

A reported archaeal mechanosensitive channel is a structural homolog of MarR-like transcriptional regulators

Zhenfeng Liu,^{1,2} Troy A. Walton,^{1,2} and Douglas C. Rees^{1,2*}

¹Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

²Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

Received 27 December 2009; Revised 23 January 2010; Accepted 25 January 2010

DOI: 10.1002/pro.360

Published online 16 February 2010 proteinscience.org

Abstract: Several archaeal mechanosensitive (MS) channels have been reported, including one from *Thermoplasma volcanium* designated MscTV. Here, we report the crystal structure of MscTV at 1.6-Å resolution. Unexpectedly, MscTV was found to be a water-soluble protein exhibiting a winged helix-turn-helix (wHTH) motif, which is the signature of the MarR (multiple antibiotic resistance regulator) family of transcriptional regulators. A cell-based osmotic downshock functional assay demonstrated that MscTV was unable to protect a knockout strain of *Escherichia coli* from hypoosmotic shock, further indicating that it does not function as a MS channel. We propose this protein be renamed MLPTv for MarR-like protein from *T. volcanium*.

Keywords: mechanosensitive channel; multiple antibiotic resistance regulator; transcriptional regulator; winged helix-turn-helix motif

Introduction

Mechanosensitive (MS) channels in prokaryotes have been implicated in the response to osmotic downshock conditions. The study of MS channels has been extended to archaea since their discovery in *Haloferax volcanii*.¹ Two MscS (small-conductance MS channel) homologs from *Methanococcus jannaschii* have been identified and functionally characterized.^{2,3} Two more MS channels were identified in *Thermoplasma volcanium* and *Thermoplasma acidophilum*,⁴ through a functional approach analogous to the one leading to the original identification of the *E. coli* MS channel of large-conductance MscL.⁵ They have been named MscTV and MscTA, respectively. Both proteins were reported to exhibit electrophysiological properties typical of MS channels.

As part of our interest in the structural analysis of MS channels, we determined the crystal structure of MscTV at 1.6-Å resolution. Instead of a membrane protein, however, this protein is water soluble, and the four subunits in the crystallographic asymmetric unit do not generate a symmetric oligomer larger than a dimer. Moreover, MscTV fails to rescue the triple knockout strain (*mscL*⁻, *mscS*⁻, and *mscK*⁻) of *E. coli* during osmotic downshock. A search for homologous structures in the Protein Data Bank (PDB) revealed that MscTV adopts the fold characteristic of the multiple antibiotic resistance regulator (MarR) family of winged helix-turn-helix (wHTH) transcriptional factors. Here, we rename this protein as MLPTv for MarR-like protein from *T. volcanium*.

Results and Discussion

Overall structure

The structure of MLPTv was initially solved at 2.5-Å resolution by the single isomorphous replacement with anomalous scattering (SIRAS) method and subsequently refined to 1.6-Å resolution. The statistics of diffraction data processing, phasing, and structure refinement are listed in Table I. There are four

Grant sponsor: National Institutes of Health; Grant number: GM084211; Grant sponsor: Howard Hughes Medical Institute.

*Correspondence to: Douglas C. Rees, Division of Chemistry and Chemical Engineering, Mail Code 114-96, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125. E-mail: dcrees@caltech.edu

