

# Conserved motifs in mechanosensitive channels MscL and MscS

Daniel Balleza · Froylan Gómez-Lagunas

Received: 28 November 2008 / Revised: 27 March 2009 / Accepted: 8 April 2009 / Published online: 8 May 2009  
© European Biophysical Societies' Association 2009

**Abstract** Mechanosensitive (MS) channels play a major role in protecting bacterial cells against hypo-osmotic shock. To understand their function, it is important to identify the conserved motifs using sequence analysis methods. In this study, the sequence conservation was investigated by an *in silico* analysis to generate sequence logos. We have identified new conserved motifs in the domains TM1, TM2 and the cytoplasmic helix from 231 homologs of MS channel of large conductance (MscL). In addition, we have identified new motifs for the TM3 and the cytoplasmic carboxy-terminal domain from 309 homologs of MS channel of small conductance (MscS). We found that the conservation in MscL homologs is high for TM1 and TM2 in the three domains of life. The conservation in MscS homologs is high only for TM3 in Bacteria and Archaea.

**Keywords** Mechanosensitive channels · Motifs · Sequence logos

## Abbreviations

MS Mechanosensitive ion channels  
EcoMscL *Escherichia coli* MscL

TbMscL *Mycobacterium tuberculosis* MscL  
EcoMscS *Escherichia coli* MscS  
WT Wild type

## Introduction

All living cells contend with differences of water activity across their membranes and osmoregulation is found to be universal. Bacterial cells maintain an outwardly directed turgor pressure through the accumulation of solutes against their chemical gradient, leading to water flow into the cytoplasm (Epstein 1986). When subjected to hypo-osmotic stress (downshock), bacteria release solutes from the cytoplasm, thus preventing excessive water inflow (Berrier et al. 1992; Levina et al. 1999). The major proteins involved in fast solute release are the Mechanosensitive (MS) channels (Martinac et al. 1987, 1990; Sukharev et al. 1994). In *Escherichia coli*, rapid release of osmolytes is required for survival of extreme turgor by activation of MscL and MscS channels (Levina et al. 1999; Booth et al. 2007). These channels are the main signalling molecules of mechanosensory transduction, because they convert mechanical forces in the lipid bilayer of cellular membranes into electrical and/or biochemical signals (Perozo 2006).

Mechanosensitive channel of large conductance (MscL) exhibits the largest conductance (greater than 2.5 nS), opens at high membrane tension (close to the lytic limit of the bilayer) and has fast kinetics (Perozo 2006). Each MscL subunit is a 136 amino acid polypeptide predicted to have two transmembrane helices TM1 and TM2 (Sukharev et al. 1994; 1999). The crystal structure of TbMscL at 3.5 Å of the *Mycobacterium tuberculosis* ortholog (151 amino acids) revealed a pentameric channel in an apparently closed state (Chang et al. 1998). The core of the protein is

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00249-009-0460-y) contains supplementary material, which is available to authorized users.

---

D. Balleza (✉)  
Laboratory of Molecular Biology,  
University of Wisconsin, 1525 Linden Drive,  
Madison, WI 53706, USA  
e-mail: dballeza.uwisc@gmail.com

D. Balleza · F. Gómez-Lagunas  
Departamento de Fisiología, Facultad de Medicina,  
UNAM. México. Cd. Universitaria, A.P. 70-250,  
México DF 04510, Mexico

formed by tightly packed TM1 (residues 15–45) and TM2 (residues 76–100) transmembrane helices from all five subunits. The TM1 and TM2 helices are linked by a periplasmic loop. The COOH-terminal part of the protein (residues 101–110) adopts an  $\alpha$ -helical bundle conformation and reside in the cytoplasm (Sukharev et al. 2001a, b; Gullingsrud and Schulten 2003; Maurer et al. 2008), while the segment S1 at the N-terminus is in close proximity to the cytoplasmic interface (Iscla et al. 2008; Steinbacher et al. 2007). On the other hand, MS channel of small conductance (MscS) is a 0.8–1-nS channel opened by moderate pressure (Perozo 2006). The structure of the EcoMscS channel from *E. coli* (286 amino acids) was originally solved at 3.9 Å (Bass et al. 2002) and recently at 3.45 Å in an open conformation (Wang et al. 2008). However, unlike the more uniform MscL family, the MscS homologs form a much larger family with more than 18 subfamilies of different peptide lengths, and it extends from Archaea to plants (Pivetti et al. 2003). The EcoMscS is a homoheptamer, each subunit consisting of two major domains: the amino-terminal membrane domain and the carboxy-terminal cytoplasmic domain. Each subunit has three TM helices (TM1–TM2–TM3). TM3 (residues 96–127) lines the channel pore and contains predominantly hydrophobic residues (Bass et al. 2002; Miller et al. 2003a). The large (~17 kDa) cytoplasmic carboxy-terminal domain exhibits three major sub-domains. A  $\beta$  sub-domain (residues 132–177) is linked to a mixed  $\alpha\beta$  sub-domain (residues 188–265) and the protein terminates in a  $\beta$  barrel (residues 271–280). The cytoplasmic domain creates a large vestibule that is perforated by seven  $10 \times 8$  Å lateral portals created by the junctions of the  $\beta$  and  $\alpha\beta$  sub-domains and an axial portal formed by the  $\beta$  barrel (Bass et al. 2002; Wang et al. 2008).

Herein, we have used a profile-to-profile alignment strategy to align the TM1, TM2 and cytoplasmic helix of 231 MscL homologs and separately the TM3 and  $\alpha\beta$  sub-domain of 309 MscS homologs. Using this approach, we have precisely aligned these repeated regions, and identified new conserved residues and consensus motifs. Sequences from Archaea and Eukarya were included, since the mechanism of mechanotransduction is essential for osmoregulation, and therefore seems ubiquitous in life (Kloda and Martinac 2001, 2002).

## Methods

### Data collection

MS sequences were collected from the NCBI (<http://www.ncbi.nlm.nih.gov>) database with a keyword search by hand using the sequence of the whole protein for the search. We select protein sequences based in the original annotation as

MscL or MscS homologs and confirmed by BLAST. To obtain a more comprehensive representation of the conserved motifs of these channels, a representative sequence from every genus available in the NCBI database was selected and aligned. In a few cases, sequences from species within the same genus were included in the analysis. For MscL family, we selected sequences from 143 gram-negative Bacteria, 66 gram-positive Bacteria, 8 Archaea and 14 Eukarya (see Supplementary Table S1). In the case of MscS family, we examined 204 sequences from gram negative and 56 from gram-positive Bacteria, 34 Archaea and 15 Eukarya (see Supplementary Table S2).

### Building alignments and identification of conserved motifs

Data analysis was carried out separately for each taxonomic group, and we generated consensus sequences for each domain using multiple alignments with Clustal-W (Thompson et al. 1994). Limited regions of MscL (TM1–TM2–Cyt-H) and MscS (TM3 and Cyt-terminus) were determined based on the crystal structure of both channels (Chang et al. 1998; Bass et al. 2002; Wang et al. 2008) and taking into account the more conserved regions for each alignment. Sequence logos were generated using the WebLogo program (<http://weblogo.berkeley.edu/logo.cgi>) (Crooks et al. 2004). Consensus motifs were determined from multiple alignments, presented in the PROSITE format (<http://www.expasy.org/prosite>) (Sigrist et al. 2002), and were determined by the  $\geq 50\%$  of conservation threshold in each sequence logo. For the logos from Bacteria, we follow the numbering of *E. coli* (MscL and MscS); for the logos from Archaea, we used the numbering of *Methanoregula boonei* (MscL) and *Methanocaldococcus jannaschii* (MscS). For the logos from Eukarya, we follow the numbering of *Aspergillus clavatus* for MscL and *Arabidopsis thaliana* (AtMSL3) for MscS. However, all our analysis was compared to *E. coli* numbering. Results from structure–function and mutational studies were collected from PubMed literature searches. It is important to note that in relation to the limited number of sequences from Archaea and Eukarya obtained in the protein database, our analysis can result in an underestimation of the logos founded. However, the method incorporates a sample correction, which can diminish this problem (Crooks et al. 2004).

## Results and discussion

In this study, a sequence logo generator was used to obtain sequence logos from the homologs of MscL and MscS channels. The logo generator program takes the consensus sequence and adds statistical information about the relative frequency, conservation of each residue and the chemical

composition of each motif. Thus, each logo gives a richer and more accurate description of sequence similarity. At present, there are some reports where multiple alignments for MscL (Chang et al. 1998; Pivetti et al. 2003; Moe et al. 1998; Oakley et al. 1999; Spencer et al. 1999) and MscS channels (Pivetti et al. 2003; Miller et al. 2003a; Koprowski and Kubalski 2003) have been done. Here, the conserved residues from MscL and MscS channels have been systematically identified and compared. To explore the homology of sequences before the generation of sequence logos separately, we compare the sequences of each channel from representative sequences of Bacteria, Archaea and Eukarya at the same time. This combined analysis shows that some residues are well conserved throughout the three domains of life (Figure S1). The most salient feature of this comparison is that the motifs analyzed are similar within each major group and shows good conservation between the three domains of life. Our comparative approach illustrates the high conservation and relative importance of TM1–TM2 for MscL and TM3– $\beta$ -subdomain for MscS. This information is important to obtain a more precise representation of conserved motifs that might help, for instance, for mutational or structural studies. Out of these results, several common characteristics to all MscL and MscS homologs can be identified in future studies.

