



UNIVERSITÀ DEGLI STUDI  
DI PERUGIA



# Force Field Assessment and Molecular Dynamics of Na<sup>+</sup>/K<sup>+</sup> Conduction Through a Human Voltage-Gated Sodium Channel

HANNAH POLLAK

SUPERVISORS:

Juan José Nogueira  
Noelia Faginas Lago  
Andrea Lombardi

Master Thesis

ERASMUS MUNDUS MASTER IN THEORETICAL CHEMISTRY AND COMPUTATIONAL  
MODELLING

July 2020





**UNIVERSITÀ DEGLI STUDI DI PERUGIA**

**DIPARTIMENTO DI CHIMICA, BIOLOGIA E BIOTECHNOLOGIA**

**CORSO DI LAUREA MAGISTRALE IN SCIENZE CHIMICHE**

**TESI DI LAUREA MAGISTRALE**

**Force Field Assessment and Molecular Dynamics of Na<sup>+</sup>/K<sup>+</sup>  
Conduction Through a Human Voltage-Gated Sodium  
Channel**

**LAUREANDA:**  
Hannah Pollak

**RELATORI:**  
Noelia Faginas Lago  
Andrea Lombardi  
Juan José Nogueira

**ANNO ACCADEMICO 2019/20**



## ABSTRACT

---

Voltage-gated sodium ( $\text{Na}_v$ ) channels are responsible for the generation and propagation of action potentials in neurons. They are membrane proteins which, in their open state, permit a depolarising inward current of  $\text{Na}^+$  from the extracellular space into the cytosol. This process is highly discriminative. In mammalian  $\text{Na}_v$  channels, selectivity arises from the so-called DEKA motif which is constituted of the conserved Asp, Glu, Lys, Ala residues which define the pore's constriction site. Experimentally determined eukaryotic  $\text{Na}^+/\text{K}^+$  conduction ratios range from 10 to 30.

Previous molecular dynamics studies were not able to reproduce these experimentally observed selectivities and up to date, a mechanism of ion transport in eukaryotic  $\text{Na}_v$  channels has yet to be confirmed.

In this thesis we used classical molecular dynamics to study ion conduction and selectivity in a truncated pore model representing an open human  $\text{Na}_v1.4$  channel. The results from our simulations clearly suggest higher  $\text{K}^+$  than  $\text{Na}^+$  conductance and  $\text{Na}^+/\text{K}^+$  conduction ratios decreased in the presence of increasing external electric fields and was highest at physiological conditions of 100 mV. An accurate representation of selectivity is likely to require a more elaborate description of intermolecular interactions than it can be achieved employing a simple pairwise fixed-charge interaction potential. Removing a fraction of the positive charge from the selectivity filter's lysine residue should provide a rough approximation of polarisation effects. However, this modification demonstrated to not increase ion discrimination in the selectivity filter.

The present results suggest that the employed force field parametrisation is not adequate for modelling the complex electronic and solvation effects which are responsible for selectivity in eukaryotic  $\text{Na}_v$  channels.

In four identified binding pockets at the entrance and exit of the channel pore, negative binding energies coincide with reduction of water molecules in the first hydration shell. In all sites, binding of  $\text{Na}^+$  is favoured over  $\text{K}^+$  and greater extent of dehydration is observed for the sodium ion.

## SOMMARIO

---

I canali ionici voltaggio-dipendenti per il sodio ( $\text{Na}_v$ ) sono responsabili della generazione e della propagazione dei potenziali d'azione nei neuroni. Essi sono proteine di membrana che, nello stato aperto, permettono una corrente in ingresso depolarizzante di  $\text{Na}^+$  dallo spazio extracellulare al citosolo. Questo processo è molto selettivo. Nei canali  $\text{Na}_v$  dei mammiferi, la selettività si origina da un motivo DEKA che è costituito da residui altamente conservati di Asp, Glu, Lys, Ala i quali formano il sito di costrizione. I raggi di conduzione  $\text{Na}^+/\text{K}^+$  negli eucarioti, determinati sperimentalmente, variano tra 10 e 30.

In passato, vari studi in dinamica molecolare non sono stati in grado di riprodurre queste selettività osservate negli esperimenti e, ad oggi, non è ancora stato chiarito il meccanismo di trasporto ionico nei canali  $\text{Na}_v$  degli eucarioti.

