doi:10.1016/j.jmb.2004.04.023 J. Mol. Biol. (2004) 339, 981–1000

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Electrostatics of Ion Stabilization in a ClC Chloride Channel Homologue from *Escherichia coli*

José D. Faraldo-Gómez and Benoît Roux*

Department of Biochemistry Weill Medical College of Cornell University, Whitney Building, 1300 York Avenue New York, NY 10021, USA

*Corresponding author

The structural determinants of electrostatics of ion stabilization within EcClC, a ClC-type chloride channel homologue from *Escherichia coli*, are studied using a continuum dielectric approximation. Specifically, the ion occupancy is investigated in the wild-type protein and a mutant thereof, and the contribution to the electrostatic binding free energy of local and non-local interactions is characterized at the single-residue level. This analysis shows that, in spite of the desolvation cost and the strong ion–ion repulsion, all previously reported binding sites can be occupied simultaneously. The stabilizing effect of the protein arises from hydrogen bonding as well as from longer-range favorable interactions, such as with the strictly conserved Lys131 side-chain. The latter is involved in the stabilization of the conserved GSGIP motif that delimits two of the binding sites. Interestingly, an additional low-affinity binding site, mediated by a structurally analogous motif including the side-chain of Arg340, can be identified on the extracellular side of the permeation pathway. Finally, it is shown that, in contrast to K-channels, and in analogy to the SBP/PBP sulfate/phosphate-binding proteins, the contribution of helix macrodipoles to chloride binding in EcClC is only marginal.

Keywords: ion channel; ClC channel; electrostatics; Poisson–Boltzmann; helix dipole effect

Introduction

The ClC family is the most ubiquitously expressed of all classes of chloride channels established thus far.[1–3] To date, up to nine ClC channels involved in a wide range of physiological functions have been identified in mammalian cells, both in the cell membrane and in the membranes of intracellular organelles. In particular, ClC-K channels regulate transepithelial transport of ions in the inner ear and in the nephrons, and ClC-1 regulates the electrical excitability of the plasma membrane in skeletal muscle cells; ClC-5 is believed to control the acidification of intracellular endosomes, for instance in the kidney, and ClC-7 channels are essential to sustain the acidic environment required for bone degradation by osteoclasts. Consistent with these physiological roles, several genetic diseases in humans are associated with functionally defective ClC mutants. These include Bartter's syndrome (a salt-wasting disorder), both Becker's and Thomsen's myotonias (muscle stiffness), Dent's disease (formation of kidney stones) and osteopetrosis (excessive bone density).[1–3] Several ClC-type homologues have been identified in bacteria and archea; for example, in *Escherichia coli* the proteins encoded by the *eriC* and *mriT* genes have been shown to be involved in the extreme acid-resistance response that allows benign and pathogenic bacteria to survive in highly acidic environments, such as in the stomach.[4]

ClC channels have proved to be a continuous source of unprecedented observations for researchers in the ion channels field. For instance, their doubled-barrelled molecular architecture, with two off-axis, independently conducting pores, is more similar to that of multimeric outer membrane porins or of aquaporins than to the structure of other known ion channels, such as the potassium channels and the nicotinic acetylcholine receptor.[5] The opening and closing of each of these pores on the millisecond timescale, or fast-gating,

Supplementary data associated with this article can be found at doi: 10.1016/j.jmb.2004.04.023

Abbreviations used: SBP, sulfate-binding protein; PBP, phosphate-binding protein.

E-mail address of the corresponding author: benoit.roux@med.cornell.edu

0022-2836/$ - see front matter © 2004 Published by Elsevier Ltd.

is uncorrelated, although a common, much slower gating mechanism that affects both pores simultaneously has been detected in some of the eukaryotic ClC channels. Furthermore, fast-gating in many of the eukaryotic ClC channels is dependent on both pH and transmembrane voltage. However, in contrast to other known voltage-gated ion channels, the gating charge appears not to be carried by side-chains in the protein, but rather by the permeant chloride ions.[1,2] Finally, and strikingly, a recent study has demonstrated that one of the ClC homologues in *E. coli* is not a channel, but a proton–chloride antiporter with a likely stoichiometric ratio of two Cl^- per proton.[6]

From a structural and topological viewpoint, ClC channels are arguably among the most complex membrane proteins known. As revealed by the crystal structures of the bacterial ClC homologues EcClC (or EriC) and StClC,[7,8] the transmembrane domain of each of the subunits or monomers is made up of 17 α-helices arranged in two distinct antiparallel domains. These helices vary in length (up to ~40 residues) and tilt relative to the membrane plane (up to ~45°), and several of them penetrate the membrane only partially. The interface between the subunits is formed primarily within the membrane, but an additional N-terminal helix "domain-swaps" so as to associate with the C-terminal helix of the opposite subunit at the cytoplasmic side of the protein.

A crucial contribution made by the crystallographic studies of EcClC and StClC has been to locate the conduction pore, within which three binding sites for chloride ions have been identified (Figure 1A). Two of these, termed S_{cen} and S_{ext}, are found within the transmembrane domain, approximately 4 Å apart (Figure 1B). A third binding site, S_{int}, is found on the intracellular face of the protein, only about 7 Å away from S_{cen}. Nonetheless, it is unclear how ions move between the S_{int} and S_{cen} sites, since the pathway between these two sites appears to be narrowed significantly by neighboring hydrophobic side-chains and by the protein's backbone. Mutagenesis combined with structural studies (when Glu148 is mutated to Ala or Gln) suggest that S_{ext} can be occupied by the side-chain of Glu148 or by a chloride ion. In the former case, permeation towards the extracellular side from S_{cen} appears to be prevented by Glu148 itself, and thus this highly conserved residue has been proposed as the fast gate in ClC channels.[8] This hypothesis seems to be consistent with the pH and voltage-dependence of the gating properties of some ClC channels and their corresponding mutants,[8–11] and by the fact that in EcClC, Glu148 is required for H^+ antiport.[6] However, the actual mechanism of gating, and its coupling to the translocation of chloride ions, pH and applied voltage, is yet to be elucidated.

Thus, in spite of the ground-breaking insights provided by the X-ray structures, the molecular mechanism by which ClC channels mediate permeation of chloride ions, whether driven solely by electrochemical diffusion or coupled to H^+ transport, remains to be understood. As shown for the potassium channel KcsA,[12–17] electrostatic effects are expected to play a prominent role. The purpose of this investigation is to identify the structural

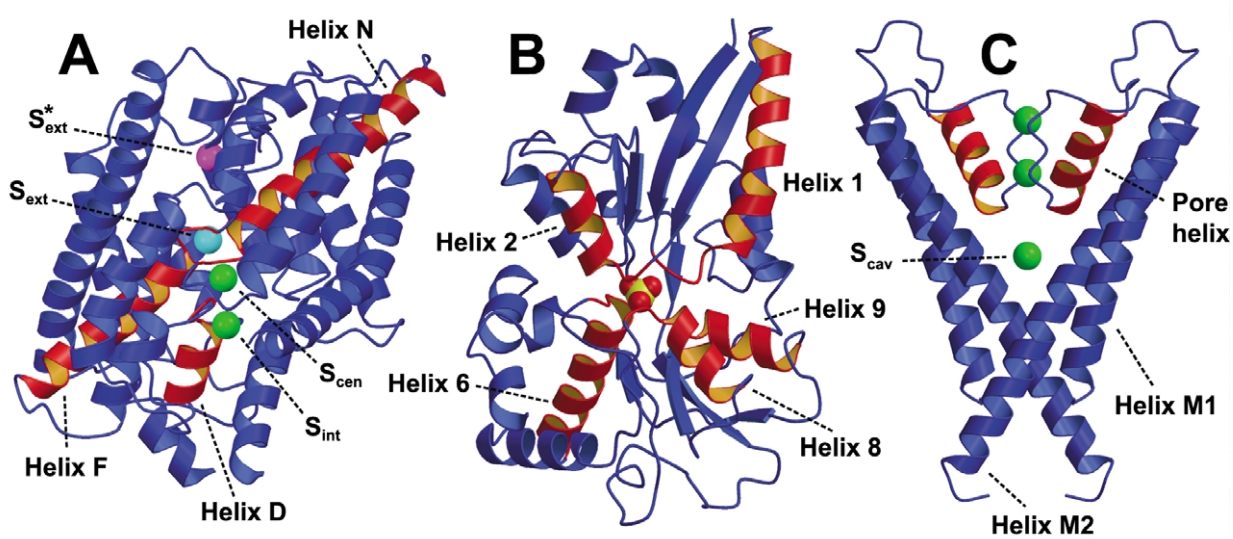

Figure 1. Ribbon representation of (A) the EcClC chloride channel homologue (one monomer),[8] (B) the sulfate-binding protein SBP,[32] and (C) the potassium channel KcsA.[24] For clarity, EcClC residues 81–95/285–299 and two KcsA subunits are not shown. Chloride, potassium and sulfate ions are depicted as van der Waals spheres; in EcClC, the cyan sphere corresponds to C^{δ} in the Glu148 side-chain, which is replaced by a third chloride ion in EcClC-E148Q; the magenta sphere represents an additional chloride ion in the putative binding site S^*_{ext}, undetected in the crystallographic studies. Both EcClC and KcsA are depicted from the membrane plane, with the extracellular side at the top of the Figure. This Figure was created with MOLSCRIPT[62] and rendered with Raster3D.[63]

elements that are responsible for the electrostatic stabilization of chloride ions within the conduction pore of the EcClC chloride channel homologue, and to characterize their influence. Through these calculations we hope to provide clues for future analyses of the actual permeation and gating processes, as well as for the interpretation or design of experimental work.

A special focus has been placed on the analysis of ion-helix interactions in EcClC; namely, on the importance of the so-called helix dipole effect. In particular, these interactions are compared to those in potassium channels and the anion-binding proteins sulfate-binding protein (SBP) and phosphate-binding protein (PBP). The notion that the intrinsic electrostatic macrodipole of protein α-helices can contribute importantly to the energetic stabilization of ionic species (ions, ligands or charged side-chains) in the vicinity of their N or C terminus is still often quoted in protein structure studies, as well as in recent editions of structural biology text books.[18] This perspective is somewhat striking, in view of theoretical work conducted more than a decade ago, on SBP as well as on barnase, which indicated that such stabilization is due predominantly to local interactions occurring at the helix termini.[19] Nonetheless, subsequent electrostatics calculations on the transmembrane potassium channel KcsA showed that the so-called pore helices have a critical role in stabilizing a potassium ion in a water-filled cavity, even though these helices are not in direct contact with the ion.[16] The reasons behind this apparent discrepancy are not fully clear; differences might arise from the choice of methodology, or from the proteins' distinct topologies or environment. For example, it was shown that the low dielectric environment provided by the membrane helps to magnify the stabilizing effect of the pore helices in KcsA,[16] but that this effect might not be as significant in other membrane proteins. As for KcsA, helix macrodipoles have been proposed to be involved in ion binding in the EcClC protein.[7,20] However, EcClC is topologically more similar to SBP and other soluble ion-binding proteins, whereas its environment is analogous to that of KcsA. Therefore, it is not clear whether the macrodipoles will play a crucial role when compared to the effect of side-chains or other charges that are more proximal to the ion-binding site. In order to clarify this issue, we have adopted a common framework in which we compare the helix dipole effect in EcClC, in SBP and PBP, and in the KcsA, KirBac, and MthK potassium channels.

Results and Discussion

Ion occupancy and overall electrostatics in EcClC

In order to characterize the energetics governing ion permeation in ClC channels, it is essential to determine what is the ion occupancy of the conduction pore. In the absence of crystallographic data specifically addressing this question, such as that for the KcsA potassium channel,[21,22] we have relied on calculations of the electrostatic free energy of binding, ΔG_B, to elucidate the occupancy of the pore. It should be noted, though, that under the Poisson–Boltzmann framework adopted here, electrostatic binding free energies may be markedly dependent on methodological uncertainties, such as the protein's dielectric constant, ϵ_p, the representation of the membrane and the protonation states of ionisable residues. Thus, it is important to ensure, first, that reasonable choices are made in the set-up of the calculations, and second, that any conclusions based on those calculations hold valid over a range of realistic alternatives.

The results of the calculations of ΔG_B for each of the sites along the conduction pore of EcClC are summarized in Table 1. Binding of chloride ions simultaneously at sites S_{int} and S_{cen} in EcClC-WT was found to be energetically favorable, despite the presence of the negatively charged Glu148 side-chain in the S_{ext} site, provided that the side-chain of Glu113 is protonated. For example, for $\epsilon_p = 4$, ΔG_B is -9.2 kcal/mol for S_{cen} and -2.2 kcal/mol for S_{int}. Consistently, our calculations for the E148Q mutant indicate that three ions can bind to the channel concurrently, e.g. ΔG_B is -7.8 kcal/mol for S_{ext}, -7.0 kcal/mol for S_{cen}, and -1.9 kcal/mol for S_{int}.

