

## Understanding the EF-hand closing pathway using non-biased interatomic potentials

L. Dupuis and Normand Mousseau

Citation: *J. Chem. Phys.* **136**, 035101 (2012); doi: 10.1063/1.3671986

View online: <http://dx.doi.org/10.1063/1.3671986>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v136/i3>

Published by the [American Institute of Physics](#).

---

### Related Articles

Understanding the EF-hand closing pathway using non-biased interatomic potentials

*JCP: BioChem. Phys.* **6**, 01B609 (2012)

Amino acid analogues bind to carbon nanotube via  $\pi$ -interactions: Comparison of molecular mechanical and quantum mechanical calculations

*JCP: BioChem. Phys.* **6**, 01B607 (2012)

Amino acid analogues bind to carbon nanotube via  $\pi$ -interactions: Comparison of molecular mechanical and quantum mechanical calculations

*J. Chem. Phys.* **136**, 025103 (2012)

Double-functionalized nanopore-embedded gold electrodes for rapid DNA sequencing

*Appl. Phys. Lett.* **100**, 023701 (2012)

An immersed boundary method for Brownian dynamics simulation of polymers in complex geometries: Application to DNA flowing through a nanoslit with embedded nanopits

*J. Chem. Phys.* **136**, 014901 (2012)

---

### Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: [http://jcp.aip.org/about/about\\_the\\_journal](http://jcp.aip.org/about/about_the_journal)

Top downloads: [http://jcp.aip.org/features/most\\_downloaded](http://jcp.aip.org/features/most_downloaded)

Information for Authors: <http://jcp.aip.org/authors>

### ADVERTISEMENT

**AIP**Advances

*Submit Now*

Explore AIP's new  
open-access journal

- Article-level metrics now available
- Join the conversation! Rate & comment on articles

## Understanding the EF-hand closing pathway using non-biased interatomic potentials

L. Dupuis<sup>1</sup> and Normand Mousseau<sup>2,a)</sup>

<sup>1</sup>Département de Biochimie, Centre Robert-Cedergren and GEPRM, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada

<sup>2</sup>Département de Physique, Centre Robert-Cedergren and GEPRM, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada

(Received 3 August 2011; accepted 4 December 2011; published online 17 January 2012)

The EF-hand superfamily of proteins is characterized by the presence of calcium binding helix-loop-helix structures. Many of these proteins undergo considerable motion responsible for a wide range of properties upon binding but the exact mechanism at the root of this motion is not fully understood. Here, we use an unbiased accelerated multiscale simulation scheme, coupled with two force fields — CHARMM-EEF1 and the extended OPEP — to explore in details the closing pathway, from the unbound holo state to the closed apo state, of two EF-hand proteins, the Calmodulin and Troponin C N-terminal nodules. Based on a number of closing simulations for these two sequences, we show that the EF-hand  $\beta$ -scaffold, identified as crucial by Grabarek for the EF-hand opening driven by calcium binding, is also important in closing the EF-hand. We also show the crucial importance of the phenylalanine situated at the end of first EF-hand helix, and identify an intermediate state modulating its behavior, providing a detailed picture of the closing mechanism for these two representatives of EF-hand proteins. © 2012 American Institute of Physics. [doi:10.1063/1.3671986]

### I. INTRODUCTION

The EF-hand motif plays a crucial role in eukaryotic cellular signaling. Protein members of this large family are characterized by a helix-loop-helix structure. Upon binding  $\text{Ca}^{2+}$  ions these motifs can act as simple calcium buffers as well as undergo large motion associated with a number of processes.<sup>1-3</sup> Among these EF-hand proteins, Calmodulin has probably attracted the most attention. Using calcium as trigger token, this small 148 residues protein opens its hydrophobic cores<sup>4</sup> and binds to hundred of different protein targets,<sup>1,5,6</sup> participating into a wide range of processes such as signal transduction,<sup>7-9</sup> endocytosis,<sup>10,11</sup> muscular contraction,<sup>12</sup> cell cycle regulation and mitosis,<sup>13-15</sup> genes regulation<sup>16,17</sup> and neural plasticity.<sup>18,19</sup> Calmodulin adopts an overall dumbbell shape, with a N-terminal nodule made of two EF-hands motifs at one end (helix pairs A-B and C-D), linked with a long and highly flexible  $\alpha$ -helix to a C-terminal nodule with a similar two EF-hands shape. Both ends may attach to different sites of a protein channel or grip around the same protein domain.<sup>17,20</sup> The CT domain is considered to be more rigid, binding structurally to its target, while the NT domain, which has less affinity to calcium, is considered the regulatory domain.<sup>21</sup>

Each EF-hand is composed of two typical helix- $\text{Ca}^{2+}$  chelation loop-helix units. It is this pair of turn-loop structures that provides the rich conformational space needed to accommodate the large scope of protein targets. To ensure its  $\text{Ca}^{2+}$  binding role, the canonical loop sequence in EF-hands is mostly preserved with a number of crucial residues at fixed positions.<sup>3</sup> While the role of each residue in binding this el-

ement has been extensively studied,<sup>3</sup> their respective motion on the pathway connecting the holo to apo conformations is much less understood.

The full structure of the apo and the calcium-bound states of CaM is well known.<sup>22,23</sup> It has served as a basis for characterizing the overall kinetics of the CaM EF-hands. This work is simplified somewhat by the fact that NT and CT domains are mostly independent from each other, so that it is appropriate to consider their properties separately.<sup>24,25</sup> A number of experiments confirm the large conformational space sampled by the unbound sequence. NMR relaxation experiments have shown that the kinetics is rapid for calcium-free CaM, suggesting important conformational changes over a 10-ns time scale.<sup>24</sup> However, the flexibility of the NT domain is not necessarily sufficient to span the full apo-holo conformational space.<sup>26</sup> In this case, the addition of calcium could force both NT and CT domains to stay open in absence of protein target, facilitating binding.<sup>25,27,28</sup> Other works suggest rather that Calmodulin's inherent flexibility is sufficient to populate the closed form and the open form in presence of calcium and that the calcium loaded open form would only be stabilized by protein target presence.<sup>29</sup>

Much less is known, however, about the exact pathway for going from one state to the other and numerous computational studies have attempted to answer part of the question. Confirming experimental observations, molecular dynamical simulations with local constraints have shown that the bound state of the CT domain displays large thermal fluctuations.<sup>30</sup> A number of studies have also focused directly on the apo to (unbound) holo pathway. Because of the difficulty associated with representing accurately both states, all simulations impose a bias in the force field to guarantee convergence to experimentally-derived states. Steered molecular dynamics,

<sup>a)</sup>Electronic mail: Normand.Mousseau@umontreal.ca. Tel: +1 (514) 343-6614. Fax: +1 (514) 343-2071.

for example, was used to force  $\text{Ca}^{2+}$  binding and dissociation from various loops in the CT domain.<sup>21,31</sup> Zhang *et al.*, for their part, used a double-Gō forcefield to study the unbound NT domain.<sup>32</sup> A similar approach was followed by Chen and Hummer, this time on the CT domain,<sup>33</sup> allowing them to identify three phases along the pathway — apo-folded, apo-unfolded and holo-like structures — as well as by Tripathi and Portman who study in details the structural changes taking place during the transformation and find a significant rigidification of the loops in the NT domain as the protein moves from apo to holo conformations.<sup>34,35</sup> These results generally support the experimentally-derived two-step Grabarek model for the  $\text{Ca}^{2+}$  binding mechanisms centered on a structure named EF $\beta$ -scaffold, associated with the residues forming the  $\beta$ -sheet in the apo conformation, which defines the ion's position and allows the repositioning of the helices to ensure binding.<sup>36,37</sup>

In this study, we revisit this question with an unbiased approach that allows us to characterize finely the pathway from the holo to the apo form. Following previous simulations, we first focus on the unbound two EF-hand motif of the N-terminal nodule of Calmodulin which has been extensively characterized. For comparison, we also simulate the holo to apo pathway for the EF-hand motif of the Troponin C N-terminal, for which little numerical work exists. For this, we apply a multiscale scheme that samples the phase space efficiently while providing a clear description of intermediate states. This method is combined with two potentials — CHARMM19 (Ref. 38) with EEF1 solvation model<sup>39</sup> and an extended version of the OPEP forcefield<sup>40</sup> (EOPEP).

While the details of the closing pathway differ slightly between the two sequences, our results confirm the EF-hand  $\beta$ -scaffold as a conformation switch, as proposed by Grabarek.<sup>37,41</sup> While Grabarek principally explored the conformational change towards the open protein form upon ion binding, our study focus on the closing pathway upon ion unbinding. We observe that the reverse pathway is initiated by the reverse reorganization of the  $\beta$ -scaffold. Our simulations also underline the importance of the phenylalanine situated at the end of first EF-hand helix of Calmodulin and Troponin C, providing a more complete picture of the closing mechanism for these representatives of this important class of proteins.

## II. METHODOLOGY

The numerical study of closing pathways in flexible proteins is challenging and most simulations of EF-hand dynamics have used Gō-like biased potentials to ensure successful transitions back and forth between the two states.<sup>32–35</sup> Such a bias accelerates the effective dynamics allowing the determination of a free energy surface while ensuring that both terminal states visited correspond to the experimentally-derived structure. By removing possible metastable states, however, the bias can also affect the closing pathway in uncontrolled ways. To complement these works, we follow here an unbiased approach using a multiscale activated algorithm based on the activation-relaxation technique (ART nouveau)<sup>44,45</sup> that has been applied successfully to protein folding,<sup>46,47</sup> pro-

tein aggregation<sup>48,49</sup> and protein flexibility.<sup>50,52</sup> The holographic multiscale approach, which is presented in details in Ref. 51, allows to concentrate the conformational sampling on the flexible angles, which increases significantly the efficiency of the method. For completeness, we summarize the algorithm below.

This sampling method is coupled to two all-atom potentials with implicit solvent: CHARMM19 (Ref. 38) with EEF1 solvation model<sup>39</sup> and an extended version of the OPEP model.<sup>40</sup> We describe briefly below the various elements of our approach.

