

# Influence of electronic polarization on the binding of anions to a chloride-pumping rhodopsin

Linda X. Phan,<sup>1,3</sup> Victor Cruces Chamorro,<sup>2</sup> Hector Martinez-Seara,<sup>2</sup> Jason Crain,<sup>3,4</sup> Mark S. P. Sansom,<sup>3</sup> and Stephen J. Tucker<sup>1,5,\*</sup>

<sup>1</sup>Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, UK; <sup>2</sup>Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 6, Czech Republic; <sup>3</sup>Department of Biochemistry, University of Oxford, Oxford, UK; <sup>4</sup>IBM Research Europe, Hartree Centre, Daresbury, UK; and <sup>5</sup>Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford, UK

**ABSTRACT** The functional properties of some biological ion channels and membrane transport proteins are proposed to exploit anion-hydrophobic interactions. Here, we investigate a chloride-pumping rhodopsin as an example of a membrane protein known to contain a defined anion binding site composed predominantly of hydrophobic residues. Using molecular dynamics simulations, we explore Cl<sup>-</sup> binding to this hydrophobic site and compare the dynamics arising when electronic polarization is neglected (CHARMM36 [c36] fixed-charge force field), included implicitly (via the prosECCo force field), or included explicitly (through the polarizable force field, AMOEBA). Free energy landscapes of Cl<sup>-</sup> moving out of the binding site and into bulk solution demonstrate that the inclusion of polarization results in stronger ion binding and a second metastable binding site in chloride-pumping rhodopsin. Simulations focused on this hydrophobic binding site also indicate longer binding durations and closer ion proximity when polarization is included. Furthermore, simulations reveal that Cl<sup>-</sup> within this binding site interacts with an adjacent loop to facilitate rebinding events that are not observed when polarization is neglected. These results demonstrate how the inclusion of polarization can influence the behavior of anions within protein binding sites and can yield results comparable with more accurate and computationally demanding methods.

**SIGNIFICANCE** Directly visualizing the behavior of water and ions within channels remains a challenge. Nonetheless, we can simulate it using the laws of physics. Current methods based on classical (Newtonian) mechanics are computationally efficient compared with more advanced methods; however, due to approximations, they may not accurately capture these complex and dynamic interactions. Here, we use the prosECCo force field that offers an improved electronic description while maintaining computational efficiency. We show that including polarization greatly influences the binding dynamics of anions to a protein binding site. This has implications for our understanding of permeation in anion-selective ion channels.

## INTRODUCTION

Halide ions (F<sup>-</sup>, Cl<sup>-</sup>, Br<sup>-</sup>, I<sup>-</sup>) are ubiquitous in biological systems (1). Among these, Cl<sup>-</sup> is the most abundant anion and is responsible for a range of physiological processes from cell sensing and signaling to regulation of membrane potential (2). These functions are performed through a diverse range of ion channels and membrane transport proteins that preferentially bind anions, in particular Cl<sup>-</sup>, to facilitate their movement across the membrane (3,4). Dysfunctional Cl<sup>-</sup> channels and transporters are known to result in a variety

of disease states (channelopathies) (5) and many represent attractive therapeutic targets. However, several aspects of how anions interact with these proteins remain poorly understood.

Many studies suggest that Cl<sup>-</sup> can form favorable interactions with a variety of hydrophobic interfaces, from simple aqueous/air interfaces (6) to more complex interfaces with proteins (7). In these circumstances, Cl<sup>-</sup> exhibits preferential adsorption to the interfacial layer where an electrolyte solution and hydrophobic medium meet. At such an interface, Cl<sup>-</sup> becomes polarized due to the anisotropy of the interface, thereby inducing a dipole not present in bulk solution. Interactions between the dipole and surrounding water molecules then compensate for the partial reduction in hydration in the interfacial layer, allowing Cl<sup>-</sup> to come into direct contact

Submitted January 3, 2023, and accepted for publication March 16, 2023.

\*Correspondence: [stephen.tucker@physics.ox.ac.uk](mailto:stephen.tucker@physics.ox.ac.uk)

Editor: Marta Filizola.

<https://doi.org/10.1016/j.bpj.2023.03.026>

© 2023 Biophysical Society.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



with the hydrophobic interface (8). Partial removal of water from the first hydration shell of  $\text{Cl}^-$  is considered energetically favorable (7,9); this phenomenon not only applies to  $\text{Cl}^-$  but also to ions across the Hofmeister series, i.e.,  $\text{F}^- < \text{Cl}^- < \text{Br}^- < \text{I}^-$ . The smaller, less polarizable  $\text{F}^-$  generally retains its hydration shell and is unlikely to be found at aqueous/hydrophobic interfaces, whereas larger, more polarizable anions such as  $\text{Br}^-$  and  $\text{I}^-$  are more easily dehydrated and localize at the interface (9,10). This is supported by experimental studies that indicate that the larger the halide ion, the larger the magnitude of surface adsorption to hydrophobic surfaces and hence change in water structure (11,12).

A rapidly increasing number of high-resolution structures of channels and transporters now exist, many of which exhibit preferential interactions with anions. These include the CFTR  $\text{Cl}^-$  channel that is defective in cystic fibrosis (13), and many anion transporters (14,15) including the NTQ chloride-pumping rhodopsin (CIR) (16). However, despite the availability of such structures, the dynamics of their interactions with anions remains underexplored in part due to the computational challenges faced in the molecular modeling of polarizable components.

The majority of molecular dynamics (MD) simulations utilize classical pairwise additive force fields with fixed point charges (17). Despite their general applicability and computational efficiency, such force fields fail to accurately capture the effects of induced polarization. Consequently, this becomes problematic for the description of systems that depend upon the behavior of anions.

The role of hydrophobic contacts with anions has been investigated in several recent simulation studies using a range of force fields (7,9,18). In particular, ongoing work on force field design has yielded hybrid methods, namely charge scaling approaches, which possess the computational efficiency of standard nonpolarizable force fields yet contain parameterizations that better represent electronic responses within a mean field framework (19). One such approach is the polarization reintroduced by the optimal scaling of electronic continuum correction origin (proECCo) method (20).

proECCo is based on the initial electronic continuum correction (ECC) method, which accounts for polarizability implicitly through rescaling highly charged groups by a factor of  $1/\epsilon_{el}^{1/2}$ , where  $\epsilon_{el}$  represents the electronic component of the dielectric constant and is estimated as the high-frequency dielectric constant ( $\epsilon_{el} = 1.78$  for water and  $\epsilon_{el} = 2$  for proteins) (21–23). Furthermore, the Lennard-Jones parameters require slight adjustments as charge scaling affects the ion-water interactions, therefore a decrease of 5–10% in radius is recommended to recover the correct hydration structure (22). The proECCo force field is based on CHARMM36 (c36) and, when required, incorporates these concepts systematically for a wide range of ions, proteins, lipids, and sugars (20).

Another method to account implicitly for polarization in a nonpolarizable force field is the empirical nonbonded fix (NBFIX) corrections applied to the CHARMM force fields that yields a more accurate model for interacting charges (22,24) and has been shown to influence the dynamics of a  $\text{Cl}^-$ -specific transporter, CLC-ec1 (18).

In this study, we chose to examine a microbial CIR (Fig. 1 A) (PDB: 5G28, 1.6 Å resolution) (16). CIR functions as a light-driven inward  $\text{Cl}^-$  pump. Upon photoactivation, retinal isomerization induces a conformational change that facilitates  $\text{Cl}^-$  transport through the protein (25). CIR contains two  $\text{Cl}^-$  binding sites: the first ( $\text{Cl}^-$  no. 1) is located near the retinal while the second ( $\text{Cl}^-$  no. 2) is on a cytoplasmic loop (Fig. 1 A), where its role in the  $\text{Cl}^-$  transfer pathway is to facilitate  $\text{Cl}^-$  release into the cytosol (16). It is this binding site that is of particular interest because  $\text{Cl}^-$  interacts primarily with the hydrophobic moieties within this site (Fig. 1 C).

