strongly anharmonic [11,12], a property which has to be taken into account in order to rationalize crucial biological processes such as energy storage and transfer upon ligand binding, chemical reaction, etc. [13,14]. Yet, even though many theoretical studies suggest that nonlinear excitations may play an active role in protein functioning [15–17], the rich phenomenology residing in the interplay between protein topology and nonlinearity still remains widely unexplored. Along these lines, we have recently introduced the Nonlinear Network Model (NNM), showing how known nonlinear effects can be modulated by the underlying non-regular topology of protein systems. For instance, within a large collection of enzyme structures, the formation of localized, robust nonlinear modes appears strongly favored at few specific sites, that often lie in close proximity of known catalytic sites [18,19].

In this paper we examine the effects of the nonlinearity/topology interplay on energy transfer phenomena across protein structures. Within the NNM framework a protein is represented by N fictive particles (amino acids) of identical mass $M = 110$ a.m.u., at equilibrium at the corresponding C_α site as specified in the experimentally determined structures (X-ray or NMR). By imposing a fixed cutoff R_c on the latter set of coordinates, a protein is mapped onto a network of nonlinear oscillators, whose potential energy reads

$$U = \sum_{i=1}^{N} u_i \stackrel{\text{def}}{=} \sum_{i=1}^{N} \left[\sum_{j=1}^{N} c_{ij} \sum_{p=2,4} \frac{k_p}{2p} (r_{ij} - R_{ij})^p \right], \quad (1)$$

where r_{ij} is the distance between residues i and j, R_{ij} their distance in the equilibrium structure and $c_{ij} = \{1 \text{ if } R_{ij} \leqslant R_c, 0 \text{ otherwise}\}$ is the connectivity matrix. As in previous studies [18], we take $R_c = 10\,\text{Å}$, $k_4 = 5\,\text{kcal/mol/Å}^4$ and fix k_2 so that the low-frequency part of the linear spectrum match actual protein frequencies, as calculated through realistic force fields [20–22]. This gives $k_2 = 5\,\text{kcal/mol/Å}^2$. The case $k_4 = 0$ corresponds to the Anisotropic Network Model (ANM) [3–5].

Our aim is to investigate how energy initially imparted at a specified site i redistributes across a given structure. To do this, we perform microcanonical simulations with all residues initially at rest at their equilibrium position but for a kinetic energy kick at site i of magnitude E_0. Sites in a 3D protein network are not equivalent, featuring

(a)E-mail: Francesco.Piazza@epfl.ch
(b)E-mail: Yves-Henri.Sanejouand@univ-nantes.fr

e.g. varying connectivity, clustering coefficient and bond directions. Thus, in order to allow for a comparison of energy relaxation from all sites in a given structure, the initial kick direction ought to be specified by a unique protocol. We chose to calculate the directions of the initial velocities $\vec{v}_i(0)$ through the Sequential-Maximum-Strain (SMS) algorithm [19], which provides an unbiased measure of the maximum-strain direction at site i for a fixed displacement (here $1\,\text{Å}$), $\hat{e}_{SMS}$. During the simulation, we record at regular intervals t_k, $k = 1, 2, \ldots, N_s$ the site that carries the highest energy, n_k, and the value of the latter, $e_{n_k} = Mv_{n_k}^2/2 + u_{n_k}$. Corresponding to a fixed simulation time (about $500\,\text{ps}$), we define a transfer probability from site i (the kicked one) to site j and the fraction of energy transferred as

$$\mathcal{P}_{i\rightarrow j} = \frac{1}{N_s}\sum_{k=1}^{N_s}\delta_{n_k,j}, \quad f_{i\rightarrow j}^e = \frac{1}{N_s\mathcal{P}_{i\rightarrow j}}\sum_{k=1}^{N_s}\frac{e_j\delta_{n_k,j}}{E_0}. \quad (2)$$

The first striking result comes from the calculation of *average* transfer probabilities. These gauge the mean transfer to a given site from kicks at all other sites, $\langle \mathcal{P}_j \rangle = \sum_{i\neq j}\mathcal{P}_{i\rightarrow j}/(N-1)$, obtained from N independent simulations. A typical probability transfer plot is shown in fig. 1. The first notable feature is that the effect of nonlinearity is to substantially increase the probability of energy funneling to a few selected sites, while depressing transfer to all other locations with respect to the harmonic (ANM) case. Remarkably, the preferred target sites lie in close proximity to the known catalytic sites, within the stiffest regions[1]. Thus, topology and nonlinearity team in this case together to sharpen energy funneling to specific functional regions.