By contrast, default protonation states in all ionizable residues were found to be incompatible with simultaneous occupancy of the pore. However, this is a highly improbable configuration, according to our analysis of the pK_A shifts in all EcClC structures (see Methods), which indicates strongly that the side-chains of Glu113 and Glu203 are very likely to share a proton within the physiological range of pH values. Differences in the description of the protein–membrane interface (having already optimized their relative location; see Methods) had only a marginal effect on the calculated binding energies, with the flat and curved membrane models giving the same results within 0.1 kcal/mol, compared to within 1 kcal/mol when no membrane model was used at all. This is consistent with the fact that the binding sites are more than 20 Å away from the protein–membrane interface, either buried within the protein or on its intracellular face.

As expected, binding free energies were found to be very sensitive to the value of ϵ_p, since this parameter determines the strength of the electrostatic interactions at a given range, as well as the magnitude of the desolvation energy. Nevertheless, simultaneous occupancy was consistently favored across the range of values considered, with the exception of S_{int} when $\epsilon_p = 2$. It should be noted, however, that these calculations used the crystallographic atomic coordinates of protein and ions, and only the position of hydrogen atoms in

Table 1. Electrostatic contribution to the binding free energy of chloride ions to the S_{cen}, S_{int}, and S_{ext} binding sites in EcClC (wild-type and E148Q mutant)

Ion/ protein	ΔG_B (kcal/mol)								
	Default protonation states, $\epsilon_p = 4$			Adjusted protonation states, flat membrane					
	Crystal structure			Crystal structure				After minimisation	
	Curved membrane	Flat membrane	No membrane	$\epsilon_p = 2$	$\epsilon_p = 4$	$\epsilon_p = 6$	$\epsilon_p = 8$	$\epsilon_p = 2$	$\epsilon_p = 4$
S_{cen}									
EcClC-WT	−4.6 (0.5)	−4.7 (0.5)	−3.7 (0.5)	−14.6 (0.9)	−9.2 (0.4)	−7.4 (0.3)	−6.5 (0.2)	−14.5 (0.6)	−9.5 (0.3)
EcClC-EQ	−2.4 (0.8)	−2.5 (0.8)	−1.7 (0.8)	−10 (1)	−7.0 (0.8)	−6.0 (0.6)	−5.4 (0.5)	−16.7 (0.4)	−10.5 (0.2)
S_{int}									
EcClC-WT	0.5 (0.4)	0.5 (0.4)	1.4 (0.3)	−0.3 (0.6)	−2.2 (0.4)	−2.8 (0.3)	−3.2 (0.3)	−2.9 (0.3)	−4.4 (0.2)
EcClC-EQ	0.9 (0.8)	0.9 (0.8)	1.8 (0.8)	0 (1)	−1.9 (0.8)	−2.6 (0.5)	−3.0 (0.4)	−3 (1)	−4.6 (0.8)
S_{ext}									
EcClC-EQ	−4.8 (0.5)	−4.9 (0.5)	−4.3 (0.4)	−12 (1)	−7.8 (0.6)	−6.5 (0.4)	−5.8 (0.3)	−13.2 (0.4)	−8.8 (0.2)

In all cases, averages were taken over the two protein subunits (in addition to the grid translational averaging); the corresponding standard deviations are given within parentheses. The results from calculations that used energy-minimised structures are shown in order to illustrate the effect of small variations in the atomic coordinates in the electrostatic binding energies. Energy minimisations were conducted with CHARMM, using 500 steps of the conjugate gradients algorithm with a distance-dependent dielectric constant and a 14 Å cut-off for all non-bonded interactions. The position of C^{α} atoms, chloride ions and Glu148 was kept fixed during the minimisations; for other atoms, RMS displacements of less than 0.7 Å were obtained, both in the vicinity of the binding sites and elsewhere in the protein. This fluctuation corresponds to a Debye–Waller *B*-factor of 13 $Å^2$, a value much smaller than those obtained for the crystal structures of EcClC. *B*-factors are related to atomic displacements Δr *via* the expression $B = 1/3\ 8\pi^2\ (\Delta r)^2$.

hydroxyl groups were energy-minimised. This is a conservative choice, in that crucial ion–protein interactions, such as with backbone amide dipoles, may be underestimated significantly, since electrostatic interactions were not optimised, or even included, during the construction and refinement of the crystallographic structures. Indeed, short energy minimisations of the structures of EcClC-WT and EcClC-E148Q, resulting in mean displacements of less than 0.7 Å in atoms within 10 Å from the chloride ions (see the legend to Table 1 for more details), yield electrostatic binding energies that are substantially more favorable, including that to the S_{int} site when $\epsilon_p = 2$. Such displacements are certainly within the experimental uncertainty in the atomic coordinates; the average *B*-factor in the same subset of atoms is 63 $Å^2$ in EcClC-WT and 75 $Å^2$ in EcClC-E148Q, which correspond to atomic fluctuations of 1.5 Å and 1.7 Å, respectively.

In summary, our analysis of the electrostatic binding free energies indicates strongly that the S_{ext}, S_{cen} and S_{int} sites in the conduction pore of EcClC can be occupied concurrently. For conciseness, we will restrict the discussion to calculations using the crystallographic coordinates, adjusted protonation states, $\epsilon_p = 4$ and a flat membrane model.

The stabilization of ions within a protein arises from favorable ion–protein interactions that counter-balance the desolvation energy of the ion (and, to a lesser degree, of the protein, too) and the ion–ion electrostatic repulsion. The magnitude of each of these contributions in the current case is given in Table 2. For example, for the ion in S_{cen} in EcClC-WT, the desolvation penalty is 17.6 kcal/mol, and the ion–ion repulsion within the same subunit is 5.8 kcal/mol (the interactions of ions across subunits are less than 0.4 kcal/mol). These unfavorable contributions are sufficiently compensated by a net favorable protein–ion interaction energy; namely, −29.9 kcal/mol within the same subunit, in addition to −3.1 kcal/mol arising from the other subunit (specifically from residues in the N-terminal domain-swapping helix). The largest contribution to the intra-subunit protein interactions between the ion in S_{cen} and EcClC-WT comes from the backbone, with −25.0 kcal/mol. Nonetheless, crucial side-chain interactions exist that contribute to counter-balance unfavorable repulsions such as with the Glu148 side-chain in S_{ext}. This is clear from the analysis of the E148Q mutant, where the side-chain contribution amounts to about 40% of the total intra-subunit interaction energy.

The energetics of binding to S_{int} differs firstly in that this site is on the surface of the protein, which reduces the energy cost of desolvation, and second in that the protein–ion interactions are significantly weaker (which is also related, to some degree, to the larger exposure to solvent). Nevertheless, the S_{int} site is electrostatically stable in the presence of negative charges at S_{cen} and S_{ext}. Interestingly, side-chain interactions appear to be even more crucial for this ion than for that in S_{cen}, adding up to 50%–60% of the total ion–protein interaction energy. By contrast, backbone interactions contribute predominantly to chloride binding to the S_{ext} site in EcClC–E148Q, specifically about 80% of the ion–protein interaction energy,

Table 2. Breakdown of the electrostatic contribution to the binding free energy for the S_{cen}, S_{int}, and S_{ext} sites in EcClC (wild-type and E148Q mutant)

Ion/protein	Breakdown of ΔG_B (kcal/mol), adjusted protonation states, flat membrane, $\epsilon_p = 4$						
	ΔG_{rf}		ΔG_{int} (same subunit)			ΔG_{int} (across subunits)	
	Ion	Rest	Other ions	Protein	Backbone	Other ions	Protein
S_{cen}							
EcClC-WT	17.6 (0.3)	0	5.82 (0.02)	−29.9 (0.3)	−25.0 (0.2)	0.34 (0.02)	−3.1 (0.1)
EcClC-EQ	17.5 (0.3)	0	22.7 (0.1)	−44 (1)	−26.6 (0.4)	0.53 (0.01)	−3.4 (0.2)
S_{int}							
EcClC-WT	9.6 (0.4)	0.94 (0.2)	5.82 (0.02)	−16.1 (0.3)	−9.2 (0.2)	0.24 (0.02)	−2.8 (0.1)
EcClC-EQ	9.7 (0.4)	1.0 (0.2)	7.8 (0.2)	−17.5 (0.2)	−7.2 (0.2)	0.37 (0.01)	−3.0 (0.1)
S_{ext}							
EcClC-EQ	17.1 (0.4)	0	20.30 (0.2)	−43.3 (0.6)	−31.6 (0.7)	0.51 (0.01)	−2.52 (0.2)

ΔG_{rf} denotes the change in the reaction field energy upon association, either of the ion or of the rest of the protein–ion complex. ΔG_{int} is the electrostatic interaction energy of the ion with the rest of the protein–ion complex, either within the same subunit or across subunits. The backbone contribution to the ion–protein interaction energy within the same subunit is given. Averages were taken over the two protein subunits (in addition to the grid translational averaging); the corresponding standard deviations are given within parentheses.

or by analogy, to the stabilization of the charged form of the Glu148 side-chain in EcClC-WT (calculations of pK_A shifts of both carboxyl oxygen atoms in the Glu148 side-chains in EcClC-WT yield $pK_A^{max} \ll pK_A^{sol} \approx 4$ in the presence of ions at S_{cen} and S_{int}; see Methods).

Finally, it is worth noting the distinct asymmetry of the charge distribution around the binding sites, notably S_{int}, in marked contrast to potassium channels where the conduction pore itself runs along the 4-fold symmetry axis of the protein.[23,24] In the current theoretical framework, this asymmetry has no effect on the binding energetics, since the electrostatic energy depends on the electrostatic potential, which is a scalar quantity. However, given that chloride is a highly polarizable ion,[25–27] it seems appropriate to assess, even to an approximate level, the importance of the energetic contribution arising from the interaction between the electric field generated by the protein's charge distribution and the induced dipole on the ion. To this end, we computed the magnitude of the electric field, E, on the position of each of the ions in each subunit using the same methodology as for the calculation of the electrostatic energies (see Methods), except that the grid-point density was increased 1000-fold to improve the accuracy of the results. Using $\epsilon_p = 4$, the resulting values range between 170 mV/Å and 210 mV/Å, while for $\epsilon_p = 2$, we obtained values between 330 mV/Å and 430 mV/Å. The corresponding energies, $U = -1/2\alpha E^2$, where $\alpha \sim 4$ is the chloride ion polarizability in the aqueous phase,[25] are around −0.15 kcal/mol and −0.50 kcal/mol, respectively. Therefore, we conclude that, at this level of approximation, the contribution associated with chloride ion polarizability to binding to EcClC is not quantitatively significant, albeit favorable, when compared with the corresponding coulombic interaction energies as calculated *via* Poisson's equation.

The energetic and structural role of Lys131 in chloride binding

A rather interesting characteristic of EcClC channels is the extent to which side-chains are seemingly involved in the permeation process. In addition to Glu148, which itself competes with chloride ions for one of the binding sites and is thought to act as the channels' gate, the hydroxyl groups of side-chains Ser107 and Tyr445 participate, along with various backbone amine dipoles, in the coordination of chloride ions.[7,8] According to the current calculations, however, the single most important favorable ion-side-chain electrostatic interaction in EcClC originates not from Ser107 or Tyr445, but from Lys131 (Figure 2). This side-chain is located in transmembrane helix E, completely buried within the protein, with its positively charged amino group pointing towards the chloride-binding sites, at a distance of 7–9 Å. Specifically, the interaction energy of Lys131 with the ion in S_{cen} is −5.9 kcal/mol, −4.7 kcal/mol with the ion in S_{int}, and in the E148Q mutant, −5.0 kcal/mol with that in S_{ext} (using $\epsilon_p = 4$).

