JMB

Dynamical Properties of the MscL of *Escherichia coli*: A Normal Mode Analysis

H. Valadié[1], J. J. Lacapčre[2], Y.-H. Sanejouand[3] and C. Etchebest[1]*

[1]*Equipe de Bioinformatique Génomique et Moléculaire EMI03-46, Université Paris 7 2, place Jussieu, case 7113 75251 Paris Cedex 05, France*

[2]*INSERM U410, Faculté de Médecine Xavier Bichat, 16 rue Henri Huchard, B.P. 416 75870 Paris Cedex 18, France*

[3]*Laboratoire de Physique, UMR 5672 du CNRS, Ecole Normale Supérieure, 46 allées d'Italie 69364 Lyon Cedex 07, France*

*Corresponding author

The mechanosensitive channel (MscL) is an integral membrane protein which gates in response to membrane tension. Physiological data have shown that the gating transition involves a very large change in the conformation, and that the open state of the channel forms a large non-specific pore with a high conductance. The *Escherichia coli* channel structure was first modeled by homology modeling, starting with the X-ray structure of the homologous from *Mycobacterium tuberculosis*. Then, the dynamical and conformational properties of the channel were explored, using normal mode analysis. Such an analysis was also performed with the different structures proposed recently by Sukharev and co-workers. Similar dynamical behaviors are observed, which are characteristic of the channel architecture, subtle differences being due to the different relative positioning of the structural elements. The ability of particular regions of the channel to deform is discussed with respect to the functional and structural properties, implied in the gating process. Our results show that the first step of the gating mechanism can be described with three low-frequency modes only. The movement associated to these modes is clearly an iris-like movement involving both tilt and twist rotation.

Introduction

Mechanosensitive channels are ubiquitous proteins, located in the inner membrane of bacteria.[1] The presence of highly conserved homologues in various bacteria (Gram+ and Gram−) suggests that they play an important role in diminishing turgor pressure upon osmotic downshock,[2] when bacteria are shifted from an high to a low osmolarity environment. Indeed, these channels have remarkable mechanical properties: they open in response to membrane strain, preventing cell lysis and the breakdown of the membrane. They act as "safety valves" for the bacteria cell. Small solutes (ions and water molecules) and much larger enti-

ties (proteins like thioredoxin-12 kD) can pass through the pore with no ionic selectivity.[1,3]

Three types of mechanosensitive channels have been characterized based on their physical properties (conductance, time of opening, pressure sensitivity): mechanosensitive channel of mini conductance "MscM", small conductance "MscS" and large conductance "MscL" (*C* > 2.5 nS). The MscL from bacteria *Escherichia coli* has been the focus of many electrophysiological, mutational and biophysical studies.[4,5] One of these studies proposed that the fully open state of the MscL consists of a large non-specific pore, with a large conductance, which deduced diameter is about 30−40 Å.[6] However, the three-dimensional structure of the open form remains to be determined. Rees and co-workers[7] have recently solved by X-ray crystallography the structure of a homologous mechanosensitive channel in its closed state, from *Mycobacterium tuberculosis*. This structure is consistent with the major structural properties expected for the MscL of *E. coli*. The channel consists of five identical subunits arranged around a central pore. Each subunit is composed of a transmembrane and a cytoplasmic domain. The transmembrane

region is defined by two helices called TM1 and TM2. The TM1 (the "inner" helix) lines the permeation pathway, whereas the TM2 (the "outer" helix) is positioned on the outside of the channel, facing the lipids. The diameter at the entrance of the pore is approximately 35 Å, whereas, in its narrowest part, the minimum diameter is only 2 Å. Now, the main goal of structural studies is to elucidate the major conformational change occurring during the gating mechanism and to explain how tension of the surrounding lipid bilayer causes the channel opening.

On the basis of experimental data, two models for the gating mechanism and the plausible open structure were proposed. The first model (called the "10 helix pore" model) is based on the assumption that, in the open state, the pore is lined by the ten transmembrane helices of alternating TM1 and TM2 helices.[7–10] This model was elaborated on the basis of mutagenesis of amino acid residues located at the interface of TM1 facing the pore, which affected channel opening considerably. Some biophysical data are rather in conflict with the structural properties of the postulated open state of this model, namely: (i) the deduced diameter of the pore in the open state; (ii) the separate reconstruction of the two "domains" of the MscL, each domain containing either the TM1 or the TM2 helix. The TM2-domain has no electrophysiological activity, while the TM1-domain is able to form channels, but has no mechanosensitivity. These results suggest that the TM1 helices, and not the TM2 helices, participate in the structure of the open pore (A. Ghazi, personal communication).

In the second model proposed by Sukharev & Guy,[11] the open state is obtained through a twist and a tilt of the TM1 helices with no major rearrangement of the TM2 helices. This model is nowadays the most studied and the more consistent with most recent experimental data (electrophysiological measurements, cross-linking with cysteine residues). The proposed gating mechanism is more complex than the first one, since a two-step process with different subconductance states is assumed. One main feature of this detailed model of the opening motion is the hypothesis that the first 12 residues, in each monomer, unresolved in the crystal structure, fold into an α-helix (called S1), the five S1 helices forming an N_{term} S1 helical bundle, largely involved in the gating mechanism.[11] In the early stage of the gating mechanism itself, the channel undergoes a conformational transition which implies a surface expansion of the protein in the bilayer with no significant increase in the size of the inner pore. This first, tension-dependent step leads to a closed-expanded structure (CE), which is still occluded by a ring formed by the amino acid side-chains of the S1 helix-bundle. The second step leads to the fully open state. Sukharev and co-workers also modeled a sequence of states along the opening pathway[11] in order to describe as best as possible the different structural states and their conformational changes for the MscL of *M. tuberculosis* and *E. coli.*

In parallel with experiments devoted to the understanding of the gating mechanism, molecular modeling approaches are used in order to study protein conformational dynamics, mainly: (1) molecular dynamics (MD) simulations, with numerical integration of the classical Newton equations of motion; (2) targeted dynamics simulations; and (3) normal mode analysis. Due to the small value of the timestep required for numerical integration, i.e. 1–2 fs, simulations of protein conformational changes occurring on a ms timescale remains a challenge from a MD simulation perspective, especially if the simulation is performed for the case of a membrane protein in an explicit medium (lipids and water molecules), using classical parameters ($T = 300$ K). Nevertheless, MD simulations of the *M. tuberculosis* MscL channel have been performed.[12,13] These studies have given useful insights about the regions of relative structural stability and instability in the structure. For the second method, the steered dynamics, two different structures are needed, one for each endpoint of the conformational pathway. Since the crystallographic structure of the open state is lacking, Ma and co-workers[14] explored the transition pathway between the open and closed modeled structures recently proposed by Sukharev and co-workers. Very recently, an interesting work has been realized without any hypothesis about the different steps of the gating mechanism. In order to simulate the effect of the membrane tension, the authors have performed MD simulations, for the *M. tuberculosis* MscL, with different pressures in the presence of a membrane model. Their results propose a third possibility for the channel gating.[15] Nevertheless, the authors admit that the results have been obtained with unrealistic pressure conditions. These conditions have the advantage to accelerate the conformational changes not currently accessible in the classical simulation time scale but they could provide important artifacts in the results obtained. Normal mode calculations provide an alternative to MD simulations for studying collective motions in macromolecules.[16–18] Normal mode analysis is based on the diagonalization of the second derivatives of the mass-weighted energy matrix. The global motion of the system is then expressed as a superposition of collective variables, called the vibrational normal modes. High frequency modes are highly localized motions, of a few side-chains, of pairs of bonded atoms, etc., while the lowest frequency ones are collective motions of large groups of atoms, usually whole structural domains. These later modes mainly depend on the shape of the molecule, as shown for instance by Bahar and collaborators[19] in a study of four structurally similar but functionally different proteins. Moreover, it was found that a handful of such collective motions, corresponding to a small subset of the lowest-frequency modes, often compare well with the conformational change observed upon ligand binding,[20–23] especially when the considered

conformational change has an highly collective character.[24,25] Thus, normal mode analysis seems to be the best suited theoretical method for studying collective motions in proteins, in particular when a large modification of the structure is expected.

Classical normal mode analysis is limited by the energy minimization step, which can be very long for complex macromolecules, and by the diagonalization of the $3N$ dimensions Hessian matrix, where $3N$ is the number of degrees of freedom of the system. Recent developments permit to circumvent the second limiting factor using either an iterative diagonalization method[26] or an approximation resting on the fact that low-frequency motions are still well described when amino acid residues are assumed to behave like rigid bodies,[27–29] or both factors by means of a simplified potential[30] and chain representation (only one particular per residue).[31–33] Normal mode analysis were preformed on a set of proteins[24,31] using such a simplified method, and, in parallel, the same calculations were done using a standard semi-empirical potential. The results obtained with the approximate normal mode calculations are in agreement with those obtained with the classical method. These studies demonstrate that such approximations (force field and simplified model) provide a good description of domain motions, as described by the low frequency normal modes calculated with classical methods. Moreover, these results show that such a simplified approach is sufficient for studying backbone protein motions of large systems, as far as the low frequency modes are concerned.

The aim of the present study is to provide further insights, at the residue level, of the gating mechanism. To make good use of the numerous biological data related to the MscL of *E. coli*, the first step of the study consists in elaborating a structural model of this channel in its closed state (MscL of *E. coli*), based on the X-ray structure, by homology modeling. In a second stage of the study, the normal mode analysis method is applied to this model, as a predictive tool. Our work is based on the assumption that the zones that are easy to deform in the absence of a mechanical constraint are those involved in the response to membrane strain. The normal mode analysis is expected to provide new insights about the flexible and rigid parts of the protein. It could also anticipate the preferential conformational change occurring in the MscL during the gating mechanism. We use the method developed by Tirion[30] to perform the normal mode calculations. On the basis of a homology modeling model that could be rather inaccurate at the atomic level, the use of such a simplified potential is the most appropriate.

