Structures of wild-type and a constitutively closed mutant of connexin26 shed light on channel regulation by CO_2

Deborah H. Brotherton[1], Sarbjit Nijjar[1], Christos G. Savva[2], Nicholas Dale[1]*, Alexander D. Cameron[1]*

[1]School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, U.K.

[2]Leicester Institute of Structural and Chemical Biology, Department of Molecular and Cell Biology, University of Leicester, Lancaster Road, Leicester, LE1 7HB, UK.

*Correspondence:

a.cameron@warwick.ac.uk

n.e.dale@warwick.ac.uk

Abstract

Connexins allow intercellular communication by forming gap junction channels (GJCs) between juxtaposed cells. Connexin26 (Cx26) can be regulated directly by CO_2. This is proposed to be mediated through carbamylation of K125. We show that mutating K125 to glutamate, mimicking the negative charge of carbamylation, causes Cx26 GJCs to be constitutively closed. Through cryo-EM we observe that the K125E mutation pushes a conformational equilibrium towards the channel having a constricted pore entrance, similar to effects seen on raising the partial pressure of CO_2. In previous structures of connexins, the cytoplasmic loop, important in regulation and where K125 is located, is disordered. Through further cryo-EM studies we trap distinct states of Cx26 and observe density for the cytoplasmic loop. The interplay between the position of this loop, the conformations of the transmembrane helices and the position of the N-terminal helix, which controls the aperture to the pore, provides a mechanism for regulation.

Introduction

Connexins form hexameric channels in the plasma membrane known as hemichannels, which can either function as regulated passageways between the cell and its environment, or dock with a hemichannel from another cell to form a dodecameric intercellular channel, or gap junction channel (GJC). Connexins have been shown to be directly regulated by various stimuli such as voltage [1,2], pH [3-5] or indirectly via intracellular calcium ion concentrations [6]. Recent reports based on structural data also suggest that lipids may be involved in regulation [7-9]. We have shown, however, that connexin26 (Cx26) and other similar β-connexins (Cx30, Cx32), can be regulated by the direct action of physiological concentrations of carbon dioxide independently of pH [10,11]. Mutants of Cx26 are a leading cause of congenital deafness [12]. While many of the mutations are non-syndromic, others lead to severe diseases such as keratitis ichthyosis deafness syndrome (KIDS) [12]. Hemichannels are known to have different properties to GJCs [13] and our previous results show that elevated levels of the partial pressure of CO_2 (PCO_2) open Cx26 hemichannels [10], which are closed in their resting state[14], but close Cx26 GJCs [15], which are naturally open.

There are 20 connexin genes in the human genome [16] and several structures have now been published [4,7,8,17-23]. The connexin subunit, which is common to all, consists of four transmembrane helices (TMs) with a cytoplasmic N-terminal helix that in the hexameric arrangement of the hemichannel points towards the central pore [23]. In structures of the dodecameric GJC, the extracellular part, involved in docking is well defined, whereas the cytoplasmic region is much more variable. A large cytoplasmic loop between TM2 and TM3, shown to be involved in regulation, has not been visible in any structure. Structures of the connexins either have the N-terminal helices tucked back against the wall of the channel [7,8,20,21], in a raised position [8,17,19], in an intermediate position [18,23] or not well-defined in the density [4,7,18,22] and the conformation of the helix is thought to be important in regulation. We have shown previously for human Cx26 GJCs, that the position of the N-terminus is dependent on the partial pressure of CO_2 (PCO_2) at constant pH [18]. Examining structures from protein vitrified at different levels of PCO_2, we observed that under conditions of high PCO_2 the conformation of the protein is biased towards a conformation where the N-terminus protrudes radially into the pore to form a constriction at the centre (Pore-Constricting N terminus; PCN). On the other hand, at low PCO_2 the N-terminus is less well-defined, appearing to have pulled back from the centre to give a more open channel (Pore Open Flexible N-terminus; POFN).

Based on a wealth of mutational data it has been hypothesised that the regulation of hemichannel opening by CO_2 is through a carbamylation reaction of a specific lysine [11,15,24]. This post-translational modification is a reversible and highly labile reaction of CO_2 [25] that effectively changes the charge of a neutral lysine residue to make it negative. A so-called "carbamylation motif" K_{125}VRIEG_{130} was identified in CO_2-sensitive connexins that when introduced into a related CO_2-insensitive connexin rendered the protein CO_2-sensitive [11]. In the crystal structure of the Cx26 GJC that was published in 2009 [23], Lys125, which is conserved amongst β-connexins that are known to be modulated by CO_2, is positioned near to the N-terminus of TM3 within ~6Å of Arg104 of TM2 of the neighbouring subunit, at either side of the disordered cytoplasmic loop. It was suggested that upon carbamylation, the negative charge of the modified lysine would attract Arg104 causing a conformational change [11]. In hemichannels, mutation of Lys125 to glutamate, so mimicking the charge of the carbamylated lysine (K125E) results in constitutively open hemichannels consistent with elevated PCO_2,

whereas the corresponding K125R mutation results in hemichannels that cannot be opened by CO_2 [11]. In GJCs, the K125R mutation desensitises the channels to CO_2 so that they do not shut, though importantly, does not prevent closure by acidification [15].

While our previous structures [18] demonstrated an effect of PCO_2 on the conformation of the protein, neither Lys125 nor Arg104 were visible in the density. Here we probe this further, with structures of mutants of Lys125. We show that the structure with the K125E mutation has a distinct conformational change consistent with the closure mechanism of the protein that we reported previously. By solubilising the protein in the detergent lauryl maltose neopentyl glycol (LMNG) rather than dodecyl β-D-maltoside (DDM) we obtain much more defined density for the cytoplasmic region of the protein. We refine two conformationally diverse structures of the protein, where we see profound differences in the cytoplasmic region of the protein. Together our data suggests a mechanism for closure involving the concerted movements of TM2 and the KVRIEG motif. The introduction of a negative charge at K125 favours the closed conformation of the protein.

Results

Mutation of K125 to glutamic acid results in constitutively closed GJCs.

Based on our experience with mutations of Cx26 we hypothesised that, if K125E results in constitutively open hemichannels, the same mutation would result in constitutively closed GJCs. To verify the effect of mutating Lys125 to a glutamic acid, we used an established dye transfer assay between coupled cells to assess gap junction function [15]. For wild type Cx26, gap junctions readily form between cells and allow rapid transfer of dye from a donor (filled via a patch pipette) to a coupled acceptor cell at a PCO_2 of 35 mmHg (Fig 1a). $Cx26^{K125E}$ forms structures that resemble wild type gap junctions (Fig 1b). However, these gap junctions appear to be shut and do not permit dye transfer at a PCO_2 of 35 mmHg (Fig 1b,c). As the action of an increase in PCO_2 is to close the wild type Cx26 gap junction, unsurprisingly $Cx26^{K125E}$ gap junctions remain closed at a PCO_2 of 55 mmHg.

The K125E mutation biases the conformational equilibrium to a more closed structure

Given that the K125E mutant resulted in constitutively closed channels, we set out to solve its structure. With respect to the wild-type construct, purification of this protein resulted in higher yields, consistent with a more stable protein. Cryo-EM data were collected from protein vitrified in CO_2/HCO_3^- buffers corresponding to a PCO_2 of 90mmHg with the pH maintained at pH 7.4 as was done previously [18]. Refinement with D6 symmetry imposed resulted in a map with a nominal resolution of 2.2Å as defined by gold standard Fourier Shell Correlations [26,27] (Supplementary Table 1, Supplementary Figure 1). Superposition of this map onto the equivalent map from the wild-type structure at the same PCO_2 (WT_{90}; PDB ID 7QEQ) showed a clear change in the position of the cytoplasmic portion of TM2 (Figure 2a-b, Supplementary Movie S1).

Previously, when analysing the WT_{90} data, we used a classification procedure involving particle expansion and signal subtraction to focus on just one of the two docked hemichannels of each gap junction. This resulted in two distinct conformations of the protein ($WT_{90\text{-}Class1}$ with 40% of particles, (PDB ID 7QEY) and $WT_{90\text{-}Class2}$ with 21% of particles (7QEW)) with the predominant difference between the two in the cytoplasmic portion of TM2 [18]. In $WT_{90\text{-}Class2}$, compared to $WT_{90\text{-}Class1}$, an anticlockwise rotation of TM2 correlated with more definition of the density for the N-terminus. The overall conformation of the $K125E_{90}$ reconstruction was very reminiscent of the $WT_{90\text{-}Class2}$ conformation, indicating a clear bias towards this conformation. To improve the clarity of the cytoplasmic portion of the maps we used the same classification procedure on the $K125E_{90}$ data. Reconstructions from the most populated class (43% of the particles ($K125E_{90\text{-}1}$) have a nominal resolution of 2.5Å (Supplementary Figure 1) with a conformation that most resembles the $WT_{90\text{-}Class2}$. In the density map deriving from this class, the position of the N-terminus is reasonably well-defined as being in the PCN conformation (Supplementary Figure 2) but, as we found in our previous WT_{90} structures, the first three residues before Gly4 are difficult to position in the averaged density.

Relative to the $WT_{90\text{-}Class2}$ structure there are two main regions which have been modelled differently in the $K125E_{90\text{-}1}$ structure after refinement (Supplementary Table 2). This may be because a more optimal separation of the particles during classification has resulted in a clearer map. The first difference is at the cytoplasmic side of TM3. In all structures of connexins solved to date the cytoplasmic loop has been disordered. For Cx26 this region

extends from approximately residue Arg104 at the C-terminus of TM2 to Glu129 at the N-terminus of TM3. In the crystal structure [23] residues 125 to 129 have been modelled as part of TM3, whereas in our previous structures solved by cryo-EM [18] there was no evidence of the helix extending beyond Glu129. Instead, when analysing the previously reported cryo-EM maps it was noted that unassigned density protruded from near to the top of TM3 towards the loop between the N-terminus and TM1 in a manner that resembles models from Alphafold [28,29]. In the $K125E_{90\text{-}1}$ map this density is much more clearly defined as the C-terminal end of the cytoplasmic loop as it joins onto TM3 (Figure 2c) and residues Gln124 to Ile128, overlapping with the $K_{125}VRIEG_{130}$ motif were tentatively modelled into it. Though the density for the side chains is poor and there is no definitive interaction, Lys125 in this position is relatively close to the side chain of Arg104 of the neighbouring subunit, with which it had been proposed to form a salt bridge (Figure 2d).

The second region that differs is in TM1. Previously we noted a difference between the Cx26 crystal structures and our cryo-EM structures in the position of the residues between Val37 and Glu42 [18]. In the $K125E_{90\text{-}1}$ structure, the N-terminal region of TM1, comprising residues Gly21 to Phe31, is rotated with respect to that modelled in the WT structures, changing the position of the π-helix in TM1 from residues Ile20-Leu25 to residues Phe29-Val38. This results in the side-chain of Trp24 rotating by ~90° between the two extreme positions (Figure 2e). At one of these extremes, it faces the exterior of the protein, and nestles a detergent or lipid tail. At the other extreme, it is within the core of the protein next to Arg143 and Ala88 (Figure 2e). This conformation would not be possible with TM2 as modelled in the D6 averaged structure of the WT_{90} data set (7QEQ) because Thr86 and Leu89 would clash with Phe31 and Ile30 on TM1.

The difference of the $K125E_{90}$ maps from the WT_{90} maps indicated a clear bias towards a more closed conformation of the protein in the CO_2/HCO_3^- buffers. To understand the contribution of the K125E mutant, independently of any effect of CO_2 we also reanalysed data collected from both WT and K125E protein vitrified in HEPES buffer at pH 7.4. While the resolution of the maps was much lower for these reconstructions (Supplementary Table 1, Supplementary Figure 3) and better for the K125E mutant (4.2Å) than the WT protein (4.9Å) superposition of the respective maps again showed a movement of the cytoplasmic portion of TM2 together

with a change in the N-terminus (Supplementary Figure 4 and Supplementary Movie S2). Thus, it appears that the K125E mutant is sufficient by itself to bias the conformation in the absence of CO_2. However, CO_2 may have other effects on the protein, for example multiple carbamylation events[30], to give the higher resolution and more defined conformation seen in the CO_2/HCO_3^- buffers.

In contrast to the K125E mutant, mutating the same lysine to an arginine, which would prevent it from being carbamylated, results in GJCs that do not close [15]. We collected data from the protein with this mutation introduced under the same high PCO_2 conditions in CO_2/HCO_3^- buffer. Overall refinement in D6 resulted in a map with a nominal resolution of 2.1Å. With respect to the K125E data, classification resulted in a much smaller proportion of particles with an appearance of the PCN conformation (c.f. Supplementary Figures 1 and 5) and the quality of the maps from the resulting 3D reconstruction (2.8Å) was poorer, especially towards the cytoplasmic end. With respect to the data from the wild-type protein [18], however, the differences are subtle, and it is difficult to draw any concrete conclusions.