In view of the magnitude of the ion–protein binding energies reported previously, these results suggest that Lys131 is crucial from an energetics standpoint alone. This appears to be particularly true for the binding of chloride ions to S_{int}, which lacks the hydroxyl interactions with Ser107 and Tyr445. But, in addition, Lys131 seems to have an important structural role in stabilizing the geometry of the S_{cen} and S_{int} binding sites, by simultaneously hydrogen bonding with the carbonyl oxygen atoms of Gly106 and Ser107 in the GSGIP motif that follows the transmembrane helix C and precedes the re-entrant helix D. This arrangement fixes the orientation of the associated N–H backbone dipoles, i.e. those in Ser107 and Gly108, which coordinate the ion in S_{int}, as well as the conformation of the Ser107 side-chain, to which the

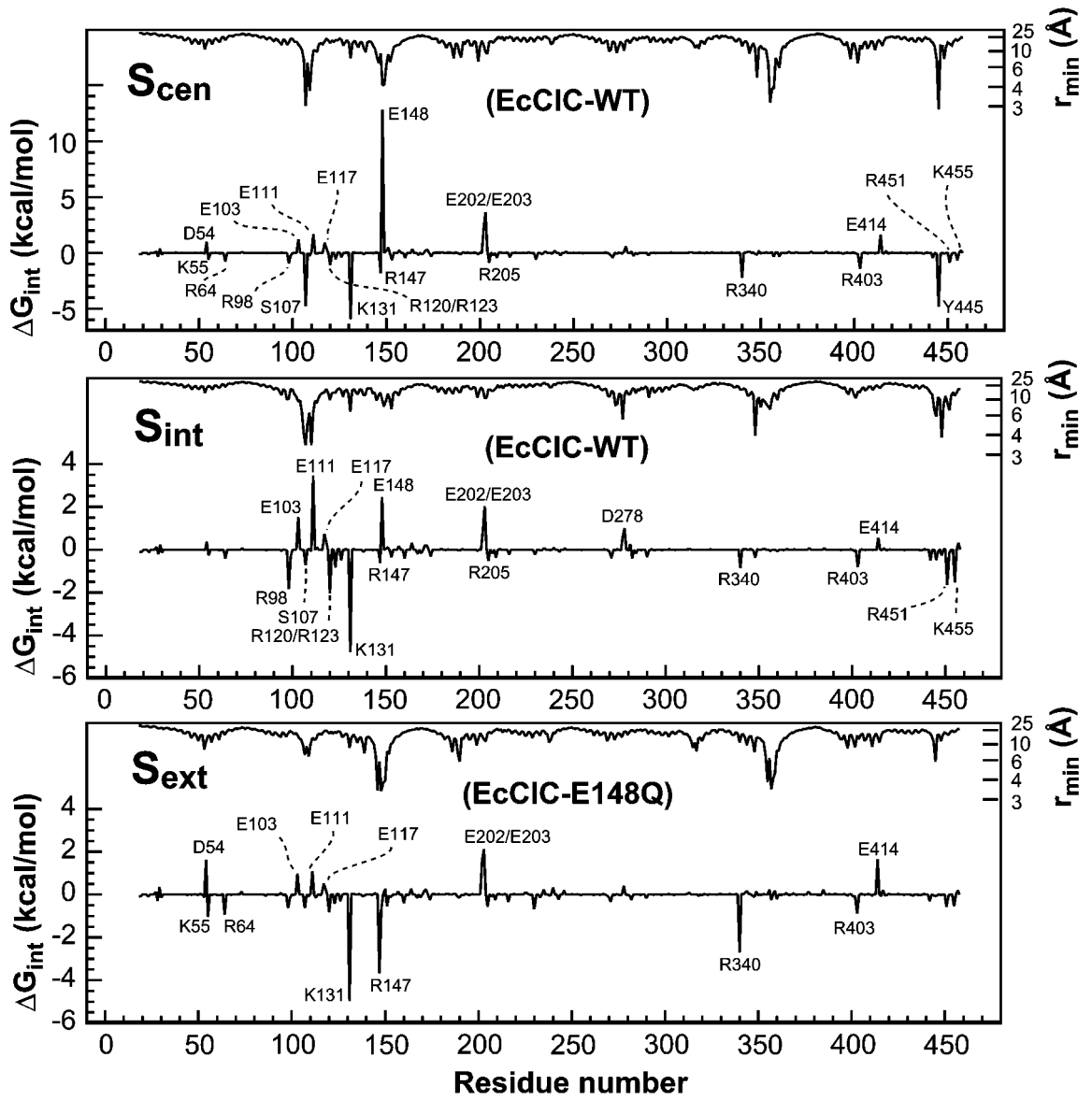

Figure 2. Ion-side-chain electrostatic interactions in the EcClC chloride channel homologue, *versus* residue number, alongside the corresponding ion-side-chain minimum distances. Favorable interactions leading to ion stabilization are discussed in the text. Among the unfavorable side-chain interactions, aside from Glu148, the most notable originate from Glu111 and Glu203, which add up to 5–6 kcal/mol, depending on the ion considered. Glu111 is located in the vicinity of the intracellular binding site, and is conserved across the ClC family; in the eukaryotic ClC-0, this side-chain has been shown to play a critical role in controlling channel conductance.[64] In contrast, Glu203 is often substituted by Val, and in EcClC appears to be primarily important for stabilizing the inter-subunit domain-swapping of helix A (by ion-pairing with Arg28) at the intracellular side of the channel. Glu203 also forms a hydrogen bond with the side-chain of Glu113 (neither Glu113 or Arg28 is conserved).

chloride in S_{cen} binds (Figure 3A). Both Lys131 and the GSGIP motif are strictly conserved across the entire ClC family (except ScClC in yeast, where the sequence is GSGIS), as are many of the residues around the binding sites. Consistently, analysis of homology models of several eukaryotic ClC channels yields similar observations regarding Lys131 and the GSGIP motif (our unpublished results).

This putative dual function of Lys131 is reminiscent of the carboxyl–carboxylate Glu71-Asp80 pair in the potassium channel KcsA. Hydrogen bonds between these residues and the backbone amide dipoles of the selectivity filter contribute to its structural stability.[24,28,29] Simultaneously, the singly charged Glu71-Asp80 pair provides the dominant non-backbone contribution to the electrostatic interaction energy with the permeating potassium ions (roughly -10 kcal/mol with $\epsilon_p = 4$), consistent with experimental data regarding the effect of mutations at the position of Asp80 in a mammalian K-channel.[30] Nonetheless, an important difference exists between EcClC and

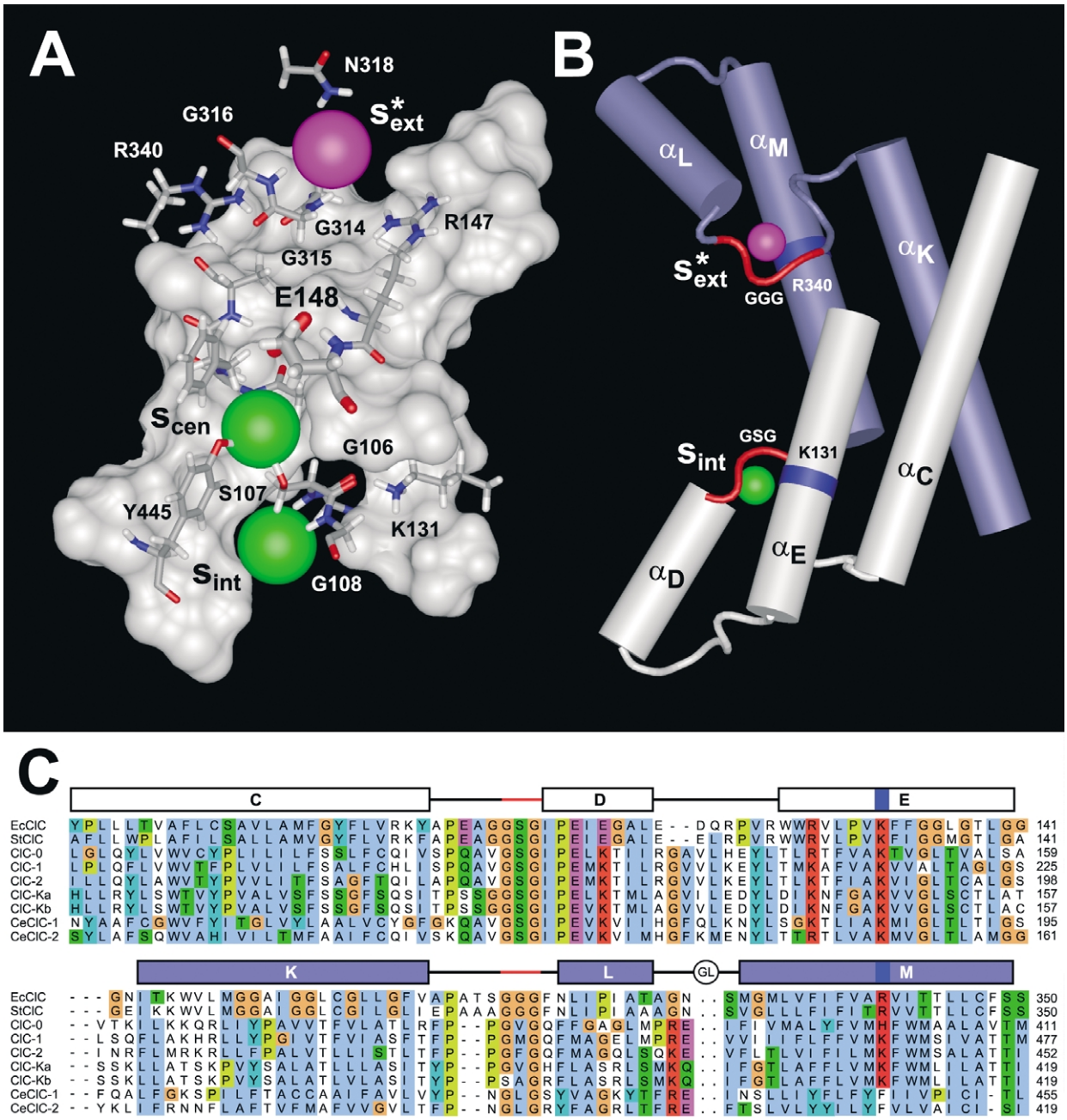

Figure 3. A, Chloride-binding sites along the conduction pore of EcClC. Chloride ions are shown in the binding sites identified in the X-ray structures (green), as well as in the putative binding site S^*_{ext} (magenta); in EcClC-WT, the Glu148 side-chain occupies the S_{ext} site (thick lines). The Lys131-GSGIP and Arg340-GGGFN motifs that mediate ion binding at the intracellular and (presumably) extracellular side are illustrated, along with other residues engaged in the electrostatic stabilization of ions within the protein. B, Antiparallel arrangement of helices C, D and E on the intracellular side, and helices K, L and M on the extracellular side, within each of the EcClC subunits. C, Multiple sequence alignment of representative members of the ClC family, for the regions comprising the S_{int} and S^*_{ext} sites. A long loop containing a glycosylation site exists between helices L and M in eukaryotic channels (dots). To facilitate the comparison, gaps have been inserted in the sequences (dashes). A and B were created with VMD[65] and Raster3D.[63] The sequence alignment in C was generated with CLUSTAL W[66] and edited with Jalview (www.jalview.org).

KcsA in this respect, in that the conserved GSGIP motif, and in particular Ser107, largely restricts the access to S_{cen} from S_{int}, whereas the movement of ions between, say, the S0 and S2 sites in the filter of KcsA is unobstructed. Given that GSGIP motif lacks secondary structure and is thus an intrinsically flexible region, it is plausible that a dynamic coupling exists between the ions, Lys131 and the GSGIP segment that might be involved in ion-exchange between these two sites.

Table 3. Electrostatic contribution to the binding free energy of chloride ions to the sites reported by Dutzler *et al.*,[8] and to the extracellular binding site S^*_{ext} proposed in the current study

Ion/protein	ΔG_B (kcal/mol)					
	Crystal structure				After minimisation	
	$\epsilon_p = 2$	$\epsilon_p = 4$	$\epsilon_p = 6$	$\epsilon_p = 8$	$\epsilon_p = 2$	$\epsilon_p = 4$
S_{cen}						
EcClC-WT	−13.8 (0.8)	−8.5 (0.4)	−6.7 (0.3)	−5.9 (0.2)	−14.5 (0.5)	−9.1 (0.2)
EcClC-EQ	−10 (1)	−6.8 (0.8)	−5.8 (0.6)	−5.3 (0.5)	−16.6 (0.5)	−10.4 (0.3)
S_{int}						
EcClC-WT	−0.1 (0.6)	−2.0 (0.4)	−2.6 (0.3)	−2.9 (0.3)	−2.8 (0.3)	−4.2 (0.3)
EcClC-EQ	0 (1)	−1.8 (0.8)	−2.5 (0.5)	−2.9 (0.4)	−3 (1)	−4.6 (0.8)
S_{ext}						
EcClC-EQ	−11 (1)	−7.6 (0.6)	−6.3 (0.4)	−5.6 (0.3)	−13.2 (0.5)	−8.7 (0.3)
S^*_{ext}						
EcClC-WT	−4.3 (0.5)	−3.8 (0.4)	−3.5 (0.4)	−3.4 (0.4)	−5.5 (0.7)	−4.2 (0.3)
EcClC-EQ	−0.95 (0.2)	−1.7 (0.1)	−1.98 (0.07)	−2.1 (0.06)	−6.2 (0.3)	−4.5 (0.1)

The electrostatic calculations and energy minimisations are analogous to those in Table 1. Averages were taken over the two protein subunits (in addition to the grid translational averaging) except for the X-ray structure of EcClC-WT, where the results shown correspond to subunit B only (see the text). In the energy-minimised structure of EcClC-WT, the side-chain of Asn318 in subunit A adopts the same rotamer as in subunit B, so as to coordinate the chloride ion in S^*_{ext}. In the X-ray structure of EcClC-E148Q, Asn318 does not contribute significantly to the binding free energy in either subunit. However, this side-chain adopts a conformation similar to that observed in subunit B of EcClC-WT after energy-minimising the structure. The displacements in the N^δ atom in Asn318 caused by the energy-minimisations are less than 1 Å, which correspond to a Debye–Waller *B*-factor of 27 Å^2. The experimental *B*-factors of the Asn318 side-chain are 65–85 Å^2 in EcClC-WT and 70–93 Å^2 in EcClC-E148Q.

