Discrete breathers in a realistic coarse-grained model of proteins

2011 Phys. Biol. 8 046008

(http://iopscience.iop.org/1478-3975/8/4/046008)

Phys. Biol. **8** (2011) 046008 (13pp) doi:10.1088/1478-3975/8/4/046008

Discrete breathers in a realistic coarse-grained model of proteins

Stefano Luccioli[1,2,3], **Alberto Imparato**[4], **Stefano Lepri**[1,2], **Francesco Piazza**[5] **and Alessandro Torcini**[1,2,3,4]

[1] CNR—Consiglio Nazionale delle Ricerche, Istituto dei Sistemi Complessi, via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy
[2] Centro Interdipartimentale per lo Studio delle Dinamiche Complesse, via G Sansone 1, I-50019 Sesto Fiorentino, Italy
[3] Istituto Nazionale di Fisica Nucleare, Sezione Firenze, via G Sansone 1, I-50019 Sesto Fiorentino, Italy
[4] Department of Physics and Astronomy, University of Aarhus, Ny Munkegade, Building 1520, DK-8000 Aarhus C, Denmark
[5] Université d'Orléans and Centre de Biophysique Moléculaire (CBM-CNRS), Rue Charles Sadron, 45071 Orléans Cedex, France

E-mail: stefano.luccioli@fi.isc.cnr.it

Received 14 February 2011
Accepted for publication 17 May 2011
Published 14 June 2011
Online at stacks.iop.org/PhysBio/8/046008

Abstract
We report the results of molecular dynamics simulations of an off-lattice protein model featuring a physical force-field and amino-acid sequence. We show that localized modes of nonlinear origin, discrete breathers (DBs), emerge naturally as continuations of a subset of high-frequency normal modes residing at specific sites dictated by the native fold. DBs are time-periodic, space-localized vibrational modes that exist generically in nonlinear discrete systems and are known for their resilience and ability to concentrate energy for long times. In the case of the small β-barrel structure that we consider, DB-mediated localization occurs on the turns connecting the strands. At high energies, DBs stabilize the structure by concentrating energy on a few sites, while their collapse marks the onset of large-amplitude fluctuations of the protein. Furthermore, we show how breathers develop as energy-accumulating centres following perturbations even at distant locations, thus mediating efficient and irreversible energy transfers. Remarkably, due to the presence of angular potentials, the breather induces a local static distortion of the native fold. Altogether, the combination of these two nonlinear effects may provide a ready means for remotely controlling local conformational changes in proteins.

1. Introduction

Biopolymers such as proteins and nucleic acids fold into complex three-dimensional structures, whose shape is strictly connected to their biological function [1]. The conformation of such molecules can change dynamically, in turn modulating the function: for example, activation or inactivation of enzymes relies on specific structural modifications occurring at specific locations [2–4]. Typically, such changes are driven by either mechanical forces or by converting chemical energy into conformational rearrangements and thus into mechanical work.

Proteins under physiological conditions are immersed in a thermal bath and therefore exhibit random thermal fluctuations. However, the biological function of a given biopolymer is often closely related to a particular kind of motion, typically involving large-scale vibrations [5–9] or the fluctuations of entire hinge-domain units [10–12].

If collective, low-frequency modes have been traditionally assumed to describe functional patterns, there is growing evidence that high-frequency vibrations contain information on protein stability [13]. Importantly, fast modes are strongly localized, due to the geometric heterogeneity of protein structures. Typically, such vibrations are localized at extremely stiff regions, such as hinges, which in turn assume a

prominent role in regulating protein stability and function, as also suggested by experiments [4]. It is also known that active sites of enzymes have a marked tendency to be located within the stiffest segments of the structures [14, 15], which provides a further intriguing motivation to investigate the connection between localized vibrations and the peculiarities of protein folds.

When nonlinear effects are considered, the connection between the dynamics of localized modes and the details of the scaffolds becomes more interesting. Recently, it has been shown that high-frequency normal modes (NMs) can be continued to nonlinear localized vibrations of high energy termed discrete breathers (DBs) [14], also known as intrinsic localized modes [16]. DBs are time-periodic, spatially localized vibrations that emerge generically in many-body nonlinear systems [17, 18]. Despite being localized in space, DBs are collective modes. Therefore, in a DB solution, all particles vibrate at the same frequency, which is non-resonant with any of the NM frequencies. Also in view of this, DBs are generally rather resilient to perturbations, including thermal noise [17]. This means that energy can be stored in a system through the excitation of a DB for longer times compared to its period. Remarkably, DB-like vibrations have been also observed experimentally in many systems [18–22].

While DBs have been widely studied in spatially homogeneous systems [18], the role played by spatial inhomogeneity on their properties remains largely unknown. However, the dynamical behaviour of DBs discovered so far in protein models is remarkable. In the context of the nonlinear network model (NNM), it has been shown that DBs emerge spontaneously, upon surface cooling, at a few specific sites, invariably within the stiffest regions [23]. Moreover, they are also able to self-excite at a target site upon injecting some energy at a different location, thus mediating high-yield energy transfer events [24–26]. In a more simplified model, it has been shown that DB excitation lowers the free-energy barrier associated with a given enzyme-catalyzed reaction [27], thus confirming the role of protein dynamics in reaction-rate enhancement by enzymes highlighted by recent experiments [28]. Furthermore, DBs have been shown to enhance noise-free escape over a barrier for a chain of coupled nonlinear oscillators [29].

However, DBs in proteins have only been found and characterized so far in extremely simplified models, either taking into account the three-dimensional folds but with no heterogeneity in the force constants [14], or with slightly more elaborate potentials but imposing crudely simplified geometries [30–32]. Importantly, all studies performed so far notably lacked (i) realistic inter-particle potentials and (ii) an explicit account of the amino acid sequence. In this paper, we make a step forward by investigating DB excitation and their properties in a realistic off-lattice model of protein dynamics with coarse-grained realistic interaction potentials and a three-code amino-acid sequence.

The paper is organized as follows. In the following section, we introduce our system and describe the model, along with an account of our simulation and analysis protocols. In section 3, we describe the emergence of DBs as the numerical continuation of NMs and provide a thorough characterization of their properties. Finally, we summarize our findings and outline possible future directions prompted by our results.

2. Model and numerical methods

The 3D off-lattice protein model studied in this paper is a modified version of the one initially introduced in [33] and successively generalized to include a harmonic interaction between next-neighbouring beads instead of rigid bonds [34]. This model has been studied to describe thermally-driven folding and unfolding [33–40] and, more recently, to reproduce mechanical manipulation experiments [41–46]. It consists of a chain of L point-like monomers mimicking the residues of a polypeptidic chain with an associated aminoacid sequence coded by a three-letter alphabet: hydrophobic (B), polar (P) and neutral (N). The intramolecular potential consists of four terms: a stiff nearest-neighbour harmonic potential, V_H, intended to maintain the bond distance almost constant, a three-body interaction V_A, which accounts for the energy associated with bond angles, a four-body interaction V_D corresponding to the dihedral angle potential and a long-range Lennard–Jones (LJ) interaction, V_{LJ}, acting on all pairs i, j such that $|i - j| > 2$, namely

$$V_H(r_{i,i+1}) = \alpha(r_{i,i+1} - r_0)^2 \tag{1}$$

$$V_A(\theta_i) = A\cos(\theta_i) + B\cos(2\theta_i) - V_0 \tag{2}$$

$$V_D(\varphi_i, \theta_i, \theta_{i+1}) = C_i[1 - S(\theta_i)S(\theta_{i+1})\cos(\varphi_i)] + D_i[1 - S(\theta_i)S(\theta_{i+1})\cos(3\varphi_i)] \tag{3}$$

$$V_{\mathrm{LJ}}(r_{i,j}) = \varepsilon_{i,j}\left(\frac{1}{r_{i,j}^{12}} - \frac{c_{i,j}}{r_{i,j}^{6}}\right). \tag{4}$$

Here, $r_{i,j}$ is the distance between the ith and the jth monomer, while θ_i and φ_i are the bond and dihedral angles at the ith monomer, respectively. The parameters α and $r_0 = 1$ fix the strength and the equilibrium distance between consecutive monomers along the backbone, respectively. The term $V_A(\theta_i)$ corresponds, up to the second order, to a harmonic term $k_\theta(\theta_i - \theta_0)^2/2$, where

$$A = -k_\theta\frac{\cos\theta_0}{\sin^2\theta_0}, \qquad B = \frac{k_\theta}{4\sin^2\theta_0},$$
$$V_0 = A\cos\theta_0 + B\cos(2\theta_0),$$

with $k_\theta = 20$, $\theta_0 = 5\pi/12$ [47].