Table I. Data Processing, Phasing, and Structure Refinement Statistics

Data processing and phasing statistics			
Data set	Native 1	KI	Native 2
Unit cell (space group P2 ₁ 2 ₁ 2 ₁) (Å)	<i>a</i> = 65.16 <i>b</i> = 88.46 <i>c</i> = 115.35	<i>a</i> = 64.94 <i>b</i> = 88.65 <i>c</i> = 115.61	<i>a</i> = 65.13 <i>b</i> = 88.34 <i>c</i> = 115.21
Wavelength (Å)	1.5418	1.5418	1.0000
Resolution (Å)	50–2.5	50–2.07	50–1.6
Unique reflections	23,672	41,575	89,296
Redundancy	5.2	17.8	8.0
Completeness (%)	99.5 (99.2)	99.5 (95.6)	100.0 (100.0)
^a <i>R</i> _{merge}	0.043 (0.103)	0.090 (0.534)	0.041 (0.497)
< <i>I</i> / <i>σ</i> (<i>I</i>)>	32.1 (14.5)	25.6 (2.4)	45.8 (4.4)
^b FOM (50–2.5 Å)	0.699/0.783 (centric/acentric)		
Refinement statistics			
Resolution (Å)	50–1.6		
Reflections (working set)	80,995		
Reflections (test set)	4300		
^c <i>R</i> _{work} / <i>R</i> _{free}	0.211/0.228		
No. atoms	4231		
Protein	3648		
Ligand/ion	33		
Water	550		
Average <i>B</i> factor (Å ²)	33.2		
Protein	31.9		
Ligand/ion	38.4		
Water	41.30		
Rmsd bond length (Å)	0.011		
Rmsd bond angle (°)	1.42		

^a $R_{\text{merge}} = \sum_j \sum_h |I_{j,h} - \langle I_h \rangle| / \sum_j \sum_h I_{j,h}$, where *h* is unique reflection indices, *I*_{*j,h*} is intensities of symmetry-related reflections, and $\langle I_h \rangle$ is the mean intensity.

^b FOM: figure of merit, determined by SIGMAA program⁶ in CCP4 suite.⁷

^c $R_{\text{work}} = \sum_{hkl} |F_{\text{obs}}| - |F_{\text{calc}}| / \sum_{hkl} |F_{\text{obs}}|$, calculated for the working set reflections, whereas *R*_{free} was calculated for the reflections in the test set (5% of total reflections).

monomers of MLPTv in the asymmetric unit (A, B, C, and D) that were continuously traced from amino acid residues 1–120 (plus two additional residues at the N-terminus from the his-tag), 5–116, 4–114, and 4–113, respectively, of 126 residues total. They form an irregular assembly [Fig. 1(A)] where the largest symmetric oligomer, taking into account the crystallographic symmetry, is a dimer. Each monomer starts with an irregularly structured region from Met1 to Ser32, followed by a globular domain from Gln33 to Thr90 and a long α helix (α 4) from Ser90 to Lys113 near the carboxyl terminus. The globular domain contains three short α helices (α 1– α 3) and a β sheet with three antiparallel strands (β 1– β 3) in the order α 1– β 1– α 2– α 3– β 2–W1(wing)– β 3 [Fig. 1(C)]. Helices α 1, α 2, and α 3 cover residue ranges 33–44, 51–57, and 63–73, respectively. The β 1, β 2, and β 3 consist of residues 48–50, 76–81, and 84–89, respectively. W1 is the loop between β 2 and β 3, making a β -hairpin motif structure. Each subunit superimposes well with the others, yielding a 0.64-Å average root mean square deviation (RMSD) in α carbon atom positions for residues 5–113.

Unlike a typical integral membrane protein with a wide layer of hydrophobic surface inaccessible

to water, MLPTv has a hydrophilic surface that is evenly hydrated by water molecules [Fig. 1(B)], suggesting that it is a water-soluble protein rather than an integral membrane protein. This is further supported by the observation that MLPTv could be purified in a detergent-free buffer from the supernatant fraction of the ultracentrifuged cell lysate and by the results of the TMHMM hydrophobicity analysis⁸ that did not identify any potential membrane spanning helices for MLPTv (not shown). Consistent with this interpretation, a size exclusion chromatography–multiangle light scattering analysis (SEC-MALS) of MLPTv demonstrates that the majority of protein is monomeric in aqueous solution in the absence of detergents (Fig. 2).

