



The harmonic analysis of cylindrically symmetric proteins: A comparison of Dronpa and a DNA sliding clamp

Guang Hu, Servaas Michielssens, Samuel L.C. Moors, Arnout Ceulemans*

Department of Chemistry and INPAC Institute for Nanoscale Physics and Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:

Accepted 22 December 2011

Available online 30 December 2011

Keywords:

Normal mode analysis

Protein topology

Symmetric proteins

ABSTRACT

The harmonic analysis of two types of proteins with cylindrical symmetry is performed by the Standard Force Field Normal Mode Analysis and by the elastic network model. For both proteins the global elastic modes are assigned to their characteristic topologies. Dronpa is a rigid β -barrel structure, presenting the twisting, bending and breathing motion of a cylindrical rod. The β sliding clamp of *Escherichia coli* is a hexagonal β -wheel, consisting of rigid segments. In its spectrum four classes of vibrations are identified which are characteristic of an elastic torus. Correlation diagrams and RMSF analysis are compared. The results provide not only a comprehensive validation of the use of both methods to describe the elastic behavior according to the low-frequency normal modes, but also depict the correlated motions of β -barrel and β -wheel proteins. The harmonic flexibility of the Dronpa protein is compared to the principal components of molecular dynamics (MD) simulation. A functionally important localized cleft opening mode is found, which is not detected by harmonic analysis.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The aim of the paper is to examine the relationship between topological structure and dynamics of proteins using harmonic analysis and MD simulations. Two representative proteins with high cylindrical symmetry, but different topologies are compared. The high symmetry greatly facilitates the application and interpretation of the mode analysis. The two selected proteins are both based on β -sheets.

The first protein is Dronpa [1]. It has the typical topology of a cylinder and is a homologue of the green fluorescent protein (GFP) [2]. It has very special luminescent properties which account for its widespread use as a biological tool. It contains eleven β -strands lining up to form a cylindrical β -barrel (Fig. 1). Two central helices, connected by the chromophore, run through the β -barrel and several interconnecting loops sit at the top and bottom to protect the chromophore.

The second representative protein is *Escherichia coli* DNA polymerase III β -subunit sliding clamp, which also has cylindrical symmetry, but toroidal topology [3]. The β clamp is a homodimer with two monomers in head-to-tail arrangement. Each monomer consists of three domains, that have almost identical tertiary structure (Fig. 2). It is homologous to the eukaryotic and archaeal proliferating cell nuclear antigen (PCNA). The ring-shaped structure

is lined by twelve α -helices in the inner surface and six β -sheets at the outer surface, forming a β -wheel [4]. The toroidal structure encircles the DNA and acts as mobile platform, which tethers the DNA polymerase to the template during replication and repair.

Several studies have reported MD simulations of GFP [5–7], Dronpa [8] and β clamps [9–11]. These studies have investigated their physicochemical properties in relation to their biological function. Here, we complement these studies by harmonic analysis [12], using two different techniques. The normal mode analysis (NMA) offers a harmonic approximation to the conformational landscape and decomposes the nuclear movements into discrete modes [13,14]. It has been shown that NMA can describe function related motions, such as channel opening motion or gating motion [15–20]. The approach, along with its modifications, has been applied for large-scale motions [21,22] and unraveling the mechanisms of protein activity [23,24]. It was also applied to explore the thermodynamic properties and atomic fluctuations of GFP [25]. The so-called standard NMA [26] is based on an extensive all-atom force field. We will denote this as FF-NMA. An alternative technique is the elastic network model (ENM), which was introduced by Tirion [27]. This is also an atomistic harmonic analysis method, but based on a single-parameter Hamiltonian. It restricts interactions to a uniform harmonic potential between all atoms within a given cutoff for describing protein motions. Hinsen [28] further showed that it was also possible to use $C\alpha$ atoms only. Two modified versions, also using $C\alpha$ coordinates and a $N \times N$ Kirchhoff matrix, are the so-called Gaussian network model (GNM) and anisotropic network model (ANM), proposed by Bahar et al. [29]

* Corresponding author.

E-mail address: arnout.ceulemans@chem.kuleuven.be (A. Ceulemans).

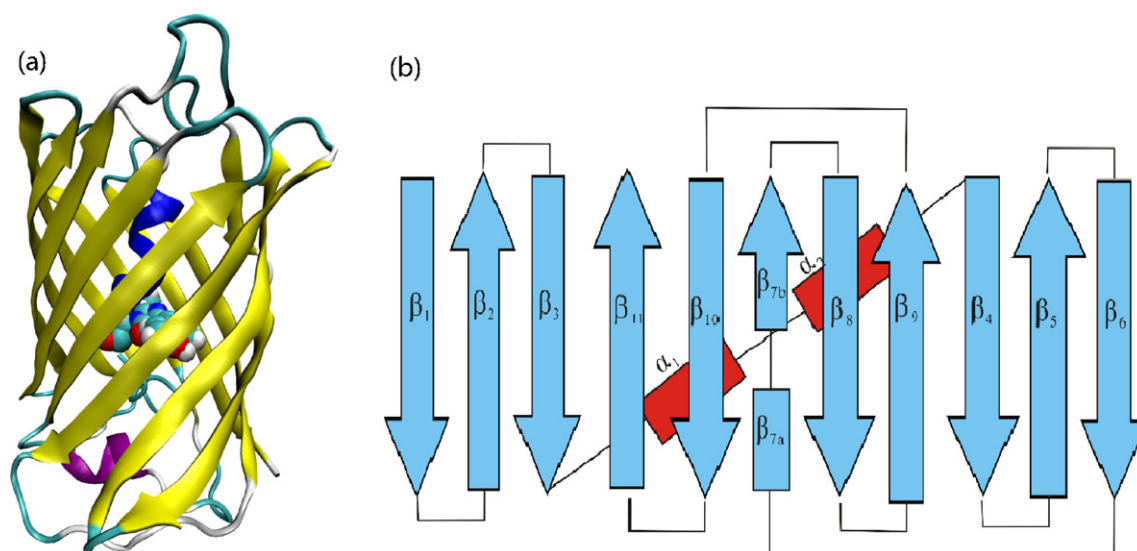


Fig. 1. The cylindrical structure of Dronpa. (a) Cartoon representation of the protein loaded with eleven β -strands and two α -helices. The internal chromophore is shown as Van der Waals spheres. (b) Schematic diagram of the secondary structure of Dronpa, where α -helices are shown as rectangles and β -sheets as arrows.

and Atilgan et al. [30]. In comparison with standard NMA, these coarse-grained models are efficient computational tools to describe the essential vibrational dynamics for larger proteins and molecular machines. In combination with MD and NMA, therefore, they are appropriate for describing the collective properties of global motions and estimating root mean square fluctuation (RMSF) of proteins [31–38].

In this paper, FF-NMA in combination with ENM is used to detect low-frequency modes of Dronpa and the β clamp. The paradigm is that low-frequency modes are mainly governed by the shape of the protein, and are of particular importance as presumable indicators of large-scale conformational changes [39–42]. In the low-frequency range one expects that proteins behave as continuous elastic bodies – a cylinder in the case of Dronpa, and a torus for

the β clamp. These topologies are furthermore common to GFP [2] and DNA polymerases [42,43], which would also explain why functional aspects are conserved. To put these accepted views to test, we will compare the results of the harmonic analysis with existing MD simulations for Dronpa.