An additional putative binding site on the extracellular side of the pore

Among the other side-chains providing favorable interactions for ion binding in EcClC, the most noteworthy is Arg147, a conserved charge in all ClC channels located towards the extracellular side of the permeation pathway, which contributes −4.8 kcal/mol and −2.0 kcal/mol to binding of chloride to S_{ext} and S_{cen}, respectively. These ions also interact significantly with Arg340, although to a lesser extent. Arg340 is of particular interest because it is structurally equivalent to Lys131, by virtue of the antiparallel symmetry of the two halves of each EcClC subunit (Figure 3B). This residue is located halfway along transmembrane helix M and interacts with the backbone of the loop connecting transmembrane helix K and re-entrant helix L, in particular with a GGGFN segment (residues 314–316) that is equivalent to the GSGIP region mentioned above. By analogy with the Lys131-GSGIP motif for the S_{int} site, it seems plausible that the Arg340-GGGFN region may serve as an extracellular binding site, where anions would be coordinated by the N–H dipoles of Gly315 and Gly316, while Arg340 fixes the orientation of the carbonyl groups of Gly314 and Gly315, and interacts favorably with the bound anion (Figure 3A). However, no ion has been detected at this location in crystallographic studies thus far, although it is worth noting that a crystallographic water molecule was reported precisely at this site in subunit B of EcClC-WT.[7,8]

To investigate this hypothesis, we conducted electrostatic calculations of EcClC-WT and EcClC-E148Q models that contain a chloride ion at this putative binding site, S^*_{ext}, in addition to those at S_{int} and S_{cen}, and with Glu148 or a chloride ion at S_{ext}. The actual coordinates of the ions at the S^*_{ext} sites were calculated by least-squares superposition of the backbone of the intracellular GSG motif in association with a chloride ion. As shown in Table 3, these calculations yielded electrostatic binding energies for S^*_{ext} that are comparable to those for S_{int} in both EcClC-WT and EcClC-E148Q, while the binding energies of S_{cen} and S_{ext} are affected only marginally (cf. Table 1). Thus, according to these results, binding of chloride to S^*_{ext} in EcClC would be energetically favored, and concurrent with chloride ions bound to all other binding sites in the conduction pore.

In addition to Arg340 and the backbone amide dipoles of Glu315 and Gly316, the stabilization of chloride at the S^*_{ext} site relies on other favorable interactions, mainly with the conserved Arg147 and, to a lesser degree, with other positively charged side-chains in helix B. In EcClC-WT the side-chain of Asn318 is involved, although it has opposite effects in each protein subunit due to alternate χ_2 rotamers in the X-ray structure, which at the current resolution are indistinguishable. The results given above correspond to the conformer in subunit B, where the NH_2 group in the side-chain participates in the coordination of the ion. In summary, although higher-resolution crystallographic data would be required to fully elucidate whether chloride ions can bind to S^*_{ext}, our calculations suggest that this is a low-affinity site similar to S_{int} in EcClC, as well as in other ClC channels, such as ClC-0, ClC-1 and ClC-2, where a

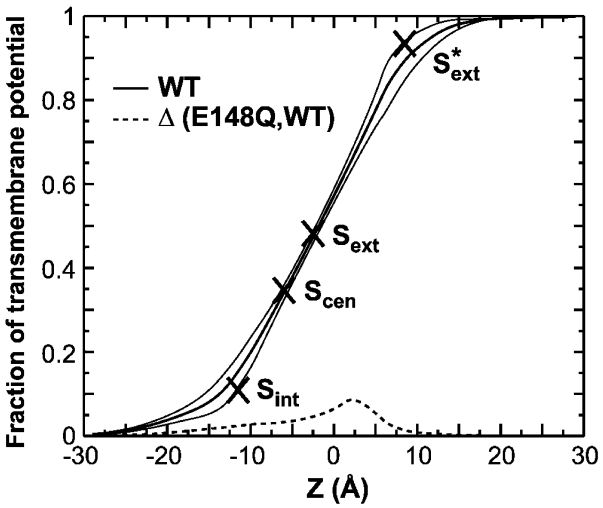

Figure 4. Fraction of the transmembrane potential along the conduction pore of EcClC, relative to the extracellular solution, calculated with the modified PB-V equation[67] with an ionic strength of 150 mM. The continuous black line corresponds to averages over a 10 Å × 10 Å area perpendicular to the *Z* axis; the standard deviations are indicated with continuous grey lines. The position and potential of each of the binding sites are marked on the plot. The broken black line is the difference between the curves for EcClC-E148Q and EcClC-WT; the peak corresponds to the side-chain of Gln148.

similar structural motif is likely to exist (Figure 3C).

Interestingly, the existence of a binding site on the extracellular entrance of the conduction pore has been proposed by Chen & Miller in a model of the mechanism of voltage-dependent fast gating in the ClC-0 channel.[31] On the basis of an analysis of the variation of channel opening rates as a function of transmembrane voltage and extracellular chloride ion concentration, they concluded that external chloride binds to the extracellular side of the channel in an essentially voltage-independent manner; subsequently, opening of the channel results from the voltage-dependent inward movement of chloride along the conduction pore. As shown in Figure 4, calculations of the electrostatic potential along the pore region of EcClC due to an external transmembrane voltage are consistent with this model. Specifically, the voltage at the proposed S^*_{ext} binding site is just 6.6% lower than that of the extracellular solution, whereas the voltage drop at the S_{ext} site is 52%. According to the model described by Chen & Miller, binding to a site in ClC-0 equivalent to S^*_{ext} in EcClC could be the concentration-dependent event, while chloride movement towards the S_{ext} binding site would be the voltage-driven step, which may then result in channel opening *via* competitive binding for this site with the side-chain of Glu148. Nevertheless, these inferences should be considered with caution as no voltage-dependent gating has been observed in EcClC itself.[9]

The helix macrodipole effect in EcClC: a comparison with K^+ channels and the PBP/SBP binding proteins

Favorable interactions with the backbone of the protein are crucial for the stabilization of ions within the EcClC channel. In this section, we analyze these interactions in more detail, specifically focussing on the importance of long-ranged contributions associated with protein α-helices. Although it is clear that an electric dipole arises from the charge distribution in the backbone of an α-helix, it is not self-evident that interactions with this macrodipole will play a crucial role when compared to the effect of side-chains or other charges more proximal to the ion-binding site. As a counter-example, let us consider the soluble anion-binding proteins PBP and SBP, which sequester and solubilize phosphate and sulfate ions, respectively (primarily HPO_4^{2-} in PBP, and SO_4^{2-} in SBP). As in ClC channels, binding of anions within PBP and SPB is largely due to favorable electrostatic interactions with the backbone of the protein, which amount to −69.6 kcal/mol in PBP (out of −92.9 kcal/mol from the whole protein) and −64.9 kcal/mol in SBP (out of −67.4 kcal/mol). These interactions counterbalance the energy cost of desolvation of the anions upon transfer into the protein (around + 53 kcal/mol), so that binding is energetically favorable, with $\Delta G_B = -40.0$ kcal/mol for phosphate and $\Delta G_B = -12.9$ kcal/mol for sulfate. (These values correspond to $\epsilon_p = 4$ and $\epsilon_w = 80$.)

In both cases, the anion-binding site is located between the N termini of three α-helices; namely, helix 2 (residues 38–46 and 45–53 in PBP and SBP, respectively), helix 6 (140–152/131–147) and helix 8 (176–186/173–181). Helix 1 (residues 13–27/14–32), and helix 9 (196–202/193–201) have a similar orientation, although these helices are further away and the anions are significantly off-axis (Figure 1B). Although this arrangement suggests that helix macrodipoles might play an important role in the energetics of ion-binding, our calculations demonstrate that this is not the case. Firstly, interactions with those residues comprising the helices contribute about 50%–60% of the total anion-backbone interaction energy (−35.8 kcal/mol in PBP and −41.4 kcal/mol in SBP); second, and more importantly, around 80% of this contribution (−29.2 kcal/mol and −33.0 kcal/mol, respectively) originates from the first turn of the helices (i.e. four N-terminal residues), while the first two turns (i.e. seven N-terminal residues) amount to about 95% (−33.7 kcal/mol and −38.4 kcal/mol, respectively) (Figure 5). Thus, long-range electrostatic effects associated with helix macrodipoles contribute to the energetics of binding in PBP and SBP only marginally. In spite of methodological differences, our conclusions agree with previous theoretical work on the SBP protein,[19] and with the observations reported by Quiocho and co-workers.[32,33]

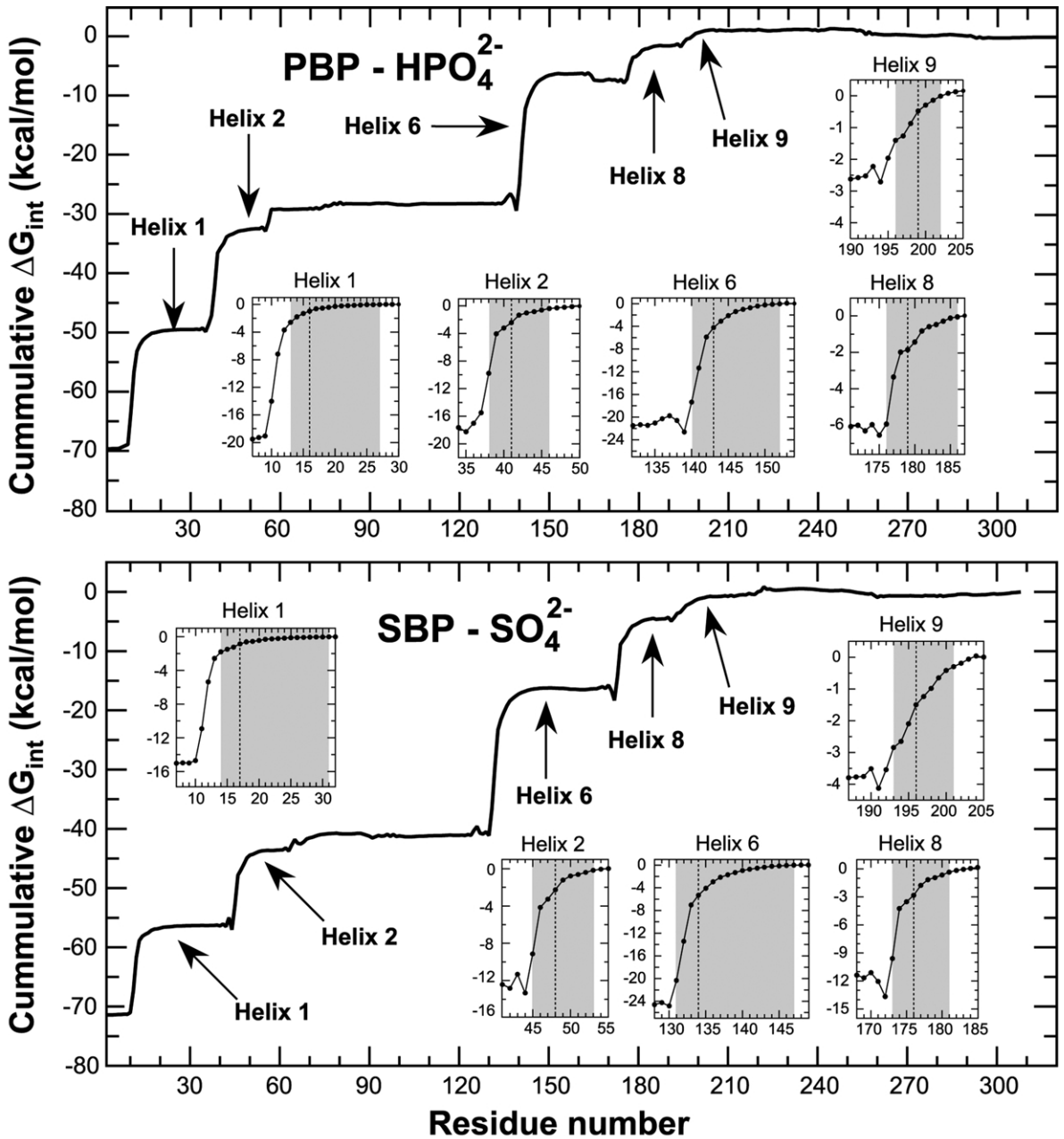

Figure 5. Cumulative anion-backbone electrostatic interaction energy in the SBP and PBP binding proteins (*versus* residue number, summed from the C terminus to the N terminus). The regions comprising α-helices that point towards the ion-binding sites are indicated. The insets show analogous data for each of these regions, with the length of the corresponding helix marked in grey; the dotted line separates the first N-terminal turn from the rest of the helix. These data, as well as those in subsequent Figures, correspond to $\epsilon_p = 4$.

Although the above observations are likely to be common for most ion-binding proteins, exceptions can occur where long-range effects might have a more prominent role. This seems to be the case in the multi-ion transmembrane channel KcsA (Figure 1C), specifically with regard to the binding of a potassium ion in the water-filled cavity region.[16] As in PBP/SBP, the presence of a monovalent cation in the KcsA cavity is primarily due to ion-backbone interactions, which add up to -11.3 kcal/mol out of -15.5 kcal/mol from the whole protein (with $\epsilon_p = 4$). More specifically, favorable interactions originate from the protein region comprising the pore helix (residues 62–74), a re-entrant helix that penetrates the membrane only partially and is oriented with its C terminus pointing towards the center of the cavity. The C-terminal halves of transmembrane helices M1 (residues 24–51) and M2 (residues 86–121) also interact favorably with the cavity ion, but their respective N-terminal halves counteract with slightly larger (by about 3 kcal/mol) unfavorable interactions (Figure 6).