In vivo osmotic downshock assay

The ability of MLPTv to function as a MS channel was evaluated using an *in vivo* assay to assess whether or not overexpression of this protein can protect the MJF465 triple knockout strain (*mscL*[−], *mscS*[−], and *mscK*[−]) of *E. coli* cells against osmotic downshock. The result clearly indicates that MLPTv is not capable of rescuing the knockout cell from the osmotic downshock, whereas the EcMscL does show

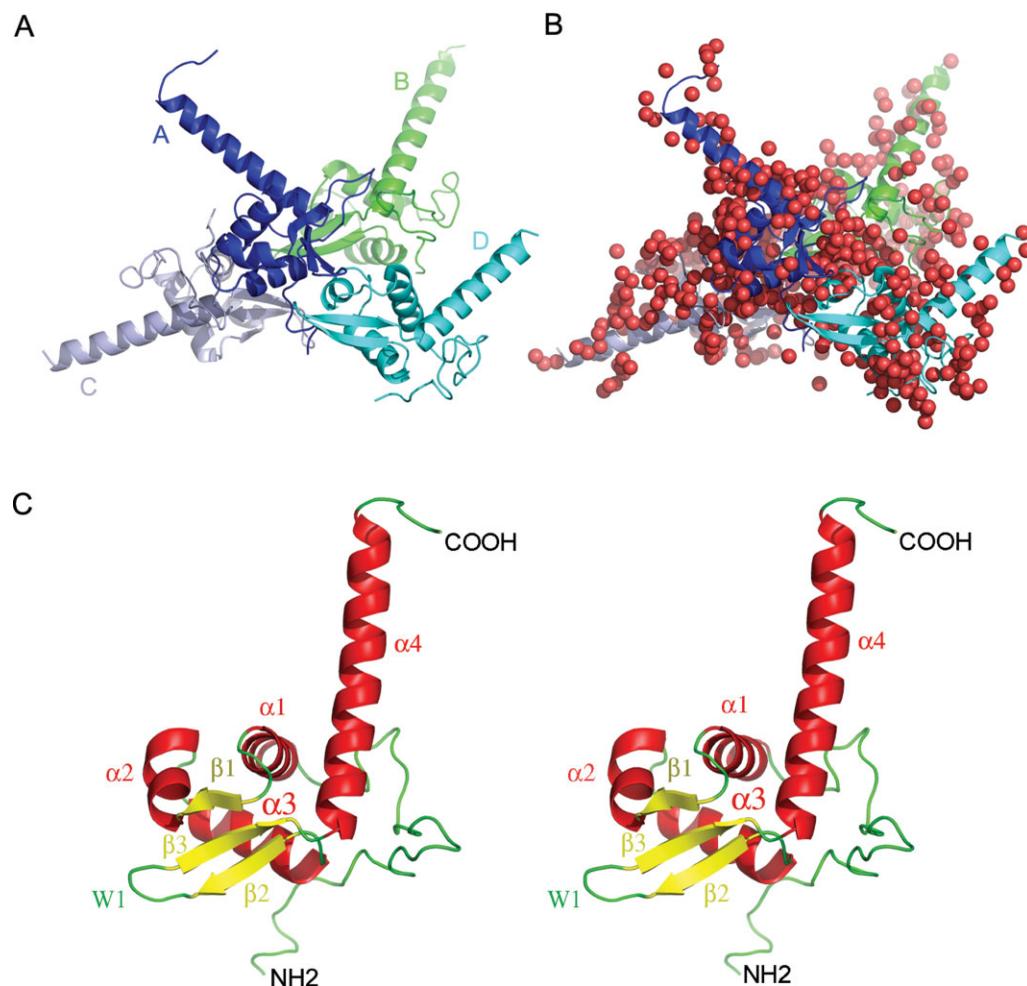


Figure 1. Structure of MLPTv (A) Overall structure of MLPTv, illustrating the four molecules in the crystallographic asymmetric unit. (B) The same view with water molecules shown as red spheres. (C) Stereoview of MLPTv monomer with α helices in red, β strands in yellow, and loops in green.

a protective effect in this assay. (Fig. 3). The expression of MLPTv or EcMscL in *E. coli* is not lethal to the cell as shown by the mock-shock data. When expressed in *E. coli*, MLPTv does not exhibit the “safety valve” function of the MS channels.