Then, the normal mode calculated were compared to those of other closed structures of the MscL, the first one corresponding to the X-ray structure and the second one to the closed state of the postulated Sukharev & Guy's model (SG model). Finally, the hypothetical pathway proposed previously by Sukharev and co-workers is submitted to normal mode calculations too, and then, discussed within the frame of this theoretical approach. Three structures were considered along this pathway: the closed state, noted C, the first open state, noted CE, and the fully open state noted O. These structures correspond to the models numbered 1, 9 and 12, respectively, of the different models proposed by Sukharev and co-workers.[11] Due to the lack of detailed knowledge about the open state, we use normal mode analysis both in order to try to predict the first step of the conformational pathway but also, as a way to evaluate the SG model.

Results

E. coli MscL homology model

All the alignments tested between *E. coli* and *M. tuberculosis* shared about 36% of sequence identity and 73% of sequence similarity. In almost all the generated models, the transmembrane helices are correctly modeled, whereas the loop regions show much more diversity. The difficulty to model these zones comes from the presence of very different amino acid residues in the two sequences, which makes alignment difficult. For the cytoplasmic parts, we observe either interactions between the backbone of the different monomers or an unstructured region. According to the criteria chosen (see Materials and Methods), the best model is obtained using the sequence alignment shown in Figure 1. In this case, the sequences of the MscL of *M. tuberculosis* and *E. coli* share 37% of sequence identity. In agreement with previous multiple alignments of MscL family,[9,7] the TM1 helix is the most conserved region.

The analysis of the geometrical parameters of the TM helices of each monomer is used for describing the general features of the MscL polymeric architecture. Overall geometrical parameters for the MscL of *M. tuberculosis* and *E. coli* are quite similar (Table 1). This means that the minimization procedure has only slightly modified the modeled structure obtained from the Modeller package. Both the range of inter-helices closest distances and tilt or Ω angles are in agreement with those observed for most transmembrane proteins.[34]

The channel constriction is formed by a ring of hydrophobic residues arising from the five subunits in both structures, Ala18 and Val21 in the *M. tuberculosis* structure, Ala20 and Val23 in the *E. coli* structure. The positioning of the periplasmic loop, longer in *E. coli* model, is quite comparable. Besides these general common features, structural differences exist between the *M. tuberculosis* MscL X-ray structure and the *E. coli* MscL homology model due to differences in their amino acid sequences. The positioning of rings of positive (Lys) or negative (Asp) charges in the pore differs

```
TbMscL 10 ARGNIVDLAVAVVIGTAFTALVTKFTDSIITPLINRIGVNAQSDVGILRIGIGGGQTI  67
          ***:*****:*:**:**   :*::::  :** *  ::  :  *::  ::  :  **  : *    ::
EcoMscL 13 RGNVVDLAVGVIIGAAFGKIVSSLVADIIMPPLGLLIGGIDFKQFAVTLRDAQGDIPAV  71


                                                                ~
TbMscL 68  DLNVLLSAAINFFLIAFAVYFLVVLPYNTLRKKGEVEQPGDTQVVLLTEIR      118
           :   :*:::  :::*::*** ::: : *   :   *** *  :   *: *  :  *******
EcoMscL 72  VMHYGVFIQNVFDFLIVAFAIFMAIKLINKLNRKKEEPAAAPAPTKEEVLLTEIR      127
                                               ~
```

Figure 1. Pairwise alignment of the MscL primary sequence from *M. tuberculosis* (TbMscL) and *E. coli* (EcoMscL). Each monomer is composed of 109 and 114 amino acid residues. The total number of residues of each monomer is 151 and 136 in the two species, respectively. Identical residues (∗) and conservative substitutions (:) illustrate the high degree of conservation between the two sequences. Sequences were aligned using the Modeller package. Locations of transmembrane helices TM1 (continuous line) and TM2 (broken line) are shown. For normal mode calculations, only amino acid residues until the one marked with the symbol ∼ were taken into account.

in the two structures, as shown in Figure 2(a) and (b).

A contact map allows us to visualize the interactions in structures of interest. The C^α contact map of our modeled structure is displayed in Figure 2(d) (in black). For a monomer *i* (left central triangle), many contacts exist inside the monomer. The interactions between monomers are limited to the two neighboring monomers ($i − 1$ to $i + 1$). The only region where the five monomers are in contact through the Val23 residues corresponds to the center of the pore.

Similar maps are observed for the *M. tuberculosis* X-ray structure (data not shown).

When we compared our homology model to the *E. coli* SG model closed structure, the overall geometrical parameters were found to be similar (Table 1). The main differences lie in the extent of the helical regions, considerably longer in the SG's model. A similar relative orientation of the two helices in both models is observed (1.2 Å for the rmsd of TM region). In contrast, the conformation and the location of periplasmic region with respect to the entrance channel is considerably different, as illustrated by the C^α contact map in Figure 2(d), in gray. In addition, the SG's model contains the $N_{terminus}$ region (1–12) folded as an α helix. The $S1_i$ region contacts the neighboring $S1_{i−1}$ and $S1_{i+1}$ regions. No contact is observed between the N_{term} and the C_{term} regions.

Normal modes calculations

Due to the ambiguity concerning the cytoplasmic region in the X-ray structure, we have examined an additional structure where residues following Pro109 are excised. This region has been shown to be non-essential for channel gating and function.[35,36] The calculation was performed on the two structures, with and without the $C_{terminus}$ region. We have verified that the deleted parts do not have a significant influence on the atomic displacements and on the global motions of the MscL, as observed with the normal mode analysis. Hereinafter, we will only discuss in detail the results obtained with the structure where residues following Pro109 are excised.

In order to measure the influence of the details of the model structure on the motions deduced from the normal analysis, we have also carried out a similar analysis starting from the X-ray structure and from different structures of the SG's model.

Our homology model

Since the largest conformational changes are known to be associated with the lowest frequency normal modes, we calculated the 106 lowest frequency normal modes for our structural model. The first six modes correspond to global movements of the molecule (translations and rotations,

Table 1. Geometrical parameters for the transmembrane helices (TM1 and TM2) for each studied system

		Crystallographic structure *M. tuberculosis*	Homology modeling model *E. coli*	SG *E. coli*
Length	TM1	19 [22−40]	20 [21−40]	36 [15−50]
	TM2	22 [71−90]	20 [77−96]	31 [76−107]
Dist$_{G-G}$	TM1$_i$−TM1$_{i+1}$ (P)	14.3	12.5	13.3
	TM2$_i$−TM2$_{i+1}$ (P)	20.4	21.1	19.9
	TM1$_i$−TM2$_{i+1}$ (A)	9.8	10.7	14.8
Dist$_{sa}$	TM1$_i$−TM1$_{i+1}$ (P)	9.6	7.9	6.7
	TM2$_i$−TM2$_{i+1}$ (P)	18.7	18.3	18.0
	TM1$_i$−TM2$_{i+1}$ (A)	9.0	9.0	10.2
τ	TM1	37	37	32
	TM2	152	146	152
Ω	TM1$_i$−TM2$_i$ (A)	141	134	143

Length, length of the helix (number of amino acid residues), the first and the last residues are given between brackets. Dist$_{G-G}$, distance between the barycenters of the helices (Å). Dist$_{sa}$, smallest distance approach as defined by Chothia (Å). τ and Ω, tilt angle and inter-axial angle, respectively (°). *i* refers to a monomer, $i − 1$ and $i + 1$ to the previous and the following monomers, respectively. A refers to anti-parallel helices and P to parallel helices.
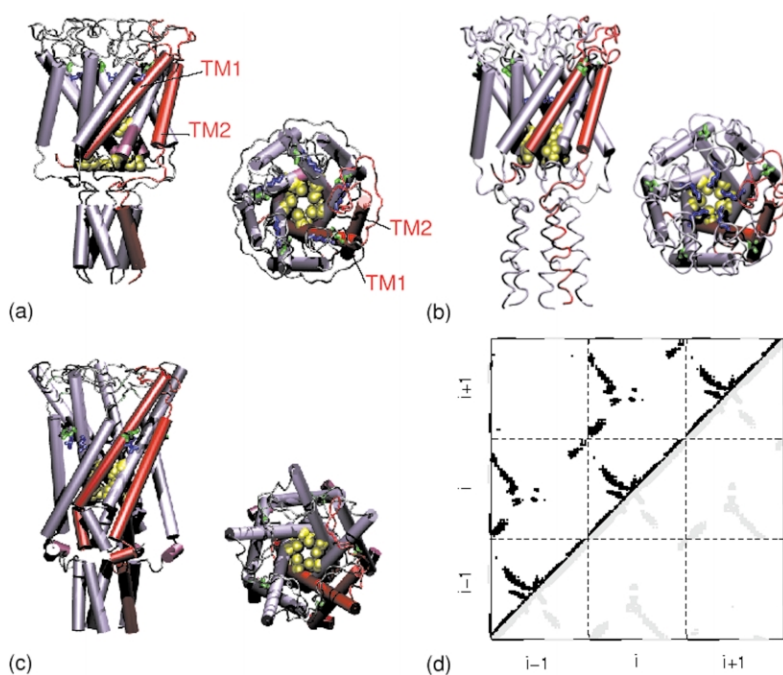
Figure 2. (a)–(c) Two molecular graphic views (side view and top view) of: (a) the crystallographic structure of the closed form of the MscL of *M. tuberculosis* (residues 10:118); (b) our homology modelling structure of the closed form of the MscL of *E. coli* (residues 13:126); and (c) the SG's modeling structure of the closed form of the MscL of the *E. coli* (residues 1:133). Rings of positive charges are shown in blue (stick), of negative charges in green (stick), and hydrophobic residues occluding the pore in yellow (spacefill). In (a), one monomer is highlighted in red. (d) Map of minimum atomic (C^α) distance between residues for three consecutive monomers ($i − 1$, i and $i + 1$) in the structure (from residue 13 to residue 126). The upper left part of the contact map corresponds to our homology modeling structure. The lower right part of the map corresponds to the SG's structure. The transmembrane helical segments (TM1 and TM2) are annotated in black and gray lines, respectively. Figures were made with VMD.[65]

with a zero frequency) and will be ignored in the following discussion. Hereinafter, the following 100 modes are thus renumbered, starting from one.