### A. ART

Sampling of the closing trajectory is performed using ART nouveau.<sup>44,45,53</sup> Starting from a conformation in a local minimum, the protein is deformed in a random direction according to the prescription described in Sec. II B until a direction of negative curvature appears on the energy surface. The conformation is then pushed along this direction while the force is minimized in the perpendicular hyperplane, bringing the system onto a first-order saddle point corresponding to a local transition state. From there, the conformation is brought into a new local minimum. This set of conformations — initial minimum, saddle point, final minimum — forms an event. The full move is accepted using a Metropolis criterion based on the energy difference between the final and initial minima and a temperature of 300 K.

This method has been successfully applied to characterize the folding of a 16-residue  $\beta$ -hairpin<sup>46</sup> and the 60-residue protein A,<sup>47</sup> as well as the aggregation of amyloid peptides.<sup>48,54</sup>

### B. Holographic multiscale algorithm

The critical part for the efficiency of ART nouveau lies in the initial deformation. Since protein motions are based on angular deformations, it is much more efficient to work within internal coordinates.<sup>55–57</sup> Even in this reduced coordinate space, a dominant fraction of possible moves involves breaking the secondary structure and lead to high-energy structures that would be rejected with the ART nouveau algorithm. Since non-local motions necessarily involve rotation of the  $\psi$  and  $\phi$  dihedral angles on each side of the main-chain  $\text{C}\alpha$ 's, our strategy is to constrain the initial deformation to loop  $\psi$  and  $\phi$  angles, with  $\alpha$ -helices able to react to the deformation, but not initiate it.

In the holographic multiscale approach, each main-chain movable angle is treated as a pivot around which the protein, separated as two blocks, and the lateral chain can rotate and swivel. As a consequence, the definition of flexible and rigid regions is dynamical: the set of splitting points change during one transition event, avoiding large block collisions, and increasing the flexibility of the approach. At each step during the activation phase, forces are evaluated in real space and projected on each pivot, allowing a local elastic response to the large scale motion generated by the move.<sup>51</sup>

For the N-terminal domains of Calmodulin and Troponin C, we allow all  $C\alpha$  on the residues forming both calcium loops as well as those borderin them to act as eventual pivots. During the activation step, we randomly choose a few of them as initial pivots, with random rotating directions. These deformations, which do not have to respect biological rules, are corrected naturally during activation as all angles are allowed to adapt as the system converges to the transition state. To remove any bias imposed by the presence of rigid secondary-structure regions, the relaxation part of each event takes place in real-space representation using a damped-MD algorithm and applied to all degrees of freedom.

### C. Forcefields

Until now, most EF-hand closing simulations have used biased potential. We focus rather on two non-biased implicit-solvent force fields coupled with our sampling technique: our in-house implementation of CHARMM19 (Ref. 38) with EEF1 solvation model<sup>39</sup> and an extended potential based on the OPEP model.<sup>40</sup>

The implicit water solvent CHARMM19 potential is designed to be used in conjunction with different solvation models. We implement our own version of CHARMM19, as described in Ref. 38. Following the approach of Ponder with Tinker,<sup>58</sup> we use our own torsion angle description, based on OPEP's,<sup>40</sup> but adapting the prefactors to those of CHARMM19. This potential is combined with the EEF1 solvation model volume exclusion terms.<sup>39</sup> We applied ionic neutralization of the side chains, the N-terminal and the C-terminal, as specified in EEF1.

We also use an extended version of OPEP potential,<sup>40</sup> generated by developing C, O, N and H atoms of the side chains. The coarse grain OPEP forcefield is based on a statistical approach for the evaluation of the side chains interactions, which are represented as unique spheres of different size, and introduces a cooperativity for main chain hydrogen bonds. It has proven very reliable for a large number of sequences and proteins of about 70 residues or less, including amyloids.<sup>47,50,59</sup> We find, however, that it is less appropriate for highly flexible sequence such as CaM and TnC, where a detailed description of the side-chain packing is necessary to ensure the stability of the native state. In order to maintain the OPEP's advantages, we develop here a new all-atom version of the potential, the extended OPEP maintaining a simple Lennard Jones 6-12 energy to compute the interaction between the different types of carbon, oxygen and nitrogen. As for the CHARMM19 forcefield, all polar hydrogens are explicit

in EOPEP. Hydrogen bridges are also computed between side chain/side chain and main chain/side chain in addition to main chain/main chain, using the traditional 10-12 force term of OPEP. Main chain hydrogen bridge cooperativity is still evaluated as in OPEP. The dimensions adopted for the different side chain atom kinds (including the different kind of carbon implicitly attached to different number of hydrogens) have been taken from the CHARMM19 van der Waals parameters. Parameters are adjusted to generate a stable structure for the NT domain of Calmodulin and protein A, but not for Troponin C.

An additional advantage of EOPEP is that it is possible, with some adaptation, to couple it with the coarse-grained OPEP to provide a multilevel description inside a single protein, decreasing significantly the computational costs while adding the appropriate degree of complexity where needed. Because of their small size, however, both proteins studied here are described fully by EOPEP.

### D. Sequences studied

We focus on two protein domains: the N-terminal nodule of Calmodulin and the N-terminal nodule of the muscular Troponin C (Fig. 1). Each domain consists of two EF-hands whose cooperative binding to two  $Ca^{2+}$  exposes a hydrophobic core. Each EF-hand is defined by an  $\alpha$ -helix-loop- $\alpha$ -helix sequence. This motif is found in several other calcium binding proteins, with some variations. In Calmodulin, each of the EF-hand motif central loop contains 12 residues with one glutamate and several aspartates, whose carboxyl group oxygen atoms that bind to a  $Ca^{2+}$  ion.<sup>60,61</sup> Binding to calcium forces a significant rearrangement of the two helices as they move from a rather parallel orientation in the apo form to a mostly perpendicular orientation.<sup>4</sup> In the closed form, the two calcium binding loops are interrelated by the formation of an interloop short  $\beta$ -sheet stabilized by three hydrogen bonds<sup>22</sup> and identified as the EF-hand  $\beta$ -scaffold by Grabarek.<sup>36,37</sup>

The 76-residues NT nodule CaM apo model used is based on the NMR resolved structures of Ref. 22 (PDB: 1CFD) while the calcium bound NT model comes from the crystallographic characterization of Ref. 23 (PDB: 1CLL). We added the 3 missing first residues to the later model, and used residues of the NT part, up to the 76th. Both NT configurations are illustrated in Fig. 4 panels (f) and (a), respectively.

The Troponin C also presents a dumbbell shape with two E-hands motifs at each end. This protein is dedicated to the activation of actin and myosin muscle microfibers components, by blocking the inhibitor, Troponin I.<sup>62</sup> It is known that

```

CaM NT
EFhand1  ADQLTEEQIAEFKEAFSLFDKDDGGIITTKELGTVMRSL
EFhand2  GQNPTAEALQDMINEVDADGNGTIDFPFLTMMARKM
TpC NT
EFhand1  ASMTDQQAEARAFLEEMIAEFKAAFDMFDADGGGDIKSTELGTVMRMLG
EFhand2  QNPTKEELDAIIEVDEDGSGTIDFEFLVMMVRQMKEDA

```

FIG. 1. Sequences for the NT domains of Calmodulin and Troponin C. Calcium binding residues are in red;  $\alpha$ -helices are in green and the EF-hand  $\beta$ -scaffold residues are underlined. The residues at loop position 7 bind using the MC oxygen.

the CT nodule of Troponin C is structurally linked to Troponin I, while the action of NT nodule is triggered by the calcium ion.<sup>63</sup> The NT domain is composed of 5 helices, as helices 2-5 compose the 2 EF-hand motifs (helix pairs A-B and C-D), while the first helix is denominated “N.” Not counting helix N, the N-terminal domain has 70 percent sequence identity with the Calmodulin NT. Each Troponin binding loops also contains 12 residues, but with a sequence that differs slightly from that of Calmodulin. The first binding loop of the NT has lower affinity to calcium,<sup>64</sup> presenting a glycine at position 5 instead of an aspartate. For the simulations, we used the entire 90 residues of model 01 presented in PDB 1TNQ (Ref. 43) without the  $\text{Ca}^{2+}$  ions. We compared the configurations generated by the simulations with each of the 40 closed models included in PDB 1TNP.<sup>43</sup>

## E. Analysis

For all sequences RMSD is measured on  $\text{C}\alpha$ 's only. For Calmodulin, RMSD includes residues 3 to 76, leaving aside residues 1-3, which are disordered and highly flexible. For Troponin C, we omit the unstable extremities — 1-15 and 86-90 — from the RMSD calculations. RMSD are measured with respect to the non-relaxed original experimentally-derived pdb files. Angles between helices are computed using the principal axis method. The inertia tensor, which includes all main chain and the  $\text{C}\beta$  atoms on each helix, is diagonalized and the principal axis is used as reference. Our analysis also includes individual amino acid energy difference, underlying cooperativity between specific groups. Energy contribution by amino acid is evaluated by adding contribution of every main chain and side chain atom composing this amino acid. Individual atom energy statistic comes from the different forcefield terms where two, three and four-body contributions are split evenly between atoms involved.

## III. RESULTS AND DISCUSSION

### A. Simulations on calmodulin NT

#### 1. Stability of the native apo conformation

We first apply CHARMM19-EEF1 and EOPEP to the native unbound closed model of CaMnt to assess its stability and the size of the apo basin under each potential. For both forcefields, we perform a local unconstrained minimization from the experimental apo form at  $T = 0$  followed by eight independent ART simulations with a 300 K Metropolis temperature and a minimum of 100 accepted events. This is a stringent test as displacement per event can reach 1 Å RMSD.