Here, we performed MD simulations of CIR embedded in a lipid bilayer with the c36 and proECCo force fields. Potential of mean force (PMF) calculations enabled us to examine the free energy landscapes of a  $\text{Cl}^-$  moving out of the  $\text{Cl}^-$  no. 2 site into bulk. We also simulated a smaller protein fragment that mimics this site to probe the dynamics of binding at the  $\text{Cl}^-$  no. 2 site using the fully polarizable force field, AMOEBA, and compared the behavior of  $\text{Cl}^-$  using the proECCo and c36 force fields. Finally, we explored how residues in a neighboring loop facilitate the rebinding of  $\text{Cl}^-$ . Our results demonstrate that accounting for polarizability yields fundamentally different  $\text{Cl}^-$  interactions with CIR, thereby highlighting how even simplified descriptions can replicate the behavior of anions in simulations.

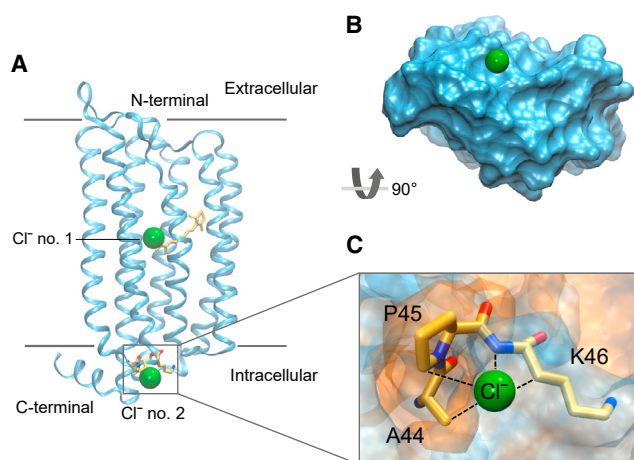
## METHODS

### Structural model and system preparation

Starting from the experimentally determined protein structure (PDB: 5G28) (16), all ligands including retinal, were removed. It was not necessary to parameterize and simulate retinal for the purposes of this study. Systems with the whole-protein and/or reduced fragments containing only the binding site of interest were prepared. Simulations involving the whole protein were embedded in a 1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phosphocholine (POPC) bilayer and prepared using the CHARMM-GUI protocol (26–28). We chose not to apply position restraints to the whole-protein simulations, whereas position restraints with a force constant of 1000 kJ/mol/nm<sup>2</sup> were applied to the  $\text{C}_\alpha$  atoms for the reduced protein simulations. Analysis of RMSDs suggested that the protein simulations without backbone position restraints deviated no more than 1.5 Å from the crystal structure (Fig. S1).

### MD simulations

For comparison, nonpolarizable atomistic simulations were performed using c36 with associated lipid parameters where relevant (29). Simulations with the ECC scaling used the proECCo force field, a variant of the same c36 force field but with scaled charges and compatible ion types (<https://gitlab.com/sparkly/prosecco/prosecco/proECCo75>) (20). While the theoretical framework of NBFIX has proven to successfully reproduce experimental observations through preventing overbinding (18), we chose



**FIGURE 1** (A) Crystal structure of Cl<sup>-</sup>-pumping rhodopsin (CIR) (PDB: 5G28). The first Cl<sup>-</sup> binding site (Cl<sup>-</sup> no. 1) is located in the center of the protein and is in contact with the all-*trans* retinal illustrated by the yellow stick representation in the center of the protein. The second site (Cl<sup>-</sup> no. 2) is located on the intracellular side (gray box) and is the focus of this study. The gray lines depict the extent of the lipid bilayer and Cl<sup>-</sup> ions are represented by the green spheres. (B) A surface representation of the protein rotated by 90° to illustrate the bottom face where the Cl<sup>-</sup> no. 2 site is located. The Cl<sup>-</sup> ion is situated outside the protein. (C) A close-up illustration of the Cl<sup>-</sup> no. 2 binding site. Cl<sup>-</sup> no. 2 is comprised of the hydrophobic residues A44 and P45 (orange) and polar residue K46 (blue). The black dashed lines indicate the known contacts with Cl<sup>-</sup> in the site. To see this figure in color, go online.

prosECCo because it does not require additional pair-specific parameterizations of Lennard-Jones parameters and has previously been applied to the study of halides (23,30).

The systems were equilibrated through a staged protocol whereby structural restraints on the protein backbone and Cl<sup>-</sup> were gradually weakened over a 15 ns equilibration period. Subsequently, all production runs of 55 ns were performed with an integration time step of 2 fs. All systems were solvated with ~0.15 M NaCl using the SPC/E water model (31). The temperature was maintained at 303.15 K with a coupling constant of 1.0 ps using the Nose-Hoover thermostat. Pressure was maintained at 1 bar using the Parrinello-Rahman barostat with a coupling constant of 5.0 ps. Semi-isotropic pressure coupling was used for systems containing a bilayer and isotropic pressure coupling for systems of proteins in solvent. The Verlet cutoff scheme (32) was applied and electrostatics were treated with the particle mesh Ewald method (33). The LINCS algorithm was used to constrain H-bonds only (34). Three independent repeats were carried out for each system. These simulations were performed in GROMACS (35) version 2020 ([www.gromacs.org](http://www.gromacs.org)) and analyzed using MDAnalysis (36,37).

### AMOEBA force field simulations

Polarizable atomic multipole simulations were carried out in OpenMM 7.4.2 ([www.openmm.org](http://www.openmm.org)). All components were modeled with the AMOEBA polarizable force field using the AMOEBA13 protein parameter set (38) and the AMOEBA03 water model (39). Starting configurations for these simulations were taken from the final frames after 15 ns of equilibration using c36 as described above. Simulations were then setup using a similar method to that described previously (<https://github.com/Inniag/openmm-scripts-amoeba>) (7). Simulations using the AMOEBA force field were performed only for the small and large fragments in solvent; therefore, all C<sub>α</sub> atoms were placed under a harmonic restraint with force constant 1000 kJ/mol/nm<sup>2</sup> to prevent the reduced structures deviating from the

experimentally determined structure. Production runs lasted 55 ns and time integration was achieved using the r-RESPA method with an outer time step of 2 fs and an inner time step 0.25 fs. The temperature was maintained at 303.15 K using the Andersen thermostat and the pressure was maintained at 1 bar using the isotropic Monte Carlo barostat. Three independent repeats were performed for each system.

### Umbrella sampling

This was performed to obtain one-dimensional PMF profiles for a Cl<sup>-</sup> ion moving from the binding site into bulk solution. These simulations were carried out on the whole protein embedded in a POPC bilayer and solvated with ~0.15 M NaCl solution using the c36 and prosECCo force fields and the SPC/E water model. Systems were simulated with and without the use of backbone restraints. The collective variable (CV) was defined as the distance between the ion and the center of mass of the protein binding site (residues 44–46) in the negative direction parallel to the *z* axis of the simulation box. Starting configurations for the umbrella windows were obtained from equilibrated simulations. The target ion was relocated to subsequent positions parallel to the *z* axis and the lateral positions of the ion were set to that of the binding site center of mass. During equilibration and sampling, a harmonic biasing potential of 1000 kJ/mol/nm<sup>2</sup> was applied to restrain the CV. Umbrella windows covered the distance from the ion binding site into the bulk water regime. This corresponded to 50 windows with a 0.5 Å distance interval between two successive windows. After ion relocation, 10 steps of energy minimization were performed to remove any steric clashes between the target ion and surrounding atoms. A further 1 ns of isothermal-isobaric equilibration was simulated per window. Each umbrella window was simulated for 1 ns with c36 and 2.5 ns with prosECCo. Simulation details were similar to those detailed above for MD simulations. Unbiasing was performed through WHAM: the weighted histogram analysis method using the Grossfield lab implementation in version 2.0.9 ([http://membrane.urmc.rochester.edu/wordpress/?page\\_id=126](http://membrane.urmc.rochester.edu/wordpress/?page_id=126)). PMF profiles were shifted so that the ion in bulk solution corresponded to 0 kJ/mol.

## RESULTS AND DISCUSSION

### Protein binding site and simulations

In selecting a system to study, we wished to explore anion-hydrophobic interactions within a biologically relevant protein structure. The second Cl<sup>-</sup> binding site (Cl<sup>-</sup> no. 2) of CIR is located on the cytosolic surface (Fig. 1 B) where it facilitates Cl<sup>-</sup> release into the cytosol and has unobstructed access to the bulk cytoplasm (16). In this site, Cl<sup>-</sup> is coordinated by an aliphatic hydrogen of A44, a hydrogen from the aromatic ring of P45, a hydrogen from the side chain of K46 and is considered to be hydrogen-bonded to the backbone nitrogen of K46 (Fig. 1 C).