#### MscL protein: conserved motifs in TM1

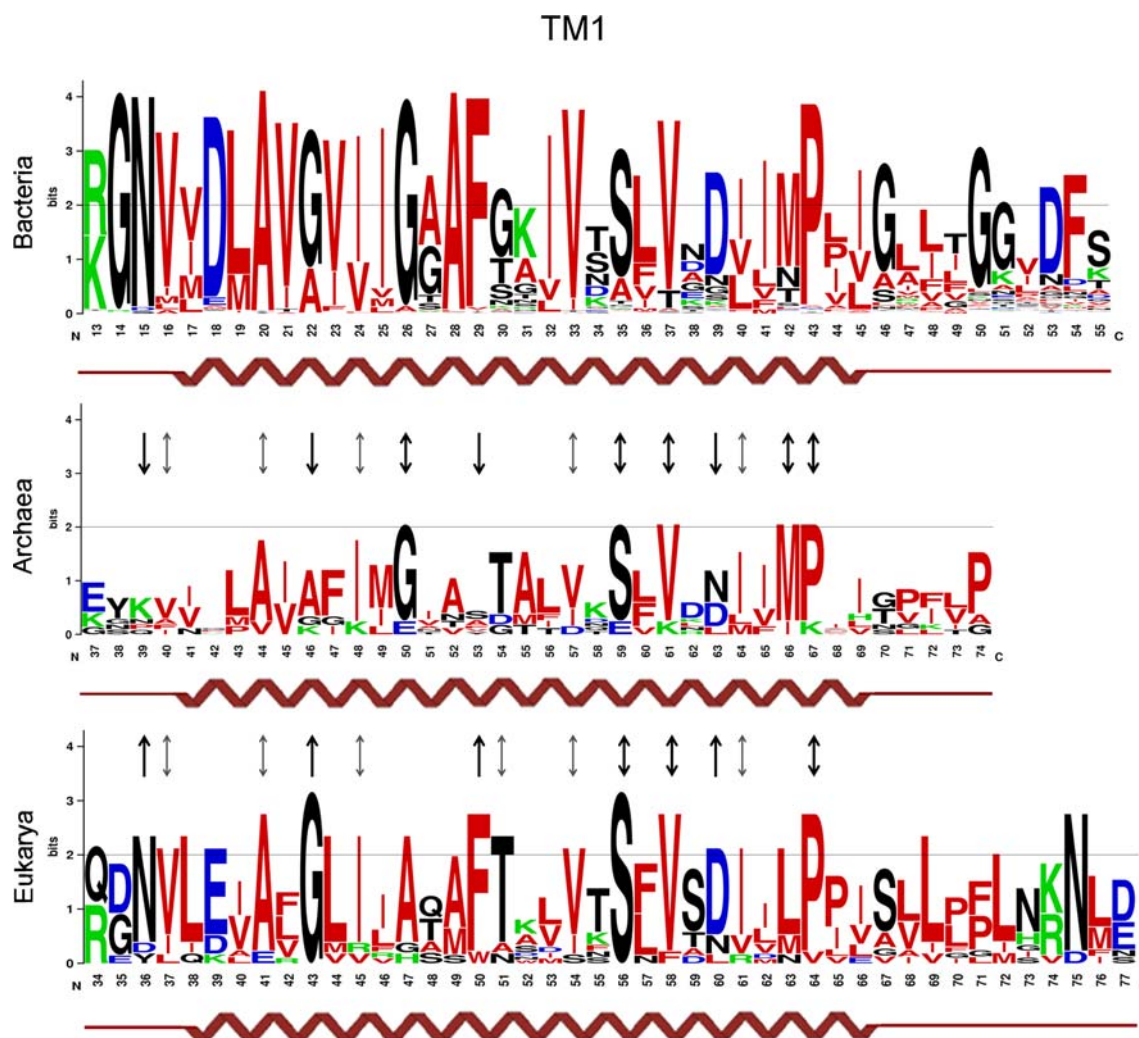
In 209 bacterial sequences compared (*E. coli* numbering), the highly conserved residues ( $\geq 90\%$ ) for this segment are aspartate/glutamate (D/E18), alanine (A) at position 20 and 28, glycine G26, phenylalanine F29, valine V33/V37 and proline P43. We found the following consensus motif for TM1 of MscL from Bacteria: [VIML]-D-[LM]-A-[VI]-[GA]-[VI]-[IV]-I-G-[AG]-A-F-[GTS]-x-I-V-x-[SA]-[LFV]-[VT]-x-D-[IVL]-[ILVF]-[MNT]-P-x-[IVL]. Similarly, for eight sequences from Archaea, we identified the following conserved residues ( $\geq 50\%$ ) (*Methanoregula boonei* numbering): G50, S59, V61, M66 and P67. In this segment we found the consensus motif G-x(8)-S-x-V-x(4)-M-P, where G50, S59, V61, M66 and P67 are equivalent to G26, S35, V37, M42 and P43 in Bacteria respectively. In Eukarya (14 sequences compared), we found six well-conserved residues ( $\geq 70\%$ ): A41, G43, F50, S56, V58 and P64 (*Aspergillus clavatus* numbering). The consensus sequence found here is: [AE]-x-G-L-I-x-A-x(2)-F-T-x(2)-V-x-S-x-V-x-[DN]-[IV]-x-[LM]-P. Again, we found here a good correspondence within the sequence from Bacteria, where A41, G43, F50, S56, V58 and P64 correspond to A20, G22, F29, S35, V37 and P43 in *E. coli* respectively (Fig. 1; Table 1). We found that the most conserved residues in TM1 present in all the three domains are the equivalent to: A20, I24, V33, S35, V37, I40 and P43 from *E. coli* (see Figure S2).

#### MscL protein: conserved motifs in TM2

We found that TM2 from the sequences of the three domains of life are rich in hydrophobic residues, principally F, I, A and V. In bacterial sequences, the highly conserved residues ( $\geq 80\%$ ) are G76, F78, F85, I87, A89 and F93. We identified the following consensus motif for TM2 from Bacteria: Y-G-x-F-[IL]-x(3)-[IVLF]-[ND]-F-[LVI]-[IL]-[IVL]-A-F-x-[IVL]-[FY]. For Archaea, we found two highly conserved residues ( $\geq 70\%$ ): G93 and F95, where G93 and F95 correspond to G76 and F78 respectively (*M. boonei* to *E. coli* numbering). In addition, we found four moderately conserved residues ( $\geq 50\%$ ): N/E101, F/R102, A/K106 and F/T110, where that positions are parallel to N/D84, F85, A89 and F93 in the bacterial sequences. Thus, the motif found for Archaea is: G-x-F-x(5)-N-F-x(3)-A-x(3)-F. In Eukarya, we found that G112 (*A. clavatus* numbering) is the most conserved residue ( $\geq 79\%$ ) and it is parallel to G76 in Bacteria and G93 in Archaea. In addition, Y111 and F114 show good conservation ( $\geq 60\%$ ) and the equivalent residues are Y75 and F78 in Bacteria. The consensus founded here is: A-x(2)-D-G-A-x-V-x-A-Y-G-x-F (Fig. 2; Table 1). We found that the most conserved residues in TM2 present in all the three domains are the equivalent to: G76, F78 and F85 from *E. coli* (see Figure S2).

#### MscL protein: conserved motifs in the cytoplasmic helix

This part of the protein is not well conserved ( $< 52\%$  of conservation threshold in each sequence logo). However, we found that  $\sim 45\%$  of the bacterial sequences studied have the motif: R-K-K-E-E-x(2)-P-x-P-P, where the second glutamate residue is frequently conserved and corresponds to E108 in EcoMscL. For Archaea, we found a putative truncated segment in seven of the eight studied sequences and only the homolog from *Thermoplasma acidophilum* looks complete. In this case, we found in a  $\sim 40\%$  the following motif: V-x(2)-E-E-K-x(2)-K. It is notable that deletion of residues 110–136 (C-terminus) in EcoMscL is tolerated, producing only a slight effect on gating (Blount et al. 1996; Hase et al. 1997). This mutant channel significantly increases the osmotic ATP release (Anishkin et al. 2003). This is interesting in the context of the MscL channels from Archaea, since there is evidence suggesting that this part of the protein is probably deleted naturally. On the other hand, the Cyt-H has a ‘prefilter’ function (Anishkin et al. 2003), and a possible role for the pH regulation of EcoMscL activity by the charged cluster RKKEE (residues 104–108) has been proposed (Kloda et al. 2006). It will be interesting to probe a similar mechanism in the archaean homologs. In eukaryotes (*A. clavatus* numbering), we observed only T148 and I158 conserved ( $\geq 60\%$ ) but not a clear motif (Fig. 3; Table 1).



**Fig. 1** Conserved motifs in the TM1 modules of MscL from Bacteria (numbering of *E. coli*), Archaea (numbering of *M. boonei*) and Eukarya (numbering of *A. clavatus*). Residues are distinguished by colors: AVLIPWFM (hydrophobic, red), KRH (positively charged, green), DE (negatively charged, blue) and CGNQSTY (polar but uncharged, black). **Bold double-arrows** show high conservation ( $\geq 50\%$ ) of the

identical residue in the three domains. *Gray double-arrows* show high conservation of the same residue in two domains and low conservation ( $< 50\%$ ) for the same residue in the third domain. *Single arrows* show high conservation ( $\geq 50\%$ ) of the same residue only in two domains. The data for these logos consists of 209 sequences from Bacteria, 8 from Archaea and 14 from Eukarya

### MscS protein: conserved motifs in TM3

The MscS family has good conservation throughout the TM3 helix, in contrast to the TM1 and TM2 (Pivetti et al. 2003; Miller et al. 2003a), which are relatively poorly conserved, and we confirm these observations. A previous study of Koprowski and Kubalski (2003) compares 36 sequences of MscS homologs and found that TM3 shows the highest level of identity as a whole (44% as compared to 30 and 23% estimated for TM1 and TM2, respectively). This observation clearly indicates that TM3 is important for MscS channel function and suggests that TM3 may be the functional equivalent of TM1 in MscL protein. It has been proposed that both TMs serve as channel-lining helices with a similar generalized structure and possibly a common

evolutionary origin (Pivetti et al. 2003; Kloda and Martinac 2001; 2002), although the overall structure of both proteins are quite different, indicating that these channels do not share a common evolutionary ancestor (Steinbacher et al. 2007). This diversity also reflects a much wider distribution of MscS than for MscL (see Supplementary Tables S1 and S2) with homologs found in Bacteria, Archaea, yeast, *Entamoeba*, *Chlamydomonas* and higher plants.