In questa tesi abbiamo usato la dinamica molecolare classica per studiare la conduttività ionica e la selettività tramite un modello a canale troncato che è rappresentativo di un canale umano  $\text{Na}_v1.4$  aperto. Le nostre simulazioni suggeriscono una conduttanza più alta per il  $\text{K}^+$  rispetto al  $\text{Na}^+$  e si è osservato che la selettività per il sodio diminuisce all'aumentare di un campo elettrico esterno ed è risultata massima alle condizioni fisiologiche di 100 mV. Una più accurata rappresentazione della selettività probabilmente richiederebbe una descrizione più elaborata delle interazioni intermolecolari di quella che può essere raggiunta impiegando un semplice potenziale di interazioni tra coppie a carica fissa. Nelle simulazioni, la rimozione di una frazione della carica positiva dal residuo di lisina del filtro di selettività è stata applicata per ottenere una approssimazione grezza degli effetti di polarizzazione. In ogni caso, questa modifica non ha dimostrato di aumentare la discriminazione degli ioni nel filtro di selettività.

I risultati suggeriscono che la parametrizzazione del campo di forza impiegata non è adeguata a modellare i complessi effetti elettronici e di solvatazione che sono responsabili della selettività nei canali di  $\text{Na}_v$  degli eucarioti.

Nei quattro siti di legame individuati all'ingresso e all'uscita del canale, energie di legame negative coincidono con una riduzione delle molecole di acqua nel primo strato di idratazione. In tutti i siti, il legame di sodio è favorito rispetto a quello di potassio e una maggiore deidratazione si osserva nel caso del sodio.

## ACKNOWLEDGMENTS

---

I feel very fortunate for having been given the opportunity to take part in the Erasmus Mundus TCCM Joint Master programme and would, therefore, like to thank everyone involved in this project.

I owe immense gratitude to my supervisor team from the University of Perugia, Noelia and Andrea, who always had open ears for my questions. I would like to particularly thank Noelia for her guidance as my local coordinator.

For the creation of this thesis, first and foremost, I am deeply grateful to my supervisor Juanjo from the Autonomous University of Madrid who never ceased to encourage and support me. In addition to the incredible learning experience, working in the MoBioChem group has always been very inspiring and delightful for me.

In this spirit, I would also like to acknowledge the other group members and the rest of the TCCM family who made our intensive courses become memorable and fun.

Finally, I would like to state my heartfelt appreciation to all my family and friends, near and far, for their relentless support during the past two years.



## CONTENTS

---

1	THEORETICAL BACKGROUND	1
1.1	Ion Channels . . . . .	1
1.2	Biological Role and Function of Ion Channels . . . . .	1
1.3	Ion Fluxes and Conductance . . . . .	2
1.4	Structure of Human Voltage-Gated Sodium Channels . . . . .	3
1.5	Motivation . . . . .	5
1.6	Statistical Averages and the Ergodic Hypothesis . . . . .	6
1.7	Equations of Motion . . . . .	7
1.8	The Virial Equation of State, Temperature and Pressure . . . . .	8
1.9	Steps in an Molecular Dynamics Simulation Cycle . . . . .	11
1.10	Periodic Boundary Conditions . . . . .	14
1.11	Temperature Control in the Isothermal Ensemble . . . . .	15
1.12	Pressure Control and the Isothermal-Isobaric Ensemble . . . . .	16
1.13	Force Fields . . . . .	17
1.13.1	Bonded Interactions . . . . .	18
1.13.2	Non-Bonded Interactions . . . . .	19
1.14	Polarisability and Higher Order Electrostatic Interactions . . . . .	21
1.14.1	Polarisable Force Fields . . . . .	23
2	COMPUTATIONAL DETAILS	25
2.1	Model Construction . . . . .	25
2.2	Molecular Dynamics Simulations . . . . .	26
2.2.1	Equilibration Runs . . . . .	27
2.2.2	Production Runs . . . . .	27
2.2.3	Constant Electric Field Simulations and Decreased Lys1244 Charge . . . . .	28
3	RESULTS	29
3.1	Pore Models . . . . .	29
3.2	Quantification of Ion Flux and Conductance . . . . .	32
3.3	External Field Strength and Lys1244 charge . . . . .	40
3.4	Binding Pockets and Dehydration . . . . .	45
4	CONCLUSIONS	53
	BIBLIOGRAPHY	55
	APPENDICES	57
A	PORE MODELS	59
B	FORCE CONSTANTS OF CONSTRAINTS AND RESTRAINTS	61
C	PROTEIN GEOMETRIES	63
D	CONDUCTANCE DATA	65
E	ION DENSITIES, BINDING FREE ENERGIES AND COORDINATION NUMBERS	69
F	DENSITY PROFILES OF THE SELECTIVITY FILTER RESIDUES	75