A structural homolog of MarR-like transcriptional regulators

The only homolog of MLPTv identified through sequence comparisons using NCBI-Blast⁹ in all databases presently available is CAC11636 (Genbank code) from *Thermoplasma acidophilum*, with *E* value of $3e^{-34}$ and 78% sequence identity. This protein has the same sequence as the 17–126 region of MscTA reported previously.⁴ Searching through the Dali server¹⁰ in the PDB with the monomeric structure of MLPTv yielded more than 100 structures with *Z* scores over 8.0, most of which belong to the MarR family of transcriptional factors. Two representative solutions (1JGS/1Z9C, *Z* score = 10.7/9.5) aligned well to the structure of MLPTv monomer [Fig. 4(A,B)] with 1.6 and 1.6 Å RMSD for 86 and 74

matched residues, respectively. These two proteins are well-studied members of the MarR family, namely MarR from *E. coli*¹¹ and the organic hydroperoxide resistance protein regulator (OhrR) from *Bacillus subtilis* in complex with *ohrA* operator¹² [The DNA has been omitted in Fig. 4(A)], respectively. As shown in Figure 4, the α 1- β 1- α 2- α 3- β 2-W1(wing)- β 3 motif, previously defined as the winged helix DNA-binding motif,¹³ superimposes particularly well among the three structures. The α 4 of OhrR deviates from the other two because of the presence of a kink in the middle. Although the percent sequence identity between MLPTv and MarR or between MLPTv and OhrR is low (12 or 18%, respectively), there are several conserved regions among them that are mainly located within the wHTH domain. Relative to MLPTv, MarR and OhrR each has two additional helices, one at the N-terminus (α 0) and the other at the C-terminus (α 5), respectively. These two helices mediate the formation of functional dimers of MarR and OhrR, which bind the operator DNA (with twofold symmetry-related

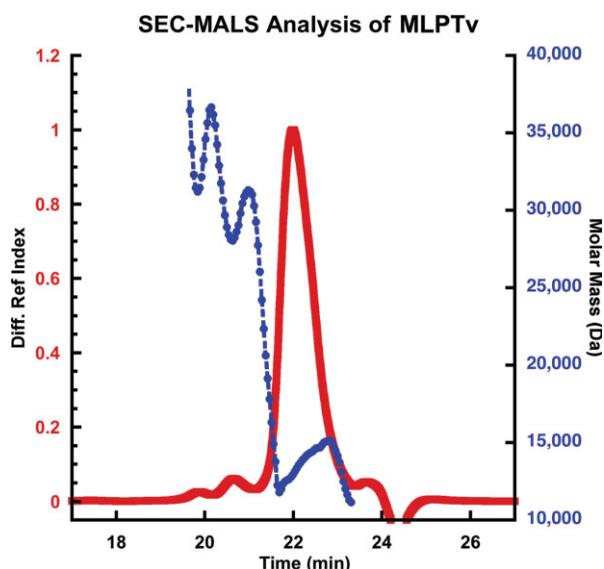


Figure 2. SEC-MALS of MLPTv. The average molar mass for the major peak fraction was calculated to be 13.5 kDa, indicating the dominance of monomer in solution.

sequence) on two sides.¹² In MLPTv, the amino- and carboxy-terminal regions form irregular structures instead. The one symmetric dimer of MLPTv observed in this crystal structure, formed between monomer B and a crystallographic symmetry-related molecule of monomer C (C'), does not superimpose well with the OhrR dimer in complex with *ohrA* operator DNA.¹² Curiously, a molecule of citric acid is bound at the BC' interface, stabilizing the dimerization by forming hydrogen bonds and salt bridges with residues from monomers B (Leu40, Thr44) and C' (Arg101, Asp104, Arg108).

Conclusions

The crystal structure of an archaeal protein designated MLPTv, previously known as MscTV and characterized as a MS channel from *T. volcanium*, has been solved at 1.6-Å resolution. The structural and biochemical data indicate that this is a water-soluble protein, rather than an integral membrane protein, with a polypeptide fold belonging to the MarR family of transcriptional factors. During purification, we have also noticed that a fraction of MLPTv protein does attach to the membrane surface, presumably through the interaction of its surface basic residues with the phosphate head group of the phospholipids. In the original report,⁴ this protein might have been enriched in the native *T. volcanium* membrane for the same reason and copurified with a MS channel activity, leading to the identification of it as the one responsible for the activity. The biological function of MLPTv remains to be established.