The dihedral-angle potential is characterized by three minima for $\varphi = 0$ (associated with the so-called trans state) and $\varphi = \pm 2\pi/3$ (corresponding to *cis* states). This term is mainly responsible for the formation of secondary structures. In particular, large values of the parameters C_i, D_i favour the formation of the trans state and therefore of β-sheets, while when cis states prevail, α-helices are formed. The parameters (C_i, D_i) have been chosen as in [37], i.e. if two or more beads among the four defining φ are neutral (N), then $C_i = 0$ and $D_i = 0.2$; in all the other cases, $C_i = D_i = 1.2$. The

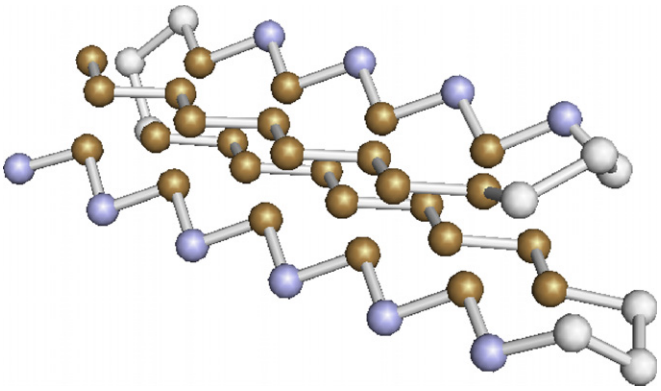


Figure 1. NC of the protein model. Different colours represent the three different types of beads: white (neutral), light blue (polar) and sand (hydrophobic). Parameters are $L = 46$, $\alpha = 1000$.

tapering function $S(\theta_i) = 1 - \cos^{32}(\theta_i)$ has been introduced in the expression of V_D in order to smooth off the singularities introduced by the derivatives of the dihedral angle cosines (for an exact definition, see [44, 48]).

The LJ term $V_{\rm LJ}$ describes the interactions with the solvent, which depend on the nature of the interacting residues as follows: if any of the two monomers is neutral, the potential is repulsive $c_{N,X} = 0$ and its energy scale is fixed by $\varepsilon_{N,X} = 4$; for interactions between hydrophobic residues, $c_{B,B} = 1$ and $\varepsilon_{B,B} = 4$; for any polar–polar or polar–hydrophobic interaction, $c_{P,P} \equiv c_{P,B} = -1$ and $\varepsilon_{P,P} \equiv \varepsilon_{P,B} = 8/3$.

According to the above definitions, the Hamiltonian of the system reads

$$H = K + V = \sum_{i=1}^{L} \frac{|\vec{p}(i)|^2}{2} + \sum_{i=1}^{L-1} V_H(r_{i,i+1}) + \sum_{i=2}^{L-1} V_A(\theta_i) + \sum_{i=2}^{L-2} V_D(\varphi_i, \theta_i, \theta_{i+1}) + \sum_{i=1}^{L-3} \sum_{j=i+3}^{L} V_{\rm LJ}(r_{ij}) \qquad (5)$$

where all monomers are assumed to have the same unitary mass.

In this paper, we consider the following sequence of $L = 46$ monomers: $B_9N_3(PB)_4N_3B_9N_3(PB)_5P$. This sequence has been widely analysed in the past for thermal folding [33–40] as well as for mechanically induced unfolding and refolding [41–44]. Here, we adopt the same potential and parameters as in [37, 43, 44], except for a stiffer harmonic constant α ($50 \leqslant \alpha \leqslant 1000$). The simulations reported henceforth refer to $\alpha = 1000$, except when otherwise indicated. We have verified that this choice does not affect the thermodynamic properties of the model (i.e. the *folding temperature* and the *hydrophobic collapse temperature*) found in [44] for $\alpha = 50$. This result could be expected, since the harmonic bead–bead interaction introduced in [34] has been already shown not to alter the folding properties of the original Honeycutt–Thirumalai model [33].

With the above choice of parameters, the heteropolymer exhibits a four-stranded β-barrel native configuration (NC), described by the coordinates $\vec{q}_{\rm NC}(i)$, $i = 1, \ldots, L$, and shown in figure 1. The NC corresponds to the absolute minimum of the potential energy, whose value for $\alpha = 1000$ is $V_{\rm NC} \approx -48.94$. The native structure is stabilized by the attractive hydrophobic interactions among the B residues; in particular, the first and third B_9 strands, forming the core of the NC, are parallel to each other and anti-parallel to the second and fourth strand $(PB)_4$ and $(PB)_5P$. The latter are exposed to the solvent due to the presence of polar residues. Overall, the four strands are separated by stretches of three consecutive neutral beads, forming three turns (see figure 1). These involve the following sites: 10–12 (first turn), 21–23 (second turn) and 33–35 (third turn).

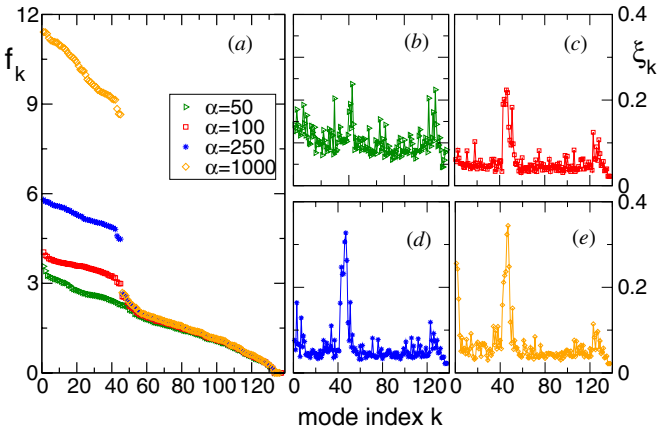


Figure 2. (*a*) Frequencies of the NMs, f_k, indexed from high-frequency modes. (*b*)–(*e*) Inverse of the participation ratio for the corresponding eigenvectors for the increasing value of the stiffness of the harmonic bonds. The triangles, squares, stars and circles refer to $\alpha = 50$, 100, 250 and 1000, respectively.

In the following, we report equilibrium microcanonical results obtained by integrating Hamilton's equations by means of a fourth-order symplectic integrator [49] with a time-step of 10^{-3} time units (ensuring a relative energy conservation of $\sim 10^{-7}$). In all cases, the initial coordinates of the beads were taken to correspond to the NC (zero displacement).

2.1. NMs' analysis

The properties of nonlinear modes in spatially and force-heterogeneous systems strongly depend on the features of the linear spectrum: in particular, non-resonant nonlinear modes may emerge within inter-mode gaps and highly localized vibrations can be associated with spectrum edges [14]. Hence, a detailed understanding of the linear spectrum is an essential pre-requisite for our analysis. More precisely, we wish to investigate how the NM frequencies and eigenvectors depend on the harmonic bond stiffness α.

The NM spectrum associated with the NC of our model protein is composed by $3L - 6$ non-zero frequencies $\{f_k = \sqrt{\lambda_k}/2\pi\}$, with λ_k being the eigenvalues of the Hessian matrix [50]. The NM spectra are reported in figure 2(*a*) for values of α ranging from 50 to 1000. The most important observation for our purposes is that, for large enough values of α, the spectrum splits into two sets separated by a gap. The set of *high-frequency modes* (or *bond-stretch modes*) comprises $L - 1$ modes associated with the stiffest force constants, that is, bond (backbone) distortions along the chain, while the set of *low-frequency modes* is composed by the

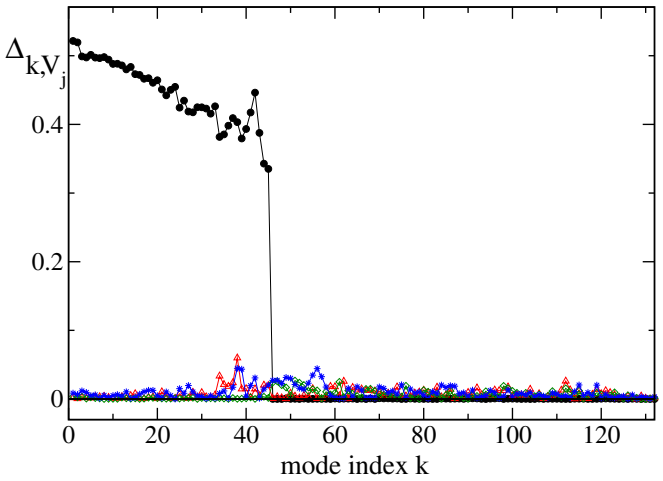


Figure 3. Relative variations of the energetic contributions due to a perturbation of the NC along the direction of the kth NM, as given by equation (6), for $\alpha = 1000$. The circles, triangles, diamond and stars correspond to Δ_{k,V_H}, Δ_{k,V_A}, Δ_{k,V_D} and $\Delta_{k,V_{LJ}}$, respectively.