2. Theory and methods

2.1. Normal mode analysis

Normal mode theory is based on the harmonic approximation of the potential energy surface around an energy minimized structure [12]. This approximation can be obtained by Taylor expansion of the potential energy function about the minimum energy up to

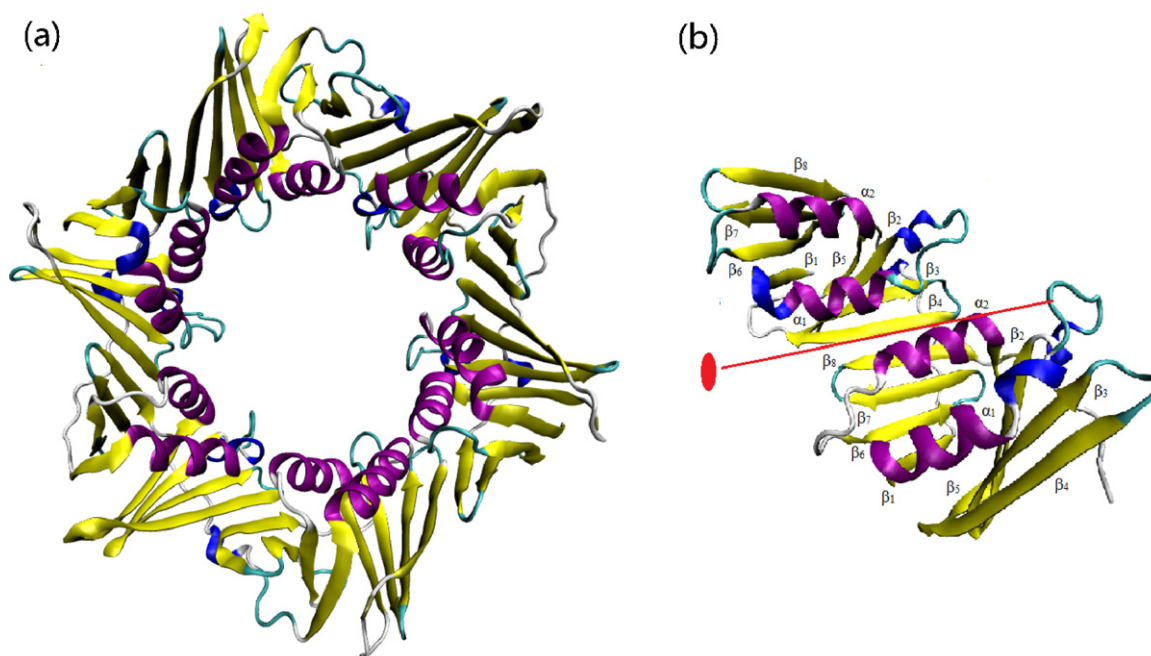


Fig. 2. The toroidal structure of the β clamp. (a) Cartoon representation of the protein, with 12 α -helices shown as spirals and 6 β -sheets as flat ribbons. (b) A view of two β -sheets at the interdomain interface. The ellipse indicates the twofold symmetry of two domains, with each containing 8 β -strands.

second-order in the displacements. The normal mode vectors are obtained by the eigenvalue equation:

$$\mathbf{W}^T \mathbf{H} \mathbf{W} = \mathbf{\Omega} \quad (1)$$

where \mathbf{H} is the mass-weighted Hessian, the l th column of the eigenvector matrix \mathbf{W} denotes the l th normal mode, and $\mathbf{\Omega}$ is a diagonal matrix with elements:

$$\Omega_{ll} = \omega_l^2 = (2\pi\nu_l)^2 \quad (2)$$

where ν_l represents the frequency.

2.2. Elastic network model

As described by Tirion's ENM [27], the "classical" molecular mechanics semi-empirical potential energy function used in the all-atom force field is replaced by an all-atom Hookean potential:

$$E_p = \frac{1}{2} \sum_{\substack{r_{ij} < R_c \\ r_{ij}^0 < R_c}} \gamma(r_{ij} - r_{ij}^0)^2 \quad (3)$$

where r_{ij} is the distance between two atoms i and j , r_{ij}^0 is the distance between these atoms in the initial three-dimensional structure, γ is a force constant assumed to be the same for all interacting pairs which refers to the spring constant of the potential, and finally R_c is an arbitrary cutoff parameter, beyond which interactions are not taken into account. The normalized eigenvectors $\{U_{il}\}_{i=1,3N}$ of its Hessian matrix are the related normal modes.

2.3. Principal components analysis

Principal component analysis (PCA) is a mathematical technique that uses an orthogonal transformation to convert a correlated dataset into a set of linearly uncorrelated vectors called principal components [12,26,38]. To accomplish this, the covariance matrix (\mathbf{Cov}) is constructed. In the context of molecular simulations the Cartesian position (q) is used to construct the covariance matrix:

$$Cov_{ij} = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle \quad (4)$$

The principal components are determined by diagonalizing the covariance matrix:

$$\mathbf{V}^T \mathbf{Cov} \mathbf{V} = \mathbf{\Lambda} \quad (5)$$

The resulting diagonal elements of $\mathbf{\Lambda}$ are the variance eigenvalues.

2.4. Comparison of ENM and FF-NMA

To quantify the similarity between the normal mode vectors and elastic network normal mode vectors, the overlap between the k th ENM mode and the l th FF-NMA mode was calculated by inner-product [44,45]:

$$S_{kl} = \sum_{i=1}^{3N} U_{ik} W_{il} \quad (6)$$

where the columns of \mathbf{U} and \mathbf{W} are the respective eigenvectors. In addition, the cumulative overlap for n normal modes was calculated as the sum over squared overlaps, starting with the first non-trivial normal mode:

$$CO_n = \sum_{k=7}^n S_{kl}^2 \quad (7)$$

2.5. Correlation analysis

The correlation matrix, \mathbf{C} , can be readily calculated based on the eigenvalues and eigenvectors [46],

$$C_{ij} = \frac{\langle \Delta x_i \Delta x_j \rangle}{\langle \Delta x_i \Delta x_i \rangle^{1/2} \langle \Delta x_j \Delta x_j \rangle^{1/2}} = \frac{\sum_{k=1}^n (W_{ik} W_{jk}) / \Omega_{kk}}{\left(\sum_{m=1}^n (W_{im} W_{im}) / \Omega_{mm} \right)^{1/2} \left(\sum_{l=1}^n (W_{jl} W_{jl}) / \Omega_{ll} \right)^{1/2}} \quad (8)$$

where Δx_i and Δx_j indicate components of the mass-weighted displacement vectors of any two atoms or residues i and j , Ω_{ll} denotes the Hessian eigenvalue of the l th mode, and W_{il} is the (normalized) eigenvector coefficient of atom or residue i in mode l . The summation is over a given n -dimensional subspace of the normal mode spectrum. The correlation coefficient r between two correlation matrices was calculated as:

$$r = \frac{\sum_i \sum_j (C_{ij} - \bar{C})(C'_{ij} - \bar{C}')}{\left(\sum_i \sum_j (C_{ij} - \bar{C})^2 \right)^{1/2} \left(\sum_i \sum_j (C'_{ij} - \bar{C}')^2 \right)^{1/2}} \quad (9)$$

where \bar{C} and \bar{C}' are the mean values of matrices \mathbf{C} and \mathbf{C}' .

2.6. Flexibility analysis

The flexibility was described in terms of the root mean square fluctuations (RMSFs) [47]. The fluctuation of the i th atomic coordinate, Δq_i , can be calculated from the eigenvectors and eigenfrequencies,

$$\langle \Delta q_i^2 \rangle = \frac{k_B T}{m_i} \sum_{l=1}^n \frac{W_{il}^2}{\omega_l^2} \quad (10)$$

where m_i is the mass of the atom corresponding to the i th atomic degree of freedom, k_B is Boltzmann's constant, the temperature T in the simulation is 300 K, n is the number of normal modes included, and W_{il} is the i th component in the l th eigenvector with angular frequency ω_l . On the other hand, the RMSF based on MD is defined as:

$$\langle (\Delta \mathbf{R}_i)^2 \rangle = \frac{1}{N} \sum_{m=1}^N (\mathbf{R}_i(t_m) - \langle \mathbf{R}_i \rangle)^2 \quad (11)$$

where $\mathbf{R}_i(t_m)$ is the position vector of atom i at time step m , and N is the number of samples. Note that Δq_i in Eq. (10) refers to individual Cartesian components, while \mathbf{R}_i denotes the position in coordinate space.

