Structural variability of C3larvin toxin. Intrinsic dynamics of the $\alpha/\beta$ fold of the C3-like group of mono-ADP-ribosyltransferase toxins

Miguel R. Lugo, Ravikiran Ravulapalli, Debjayoti Dutta and A. Rod Merrill

*Department of Molecular and Cell Biology, University of Guelph, Science Complex, Guelph, ON N1G2W1, Canada; Faculty of Medicine, Department of Biochemistry, University of Alberta, Edmonton, AB T6G 2H7, Canada

Communicated by Ramaswamy H. Sarma

Received 7 October 2015; accepted 17 November 2015

C3larvin toxin is a new member of the C3 class of the mono-ADP-ribosyltransferase toxin family. The C3 toxins are known to covalently modify small G-proteins, e.g. RhoA, impairing their function, and serving as virulence factors for an offending pathogen. A full-length X-ray structure of C3larvin (2.3 Å) revealed that the characteristic mixed $\alpha/\beta$ fold consists of a central $\beta$-core flanked by two helical regions. Topologically, the protein can be separated into N and C lobes, each formed by a $\beta$-sheet and an $\alpha$-motif, and connected by exposed loops involved in the recognition, binding, and catalysis of the toxin/enzyme, i.e. the ADP-ribosylation turn–turn and phosphate–nicotinamide PN loops. Herein, we provide two new C3larvin X-ray structures and present a systematic study of the toxin dynamics by first analyzing the experimental variability of the X-ray data-set followed by contrasting those results with theoretical predictions based on Elastic Network Models (GNM and ANM). We identify residues that participate in the stability of the N-lobe, putative hinges at loop residues, and energy-favored deformation vectors compatible with conformational changes of the key loops and 3D-subdomains (N/C-lobes), among the X-ray structures. We analyze a larger ensemble of known C3bot1 conformations and conclude that the characteristic ‘crab-claw’ movement may be driven by the main intrinsic modes of motion. Finally, via computational simulations, we identify harmonic and anharmonic fluctuations that might define the C3larvin ‘native state.’ Implications for docking protocols are derived.

Keywords: mono-ADP-ribosyltransferase toxins; NAD$^+$-binding protein; Paenibacillus larvae; Elastic Network Models; principal component analysis; molecular dynamics

1. Introduction

Bacterial exotoxins are an important class of virulence factors (Popoff, 2005; Yates, Jørgensen, Andersen, & Merrill, 2006) produced by bacterial pathogens that are responsible for many effects in eukaryotic cells observed upon bacterial invasion and infection. Notably, diphtheria toxin (DT) is a classic example of a single toxin that causes a disease (diphtheria). Many mART toxins are AB toxins that bind to the eukaryotic target cell membrane with a receptor-binding domain (B-subunit) and deliver a second moiety (A-subunit) into the host cytoplasm; others consist of an A-only domain that must be delivered via a bacterial secretion system. A large group of exotoxins are known as the mono-ADP-ribosyltransferase (mART) toxin (Deng & Barbieri, 2008). These toxins are bacterial enzymes that act to kill/maim target eukaryotic cells by covalent modification of an essential protein (usually) within the host organism (Simon, Aktories, & Barbieri, 2014). In general, the cellular targets for mART toxins are often key regulators of cell function and include (i) G proteins, including eEF2, EF-TU, Ras, Rho, and Gs$\alpha$, (ii) actin, (iii) kinase regulators, and (iv) RNA recognition domains. Remarkably, some binary toxins from *Clostridium* species have even been used as ‘molecular Trojan horses’ to deliver cargo into eukaryotic cells (Barth & Stiles, 2008).

These enzymes bind NAD$^+$, facilitate the scission of the glycosidic bond (C–N) between nicotinamide and the N-ribose of NAD$^+$, and transfer the ADP-ribose group to a specific target protein (or suitable macromolecule) (Yates et al., 2006; Zhou et al., 2004). In addition, this enzyme family also possesses NAD$^+$ase or glycohydrolysis activity (GH), but the physiological relevance of this activity is not known (Beattie, Prentice, & Merrill, 1996; Collier, 2001; Yates & Merrill, 2005). The covalent modification of the target protein often has a dramatic effect on its function and can act as an ‘on’ or ‘off’ switch for activity. For example, DT modifies elongation factor-2 (eEF2), which blocks its function on the ribosome, thereby inhibiting protein synthesis in the host cell (Collier, 2001). Cholera toxin acts on the Gs$\alpha$ protein, trapping this protein in its active conformation, perpetually activating a signaling pathway, causing massive loss of body fluids and often death (Fan et al., 2004).

*Corresponding author. Email: rmerrill@uoguelph.ca

© 2016 Informa UK Limited, trading as Taylor & Francis Group
The C3 toxin class within the mART toxin family is known to modify the small G-proteins, RhoA, RhoB, and RhoC, at the Asn41 residue (Vogelgesang, Pautsch, & Aktories, 2007). Recently, C3larvin toxin was characterized as a new member of the C3 class of mART toxins and a possible virulence factor from Paenibacillus larvae, which is the causative agent of American Foulbrood in honey bees (Krska, Ravulapalli, Fieldhouse, Lugo, & Merrill, 2015). C3larvin was shown to be cytotoxic to yeast when expressed in the cytoplasm and catalytic variants of the enzyme lost the ability to kill the yeast host. C3larvin targets RhoA as a substrate for its transferase reaction, with $K_M = 34 \pm 12 \mu M$ and $17 \pm 3 \mu M$ for NAD+ and RhoA, respectively, and with $K_i = 11 \pm 2 \mu M$ for a small competitive inhibitor, M3 (Krska et al., 2015). The crystal structure of C3larvin was solved to 2.3 Å resolution and it was shown to have a different mechanism of cell entry from other C3 toxins. C3larvin toxin possesses a mixed α/β-fold with a central β-sheet core formed by a five-stranded βα-sheet ($\beta_1$, $\beta_2$, $\beta_4$, $\beta_7$, and $\beta_8$) in a perpendicular and antiparallel disposition to a three-stranded βα-sheet ($\beta_3$, $\beta_5$, and $\beta_6$) — typically seen for catalytic domains of ADP-ribosyltransferase toxins (Figure 1). In this toxin, both β-sheets are topologically flanked by an N-terminal helical bundle of four α-helices (α1–α4), and by a helical motif (H-motif) with two additional α-helices (α5α and α6) that connect $\beta_1$ and $\beta_2$. Based on the topological similarity of C3larvin among C3-like toxins, the NAD+–binding pocket is formed in the cleft between both β-sheets, and delimited by the $\alpha_3$ helix and a $\alpha_5\alpha$ at the H-motif. In addition, two exposed loops contact and stabilize the interaction with NAD+ and the protein substrates: (i) the ADP-ribosylation turn–turn (ARTT) loop, which joins the $\beta_5$ and $\beta_6$ strands; and (ii) the phosphate–nicotinamide (PN) loop, which joins the $\beta_3$ and $\beta_4$ strands (Figure 1).

Due to thermal fluctuations, a protein undergoes a broad spectrum of motions on different timescales: from femtoseconds (bond vibrations) to microseconds (loop and subdomain motions), exhibiting ‘equilibrium dynamics’ that sample all the accessible microstates under physiological conditions (Bahar, Lezon, Bakan, & Shrivastava, 2010). The random conformational changes that typify a protein in its ‘native state’ reflect mainly cooperative motions ‘intrinsically accessible’ to the protein (Hall, Kaye, Pang, Perera, & Biggin, 2007; Hayward & Go, 1995; Kitao & Go, 1999). These modes of fluctuations are defined largely by the structure dynamics encoded in the folded structure (Bahar, Lezon, Yang, & Eyal, 2010). The propensity of a protein to sample its
‘equilibrium ensemble,’ by particular collective fluctuations, has been related to its ability to achieve its ‘biological function’ (Bahar, Attilgan, Demirel, & Erman, 1998). Thus, describing the intrinsic dynamics of a protein in its native state is a first and important step in the goal of understanding how a protein achieves its function. In this sense, (i) principal component analysis (PCA) on an ensemble of experimental X-ray or NMR structures (Yang, Eyal, Bahar, & Kitao, 2009), (ii) essential dynamic analysis (EDA) on an ensemble of conformations originating from any computational simulations (Amadei, Linssen, & Berendsen, 1993; Daidone & Amadei, 2012), or (iii) theoretical approaches such as the anisotropic network model (ANM) on a reference structure (Bahar & Rader, 2005) is able to characterize the equilibrium dynamics of a protein by identifying the dominant features of the structural variability.

Structural deformations observed in some X-ray and NMR structures between bound and unbound forms of the receptor provide an invaluable opportunity to assess structural changes associated with or induced by ligand binding with the receptor (Bahar, Chennubhotla, & Tobi, 2007; Meireles, Gur, Bakan, & Bahar, 2011). However, this needs to be carefully considered by working with an unique or limited number of receptor structures since it is easy to misconstrue the conformational effects induced by substrate(s) and inhibitor(s) upon binding. For example, based on X-ray structures of free and NAD⁺-bound forms of C3bot1 toxin from Clostridium botulinum, it was concluded that NAD⁺ triggers a particular conformational change between two states of the ARTT loop (Ménétrey et al., 2002). However, more recent structures revealed that the loop can adopt both states, irrespective of its ligation state (Ménétrey, Flatau, Boquet, Ménez, & Stura, 2008). Separately, the authors observed a variable ‘flexibility’ between apo and NAD⁺-bound C3bot1 structures (e.g. 1GZE and 1GZF) (Ménétrey et al., 2002). However, differences were also shown for molecule D with respect to other molecules in the asymmetric unit of apo C3bot1 (e.g. 2C89) (Ménétrey et al., 2008). Moreover, it was found that the flexibility of three C3bot1 molecules of a different crystal form was quite variable, irrespective of the NAD-bound state (e.g. 2C8E and 2C8F) (Ménétrey et al., 2008). The authors speculated that the magnitude of this movement is probably higher in solution and might be related to clasp-and-release cycles of substrate(s) and product(s), and therefore influenced by contact with the protein target.