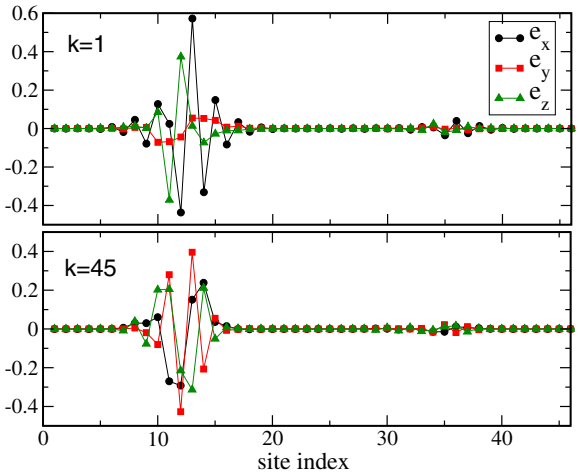


Figure 4. Cartesian components of the eigenvectors $\vec{e}_{\,1}$ (top) and $\vec{e}_{\,45}$ (bottom) corresponding to the frequencies at the upper and lower edges of the bond-stretch modes, respectively, for $\alpha = 1000$.

remaining $2L - 5$ modes that represent collective motions of the beads, which are hardly affected by changes in α. This interpretation can be substantiated by the following analysis. We have first slightly perturbed the NC along the direction of each eigenvector with an arbitrary amplitude ε, i.e. $\vec{q}_{\rm NC}(i) \to \vec{q}_k(i) = \vec{q}_{\rm NC}(i) + \varepsilon\, \vec{e}_{\,k}(i)$. Then we have evaluated the relative variations of the potential energy components V_j, j = H (harmonic), A (three-body angular), D (dihedral), LJ (Lennard–Jones), with respect to their values in the NC according to the following definition:

$$\Delta_{k,V_j} = \frac{|V_j(\vec{q}_k) - V_j(\vec{q}_{\rm NC})|}{|V_j(\vec{q}_k)|}. \tag{6}$$

As is clear from figure 3, the perturbation along the first 45 eigenvectors is only restricted to the degrees of freedom associated with bond deformation, i.e. to the harmonic contribution to the potential energy.

Furthermore, the formation of the gap is accompanied by a localization of the eigenvectors around the edges of the two sets of modes facing the gap. In order to characterize the degree of localization of the kth NM $\vec{e}_{\,k}(i)$ $(i = 1, 2, \ldots, L)$, we measured its *inverse participation ratio* [51],

$$\xi_k = \sum_{i=1}^{L} |\vec{e}_{\,k}(i)|^4 \tag{7}$$

where the eigenvectors are normalized to unity. For an eigenvector localized on a single site, $\xi \simeq 1$, while for a completely delocalized state, $\xi \simeq 1/L$. Therefore, the more the eigenvector is localized, the larger is ξ. The inverse participation ratio is plotted in figures 2(*b*)–(*e*) for different values of α. In particular, for $\alpha = 50$ all the eigenvectors are extended (see figure 2(*b*)). However, by increasing α, the degree of localization of NMs at the lower edge of the high-frequency set ($k = 43, 44, 45$) and of the first three NMs at the edge of the low-frequency set ($k = 46, 47, 49$) increases (figures 2 (*c*)–(*e*)). A further increase in stiffness does not only lead to an enhancement of localization in the proximity of the gap, but also determines the localization of the first three bond-stretch eigenvectors ($k = 1, 2, 3$), see figure 2(*e*). It is also interesting to observe that for increasing α the first and last three highest frequencies detach more and more from the core of the set (see again figure 2(*a*)). In the following, we will fix $\alpha = 1000$.

We note that the most localized modes of the high-frequency set are localized at the three turns of the NC, more precisely, modes $k = 1, 45$ at the first turn (figure 4), NMs $k = 2, 44$ at the third turn and modes $k = 3, 43$ at the second one. The modes at both edges of the high-frequency set are characterized by single or groups of neighbouring oscillators in opposition of phase, as shown in figure 4. Qualitatively, one may interpret them as impurity modes with the turns acting as structural defects [52]. Interestingly, for a toy model reproducing a single bent $2D$ chain with fixed curvature, Archilla *et al* [31] also reported a pair of localized modes lying at the set edge, with frequencies slightly detached from the core of the set itself. Concerning $k = 46, 47, 48$ at the edge of the low-frequency set, only the y-component appears localized on the first and third β-strand B_9, with the oscillators on the first and third strands in opposition of phase. Since these strands represent the core of the protein, we expect that the excitation of these modes should cause large rearrangements in the structure, leading to protein unfolding.

3. Emergence of DBs

In this section, we show how DBs can be created by exciting the NC along the direction of certain NMs. In these simulations, we initialize the beads' positions in the NC and assign the initial velocities proportional to the pattern of the selected NM. The amplitudes of the kinetic energy perturbation K_0 will be henceforth measured in temperature units $T = 2K_0/3L$, where a unitary Boltzmann constant is assumed.

As we shall show, the excitation of a NM provides an effective means of feeding energy to a DB. However, due to specific selection rules matching spatial overlap between NMs [62], a portion of the initial energy necessarily flows into a number of other modes. Such *background radiation* competes

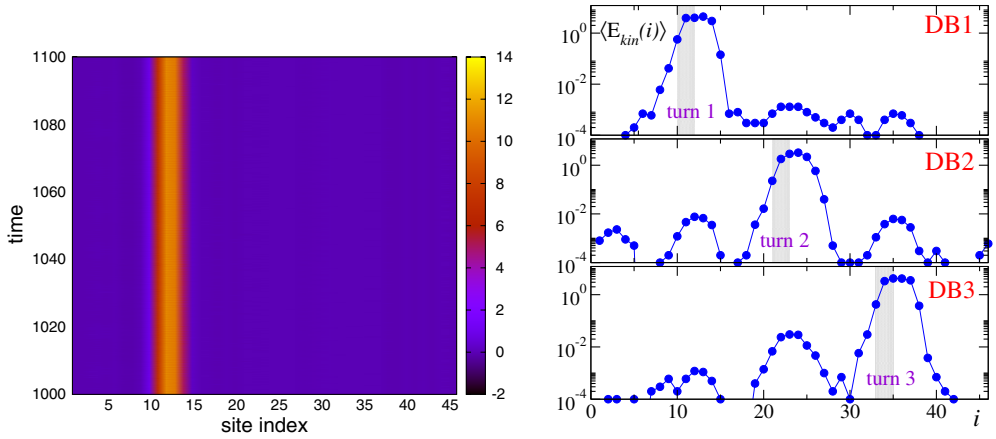


Figure 5. (*a*) Total energy difference per site $\Delta E_{\rm tot}(i)$ with $i = 1, \ldots, L$, as a function of the time, for DB1 with an initial temperature $T = 0.63$. (*b*) Kinetic site energy $\langle E_{\rm kin}(i)\rangle$ averaged over a time span of the order of 50–100 breather's periods in order to smooth out fluctuations. From top to bottom DB1 with initial temperature $T = 0.63$, DB2 at $T = 0.42$ and DB3 at $T = 0.63$. Cooling parameters are $\gamma = 0.001$, $\Delta t = 0.02$ (*a*) and $\Delta t = 0.005$ (*b*). The grey regions mark the protein turns (see the text).

with the nonlinear localization process. This in turn, as we shall see in the following, accelerates the collapse of nonlinear modes. In order to analyse DB properties, it proves useful to get rid of background radiation through *surface damping* [23]. To this end, in the first part of the work we shall cool down the velocities of the beads at the chain terminals. More precisely, we integrate Hamilton's equations of motion while rescaling, at regular time intervals Δt, the momenta of the beads located at the chain terminals $i = 1, \ldots, 4$ and $i = 43, \ldots, 46$), that is,

$$\vec{p}(i) \to \vec{p}'(i) = (1-\gamma)\vec{p}(i) \quad \text{with} \quad \gamma \leqslant 1. \tag{8}$$

Unless otherwise specified, the cooling parameters are fixed to $\gamma = 0.001$ and $\Delta t = 0.02$. In section 3.1.1, a comparison between the dynamics with and without cooling is presented.

In order to visualize the energy localization along the chain highlighting the emergence of a DB, we recorded the site kinetic energy $E_{\rm kin}(i)$ and the total excitation energy per site $\Delta E_{\rm tot}(i) = E_{\rm kin}(i) + V(i) - V_{\rm NC}(i)$, where $V(i)$ is the potential energy contribution which can be associated with the ith bead (see the appendix for the exact definition) and $V_{\rm NC}(i)$ is the corresponding value in the NC. The presence of a DB is detected as a protracted concentration of energy on a limited number of sites.

Our first important observation is that a stable DB can be created by perturbing the NC along the direction of *each of the first 45 modes*, or along a linear combination of them. An example is reported in figure 5(*a*) for the excitation of the lower edge mode of the high-frequency set. Remarkably, only three distinct breathers were observed to emerge despite the different directions of the employed perturbations. They are located close to the three turns and we thus denote them as DB1, DB2 and DB3, after the index of the corresponding turn (figure 5(*b*)).