2.7. Computational details

The supplementary information provides a description of the input data and the programs, that were used [48–52].

3. Results

3.1. Correlation between ENM and FF-NMA

The first six normal modes are spurious rotations and translations, with frequencies which are nearly zero, with an accuracy of at least 10^{-3} cm^{-1} . These modes are skipped and the spectrum is renumbered from one. The results of vibrational frequencies and normal modes from FF-NMA and ENM were compared in terms of cross plots of eigenfrequencies and the inner-products between low-frequency (the first 100) eigenvectors. A nearly straight line is observed between the frequencies of the first 100 modes of

ENM and FF-NMA for both proteins. The correlation coefficients for Dronpa and the β clamp are 0.9897 and 0.9968, respectively. This plot is used to rescale the ENM parameter γ , so that the slope attains 45° . The rescaled value for the γ force constant amounts to 14 and 11 kcal/(mol \AA^2) for Dronpa and the β clamp, respectively. After rescaling, the values of the FF-NMA and ENM frequencies are in excellent agreement for the first mode (6.9 cm^{-1} vs 7.0 cm^{-1} for Dronpa; 1.53 cm^{-1} vs 1.60 cm^{-1} for β clamp) till the 100th mode (27.7 cm^{-1} vs 26.0 cm^{-1} for Dronpa; 13.2 cm^{-1} vs 13.1 cm^{-1} for β clamp). In addition, a direct view of the overlap was obtained from a plot of the inner-product between corresponding pairs of eigenvectors (Fig. S1 in the supplementary material). In the case of the β clamp there is a close agreement between the first 30 eigenvectors; for Dronpa the diagram is much more scattered. This is explained by the higher onset of the Dronpa spectrum (vide infra).

3.2. The global motions

The lowest frequency modes often show the highest correlations and collectivity. Therefore, they usually describe highly collective motions which reflect the most flexible distortions of the protein. It is very interesting to visualize these modes, not only because they are highly symmetric but also are potentially relevant to biological function. Animations of the most important modes for Dronpa and the β clamp are available on line.

3.2.1. The global motions of Dronpa

Dronpa is a quite rigid protein, which is also indicated by the relatively high frequency of the first normal mode (6.9 cm^{-1}). In this frequency range, loop flexibility and 'in vacuo' side chain rotations are convoluted in the FF-NMA spectrum with the global modes. This contrasts with ENM, where the onset of the spectrum offers a neat view of the elastic modes of the cylinder. The ENM mode 1 is the twisting motion of a cylinder (Fig. S2(a)). This is a nearly pure rigid-body rotation around the pore axis, with the two ends of the cylinder molecule rotating in opposite directions. This shear motion is comparable to the lowest mode of gramicidin A, which is crucial to the gating mechanism of this membrane protein [18,24]. Modes 2 and 3 are degenerate modes, which show bending motions of the cylinder (Fig. S2(b)). Mode 7 corresponds to the breathing motion, which describes the expansion-contraction of the cylinder (Fig. S2(c)). The Poisson ratio of longitudinal elongation over transversal contraction is approximately 0.26. For polymers this ratio is usually assumed to be close to 0.33. The deviation from this value is due to the chiral β -barrel architecture of Dronpa.

It is worthwhile to determine how well these cylindrical motions can be represented by the low-frequency FF-NMA modes. To quantitatively compare the ENM modes to the FF-NMA modes, we calculated the cumulative overlap for each cylindrical mode in ENM with the subspace of the first 100 FF-NMA modes (Fig. S3). For the twisting motion of ENM mode 1, the corresponding FF-NMA modes are mode 1, 4, and 6, with squared overlaps of 0.148, 0.146 and 0.129, respectively. For the bending ENM mode 2, the corresponding mode is also FF-NMA mode 1, with a squared overlap of 0.351, and the dominant contribution to ENM mode 3 is FF-NMA mode 3, with a squared overlap of 0.457. The breathing motion of ENM mode 7 is smeared out over several FF-NMA modes; the most important ones are FF-NMA modes 5, and 9 with quite small squared overlaps of 0.150 and 0.140 respectively. In general, even the largest overlaps between these modes are small (below 50%). So, these cylindrical modes are convoluted in FF-NMA with loop motions and side-group rotations. This mixing occurs because the collective elastic modes are situated at relatively high frequency, which is indicative for the high rigidity of Dronpa. This is particularly true for the twisting and breathing. The first modes in both methods are also very special. The overlap between mode

1 in ENM and mode 5 in FF-NMA is zero, because the FF-NMA mode 5 displays the largest deformation of the waist of the cylinder while its two poles are relatively rigid. The FF-NMA mode 1 has the largest overlap both with mode 1 and mode 2 in ENM, so the motion of the first mode in FF-NMA is the mixture of twisting and bending motion. It should also be remarked that the cumulative overlaps over the spectral region considered all remain below 80%. Most of the remaining 20% is gained in the frequency region above 1200 cm^{-1} , which corresponds to the stretching vibrations. This is explained by the uniform treatment of all distance variations in ENM.

3.2.2. The global motions of β clamp

Unlike Dronpa, the β clamp has an open wheel like structure, and the β -sheet does not form a single interconnected structure but is fragmented in six rigid domains. The whole structure is much less rigid. As a consequence the spectrum starts off at a lower frequency (1.53 cm^{-1}) and the overlap between FF-NMA and ENM collective modes is more pronounced. This is also in line with our previous results on a similar toroidal protein, the trp RNA binding attenuation protein [22].

Both symmetry and topology [53,54] play a dominant role in this region of the spectrum. Because of the almost perfect sixfold symmetry of the force field, we can base the analysis of the circular modes on the vibrations of a hexagonal ring, as exemplified by the ring vibrations of benzene. There are three classes of vibrations [54], denoted as ω_1 , ω_2 , ω_3 . The first and most energetic class ω_1 contains the extensional in-plane stretchings, starting with the uniform breathing mode, which has ring quantum number $\Lambda = 0$. This mode is retrieved as ENM mode 10. The second class, ω_2 , concerns the flexural in-plane bending modes. Monopolar and dipolar bending modes are spurious since they correspond to in-plane rotation and translations, respectively. Hence the ω_2 class starts at ring quantum number $\Lambda = 2$, giving rise to quadrupolar elongations and contractions (Fig. 3(b)). These modes correspond to the degenerate pair of ENM modes 3–4. Quite remarkably one can also detect in the β clamp the next octupolar member of the flexural class, corresponding to $\Lambda = 3$. In benzene this mode is the $b_{1u}(\beta_{CC})$ vibration at 1010 cm^{-1} (Fig. 3(c)). It nicely corresponds to mode 11 in the ENM analysis. The third and most flexible class, ω_3 , contains the out-of-plane vibrations. The monopole and dipole members of this class are the out-of-plane translation and rotations, respectively, so that the vibrational modes also start at the quadrupolar out-of-plane vibrations. They are also the lowest modes of the ENM analysis, corresponding to the degenerate pair 1–2 (Fig. 3(a)). Interestingly, here too an octupolar excitation of the ring is observed, corresponding to $\Lambda = 3$. It is ENM mode 6. In benzene this mode is the $b_{2g}(\gamma_{CC})$ vibration at 685 cm^{-1} (Fig. 3(c)). In this way we have been able to account for all low-frequency modes, except for modes 5, 7, and 8. These modes do not change the circular shape, but are characteristic of toroidal topology. They thus constitute a separate class, which we will denote as ω_4 . The modes in this class correspond to internal whirling motions around the spine of the torus. The first member of this class corresponds to uniform motion where all parts of the torus rotate with the same handedness (Fig. 4(a)). This mode is uniform and therefore has ring quantum number $\Lambda = 0$. It is ENM mode 5, and is called the anapolar mode [53]. It was also observed in our previous study on the TRAP protein [22]. However quite remarkably, in the present case the almost perfect sixfold symmetry and the compact structure of the domains allows to detect also the first excitations of the anapolar mode: these correspond to the two degenerate modes 7 and 8 (Fig. 4b and c). These too are toroidal vibrations, but while in the ground anapole mode all parts of the torus rotate with the same handedness, in the first excitation opposite parts rotate with the opposite handedness, hence one part turns inwards and the opposite part turns outward.