Molecule trapped under N-terminal helix likely to be lipid

In all the cryo-EM maps from both this study of mutant Cx26 and our previous study of the wild-type protein we observe density consistent with a hydrophobic molecule in the pore of the protein that could be a lipid or detergent (Supplementary Figure 6a) [18]. Lipids have been observed in other connexins and related large pore channels in the pore of the protein and have been suggested to be part of the mechanism [7,8,19,31]. As the artefactual presence of detergent could potentially interfere with the position of the N-terminus, we decided to solubilise the protein using lauryl maltose neopentyl glycol (LMNG), a much larger detergent than the dodecyl β-D-maltoside (DDM) used previously. Thus, if the density was due to a detergent trapped during the solubilisation process it should not be able to bind. We also reasoned that, if the high PCO_2 buffers were biasing the protein to a more closed conformation, carrying out all steps in this buffering system would reduce the entry of any molecule into the pore. Data were collected from the wild type protein solubilised in LMNG at a PCO_2 corresponding to 90mmHg as before. Refinement with D6 symmetry imposed resulted in a map with a nominal resolution of 2.0Å as defined by gold standard Fourier Shell Correlations [26,27] (Supplementary Figure S7, Supplementary Table 1). Like the maps deriving

from protein solubilised in DDM clear density was observed for a similar hydrophobic molecule in the pore of the protein. (Supplementary Figure 6b). Though the presence of the molecule may still be an artefact of either the solubilisation process or heterologous expression in insect cells, it strongly suggests that this molecule cannot be detergent and is more likely to be a lipid molecule.

The cytoplasmic side of the protein is much better defined in the presence of LMNG

After classification as above, two classes were selected for further refinement based on the two most extreme positions of TM2 in the POFN ($LMNG_{90\text{-}class1}$) and PCN ($LMNG_{90\text{-}class2}$) conformations, respectively (Fig 3, Supplementary Figure 7, Supplementary Table 2; Supplementary Movie S3). Though the nominal resolution of the $LMNG_{90\text{-}class2}$ map (2.3Å) where the protein is in the PCN conformation, is similar to the equivalent $WT_{90\text{-}class2}$ map, the cytoplasmic region of the protein is better defined (Fig 3, Supplementary Figures 2 and 7). Clear density is compatible with the KVRIEG motif in a similar position to that modelled in the $K125E_{90\text{-}1}$ structure (Fig 3b). As seen for the $K125E_{90\text{-}1}$ structure, however, neither the side-chains of Lys125 nor Arg104 are clearly defined. The density for the side-chains of the residues of the N-terminus and the following link is much clearer than seen in the other maps. His16 nestles up to Ala96 of TM2 and Val126 of the KVREIG motif is situated just above (Fig 3b). Again, though, it is difficult to place the first three residues of the N-terminus unambiguously.

In contrast, in the $LMNG_{90\text{-}class1}$ map (POFN conformation) neither the N-terminus nor the KVRIEG motif are well defined and the map is reminiscent of the map derived from particles vitrified under low PCO_2 conditions [18] (Fig 3c). In the associated structure the conformation of TM1 is the same as was modelled for the WT_{90} structure. As discussed above for the K125E structure, this conformation would not be compatible with the position adopted by TM2 in the $LMNG_{90\text{-}class2}$ structure (Fig 4b, Supplementary Movie 4a, 4b). The rotation of TM1 changes not only how TM1 interacts with the N-terminus, but also the conformation of the linker between the N-terminal helix and TM1 (Fig 4a). The position of His16 on this linker is not compatible with the conformation of the KVRIEG motif in $LMNG_{90\text{-}class2}$ (Fig 4b).

Density for cytoplasmic loop compatible with models from Alphafold

Despite the $LMNG_{90\text{-}class2}$ map being much clearer for the cytoplasmic region of the protein residues from 107 to 123 were still missing. We, therefore, carried out further classification of the particles focussed on the cytoplasmic region of a single subunit (see methods). As above, this resulted in a range of maps showing varying positions of the transmembrane helices and clarity of the N-terminal helix (Supplementary Figure 8). Importantly, in one case and where the N-terminus was clearly defined, extra density was also seen for the cytoplasmic loop, albeit at low resolution. Following reconstruction of the full dodecamer from the particles corresponding to this class ($LMNG_{90\text{-}mon\text{-}pcn}$ Supplementary Figure 8), the structure was tentatively built into the density with the cytoplasmic loop of the classified subunit in a conformation resembling models from Alphafold [28] (Fig 5) and with a complete N-terminus. We note, however, density near to Lys125, between Ser19 in the TM1-N-term linker, Tyr212 of TM4 and Tyr97 on TM3 of the neighbouring subunit, which we have been unable to explain with our modelling. Overall, the conformation of the subunit is very similar to the $LMNG_{90\text{-}class2}$ (PCN) structure with an RMSD of 0.38Å for 198 Cα atoms. Rather surprisingly, only the subunit upon which the focussed classification had been carried out had this conformation, with the density from the other subunits appearing more like the map before focussed classification. When hexameric symmetry is applied to the subunit, though the conformation of the N-terminus causes the aperture of the pore to appear closed, steric clashes suggest a symmetric arrangement of this conformation would not be possible.

Discussion

Our data show that there is a clear effect on the structure of Cx26 in changing K125 to a glutamate both in CO_2/HCO_3^- and in HEPES buffer. The biasing of the protein towards a conformation with a more ordered N-terminus with a narrow constriction is consistent with the closure of gap junctions that we observe in the dye transfer assays. It is also consistent with the hypothesis that the CO_2 mediated closure of the gap junctions is caused by the carbamylation of K125. Exactly why this happens however, is less clear. The hypothesis has been that the negative charge introduced onto the lysine side-chain would enable it to form a salt bridge with Arg104 from the neighbouring subunit[11]. Consistent with this, in models from Alphafold, which differ in conformation from the structure upon which the hypothesis was based, the lysine is located next to the arginine (Fig 5a). In the $K125E_{90\text{-}class1}$ and $LMNG_{90\text{-}}$

$_{class2}$ structures (PCN conformation) the position of the K/E125 main chain is also near to Arg104. Though there is no clear interaction between them, with minor adjustments of the side-chains, the residues could be made to interact and the absence of a definite interaction may be due to distortions during the vitrification process or radiation damage to which acidic residues are more prone. On the other hand, TM2, on which Arg104 is located, moves away from, rather than towards the lysine in going from the $LMNG_{90\text{-}class1}$ structure (POFN conformation) to the constricted structures and we note there is unexplained density between the subunits. It is possible therefore that, though the negative charge is important, and the mutational data of the hemichannel would strongly indicate an interaction between the two residues, attributing the induction of the conformational change to a single salt-bridge may be an over-simplification. Nevertheless, it is clear that the conformations of TM2 and the KVRIEG motif are linked to those of TM1 and the N-terminus. The conformation of TM1 seen in the constricted structures, for instance, is not compatible with the conformation of TM2 in the open structure. Likewise, the position of the link between the N-terminus and TM1 in the open structure could not be adopted if the KVRIEG motif were to be as in the constricted structures.

The results are also consistent with our previous study investigating the effect of PCO_2 on the structure [18]. At low PCO_2, where there should be little carbamylation, we see very little protein in the constricted conformation. The difference between the WT_{90} samples and the K125E samples would suggest that the WT protein is not completely carbamylated during vitrification. This may also explain why the effects of the mutation of K125 to arginine, which cannot be carbamylated, are more subtle than might be expected. Our data would be consistent with a conformationally flexible protein, in which the introduction of a negative charge would stabilise one particular conformation, rather than causing a conformational change *per se*.

It is interesting to note that the density we observe for the cytoplasmic loop in the focussed refinement of a single subunit is consistent with Alphafold models despite no other structure in the database having this conformation. Although the resolution of the density is poor in this region, the combination of our structure and the predicted models paves the way for further studies of the importance of particular interactions in the loop. Rather surprisingly

given that we can isolate a symmetrical conformation of the protein using C6 symmetry, reconstructions of the full dodecamer following the focussed classifications of the single subunit do not show an influence of the conformation of that subunit on its neighbours. While this is consistent with structural studies on Cx43 gap junctions [8] the stochastic nature of the subunit conformations contrasts with studies of Cx26 hemichannels in cell membranes where significant cooperativity has been observed [32]. This may be due to the protein being in a non-native environment, but would also support the above hypothesis that introduction of negative charge (from carbamylation of K125 or the K125E mutation) would trap one particular conformation rather than inducing the conformational change.

Recently, structures of gap junction channels have been reported for the alpha connexins Cx43[8,9] and Cx36[7] in multiple conformations. As we have observed here, in these studies, π to α transitions in the conformation of TM1, are coordinated with the position of the N-terminus. A consensus appears to be emerging in which the N-terminus being tucked back against the pore (pore lining) as first seen for Cx46/50 [20,21] has been described as an open conformation and one in which the N-terminus is more lifted as initially described for the Cx31.3 hemichannel [19] is the closed conformation. Due to the presence of lipids within the pores in these structures, it has also been inferred that lipids are important in the gating mechanism, sealing the pore where the N-terminus is in the raised position [7-9]. In direct contrast to this consensus, the constricted state that we observe contains the N-terminus pointing into the pore with TM1 in a similar conformation to the alpha connexin structures with the N-terminus tucked back against the pore. The N-terminus in this position would not be possible with the presence of the lipid that we observe lining the pore. It is interesting to note that a lipid-like molecule in this position appears to be a constant feature of all the high resolution cryo-EM maps associated with the structures of connexins where the N-terminus is not tucked back against the pore, including both gap junction channels[7-9,33] and hemichannels[19,33]. This is irrespective of the solubilisation method or whether the structure has been solved in the presence of detergent or nanodiscs.

More importantly, however, the pore lining position of the N-terminus seen in the alpha connexins would not be possible with the conformation of TM2 in the constricted structures, because Leu89, Ala92 and Ala96 on TM2 would push the N-terminal helix outwards. A

comparison of all the connexin structures that have been published shows that the position of TM2 seen in the PCN conformation of Cx26 is an outlier, with an apparent anticlockwise rotation with respect to the pore (Figure 6). The flexing of TM2 is near to Pro87. Mutation of Pro87, which is conserved in the connexins, causes a change in voltage gating, but not to the response to pH [34].

The conformation of the KVRIEG motif is interesting. In most structures TM3 has not been modelled before the equivalent of Gly130. It does, however, take a similar though shifted conformation in both the structures of Cx50 and in one of the hemichannel focussed classifications of Cx43 (7xqj) where the N-terminus is in the pore lining conformation. This shifted conformation, as seen for the Alphafold structures shown in Figure 5 would not be possible in the occluded structures here, due to the position of TM2. In the original low resolution crystal structure of Cx26 [23], the KVRIEG motif is modelled as an extension of TM3. In this structure the conformation of TM1 is more similar to the $LMNG_{90\text{-class1}}$ (OPFN) structure. Although, the crystal structure contains the N-termini pointing into the pore the helices are pulled further back than seen for our constricted structures. Clearly, the position of the KVRIEG motif is linked to that of the N-terminus.

It must be asked whether the PCN structures represent the closed state as seen in the dye-transfer experiments. It is difficult to model the first three residues of the N-terminus unambiguously, presumably due to the breakdown in 6-fold symmetry at this point, but with minor changes to the side chains, the centre of the pore could be closed. Closure of the centre of the pore would leave a potential lateral passageway into the pore. This would be blocked, however, by the lipid that is common to all the connexin structures packing into the pore.

Mutations in Cx26 lead to both syndromic and non-syndromic deafness [12]. While these mutations have been mapped on the structure previously, the position of Ala88, mutation of which causes KIDS, is interesting with regard to the new conformations. The mutation of Ala88, which lies at the flexion point of TM2, to valine has been shown to prevent CO_2 sensitivity either to the gap junctions, in closing the channels [15], or to the hemichannels in opening the protein [35]. In the constricted structures the C_β of Ala88 lies within 4.2Å of Trp24, which moves during the conformational change of TM1. Replacement of the alanine with the

bulkier valine would prevent this conformation being adopted. Interestingly the alanine and the tryptophan are located next to Arg143. Mutation of Arg143 to tryptophan is a very common mutation that leads to non-syndromic hearing loss [12].

Previously we showed that increasing PCO_2 biased the protein towards a conformation with a more ordered N-terminus. Our results, summarised in Figure 7, are consistent with this effect being mediated through K125. The movements of the transmembrane helices in response to the introduction of a negative charge provide a simple mechanism for closure through the modulation of the N-terminal helix. Though further studies are clearly required to establish the link between carbamylation and conformational change, the result is consistent with electrophysiological measurements.

Methods

Mutant preparation:

K125R and K125E mutations of human connexin 26 were prepared using the QuikChange II mutagenesis kit (Agilent) and the following primers: K125R forward: 5'-tcgaggagatcaaaacccagagggtccgcatcg-3', K125R reverse: 5'-cgatgcggaccctctgggttttgatctcctcga-3', K125E forward: 5'-gagatcaaaacccaggaggtccgcatcgaa-3', K125E reverse: 5'-ttcgatgcggacctcctgggttttgatctc-3' (Sigma) with the wild-type human connexin 26 pFast construct used for previous studies as the template for mutagenesis [18]. Viruses harbouring the connexin constructs were prepared and protein expressed in *Sf9* cells.