In this study, we performed simulations of the whole-protein structure with nonpolarizable force fields (c36 and prosECCo). However, due to the methodological and computational demands of implementing a fully polarizable force field (AMOEBA) on these larger simulations, we therefore designed a reduced system containing only the binding site of interest and a few surrounding residues that neither interact with nor obstruct the Cl<sup>-</sup> pathway. We truncated the structure to form a smaller reduced fragment consisting of only the binding site of interest (residues 44–46) (Fig. 1 C) and a few extra surrounding residues (residues 42–49). We refer

to this model as the “smaller fragment.” A larger reduced protein system also included additional nearby residues of interest (residues 252–256) (Fig. 5 A) and is referred to as the “larger fragment.” For these reduced protein systems, position restraints were applied to the  $C_\alpha$  atoms to ensure the structures did not significantly deviate from the experimental coordinates.

### Influence of effective polarization on the energetics of $\text{Cl}^-$ ion binding

We first examined how the energetics of  $\text{Cl}^-$  were impacted by the inclusion of effective polarization by means of the prosECCo force field compared with standard c36. We define the  $CV$  as the  $z$  distance from the center of geometry of the binding site. In addition, we compared how the energetics vary when applying backbone restraints. Examination of the umbrella histograms revealed good overlap between windows, indicating that the resulting PMF profiles had converged (Fig. S2).

The PMF profiles with backbone restraints using the prosECCo force field reveal a free energy minimum of  $\sim -1.9$  kJ/mol at  $CV \sim 3$  Å, a distance of  $\sim 0.6$  Å further from the binding location of the  $\text{Cl}^-$  ion in the crystal structure (Fig. 2 A). With c36, there is an energy well of  $\sim -1.7$  kJ/mol situated at  $CV \sim 4$  Å, which corresponds to a distance of  $\sim 1.3$  Å relative to the experimentally determined ion location. Furthermore, the prosECCo free energy profile exhibits a distinctive second local free energy minimum at  $CV \sim 5$  Å with a free energy  $\sim -1.5$  kJ/mol. This second local minimum is less pronounced when backbone restraints are not implemented (Fig. 2 B) yet is within error of the PMF profile with backbone restraints and is not observed with c36. On closer inspection of the  $\text{Cl}^-$  ion at this position, this second energetic well (Fig. 2 A) appears to be associated with the displacement of the  $\text{Cl}^-$  ion from the binding site by a water molecule. The ion then proceeds to form contacts with the methyl group of A253 and hydrogen atoms from the aromatic ring of Y255.

When the backbone restraints are relaxed to enable protein flexibility and unrestrained dynamics, different free energy landscapes are observed. For the prosECCo force field, a free energy minimum of  $\sim -3.7$  kJ/mol relative to bulk solution at  $CV \sim 3$  Å can be observed, which corresponds to a distance of  $\sim 0.6$  Å relative to the crystal structure (Fig. 2 B). However, this binding position is consistent with the location of the first minima of the PMF profile with backbone restraints, which suggests that the shape of the binding site does not deviate much from the crystal structure when the ion is bound. A profile shift can be seen with the c36 force field, which exhibits an energy well of  $\sim -3.3$  kJ/mol relative to bulk solution at  $CV \sim 4$  Å (Fig. 2 B),  $\sim 1.6$  Å from the crystal structure. Here, the location of the free energy minimum corresponds to a dissociated  $\text{Cl}^-$  ion forming favorable interactions with a nearby

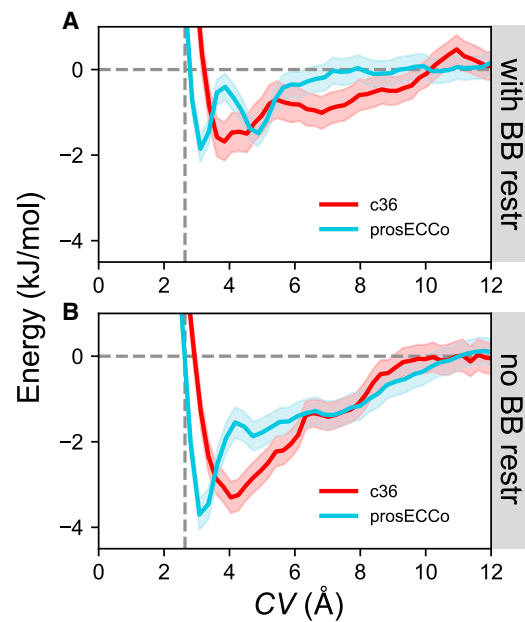


FIGURE 2 Potential of mean force profiles of a  $\text{Cl}^-$  ion moving away from the binding site into bulk solution. The collective variable ( $CV$ ) is defined as a straight line parallel to the  $z$  axis starting from the center of geometry of the binding site and into bulk solution. Free energy profiles from simulations with (A) and without (B) the application of protein backbone restraints are shown in red and blue for the CHARMM36 and prosECCo force fields, respectively. Confidence bands were obtained by calculating the standard error over 200 and 500 ps sampling blocks during the same period for c36 and prosECCo, respectively. The horizontal gray dashed line represents the free energy of a  $\text{Cl}^-$  ion in bulk solution. The vertical gray dashed line indicates the location of the  $\text{Cl}^-$  ion bound in the experimentally determined crystal structure in terms of  $CV$ . To see this figure in color, go online.

lipid headgroup, hence the crystallographic binding site goes undetected with the c36 non polarizable force field.

In both cases (with/without backbone restraints), the interactions with  $\text{Cl}^-$  occur more closely to the crystal structure ion interaction site when employing prosECCo than c36, which could be associated with partial ion dehydration in the site (discussed later, Table 1). However, we find approximately twofold stronger binding when no backbone restraints are applied (Fig. 2 B), an effect that occurs with both force fields.

CIR exhibits similar pumping activity with  $\text{Cl}^-$  and  $\text{Br}^-$ , and so the anomalous signal of  $\text{Br}^-$  was used to locate the ion-binding sites (16). We therefore also performed PMF calculations for the  $\text{Br}^-$  ion moving out of the binding site into bulk using prosECCo. Comparing these PMF profiles (Fig. S3), we observe that the larger, more polarizable  $\text{Br}^-$  ion binds in similar locations as the  $\text{Cl}^-$  ion and with similar relative free energies of binding, thus supporting these experimental observations. In contrast, PMF profiles for cations indicate that a smaller, positively charged  $\text{Na}^+$  ion does not interact with the same binding site, thus explaining the selective nature of the interactions involved in this anion binding site (Fig. S4).



**TABLE 1** Average hydration number for Cl<sup>-</sup> in proximity of interaction site contacts

Force field	Average hydration number				
	H-A44	H-P45	N-K46	H-K46	Bulk
c36	6.9	6.7	6.3	6.7	7.4
prosECCo	5.0	4.9	3.9	4.3	6.4
AMOEBa	5.2	5.0	4.3	5.5	7.3

### Influence of polarization on ion binding dynamics

We next examined the dynamics of the ion binding using the c36, prosECCo, and AMOEBA force fields applied to the smaller protein fragment. We analyzed the distances between Cl<sup>-</sup> and selected atoms in the binding site that are considered to stabilize binding (Fig. 3 A) (16). Ions within an initial radius of 5 Å of the center of geometry of the binding site were first selected and then distances to binding site atoms of selected ions were then calculated. This initial distance was appropriately chosen based on halide interactions within other protein binding sites (1). We considered Cl<sup>-</sup> to be bound if it resided within the distance defined by the first hydration shell (~3.9 Å for all force fields). This was realized through calculating the radial distribution functions between the Cl<sup>-</sup> and surrounding oxygen atoms from water molecules in bulk solvent ( $g_{Cl-O}(r)$ ) (Fig. S5). In this section we focus only on distances between Cl<sup>-</sup> and an aliphatic hydrogen atom from A44 (Fig. 3 A).