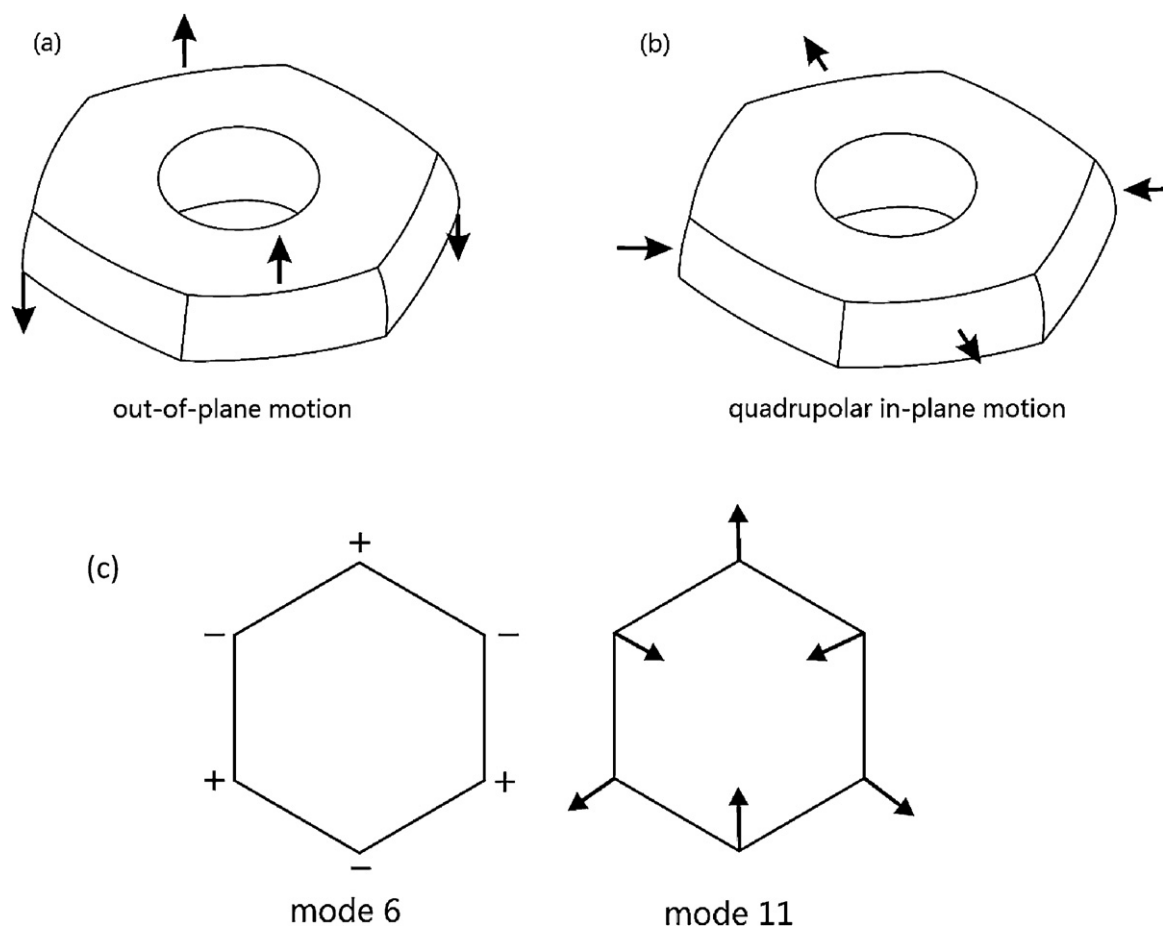


Fig. 3. The quadrupolar out-of-plane motion (a) and in-plane (b) motion of the ring, and the octupolar modes of benzene (c). Mode 6 is the b_{2g} mode, belonging to class ω_3 , mode 11 is the flexural b_{1u} mode, and belongs to class ω_2 .

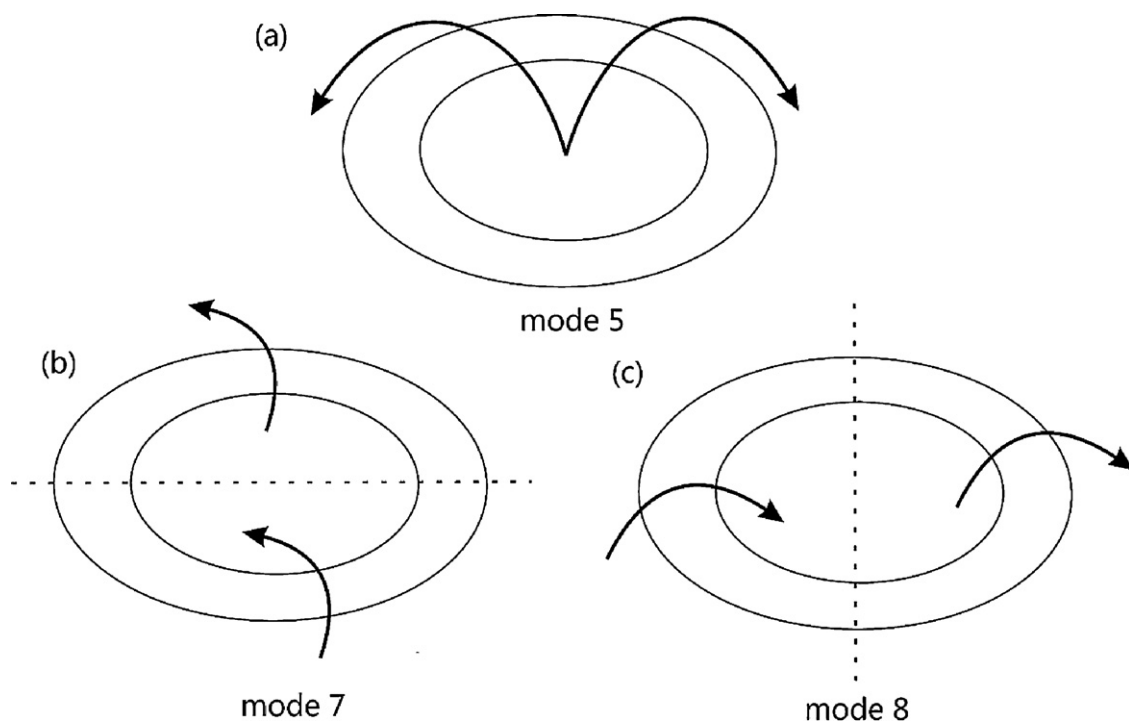


Fig. 4. The toroidal motions of the β clamp.