The frequencies lie between $65 \, cm^{-1}$ and $315 \, cm^{-1}$. Such frequency values are large compared to values generally obtained with standard normal mode analysis. This is due to the fact that the potential energy is determined apart from a constant c which refers to the strength of the potential (see Materials and Methods). In our case, the constant c was arbitrarily fixed to a value of $10 \, kcal/Å^2 \, mol$. Such an initial spring constant value is often modified in order to best fit the mean quadratic displacement values of each C^α, as deduced from the X-ray crystallographic *B*-factors.[33] In the present case, the poor resolution of the crystal structure ($R = 3.5 \, Å$) impaired the

B-factor determination and prevented the c fit. Thus, here, frequency values will not be discussed in a quantitative way and we will refer to each normal mode by its mode number rather than by its frequency value.

Many modes are degenerated, exhibiting identical frequencies. For instance, pairs of modes have frequencies that are identical with each other. This degeneracy is due to the $C5$ symmetry[37,38] of the system. Such a feature was also observed previously, in the cases of double-stranded DNA,[39] (Dialanine)$_{60}$,[40] and a virus.[41] Other modes, for example, modes 1, 10, 15, 18 and 21, have non-degenerate frequencies.

For each low-frequency normal mode, different structures were generated, through a displacement along the normal mode considered. The overall C^α
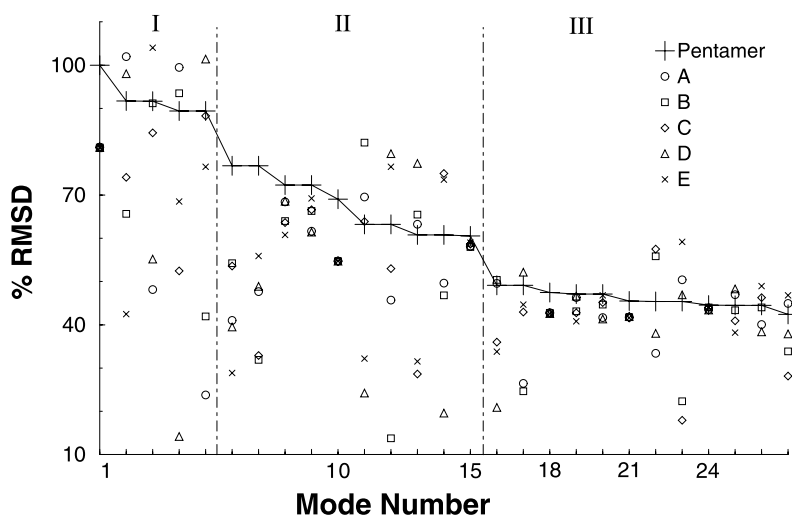

Figure 3. Percentage of root-mean-square displacements relative to the rmsd associated with the slowest frequency mode as a function of mode number (along the modes 1 to 27). The line represents the overall rms displacements for the pentamer. The intra-monomer displacements are refered by symbols (○, □, ◇, △, × for monomers A to E, respectively). In this case, only displacements within one monomer are considered, whereas its orientation is kept fixed. Vertical lines separate the different groups (I to III).

rmsd, as well as the C^α rmsd for each monomer, with respect to the initial structure, were calculated using ProFit (A.C. Martin ProFitV1.8). Next, all rmsd values were normalized with respect to the overall rmsd of the lowest frequency mode. Three main groups of rmsd values are observed (see Figure 3): (1) the first group corresponds to the five lowest frequency modes, which contribute the most to structural deviations; (2) for the second group, modes 6 to 15, deviations are significantly smaller; and (3) for the last group, the displacements are reduced by an half and vary smoothly from a mode to another, until mode 100. The atomic displacements of the different monomers for modes 1 and 10 are identical but the overall displacement for these two modes is larger than the sum of the intra-monomer displacements. This is due to a relative motion between the monomers. For the other modes, the relative displacements between the monomers may occur but their contribution to the overall motion is not additive. Examination of the motions associated with the different modes reveals two types of movements. The first type (modes 1, 10, 15, 18) is symmetrical and corresponds to an overall iris-like movement. In the second type, the different monomers behave differently, producing an asymmetrical movement. In this case, bending (modes 2, 3) or translating motions (modes 6, 7) involving generally two monomers are observed. These particular behaviors reflect the symmetry of the system. All the modes involved in symmetrical movements are non-degenerate modes, whereas the others are degenerate ones.

The relative displacements of the C^α corresponding to the lowest frequency mode of the closed structure are depicted in Figure 4(a). The relative displacements associated with each monomer are perfectly identical and, for the sake of clarity, only displacements associated with one monomer are represented. The Figure shows a zone which exhibits significant relative displacements, with a high value centered on residue 99. The zones involving the N_{term} part, the periplasmic loop, and the first part of the TM2 helix exhibit less significant displacements. Finally, the more hidden region constituting the pore (Val23–Lys31), is deprived of fluctuations. Along the TM2 helix, a small region located in the vicinity of residue 90 presents very small fluctuations. These zones with no fluctuation behave like pivots around which the other regions are moving. Similar profiles are observed for modes 1, 10 and 15 (see inset to Figure 4(a)): the peaks are located in the same region, the relative displacements of the periplasmic loop are identical. The major differences correspond to slight shifts of the pivot regions but they remain located in the first helical turns of TM1 and in the last helical turn of TM2. For the second group of low-frequency modes (modes 6 to 15), we observe the same overall characteristics although the five monomers do not behave identically. Common characteristics associated to the low frequency

modes may be summarized by: (1) the residues which constitute the first part of the TM1 helix and forms the most inner part of the pore channel do not move; (2) the last turn of the TM2 helix (Ala89–Phe93) delimits two regions with very different relative displacements and do not move too; and (3) from residue 99 to the end of the sequence, the highest atomic displacements are observed. It has to be noted that the two pivot regions (around Val23 and Phe93, respectively) are located roughly in the same plane, perpendicular to the channel axis.

Direction of movements

In order to describe the relative movements of residues in each normal mode, residue–residue distance variation maps are displayed. Figure 5(a) shows such a map, for the slowest mode. A high similarity of the different squares corresponding to the different monomers is clearly observed along the diagonal. This illustrates the symmetrical behavior of the system. The relative motions of the amino acid residues are clearly localized in the terminus regions of the monomers.

Figure 5(b) focuses on intra-monomeric distance variations. Four regions (numbered I to IV) represented by large white areas around the matrix diagonal may be distinguished. Each of them represents a rigid body region. They are delimited, respectively, by residues Arg13 to Ala20 (zone I: N_{term} part), residues Val23 to Asp84 (zone II: the TM1 helix, the periplasmic loop and the first part of the TM2 helix), residues Ala89 to Leu98 (zone III: the end of the TM2 helix) and residues Arg104 to the end of the sequence (zone IV: the C_{term} part) (see Figure 5(b)). Zones II and IV behave dynamically as a unit, as well as the first rigid segment (Arg13 to Ala20) and the region of the periplasmic loop which dips into the pore (the white blocks in the matrix). All zones considered as rigid body segments are connected by residues which can be interpreted as pivots in the structure. In addition to internal monomer motions, relative motions involving regions of different monomers are also observed (Figure 5(a)). For instance, between two close monomers (C and D as for example), zones II of the two monomers behave dynamically as a unit. The C_{term} part of monomer D moves away from monomer C. Similarly, the zone I of monomer C moves away from zone II of monomer D. On the opposite, the zone I of monomer D gets closer to the periplasmic loop and the first part of TM2 of monomer C. Despite their localized nature, these different movements result in a collective motion describing an iris-like movement. This results in twisting motions around the channel axis of the upper part of the molecule concomitant to a twisting of the lower part in the opposite direction (Figure 5(c)). The amplitude motions of the region towards the cytoplasmic medium are much larger than in the upper part of the structure. This type of motion tends to slightly incline or kink the

helices, as illustrated, in Figure 6, by the kink of the TM2 helix for the lowest frequency mode. The gray structure refers to an intermediate built along the normal mode direction. For the other slow modes which are degenerate, the relative motions of each monomer in the structure are not symmetric. Even if the movements are specific for each monomer, the previous segmentation in rigid bodies is also observed and, residues Ala20-Val23, Ile96-Leu98, act as hinges too. Other key regions are suggested, defined by residues Gly30-Lys31 and around residue Ile49. For the modes of higher frequency, the global motion is more complex, the segmentation in rigid body becomes fuzzier and the movements of the loops are much more disparate.

Motions of the *M. tuberculosis* MscL

To evaluate the influence of the sequence on the global motion of the system, we performed a normal mode analysis on the crystal structure of MscL of *M. tuberculosis*. It is clear that at the level of simplification used, we are not able to directly describe the influence of the amino acid type. Nevertheless, differences exist in the two structures. Indeed, in order to accommodate the presence of different side-chains, the C^α are slightly displaced in the two structures. Thus, differences observed constitute an indirect evaluation of the influence of the sequence.

To be in agreement with previous calculations, the C_{term} part of each monomer was deleted, from residues Thr109 to Arg118.[13] The crystallographic structure of the channel shows the same values regarding the range of vibrational frequencies as with our *E. coli* model.

As previously, the five first lowest modes are the most collective (data not shown). Beyond mode 16, rmsd values vary smoothly from one mode to another, while the movements are more localized.

Except for the lowest frequency mode and modes 8 and 11, where the rmsd values associated with each of the five monomers are identical, the movements of the monomers are different.

The relative displacements of the C^α for the lowest frequency mode (Figure 4(b)) is strongly similar to the one of our model. The same key zones are observed for both structures: located around residues Leu17-Val31 and Phe84 in *M. tuberculosis*. The rigid bodies regions are in the same structural parts, with a pronounced flexibility near Phe84, the Ccap of the TM2 helix for *M. tuberculosis*. As for our model, the prevalent movement for the MscL of *M. tuberculosis* is in corolla.