## LIST OF FIGURES

---

Figure 1.1	Schematic representation of the $\text{Na}_v\text{-}\alpha$ -subunit	4
Figure 1.2	Top view on selectivity filter residues. . . . .	5
Figure 1.3	Steps in a Molecular Dynamics simulation. . .	14
Figure 1.4	Van-der-Waals', repulsive, and Lennard-Jones potential energy functions. . . . .	22
Figure 2.1	Top view on cryo-electron microscopy structure of the human $\text{Na}_v1.4$ channel's $\alpha$ -subunit . . . .	25
Figure 2.2	Pore model systems A, B, C, D. . . . .	26
Figure 2.3	Side view of selectivity filter with indicated direction of external field and top view on the Lys1244 $\epsilon$ -ammonium group in pore model C.	28
Figure 3.1	Probability distributions of the DEKA residues in pore models A, B, C, D. . . . .	30
Figure 3.2	Average $\text{Na}^+$ and $\text{K}^+$ current, in the simulation box in function of different external electric potentials, $U$ . . . . .	33
Figure 3.3	Distances used for determining the size of the pore region. . . . .	34
Figure 3.4	Top and side view on Z1 volume element . . .	34
Figure 3.5	Average $\text{Na}^+$ and $\text{K}^+$ current in cylindrical volume element Z1 in function of different external electric potentials. . . . .	35
Figure 3.6	Truncated cylindrical volume elements representing protein and selectivity filter volume elements. . . . .	36
Figure 3.7	Average $\text{Na}^+$ and $\text{K}^+$ current in cylindrical volume element Z2 in function of different external electric potentials. . . . .	36
Figure 3.8	Average $\text{Na}^+$ and $\text{K}^+$ current in cylindrical volume element Z3 in function of different external electric potentials. . . . .	37
Figure 3.9	Voltage dependence of $q_{\text{Lys1244}} = 0.5 e / q_{\text{Lys1244}} = 1.0 e$ conductance ratios for $\text{Na}^+$ and $\text{K}^+$ and $\text{Na}^+/\text{K}^+$ conductance ratios for $q_{\text{Lys1244}} = 0.5 e$ and $q_{\text{Lys1244}} = 1.0 e$ in Z2. . . . .	38
Figure 3.10	Voltage dependence of $q_{\text{Lys1244}} = 0.5 e / q_{\text{Lys1244}} = 1.0 e$ conductance ratios for $\text{Na}^+$ and $\text{K}^+$ and $\text{Na}^+/\text{K}^+$ conductance ratios for $q_{\text{Lys1244}} = 0.5 e$ and $q_{\text{Lys1244}} = 1.0 e$ in Z3. . . . .	39