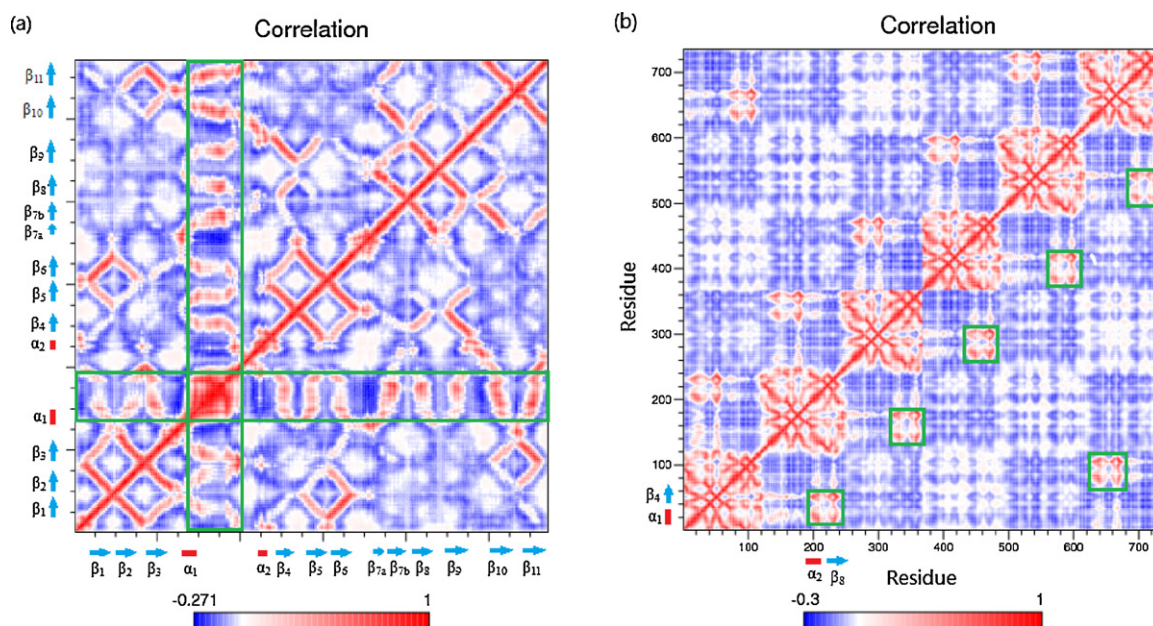


Fig. 5. Correlation maps of Dronpa (a) and β clamp (b) based on the first 100 ENM modes. In the map, red indicates the highly correlated motions, blue represents the anti-correlated motions and white is for the uncorrelated motions. The green boxes in (a) highlight the correlation between the chromophore and the other residues. In (b) they highlight the interdomain correlations between α_1 and α_2 , and between β_4 and β_8 . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

This gives rise to a nodal plane in between, which is characteristic of a ring quantum number $\Lambda = 1$. There are two solutions with perpendicular nodal planes, as shown in the figure. Interestingly for one of these modes the nodal plane runs through the dimer contacts. These out-of-plane motions may induce significant conformational flexibility [10]. The overlaps of these collective modes with the subspaces of the first 100 FF-NMA modes are shown in Fig. S4. We find mostly a one-to-one correspondence between modes calculated with ENM and with FF-NMA. ENM mode 1 matches FF-NMA mode 2 (the squared overlap is 0.587), while ENM mode 2 pairs up with FF-NMA mode 1 (the squared overlap is 0.519). ENM mode 5 and FF-NMA mode 5 pair up (the squared overlap is 0.415) while ENM mode 3 corresponds to FF-NMA mode 4 and vice versa (the squared overlaps are 0.623 and 0.574). The breathing mode is more convoluted; the corresponding modes are FF-NMA mode 9 and mode 12, with the squared overlap 0.165 and 0.117.

3.3. Correlations

Previous studies [22,55] argued that the first 100 normal modes are sufficient to assess the essential protein dynamics, even for large systems. While such a cutoff works well in practice, it is to a certain extent arbitrary, and the use of a complete energy-weighted normal mode space has been proposed recently [56]. In the present correlation maps limitation of the normal mode space to the 100 lowest modes is sufficient.

The ENM results for Dronpa are shown in Fig. 5(a). Although there is no close match between individual modes in ENM and FF-NMA, the collective characteristics are very similar using the two approaches (Fig. S5). The correlation coefficient between the two correlation matrices is 0.9168. When plotting these correlation maps, we replace the interior chromophore by all its carbon atoms, and represent the other residues by their C_α atoms. The correlation map well reflects the hydrogen network of the cylinder, which connects neighboring β -strands. While the all-atom ENM does not consider H-bonds explicitly, it nevertheless probes well the beta-barrel structures due to the use of a tolerant cutoff distance.

The red lines perpendicular to the main diagonal represent the correlation caused by the hydrogen bonds between neighboring anti-parallel β -strands, while the red lines parallel to the main diagonal represent the correlation caused by the hydrogen bonds between the parallel neighbor strands β_1 and β_6 , as well as the weaker correlation caused by indirect hydrogen bonds between next nearest neighbor β -strands. In addition, the correlations also link α -helices to β -strands, which may be caused by other weak bonds, such as hydrophobic interaction. The green boxes highlight the correlation between the chromophore and the other residues; it shows correlations with all 11 β -strands. Fig. S5(c) shows the correlation difference map obtained by the correlation matrix of ENM minus the correlation matrix of FF-NMA. In the ENM method, the effect of correlation between β -strands and α -helices is more pronounced.

In Fig. 5(b) similar ENM results are shown for the β clamp. The similarity of the correlation between ENM and FF-NMA is also reflected in the correlation coefficient between the two correlation matrices, which amounts to 0.9496 (Fig. S6). The correlation panel clearly shows that the β clamp has six modules, which each relate to one domain. Thus, the remarkable D_6 symmetry is evident in the correlation maps. In a correlation map, the intra-monomer correlations are visible on diagonal blocks, while the inter-monomer correlations are represented on off-diagonal blocks. Of all correlated motions, the intra-domain and – to a lesser extent – nearest neighbor correlations are the strongest. The reason for this lies in the way a domain is constructed. The hydrophobic core located between the two β -sheets keeps the overall shape of each domain intact, whereas adjacent domains are connected by strong hydrogen bonds between two neighboring β -strands, thereby allowing limited flexibility between them. The green boxes in Fig. 5(b) and Fig. S6 highlight the inter-domain correlation, namely, between α_1 and α_2 , and between β_4 and β_8 . In Fig. S6(c), one notes the stronger anti-correlations between the neighboring domains in ENM, and weaker anti-correlations between distant domains, as highlighted in the green box.

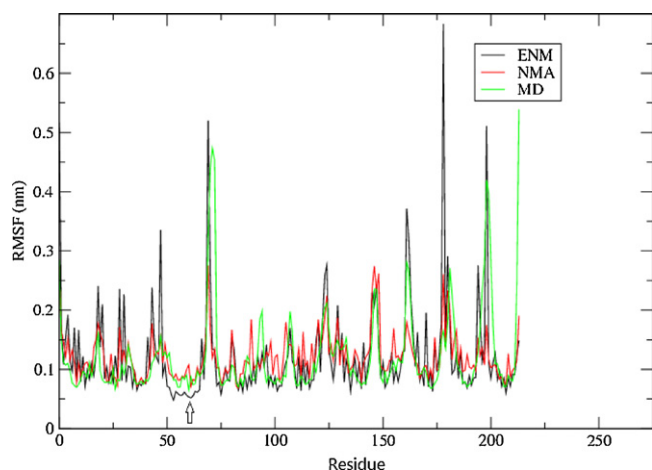


Fig. 6. The backbone RMS fluctuations of Dronpa by ENM (the black line), NMA (the red line) and MD (the green line). The RMSFs based on ENM and MD are rescaled so that the average of the fluctuations matches the average of the NMA fluctuation. The arrow indicates the location of the chromophore. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.4. RMS fluctuations

The comparison of backbone RMSF due to the lowest 100 modes of ENM and FF-NMA for Dronpa is shown in Fig. 6 to estimate the flexibility of proteins. Both profiles are comparable. The ENM force constant parameters were determined beforehand by fitting the spectral frequencies, but have to be rescaled again to reproduce well the vibrational amplitudes. The optimal value for the γ force constant of Dronpa is 22.7 as compared to 14 kcal/(mol Å²) from the spectral fit. For the case of Dronpa, we have also computed the dynamic RMSF values based on the last 12.5 ns of a 37.7 ns replica exchange MD simulation [8]. These results have been reduced by a factor of 1.77 to match the average FF-NMA fluctuation. Overall almost every peak in the FF-NMA data has counterparts both in the ENM and MD curves. Usually ENM fluctuations are most pronounced. The most flexible residues are located in the loops connecting the β -strands, which is to be expected. In contrast, the residues involved in the β -sheets of the Dronpa cylinder have small fluctuations due to tight H-bonding between the strands of β -barrel. The arrow in Fig. 6 indicates the location of the chromophore. In ENM, this appears to be the most rigid part of the protein, which is indeed a crucial factor to favor emission over non-radiative relaxation processes.