Altogether, there is a need for compiling a greater number of structures, either WT or catalytic variants, from different experimental sources and/or crystallization conditions and/or ligation states in order to understand the structural variability of the toxin in the free and bound states. In this work, we provide two new experimental conformations of C3larvin toxin. We perform a systematic study of the dynamics of the C3larvin toxin by first analyzing the experimental variability followed by making theoretical predictions and computational simulations in order to define the C3larvin ‘native state.’ We also analyze C3larvin structures in the context of the known C3bot1 conformations, obtaining insights concerning the origin and characteristics of the protein flexibility. Implications for docking protocols are derived.

2. Methodology

2.1. C3larvin crystallization, data collection, and refinement

The C3larvin recombinant protein was expressed and purified as described earlier (Krska et al., 2015). The purified protein was concentrated to 12-mg/ml and was dialyzed against 50 mM Tris HCl (pH 7.5) and 150 mM NaCl. Prior to crystallization, the protein sample was mixed with 400 μM M3 inhibitor (N-[3R]-1-(1H-pyrazolo[3,4-d] pyrimidin-4-yl)piperidin-3-yl-methyl] methane-sulfonamide) and incubated for 1 h. Crystallization trials were performed by the hanging-drop vapor diffusion method with commercial screens from Hampton Research (USA). Diffraction-quality crystals were obtained in .2 M MgCl₂:6H₂O, .1 M Tris HCl (pH 8.5), and 30% w/v PEG 4000. Prior to flash-freezing, crystals were briefly soaked with a cryoprotectant solution that contained mother liquor supplemented with 15% glycerol.

Crystals were sent to the Canadian Light Source at the Canadian Macromolecular Crystallography Facility (beamline 08ID-1) for X-ray diffraction data collection. A total 360° diffraction images were collected with a .5° oscillation range. Images were processed in the P212121 space group with XDS (Kabsch, 2010). Phases were obtained using the C3larvin structure (PDB 4TR5) with MOLREP (Vagin & Teplyakov, 2010). The resulting electron density was weak and hence was subjected to density modification and solvent flattening using PARROT (Zhang, Cowtan, & Main, 1997). The atomic structure was further refined against 1.9 Å data, with initial cycles of two-fold NCS restraints in REFMAC (Winn et al., 2011). The final model was built using iterative cycles of model building in COOT (Emsley & Cowtan, 2004), and refinement in Phenix with TLS restraints (Adams et al., 2002). The refined structure was deposited in the Protein Data Bank with the code 5DZQ; Table S1.

For the analysis of 5DZQ in this paper, the backbone trace of the dimer in the asymmetric unit (chains A and O) was completed in COOT (Emsley & Cowtan, 2004) as follows: in chain O, although there is a weak continuous electron density, Tyr153 was built based on the allowable peptide chain geometry. In the same chain, the backbone containing residues 1–5 was built based on continuous tracing of the weak electron density. In chain A, residues Lys149–Tyr157 were fit into weak electron
density based on its NCS partner. Residues 1–4 were not built because of the ambiguous electron density. All fitted residues were further validated for clashes and proper stereochemical geometry.

2.2. Force field setting and structure preparation

Protein preparation, molecular mechanics calculations, and protein rendering were performed using the computational suite Molecular Operative Environment (MOE) release 2014.08 (Chemical Computing Group Inc, Montreal, CA). The force field employed was the AMBER12 parameters set (ff12). For the implicit solvent model, the generalized Born-volume integral (GB-VI) formalism was employed, with dielectric \( \epsilon_{\text{sol}} = 2 \) for the interior of the protein and \( \epsilon_{\text{sol}} = 80 \) (water) for the exterior. A switching, cut off function was defined between 10 and 12 Å for non-bonded interactions. The apo C3larvin structure (PDB: 4TR5) was loaded and truncated by deleting the first 22 residues, leaving the full-length, wild-type toxin (Met\(^1\)–Pro\(^{19}\)), ions, and crystallographic waters molecules. The MOE Protonate3D module was used to assign the ionization states and tautomers of waters molecules. The MOE Protonate3D module was used to assign the ionization states and tautomers of waters molecules. The energy function was updated to vacuum (\( \epsilon_{\text{sol}} = 1 \)), under periodic boundary conditions (chain A and O) of the asymmetric unit of a new X-ray structure herein reported (PDB 5DZQ); (ii) a C3bot1 X-ray ensemble from 40 C3bot1 X-ray structures in 12 PDB files of WT (PDB: 1G24, 1GZF, 2C89, and 2C8A) and diverse catalytically altered variants (PDB: 1GZE, 2C8B, 2C8C, 2C8D, 2C8E, 2C8F, 2C8G, and 2C8H), for apo and holo forms; and (iii) a simulated ensemble of 18,000 conformations resulting from snapshots taken at 5 ps intervals during a 90 ns of the MD trajectory for the 4TR5 structure (see MD Simulation). This ensemble was split into a building model ensemble with 16,000 conformations (80 ns), and a cross-validation ensemble with 2000 conformations (10 ns).

For the optimized superposition of the ensemble of a particular set of \( N \) C\(_{\alpha}\)-atoms, an iterative superposition, tierpose, was implemented that gave an unique solution that minimizes the average RMSD in the ensemble. Briefly, the \( i \)th member of the ensemble, i.e. \( i \)th conformation, is represented by a 3\( N \)-dimensional column vector as

\[
R_i = [r_{1,i} r_{2,i} r_{3,i} \ldots r_{N,i}]^T
\]

with \( r_{p,i} = [x_p y_p z_p]_i \) being the 3-D position vector of the \( p \)th C\(_{\alpha}\)-atom in the \( i \)th configuration. For an ensemble of \( M \) configurations, and average 3\( N \)-dimensional column vector is defined as

\[
\bar{R} = \frac{1}{M} \sum_{i=1}^{M} R_i = [\bar{r}_1 \bar{r}_2 \bar{r}_3 \ldots \bar{r}_p]^T
\]

The RMSF for the \( p \)th position is calculated over all the \( M \) conformations by

\[
\text{RMSF}_p = \langle \Delta r_p^2 \rangle^{1/2} = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (r_{p,i} - \bar{r}_p)^2}
\]
The RMSD for the $i$th configuration is calculated over all the $N$ atoms by

$$\text{RMSD}_i = \langle \Delta r_i^2 \rangle^{1/2} = \sqrt{\frac{1}{N} \sum_{p=1}^{N} (r_{p,i} - r_{p})^2}$$

while the average RMSD or RMSF for the ensemble is defined by

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \langle \Delta r_i^2 \rangle} = \text{RMSF} = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \langle \Delta r_i^2 \rangle}$$

The interpose routine proceeded as follows: (i) firstly, a random reference set of coordinates was selected; (ii) each member of the ensemble was then pairwise superimposed onto the reference structure by a rigid body translation and rotation to minimize the average RMSD [Equation (5)]; and (iii) the average coordinates were then calculated and set as a new reference structure. Steps (ii) and (iii) were iteratively performed until the RMSD between two successive average coordinates was lower than an arbitrary threshold (usually 0.001 Å).

The conformational change, $\Delta \mathbf{R}_j$, on the $\alpha$ trace of two conformers, $\mathbf{R}_i$ and $\mathbf{R}_j$, is defined by the deformation $3N$-dimensional vector

$$\Delta \mathbf{R}_j = \mathbf{R}_j - \mathbf{R}_i = [\Delta r_1, \Delta r_2, \ldots, \Delta r_N]^T$$

with $\Delta r_p = [x_p y_p z_p - [x_p y_p z_p]]_j$, for the $p$th residue or node.

The overlap between the $k$th ANM mode and the $c$th PCA mode, $O_{k,c}$, is given by the correlation cosine defined by the respective eigenvectors

$$O_{k,c} = \frac{|v_k \cdot u_c|}{\left| \sum_{p=1}^{N} v_{p,k}^2 \right|^{1/2} \left| \sum_{p=1}^{N} u_{p,c}^2 \right|^{1/2}}$$

The cumulative overlap between $n$ ANM modes and the $c$th PCA mode, $O_{\sum n,c}$, measures the degree of concordance between the subset of low-frequency ANM modes and a particular $c$ PCA mode by

$$O_{\sum n,c} = \left[ \sum_{k=1}^{n} (O_{k,c})^2 \right]^{1/2}$$

The generation of alternative conformations from an original coordinate $\mathbf{R}^0$, along the normalized directional vector of the $k$th ANM mode, can be calculated by

$$\mathbf{R}_i = \mathbf{R}^0 \pm s_i \left( \frac{k_BT}{k_c} \right)^{1/2} u_k$$

with $s_i$ a scalar to reproduce a particular RMSD.

3. Results

3.1. Heterogeneity of C3larvin X-ray data-set

A new crystal form of C3larvin was obtained with two molecules (chains A and O) in the asymmetric unit (PDB: 5DZQ, Table S1). $\alpha$-traces of both chains were made complete to 187 residues for chain A (Lys$^5$–Pro$^{191}$) and to 190 residues for chain O (Met$^1$–Gln$^{190}$). Figure 1 depicts the overall topology of the C3larvin X-ray structure based on the PDB entry 4TR5 (Krska et al., 2015). The superposition of two new conformations onto the 4TR5 structure, based on the common $\alpha$ trace solved for the three molecules (segment Lys$^5$–Gln$^{190}$), reports an average $\alpha$RMSD of 1.28 Å for 186 residues. The chains A and O superimpose well ($\alpha$RMSD = 1.67 Å), with only a slight difference at the N-terminus of $\alpha_1$ and at the ARTT loop (Figure 2(a)). The heterogeneity revealed by this small set of structures includes, in addition to the expected exposed loop(s) and terminal segments, divergence at the $\alpha_3$-$\alpha_4$ helices and at the H-motif since the 4TR5 conformation is more open than the new structures. This relative subdomain shift deserves attention since these motifs are part of the NAD$^+$-binding pocket (Krska et al., 2015), and it has been described as a ‘crab-claw’ movement on the C3bot1 toxin (Ménétrey et al., 2002, 2008). In this sense, it is important to understand the origin and characteristics of the flexible structural elements within the C3-like mART toxins.