Our analysis corroborate that TM3 is G-rich, showing the characteristic periodicity of this residue (Edwards et al. 2005), and additionally we found that TM3 is rich in A, L, F and V residues with different level of conservation between the three domains of life. In Bacteria, the highly conserved residues ( $\geq 60\%$ ) that have shown periodicity in this domain are G101, G104, A106, G108, L/F109,

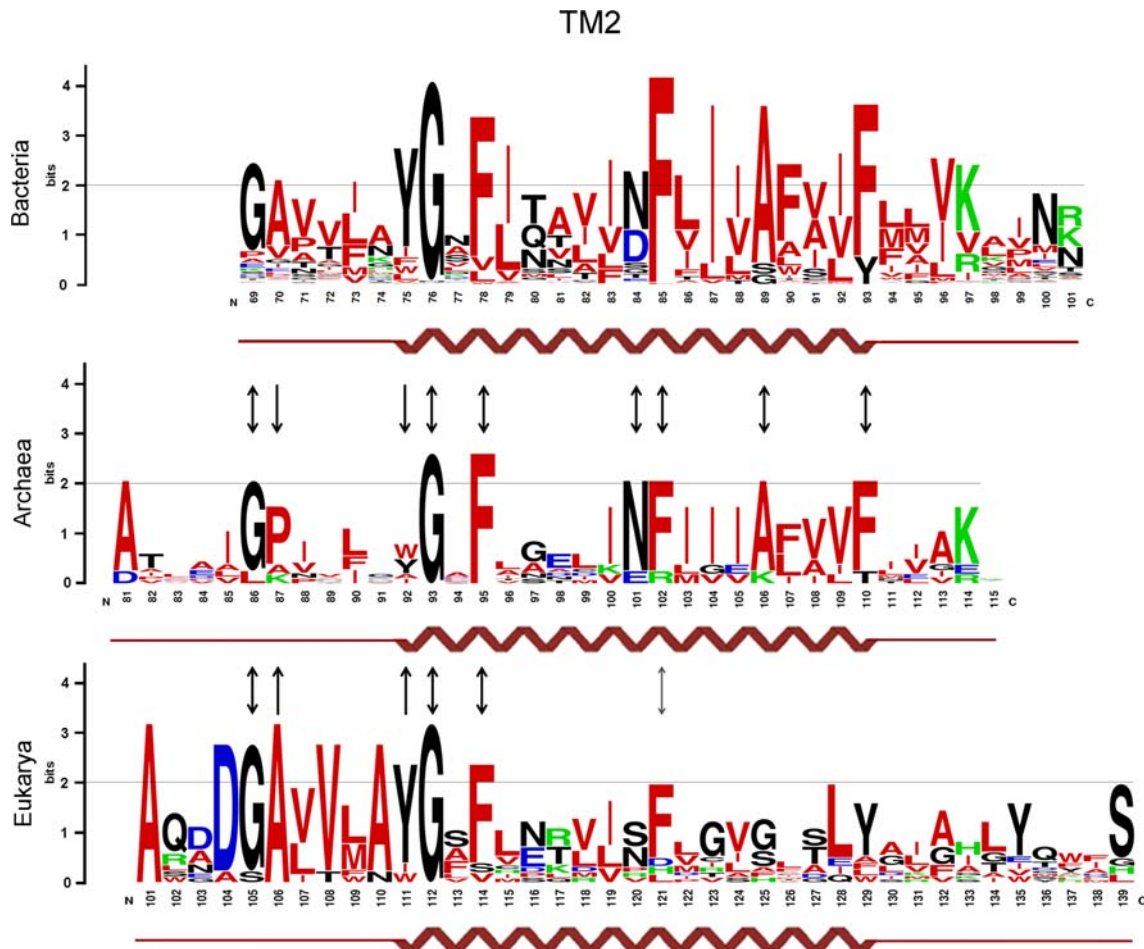
**Table 1** Conserved residues and motifs in MS channels from Bacteria, Archaeobacteria and Eukarya

| Channel                                  | Conserved residues (as a minimum 50%)  | Consensus motif   |
|--|--|---|
| <b>Bacteria<sup>a</sup></b>              |  |   |
| MscL                                     |  |   |
| TM1                                      | V16, V/I17, D18, L19, A20, V21, G/A22, V23, I/V24, I25, G26, A/G27, A28, F29, I32, V33, S35, L36, V37, D39, I/V/L40, I41, M42, P43, I/V/L45, G46             | [VIML]-D-[LM]-A-[VI]-[GA]-[VI]-[IV]-I-G-[AG]-A-F-[GTS]-x-I-V-x-[SA]-[LFV]-[VT]-x-D-[IVL]-[ILVF]-[MNT]-P-x-[IVL]         |
| TM2                                      | Y75, G76, F78, I/L79, I/V83, N/D84, F85, L/V86, I87, I/V/L88, A89, F90, V/A/I91, I/V/L92, F/Y93  | Y-G-x-F-[IL]-x(3)-[IVLF]-[ND]-F-[LVI]-[IL]-[IVL]-A-F-x-[IVL]-[FY]   |
| Cyt-Helix                                | E108   | –   |
| MscS                                     |  |   |
| TM3                                      | L/F/I/V96, A/T/S98, G101, A/I/V102, G/S/T/A104, L/V/I/A105, A/V/G106, I/V/L107, G/A/M108, L/F109, A/G110, Q/K112, L/V/I115, N/D117, A/S/G120, G/S121, I/L124 | [LFIV]-x-[ATS]-x(2)-G-x(2)-[GSTA]-[LVIA]-[AVG]-[IVL]-[GAM]-[LF]-[AG]-x-[QK]-x(2)-[LVI]-x-[ND]-x(2)-[ASG]-[GS]-x(2)-[IL] |
| $\alpha\beta$ -subdomain (Cyt-vestibule) | I/L/V261, I/L/V270, P/A273, P275   | [ILV]-x(8)-[ILV]-x(2)-[PA]  |
| <b>Archaea</b>                           |  |   |
| MscL <sup>b</sup>                        |  |   |
| TM1                                      | G50, S59, V61, M66, P67  | G-x(8)-S-x-V-x(4)-M-P   |
| TM2                                      | G93, F95, N101, F102, A106, F110   | G-x-F-x(5)-N-F-x(3)-A-x(3)-F  |
| Cyt-Helix                                | K121   | –   |
| MscS <sup>c</sup>                        |  |   |
| TM3                                      | L/A/I/V162, G/A/T166, G/A/T169, I/L/A170, L/I/V172, G/A/S173, F/L174, A/G175, Q/K/R177, N/D182, A/S/G185, G186, I/L/M189                                     | [GAT]-x(2)-[GAT]-[ILA]-x-[LIV]-[GAS]-[FL]-[AG]-x-[QKR]-x(4)-[ND]-x(2)-[ASG]-G-x(2)-[ILM]                                |
| $\alpha\beta$ -subdomain (Cyt-vestibule) | L/I/F/V301, I/V/R/L326   | –   |
| <b>Eukaryota</b>                         |  |   |
| MscL <sup>d</sup>                        |  |   |
| TM1                                      | V37, L38, E/D39, A41, G43, L44, I45, A47, F50, T51, V54, S56, V58, D/N60, I/V61, L/M63, P64  | [AE]-x-G-L-I-x-A-x(2)-F-T-x(2)-V-x-S-x-V-x-[DN]-[IV]-x-[LM]-P   |
| TM2                                      | A110, Y111, G112, F114, F121, L128   | A-x(2)-D-G-A-x-V-x-A-Y-G-x-F  |
| Cyt-Helix                                | T148, I158   | –   |
| MscS <sup>e</sup>                        |  |   |
| TM3 <sup>f</sup>                         | G/E/S/C271, A/I280, N287, G/S291   | –   |
| $\alpha\beta$ -subdomain (Cyt-vestibule) | –  | –   |

<sup>a</sup> Numbering of *E. coli*<sup>b</sup> Numbering of *M. boonei*<sup>c</sup> Numbering of *M. jannaschii*<sup>d</sup> Numbering of *A. clavatus*<sup>e</sup> Numbering of AtMSL3 from *A. thaliana*<sup>f</sup> 48% as a minimum

A/G110, Q112, N117 and G121. In the crystal structure, the TM3 helix bends at G113 and this residue is very important in the gating of the channel (Edwards et al. 2005; Akitake et al. 2007). Curiously, after our analysis, G113 is not well conserved in Bacteria and Asp mainly occupies such position. This observation suggests that the particular

amino acid at position 113 is significant but not restricted to a G residue. Nevertheless, we found high conservation for G121, which is frequently used for the same function of flexibility and alternative kink positions could potentially correspond to different functional states of the channel (e.g. closed, desensitized and inactivated) (Akitake et al. 2007;