Motion of the Sukharev & Guy's modeled MscL

In order to compare our *E. coli* model with that of SG model and to study the influence of the different parts of the structure especially the S1 helices, two SG structures were explored, a first one, noted SG(+S1), including all the parts except

the C_{term} helix, a second one, noted SG(−S1), excluding both S1 helices and the C_{term}. The normal modes calculations were performed for both structures. The two sets of frequencies are slightly lower compared to those observed in our model. Whatever the presence or not of the S1 helices, the rmsd values spectrum associated with the different lowest frequency modes is quite similar to that observed previously (data not shown). The relative displacements corresponding to the slowest non-degenerate normal mode for the two structures are displayed in Figure 4(c). The profiles with and without S1 are very similar in the common region (residues 13–109) with two zones deprived of fluctuations, the first one in the Ncap of TM1 (around residue 21) and the second one in the Ccap of TM2 (around residue 101). The absence of S1 yields only a small shift in the location of the hinge regions (no more than 3–4 residues). Compared to the corresponding profile for our model, two main differences are observed: (i) in the SG(−S1), the hinge region in the TM2 is located two helical turns further than in our model, that may be partly explained by the lengthening of the helices, (ii) in the periplasmic region, in a zone including residues 61–70, the behavior is different. This may be due to the different relative positioning of the middle of the periplasmic loop in the SG's model and our model. As far as the relative movements of the residues are concerned, the presence or the absence of the S1 helical bundle does not modify the global motion. The motion associated with the SG's model (see Figure 7(a) and (b)), for the lowest mode, is very close to the one previously observed (see Figure 5(c)) with our model. We observe the same opposite twist of the channel around the helical axis, on each side of the plane defined by the two pivot regions, the lower part (towards the cytoplasm) being much more mobile than the upper part (towards the periplasm). For the other non-degenerate modes, the Ccap of the TM2 helices is animated by the largest motion. Other non-degenerate modes are animated by singular motions, not yet observed. For modes 11 (Figure 7(c)) and 31 (Figure 7(d)), the twisting motion is weaker than in the lowest mode (Figure 7(b)), and is accompanied by an expansion/compression movement of the ensemble of the structure along the pentamer helical axis. However, whatever the presence or not of the S1 helices, the structure remains closed.

Sukharev & Guy's model of gating mechanism

In order to identify the contribution of each normal mode to the transition pathway between the open and closed conformations of the SG's model, the projections of each normal mode onto the vector describing the difference between the open and closed forms of MscL were calculated (see Materials and Methods). As shown in previous studies,[23,24] the open structure generally seems to be more informative than the closed structure for
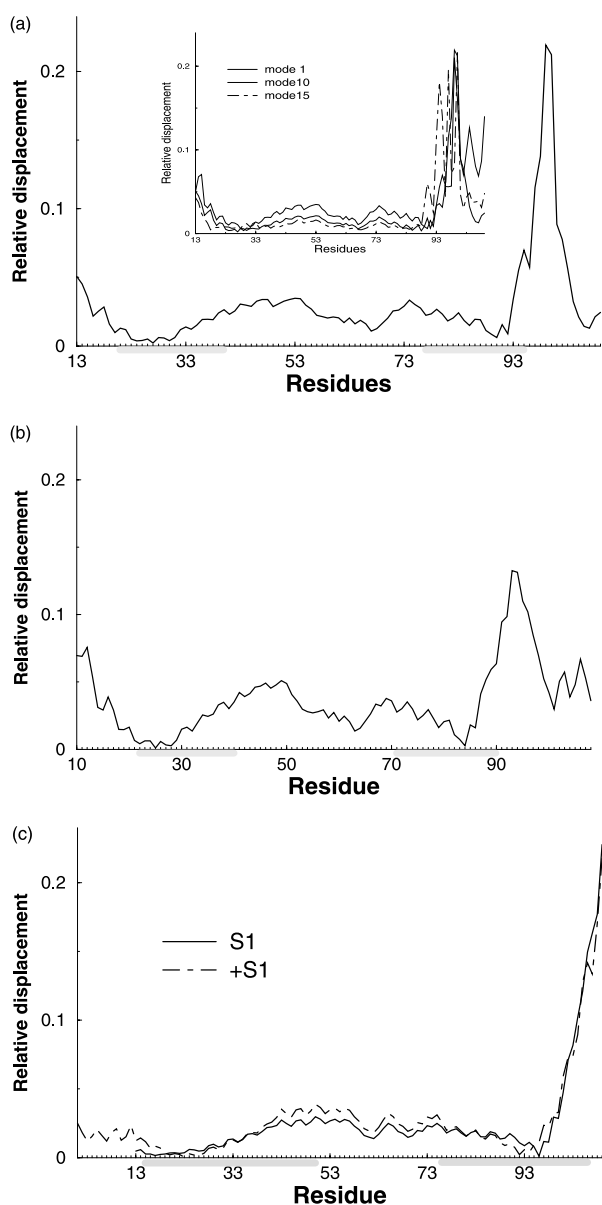
Figure 4. Relative atomic displacements of the C$^\alpha$ of one monomer as a function of residue number, for the lowest non-degenerate normal mode of (a) our *E. coli* MscL, the inset corresponds to the non-degenerate low frequency modes (modes 1, 10 and 15), (b) the *M. tuberculosis* MscL, (c) the *E. coli* SG model. Plain line: the SG's model (−S1) without the S1 helix (mode 3), broken line: the SG's model (+S1) with the S1 helix (mode 1). The transmembrane helical segments (TM1 and TM2) are annotated in gray.

such transitions. More specifically, the transition is usually better described with a small subset of low-frequency modes, when modes obtained starting from then open structure than from the closed one. We performed the computations for both cases, in order to estimate the influence of such a choice in the present case. Three transitions were considered: (i) from the closed form to the last fully open form (C $\rightleftharpoons$ O); (ii) from the closed form to the first open state, i.e. the closed-expanded

structure (C $\rightleftharpoons$ CE); and (iii) from the closed expanded form to the fully open state (CE $\rightleftharpoons$ O) and *vice versa*. All the structures analyzed in this part have the S1 helices, because of their postulated primordial role in the gating process;[42] the cytoplasmic part was deleted as previously. The three structures, closed (C), closed-expanded (CE) and open (O), correspond to the structural models 1, 9 and 12, respectively, proposed in the work of Sukharev and co-workers.[42] Figure 8 shows the cumulative sum of the squares of the overlaps between the set of modes and the vector difference considered. Note that if all modes were taken into account, in the three cases, the last value of the sum would be equal to 1.0, since a perfect description of any kind of motion is expected with a complete coordinate set (here, the so-called normal coordinates).

From the closed state to the fully open state and vice versa

Figure 8(a) shows that 76% of the C $\rightarrow$ O conformation change can be described with the 100 first normal modes. Three symmetrical, non-degenerate modes (11, 31 and 64) participate to 65% of the global motion. Similarly, the 100 first modes participate to 75% of the O $\rightarrow$ C transition but in that case, five non-degenerate modes (5, 17, 20, 29, 37) are needed to recover 65% of the conformational change. Each of these modes contributes 10–30% of the global conformational change. Large overlap values (>0.4) are observed for modes 11 and 31 in the C $\rightarrow$ O transition and for mode 29 in the O $\rightarrow$ C transition. In the C $\rightarrow$ O transition, the motions involved by the two modes are described in details in the previous section. Indeed, they correspond to a stretching/compression movement along the pentamer helical axis coupled with a weak twisting motion around the helical axis (Figure 5). For the O $\rightarrow$ C transition, the motion associated to mode 29 is completely different: it involves an expansion/compression movement of the ensemble of the structure in the plane perpendicular to the channel axis. In both transitions, we further calculated the collectivity values and the C$^\alpha$ displacements between the two structures. The global motions are highly collective, with collectivity values greater than 0.6. The profile of the relative atomic displacements computed for mode 11 (C $\rightarrow$ O transition) is globally similar to that observed for the C$^\alpha$ displacements between the two structures. Conversely, the profile for mode 29 in the O $\rightarrow$ C transition is different to the calculated C$^\alpha$ displacements between the two structures. The main differences are located in the connection between S1 and TM1 and at the end of TM1.

From the closed state to the closed-expanded state and vice versa

The cumulative sum of the squares of the overlaps between the set of modes and the vector

difference considered shows that 74% of the C → CE conformational change an 82% of the CE → C of the conformational change may be described with the 100 first normal modes (see Figure 8(b)). In the C → CE transition and *vice versa*, only three modes (8, 11, 29 and 3, 27, 32) contribute, respectively, to 50% and 60% of the displacement. In the C → CE transition, mode 11 is the only mode with an overlap value greater than 0.3. In the reverse transition, two modes are significant, mode 27 corresponding to an opening of the channel on the periplasmic side, like a corolla, and mode 32, which is quite similar to the motion of mode 29 described in the previous section.

From the open state to the closed-expanded state and *vice versa*

In that case, only 63% of the CE → O and O → CE conformational changes are described with the 100 first normal modes; ten and seven modes are necessary for describing 50% of both transitions, CE → O and O → CE, respectively (see Figure 8(c)). The motion associated with these modes are expansion/compression movements, coupled with twisting motion of the Ccap of TM2, or less collective motion located essentially in the periplasmic loop and the Ccap of TM1.

Thus, unexpectedly, better results are not obtained with the most open structure analyzed. In the present case, the closed model seems to be the most informative one, in order to extract some knowledge about the ability of protein regions to deform during the functional motion.

Discussion

Different structures of MscL have been examined, using the normal mode analysis. The main goal of our study was to try to understand how this channel can gate, and identify which parts of the protein could be involved in the conformational change. The main features that emerge from the present study concern different and essential points: the global dynamics of the system, in close relationship with the topology of the structure and its flexibility, the role of the N_{term} and C_{term} parts, and the opening channel mechanism. These points are discussed below.

Global dynamics

The global motions observed in the three examined closed structures exhibit some similarities. For the non-degenerate lowest frequency modes, the three models have a common behavior: an iris-like movement that consists in a twist and tilt of the helices. The channel can be divided into two parts relative to a pivot plane perpendicular to the channel axis, passing through the channel gate (Va123). The relative position of the pivot plane, on both sides of which the torsion move-

ment occurs, can be slightly shifted depending of the structural model examined. The upper part (periplasmic side) is animated by weaker amplitude motions than the lower part, as also observed in the course of MD simulations.[43] Indeed, the authors noticed that the extracellular end of the helices did not move outward as much as the intracellular end.