HeLa cell culture and transfection:

HeLa DH (ECACC) cells were grown in DMEM supplemented with 10% fetal bovine serum, 50 µg/mL penicillin/streptomycin and 3 mM $CaCl_2$. For electrophysiology and intercellular dye transfer experiments, cells were seeded onto coverslips in 6 well plates at a density of $2x10^4$ cells per well. After 24 hours, the cells were transiently transfected with Cx26 constructs tagged at the C-terminus with a fluorescent marker (mCherry) according to the GeneJuice Transfection Reagent protocol (Merck Millipore).

Patch clamp recording and gap junction assay:

2-Deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-glucose, NBDG, was included at 200 µM in the patch recording fluid, which contained: K-gluconate 130 mM; KCl 10 mM; EGTA 5 mM; $CaCl_2$ 2 mM, HEPES 10 mM, pH was adjusted to 7.3 with KOH and a resulting final osmolarity of 295 mOsm. A coverslip of cells was placed in the recording chamber and superfused with a control saline (124 mM NaCl, 3 mM KCl, 2 mM $CaCl_2$, 26 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 1 mM $MgSO_4$ and 10 mM D-glucose saturated with 95% O_2/5% CO_2, pH 7.4, PCO_2 35 mmHg). Cells were imaged on a Cleverscope (MCI Neuroscience) with a Photometrics Prime camera under the control of Micromanager 1.4 software. LED illumination (Cairn Research) and an image splitter (Optosplit, Cairn Research) allowed simultaneous imaging of the mCherry-tagged Cx26 subunits and the diffusion of the NBDG into and between cells. Coupled cells for intercellular dye transfer experiments were selected based on tagged Cx26 protein expression and the presence of a gap junctional plaque, easily visible as a band of

mCherry fluorescence. After establishing the whole cell mode of recording, images were collected every 10 seconds.

Protein Production, Purification and Grid Preparation:

Purification of all proteins were performed as previously described [18], briefly described here for each protein sample.

K125E in HEPES buffer

Protein Production and purification

Sf9 cells harbouring the Cx26 virus were harvested at 72 hours post infection at 2500 x g in a Beckmann JLA 8.1000 rotor, cell pellets were snap frozen in liquid nitrogen, and stored at -80°C until purification. All purification steps were performed on ice, or at 4°C. Cells were thawed in hypotonic lysis buffer (10 mM sodium phosphate, 10 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, pH 8.0- DNAse I, cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) and AEBSF) for 30 minutes before breakage using a dounce homogeniser. Membranes were separated by ultracentrifugation for 1 hour at 4 °C, 158000 x g. After resuspending the membranes in membrane resuspension buffer (25 mM sodium phosphate, 150 mM NaCl, 5 % glycerol, 1 mM DTT, pH 8.0- DNAse I, cOmplete™ EDTA-free Protease Inhibitor Cocktail and AEBSF) solubilisation was carried out in membrane solubilisation buffer (10 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, 1 mM DTT, 1% DDM (Glycon Biochemicals GMBH), pH 8.0) for 3-4 hours, and insoluble material removed by a further 1 hour ultracentrifugation at 4 °C, 158000 x g. Soluble protein was batch-bound to pre-equilibrated HisPur Ni-NTA resin (Thermo Scientific) overnight and then poured into an Econo-Column for subsequent manual washing and elution steps. Resin was washed with 5x CV wash buffer (10 mM sodium phosphate, 500 mM NaCl, 10 mM histidine, 5 % glycerol, 1 mM DTT, 0.1 % DDM, pH 8.0) before eluting hCx26 with elution buffer (10 mM sodium phosphate, 500 mM NaCl, 200 mM histidine, 5 % glycerol, 1 mM DTT, 0.1 % DDM, pH 8.0). hCx26- containing fractions were dialysed (20 mM HEPES, 500 mM NaCl, 5 % glycerol, 1 mM DTT, 0.03 % DDM, pH 8.0) overnight with thrombin at (a 1:1 w/w ratio). The hCx26 was then passed through a 0.2 μm filter, concentrated using a Vivaspin concentration with 100,000 MWCO and loaded onto a Superose 6 Increase 10/300 size exclusion chromatography column (GE Healthcare Lifescience) equilibrated with the same HEPES-dialysis buffer to remove thrombin. The protein was subsequently concentrated

to ~ 3 mg/ml. The concentrated protein was then dialysed for a minimum of 3 hours prior to grid preparation against 20mM HEPES, 250mM NaCl, 2.5% Glycerol, 5mM DTT, 0.03% DDM, 1mM $CaCl_2$, pH 8.0.

Grid preparation

Protein (3.5mg/ml) was centrifuged at 17,200 g for 5mins at 4°C. Grids (0.6/1 quantifoil AU 300) were glow discharged for 30 seconds at 30mA. Vitrification of the protein in liquid ethane at -180°C was carried out with a Vitrobot MKIV with 3 μl protein per grid at 4°C, 100 % humidity, blot force 10, 3 seconds blotting.

Data Collection and Processing

Data were collected using a Titan Krios G3 on a Falcon 3 detector. Data processing was performed in Relion 3 [36]. Movies were motion corrected with MotionCor2 [37] and the CTF parameters estimated with CTFfind-4.1 [38], both implemented in Relion 3. Particles were picked from selected images using the Laplacian-of-Gaussian (LoG) picker, and serial rounds of 2D classifications on binned particles were used to filter out junk and poor particles. An initial model was generated using stochastic gradient descent, and this was used for further cleaning of particles via 3D classifications. Exhaustive rounds of 3D refinement, CTF refinement and polishing were performed on unbinned particles until no further improvement of the resolution for the Coulomb shell was gained. The resolution was estimated based on the gold standard Fourier Shell Coefficient (FSC) criterion [26,27] with a soft solvent mask. All masks for processing were prepared in Chimera [39,40]. All processing was carried out without imposed symmetry until the final stage, where tests with C2, C3, C6 and D6 for refinement were carried out to look for improvements in resolution.

<u>WT in HEPES buffer</u>

All methods are as above, with the following changes: the final dialysis prior to freezing was against 20 mM HEPES, 200 mM NaCl, 1 % glycerol, 1 mM DTT, 1mM $CaCl_2$, 0.03 % DDM, pH 8.0. Freezing concentration was 3mg/ml WT, and data collection was carried out using a Volta phase-plate.

<u>K125E in αCSF90 buffer</u>

All methods are as for K125E in HEPES buffer, except for the following changes: Fractions eluted from the NiNTA containing hCx26 were dialysed overnight at 4 °C against 10 mM sodium phosphate, 500 mM NaCl, 5 % glycerol, 1 mM DTT, 0.03 % DDM, pH 8.0. A Superose 6 Increase 5/150 size exclusion chromatography column (GE Healthcare Lifescience) was used to remove thrombin and exchange the buffer to αCSF90 buffer (70 mM NaCl, 5 % glycerol, 1 mM DTT, 0.03 % DDM, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$.) K125E (3.4mg/ml) was gassed with 15% CO_2 (3 x 12 seconds) followed by centrifugation at 17,200g for 5 mins at 4°C. Grids (0.6/1 quantifoil AU 300) were glow discharged for 1min at 30mA. Vitrification of the protein in liquid ethane/propane at -180°C was carried out with a Leica GP2 automated plunge freezer with 3 µl protein per grid at 4°C, 100 % humidity, 7 seconds blotting without sensor-blot in a 15% CO_2 atmosphere. Data were collected using a GATAN K3 detector in super-resolution mode, and were processed using Relion 4.

Particle subtraction and masked classification focussed on the hemichannel

Hemichannel classifications with C6 imposed symmetry was carried out as described previously [18]. The particles from the top class were unsubtracted, and the particles were refined with C6 symmetry, using a hemichannel mask and limited angular sampling. Local Resolution estimation was carried out in Relion.

K125R in αCSF90 buffer

All methods are as for K125E in αCSF90 buffer, except for the following changes: Grids (1.2/1.3 Aultrafoil Au300) were glow discharged at 30mA for 30 seconds. Vitrification of the protein in liquid ethane at -160°C was carried out with a Vitrobot with 3 µl protein per grid at 4°C, 100 % humidity, 3 seconds blotting (force 10, 1 blot, skip transfer) in a 15% CO_2 atmosphere. Data were collected using a K3 detector in Counting bin 1 mode. Data processing was carried out in Relion 4.

$LMNG_{90}$ hCx26 WT

Preparation of $LMNG_{90}$ hCx26 WT protein was carried out as for K125E in αCSF90 buffer, with the following changes: *Sf9* cells were lysed in αCSF90 buffer (70 mM NaCl, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2) and membranes were resuspended in (110 mM NaCl, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2) and solubilised in (500 mM NaCl, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2). Samples were taken periodically to check the pH, and re-adjusted by further addition of CO_2 when necessary to keep the pH constant. Wash buffer for NiNTA resin (500 mM NaCl, 10mM Histidine, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2) and elution buffer (500 mM NaCl, 200mM Histidine, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2) were prepared and the pH checked just prior to use, to ensure no drifting of pH before interaction with the connexin. Selected fractions eluted from NiNTA were dialysed against (500 mM NaCl, 5 % glycerol, 1 mM DTT, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$, pH corrected to 7.4 by addition of CO_2). The final size exclusion step was performed in αCSF90 buffer (70 mM NaCl, 5 % glycerol, 1 mM DTT, 0.03 % DDM, 80 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 3 mM KCl, 1 mM $MgSO_4$, 4 mM $MgCl_2$.) without additional CO_2. The concentrated, pooled samples were gassed to pH to 7.4 both prior to freezing as described previously [18]. Vitrification was carried out at 3.7mg/ml on 0.6/1 Aultrafoil grids using the method described for K125R in αCSF90 buffer. Data were collected using the K3 detector in super-resolution bin 2 mode.

Particle subtraction and masked classification focussed on a single subunit

Following particle expansion with D6 symmetry and particle subtraction with a mask encompassing a single subunit, masked, fixed angle classification was carried out in Relion 4. Following unsubtraction of particles, refinement of the particle positions was carried out as above, with a hemichannel mask.

Model building and Refinement

Model building was carried out in Coot [41] with real space refinement in Phenix [41,42] using maps that had been sharpened using model free local sharpening in Phenix. For the $LMNG_{90}$ hemichannel-based classification two maps were selected for refinement. The first of these ($LMNG_{90\text{-class2}}$) was chosen because the density of the cytoplasmic region was the most defined. The second ($LMNG_{90\text{-class1}}$) was chosen as the highest resolution map with TM2 in the most diverse position. A similar selection was made for the single subunit-based classification. In building the cytoplasmic region of the protein reference was made to both ModelAngelo [43] and Alphafold [44].

Structural Analysis

All structural images shown in this paper were generated in PyMol [45] or Chimera [39,40]. Superpositions were carried out in Chimera such that only matching C_α pairs within 2Å after superposition were included in the matrix calculation.

Figures

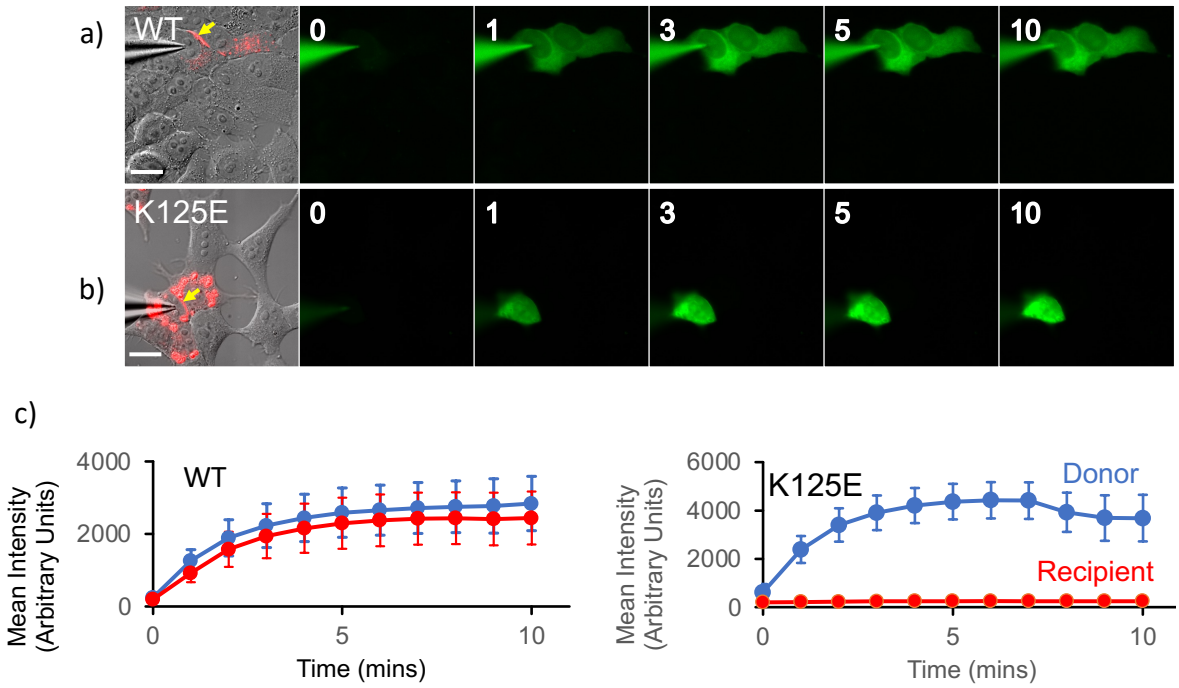

Figure 1: Cx26^{K125E} gap junctions are constitutively closed at a PCO_2 of 35 mmHg. a,b) Montages each showing bright field DIC image of HeLa cells with mCherry fluorescence superimposed (leftmost image) and the permeation of NBDG from the recorded cell to coupled cells. Yellow arrow indicates the presence of a gap junction between the cells; scale bar, 20 µm. The numbers are the time in minutes following the establishment of the whole cell recording. In Cx26^{WT} expressing cells (a), dye rapidly permeates into the coupled cell. For Cx26^{K125E} expressing cells (b), no dye permeates into the neighbouring cell even after 10 minutes of recording despite the presence of a morphological gap junction. **c)** Quantification of fluorescence intensity in the recorded cell (donor) and the potentially coupled cell (recipient) for both Cx26^{WT} and Cx26^{K125E} (7 pairs of cells recorded for WT and K125E, data presented as mean ± SEM). While dye permeation to the recipient cell follows the entry of dye into the donor for Cx26^{WT}, no dye permeates to the acceptor cell for Cx26^{K125E}. Note that the fluorescent intensity in the donor cell for Cx26^{K125E} is higher than for Cx26^{WT}, presumably because the dye remains trapped in the donor cell rather than diffusing to the acceptor cell.

