Structure-based analysis of CysZ-mediated cellular uptake of sulfate

Authors and Affiliations:

Zahra Assur Sanghai[1,2], Qun Liu[3], Oliver B. Clarke[2], Edgar Leal Pinto[4], Brian Kloss[5], Shantelle Tabuso[5], James Love[5], Marco Punta[6], Surajit Banerjee[7], Kanagalaghatta R. Rajashankar[7], Burkhard Rost[8], Diomedes Logothetis[4], Matthias Quick[9], Wayne A. Hendrickson[1,2,6] and Filippo Mancia[1*]

[1] Department of Physiology and Cellular Biophysics, [2] Department of Biochemistry and Molecular Biophysics, and [9] Department of Psychiatry, Center for Molecular Recognition, Columbia University, New York, NY 10032, USA.

[3] Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA.

[4] Department of Physiology and Biophysics, Virginia Commonwealth University School of Medicine, Richmond, VA 23298, USA.

[5] New York Structural Biology Center, New York, NY 10027, USA.

[6] Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK.

[7] Department of Chemistry and Chemical Biology, Cornell University, NE-CAT, Advanced Photon Source, Argonne, IL 60439, USA.

[8] Department of Informatics, Technical University of Munich, Munich 85748, Germany

* To whom correspondence should be addressed: E-mail fm123@cumc.columbia.edu (F.M.)

ABSTRACT

Sulfur, most abundantly found in the environment as sulfate (SO_4^{2-}), is an essential element in metabolites required by all living cells, including amino acids, co-factors and vitamins. Current understanding of the cellular delivery of SO_4^{2-} at the molecular level is limited however. CysZ has been described as a SO_4^{2-} permease, but its sequence family is without known structural precedent. Based on crystallographic structure information, SO_4^{2-} binding and uptake experiments in cells and proteoliposomes, and single-channel conductance measurements, we provide insight into the molecular mechanism of CysZ-mediated translocation of SO_4^{2-} across membranes. CysZ properties differ markedly from those of known transporters and ion channels. The structures display a hitherto unknown fold with dual topology, assembling in CysZ from *Pseudomonas denitrificans* as a trimer of antiparallel dimers in the membrane. CysZ structures from two other species recapitulate dimers from this assembly. Mutational studies highlight the functional relevance of conserved CysZ residues.

Sulfur has a central role in many cellular processes across all kingdoms of life. It is a vital component of several essential compounds, including the sulfur-containing amino acids cysteine and methionine, in prosthetic groups such as the Fe-S clusters, as well as vitamins and micronutrients such as biotin (vitamin H), thiamine (vitamin B1) and lipoic acid, and in coenzymes A and M (Barton, 2005). Whilst mammals obtain the majority of the necessary sulfur-containing metabolites directly from the diet, plants, fungi, and bacteria are able to assimilate and utilize sulfur from organic and inorganic sources (Barton, 2005). Sulfate (SO_4^{2-}) is the most abundant source of sulfur in the environment and its utilization is contingent upon its

46 entry into the cell (Kertesz, 2000). In certain fungi, and prokaryotes, once internalized, SO_4^{2-} is
47 first reduced to sulfite (SO_3^{2-}), and then further to sulfide (S^{2-}), a form that can be used by the
48 cell (Kredich, Hulanicka, & Hallquist, 1979) (Fig. 1-S1). In *Escherichia coli* and other gram-
49 negative bacteria, the culmination of the aforementioned sulfate assimilatory (also known as
50 reductive) pathway is the formation of cysteine by the addition of S^{2-} to O-acetylserine by
51 cysteine synthase, followed by the synthesis of methionine from homocysteine (Kredich, 1971)
52 (Fig. 1-S1).
53
54 In prokaryotes, the entry of SO_4^{2-} into the cell is mediated by four known families of dedicated
55 transport systems: the ABC sulfate transporter complexes SulT or CysTWA, the SulP family of
56 putative SLC13 sodium:sulfate or proton:sulfate symporters or SLC26 solute:sulfate exchangers,
57 the phosphate transporter-like CysP/PitA family, and the CysZ family classified as SO_4^{2-}
58 permeases (Aguilar-Barajas, Diaz-Perez, Ramirez-Diaz, Riveros-Rosas, & Cervantes, 2011;
59 Hryniewicz, Sirko, Palucha, Bock, & Hulanicka, 1990; Kertesz, 2001; Loughlin, Shelden,
60 Tierney, & Howitt, 2002; Mansilla & de Mendoza, 2000; Sirko, Zatyka, Sadowy, & Hulanicka,
61 1995). CysZ family members are 28-30 kDa bacterial inner-membrane proteins found
62 exclusively in prokaryotes with no apparent homology to any of the established channel or
63 transporter folds, and are scarcely studied in the literature (Zhang, Jiang, Nan, Almqvist, &
64 Huang, 2014). The *cysZ* gene owes its name to its presence in the cysteine biosynthesis regulon.
65 In two reports from thirty years ago, an *E. coli* K12 strain with a *cysZ* deletion showed a severe
66 impairment in its ability to accumulate SO_4^{2-} and was not viable in sulfate-free media without an
67 alternate sulfur source such as thiosulfate ($S_2O_3^{2-}$) (Britton et al., 1983; Parra, Britton, Castle,
68 Jones-Mortimer, & Kornberg, 1983). Recently, a third report studying the functional properties

69 of CysZ, concluded that the protein from *E. coli* functions as a high affinity, highly specific pH-
70 dependent SO_4^{2-} transporter, directly regulated by the toxic, assimilatory pathway intermediate,
71 SO_3^{2-} (Zhang et al., 2014).
72
73 To investigate the role of CysZ in cellular sulfate uptake at a molecular level, we have
74 undertaken an approach that combines structural and functional studies. To this end, we
75 determined the crystal structures of CysZ from three species, *Idiomarina loihiensis (Il; IlCysZ)*,
76 *Pseudomonas fragi (Pf; PfCysZ),* and *Pseudomonas denitrificans (Pd; PdCysZ)*, and
77 characterized CysZ function in purified form, in reconstituted proteoliposomes, in planar lipid
78 bilayers, and in cells. Combining the structural information from the three orthologs reveals that
79 CysZ features a novel protein fold that assembles as oligomers with a dual topology. This
80 arrangement can be understood as being derived from trimers of dimers akin to the hexameric
81 assembly captured in one of the structures. Interpreting the functional data in a structural context
82 has allowed us to formulate a mechanistic model for CysZ-mediated SO_4^{2-} translocation across
83 the bacterial cytoplasm membrane.
84
85 Both the structures and the functional properties of CysZ proteins are distinct from those of any
86 known membrane transporter or ion channel. Besides not resembling other transporter structures,
87 CysZ mediates sulfate uptake into cells or proteoliposomes without coupling to ion gradients,
88 partner proteins, or exogenous energy sources such as ATP. Besides differing from other ion
89 channel structures, CysZ generates sulfate currents across lipid bilayers with unusual
90 characteristics. These distinctive properties make CysZ appealing as a model system for studies
91 of biophysical principles of membrane protein biogenesis and transmembrane ion passage.

92

93 **RESULTS**

94 **Structure determination of CysZ**

95 Following a structural genomics approach aimed at crystallization for structural analysis, we
96 cloned and screened a total of 63 different bacterial homologs of CysZ for high-level expression
97 and stability in detergents (Love et al., 2010; Mancia & Love, 2011). Crystal structures were
98 determined for CysZ from three organisms. Chronologically, the structure of *Il*CysZ was the first
99 solved, to 2.3 Å resolution in space group *C2* by SAD, initially based on a single selenate ion
100 bound to the protein and subsequently also by selenomethione derivatization (SeMet) SAD, and
101 multi-crystal native SAD (Liu et al., 2012). The structure of *Pf*CysZ was solved second, to 3.5 Å
102 resolution, with crystals also belonging to space group *C2*. Although *Pf*CysZ and *Il*CysZ share
103 42% sequence identity, molecular replacement failed to find a convincing solution, and we
104 instead used SeMet-derivatized *Pf*CysZ to obtain phase information by multi-crystal SeMet
105 SAD. Third, we determined the structure of *Pd*CysZ, which crystallized in multiple forms
106 belonging to space groups $P6_3$, $P4_122$ and $P2_12_12_1$, revealing the same architecture and
107 oligomeric assembly each time (Fig. 1-S2b). *Pd*CysZ structures in the $P6_3$ and $P2_12_12_1$ lattices
108 each contain an entire hexamer in their asymmetric units, whereas a molecular diad coincides
109 with a crystallographic axis in the $P4_122$ lattice. We focused our analysis on the best of these
110 (3.4 Å resolution in $P6_3$). The location of the selenium sites was obtained from a SeMet SAD
111 data set, and the structure was solved by combining the resulting SAD phases with those from a
112 molecular replacement solution obtained by positioning the *Pf*CysZ model (75% sequence
113 identity) onto SeMet fiducials in the initial electron density map (Table 1, Fig. 1-S2a).

114

115 **The hexameric structure of *Pd*CysZ**

116 The refined structure of *Pd*CysZ comprises an entire hexamer of near-perfect D3 symmetry.
117 Antiparallel pairs of protomers arrange together as a trimer of dimers (Fig. 1a), with the three-
118 fold axis oriented perpendicular to the plane of the putative membrane and three two-fold axes
119 between dimers of the hexamer. Both the periplasmic and cytoplasmic faces of the hexamer are
120 essentially identical by symmetry, resulting in a dual-topology assembly for *Pd*CysZ. The
121 hexamer has a triangular face of equal sides measuring approximately 75 Å, with the
122 perpendicular span of about 65 Å. The interaction of the six protomers results in a total buried
123 surface area (Krissinel & Henrick, 2007) of 5,700 $Å^2$. A surface electrostatic representation
124 reveals a hydrophobic belt along the mid-section of the hexamer when viewed from its side,
125 outlining the orientation of CysZ in the lipid bilayer (Czodrowski, Dramburg, Sotriffer, & Klebe,
126 2006; Dolinsky et al., 2007; Dolinsky, Nielsen, McCammon, & Baker, 2004) (Fig. 1b). A surface
127 representation of the sequence conservation, calculated by analysis of multiple sequence
128 alignments (MSA) (Ashkenazy et al., 2016; Glaser et al., 2003) highlights the regions of
129 invariance in the sequence and in turn, the areas on the molecule that are most likely to have
130 structural and functional importance (Fig. 1c).

131

132 The CysZ protomer is an alpha-helical integral membrane protein with two long transmembrane
133 (TM) helices (H2b and H3a) and two pairs of shorter helices (H4b-H5a and H7-H8) that insert
134 only partially into the membrane (hemi-penetrating), forming a funnel or tripod-like shape within
135 the membrane (Fig. 2a, b). The protein has an extra-membranous hydrophilic 'head', comprising
136 an iris-like arrangement of the two short helices, H1 and H6, and kinked helices H3b, H4a, and
137 H5b. The amino and carboxyl termini are also located in this region (Fig. 2a, b). H4b and H5a lie

138 partly inserted in the membrane, with the turn between the two helices pointing in towards the
139 three-fold axis at the center of the hexamer. CysZ amino-acid sequences from different
140 organisms are very similar; for example, those of *Pd*CysZ and *E. coli* CysZ (*Ec*CysZ) are 40.5%
141 identical and those of *Ec*CysZ and our three structures have 30.0% of their residues exactly in
142 common (Fig. 2c). Helix boundaries are also essentially the same in the structures of *Pd*CysZ,
143 *Il*CysZ and *Pf*CysZ (Fig. 2c).
144

145 **The dimeric assembly of *Il*CysZ and *Pf*CysZ**

146 Unlike *Pd*CysZ, both *Il*CysZ and *Pf*CysZ crystallize as dimers (Fig. 3a, b), in agreement with
147 their different behavior in detergent-containing solution. Indeed, size-exclusion chromatography
148 runs of the three species of CysZ in the same buffer and detergent conditions, show mono-
149 disperse peaks eluting at 13.42 ml (*Il*CysZ), 13.95 ml (*Pf*CysZ) and 12.56 ml (*Pd*CysZ) (Fig. 3-
150 S1). This result is consistent with a different and smaller oligomeric state of *Il*CysZ and *Pf*CysZ
151 (similar retention volume) compared to *Pd*CysZ (eluting 1 ml ahead).
152

153 The protomers of the *Il*CysZ dimer are arranged in a head-to-tail antiparallel association, with
154 H4b-H5a protruding at an angle that is nearly parallel to the putative plane of the membrane
155 (Fig. 3a, 3-S2a, b). The *Pf*CysZ dimer is also arranged in an antiparallel orientation but with a
156 different dimer interface (Fig. 3b). In the *Pf*CysZ structure, H4b-H5a together form a narrower
157 angle with H2 and H3, and tuck-in closer to the rest of the molecule (Fig. 3b). The resulting
158 dumbbell-shaped *Pf*CysZ dimer is predicted, by OPM/PPM (Orientation of Proteins in
159 Membranes) (Lomize, Pogozheva, Joo, Mosberg, & Lomize, 2012), to lie in the membrane at a
160 31° tilt to the perpendicular, in agreement with the position of its central hydrophobic belt, as

161 revealed by surface-electrostatics calculations (Fig. 3- S2e, f). The dimer interfaces of *IlCysZ* and
162 *PfCysZ* bury 780 Å^2 and 1136 Å^2 of surface area respectively (Krissinel & Henrick, 2007)
163
164 The individual protomers of CysZ from all three species adopt the same topology and fold, and
165 superpose well with an overall pairwise root mean squared deviation (r.m.s.d) of ~2.5 Å (Fig.
166 3c). The greatest variation between the protomers of each structure is seen in the orientation of
167 H4b-H5a with respect to the TM helices, H2 and H3 (Fig. 3c), which seem to be the most
168 conformationally flexible with respect to the rest of the molecule.
169
170 Comparison of the structures of *Il*CysZ and *Pf*CysZ with that of *Pd*CysZ revealed that these two
171 distinctive dimeric structures are both represented in the hexameric one. Indeed, the *Il*CysZ
172 dimer (Fig. 3a) resembles the vertically arranged pair of protomers in the *Pd*CysZ hexamer (Fig.
173 3d) at each of the vertices of the triangular structure. On the other hand, the *Pf*CysZ dimer (Fig.
174 3b) resembles the transverse pair of protomers lying at a tilt, just as was predicted by OPM
175 (Lomize et al., 2012), along the side of the *Pd*CysZ hexamer when viewed from inside the plane
176 of the membrane (Fig. 3e). In essence, the *Pd*CysZ hexamer can be seen as a trimer of either
177 *Il*CysZ or *Pf*CysZ dimers (Fig. 3f).
178
179 **Dual-topology assembly of CysZ**
180 To validate the dual-topology assembly of CysZ observed in all our structures, we performed
181 disulfide crosslinking assays on engineered cysteine mutants of CysZ designed to capture the
182 antiparallel dimer, utilizing isolated membranes. Disulfide-trapping experiments of the
183 transverse dimer, performed by crosslinking a pair of mutants (L161C-A164C) on H5 of *Pf*CysZ,

184 as well as the corresponding pair in *Il*CysZ (V157C-Q163C) confirm the dimer interface
185 observed in the structures of *Pf*CysZ and *Il*CysZ, and, as a consequence, the dual-topology
186 assembly of the two proteins (Fig. 3-S2f, g, h).
187
188 To further validate the dual-topology orientation of CysZ in the membrane, we performed a
189 cysteine accessibility scan experiment, mapping residues expected to be located outside the lipid
190 bilayer, by fluorescence labeling of the thiol groups of various single cysteine mutants with a
191 membrane-impermeable dye. Membrane fractions isolated from recombinant cultures expressing
192 *Il*CysZ cysteine mutants introduced at positions along the edge of helix H4 predicted to be
193 solvent accessible based on our structure, were indeed labeled with a membrane-impermeable
194 fluorescent thiol-specific maleimide dye (Fig. 3-S2c, d).
195
196 **Functional characterization of CysZ**
197 To characterize the functional properties of CysZ, we used a four-pronged approach: (i)
198 radiolabeled $[^{35}S]O_4^{2-}$ uptake experiments in cells expressing plasmid-encoded CysZ, (ii)
199 radiolabeled ($[^{35}S]O_4^{2-}$) binding to purified CysZ in detergent solution, (iii) $[^{35}S]O_4^{2-}$ uptake in
200 proteoliposomes reconstituted with purified CysZ, and (iv) single-channel electrophysiological
201 recordings in a planar lipid bilayer reconstituted with CysZ.
202 First, we compared the time course of SO_4^{2-} accumulation in an *E. coli cysZ* knockout strain (*E.*
203 *coli* K12 JW2406-1, CysZ$^-$) (Baba et al., 2006) with that in the wild-type (WT, *E. coli* K12
204 BW25113, *cysZ*$^+$) strain. CysZ$^-$ cells, after growth in minimal media and 12 hours of sulfate
205 starvation, showed significantly diminished SO_4^{2-} uptake when compared to the WT strain,
206 consistent with previous results (Fig. 4a) (Parra et al., 1983). The CysZ$^-$ strain still showed a low

207 level of sulfate accumulation, which could be attributed to the other endogenous sulfate transport
208 systems present in the bacteria (for example ABC transporter and SulP). This uptake-deficient
209 phenotype of the CysZ$^-$ cells could be rescued by plasmid-driven expression of *Pd*CysZ (Fig.
210 4b), as well as *Il*CysZ and *Pf*CysZ (data not shown). Consistent with previously reported data
211 (Zhang et al., 2014), we observed that SO_3^{2-} ions severely impede or block the accumulation of
212 $[^{35}S]O_4^{2-}$ by CysZ-expressing cells (Fig. 4b).
213
214 To further characterize the interaction of CysZ with both SO_4^{2-} and SO_3^{2-}, we performed binding
215 experiments on purified *Pd*CysZ using the scintillation proximity assay (SPA) (Quick & Javitch,
216 2007). To determine the concentration of half-maximal binding (EC_{50}), we isotopically diluted
217 $[^{35}S]O_4^{2-}$ with non-labeled SO_4^{2-}, obtaining an EC_{50} of 0.81 ± 0.045 μM (Fig. 4c). Competing
218 binding of $[^{35}S]O_4^{2-}$ with SO_3^{2-} revealed that *Pd*CysZ binds SO_3^{2-} with greater affinity as reflected
219 by a half-maximum inhibition constant (IC_{50}) of 0.12 ± 0.034 μM.
220
221 Next, we conducted $[^{35}S]O_4^{2-}$ uptake experiments on detergent solubilized and purified CysZ
222 reconstituted in proteoliposomes. Proteoliposomes provide a means for assessing the activity of a
223 membrane protein independent of other proteins found natively in cellular expression systems.
224 Such binding partners could be involved in SO_4^{2-} uptake as measured in intact *E. coli* cells (Fig.
225 4a). Furthermore, in addition to the SPA measurements, these experiments allowed us to confirm
226 that the detergent-based extraction from the membrane and purification of CysZ used for our
227 crystallization approaches did not compromise its activity. Acting alone, *Pd*CysZ in
228 proteoliposomes mediated the accumulation of SO_4^{2-} in a time-dependent manner; the uptake
229 occurred rapidly, peaking at about 30 seconds after incubation with the radioligand (Fig. 4d).