Some modes, not among the lowest-frequency ones through, are found to be model-dependent. As an example, the SG's model exhibits a particular motion, which is an expansion/compression of the structure along the protein principal axis. In this model, the periplasmic loop adopts a distinct position compared to the two other structures, and the transmembrane helices are longer. This particular motion could be a consequence of these structural differences. Such results suggest that the global movements may also be closely related to the fine geometry of the structure (length of the secondary structure elements, relative position of these elements within the complete structure, positioning of the loops).

May α-helices always be considered as rigid bodies?

For all normal modes analyzed, a given monomer does not behave as a rigid body, but rather as an assembly of rigid bodies. These rigid bodies do not inevitably include complete secondary structure elements, but rather parts of them. In the MscL, the hinge points are mainly localized within secondary structural elements: the different parts of an helix can be animated by motions of different amplitude and direction. This contrasts with normal mode results for enzymes or bigger molecular complexes where the hinge points are usually found in unstructured regions linking the different domains.

Irrespective of the structure examined, the lowest frequency modes show that in the TM2 helices two regions can be distinguished, which are delimited by the pivot plane. The Ncap of the TM2 is animated by motions of smaller amplitude, as compared to the Ccap. This results in a kink in the helix. Such a kink was also found in a recent targeted MD study[14] between the two modeled states (C and O) proposed by Sukharev and co-workers. While in both structures, the TM2 helix is perfectly straight, in the structural intermediates along the conformational pathway, they are kinked.

Interestingly, this feature is sometimes encountered in membrane proteins,[44,45] and is supposed to play a role in their structural or functional properties. Two types of kink in transmembrane α-helices are observed. The most frequent one is proline-based. However, hinges in α-helices are not found only near proline residues.[45] An example of non-proline distortion has been observed at 1.5 Å of resolution for the bacteriorhodopsin. This distortion is due to the presence of a π-bulge
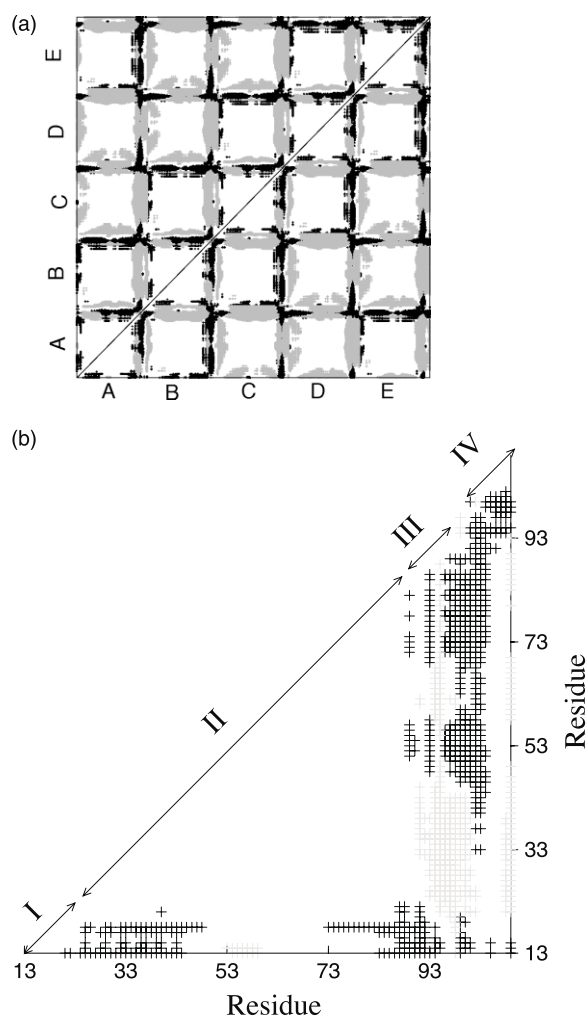
Figure 5 (*legend opposite*)

functionally important. The same behavior is also observed in the open form of a potassium channel (the MthK channel) for the inner transmembrane helices[46] with a glycine residue. In those cases, the distortions occur near residues close to the "proteins active sites", and these residues are not proline. In our case, the kink appears in the vicinity of a phenylalanine residue (Phe90). Interestingly, the residues around Phe90 belong to a highly conserved motif between the MscL from different species. In the crystal form of *M. tuberculosis*, a small deformation is observed in this zone: the value of the twist helical angle in this region is larger (105°), as compared to an ideal value (100°). Surprisingly, despite of the absence of such a distortion in the SG's model (the twist angle value is 101°), the movement described by the normal mode analysis involves a deformation of the same zone. Thus, the motion observed is not due to a weakness of the TM2 helix in the crystal structure.

In our work, the membrane environment is omitted. The presence of lipids may well influence the amplitude and the direction of the motions of the helices. However, a MD study performed in an explicit medium (the channel embedded in a hydrated bilayer) exhibits the same property (work in progress), namely the presence of a hinge in the Phe90 region. All these results suggest that this particular motion is an intrinsic structural property of the TM2 helix. The particular motif conserved in the TM2 may be related to a dynamical specificity of this part, and our results suggest that this region could play a role in the gating mechanism of the channel.

Recent experiments led by Perozo and collaborators[47] provide interesting results concerning the behavior of the TM2 helix during the gating mechanism. They tested numerous mutants distributed along the TM2 helix and the accessibility. Moreover, using various lipids environments, they have been able to trap different states of the channel. Two states were characterized; the first one corresponds roughly to the first intermediate open state (CE) and the second one to the fully open state (O). From the closed to the intermediate
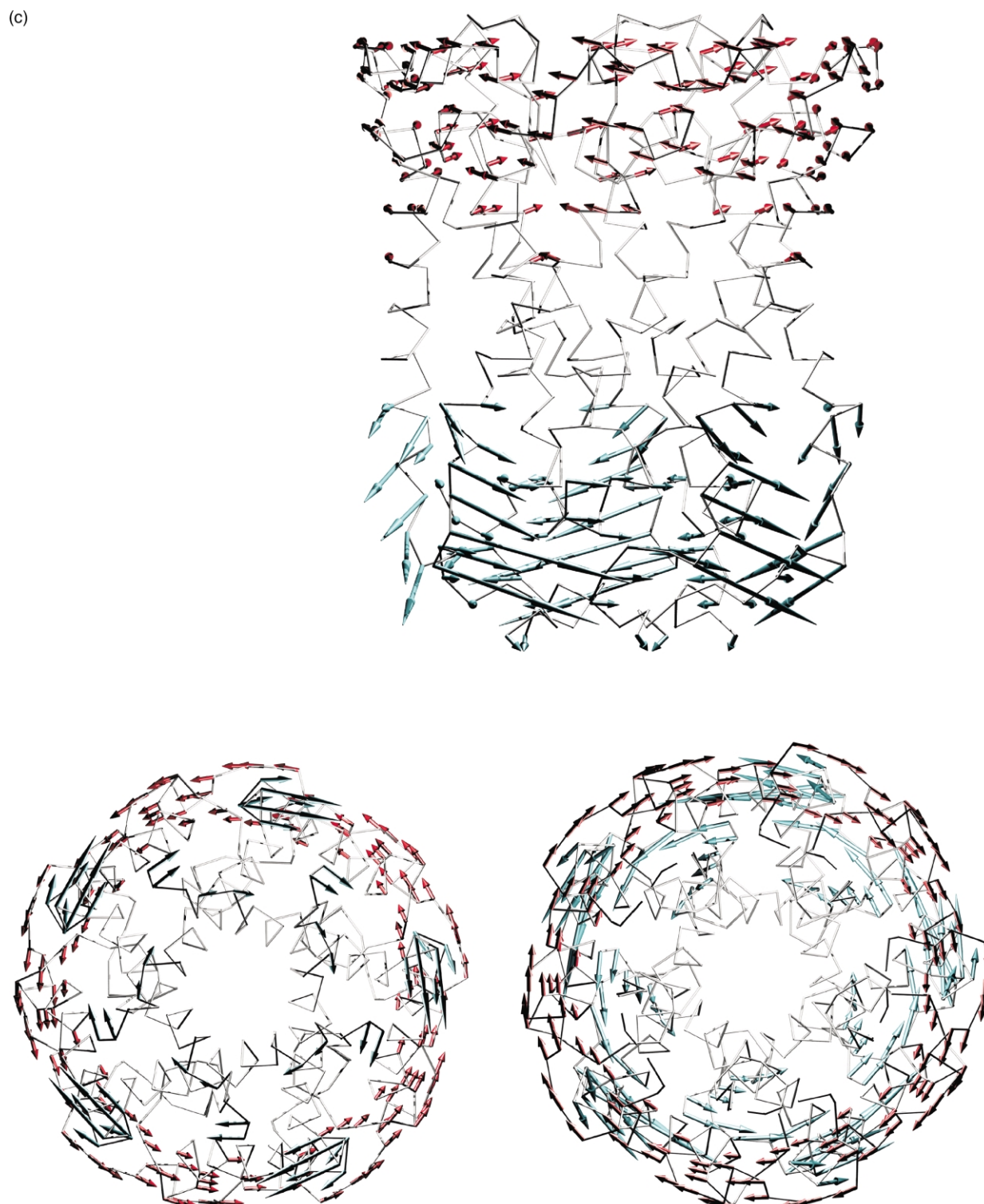
(c)

Figure 5. (a) Map of variation of the C^α distance between pairs of residues during the motion in the slowest mode for the overall structure (monomers A to E). The relative degrees of variations of distances between pairs of atoms are indicated in dark (increasing), in gray (decreasing) and in white (non-significant), regarding to the average distance variation. (b) Map of intra-monomeric distance variation (a zoom of a diagonal square from (a)). (c) Vector representation of the amplitude and direction of the C^α displacements along the mode (side, top (left) and bottom (right) views).

open state, no change along the TM2 helix is observed for the parameters examined (lipid accessibility profile, water accessibility). In contrast, in the fully open state, a notable change in the accessibility to lipids profile appears in the Ccap of the TM2 helices: the 83–96 residues are more accessible to lipids in the open state than in the closed state.