Both simulation sets lead to stable apo structures, with unmodified  $\beta$ -sheet and  $\alpha$ -helices. With CHARMM19-EEF1, the CaMnt apo form stabilises within a region of about 1.5 to 2.5 Å RMSD away from the experimentally-derived structure, with the lowest-energy structures fluctuating between 2.0 and 2.5 Å (Figs. 2(a) and 2(b)). Using EOPEP, the Calmodulin NT apo form evolves to a region of about 2.5 Å (Figs. 2(c) and 2(d)). Trajectories visiting the lowest-energy structures tend to remain slightly closer to the experimental-derived struc-

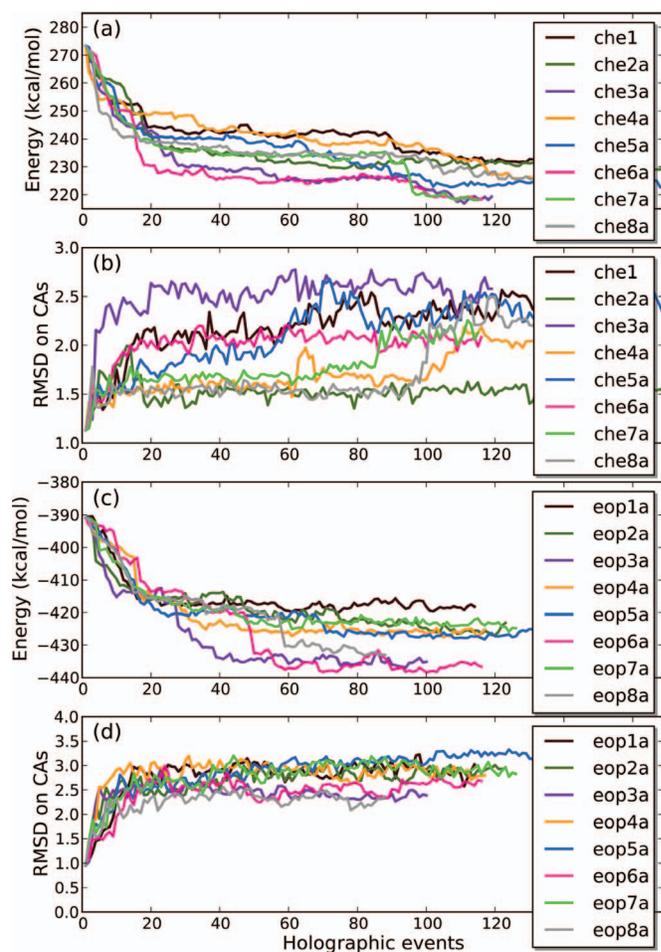


FIG. 2. Energy and RMSD measured from the experimentally-derived closed Calmodulin NT structure as a function of accepted ART-event for eight independent simulation started from the experimental apo conformation. (a) and (b) Energy and RMSD results, respectively, for CHARMM19-EEF1; (c) and (d) same for EOPEP. RMSD are computed on  $\text{C}\alpha$ , excluding unstable residues 1-3.

ture, at about 2.3 Å. In all cases, the displacement away from the experimental data is due to fluctuations in the tertiary structure, with unchanged secondary structure.

#### 2. Folding from the unbound holo conformation

Having established the stability zone of the apo conformation, we now turn to characterizing the closing pathway. For each forcefield, 24 independent simulations with a minimum of 100 accepted events are launched starting from the open model of CaMnt with the calcium ions removed, at 5 Å RMSD of the apo structure, and without imposed biased, with a 300 K Metropolis temperature.

Table I reports the averaged RMSD between initial and final states of events for various holographic ART simulations on the open CaMnt model. The generated events were typically between 1.2 and 1.7 Å. The largest moves are unlikely to lead to low energy structures, however, and at low Metropolis temperature, most moves tend to be rejected. Statistics for accepted events at 300 K show that most lead to displacements are around 0.6 Å.

TABLE I. RMSD statistics for selected trajectories. Averaged displacement measured from the initial minimum to the saddle point and the final minimum for various simulations on both the Calmodulin and Troponin C using OPEP char CHARMM/EEF1 potentials. Statistics are given both for the full sets of events and those accepted only.

Model	Forcefield	Sim	Averaged RMSD in (Å)			
			saddle all events	final min all events	saddle accep.	final min accep.
CAM	CHARMM	14	1.53	1.24	0.73	0.44
CAM	CHARMM	18	1.77	1.64	0.79	0.61
CAM	EOPEP	5	1.78	1.70	0.81	0.62
CAM	EOPEP	9	1.73	1.61	0.83	0.65
TPC	CHARMM	11	1.24	1.09	0.72	0.53
TPC	CHARMM	17	1.35	1.15	0.79	0.58
TPC	EOPEP	16	1.88	1.84	0.83	0.71
TPC	EOPEP	21	1.69	1.63	0.87	0.71

Using this protocol with CHARMM19-EEF1, only one pathway out of 24 reaches the apo state, passing within a minimum distance of 2.2 Å and stabilizing at 2.4 Å. Most other runs visit higher-energy structures at 3.5 to 4.0 Å RMSD from the apo state. For legibility, we show respectively in Fig. 3 the configurational energy (panel (a)) as well as the RMSD (panel (b)) for the four simulations leading to the lowest energy as well as the four with the smallest RMSD with respect to the target CaMnt apo form, for a total of six curves (trajectories che12h and che14h belong to both groups). We see that, for the six trajectories, the energy goes down systematically, reaching a plateau during the runs in all but one case (simulation che12h). RMSD fluctuations are large except for simulation che14h, that reaches the apo state within 20 events.

Folding is easier with EOPEP, and ten out of the 24 runs performed with this forcefield sample within 2.2 Å of the experimentally-derived closed structure, reaching a distance identical to of the closed-state basin in Sec. III A 1. As with CHARMM19-EEF1, Figures 3(c) and 3(d) show the energy and RMSD as a function of accepted even for the four trajectories with lowest energy and RMSD. Again, the minimum RMSD, 1.9 Å to 2.3 Å, is obtained very quickly, typically within less than 30 events. Not all of these apo-like structures are stable; however in some cases, the trajectory evolves away from this conformation as the energy goes down. Here again, an energy plateau is generally attained within 80 to 90 events.

Before looking into the successful closing trajectories, we first summarize the critical closing steps extracted from our simulations, represented in Fig. 4:

1. Increase in the distance between the side chain oxygen atoms of the calcium binding loops.
2. Consolidation/elongation of the  $\beta$ -sheet linking the two calcium loops: a bond is formed between THR29 and GLY61.
3. Repositioning of PHE19 at the outside of the NT node, allowing a tighter closure of helices B-C toward helices A and D. During this process, the PHE19 must pass through a three-residue pocket, LEU-MET-VAL, positioned at the bottom of helix B and where it can be

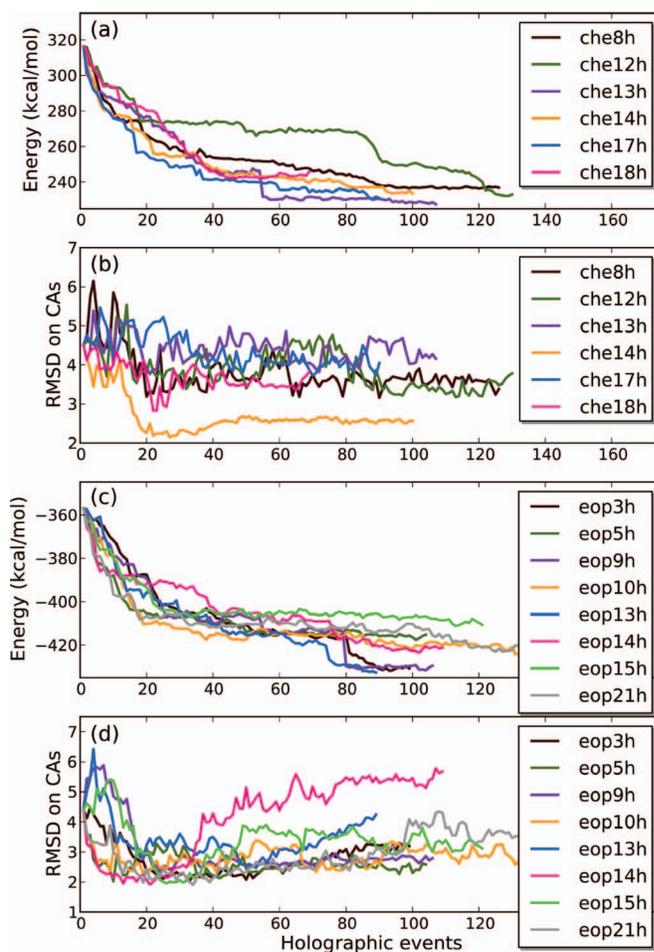


FIG. 3. Energy and RMSD measured from the experimentally-derived closed CaMnt structure as a function of accepted ART-event for independent simulations started from the experimentally-derived open conformation with Ca ion removed. From a set of 24 trajectories, we present the four simulations with the lowest-energy and RMSD structures, respectively. Panels (a) and (b) show the evolution of the energy and RMSD for CHARMM-EEF1 simulations, respectively; panels (c) and (d) present results for EOPEP simulations.

trapped as a metastable state before being ejected from the hydrophobic core.

4. General increase of hydrophobic contacts as each EF-hand helices pair falls into a stable parallel conformation.

Of the 24 EOPEP simulations, 5 fold into the apo state, all following the above mechanisms. For clarity, we focus on the simulation eop9h, which flows near the experimentally-derived closed apo form but also adopts all the important structural elements of the apo state, including the  $\beta$  strand, the exit of PHE19, and the closing of the hydrophobic core.

### 3. Oxygen carrier binding residues

Simulations start from the open model with  $\text{Ca}^{2+}$  removed. The calcium binding residues are located at position 1, 3, 5, 7, and 12 of each of the two binding loops. At position 12, we have a glutamate, GLU31, for the first EF-hand, positioned at the beginning of the EF-hand exiting helix. In the presence of  $\text{Ca}^{2+}$  ions, GLU31 maintains the bottom of helix

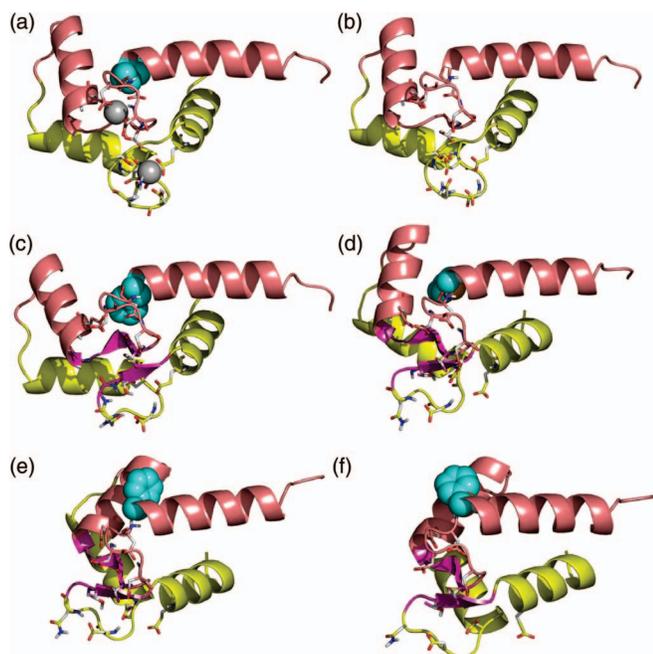


FIG. 4. Folding pathway for simulation eop9h. Each EF-hand pair goes from a rather perpendicular arrangement in the holo state toward a more parallel state in the apo form. (a) 1CLL.pdb model up to residue 76, residues 1-3 added. (b) After a first minimization without  $\text{Ca}^{2+}$ , we observe an increase in the distance between oxygen carrier residues of the loop. (c) Event 3: consolidation of the  $\beta$ -sheet (magenta) linking the 2 calcium binding loops. A hydrogen bridge is formed between THR29 and ILE61. (d) Event 16: Increase in the number of hydrophobe interactions, the core begins to close, but the presence of PHE19 (cyan) prevents a better packing. (e) Event 19: PHE19 gets out of the hydrophobic core, allowing a better closing of the core. (f) Event 24: The helices pack together near the apo form. The end helices (A and D) are held relatively fixed in these views, in order to show the cooperative motion of helices B and C. For ease of viewing, helices A-B (EF-hand 1) are colored in peach and helices C-D (EF-hand 2) in yellow.