A peculiarity of such localized modes is that they feature patterns that are not simply localized over a few adjacent sites. Rather, they display non-negligible components localized elsewhere in the chain. This can be appreciated by looking to

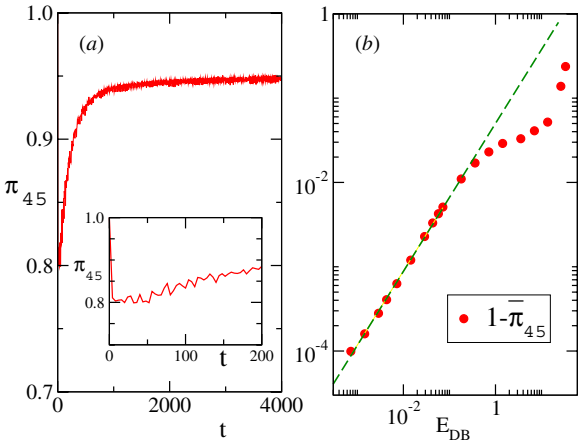
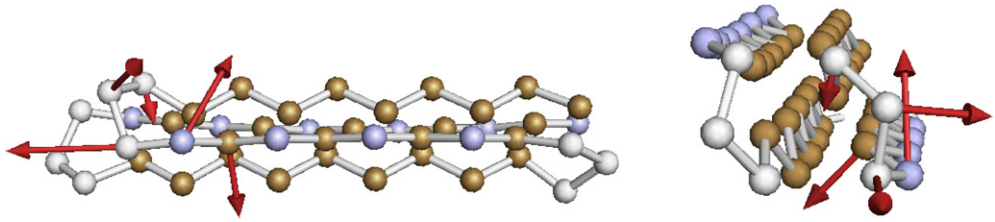


Figure 6. DB obtained as a continuation of NM $k = 45$. (*a*) Time evolution of the projection π_{45} for $T = 0.21$. The inset shows a close-up of the short-time dynamics. (*b*) $1 - \overline{\pi}_{45}$ as a function of the total energy $E_{\rm DB}$. The dashed line represents a power law of the type $1 - \overline{\pi}_{45} \propto E_{\rm DB}^{\eta}$ with $\eta \approx 0.88$.

the average site kinetic energy reported in figure 5(*b*) for the three breathers. The largest contribution to the DB total energy comes indeed from the few sites located at the corresponding turn. However, energy components two-orders of magnitude smaller can be found over the other turns too. The origin of these peculiar localization patterns can be found in the three-dimensional structure of the protein, and it is induced by the requirement of momentum conservation. In other words, the special built-in momentum-conservation properties of NM patterns are reproduced by DBs. As a matter of fact, it turns out that a mode that is localized at a few sites conserves its momentum by fractioning an equal and opposite amount of momentum among a few other locations, instead of spreading it over the whole structure. The interesting question arises whether this is a generic feature of NMs in proteins or it is rather the expression of the peculiar native fold considered here.



(*a*) (*b*)

Figure 7. Displacement field of DB1. Side (*a*) and front (*b*) views (only displacements larger than 1.0 have been shown as arrows). The initial condition corresponds to $T = 0.63$.

The three DBs can be considered as nonlinear continuations of three corresponding NMs, which are localized at the same turns and whose vibration patterns can be considered as precursors of the DB displacement fields. This is confirmed by looking at the normalized projections π_k of the breathers' velocity fields on the corresponding modes, defined as

$$\pi_k = \left\langle \sum_{i=1}^{L} \frac{\vec{p}(i)}{\sqrt{2K}} \cdot \vec{e}_k(i) \right\rangle_t \tag{9}$$

where the angular brackets denote the average over a window of w time units ($w = 4$ in the following). Figure 6(*a*) shows that, for DBs originating from NM 45, after an initial transient π_{45} shows a stable trend with very small residual oscillations around the mean value $\overline{\pi}_{45}$ evaluated in the last stage of the simulation ($3000 \leqslant$ time $\leqslant 4000$). The projections on the other modes (not shown in the figure) are smaller than 0.015. In figure 6(*b*), we plot $1 - \overline{\pi}_{45}$ as a function of the DB total energy, $E_{\rm DB} = K + V - V_{\rm NC}$, measured in the last stage. The figure shows that $1-\overline{\pi}_{45}$ vanishes when the energy is decreased following a nearly linear trend.

These results confirm and generalize what was reported in the context of the NNM, where gap-less breathers were shown to arise as continuations of edge NMs [14]. Also in the framework of the NNM, a few special regions were shown to act as energy-accumulating centres upon generic excitation of the system, exactly as we observed here [23].

Perturbations along all the bond-stretch modes of amplitude $T = 0.63$ resulted in the excitation of DB1 or DB3, while DB2 was observed only in a few cases. The reason why DB2 seems somehow more difficult to excite is likely to depend upon the peculiar structural neighbourhood of the second turn. This region lies deeper inside the core of the protein, where beads are more constrained, thus hindering the DB oscillations (see also the sketch showing DB patterns in figure 7).

The DB frequencies $f_{\rm DB1}$, $f_{\rm DB2}$ and $f_{\rm DB3}$ are computed from the power spectrum of the displacements of the most energetic bead within each DB. Typical DB spectra feature a main peak and minor (several orders of magnitude smaller) peaks signalling the residual excitation of a few NMs. The measured frequencies are shown in figure 8 as functions of the DB total energy $E_{\rm DB}$; all three DBs exhibit an almost linear decrease characterized by the same dispersion relation $f_{\rm DB}(E_{\rm DB}) = f_k(1 - E_{\rm DB}/\varepsilon)$, where f_k is the frequency of the

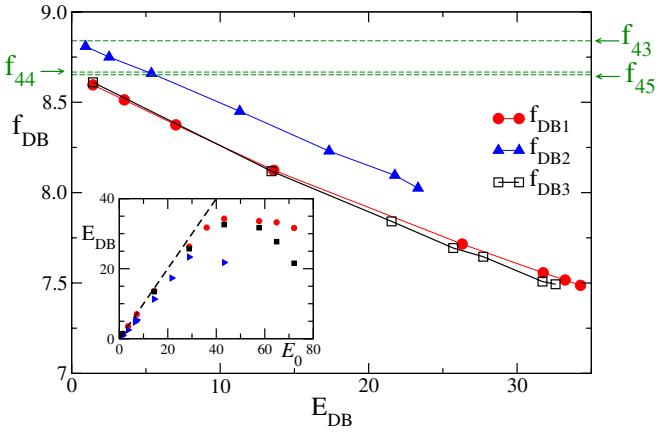


Figure 8. DBs originating from $k = 45$ (DB1), $k = 43$ (DB2) and $k = 44$ (DB3). Frequencies of DBs as a function of their total energy $E_{\rm DB}$. The horizontal dashed lines flag the frequencies of the corresponding NMs. The inset displays $E_{\rm DB}$ as a function of the initial energy $E_0 = K_0 - V_{\rm NC}$ for the three DBs: the circles, triangles and squares refer to DB1, DB2 and DB3, respectively. The dashed line marks the *full-efficiency* regime $E_{\rm DB} = E_0$.

NM from which the DB originates and $\varepsilon \approx 250$. We note that ε could be in principle calculated through Lindstedt–Poincaré perturbation theory as done, for example, in [14].

The decrease of the DB frequency with the energy $E_{\rm DB}$ suggests that the relevant nonlinearity is of the *soft* type [17].[6] This explains why the presence of a gap is necessary for the formation of localized nonlinear modes. Indeed, we were unable to excite DBs in cases where the spectrum is gap less.

The frequencies $f_{\rm DB1}$ and $f_{\rm DB3}$ originate from the bottom of the bond-stretch modes ($f_{44} \approx f_{45}$) and enter the gap for arbitrary small energies. In contrast, $f_{\rm DB2}$ lies between f_{43} and f_{44} for small energies, staying manifestly non-resonant with the neighbouring high-frequency modes. This behaviour is reminiscent of intra-band DBs predicted by Kopidakis and Aubry in nonlinear disordered chains [53, 54] and found analytically in NNMs of proteins [14]. For larger energies, the frequency decreases further and also these DBs eventually enter the gap.

[6] Following standard convention in the domain of nonlinear physics, hard and soft types refer here to the exponent of the leading power-law nonlinearity in the effective potential term responsible for the emergence of the DB. Hard refers to an even exponent, while soft to an odd exponent. Soft (resp. hard) nonlinearities are associated with a decrease (resp. increase) of the DB frequency for increasing energy ($E_{\rm DB}$).

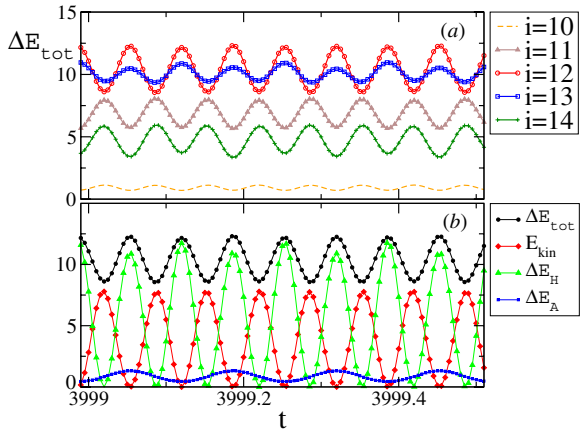


Figure 9. (*a*) Total site energies for DB1 and (*b*) the different energy contributions at site $i = 12$ versus time ($\Delta E_j = V_j - V_{\rm NC}(j)$ where $j = H, A$, the dihedral and LJ contributions being negligible). Parameters are $T = 0.63$, $\gamma = 0.001$, $\Delta t = 0.005$.

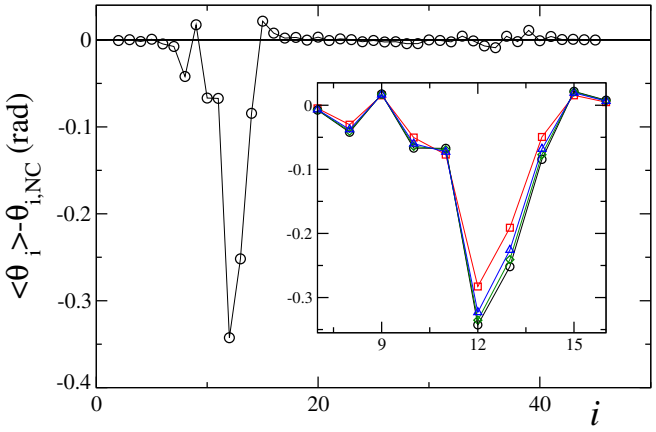


Figure 10. Difference between the average values of the bond–bond angles θ_i when the DB1 is present and the corresponding equilibrium values (in the NC), $\theta_{i,\rm NC}$, when the total energy of DB1 is $E_{\rm DB} = 34.6$. The inset shows a close-up of the kink region for increasing energies: the (red) squares, (blue) triangles, (green) diamonds and (black) circles refer to $E_{\rm DB} = 26.3$, 31.68, 33.61 and 34.6, respectively.