230 Consistent with our observations on cells, the influx of sulfate was inhibited by the presence of
231 SO_3^{2-} at micromolar concentrations. *Il*CysZ and *Pf*CysZ exhibited similar uptake profiles in
232 liposomes, suggesting similar functional properties (Fig. 4-S1a).
233
234 To determine whether sulfate uptake by CysZ was driven by the concentration gradient of sulfate
235 itself, or if it was coupled to a secondary driving force such as an ion (e.g., H^+, Na^+) gradient, we
236 tested the effect of the uncouplers CCCP (carbonyl cyanide m-chlorophenyl hydrazone), and
237 oligomycin, as well as that of the ionophores valinomycin and gramicidin on CysZ-mediated
238 $[^{35}S]O_4^{2-}$ accumulation in our cell-based assay (Buckler & Vaughan-Jones, 1998). Since none of
239 these uncoupling agents had a significant effect on the accumulation of $[^{35}S]O_4^{2-}$, these
240 experiments suggest that CysZ-mediated $[^{35}S]O_4^{2-}$ flux is not dependent on the 'classical' H^+ or
241 Na^+ electrochemical transmembrane gradient, thus indicating that the observed uptake solely
242 depends on the SO_4^{2-} transmembrane concentration gradient (Fig. 4-S1b), a notion that is
243 consistent with a channel-like mechanism for CysZ-mediated SO_4^{2-} translocation.
244
245 Finally, we employed the use of single-channel electrophysiological current recordings on CysZ
246 (*Pd*CysZ and *Il*CysZ) reconstituted in a planar lipid bilayer (Fig. 4-S2). Consistent with our
247 $[^{35}S]O_4^{2-}$-based flux assays, these electrophysiological measurements showed that reconstituted
248 CysZ exhibits channel-like properties when a membrane potential is applied in symmetrical
249 Na_2SO_4 solutions on both sides of the bilayer. The *Il*CysZ protein exhibits an average open
250 probability (P_{OPEN}) of 0.8 and conductance estimated as 92 ± 12 pS. However, unitary
251 conductance levels varied, showing a dependence on the amount of protein reconstituted in the
252 planar lipid bilayer. The *Pd*CysZ protein passes a broad distribution of currents at fixed voltages.

253 For both proteins, the SO_4^{2-} currents are completely abolished by the presence of low micromolar
254 SO_3^{2-} concentrations.
255
256 **Sulfate Binding Site**
257 A bound SO_4^{2-} ion was observed in the structure of *Il*CysZ. This binding site was confirmed by
258 purifying and crystallizing CysZ in the presence of the heavier SO_4^{2-}-analog selenate (SeO_4^{2-}). In
259 the SO_4^{2-} bound structure, each protomer in the dimer showed electron density consistent with a
260 bound sulfate, but only one of the two refined to full occupancy, hence only one is shown in the
261 refined model of *Il*CysZ (Fig. 5a). The location of the bound SO_4^{2-} is close to the putative
262 membrane interface where it is coordinated by two arginine residues (R27 and R28 in *Il*CysZ)
263 and the backbone amides of residues G25 and L26. This SO_4^{2-}-binding site is in the loop between
264 H1 and H2, with the motif GLR(R) being well conserved among the CysZ family. Consistent
265 with this observation, crystals of *Il*CysZ in which R27 and R28 were replaced with alanines did
266 not show interaction with SO_4^{2-} in that site based on i) the lack of any density in the refined
267 mutant structure at 2.3 Å and ii) functional studies (see below). We could not confirm the
268 presence of SO_4^{2-} in either of the other structures of *Pf*CysZ and *Pd*CysZ, owing, at least in part,
269 to their lower resolution limit (~3.5 Å). Soaking and co-crystallization attempts on both *Pf*CysZ
270 and *Pd*CysZ with SeO_4^{2-} resulted in cracking and destabilization of crystals, and loss of
271 measurable diffraction. In this vein, *Il*CysZ R27A/R28A, and the corresponding *Pf*CysZ R25A
272 and *Pd*CysZ R23A mutants all have severely impaired sulfate uptake capability, as demonstrated
273 in cells and proteoliposomes (Fig. 5b, 4-S1a), consistent with a conserved role for these targeted
274 arginine residues in all CysZ variants.
275

276 **Conserved core in the hydrophilic head**

277 Each apex of the triangular faces of the *Pd*CysZ hexamer comprises an extra-membranous
278 hydrophilic head of a CysZ protomer (Fig. 1a and 5c). The central core of this hydrophilic head
279 consists of the ends of H3 and H5 and the start of H4, with their residues forming an intricate
280 network of hydrogen bonds and salt bridges that hold this helical bundle together (Fig. 5c). Two
281 highly conserved motifs in this region, ExVE and QYxDYPxDNHK (Fig. 2c), likely play critical
282 roles in this network of interactions. The two aspartates E103 and E106 (*Pd*CysZ) in the first
283 motif interact with R129, N184 and R110, H185, respectively (Fig. 5c). The conserved R129 and
284 R133 on H4b in turn also interact with D183 and N184 in the DNHK sequence stretch of H5b in
285 the second motif. In the second motif, a conserved tyrosine (Y177) lies below the membrane
286 interface towards the core of the molecule, along H5.

287

288 Mutational studies on a subset of these conserved hydrophilic-head residues highlight their
289 functional relevance (Fig. 5d). R129 mutants (R129E or Q) exhibited a severe loss of SO_4^{2-}
290 binding and uptake, and mutations made to Q176 and Y177 also showed loss in function, albeit
291 to a lesser extent. Charge reversal mutations made to E103 and E106 (to K or R) or the
292 equivalent *Il*CysZ E107, E110 resulted in very poor expression and stability levels, suggesting
293 that the disruption of the interaction network in this region may destroy the structural integrity of
294 the protein.

295

296 **Putative pore and sulfate translocation pathway of CysZ**

297 It was already evident from the initial *Il*CysZ structure that the hydrophilic head presented an
298 incipient opening into the membrane, as seen in the center of the Fig. 5c inset, but it was quite

299 unclear how this opening might relate to transmembrane sulfate translocation. The 'transverse'
300 dimer structure of *Pf*CysZ clarified the possibility for ion permeation by showing that openings
301 in its two protomers aligned across the putative membrane; however, the prospective
302 transduction pathway would then be open on one side to the bilayer. The *Pd*CysZ structure
303 showed that the open sides of transverse dimers line a central cavity in the *Pd*CysZ hexamers.
304 Thus, plausible transduction pathways became more evident.

305

306 The view into a protomer surface along the direct pathway between transverse dimer apices (e.g.
307 green to orange in Fig. 3e) displays the pattern of exceptionally high conservation associated
308 with the entrance to a putative pore for sulfate translocation (Fig. 6a, b). This putative entrance
309 or 'pore' lies in the midst of a tight network of interacting residues, which is seen detailed in the
310 inset of Fig. 5c in the very same view as for Fig. 6a. These interactions close the incipient pore
311 from ion conduction in this conformation. The electrostatic potential surface of PdCysZ (Fig. 6c)
312 shows striking and puzzling electronegativity at the conserved pore entrance (Fig. 6d compared
313 with Fig. 6b). Because of the conservations, similar electrostatics pertain to the two other
314 homologs. Moreover, we observe the same helix dispositions, pore shape and charge interactions
315 in all three structures, evident when the protomers are superimposed (Fig. 3c, top view).

316

317 Pathway prediction algorithms (namely, PoreWalker (Pellegrini-Calace, Maiwald, & Thornton,
318 2009)) performed on the CysZ protomer revealed a putative ion translocation pathway that
319 begins at the entrance to the putative pore and ends in the large central cavity of the CysZ
320 hexamer that lies within the membrane (Fig. 7a, b). In the structure of *Pd*CysZ, the entrance of
321 the putative pore appears to be closed by the network of charge-interactions by conserved

322 residues, R110, E106, N184, H185 and W235 (Fig. 6a, 7c-1). These residues interact to tightly
323 restrict access to the putative translocation pathway and constrict the entry of ions, likely
324 selected for size and charge (Fig. 7-S1a). After the narrow entrance, the putative pathway
325 broadens, surrounded by a ring of conserved asparagine and glutamine residues, N33, N91, N98,
326 Q176, and N220 (Fig. 7c-2). Mutations to N98 (N98A, N98D) led to impaired sulfate uptake
327 (Fig. 5d). Following this polar environment, the putative pathway widens even further leading
328 into a large, primarily hydrophobic, internal cavity encapsulated by the TM helices of the
329 *Pd*CysZ hexamer, located in the plane of the lipid bilayer (Fig. 7b, 7-S1c, d). The hydrophobic
330 cavity is enclosed on both the top and bottom by pairs of helices H4-H5, three on each side from
331 the six protomers, pointing in towards the three-fold axis, with each side of the cavity, at its mid-
332 section, measuring ~50 Å, when viewed from above or outside the membrane (Fig. 7-S1c, d).
333 Given the symmetric, dual topology nature of the CysZ assembly, the ions could then exit the
334 cavity via the same pathway as they entered, but traveling through CysZ protomers located on
335 the opposite side of the membrane.
336

337 **DISCUSSION**

338 There are four known families of dedicated SO_4^{2-} transport systems in prokaryotes[6], of which
339 CysZs are the least studied (Zhang et al., 2014). The sequence of CysZ, coding for an integral
340 membrane protein with four predicted TM segments, shows no resemblance to any other known
341 protein. This, and the quest to set the basis for a mechanistic understanding of function,
342 prompted us to investigate the structure of CysZ.
343

344 We present here three structures of CysZ from different species, all determined by x-ray

345 crystallography. These structures are all similar, showing a novel fold comprising two extended
346 TM helices and two hemi-penetrating helical hairpins, giving rise to a tripod-like shape within
347 the membrane, and a hydrophilic head (Fig. 2a, b, 3c). Two of the structures (*Il*CysZ and
348 *Pf*CysZ) show dimers in the crystals, albeit with different interfaces (Fig. 3a, b), and the third
349 structure (*Pd*CysZ) displays a hexameric assembly in multiple crystal forms (Fig. 1a).
350
351 All three structures have a dimeric component, each of which is present in the hexameric
352 arrangement present in all the crystal forms of *Pd*CysZ (Fig. 3a, b), showing how they are all
353 related. Furthermore, *Il*CysZ can be crosslinked in membranes to the alternative *Pf*CysZ dimer
354 by engineering of disulfide cysteine mutants (Fig. 3-S2), which suggests that *Il*CysZ can adopt
355 both of the dimer assemblies observed in the *Pd*CysZ hexamer. Thus, we hypothesize that CysZ
356 is a hexamer in nature, as observed in *Pd*CysZ, and that in the case of *Il*CysZ and *Pf*CysZ, this
357 assembly may have come apart into the subsequently crystallized dimeric components during the
358 process of detergent extraction from the membrane and purification. This could be explained by
359 the unusually labile, and conformationally-flexible, nature of the H4b-H5a helical hairpin of
360 CysZ, with only one pair of stabilizing TM helices per protomer (H2-H3). We presume that the
361 hexamer is maintained with the support and scaffolding of the lipid bilayer.
362
363 CysZ shows a clear dual topology arrangement, which we confirmed with cross-linking
364 experiments on membranes (Fig. 3-S2). Dual-topology insertion is fairly uncommon in
365 membrane proteins; however, reported cases include the well-documented EmrE, a multi-drug
366 resistant export protein that inserts into the membrane as an anti-parallel dimer, and the more
367 recently discovered and studied family of double-barreled 'Fluc' fluoride channels (Amadi,

368 Koteiche, Mishra, & McHaourab, 2010; Korkhov & Tate, 2009; Rapp, Granseth, Seppala, & von
369 Heijne, 2006; Stockbridge et al., 2015; Stockbridge, Robertson, Kolmakova-Partensky, & Miller,
370 2013).
371
372 Functional characterizations of the three CysZs for which we have obtained structural
373 information show that all three mediate SO_4^{2-} uptake and that this uptake is inhibited by SO_3^{2-},
374 for which the protein has higher affinity. Comparable results were obtained in cell-based as well
375 as in proteoliposome-based uptake assays (Fig. 4b, d, 4-S1a). Affinities for SO_4^{2-} and SO_3^{2-} were
376 measured by SPA (Fig. 4c). These results are consistent with previously reported data on *E. coli*
377 CysZ (Zhang et al., 2014). Experiments performed on cells with uncoupling agents (Fig. 4-S1b)
378 showed that CysZ-mediated SO_4^{2-} uptake was independent of a classical electrochemical
379 gradient, thus suggestive of the driving force being a concentration gradient of SO_4^{2-} across the
380 membrane. This hypothesis is supported by our single-channel bilayer experiments performed
381 with *Pd*CysZ and *Il*CysZ, which show that CysZ has channel-like properties (Fig. 4-S2).
382
383 Building on our functional and structural discoveries, our results suggest a fascinating hypothesis
384 for mechanisms of SO_4^{2-} transfer and regulation. There is a conserved central core (Fig. 6a),
385 which is likely to have a structural role. Close to this lies a sulfate-binding site, with functional
386 implications (Fig. 5a), and the entrance to a putative pore (Fig. 5c). The entrance of this putative
387 pore is delineated by a hydrophilic network of conserved residues that form a tight constriction
388 in our observed conformation. Following this hypothetical route, sulfate ions that might enter
389 through the three separate pores, one per CysZ dimer pair, would converge into a central
390 hydrophobic cavity (Fig. 7b, 7-S1c, d). Once sulfate ions enter this central cavity, due to the

391 unfavorable environment, these are likely to exit it rather rapidly through one of the three
392 available exit pores on the cytoplasmic side of the hexamer. Hydrophobic cavities and pores are
393 seen commonly in ion channels, with examples ranging from the well documented hydrophobic
394 inner pores of the various potassium channels (Doyle et al., 1998) to the SLAC1 (Chen et al.,
395 2010) and bestrophin anion channels (Yang et al., 2014), and the MscS and MscL
396 mechanosensitive channels (Anishkin, Akitake, Kamaraju, Chiang, & Sukharev, 2010; Bass,
397 Strop, Barclay, & Rees, 2002; Birkner, Poolman, & Kocer, 2012; Chen et al., 2010; Doyle et al.,
398 1998), facilitating the rapid passage of ions due to the unfavorable environment, as well as
399 potentially providing a means of 'hydrophobic gating' (Aryal, Sansom, & Tucker, 2015).

400

401 The surface electrostatics of CysZ (Fig. 1b and Fig. 6c, d) draw attention to the negative
402 potential of the conserved core of each protomer, which is then surrounded by a more neutral
403 annulus. While this feature seems contradictory to admission of SO_4^{2-} ions at the extracellular
404 side, it could be advantageous for expulsion into the cytoplasm on the opposite side. In any case,
405 the structures that we have determined are evidently in a closed state, implying that a
406 conformational change would have to occur to allow passage of SO_4^{2-}, likely modifying the
407 surface electrostatics of the protein. It is tempting to speculate that the regulation of the opening
408 of the pore could be modulated by the binding of sulfate ions to the identified sulfate-binding
409 site, as it is near the entrance of the putative pore. L22 lies in proximity of the entrance of the
410 putative pore, and its backbone amide (along with G21) coordinates the sulfate ion in the GLR
411 motif of the sulfate-binding site. Thus, the binding of sulfate to the GLR motif could trigger a
412 conformational change needed to displace L22, allowing for a wider opening for the sulfate ions

413 to enter the pore. Sulfite could hypothetically exert its inhibitory effect on CysZ function by
414 binding to this site.