The case shown in fig. 1 is not a singular one. In fig. 2 we show the stiffness patterns for four other enzymes along with the sites ranking first to tenth as to the energy delivered on average to spherical shells with $6\,\text{Å}$ radius around each site. For residue j, this amounts to further averaging the mean energy deposited at sites within the j-th ball $\mathcal{B}(j)$, *i.e.* $\langle\langle f_j^e \rangle\rangle = \langle \sum_{i\neq j} f_{i\rightarrow j}^e/(N-1)\rangle_{\mathcal{B}(j)}$. As it shows, the sites around which most of the energy is deposited invariably spotlight the stiffest regions, at the same time identifying functionally relevant locations (see catalytic sites). Moreover, the same locations clearly attract substantial fractions of the initial excitation energy, as revealed by surveying the maximum transferred energies to each ball $\mathcal{B}(j)$, that is $f_{\max}(j) = \langle\max_{i\neq j} f_{i\rightarrow j}^e\rangle_{\mathcal{B}(j)}$ (empty circles). Many events featuring transfers of energy fractions in the range 20 to 25% were indeed observed.

We can learn more on the mechanisms underlying the energy transfer process by examining in detail the outcome of a single kick. Figure 3 pictures a long-range

---

[1]We measure the local stiffness s_i as a sum over the set $\mathcal{S}$ of ten highest normal modes $\vec{\xi}^k$, that is the eigenvectors of the Hessian of the ANM total potential energy, $s_i = \sum_{j,\alpha}\sum_{k\in\mathcal{S}} c_{ij}[\xi_{j\alpha}^k]^2/\mathcal{N}_i$, with $\mathcal{N}_i = \sum_j c_{ij}$ [18].
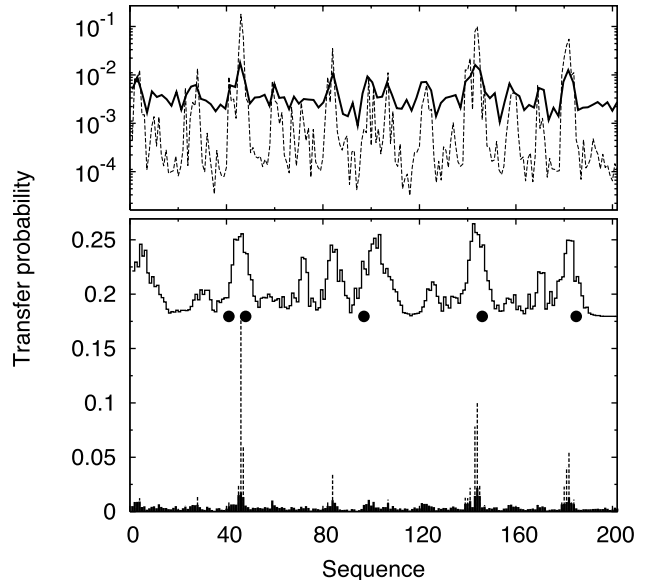

Fig. 1: Average transfer probability in Riboflavin Synthase (PDB id. 1KZL, $N = 202$) on logarithmic (upper panel) and linear (lower panel) scale. Thick solid line: ANM. Dashed line: NNM. The staircase plot in the lower panel reproduces the stiffness pattern (arbitrary units). Filled circles flag catalytic sites. $E_0 = 75\,\text{kcal/mol}$.

transfer event occurring when kicking at site LEU 42 in the enzyme Subtilisin. The middle-lower panel (a) shows a plot of the most energetic site n_k as a function of time, clearly illustrating the transfer to site VAL 177, some $23\,\text{Å}$ away, occurring at $t_* \approx 275\,\text{ps}$. The transfer process also involves site ALA 85, as a *passage* site. Remarkably, a plot of the energy $e_{\max}(t)$ of the most energetic site at time t clearly shows that such passage coincides with a redistribution of energy across the structure (see middle panel (c)). Subsequently, energy is garnered from the neighborhood and stabilized in a localized mode centered at VAL 177, finally carrying about 20% of the total energy. This marks the true transfer event. Such energy-harvesting, self-localized vibrations are generic in discrete nonlinear systems and are well known as Discrete Breathers (DB) [23]. These are robust, time-periodic exponentially localized modes, whose vibrational frequency lies outside the linear spectrum of the system. In the context of the NNM, we have shown how accurate approximations of such periodic orbits can be calculated analytically, reproducing the marked affinity of DB self-localization in topologically disordered media for the stiffest spots [18,19]. Here we have shown that DBs may also be excited as a consequence of localized impulses at considerable distances from the excitation, playing the role of energy-accumulating transfer vectors. Incidentally, this is why in fig. 2 we measure the energies delivered to local spherical neighborhoods around target sites.