To characterize the conformational variation among these structures, a PCA was performed on the superimposed structures. The PCA reported two principal collective variables (principal component 1 and 2, pc1 and pc2) that partition the total variance of the $\alpha$ displacement by 80 and 20%, respectively. The pc1 segregates the 4TR5 molecule from the A/O chains by an average $\alpha$RMSD of 1.5 Å, while pc2 discriminates the chain A from the chain O by ~1.0 Å (Figure 2(b)). In absolute terms, 4TR5 and chain A are spanned over the pc1 axis by 22.25 Å (i.e. cumulative displacement along the pc1), but only 3.32 Å between chain A and chain O; these two latter chains are spanned along the pc2 with a total displacement of 11.86 Å. Figure 2(c) shows the mean square fluctuation of each $\alpha$ for the data-set according to the two modes of deformation, with an average $\alpha$RMSF of .72 and .36 Å for pc1 and pc2, respectively. Both modes describe spikes of variation at the N-terminus of $\alpha_1$ and at the ARTT loop; but only pc1 contains a significant variation at the $\alpha_3$-$\alpha_4$ segment, at the H-motif, and at the $\beta_7$-$\beta_8$ segment. Figure 2(d) shows the structural changes associated with each PC mode of deformation by depicting the 3-D directional vectors on the $\alpha$ trace of 4TR5 as the template. A question that arises is how heterogeneous are these three structures? In other
words, is this data-set representative of the structural variability of the protein, or rather, are they simply just three members of a larger set of conformations that might define the ‘native state’ of the C3larvin toxin?

3.2. Thermal fluctuations of the C3larvin structures
The three C3larvin X-ray structures gave various B-factor profiles (Figure 3(a)). Table 1 shows their correlation, with chain A the most divergent in terms of the B-factor distribution. The low correlation in the mobility between chain A and O is remarkable, considering the low αRMSD between both structures. Furthermore, if residues mainly responsible for the structural variability between chain A and O (i.e. the ones with the highest displacement along pc2) are removed, it is surprising that the correlation for the B-factors between these two structures is almost nil. Table 1 includes the correlation coefficient between the experimental B-factor profiles (Figure 3(a), colored lines) and the total fluctuation profile in the X-ray data-set (Figure 3(a), black line). B-factors of chain O correlate best \((r = .65)\), following closely the fluctuations of the H-motif, PN, and ARTT loops. Indeed, disregarding the first 50 residues \((\alpha_1-\alpha_4\text{ segment})\), the correlation increases to \(r = .82\). In contrast, the X-ray variability found in the first 50 residues correlates better with the B-factors of chain A \((r = .65)\). The lowest overall correlation corresponds to the 4TR5 structure.

The GNM was implemented on 186 nodes corresponding to the \(C_\alpha\)-atoms of each structure of the ensemble (Figure 3(b)). The mean square fluctuation distribution calculated by the GNM for each structure is...
plotted in Figure 3(c). The high pairwise correlation (all \( r > .94 \), Table 2) reflects the dependence of the calculated fluctuations on the overall folded topology rather than on local structural details. For example, it is notable that there are structural differences in the \( \alpha_2-\alpha_4 \) segment and the ARTT loop backbone among the X-ray structures (residues 25–50 and around 160; Figure 3(a), colored lines), while a much better match was observed among the GNM fluctuation profiles. The highest correlation between fluctuations by GNM and the experimental \( B \)-factors corresponds for the 4TR5 structure, with \( r = .56 \) (Table 2, last column). Thus, considering also the low correlation with the experimental fluctuation (Table 1), the 4TR5 conformation was selected for a
further detailed analysis. Figure 3(d) shows the 4TR5 GNM variance distribution, where the first 25 (out of 185) modes were able to account for ~60% of the total variance in the positional fluctuation.

Figure 4(a) shows the overlay of the GNM fluctuations and the B-factor profiles of the 4TR5 structure. A closer inspection reveals regions where the GNM fluctuations are overestimating the mobility – at the \( \alpha_{3/4} \)–PN, and \( \beta_{7/8} \) loops. In particular, the \( \alpha_{3/4} \) loop is restricted in the 4TR5 crystal structure partially due to an H-bond between the C–O backbone of Gly44 with a Lys residue of a symmetry-related molecule (Figure 4(b)). Similarly, the complete PN loop (which adopts a \( \beta \)-strand configuration in the crystal lattice) and the C-terminus of the \( \beta_7 \) strand are stabilized reciprocally by intermolecular contacts with a symmetry-related molecule (Figure 4(b)).

Some residues with low-experimental B-factors belong to a hydrophobic core formed at the interface between the \( \alpha_2 \)–\( \alpha_4 \) segment and the \( \beta_6 \)-sheet. This core includes two conserved aromatics in the C3-group, Tyr25 and Tyr116, which in combination with Phe52 form a stable coordinated triad that, along with the hydrophobic residues Val22, Leu36, Ile53-Leu56, Leu60, and Leu162, forms an interacting cluster (Figure 4(c)). In addition, three conserved residues Arg163, Asp57, and Gly41 make a network of H-bonds which stabilize the \( \alpha_{3/4} \) loop with \( \alpha_4 \), while the short \( \beta_{2/3} \) loop is stabilized by an H-bond between the backbone of Tyr116 and a conserved Asn33 at \( \alpha_3 \).

Most of the residues with lower B-factors (\( B_f < 30 \text{ Å}^2 \)) are found within structured segments with low mobility as predicted by the GNM. The case of...
Gly^{115} is interesting due to its location at an unstructured location (at the β_{2/3} loop), which becomes a global hinge that would allow pivoting of the N-lobe and the C-lobe (Figure 4(d)). In addition, the GNM fluctuation profile shows local minima compatible with low-experimental B-factors at loops or terminal residues of secondary structures, and are: Leu^{14} (at the α_{1/2} loop), Ala^{29} (at the α_{2/3} loop), Gly^{41} (at the α_{3/4} loop), Asp^{76} and Gly^{108} (both flanking the H-motif), Ile^{121} and Pro^{130} (both flanking the PN loop), and Asp^{148} and Gln^{156} (both flanking the ARTT loop), among others. These may become local hinges that offer independent mobility to individual helices, loops, and subdomains (Figure 4(d)). However, whether these loops fluctuate independently or are coupled to other moieties is not reflected in the fluctuation profiles shown in Figure 4(a); in this case, a residue–residue fluctuation correlation map would be required. Figure 5(a)–(c) shows the fluctuation profile along the dominant GNM mode, gnm1, which accounts for ~17% of the total variance. According to this mode, the protein fluctuates in two anticorrelated (coupled, but in opposite directions) regions that closely correspond to the topological division in N and C lobes (Figure 4(d)), with five hinges (gnm1 = 0) at the interface between both subdomains. The second GNM mode, gnm2, with a variance contribution of ~6%, represents mainly the correlation (coupled, and in the same direction) between all of the β-strands with the ARTT and the PN loop fluctuations. These occur along with α_{1}, but are anticorrelated with the rest of helical segments (α_{2–α_{4}} and H-motif) and the β_{7/8} loop (Figure 5(d)–(f)). The broader fluctuation profiles of these two lowest indexed GNM modes reflect their high degree of cooperativity (i.e. the participation of a large number of residues). In contrast, the highest indexed modes are localized fluctuations in only a few residues, reflected in the fluctuation distribution as sharp spikes of low amplitude (Figure S1). Some are located in hinge regions of low mobility predicted by the GNM, e.g. Ala^{124}–Gly^{145} at the β_{4/5} loop; others are constrained residues responsible for the overall stability of subdomains, e.g. Leu^{60}, Leu^{132}, etc.

### 3.3. Conformational fluctuations by the ANM

To further analyze the intrinsic dynamics of the C3larvin toxin, the ANM was implemented for the Lys^{5}–Gln^{190} C_{α} trace of the 5TR4 structure (Figure 6(a)). The square displacements profile predicted by the ANM resembles the one by the GNM, with r = .76 (Figure 6(b)). However, neither of the previous profiles reflects the absolute direction and sense of the motifs, loops, and subdomains motions, while the ANM eigenvectors indeed contain information about the directions of the structural deformations (Figure 7). The comparison between the dominant ANM and PCA modes, anm1 and pc1, report a moderate overall overlap (O_{1,1} = .38). However, by excluding the α_{1} and ARTT residues, the overlap rises to O_{1,1} = .68. Hence, the anticorrelated movement between the N-lobe in relation to the H-motif (as part of the C-lobe) is described experimentally by the pc1 and theoretically by the anm1 (Figure 7). In regard to the pc2, the overlap with the intrinsic modes of deformation is significantly lower, even by considering the first 20 ANM modes (O_{2/20} = .41).

The inter-residue correlation map by the first ANM mode of fluctuation is complex (Figure 8(b)), and shows that the coupling among sections is not evident from the contact map (Figure 8(a)). The helical α_{1–α_{4}} segment of the N-lobe is partitioned into two fragments with anticorrelated mobility: (i) a proximal segment formed by the α_{1–α_{3}} helices, which is expected to be coupled – according to the high packing – to other elements of the N-lobe (e.g. β_{3} and the β_{5}–ARTT–β_{6} segment), while its coupling to the H-motif at the C-lobe is less evident; (ii) a terminal self-correlated α_{3–α_{4}} segment is also seen that fluctuates upon correlation with β_{1}, the β_{2/3} loop, the PN–β_{4} segment, and with the β_{7–β_{8}} motif (all belonging to the C-lobe); however, it is anticorrelated with its flanking sequences and with the H-motif. The H-motif is moderately isolated in terms of the contact map, and is only close to the β_{7–β_{8}} strands. However, in terms of mobility, it is partitioned into two well, self-correlated segments (red squares into the H-motif in Figure 8(b)), but coupled differently with the rest of the protein: (i) a N-terminal segment (α_{5α}–loop segment) is coupled to α_{1}, moderately independent from the PN and ARTT loops, and anticorrelated to the α_{3–α_{4}} segment; (ii) a C-terminal segment (α_{5α}–loop segment) is correlated to the ARTT, anticorrelated to the PN loop, and moderately independent from the α_{3–α_{4}} segment. In summary, there is an anticorrelated fluctuation of the N and C lobes, along a slightly backward movement of the β-core (Figure 7). The pivoting of the N-lobe displaces α_{1} in the same direction as the H-motif, which displaces the β_{5}–ARTT–β_{6} segment. The β_{7–β_{8}} segment is drawn inward by the H-motif.