415

416 There is an overall electropositive region in the center of the hexameric molecule, when viewed
417 from the top (Fig. 1b, 6c, d). This central region is lined by conserved residues along helices
418 H4b-H5a, namely, R129, R133 and K137. The hydrophobic tips of helices H4b-H5a of the three
419 protomers on each side of the membrane then converge in the center of the hexamer. A pore
420 through the three-fold axis of the hexamer could provide an alternative passageway for SO_4^{2-}
421 ions through this assembly. Although we cannot exclude it, this possibility seems less plausible
422 because of the poor conservation at the tips of helices H4b-H5a and the strictly hydrophobic
423 nature of this region.

424

425 We observe, in agreement with previous data, that SO_3^{2-} inhibits CysZ-mediated SO_4^{2-} flux (Fig.
426 4b, d, 4-S2a, f) (Zhang et al., 2014). The dual topology nature of CysZ could provide a means for
427 internal as well as external regulation. However, experiments in solutions – such as
428 crystallizations – where all molecules are exposed to the same chemical environment, makes
429 capturing such a state in an open or SO_3^{2-} blocked conformation challenging. Despite this
430 limitation, our CysZ structures and associated functional experiments have allowed us to make
431 substantial progress in the understanding SO_4^{2-} uptake by these membrane permeases. This work
432 sets the framework for future experiments aimed at unraveling the molecular details of how
433 SO_4^{2-} is translocated across the membrane by CysZ and how this process is regulated.

434

435 **PDB Accession codes**: *Idiomarina loiheinsis* CysZ: **3TX3**, *Pseudomonas fragi* CysZ: **TBD,**
436 *Pseudomonas denitrificans* CysZ: **TBD**.
437
438 **Materials and Methods**
439
440 **Ortholog selection and cloning**
441 A total of 63 *cysZ* candidate genes were selected by a bioinformatics approach implemented by
442 the New York Consortium of Membrane Protein Structure (NYCOMPS), as previously described
443 (Punta et al., 2009). The majority of the genes (including *Il*CysZ, uniprot ID: Q5QUJ8) were
444 PCR-amplified from fully sequenced prokaryotic genomic DNA (obtained from ATCC®) (Love
445 et al., 2010). CysZ genes from certain species, such as *Pf*CysZ (uniprot ID: A0A0X8F058) and
446 *Pd*CysZ (uniprot ID: M4XKU7) were chemically synthesized by GenScript®, with codon-
447 optimization for protein expression. All genes were cloned by ligation-independent cloning
448 (LIC) (Aslanidis & de Jong, 1990) into an IPTG (isopropyl β-D-1-thiogalactopyranoside)
449 inducible, kanamycin-resistant pET derived plasmid (Novagen), with an N-terminal deca-
450 histidine tag (His10) and a TEV (tobacco etch virus) protease site to allow for tag cleavage upon
451 purification.
452
453 **Protein expression and purification**
454 Expression plasmids bearing the *cysZ* genes were transformed into BL21(DE3)pLysS cells using
455 standard protocols, and grown at 37 °C in 2XYT media supplemented with 50 µg/ml kanamycin
456 and 50 µg/ml chloramphenicol in an orbital shaker at 250 rpm. Protein expression was induced
457 for ~16 hours at 22 °C with 0.2 mM IPTG once an absorbance ($A_{600\ nm}$) of 0.8-1.0 was reached.

458 Selenomethione (Se-Met)-incorporated proteins were expressed in BL21(DE3)pLysS cells
459 grown using an M9 minimal media kit (Shanghai Medicilon Inc.) supplemented with the
460 necessary minerals, vitamins and non-inhibitory amino acids. Se-Met was added prior to IPTG
461 induction at an $A_{600\ nm}$ of 1.2. The Se-Met-incorporated protein was purified using the same
462 procedures as the native protein. Once harvested, the cells were resuspended at 0.2 g/ml in lysis
463 buffer containing 20 mM Na-Hepes pH 7.5, 200 mM NaCl, 20 mM MgSO4, DNase I and RNase
464 A, 0.5 mM PMSF (phenylmethylsulfonyl fluoride), EDTA-free Complete protease inhibitor
465 cocktail (Roche) and 1 mM TCEP-HCl (Tris (2-carboxyethyl) phosphine hydrochloride) as a
466 reducing agent. Initial small-scale expression and detergent screening was performed on 80 mg
467 of pelleted cells (wet weight), and 7-10 g of cells for large-scale protein purification. Cells were
468 lysed using an Avestin® EmusiFlex-C3 homogenizer, followed by protein solubilization with 1%
469 (w/v) decyl maltopyranoside (DM) (Anatrace® Affymetrix) for 1 hour at 4 °C, after-which
470 insoluble material was removed by ultra-centrifugation at 100,000 x g. The solubilized protein
471 was applied to Ni-NTA Sepharose (Qiagen) in batch, washed with lysis buffer containing 0.2%
472 DM and 40 mM imidazole and eluted in buffer containing 250 mM imidazole. Upon elution,
473 CysZ was dialyzed overnight with His-tagged TEV protease at 4°C, against a buffer containing
474 20 mM Na-Hepes pH 7.0, 200 mM NaCl, 0.2% DM, 1 mM TCEP-HCl and 20 mM Na2SO4,
475 allowing for the cleavage of the His10 tag and removal of the imidazole. Tagless CysZ was then
476 re-passaged over Ni-NTA sepharose to re-bind of any uncleaved CysZ, TEV protease and the
477 cleaved His10 tag. The protein was then subjected to size-exclusion chromatography (Superdex
478 200 10/30 HR, GE Healthcare) in 20 mM Na-Hepes pH 7.0, 200 mM NaCl, 1 mM TCEP-HCl,
479 20 mM Na2SO4 and appropriate detergent for crystallization, for *Il*CysZ: 0.06% Lauryl
480 dimethylamine oxide (LDAO) and for *Pf*CysZ and *Pd*CysZ: 1% β-octyl glucopyranoside (β-

481 OG). The choice of detergent was made based on protein yield, stability and mono-disperse gel-
482 filtration peaks obtained in the initial small-scale detergent screening. A yield of ~1.5 mg of
483 purified CysZ was typically obtained from a cell pellet of 7-8 grams (1 liter of culture).
484

485 **Protein Crystallization**

486 *Il*CysZ: Crystals of *Il*CysZ in LDAO were obtained by vapor diffusion at a protein concentration
487 of 6-8 mg/ml at 4°C, in a 1:1 v/v ratio against a precipitant of 28-32% PEG400, 0.1M Tris-HCl
488 pH 8.0, with salt additive of 0.1 M NaCl or 0.1 M $MgCl_2$. The crystals appeared overnight,
489 continued to grow in size over the course of 2-4 days after set-up. After optimization, the crystals
490 grew to a maximum size of ~200 μm x 100 μm x 50 μm with a rhomboid or cuboid shape. The
491 crystals were harvested directly without the addition of a cryo-protectant and flash-frozen into
492 liquid nitrogen, for data collection on the X4A/X4C beamlines at National Synchrotron Light
493 Source (NSLS), Brookhaven National Labs (Upton, NY). SeMet derivatized and selenate co-
494 crystals were obtained from the same conditions as the native crystals.

495 *Pf*CysZ: Crystals of *Pf*CysZ in β-OG at 5 mg/ml were initially obtained at 4 °C by vapor
496 diffusion, in a 1:1 protein to precipitant ratio, after 1-2 days against 28% PEG400, 0.1 M MES
497 pH 6.0. They were cuboid in shape and grew in clusters of multiple crystals originating from a
498 common locus. The crystals were optimized to a maximal size of ~150-200 μm x 50 μm x 50
499 μm, with the best diffracting crystals grown under silicone oil (visc. 500) in microbatch Terazaki
500 plates. Crystals were directly flash-frozen into liquid nitrogen without the use of a cryo-
501 protectant and were exposed to X-rays at the NE-CAT (241DC and E) beamlines at APS,
502 Argonne National Lab (Argonne, IL) for data collection. SeMet crystals were obtained from the
503 same conditions as the native protein.

504 *Pd*CysZ: Crystals of *Pd*CysZ in β-OG at 5-8 mg/ml were initially obtained at 4°C by vapor
505 diffusion, in a 1:1 protein to precipitant ratio, after 2-3 days against 22-30% PEG550MME, 0.1
506 M Na-Hepes pH 7.0. The rod-like crystals were hexagonal on one face, and grew in clusters
507 originating from a common locus as well as on the edge of the drop. The multiple crystal forms
508 observed were all obtained in the same crystallization conditions. The crystals were optimized to
509 a maximal size of ~200-250 μm x 25 μm x 50 μm. With the addition of 20% glycerol (w/v) as a
510 cryo-protectant, the crystals were flash-frozen into liquid nitrogen and were exposed to X-rays at
511 the NE-CAT (24-IDC and 24-IDE) beamlines at APS, Argonne National Lab (Argonne, IL) for
512 data collection.

513 **Data collection and structure determination**

514 *Il*CysZ: The structure of CysZ was determined by the single-wavelength anomalous diffraction
515 (SAD) method from anomalous diffraction of a selenate (SeO_4^{2-}) derivative crystal. The
516 anomalous signals were measured at the Se K-edge peak wavelength, which was determined
517 experimentally from fluorescence scanning of the crystal prior to data collection. All diffraction
518 data were recorded at 100K using an ADSC Q4R CCD detector at the NSLS X4 beamline.
519 Diffraction data were indexed, integrated, scaled, and merged by HKL2000 (Otwinowski &
520 Minor, 1997). Selenate substructure determination was performed with the SHELXD program
521 through HKL2MAP (Pape & Schneider, 2004). A resolution cut-off at 2.6 Å was used for
522 finding Se sites by SHELXD. A strong peak found by SHELXD was used to calculate initial
523 SAD phases, which were improved by density modification by SHELXE (Sheldrick, 2010). With
524 a solvent content of 65% corresponding to 2 molecules in the asymmetric unit, 50 cycles of
525 density modification resulted in an electron density map of sufficient quality for model building.
526 The initial polypeptide chain was built by Arp/Warp (Langer, Cohen, Lamzin, & Perrakis, 2008),

527 at 2.1 Å by using experimental phases. Further cycles of model building were performed
528 manually using COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010) and all rounds of
529 refinements were performed with PHENIX (Adams et al., 2010). The native structure of CysZ
530 with bound sulfate was determined both by multi-crystal native SAD (Liu et al., 2012) (final
531 resolution of 2.3 Å) and by molecular replacement with the selenate bound model (final
532 resolution of 2.1 Å). In addition, phase information obtained from Se-Met derivatized protein
533 with 9 Se sites per CysZ molecule, verified our model obtained from the selenate data.
534
535 *Pf*CysZ: Multi-crystal SeMet-SAD data sets were collected at APS beamline 24-IDC with a
536 Pilatus 6M pixel array detector under a cryogenic temperature of 100 K. To enhance anomalous
537 signals from Se atoms for phasing (Liu, Zhang, & Hendrickson, 2011), the X-ray wavelength
538 was tuned to the Se-K edge (λ=0.9789 Å). The orientation of crystals was random without
539 special consideration of crystal alignment, and beam size was adjusted to match the crystal size.
540 A total of 22 data sets were collected, each from a single crystal. An oscillation angle of 1° was
541 used for data collection with a total of 360 frames for each data set. The beam size was adjusted
542 to match the crystal size. The 22 single-crystal data sets were processed individually by using
543 XDS (Kabsch, 2010) and CCP4 packages (Winn et al., 2011). For phasing purposes, the low-
544 resolution anomalous signals were enhanced by increased multiplicity. By rejection of 7 outlier
545 crystals (Liu et al., 2012), anomalous diffraction data from 15 statistically-compatible crystals
546 were scaled and merged for phasing. For outlier rejection, a unit-cell variation of 1.0σ was used.
547 CCP4 program POINTLESS and SCALA (Evans, 2006) were used for data combining; and
548 Bijvoet pairs were kept separately throughout the data flow. For refinement purposes, keeping
549 the high angle data was important, and done by limiting radiation damage as well as by

550 increasing multiplicity. Although most *Pf*CysZ crystals diffracted to only about 3.5 Å spacings
551 or poorer, we intentionally set the detector distance to include higher spacings. Higher resolution
552 data were retained through a data merging procedure that is described as follows: 1) The 22
553 individually processed data sets were analyzed by diffraction dissimilarity analysis by using only
554 high angle data between 3.5-3.0 Å, resulting in three subsets. 2) The data statistics of members in
555 each subset were checked manually and the subset that contained the highest angle data set, e.g.
556 data set 6, was selected for further analyses and data combination. 3) Each data set within the
557 selected subset was compared with data set 6 by high-angle intensity correlation. Six of the
558 highest resolution data sets were statistically comparable and therefore were selected for
559 merging. For phasing, substructure solutions were found by SHELXD (Sheldrick, 2010) and
560 were further refined and completed by PHASER (McCoy et al., 2007) and then used to compute
561 initial SAD phases at the data limit by SAD phasing with PHENIX (Adams et al., 2010). Phases
562 were density modified with solvent flattening and histogram matching as implemented in CCP4
563 program DM (Cowtan & Zhang, 1999) to improve phases and also to break phase ambiguity.
564 The estimated solvent contents of 71% were used for density modification. The model was
565 initially built into the experimental electron density map by COOT (Emsley et al., 2010),
566 followed by iterative refinement by PHENIX and model building in COOT. The refined model
567 does not contain solvent molecules at this resolution.
568
569 <u>*Pd*CysZ:</u>
570 Native crystal data were collected at the APS beamline 24-IDC with a Pilatus 6M pixel array
571 detector at a cryogenic temperature of 100 K at an X-ray wavelength of λ=1.023 Å. The sample-
572 to-detector distance was set to 500 mm. An oscillation angle of 0.5° was used for data collection.

573 The beam size was adjusted to match the crystal size. Molecular replacement was attempted
574 using a variety of search models (*Pf*CysZ and *Il*CysZ monomer/dimer models, with various
575 degrees of truncation). Success was achieved by searching for six copies of a search model
576 consisting of the *Pf*CysZ monomer, with residues 36-53 deleted and the sequence adjusted using
577 CHAINSAW (Stein, 2008) (pruning non-conserved residues to the gamma carbon). Density
578 modification of the initial map was performed in PARROT, incorporating solvent flattening,
579 histogram matching and NCS-averaging. An initial round of model building was performed in
580 COOT (Emsley et al., 2010) into this map, followed by further phase improvement and bias-
581 removal using *phenix.prime_and_switch* (Adams et al., 2010*)*. The improved map was used for a
582 second round of model building, followed by iterative cycles of reciprocal space refinement
583 using *phenix.refine*, and real-space refinement and correction in COOT.
584 All graphical representations and figures of our structural models were made in either PyMOL
585 (Schrodinger, 2010) or Chimera (Pettersen et al., 2004).
586

587 **Radioligand Binding by Scintillation Proximity Assay (SPA)**

588 CysZ was purified by standard protocols as described above, with the exception of leaving the
589 histidine-tag intact without cleavage by the TEV protease. The imidazole was removed by
590 dialysis and the purified protein was not run over the size exclusion column, and instead was
591 directly concentrated to 2 mg/ml for the SPA experiment. $[^{35}S]O_4^{2-}$ obtained in the form of
592 sulfuric acid (American Radiolabeled Chemicals (ARC)) was used as the radioligand. 100 ng –
593 250 ng of CysZ was used per assay point, diluted in 100 µl of assay buffer containing 20 mM
594 Hepes pH 7.0, 200 mM NaCl, 0.2% DeM, 20% glycerol, 0.5 mM TCEP, with 12.5 µl of Copper
595 YSi beads (Perkin Elmer). A trace amount of $[^{35}S]O_4^{2-}$ (~10-30 nM) mixed with cold ligand

596 (Na_2SO_4) was added to each well of a 96 well clear bottom plate along with the mix of protein,
597 buffer and beads. The plate was agitated for a minimum of 30 minutes to as long as overnight,
598 and measured in a scintillation counter the following morning (MicroBeta, Perkin Elmer).
599 Competitive binding experiments with a gradient of a competing cold ligand ranging from (10
600 nM – 100 mM) were performed in triplicate, with the control measured at each concentration by
601 adding 1 M Imidazole to the reaction buffer, to prevent binding of the protein to the scintillation
602 beads. All data were analyzed and graphically represented with GraphPad™ Prism6 software.
603

604 **[^{35}S]O_4^{2-} Uptake Experiments**

605 *In whole cells:* The *cysZ* gene knockout strain in *E. coli* K-12 was obtained from the Coli
606 Genetic Stock Center at Yale University (http://cgsc.biology.yale.edu/), originally found in the
607 Keio knockout collection (Baba et al., 2006). Genotype: F-, Δ(araD-araB)567, Δ
608 lacZ4787(::rrnB-3), λ-, ΔcysZ742::kan,rph-1, Δ(rhaD- rhaB)568, hsdR514. The knockout
609 strain was made competent by standard protocols (Hanahan, 1983) to allow for the
610 transformation and expression of *cysZ*-containing plasmids, for rescue experiments. Wild-type
611 parental *E. coli* K-12 was used as a control strain. The strains were grown in LB (Luria Broth)
612 without any antibiotic (for the WT cells) and with 50 µg/ml of kanamycin for the knockout cells
613 at 37 °C overnight to saturation. The culture was spun down at 3000 x g the next morning and
614 the cells were resuspended in Davis-Mingioli (DM) minimal media without sulfate ($MgSO_4$ was
615 replaced with $MgCl_2$ and $(NH_4)_2SO_4$ was replaced by NH_4Cl), supplemented with 0.63 mM L-
616 cysteine to allow the cells to grow in the absence of sulfate (Davis & Mingioli, 1950). The cells
617 were incubated in DM media for a minimum of 9 hours, maximum overnight, at 37°C after
618 which they were spun down and washed 3 times in 5 mM Na-Hepes pH 7.0. This ensured that

619 the cells were starved of SO_4, to deplete the sulfate stores in the cell, enhancing sulfate uptake
620 measured (F. Parra, personal communication). The cells were resuspended in the same buffer at
621 0.7 g/ml at room temperature, and 10 µl of the cell suspension was used for each point in the
622 uptake experiment. A final concentration of 320 µM of SO_4 (Na_2SO_4 with $[^{35}S]O_4^{2-}$ as a tracer)
623 was used outside. The uptake experiments were performed in triplicate, measured over a time
624 course of typically 0 - 300 seconds. The reaction was stopped by dilution, by adding 1.5 ml of
625 ice-cold buffer (5 mM Na-Hepes pH 7.0) to the cells in the tube. The reaction mix was
626 immediately poured onto a vacuum filtration device, with an individual glass-fiber filter (0.75
627 µm, GF/F) per assay point, after which the filters are washed once more with 1.5 ml of buffer.
628 The dry filters, with the cells attached to their surface, were then moved to scintillation vials,
629 containing 4 ml of EconoSafe® scintillation cocktail, to be counted the next morning. For the
630 *CysZ*⁻ cell rescue experiments, a similar protocol was used, transformed with an ampicillin
631 resistant expression vector containing the CysZ gene, mutant or empty vector (as the control).
632 Upon transformation, the cells were grown overnight in a 4 ml starter culture of LB with 50
633 µg/ml kanamycin and 100 µg/ml ampicillin. The next morning the entire 4 ml was used to
634 inoculate 75 ml of LB (Kan, Amp), and grown at 37 °C for 2.5-3 hours until an OD_{600} of 0.6-0.8
635 was reached. Protein expression was then induced at 37 °C for 4 hrs with 0.2 mM IPTG. After
636 4hrs, the cells were spun down and resuspended in the DM minimal media without sulfate, 0.63
637 mM cysteine, Amp, Kan and 0.2 mM IPTG, and incubated overnight at 22 °C. The next morning
638 the cells were spun down, washed 3 times with 5 mM Na-Hepes 7.0 and resuspended at final
639 concentration of 0.7 g/ml for uptake.