Materials and Methods

Cloning, expression, and purification

The sequences of MLPTv and MLPTa (formerly MscTV and MscTA, respectively) are available under Genbank accession codes of BAB59904 and CAC11636. The genes were cloned from the genomic DNA and inserted into pET15b vector between Nde I and BamH I sites. Consequently, the sequence MGSHHHHHHSSGLVPRGSH is fused to the N-terminus of each protein, so that the complete sequence expressed for MLPTv is MGSHHHHHHSSGLVPRGSHMSAMAESKVLVKGTPFNKPKVIKGLKENNYDM SQDEVSLLLFLKTHGGKIPLYRIKNETGLKDPES VLKNLMDYGFALDKERLGEKIVLTSEGEFVAQ AIRVRDEELRLKEMKQKKNVNRSSAPPQ (where the true N-terminal residue is underlined, and the C-terminal residue is Gln126). The plasmid was transformed into BL21-Gold (DE3) competent cell (Stratagene) for protein expression. When the OD_{600nm} reached 2.0 in Terrific Broth media, the expression was induced with 1 mM IPTG for 2 h. The cells were harvested by centrifugation and

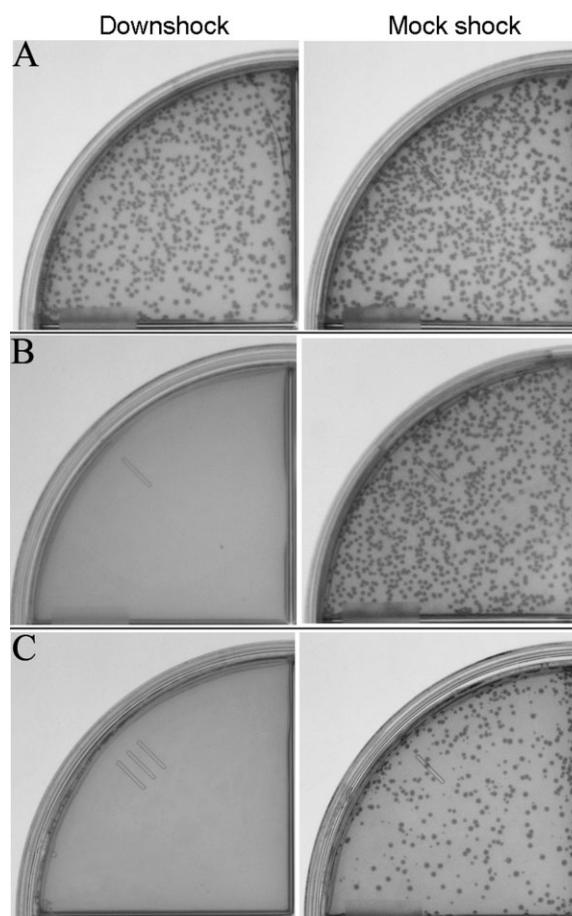


Figure 3. Osmotic downshock assay. (A, B, and C) Plates of cells hosting EcMscL, MLPTv, and empty vector, respectively, challenged with 0.5M osmotic downshock (left) or mock shock (right).