In order to substantiate the above interpretation, we have performed Principal Component Analysis (PCA) on
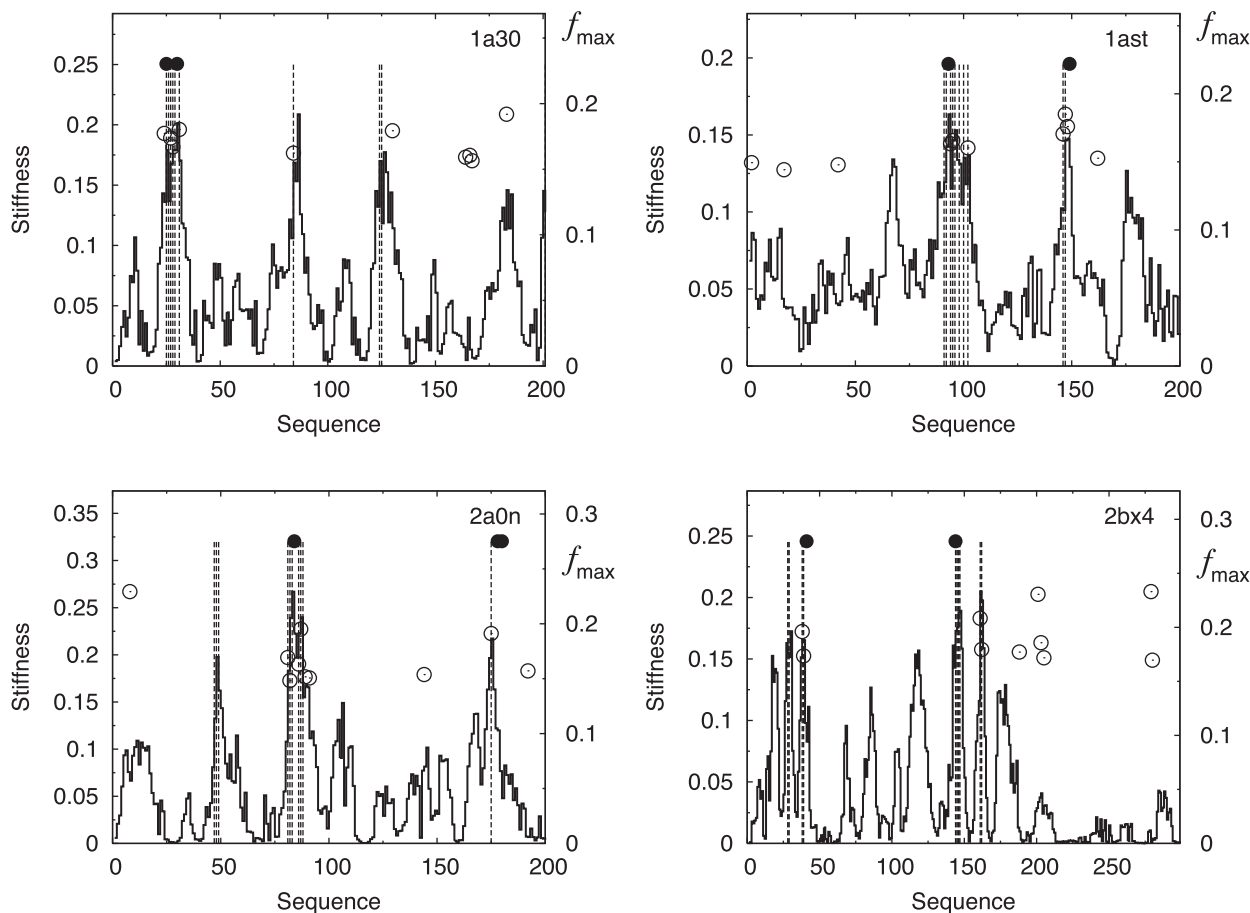
Fig. 2: Stiffness plots (staircases) and first ten sites in the ranked list of average energies delivered to 6 Å-balls around each site (dashed impulses) for four enzymes: HIV-I protease (PDB id. 1A30, $N = 201$), Astacin (PDB id. 1AST, $N = 200$), Imidazole glycerol phosphate synthase subunit hisF (PDB id. 2A0N, $N = 200$), SARS coronavirus main proteinase (PDB id. 2BX4, $N = 299$). Filled circles flag catalytic sites. Right axes report the ten largest average fractions of the initial energy transferred to 6 Å balls (empty circles). The abscissas denote the centers of the 6 Å-radius spheres. $E_0 = 75$ kcal/mol.