For simulations with the c36 force field (Fig. 3 B), Cl<sup>-</sup> occasionally approached the binding site to interact momentarily before dissociating back into solution. These interactions were brief, with an average duration of <0.5 ns and a minimum interaction distance of ~3.0 Å, just within the first hydration shell distance (Fig. 3 B, gray line). In comparison, significantly more Cl<sup>-</sup> ions come into proximity of the binding site when prosECCo is used. The dynamics remain relatively transient, with Cl<sup>-</sup> spending ~0.5 ns interacting with A44 in the binding site at a minimum distance of ~2.9 Å. With the AMOEBA force field, a similar number

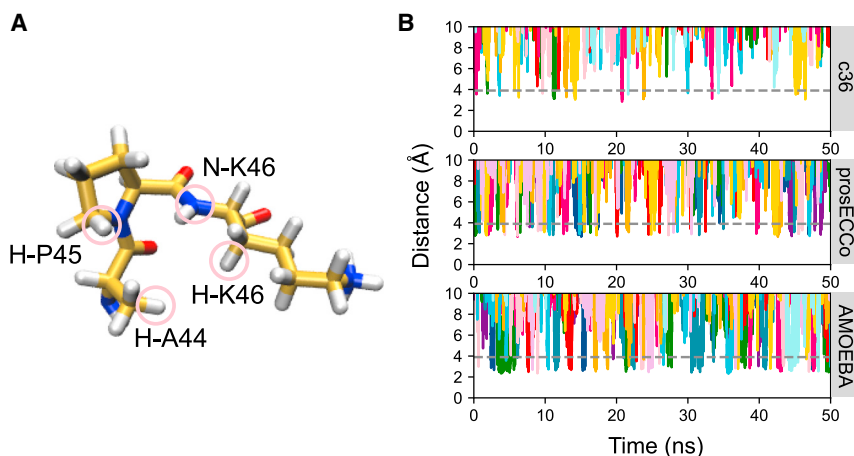
of ions come into proximity of the binding site to interact with an even closer minimum distance of ~2.5 Å and longer duration of occupation ~2 ns.

These data suggest that polarizability leads to stronger interactions between Cl<sup>-</sup> ions and A44 in comparison with standard fixed-charge models. The short interaction times observed here are expected because the functional role of this binding site is to aid ion release into the cytosol (16). Simulations and analysis involving the whole protein embedded in a POPC bilayer with the c36 and prosECCo force fields qualitatively support these conclusions (Fig. S6).

### Cl<sup>-</sup>-hydrophobic interactions within the binding site

We expanded the analysis of the small fragment simulations further to examine the interactions of Cl<sup>-</sup> with the whole binding site by considering the other atoms that contribute to this site (16). Therefore, we computed the distances between the ion and hydrogen atoms from A44 (H-A44), P45 (H-P45), and K46 (H-K46) as well as the backbone amide nitrogen from K46 (N-K46) (Fig. 3 A). This analysis was prepared for each force field and enabled us to gain a more comprehensive understanding of the mechanisms behind Cl<sup>-</sup> binding to this hydrophobic site.

Overall, a notable trend was found across all force fields: the Cl<sup>-</sup> comes into closest proximity with the hydrophobic contacts of the binding site, namely H-A44 and H-P45 (Fig. 4), which are approximately equidistant. This suggests that it is more favorable for Cl<sup>-</sup> to associate with these hydrophobic contacts (8). These Cl<sup>-</sup> preferences were achieved through the partial loss of hydration shell; an effect that was observed previously for Cl<sup>-</sup> ions in a model hydrophobic nanopore (9). As a consequence of including polarization and the resulting induced dipoles,



**FIGURE 3** (A) Schematic diagram of hydrophobic binding site (Cl<sup>-</sup> no. 2). The atoms circled in pink correspond to the atoms used for Cl<sup>-</sup> ion-distance calculations. (B) Ion-distance plots from Cl<sup>-</sup> ions to an aliphatic hydrogen atom of A44 over 55 ns within simulations of the smaller protein fragment. Each color represents an individual ion trajectory, plotting only the ions that initially come into proximity of 5 Å to the binding site. The gray dashed line indicates the binding distance taken as the Cl<sup>-</sup>-oxygen atom of water, i.e., the distance of the first hydration shell ~3.9 Å. To see this figure in color, go online.

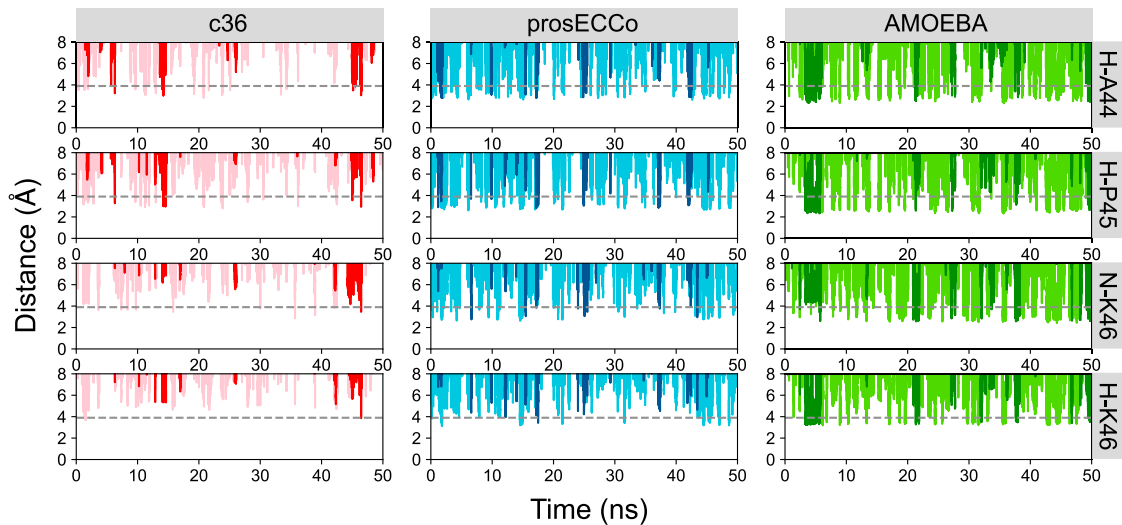


FIGURE 4 Examples of single  $\text{Cl}^-$  ion-distance plots (darker colors). Other  $\text{Cl}^-$  ion-distance plots are also shown as a function of time (lighter colors). The ion distances are measured between a  $\text{Cl}^-$  ion to an aliphatic hydrogen atom of A44 (H-A44), a hydrogen of the aromatic ring of P45 (H-P45), the backbone amide nitrogen of K46 (N-K46), and a hydrogen from the side chain of K46 (H-K46) (see Fig. 3 A). The gray dashed line indicates the binding distance taken as the  $\text{Cl}^-$ -oxygen atom of water, i.e., the distance of the first hydration shell  $\sim 3.9$  Å. To see this figure in color, go online.

the balance between ion-water and ion-protein interactions therefore shifts toward favoring the interactions with the hydrophobic protein contacts over water. This can be realized by calculating the average hydration numbers around the  $\text{Cl}^-$  ion within  $\sim 3.9$  Å of the selected binding site atoms for each given force field (Table 1).

For c36, the average hydration number remains similar for each hydrophobic contact. A slight reduction is observed compared with  $\text{Cl}^-$  in bulk solution. This retention of the hydration shell correlates with the ion-distance plots (Fig. 4) where  $\text{Cl}^-$  ions do not come into as close proximity with the binding site contacts compared with prosECCo and/or AMOEBA. With the prosECCo force field, a substantial  $\text{Cl}^-$  ion dehydration occurs in the binding site relative to bulk solution. Based on previous studies using the ECC method (9), this effect is anticipated and is again reflected in the  $\text{Cl}^-$  approach distance illustrated in Fig. 4. Moreover, this dehydration is energetically advantageous as it promotes ion binding (Fig. 2). The largest dehydration is observed in simulations with the AMOEBA force field where  $\text{Cl}^-$  loses  $\sim 2$  water molecules relative to bulk when in the vicinity of hydrophobic contacts. This enables  $\text{Cl}^-$  to exist in a more stably bound state reflected by the longer binding durations (Fig. 4) and increased ion occupancy (Fig. 3 B). This is also consistent with the interactions and behaviors associated with anion hydration that have been more thoroughly explored in studies by Rempe and co-workers (40,41).

The ion-distance plots (Fig. 4) suggest that N-K46 is further than H-A44 and H-P45 from  $\text{Cl}^-$  within the site. This is a result of the tendency for the tail of K46 to enclose the anion such that N-K46 is then positioned deeper inside the concave cavity formed by the binding site as a whole (Fig. S7).