B near the calcium loop. As the calcium is removed, it leaves behind several negatively charged oxygen within a close distance in the loop area, a very unfavorable cohabitation.<sup>42</sup> Not surprisingly, this unstable conformation relaxes instantaneously in our simulations; this step has already taken place by event number one. Figure 6(a) shows, however, that not all side chains on the binding loop reach equilibrium at once and a few residues, such as GLU20, GLU56 and GLU58, continue their reorganization. This allows, for example, GLU20 to form several hydrogen bridges with the main chain nitrogen of the loop, acting as an elbow that stabilizes the motion between helices A and B in the apo conformation.

#### 4. $\beta$ -sheet scaffold consolidation

After a few events, a number of EOPEP simulations show the formation of the  $\beta$ -sheet scaffold between the binding loops. This event generally takes place early in simulation (event 22 of eop3h, event 11 of eop5h and eop13h, event 3 of eop9h, event 10 of eop10h, event 4 of eop15h, event 9 of eop21h), except for simulation eop14h (event 36). Fig. 6(b), which shows the four residues involved in this short  $\beta$ -sheet, illustrates the precocity of this event in simulation 9, with a concomitant and abrupt energy drop for THR29 and ILE61.

In parallel, we see that ILE27 and ILE63 consolidate their double hydrogen bridge. These results support the EF-hand  $\beta$ -scaffold model of Grabarek<sup>36,37</sup> which strongly correlates the ILE27 Phi and Psi conformation changes with the angular transition of helices A and B between the CaM closed and open conformations. Interestingly, these first reorganization steps often bring the structure away from the final closed form before converging towards it, as can be seen in Fig. 3 for EOPEP simulations 9, 13 and 15. Among the 10 EOPEP simulations closing within 2.25 Å of the apo state, only simulation eop14h manages to close the protein before zipping the  $\beta$ -sheet. This near native-state structure is unstable, however, and further relaxation pushes the simulation away from the apo state as can be seen in Fig. 3. This mispacking correlates with a delayed  $\beta$ -sheet zipping that occurs at event 36.

None of the CHARMM-EEF1 simulations manages to close the  $\beta$ -sheet scaffold, which explains why only one trajectory out of 24 reaches the apo conformational basin. However, we observe the  $\beta$ -sheet zipping with CHARMM19 without solvation terms (data not shown), thus enhancing the strength of main chain hydrogen bridge. As expected, the inappropriate attractions from charged side chains resulting from the absence of a solvent prevent the molecule from performing a proper closing.

In addition, we also tested for 24 simulations on Calmodulin NT using EOPEP without using the main chain cooperativity term. None of these were able to close the  $\beta$ -sheet scaffold, nor complete the closing transformation. Clearly, as pointed out by Grabarek, the EF-hand  $\beta$ -scaffold is a crucial step on the pathway connecting holo and apo. This transition is schematized by the first two panels of the diagram in Fig. 5.

#### 5. Exit of phenylalanine 19

$\beta$ -sheet zipping is generally followed by the formation of hydrophobic contacts between helices. Examination of Fig. 6(c) indicates that these contacts appear in two phases during simulation eop9h.

First, as ILE27 and ILE63 adopt a  $\beta$ -conformation, a few contacts are initiated, involving PHE16, VAL35 and VAL55 (events 1-10, Fig. 6(c)). Second, around event 19, we observe a more general energy relaxation for hydrophobic core residues. To improve packing, two residues, PHE12 and PHE19, move away from the core. PHE19 escapes the hydrophobic core in only four simulations. From this and other simulations, it seems that an early occurrence helps to further stabilize the structure in the apo conformational basin (event 19 of eop9h, event 15 of eop10h, as well as events 10-14 of che14h). While the RMSD of several other simulations can also reach values below 2.5 Å, their final states do not display the correct position for the PHE19, which should be outside the hydrophobic core (see Figs. 4(e), 4(f), and 7(a)). This suggests that both the  $\beta$ -sheet formation and the PHE19 escape must occur in a specific order during closing to ensure that the protein reaches the correct apo state. The PHE19 has a structurally analog in the CT lobe, PHE 92. In both domains, this phenylalanine is located at the end of the first EF-hand

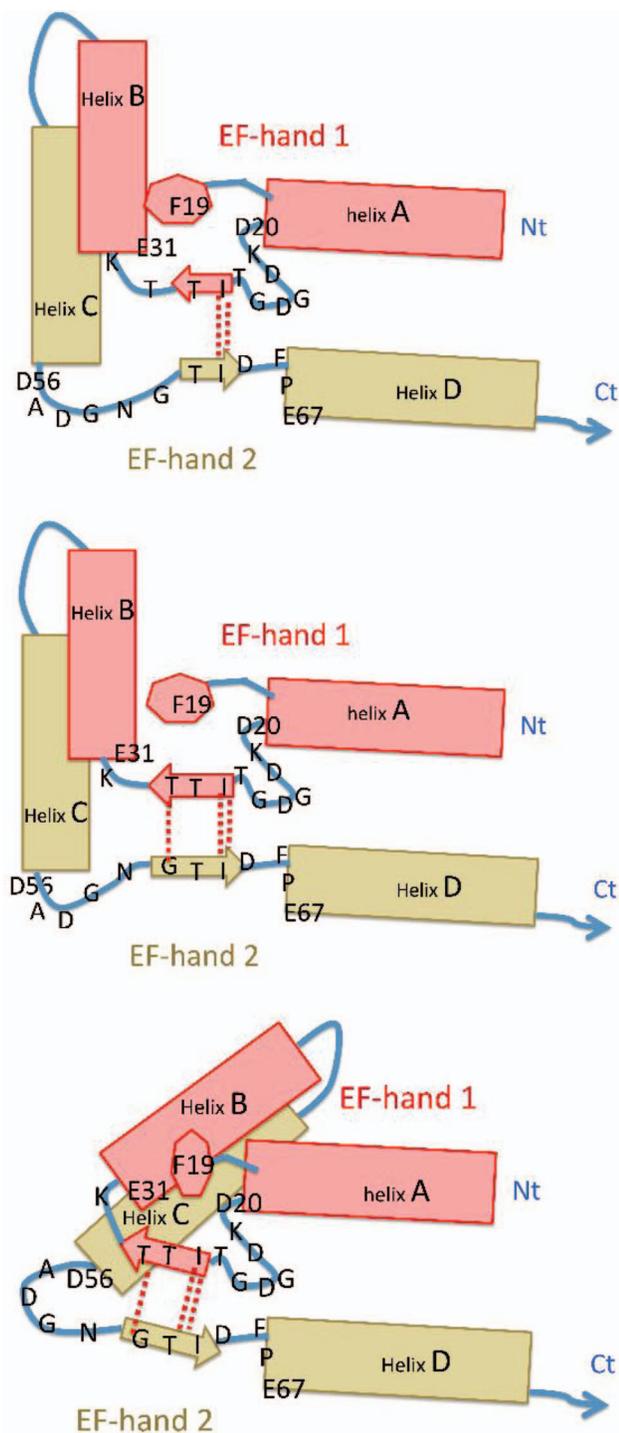


FIG. 5. Diagrammatic representation of the Calmodulin NT mechanism. Top: Initial open state. Middle: Consolidation of the  $\beta$ -sheet, widening out the space between helices, freeing PHE19 latch. Bottom: Closing of the helices in parallel configuration, PHE19 getting outside the hydrophobic core.

helix, just before the first aspartate of the first binding loop. Experimental mutation studies of this position in the CT nodule revealed an instability of the closed CT nodule form when PHE is mutated to ALA. This instability is associated with a decreased in the accessibility of several side chains of the hydrophobic core in the calcium bound open form.<sup>41</sup> From our simulations, PHE19 appears to play a role in the open architecture, behaving as a lock of the EF-hand into the open