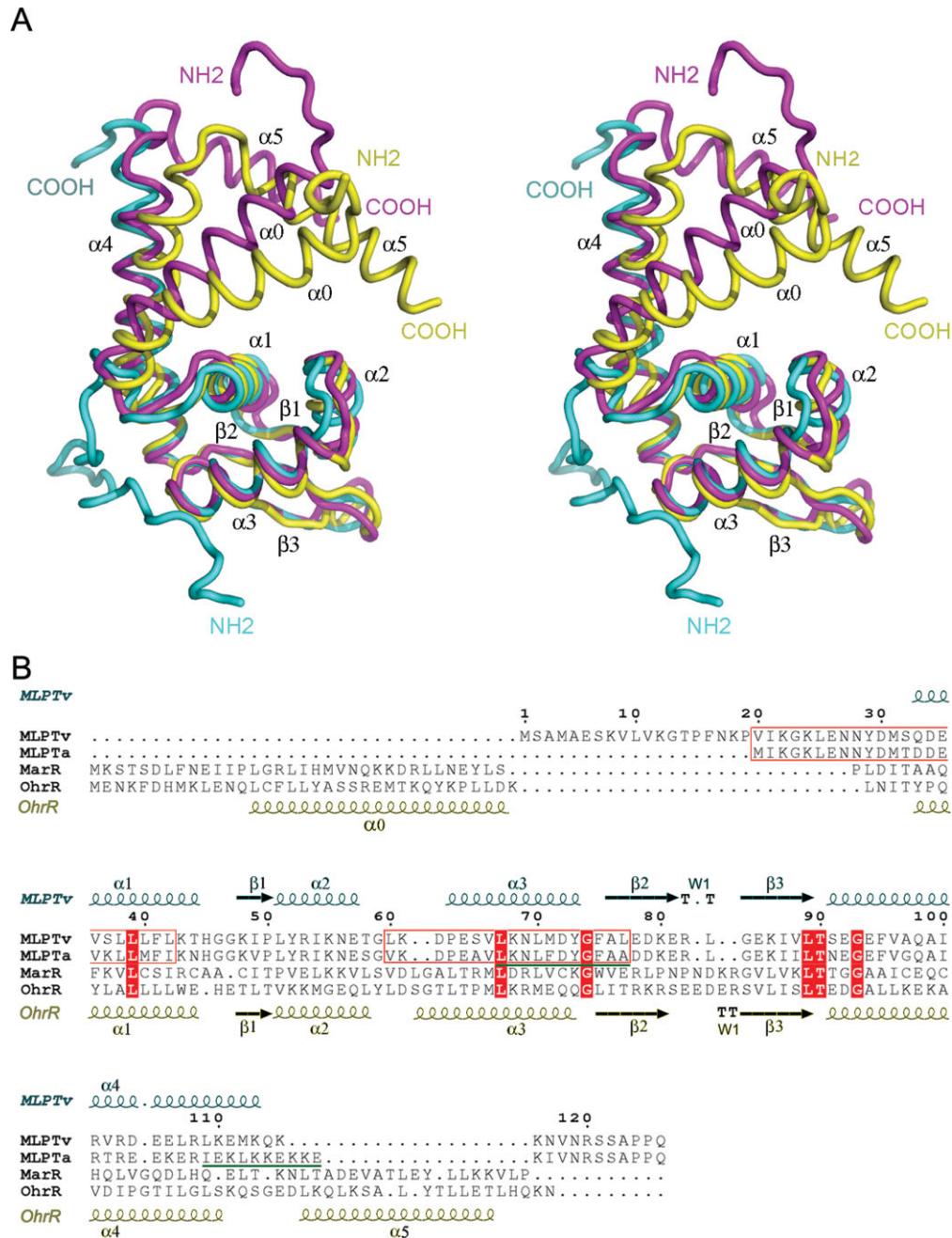


Figure 4. Alignment of MLPTv, MarR, and OhrR on their three-dimensional structures (A) and amino acid sequences (B), respectively. Cyan, MLPTv; magenta, MarR; Yellow, OhrR generated from PDB codes 3LFLK, 1JGS and 1Z9C, respectively. Stereoview image is presented in (A). The red-line framed area of MLPTv and MLPTa (formerly MscTA) sequences in (B) highlights the putative transmembrane helical regions predicted in the previous report,⁴ and the green line-underscored segments in MLPTa were indicated to be homologous to IQNVDFLIVA and INKLNRRKKEE in EcMscL, respectively.⁴

resuspended in lysis buffer (50 mM Tris-HCl, 200 mM NaCl, pH 8.0). Initially, this protein was treated as a membrane protein and purified in the presence of detergent dodecyl- β -D-maltopyranoside (DDM). Later, it was found that MLPTv is present in both the supernatant and membrane fraction of the ultracentrifuged lysate (100,000g). The protein in the membrane fraction can be readily stripped off the membrane by washing with either 200 mM NaCl or 4 M urea, indicating that it is not embedded in the

membrane but rather attached to the surface. Both the membrane and soluble fractions yielded pure MLPTv protein with similar quality and crystallizability. The protein was purified by affinity chromatography on a Ni-NTA column, utilizing an initial washing step with 75 mM imidazole, followed by elution with 300 mM imidazole. To remove the his-tag, thrombin was added to the eluate at 3 U/mg protein, and the reaction was dialyzed in 10 kD cutoff Slide-A-Lyzer cassette (Pierce) against 3 L reservoir buffer