Figure 6. Top and side views of the motion of the transmembrane helices, in the slowest mode. (a) A part of N_{term} and TM1, from residues 17–40. (b) TM2, from residues 77–96. The initial structure is drawn in black. An intermediate structure displaced along the normal mode is drawn in gray.

Role of the end parts

The motions associated with the complete structure or with the structure without the C_{term} part are very similar. The deletion of the C_{term} part of the structure does not have significant effect on the dynamical behavior. Moreover, a series of molecular simulations[43] in which a force mimicking the stretch of the membrane is applied, shows a significant disorder of the C_{term} tails. All these results suggest that this part of the protein does not play a major role in the gating mechanism, in agreement with experimental data.

In our analysis performed on the SG's model, the presence or the absence of the S1 helices do not modify the global motion of the system (twisting motion). No low frequency normal modes among the 100 first modes calculated is associated with a particular behavior of this region in the protein. These results are in agreement with those observed by Ma and co-workers[14] who did not observe any coupling between the S1 helical bundle and the TM2 helices. In addition, they suggest that the movement of the S1 helices, suggested by the Sukharev's pathway, would simply be a consequence of the pulling movement of the Ncap of the TM1 outward the membrane plane.

The N_{term} part is mobile in the different normal mode analyzed (see Results). This may be linked to a study dealing with the contributions of the different extramembranous domains of the MscL to its response to membrane tension. This study showed that the N_{term} and C_{term} parts of the channel essentially regulate the mechanosensitivity of the channel without altering its functionality.[36]

Key and functional region

In the case of MscL, the region located around residue Val23 has been shown to be functionally important. Numerous mutagenesis studies[47–49] in this region highlight the structural and functional importance of residues around the Val23, i.e. Ala20, Val23 and Gly26. Dramatic lack of function occurs if one residue in this region is mutated. Moreover, these residues are highly conserved in other bacteria, and illustrate a "signature in the sequence".[9] This region is considered as a key region, both for the gating process and the stability of the three-dimensional structure. Our normal mode analysis shows that this region which occludes the pore is always deprived of fluctuations. Interestingly, in many cases,[21,22,50,51] it has been observed that the regions deprived of fluctuation in the normal mode analysis refer to active
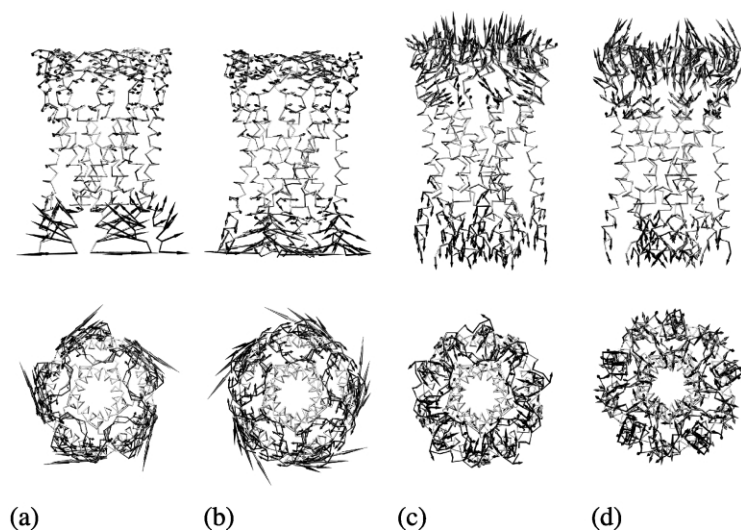

Figure 7. Vector representation (side and top views) of the C^{α} motions associated with (a) the SG(−S1) model for the lowest non-degenerate mode (mode 3); (b) the SG(+S1) model for the lowest non-degenerate mode (mode 1); (c) the SG(+S1) model for the mode 11; (d) the SG(+S1) model for the mode 31. For clarity, only the major displacements are represented.
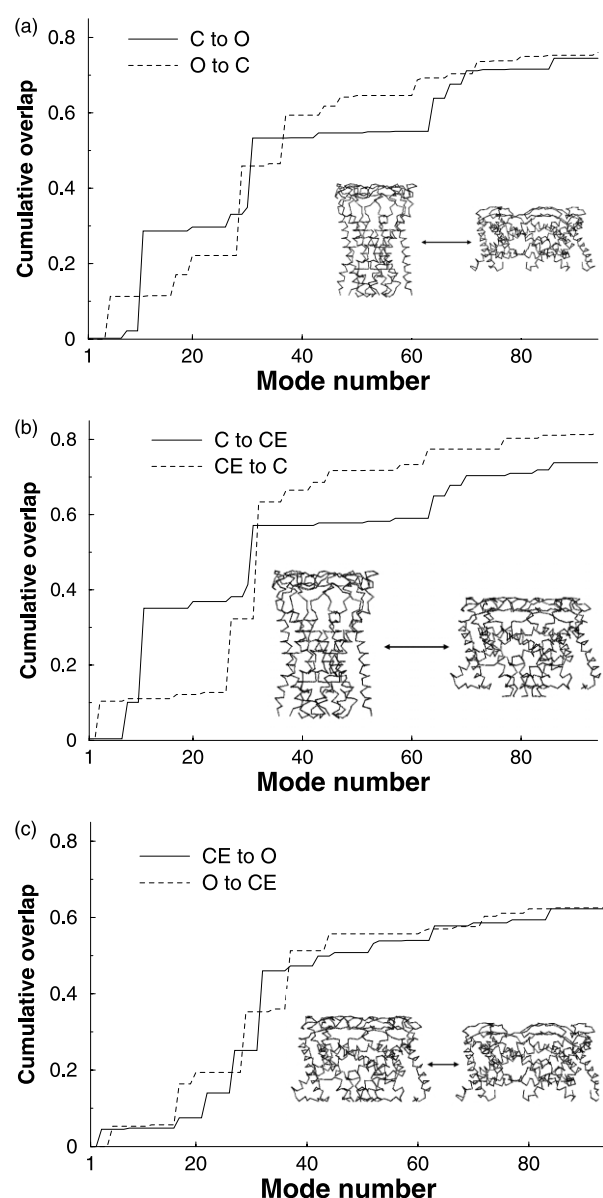
Figure 8. Cumulative square of the overlap between the mode and the vector difference as a function of the normal mode number. The vector difference describes the transition between (a) the closed and the open form (C ⇌ O), (b) the closed and the first open, closed-expanded form (C ⇌ CE), (c) the first open and the fully open form (CE ⇌ O).

sites or hinge regions of the protein. The recent experiments on the MscL also seem to indicate that the region which occludes the pore remains firmly immobilized during the transition from the closed to the closed-expanded state.[47]

The gating process

The normal mode analysis predicts a tilt of the transmembrane helices, and more particularly, a tilt of the TM1 with respect to its principal axis. This is in agreement with previous hypothesis about the gating mechanism. The dynamical prop-

erty associated with the Ccap of the TM2 is remarkable. This region may be the first part of the protein which reacts to the strain experienced by the membrane. Nevertheless, the global motion of the protein can be well described by a "twist to open" mechanism.

In the normal mode analysis performed, no individual normal mode leads to the opening of the channel. Even if twisting motion along the protein axis, tilting motion of transmembrane helices appear, the pore remains occluded. In fact, the normal mode analysis gives information about the preferential direction of movement along "easy" directions. Moreover, the fact the channel remains closed would be comprehensible if we take into account the mechanosensitivity of the channel. An external energy, brought by the membrane strain in the cell, must be given to the channel in order to open. The normal mode method does not take into account this constraint, but provides new insights on the zones of the protein which are able to participate to the gating mechanism. In addition, the gating mechanism involves an obviously large structural change of the channel, with for instance a rmsd value of 17 Å between the two SG's models (open and closed). Therefore, at variance with most cases previously considered when low-frequency normal modes were compared to conformational transitions, such a conformational change may well occur in several, largely independent steps of different amplitudes and directions. Interestingly, the cumulative overlaps value describing the CE → C pathway was found to be the highest, as compared to the other ones. The CE → C transition is thus the best described pathway by the normal mode analysis, among all the proposed ones. From the normal mode analysis point of view, this structural intermediate is thus the best candidate, the one that is the more likely to lie close to the correct conformational pathway. On the opposite, the cumulative overlaps value is lower for the CE → O transition. This suggests that the fully open model structure considered in the present study may not be a good enough description of the correct open state.

Lastly, we have to address additional questions concerning, in particular, the role of the membrane. Many experiments indicate that the channel is gated by the tension transmitted through the bilayer alone. Most of these experiments as well as theoretical works focus on the structural changes undergone by the channel itself but little information exists about the way, at a molecular level, the membrane accommodates the channel deformation upon gating. In this field, the recent work of Perozo and co-workers[52] tries to answer the question still open about the respective role of the hydrophobic mismatch and/or the intrinsic curvature of the membrane. Studying different lengths of lipid acyl chains and introducing lysophospholipids to control the curvature, this work provides very interesting information concerning the behavior of the membrane itself. On the basis of EPR

approaches and patch–clamp experiments, their results show that the reduction of the membrane thickness, even not sufficient to open the channel, lowers the activation threshold and simultaneously the pipette pressure required to fully open the channel. Reversely, asymmetrical lysophosphatidylcholine–phosphatidylcholine mixtures appear to trap MscL in its fully open state. Our results show that the structural changes related to the lower modes of the normal modes analysis, that tilt the helices and thus reduce the long axis component of the channel, are quite compatible with a reduction of the membrane thickness. Nevertheless, to reach a detailed description of the membrane behavior, it would be necessary to introduce explicitly a large number of lipid molecules to represent the physical properties of a membrane. These calculations are computationally expensive, or need appropriate parametrization if approximations are introduced. Work in this field is currently in progress.