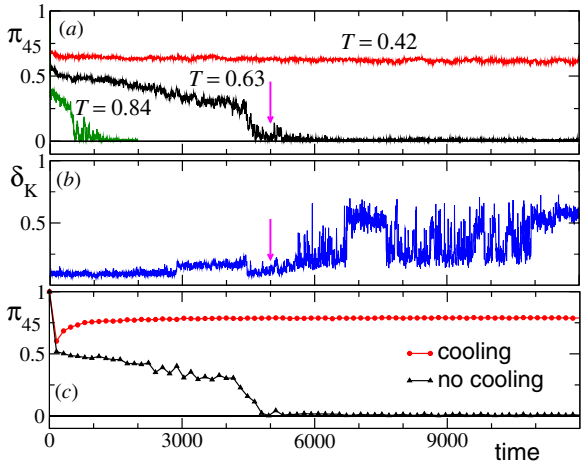


Figure 11. (*a*) Projection of the DB1 velocity field on the $k = 45$ NM, π_{45}, as a function of the time for increasing values of the initial energy. From the top to the bottom curve, $T = 0.42$, 0.63, 0.84 (non-dissipative dynamics). (*b*) Kabsch distance δ_K for $T = 0.63$ (non-dissipative dynamics). The vertical arrows mark the time at which the DB collapses. (*c*) Comparison of π_{45} for dissipative (upper red line), and non-dissipative (lower black line) dynamics for $T = 0.63$.

It is instructive to consider whether a given DB can be created with arbitrary energy by exciting a given NM. The inset in figure 8 shows the relation between the initial total energy $E_0 = K_0 - V_{\rm NC}$ and the energy that is found stored in the DB. The ratio of the two values can be interpreted as an estimate of *transfer efficiency*, i.e. the fraction of the initial excitation energy channelled into the DB and consequently pinned down at a very specific location. Remarkably, we see that all the three DBs are characterized by a nearly unitary efficiency, up to a point where the curve saturates and then starts bending down. The transfer of energy from the NM to the DB is nearly optimal up to initial energies of the order $E_0 = K_0 - V_{\rm NC} \approx 30$. This indicates the existence of a sort of maximal 'DB capacity'—if the NM is fed with larger energies, the excess energy cannot be further injected in a localized mode and it is necessarily spread across the whole structure.

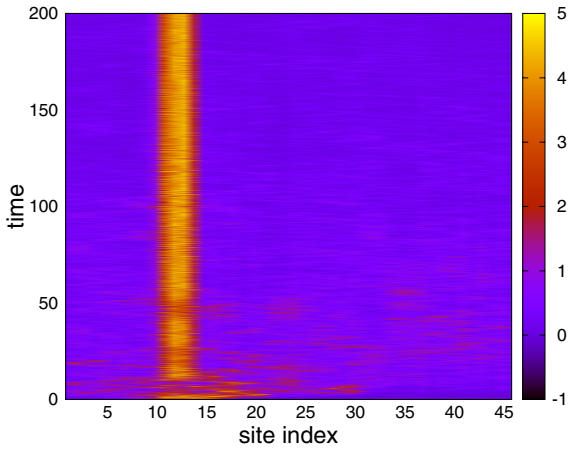
Finally, no DB could be excited by perturbing the structure along NMs belonging to the low-frequency set. This is most likely a consequence of the soft nonlinearity. Moreover, we were also unable to excite DBs through arbitrary localized perturbations of the NC. Indeed, initial conditions of the latter type have non-zero components over many NMs, including the ones in the low-frequency set.

3.1. DBs originating from the lower edge mode of the high-frequency set

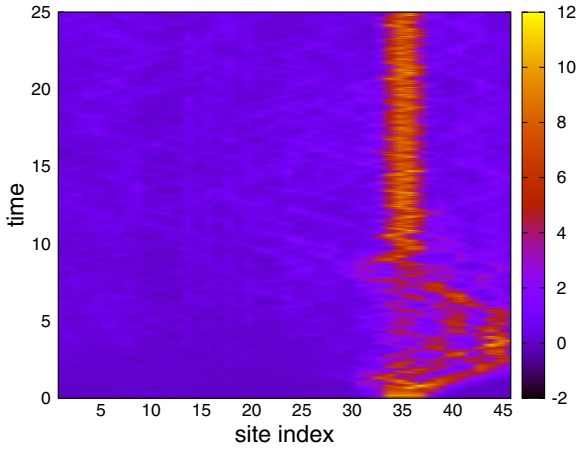
In this section, we present a more detailed analysis of the properties of DB1 originating from the bottom bond-stretch mode $k = 45$. This breather is localized on the first turn, from site 10 to site 14, as is shown in figure 5. The energy redistribution dynamics within the DB can be appreciated from figure 9(*a*). In particular, it is clear that the two site pairs (12,13) and (11,14) exhibit approximately phase-locked oscillations, while the two pairs are anti-phase locked among them. This is another illustration of the fact that the DB exchanges little or no energy with the surrounding.

It is instructive to analyse the different contributions to the total energies when the DB is present, in order to quantify to what extent the different degrees of freedom participate to the breather dynamics. As can be seen from figure 9(*b*), at one of the most energetic sites, $i = 12$, the DB vibration involves essentially the backbone bonds (≈84 %) with a small contribution coming from the angular terms (≈13 %), while the dihedral and LJ contributions are negligible. The observed ratios between the different energetic contributions do not change by varying the initial energy. We observe the same behaviour for all the sites in the interval (11–13).

A closer inspection of figure 9(*b*) shows that the harmonic bond component oscillates at frequencies $2f_{\rm DB}$ but also has a sizeable component at $f_{\rm DB}$. This is due to the soft nonlinearity which is known to induce a DC component in the displacement patterns of localized modes [18, 55]. To understand this in



(*a*)



(*b*)

Figure 12. Time evolution of total site energies starting from excitation of the upper-edge modes of the high-frequency set: (*a*) emergence of DB1 from $\vec{e}_1$, $T = 0.63$ and (*b*) of DB3 from $\vec{e}_2$, $T = 0.84$.

a simple way, let us consider two bond lengths along the backbone at site n , $r_1 \equiv r_{n-1,n}$ and $r_2 \equiv r_{n,n+1}$. We may write their evolution as an oscillating part plus a static distortion of the bond lengths,

$$r_1(t) = A_1 \sin(\omega t) + \langle r_1 \rangle$$
$$r_2(t) = A_2 \sin(\omega t) + \langle r_2 \rangle$$

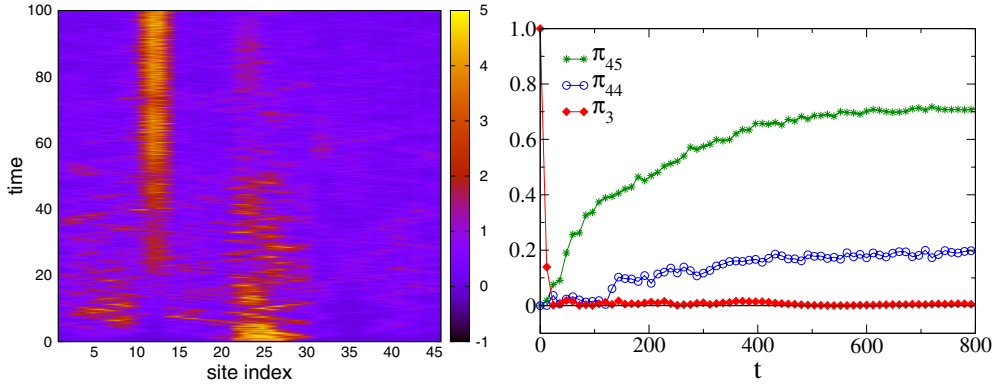
where $\omega = 2\pi f_{\mathrm{DB}}$ and the angular brackets represent a time average. Using the definition reported in the appendix, it can be shown that the contribution of the two bonds to the local harmonic energy is

$$V_{\mathrm{H}}(n) \propto \frac{\left(A_1^2 + A_2^2\right)}{2} \sin^2(\omega t) + [A_1(\langle r_1 \rangle - r_0) + A_2(\langle r_2 \rangle - r_0)] \sin(\omega t) + \frac{(\langle r_1 \rangle - r_0)^2 + (\langle r_2 \rangle - r_0)^2}{2}. \quad (10)$$

Thus, the presence of a component at frequency ω implies that $\langle r_1 \rangle$ and/or $\langle r_2 \rangle$ must be different from their equilibrium value r_0. Accordingly, we conclude that the emergence of a DB also causes a stable structural distortion of the protein. Given that $A_1 \approx A_2 \approx 0.1$ (see again figure 9), also tiny differences $\langle r_i \rangle - r_0 \approx 10^{-3}$, as we recorded in our simulations, are able to affect the local bond energies.