A similar comparison of the β clamp backbone RMSF derived by ENM and FF-NMA is shown in Fig. S7. The γ constant was rescaled to 20.8 as compared to 11 kcal/(mol Å²) from the spectral fit. Clearly the β clamp is the more flexible protein of the two: the RMSF baseline in Dronpa is at about 0.9 Å, as compared to nearly 1.6 Å for the β clamp. In Fig. S7 the RMSFs are averaged over the two monomers, and the fluctuation profiles of both methods are very similar. The most flexible residues are located at five of the six small connecting α -helices, the exception being the one at residues 143–145. In the ENM the amplitudes of peak fluctuations are mostly more pronounced, except for the peaks at two connecting loops (residues 94 and 175). In Ref. [57] the crystal structure of the β clamp subunit of *E. coli* DNA polymerase III was revisited at higher resolution of 1.85 Å, as compared to 2.4 Å in the original study of Kong et al. [3]. The comparison of the calculated RMSF to the B-factor derived from this more recent X-ray crystal structure is interesting. The RMSF analysis mostly validates the X-ray experiment, but also reveals additional dynamic regions not seen in the crystal study, such as the α -helix (residue 73) and the connecting loop (residue 277).

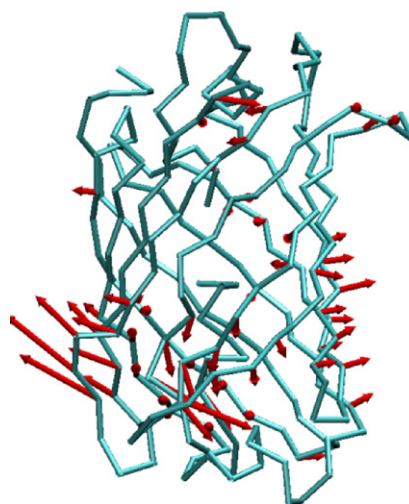


Fig. 7. Vector representation of the amplitude and direction of the C α displacements along the first PCA mode as a gating motion.

3.5. PCA of the dynamics of Dronpa

Although the close relationship between harmonic analysis in the low-frequency range and functional dynamics of proteins has been illustrated in several studies [12,24,39,55] it remains a remarkable coincidence since the normal modes are obtained in the strict harmonic approximation, while MD explores the full anharmonic potential energy surface. In a previous study of Dronpa, we performed a 37.5 ns replica exchange MD run [8]. From this run the last 12.5 ns were used to obtain the correlated motions by applying a PCA. In contrast to ENM, the all atom PCA spectrum starts off with anharmonic motions of the loops at the poles of the cylinder. These motions have the largest variance. They are also retrieved in the FF-NMA analysis. The global motions are found at lower variance eigenvalues where they are mixed with side chain movements. We computed the variance along the directions of the ENM-derived global motions and found 0.34, 0.45, 1.15 and 0.33 nm² for twisting, bending1, bending2, and breathing, respectively. For Dronpa, PCA eigenvalues ranging from 0.45 to 0.33 are found between modes 20 and 30. The large variance in bending mode 2 is noteworthy. While the bending motions are degenerate for the perfect cylinder and have nearly the same frequency in ENM, the fluctuations along these two modes in the MD simulation are quite different. This is seen in Fig. S8, where we plot the free energy along the twisting, bending, and stretching modes. The bending2 mode clearly has a much higher flexibility than the bending1 counterpart. On closer inspection this bending2 mode is found to be in the plane of strand β_7 , which is interrupted in the middle (Fig. 1), and might therefore bend more easily. While ENM is unable to distinguish the two bending modes to this extent, it nevertheless reproduced the correct resolution of the degeneracy level for the ideal cylinder. This indicates that ENM incorporated the small symmetry breaking due to the interruption of strand β_7 .

In order to detect the global modes of the cylinder, the PCA analysis was repeated on the C α atoms of the sheets and helices only, i.e. by eliminating loops and sidechains. In this case the PCA mode with the greatest variance turns out to be a localized gating motion which opens the barrel at the wedge, as shown in Fig. 7. This is of great functional importance, because it forms a communication channel between the chromophore and the solution, and explains how the environment could influence the luminescence. However, this barrel opening mode is not picked up by the FF-NMA analysis. Overlap calculations with a normal mode space which is projected onto the C α atoms, indicate that this mode is indeed dispersed over

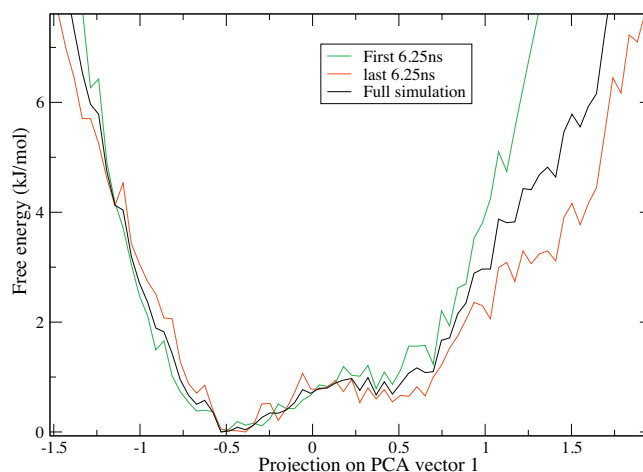


Fig. 8. The free energy along the PCA component corresponding to the cleft opening. To illustrate that the free energy is reasonable converged in the low energetic region, the simulation is divided into two.

a large spectral region (see Fig. S9). Additional to the fact that this motion is localized, the reason for this limitation of the FF-NMA method might be found in the anharmonic nature of the barrel opening motion. This is illustrated by the free energy plot along this mode in Fig. 8.

4. Discussion

4.1. Dronpa

Undoubtedly, in both proteins we have examined, there is a clear link between structure and global motions. Key elements in establishing this link were the presence of a high symmetry, and the relative rigidity of β -sheet segments. The natural accordion motions of these sheets usually occur at frequencies above 20 cm^{-1} [58], which implies that at lower frequencies these segments are moving as separate blocks. In Dronpa the β -sheet is rolled up to a cylindrical β -barrel, stabilized by a network of backbone hydrogen bonds, which is a common structure of GFP. The barrel forms one interconnected domain, which explains why the spectral onset is relatively high (7 cm^{-1}), as compared to protein structures with domain motions over hinges. In this spectral region, ENM clearly detects the elastic cylindrical rod behavior: bending, twisting and breathing motion, which are general features of cylindrical molecules [59]. They are also encountered in other polymers such as a single α -helix and DNA [60,61]. A distinctive feature of Dronpa though is the position of the twisting mode as the lowest frequency vibration. We attribute this to the chiral architecture of Dronpa.

Moreover when comparing the harmonic analysis to the dynamic behavior as revealed by a PCA analysis of the barrel MD, it must be concluded that the functionally important gating motion is not detected by the global analysis. This motion is quite localized at the cleft region in the barrel, and therefore will be masked by the global motions.

4.2. β clamp

As compared to Dronpa, the sliding clamp is a β -wheel, which is obtained by polymerization of six very similar domains. As a symmetric assembly built of identical subunits, a β -wheel is a toroidal structure with each subunit tied to its neighbors by hydrogen bonds. Such β -wheels are also a common feature of protein architecture, particularly in DNA polymerases. In the low-frequency modes, the β -wheel vibrates as an assembly of rigid bodies, in

marked contrast with the β -barrel. The reason is the existence of hydrogen bonds. As the analysis shows the low frequency region of the β clamp exhibits the elastic characteristics of a torus, for which four classes of vibrations can be distinguished. Three relate to deviations from the ring-like shape, while one is typical for toroids [62].