### 3.4. C3larvin in the context of the C3bot1 structural variability

The similarity between C3larvin and C3bot1 based on the fraction of identical residues (~32%) and the overall folding (r_{RMSD} < 2.5 Å) offers an excellent opportunity to assess the dynamics of C3larvin at the level of the structural variability found in the large number of C3bot1 X-ray structures (Han, Arvai, Clancy, & Tainer, 2001; Ménetrey et al., 2002, 2008). In this sense, an ensemble of 40 C3bot1 structures of WT and various catalytically impaired variants for either apo or bound forms with substrates or ligands (e.g. nicotinamide and...
Figure 5. Structural fluctuations along the dominant GNM modes. Upper panel for the gnm1, lower panel for the gnm2. (a) and (d): mode shapes for the 191 Cα atoms of 4TR5. (b) and (e): tube representation of C3larvin colored according to a block of correlated movements from the mode shape. Rendered by NMWiz in VMD. (c) and (f): residue–residue correlation map from the sign in the mode shape.

Figure 6. (a) Tube representation of the spring network among the Cα as nodes (in ball format) of the 4TR5 structure, with a cutoff of 15 Å and colored spectrally by the experimental B-factors from blue (low) to red (high). Rendered by NMWiz in VMD. (b) Theoretical mobility profiles according to GNM and ANM. Normalized mean square fluctuations by the GNM (blue) and the ANM (red), for 186 Cα of the 4TR5 structure.
NAD\(^+\)) was formed with the C\(\alpha\) coordinates that superimpose with most C3larvin residues. After optimal superposition of the data-set (average \(\alpha\)RMSD = .48 Å for 181 residues), the PCA reported 39 modes of structural variation, with the first 4 modes accumulating \(\sim\)75\% of the total variance in the ensemble. Figure 9 shows the projection of the 40 C3bot1 conformations onto the first three PCA axes.

The pc1 separates the C3bot1 structures into two distinct sets with an average cumulative displacement of \(\sim\)11 Å between them, regardless of the ligation state. The major or unique deformation represented by this dominant mode of movement occurs at the ARTT loop, with an average \(\alpha\)RMSD of 1.88 Å for nine residues (Figure 10(a)). The groups with negative and positive pc1 values correspond, respectively, to the ‘in’ and ‘out’ conformations reported previously (Ménétrey et al., 2008) (Figure 10(b)). The ARTT loop contains the conserved motif Gln\(^{212}\)–X–Glu\(^{214}\), which is rearranged between these two states from a solvent exposed to a buried
conformation, respectively. In particular, associated with the in-to-out reconfiguration, the catalytic Gln\textsuperscript{212} switches toward the interior of the NAD\textsuperscript{+}-binding cleft, able to interact with the O-2\textsuperscript{′}-hydroxyl of the nicotinamide-ribose when NAD\textsuperscript{+} or NMP\textsuperscript{+} is bound (Ménétrey et al., 2002, 2008). Furthermore, the ‘in’ set is segregated into apo (black) and liganded (yellow) subgroups along the pc2 and pc3 axes (Figure 9). Along the pc3, for the ‘in’ set, the structures are clustered at the negative values when bound to ligand, while at positives values for the apo structures (Figure 9(b)). For the ‘out’ set, the apo structures span the whole pc3 range, while the liganded ones are mainly at the negative pc3 values (as for the ‘in’ set), except for 3 out of 10 structures with positive values (which will be considered later). By projecting the structures onto the pc1–pc2 plane, there is no a better distinction between apo and liganded forms (Figure 9(a)). The pc2 includes deformations at the α\textsubscript{3}–α\textsubscript{4} segment (indexes 25–45, Figure 10(c)), while both pc2 and pc3 describe fluctuations along the H-motif (indexes 60–100) and the β\textsubscript{7/8} loop (indexes around 170), and mainly show fluctuations in the PN loop (index at 122), with an average σRMSD of 1.06 Å for 10 residues. This loop is known to interact with NAD\textsuperscript{+}, and its conformation may be affected by the presence or absence of a ligand (Ménétrey et al., 2008). Indeed, for the ‘in’ set, the segregation in ‘liganded’ and ‘unliganded’ groups along the pc2 and pc3 is evidenced in the offset conformation of the PN loop, and particularly in the conformations of Gln\textsuperscript{182}, Phe\textsuperscript{183}, and Arg\textsuperscript{186} – shown to interact with NAD\textsuperscript{+} (Figure 10(d)).

Due to the fact that the exposed loops and motifs in each molecule are affected differently by side-chain intramolecular interactions, interaction with bound ligand, and crystal contacts, depending on the crystal form and position/orientation of the molecule into the asymmetric unit (e.g. the 3 apo outliers in the ‘out’ set along the pc3 include molecule B of the crystal units), is required to assess the structure-encoded ability of C3bot1 toxin to explore deviations compatible with the X-ray ensemble. ANM modes were calculated for representative unliganded WT C3bot1 structures. These included molecule A (‘in’ configuration) and molecule B (‘out’ configuration) of 2C89. The cumulative overlap between their slowest 20 ANM modes with the pc1 are relatively small as $O_{20,1} = .54$ and $O_{20,1} = .45$, respectively. Distinctively, the apo C3bot1 forms are readily prone to deformation according to the pc2 and pc3 axes by means of the joined effect of the first 3 ANM modes, based on the higher overlap: $O_{23,2} = .82$ and $O_{23,3} = .59$ for the molecule A, while $O_{23,2} = .83$ and $O_{23,3} = .60$ for the molecule B, with pc2 and pc3, respectively. Figure 11(a) shows good correlation ($r = .96$) on the projection of the ensemble onto the pc2 mode and the joined first three ANM modes. In other words, the interconversion unliganded $\Leftrightarrow$ liganded defined by the pc2 and pc3 modes of deformation is compatible with the slowest ANM modes of fluctuation (Figure 11(b)). Individually, anm2 and anm3 present the largest overlap with pc2 (Figure 12), while the anm1 with pc6 (e.g. $O_{1,6} = .42$ for molecule B of 2C89). This latter PCA mode also has a fluctuation peak at Phe\textsuperscript{183} that contributes to the PN loop deformation (not shown).

The C3larvin set, all unliganded structures, reveals an average σRMSD of .78 Å (for 181 residues) when superimposed onto the average C3bot1 coordinates, which contrasts with the lower .48 Å among the entire C3bot1 ensemble that includes apo and diverse liganded
forms. The C3larvin structures fall into the ‘in’ conformation when plotted along the pc1 axis (Figure 9). Consequently, other C3larvin structures with variations according to the pc1 mode of deformation would be expected to be found under a more extensive sampling (e.g. a larger X-ray data-set). However, the C3larvin ensemble shows a higher variation for the ARTT loop with respect to the C3bot1 structures (Figure 10(b)), with \( \alpha \)RMSD of 9.2 Å (for 8 residues) for the 4TR5 structure. Obviously, the ARTT loop is not the same between these two sets and is also quite variable in the C3larvin set. Moreover, even though the C3larvin structures contain no bound ligand, they display a mixed behavior – grouped as apo structures according to pc2 but as holo structures according to pc3 (Figure 9). However, with respect to the pc2 axis, the C3larvin 4TR5 structure is an outlier (~8 Å offset), with the closest C3bot1 structure corresponding to the apo 2C8B (an ARTT loop variant). The fact that the C3larvin 4TR5 and C3bot1 2C8B structures are the only two X-ray models with one monomer in their asymmetric units points toward the influence of packing forces in the adopted conformations – recall, this PCA mode includes PN loop fluctuations, which in 4TR5 are highly affected by crystal contacts (Figure 4(b)).

3.5. Simulated dynamics of C3larvin toxin

To further explore the dynamics of the C3larvin toxin, a simulated ensemble was generated from MD simulations using the full-length (191 residues) 4TR5 structure. Accordingly, 16,000 snapshots from the first 80 ns of the production phase (model-building set) showed stability after 20 ns (over the index ~4000), according to the
RMSD time course with respect to the original frame (Figure 13(a)). A total of 2000 extra snapshots (10 ns) were used for cross-validation of the model (cross-validation set). When the model-building ensemble was interposed and an EDA performed, the first 20 EDA modes (out of 567 modes of deformation) described the essential subspace of the toxin dynamics since these can explain ~80% of the structural variance (Figure 13(b)), with only the ed1 and ed2 modes accounting for 32% of the configurational changes in the trajectory. The projection of the ensemble onto the ed1–ed2 plane shows a broad distribution of the conformations, with a shift from the right (blue conformations) to the left (red conformations) along the ed1 axis (Figure 13(c)). For this ensemble, the overall transition along the ed1 mode seems to have a significant anharmonic component responsible for the large-scale motion in the configurational space. Figure 14(a) shows the distribution for the 2000 first and 2000 last snapshots (10 ns each) of the model-building trajectory along the ed1 axis. Their distributions contrast with the overlapping character exhibited for higher modes, for example, for the 30th EDA mode, ed30 (Figure 14(b)). The fluctuation profile driven by the lowest EDA modes (ed1) and by the combined effect of the next 29 lowest modes (ed2–ed30) is plotted in Figure 13(d). The fluctuations along ed1 show defined peaks at the N-terminus of α1 and at residues 122, 151, and 177, with amplitudes comparable to the combined effect of the rest of the considered modes. Thus, the anharmonic deformation along the ed1 mode does not present significant structural changes at an unique residue or set of residues, rather, a particular directionality associated with its orthogonal character with respect to rest of the modes. On the contrary, ed1 shows reduced fluctuations on relevant structures such as the α2–α3 segment, and mainly at the H-motif (Figure 13(d)).