Figure 3.11	Density profiles of the Center of Geometry (COG) of the terminal selectivity filter DEKA residue sidechains along the $z$ -axis with respect to $\text{COG}_{\text{DEKA}}$ . . . . .	41
Figure 3.12	Position distribution of the COG of the terminal selectivity filter DEKA residue sidechains within the $xy$ -plane with respect to $\text{COG}_{\text{DEKA}}$ . . . . .	43
Figure 3.13	Lys1244 $x$ , $y$ and $z$ densities, $\rho$ , at $q_{\text{Lys1244}} = 1.0 e$ and $U_z = 100 \text{ mV}$ and $\text{Na}^+$ or $\text{K}^+$ within $10 \text{ \AA}$ of $\text{COG}_{\text{Lys1244 } \epsilon\text{-NH}_3^+}$ . . . . .	44
Figure 3.14	Probability densities, $\rho(z)$ , relative binding free energies $G(z)$ and hydration numbers for $\text{Na}^+$ and $\text{K}^+$ ions along the $z$ -axis during simulations with $U = 100 \text{ mV}$ . . . . .	46
Figure 3.15	Top view and side view of binding pockets and selection cylinders. . . . .	49
Figure 3.16	Probability densities, relative binding free energies and hydration numbers for $\text{Na}^+$ and $\text{K}^+$ ions along the $x$ - and $y$ -axes during simulations with $U = 100 \text{ mV}$ and $q_{\text{Lys1244}} = 1.0 e$ . . . . .	50
Figure 3.17	Selection cylinder and residues around binding site P1. . . . .	51
Figure 3.18	Selection cylinder and residues around binding site P2. . . . .	51
Figure 3.19	Selection cylinder and residues around binding sites P3 and P4. . . . .	52
Figure E.1	Probability densities, relative binding free energies and coordination numbers for $\text{Na}^+$ and $\text{K}^+$ ions along the $x$ , $y$ and $z$ -axes during simulations with $U = 0 \text{ mV}$ and $q_{\text{Lys1244}} = 0.5 e$ , $q_{\text{Lys1244}} = 1.0 e$ . . . . .	69
Figure E.2	Probability densities, relative binding free energies and coordination numbers for $\text{Na}^+$ and $\text{K}^+$ ions along the $x$ , $y$ and $z$ -axes during simulations with $U = 100 \text{ mV}$ and $q_{\text{Lys1244}} = 0.5 e$ , $q_{\text{Lys1244}} = 1.0 e$ . . . . .	70
Figure E.3	Probability densities, relative binding free energies and coordination numbers for $\text{Na}^+$ and $\text{K}^+$ ions along the $x$ , $y$ and $z$ -axes during simulations with $U = 300 \text{ mV}$ and $q_{\text{Lys1244}} = 0.5 e$ , $q_{\text{Lys1244}} = 1.0 e$ . . . . .	71

Figure E.4	Probability densities, relative binding free energies and coordination numbers for Na <sup>+</sup> and K <sup>+</sup> ions along the $x$ , $y$ and $z$ -axes during simulations with $U = 500$ mV and $q_{Lys1244} = 0.5 e$ , $q_{Lys1244} = 1.0 e$ . . . . .	72
Figure F.1	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 0.5 e$ and $U_z = 0$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	75
Figure F.2	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 0.5 e$ and $U_z = 100$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	76
Figure F.3	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 0.5 e$ and $U_z = 300$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	77
Figure F.4	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 0.5 e$ and $U_z = 500$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	78
Figure F.5	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 1.0 e$ and $U_z = 0$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	79
Figure F.6	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 1.0 e$ and $U_z = 300$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	80
Figure F.7	Lys1244 $x$ , $y$ and $z$ densities at $q_{Lys1244} = 1.0 e$ and $U_z = 500$ mV and Na <sup>+</sup> or K <sup>+</sup> within 10 Å of COG <sub>Lys1244 ε-NH<sub>3</sub><sup>+</sup></sub> . . . . .	81

## LIST OF TABLES

---

Table 1.1	Intra- and extracellular concentrations of Na <sup>+</sup> and K <sup>+</sup> ions. . . . .	3
Table 3.1	Average relative binding free energies and coordination numbers for Na <sup>+</sup> and K <sup>+</sup> within $z = 8 \pm 2.5$ Å. . . . .	48
Table A.1	Selected residues for the pore models A and D. . . . .	59
Table A.2	Selected residues for the pore model B. . . . .	59
Table A.3	Selected residues for the pore model C. . . . .	59
Table A.4	Numbers of atoms of the different pore models. . . . .	60
Table A.5	Simulation box dimensions of the different pore models. . . . .	60

Table B.1	Constraints and restraints placed on the protein and lipid membrane in system A and on the protein in system D during the equilibration runs. . . . .	61
Table C.1	Distances between the selectivity filter's Asp406 and Lys1244 C $\alpha$ atoms. . . . .	63
Table C.2	Coordinates of the identified binding pockets P1-P4. . . . .	63
Table C.3	Relative z-coordinates of the selectivity filter region. . . . .	64
Table C.4	Relative z-coordinates of the protein. . . . .	64
Table D.1	Average Na $^+$ and K $^+$ currents in the entire simulation box. . . . .	65
Table D.2	Average Na $^+$ and K $^+$ currents in cylinder selection Z1. . . . .	66
Table D.3	Average Na $^+$ and K $^+$ currents in cylinder selection Z2. . . . .	67
Table D.4	Average Na $^+$ and K $^+$ currents in cylinder selection Z3. . . . .	68
Table E.1	Average relative binding free energies and coordination numbers for Na $^+$ and K $^+$ within $z(\text{COG}_{DEKA}) \pm 2.5 \text{ \AA}$ . . . . .	73