Dynamics studies of the processivity clamps, both in prokaryotes and eukaryotes, have concentrated on the ring opening process which is required to encircle the DNA chain [9–11]. According to simulations, the isolated ring in water seems to remain stable [9]. However in a recent study based on hydrogen exchange mass spectrometry, gradual deuteration of backbone proton sites in the hydrophobic clefts was observed [63]. This suggests that spontaneous ring opening may take place in solution, but with a rather long half life of 4 h. The ring opening could be triggered at the dimer interfaces. The excessive time scale for spontaneous ring opening prevents simulation of the dynamics. For the case of the eukaryotic variants, MD studies were carried out on open ring fragments [10,11], obtained by omitting one of the subunits. This approach was preferred to a biased molecular dynamics where the ring is driven to the open state by some external force. The ring opening dynamics, based on this strategy, are not directly comparable to the results of the present detailed elasticity study of the stable complex. Perhaps the most interesting outcome is the almost perfect hexagonal symmetry of the correlation plot in Fig. 5(b). The distinctions between the three domains and between the intramonomer and the dimer contact terms are almost absent in these plots. This confirms the symmetry of the tertiary structure, which is a highly conserved property of the clamps.

4.3. The relationship between ENM and FF-NMA

Both ENM and FF-NMA are based on the harmonic approximation to model interactions. Both can be used to identify the global motions. ENM is a helpful tool to estimate the distribution of global motions in FF-NMA by calculating overlaps. FF-NMA takes into account the detailed chemical structures, and is therefore less suitable to detect global modes because of mode mixing. This was especially true for Dronpa, where the global modes require more energy. Since ENM is an approximate model, it removes the atomistic details of FF-NMA and enhances the collective character of the soft modes. In contrast to ENM methods based on the C α atoms, the Tirion ENM method used here is all-atom based. It is therefore suitable to detect the detailed information of the flexibility and correlation, although it includes some more approximations. The

amplitude of fluctuations of loops, however, becomes larger when implementing the ENM approach. This is quite rational because Tirion's ENM is based on the elimination of entropic springs corresponding to the side chains, and consequently, the coarse-grained structures are more flexible than the original structure.

5. Conclusion

A comparative study of ENM and FF-NMA was used to reveal the harmonic properties of Dronpa and the β clamp protein. Families of β -barrel and β -wheel proteins incorporate two particular topologies of cylinder and torus. ENM can demonstrate the global motion of Dronpa as a chiral cylinder and the clamp as a toroidal ring in the low-energy end of the spectrum. The correlation diagrams for Dronpa and β clamp mainly highlight the intra-domain H-bonds, and – to a somewhat lesser degree – also the inter-domain H-bonds. The evident similarity of the correlation and RMSF between ENM and FF-NMA shows that the all-atom ENM method used here is able to describe the internal protein flexibility. A clear exception is the cleft opening mode of the Dronpa barrel which is an important functional feature of protein dynamics, but cannot be recognized in a harmonic approximation.

Although this study focuses on two specific proteins, β -barrels and β -wheels are common structures in GFP and DNA polymerases, which makes our results useful for the harmonic description of a large class of proteins.

Acknowledgements

Financial support from the Flemish Government through the concerted action scheme is gratefully acknowledged. G.H. thanks INPAC for a postdoctoral fellowship. S.M. is doctoral fellow of the FWO Science Fund.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmngm.2011.12.005](https://doi.org/10.1016/j.jmngm.2011.12.005).

References

- [1] P.G. Wilmann, K. Turcic, J.M. Battad, M.C.J. Wilce, R.J. Devenish, M. Prescott, J. Rossjohn, The 1.7 Å crystal structure of Dronpa: a photoswitchable green fluorescent protein, *J. Mol. Biol.* 364 (2006) 213–224.
- [2] K. Brejc, T.K. Sixma, P.A. Kitts, S.R. Kain, R.Y. Tsien, M. Ormo, S.J. Remington, Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 2306–2311.
- [3] X.-P. Kong, R. Onrust, M. O'Donnell, J. Kuriyan, Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp, *Cell* 69 (1992) 425–437.
- [4] Z. Kelman, PCNA: structure functions and interactions, *Oncogene* 14 (1997) 629–640.
- [5] V. Helms, T.P. Straatsma, J.A. McCammon, Internal dynamics of green fluorescent protein, *J. Phys. Chem.* 103 (1999) 3263–3269.
- [6] N. Reuter, H. Lin, W. Thiel, Green fluorescent proteins: empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulations of the wild type and S65T mutant, *J. Phys. Chem. B* 106 (2002) 6310–6321.
- [7] R. Nifosi, V. Tozzini, Molecular dynamics simulations of enhanced green fluorescent proteins: effects of F64L, S65T and T203Y mutations on the ground-state proton equilibria, *Proteins: Struct. Funct. Genet.* 51 (2003) 378–389.
- [8] S.L.C. Moors, S. Michielssens, C. Flors, P. Dedeker, J. Hofkens, A. Ceulemans, How is cis-trans isomerization controlled in Dronpa mutants? A replica exchange molecular dynamics study, *J. Chem. Theory Comput.* 4 (2008) 1012–1020.
- [9] D. Jeruzalmi, O. Yurieva, Y. Zhao, M. Young, J. Stewart, M. Hingorani, M. O'Donnell, J. Kuriyan, Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III, *Cell* 106 (2001) 417–428.
- [10] S.L. Kazmirski, Y. Zhao, G.D. Bowman, M. O'Donnell, J. Kuriyan, Out-of-plane motions in open sliding clamps: molecular dynamics simulations of eukaryotic