Despite the observation that the stability in the overall structural deformation is reached after 20 ns of simulation (Figure 13(a)), the motions described by the first 4
Figure 13. (a) RMSD (in Å) time profile of 16,000 snapshots from a trajectory of 80 ns of production phase, based on the superposition to the initial conformation of the phase (at $t = 5$ ns). The arrow represents the frame at time = 20 ns. (b) Fractional variance (blue bars) and cumulative fractional variance (solid line) for the first 20 PCA modes describing the structural fluctuations in the MD ensemble. (c) Distribution of 16,000 conformation of the MD ensemble onto the ed1–ed2 axes (in RMSD, Å) after optimal superposition. First 2000 snapshots (in blue), last 2000 (in red). The schematic subaxes represent a new basis set from EDA over the last 4000 conformation (see Figure 15(d)). (d) Mean square fluctuations driven by the ed1 mode (blue) and by the ed2–ed30 modes (red) from the MD ensemble.

Figure 14. (a) Distribution of the first (blue) and last (red) 2000 conformations of the trajectory onto the ed1 mode. (b) Distribution of the first (blue) and last (red) 2000 conformations of the trajectory onto the ed30 mode.
EDA modes do not reach their stationary phase during the first 60 ns or 12,000 snapshots (Figure 15(a)). The transition along ed1 seems to occur coupled to ed2, as can be observed in the short-range subspace with a ‘diagonal trend’ between both EDA variables (Figure 13(c)). This suggests that the short-range transition along the ed1–ed2 axes might indeed be harmonic. To test this hypothesis, ANM modes for the full 4TR5 structure were calculated and contrasted with the EDA modes. In effect, the projection of the MD ensemble onto the anm2 (ed1+ed2) plane (Figure 15(c)) shows the agreement with the hypothesis that the coupled fluctuations along ed1 and ed2 can be driven by one of the more favorable ANM modes, i.e. anm2. Furthermore, an EDA for the last 20 ns range of the building model trajectory reports a new basis set where the two first EDA modes (edn1 and edn2) present their axes rotated in relation to the original ed1 and ed2 axes. The edn1 mode follows the direction of the diagonal trends, hence the large overlap between edn1 and anm2 ($O_{2,1n} = .66$), which is compatible with the high correlation between these modes when the last 4000 conformations are projected on the edn1–anm2 plane (Figure 15(d)). The calculation of the EDA modes for the MD ensemble over 186 Cα atoms, EDAt modes (‘t’ from truncated), allows comparison with the C3larvin X-ray PCA modes. In effect, their projection onto the ed1–ed2 plane reveals that the X-ray conformations are included in the MD subspace, spanned over the major two EDA modes (Figure 15(e)). The calculated fluctuation profiles for the simulated ensemble, considering all the 552 orthogonal modes, resemble the variability found in the X-ray data-set with a correlation of $r = .70$ (Figure 15(f)).

From the MD trajectory, the described ‘crab-claw’ movement can be seen by measuring the time evolution of the distance between Cα of Thr34 (at $\alpha_3$) and Ser78 (at $\alpha_{5a}$) (Figure 16(a)). Figure 16(b) shows variation on the distance, $\Delta d_{T34-S78}$, with fast transitions of ~2.5 Å (comparable to the difference between 4TR5 and chain A/O distances), and slower transitions with amplitudes as high as ~7.0 Å. On the contrary, the N-lobe and the H-motif remain compact, as evidenced by the changes in the distance between (i) Thr34 (at $\alpha_3$) and Ile53 (at $\alpha_4$) and (ii) Ser78 (at $\alpha_{5a}$) and Lys101 (at $\alpha_3$), respectively, with $\Delta d_{T34-I53} \approx \Delta d_{S78-L101} < 2.0$ Å. Furthermore, changes in the cavity formed by the β-fold can be observed by vari-

![Figure 15](image_url)

**Figure 15.** (a) Time course of the 16,000 snapshots of the trajectory projected onto the ed1 (blue), ed2 (green), ed3 (red), and ed4 (cyan) axes. The projection onto ed2–ed4 is shifted for clarity. (b) Projection of the MD ensemble onto the anm1 (red), anm2 (blue), and anm3 (green). For anm2 and anm3, every other 2 (8000 snapshots) and every other 4 (4000 snapshots) conformers were taken to allow the simultaneous plots. (c) Projection of the ensemble onto the joined ed1 and ed2 modes (absissa), and the anm2 (ordinate). (d) Projection of last 4000 snapshots onto the ed1 and anm2 axes. (e) Projection of MD ensemble onto the ed1 and ed2 axes. In color are the C3larvin X-ray structures. (f) Square fluctuation profiles (in Å²) represented by the two PCA modes of the C3larvin X-ray ensemble (blue), by the 552 EDA modes from the MD ensemble (red), and by the 552 ANM modes (green). This latter was scaled and shifted by the transformation ‘3y→0’ for convenience.
ations in the angle formed by Ser\textsuperscript{120} (at $\beta_3$), Thr\textsuperscript{165} (vertex at $\beta_7$), and Met\textsuperscript{172} (at $\beta_7$) (Figure 16(a)). A variance of 3.8$^\circ$, with fluctuations between 58.5$^\circ$ and 68.0$^\circ$, provides evidence of the relative mobility between the $\beta_A$-sheet and the $\beta_B$-sheet (Figure 16(c)). This occurs while the $\beta_B$-sheet expands and then contracts as seen in changes in the distance between Arg\textsuperscript{73} (at $\beta_1$) and Thr\textsuperscript{165} (at $\beta_7$) from 12.9 and 16.2 Å (Figure 16(d)).

4. Discussion

To date, three distinct C3larvin conformations identified by X-ray crystallography (4TR5, chain A, and chain O) have been described. Although identical in the overall fold (Figure 1), these conformations show differences in their relative position/orientation of subdomains and loops’ configurations (Figure 2(a)). The PCA of this ensemble defined the two composite modes that summarize the conformational changes (Figure 2(b) and (d)). Both modes describe major fluctuations between these structures at the ARTT loop and the N-terminus of $\alpha_1$ (Figure 2(c)). However, due to the orthogonal nature for both PCA axes, the mode of deformability of this loop is unique between 4TR5 and any of the chain O/A conformers, in comparison to the deformation between the latter (Figure 2(d)). In addition, the pc1 mode of deformation also includes the variation in the distance between the $\alpha_3$–$\alpha_4$ segment and the H-motif (Figure 2(c)–(d)), as observed between the chain A/O and the 4TR5 structure (Figure 2(a)), as a part of an overall ‘flexibility’ that also includes changes in the $\beta$-sheet core. Incidentally, these structures arise from different crystal forms that have a contrast in the number of molecules in their asymmetric units. Hence, the structural

Figure 16. (a) C3larvin 4TR5 structure showing distances and angles between reference atoms rendered by MOE. (b) Variations in the inter-motif distance (in Å) by measuring the distance between Thr\textsuperscript{34} and Ser\textsuperscript{78} (cyan), compactness of the N-lobe by the distance between Thr\textsuperscript{34} and Ile\textsuperscript{53} (green), and compactness of the H-motif by the distance between Ser\textsuperscript{78} and Lys\textsuperscript{101} (red). (c) Variations in the $\beta$-sheet angles (in degrees) by measuring the angle formed by Ser\textsuperscript{120}, Thr\textsuperscript{165} (vertex), and Met\textsuperscript{172}. (d) Variations in the $\beta_B$-sheets expansion (in Å) by measuring the distance between Arg\textsuperscript{73} and Thr\textsuperscript{165}. 

Molecular dynamics of C3 mART toxins
variability might arise from variable packing and crystal forces. However, a distinct aspect is the fact that the chain O/A was incubated with the M3 compound prior to crystallization, which is known to bind tightly to the protein and inhibits the mART activity (Krska et al., 2015), although no electronic density corresponding to the M3 molecule was detected in the final crystal. We do not know whether the ligand leached from the active site during the crystallization or if it is in the binding pocket, but simply exhibits high mobility. However, the presence of crystal water molecules in the NAD$^+$ active site and the reduced mobility displayed by side-chain pocket residues argue against this latter hypothesis. In any case, it is possible that the toxin structure observed in the chain O/A might represent an ‘active’ conformation resulting from the interaction with M3. Nevertheless, this flexibility in the C3larvin structures has also been reported for the C3bot1 toxin, regardless of the crystal form and ligation state (Ménétrey et al., 2002, 2008). The authors ascribed functionality of these fluctuations in the catalytic cycle with substrate binding and release of products. However, currently, it is unclear as to the molecular details that facilitate or drive this plasticity, and the present work offers insight(s) into the possible mechanism in this regard. Similarly, the defined conformations reported for the ARTT loop among the C3bot1 structures apparently do not correspond to the ligation state or to the crystal packing and hence, must be part of the native fluctuations of the toxin in solution.

To determine if the variability of these C3larvin conformations is the result of crystal artifacts, induced by the ligand or belong to its equilibrium dynamics, several analyses were performed. First, the information contained into the experimental B-factors was considered for analysis. For this, the residue profiles of the B-factors were correlated with experimental structural deformations (Figure 3(a)) and with calculated atomic fluctuations using a physically based model (Figure 3(b)–(c)). The underlying reason was that the crystallographic temperature factors depend on: (i) thermal vibrations plus harmonic collective motions (dynamic component), which are affected by specific/strong intramolecular and intermolecular interactions and (ii) rigid body motions of the molecule into the crystal unit (disorder) and alternative conformations (static component). For the latter, the distribution of fluctuations obtained from different molecules in a crystal and/or from different crystal forms may estimate the static components. In this sense, a different level of correlation was seen between the experimental B-factor profiles of full length or segments of different chains and the structural variability of the data-set (Table 1). The small correlation exhibited by the 4TR5 structure pointed toward a minor amount of static component in its overall crystal dynamics.