The stabilization of the cavity ion in KcsA differs

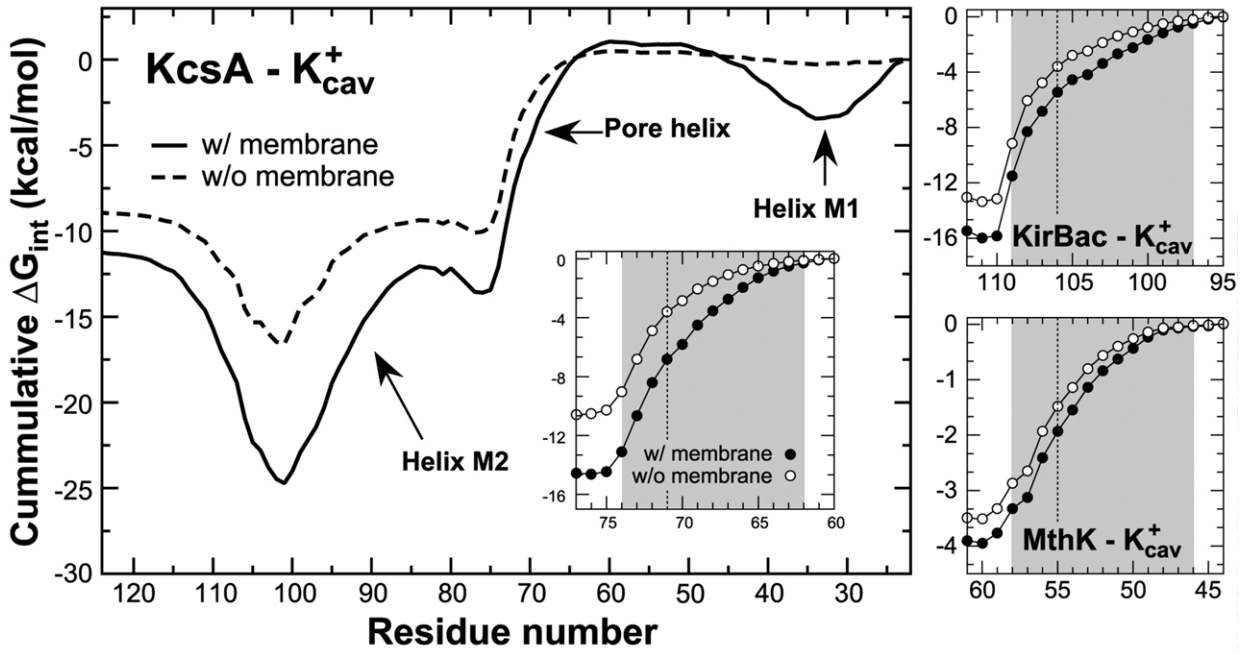

Figure 6. Left, cumulative electrostatic interaction energy between the cavity ion in the potassium channel KcsA and the protein's backbone (*versus* residue number, summed from the N terminus to the C terminus), with and without a low-dielectric membrane model ($\epsilon_m = 2$). The inset shows analogous data for the region comprising the pore helix, with the helix length marked in grey; the dotted line separates the first C-terminal turn from the rest of the helix. The corresponding binding energy is -1.3 kcal/mol, using $\epsilon_m = 4$. Right, results from an analogous analysis of the potassium channels MthK and KirBac.

from that of phosphate and sulfate in the PBP and SBP proteins, firstly in that the C terminus of the pore helices is approximately 8 Å away from the cation-binding site rather than in direct contact with water occupying the space between them. In spite of this, the contribution to the ion-backbone interaction energy from the pore helices adds up to -12.9 kcal/mol, which is larger than the total backbone interaction energy by 15%. In addition, the actual ion–helix interactions differ in the distribution of per-residue contribution along the helices. Specifically, the first C-terminal turn now accounts for 56% of the pore helix contribution (-7.2 kcal/mol), and the first two turns amount to 80% (-10.2 kcal/mol).

A similar case can be made for the homologous potassium channels KirBac[34] and MthK,[35] although given the moderate resolution of the current X-ray structures, it remains unclear whether an ion can be stabilized in their cavities. For KirBac (Figure 6), the effect of the pore helices is quantitatively and qualitatively analogous to that in KcsA, in spite of the fact that their axis does not point exactly towards the center of the cavity. This subtle difference in the orientation of the helices appears to have only a marginal effect from an energetic point of view, probably because the distances between the cation and the partial charges in the helices are approximately the same as in KcsA. In the case of MthK, the geometry of the cavity is changed dramatically by the wide opening of the helix bundle on the intracellular side of the protein. This results in a greater exposure of the cavity ion to the highly dielectric solvent, which attenuates significantly the favorable electrostatic field from the pore helices. Nonetheless, the qualitative long-range effect of the pore helices is clearly comparable to that in KcsA and KirBac (Figure 6).

The different relative importance of helical macrodipoles for the energetic stabilization of ions in K-channels and the SBP/PBP proteins can be understood on the basis of the distinct architecture and environment of these proteins. Firstly, any distant ion-backbone interactions with the cavity ion in K-channels, which as it happens come mainly from the pore helices, are prominent simply because more proximal interactions do not exist. In contrast, the helices' N termini and other unstructured backbone fragments come into direct contact with the anions in SBP and PBP, overriding other interactions (such as with helices 1 and 9, which actually resemble those between the cavity ion in KcsA and one of the pore helices). More quantitatively, while in SBP and PBP interactions beyond 12 Å from the binding site result in only about 10% the total backbone interaction energy, in KcsA the analogous contribution is as much as 70% (Figure 7A). In addition to this purely structural argument, the effect of the low-dielectric membrane environment, which allows interactions to have a longer range, should be considered.[16] Indeed, in the absence of the

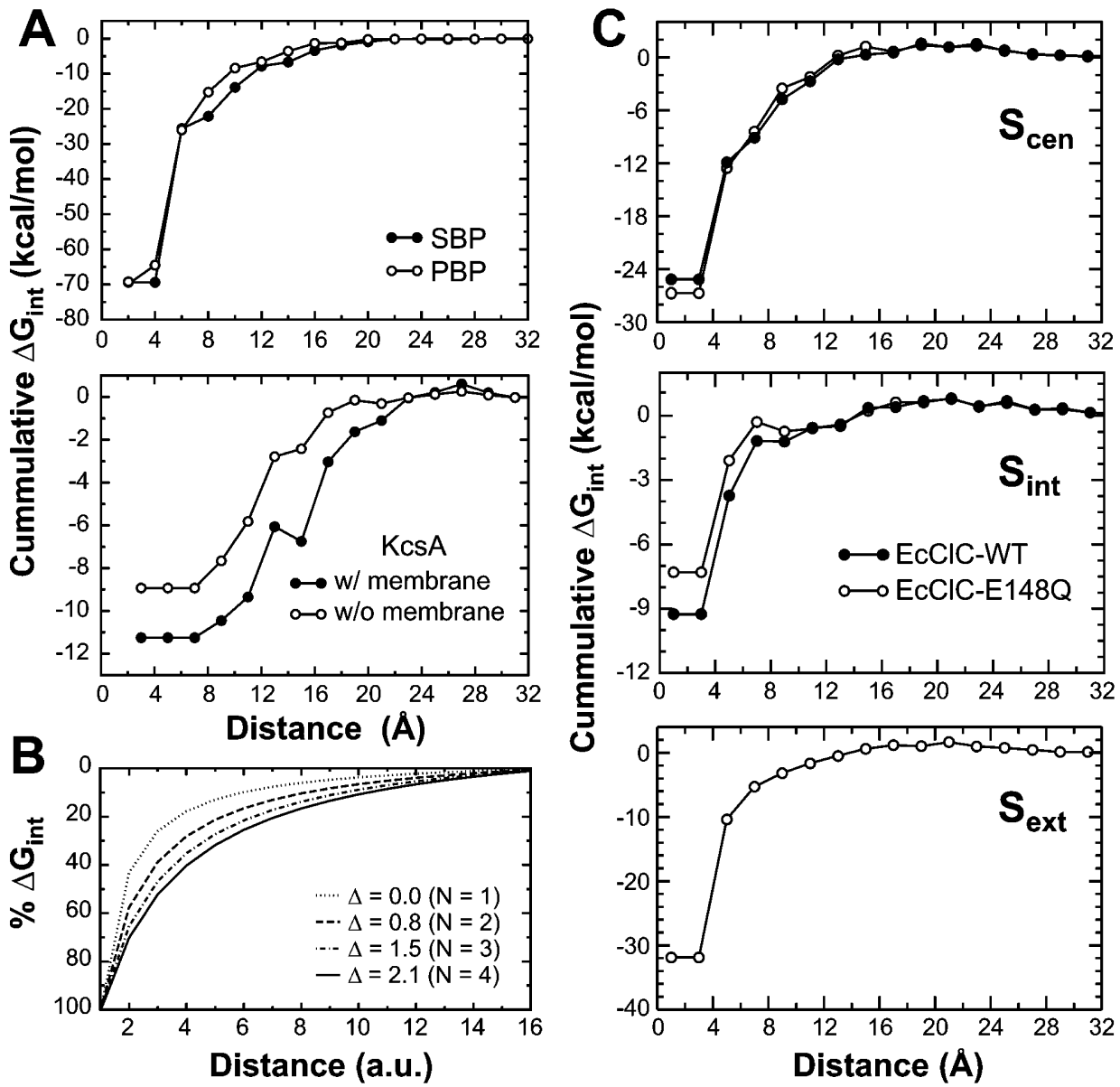

Figure 7. A, Cumulative ion-backbone electrostatic interaction energy, *versus* distance to the ion, in the anion-binding protein SBP and PBP, and in the potassium channel KcsA (with and without the membrane model). B, Cumulative interaction energy *versus* distance, for a simplified ion-helix system; the ion is represented by a point charge, with $q = 1$, and the helix by a set of aligned, equally spaced microdipoles, with $q = \pm 0.5$. The distance between the ion and the helix is given by Δ; a factor N is used to normalize each curve so as to produce the same total interaction (e.g. $N = 4$ would represent the effect of four helices). C, Analogous data to A for each of the chloride ions bound to EcClC-WT and EcClC-E148Q (the corresponding profiles for calculations without a membrane model differ by less than 0.25 kcal/mol).

membrane model for KcsA, the contribution of the backbone interactions beyond 12 Å drops to less than 50% of the total.

In these systems, the ion–helix backbone interaction differs also in the way in which the total interaction energy is distributed along the helix. To rationalize this, let us consider a simplified system in which a helix is represented by an aligned series of equally spaced, alternate positive and negative charges, which interact with a single negative charge located at the positive end, with $\Delta G_{int} \sim 1/d$. As illustrated in Figure 7B, as the distance between the helix model and the single charge increases, the total interaction energy becomes distributed more evenly along the entire helix. Similarly, the longer distance separating the cavity ion in KcsA from the C terminus of the pore helices, compared to that between the sulfate/phosphate anion and the N terminus of helix 2, helix 6 and helix 8 in SBP/PBP, results in a non-negligible contribution from, say, the furthest half of the helices.