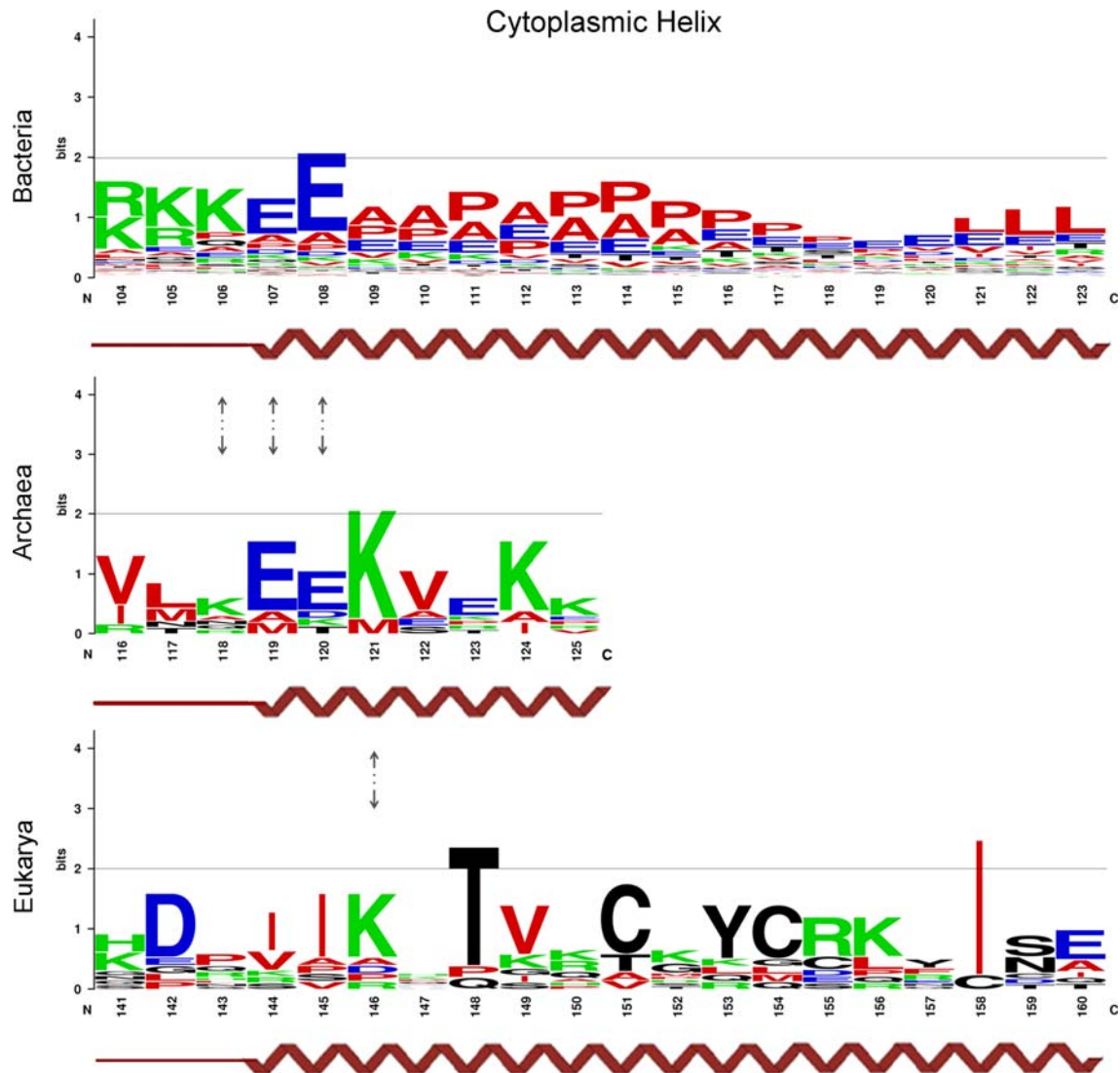
**Fig. 2** Conserved motifs in the TM2 modules of MscL from Bacteria, Archaea and Eukarya. See Fig. 1 for details

Edwards et al. 2008). For TM3 domain from Bacteria, we found the consensus motif: [LFIV]-x-[ATS]-x(2)-G-x(2)-[GSTA]-[LVIA]-[AVG]-[IVL]-[IGAM]-[LF]-[AG]-x-[QK]-x(2)-[LVI]-x-[ND]-x(2)-[ASG]-[GS]-x(2)-[IL]. For Archaea, we observed eight well-conserved residues ( $\geq 63\%$ ): G166, G169, G173, F/L174, A175, Q177, N182 and G186. These residues are equivalent to G101, G104, G108, L/F109, A/G110, Q112, N117 and G121, respectively (*M. jannaschii* to *E. coli* numbering). It is important to highlight that D/T/S178 in Archaea is equivalent to the residue G113 and the well-conserved G186 corresponds to G121 from Bacteria. For TM3 domain from Archaea, we found the following consensus motif: [GAT]-x(2)-[GAT]-[ILA]-x-[LIV]-[GAS]-[FL]-[AG]-x-[QKR]-x(4)-[ND]-x(2)-[ASG]-G-x(2)-[ILM]-x(3)-[RKQ]-[PFTS]-x(3)-[GN]. Concerning TM3 from Eukarya, we observed poor conservation but some periodicity for the G residues in comparison with prokaryotes. We found nine residues weakly conserved ( $\geq 39\%$ ): L267, G271, G/L274, L/F279, A/I280, N287, S/E290, G/S291 and L292 (numbering of ‘MscS-like protein’ AtMSL3 from *A. thaliana*). It is important to note here that G271, G/L274

and G/S291 are equivalent to G101, G104 and G121 from Bacteria and G166, G169 and G186 from Archaea, respectively (Fig. 4; Table 1). We found that the most conserved residues in TM3 present in all the three domains are the equivalent to G101, G104, G108, A110, N117 and G121 from *E. coli* (see Figure S3).

MscS protein: conserved motifs in the C-terminal domain

In *E. coli*, W240 and W251 lie at the  $\beta$  and  $\alpha\beta$  interfaces that create the seven lateral portals, and both residues are located close to the subunit interfaces in the C-terminal domain, and therefore they are close to the portals that penetrate the vestibule (Bass et al. 2002; Steinbacher et al. 2007). W240 has an essential role in MscS oligomer stability and assembly, but W251 has a less important role in such functions. It is located at the side portal and is in contact with the neighboring subunit (Rasmussen et al. 2007). According to a previous study, aromatic residues are required at both positions. W240 is conserved in almost 100% of the top 100 proteins showing strong overall similarity to

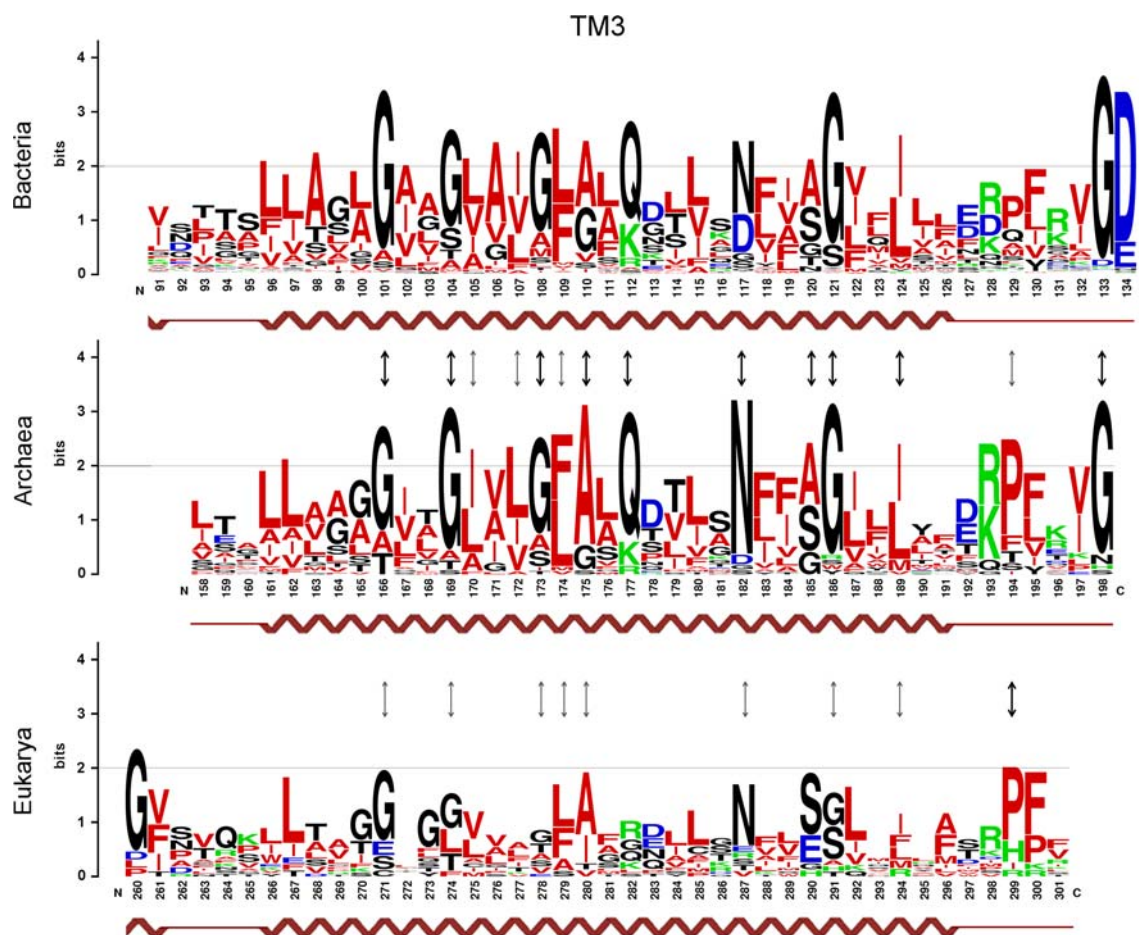


**Fig. 3** Conserved motifs in the cytoplasmic helix of MscL from Bacteria, Archaea and Eukarya. Dotted arrows show reduced conservation (<50%) of the same residue only in two domains

EcoMscS by BLAST, and W251 is more frequently founded as a F (57%) and only 29% have W at this position (Altschul et al. 1997). Nevertheless, in accordance with our analysis, after comparing 260 MscS homologs from Bacteria we found that W240 is weakly conserved (40.6%), Tyr can be found usually in this position (16.5%) and occasionally Phe (7.1%). On the other hand, at position 251 we found W (7.1%) or L/F (9.1%). However, it is important to note that when W is present in position 251, W at position 240 is constantly present. Indeed, in Bacteria we found poor conservation for this region and only five positions are relatively conserved ( $\geq 48\%$ ): V/L/I237, I/L/V261, I/L/V270, P/A273 and P/F275 (*E. coli* numbering). These residues are located between  $\alpha\beta$  subdomain and the  $\beta$  barrel. For Archaea, there are five positions conserved ( $\geq 46\%$ ): V/L/I299, L/I/F/V301, W306, V/A307 and I/V326 (*M. janna-*

*schii* numbering). After comparing 34 MscS sequences, we found that W306 from Archaea is equivalent to W240 in Bacteria. Finally, when 15 MscS-like sequences were compared from eukaryotes, a significant conservation in this region was not found, showing a high variability of sequence (Fig. 5; Table 1).