an extended portion of the post-transfer dynamics. The power spectrum of the system trajectory projected on the first principal mode (PM1) is shown in the upper left panel of fig. 3, clearly revealing the existence of a nonlinear, time-periodic excitation, reminiscent of chaotic DBs that self-localize through modulational instability in nonlinear lattices [24].

More insight as to why energy is transferred from LEU 42 a long distance away can be obtained by turning to the space of Normal Modes (NM) $\vec{\xi}^k$ ($k = 1, 2, \ldots, 3N - 6$). Middle panel (b) of fig. 3 reports the mode carrying the highest energy $\varepsilon_k(t) = (\dot{Q}_k^2 + \omega_k^2 Q_k^2)/2$ at time t, where $Q_k = \sum_{j,\alpha} x_{j\alpha} \xi_{j\alpha}^k$ is the NM-transform of the system coordinates $x_{j\alpha}$ ($j = 1, 2, \ldots, N, \alpha = x, y, z$) and ω_k are the NM frequencies. Before the transfer event, energy is bounced among four high-frequency modes, NMs 1, 3, 8 and 9. This can be understood by constructing the *NM overlap network* starting from the NMs with the largest projections on the initial excitation unit vector $\hat{e}_{SMS}$ (four modes in the first column of the bottom network in

fig. 3, NMs 138, 127, 67 and 37, making up about 60% of $\hat{e}_{SMS}$). For a given NM p, two links are drawn to the two-highest ranking NMs in the ordered list of absolute overlap coefficients $t_{pq} = \sum_{j,\alpha} |\xi_{j\alpha}^p| |\xi_{j\alpha}^q|$. By doing this for the four NMs involved in the SMS vector, a closed network emerges identifying the NM3 $\rightleftarrows$ NM8 $\rightleftarrows$ NM9 loop. Thus, in the presence of nonlinearity energy is immediately directed to a reduced group of NMs via resonant overlap mechanisms. This finding agrees with results of atomistic simulations highlighting the importance of spatial overlap for NM-NM energy transfer [25].