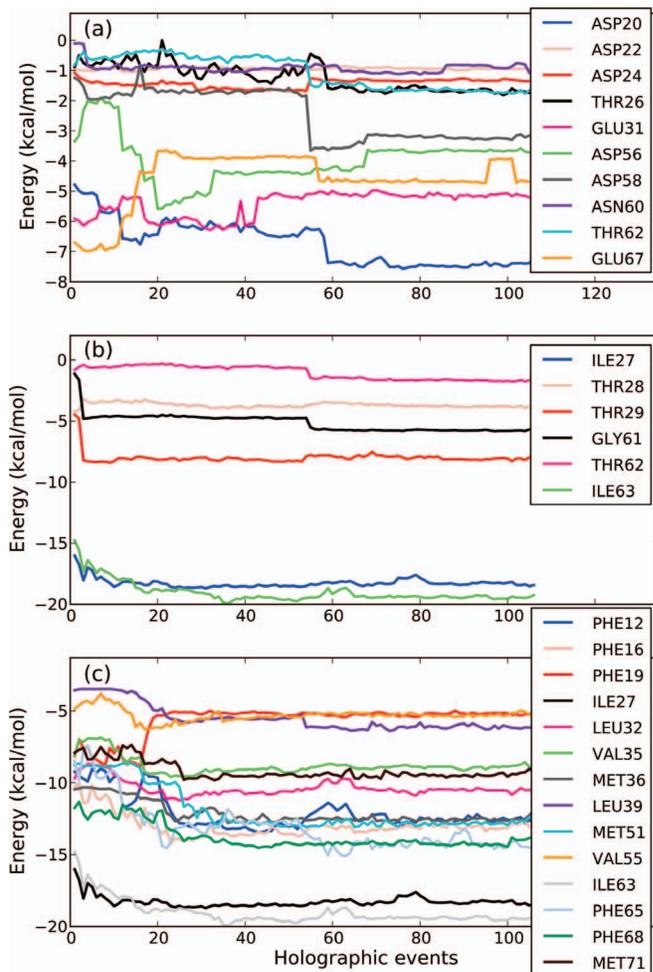


FIG. 6. Evolution of the energy as a function of accepted ART event during eop9h simulation. (a) Evolution of the calcium binding residues. Following calcium removal, the zero temperature minimization already causes a significant energy drop before the first event. (b) Evolution of the six residues participating in the formation of the short  $\beta$ -sheet associated with the closed form. (c) Evolution of the CaMnt 14 hydrophobic residues.

form (see top panel of Fig. 5 and both blue models in Fig. 7 with this position pointed by cyan arrows). Thus, the absence of the bulky phenylalanine at this strategic position may compromise the full opening of the hydrophobic core. Simulations eop9h, eop10h, eop11h and eop23h represent the ideal folding scenario as all structural details associated with the apo state are in place by the end of the simulation. It is instructive to compare with trajectory eop21h, for example, that visits low RMSD structures (2.0 Å) and manages a late expulsion of PHE19 towards the solvent at event 98. The order in which this step takes place is associated with an incorrect packing of the hydrophobic core that prevents full stabilization into the native state. Even though PHE19 is now outside of the hydrophobic core, the conformation does not correspond as well to the native apo state and it moves away from it after event 98 (Fig. 3). With further simulations, however, as discussed in Subsection III A 8, the protein reaches a stable apo state similar to the of the four other fully closed simulations.

Two more sets of 24 simulations showed a similar success rate for reaching the apo from the holo state with, respectively 4 and 5 successful PHE19 exit, after about 100 accepted

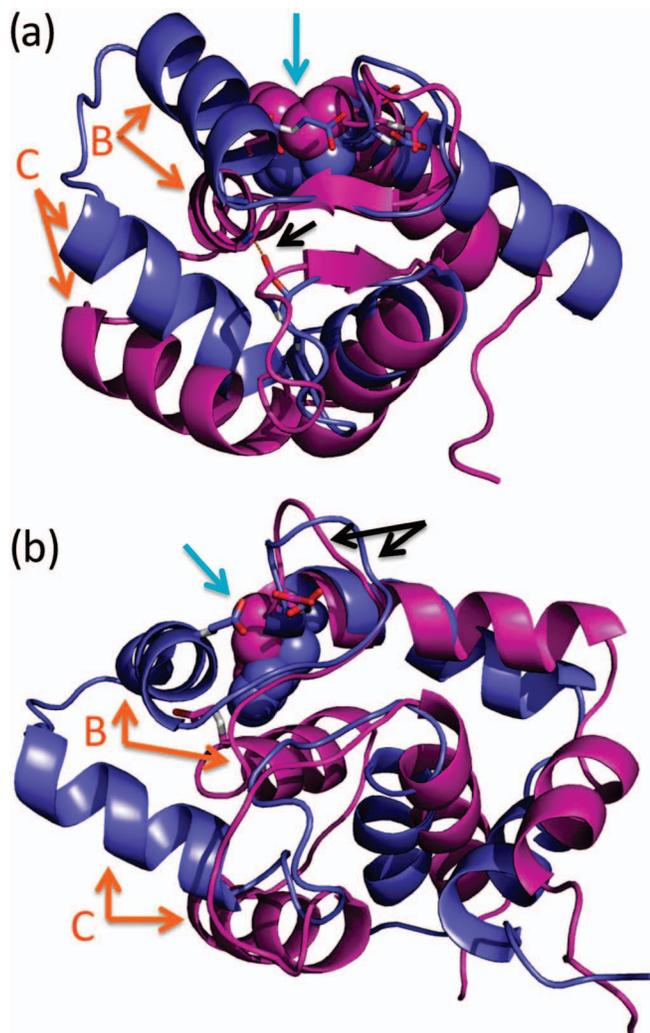


FIG. 7. Comparison between open and close models for both (a) CAMnt and (b) TpCnT. Open conformations are shown in blue and closed structure in magenta. The models are aligned on the stable part of their beta sheets. The orange arrows point out the variation of helices B and C relatively to their respective EF-hand partners A and D. Black arrows indicates the most variable location within or nearby the  $\beta$ -scaffold, not the same for both type of molecules. Cyan arrows point the pre-loop 1 phenylalanine, that undergoes an important relocalization relatively to helix B.

events (data not shown), suggesting that the results obtained here are typical. As shown in Subsection III A 8, extending the simulation time or raising the temperature can increase significantly this rate of success.

### 6. Closure of the hydrophobic core

The evolution of a selection of 14 hydrophobic residues is shown in Fig. 6(c). Simulation eop9h reaches its lowest RMSD, 2.3 Å, at event 24 but an overall energy relaxation for the hydrophobic core extends clearly up to events 85. Starting from the open state, the NT nodule undergoes a motion that first moves away from the closed model before converging back towards the apo basin. It then continues to relax through a series of smaller adjustments as the hydrophobic residues react to each other. The non-linear relaxation is clearly seen at events 17 and 18, when PHE19 leaves the hydrophobic

core for an energetically uncomfortable position that allows several other hydrophobic residues to adopt low-energy conformation, as illustrated by the third panel of Fig. 5. PHE12 initially accompanies PHE19's extraction from the hydrophobic core, but while PHE19 remains in its outside position, PHE12 later inserts back into the core.

### 7. EF-hand helices angular evolution

It is possible to characterize the closing trajectories by following the relative angle between the four helices forming the EF-hand.<sup>1,2</sup> Figure 8 shows the evolution of various angles for simulations eop9h and eop14h. These graphics underline the cooperativity between the two EF-hand pairs. For the first simulation, both pairs A-B and C-D move rapidly, within 20 and 45 events, respectively, from the open to the closed values, and oscillate around their experimental value as the protein adjusts the finer structural details. In some cases, eop14h, for example (Fig. 8, bottom), we observe an overshoot motion in the first helix pair that pushes the protein away from the native state.

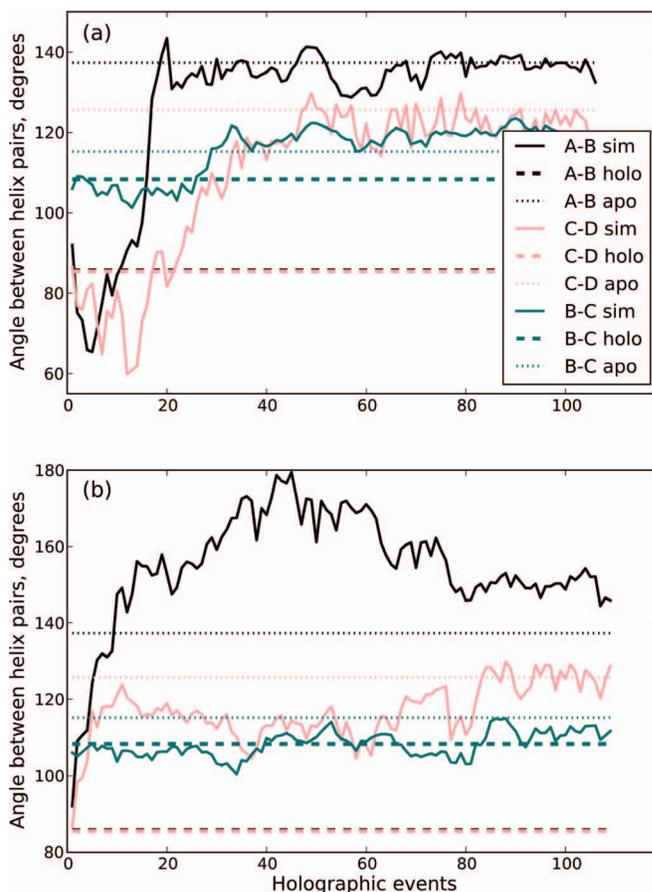


FIG. 8. Top: Evolution of the angles between various helix pairs as a function of accepted ART event for simulation eop9h. Dashed (dotted) lines indicate the value for the experimentally-derived open (closed) form. Bottom: Evolution of the angles between various pairs of helices as a function of accepted ART event for simulation eop14h. Dashed (dotted) lines indicate the value for the experimentally-derived open (closed) form.

### 8. Higher temperature simulations: Characterization of the pathways to the apo state from metastable minima

Even though the low temperature of the first simulation set allows us to clearly identify the various steps needed for bringing Calmodulin into its apo state, it leaves a large fraction of the runs into metastable, non-native conformations. To assess the validity of the model presented in the previous paragraphs, it is necessary to see how the protein manages to escape from these metastable states and find its way to the close conformation.

For this purpose, we extended the simulations on the 24 EOPEP runs at two Metropolis temperatures: 300 K, similar to the initial run, and 900 K, to allow the protein to escape local minima. Even though 900 K might appear large, this temperature is fictive as thermal fluctuations are not taken into account here and the protein still moves from local-energy minimum to local-energy minimum.<sup>47</sup> For both temperatures, 1500 events are attempted.

At 300 K, this leads to a further 2 runs (numbers 5 and 12) that manage to evict PHE19. At 900 K, closing rate is even larger, will 11 runs showing an end conformation with the PHE19 expelled and a conformation compatible with the apo state of Calmodulin, and a further 6 runs with the PHE19 in an intermediate state on the pathway to ejection. Figure 9 shows the evolution of the total energy and RMSD as measured from the experimental state for these 24 runs. For clarity, these are separated into four subgroups: the five runs already in apo state at the onset of the 900 K simulations, the 6 runs that expel PHE19 at 900 K and converged into the closed state, the 6 runs with PHE19 stabilized in an intermediate state and the 7 that do not fold.