(20 mM Tris-HCl, 150 mM NaCl, 2.5 mM CaCl₂, 0.02% DDM, pH 8.4) at room temperature for 42 h. After thrombin cleavage, the tripeptide GSH remains fused to the true N-terminus of the protein. The dialyzed sample was concentrated in 10 kD cut-off Amicon Ultra-4 concentrator (Millipore) to ~10 mg/mL and further purified by gel filtration through Superdex 200 10/30 HR column (GE Healthcare) in a buffer with 10 mM Tris-HCl, 150 mM NaCl, and 0.02% DDM. The major peak eluting around 16–17 mL was pooled and concentrated to ~10 mg/mL for crystallization. For the purification of the soluble fraction, DDM was omitted from all buffers. The molecular weight of the major species was determined by MALDI-TOF mass spectrometry to be 13991.56, consistent with the predicted M_w (13990.2) of residues 4–126 of MLPTv, while a peak of ~1/4 the height had a molecular weight of 14556.79, consistent with the full length sequence, including the three residues GSH remaining from the thrombin cleavage (predicted M_w = 14560.8).

Size exclusion chromatography and multiangle laser light scattering analysis

Size exclusion chromatography was carried out on a Shodex KW-803 column at 0.5 mL/min in 10 mM Tris-HCl, 150 mM NaCl, pH 7.5. Data were collected using Wyatt technologies HELEOS multiangle light scattering detector followed by a rEX differential refractive index detector. Calculation of molecular weights from the data was accomplished with Astra software (version 5.3.4.14). Protein samples were diluted in running buffer before injection.

Crystallization, data collection, and structure determination

MLPTv was crystallized using the hanging drop vapor diffusion method. One microliter of 10 mg/mL MLPTv was mixed with 1 μ L reservoir solution (22–26% PEG 550MME, 80 mM ammonium sulfate, 85 mM citric acid, pH 5.0) and then equilibrated against 1 mL reservoir solution at 4°C. Plate- or rod-shaped crystals appeared within a week and grew to a full size of 0.3–0.4 mm (the largest dimension) in a month or two. Three different preparations have been used for crystallization, namely a preparation with both soluble and membrane-attached protein fractions mixed together (originating from whole-cell lysate with detergent added) and the other two preparations isolated from soluble or membrane fractions separately. The protein sample purified from the soluble fraction tended to yield rod-shaped or chunky block crystals, whereas the mixed preparation and membrane fraction mostly led to plate forms and occasionally yielded rod crystals. All three preparations yielded isomorphous crystals with less than 0.5% variation on the cell dimensions, despite the difference in morphology. The data sets listed in

Table I are all collected from crystals grown out of soluble fraction as it yielded high-quality crystals more readily than the mixed preparation or membrane fraction. For data collection, crystals were soaked in a solution with 28% PEG 550MME, 200 mM ammonium sulfate, 0.02% DDM, 10% glycerol, and 85 mM citric acid (pH 5.0) and then frozen in liquid nitrogen. A native data set to 2.5-Å resolution was collected on a laboratory RaxisIV ++ system with a rotating anode generator (Rigaku). For phasing, one single crystal was soaked in the presence of 0.5 M KI for 1 min, and a 2.07-Å resolution data set was collected in house. The initial phases were determined to 2.5-Å resolution using the program Shelx CDE¹⁴ with the Native 1 and KI anomalous datasets. The starting model was built by ARP/wARP¹⁵ using the 2.5-Å resolution SIRAS phases and the 2.07-Å resolution KI data. The model was further manually adjusted and corrected in program O.¹⁶ The structure was first refined with CNS v1.2 (Ref. 17) against the 2.07-Å KI data to $R_{\text{work}} = 24.9\%$ and $R_{\text{free}} = 28.2\%$. After the high-resolution Native 2 data were collected at Stanford Synchrotron Radiation Lightsource (SSRL), the structure refinement was extended to 1.6-Å resolution with the final $R_{\text{work}} = 21.1\%/R_{\text{free}} = 22.8\%$. The data collection, phasing, and structure refinement statistics are summarized in Table I. Among all residues, 92.2 and 7.8% are the most favored and additional allowed regions in the Ramachandran plot reported by Procheck.¹⁸ No residues are within the generously allowed or disallowed regions. The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 3LFK.