640 *In proteoliposomes:* CysZ was purified by the procedure described above, and upon elution from
641 the size-exclusion column concentrated to 1 mg/ml for reconstitution into liposomes at a protein

642 to lipid ratio of 1:100. The liposomes were comprised of a 3:1 ratio of *E. coli* polar lipids and
643 phosphatidylcholine (PC) (Avanti Polar Lipids), prepared by previously described methods
644 (Rigaud, Pitard, & Levy, 1995). Typically, 10 µl of CysZ-proteoliposomes and control 'empty'
645 liposomes at a concentration of 5-10 mg/ml, were used per assay point and diluted into a reaction
646 buffer of 100 µl, containing $[^{35}S]O_4^{2-}$ (at trace amounts of ~10-15 nM) supplemented with
647 Na_2SO_4 at the desired concentration. The experiment was performed in triplicate and sulfate
648 accumulation was measured either over a time-course or at a fixed time with different amounts
649 of SO_4^{2-} or SO_3^{2-} present. The reaction was stopped by dilution, by adding 1.5 ml of ice-cold
650 buffer (5 mM Na-Hepes pH 7.0) to the cells in the tube. The reaction mix was immediately
651 poured onto a vacuum filtration device, with an individual nitrocellulose filter (0.22 µm) per
652 assay point, after which the filters are washed once more with 1.5 ml of buffer. The dry filters,
653 with the cells attached to their surface, were then moved to scintillation vials, containing 4 ml of
654 EconoSafe® scintillation cocktail, to be counted the next morning.

655 **Measurement of CysZ single-channel activity in the planar lipid bilayer**

656 A previously described method (Mueller, Rudin, Tien, & Wescott, 1962) to insert ion channels
657 incorporated in liposomes into "painted" planar lipid bilayers by vesicle fusion was used to
658 incorporate CysZ into a lipid bilayer created on a small aperture between two aqueous
659 compartments, called the cis and trans compartments (Morera, Vargas, Gonzalez, Rosenmann, &
660 Latorre, 2007). Since this system is very sensitive to contaminants, CysZ was expressed in and
661 purified from a porin- deficient strain of *E. coli* cells to prevent any carry through of
662 contaminating porins that could create large conductances and artifacts in the single-channel
663 recordings. Phosphatidylethanolamine (PE) and phosphatidylserine (PS) (Avanti Polar Lipids),
664 in a ratio of 1:1 were dissolved in chloroform to mix, and dried completely under an argon

665 stream. The mixed and dried lipids were then dissolved in n-decane to a final concentration of 50
666 μg/ml, and kept at 4 °C. The lipids are always prepared fresh, on the same day of the experiment.
667 The purified CysZ protein were incorporated into PE:PS (1:1) liposomes by brief sonication at
668 80kHz for 1 minute at 4 °C. The experimental apparatus consisted of two 1-ml buffer chambers
669 separated by a Teflon film that contains a single 20- to 50-μm hole. A lipid bilayer was formed
670 by "painting" the hole with the 1:1 mixture of PE:PS, this results in a seal between the two cups
671 formed by the lipids (Leal-Pinto, London, Knorr, & Abramson, 1995). For these studies, the cis
672 side was defined as the chamber connected to the voltage-holding electrode and all voltages are
673 referenced to the trans (ground) chamber. Stability of the bilayer was determined by clamping
674 voltage at various levels. If a resistance > 100 Ω and noise < 0.2 pA were maintained in the
675 patch, the proteoliposomes containing CysZ were added to the trans side of the chamber and
676 stirred for 1 min. The fusion event or insertion of a channel into the bilayer was assessed by the
677 presence of clear transitions from 0 current to an open state.
678
679 **Site-specific cysteine labeling experiments**
680 All functional and cysteine mutants of CysZ were generated by site-directed mutagenesis, using
681 the QuikChange® site-directed mutagenesis kit (Strategene). The sequence-verified mutants
682 were then tested for expression in comparison to the WT CysZ. To address the membrane
683 topology of *Il*CysZ, single cysteine mutants were designed to perform site-directed fluorescence
684 labeling based on the accessibility of the cysteine to the membrane impermeable thiol-directed
685 fluorescent probe (Ye, Jia, Jung, & Maloney, 2001). A set of surface-exposed residues at
686 different positions on the CysZ molecule were selected to be mutated to cysteines, based on the
687 *Il*CysZ structure. The cysteine mutants were expressed by standard protocols that were used for

688 the WT protein. The membrane fraction of each mutant was pelleted after cell lysis by
689 ultracentrifugation at 100,000 x g and resuspended at 20 mg/ml (Bradford assay) in fresh buffer
690 containing 20 mM Na-Hepes pH 7.0, 200 mM NaCl, protease inhibitors: 0.5 mM PMSF and
691 Complete protease inhibitor cocktail EDTA-free and 1 mM TCEP-HCl. 1 ml of membranes were
692 then incubated with 30 μM membrane-impermeant fluorescein-5-maleimide dye (Invitrogen, 2
693 mM stock freshly prepared in water, protected from light) for 30 minutes in the dark at room
694 temperature. The labeling reaction was stopped by the addition of 6 mM β- mercaptoethanol, and
695 the membranes were spun down and resuspended in fresh buffer to remove any remaining
696 unreacted fluorescent dye. The fluorescently labeled protein was then purified from the
697 membranes by solubilization with 1% DM, using a standard Ni-NTA purification protocol,
698 qualitatively analyzed on an SDS-PAGE, and quantitatively measured by a Tecan® fluorescence
699 plate reader at an excitation of 495 nm, emission of 535 nm.
700

701 **Crosslinking of Cysteine Mutant Pairs in CysZ**

702 Cysteine mutants of *Pf*CysZ and *Il*CysZ were designed based on pairs of residues that were in
703 close proximity (within 3-7 Å) of each other in our structure that could have the ability to
704 covalently join the 2 protomers of the dimer. The single and double cysteine mutants were made
705 by site-directed mutagenesis using the QuikChange® Site-Directed Mutagenesis Kit
706 (Strategene). Mutants and WT were expressed using the standard protocols CysZ expression.
707 Isolated membrane fraction was resuspended by homogenization in fresh lysis buffer at a
708 membrane protein concentration of ~25 mg/ml (measured by Bradford Assay). Bis-
709 methanethiosulfonate (Bis-MTS) crosslinkers (Santa Cruz Biotechnolgies) of different spacer
710 lengths were used at 0.5mM (dissolved in DMSO) added to 1ml of resuspended membranes at

711 RT for 1 hour. Bis-MTS crosslinkers are membrane permeable, highly reactive and specific to
712 sulfhydryl groups, and the covalent linkage is resistant to reducing agents like β-mercaptoethanol
713 (Akabas, Stauffer, Xu, & Karlin, 1992). The crosslinking lengths used were: 1,1-Methanediyl
714 Bismethanethiosulfonate (3.6 Å), 1,2- Ethanediyl Bismethanethiosulfonate (5.2 Å), 1,4-
715 Butanediyl Bismethanethiosulfonate (7.8 Å) and 1,6-Hexanediyl Bismethanethiosulfonate (10.4
716 Å). The reaction was quenched by the addition of 10 mM free cysteine, and the protein was then
717 extracted and purified from the membranes with 1% DM followed by Ni-NTA resin. The
718 imidazole is then diluted out in the Ni-elute, and the His-tag is cut with TEV protease in small
719 scale. The cleaved protein is passaged over Ni-NTA resin to remove contaminants and uncleaved
720 protein, and flowthrough is run on a reducing SDS- PAGE and stained with Coomassie blue to
721 analyze dimer formation. The mutants designed were L161C (single mutant), L161C-A164C and
722 N160C-L168C (double mutants) for *Pf*CysZ; and L156C-Q163C and V157C-Q163C (double
723 mutants) for *Il*CysZ.
724

725 **Acknowledgments:**

726 Crystallographic data for this study were collected on the NSLS beamline X4A at Brookhaven
727 National Laboratory and on the NE-CAT beamlines 24ID-C and E (supported by NIH-NIGMS
728 grant P41 GM103403) at the Advanced Photon Source. The Pilatus 6M detector on 24-ID-C
729 beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This work was supported by
730 NIH-NIGMS grants R01 GM098617 (F.M.) and R01 GM107462 (W.A.H.), and by grants
731 through the New York Structural Biology Center for the New York Consortium on Membrane
732 Protein Structure (NYCOMPS; U54 GM095315) and for the Center on Membrane Protein
733 Production and Analysis (COMPPÅ; P41 GM116799). O.B.C. was supported by a Charles H.

Revson Senior Fellowship. We thank John Schwanof for assistance during data collection at NSLS, Joe Mindell, Francisco Parra, Albano Meli, Giuliano Sciara and Carlos A Villalba Galea for helpful advice and contributions at the early stages of the project, and Leora Hamberger for her assistance managing the Mancia Lab.

Declaration of Competing Interests:

The authors declare that they have no financial or non-financial competing interests.

741 **Figures and Tables:**

742 **Table 1: Crystallographic data and refinement statistics.**

743

Data collection	**Native *IlCysZ***	***IlCysZ w/SeO_4^{2-} (SAD)***	**SeMet *PfCysZ* *(6 crystals, refinement)***	**SeMet *PfCysZ* *(15 crystals, SAD)***	**Native *PdCysZ* *(MR-SAD)***
Beamline	NSLS X4A	NSLS X4A/C	APS 24-ID-E	APS 24-ID-C	APS 24-ID-C
Space group	C2	C2	C2	C2	$P6_3$
Cell dimensions:					
a, b, c (Å)	128.9, 82.0, 100.4	128.9, 81.9, 100.3	172.29, 56.9, 96.17	172.35, 56.9, 96.31	225.13, 225.13, 96.62
α, β, γ (°)	90, 125.1, 90	90, 125.1, 90	90, 91.43, 90	90, 91.33, 90	90, 90, 120
Z_a	2	2	2	2	6
Wavelength	1.7432	0.96789	0.97890	0.97890	1.0230
Bragg spacings (Å)	30-2.30	50-2.10	40-3.20	40-3.50	86-3.40
R_{merge}	0.047 (0.257)	0.046 (0.417)	0.136 (6.506)	0.135 (2.13)	0.095 (2.89)
I / σ_I	29.4 (7.9)	25.0 (1.9)	12.3 (0.7)	24.2 (2.6)	22.4 (1.2)
Completeness (%)	99.9 (99.9)	99.9 (100.0)	99.8 (99.3)	99.9 (100.0)	100.0 (100.0)
Multiplicity	7.3 (7.2)	7.6 (7.6)	27.2	72.8	19.7
Refinement:					
Resolution (Å)	2.30		3.50		3.40
No. of reflections	38075		20741		37221
R_{work} / R_{free}	0.200/0.238		0.297/0.339		0.239/0.277
No. of atoms:					
Protein	3679		3428		11415
Ligand/ion	249		0		560
Water	162		0		0
Average B-factors (Å^2)	49.4		198.9		204.18
Protein	48.4		198.9		204.18
Ligand/Ion	77.3		-		-
Water	45.3		-		-
Bond Ideality (r.m.s.d.):					
Bond lengths (Å)	0.006		0.006		0.004
Bond angles (°)	0.894		1.133		0.96
Ramachandran Analysis:					
Favored (%)	99.0		99.2		98.5
PDB accession code	3TX3		TBD		TBD

744

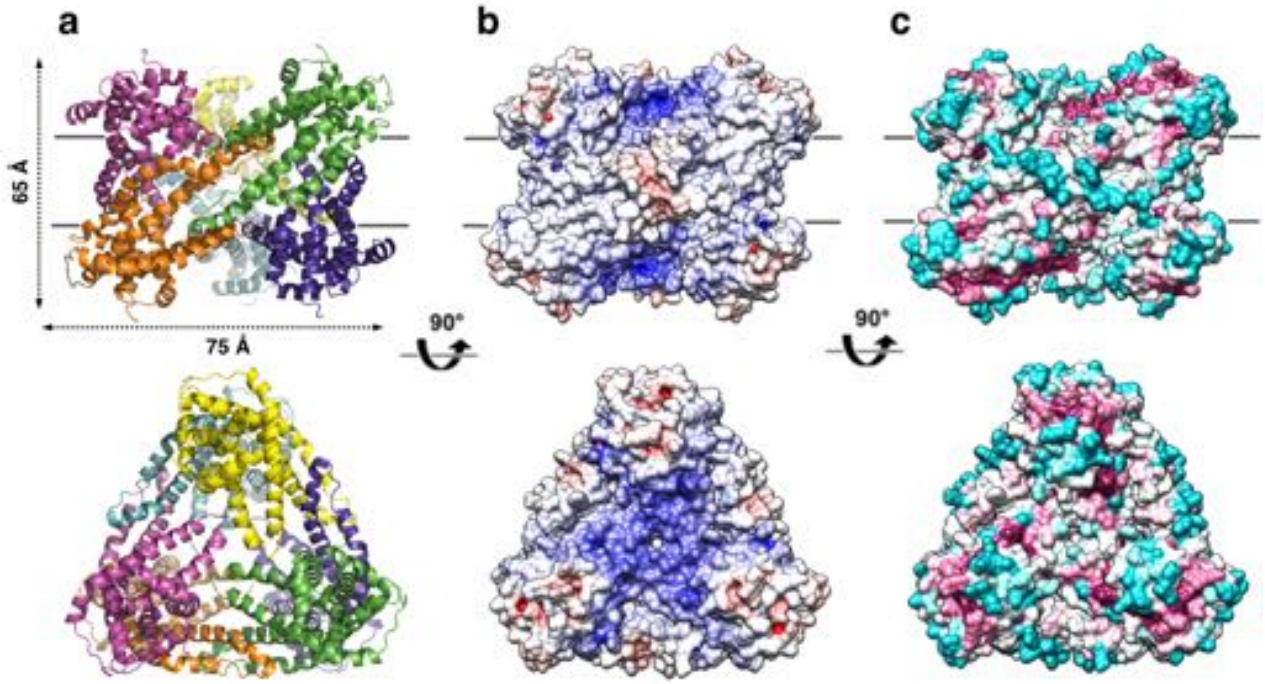
745 Values in parentheses are from the highest resolution shell. R_{free} was calculated using 5% of data

746 excluded from refinement.

747

748

749 **Figure 1.**



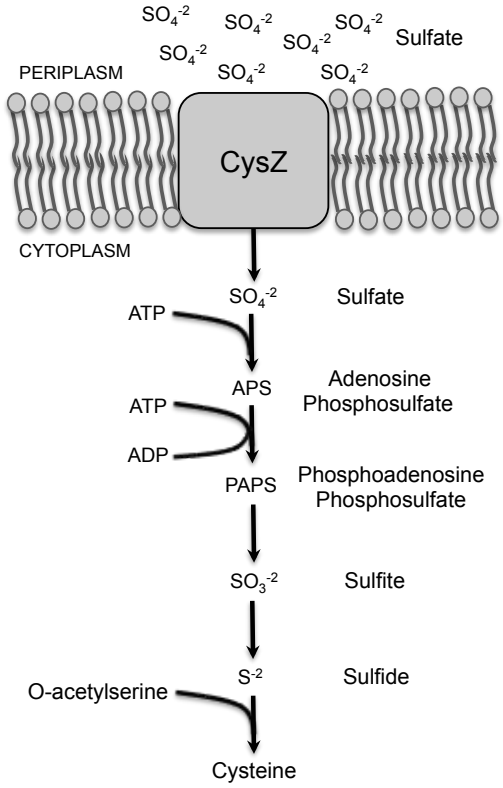
750

751

752 **Fig. 1. Overall structure of the *P. denitrificans* CysZ (*Pd*CysZ) hexamer.** a. Side and top
753 views of the hexamer as a ribbon diagram with each protomer chain colored differently. The
754 approximate dimensions of the hexamer marked in Å. b. Side and top views represented by
755 surface electrostatics as calculated by APBS, with negative and positive surface potential
756 represented in red and blue respectively. c. Side and top views representing conservation of
757 residues as calculated by ConSurf, with maroon being most conserved to cyan being least
758 conserved.