As discussed above, the intermediate state can be found on the pathway of each of the 11 simulations that reach a closed state either in the original 300 K or the new prolonged runs (Figs. 9(a) and 9(b)). In this intermediate position, PHE19 of helix A is inserted within a hydrophobic pocket localized at the beginning of helix B. This pocket is formed of LEU 32, MET 36 and VAL 35 (see Fig. 10). This position corresponds to a metastable state favorable to a further PHE19 escape. The importance of stabilizing PHE19 is confirmed by the fact that only runs that do not position PHE19 on the outside or in an intermediate position are rapidly destabilized with an RMSD that reaches between 5 and 9 Å (Fig. 9(d) and Fig. 10(f)). Those seven simulations all show PHE19 inside the hydrophobic core and away from the pocket, correlating with an unstable behavior.

Even if the 900 K conformations are somewhat disturbed compared to those at 300 K, close states with externally positioned PHE19 are stable and never bring this residue back inside the hydrophobic core. This suggests that this movement in solution should only take place after the disruption of the  $\beta$ -scaffold by an incoming Ca ion.

Extended simulations at 300 K and 900 K were also performed with CHARMM19-EEF1. Due to the observed difficulty for forming an elongated  $\beta$ -scaffold associated with this forcefield, no new simulation manages to expel PHE19. Che14t, however, the only run reaching the apo state

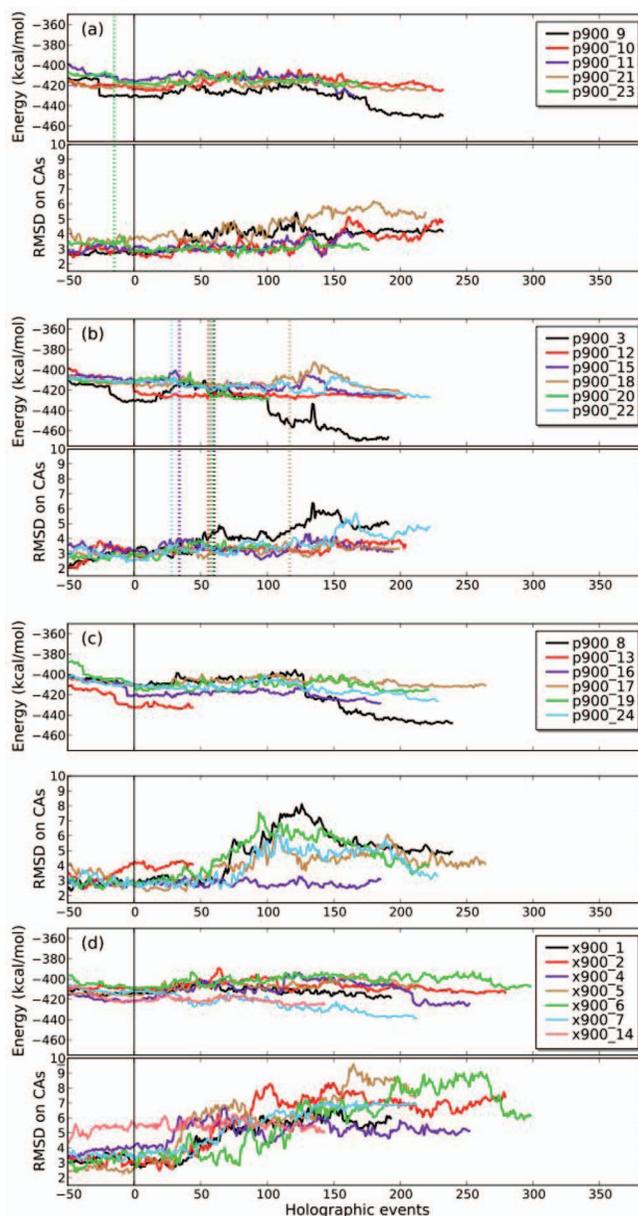


FIG. 9. Extension of the 24 EOPEP Calmodulin NT closing simulations, using a 900 K acceptance Metropolis temperature. The vertical continuous black lines represent the beginning of the runs at this temperature following the last 50 events accepted at 300 K. The energy (top) and RMSD from the experimentally-derived apo state (bottom) are shown for the: (a) 5 simulations that had already succeeded PHE19 escape at metropolis 300 K. (b) 6 additional simulations succeeding PHE19 escape; (c) 6 simulations with PHE19 in an intermediate state. (d) 7 simulations with destabilizing trajectory. The dashed vertical lines in (a) and (b) indicate the escape event of PHE19. The line is not shown for simulations 9, 10, 11 and 21 because this event occurred before the beginning of this graph.

remained very stable at 2.5 Å RMSD during the 900 K run. At 900 K, only simulation 19 manages to reach the intermediate state with PHE19 in the LEU-MET-VAL pocket. This locks the protein in a conformation near the apo state, at 2.5–3.0 Å RMSD. This observation confirms both the importance of the  $\beta$ -scaffold and the PHE19 for stabilizing the close state.

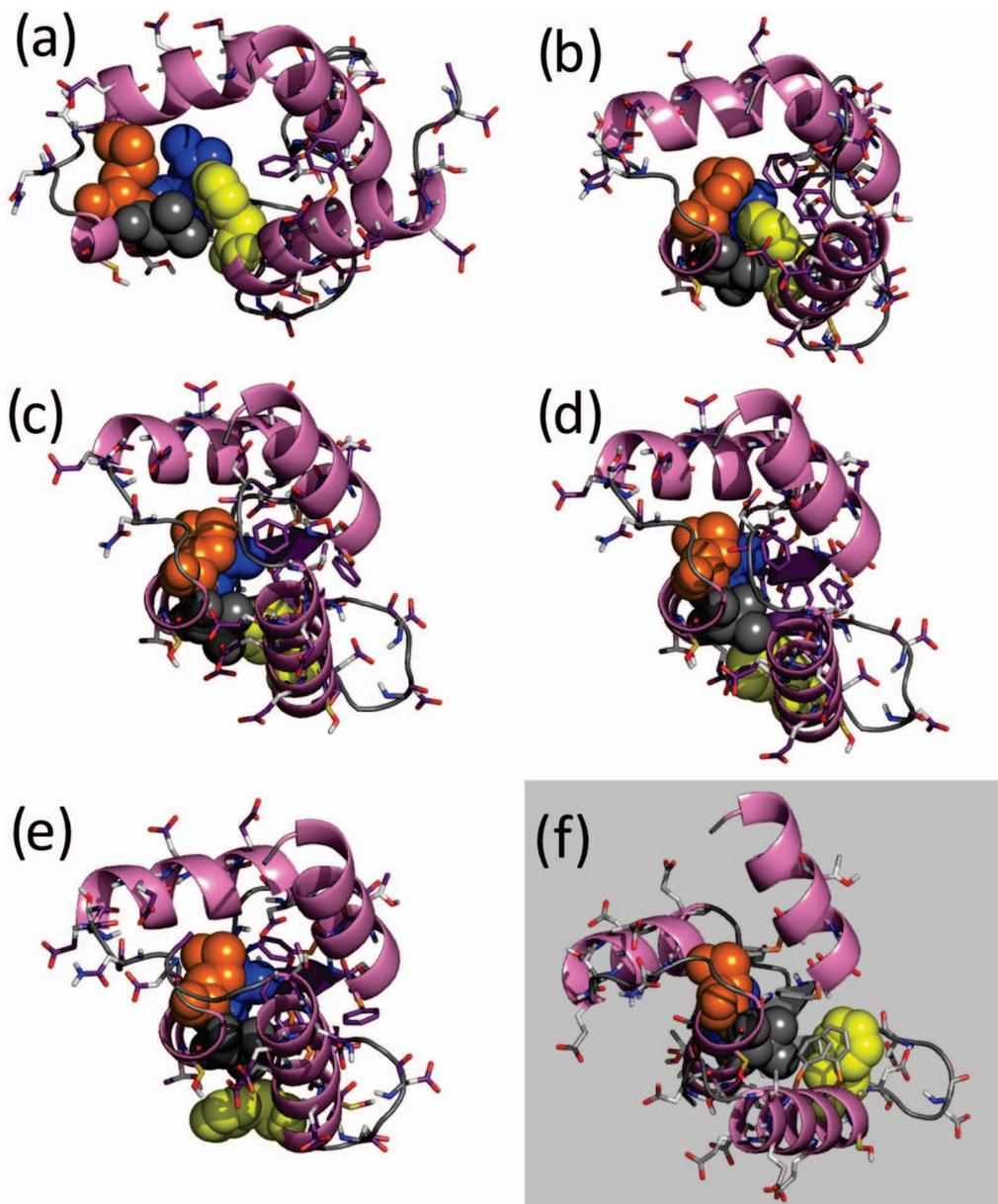


FIG. 10. Cartoon representation of the PHE19 as it enters and leaves the intermediate pocket. PHE19 of helix A and the tree pocket residues of helix B are represented as spheres: PHE19 in yellow, VAL35 in grey, MET36 in orange and LEU32 in blue. (a) The open Calmodulin NT model. (b) While closing, the LEU-MET-VAL pocket receives PHE19. (c) and (d) Valine 35 swivels and PHE19 finds its way toward (e) exit. (f) Typical unstable structures for simulation at 900 K where the PHE19 is neither trapped in the intermediate state nor ejected outside the hydrophobic core. We see the  $\beta$ -sheet in dark grey at the back of the molecule. This structure appears in (c,d,e) in dark purple.

## B. Simulations on Troponin C NT

We also performed CHARMM19-EEF1 and EOPEP simulations on the five helices of the 90-residue muscular Troponin C N-terminal nodule. We used the first model included in the NMR PDB file 1TNQ<sup>43</sup> as starting point for our stability studies. We proceed following the same protocol as for CaM with ART simulations started from the apo state to assess the sequence's stability with both potential, with the RMSD computed from the best fit of 1TNP PDB file, i.e., the 30th structure for CHARMM-EEF1 and the 3rd for EOPEP. In both cases, the apo state is stable and trajectories remain between 1.9 and 2.9 Å RMSD with CHARMM-EEF1 and between 2.3 and 3.3 Å RMSD for EOPEP $\alpha$ .