Much more prominent is the distortion effect caused by the soft-nonlinearity in the angular degrees of freedom. In particular, figure 10 reveals that the breather is also characterized by a kink-shaped angular distortion. The kink's amplitude increases with the total energy stored in the breather (see the inset in figure 10).

3.1.1. DB collapse and the effect of cooling. In the case where no cooling is applied ($\gamma = 0$), the DBs have a finite lifetime. To illustrate this, in figure 11(*a*) we report the time evolution of the projection π_{45} of the DB1, obtained as a continuation of the lower edge mode ($k = 45$) of the high-frequency set. The sudden drop of π_{45} signals the collapse of the excitation, accompanied by a rapid redistribution of energy over all the NMs (equipartition). The typical collapse time increases upon decreasing the initial energy (compare the three curves in figure 11(*a*)). This is similar to what was



(*a*) (*b*)

Figure 13. Long-range energy transfer starting from $\vec{e}_3$ at $T = 0.63$. (*a*) Time evolution of total site energies; (*b*) projections of the DB velocity field on three NMs as functions of the time.

demonstrated for one-dimensional chains where the time to equipartition scales as an inverse power of the energy density [56, 57].

The effect of the collapse of a DB on the protein structure can be appreciated by considering the Kabsch distance δ_K, which is a commonly used measure of the structural distance between two protein configurations [58]. In particular, figure 11(*b*) shows how the distance δ_K of the protein structure from the NC evolves in time in the presence of a DB of type DB1. As long as the DB is present, δ_K fluctuates around a relatively small value ($\delta_K \simeq 0.1$), meaning that the protein structure does not deviate appreciably from the NC. In fact, the static distortion effect illustrated above only concerns a small region of the fold. On the other hand, when the DB collapses, δ_K starts to fluctuate wildly around a substantially larger value ($\delta_K \approx 0.35$), thus signalling the occurrence of important conformational rearrangements.

The DB lifetime gets substantially increased by removing the background vibrations through cooling. Figure 11(*c*) clearly illustrates how dissipation renders the DB stable by quickly eliminating background radiation. The stabilizing role of the boundary cooling method is well documented [23, 59, 60] and our results confirm that it can be used conveniently also for such a complex structure.

3.2. Excitation of highest energy edge NMs: long-range energy transfer

The excitation of NMs at the upper edge of the spectrum gives rise to quite peculiar phenomena. As a first example, figure 12 depicts the emergence of a DB obtained by exciting the first two highest-frequency NMs, i.e. $\vec{e}_1$, which is localized at sites 10–12 and $\vec{e}_2$, which is localized at sites 34–36. In the first stage, the energy quickly spreads and remains confined in the vicinity of the perturbed location. After some time, a DB self-localizes, harvesting energy from the background and pinning it in the same region. The DBs obtained are again DB1 and DB3. This is reasonable in view of the similar spatial structure of $\vec{e}_1$, $\vec{e}_2$ and DB1, DB3, respectively (see again figure 4). Moreover, we found that the frequencies of DBs originating from $\vec{e}_1$ lie on the same frequency-energy curve as the breather originating from NM 45, shown in figure 8.

Remarkably, we have also observed that the excitation of an edge mode may result in a *long-range energy transfer* event. An illustration of this phenomenon is given in figure 13(*a*), for the excitation of the third highest-frequency NM. The NM $\vec{e}_3$ is localized at sites 22–27. Immediately after the perturbation, the DB appears localized in the same region, but seems to collapse rapidly afterwards, spreading its energy evenly across the structure. However, after a considerable time span, another DB emerges at a different location, namely at sites 11–13. This is the region where the lower-edge mode of the high-frequency set ($k = 45$) is localized. A substantial part of the initial energy has been transferred and pinned down irreversibly at the other end of the structure, covering the distance from a turn to the following one (see again the sketch depicting the protein structure).

Such an energy transfer phenomenon can be rationalized by analysing the projections of the DB velocity field on the

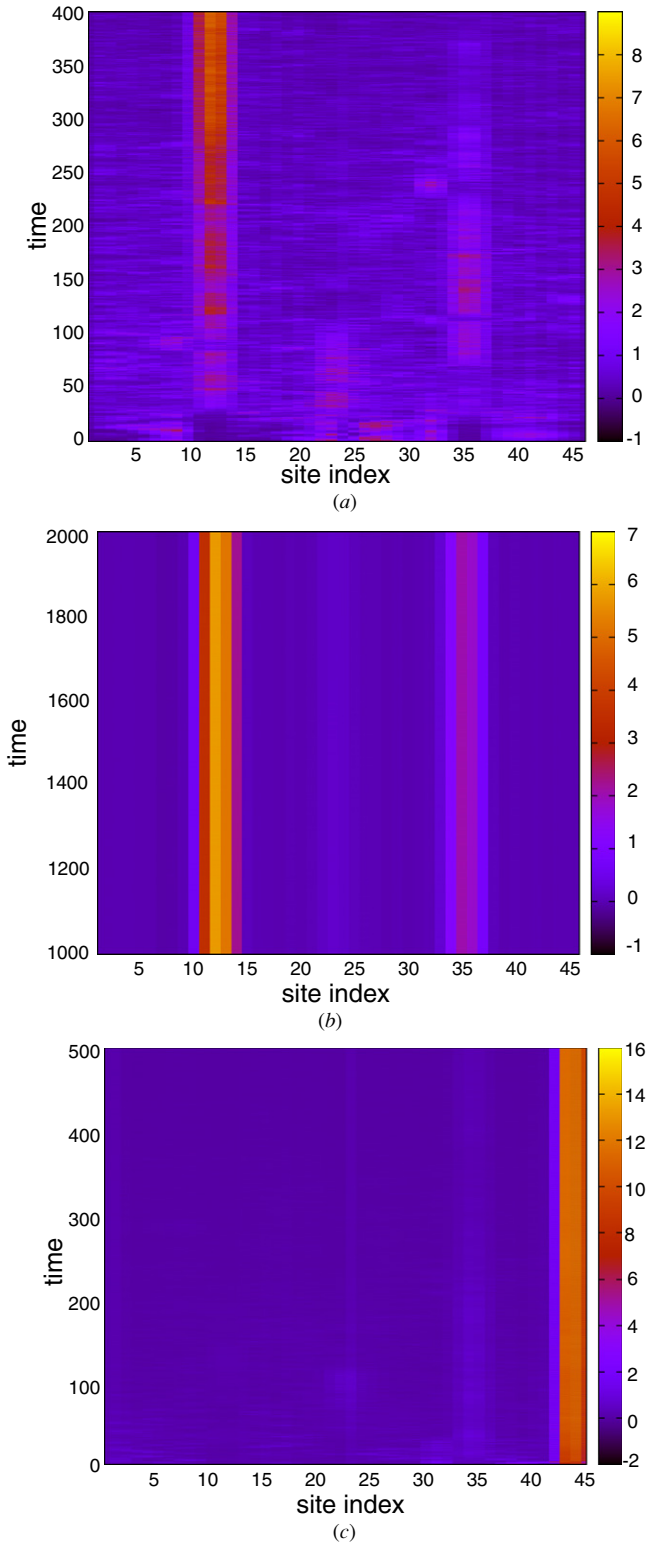


Figure 14. Evolution of total site energies starting from the excitation of the bond-stretch modes in the core of the set: (*a*) emergence of DB1 from $\vec{e}_5$, $T = 0.63$ and (*b*) the multibreather solution localized on the first and third turn starting from $\vec{e}_{33}$ at $T = 0.63$; (*c*) BB following the excitation of mode $\vec{e}_{36}$, $T = 0.84$ (non-dissipative dynamics).

NMs during the time evolution of the perturbation. From figure 13(*b*) one can clearly see that the initially excited mode is quickly emptied of its initial energy, which gradually flows

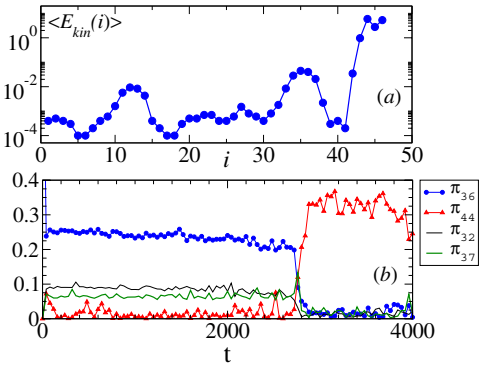


Figure 15. Emergence of a BB following the excitation of mode $k = 36$ with $T = 0.63$. (*a*) Time-averaged site kinetic energies (the cooling is applied to the first four beads); (*b*) projections of the DB normalized velocity field on the NMs as a function of the time (non-dissipative dynamics).

into other two modes. As it shows, asymptotically the two target modes almost describe the entirety ($\pi_{44} + \pi_{45} \approx 1$) of the DB, which is in fact localized at the turn opposite to the excited one, where the two target NMs are also localized. The reverse process is forbidden since the DB frequency is no longer resonating with the original mode. We remark that this phenomenon, although evocative of resonant energy transfer among a few selected NMs in proteins, is a one-way transfer, as the DB will retain the transferred energy for times comparable to its lifetime [61, 62].