### Influence of polarization on the mechanism of $\text{Cl}^-$ binding

Based on findings from the PMF profiles discussed above (Fig. 2), we further explored the mechanisms that give rise to the metastable free energy minimum observed in simulations employing the prosECCo force field (second minima at  $\sim 5$  Å, Fig. 2 A). On closer observation of the umbrella sampling windows, the residues A253 and Y255 in a nearby loop appeared to coordinate  $\text{Cl}^-$  at the location of the second free energy minimum. Notably, these residues are hydrophobic. To explore the influence of these hydrophobic loop contacts, we further simulated a larger protein fragment consisting of the previous smaller fragment with an additional fragment consisting of residues from the nearby loop (residues 252–256). This larger fragment was then simulated using the c36, proECCo, and AMOEBA force fields. We calculated the distance between  $\text{Cl}^-$  and the coordinating binding site atoms as before, but also now calculating the distance to a hydrogen from the methyl group of A253 and a hydrogen of CD1 belonging to the aromatic ring of Y255 (Fig. 5 A).

Simulations of this larger fragment with the c36 force field did not show any significant interactions with residues Y255 and A253. However, with the prosECCo force field (Fig. 5 B) we observed instances where  $\text{Cl}^-$  was initially bound to its binding site then dissociated to interact with the nearby loop (e.g., A253) and then proceeded to rebind to its site once more. Due to the intriguing nature of this observation, we therefore repeated the simulations with a fully polarizable force field.

With AMOEBA, the ion also dissociated from the binding site in exchange with a water molecule (e.g., Fig. 5 C,

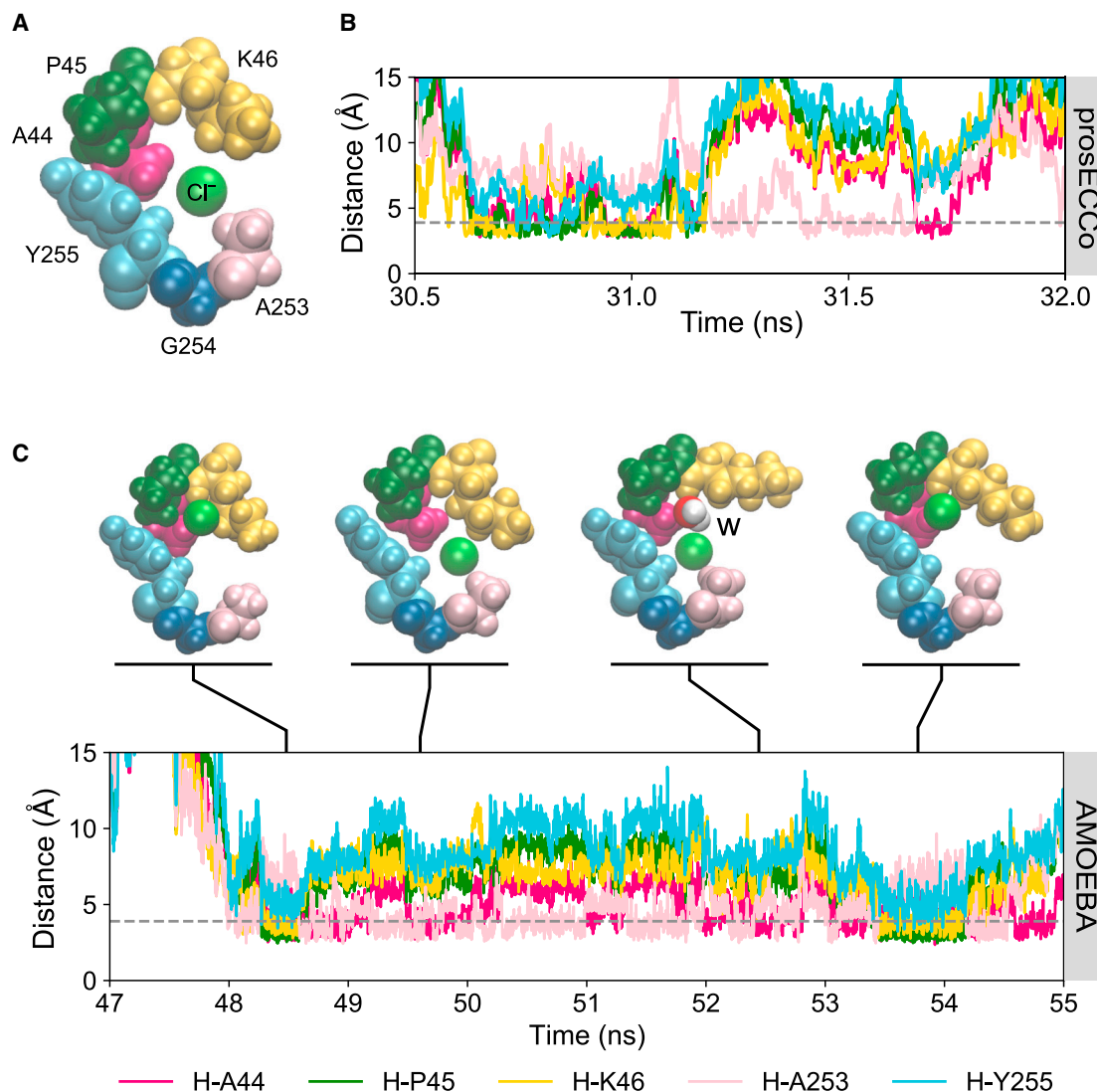


FIGURE 5 (A) Larger protein fragment containing the  $\text{Cl}^-$  no. 2 binding site and residues from a nearby loop. The binding site is composed of A44 in magenta, P45 in dark green, and K46 in yellow. The additional residues of interest from the nearby loop are illustrated by A253 in pink, G254 in navy, and Y255 in light blue. (B) Example of an ion-distance plot as a function of time for a  $\text{Cl}^-$  ion in the binding site simulated using the proSECCo force field. (C) An ion-distance plot as a function of time for a  $\text{Cl}^-$  ion in the binding site simulated using the AMOEBA force field. Snapshots of a  $\text{Cl}^-$  ion in the binding site illustrate the ion dissociation and rebinding mechanism. A water molecule, marked by “w,” can be observed to displace the  $\text{Cl}^-$  ion from  $\text{Cl}^-$  no. 2 while the ion favorably interacts with nearby hydrophobic residues A253 and Y255. To see this figure in color, go online.

snapshot marked w). From here, the ion could then form favorable interactions with A253 and/or Y255 to remain in close proximity to the binding site, thus providing the opportunity to rebind or otherwise dissociate into bulk solution. An example of such a rebinding event is illustrated in Fig. 5 C. These findings are supported by simulations of the full protein embedded in a lipid bilayer with the c36 and proSECCo force fields where the effect was not observed with c36 but was with proSECCo (Fig. S8). This mechanism is therefore only observed when polarizability is considered. It is interesting to note that the ion binding/rebinding occurs on a longer timescale with AMOEBA in comparison with c36 and proSECCo, suggesting that the

rate of transport of  $\text{Cl}^-$  may be lower than that predicted by nonpolarizable force fields.

Improving common fixed-charged methods to better capture the effects of polarization (e.g., proSECCo) preserves computational efficiency and Table S1 compares the computational performance of each force field per system size. As demonstrated here, proSECCo can act as a proxy for identifying potentially interesting anion interactions at no additional computational cost to a standard nonpolarizable force field. These interactions can then be examined further by performing simulations using explicitly polarizable models such as AMOEBA (38), CHARMM Drude (42), etc. To some degree, proSECCo can replicate the

qualitative behavior produced by more advanced force fields; however, to gain a higher level of physical accuracy, extensive simulations employing fully polarizable force fields may be required or even ab initio MD calculations.

## CONCLUSIONS

We have performed simulations of a Cl<sup>-</sup>-pumping rhodopsin to investigate the effects of polarizability on the binding of anions within a defined binding site. Our results demonstrate how the inclusion of electronic polarization leads to stronger anion binding events and longer binding durations. These effects result from energetically favorable interactions with a site-adjacent loop and partial ion dehydration to interact with the hydrophobic moieties of the binding site. Crucially, these binding mechanisms were only observed when polarization (effects) were incorporated into the simulations.