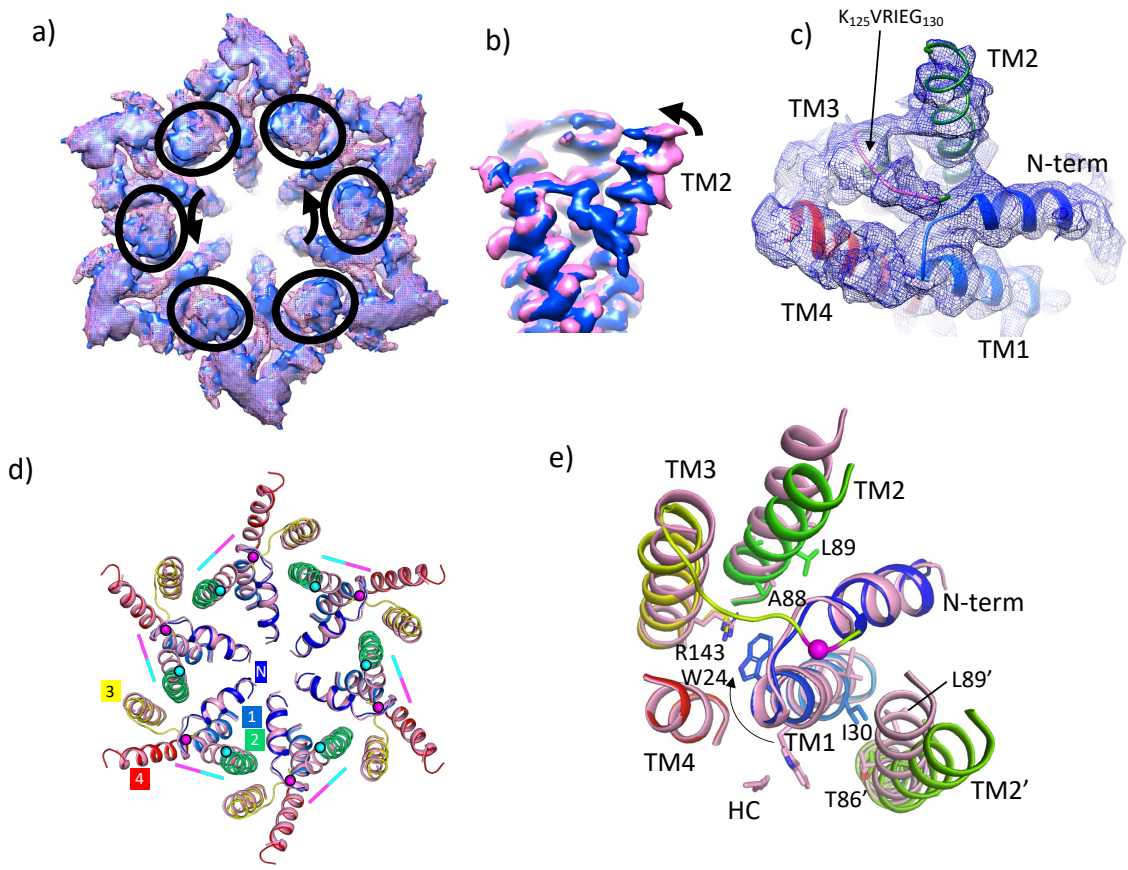

Figure 2: Comparison of WT Cx26 with the K125E mutant. a) Superposition of density for WT_{90} Cx26 (PDB ID 7QEQ; pink) on the density for $K125E_{90}$ (blue). The ovals show the position of TM2 from each subunit and the arrows show the direction of the difference between TM2 in the wild type and mutant structures. **b)** As (a) but focussed on TM2. **c)** Density associated with one subunit of the $K125E_{90}$ structure (unsharpened map). The structure has been coloured according to residue number with blue at the N-terminus and red at the C-terminus. The $K_{125}VRIEG_{130}$ motif has been depicted in magenta. **d)** Superposition of the WT_{90} structure (7QEQ, pink) on the $K125E_{90\text{-}1}$ structure (based on colouring in (c)). The positions of E125 and R104 have been depicted by magenta and cyan spheres respectively with the hypothesised salt bridge between them marked with a line. **e)** As d but focussed on TM1. The position of TM1 differs between the two structures as seen by the change in position of Trp24. (HC denotes the hydrocarbon chain from a lipid.)

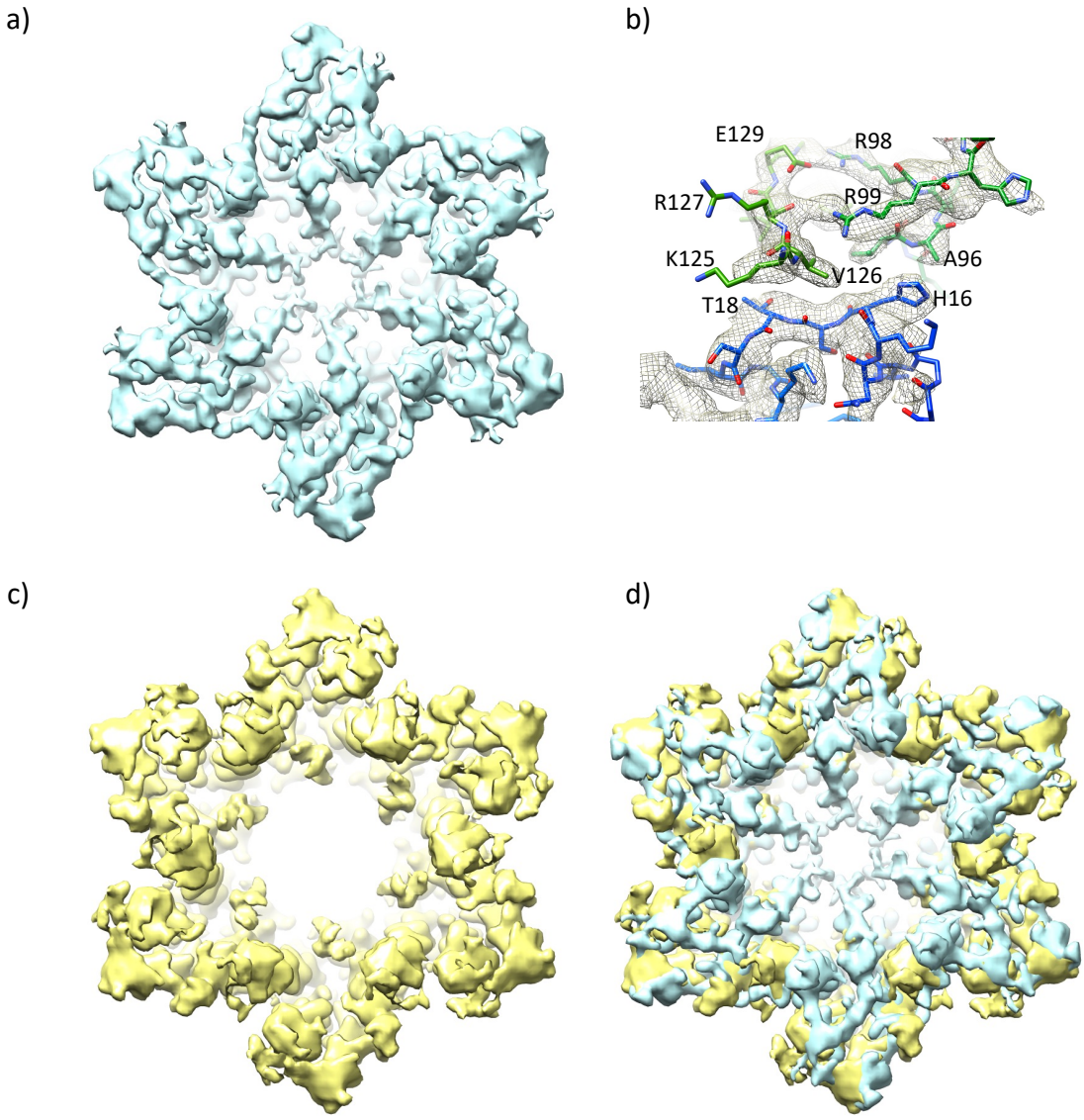

Figure 3: Distinct classes from classification of Cx26 solubilised in LMNG. **a)** Overall density associated with $LMNG_{90\text{-}class2}$. **b)** As (a) focussed on the KVRIEG motif and the link between the N-terminus and TM1. Arg98 and Arg99 of TM2 appear to stabilise the conformation of the KVREIG motif of the same subunit, with Arg98 interacting with Glu129 and Arg99 stabilising the main chain. **c)** Density associated with $LMNG_{90\text{-}class1}$. **d)** Superposition of the density from (a) and (c).

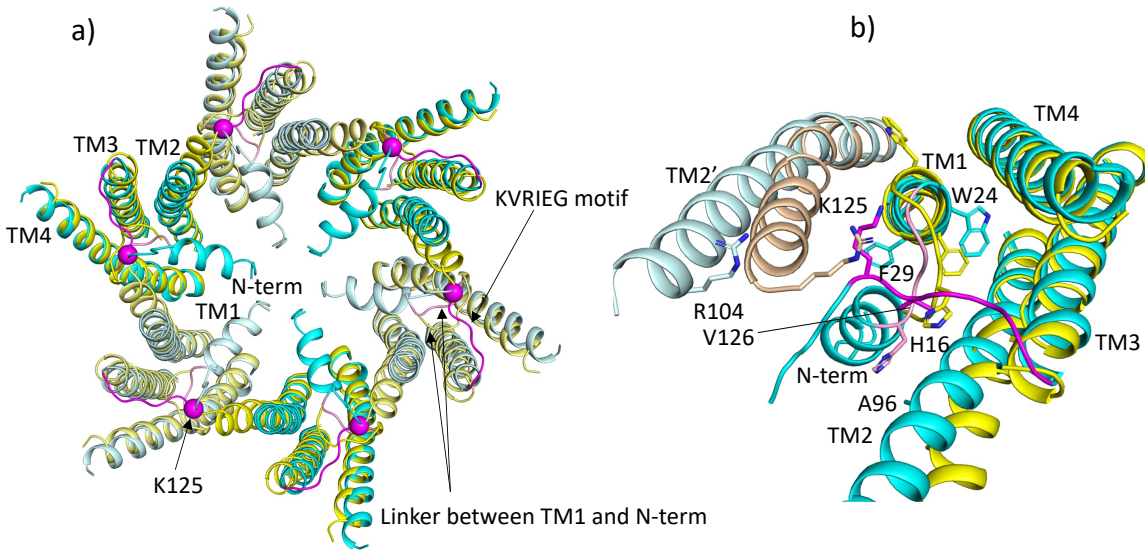

Figure 4: Comparison of structures refined from most distinct classes from LMNG classification a) Overall superposition showing the movement of TM2 and the link between the N-terminus and TM1. $LMNG_{90\text{-}class2}$ in cyan and $LMNG_{90\text{-}class1}$ in yellow (alternate subunits have been coloured in lighter shades). The KVRIEG motif has been coloured magenta with a sphere indicating the position of K125. The residues between the N-terminus and TM1 for the $LMNG_{90\text{-}class2}$ structure have been coloured pink. **b)** As (a) but focussed on TM1.