Twenty-four 100-accepted-event simulations are then launched from the unbound open structure of skeletal TnC NT with both potentials. We used the unbound first model from the PDB file 1TNQ<sup>43</sup> as the initial state. Table I also shows the averaged displacements for TpCnT computed for a number of trajectories. Here, generated events were typically between 1.1 and 1.8 Å with, again, significantly smaller displacements for accepted events at 300 K.

Figure 11 shows the four trajectories visiting, respectively, the lowest-energy structure and the smallest RMSD for CHARMM-EEF1 (top) and EOPEP (bottom). Several CHARMM19-EEF1 simulations (Fig. 11(a)) succeed in converging near 2.5 Å RMSD from the apo state, simulation che18t reaching 2.5 Å. As with CaM, the

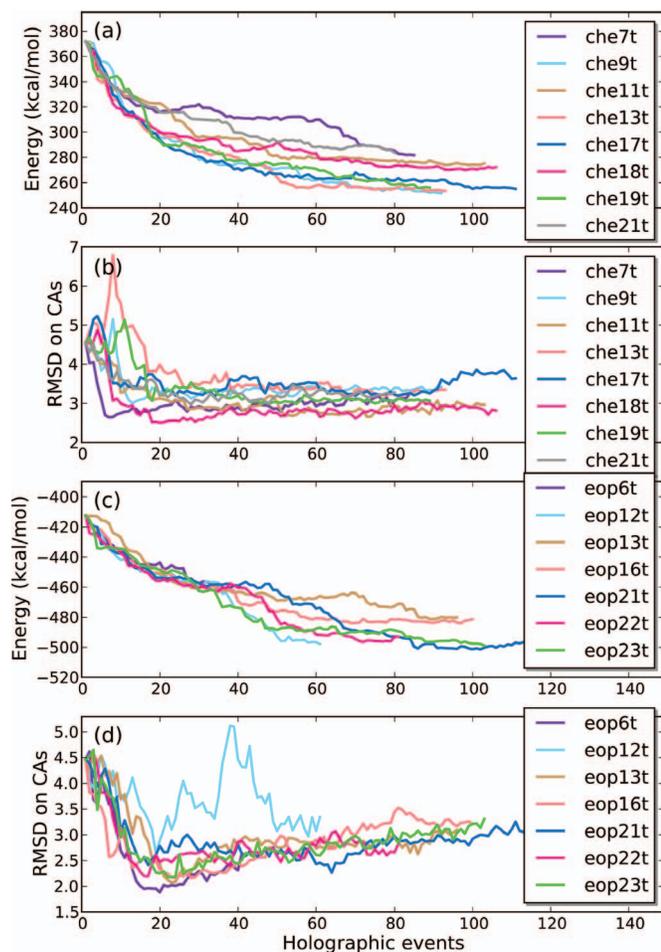


FIG. 11. Energy and RMSD measured from the experimentally-derived closed TpC structure as a function of accepted ART-event for independent simulations started from the experimentally-derived open conformation with Ca ion removed. From a set of 24 trajectories, we present the four simulations with the lowest-energy and RMSD structures, respectively. Panels (a) and (b) show the evolution of the energy and RMSD for CHARMM-EEF1 simulations, respectively. RMSD is relative to model 30 of PDB 1TNP; panels (c) and (d) present results for EOPEP simulations, with RMSD relative to model 03 of PDB 1TNP. Simulation eop22h reaches both energy and RMSD criteria.

structure also collapses rapidly with EOPEP and the trajectories visit conformations between 1.9 and 2.3 Å from the experimentally-derived apo state within the first 20 to 30 events (EOPEP $\alpha$  simulations 6, 13 and 23, Fig. 11), converging to a distance between 2.5 and 3.0 Å.

Folding of Troponin C follows basically the same main steps as Calmodulin NT, with some nuances in the  $\beta$ -scaffold behavior. The first TpCnt EF-hand motif is preceded by a phenylalanine (here at position 29) that plays a function similar to PHE19 and PHE92 in Calmodulin. Within the 24 EOPEP $\alpha$  simulations, PHE29 exit was realized in 8 simulations, including eop13t (events 23-31), eop21t (events 104-110) and eop22t (events 13-15), represented in Figs. 11(c) and 11(d). With CHARMM19-EEF1, PHE29 exit is partially achieved by simulations che17t (events 92-93) and che18t (events 78-80). With this potential, many other simulations also show a relatively low RMSD with the apo form even though PHE19 is not expelled from the hydrophobic core.

As with calmodulin, previous to PHE29 exit, we observe modifications within the Troponin C NT  $\beta$ -scaffold that sets

up the structure for its eviction. As compared with CaM, the region forming the  $\beta$ -scaffold has almost the same conformation in the TpC holo and apo forms, both models showing a single hydrogen bridge, at ILE37-ILE74. However as calcium is removed, two additional hydrogen bridges are formed during the first energy minimization: ILE37-ILE73 and GLY35-PHE75. The  $\beta$ -sheet consolidates therefore from loop 1 position 8–loop 2 position 8 toward helices A and D (loop 1 position 6–loop 2 position 10) while the CaMnt  $\beta$ -sheet, which also starts from the ILE at loop 1 and loop 2 in position 8, zips in the reverse direction, towards helices B and C, involving loop 1 at position 10 and loop 2 at position 6.

In the successful EOPEP PHE29 exit simulations, the  $\beta$ -scaffold conformation change is later followed by a reconfiguration of the 3-glycine area bordering the  $\beta$ -sheet (GLY33-34-35 at loop 1, positions 4 to 6). This zone provides a higher flexibility for TpC, and in most simulations its restructuring closely precedes the exit of PHE29. This could explain why pre-loop 1 phenylalanine exit was realized in a significantly higher proportion of folding simulations for TpCnt than for CaMnt.

The LEU-MET-VAL pocket also plays a role within the Troponin C Nt (residues 42, 46, 45), even though, in EOPEP simulations, PHE29 is less deeply engaged within it as compared with its calmodulin Nt analog. In simulations eop6t and eop16t presented Figs. 11(a) and 11(b), we observe that PHE29 is on its way to exit the pocket, positioned just under valine 45. In simulations eop12t and eop23t, PHE29 is still within the LEU-MET-VAL pocket, in the intermediated state.

In both CHARMM-EEF1 simulations that nearly succeed in expelling PHE (che17t and che18t of Figs. 11(a) and 11(b)), PHE29 is outside the LEU-MET-VAL pocket, but lies between VAL45 and helix A. In the other simulations, PHE29 stays strongly inside the LEU-MET-VAL pocket. We observe considerable instability for the  $\beta$ -scaffold with respect to the closed form model at the THR39 ending in the Troponin C Nt CHARMM-EEF1 simulations. This situation may prevent the LEU-MET-VAL pocket of helix B to gain distance from PHE29 of helix A. This points again at the importance of the  $\beta$ -scaffold formation, even though this structure is less defined in Troponin C NT than in calmodulin NT.

#### IV. CONCLUSIONS AND PERSPECTIVES

We simulated the folding pathway, from Ca<sup>2+</sup>-unbound open holo conformations to closed apo states, of two well-studied EF-hand: the NT domains of calmodulin and Troponin C. Trajectories are generated using an activated algorithm, ART nouveau, along with the multi-scale holographic method, which lead to an unbiased sampling of the conformation space with a focus on the flexible regions. This allows us to complement recent numerical work using Gō-like biased forcefield approaches.

Our results are compatible with the EF-hand  $\beta$ -sheet scaffold model proposed by Grabarek which explains the apo to holo conformation change.<sup>36,37</sup> Here we provide a complete, four-stage, picture of the reverse motion, or folding from holo toward apo conformation: (1) the removal of the Ca<sup>2+</sup> atom, forces the side-chain oxygen atom to repel each

other; (2) upon release of the oxygen carrier residues, the  $\beta$ -sheet scaffold linking the 2 binding loops is able to consolidate, widening out the loops area; (3) this allows helix B swivel towards the inner side of PHE19 (CaMnt) or PHE29 (TpCnt), expelling the residue outside the hydrophobic core after a variable time spent for the PHE19 or PHY29 into a intermediate state; (4) this motion leads to a tighter packing of helices B and C toward helices A and D, optimizing and stabilizing the closing of the hydrophobic core.

There is much interest in the reverse transition, where  $\text{Ca}^{2+}$  ions bind to the two loops. This transition may be favored by the large apo fluctuations shown by experimentalists. But the oxygen carrier residues repulsion is likely to prevent the tightening of the binding loops and, consequently, a significant occurrence of the typical open conformation. Thus, the  $\beta$ -sheet scaffold may act as a switch for both conformational direction changes. The closed form is stabilized by the  $\beta$ -sheet consolidation/elongation tendency along with the hydrophobic attractions between helices. The open form is a consequence of the  $\beta$ -sheet disruption by oxygenated side chains attraction toward calcium, leading to spatial constraints locking open the first EF-hand pair. This description of the  $\beta$ -sheet scaffold conformational switch activated by calcium is also supported by the inability of  $\text{Mg}^{2+}$  to bring up the open conformation neither for calmodulin or Troponin C.<sup>65</sup>

The expulsion of the phenylalanine at the end of helix A appears to be modulated by the presence of a LEU-MET-VAL pocket at the beginning of helix B. The valine acts as a transitional door by which the phenylalanine goes out. The  $\beta$ -scaffold elongation favors a slight rotation of helix B (Fig. 10(b)), which allow the full expulsion of PHE19 or PHE29. Due to the high stability of this state, calcium binding should be necessary to induce the reverse tendency and thus favor reentrance of the phenylalanine.

The holographic approach used in this work has proven able to compute a class of ensemble motions, suitable for the calmodulin and Troponin C NT conformations. In the future, we intend to apply this method to larger systems where the multiscale approach is essential. Here, it allows us to identify a very detailed picture of the atomic scale steps crucial for the large conformational transition observed in EF-hands.

## ACKNOWLEDGMENTS

This work was funded in part by the BiT bioinformatic scholarship program, FQRNT, NSERC and the Canada Research Chair Foundation. Calculations were done using resources from Calcul Québec.