In regard to the first effect, theoretical calculations have shown that the atomic fluctuations can be estimated accurately from elastic models based on the 3-D topology of the structure, i.e. the $C_\alpha$ trace (Bahar, Atilgan, & Erman, 1997; Tirion, 1996). In this sense, a coarse-grained Gaussian Network Model analysis was performed on the C3larvin X-ray structures (Figure 3(b)–(d)). The best correlation between the calculated and experimental fluctuation profiles was $r = .56$ for the 4TR5 structure (Table 2). This value does not deviate significantly from the average correlation of $r = .59 \pm .15$, obtained from a data-set of 64 dissimilar proteins (Yang et al., 2007), particularly when considering that the GNM modeling was performed over an N-terminal truncated version of C3larvin (186 residues). The first 4 residues of the toxin sequence, in addition to 22 residues from the expression construct, were omitted in these calculations. These residues are found in the crystalized protein and therefore interact differently with the rest of the structure. The N-terminus of the $\alpha_1$ helix, is, indeed, a variable region among the three structures with a high dispersion of experimental B-factors. Altogether, a major divergence is expected between theoretical and experimental fluctuations in this region. Indeed, when evaluating the GNM to the full-length (191 residues) 4TR5 structure, while only considering 186 residues for the correlation, the coefficient increased to $r = .755$. In summary, the 4TR5 conformation seems to be the best estimate of the ‘in solution’ C3larvin structure in the data-set. Based on the fact that the structure is complete, i.e. there are atomic coordinates for all 191 residues, it was selected for further analysis.

Nevertheless, the experimental B-factor profile of 4TR5 and the fluctuations estimated by the GNM modeling lack agreement at specific regions of crystal contacts (Figure 4(a)), in particular at the PN and the $\beta_{7/8}$ loops. Thus, although the 4TR5 structure presents intermolecular contacts with both symmetry- and asymmetry-related molecules, these are the only crystal interactions with structured regions (Figure 4(b)); hence, their mobility is significantly affected. Other minor deviations occur at the $\alpha_2$-$\alpha_4$ segment. In this regard, a core of hydrophobic residues complemented with an H-bond network (Figure 4(c)) make strong/specific interactions that overcome the ‘average’ potential based only on the 3-D topology and expressed in a single elastic constant in the GNM. A high proportion of interacting residues are conserved in the C3-group, while others are synonymous substitutions, and this clearly reveals their structural role in stabilizing the N-lobe tertiary substructure (Figure 4(d)). Effectively, some of the highest indexed GNM modes are localized fluctuations of high frequency involved in the stability of the N-lobe, as is the case in the leucine cluster: Leu$^{16}$, Leu$^{65}$, Leu$^{65}$, and Leu$^{159}$–Leu$^{161}$ (Figure S1); most are completely conserved within the C3-group. On the
contrary, at the H-motif, the correlation is good (Figure 4(a)), reflecting that this subdomain is mostly free of crystal contacts and with low structural variability (see Figure 3(a), black line). Thus, the precise nature of the residues constituting the H-motif would be irrelevant for the protein’s dynamics and hence, only its folded structure matters. Incidentally, this motif lacks identical residues and is not well conserved in the C3-group. In summary, the elastic network approach implemented here by the GNM with uniform force constants seems to be a simple but sufficient model to describe the thermal fluctuations of the C3larvin toxin.

Based on this strong correlation, the GNM was further used to assess the degree of coupling/decoupling of the fluctuations between subdomains and moieties. Firstly, the GNM modeling identified residues and short sections with low mobility that become global hinges (i.e. pivot points for quasi-rigid body movements of large subdomains) and local hinges (pivot points for quasi-rigid body movements of small motifs and loops) (Figure 4(d)). Secondly, the model revealed that the slowest and most collective mode of fluctuation, the gnm1, is in agreement with the anticorrelated movement of the N-lobe with respect to the C-lobe (Figures 4(d) and 5(a)-(c)). Incidentally, the ARTT loop aside, this large, collective movement shows the greatest differences between the 4TR5 and the chain A/O structures. Thus, in principle, the observed flexibility in the C3larvin dataset, rather than strained conformations induced by packing forces, may be favored by the intrinsic dynamics of the protein. The distinction between chains A and O, characterized mainly by variation along the pc2 axis and largely ascribed to variations at α1 and the ARTT loop (Figure 2(c), blue curve), is considered by the GNM modeling (Figure 5(d)); this might be related, according to the gnm2, to the decoupling of these two structural elements from the rest of the N-lobe (Figure 5(d)-(e)).

However, by implementing the Anisotropic Gaussian Model (Figure 6(a)), the absolute direction of the main conformational changes accessible to the toxin was revealed. The experimental fluctuations described by the main PCA mode, pc1, were seen to be compatible with the most energetically favored mode of fluctuation, anm1. Indeed, the overall plasticity of the toxin, exhibited by the anticorrelated movement of the N and the C lobes fluctuated in a ‘clamping’ fashion, while expanding/contracting the cavity formed by the β-core, can be mostly driven by the anm1 mode of fluctuation (Figures 6(b) and 7). In addition, according to the anm1, the exposed and relevant ARTT and PN loops do not appear to fluctuate independently from other structured domains and motifs (Figure 8(b)). In regard to the pc2 mode of fluctuation observed in the C3larvin data-set (Figure 2(c), blue curve), contrary to that suggested by the GNM modeling, it is not driven by harmonic modes, according to the ANM modeling. Rather, an anharmonic character might be responsible for this variation, as was revealed by MD simulation (Figure 13(d)), and which will be discussed in a later section.

The analysis of 40 C3bot1 X-ray conformers by PCA clustered the conformations into two groups according to the main mode of structural variability (Figure 9), pc1, with 42% of the total variance that corresponds exclusively to the ARTT loop deformation (Figure 10(a)), and corresponds with the ‘in’ and ‘out’ conformations defined in Ménétrey et al. (2008). Moreover, the PCA readily dissected the ‘in’ group into ‘liganded’ and ‘unliganded’ subgroups according to the second and third modes of variation, pc2 and pc3 (Figure 9). These latter modes accounted for 25% of the total variance, and represent a fluctuation peak at the PN loop, evident between the apo and holo forms (Figure 10(d)), and show fluctuations at the α3-α4 and H-motif (Figure 10(c)). The structural changes for the α1-helix observed in the C3larvin data-set are not present in the C3bot1 ensemble. Such differences might be related to the variability in the length of this helix in both toxins, which might alter its interaction with the PN and ARTT loops and therefore its stability (Krska et al., 2015).

The calculated intrinsic dynamics of two apo forms of C3bot1 by ANM modeling are compatible with the interconversion between the ‘in’ and ‘out’ conformations (in ⇔ out) along the pc1 axis. This mode exhibits a greater anharmonic character, characteristic of low-indexed PCA modes, although it might be driven by the combined effect of a relatively high number of harmonic ANM modes. Hence, the apo protein would fluctuate spontaneously with a low frequency between both discrete states due to the high energy associated with high-indexed ANM modes. Since the ‘in’ conformation must be the ‘pre-active’ conformation of the toxin — mainly due to the conformation of the catalytic QxE motif at the ARTT loop — the ‘out’ conformation corresponds to either inactive complexes (regardless of the ligation state) or post-hydrolytic complexes (e.g. molecules C and D of 1GZF, C3bot1 bound with ADP and ribose-nicotinamide, respectively).

Therefore, by exploring the ‘in’ subspace driven by the most favorable ANM modes, the apo toxin might adopt the ‘liganded’ configuration (positive pc2 and negative pc3 values, Figure 9). A question that arises is if the ‘liganded’ configuration is induced by the ligand or rather is selected by the ligand, would the toxin exhibit one or more conformation(s) from its equilibrium ensemble that are amendable for selection or recognition by the ligand? Supporting the latter possibility, for the ‘out’ conformation, the ‘unliganded’ form is able to fluctuate and adopt the ‘liganded’ configuration, as defined by the pc2 and pc3 deformation modes (Figure 9). In addition, for the ‘in’ configuration, the fact that C3bot1 variants in Molecular dynamics of C3 mART toxins
the ARTT loop (e.g. E214N and Q212A) and in the PN loop (e.g. Q186A) complexed with NAD$^+$ show the same conformation according to the pc2 and pc3 exhibited by the WT–NAD$^+$ complex (1GZF, molecule A) suggest that the specific interaction with NAD$^+$ may not be strictly necessary. Further indirect evidence appears in the C3bot1 complex with nicotinamide, 2C8A. This ligand does not have the structural or the chemical features to induce an overall conformational change in the protein comparable in magnitude—at the level of pocket topology—to the one that might be induced by NAD$^+$. For example, nicotinamide has no equivalent to the NAD$^+$ phosphate to interact with Arg$^{186}$ at the PN loop, so it is difficult to conceive that a free sulfate molecule would ‘steer’ the backbone of the PN loop to approach and bind Arg$^{186}$. Even so, another consideration is that the presence of the pyridinium ring in the binding pocket (either by NAD$^+$, NMP$^+$, or nicotinamide) might be responsible for inducing the conformation of the PN loop by stacking with Phe$^{183}$ (Ménétray et al., 2008), which led to the assignment to Phe$^{183}$ the role of ‘nicotinamide sensor.’ However, the bound conformations of Phe$^{183}$ and the complete PN loop for the uncomplexed molecules F and G of the structure 2C8F (an ARTT variant with the ‘out’ conformations) argue against such a function. Likely, the location of the crystallographic sulfate ion and the stabilization of nicotinamide (in 2C8A) are more a consequence of the pose adopted by Arg$^{186}$ and Phe$^{183}$ after a rearrangement of the PN loop backbone (followed by changes in the side-chain rotamers), rather than the cause. Nevertheless, what the mechanism proposes is that the unliganded toxin might adopt the backbone ‘liganded’ topology which binds to NAD$^+$, allowing further improvement in the electrostatic and steric complementarities by a ‘local’ induced-fit mechanism. The complex may optimize contacts through subsequent conformational changes mainly by side-chains and solvent reconfiguration around/at the active site. Effectively, the distances between the PN loop of the molecule-D of 1GZE (apo) and 2C8A (nicotinamide bound) for both WT toxins crystallized in the same form are $\Delta$RMSD = .83 Å and allRMSD = 1.74 Å, for 11 residues. When only the residues Gln$^{182}$, Phe$^{183}$, and Arg$^{186}$ are considered, the distances are $\Delta$RMSD = .98 Å and allRMSD = 2.24 Å. The small backbone RMSD reasonably suggests that the protein adopts the ‘backbone liganded’ configuration by itself—along the pc2 and pc3 axes by the first ANM modes—and then binds the small nicotinamide molecule and stabilizes the complex by optimizing the side-chain rotamers of Gln$^{182}$ and Phe$^{183}$. Incidentally, the three slowest ANM modes of this C3bot1–nicotinamide complex also exhibit a high cumulative overlap with the pc2 and pc3 modes ($O_{23,2} = .81$ and $O_{23,3} = .57$, respectively); therefore, it is prone to go back to the apo form, probably releasing the ligand. This is not the case for the NAD$^+$-bound form of the toxin (not shown).