In summary, bacterial MscL and MscS have been extensively studied, and several residues important for the channel activity have been previously described (see Tables 2, 3 and references therein). Given its high frequency in the sequences analyzed here, we confirm the importance of such residues. Even so, according our analysis, some residues have received less attention. Tables 2 and 3 shows significant residues to be considered. To our knowledge, some residues have not been reported as important mutant phenotypes or the effect of mutations produce mild phenotypes.



**Fig. 4** Conserved motifs in the TM3 modules of MscS from Bacteria (numbering of *E. coli*), Archaea (numbering of *M. jannaschii*) and Eukarya (numbering of AtMSL3 from *A. thaliana*). The data for these

logos consists of 260 sequences from Bacteria, 34 from Archaea and 15 from Eukarya. See Fig. 1 for details

On the other hand, highly conserved residues founded in archaeal and eukaryotic homologs have not been evaluated until now.

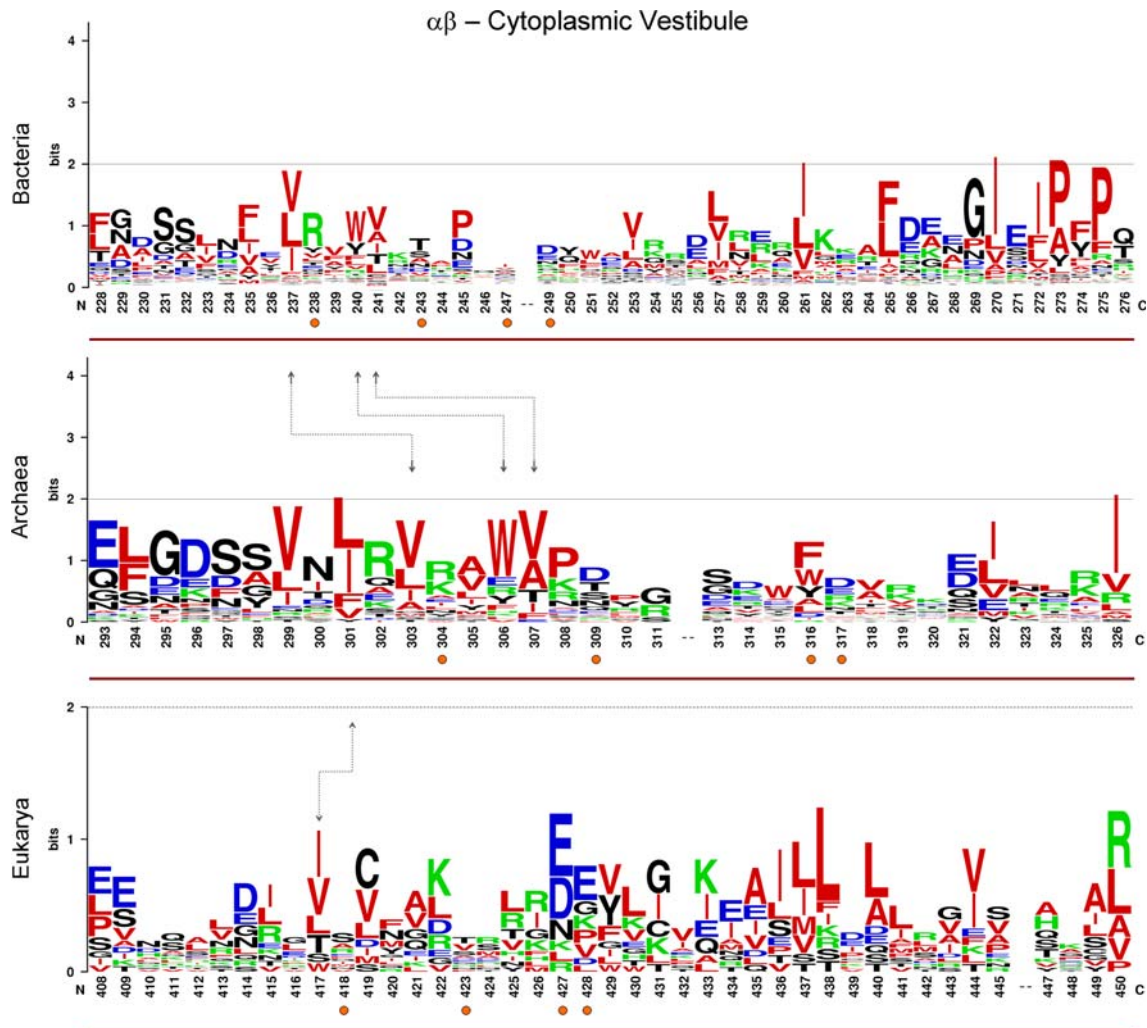
#### An evolutionary consideration

MS channels are present in organisms from all three domains of the universal tree. These channels form a superfamily of MS membrane proteins that probably have descended from a common ancestor (Pivetti et al. 2003; Kloda and Martinac 2001, 2002). This suggests that MS channels may have appeared early in the evolution of cellular life and that mechanosensation probably originates very early to fulfill crucial functions for cell survival (Martinac and Kloda 2003; Anishkin and Kung 2005; Kung 2005; Martinac et al. 2008). It has been proposed that gating of MS channels by mechanical force may have first evolved in cell-walled microbes and then has been relatively conserved in eukaryotic cells (Hamill and Martinac 2001). In general terms, this is in accordance with the good conservation between MS channels orthologs in Bacteria and

Archaea but the minor conservation in Eukarya that we found. Indeed, it is interesting to know which sequences used in our study form functional MS channels.

In particular, MscS-like proteins have been identified in some higher plants (*Arabidopsis* and *Oryza*), single-celled alga *Chlamydomonas* and fission yeast (*Schizosaccharomyces pombe*), but MscS homologs have not been found in animals (Pivetti et al. 2003) and our sequence analysis confirms this. Hence, it is notable that both MscS and MscL homologs founded in Eukarya are mostly from cell-walled organisms (Fungi and Plantae): for MscS, we found 13 homologs from plants, 1 from *S. pombe* and 1 for *Entamoeba* whereas for MscL, we found 13 homologs from Fungi and curiously 1 from *Apis*. The cell wall itself provides important mechanical protection to the membrane, and this finding suggests that MS channels may have evolved to regulate turgor pressure and perform similar functions as in prokaryotic MS channels (Kloda and Martinac 2002). In *Arabidopsis*, for example, MSL2 and MSL3 could be MS channels that control plastid size, shape and perhaps division during normal plant development releasing





**Fig. 5** Conserved motifs in the cytoplasmic vestibule of MscS channels from Bacteria, Archaea and Eukarya. Polar and charged residues that surround the lateral portals in the MscS crystal structure (Bass

et al. 2002) and homologs in Archaea are marked below the lines with a filled orange circles. Dotted arrows show poor conservation (<50%) of the same residue only in Bacteria and Archaea

solutes in response to osmotic downshock (Haswell and Meyerowitz 2006). The presence of such MscS homologs in *Arabidopsis* and *Chlamydomonas* (Nakayama et al 2007) is very interesting with respect to the symbiotic origin of the chloroplast as well as the function of this channel in the control of organelle morphology (Haswell and Meyerowitz 2006). It is possible that MscS family members have evolved new roles in plants since the endosymbiotic event. On the other hand, the presence of MscL homologs predominantly in filamentous fungi is of great significance because, as saprophyte and parasitic symbionts, they experience both slow and fast changes in extracellular water activity, and adaptation to changing environmental conditions is an essential trait for the survival in competitive habitats (Hohmann 2002). MscL channel is essential for the survival of extreme turgor and the presence of MscL homologs in fungi could be an indicative of their role in releasing

the pressure built up by hypoosmotic shocks. A basic question is whether the mechanisms of mechanical gating that originated in prokaryotic channels are conserved in the eukaryotic MS channels.