759

760 **Figure 1 – Figure Supplement 1.**

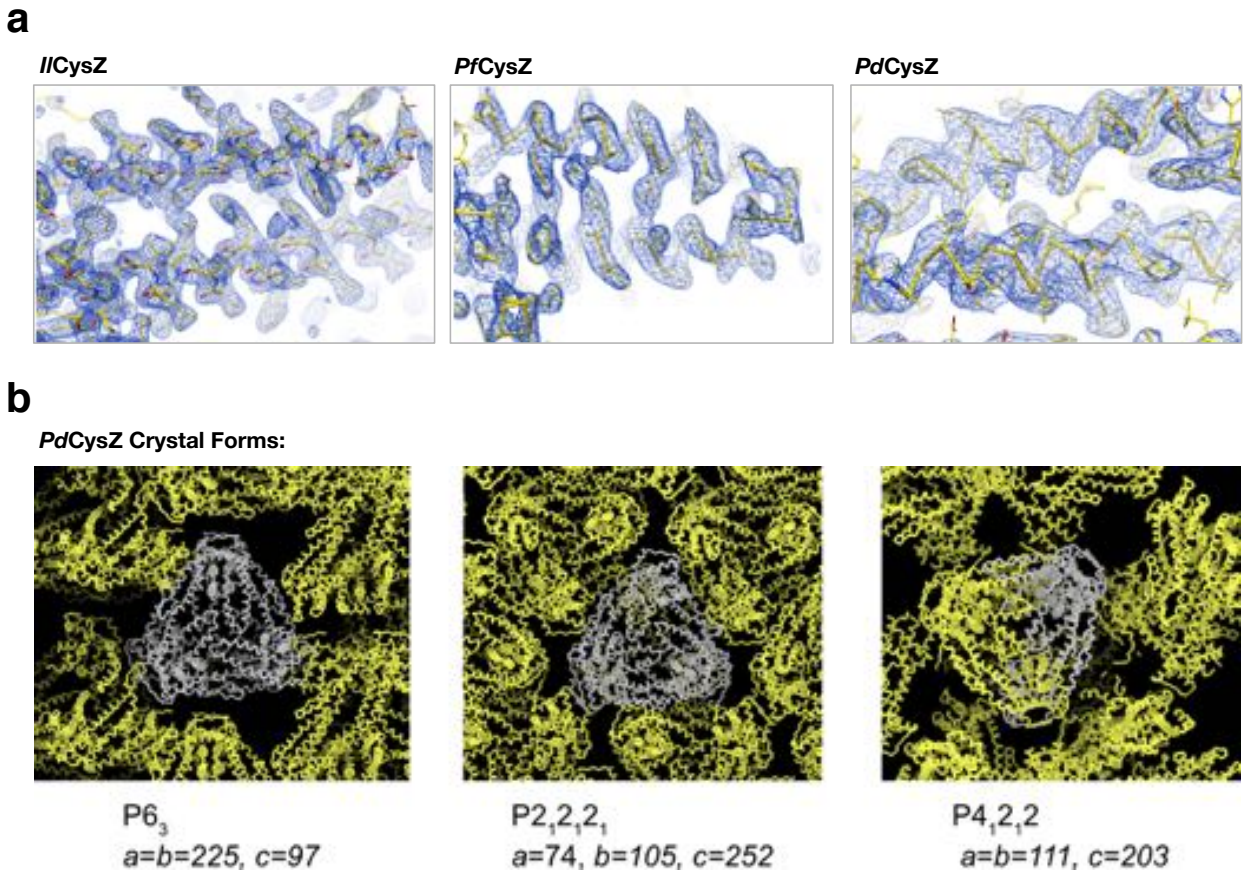


761

762

763 **Fig. 1. S1. Schematic of assimilatory sulfate reduction in bacteria for cysteine biosynthesis.**
764 Assimilatory reduction is energetically dispendious as sulfate is extremely stable. In this
765 pathway, sulfate ions enter the cell via a transporter or channel, like CysZ; this is followed by its
766 activation by ATP-sulfurylase, which utilizes ATP to form adensosine phosphosulfate (APS) and
767 inorganic pyrophosphate (PPi) as a by-product. APS is then further activated by the addition of a
768 second phosphate forming phosphoadenosine phosphosulfate (PAPS) and this high-energy
769 intermediate product is poised for sulfur removal to form sulfite ($SO3^{2-}$) by the thioredoxin
770 enzyme, which is then further reduced to sulfide ions (S^-). Sulfide, the final reduced sulfur
771 product, can then be used in a variety of ways by the cell, such as the incorporation into O-
772 acetylserine to synthesize cysteine.

773 **Figure 1 – Figure Supplement 2.**

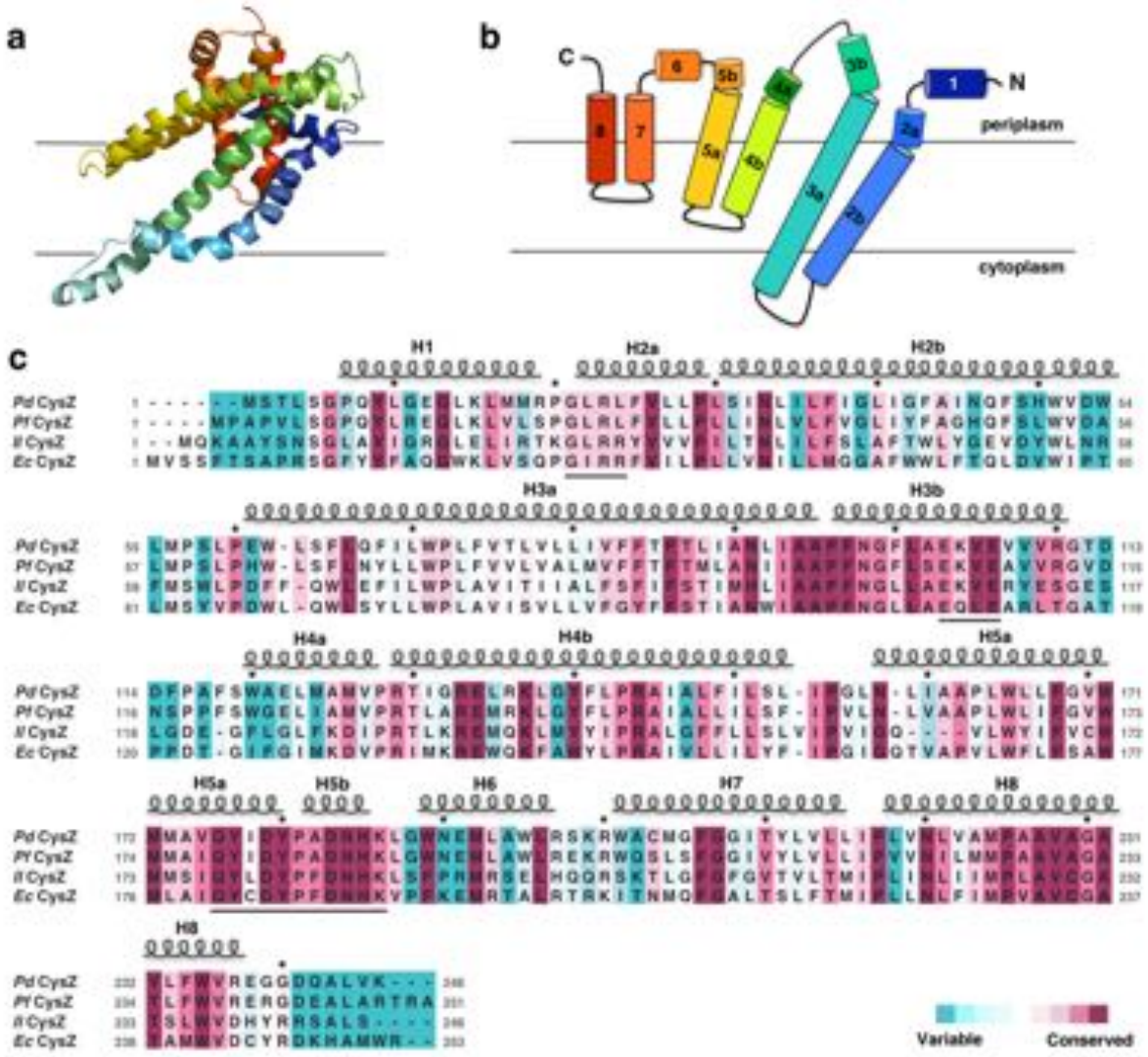


774

775

776 **Fig. 1. S2. Representative electron density of the crystal structures of *Il*CysZ, *Pf*CysZ and**
777 ***Pd*CysZ, and the different crystal forms observed for *Pd*CysZ.** a. Representative electron
778 density in blue mesh for: *Il*CysZ, contoured at 1 x r.m.s.d., *Pf*CysZ, contoured at 2 x r.m.s.d, with
779 a negative B-factor of -100 Å^2 applied to map**;** *Pd*CysZ, contoured at 1 x r.m.s.d, with a negative
780 B-factor of -100 Å^2 applied to map. Final atomic models (Cα trace with side chains depicted) are
781 shown as yellow stick representation. b. Preliminary molecular replacement analyses of the 3
782 different crystal forms obtained for *Pd*CysZ, showing the hexameric assembly of the molecule,
783 seen in each case. The Cα trace of an asymmetric unit in each case is depicted in grey, with their
784 symmetry mates are shown in yellow. Space group and unit cell dimensions are listed below.

785 **Figure 2.**



786

787

788 **Fig. 2. Structure and topology diagram of *Pd*CysZ protomer, with sequence alignment and**

789 **conservation** a. Ribbon diagram of the *Pd*CysZ protomer colored in rainbow colors from N

790 (blue) to C terminus (red), viewed from within the plane of the membrane, shown in the same

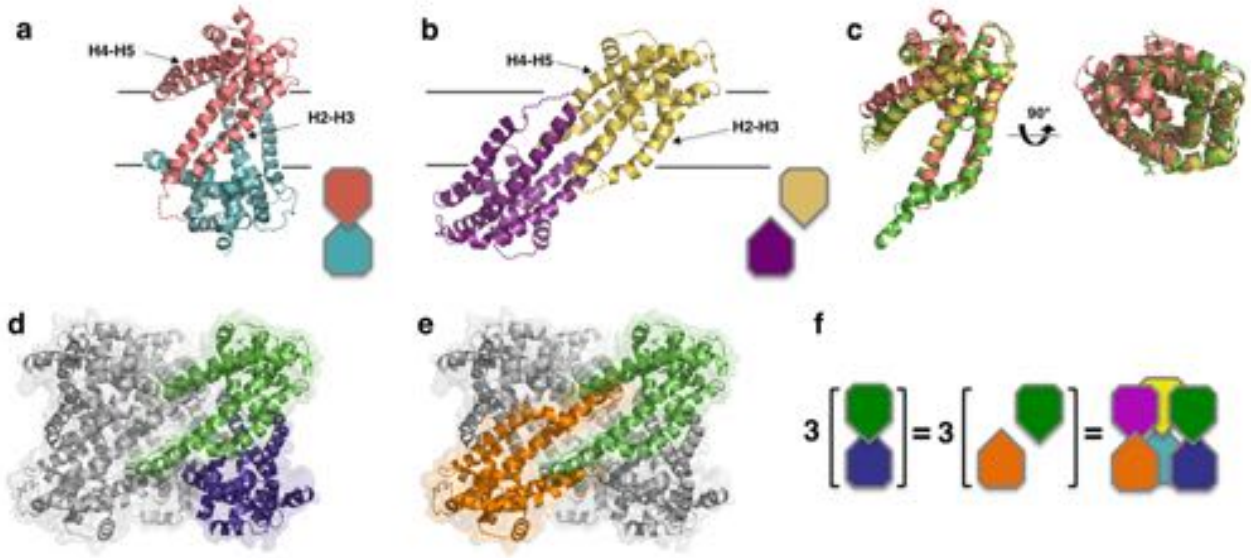
791 orientation as the protomer drawn in green in Fig. 1a. b. Topology diagram of the *Pd*CysZ

792 protomer with helices marked from 1 to 8. Helices H2b and H3a are transmembrane helices,

793 whereas H4-H5 and H7-H8 are hemi-penetrating helical hairpins, only partially inserted into the

794 membrane. c. Sequence alignment of *E. coli* CysZ (*Ec*CysZ), *P. denitrificans* CysZ (*Pd*CysZ), *P.*
795 *fragi* CysZ (*Pf*CysZ), and *I. loihiensis* CysZ (*Il*CysZ). Residues are colored based on
796 conservation, with maroon being most conserved and cyan least conserved, as calculated by
797 ConSurf using a sequence alignment of 150 non-redundant sequences from the CysZ family as
798 input. Spirals above residues mark the extent of the helical segments based on the atomic
799 structure of *Pd*CysZ with helices numbered H1-H8; letters mark residue identities; black dots
800 above identify every tenth residue (modulo 10) in the *Pd*CysZ sequence and black underlines
801 mark functionally relevant motifs discussed in the text.

802 **Figure 3.**



803

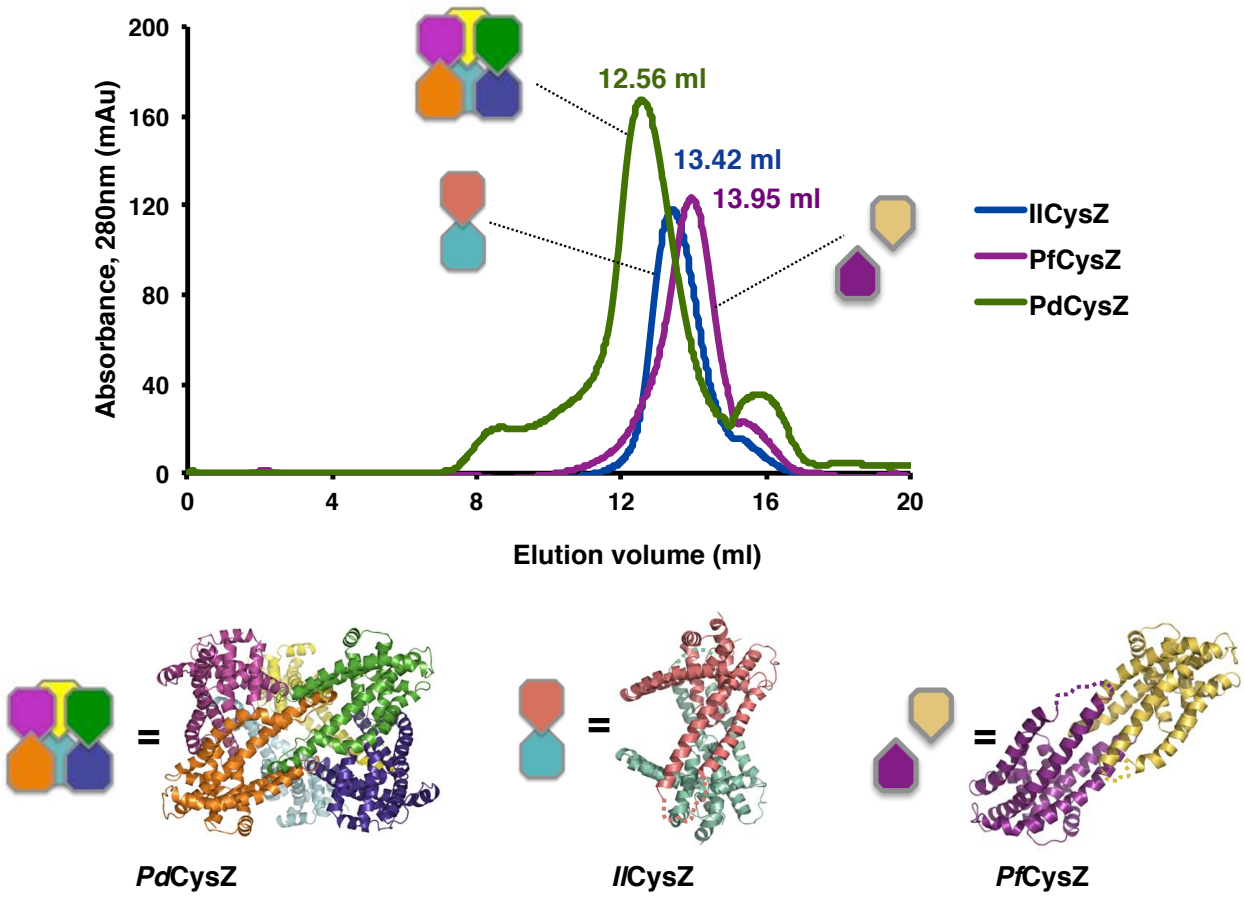
804

805 **Fig. 3. Structures of *Il*CysZ and *Pf*CysZ, with comparison to *Pd*CysZ** a. Ribbon diagram of
806 the structure of CysZ from *I. loihiensis (Il*CysZ) at 2.3 Å. Protomers of the dimer are colored in
807 salmon pink and teal blue, arranged in a head-to-tail conformation in the membrane, with helical
808 hairpins H2-3 and H4-5 labeled for clarity. The dimer interface of *Il*CysZ involves H2-3. b.
809 Ribbon diagram of the structure of CysZ from *P. fragi (Pf*CysZ*)* at 3.2 Å. The protomers of the
810 dual topology dimer in the membrane are colored gold and purple, with helical hairpins H2-3 and
811 H4-5 again labeled. The dimer interface here involves the interaction of helices H4-5 of each
812 protomer. c. Side and top views of the superposition of the three different protomers from
813 *Pd*CysZ (green), *Il*CysZ (pink) and *Pf*CysZ (yellow) after aligning H1-H3. d, e. The same dimer
814 interfaces observed in *Il*CysZ (a) and *Pf*CysZ (b) observed in the hexameric assembly of
815 *Pd*CysZ, as highlighted in green and blue (d) and green and orange (e). f. Schematic
816 representation showing how three copies of the dimeric protomers of *Il*CysZ (green and blue,

817 left) and of *Pf*CysZ (orange and green, center), can coexist in and each recapitulate the
818 hexameric assembly of *PdCysZ*.