Future work might explore more complex biological anion binding sites by applying both proECCo and AMOEBA force fields, which will help evaluate the applicability of simplified models for complex interactions arising from polarization. Structures of future interest might include the Na<sup>+</sup>/I<sup>-</sup> symporter (PDB: 7UV0) that mediates active I<sup>-</sup> transport and contains a highly conserved I<sup>-</sup> binding site consisting of predominantly hydrophobic and aromatic residues (15). Similarly, the Cl<sup>-</sup>-selective human bestrophin channels (hBest1 and hBest2) contain a set of highly conserved hydrophobic residues within the neck of the permeation pathway (43). These structures suggest not only an important role for hydrophobic contacts in the ion permeation pathway but potentially influential roles of anion- $\pi$  interactions.

## SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2023.03.026>.

## AUTHOR CONTRIBUTIONS

L.X.P., V.C.C., and H.M.-S. performed research and analyzed data. All authors designed research and wrote the paper. S.J.T., J.C., and M.S.P.S. obtained funding for the project.

## ACKNOWLEDGMENTS

This work was supported by grants from the BBSRC and the EPSRC and by an EPSRC iCASE studentship award in collaboration with IBM Research to L.X.P. It was also supported by the Hartree National Centre for Digital Innovation—a collaboration between the Science and Technologies Facilities Council and IBM. We would also like to acknowledge use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work.

## DECLARATION OF INTERESTS

J.C. is an employee of IBM Research.

## REFERENCES

- Skitchenko, R. K., D. Usoltsev, ..., A. Guskov. 2020. Census of halide-binding sites in protein structures. *Bioinformatics*. 36:3064–3071.
- Gouaux, E., and R. Mackinnon. 2005. Principles of selective ion transport in channels and pumps. *Science*. 310:1461–1465.
- Hille, B. 2001. *Ionic Channels of Excitable Membranes*, 3rd ed. Sinauer Associates Inc.
- Dutzler, R., E. B. Campbell, and R. Mackinnon. 2003. Gating the selectivity filter in CIC chloride channels. *Science*. 300:108–112.
- Ashcroft, F. 2000. *Ion Channels and Disease*. Academic Press, San Diego.
- Garrett, B. C. 2004. Ions at the air/water interface. *Science*. 303:1146–1147.
- Klesse, G., S. Rao, ..., M. S. P. Sansom. 2020. Induced polarization in molecular dynamics simulations of the 5-HT<sub>3</sub> receptor channel. *J. Am. Chem. Soc.* 142:9415–9427.
- Petersen, P. B., and R. J. Saykally. 2006. On the nature of ions at the liquid water surface. *Annu. Rev. Phys. Chem.* 57:333–364.
- Phan, L. X., C. I. Lynch, ..., S. J. Tucker. 2022. Influence of effective polarization on ion and water interactions within a biomimetic nanopore. *Biophys. J.* 121:2014–2026.
- Bajaj, P., M. Riera, ..., F. Paesani. 2019. Halide ion microhydration: structure, energetics, and spectroscopy of small halide-water clusters. *J. Phys. Chem. A*. 123:2843–2852.
- Piatkowski, L., Z. Zhang, ..., M. Bonn. 2014. Extreme surface propensity of halide ions in water. *Nat. Commun.* 5:4083.
- Petersen, P. B., and R. J. Saykally. 2006. Probing the interfacial structure of aqueous electrolytes with femtosecond second harmonic generation spectroscopy. *J. Phys. Chem. B*. 110:14060–14073.
- Linsdell, P. 2016. Anion conductance selectivity mechanism of the CFTR chloride channel. *Biochim. Biophys. Acta*. 1858:740–747.
- Jojoa-Cruz, S., K. Saotome, ..., A. B. Ward. 2022. Structural insights into the Venus flytrap mechanosensitive ion channel Flycatcher1. *Nat. Commun.* 13:850.
- Ravera, S., J. P. Nicola, N. Carrasco, ..., 2022. Structural insights into the mechanism of the sodium/iodide symporter (NIS). *Nature*. 612:795–801.
- Kim, K., S. K. Kwon, ..., H. S. Cho. 2016. Crystal structure and functional characterization of a light-driven chloride pump having an NTQ motif. *Nat. Commun.* 7:12677.
- Lopes, P. E. M., B. Roux, and A. D. MacKerell. 2009. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability: theory and applications. *Theor. Chem. Acc.* 124:11–28.
- Orabi, E. A., T. N. Öztürk, ..., J. D. Faraldo-Gómez. 2021. Corrections in the CHARMM36 parametrization of chloride interactions with proteins, lipids, and alkali cations, and extension to other halide anions. *J. Chem. Theory Comput.* 17:6240–6261.
- Chen, S., and G. A. Voth. 2023. How does electronic polarizability or scaled-charge affect the interfacial properties of room temperature ionic liquids? *J. Phys. Chem. B*. 127:1264–1275.
- Nencini, R., C. Tempra, ..., H. Martinez-Seara. 2022. Prosecco: polarization reintroduced by optimal scaling of electronic continuum correction origin in MD simulations. *Biophys. J.* 121:157a.
- Leontyev, I. v., and A. A. Stuchebrukhov. 2009. Electronic continuum model for molecular dynamics simulations. *J. Chem. Phys.* 130:085102.
- Duboue-Dijon, E., M. Javanainen, ..., H. Martinez-Seara. 2020. A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization. *J. Chem. Phys.* 153:1–15.



23. Pluhařová, E., H. E. Fischer, ..., P. Jungwirth. 2014. Hydration of the chloride ion in concentrated aqueous solutions using neutron scattering and molecular dynamics. *Mol. Phys.* 112:1230–1240.
24. Tolmachev, D. A., O. S. Boyko, M. Karttunen..., 2020. Overbinding and qualitative and quantitative changes caused by simple Na<sup>+</sup> and K<sup>+</sup> ions in polyelectrolyte simulations: comparison of force fields with and without NBFIX and ECC corrections. *J. Chem. Theory Comput.* 16:677–687.
25. Yun, J.-H., X. Li, ..., H. Liu. 2021. Early-stage dynamics of chloride ion-pumping rhodopsin revealed by a femtosecond X-ray laser. *Proc. Natl. Acad. Sci. USA.* 118:e2020486118.
26. Lee, J., X. Cheng, ..., W. Im. 2016. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* 12:405–413.
27. Brooks, B. R., C. L. Brooks, ..., M. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.
28. Jo, S., T. Kim, ..., W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.
29. Best, R. B., X. Zhu, A. D. MacKerell..., 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory Comput.* 8:3257–3273.
30. Vazdar, M., E. Pluhařová, ..., P. Jungwirth. 2012. Ions at hydrophobic aqueous interfaces: molecular dynamics with effective polarization. *J. Phys. Chem. Lett.* 3:2087–2091.
31. Fuentes-Azcatl, R., N. Mendoza, and J. Alejandre. 2015. Improved SPC force field of water based on the dielectric constant: SPC/e. *Phys. Stat. Mech. Appl.* 420:116–123.
32. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.
33. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
34. Hess, B., H. Bekker, J. G. E. Fraaije..., 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.
35. Abraham, M. J., T. Murtola, ..., E. Lindahl. 2015. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.
36. Gowers, R. J., M. Linke, ..., O. Beckstein. 2016. MDAAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. *In Proc. of the 15th Python in Science Conference*, pp. 98–105.
37. Michaud-Agrawal, N., E. J. Denning, ..., O. Beckstein. 2011. MDA-analysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 32:2319–2327.
38. Shi, Y., Z. Xia, ..., P. Ren. 2013. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* 9:4046–4063.
39. Ren, P., and J. W. Ponder. 2003. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B.* 107:5933–5947.
40. Muralidharan, A., L. R. Pratt, S. B. Rempe..., 2019. Quasi-chemical theory for anion hydration and specific ion effects: Cl<sup>-</sup>(aq) vs. F<sup>-</sup>(aq). *Chem. Phys. Lett.* 737:100037.
41. Gomez, D. T., L. R. Pratt, ..., S. B. Rempe. 2022. Hydrated anions: from clusters to bulk solution with quasi-chemical theory. *Acc. Chem. Res.* 55:2201–2212.
42. Huang, J., P. E. M. Lopes, ..., A. D. MacKerell. 2014. Recent advances in polarizable force fields for macromolecules: microsecond simulations of proteins using the classical drude oscillator model. *J. Phys. Chem. Lett.* 5:3144–3150.
43. Owji, A. P., J. Wang, ..., T. Yang. 2022. Structures and gating mechanisms of human bestrophin anion channels. *Nat. Commun.* 13:3836.