### Concluding remarks

In this paper, we identify several new conserved motifs from homologs of MscL and MscS from all three branches of the universal tree of life. A similar study for voltage-gated ion channel has been published elsewhere (Guda et al. 2007). Our study presents the first attempt to summarize very recent knowledge about the MS channels of the bacterial as well as non-bacterial homologs, and offers the opportunity to compare and contrast the structural and functional properties of these channels. Our analysis could

**Table 2** Alternative residues in conserved motifs for MscL channels from Bacteria, associated with mutational data

| Channel | Position <sup>a</sup> | Alternative residues <sup>b</sup> | Mutational data | References   |
|---------|-----------------------|-----------------------------------|-----------------|--|
| EcoMscL | R13 (0.53)            | K (0.46)                          | R13C/L          | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | G14 (0.98)            | R (<0.01)                         | G14C/E          | Ou et al. (1998) and Bartlett et al. (2004)  |
|         | N15 (0.96)            | D (0.01)                          | N15D/K/S/Y      | Ou et al. (1998) and Maurer and Dougherty (2003)   |
|         | V16 (0.85)            | I (0.04)                          | V16A/C/E/M      | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | V17 (0.43)            | I (0.27)                          | V17A/C/E        | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | D18 (0.92)            | E (0.04)                          | D18C/E/V        | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | L19 (0.79)            | M (0.19)                          | L19C/Y          | Ou et al. (1998), Batiza et al. (2002) and Bartlett et al. (2004)  |
|         | A20 (0.98)            | S (<0.01)                         | A20C/V          | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | V21 (0.86)            | I (0.11)                          | V21C            | Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | G22 (0.75)            | A (0.24)                          | G22C/D/I/N/S    | Ou et al. (1998), Maurer and Dougherty (2001), Bartlett et al. (2004), Levin and Blount (2004), Anishkin et al. (2005), Powl et al. (2008a, b) and Yoshimura et al. (2008) |
|         | V23 (0.76)            | I (0.19)                          | V23A/C/D/G/I/T  | Ou et al. (1998), Maurer and Dougherty (2003), Bartlett et al. (2004), Levin and Blount (2004) and Anishkin et al. (2005)  |
|         | I24 (0.71)            | V (0.28)                          | I24C/N/T/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | I25 (0.86)            | V (0.05)                          | I25C/S          | Maurer and Dougherty (2003) and Levin and Blount (2004)  |
|         | G26 (0.97)            | A (0.02)                          | G26C/S          | Ou et al. (1998), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | A27 (0.51)            | G (0.36)                          | A27C/V          | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | A28 (0.98)            | L (<0.01)                         | A28C            | Bartlett et al. (2004)   |
|         | F29 (0.96)            | V (0.01)                          | F29C/S/V/Y      | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | G30 (0.55)            | T (0.20)                          | G30C/E/R        | Ou et al. (1998), Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | K31 (0.45)            | A (0.23)                          | K31C/D/E/R/T    | Blount et al. (1997), Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | I32 (0.81)            | V (0.12)                          | I32C/T/V        | Maurer and Dougherty (2003) and Levin and Blount (2004)  |
|         | V33 (0.92)            | I (0.05)                          | V33A/C/D/F/I/S  | Maurer and Dougherty (2003) and Levin and Blount (2004)  |
|         | S34 (0.23)            | T (0.27)                          | S34C/L/P/T      | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | S35 (0.77)            | A (0.14)                          | S35P/T          | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | L36 (0.63)            | F (0.16)                          | L36C/P/Q/R      | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | V37 (0.87)            | T (0.09)                          | V37A/C/D        | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | A38 (0.13)            | N (0.23), D (0.15)                | A38C/T          | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | D39 (0.73)            | N (0.07)                          | D39C/E/N/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | I40 (0.45)            | V (0.26), L (0.25)                | I40C/N          | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | I41 (0.70)            | L (0.11)                          | I41C/N/T/V/W    | Maurer and Dougherty (2003), Bartlett et al. (2004) and Chiang et al. (2005)   |
|         | M42 (0.66)            | N (0.14)                          | M42C/G/I/K/T/V  | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | P43 (0.96)            | V (0.01)                          | n.d.            | This study   |
|         | P44 (0.24)            | L (0.29)                          | n.d.            | This study   |
|         | L45 (0.18)            | I (0.55), V (0.29)                | L45P/Q          | Maurer and Dougherty (2003)  |
|         | G46 (0.73)            | A (0.10)                          | G46D            | Ou et al. (1998)   |
|         | A70 (0.64)            | V (0.13)                          | A70C            | Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | V71 (0.45)            | P (0.18)                          | V71A/C/I        | Maurer and Dougherty (2003) and Bartlett et al. (2004)   |
|         | V72 (0.42)            | T (0.20)                          | V72A/C/G/I      | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)  |

**Table 2** continued

| Channel | Position <sup>a</sup> | Alternative residues <sup>b</sup>   | Mutational data | References  |
|---------|-----------------------|-------------------------------------|-----------------|---|
|         | M73 (0.07)            | L, F (0.46), I (0.29)               | M73C/K/L/T/V    | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | H74 (0.03)            | A, N (0.53)                         | H74C/L/Q/R      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | Y75 (0.72)            | I (0.07)                            | Y75C/D/E/F      | Maurer and Dougherty (2003), Bartlett et al. (2004) and Chiang et al. (2005)  |
|         | G76 (0.99)            | N (<0.01)                           | G76C/S          | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | V77 (0.09)            | N, A, S (0.68)                      | V77A/C/D        | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | F78 (0.83)            | V (0.08)                            | F78C/I/L/S/Y    | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | I79 (0.56)            | L (0.33)                            | I79C/L/N/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | Q80 (0.28)            | T (0.33)                            | Q80C/R          | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | N81 (0.10)            | A, T, V (0.77)                      | N81C/D/I/S      | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | V82 (0.37)            | I (0.2)                             | V82A/C/D/P      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | F83 (0.12)            | I, V, L (0.88)                      | F83C/G/I/L/S/Y  | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | D84 (0.28)            | N (0.52)                            | D84C/E/G/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | F85 (0.99)            | L (<0.01)                           | F85C/L          | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | L86 (0.49)            | V (0.18)                            | L86C/P/R        | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | I87 (0.88)            | L (0.09)                            | I87C/F/T/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | V88 (0.31)            | I (0.44)                            | V88C/E/G/L      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | A89 (0.88)            | S (0.06)                            | A89C            | Bartlett et al. (2004) and Levin and Blount (2004)  |
|         | F90 (0.64)            | A (0.12)                            | F90C/L/S/V      | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | A91 (0.29)            | V (0.36)                            | A91C/T          | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | I92 (0.49)            | V (0.32)                            | I92C/F/L/T/V    | Ou et al. (1998), Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)                         |
|         | F93 (0.84)            | Y (0.16)                            | F93C/L/S/Y/W    | Maurer and Dougherty (2003), Powl et al. (2003), Bartlett et al. (2004), Levin and Blount (2004) and Chiang et al. (2005) |
|         | M94 (0.23)            | L (0.32)                            | M94C/I/L/T/V    | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | A95 (0.12)            | L, M, V (0.74)                      | A95C/V          | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | I96 (0.31)            | V (0.50)                            | I96C/M/N/V      | Maurer and Dougherty (2003), Bartlett et al. (2004) and Levin and Blount (2004)   |
|         | K97 (0.55)            | V (0.24)                            | K97C/E/R        | Maurer and Dougherty (2003) and Bartlett et al. (2004)  |
|         | R104 (0.37)           | K (0.33)                            | R104Q/E/W       | Maurer and Dougherty (2003) and Kloda et al. (2006)   |
|         | K105 (0.45)           | R (0.21)                            | K105Q/E/N       | Maurer and Dougherty (2003), Kloda et al. (2006) and Powl et al. (2008b)  |
|         | K106 (0.49)           | P (0.07)                            | K106E/Q/R       | Maurer and Dougherty (2003) and Kloda et al. (2006)   |
|         | E107 (0.45)           | A (0.10)                            | E107Q/K/V       | Maurer and Dougherty (2003) and Kloda et al. (2006)   |
|         | E108 (0.63)           | A (0.10)                            | E108D/G/K/Q/R   | Maurer and Dougherty (2003) and Kloda et al. (2006)   |
|         | L121 (0.24)           | E (0.21)                            | L121C/E/I/S     | Anishkin et al. (2003) and Maurer and Dougherty (2003)  |
|         | L122 (0.37)           | E (0.18)                            | L122C/P/Q/R     | Anishkin et al. (2003) and Maurer and Dougherty (2003)  |
| TbMscL  | V15                   | [V17: <i>EcoMscL</i> ] <sup>c</sup> | V15C            | Shapovalov et al. (2003)  |
|         | L17                   | [L19: <i>EcoMscL</i> ]              | L17Y            | Maurer et al. (2000)  |
|         | A20                   | [G22: <i>EcoMscL</i> ]              | A20E/R          | Maurer et al. (2000)  |
|         | V21                   | [V23: <i>EcoMscL</i> ]              | V21A/D/K        | Maurer et al. (2000), Moe et al. (2000), Powl et al. (2005), Elmore and Dougherty (2001) and Yefimov et al. (2008)        |
|         | T28                   | [G30: <i>EcoMscL</i> ]              | T28R            | Maurer et al. (2000)  |

**Table 2** continued

| Channel | Position <sup>a</sup> | Alternative residues <sup>b</sup> | Mutational data | References  |
|---------|-----------------------|-----------------------------------|-----------------|---|
|         | N44                   | [G46: <i>EcoMscL</i> ]            | N44D            | Maurer et al. (2000)                              |
|         | R45                   | [L47: <i>EcoMscL</i> ]            | R45K/C          | Maurer et al. (2000)                              |
|         | L69                   | [Y75: <i>EcoMscL</i> ]            | L69W            | Powl et al. (2005)                                |
|         | F80                   | [L86: <i>EcoMscL</i> ]            | F80W            | Powl et al. (2003, 2005)                          |
|         | Y87                   | [I92: <i>EcoMscL</i> ]            | Y87W            | Powl et al. (2005)                                |
|         | R98                   | [R104: <i>EcoMscL</i> ]           | R98Q            | Powl et al. (2005)                                |
|         | K99                   | [K105: <i>EcoMscL</i> ]           | K99Q            | Powl et al. (2005)                                |
|         | K100                  | [K106: <i>EcoMscL</i> ]           | K100Q           | Powl et al. (2005)                                |
|         | E102                  | [E107: <i>EcoMscL</i> ]           | E102C           | Maurer et al. (2008)                              |
|         | E104                  | [P109: <i>EcoMscL</i> ]           | E104Q           | Maurer et al. (2008) and Shapovalov et al. (2003) |
|         | D108                  | [T116: <i>EcoMscL</i> ]           | D108N           | Maurer et al. (2008)                              |
|         | L113                  | [L121: <i>EcoMscL</i> ]           | L113I           | Maurer et al. (2008)                              |
|         | T115                  | [T123: <i>EcoMscL</i> ]           | T115A           | Maurer et al. (2008)                              |
|         | E116                  | [E124: <i>EcoMscL</i> ]           | E116D/Q         | Maurer et al. (2008)                              |
|         | D119                  | [D127: <i>EcoMscL</i> ]           | D119N           | Maurer et al. (2008)                              |
|         | L121                  | [L129: <i>EcoMscL</i> ]           | L121I           | Maurer et al. (2008)                              |