In conclusion, the use of the normal mode theory on a particular system (mechanosensitive channel) without adding external mechanical stretch makes possible the exploration of the channel global motions. The non-degenerate modes seem to be implied in the conformational rearrangements occurring during the gating process. This analysis provides more precise insights, at the residue level, of the regions of the protein which may participate to the deformation process. Common features have been found, concerning the relation between the global architecture and the associated motions. The normal modes analysis predicts that a large amplitude motion of the Ccap of TM2 helices may occur, coupled with the tilt of the TM helices. An interesting point is the global movement associated with all the structures analyzed: the gating process combines a tilting motion of the helices and a twisting motion of the structure along the channel axis. The "twist to open" mechanism may be a general process for this type of channel, i.e. channels where the pore architecture is like a funnel. Other channels, as nucleotide-gated channel[53] or potassium channels,[54] have a transmembrane architecture close to the structure of the MscL. They also undergo rotational movements, together with a kink of secondary structural elements associated with the channel opening, as suggested by the present study for the MscL case.

Materials and Methods

Before carrying out the normal mode analysis, we need a stable structure of the molecule of interest (the MscL of *E. coli*), and thus, we must generate a three-dimensional model of this channel.

Modeling

The structure of *E. coli* MscL has been constructed using homology modeling method. This approach requires a template structure solved with an atomic resolution by experimental methods (X-ray, NMR). Moreover, the template sequence need to share significant sequence similarity with the target sequence.

The template structure

The only structure available filling these criteria, is the MscL structure from *M. tuberculosis* solved, using X-ray crystallography, by Rees and co-workers[7] (PDB code: 1msl, 3.5 Å resolution). Only residues 10–118 were solved, the missing residues being disordered. The X-ray structure displays the MscL in its closed state. In the crystal, the MscL is a homo-pentamer. Monomers are rigorously identical. The transmembrane part is composed of two transmembrane α-helices, called TM1 and TM2. The periplasmic part is composed of a large loop connecting the TM1 and TM2 helices. The cytoplasmic part is folded into an α-helix in the crystal (see Figure 2(a)). A recent study[13] points out the influence of the pH on this part: the α-helical secondary structure is preserved at low pH ($\sim$3.6, crystal pH) while it is slightly destructured at physiological pH.

The sequence alignment

According to the internal symmetry observed for the complex in the crystal structure (5-fold symmetry axis), in a first step, the structure of one monomer was built. The sequence alignment was performed using different alignments tools (BLAST, PSI-BLAST[55], CLUSTAL W[56]). Depending on parameters (substitution matrix, gap penalties) or algorithms, alternative alignments with very close sequence identity score were obtained. The alignments differ in the position of the insertion/deletion (indel), essentially located in the periplasmic region connecting TM1 and TM2 and in the cytoplasmic region. These alignments were further submitted to the Modeller package†[57] to construct different models. For each alignment, 100 models were generated and optimized. The models differ essentially in the conformation of the indel regions. The 100 models were ranked according to the Modeller objective function and only the first ten models were selected for each alignment.

The complete structure

The second step consists in reconstructing the overall three-dimensional structure by rigid superimposition of the generated model on each monomer of the crystal structure. Due to the strategy used, the three-dimensional compatibility of the monomers in the pentamer cannot be assumed, mainly because of the poor sequence identity in loops. The superimposition was performed on the first transmembrane helix—TM1.[7] Among all the models generated, a complete structure was finally selected according to three criteria: the rmsd value between the model and the X-ray structure for aligned regions, no three-dimensional overlapping region, conserved residues in similar position as in *M. tuberculosis* MscL structure, periplasmic loops positioned in a way to lie on the top of the channel entrance. The side-chains of the complete model were repositioned in optimized conformations as predicted with the Confmat method.[58] Energy minimization (steepest descent algorithm, $\epsilon = 4$)

---

† http://salilab.org/modeller

of the final model was carried out, using the Gromos 96 force field of the Gromacs2.0 package.[59]

Normal mode calculation

The normal mode theory is based on the harmonic approximation of the potential energy function, around a minimum energy conformation. This approximation allows an analytical solution to the equations of motion by diagonalizing the Hessian matrix (mass-weighted second derivatives of the potential energy matrix). The eigenvectors of this matrix are the normal modes, and the eigenvalues are the squares of the associated frequencies. The protein motion can be represented as a superposition of normal modes, fluctuating around a minimum energy conformation. The displacements ($r_i(t)$) of atomic coordinate i can be expressed as:

$$r_i(t) = \frac{1}{\sqrt{m_i}} \sum_k^{3N} C_k a_{ik} \cos(\omega_k t + \phi_k)$$

where m_i is the mass of the atom i, C_k and ϕ_k, are the amplitude and phase of mode k, $v_k = \omega_k/2\pi$ is the vibrational frequency, and a_{ik} is the ith coordinate of the eigenvector k, i.e. the corresponding coordinate displacement. C_k and ϕ_k depend on initial conditions. For proteins, the normal modes responsible for most of the amplitude of the atomic displacements exhibit frequencies below $30–100$ cm^{-1} (low frequencies). Because of the model building strategy used and the large number of internal coordinates considered in our system, we calculate the vibrational normal modes using a simplified potential.

Simplified potential

As described by Tirion,[30] the "classical" semi-empirical potential energy function used in the all-atom force field is replaced by a single parameter potential (a Hookean potential):

$$E_p = \sum_{d_{ij}^0 < R_c} c(d_{ij} - d_{ij}^0)^2$$

where d_{ij} is the distance between two atoms i and j, d_{ij}^0 is the distance between these atoms in the three-dimensional structure, c is a constant assumed to be the same for all interacting pairs which refers to the spring constant of the potential, and finally R_c is an arbitrary cutoff parameter, beyond which interactions are not taken into account. By using this Hookean potential it is assumed that the total energy potential function of the reference three-dimensional structure is minimum ($E = 0$). The stage of energy minimization is not necessary anymore, the starting structure being used as reference structure. In the present study, only C$^\alpha$ atoms have been taken into account. Previous studies demonstrated that such a model is sufficient for studying backbone protein motions of very large systems, considering only the low-frequency normal modes. The higher frequency normal modes refer to rapid oscillations of side-chains and small groups of atoms, that requires an accurate analysis at the atomic level and could not be modeled and further analyzed using such a simplified potential. Different values for the cutoff value were tested ($R_c = 8–13$ Å), regarding to the size of our system. For each cutoff value, the vibrational normal modes were calculated on the structure, and the relative displacements associated were analyzed. From a value of 9 Å,

we observe a convergence of displacements. Moreover, if we have a look at the contact map of the structure, we can see that below $9–10$ Å, there are significant contacts between the monomers which are not taken into account. The best value kept for the cutoff is thus 10 Å.

It has to be noted that until now, the normal mode analysis has been mainly applied on globular soluble proteins. In order to assess the appropriateness of such an approximation for membrane proteins, we have tested this approach on a relatively bare helix bundle, the rhodopsin (PDB code: 1F88). Comparison of the C$^\alpha$ relative displacements computed with Tirion's model and a classical force field (charmm22 all-atom force field), performed for the energy-minimized structure, the classical normal modes being calculated with the DIMB method,[26] shows excellent agreement as well for the lowest frequency mode as well as for the two next ones (data shown in Supplementary data available on the internet).

Distance fluctuation

The map for distance fluctuations between residues i and j measures the relative movements between residues in the mode k.[60,61] In such maps, rigid and flexible blocks of amino acid residues can be identified, as well as their relative movements. The blocks of amino acid residues which behave as rigid bodies during the motion appear in white in the map, whereas flexible segments are filled with symbols (black or gray). A black symbol indicates that distance between two C$^\alpha$ atoms increases significantly, and a gray symbol that it decreases.

Overlap

The overlap I_k measures the degree of similarity between the direction of a chosen conformational change $\overrightarrow{\Delta r}$ and the one given by the normal mode considered. It has been calculated as described by Marques & Sanejouand:[21]

$$I_k = \frac{|\sum^{3N} a_{ik} \Delta r_i|}{|\sum^{3N} a_{ik}^2 \sum^{3N} \Delta r_i^2|^{1/2}}$$

A value of one for the overlap means that the direction given by the normal mode k considered is identical with the conformational change.

Collectivity

The degree of collectivity (κ_k) measures the collective protein motion in the mode k considered, i.e. the number of atoms which are significantly affected in the normal mode. The value of κ has been calculated as described by Bruschweiler:[62]

$$\kappa_k = \frac{1}{N} \exp\left(-\sum^N \alpha \Delta A_{ik}^2 \log \alpha \Delta A_{ik}^2\right)$$

where A_{ik} is the amplitude of the displacement of atom i in the mode k and α is a normalization factor chosen such that $\sum^N \alpha \Delta A_{ik}^2 = 1$. The conformational change is maximally collective for a value of one. On opposite, for local motion, where the conformational change involves only few atoms, κ is minimal ($\kappa = 1/N$).

Helical geometrical parameters

Helical regions were defined with the STRIDE program.[63] Helix parameters analyzed are the lengths of the transmembrane helices, the helix axis orientation τ relative to the protein axis (tilt angles), and interactions between helices. Interactions between helices were described by the distance of closest approach of the two helix axis, the distance between the helix barycenters, and the inter-axial angle Ω as described by Chothia.[64] The axis of the protein is defined as the 5-fold symmetry axis and can be compared to the bilayer normal.

Acknowledgements

References

1. Berrier, C., Garrigues, A., Richarme, G. & Ghazi, A. (2000). Elongation factor tu and dnak are transferred from the cytoplasm to the periplasm of *Escherichia coli* during osmotic downshock presumably *via* the mechanosensitive channel MscL. *J. Bacteriol.* **182**, 248–251.

2. Sukharev, S. I., Blount, P., Martinac, B. & Kung, C. (1997). Mechanosensitive channels of *Escherichia coli*: the MscL gene, protein, and activities. *Annu. Rev. Physiol.* **59**, 633–657.

3. Ajouz, B., Berrier, C., Garrigues, A., Besnard, M. & Ghazi, A. (1998). Release of thioredoxin *via* the mechanosensitive channel MscL during osmotic downshock of *Escherichia coli* cells. *J. Biol. Chem.* **273**, 26670–26674.