3.3. Excitation of inner NMs

NMs in the core of the high-frequency set are generally characterized by a low level of localization, as one can see in figures 2(*b*)–(*e*). As a result, when one of such modes is excited, we observe a complex localization pattern, which alternates site-hopping and energy-pinning stages to phases where the energy is more evenly distributed. After a certain time, a stable DB emerges, focusing the energy at one of the turns. A typical realization of this scenario is illustrated in figure 14(*a*) for the excitation of the fifth highest-frequency mode.

A different phenomenology is observed when we excite a mode lying deeper within the set. It may occur that the excitation energy remains trapped in two DBs localized on two different turns, realizing a state which is reminiscent of multibreather states observed in nonlinear disordered systems [53, 54]. A realization of this phenomenon for NM 33 is reported in figure 14(*b*). In particular, we note that starting from NM 33 and changing the initial energy, we can obtain different solutions. For example, for $T = 0.21$ we obtain a solution localized on the first and second turn, while the multibreather localizes on the first and third ones for $T = 0.63$.

A very peculiar localized solution develops through the excitation of NM $k = 36$, whose pattern is localized in the protein's tail. As is shown in figure 14(*c*), the energy remains almost entirely confined to the edge sites 44–46, with smaller, but non-negligible, amounts of energy also involving the turns, figure 15(*a*). We term such excitations *boundary breathers* (BB). As in the case of DB1, DB2 and DB3, the frequency of oscillation of the BB also lies in the gap ($f_{\mathrm{BB}} = 7.739$ for $E_{\mathrm{DB}} = 34.75$). At variance with DB1, DB2, DB3, the energy of BBs has not only harmonic and angular contributions but also a sizeable LJ one, resulting from a non-negligible interaction between the first and last beads, as well as the second turn region of the chain which are relatively close in the NC. We note that nonlinear surface states of different kinds are known and studied in many contexts [18, 63].

Eventually, the BBs hop to one of the three breathers described above. Figure 16 illustrates an example of this process, where the BB transfers its energy to DB3, as confirmed by the time evolution of the NM projections shown in figure 15(*b*). Asymptotically, we see that the main projection is π_{36}, while at $t \approx 2.75 \times 10^3$ an abrupt transition occurs and the leading projection becomes π_{44}, signalling the energy transfer to DB3 (corresponding to NM $k = 44$) localized at the third turn.

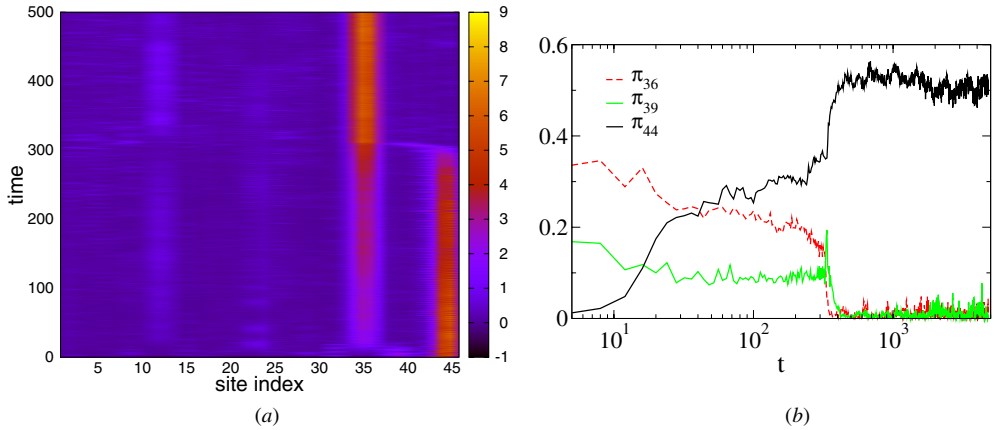


Figure 16. Transient multibreather-like mode, localized on the protein boundary and on the third turn, obtained following the excitation of the $k = 36$ NM at $T = 0.42$ (non-dissipative dynamics). (*a*) Spacetime plot of the total site energies. (*b*) Normalized projections of the DB velocity field on the NMs as a function of the time.

While at high energies, i.e. $T = 0.84, 1.05$, the lifetime of BBs before a transition to DB3 takes place is short, at lower energies, we observe the formation of a transient multibreather-like state with a progressive transfer of energy from BB to DB3. One of such examples is depicted in figure 16. Decreasing the excitation energy further, BBs are no longer observable.

We note that the end of the protein can be considered as a defect/discontinuity in the chain, since it is a free end that interrupts the chain itself. Therefore, BBs can be considered as akin to localized modes emerging in other contexts close to point defects. Furthermore, we were unable to excite stable BBs localized on the first beads of the sequence, despite NMs with significative components on that terminal do exist. A possible explanation of this apparently contradictory result is the fact that the first strand is part of the protein core. Therefore, it is a much more rigid structure with respect to the last strand, which instead is exposed to the solvent and can oscillate more freely.

4. Conclusions

In this paper, we have reported the existence of discrete breathers (DBs) in an off-lattice model of proteins with realistic interaction potentials and a three-code amino-acid sequence. To our knowledge, this is the first time that this class of peculiar vibrational modes are found in such a complex and heterogeneous dynamical system. In particular, we have shown that, due to the softness of the leading nonlinearity, a necessary condition for DB existence is that the linear spectrum exhibits a gap (at least one), as it is indeed the case in all proteins.

For the particular structure analysed in this work, we have identified three families of DBs, each one localized on a different turn of the native fold. The largest fraction of the energy of a DB typically resides on a few sites and involves essentially harmonic and angular degrees of freedom, while dihedral and LJ interactions between distant sites are practically not excited.

We have obtained DBs as continuations of the lower edge NMs of the bond-stretch set. Quite generically, however, we have shown that the members of the three different breather families can be excited by feeding energy to *any* of the NMs lying in the high-frequency set. Remarkably, the excitation of a DB obtained by feeding energy to a given NM is an extremely *efficient* process. This means that there is an extended range of initial energies that can be fed to the system and immediately channelled almost entirely into a breather, which is permanently localized at a very specific location.

Our results prove that DBs have a stabilizing effect on the protein kinetics. In fact, a broad class of perturbations of the native fold result in the formation of a DB, that is, a stable and highly localized vibration, while on the other hand long-range contacts are almost devoid of energy. This enhances the structural stability of the protein. To this regard, it would be interesting to connect the structure-dynamics relation underlying DB formation to the well-conserved patterns of nucleation sites along folding pathways [64].

We also discovered peculiar DB-assisted long-range energy transfer phenomena, whereby energy is channelled to a distant region of the structure away from the excitation site. At variance with ordinary energy exchange between resonant NMs [62], this is a one-way process, meaning that the transferred energy is never released back to the starting location. In particular, this effect may be of importance in the functioning of allosteric proteins [6, 65, 66] and surely deserves further investigation.

Our work has highlighted nontrivial correlations between structural and dynamical features of protein folds. In a simple β-barrel, as the one here considered, localization occurs on loops. Concerning more complex structures, our results raise the important question whether more complex structural selection rules exist driving nonlinear energy localization and long-range targeted transfer.

The model examined in this paper exhibits a few competing β-barrel minima of the potential energy landscape [34]. For future work, it could be worth analysing the emergence of DBs in models without frustration, better suited to reproduce naturally occurring energy landscapes, such as native-centric models or modified versions of the present model containing salt bridges (see [67, 68]).

Acknowledgments

We acknowledge useful discussions with S Flach, R Livi and L Schimansky-Geier. Thanks are due to S Allaccia and A Scarpa for inspiring the style and presentation of the paper. AT gratefully acknowledges the financial support of Aarhus Universitet Forskning Fonden (AUFF) during his stay at Aarhus University. AI was partially supported by Lundbeck Fonden. This work is part of the CNR-RSTL project no 827 *Dinamiche cooperative in strutture quasi uni-dimensionali*. Finally, we thank the Danish Centre for Scientific Computing for providing us with computational resources.

Appendix. Single-site energy

Since each potential energy term represents a many-body interaction, we have estimated the single-site potential energies $V(i)$ by equally redistributing the energy to each site involved in the different terms and then by summing up all these contributions.