*n.d.* not determined

<sup>a</sup> Numbers in parenthesis indicate relative frequency

<sup>b</sup> The letters are ordered from most to least frequent

<sup>c</sup> Equivalent residue in *EcoMscL* according to multiple alignment

**Table 3** Alternative residues in conserved motifs for MscS channels from Bacteria, associated with mutational data

| Channel | Position <sup>a</sup> | Alternative residues <sup>b</sup> | Mutational data | References  |
|---------|-----------------------|-----------------------------------|-----------------|---|
| EcoMscS | T93 (0.19)            | L (0.16)                          | T93R            | Miller et al. (2003a) and Edwards et al. (2008)   |
|         | A94 (0.15)            | T (0.33)                          | A94D/L/S/V      | Miller et al. (2003a) and Edwards et al. (2005)   |
|         | S95 (0.30)            | A (0.18)                          | S95C/P          | Miller et al. (2003a, b)  |
|         | V96 (0.12)            | L (0.51), F, I (0.30)             | <i>n.d.</i>     | This study  |
|         | I97 (0.21)            | L (0.45)                          | <i>n.d.</i>     | This study  |
|         | A98 (0.60)            | T (0.13)                          | A98G/L/S/V      | Edwards et al. (2005) and Wang et al. (2008)  |
|         | V99 (0.11)            | G, S, L (0.62)                    | <i>n.d.</i>     | This study  |
|         | L100 (0.36)           | A (0.32)                          | <i>n.d.</i>     | This study  |
|         | G101 (0.87)           | A (0.06)                          | G101A/D/S       | Edwards et al. (2005, 2008) and Wang et al. (2008)  |
|         | A102 (0.31)           | I (0.25)                          | A102G/P         | Miller et al. (2003a), Edwards et al. (2008),<br>Vásquez et al. (2008) and Wang et al. (2008)                 |
|         | A103 (0.27)           | G (0.20)                          | A103G           | Edwards et al. (2008)   |
|         | G104 (0.70)           | S (0.15)                          | G104A/S         | Edwards et al. (2005) and Wang et al. (2008)  |
|         | L105 (0.39)           | V (0.27)                          | L105C           | Vásquez et al. (2008)   |
|         | A106 (0.61)           | V (0.15)                          | A106G/L/S/V     | Edwards et al. (2005), Vásquez et al. (2008)<br>and Wang et al. (2008)  |
|         | V107 (0.31)           | I (0.36)                          | V107C           | Vásquez et al. (2008)   |
|         | G108 (0.71)           | A (0.11)                          | G108A/C/S       | Edwards et al. (2005) and Vásquez et al. (2008)   |
|         | L109 (0.52)           | F (0.40)                          | L109A/C/S       | Miller et al. (2003a), Anishkin and Sukharev (2004),<br>Edwards et al. (2005, 2008) and Vásquez et al. (2008) |
|         | A110 (0.51)           | G (0.34)                          | A110C/G/L/S/V   | Edwards et al. (2005), Vásquez et al. (2008)<br>and Wang et al. (2008)  |

**Table 3** continued

| Channel | Position <sup>a</sup> | Alternative residues <sup>b</sup> | Mutational data   | References   |
|---------|-----------------------|-----------------------------------|-------------------|--|
|         | L111 (0.34)           | A (0.34)                          | L111C             | Vásquez et al. (2008)  |
|         | Q112 (0.66)           | K (0.21)                          | Q112C/G           | Akitake et al. (2007) and Vásquez et al. (2008)                              |
|         | G113 (0.18)           | D (0.27)                          | G113A/C/D/M/P/R/S | Edwards et al. (2005, 2008), Akitake et al. (2007) and Vásquez et al. (2008) |
|         | S114 (0.17)           | L, T (0.51)                       | n.d.              | This study   |
|         | L115 (0.42)           | V (0.25)                          | L115G/I/V         | Wang et al. (2008)   |
|         | S116 (0.23)           | K, A, G (0.34)                    | n.d.              | This study   |
|         | N117 (0.54)           | D (0.29)                          | N117D             | Miller et al. (2003a)  |
|         | L118 (0.25)           | F (0.36)                          | L118P             | Miller et al. (2003a)  |
|         | A119 (0.20)           | I, V (0.48)                       | n.d.              | This study   |
|         | A120 (0.40)           | S (0.31)                          | A120G             | Akitake et al. (2007)  |
|         | G121 (0.83)           | S (0.12)                          | G121A             | Akitake et al. (2007)  |
|         | V122 (0.27)           | I (0.27)                          | n.d.              | This study   |
|         | L123 (0.15)           | F (0.21)                          | n.d.              | This study   |
|         | L124 (0.43)           | I (0.44)                          | n.d.              | This study   |
|         | V125 (0.10)           | L, I (0.55)                       | n.d.              | This study   |
|         | M126 (0.04)           | L, F, A, I, V (0.77)              | n.d.              | This study   |
|         | F127 (0.13)           | E, D (0.46)                       | n.d.              | This study   |
|         | R128 (0.36)           | D (0.23)                          | R128C/D           | Nomura et al. (2008)   |
|         | P129 (0.43)           | Q (0.16)                          | n.d.              | This study   |
|         | F130 (0.42)           | L (0.15)                          | n.d.              | This study   |
|         | R131 (0.36)           | K (0.19)                          | R131C/D           | Nomura et al. (2008)   |
|         | A132 (0.06)           | V, I, L (0.75)                    | n.d.              | This study   |
|         | G133 (0.94)           | D (0.03)                          | n.d.              | This study   |
|         | E134 (0.13)           | D (0.82)                          | n.d.              | This study   |
|         | V237 (0.36)           | L (0.30)                          | n.d.              | This study   |
|         | W240 (0.41)           | Y (0.16)                          | W240F/L           | Rasmussen et al. (2007)  |
|         | W251 (0.27)           | L, F (0.26)                       | W251F/L           | Rasmussen et al. (2007)  |
|         | F261 (0.03)           | I, L, V (0.88)                    | n.d.              | This study   |

*n.d.* not determined

<sup>a</sup> Numbers in parenthesis indicate relative frequency

<sup>b</sup> The letters are ordered from most to least frequent

provide a useful starting point for future structure–function studies of diverse MS channels and leads us to some important conclusions:

- Conserved residues and motifs identified in this report can be used for new mutational studies to better understand the function of the MS channels in homologs from Bacteria, Archaea and Eukarya.
- Specifically, the conserved residues founded in archaeal and eukaryotic sequences are important targets to be considered in future research. Their sequence similarity with bacterial homologs may reflect functional relatedness. The first objective is to establish more electrophysiological activity for those putative channels.
- The elevated homology founded for MS channels from Bacteria and Archaea suggests that archaeal MS channels could participate in similar mechanisms of osmoregulation.
- Although we found a reduced homology in each motif for eukaryotes, their presence could suggest that all three domains of life may have evolved from a common ancestor, and therefore these channels may share similar functions.

**Acknowledgments** We thank Professor Ching Kung for his critical reading of this manuscript, to W. John Haynes and to Carmen Quinto for valuable suggestions and to Karla García y García for the compilation of the sequences. DGAPA IN222208 supported this work. DB received a fellowship from DGAPA-UNAM.