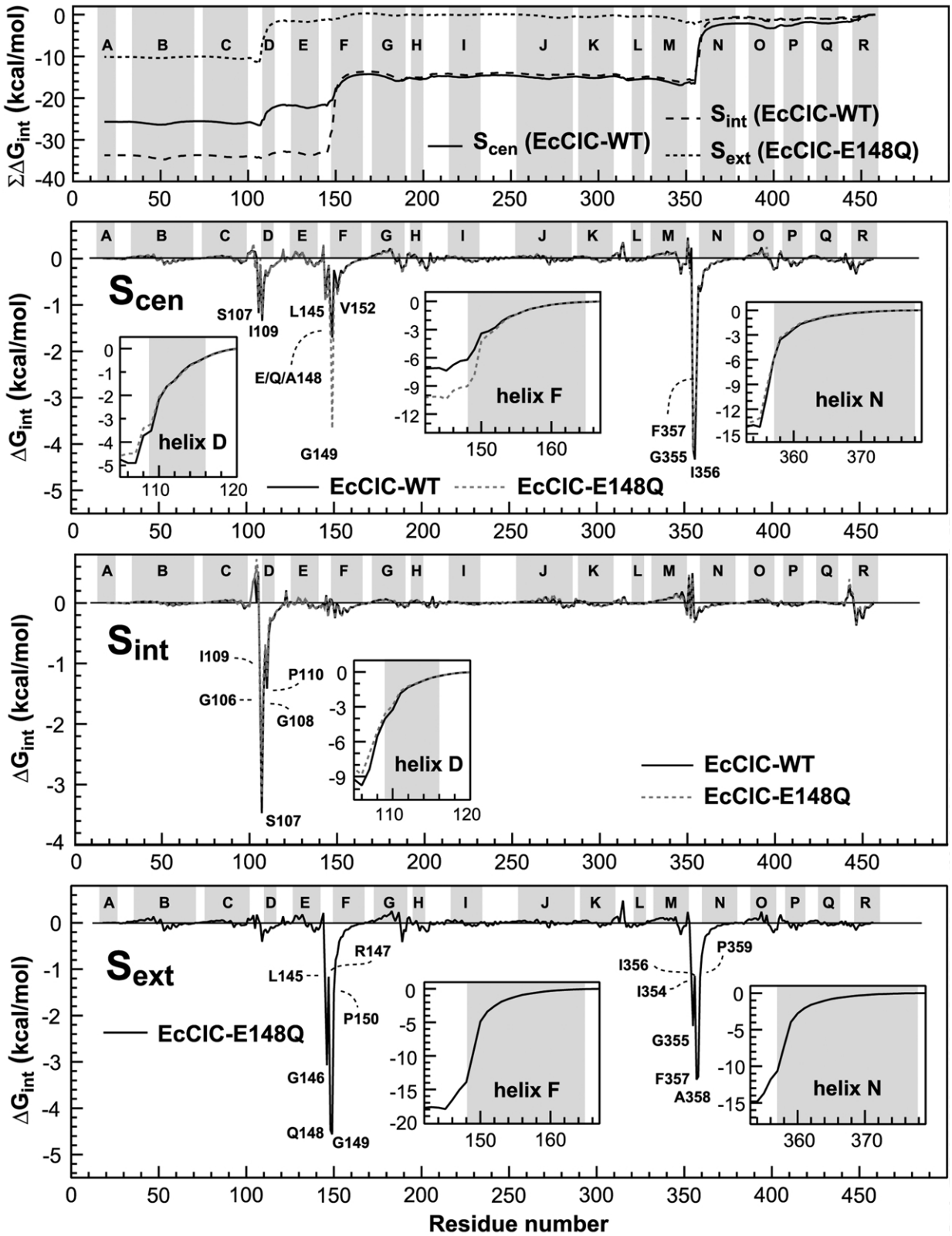

Figure 8. Ion-backbone electrostatic interactions in the EcClC chloride channel homologue *versus* residue number. The top panel shows the cumulative interaction energy, calculated from the C terminus to the N terminus. The insets show cumulative energies for the regions comprising the α-helices that point towards the chloride-binding sites; the length of each helix is marked in grey.

The stabilization of ions within the transmembrane domain of the EcClC antiporter resembles that in the PBP and SBP binding proteins. Backbone interactions in close proximity to the ion-binding sites are dominant relative to longer-range effects, although to a different extent for each of the ions associated with the protein (Figures 7C and 8). For the ion found at the S_{ext} site in the EcClC-E148Q mutant, interactions with the backbone of helix F (residues 148–165) and helix N (residues 357–378) represent around 70% of the total backbone interaction energy (−22.7 kcal/mol; see Table 1 for total energies), with their first turn contributing about 80% of the energy, and 91% by the first two turns. If, in addition to the first turn of each helix, we consider the three residues preceding their N terminus, all of which are within 8 Å from the ion, the interactions amount to up to about 90% of the total backbone interaction energy (−28.0 kcal/mol). Therefore, as in PBP and SBP, it is clear that the significance of the helical macrodipoles in the energetics of this site is only marginal.

The chloride ion bound to the S_{cen} site, present in EcClC-WT as well as in the E148Q mutant, is only about 4 Å away from S_{ext}, and therefore similar conclusions can be drawn with regard to the significance of helix macrodipoles. However, it is worth mentioning that the fact that the ion in S_{cen} is slightly further from the N terminus of helices F and N (the ion in S_{ext} is literally caged between the helices) has some effect on the relative importance of the furthest residues in these helices, consistent with the arguments given above. Specifically, the contribution of the first turn of helices F and N is now 65%–70% of the total interaction energy from the backbone of these helices (which is −11.9 kcal/mol in EcClC-WT and −14.61 kcal/mol in EcClC-E148Q), while the first two turns contribute 80%–85%. In addition, the chloride ion in S_{cen} interacts significantly with helix D (residues 109–116), whose N terminus is about 5 Å away from the binding site (the interaction of this helix with the ion in S_{ext} is less than 1.5 kcal/mol). Nevertheless, around 85% of the total ion-backbone interaction energy for this ion can be accounted for if we consider the first turn of helices D, F and N, as well as the three preceding residues in the sequence, i.e. within 10 Å of the binding site (−21.1 kcal/mol in EcClC-WT and −22.4 kcal/mol in EcClC-E148Q).

Finally, the intracellular ion bound to S_{int} interacts only weakly with helices F and N (less than −1.5 kcal/mol each), and somewhat more strongly with helix D, e.g. −3.7 kcal/mol in EcClC-WT. Yet again, the first turn of helix D (about 80% of this helix's contribution), plus three residues preceding its N terminus, account for more than 90% of the total ion-backbone interaction energy, e.g. −8.6 kcal/mol in EcClC-WT. In summary, the current analysis of the ion-backbone interactions in the EcClC protein does not support the notion that helix macrodipoles have a role in ClC channels comparable to that of the pore helices in K-channels.

Conclusions

We have characterized the electrostatic interactions governing the association of chloride ions with the EcClC chloride channel homologue from *E. coli*. This investigation is expected to provide clues for future analyses of the actual permeation and gating processes, as well as for the interpretation or design of experimental work. Under all conditions considered, our results support the view that the conduction pore of EcClC can hold chloride ions simultaneously at positions S_{int} and S_{cen}, while S_{ext} is occupied by the side-chain of Glu148. Furthermore, we propose that a third chloride ion can occupy an additional binding site located on the extracellular side of the protein, or S^{*}_{ext}. Consistently, our calculations for the EcClC-E148Q mutant show that binding of a fourth chloride ion to S_{ext} is also electrostatically favorable.

Ion binding is only possible due to stabilizing electrostatic interactions with the channel that counter-balance the energetic cost for desolvation, as well as the ion–ion repulsion. Our analysis of the EcClC structures reveals that interactions with both the backbone and side-chains of the protein are crucial in achieving binding, although the relative contributions depend on the ion considered (about 60/40 for S_{cen}, 40/60 for S_{int}, and 80/20 for S_{ext}). Among the side-chain contributions, it is worth noting the strong stabilizing interaction of all three ions with the strictly conserved Lys131, in addition to previously reported Ser107 and Tyr445 (which interact with the ion in S_{cen} only). Given the association of this side-chain with the conserved GSG motif that separates the S_{int} and S_{cen} sites (residues 106–108), Lys131 is likely to be crucial by virtue of its direct electrostatic interaction with the ions, and in regulating the structural stability of the binding sites; consequently, we speculate that Lys131 may be involved in the exchange of ions between S_{int} and S_{cen}. We propose that Arg340 and the adjacent GGG segment (residues 314–316) may constitute a low-affinity binding site at the extracellular side of the conduction pore, in addition to those already identified elsewhere in the protein.

Finally, our analysis shows that the backbone contribution to the electrostatic binding free energy originates predominantly from interactions within 10 Å from the binding sites; namely, with specific amide dipoles in segments lacking secondary structure, as well as with the N-terminal turn of helices D, F and N. The effect of the helix macrodipoles *per se* appears to be only marginal, in analogy to soluble anion-binding proteins such as SBP and PBP, and in contrast to the case of the cavity ion in KcsA.

Methods

Calculation of electrostatic energies

Calculations of the electrostatic energy throughout the current study were conducted under a semi-macroscopic framework in which an all-atom representation is used for the protein, and a continuum dielectric representation is adopted for the surrounding medium, i.e. solvent and membrane.[36,37] In this context, the electrostatic potential $\Phi(\boldsymbol{r})$ that originates from a given spatial configuration of the protein is given by Poisson's equation:

$$\nabla\cdot(\epsilon(\boldsymbol{r})\nabla\Phi(\boldsymbol{r})) = -4\pi\sum_{i}^{N} q_i\delta(\boldsymbol{r}-\boldsymbol{r}_i) \qquad (1)$$

where q_i and $\boldsymbol{r}_i$ are the charge and location of each of the N atoms in the protein, and $\epsilon(\boldsymbol{r})$ denotes the position-dependent dielectric constant across the entire system. (Note that mobile ionic charges in solution are not included.) The electrostatic energy of the system is then defined as:

$$\Delta G = \frac{1}{2}\sum_{i}^{N} q_i\Phi(\boldsymbol{r}_i) \qquad (2)$$

which can be understood as the energy required to assemble the configuration of N protein charges in the dielectric environment described by $\epsilon(\boldsymbol{r})$. In the context of biomolecular electrostatics, one is normally interested in energy differences between two states of the system, e.g. in vacuum or in solution, ligand-bound or ligand-free, etc.

In order to solve Poisson's equation, we used the finite-difference Poisson–Boltzmann (PB) equation solver PBEQ,[38,39] as implemented in CHARMM C30A1.[40] The code was modified so that a set of "dummy" atoms can be used to define a region of arbitrary shape and dielectric constant. In the finite-difference scheme, the system's charge and dielectric distributions are discretized by mapping them onto a three-dimensional grid, which allows Poisson's equation to be solved numerically. As a result, one obtains the value of the electrostatic potential for each grid-point, which is then used to calculate the electrostatic energy according to equation (2).

The dielectric distribution $\epsilon(\boldsymbol{r})$ included three different regions; namely, protein–ion complex, membrane and solvent. A dielectric constant of 2, 4, 6 or 8 was assigned to the interior of the region corresponding to the protein–ion complex, which was delimited by its solvent-excluded surface. This surface was constructed using a water probe radius of 1.4 Å and the optimized atomic radii calculated by Nina *et al.*[39] (the radius of OE2 and of OD2 in the protonated forms of Glu and Asp, respectively, was set to 1.64 Å). The membrane region was delimited by a continuous slab of overlapping, uncharged dummy atoms in which the protein–ion complex was embedded (see the next section). Within this region the dielectric constant was set to 2, so as to mimic the low-dielectric hydrocarbon core of the membrane. For the remainder of the system, or solvent, a dielectric constant of 80 was used. Atomic charges were obtained from the CHARMM27 forcefield.[41]

Grid-focusing and translational averaging were implemented in all calculations in order to minimize discretization effects without increasing computational time significantly. The dimensions of the initial grid were 140 Å × 96 Å × 100 Å (ca 200% those of the protein), with a grid-point spacing of 1.0 Å. Boundary conditions were interpolated from this calculation, and used for a second grid, of dimensions 110 Å × 66 Å × 70 Å (ca 110%), with a grid-point spacing equal to 0.5 Å. The focusing procedure was repeated for eight different locations of the protein-slab system relative to the center of the 3D-grids (shifted 0–0.25 Å in each direction), and the resulting values were averaged; the corresponding standard deviations are typically $\leq$0.5 kcal/mol.

Set-up of EcClC channel-membrane system

Atomic coordinates of the EcClC channel were obtained from the Protein Data Bank[42] entries 1OTS (EcClC-WT), 1OTU (EcClC-E148Q) and 1KPL (StClC-WT).[7,8] Missing hydrogen atoms were added using HBUILD within the CHARMM package; the position of hydroxyl hydrogen atoms in Tyr, Ser and Thr side-chains was subsequently refined by energy minimisation.

Each of the ClC structures analyzed was embedded in a face-centered cubic lattice of uncharged atoms of 2.5 Å radius, which was used to delimit a homogeneous low-dielectric region mimicking the membrane hydrophobic environment. In the plane of the membrane, the spacing between the centers of these dummy atoms was 3 Å, and the total slab dimensions of the slab were around 145 Å × 100 Å. Along the membrane normal, which is considered to be parallel with the 2-fold symmetry axis of the protein dimer, the spacing between layers was 0.75 Å, and several slabs of different width were considered initially. (Note that we approximate the width to be the distance between atom centers in the top and bottom layers of the dummy-atom lattice, plus twice the value of their radii.)

The optimal position of ClC channels within such membrane models is not obvious, due to the pronounced irregularity and variable width of the hydrophobic interface. In order to determine the optimal location of the protein–ion complex within each of the slabs, we calculated the free energy of transfer from solution, ΔG_{tot}, as a function of the position of the protein's geometric center relative to the center of the slab, Z, so that the minimum could be identified. (The crystal structure of StClC was used in these calculations.) Assuming a fixed conformation for the protein–ion complex, this free energy of transfer can be written as:

$$\Delta G_{\mathrm{tot}}(Z) = \Delta G_{\mathrm{np}}(Z) + \Delta G_{\mathrm{elec}}(Z) \qquad (3)$$

where ΔG_{np} is the change in the so-called free energy of cavity formation, or non-polar term, and ΔG_{elec} is the change in free energy of charging the protein–ion complex, upon change in its dielectric environment.[43–46] In the current non-atomistic, continuum description of a protein's environment, the latter contribution corresponds simply to the change in the electrostatic energy of the system, given by:

$$\Delta G_{\mathrm{elec}}(Z) = \frac{1}{2}\sum_{i}^{N} q_i(\Phi_i(Z) - \Phi_i(\infty)) \qquad (4)$$

where q_i denotes the charge of atom i, and $\Phi_i(Z)$ and $\Phi_i(\infty)$ are the electrostatic potentials, evaluated at the position of atom i, when the protein is at a finite distance Z from the membrane center or in solution, respectively. Concerning the non-polar term, we approximate ΔG_{np} to be proportional to the reduction of the solvent-exposed surface area of the protein upon transfer into

the membrane:

$$\Delta A_w(Z) = A_w(Z) - A$$

on the basis of the empirical observation that the free energies of transfer of hydrophobic solutes are linearly related to their surface area.[47–49] That is:

$$\Delta G_{np}(Z) \approx \gamma_{v-w} A_w(Z) + \gamma_{v-m}(A - A_w(Z)) - \gamma_{vw} A$$
$$= \gamma_{eff} \Delta A_w(Z) \quad (5)$$

where γ_{v-w}, γ_{v-m} and γ_{eff} denote the effective "microscopic surface tensions" of the vacuum–water, vacuum–alkane and water–alkane interfaces, respectively.