High-frequency NMs are strongly localized in space. In particular, ALA 85 is the NM site (the site with largest displacement) in NM3 and the second NM site in NM8, which explains the role of ALA 85 in the energy circulation process. Before transfer, however, energy also bounces back and forth from NM1, the highest-frequency mode, reflecting the nonlinear frequency shift on NM3 toward greater frequencies (see again fig. 3(b)). At $t = t_*$ energy starts departing the region around LEU 42 and a
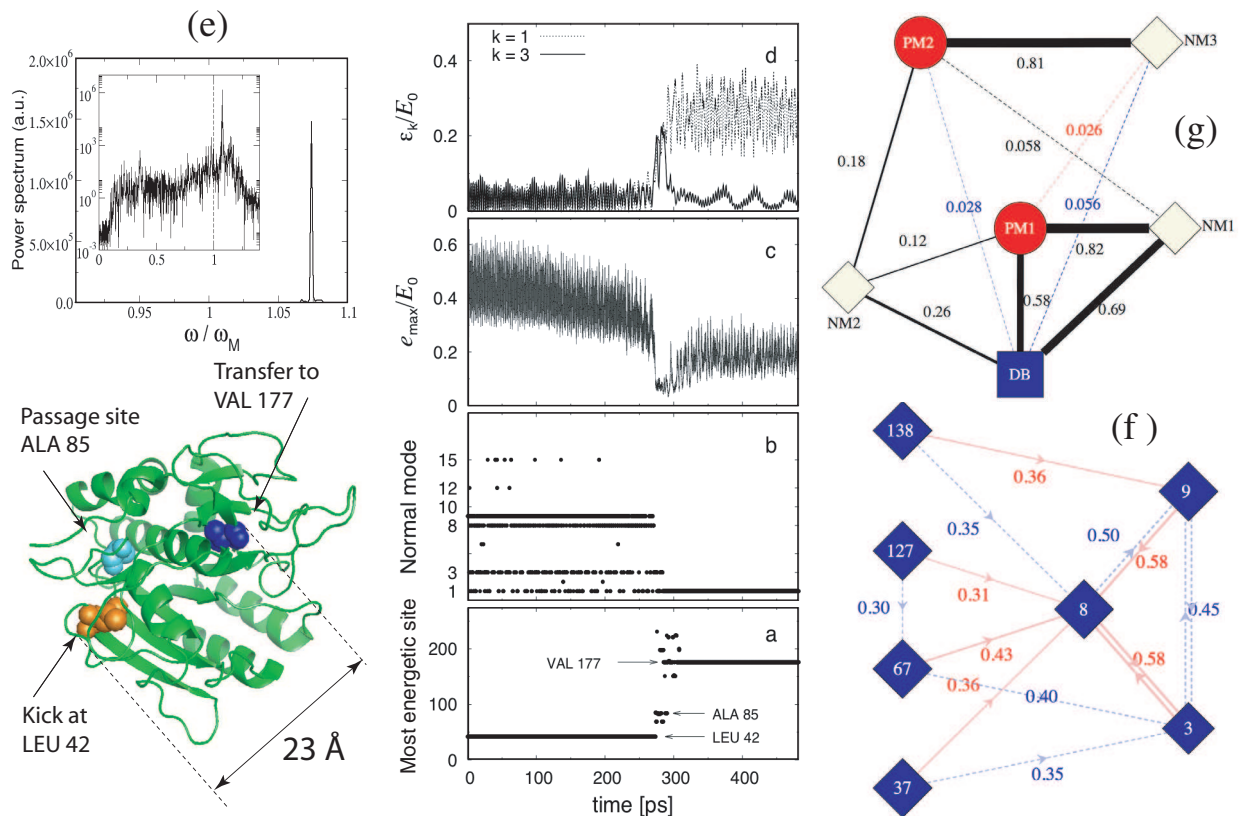
Fig. 3: (Colour on-line) Kick at site LEU 42 in Subtilisin (PDB id. 1AV7, $N = 274$). Middle panels: most energetic site *vs.* time (a), most energetic normal mode *vs.* time —NMs being ranked in order of decreasing eigenfrequency— (b), highest site energy *vs.* time (c), energies of two NMs *vs.* time (d). Power spectrum of the system trajectory for $t > t_* = 275$ ps projected on the first principal mode; $\omega_M = 101.2\,\mathrm{cm}^{-1}$ is the band-edge linear frequency (e). The inset shows the same plot in logarithmic scale. (f) NM overlap network. Nodes are NMs, red and blue links connect to nodes ranking first and second, respectively, in overlap (see text). Link weights are the overlap coefficients t_{pq}. (g) Network relating principal modes (PM), NMs and the analytical Discrete Breather pattern (DB). Link weights are the absolute cosines (normalized scalar products). In both graphs the link width is proportional to its weight. $E_0 = 100\,\mathrm{kcal/mol}$.

fluctuation pumping up NM3 occurs (fig. 3(d)), shifting its frequency upwards by virtue of nonlinearity. The energy at stake is sufficient to trigger nonlinear localization and a DB finally installs at VAL 177, the NM site of NM1, gathering vibrational energy from the background. Correspondingly, the energy on NM1 increases (see fig. 3(d)). To substantiate the above analysis, we have calculated analytically the DB mode pattern centered at site VAL 177 with the technique described in ref. [19]. Then we have built the network connecting the first two principal modes, the first three NMs and the DB, where the links are weighted by the normalized scalar products (upper network in fig. 3). As it shows, the PMs essentially reflect the underlying competition between NM1 and NM3. In particular, the first principal mode confirms the excitation of a DB emerging as a nonlinear continuation of the edge normal mode, as predicted theoretically in ref. [19]. In agreement with this picture, kicks at ALA 42 of weaker energy resulted in

a DB installing at MET 199, the NM site of NM2. That is, less energy causes a smaller frequency shift and the DB branch originating from the continuation of NM2 is excited instead. Reducing E_0 further, the transfer is observed to halt at ALA 85, as explained by the NM overlap network.