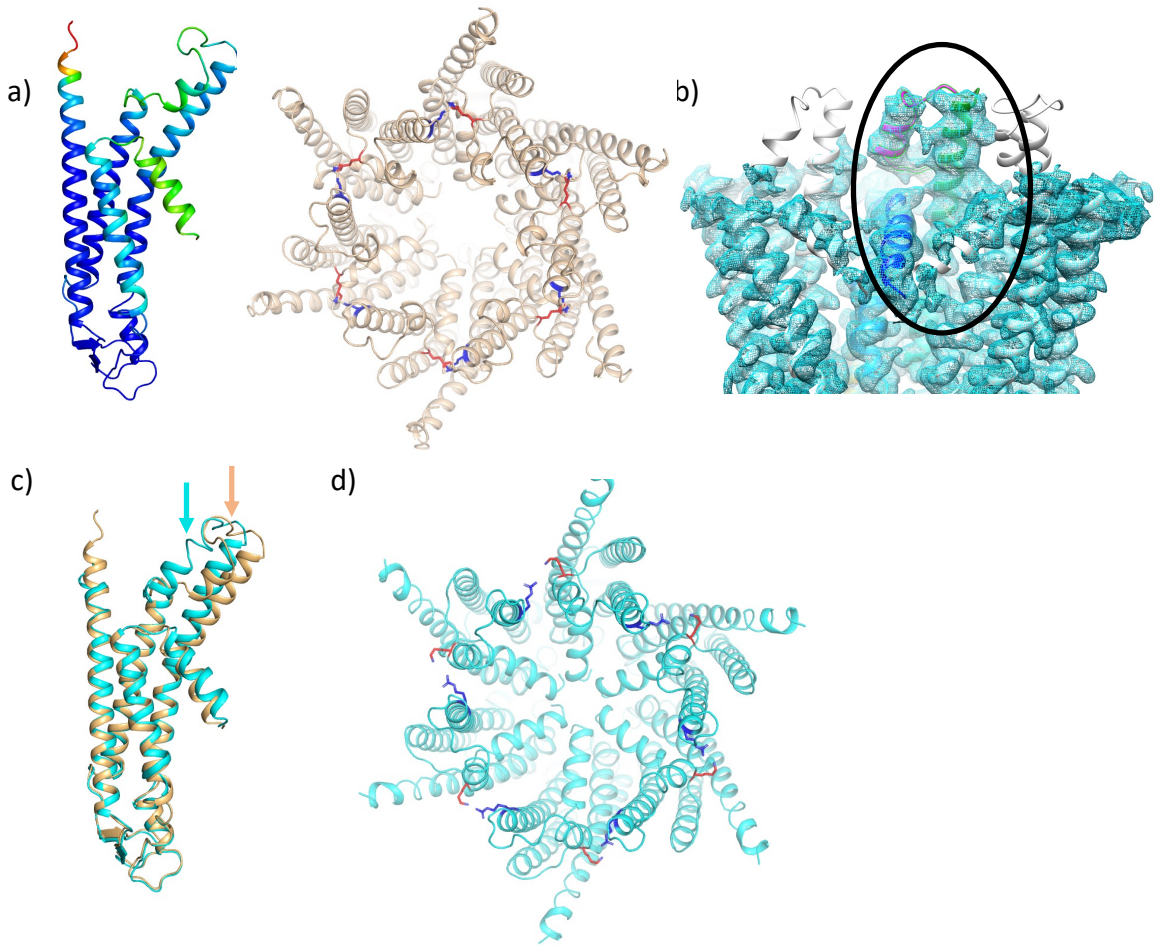

Figure 5: Focussed classification of a single subunit results in density for the cytoplasmic loop consistent with models from Alphafold. a) Models generated by Alphafold for a single subunit (left; coloured according to confidence level) and for the hexamer (right; in wheat with the position of K125 depicted by blue sticks and the position of R104 in red). **b)** Focussed classification of a single subunit (coloured as in Figure 2 and highlighted by an oval) resulted in clear density for part of the cytoplasmic loop in a conformation consistent with the models from Alphafold. This does not extend to the neighbouring subunits. **c)** Superposition of the single subunit built into the density (cyan) on the alphafold model (wheat). Showing the change in position of the helix in the cytoplasmic loop (highlighted by an arrow in the relevant colour). **d)** Reconstituting a hexamer by replicating the conformation of the subunit seen in (b) to all 6 subunits of the hexamer results in an apparently more closed conformation of the hemichannel, though there are also residue clashes, especially at the N-terminus. Lys 125 and Arg 104 are depicted with red and blue sticks respectively.

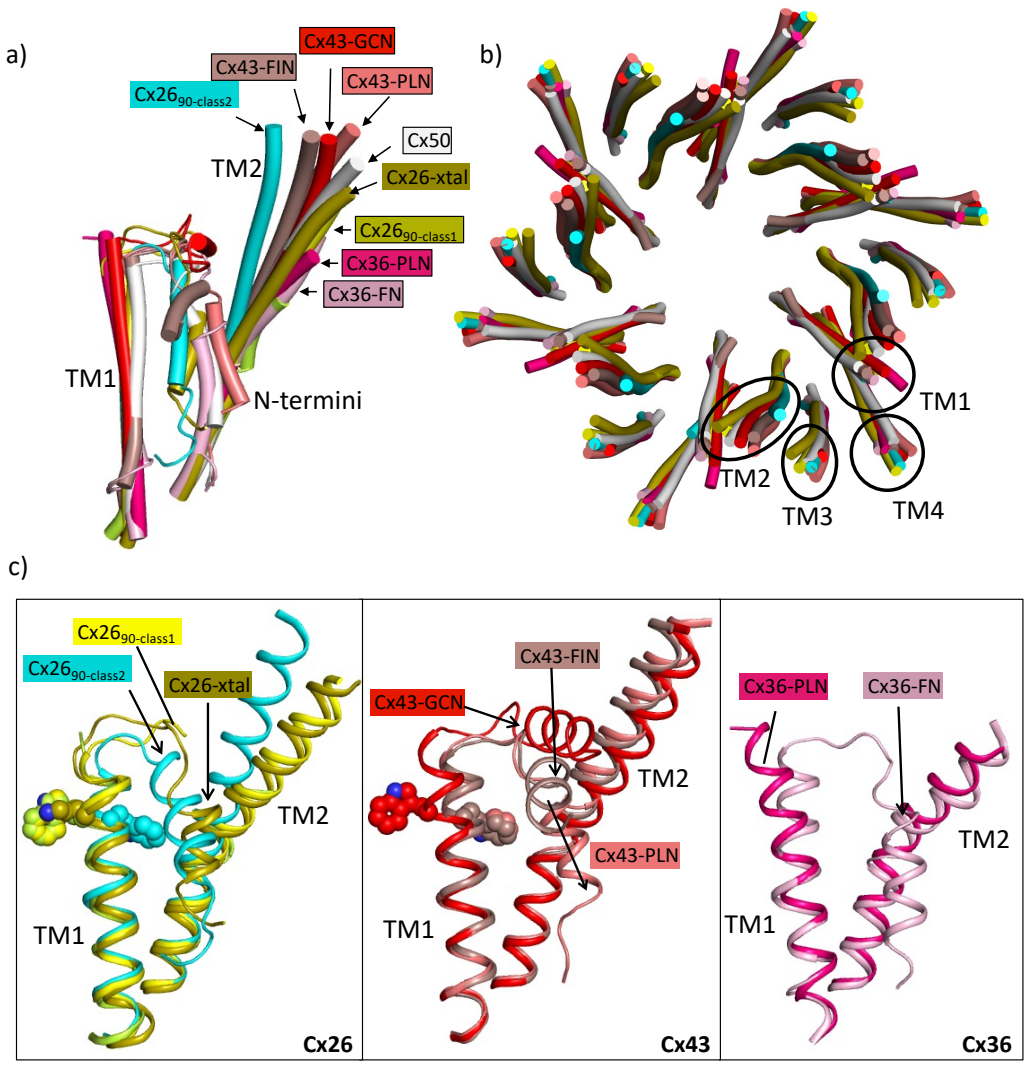

Figure 6: Comparison of the two structures derived from the LMNG classification with other structures of connexins. a) Superposition of a single subunit from the LMNG$_{90\text{-class2}}$ (cyan) and LMNG$_{90\text{-class1}}$ (yellow) structures on: Cx26 crystal structure (chartreuse, PDB ID 2ZW3); Cx50 (white, 7JJP); Cx43 in gate closing (red, 7XQF), flexible intermediate (chocolate, 7XQI) and pore lining (salmon, 7F94) conformations; Cx36 in pore lining (pink, 7XNH) and flexible N terminus (raspberry, 7XKT). The structures were superposed based on all chains of the hexamer. Only TM1, TM2 and the N-terminal helices are shown for each structure. **b)** View from the cytoplasmic side of the hemichannels from the same structures as in (a). Only the cytoplasmic part of the four transmembrane helices are shown. **c)** As (a) for the Cx26 (left), Cx43 (middle) and Cx36 (right) structures separately. Trp 24 in each of the Cx26 and Cx43 structures has been depicted with a sphere representation. The sequence identities for common residues to Cx26 is 49% for Cx50, 43% for Cx43 and 35% for Cx36.

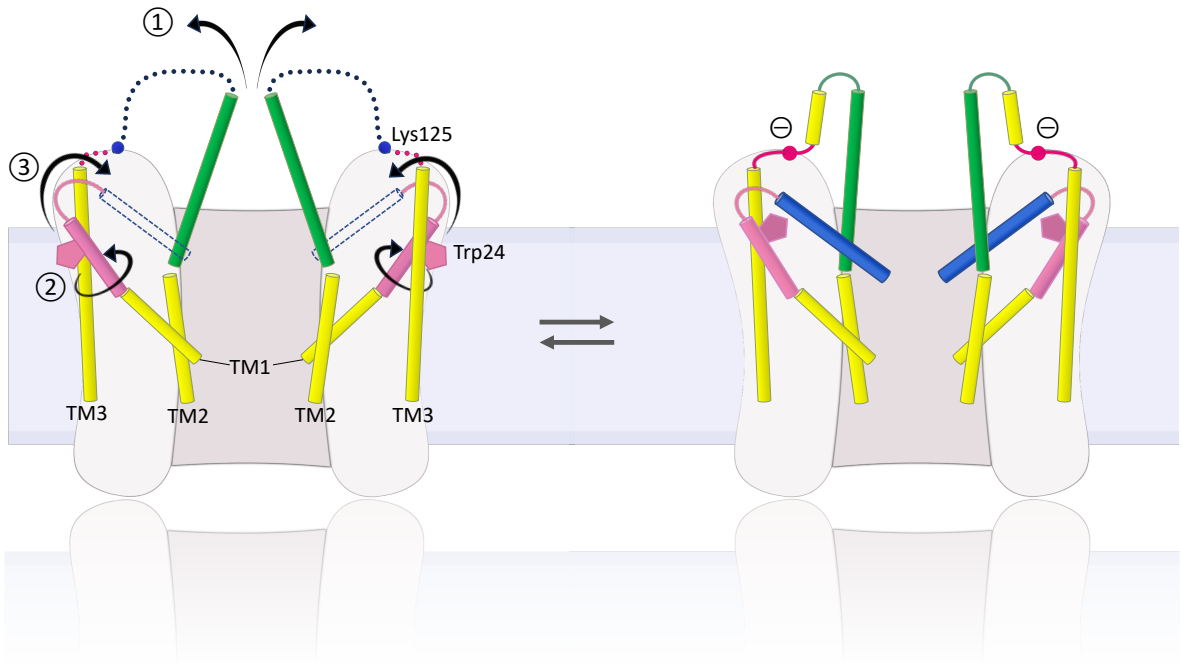

Figure 7: Schematic representation of conformational changes

Schematic view of the cytoplasmic region of two opposing subunits within one hemichannel of the gap junction. The open structure on the left and the constricted structure on the right are in a dynamic equilibrium. The introduction of a negative charge on Lys125 of the KVRIEG motif (magenta) pushes the equilibrium to the right. In going from one conformation to the other: ① the cytoplasmic region of TM2 flexes around Phe83 and the cytoplasmic loop adopts a more defined conformation; ② the cytoplasmic region of TM1 rotates, illustrated here by the movement of Trp24; ③ the N-terminal helix, which will be affected by both these conformational changes adopts a position within the pore that constricts entrance to the channel.

Acknowledgements

We thank the Leverhulme Trust (RPG-2015-090) and MRC (MR/P010393/1) for support. We acknowledge the Midlands Regional Cryo-EM Facility, hosted at the Warwick Advanced Bioimaging Research Technology Platform, for use of the JEOL 2100Plus, and the Midlands Regional Cryo-EM Facility, hosted at Leicester Institute of Structural and Chemical Biology for use of the FEI Titan Krios G3 and associated facilities, supported by MRC award reference MC_PC_17136. We thank Dr TJ Ragan for help with cryo-EM. We are grateful to the technical support in the School of Life Sciences, University of Warwick.