819 **Figure 3 – Figure Supplement 1.**

820



821

822

823 **Fig. 3. S1. Size-exclusion chromatography of CysZ shows a mono-disperse elution profile**

824 **for each of the three species purified - *Pd*CysZ, *Pf*CyZ and *Il*CysZ.** All three proteins were

825 solubilized from isolated membrane fractions purified in the presence of 0.2% decyl

826 maltopyranoside (DeM) and exchanged into buffer containing 1% β-octyl glucopyranoside (β-

827 OG) on the size-exclusion column (Superdex 200 10/30 HR). A schematic of the oligomeric

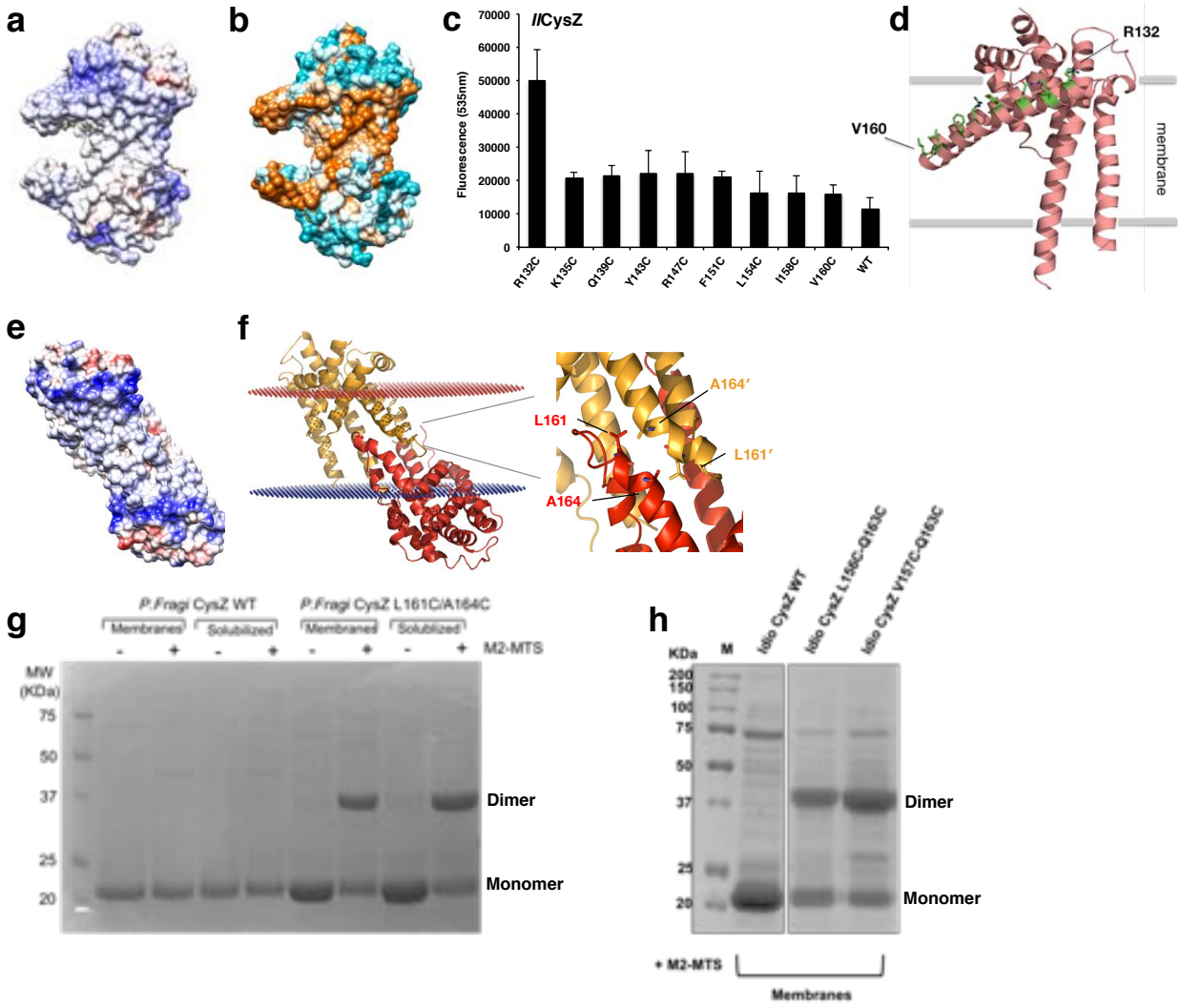
828 states observed in the three crystal structures of CysZ are shown below the elution profiles.

829 Crystallization trials were set-up directly after the size-exclusion chromatography step. A shift of

830 ~1 ml was observed for the elution of *Pd*CysZ (peak at 12.56 ml), with respect to *Pf*CysZ and
831 *Il*CysZ, indicating that the size (volume) and shape of *Pd*CysZ is significantly larger than the
832 other two species. *Il*CysZ elutes at 13.42 ml, 0.5 ml ahead of *Pf*CysZ, which could be explained
833 by the shape of *Il*CysZ, occupying more volume as compared to *Pf*CysZ, with its helices (H4-
834 H5) pointing outward, away from the body of the dimer.

835

836 **Figure 3 – Figure Supplement 2.**



837

838

839 **Fig. 3. S2. Surface electrostatics, hydrophobicity and site-directed fluorescence labeling of**

840 **cysteine mutants located on H4 of *Il*CysZ.** a. A depiction of the surface electrostatic potentials

841 of the *Il*CysZ dimer with negative surface potential represented in red, and positive potential in

842 blue as calculated by APBS. b. Surface of *Il*CysZ rendered by level of hydrophicity, using the

843 Kyte-Doolitle hydrophobicity scale with cyan being most polar, to orange being most

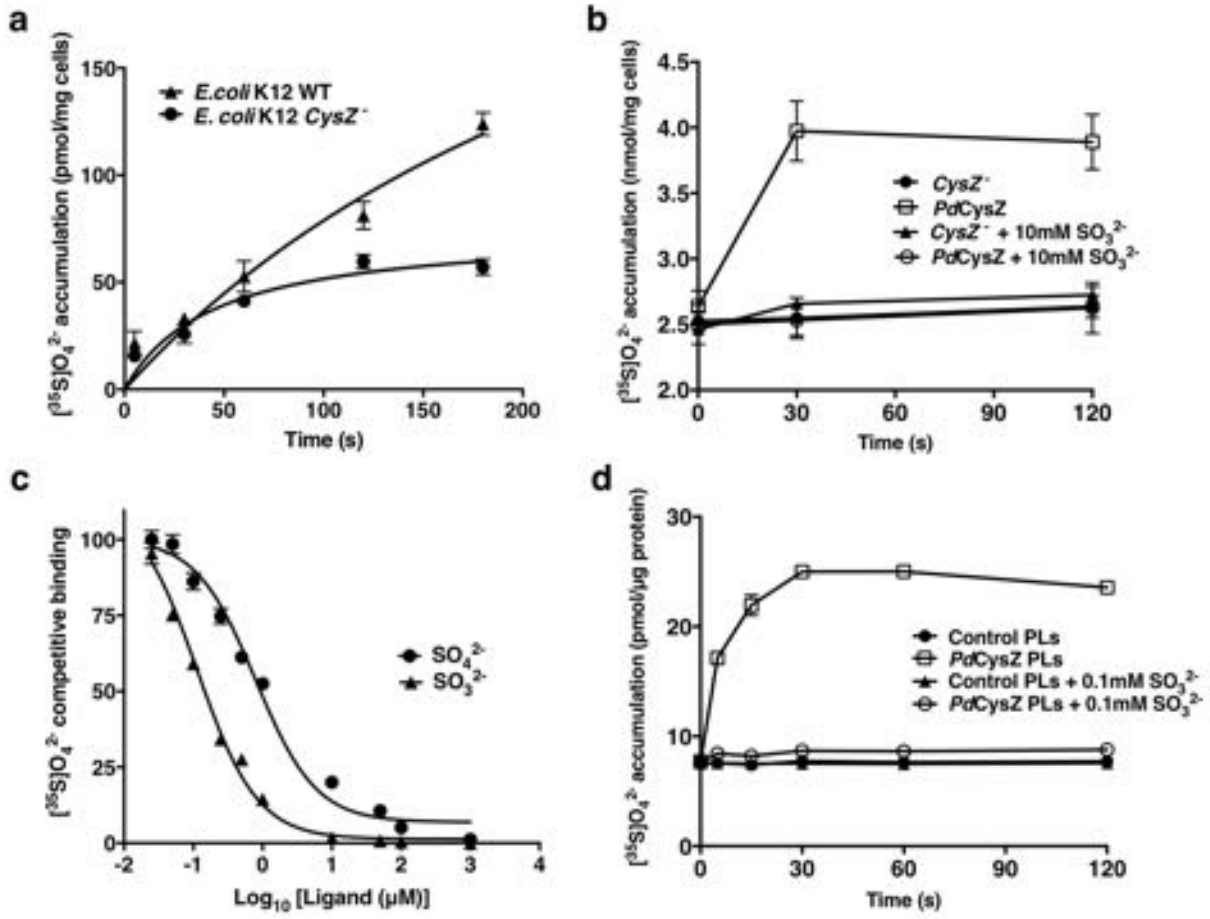
844 hydrophobic, reveals a hydrophobic belt along the center of the dimer, suggesting the orientation

845 of the dimer in the lipid bilayer. c, d. Single cysteine mutants were designed to be located along

846 the length of helix 4 of CysZ, to gauge the extent of its membrane insertion. The labeled protein
847 was extracted from the membrane and purified after quenching the labeling reaction.
848 Fluorescence intensity was measured and quantified by a Tecan fluorescence plate reader at an
849 excitation of 485nm and emission of 535nm the results of which were plotted on the left, with
850 error bars representing the standard deviation from the mean for n=3. The results show that of all
851 the residues on the helix, only R132C (the top-most residue) was accessible to the fluorophore
852 and hence exposed out of the membrane. d. Locations of the mutated residues on helix H4 are
853 marked on the model of *IlCysZ* in green. Crosslinking of L161C-A164C cysteine mutant of
854 *Pf*CysZ and *Il*CysZ exhibits dimer formation. e. A depiction of the surface electrostatic
855 potentials of the *Pf*CysZ dimer with negative surface potential represented in red, and positive
856 potential in blue as calculated by APBS, highlights hydrophobic belt marking the orientation of
857 the *Pf*CysZ dimer in the membrane. f. Membrane orientation of *Pf*CysZ predicted by OPM/PPM
858 server shows its 31-degree tilt to the perpendicular, with a zoomed in view of the location of the
859 cysteine mutants at the dimer interface, used for crosslinking of the dimer. g. h. Dimer formation
860 by the crosslinking of *Pf*CysZ (g) and *Il*CysZ (h) with introduced structure-based cysteine
861 substitution mutations in the H4-H5 dimer interface. Analogous pairs of cysteine mutants from
862 each species are shown here: *Pf*CysZ L161C-A164C and *Il*CysZ L156C-Q163C and V157C-
863 Q163C. Sulfydryl specific Bis-MTS crosslinkers of a certain spacer length (in this case 5.2Å)
864 were used at 0.5mM for 1hour at room temperature to crosslink the protein. Experiment was
865 performed on protein in the membrane, as well as on solubilized protein. WT protein (cysteine-
866 less) was used as a control. Proteins were purified after stopping the reaction, and run on an SDS
867 PAGE with reducing dye, and stained with Coomassie blue.
868

869 **Figure 4.**

870



871

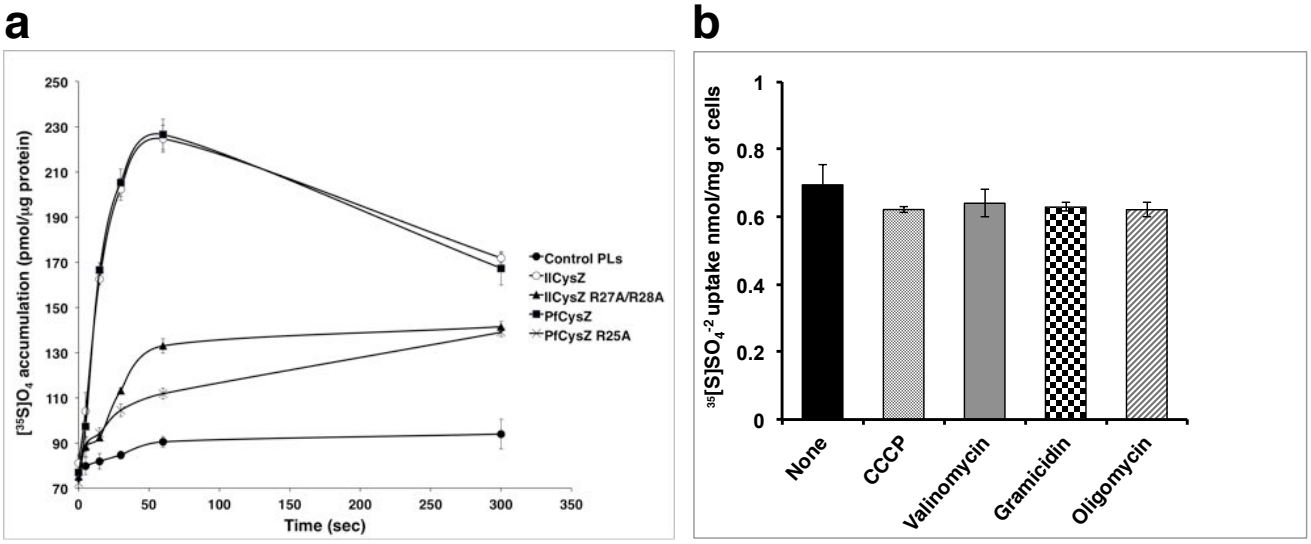
872

873 **Fig. 4. Functional characterization of CysZ.** a. Time course of $[^{35}S]O_4^{2-}$ uptake (320 μM) by
874 cells of WT *E. coli K12* (strain BW25113) and by CysZ$^-$ (strain JW2406-1) cells (n=3). b. Sulfite
875 inhibition of sulfate uptake. Radiolabeled sulfate (320 μM $[^{35}S]O_4^{2-}$) uptake is rescued in the
876 same *cysZ*$^-$ knockout strain by transformation with an expression vector for *Pd*CysZ whereas the
877 control cells were transformed with an empty vector. Sulfate uptake is inhibited by the addition
878 of 10 mM SO_3^{2-} (n=3). c. Competitive binding of labeled sulfate in presence of unlabeled
879 ligands. The inhibition of labeled sulfate (7 nM $[^{35}S]O_4^{2-}$) binding by non-labeled sulfate or

880 sulfite is measured by the scintillation proximity assay using purified *Pd*CysZ (n=3). *Pd*CysZ
881 binds SO_4^{2-} with an EC_{50} of 0.81 ± 0.045 μM and exhibits an IC_{50} of 0.12 ± 0.034 μM for SO_3^{-2}.
882 d. Radiolabeled sulfate (10 μM $[^{35}S]O_4^{2-}$) accumulation into proteoliposomes reconstituted with
883 detergent-purified *Pd*CysZ. Uptake into reconstituted proteoliposomes (open squares) is shown
884 in comparison with control (empty) liposomes (solid discs). Uptake is inhibited by the presence
885 of 0.1 mM SO_3^{2-} (n=3).
886