In particular, the harmonic term is a two-body potential involving nearby sites; therefore, the contribution is estimated as follows:

$$V_H(i) = \begin{cases} \sum_{k=i}^{i+1} V_H(r_{k-1,k})/2 & \text{for} \quad i = 2, L-1 \\ V_H(r_{1,2})/2 & \text{for} \quad i = 1 \\ V_H(r_{L-1,L})/2 & \text{for} \quad i = L \end{cases} .$$

The angular term is a three-body interaction involving three consecutive sites, and therefore

$$V_A(i) = \begin{cases} \sum_{k=i-1}^{i+1} V_A(\theta_k)/3 & \text{for} \quad i = 3, L-2 \\ V_A(\theta_2)/3 & \text{for} \quad i = 1 \\ \sum_{k=2}^{3} V_A(\theta_k)/3 & \text{for} \quad i = 2 \\ \sum_{k=L-2}^{L-1} V_A(\theta_k)/3 & \text{for} \quad i = L-1 \\ V_A(\theta_{L-1})/3 & \text{for} \quad i = L \end{cases} .$$

The dihedral angle φ is the angle between two nearby planes each containing three consecutive sites, the corresponding potential term is therefore a four-body term and the single-site contribution can be evaluated as follows:

$$V_D(i) = \begin{cases} \sum_{k=i-2}^{i+1} V_D(\varphi_k, \theta_k, \theta_{k+1})/4 & \text{for} \quad i = 4, L-3 \\ V_D(\varphi_2, \theta_2, \theta_3)/4 & \text{for} \quad i = 1 \\ \sum_{k=2}^{3} V_D(\varphi_k, \theta_k, \theta_{k+1})/4 & \text{for} \quad i = 2 \\ \sum_{k=2}^{4} V_D(\varphi_k, \theta_k, \theta_{k+1})/4 & \text{for} \quad i = 3 \\ \sum_{k=L-4}^{L-2} V_D(\varphi_k, \theta_k, \theta_{k+1})/4 & \text{for} \quad i = L-2 \\ \sum_{k=L-3}^{L-2} V_D(\varphi_k, \theta_k, \theta_{k+1})/4 & \text{for} \quad i = L-1 \\ V_D(\varphi_{L-2}, \theta_{L-2}, \theta_{L-1})/4 & \text{for} \quad i = L \end{cases} .$$

The LJ term is a two-body interaction involving sites separated by more than two bonds along the chain

$$V_{\mathrm{LJ}}(i) = \sum_j V_{\mathrm{LJ}}(r_{ij})/2 \quad \text{for} \quad |i-j| > 2 .$$

The total site potential energies read

$$V(i) = V_H(i) + V_A(i) + V_D(i) + V_{\mathrm{LJ}}(i).$$

References

[1] Creighton T E 2002 *Proteins: Structures and Molecular Properties* (New York: Freeman)
[2] Yang L and Bahar I 2005 *Structure* **13** 893–904
[3] Bahar I, Erman B, Jernigan R L, Atilgan A R and Covell D G 1999 *J. Mol. Biol.* **285** 1023–37
[4] Eisenmesser E Z, Bosco D A, Akke M and Kern D 2002 *Science* **295** 1520–3
[5] Kondo J, Urzhumtsev A and Westhof E 2006 *Nucleic Acids Res.* **34** 676–85
[6] del Sol A, Tsai C J, Ma B Y and Nussinov R 2009 *Structure* **17** 1042–50
[7] Delarue M and Sanejouand Y H 2002 *J. Mol. Biol.* **320** 1011–24
[8] Zheng W and Brooks B R 2005 *Biophys. J.* **88** 3109–17
[9] Elezgaray J, Marcou G and Sanejouand Y H 2002 *Phys. Rev.* E **66** 31908–15
[10] Brooks B R and Karplus M 1985 *Proc. Natl Acad. Sci. USA* **82** 4995–9
[11] Marques O and Sanejouand Y H 1995 *Proteins* **23** 557–60
[12] Murray J, Hu Q, Navenot J and Peiper S 2002 *Biochem. Biophys. Res. Commun.* **292** 449–55
[13] Bahar I, Atilgan A R, Demirel M C and Erman B 1998 *Phys. Rev. Lett.* **80** 2733–6
[14] Piazza F and Sanejouand Y H 2008 *Phys. Biol.* **5** 026001
[15] Sacquin-Mora S, Laforet E and Lavery R 2007 *Proteins* **67** 350–9
[16] Sievers A J and Takeno S 1988 *Phys. Rev. Lett.* **61** 970–3
[17] Flach S and Willis C R 1998 *Phys. Rep.* **295** 181–264
[18] Flach S and Gorbach A V 2008 *Phys. Rep.* **467** 1–116
[19] Trias E, Mazo J J and Orlando T P 2000 *Phys. Rev. Lett.* **84** 741–4
[20] Binder P, Abraimov D, Ustinov A V, Flach S and Zolotaryuk Y 2000 *Phys. Rev. Lett.* **84** 745–8
[21] Sato M and Sievers A 2009 *J. Biol. Phys.* **35** 57–72
[22] Manley M E, Sievers A J, Lynn J W, Kiselev S A, Agladze N I, Chen Y, Llobet A and Alatas A 2009 *Phys. Rev.* B **79** 134304
[23] Juanico B, Sanejouand Y H, Piazza F and De Los Rios P 2007 *Phys. Rev. Lett.* **99** 238104
[24] Piazza F and Sanejouand Y H 2009 *Phys. Biol.* **6** 046014
[25] Piazza F and Sanejouand Y H 2009 *Europhys. Lett.* **88** 68001
[26] Piazza F and Sanejouand Y H 2010 *Discrete Contin. Dyn. Syst.* S **4** 1247–66
[27] Sitnitsky A E 2006 *Physica* A **371** 481–91
[28] Agarwal P K 2005 *J. Am. Chem. Soc.* **127** 15248–56
[29] Hennig D, Schimansky-Geier L and Hänggi P 2007 *Europhys. Lett.* **78** 20002
[30] Salerno M and Kivshar Y S 1994 *Phys. Lett.* A **193** 263–6
[31] Archilla J F R, Gaididei Y B, Christiansen P L and Cuevas J 2002 *J. Phys. A: Math. Gen.* **35** 8885–902
[32] Sánchez-Rey B, Archilla J F R, Palmero F and Romero F R 2002 *Phys. Rev.* E **66** 017601
[33] Honeycutt J D and Thirumalai D 1990 *Proc. Natl Acad. Sci. USA* **87** 3526–9
[34] Berry R S, Elmaci N, Rose J P and Vekhter B 1997 *Proc. Natl Acad. Sci. USA* **94** 9520–4
[35] Guo Z and Thirumalai D 1995 *Biopolymers* **36** 83–102
[36] Guo Z and Brooks C L 1997 *Biopolymers* **42** 745–57
[37] Veitshans T, Klimov D and Thirumalai D 1997 *Fold. Des.* **2** 1–22
[38] Evans D A and Wales D J 2003 *J. Chem. Phys.* **118** 3891–7
[39] Kim J, Straub J E and Keyes T 2006 *Phys. Rev. Lett.* **97** 050601
[40] Kim J M and Keyes T 2007 *J. Phys. Chem.* B **111** 2647
[41] Li F Y, Yuan J M and Mou C Y 2001 *Phys. Rev.* E **63** 021905
[42] Lacks D J 2005 *Biophys. J.* **88** 3494–501
[43] Imparato A, Luccioli S and Torcini A 2007 *Phys. Rev. Lett.* **99** 168101
[44] Luccioli S, Imparato A and Torcini A 2008 *Phys. Rev.* E **78** 031907
[45] Imparato A, Luccioli S and Torcini A 2010 *Prog. Theor. Phys. Suppl.* **184** 339–50
[46] Luccioli S, Imparato A, Mitternacht S, Irbäck A and Torcini A 2010 *Phys. Rev.* E **81** 010902
[47] Blondel A and Karplus M 1996 *J. Comput. Chem.* **17** 1132–41
[48] Rampioni A 2005 Caratterizzazione del panorama energetico di piccoli peptidi al variare della loro lunghezza *PhD Thesis* University of Florence, Italy
[49] Atela P and McLachlan R I 1992 *Nonlinearity* **5** 541
[50] Wales D J 2003 *Energy Landscapes* (Cambridge: Cambridge University Press)
[51] Aoki H 1983 *J. Phys. C: Solid State Phys.* **16** L205
[52] Forinash K, Peyrard M and Malomed B 1994 *Phys. Rev.* E **49** 3400–11

[53] Kopidakis G and Aubry S 2000 *Physica* D **139** 247–75
[54] Kopidakis G and Aubry S 1999 *Physica* D **130** 155–86
[55] Bickham S R, Kiselev S A and Sievers A J 1993 *Phys. Rev.* B **47** 14206
[56] Cretegny T, Dauxois T, Ruffo S and Torcini A 1998 *Physica* D **121** 109–26
[57] Kosevich Y and Lepri S 2000 *Phys. Rev.* B **61** 299–307
[58] Kabsch W 1976 *Acta Crystallogr.* A **32** 922–3
[59] Tsironis G P and Aubry S 1996 *Phys. Rev. Lett.* **77** 5225–8
[60] Piazza F, Lepri S and Livi R 2003 *Chaos* **13** 637
[61] Moritsugu K, Miyashita O and Kidera A 2003 *J. Phys. Chem.* B **107** 3309–17
[62] Moritsugu K, Miyashita O and Kidera A 2000 *Phys. Rev. Lett.* **85** 3970–3
[63] Lazarides N, Tsironis G P and Kivshar Y S 2008 *Phys. Rev.* E **77** 065601
[64] Mirny L and Shakhnovich E 2001 *J. Mol. Biol.* **308** 123–9
[65] Tsai C J, Del Sol A and Nussinov R 2009 *Mol. Biosyst.* **5** 207–16
[66] Gunasekaran K, Ma B and Nussinov R 2004 *Proteins: Struct. Funct. Bioinformatics* **57** 433–43
[67] Miller M A and Wales D J 1999 *J. Chem. Phys.* **111** 6610–6
[68] Wales D J and Dewsbury P E J 2004 *J. Chem. Phys.* **121** 10284–90