Author Contributions

The project was initiated and supervised by ND and AC. Cloning, expression, purification and grid preparation were carried out by DB. Data collection was performed by DB and CS. Data processing was done by DB and AC. DB and AC refined and analysed the structures. SN carried out the dye transfer assays. AC, DB and ND wrote the manuscript.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Data Availability

Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-18290 ($K125E_{90\text{-}1}$), EMD-18295 ($K125R_{90}$), EMD-18296 ($K125E_{HEPES}$), EMD-18297 (WT_{HEPES}), EMD-18292 ($LMNG_{90\text{-}class1}$), EMD-18291 ($LMNG_{90\text{-}class2}$), EMD-18293 ($LMNG_{90\text{-}mon\text{-}pcn}$), EMD-18294 ($LMNG_{90\text{-}mon\text{-}fn}$). Structure models have been deposited in the RCSB Protein Data Bank under accession numbers 8Q9Z, 8QA1, 8QA0, 8QA1, 8QA3 as listed in Supplementary Table 2.

References

1. Valiunas, V. Biophysical properties of connexin-45 gap junction hemichannels studied in vertebrate cells. *Journal of General Physiology* **119**, 147-164 (2002).
2. Young, K.C. & Peracchia, C. Opposite Cx32 and Cx26 voltage-gating response to CO2 reflects opposite voltage-gating polarity. *J Membr Biol* **202**, 161-70 (2004).
3. Bevans, C.G. & Harris, A.L. Regulation of connexin channels by pH. Direct action of the protonated form of taurine and other aminosulfonates. *J Biol Chem* **274**, 3711-9 (1999).
4. Khan, A.K. et al. A Steric "Ball-and-Chain" Mechanism for pH-Mediated Regulation of Gap Junction Channels. *Cell Rep* **31**, 107482 (2020).
5. Yu, J., Bippes, C.A., Hand, G.M., Muller, D.J. & Sosinsky, G.E. Aminosulfonate modulated pH-induced conformational changes in connexin26 hemichannels. *J Biol Chem* **282**, 8895-904 (2007).
6. Peracchia, C. Chemical gating of gap junction channels; roles of calcium, pH and calmodulin. *Biochim Biophys Acta* **1662**, 61-80 (2004).
7. Lee, S.N. et al. Cryo-EM structures of human Cx36/GJD2 neuronal gap junction channel. *Nat Commun* **14**, 1347 (2023).
8. Lee, H.J. et al. Conformational changes in the human Cx43/GJA1 gap junction channel visualized using cryo-EM. *Nat Commun* **14**, 931 (2023).
9. Qi, C. et al. Structure of the connexin-43 gap junction channel in a putative closed state. *eLife* (2023).
10. Huckstepp, R.T., Eason, R., Sachdev, A. & Dale, N. CO2-dependent opening of connexin 26 and related beta connexins. *J Physiol* **588**, 3921-31 (2010).
11. Meigh, L. et al. CO(2)directly modulates connexin 26 by formation of carbamate bridges between subunits. *Elife* **2**, e01213 (2013).
12. Xu, J. & Nicholson, B.J. The role of connexins in ear and skin physiology - functional insights from disease-associated mutations. *Biochim Biophys Acta* **1828**, 167-78 (2013).
13. Stout, C., Goodenough, D.A. & Paul, D.L. Connexins: functions without junctions. *Curr Opin Cell Biol* **16**, 507-12 (2004).

14. Evans, W.H., De Vuyst, E. & Leybaert, L. The gap junction cellular internet: connexin hemichannels enter the signalling limelight. *Biochem J* **397**, 1-14 (2006).
15. Nijjar, S. et al. Opposing modulation of Cx26 gap junctions and hemichannels by CO2. *J Physiol* **599**, 103-118 (2021).
16. Abascal, F. & Zardoya, R. Evolutionary analyses of gap junction protein families. *Biochim Biophys Acta* **1828**, 4-14 (2013).
17. Qi, C. et al. Structure of the connexin-43 gap junction channel in a putative closed state. *elife*, RP87616 (2023).
18. Brotherton, D.H., Savva, C.G., Ragan, T.J., Dale, N. & Cameron, A.D. Conformational changes and CO(2)-induced channel gating in connexin26. *Structure* **30**, 697-706 e4 (2022).
19. Lee, H.J. et al. Cryo-EM structure of human Cx31.3/GJC3 connexin hemichannel. *Sci Adv* **6**, eaba4996 (2020).
20. Flores, J.A. et al. Connexin-46/50 in a dynamic lipid environment resolved by CryoEM at 1.9 A. *Nat Commun* **11**, 4331 (2020).
21. Myers, J.B. et al. Structure of native lens connexin 46/50 intercellular channels by cryo-EM. *Nature* **564**, 372-377 (2018).
22. Bennett, B.C. et al. An electrostatic mechanism for Ca(2+)-mediated regulation of gap junction channels. *Nat Commun* **7**, 8770 (2016).
23. Maeda, S. et al. Structure of the connexin 26 gap junction channel at 3.5 A resolution. *Nature* **458**, 597-602 (2009).
24. Meigh, L., Cook, D., Zhang, J. & Dale, N. Rational design of new NO and redox sensitivity into connexin26 hemichannels. *Open Biol* **5**, 140208 (2015).
25. Lorimer, G.H. Carbon dioxide and carbamate formation: the makings of a biochemical control system. *Trends in Biochemical Sciences* **8**, 65-68 (1983).
26. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-30 (2012).
27. Rosenthal, P.B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721-45 (2003).
28. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

29. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022).

30. Brotherton, D.H. et al. Conformational changes and channel gating induced by CO2 binding to Connexin26. *bioRxiv* (2020).

31. Burendei, B. et al. Cryo-EM structures of undocked innexin-6 hemichannels in phospholipids. *Sci Adv* **6**, eaax3157 (2020).

32. de Wolf, E., Cook, J. & Dale, N. Evolutionary adaptation of the sensitivity of connexin26 hemichannels to CO2. *Proc Biol Sci* **284**(2017).

33. Qi, C. et al. Structures of wild-type and selected CMT1X mutant connexin 32 gap junction channels and hemichannels. *Preprint at https://www.biorxiv.org/content/10.1101/2023.03.08.531661v1* (2023).

34. Suchyna, T.M., Xu, L.X., Gao, F., Fourtner, C.R. & Nicholson, B.J. Identification of a proline residue as a transduction element involved in voltage gating of gap junctions. *Nature* **365**, 847-9 (1993).

35. Meigh, L., Hussain, N., Mulkey, D.K. & Dale, N. Connexin26 hemichannels with a mutation that causes KID syndrome in humans lack sensitivity to CO2. *Elife* **3**, e04249 (2014).

36. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**(2018).

37. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

38. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-21 (2015).

39. Goddard, T.D., Huang, C.C. & Ferrin, T.E. Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**, 281-7 (2007).

40. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).

41. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).

42. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
43. Jamali, K. et al. Automated model building and protein identification in cryo-EM maps. *bioRxiv* (2023).
44. Jumper, J. & Hassabis, D. Protein structure predictions to atomic accuracy with AlphaFold. *Nat Methods* **19**, 11-12 (2022).
45. Delano, W.L. The PyMOL Molecular Graphics System. *DeLano Scientific, Palo Alto, CA, USA* (2002).

Supplementary Material

Table S1: Cryo-EM data collection and processing statistics

	$K125E_{90}$	**$K125R_{90}$**	**$LMNG_{90}$**	**$K125E_{hepes}$**	**WT_{hepes}**
Voltage (kV)	300	300	300	300	300
Magnification (x1000) X	105	105	105	75	75
Camera	K3	K3	K3	Falcon 3	Falcon 3
Frame alignment on Falcon 3	-		-	yes	yes
Camera mode	Super-Resolution	Counting bin1	Super-Resolution bin2	counting	counting
Energy filter (eV)	20	20	20	20	20
Defocus range (μm)	-0.8 to -2.0	-0.8 to -2.0	-0.8 to -2.3	-0.3 to -1.7	-0.5 with Volta phaseplate
Pixel size (Å/pix)	0.85	0.835	0.835	1.08	1.08
Dose on detector (e^-/pix/sec)	10	15	18.77		0.7
Dose on sample (e^-/pix/sec)	11.4	18		1.06	0.69
Exposure time (sec)	3	2	2	44.01	60
No. of images	4731	10044	11362	2573	2436
Frames per image	45	50	50	40	75
Final particle number	189887	222622	204438	147546	60995
Resolution[a]					
masked D6 (Å)	2.2	2.1	2.0	4.3	4.9

a) From Relion_postprocess [26]

Table S2 Cryo-EM refinement and validation statistics

	K125E$_{90}$	**LMNG$_{90}$**			
	class1	class1	class2	mon-PCN	mon-OPFN
Deposited Structure PDB ID	8Q9Z	8QA1	8QA0	8QA2	8QA3
Final particle number	161625	59005	35007	357859	240137
Map resolution					
FSC threshold	0.143	0.143	0.143	0.143	0.143
Symmetry imposed[a]	C6	C6	C6	C1	C1
Unmasked (Å)	2.3	2.2	2.4	2.6	2.4
Masked (Å) (masked)	2.4	2.0	2.1	2.3	2.2
Refinement					
Initial model (PDB code)	7QEQ	7QEQ	K125E$_{90}$	Alphafold	mon-PCN
Resolution (Å; FSC=0.5) Model	2.6	2.4	2.6	2.6	2.6
Sharpening B factor (Å^2)	local	local	local	local	local
Model composition	12 chains	12 chains	12 chains	12 chains	12 chains
Non-hydrogen atoms	19704	18720	19878	18943	18580
Protein residues	2316	2190	2340	2260	2226
Water	348	318	342	322	254
Ligand: lipid/detergent	48	24	24	13	12
B factor (Å^2)					
protein	66	54	77	68	73
water	42	33	56	45	44
Lipid/detergent	56	60	85	69	70
R.m.s. deviations					
Bond lengths (Å)	0.004	0.002	0.004	0.002	0.003
Bond angles (°)	0.516	0.484	0.593	0.469	0.544
Validation					
MolProbity score	1.54	1.25	1.55	1.59	1.74
Clashscore	5.66	3.46	6.50	4.79	5.20
Rotamer outliers (%)	1.83	1.42	1.05	2.86	3.71
Ramachandran plot					
Favoured (%)	97.84	98.97	96.99	98.23	97.89
Allowed (%)	2.16	1.03	3.01	1.77	2.11
Disallowed (%)	0	0	0	0	0
CaBlam Outliers (%)	1.53	1.24	1.6	0.97	1.08
Correlation coefficients					
CC (mask)	0.86	0.87	0.85	0.87	0.87
CC (box)	0.7	0.71	0.66	0.68	0.70
CC (peaks)	0.68	0.70	0.63	0.67	0.68
CC (volume)	0.84	0.85	0.84	0.85	0.85
Mean CC for ligands	0.71	0.73	0.61	0.68	0.70

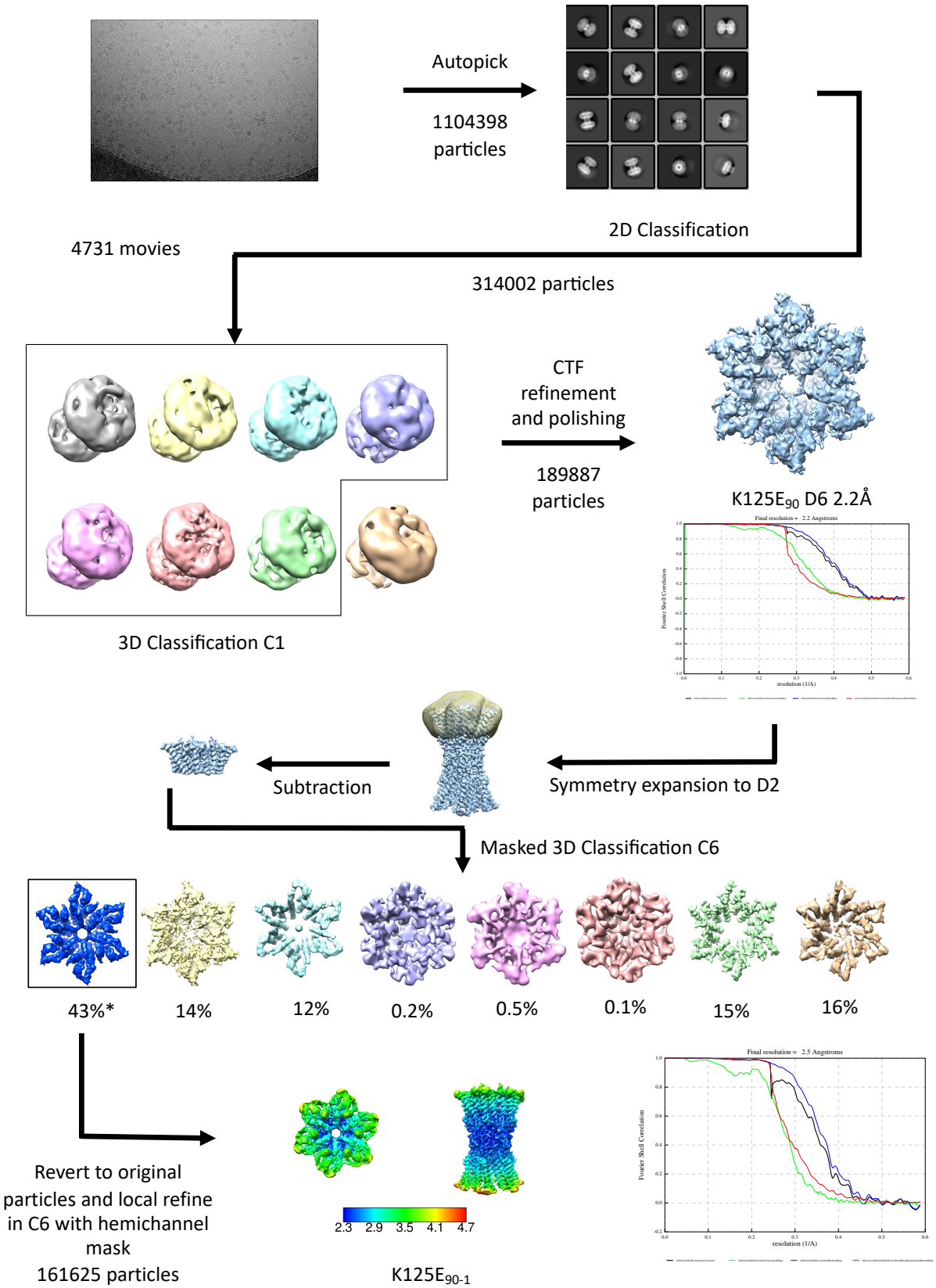

Figure S1: Workflow for processing of cryo-EM data for K125E sample in CO_2/$HCO3^-$ buffer. The star denotes the classifications with the appearance of the PCN conformation that refine to a resolution greater than 4Å.