In this paper we have shown how nonlinearity in a topologically non-regular system boosts energy transfer to few specific locations. In enzyme structures, these coincide invariably with the stiffest regions, also hosting the functionally relevant sites. Nonlinearity sharpens the transfer selectivity, by reducing at the same time the transfer probability to generic locations. The energy transferred by virtue of nonlinearity may be a conspicuous portion of the initial excitation, in which cases localized vibrations akin to Discrete Breathers self-localize as energy-collecting centers, often realizing amazingly efficient energy transfer channels across considerable distances.

REFERENCES

[1] BEST R. B., MERCHANT K. A., LOUIS J. M., GOPICH I. V. and EATON W. A., *Structure and dynamics of unfolded proteins from single molecule FRET spectroscopy and molecular simulation, Biophys. J.*, 534A-534A (2007) Meeting Abstract.

[2] TORCHIA D. A. and ISHIMA R., *Pure Appl. Chem.*, **75** (2003) 1371.

[3] TIRION M. M., *Phys. Rev. Lett.*, **77** (1996).

[4] BAHAR I., ATILGAN A. R. and ERMAN B., *Fold. Des.*, **2** (1997) 173.

[5] HINSEN K., *Proteins*, **33** (1998) 417.

[6] TAMA F. and SANEJOUAND Y.-H., *Protein Eng.*, **14** (2001) 1.

[7] MICHELETTI C., LATTANZI G. and MARITAN A., *J. Mol. Biol.*, **321** (2002) 909.

[8] DELARUE M. and SANEJOUAND Y.-H., *J. Mol. Biol.*, **320** (2002) 1011.

[9] KREBS W. G., ALEXANDROV V., WILSON C. A., ECHOLS N., YU H. and GERSTEIN M., *Proteins*, **48** (2002) 682.

[10] KONDRASHOV D., VAN WYNSBERGHE A., BANNEN R., CUI Q. and PHILLIPS G., *Structure*, **15** (2007) 169.

[11] XIE A. H., VAN DER MEER L., HOFF W. and AUSTIN R. H., *Phys. Rev. Lett.*, **84** (2000) 5435.

[12] EDLER J., PFISTER R., POUTHIER V., FALVO C. and HAMM P., *Phys. Rev. Lett.*, **93** (2004) 106405.

[13] SAGNELLA D., STRAUB J. and THIRUMALAI D., *J. Chem. Phys.*, **113** (2000) 7702.

[14] LEITNER D. M., *Annu. Rev. Phys. Chem.*, **59** (2008) 233.

[15] KOPIDAKIS G., AUBRY S. and TSIRONIS G. P., *Phys. Rev. Lett.*, **87** (2001) 165501.

[16] ARCHILLA J. F. R., GAIDIDEI Y. B., CHRISTIANSEN P. L. and CUEVAS J., *J. Phys. A: Math. Gen.*, **35** (2002) 8885.

[17] CUEVAS J., STARIKOV E. B., ARCHILLA J. F. R. and HENNIG D., *Mod. Phys. Lett. B*, **18** (2004) 1319.

[18] JUANICO B., SANEJOUAND Y.-H., PIAZZA F. and DE LOS RIOS P., *Phys. Rev. Lett.*, **99** (2007) 238104.

[19] PIAZZA F. and SANEJOUAND Y.-H., *Phys. Biol.*, **5** (2008) 026001.

[20] BROOKS B. R. and KARPLUS M., *Proc. Natl. Acad. Sci. U.S.A.*, **82** (1985) 4995.

[21] MARQUES O. and SANEJOUAND Y.-H., *Proteins*, **23** (1995) 557.

[22] PERAHIA D. and MOUAWAD L., *Comput. Chem.*, **19** (1995) 241.

[23] FLACH S. and WILLIS C. R., *Phys. Rep.*, **295** (1998) 181.

[24] CRETEGNY T., DAUXOIS T., RUFFO S. and TORCINI A., *Physica D: Nonlinear Phenom.*, **121** (1998) 109.

[25] MORITSUGU K., MIYASHITA O. and KIDERA A., *Phys. Rev. Lett.*, **85** (2000).