Furthermore, there is increasing evidence in support of the relationship between functional changes observed experimentally and the global motions predicted by theoretical physically based models (Bahar, Lezon, Bakan, et al., 2010; Berendsen & Hayward, 2000). We have shown that the overall flexibility in the C3bot1 X-ray structures, which is represented by structural variations along the pc2 and pc3 axes (Figure 10(c)), is compatible or might be driven by the most collective and lowest energy modes of intrinsic deformation of the protein (Figures 11–12). This represents an important observation since it helps to explain why the ‘clamping/unclamping’ movements occur for the apo form of the C3bot1 toxin, and validates the putative functional role of the ‘crab-claw’ topology likely associated with the clamping of the NAD$^+$ substrate and release of product(s).

Because of the limited number of C3larvin structures, simulated ensembles are mandatory to obtain an extended conformational space that might define the ‘native state.’ Thus, a long MD simulation with explicit solvent generated a collection of conformations with harmonic and anharmonic fluctuations. The EDA revealed that the essential dynamics of the protein consisted in the coupling of the two major fluctuation modes, ed1 and ed2 (Figure 13(c)). The sampled conformational space by the MD simulation is greater than its counterpart generated by experimental structures (Figure 15(e)), which is expected since large-scale motions (anharmonic motions) can be frozen due to crystal packing. In particular, movement along ed1 seems to be attenuated in the crystal environment. Accordingly, this EDA mode reflects variation at the $\alpha_1$ and ARTT loop (Figure 13(d)), which in turn represents the main experimental variations between chain A and chain O (Figure 2(c)), and cannot be sufficiently explained by the harmonic approach of the ANM modeling, as was already discussed. The fact that the ARTT loop presents reciprocal crystal contacts in the asymmetric unit of this crystal form (Figure S2) supports the altered mobility that these elements exhibit in the crystal lattice.

The MD ensemble follows Gaussian distributions when projected onto the dominant ANM modes (Figure 15(b)), which is in agreement with the idea that the MD trajectory—generated by an all-atom force field—can be determined by effective forces represented by a simple elastic network with uniform force constants. In particular, the coupled ed1 and ed2 fluctuations, at the stable phase or at short range, can be easily driven by one of the more favorable ANM modes, anm2 (Figure 15(c)–(d)). Nevertheless, questions arise as to whether the sampled subspace by the MD simulation is representative of the toxin dynamics and is the MD
subspace statistically significant of the configurational space accessible for the toxin. It has been reported that for a single-domain protein, the EDA eigenvectors converge on the nanosecond timescale (Daidone & Amadei, 2012). Thus, the 80-ns simulation period for the mid-size C3larvin toxin adequately satisfies these criteria. The strong correlation \((r = .96)\) of the projection of the extra 10 ns of simulation (not used for the EDA calculation) onto the amn2-(ed1+ed2) cross-validates the EDA (Figure S3). Accordingly, a EDA analysis over an independent simulation, using the chain O structure of 5DZQ as initial coordinates, shows compatible results, with high correlation between the main EDA and ANM modes (Figure S4).

The structural deformation associated with the ‘crab-cab’ movement is observed in the MD simulation by fluctuations in the distance of reference points ascribed to N and C lobes, \((d_{T34-S78}, \text{ for example})\), as well as changes in inter \(\beta\)-sheet distances and angles (Figure 16). A fast flickering of the inter-lobe distance (~1 Å of amplitude) occurs on tens of picoseconds timescales. Larger fluctuations, with amplitudes comparable to the difference between the 4TR5 and the chain A/O structures (~3 Å), occur on the nanosecond timescale, while significant deviations (up to 7–8 Å) are exhibited in tens of nanoseconds (Figure 16(a)). Hence, the ‘crab-cab’ movement corresponds to a ‘natural’ mode of motion with cycles spanning frequencies of various decades, and therefore displayed in its equilibrium dynamics, rather than a discrete deviation observed in the MD between the original (from the crystal lattice) and the stationary conformations. In summary, to describe the ‘native state’ of the C3larin toxin, the MD simulation might provide the long-range variability along the ed1 mode, while the slowest 20 ANM modes may provide the most favorable intrinsic harmonic modes of fluctuations – which are in turn compatible with the short-range coupling between ed1 and ed2 in the MD simulation. The underlying reasoning is that the ‘diagonal-trend’ subspaces in the ed1–ed2 plane would correspond to fluctuations around minima configurations in the energy landscape, where the harmonic approximation by the NMA modeling is effectively fulfilled. Thus, the protein samples the conformational subspace around each conformational minimum by fluctuations driven mainly by amn2. Eventually, the energy barrier along the eda1 axis would be overcome. In this sense, a component of the amn1 mode might drive the net shift between conformational minima along the direction of ed1 by means of fluctuations that follow the orthogonal edn2 and edn3 deformation axes (Figure S5), thus yielding the zigzag pattern observed in the ed1–ed2 projection plane.

With the N-terminus of \(\alpha_1\) aside, the high correlation between experimental and calculated or simulated fluctuation profiles for the 4TR5 structure \((r_{\text{PCA-EDA}} = .70\) and \(r_{\text{PCA-ANM}} = .62)\) supports – in principle – both \textit{in silico} methods (Figure 15(f)). The main outliers correspond to fluctuations at the PN and the \(\beta_7/8\) loops, which were previously reported to be constrained in the crystal by packing forces, i.e. crystal artifacts (Figure 4(b)). In
In addition, the difference between the C3larvin chain A and chain O corresponds to variations at the ARTT loop and α₁-helix that contrast the conformation of the 4TR5 structure, and do not seem to be driven by the energy-favored modes of deformation – they appear as short-scale ‘anharmonic’ movements. Deformation at the ARTT loop has also been observed among the C3bot1 structures, corresponding to the ‘in’ and ‘out’ conformations, with an anharmonic character for this interconversion. Furthermore, a long MD simulation of the 4TR5 structure revealed, in agreement with the previous results, that these changes occur – although not exclusively – along the first EDA mode, ed1, which behaves in an anharmonic manner, and is coupled to the ed2 mode. The large-scale variation observed along ed1 might be achieved by means of ‘jumping’ between different minima in a progression of structural changes, which can be driven or favored by the intrinsic dynamics. Altogether, according to experimental data and simulations, both anharmonic and harmonic modes of fluctuation are needed to adequately describe the equilibrium dynamics of C3larvin, and, in general, the C3-like toxins.

We can conclude that the 4TR5 structure is the best representative, so far, of the C3larvin toxin in solution. However, the current understanding is that the native state of a protein, rather than a ‘unique structure,’ is an ‘ensemble of microstates’ (Daniel, Dunn, Finney, & Smith, 2003; Yang et al., 2007). These microstates feature the same overall 3° and mostly 2° structures, but may also contain structural variations in the packing due to the relative domain or subdomain disposition, or variable motif exposure to solvent, loop conformations, and side-chain geometry (i.e. rotational conformers with bond angle/length variations). In this sense, the ‘equilibrium dynamics’ of the 4TR5 structure would give a better picture of C3larvin toxin in solution since it contains all the accessible microstates under physiological condition. To accomplish this, we were able to implement a mixed procedure that takes advantage of the information derived from the MD and ANM approaches – first, applying the MD lowest indexed modes to account for the anharmonic and coupled component(s) of the ‘equilibrium’ motion; then, using a subset of the lowest indexed ANM modes. In the latter step, there is no ambiguity since it represents a unique solution derived solely from the original coordinates.

For this purpose, the ed1 mode was extended to the all backbone atoms of the 4TR5 as reference coordinates in order to generate an ensemble of the backbone traces with fluctuations, with an average $\bar{\sigma}$RMSF that reproduces the average $\bar{\sigma}$RMSF along the ed1 mode (~5 Å). Figure 17(a) shows a synthetic ensemble of 16 uniform decoys generated from the 4TR5 backbone with fluctuations driven by ed1. From this ensemble, the two most
different decoys were extracted according to the pairwise \( \text{bbRMSD} \) to use each as reference coordinate to generate additional subensembles based on the 20 dominant ANM modes of the 4TR5 structure. Figure 17(b) shows the output of the previous modeling by superposing the two subensembles. The conformations of the two subensembles onto the \( \text{anm1-anm3} \) axes are well defined by the subspace spanned by the MD ensemble (Figure S6).

6. Future directions
We recently reported models of the complex of C3larvin with \( \text{NAD}^+ \) and the inhibitor M3, derived from the apo 4TR5 coordinates (Krska et al., 2015). The potential binding poses of M3 were obtained by flexible docking allowing the pocket side-chains to relax using reaction-field electrostatics, while constraining the M3 ligand to a volume defined by pharmacophore features derived from the modeled \( \text{NAD}^+ \) binding pose. The top-scoring pose located the M3 molecule at the A-side (adenine site) of the binding pocket, where the ring system of M3 surrogates the adenine moiety of the bound \( \text{NAD}^+ \). In this docked complex, M3 is clamped by \( \alpha_3 \) and \( \delta_3 \) at the H-motif, reflecting the relevance of the pocket topology at the A-site to the ability of the docking program to generate and score active conformations of putative toxin–inhibitor complexes.