Across a range of values of γ_{eff} between 0.02 kcal/mol Å^2 and 0.05 kcal/mol Å^2, and for slab-widths between 15 Å and 35 Å, the minimum transfer free energy, ΔG_{min}, was consistently found at $Z = -2.5(\pm 1)$ Å in ~70% of the cases, and at $Z = -4.5(\pm 1)$ Å otherwise. Nonetheless, the actual width for which ΔG_{min} was the lowest was observed to be sensitive to the value of γ_{eff} chosen. This is expected, since this parameter determines the relative weight of the non-polar term ($\Delta G_{np} \leq 0$) relative to the electrostatic contribution ($\Delta G_{elec} \geqq 0$). For example, for the minimum and maximum values of γ_{eff} in the range analyzed, the lowest value of ΔG_{min} was found for slab-widths of 15 Å and 27 Å, respectively. Given the uncertain free energy cost associated with varying the membrane thickness, we finally opted for a 23 Å wide slab at $Z \sim -2.5$ Å, for which ΔG_{min} was the lowest when $\gamma_{eff} = 0.03$ kcal/mol Å^2, because these values are comparable to those estimated from experimental and theoretical studies of lipid bilayer structure and dielectric properties,[50–52] as well as of hydrocarbon-to-water transfer free energies of non-polar compounds.[47–49,53]

In addition to the above "flat" membrane models, a curved slab was constructed so as to examine the influence of a variable width in the interface between the protein and the membrane. This slab was built *ad hoc*, so as to minimize the solvent exposure of hydrophobic regions on the protein–membrane interface without burying charged side-chains. Indeed, the calculated transfer free energies for this membrane model were lower than ΔG_{min} for any of the flat slabs within the range of values of γ_{eff} considered, ignoring any energetic cost associated with membrane bending. Both the 23 Å-wide flat slab and this "optimized" curved membrane were used throughout the analysis of EcClC, in order to determine whether the shape of the membrane model was critically relevant to the electrostatic energies calculated herein.

Protonation states in EcClC and pK_A calculations

All electrostatic calculations on the EcClC channels used default ionization states for Arg, Lys and Tyr side-chains; for Glu, Asp and His side-chains, pK_A calculations were carried out using the Poisson–Boltzmann framework in order to determine whether the non-default, protonated state was favorable within the protein–ion-membrane environment (i.e. whether pK_A > pH). Both possible protonation modes were analyzed in each of the Glu, Asp and His side-chains, and equivalent side-chains in each protein monomer were studied independently. It is worth noting that under this approach we can provide only a semi-quantitative assessment of the energetics of protonation, primarily because the protein's flexibility is not included in the model,[54] but also due to the limitations of a continuum description.[55] Nonetheless, it is expected that the calculated pK_A shifts will be in qualitative accordance with more sophisticated methodologies such as those based on the calculation of free energies by means of molecular dynamics simulations.[14,56]

The pK_A of a given titratable protein side-chain can be determined using the expression:

$$pK_A = pK_A^{sol} + \frac{1}{\ln 10 k_B T}(\Delta G^u_{trans} - \Delta G^p_{trans}) \quad (6)$$

where pK_A^{sol} is the pK_A of the titratable residue in solution (4.3 for Glu, 3.9 for Asp, and 6.0 for His), k_B is the Boltzmann constant, T is the temperature, and ΔG^u_{trans} and ΔG^p_{trans} are the free energies of transferring the titratable residue from solution to the protein–ion complex in its unprotonated and protonated forms, respectively. The calculation of these quantities is, however, non-trivial, mainly because it requires accounting for the existence of multiple titratable sites in the protein. To circumvent this problem, given that our purpose is simply to determine the most likely protonation state at a given pH, we have aimed to calculate the upper bound for the pK_A of Glu, Asp or His side-chains, within the reasonable assumption that all Arg, Lys and Tyr side-chains in the protein will be always protonated under physiological conditions. Under this approximation, pK_A^{max} for a given Glu, Asp or His side-chain is obtained in the hypothetical situation where all other Glu, Asp, and His side-chains exist in their default, unprotonated states, thus maximizing ΔG^u_{trans} and minimizing ΔG^p_{trans}. If, for instance, $pK_A^{max} < 7$ even in this extreme configuration of the protein, one can be certain, within the limits of the methodology and the approximations taken, that the side-chain's most likely state at pH 7 will be unprotonated.

The calculation of ΔG^u_{trans} and ΔG^p_{trans} is analogous to that in the previous section, except that only the electrostatic contribution in equation (3) is required, since $\Delta G^u_{np} \approx \Delta G^p_{np}$ (again assuming that all conformations remain unchanged upon transfer). Therefore, we calculate:

$$\max(\Delta G^u_{trans}) = \frac{1}{2}\left(\sum_i^{N_r} Q_i^u + \sum_i^{N-N_r} q_i^u\right)\Phi_i^{uu} - \frac{1}{2}\sum_i^{N_r} Q_i^u \Phi_i^u \quad (7)$$

$$\min(\Delta G^p_{trans}) = \frac{1}{2}\left(\sum_i^{N_r} Q_i^p + \sum_i^{N-N_r} q_i^u\right)\Phi_i^{pu} - \frac{1}{2}\sum_i^{N_r} Q_i^p \Phi_i^p \quad (8)$$

where Q_i^x denotes the charge of each of the N_r atoms in the residue analyzed, either in its unprotonated (u) or protonated (p) form; q_i^u represents the atomic charge of the rest of the atoms in the protein–ion complex, with unprotonated Glu, Asp, and His side-chains; and Φ_i^{xy} and Φ_i^x are the corresponding electrostatic potentials originating from the protein–ion–membrane system and from the isolated residue in solution, respectively.

Following this approach, large pK_A shifts (i.e. larger values than pK_A^{sol}) were found for Glu113 and Glu203 in both monomers of EcClC-WT and EcClC-E148Q, using either 2 or 4 as the protein dielectric constant (ϵ_p). These

shifts were especially marked for the protonation of the carboxylate $O^{\epsilon 2}$ atom in both Glu113 and Glu203. Smaller shifts in the pK_A were found for $O^{\delta 2}$ as well as in Glu103, Glu111, Glu117 and Asp118, although in the latter cases pK_A shifts were found only when $\epsilon_p = 2$, which leads to larger shifts in general, and mainly in EcClC-WT.

These results can be rationalized easily by inspecting the EcClC atomic structure: Glu203 is found in a region of reduced solvent-accessibility, forming an ion-pair with Arg28, *via* the carboxylate $O^{\epsilon 1}$ atom (nomenclature from PDB entry 1OTS). Towards the intracellular side of the protein, the nearest neighbor of Glu203 is Glu113, with their respective $O^{\epsilon 2}$ atoms only 2.5 Å apart. This seemingly unfavorable charge configuration is not balanced by the protein environment. Indeed, Glu113 is surrounded by hydrophobic residues, except for Glu203, and lacks any direct interaction with positively charged side-chains; furthermore, although some crystallographic water molecules can be seen in close contact with Glu203, none exists in the vicinity of Glu113. This hypothetical charge excess would also have some effect on the pK_A of Glu117 and Asp118, which are found along the same direction, approximately 8 Å and 13 Å away from Glu113. However, these side-chains are significantly more solvent-exposed than Glu113 or Asp203, and are engaged in ion-pairs; namely, *via* $O^{\epsilon 1}$ with Arg209 and Arg174, respectively. Therefore, it is expected that their energetic coupling with Glu203/Glu113 will be weak, as our results indicate. On the basis of these observations and our calculations, we conclude that Glu113 exists preferably in its protonated form at neutral pH or lower, and that a hydrogen bond exists between Glu113 and Glu203 side-chains *via* their respective $O^{\epsilon 2}$ carboxylate atoms. As was discussed previously, the multiple ion occupancy of the conduction pore of EcClC is critically dependent on the protonation state of Glu113, since this side-chain is only ~10 Å away from the chloride-binding sites.

Asp417 is located at the homodimer interface at the extracellular side of the protein, and interacts primarily with its counterpart in the adjacent monomer. Although these side-chains are solvent-exposed and neighbored by Lys side-chains, their close proximity appears to favor the protonated state at pH 7, especially for the $O^{\delta 2}$ atoms, which are only 4 Å apart. However, because the pK_A shifts are induced by the interaction with each other, it is unlikely that both side-chains will be protonated at a given time at neutral pH. Consequently, we have used the neutral state for Asp417 only in monomer A. The choice of subunit has essentially no effect on the current study.

Concerning His side-chains, no significant pK_A shift was detected that could be critical for the ion–protein association, and therefore neutral states were used in this study. It should be noted, though, that according to our calculations most of the His side-chains will be positively charged at low pH values, which in turn will favor chloride binding to the protein. Comparative analysis of the pK_A^{max} values obtained for each possible protonation mode (that is, protonation of N^{ϵ} with N^{δ} already protonated, or *vice versa*) was useful to identify the tautomerisation state of the side-chains. In this respect, His70, His234 and His383 were found to favor protonation at N^{δ}, whereas in His175, His281 and His284 the protonation of N^{ϵ} was preferred (except in EcClC-E148Q, where the rings of His281 and His284 are flipped over and thus favor the protonation of N^{δ}).

In order to test whether our conclusions were plausible, pK_A^{max} values were re-calculated after adjusting the protonation states of Glu113, Asp417A, His175, His281 and His284 according to the results described above. These calculations yielded no significant difference with respect to the initial calculations, except that now $pK_A^{max} \ll 7$ for Glu103, Glu111, Glu117, Asp118, Glu203 and Asp417B, thus supporting our choice of protonation states at physiological pH. Finally, we analyzed the protonation state of K131, which, as has been discussed, is crucial in the energetics of ion–protein association. In the presence of ions in all binding sites and for $\epsilon_p = 4$, its pK_A is larger than $pK_A^{sol} \approx 11$ by about four units; in the absence of ions in the channel, the pK_A of K131 shifts towards lower values ($pK_A \sim 8$), but would remain charged at neutral or acidic pH.

Electrostatic free energy of binding and interactions

The electrostatic contribution to the free energy of association of an ion i with a protein p can be expressed as: $\Delta G_B = \Delta G^{ip} - (\Delta G^{p} + \Delta G^{i})$, where ΔG^{ip}, ΔG^{p} and ΔG^{i} represent the electrostatic energies of the protein–ion complex, the isolated protein and the isolated ion, respectively (with the protein being embedded in a membrane). The above expression for ΔG_B can be re-written as:

$$\Delta G_B = \Delta G^{i}_{rf} + \Delta G^{p}_{rf} + \Delta G_{int} \tag{9}$$

where ΔG^{i}_{rf} and ΔG^{p}_{rf} now denote the changes in the reaction field energy of the ion and of the protein, respectively (the latter is usually negligible), and ΔG_{int} is the ion–protein interaction energy upon association. Note that in the current case, p denotes the protein channel and the membrane, as well as the other bound ions. Thus:

$$\Delta G^{i}_{rf} = \frac{1}{2} q^{i} \{\Phi(q^{i}, \epsilon^{ip}) - \Phi(q^{i}, \epsilon^{i})\} \tag{10}$$

$$\Delta G^{p}_{rf} = \frac{1}{2} \sum_{n} q^{p}_{n} \{\Phi_n(q^{p}, \epsilon^{ip}) - \Phi_n(q^{p}, \epsilon^{i})\} \tag{11}$$

$$\Delta G_{int} = q^{i} \Phi(q^{p}, \epsilon^{ip}) \tag{12}$$

where the electrostatic potentials $\Phi(q^{x}, \epsilon^{y})$ originate either from the charge of ion $i(q^{i})$ or from those in the rest of the protein–ion complex (q^{p}_{n}), when the dielectric distribution is either that of ion i in solution (ϵ^{i}), or that of the membrane-embedded protein–ion complex, in association with ion (ϵ^{ip}) or in its absence (ϵ^{p}).

Because the ion is transferred from bulk solution to the interior of a low-dielectric region, $\Delta G_{rf} \geqq 0$, and the energetics of the association of ion and protein will depend critically on the extent to which ΔG_{int} counteracts the cost of desolvation. Therefore, it is of interest to identify which are the structural elements that contribute the greatest stabilizing effects. To this end, we take advantage of the fact that equation (12) can be re-written as a sum of interactions between individual atomic charges in the protein–ion complex and the electrostatic potential generated by ion i:

$$\Delta G_{int} = \sum_{n} q^{p}_{n} \Phi_n(q^{i}, \epsilon^{ip}) \tag{13}$$

This expression allows one to "dissect" the ion–protein interaction energy term according to any criterion, without having to recompute the electrostatic potential. For example, in order to analyze the electrostatic stabilization of ion i by a given helix macrodipole, the q^{p}_{n} charges

in equation (13) would represent only those in the helix backbone. As has been shown, by varying the number of amino acids used in such a calculation, it is possible to quantify the importance of long-range effects due to the macrodipole relative to the local interactions at the helix terminus.