4. Blount, P., Sukharev, S. I., Moe, P. C., Marinac, B. & Kung, C. (1999). Mechanosensitive channels of bacteria. *Methods Enzymol.* **294**, 458–482.

5. Hamill, O. P. & Martinac, B. (2001). Molecular basis of mechanotransduction in living cells. *Physiol. Rev.* **81**, 685–740.

6. Cruickshank, C., Minchin, R., LeDain, A. & Martinac, B. (1997). Estimation of the pore size of the large-conductance mechanosensitive ion channel of *Escherichia coli*. *Biophys. J.* **73**, 1925–1931.

7. Chang, G., Spencer, R., Lee, A., Barclay, M. & Rees, D. C. (1998). Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive channel. *Science*, **282**, 2220–2226.

8. Batiza, A. F., Rayment, I. & Kung, C. (1999). Channel gate! Tension, leak and disclosure. *Structure*, **7**, 99–103.

9. Spencer, R. H., Chang, G. & Rees, D. C. (1999). Feeling the pressure: structural insights into a gated mechanosensitive channel. *Curr. Opin. Struct. Biol.* **9**, 448–454.

10. Yoshimura, K., Batiza, A. & Schroeder, M., Blount, P. & Kung, C. (1999). Hydrophilicity of a single residue within MscL correlates with increased channel mechanosensitivity. *Biophys. J.* **77**, 1960–1972.

11. Sukharev, S., Durell, S. R. & Guy, H. R. (2001). Structural models of the MscL gating mechanism. *Biophys. J.* **81**, 917–936.

12. Gullingsrud, J., Kosztin, D. & Schulten, K. (2001). Structural determinants of MscL gating studied by molecular dynamics simulations. *Biophys. J.* **80**, 2074–2081.

13. Elmore, D. E. & Dougherty, D. A. (2001). Molecular dynamics simulations of wild-type and mutant forms of the *Mycobacterium tuberculosis* MscL channel. *Biophys. J.* **81**, 1345–1359.

14. Kong, Y., Shen, Y., Warth, T. E. & Ma, J. (2002). Conformational pathways in the gating of *Escherichia coli* mechanosensitive channel. *Proc. Natl Acad. Sci. USA*, **99**, 5999–6004.

15. Colombo, G., Marrink, S. J. & Mark, A. E. (2003). Simulation of MscL gating in a bilayer under stress. *Biophys. J.* **84**, 2331–2337.

16. Go, N., Noguti, T. & Nishikawa, T. (1983). Dynamics of a small globular proteins in terms of low-frequency vibrational modes. *Proc. Natl Acad. Sci. USA*, **80**, 3696–3700.

17. Brooks, B. & Karplus, M. (1983). Harmonic dynamics of proteins: normal mode and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl Acad. Sci. USA*, **80**, 6571–6575.

18. Levitt, M., Sander, C. & Stern, P. S. (1985). Protein normal-mode dynamics: trypsin inhibitor, crambin, tibonuclease and lysozyme. *J. Mol. Biol.* **181**, 423–447.

19. Keskin, O., Jernigam, R. L. & Bahar, I. (2000). Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* **78**, 2093–2106.

20. Harrison, W. (1984). Variational calculation of the normal modes of a large macromolecules: methods and some initial results. *Biopolymers*, **23**, 2943–2949.

21. Marques, O. & Sanejouand, Y. H. (1995). Hinge-bending motion in citrate synthase arising from normal mode calculations. *Proteins: Struct. Funct. Genet.* **23**, 557–560.

22. Sanejouand, Y. H. (1996). Normal mode analysis suggests important flexibility between the two N-terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. *Protein Eng.* **9**, 671–677.

23. Delarue, M. & Sanejouand, Y. H. (2002). Simplified normal mode analysis of conformational transitions in DNA-dependant polymerases: the elastic network model. *J. Mol. Biol.* **320**, 1011–1024.

24. Tama, F. & Sanejouand, Y. H. (2001). Conformational change of protein arising from normal mode calculations. *Proteins Eng.* **14**, 1–6.

25. Krebs, W., Alexandrov, V., Wilson, C., Echols, N., Yu, H. & Gerstein, M. (2002). Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins: Struct. Funct. Genet.* **48**, 682–695.

26. Perahia, D. & Mouawad, L. (1995). Computation of low-frequency normal modes in macromolecules: improvements of the method of diagonalization in a mixed basis and applications to hemoglobin. *Comput. Chem.* **19**, 241–246.

27. Durand, P., Trinquier, G. & Sanejouand, Y. H. (1994). New approach for determining low-frequency normal modes of macromolecules. *Biopolymers*, **34**, 759–771.

28. Tama, F., Gadea, F. X., Marques, O. I. & Sanejouand, Y. H. (2000). Building-block approach for determining low frequency normal modes of macromolecules. *Proteins: Struct. Funct. Genet.* **41**, 1–7.

29. Li, G. & Cui, Q. (2002). A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to ca(2 +)-atpase. *Biophys. J.* **83**, 2457–2474.

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

31. Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. *Proteins: Struct. Funct. Genet.* **33**, 417–429.

32. Hinsen, K. & Kneller, R. (1999). A simplified force field for decribing vibrational protein dynamics over the whole frequency range. *J. Comp. Chem.* **111**, 10766–10769.

33. Bahar, I., Atilgan, A. & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* **17**, 412–425.

34. Bowie, J. U. (1997). Helix packing in membrane proteins. *J. Mol. Biol.* **272**, 780–789.

35. Blount, P., Sukharev, S. I., Schroeder, M., Nagle, S. & Kung, C. (1996). Single residue substitutions that change the gating properties of a mechanosensitive channel in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **93**, 11652–11657.

36. Ajouz, B., Berrier, C., Besnard, M., Martinac, B. & Ghazi, A. (2000). Contributions of the different extra-membranous domains of the mechanosensitive ion channel Mscl to its response to membrane tension. *J. Biol. Chem.* **275**, 1015–1022.

37. Hargittai, I. & Hargattai, M. (1995). *Symmetry Through the Eyes of a Chemist*, Plenum Press, New York.

38. Labarre, J.F. (1978). La théorie des groupes. Presses universitaires de France.

39. Matsumoto, A. & Go, N. (1999). Dynamic properties of double-stranded DNA by normal mode analysis. *J. Chem. Phys.* **110**, 11070–11075.

40. van Vlijmen, H. W. T. & Karplus, M. (2001). Normal mode analysis of large systems with icosahedral symmetry: application to (Dialanine)60 in full and reduced basis set implementations. *J. Chem. Phys.* **115**, 691–698.

41. Tama, F. & Brooks, C. L., III (2002). The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus. *J. Mol. Biol.* **318**, 733–747.

42. Sukharev, S. I., Betanzos, M., Chiang, C. S. & Guy, H. R. (2001). The gating mechanism of the large mechanosensitive channel MscL. *Nature*, **409**, 720–724.

43. Bilston, L. E. & Mylvaganam, K. (2002). Molecular simulations of the large conductance mechanosensitive (MscL) channel under mechanical loading. *FEBS Letters*, **25743**, 1–6.

44. Barlow, D. J. & Thornton, J. M. (1988). Helix geometry in proteins. *J. Mol. Biol.* **201**, 601–619.

45. Sansom, M. S. P. & Weinstein, H. (2000). Hinges, swivels and switches: the role of proline in signalling *via* transmembrane α-helices. *Trends in Pharmacological Sciences*, **21**, 445–451.

46. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T. & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature*, **417**, 515–522.

47. Perozo, E., Cortes, D. M., Sompornpisut, P. & Martinac, B. (2002). Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature*, **418**, 942–948.

48. Ou, X., Blount, P., Hoffman, R. J. & Kung, C. (1998). One face of a transmembrane helix is crucial in mechanosensitive channel gating. *Proc. Natl Acad. Sci. USA*, **95**, 11471–11475.

49. Maurer, J. A., Elmore, D. E., Lester, H. A. & Dougherty, D. A. (2000). Comparing and contrasting *E. coli* and *M. tuberculosis* mechanosensitive channels (MscL). New gain of function mutations in the loop region. *J. Biol. Chem.* **275**, 22238–22244.

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

51. Bahar, I., Erman, B., Jernigan, R. L., Atilgan, A. & Covell, D. G. (1999). Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J. Mol. Biol.* **285**, 1023–1037.

52. Perozo, E., Kloda, A., Cortes, D. M. & Martinac, B. (2002). Physical underlying the transduction of bilayer deformation forces during mechanosensitive channel gaing. *Nature Struct. Biol.* **9**, 696–703.

53. Johnson, J. P. & Zagotta, W. N. (2001). Rotational movement during cycloc nucleotide-gated channel opening. *Nature*, **412**, 917–921.

54. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T. & MacKinnon, R. (2002). The open pore conformation of potassium channels. *Nature*, **417**, 523–526.

55. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, M. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.

56. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W.: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weigh matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.

57. Sali, A. & Blundell, T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.

58. Koehl, P. & Delarue, M. (1994). Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy. *J. Mol. Biol.* **239**, 249–275.

59. Berendsen, H., van der Spoel, D. & van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Commun.* **91**, 43–56.

60. Nishikawa, T. & Go, N. (1987). Normal modes of vibrations in bovine pancreatic trypsin inhibitor and its mechanical property. *Proteins: Struct. Funct. Genet.* **2**, 308–329.

61. Seno, Y. & Go, N. (1990). Deoxymyoglobin studied by the conformational normal mode analysis. *J. Mol. Biol.* **216**, 95–126.

62. Bruschweiler, R. (1995). Collective protein dynamics and nuclear spin relaxation. *J. Chem. Phys.* **102**, 3396–3403.

63. Frishman, D. & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Struct. Funct. Genet.* **23**, 566–579.

64. Chothia, C., Levitt, M. & Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol.* **145**, 215–250.

65. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD—visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.

Edited by D. Rees

SCIENCE @ DIRECT®

www.sciencedirect.com

Supplementary Material is available on Science Direct