- and archaeal proliferating cell nuclear antigen, *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 13801–13806.
- [11] J.L. Adelman, J.D. Chodera, I.F. Kuo, T.F. Miller, D. Barsky, The mechanical properties of PCNA: implications for the loading and function of a DNA sliding clamp, *Biophys. J.* 98 (2010) 3062–3069.
- [12] B.R. Brooks, D. Janezic, M. Karplus, Harmonic analysis of large systems. I. Methodology, *J. Comp. Chem.* 16 (1995) 1522–1542.
- [13] Q. Cui, I. Bahar, Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems, Chapman and Hall/CRC, London, 2006.
- [14] L. Skjaerven, S.M. Hollup, N. Reuter, Normal mode analysis for proteins, *J. Mol. Struct. Theorchem.* 898 (2009) 42–48.
- [15] H. Valadie, J.J. Lacapere, Y.H. Sanejouand, C. Etchebest, Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis, *J. Mol. Biol.* 332 (2003) 657–674.
- [16] X. Cheng, B. Lu, B. Grant, R.J. Law, J.A. McCammon, Channel opening motion of alpha7 nicotinic acetylcholine receptor as suggested by normal mode analysis, *J. Mol. Biol.* 355 (2006) 310–324.
- [17] E.J. Bertaccini, E. Lindahl, T. Sixma, J.R. Trudell, Effect of cobratoxin binding on the normal mode vibration within acetylcholine binding protein, *J. Chem. Inf. Model.* 48 (2008) 855–860.
- [18] G. Miloshevsky, P. Jordan, The open state gating mechanism of Gramicidin A requires relative opposed monomer rotation and simultaneous lateral displacement, *Structure* 14 (2006) 1241–1249.
- [19] A. Szarecka, Y. Xu, P. Tang, Dynamics of heteropentameric nicotinic acetylcholine receptor: implications of the gating mechanism, *Proteins* 68 (2007) 948–960.
- [20] A. Taly, M. Delarue, T. Grutter, M. Nilges, N. Le Novère, P.J. Corringer, J.P. Changeux, Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism, *Biophys. J.* 88 (2005) 3954–3965.
- [21] N. Go, T. Noguti, T. Nishikawa, Dynamics of a small globular proteins in terms of low frequency vibrational modes, *Proc. Natl. Acad. Sci. U.S.A.* 80 (1983) 3696–3700.
- [22] G. Hu, S. Michielssens, S.L.C. Moors, A. Ceulemans, Normal mode analysis of trp RNA-binding attenuation protein: structure and collective motions, *J. Chem. Inf. Model.* 51 (2011) 2361–2371.
- [23] B.R. Brooks, M. Karplus, Harmonic dynamics of proteins: normal mode and fluctuations in bovine pancreatic trypsin inhibitor, *Proc. Natl. Acad. Sci. U.S.A.* 80 (1983) 6571–6575.
- [24] I. Bahar, T.R. Lezon, A. Bakan, I.H. Shrivastava, Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins, *Chem. Rev.* 110 (2009) 1463–1497.
- [25] V. Krasnenko, A.H. Tkaczyk, E.R. Tkaczyk, Ö. Farkas, K. Mairing, Vibrations-determined properties of green fluorescent protein, *Biopolymers* 78 (2005) 140–146.
- [26] S. Hayward, B.L. de Groot, Normal modes and essential dynamics, *Methods Mol. Biol.* 443 (2008) 89–106.
- [27] M.M. Tirion, Large amplitude elastic motions in proteins from a single-parameter atomic analysis, *Phys. Rev. Lett.* 77 (1996) 1905–1908.
- [28] K. Hinsen, Analysis of domain motions by approximate normal mode calculations, *Proteins* 33 (1998) 417–429.
- [29] I. Bahar, A.R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, *Fold. Des.* 1 (1997) 357–370.
- [30] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, *Biophys. J.* 80 (2001) 505–515.
- [31] E.J. Bertaccini, J.R. Trudell, E. Lindahl, Normal-mode analysis of the glycine alpha1 receptor by three separate methods, *J. Chem. Inf. Model.* 47 (2007) 1572–1579.
- [32] I. Bahar, Perspectives on: molecular dynamics and computational methods: on the functional significance of soft modes predicted by coarse-grained models for membrane proteins, *J. Gen. Physiol.* 135 (2010) 563–573.
- [33] I. Bahar, T.R. Lezon, L.W. Yang, E. Eyal, Global dynamics of proteins: bridging between structure and function, *Annu. Rev. Biophys.* 39 (2010) 23–42.
- [34] A. Hung, K. Tai, M.S.P. Sansom, Molecular dynamics simulation of the M2 helices within the nicotinic acetylcholine receptor transmembrane domain: structure and collective motions, *Biophys. J.* 88 (2005) 3321–3333.
- [35] W. Zheng, Anharmonic normal mode analysis of elastic network model improves the modeling of atomic fluctuations in protein crystal structures, *Biophys. J.* 98 (2010) 3025–3034.
- [36] S. Cheng, M.Y. Niv, Molecular dynamics simulations and elastic network analysis of protein kinase B (Akt/PKB) inactivation, *J. Chem. Inf. Model.* 50 (2010) 1602–1610.
- [37] S.A. Wieninger, E.H. Serpersu, G.M. Ullmann, ATP binding enables broad antibiotic selectivity of aminoglycoside phosphotransferase(3)-IIIa: an elastic network analysis, *J. Mol. Biol.* 409 (2011) 450–465.
- [38] L. Skjaerven, A. Martinez, N. Reuter, Principal component and normal mode analysis of proteins: a quantitative comparison using the GroEL subunit, *Proteins: Struct. Funct. Bioinf.* 79 (2011) 232–243.
- [39] E. Marcos, R. Crehuet, I. Bahar, On the conservation of the slow conformational dynamics within the amino acid kinase family: NAGK the paradigm, *PLoS Comput. Biol.* 6 (2010) e1000738.
- [40] M. Lu, J. Ma, The role of shape in determining molecular motions, *Biophys. J.* 89 (2005) 2395–2401.
- [41] F. Tama, C.L. Brooks, Symmetry, form, and shape: guiding principles for robustness in macromolecular machines, *Annu. Rev. Biophys. Biomol.* 35 (2006) 115–133.

- [42] M. Delarue, Y.H. Sanejouand, Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model, *J. Mol. Biol.* 320 (2002) 1011–1024.
- [43] M.M. Hingorani, M. O'Donnell, Toroidal proteins: running rings around DNA, *Curr. Biol.* 8 (1998) R83–R86.
- [44] S. Hayward, A. Kitao, N. Go, Harmonic and anharmonic aspects in the dynamics of PBT1: a normal mode analysis and principal component analysis, *Protein Sci.* 3 (1994) 936–943.
- [45] L. Yang, G. Song, A. Carriquiry, R.L. Jernigan, Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes, *Structure* 16 (2008) 321–330.
- [46] A.W. van Wynsberghe, Q. Cui, Interpreting correlated motions using normal mode analysis, *Structure* 14 (2006) 1647–1653.
- [47] A. van Wynsberghe, G.H. Li, Q. Cui, Normal-mode analysis suggests protein flexibility modulation throughout RNA polymerase's functional cycle, *Biochemistry* 43 (2004) 13083–13096.
- [48] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient load-balanced, and scalable molecular simulation, *J. Chem. Theory Comput.* 4 (2008) 435–447.
- [49] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G.M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J.M. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, *J. Comp. Chem.* 24 (2003) 1999–2012.
- [50] T.E. Cheatham 3rd, P. Cieplak, P.A. Kollman, A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat, *J. Biomol. Struct. Dyn.* 16 (1999) 845–862.
- [51] E. Lindahl, C. Azuara, P. Koehl, M. Delarue, NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis, *Nucleic Acids Res.* 34 (2006) 52–56.
- [52] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, *Proteins Struct. Funct. Genet.* 17 (1993) 412–425.
- [53] A. Ceulemans, L.F. Chibotaru, P.W. Fowler, Molecular anapole moments, *Phys. Rev. Lett.* 80 (1998) 1861–1864.
- [54] A. Ceulemans, I. Vos, The vibrations of annular and globular molecules – theory, *Mol. Phys.* 72 (1991) 1051–1080.
- [55] M.Y. Niv, M. Filizola, Influence of oligomerization on the dynamics of G-protein coupled receptors as assessed by normal mode analysis, *Proteins: Struct. Funct. Bioinf.* 71 (2008) 575–586.
- [56] D.S. Palmer, F. Jensen, Predicting large-scale conformational changes in proteins using energy-weighted normal modes, *Proteins* 79 (2011) 2778–2793.
- [57] A.J. Oakley, P. Prosser, G. Wijffels, J.L. Beck, M.C.J. Wilce, N.E. Dixon, Flexibility revealed by the 1.85 Å crystal structure of the sliding-clamp subunit of *Escherichia coli* DNA polymerase III, *Acta Crystallogr. D* 59 (2003) 1192–1199.
- [58] K.C. Chou, Low-frequency motions in protein molecules: beta-sheet and beta-barrel, *Biophys. J.* 48 (1985) 289–297.
- [59] T. Mory, H. Kokubo, H. Shimizu, M. Iwamoto, S. Oiki, Y. Okamoto, Normal mode analysis of polytheonamide B, *J. Phys. Soc. Jpn.* 76 (2007), 094801–1–094801–10.
- [60] A. Matsumoto, W.K. Olson, Sequence-dependent motions of DNA: a normal mode analysis at the base-pair level, *Biophys. J.* 83 (2002) 22–41.
- [61] I. Adamovich, S.M. Mijailovich, M. Karplus, The elastic properties of the structurally characterized myosin II S2 subdomain: a molecular dynamics and normal mode analysis, *Biophys. J.* 94 (2008) 3779–3789.
- [62] P.W. Fowler, A. Rassat, A. Ceulemans, Symmetry generalization of the Euler–Schlafli theorem for multi-shell polyhedra, *J. Chem. Soc. Faraday Trans.* 92 (1996) 4877–4884.
- [63] J. Fang, J.R. Engen, P.J. Beuning, *Escherichia coli* processivity clamp β from polymerase III is dynamic in solution, *Biochemistry* 50 (2011) 5958–5968.