Electrostatic calculations on KcsA, KirBac, MthK, SBP and PBP

The analysis of the electrostatic interactions in the potassium channels KcsA, KirBac and MthK, and in the proteins SBP and PBP, was analogous to that carried out for the bacterial ClC homologue. The structure of KcsA corresponds to the PDB entry 1K4C,[24] that of KirBac to 1P7B,[34] and that of MthK to 1LNQ.[35] Atomic coordinates for SBP and PBP were obtained from the PDB entries 1IXH[57] and 1SBP,[32] respectively. In all proteins, missing hydrogen atoms were added as described; other missing atoms (notably all side-chains in MthK) were added with CHARMM and SCWRL 2.95[58] when their coordinates could not be deduced by sequence similarity.

Calculations on the potassium channels included potassium ions both in the selectivity filter (positions S1 and S3) and in the cavity (as in KcsA), and a 25 Å wide membrane slab centered approximately at the level of the cavity region.[16] Calculations on SBP included a sulfate ion, and those of PBP included a phosphate ion, in accordance with Figure 4 of Vyas *et al*.[33] Atomic charges and radii for these anions were adapted from the CHARMM27 force field[41] and from Banavali *et al*.[59] The corresponding solvation free energies (with $\epsilon_p = 1$) are -266 kcal/mol for sulfate (cf. experimental value of -260 kcal/mol)[60] and -276 kcal/mol for phosphate (cf. theoretical estimate of $-245(\pm 15)$ kcal/mol).[61]

Acknowledgements

We are thankful to R. Duztler, C. Miller and T. Y. Chen for stimulating conversations; to the former two, respectively, also for sending us the atomic coordinates of EcClC before their public release, and for allowing us to read an unpublished manuscript on the EcClC chloride–proton antiport. We thank L. R. Forrest for helpful discussions throughout this project and for comments on the current manuscript. This work was supported by NIH grant R01-GM62342.

References

1. Jentsch, T. J., Stein, V., Weinreich, F. & Zdebik, A. A. (2002). Molecular structure and physiological function of chloride channels. *Physiol. Rev.* **82**, 503–568.
2. Maduke, M., Miller, C. & Mindell, J. A. (2000). A decade of CLC chloride channels: structure, mechanism, and many unsettled questions. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 411–438.
3. Wills, N. K. & Fong, P. (2001). ClC chloride channels in epithelia: recent progress and remaining puzzles. *News Physiol. Sci.* **16**, 161–166.
4. Iyer, R., Iverson, T. M., Accardi, A. & Miller, C. (2002). A biological role for prokaryotic ClC chloride channels. *Nature*, **419**, 715–718.
5. Mindell, J. A., Maduke, M., Miller, C. & Grigorieff, N. (2001). Projection structure of a ClC-type chloride channel at 6.5 Å resolution. *Nature*, **409**, 219–223.
6. Accardi, A. & Miller, C. (2004). Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature*, **427**, 803–807.
7. Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. & MacKinnon, R. X-r. (2002). ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature*, **415**, 287–294.
8. Dutzler, R., Campbell, E. B. & MacKinnon, R. (2003). Gating the selectivity filter in ClC chloride channels. *Science*, **300**, 108–112.
9. Accardi, A., Kolmakova-Partensky, L., Williams, C. & Miller, C. (2004). Ionic currents mediated by a prokaryotic homologue of ClC Cl^- channels. *J. Gen. Physiol.* **123**, 109–119.
10. Chen, M. F. & Chen, T. Y. (2001). Different fast-gate regulation by external Cl^- and H^+ of the muscle-type ClC chloride channels. *J. Gen. Physiol.* **118**, 23–32.
11. Waldegger, S. & Jentsch, T. J. (2000). Functional and structural analysis of ClC-K chloride channels involved in renal disease. *J. Biol. Chem.* **275**, 24527–24533.
12. Bernèche, S. & Roux, B. (2001). Energetics of ion conduction through the K^+ channel. *Nature*, **414**, 73–76.
13. Guidoni, L., Torre, V. & Carloni, P. (1999). Potassium and sodium binding in the outer mouth of the K^+ channel. *Biochemistry*, **38**, 8599–8604.
14. Luzhkov, V. & Åqvist, J. (2000). A computational study of ion binding and protonation states in the KcsA potassium channel. *Biochim. Biophys. Acta*, **1481**, 360–370.
15. Ranatunga, K., Shrivastava, I. H., Smith, G. R. & Sansom, M. S. P. (2001). Sidechain ionisation states in a potassium channel. *Biophys. J.* **80**, 1210–1219.
16. Roux, B. & MacKinnon, R. (1999). The cavity and pore helices in the KcsA K^+ channel: electrostatic stabilization of monovalent cations. *Science*, **285**, 100–102.
17. Roux, B., Bernèche, S. & Im, W. (2000). Ion channels, permeation and electrostatics: insight into the function of KcsA. *Biochemistry*, **39**, 13295–13306.
18. Branden, C. & Tooze, J. (1999). *Introduction to Protein Structure*, Garland, New York.
19. Åqvist, J., Luecke, H., Quiocho, F. A. & Warshel, A. (1991). Dipoles located at helix termini of proteins stabilise charges. *Proc. Natl Acad. Sci. USA*, **88**, 2026–2030.
20. Booth, I. R., Edwards, M. D. & Miller, S. (2003). Bacterial ion channels. *Biochemistry*, **42**, 10045–10053.
21. Morais-Cabral, J. H., Zhou, Y. & MacKinnon, R. (2001). Energetic optimization of ion conduction rate by the K^+ selectivity filter. *Nature*, **414**, 37–42.
22. Zhou, Y. F. & MacKinnon, R. (2003). The occupancy of ions in the K^+ selectivity filter: charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates. *J. Mol. Biol.* **333**, 965–975.
23. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Cahit, B. T. & MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science*, **280**, 69–77.
24. Zhou, Y. F., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K^+ channel–Fab complex at 2.0 Å resolution. *Nature*, **414**, 43–48.

25. Jungwirth, P. & Tobias, D. J. (2002). Chloride anion on aqueous clusters, at the air–water interface, and in liquid water: solvent effects on Cl^- polarizability. *J. Phys. Chem. ser. A*, **106**, 379–383.
26. Stuart, S. J. & Berne, B. J. (1996). Effects of polarizability on the hydration of the chloride ion. *J. Phys. Chem.* **100**, 11934–11943.
27. Yoo, S., Lei, Y. A. & Zeng, X. C. (2003). Effect of polarizability of halide anions on the ionic solvation in water clusters. *J. Chem. Phys.* **119**, 6083–6091.
28. Åqvist, J. & Luzhkov, V. (2000). Ion permeation mechanism of the potassium channel. *Nature*, **404**, 881–884.
29. Bernèche, S. & Roux, B. (2000). Molecular dynamics of the KcsA K^+ channel in a bilayer membrane. *Biophys. J.* **78**, 2900–2917.
30. Chapman, M. L., Krovetz, H. S. & VanDongen, A. M. (2001). GYGD pore motifs in neighbouring potassium channel subunits interact to determine ion selectivity. *J. Physiol. London*, **530**, 21–33.
31. Chen, T. Y. & Miller, C. (1996). Non-equilibrium gating and voltage dependence of the ClC-0 Cl^- channel. *J. Gen. Physiol.* **108**, 237–250.
32. He, J. J. & Quiocho, F. A. (1993). Dominant role of local dipoles in stabilizing uncompensated charges on a sulfate sequestered in a periplasmic active transport protein. *Protein Sci.* **2**, 1643–1647.
33. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (2003). Crystal structure of *M. tuberculosis* ABC phosphate transport receptor: specificity and charge compensation dominated by ion–dipole interactions. *Structure*, **11**, 765–774.
34. Kuo, A., Gulbis, J. M., Antcliff, J. F., Rahman, T., Lowe, E. D., Zimmer, J. *et al.* (2003). Crystal structure of the potassium channel KirBac1.1 in the closed state. *Science*, **300**, 1922–1926.
35. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T. & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature*, **417**, 512–522.
36. Fogolari, F. & Molinari, H. (2002). The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *J. Mol. Recogn.* **15**, 377–392.
37. Simonson, T. (2003). Electrostatics and dynamics of proteins. *Rep. Prog. Phys.* **66**, 737–787.
38. Im, W., Beglov, D. & Roux, B. (1998). Continuum solvation model: electrostatic forces from numerical solutions to the Poisson–Boltzmann equation. *Comput. Phys. Commun.* **111**, 59–75.
39. Nina, M., Beglov, D. & Roux, B. (1997). Atomic radii for continuum electrostatic calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. ser. B*, **101**, 5239–5248.
40. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.
41. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J. *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. ser. B*, **102**, 3586–3616.
42. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucl. Acids Res.* **28**, 235–242.
43. Ben-Tal, N., Ben-Shaul, A., Nicholls, A. & Honig, B. (1996). Free-energy determinants of α-helix insertion into lipid bilayers. *Biophys. J.* **70**, 1803–1812.
44. Bernèche, S., Nina, M. & Roux, B. (1998). Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane. *Biophys. J.* **75**, 1603–1618.
45. Nina, M., Bernèche, S. & Roux, B. (2000). Anchoring of a monotopic membrane protein: the binding of prostaglandin H2 synthase-1 to the surface of a phospholipid bilayer. *Eur. Biophys. J.* **29**, 439–454.
46. Roux, B. & Simonson, T. (1999). Implicit solvent models. *Biophys. Chem.* **78**, 1–20.
47. Hermann, R. B. (1972). Theory of hydrophobic-bonding. II. The correlation between hydrocarbon solubility in water with solvent cavity surface area. *J. Phys. Chem.* **76**, 2754–2759.
48. Hermann, R. B. (1977). Use of solvent cavity area and number of packed solvent molecules around a solute in regard to hydrocarbon solubilities and hydrophobic interactions. *Proc. Natl Acad. Sci. USA*, **74**, 4144–4145.
49. Reynolds, J. A., Gilbert, D. B. & Tanford, C. (1974). Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proc. Natl Acad. Sci. USA*, **71**, 2925–2927.
50. Nagle, J. F. & Tristam-Nagle, S. (2000). Structure of lipid bilayers. *Biochim. Biophys. Acta*, **1469**, 159–195.
51. Stern, H. A. & Feller, S. (2003). Calculation of the dielectric permitivity profile for a nonuniform system: application to a lipid bilayer simulation. *J. Chem. Phys.* **118**, 3401–3412.
52. White, S. H. & Wimley, W. C. (1998). Hydrophobic interactions of peptides with membrane interfaces. *Biochim. Biophys. Acta*, **1376**, 339–352.
53. Simonson, T. & Brünger, A. T. (1994). Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. *J. Phys. Chem.* **98**, 4683–4694.
54. Alexov, E. G. & Gunner, M. R. (1997). Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* **74**, 2075–2093.
55. Schutz, C. N. & Warshel, A. (2001). What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins: Struct. Funct. Genet.* **44**, 400–417.
56. Bernèche, S. & Roux, B. (2002). The ionization state and the conformation of Glu-71 in the KcsA K^+ channel. *Biophys. J.* **82**, 772–780.
57. Wang, Z., Luecke, H., Yao, N. & Quiocho, F. A. (1997). A low energy short hydrogen bond in very high resolution structures of protein receptor–phosphate complexes. *Nature Struct. Biol.* **4**, 519–522.
58. Bower, M. J., Cohen, F. E. & Dunbrack, R. L. (1997). Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *J. Mol. Biol.* **267**, 1268–1282.
59. Banavali, N. K. & Roux, B. (2002). Atomic radii for continuum electrostatics calculations on nucleic acids. *J. Phys. Chem. ser. B*, **106**, 11026–11035.
60. Marcus, Y. (1985). *Ion Solvation*, Wiley, Chichester, UK.
61. Florian, J. & Warshel, A. (1997). Langevin dipoles model for *ab initio* calculations of chemical processes in solution: parametrization and application to hydration free energies of neutral and ionic solutes and conformational analysis in aqueous solution. *J. Phys. Chem. ser. B*, **101**, 5583–5595.
62. Kraulis, P. J. (1991). MOLSCRIPT: a program to

produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
63. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.
64. Chen, M. F. & Chen, T. Y. (2003). Side-chain charge effects and conductance determinants in the pore of ClC-0 chloride channels. *J. Gen. Physiol.* **122**, 133–145.
65. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.
66. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W.: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.
67. Roux, B. (1999). Statistical mechanical equilibrium theory of selective ion channels. *Biophys. J.* **77**, 139–153.

Edited by B. Honig

(Received 14 January 2004; received in revised form 2 April 2004; accepted 6 April 2004)

Supplementary Material comprising two Figures and two Tables is available on Science Direct