**Biophysical Journal, Volume 122**

**Supplemental information**

**Influence of electronic polarization on the binding of anions to a chloride-pumping rhodopsin**

**Linda X. Phan, Victor Cruces Chamorro, Hector Martinez-Seara, Jason Crain, Mark S.P. Sansom, and Stephen J. Tucker**

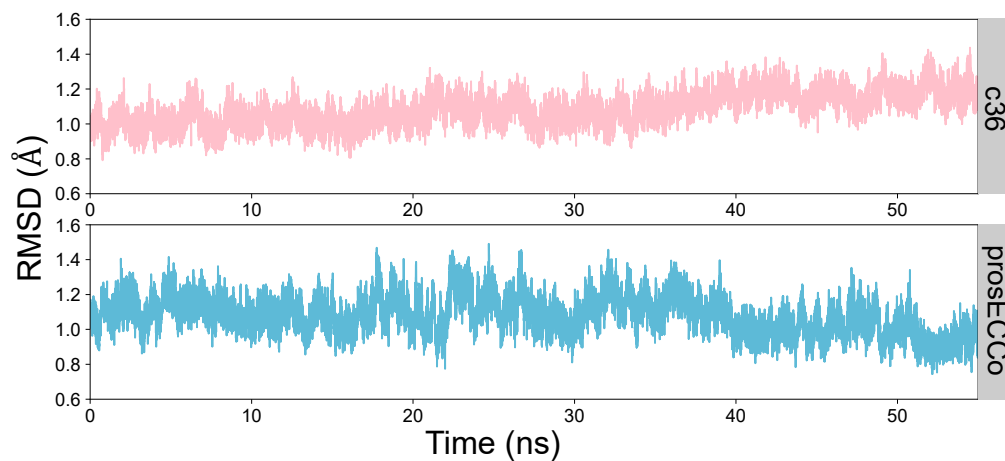
**Biophysical Journal, Volume 122**

**Supplemental information**

**Influence of electronic polarization on the binding of anions to a chloride-pumping rhodopsin**

**Linda X. Phan, Victor Cruces Chamorro, Hector Martinez-Seara, Jason Crain, Mark S.P. Sansom, and Stephen J. Tucker**

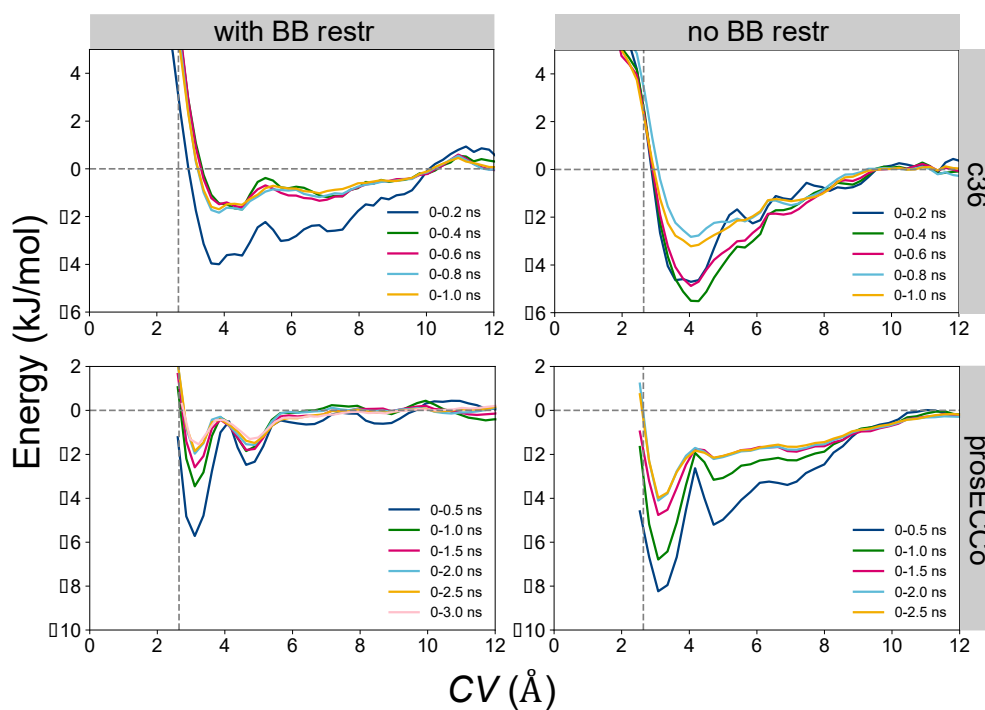
**Figure S1**



**Figure S1:** RMSDs of the protein backbone in simulations of the whole protein embedded in a lipid bilayer without backbone position restraints applied. The protein deviates by  $< 2.0$  Å compared to the crystal structure and therefore justifies omitting the use of backbone restraints in simulations of the whole protein.

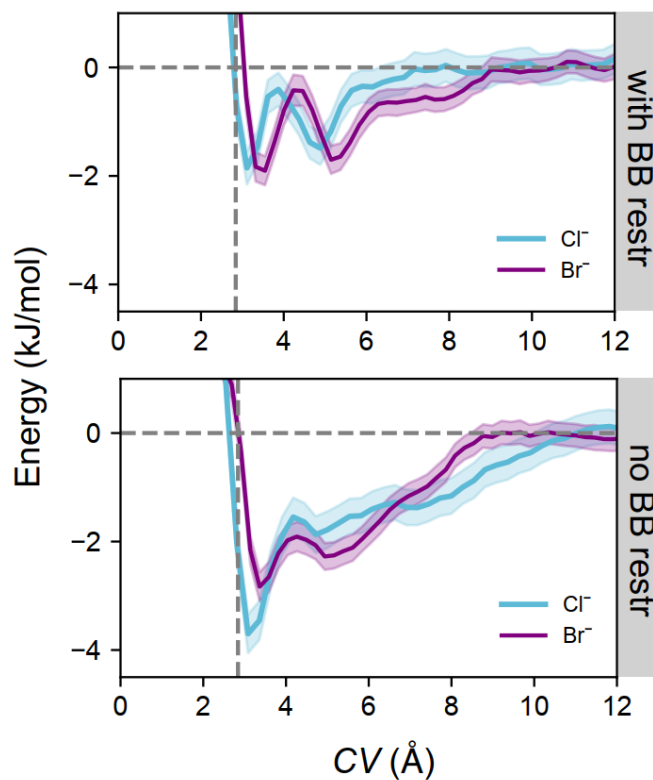


Figure S2



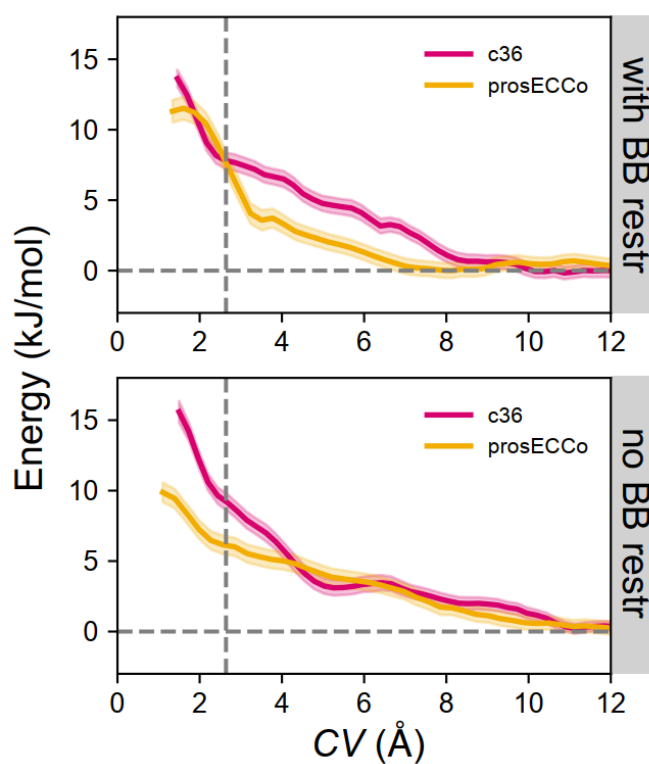
**Figure S2:** Cl<sup>-</sup> PMF convergence analysis using the CHARMM36 (c36) and prosECCo force fields with and without the use of protein backbone restraints. Convergence analysis was performed by calculating 0.2 ns and 0.5 ns sampling blocks over the sampling time for of 1.0 ns and 2.5 ns for c36 and prosECCo respectively. The vertical grey dashed line represents the position of the crystal structure binding site in terms of *CV*.