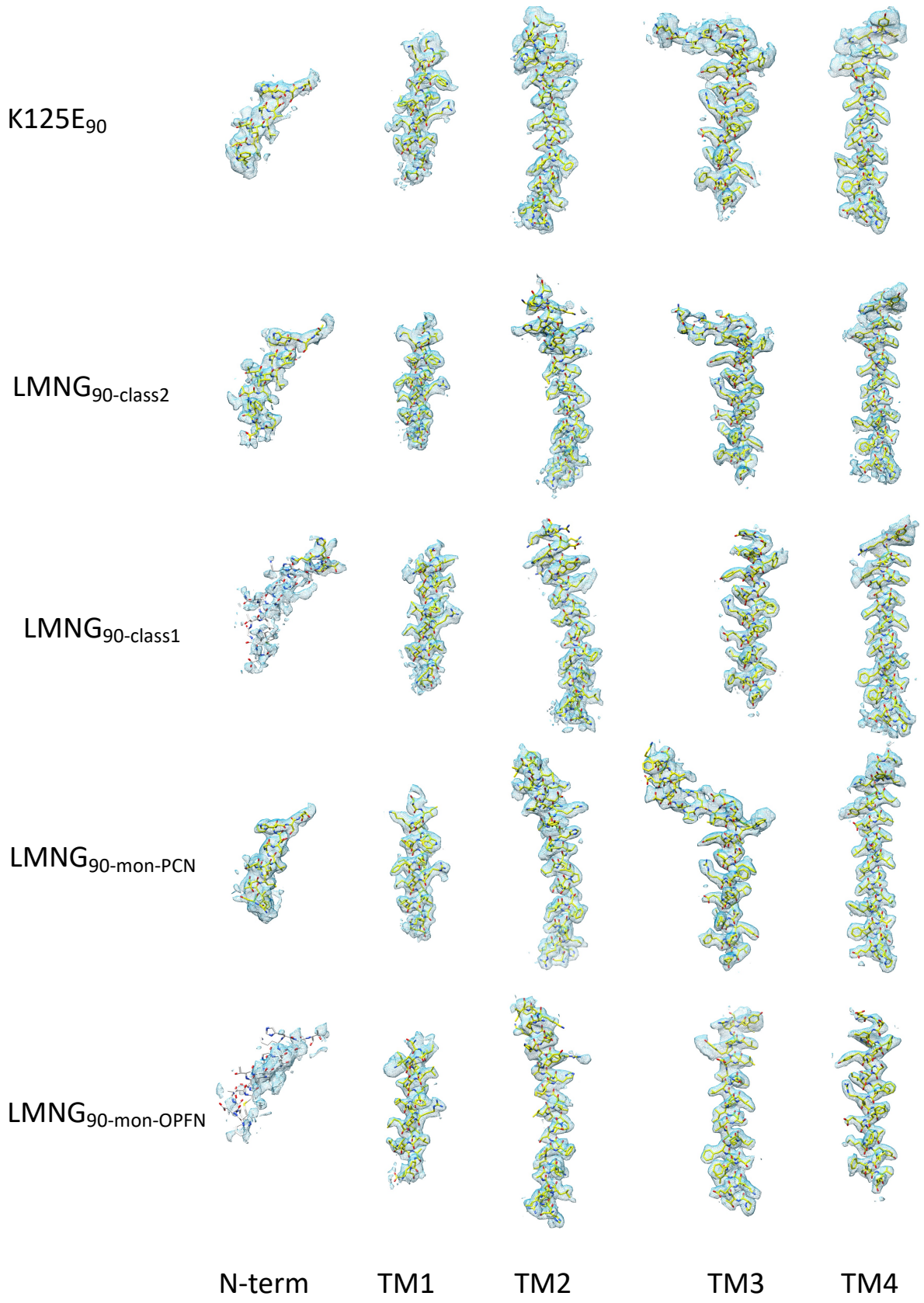

Figure S2: Density for the transmembrane and N-terminal helix associated with each structure.

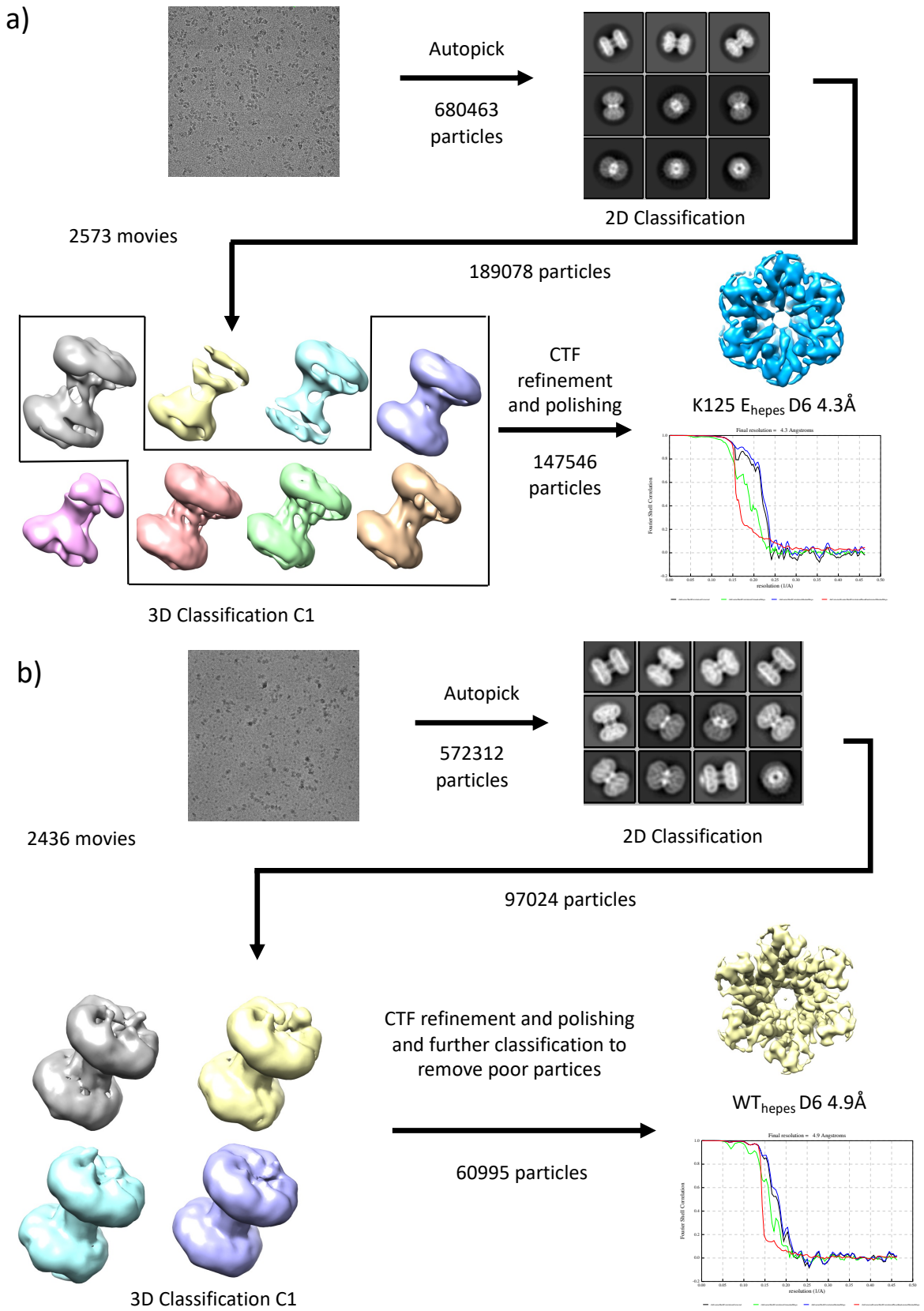

Figure S3 Workflows for processing of cryo-EM data for samples in HEPES buffer

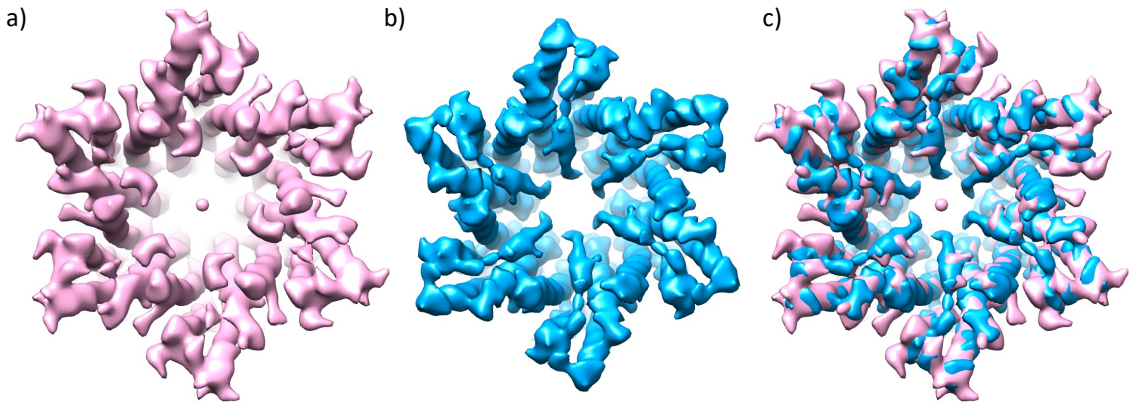

Figure S4: Comparison of density maps from wild type and K125E Cx26 purified in HEPES buffer at pH 7.4: a) WT Cx26 at 4.9Å resolution sharpened with a B-factor of -100. **b)** K125E sharpened with B-factor of -273 and low pass filtered to 5Å. **c)** Superposition of the two maps.