- <sup>1</sup>K. L. Yap, J. B. Ames, M. B. Swindells, and M. Ikura, *Proteins: Struct., Funct., Genet.* **37**, 499 (1999).
- <sup>2</sup>E. Babini, I. Bertini, F. Capozzi, C. Luchinat, A. Quattrone, and M. Turano, *J. Proteome Res.* **4**, 1961 (2005).
- <sup>3</sup>J. L. Gifford, M. P. Walsh, and H. J. Vogel, *Biochem. J.* **405**, 199 (2007).
- <sup>4</sup>T. M. Zhang and M. Ikura, *Nat. Struct. Biol.* **2**, 758 (1995).
- <sup>5</sup>K. L. Yap, J. Kim, K. Truong, M. Sherman, T. Yuan, and M. Ikura, *J. Struct. Funct. Genomics* **1**, 8 (2000).
- <sup>6</sup>X. Shen, C. A. Valencia, J. Szostak, B. Dong, and R. Liu, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5969 (2005).
- <sup>7</sup>D. J. O'Connell, M. C. Bauer, J. O'Brien, W. M. Johnson, C. A. Divizio, S. L. OKane, T. Berggrd, A. Merino, K. S. Okerfeldt, S. Linse, and D. J. Cahill, *Mol. Cell Proteomics* **9**, 1118 (2010).

- <sup>8</sup>M. D. Feldkamp, S. E. O'Donnell, L. Yu, and M. A. Shea, *Proteins: Struct., Funct., Bioinf.* **78**, 2265 (2010).
- <sup>9</sup>J. T. Goncalves and W. Sthmer, *PLoS ONE* **5**, e10873 (2010).
- <sup>10</sup>H. Grotsch, J. P. Giblin, F.-Z. Idrissi, I.-M. Fernandez-Golbano, J. R. Collette, T. M. Newpher, V. Robles, S. K. Lemmon, and M.-I. Geli, *EMBO J.* **29**, 2899 (2010).
- <sup>11</sup>T. Sun, X.-S. Wu, J. Xu, B. D. McNeil, Z. P. Pang, W. Yang, L. Bai, S. Qadri, J. D. Molkenin, D. T. Yue, and L.-G. Wu, *J. Neurosci.* **30**, 11838 (2010).
- <sup>12</sup>G. Smith and M. MacQuaide, *J. Mol. Cell. Cardiol.* **41**, 416 (2006).
- <sup>13</sup>G. Papoff, N. Trivieri, R. Crielesi, F. Ruberti, S. Marsilio, and G. Ruberti, *Biochim. Biophys. Acta: Mol. Cell Res.* **3**, 898 (2010).
- <sup>14</sup>K. Torok, M. Wilding, L. Groigno, R. Patel, and M. Whitaker, *Curr. Biol.* **8**, 692 (1998).
- <sup>15</sup>M. V. Vinogradova, V. S. Reddy, A. S. N. Reddy, E. P. Sablin, and R. J. Fletterick, *J. Biol. Chem.* **279**, 23504 (2004).
- <sup>16</sup>J. Onions, S. Hermann, and T. Grundstrom, *J. Biol. Chem.* **272**, 23930 (1997).
- <sup>17</sup>J. Onions, S. Hermann, and T. Grundstrom, *Biochemistry* **39**, 4366 (2000).
- <sup>18</sup>R. Andersen, Y. Li, M. Resseguie, and J. E. Brenman, *J. Neurosci.* **25**, 8878 (2005).
- <sup>19</sup>H. M. Gradin, U. Marklund, N. Larsson, T. A. Chatila, and M. Gullberg, *Mol. Cell. Biol.* **17**, 3459 (1997).
- <sup>20</sup>H. Asmara, E. Minobe, Z. A. Saud, and M. Kameyama, *J. Pharmacol. Sci.* **112**, 397 (2010).
- <sup>21</sup>Y. Zhang, H. Tan, Y. Lu, Z. Jia, and G. Chen, *FEBS Lett.* **582**, 1355 (2008).
- <sup>22</sup>H. Kuboniwa, N. Tjandra, S. Grzesiek, H. Ren, C. B. Klee, and A. Bax, *Nat. Struct. Biol.* **2**, 768 (1995).
- <sup>23</sup>R. Chattopadhyaya, W. E. Meador, A. R. Means, and F. A. Quiocho, *J. Mol. Biol.* **228**, 1177 (1992).
- <sup>24</sup>N. Tjandra, H. Kuboniwa, H. Ren, and A. Bax, *Eur. J. Biochem.* **230**, 1014 (1995).
- <sup>25</sup>B. E. Finn, J. Evenas, T. Drakenberg, J. P. Waltho, E. Thulin, and S. Forsén, *Nat. Struct. Biol.* **2**, 777 (1995).
- <sup>26</sup>V. Borsi, C. Luchinat, and G. Parigi, *Biophys. J.* **97**, 1765 (2009).
- <sup>27</sup>J. J. Chou, S. Li, C. B. Klee, and A. Bax, *Nat. Struct. Biol.* **8**, 990 (2001).
- <sup>28</sup>Y. Komeiji, Y. Yutaka Ueno, and M. Uebayasi, *FEBS Lett.* **521**, 133 (2002).
- <sup>29</sup>D. Vigil, S. G. Gallagher, J. Trewella, and A. E. Garcia, *Biophys. J.* **80**, 2082 (2001).
- <sup>30</sup>J. Elezgaray, G. Marcou, and Y. H. Sanejouand, *Phys. Rev.* **66**, 0319081 (2002).
- <sup>31</sup>C. Kobayashi and S. Takada, *Biophys. J.* **90**, 3043 (2006).
- <sup>32</sup>Y. Zhang, D. Jasnow, and M. Zuckerman, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18043 (2008).
- <sup>33</sup>Y. G. Chen and G. Hummer, *J. Am. Chem. Soc.* **129**, 2414 (2007).
- <sup>34</sup>S. Tripathi and J. J. Portman, *J. Chem. Phys.* **128**, 205104 (2008).
- <sup>35</sup>S. Tripathi and J. J. Portman, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2104 (2009).
- <sup>36</sup>Z. Grabarek, *J. Mol. Biol.* **346**, 1351 (2005).
- <sup>37</sup>Z. Grabarek, *J. Mol. Biol.* **359**, 509 (2006).
- <sup>38</sup>E. Neria, S. Fischer, and M. Karplus, *J. Chem. Phys.* **105**, 1902 (1996).
- <sup>39</sup>T. Lazaridis and M. Karplus, *Proteins: Struct., Funct., Genet.* **35**, 133 (1999).
- <sup>40</sup>J. Maupetit, P. Tuffery, and P. Derreumaux, *Proteins* **69**, 394 (2007).
- <sup>41</sup>D. F. Meyer, Y. Mabuchi, and Z. Grabarek, *J. Biol. Chem.* **271**, 11284 (1996).
- <sup>42</sup>S. Strobl, C. Fernandez-Catalan, M. Braun, R. Huber, H. Masumoto, K. Nakagawa, A. Irie, H. Sorimachi, G. Bourenkow, H. Bartunik, K. Suzuki, and W. Bode, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 588 (2000).
- <sup>43</sup>S. Gagné, S. Tsuda, M. Li, L. Smillie, and B. Sykes, *Nat. Struct. Biol.* **2**, 784 (1995).
- <sup>44</sup>G. T. Barkema and N. Mousseau, *Phys. Rev. Lett.* **77**, 4358 (1996).
- <sup>45</sup>R. Malek and N. Mousseau, *Phys. Rev.* **62**, 7723 (2000).
- <sup>46</sup>G. Wei, P. Derreumaux, and N. Mousseau, *J. Chem. Phys.* **119**, 6403 (2004).
- <sup>47</sup>J.-F. St-Pierre, N. Mousseau, and P. Derreumaux, *J. Chem. Phys.* **128**, 0145101 (2008).
- <sup>48</sup>S. Santini, G. Wei, N. Mousseau, and P. Derreumaux, *Structure* **12**, 1245 (2004).
- <sup>49</sup>G. Wei, N. Mousseau, and P. Derreumaux, *Biophys. J.* **87**, 3648 (2004).
- <sup>50</sup>M.-R. Yun, N. Mousseau, and P. Derreumaux, *J. Chem. Phys.* **126**, 105101 (2007).

- <sup>51</sup>L. Dupuis and N. Mousseau “Holographic multiscale method used with non-biased atomistic forcefields, for simulation of large transformations in protein,” *J. Phys.: Conf. Ser.* (in press).
- <sup>52</sup>H. Gohlke and M. Thorpe, *Biophys. J.* **91**, 2115 (2006).
- <sup>53</sup>N. Mousseau, P. Derreumaux, G. T. Barkema, and R. Malek, *Mol. Graph. Model.* **19**, 78 (2001).
- <sup>54</sup>G. Wei, W. Song, P. Derreumaux, and N. Mousseau, *Front. Biosci.* **13**, 5681 (2008).
- <sup>55</sup>R. Lavery, I. Parker, and J. Kendrick, *J. Biomol. Struct. Dyn.* **4**, 443 (1986).
- <sup>56</sup>R. Lavery, K. Zakrzewska, and H. Sklenar, *Comput. Phys. Commun.* **91**, 135 (1995).
- <sup>57</sup>M.-R. Yun, R. Lavery, N. Mousseau, K. Zakrzewska, and P. Derreumaux, *Proteins: Struct., Funct., Genet.* **63**, 967 (2006).
- <sup>58</sup>J. W. Ponder, <http://dasher.wustl.edu/tinker> for the Tinker software.
- <sup>59</sup>R. Laghaei, N. Mousseau, and G. Wei, *J. Phys. Chem.* **114**, 7071 (2010).
- <sup>60</sup>A. Lewit-Bentley and S. Réty, *Curr. Opin. Struct. Biol.* **10**, 637 (2000).
- <sup>61</sup>M. Lepsik and M. J. Field, *J. Phys. Chem. B* **111**, 10012 (2007).
- <sup>62</sup>C. S. Farah and F. C. Reinach, *FASEB J.* **9**, 755 (1995).
- <sup>63</sup>C. S. Farah, C. A. Miyamoto, C. H. I. Ramos, A. C. R. da Silva, R. B. Quaggio, K. Fujimoi, L. B. Smillie, and F. C. Reinach, *J. Biol. Chem.* **269**, 5230 (1994).
- <sup>64</sup>S. Gagné, L. Tsuda, S. Spyropoulos, L. E. Kay, and B. D. Sykes, *J. Mol. Biol.* **278**, 667 (1998).
- <sup>65</sup>Z. Grabarek, *Biochim. Biophys. Acta* **1813**, 913 (2011).