## References

- Akitake B, Anishkin A, Liu N, Sukharev S (2007) Straightening and sequential buckling of the pore-lining helices define the gating cycle of MscS. *Nat Struct Mol Biol* 14:1141–1149
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Anishkin A, Kung C (2005) Microbial mechanosensation. *Curr Opin Neurobiol* 15:397–405
- Anishkin A, Sukharev S (2004) Water dynamics and dewetting transitions in the small mechanosensitive channel MscS. *Biophys J* 86:2883–2895
- Anishkin A, Gendel V, Sharifi NA, Chiang C-S, Shirinian L, Guy HR, Sukharev S (2003) On the conformation of the COOH-terminal domain of the large mechanosensitive channel MscL. *J Gen Physiol* 121:227–244
- Anishkin A, Chiang C-S, Sukharev S (2005) Gain-of-function mutations reveal expanded intermediate state and a sequential action of two gates in MscL. *J Gen Physiol* 125:155–170
- Bartlett JL, Levin G, Blount P (2004) An in vivo assay identifies changes in residue accessibility on mechanosensitive channel gating. *Proc Natl Acad Sci USA* 101:10161–10165
- Bass RB, Strop P, Barclay M, Rees DC (2002) Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science* 298:1582–1587
- Batiza AF, Kuo MM-C, Yoshimura K, Kung C (2002) Gating the bacterial mechanosensitive channel MscL in vivo. *Proc Natl Acad Sci USA* 99:5643–5648
- Berrier C, Coulombe A, Szabo I, Zoratti M, Ghazi A (1992) Gadolinium ion inhibits loss of metabolites induced by osmotic shock and large stretch-activated channels in bacteria. *Eur J Biochem* 206:559–565
- Blount P, Sukharev S, Schroeder MJ, Nagle SK, 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
- Blount P, Schroeder M, Kung C (1997) Mutations in a bacterial mechanosensitive channel change the cellular response to osmotic stress. *J Biol Chem* 272:32150–32157
- Booth IR, Edwards MD, Black S, Schumann U, Miller S (2007) Mechanosensitive channels in bacteria: signs of closure? *Nature Rev* 5:431–440
- Chang GRH, Spencer AT, Lee M, Barclay T, Rees DC (1998) Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science* 282:2220–2226
- Chiang C-S, Shirinian L, Sukharev S (2005) Capping transmembrane helices of MscL with aromatic residues changes channel response to membrane stretch. *Biochemistry* 44:12589–12597
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190
- Edwards MD, Li Y, Kim S, Miller S, Bartlett W, Black S, Dennison S, Iscla I, Blount P, Bowie JU, Booth IR (2005) Pivotal role of the glycine-rich TM3 helix in gating the MscS mechanosensitive channel. *Nature Struct and Mol Biol* 12:113–119
- Edwards MD, Bartlett W, Booth IR (2008) Pore mutations of the *Escherichia coli* MscS channel affect desensitization but not ionic preference. *Biophys J* 94:3003–3013
- Elmore DE, Dougherty DA (2001) Molecular dynamics simulations of wild-type and mutant forms of the *Mycobacterium tuberculosis* MscL channel. *Biophys J* 81:1345–1359
- Epstein W (1986) Osmoregulation by potassium transport in *Escherichia coli*. *FEMS Microbiol Lett* 39:73–78
- Guda P, Bourne PE, Guda C (2007) Conserved motifs in voltage-sensing and pore-forming modules of voltage-gated ion channel proteins. *Biochem and Biophys Res Comm* 352:292–298
- Gullingsrud J, Schulten K (2003) Gating of MscL studied by steered molecular dynamics. *Biophys J* 85:2087–2099
- Hamill OP, Martinac B (2001) Molecular basis of mechanotransduction in living cells. *Physiol Rev* 81:685–740
- Hase CC, Le Dain AC, Martinac B (1997) Molecular dissection of the large mechanosensitive ion channel (MscL) of *E coli*: mutants with altered channel gating and pressure sensitivity. *J Membr Biol* 157:17–25
- Haswell ES, Meyerowitz EM (2006) MscS-like proteins control plasmid size and shape in *Arabidopsis thaliana*. *Current Biol* 16:1–11
- Hohmann S (2002) Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol Mol Biol Rev* 66:300–372
- Iscla I, Wray R, Blount P (2008) On the structure of the N-terminal domain of the MscL channel: helical bundle or membrane interface. *Biophys J* 95:2283–2291
- Kloda A, Martinac B (2001) Mechanosensitive channels in Archaea. *Cell Biochem Biophys* 34:349–381
- Kloda A, Martinac B (2002) Common evolutionary origins of mechanosensitive ion channels in Archaea, Bacteria and cell-walled Eukarya. *Archaea* 1:35–44
- Kloda A, Ghazi A, Martinac B (2006) C-terminal charged cluster of MscL, RKKEE, functions as a pH sensor. *Biophys J* 90:1992–1998
- Koprowski P, Kubalski A (2003) C termini of the *Escherichia coli* mechanosensitive ion channel (MscS) move apart upon the channel opening. *J Biol Chem* 278:11237–11245
- Kung C (2005) A possible unifying principle for mechanosensation. *Nature* 436:647–654
- Levin G, Blount P (2004) Cysteine scanning of MscL transmembrane domains reveals residues critical for mechanosensitive channel gating. *Biophys J* 86:2862–2870
- Levina N, Töttemeyer S, Stokes NR, Louis P, Jones MA, Booth IR (1999) Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J* 18:1730–1737
- Martinac B, Kloda A (2003) Evolutionary origins of mechanosensitive ion channels. *Prog Biophys Mol Biol* 82:11–24
- Martinac B, Buechner M, Delcour AH, Adler J, Kung C (1987) Pressure-sensitive ion channel in *Escherichia coli*. *Proc Natl Acad Sci USA* 84:2297–2301
- Martinac B, Adler J, Kung C (1990) Mechanosensitive ion channels of *E. coli* activated by amphipaths. *Nature* 348:261–263
- Martinac B, Saimi Y, Kung C (2008) Ion channels in microbes. *Physiol Rev* 88:1449–1490
- Maurer JA, Dougherty DA (2001) A high-throughput screen for MscL channel activity and mutational phenotyping. *Biochim Biophys Acta* 1514:165–169
- Maurer JA, Dougherty DA (2003) Generation and evaluation of a large mutational library from the *Escherichia coli* mechanosensitive channel of large conductance, MscL: implications for the channel gating and evolutionary design. *J Biol Chem* 278:21076–21082
- Maurer JA, Elmore DE, Lester HA, Dougherty DA (2000) Comparing and contrasting *Escherichia coli* and *Mycobacterium tuberculosis* mechanosensitive channels (MscL). New gain of function mutations in the loop region. *J Biol Chem* 275:22238–22244
- Maurer JA, Elmore DE, Clayton D, Xiong L, Lester HA, Dougherty DA (2008) Confirming the revised C-terminal domain of the MscL crystal structure. *Biophys J* 94:4662–4667
- Miller S, Bartlett W, Chandrasekaran S, Simpson S, Edwards M, Booth IR (2003a) Domain organization of the MscS mechanosensitive channel of *Escherichia coli*. *EMBO J* 22:36–46

- Miller S, Edwards MD, Ozdemir C, Booth IR (2003b) The closed structure of the MscS mechanosensitive channel. Cross-linking of single cysteine mutants. *J Biol Chem* 278:32246–32250
- Moe PC, Blount P, Kung C (1998) Functional and structural conservation in the mechanosensitive channel MscL implicates elements crucial for mechanosensation. *Mol Microbiol* 28:583–592
- Moe PC, Levin G, Blount P (2000) Correlating a protein structure with function of a bacterial mechanosensitive channel. *J Biol Chem* 275:31121–31127
- Nakayama Y, Fujii K, Sokabe M, Yoshimura K (2007) Molecular and electrophysiological characterization of a mechanosensitive channel expressed in the chloroplasts of *Chlamydomonas*. *Proc Natl Acad Sci USA* 104:5883–5888
- Nomura T, Sokabe M, Yoshimura K (2008) Interaction between the cytoplasmic and transmembrane domains of the mechanosensitive channel MscS. *Biophys J* 94:1638–1645
- Oakley AJ, Martinac B, Wilce MCJ (1999) Structure and function of the bacterial mechanosensitive channel of large conductance. *Protein Sci* 8:1915–1921
- Ou X, Blount P, Hoffman RJ, Kung C (1998) One face of a transmembrane helix is crucial in mechanosensitive channel gating. *Proc Natl Acad Sci USA* 95:11471–11475
- Perozo E (2006) Gating prokaryotic mechanosensitive channels. *Nature Rev* 7:109–119
- Pivetti CD, Yen M-R, Miller S, Busch W, Tseng Y-H, Booth IR, Saier MH (2003) Two families of mechanosensitive channel proteins. *Microbiol Mol Biol Rev* 67:66–85
- Powl AM, East JM, Lee AG (2003) Lipid-protein interactions studied by interactions studied by introduction of a tryptophan residue: the mechanosensitive channel MscL. *Biochemistry* 42:14306–14317
- Powl AM, East JM, Lee AG (2005) Heterogeneity in the binding of lipid molecules to the surface of a membrane protein: hot spots for anionic lipids on the mechanosensitive channel of large conductance MscL and effects on conformation. *Biochemistry* 44:5873–5883
- Powl AM, East JM, Lee AG (2008a) Anionic phospholipids affect the rate and extent of flux through the mechanosensitive channel of large conductance MscL. *Biochemistry* 47:4317–4328
- Powl AM, East JM, Lee AG (2008b) Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. *Biochemistry* 47:12175–12184
- Rasmussen A, Rasmussen T, Edwards MD, Schauer D, Schumann U, Miller S, Booth IR (2007) The role of tryptophan residues in the function and stability of the mechanosensitive channel MscS from *Escherichia coli*. *Biochemistry* 46:10899–10908
- Shapovalov G, Bass R, Rees DC, Lester HA (2003) Open-state disulfide crosslinking between *Mycobacterium tuberculosis* mechanosensitive channel subunits. *Biophys J* 84:2357–2365
- Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* 3:265–274
- Spencer RH, Chang G, Rees DC (1999) ‘Feeling the pressure’: structural insights into a gated mechanosensitive channel. *Curr Opin Struct Biol* 9:448–454
- Steinbacher S, Bass R, Strop P, Rees DC (2007) Structures of the prokaryotic mechanosensitive channels MscL and MscS. *Curr Topics Membr* 58:1–24
- Sukharev S, Blount P, Martinac M, Blattner FR, Kung C (1994) A large-conductance mechanosensitive channel in *E. coli* encoded by mscL alone. *Nature* 368:265–268
- Sukharev S, Sigurdson WJ, Kung C, Sachs F (1999) Energetic and spatial parameters for gating of the bacterial large conductance mechanosensitive channel. *MscL J Gen Physiol* 113:525–540
- Sukharev S, Betanzos M, Chiang CS, Guy HR (2001a) The gating mechanism of the large mechanosensitive channel MscL. *Nature* 409:720–724
- Sukharev S, Durell SR, Guy HR (2001b) Structural models of the MscL gating mechanism. *Biophys J* 81:917–936
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Vásquez V, Sotomayor M, Cordero-Morales J, Schulten K, Perozo E (2008) A structural mechanism for MscS gating in lipid bilayers. *Science* 321:1210–1214
- Wang W, Black SS, Edwards MD, Miller S, Morrison EL, Bartlett W, Dong C, Naismith JH, Booth IR (2008) The structure of an open form of an *E. coli* mechanosensitive channel at 3.45 Å resolution. *Science* 321:1179–1183
- Yefimov S, Giessen E, Onck PR, Marrink SJ (2008) Mechanosensitive membrane channels in action. *Biophys J* 94:2994–3002
- Yoshimura K, Usukura J, Sokabe M (2008) Gating-associated conformational changes in the mechanosensitive channel MscL. *Proc Natl Acad Sci USA* 105:4033–4038