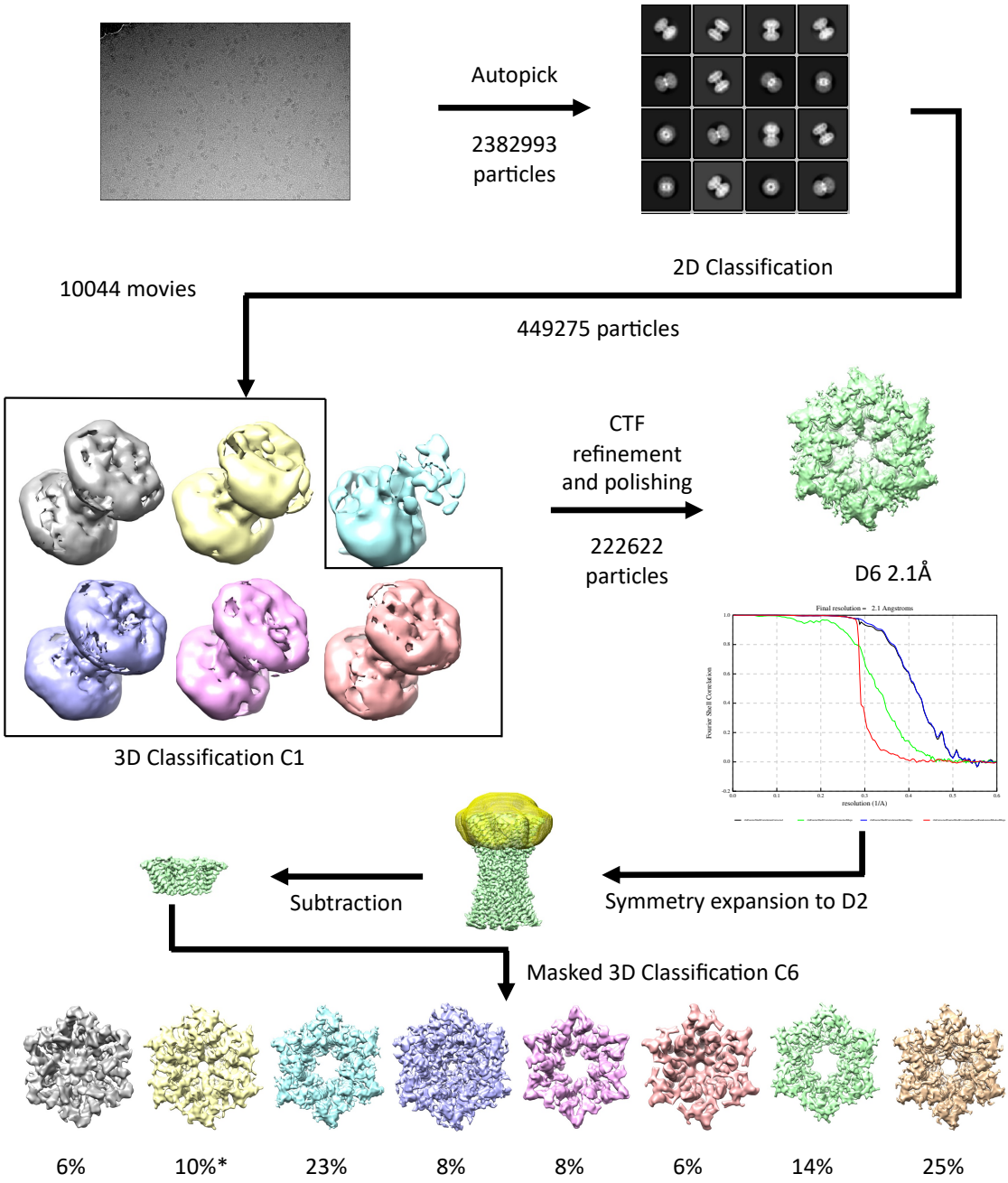

Figure S5: **Workflow for processing of cryo-EM data for K125R sample in CO_2/$HCO3^-$ buffer**

The star denotes the classifications with the appearance of the PCN conformation that refine to a resolution greater than 4Å.

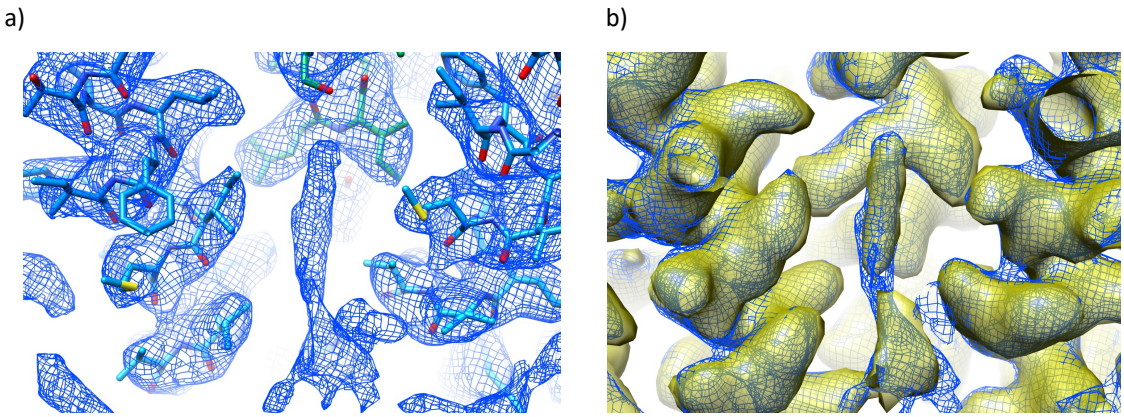

Figure S6: Non-protein density in pore of Cx26. a) Density for $K125E_{90}$. **b)** $LMNG_{90}$ D6 averaged map superposed on the map for $K125E_{90}$.

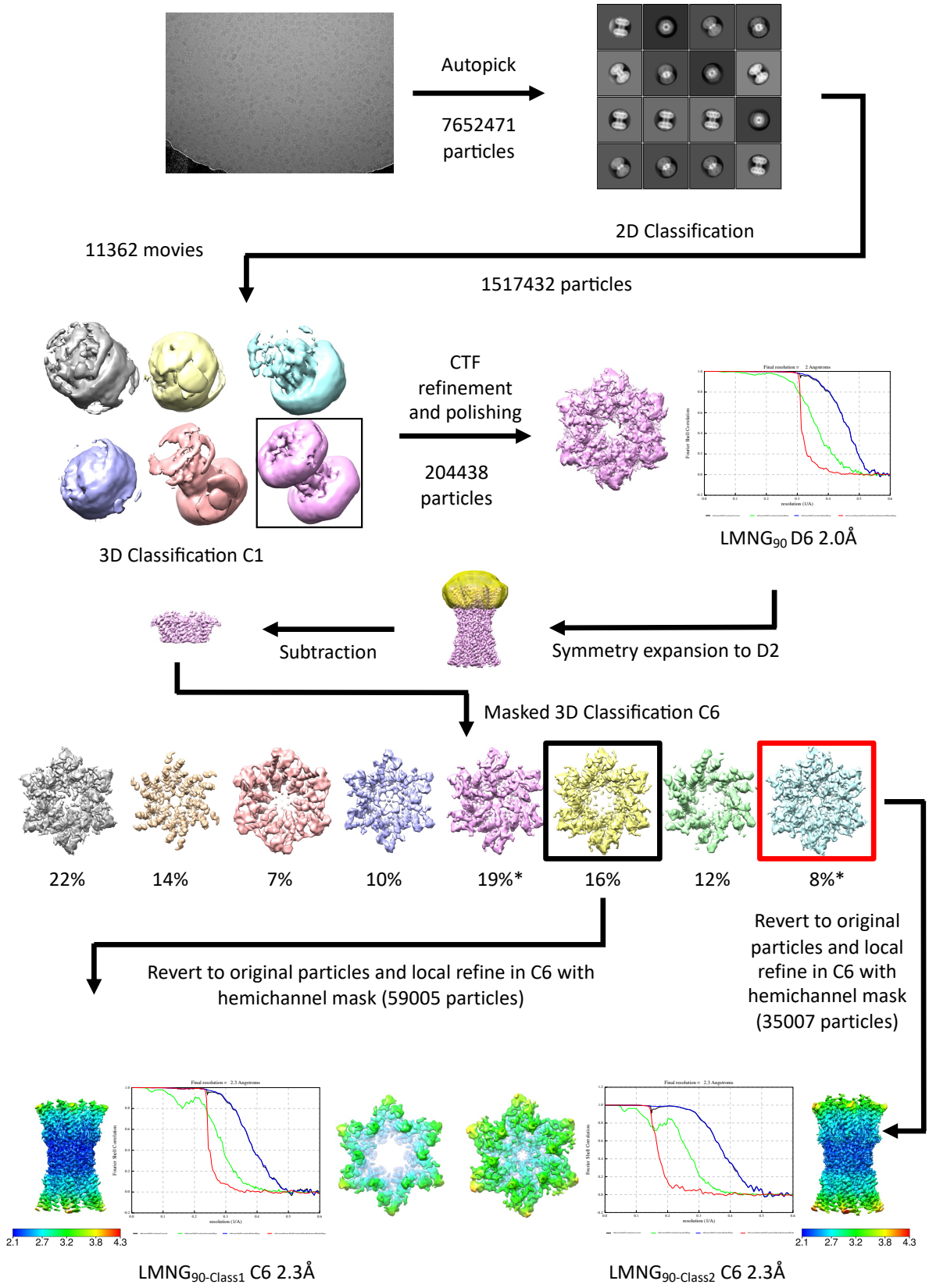

Figure S7: Workflow for initial processing of cryo-EM data for WT sample, solubilised in LMNG in CO_2/$HCO3^-$ buffer. The star denotes the classifications with the appearance of the PCN conformation that refine to a resolution greater than 4Å.

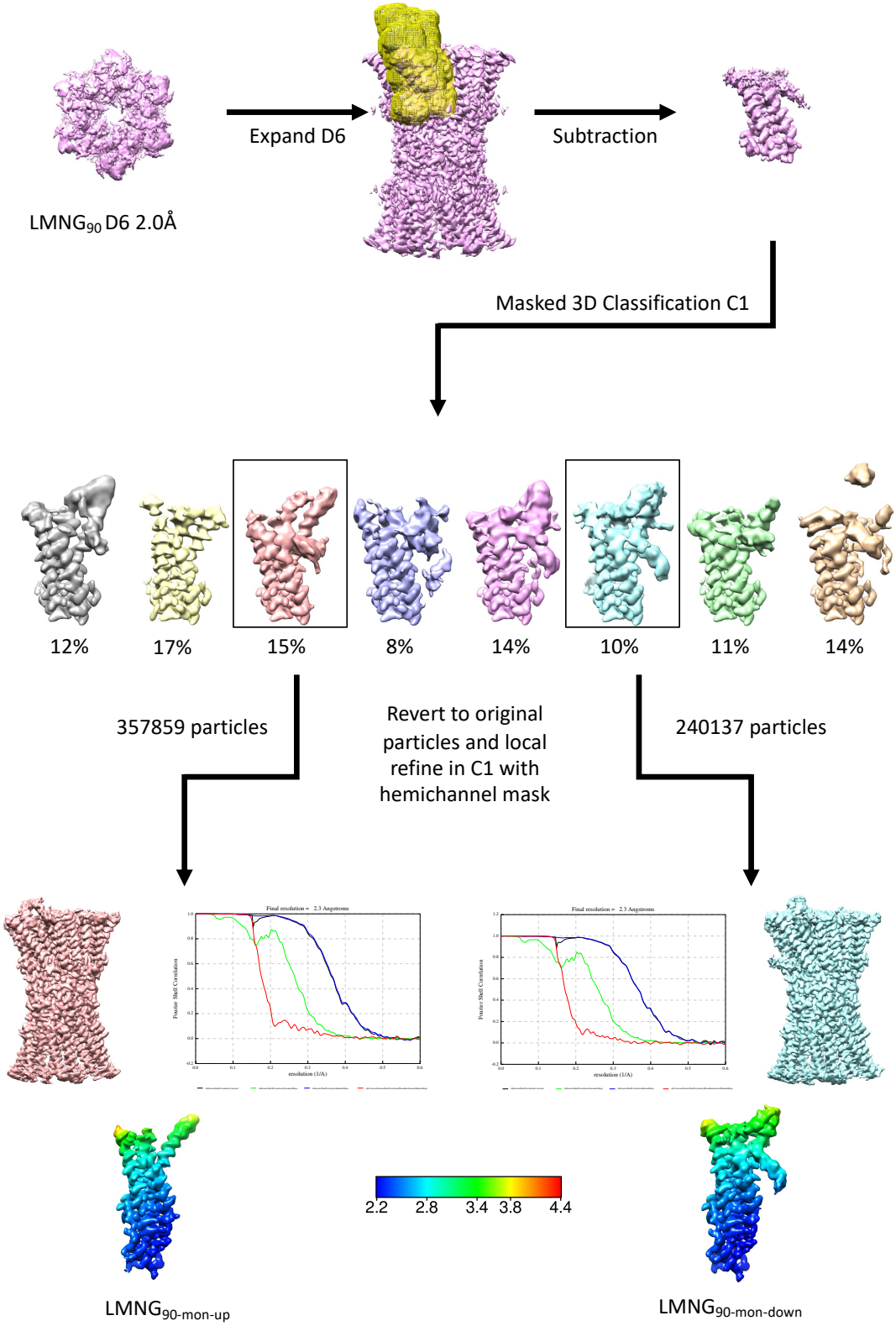

Figure S8: Workflow for single subunit classification for LMNG solubilised sample

Movie S1: Morph showing the conformational difference between D6 refined reconstructions of WT and K125E. WT_{90} (pink) K125E (blue) The position of TM2 is highlighted by an oval in one of the subunits.

Movie S2: Morph showing the conformational differences between D6 refined reconstructions of WT and K125E in HEPES buffer. WT (pink) K125E (blue) The position of TM2 is highlighted by an oval in one of the subunits.

Movie S3: Morph showing the conformational differences between reconstructions of $LMNG_{90\text{-class1}}$ and $LMNG_{90\text{-class2}}$. $LMNG_{90\text{-class2}}$ (pale blue) $LMNG_{90\text{-class1}}$ (yellow).

Movie S4: Morph showing the conformational differences between structures of $LMNG_{90\text{-class1}}$ and $LMNG_{90\text{-class2}}$. a) Overall. As the N-terminus and KVRIEG motif are not visible in the class 6 structure these residues are not present in the morph. Residues linking the N-terminal helix to TM1 are shown in pink. Trp24 is depicted by pink sticks. **b)** Focussed on TM1 with TM1, 2, 3, 4 of one subunit depicted in cyan, green, yellow and red respectively and TM2 of a second subunit in darker green. The movement of TM2 stems from the region around Phe 83 so that the position of the C_α atom of Lys103 near the top of TM2 of the respective structures differs by ~8.5Å