887 **Figure 4 – Figure Supplement 1.**

888



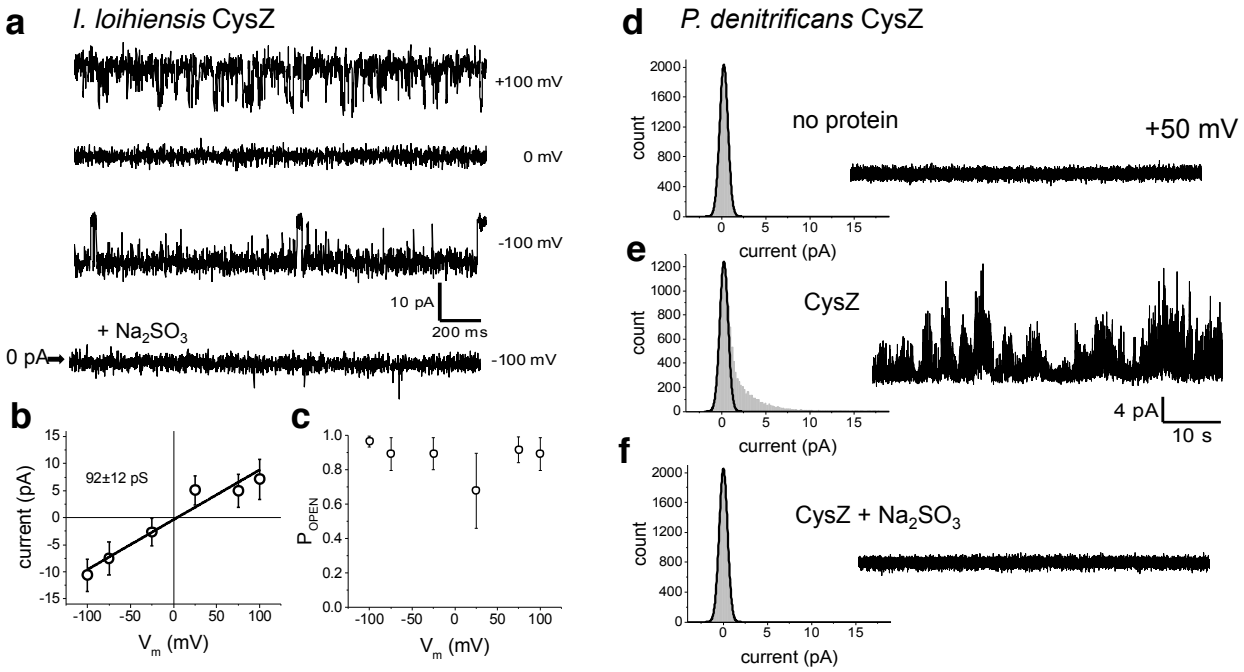
889

890

891 **Fig. 4. S1. a. Radiolabeled sulfate (10 μM [^{35}S]O_4^{2-}) accumulation is measured from**
892 **detergent purified *Il*CysZ and *Pf*CysZ reconstituted in proteoliposomes, in comparison to**
893 **the control (empty) liposomes.** Sulfate-binding site alanine mutants (*Il*CysZ R27A/R28A and
894 *Pf*CysZ R25A) display a significantly diminished sulfate uptake capability. Error bars represent
895 the standard deviation from the mean (SEM), for n=3. b. Effect of uncouplers and ionophores on
896 the sulfate accumulation by CysZ. *Il*CysZ expressed in the *E. coli* knockout strain shows
897 unchanged sulfate uptake levels in the presence of the uncouplers CCCP and oligomycin or the
898 ionophores gramicidin and valinomycin. Sulfate accumulation was measured in triplicate at pH 7
899 and total accumulation was plotted after a 10 s incubation. Error bars were calculated by standard
900 deviation from the mean, for n=3.

901

902 **Figure 4 – Figure Supplement 2.**



903

904

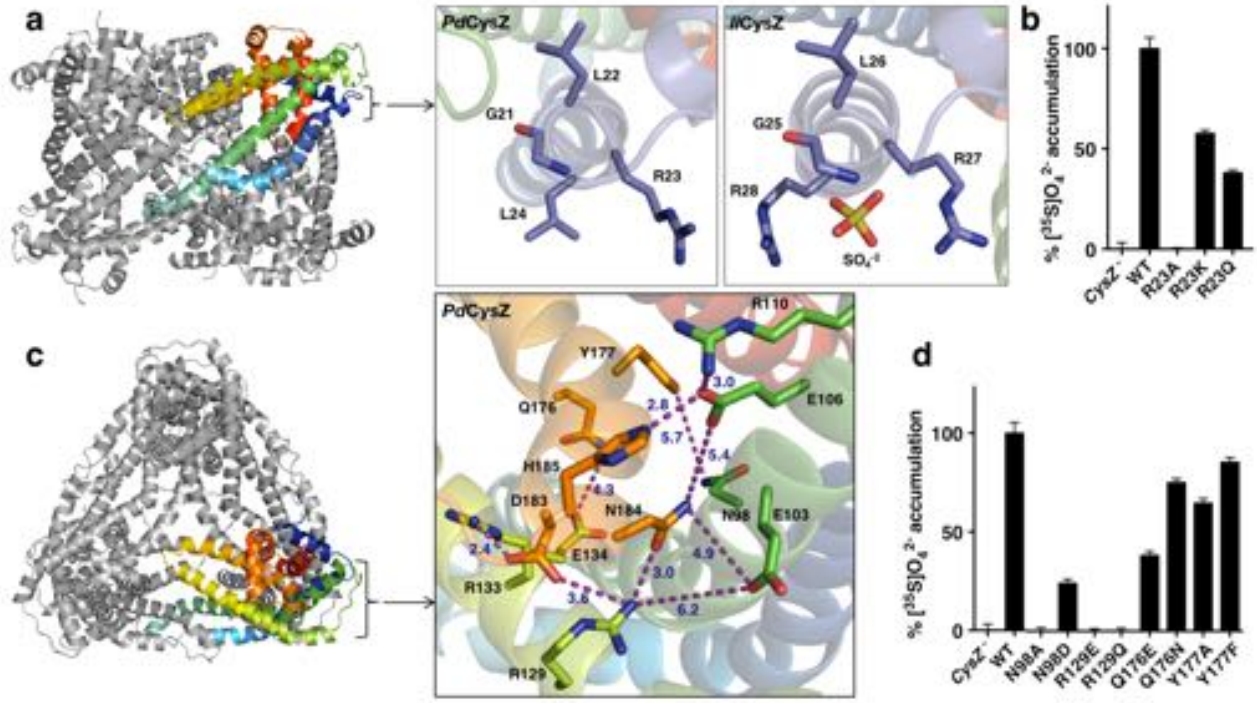
905 **Fig. 4. S2. Reconstitution of CysZ in planar lipid bilayers yields sulfite-sensitive**
906 **conductances.** a. Sample traces of *Il*CysZ under 150 mM symmetrical Na_2SO_4 solutions at the
907 indicated membrane potentials and pH = 5.4. The lowest trace shows inhibition of activity by 15
908 μM Na_2SO_3. b. Unitary conductance estimated at 92 ± 12 pS with a reversal potential (E_{rev}) at 0
909 mV. c. Unitary open probability does not change as a function of membrane potential, and is
910 stable at an average P_{OPEN}= 0.8. d. Stable bilayer recording at a membrane potential V_m = +50
911 mV, with no protein reconstituted. An all point current amplitude histogram is plotted on the left
912 of the trace. The peak of the normal distribution denotes the baseline zero current. e.
913 Reconstitution of *Pd*CysZ under 150 mM symmetrical Na_2SO_4 solutions at +50 mV. An all point
914 current amplitude histogram is plotted on the left of the trace. The high frequency activity shows
915 a poorly resolved distribution of unitary currents above baseline. f. Inhibition of activity by 5 μM

916 Na_2SO_3. An all point current amplitude histogram is plotted on the left of the trace. Sulfite

917 inhibited activity back to baseline levels.

918

919 **Figure 5.**



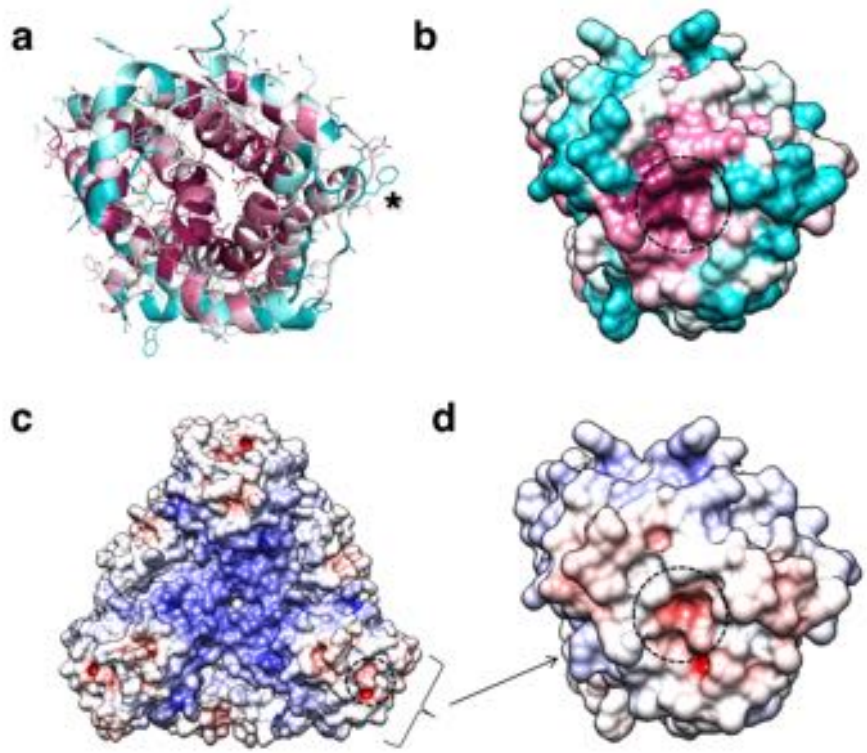
920

921

922 **Fig. 5. Functionally relevant features of the CysZ molecule** a. Side-view of a ribbon diagram
923 of the *Pd*CysZ hexamer, with one chain rainbow-colored from N- (blue) to C-termini (red).
924 Insets show the sulfate-binding site located at the start of H2a in *Pd*CysZ (left), with conserved
925 residues G21, L22, R23 and L24 labeled, and the corresponding site in *Il*CysZ (right), with
926 residues G25, L26, R27, and R28 labeled and showing the SO_4^{-2} ion as bound in the crystal
927 structure. b. Sulfate uptake by sulfate-binding site mutants of *Pd*CysZ. *E. coli K12* CysZ$^-$ cells
928 transformed with the listed *Pd*CysZ R23 mutants were used to measure $[^{35}S]O_4^{-2}$ uptake (n=3).
929 Sulfate uptake was abolished for R23A, and rescued to 50% and 40% of the wild type levels for
930 the R23K and R23Q mutants respectively. c. A top-view of the *Pd*CysZ hexamer, colored as in a.
931 The inset magnifies the central core of CysZ to show the associated network of hydrogen bonds
932 (R129-N184, E106-H185), van der Waals interactions (E103-N184, E106-N184, Q176-E134)

933 and salt bridges (R129-D183, R133-D183, R110-E106) between pairs of highly conserved
934 residues. Interatomic contacts are shown as purple dotted lines with distances (in Å) marked in
935 blue. d. Sulfate uptake by central-core mutants of *Pd*CysZ. *E. coli K12* CysZ$^-$ cells transformed
936 with the listed *Pd*CysZ mutants were used to measure [^{35}S]O_4^{-2} uptake (n=3). N98A and R129E
937 and R129Q showed severely impaired sulfate uptake, whereas more conservative substitutions
938 such as N98D, Q176E and Q176N had less of a negative effect on function. Y177A and Y177F
939 do not show any impaired function.
940

941 **Figure 6.**



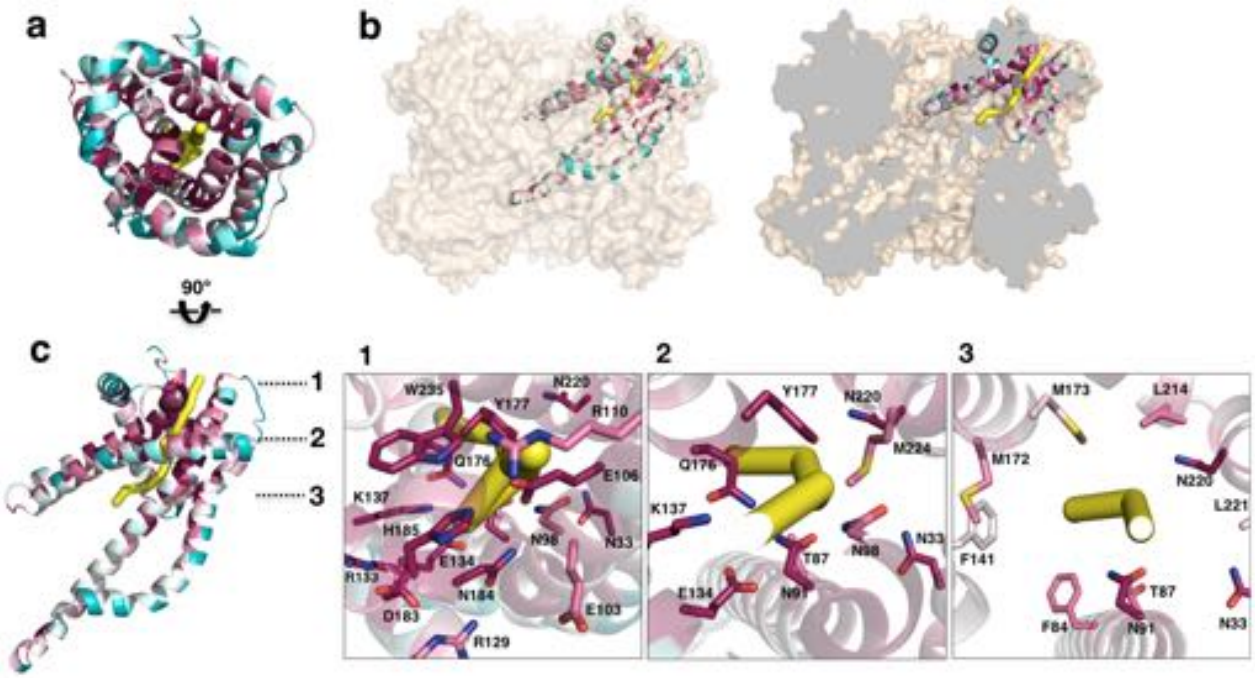
942

943

944 **Fig. 6. Conservation and entrance of putative pore of *Pd*CysZ.** a. Ribbon diagram colored by
945 conservation with residues in maroon being most conserved to cyan being least conserved
946 (calculated by ConSurf) to highlight the entrance to the putative pore; an asterisk (*) marks the
947 location of the sulfate-binding site (GLR motif) at top of H2a. b. Same view and coloring scheme
948 as in a, but now shown in surface representation. c. Electrostatic representation of hexameric
949 *Pd*CysZ as viewed from the top, with negative surface potential represented in red, and positive
950 potential in blue as calculated by APBS, with location of the putative pore marked by a dashed
951 circle. d. Close-up view in electrostatic representation of the putative pore within a *Pd*CysZ
952 protomer, surface and orientation as in b.

953

954 **Figure 7.**



955

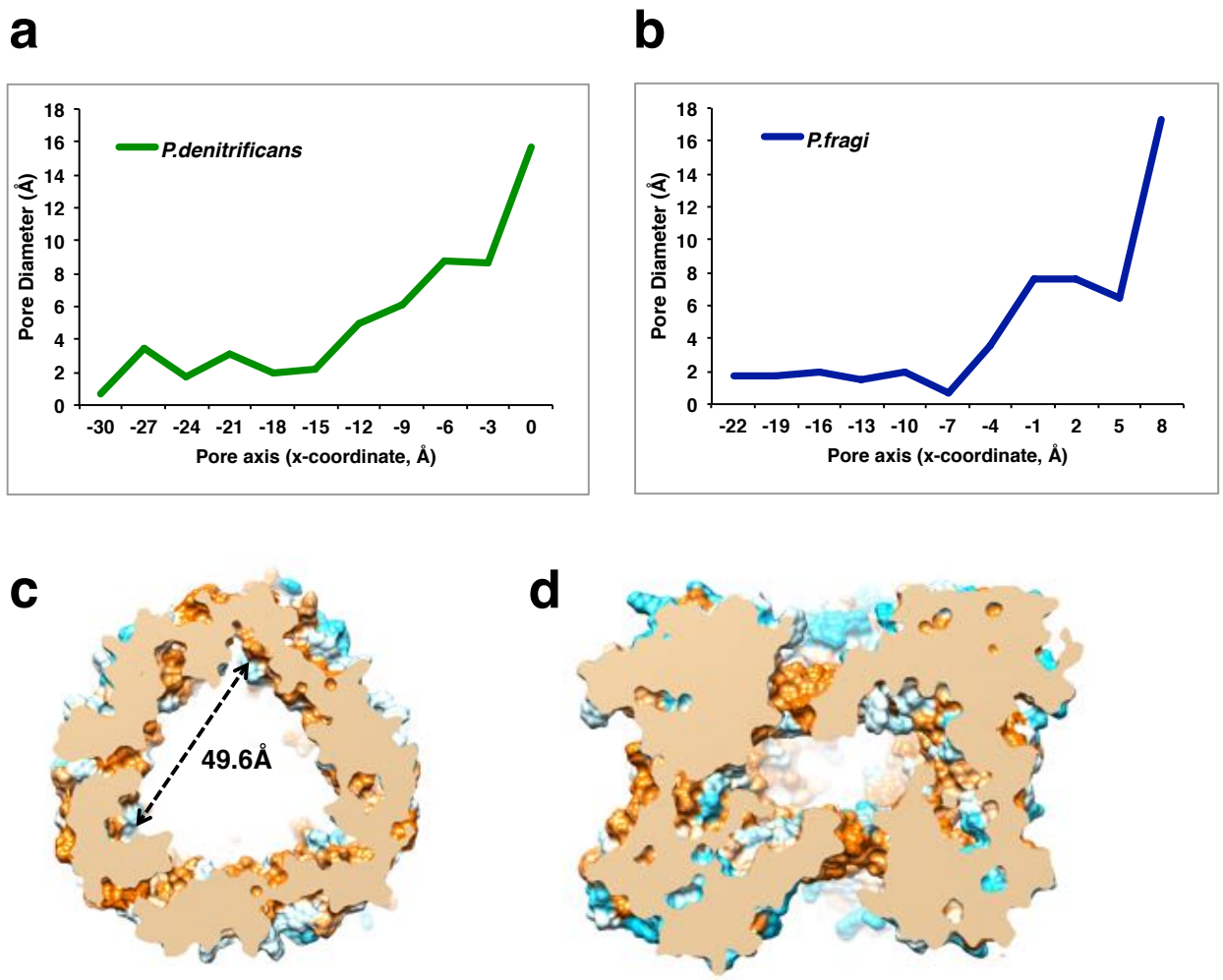
956

957 **Fig. 7. Putative ion conductance pathway of CysZ.** a. A *Pd*CysZ protomer looking into the
958 incipient pore entrance. The polypeptide ribbon is oriented as in 6a and colored by level of
959 sequence conservation (calculated by ConSurf, with maroon being most conserved and cyan
960 least). The putative pore and ion conduction pathway is shown as a yellow tube of 1 Å diameter,
961 calculated by PoreWalker. b. Surface representation of the *Pd*CysZ hexamer, as side views, both
962 semi-transparent (left) and with surface clipped (right, cut surfaces colored in grey) to allow for
963 the internal visualization of the pathway leading to the central cavity. c. Side view of a *Pd*CysZ
964 protomer left, viewed as in b and rotated 90° from a. Insets show magnified views of cross-
965 sections along the pathway: At level 1, the entrance to the pore consists of a narrow constriction
966 created by a network of highly conserved polar and charged residues (E106, R110, E134, N184,
967 H185) tightly interacting with one another. At level 2, the pathway broadens and becomes less