Figure S3



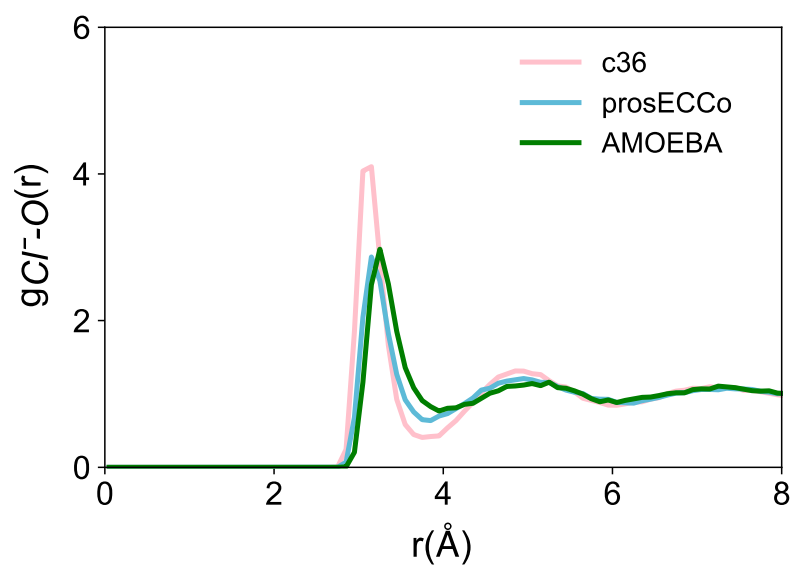
**Figure S3:** Potential of mean force profiles for a  $\text{Cl}^-$  and  $\text{Br}^-$  ion moving away from the binding site into bulk solution employing the prosECCo force field. Protein backbone restraints were not applied to the simulations. The horizontal grey dashed line represents the free energy of the ion in bulk solution and the vertical grey dashed line is representative of the location of the corresponding ion in the resolved crystal structure.

Figure S4



**Figure S4:** Potential of mean force profiles for a  $\text{Na}^+$  ion moving away from the binding site into bulk solution with the c36 and prosECCo force fields.  $\text{Na}^+$  does not appear to bind to the defined binding site. The horizontal grey dashed line represents the free energy of the ion in bulk solution and the vertical grey dashed line is representative of the location of the  $\text{Cl}^-$  ion in the resolved crystal structure in terms of  $CV$ . Samples over 1 ns per umbrella window were used in unbiasing for both force fields.

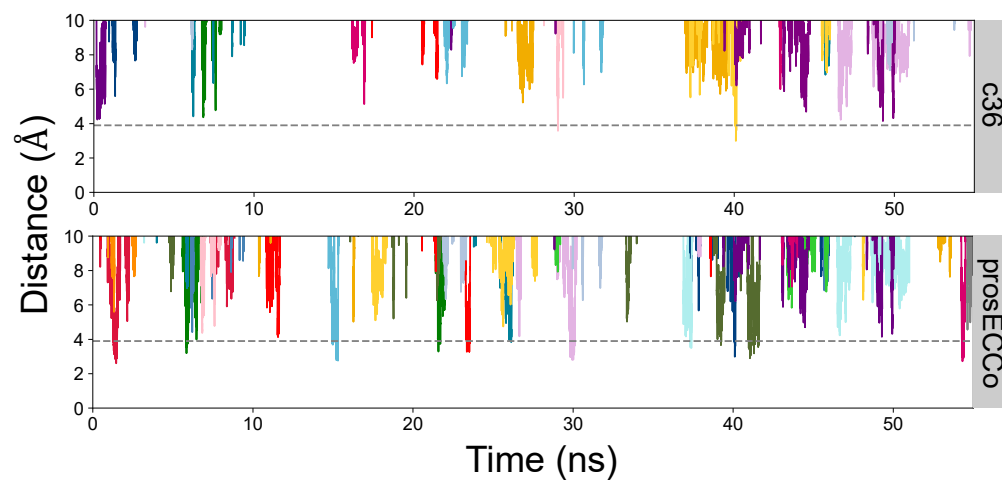
**Figure S5**



**Figure S5:** Radial distribution functions (RDFs),  $g_{\text{Cl}^--\text{O}}(r)$ , of water oxygen atoms around a  $\text{Cl}^-$  ion in bulk solution with the c36, prosECCo and AMOEBA forcefields.

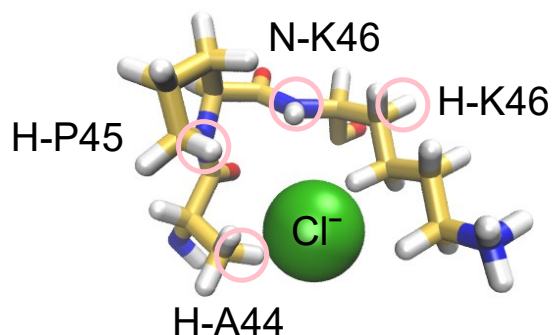


**Figure S6**



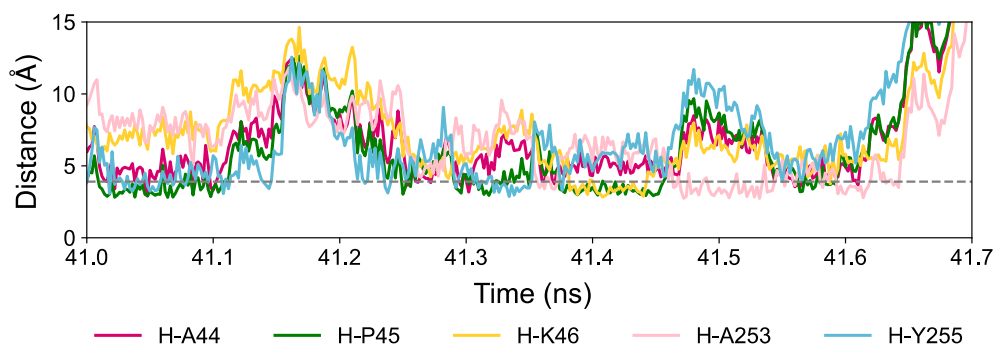
**Figure S6:** Ion-distance plots between  $\text{Cl}^-$  ions to an aliphatic hydrogen atom of A44 within simulations of the whole protein embedded in a lipid bilayer with the c36 and proECCo forcefields. Each color represents an individual ion trajectory, plotting only the ions that initially come within a distance of 5 Å to the binding site.

**Figure S7**



**Figure S7:** Snapshot of a  $\text{Cl}^-$  ion in the binding site. This is an example of where the K46 tail encloses the ion to form a deeper concave shape that facilitates ion binding.

**Figure S8**



**Figure S8:** Ion-distance plot as a function of time for a  $\text{Cl}^-$  ion in the binding site and residues from a nearby loop. The ion distances were measured from simulations of the whole protein embedded in a lipid bilayer using the proECCo force field. The  $\text{Cl}^-$  is bound to the site (contacts H-A44, H-P45 & H-K46) then dissociates (at  $\sim 41.5$  ns) and forms interactions with residues from the nearby loop (H-A253 & H-Y255).  $\text{Cl}^-$  rebinding can be observed (at  $\sim 41.6$  ns) and the ion is then released into bulk solution (at  $\sim 41.7$  ns).

**Table S1**

<b>System Size</b>	<b>Force field</b>	<b>ns/day</b>
Whole Protein ~ 82300 atoms	c36	110
	prosECCo	107
	AMOEBA	N/A
Fragments ~ 12100 atoms	c36	690
	prosECCo	697
	AMOEBA	28

**Table S1:** Computational performance comparison of different system sizes per force field. All systems were benchmarked using 1 GPU (NVIDIA Quadro RTX8000), 1 node with 6 cores (CPUs: Intel Xeon Platinum 8268). c36 and prosECCo have very similar computational costs by running on GROMACS 2020. AMOEBA was run with the OpenMM 7.4.2 implementation which enables GPU acceleration and is considerably more computationally expensive relative to c36 and prosECCo.