The structural figures were produced using PyMol, DeLano Scientific (<http://www.pymol.org>). The hydrophathy plots were produced by TMHMM program.⁸ DaliLite v.3 server¹⁰ was used to search for structures in PDB with similarity to MLPTv. LSQMAN¹⁹ was used to superimpose structures. Sequence alignment image in Figure 4(B) was generated with ESPript program.²⁰

Osmotic downshock assay

MJF465 (*mscL*⁻, *mscS*⁻, and *mscK*⁻)²¹ (DE3) cells transformed with pET15b plasmids carrying genes of EcMscL (positive control), MLPTv, or empty vector (negative control) were grown on LB-ampicillin (Amp) plates overnight. The protocol for the assay was modified from Refs. 21 and 22. For the osmotic shock, the cells were challenged by 1:500 dilution into sterile double-distilled H₂O. For the mock shock, LB media with 500 mM NaCl were used instead of H₂O. The cells were allowed to recover at 37°C for 20 min, and then 20 μ L of the downshock/mock-shock culture was combined with 80 μ L LB-500 and plated onto LB-Amp plates, followed by overnight incubation at 37°C. One of the four

repeats was shown for each downshock/mock shock experiment.

Acknowledgments

The authors thank A. Shih for cloning work, Y. Poon and J. Lai for protocols of osmotic downshock assay and lysogenizing MJF465 cells with λ DE3, Y. Liu and M. Shahgholi for the mass spectrometry analysis, B. Martinac, H. Pinkett, and C. Gandhi for discussions, P. Blount for the MJF465 cells, and the staff at SSRL for technical support during data collection. The authors acknowledge the Gordon and Betty Moore Foundation for support of the Molecular Observatory at Caltech. Operations at SSRL are supported by the U.S. Department of Energy and NIH. D. C. R. is an Investigator in the Howard Hughes Medical Institute.

References

1. Le Dain AC, Saint N, Kloda A, Ghazi A, Martinac B (1998) Mechanosensitive ion channels of the archaeon *Haloflex volcanii*. *J Biol Chem* 273:12116–12119.
2. Kloda A, Martinac B (2001) Molecular identification of a mechanosensitive channel in archaea. *Biophys J* 80: 229–240.
3. Kloda A, Martinac B (2001) Structural and functional differences between two homologous mechanosensitive channels of *Methanococcus jannaschii*. *EMBO J* 20: 1888–1896.
4. Kloda A, Martinac B (2001) Mechanosensitive channel of Thermoplasma, the cell wall-less archaea. *Cell Biochem Biophys* 34:321–347.
5. Sukharev SI, Blount P, Martinac B, Blattner FR, Kung C (1994) A large-conductance mechanosensitive channel in *E. coli* encoded by *mscL* alone. *Nature* 368: 265–268.
6. Read R (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A* 42:140–149.
7. Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
8. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580.
9. Altschul SF, Madden TL, Schaffer AA, Zhang J, 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.
10. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24:2780–2781.
11. Alekshun MN, Levy SB, Mealy TR, Seaton BA, Head JF (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution. *Nat Struct Biol* 8:710–714.
12. Hong M, Fuangthong M, Helmann JD, Brennan RG (2005) Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. *Mol Cell* 20:131–141.
13. Gajiwala KS, Burley SK (2000) Winged helix proteins. *Curr Opin Struct Biol* 10:110–116.
14. Sheldrick G (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
15. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3:1171–1179.
16. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
17. Brunger AT (2007) Version 1.2 of the crystallography and NMR system. *Nat Protoc* 2:2728–2733.
18. Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
19. Kleywegt GJ (1996) Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr D* 52:842–857.
20. Gouet P, Courcelle E, Stuart DI, Metz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15:305–308.
21. Levina N, Totemeyer 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.
22. 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.