968 charged, but yet polar in nature as lined by conserved asparagine, tyrosine and threonine residues
969 (N33, T87, N91, N98, Q176, Y177). From level 3, the pathway widens further and ultimately
970 leads into the large central hydrophobic cavity.
971

972 **Figure 7 – Figure Supplement 1.**

973



974
975

976 **Fig. 7. S1. Plot of the pore diameter of the putative ion conduction pathway.** a. *Pd*CysZ b.
977 *Pf*CysZ, showing a very narrow constriction of under 2 Å in diameter near the entrance of the
978 pore, which then widens as the ions move past the core of the protomer to the inner central cavity
979 of the hexameric channel (as calculated by PoreWalker, measured at 3 Å steps along the x-axis)
980 c. Central hydrophobic cavity of *Pd*CysZ hexamer, cross-section of central cavity of the
981 hexameric *Pd*CysZ (top view) and d. side view, shows a triangular-shaped cavity with each side

982 measuring 49.6 Å in length. The cavity is primarily hydrophobic, as depicted and colored with
983 the Kyte-Doolitle hydrophobicity scale from most polar colored in cyan to most hydrophobic
984 colored in orange.

985

986 **References:**
987
988 Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J.,
989 Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W.,
990 Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., &
991 Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for
992 macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 213-
993 221. doi:S0907444909052925 [pii]
994 10.1107/S0907444909052925
995 Aguilar-Barajas, E., Diaz-Perez, C., Ramirez-Diaz, M. I., Riveros-Rosas, H., & Cervantes, C.
996 (2011). Bacterial transport of sulfate, molybdate, and related oxyanions. *Biometals, 24*(4),
997 687-707. doi:10.1007/s10534-011-9421-x
998 Akabas, M. H., Stauffer, D. A., Xu, M., & Karlin, A. (1992). Acetylcholine receptor channel
999 structure probed in cysteine-substitution mutants. *Science, 258*(5080), 307-310.
1000 Amadi, S. T., Koteiche, H. A., Mishra, S., & McHaourab, H. S. (2010). Structure, dynamics, and
1001 substrate-induced conformational changes of the multidrug transporter EmrE in
1002 liposomes. *J Biol Chem, 285*(34), 26710-26718. doi:10.1074/jbc.M110.132621
1003 Anishkin, A., Akitake, B., Kamaraju, K., Chiang, C. S., & Sukharev, S. (2010). Hydration
1004 properties of mechanosensitive channel pores define the energetics of gating. *J Phys*
1005 *Condens Matter, 22*(45), 454120. doi:10.1088/0953-8984/22/45/454120
1006 Aryal, P., Sansom, M. S., & Tucker, S. J. (2015). Hydrophobic Gating in Ion Channels. *Journal*
1007 *of Molecular Biology, 427*(1), 121-130. doi:10.1016/j.jmb.2014.07.030
1008 Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016).
1009 ConSurf 2016: an improved methodology to estimate and visualize evolutionary
1010 conservation in macromolecules. *Nucleic Acids Res, 44*(W1), W344-350.
1011 doi:10.1093/nar/gkw408
1012 Aslanidis, C., & de Jong, P. J. (1990). Ligation-independent cloning of PCR products (LIC-
1013 PCR). *Nucleic Acids Res, 18*(20), 6069-6074.
1014 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita,
1015 M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame,
1016 single-gene knockout mutants: the Keio collection. *Mol Syst Biol, 2*, 2006 0008.
1017 doi:10.1038/msb4100050
1018 Barton, L. L. (2005) *Structural and Functional Relationships in Prokaryotes* (pp. 680-720):
1019 Springer.
1020 Bass, R. B., Strop, P., Barclay, M., & Rees, D. C. (2002). Crystal structure of Escherichia coli
1021 MscS, a voltage-modulated and mechanosensitive channel. *Science, 298*(5598), 1582-
1022 1587. doi:10.1126/science.1077945
1023 Birkner, J. P., Poolman, B., & Kocer, A. (2012). Hydrophobic gating of mechanosensitive
1024 channel of large conductance evidenced by single-subunit resolution. *Proc Natl Acad Sci*
1025 *U S A, 109*(32), 12944-12949. doi:10.1073/pnas.1205270109
1026 Britton, P., Boronat, A., Hartley, D. A., Jones-Mortimer, M. C., Kornberg, H. L., & Parra, F.
1027 (1983). Phosphotransferase-mediated regulation of carbohydrate utilization in
1028 Escherichia coli K12: location of the gsr (tgs) and iex (crr) genes by specialized
1029 transduction. *J Gen Microbiol, 129*(2), 349-356.

1030 Buckler, K. J., & Vaughan-Jones, R. D. (1998). Effects of mitochondrial uncouplers on
1031 intracellular calcium, pH and membrane potential in rat carotid body type I cells. *J*
1032 *Physiol, 513 (Pt 3)*, 819-833.
1033 Chen, Y. H., Hu, L., Punta, M., Bruni, R., Hillerich, B., Kloss, B., Rost, B., Love, J.,
1034 Siegelbaum, S. A., & Hendrickson, W. A. (2010). Homologue structure of the SLAC1
1035 anion channel for closing stomata in leaves. *Nature, 467*(7319), 1074-1080.
1036 doi:10.1038/nature09487
1037 nature09487 [pii]
1038 Cowtan, K. D., & Zhang, K. Y. (1999). Density modification for macromolecular phase
1039 improvement. *Prog Biophys Mol Biol, 72*(3), 245-270.
1040 Czodrowski, P., Dramburg, I., Sotriffer, C. A., & Klebe, G. (2006). Development, validation, and
1041 application of adapted PEOE charges to estimate pKa values of functional groups in
1042 protein-ligand complexes. *Proteins, 65*(2), 424-437. doi:10.1002/prot.21110
1043 Davis, B. D., & Mingioli, E. S. (1950). Mutants of Escherichia coli requiring methionine or
1044 vitamin B12. *J Bacteriol, 60*(1), 17-28.
1045 Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A.
1046 (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular
1047 structures for molecular simulations. *Nucleic Acids Res, 35*(Web Server issue), W522-
1048 525. doi:10.1093/nar/gkm276
1049 Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: an
1050 automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations.
1051 *Nucleic Acids Res, 32*(Web Server issue), W665-667. doi:10.1093/nar/gkh381
1052 Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B.
1053 T., & MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of
1054 K+ conduction and selectivity. *Science, 280*(5360), 69-77.
1055 Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of
1056 Coot. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 4), 486-501. doi:S0907444910007493
1057 [pii]
1058 10.1107/S0907444910007493
1059 Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr,*
1060 *62*(Pt 1), 72-82. doi:10.1107/S0907444905036693
1061 Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., & Ben-Tal, N. (2003).
1062 ConSurf: identification of functional regions in proteins by surface-mapping of
1063 phylogenetic information. *Bioinformatics, 19*(1), 163-164.
1064 Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of*
1065 *Molecular Biology, 166*(4), 557-580.
1066 Hryniewicz, M., Sirko, A., Palucha, A., Bock, A., & Hulanicka, D. (1990). Sulfate and
1067 thiosulfate transport in Escherichia coli K-12: identification of a gene encoding a novel
1068 protein involved in thiosulfate binding. *J Bacteriol, 172*(6), 3358-3366.
1069 Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 125-132.
1070 doi:10.1107/S0907444909047337
1071 Kertesz, M. A. (2000). Riding the sulfur cycle--metabolism of sulfonates and sulfate esters in
1072 gram-negative bacteria. *FEMS Microbiol Rev, 24*(2), 135-175.
1073 Kertesz, M. A. (2001). Bacterial transporters for sulfate and organosulfur compounds. *Res*
1074 *Microbiol, 152*(3-4), 279-290.

1075 Korkhov, V. M., & Tate, C. G. (2009). An emerging consensus for the structure of EmrE. *Acta*
1076 *Crystallogr D Biol Crystallogr, 65*(Pt 2), 186-192. doi:10.1107/S0907444908036640
1077 Kredich, N. M. (1971). Regulation of L-cysteine biosynthesis in Salmonella typhimurium. I.
1078 Effects of growth of varying sulfur sources and O-acetyl-L-serine on gene expression. *J*
1079 *Biol Chem, 246*(11), 3474-3484.
1080 Kredich, N. M., Hulanicka, M. D., & Hallquist, S. G. (1979). Synthesis of L-cysteine in
1081 Salmonella typhimurium. *Ciba Found Symp*(72), 87-99.
1082 Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline
1083 state. *Journal of Molecular Biology, 372*(3), 774-797 'Protein interfaces, surfaces and
1084 assemblies' service PISA at the European Bioinformatics Institute.
1085 (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). doi:10.1016/j.jmb.2007.05.022
1086 Langer, G., Cohen, S. X., Lamzin, V. S., & Perrakis, A. (2008). Automated macromolecular
1087 model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc, 3*(7),
1088 1171-1179. doi:nprot.2008.91 [pii]
1089 10.1038/nprot.2008.91
1090 Leal-Pinto, E., London, R. D., Knorr, B. A., & Abramson, R. G. (1995). Reconstitution of
1091 hepatic uricase in planar lipid bilayer reveals a functional organic anion channel. *J*
1092 *Membr Biol, 146*(2), 123-132.
1093 Liu, Q., Dahmane, T., Zhang, Z., Assur, Z., Brasch, J., Shapiro, L., Mancia, F., & Hendrickson,
1094 W. A. (2012). Structures from anomalous diffraction of native biological
1095 macromolecules. *Science, 336*(6084), 1033-1037. doi:10.1126/science.1218753
1096 Liu, Q., Zhang, Z., & Hendrickson, W. A. (2011). Multi-crystal anomalous diffraction for low-
1097 resolution macromolecular phasing. *Acta Crystallogr D Biol Crystallogr, 67*(Pt 1), 45-59.
1098 doi:10.1107/S0907444910046573
1099 Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I., & Lomize, A. L. (2012). OPM
1100 database and PPM web server: resources for positioning of proteins in membranes.
1101 *Nucleic Acids Res, 40*(Database issue), D370-376. doi:10.1093/nar/gkr703
1102 Loughlin, P., Shelden, M. C., Tierney, M. L., & Howitt, S. M. (2002). Structure and function of a
1103 model member of the SulP transporter family. *Cell Biochem Biophys, 36*(2-3), 183-190.
1104 doi:CBB:36:2-3:183 [pii]
1105 10.1385/CBB:36:2-3:183
1106 Love, J., Mancia, F., Shapiro, L., Punta, M., Rost, B., Girvin, M., Wang, D. N., Zhou, M., Hunt,
1107 J. F., Szyperski, T., Gouaux, E., MacKinnon, R., McDermott, A., Honig, B., Inouye, M.,
1108 Montelione, G., & Hendrickson, W. A. (2010). The New York Consortium on Membrane
1109 Protein Structure (NYCOMPS): a high-throughput platform for structural genomics of
1110 integral membrane proteins. *J Struct Funct Genomics, 11*(3), 191-199.
1111 doi:10.1007/s10969-010-9094-7
1112 Mancia, F., & Love, J. (2011). High throughput platforms for structural genomics of integral
1113 membrane proteins. *Curr Opin Struct Biol, 21*(4), 517-522. doi:10.1016/j.sbi.2011.07.001
1114 Mansilla, M. C., & de Mendoza, D. (2000). The Bacillus subtilis cysP gene encodes a novel
1115 sulphate permease related to the inorganic phosphate transporter (Pit) family.
1116 *Microbiology, 146 (Pt 4)*, 815-821.
1117 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R.
1118 J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography, 40*(Pt 4),
1119 658-674. doi:10.1107/S0021889807021206

Morera, F. J., Vargas, G., Gonzalez, C., Rosenmann, E., & Latorre, R. (2007). Ion-channel reconstitution. *Methods Mol Biol, 400*, 571-585. doi:10.1007/978-1-59745-519-0_38

Mueller, P., Rudin, D. O., Tien, H. T., & Wescott, W. C. (1962). Reconstitution of cell membrane structure in vitro and its transformation into an excitable system. *Nature, 194*, 979-980.

Otwinowski, Z., & Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology, 276*, 307-326.

Pape, T., & Schneider, T. R. (2004). HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Cryst., 37*, 843-844.

Parra, F., Britton, P., Castle, C., Jones-Mortimer, M. C., & Kornberg, H. L. (1983). Two separate genes involved in sulphate transport in Escherichia coli K12. *J Gen Microbiol, 129*(2), 357-358.

Pellegrini-Calace, M., Maiwald, T., & Thornton, J. M. (2009). PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure. *PLoS Comput Biol, 5*(7), e1000440. doi:10.1371/journal.pcbi.1000440

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem, 25*(13), 1605-1612. doi:10.1002/jcc.20084

Punta, M., Love, J., Handelman, S., Hunt, J. F., Shapiro, L., Hendrickson, W. A., & Rost, B. (2009). Structural genomics target selection for the New York consortium on membrane protein structure. *J Struct Funct Genomics, 10*(4), 255-268. doi:10.1007/s10969-009-9071-1

Quick, M., & Javitch, J. A. (2007). Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc Natl Acad Sci U S A, 104*(9), 3603-3608. doi:0609573104 [pii]
10.1073/pnas.0609573104

Rapp, M., Granseth, E., Seppala, S., & von Heijne, G. (2006). Identification and evolution of dual-topology membrane proteins. *Nat Struct Mol Biol, 13*(2), 112-116. doi:10.1038/nsmb1057

Rigaud, J. L., Pitard, B., & Levy, D. (1995). Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochim Biophys Acta, 1231*(3), 223-246.

Schrodinger, LLC. (2010). *The PyMOL Molecular Graphics System, Version 1.3r1*.

Sheldrick, G. M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 4), 479-485. doi:S0907444909038360 [pii]
10.1107/S0907444909038360

Sirko, A., Zatyka, M., Sadowy, E., & Hulanicka, D. (1995). Sulfate and thiosulfate transport in Escherichia coli K-12: evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *J Bacteriol, 177*(14), 4134-4136.

Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography, 41*, 641-643. doi:10.1107/S0021889808006985

1164 Stockbridge, R. B., Kolmakova-Partensky, L., Shane, T., Koide, A., Koide, S., Miller, C., &
1165 Newstead, S. (2015). Crystal structures of a double-barrelled fluoride ion channel.
1166 *Nature, 525*(7570), 548-551. doi:10.1038/nature14981
1167 Stockbridge, R. B., Robertson, J. L., Kolmakova-Partensky, L., & Miller, C. (2013). A family of
1168 fluoride-specific ion channels with dual-topology architecture. *Elife, 2*, e01084.
1169 doi:10.7554/eLife.01084
1170 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R.
1171 M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N.,
1172 Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S.
1173 (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol*
1174 *Crystallogr, 67*(Pt 4), 235-242. doi:10.1107/S0907444910045749
1175 Yang, T., Liu, Q., Kloss, B., Bruni, R., Kalathur, R. C., Guo, Y., Kloppmann, E., Rost, B.,
1176 Colecraft, H. M., & Hendrickson, W. A. (2014). Structure and selectivity in bestrophin
1177 ion channels. *Science, 346*(6207), 355-359. doi:10.1126/science.1259723
1178 Ye, L., Jia, Z., Jung, T., & Maloney, P. C. (2001). Topology of OxlT, the oxalate transporter of
1179 Oxalobacter formigenes, determined by site-directed fluorescence labeling. *J Bacteriol,*
1180 *183*(8), 2490-2496. doi:10.1128/JB.183.8.2490-2496.2001
1181 Zhang, L., Jiang, W., Nan, J., Almqvist, J., & Huang, Y. (2014). The Escherichia coli CysZ is a
1182 pH dependent sulfate transporter that can be inhibited by sulfite. *Biochim Biophys Acta,*
1183 *1838*(7), 1809-1816. doi:10.1016/j.bbamem.2014.03.003
1184