The utility of the dominant modes of fluctuation consists in their ability to generate alternative conformations from a reference structure. This is particularly pertinent for docking of small ligands in virtual screening protocols, where the overall flexibility of the receptor may be taken into account beforehand using an ensemble of receptor conformations instead of just a single structure. In this sense, the decoys depicted in Figure 17(b) would only require refinement by an all-atom force field to provide realistic, alternative conformations of the complete protein for use as synthetic ‘receptors’ of C3larvin toxin. These subensembles will indeed be used in virtual screening protocols as part of our drug discovery working strategy.

Abbreviations

- **mART**: mono-ADP-ribosyltransferase
- **ANM**: anisotropic network model
- **anm**: anisotropic network mode
- **ARTT**: ADP-ribosylation turn-turn loop
- **EDA**: essential dynamic analysis
- **EHT**: extended Hückel theory
- **GB-VI**: generalized Born-volume integral
- **GNM**: Gaussian network model
- **gmm**: Gaussian network mode
- **MD**: molecular dynamics
- **MM**: molecular mechanics
- **M3**: \( N'[(1-[1H-pyrazolo[3,4-d]pyrimidin-3-yl]piperidin-3-yl)methyl]methanesulfonamide \)
- **MMFF94**: Merck molecular force field
- **NMN+**: nicotinamide mono nucleotide
- **PCA**: principal component analysis
- **PDB**: protein data bank
- **PN**: phosphate-nicotinamide loop
- **RMS**: root mean square
- **RMSF**: root mean square fluctuations
- **RMSD**: root mean square deviation
- **\( \text{RMSD}_{\alpha} \)**: \( \alpha \)-root mean square deviation
- **\( \text{RMSD}_{\text{all}} \)**: all heavy-atom root mean square deviation
- **\( \text{bbRMSD} \)**: backbone root mean square deviation
- **VS**: virtual screening

Supplementary material
The supplementary material for this paper is available online at [http://dx.doi.org/10.1080/07391102.2015.1123189](http://dx.doi.org/10.1080/07391102.2015.1123189).

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This research was funded by the Institute of Infection and Immunity [grant number MOP-106661] of the Canadian Institutes of Health Research (to A.R.M.).

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Iloerger, T. R., McCoy, A. J., Moriarty, N. W., … Terwilliger, T. C. (2002). PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallographica Section D, Biological Crystallography*, 58, 1948–1954.

Amadei, A., Linsen, A. B. M., & Berendsen, H. J. C. (1993). Essential dynamics of proteins. *Proteins: Structure, Function, and Genetics*, 17, 412–425.

Bahar, I., Atilgan, A. R., Demirel, M. C., & Erman, B. (1998). Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Physical Review Letters*, 80, 2733–2736.

Bahar, I., Atilgan, A. R., & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding & Design*, 2, 173–181.

Bahar, I., Chennubhotla, C., & Tobi, D. (2007). Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Current Opinion in Structural Biology*, 17, 633–640.

Bahar, I., Lezon, T. R., Bakan, A., & Shrivastava, I. H. (2010). Normal mode analysis of biomolecular structures: Functional mechanisms of membrane protein. *Chemical Reviews*, 110, 1463–1497.

Bahar, I., Lezon, T. R., Yang, L. W., & Eyal, E. (2010). Global dynamics of proteins: Bridging between structure and function. *Annual Review of Biophysics*, 39, 23–42.

Bahar, I., & Rader, A. J. (2005). Coarse-grained normal mode analysis in structural biology. *Current Opinion in Structural Biology*, 15, 586–592.
Bakan, A., Meireles, L. M., & Bahar, I. (2011). ProDy: Protein dynamics inferred from theory and experiments. Bioinformatics, 27, 1575–1577.

Barth, H., & Stiles, B. G. (2008). Binary actin-ADP-ribosylating toxins and their use as molecular Trojan horse for drug delivery into eukaryotic cells. Current Medicinal Chemistry, 15, 459–469.

Beattie, B. K., Prentice, G. A., & Merrill, A. R. (1996). Investigation into the catalytic role for the tryptophan residues within domain III of Pseudomonas aeruginosa exotoxin A. Biochemistry, 35, 15134–15142.

Berendsen, H. J., & Hayward, S. (2000). Collective protein dynamics in relation to function. Current Opinion in Structural Biology, 10, 165–169.

Chemical Computing Group Inc. (2012). Molecular operating environment (2014.08) [Computer software]. Montreal.

Collier, R. J. (2001). Understanding the mode of action of diphtheria toxin: A perspective on progress during the 20th century. Toxicol, 39, 1793–1803.

Daidone, I., & Amadei, A. (2012). Essential dynamics: Foundation and applications. WIREs Computational Molecular Science, 2, 762–770. doi:10.1002/wcms.1099

Daniel, R. M., Dunn, R. V., Finney, J. L., & Smith, J. C. (2003). The role of dynamics in enzyme activity. Annual Review of Biophysics and Biomolecular Structure, 32, 69–92.

David, C. C., & Jacobs, D. J. (2014). Principal component analysis: A method for determining the essential dynamics of protein. In Liversay, D. R. (Ed.), Protein dynamics: methods and protocols. Methods in molecular biology 1084, pp. 193–226. New York, NY: Springer.

Deng, Q., & Barbieri, J. T. (2008). Molecular mechanisms of the cytotoxicity of ADP-ribosylating toxins. Annual Review of Microbiology, 62, 271–288.

Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallographica Section D, Biological Crystallography, 60, 2126–2132.

Fan, E., O’Neal, C. J., Mitchell, D. D., Robien, M. A., Zhang, Z., Pickens, J. C., … Hol, W. G. (2004). Structural biology and structure-based inhibitor design of cholera toxin and heat-labile enterotoxin. International Journal of Medical Microbiology, 294, 217–223.

Hall, B. A., Kaye, S. L., Pang, A., Perera, R., & Biggin, P. C. (2007). Characterization of protein conformational states by normal-mode frequencies. Journal of the American Chemical Society, 129, 11394–11401.

Han, S., Arvai, A. S., Clancy, S. B., & Tainer, J. A. (2001). Crystal structure and novel recognition motif of Rho ADP-ribosylating C3 exoenzyme from Clostridium botulinum: Structural insights for recognition specificity and catalysis. Journal of Molecular Biology, 305, 95–107.

Hayward, S., & Go, N. (1995). Collective variable description of native protein dynamics. Annual Review of Physical Chemistry, 46, 223–250.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD-visual molecular dynamics. Journal of Molecular Graphics, 14, 33–38.

Kabsch, W. (2010). XDS. Acta Crystallographica Section D, Biological Crystallography, 66, 125–132.

Kitaou, A., & Go, N. (1999). Investigating protein dynamics in collective coordinate space. Current Opinion in Structural Biology, 9, 164–169.

Krska, D., Ravulapalli, R., Fieldhouse, R. J., Lugo, M. R., & Merrill, A. R. (2015). C3larvin toxin, and ADP-ribosyltransferase from Paenibacillus larvae. Journal of Molecular Biology, 290, 1639–1653.

Meireles, L., Gur, M., Bakan, A., & Bahar, I. (2011). Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins. Protein Science, 20, 1645–1658.

Ménétrey, J., Flatau, G., Boquet, P., Ménez, A., & Stura, E. A. (2008). Structural basis for the NAD-hydrolysis mechanism and the ARTI-loop plasticity of C3 exoenzymes. Protein Science, 17, 878–886.

Ménétrey, J., Flatau, G., Stura, E. A., Charbonnier, J. B., Gas, F., Teulon, J. M., … Menez, A. (2002). NAD binding induces conformational changes in rho ADP-ribosylating Clostridium botulinum C3 exoenzyme. Journal of Biological Chemistry, 277, 30950–30957.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten, K. (2005). Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26, 1781–1802.

Popoff, M. R. (2005). Bacterial exotoxins. Contributions to Microbiology, 12, 28–54.

Simon, N. C., Aktories, K., & Barbieri, J. T. (2014). Novel bacterial ADP-ribosylating toxins: Structure and function. Nature Reviews. Microbiology, 12, 559–611.

Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Physical Review Letters, 77, 1905–1908.

Vagin, A., & Tepljakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallographica Section D, Biological Crystallography, 66, 22–25.

Vogelsang, M., Pautsch, A., & Aktories, K. (2007). C3 exoenzymes, novel insights into structure and action of Rho-ADP-ribosylating toxins. Naunyn-Schmiedebergs Archives of Pharmacology, 374, 347–360.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. P., …& Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. Acta Crystallographica Section D, Biological Crystallography, 67, 235–242.

Yang, L. W., Eyal, E., Bahar, I., & Kitao, A. (2009). Principal component analysis of native ensembles of biomolecular structures (PCA NEST): Insights into functional dynamics. Bioinformatics, 25, 606–614.

Yang, L. W., Eyal, E., Chennubhotla, C., Jee, J. G., Gronenborn, A. M., & Bahar, I. (2007). Insights into equilibrium dynamics of proteins from comparison of NMR and X-Ray data with computational predictions. Structure, 15, 741–749.

Yates, S. P., Jørgensen, R., Andersen, G. R., & Merrill, A. R. (2006). Stealth and mimicry by deadly bacterial toxins. Trends in Biochemical Sciences, 31, 123–133.

Yates, S. P., & Merrill, A. R. (2005). Characterization of oxidized nicotinamide adenine dinucleotide (NAD+) analogues using a high-pressure-liquid-chromatography-based NAD-glycohydrolase assay and comparison with fluorescence-based measurements. Analytical Biochemistry, 340, 41–51.

Zhang, K. Y., Cowtan, K., & Wilson, K. S. (2011). Overexpression of human C3 exoenzyme from Pseudomonas aeruginosa and purification of a functional C3 exoenzyme. Journal of the American Chemical Society, 133, 5690–5698.