## ABBREVIATIONS

---

MD	Molecular Dynamics
RMSD	Root Mean Square Deviation
CHARMM	Chemistry at Harvard Macromolecular Mechanics
NAMD	Nanoscale Molecular Dynamics
VMD	Visual Molecular Dynamics
AMBER	Assisted Model Binding with Energy Refinement
POPC	1-palmitoyl-2-oleoyl- <i>sn</i> -glycero-3-phosphoholine
COG	Center of Geometry
OPM	Operations of Protein in Membranes
PPM	Positioning of Proteins in Membranes
VSD	Voltage-Sensing Domain
DFT	Density-Functional Theory



## THEORETICAL BACKGROUND

---

This chapter aims to outline the physical grounds of ion transport across cellular membranes and provide an introduction to the basics of theoretical and computational methods which have been used to model these phenomena.

### 1.1 ION CHANNELS

Ion channels are molecular pores allowing the displacement of  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$  and  $\text{Cl}^-$  ions across cell plasma membranes which are key events in many biological processes. Ion transport occurs in a passive fashion, meaning that the driving force leading to a flux across the membrane is an electrochemical gradient between the intra- and the extracellular space.

A grouping can be made based on the stimuli which induce a channel's opening or closing. Those can be ligand binding, in ligand-gated ion channels, mechanical deformation, in mechanosensitive channels, or a change in transmembrane potential, in voltage-gated ion channels.

Further categorisation can be established from the channel's discriminatory properties. The pore through which ions pass contains a narrow region, the selectivity filter, which gives rise to a preference for allowing the passage of a certain ion species. During transport, cations are stabilised by negative charges of the selectivity filter residues which compensates for the loss of water molecules in their hydration shell. This is required since the channel pore is often just narrow enough to permit the bare or partially desolvated ion to pass implying the need to shed hydration layers.

Selectivities vary depending on the architecture of the selectivity filter region. [1, 2] Potassium channels typically permit the displacement of one  $\text{Na}^+$  in 1000  $\text{K}^+$  ions whereas sodium channels have a 8:1 preference for  $\text{Na}^+$  over  $\text{K}^+$ . [3] Once activated and open, ion fluxes through  $\text{Na}^+$  and  $\text{K}^+$  channels amount to approximately 2 to 10 pA. This generates a conductance which can be measured or simulated and, therefore, used to quantify selectivity.

### 1.2 BIOLOGICAL ROLE AND FUNCTION OF ION CHANNELS

Voltage-gated sodium ( $\text{Na}_v$ ) channels serve an essential purpose in the propagation and generation of action potentials. At resting state, the

cytoplasm is approximately 80 mV more electrically negative than the extracellular space. This gives rise to a negative resting potential. [2, 4] In the absence of external stimuli, open  $K^+$  channels allow positively charged  $K^+$  ions to flow from the cytoplasm into the extracellular fluid. This movement is driven by a concentration gradient. Negatively charged  $Cl^-$  ions cannot permeate the membrane and, therefore, are bound to remain inside the cell. As the positive charge outside the cell accumulates, the electric force arising from the electric potential difference impedes  $K^+$  displacement and leads to the establishment of an equilibrium potential at which no net ion flux is observed. If a threshold potential of approximately  $-53$  mV is reached, firing occurs.  $Na_v$  channels open and permit the flux of  $Na^+$  from the extracellular space into the cytosol leading to further depolarisation. The generated action potential can travel along excitable cells. Once the potential surpasses 0 mV, inactivation occurs within 1 to 2 ms. This yields a positive overshoot in the transmembrane potential. At around 30 mV, the hyperpolarisation triggers the opening of voltage-gated potassium channels which allows the outward flux of  $K^+$  ions in order to restore the negative resting potential. [4, 5]

### 1.3 ION FLUXES AND CONDUCTANCE

The separation of opposite electric charges gives rise to electrical phenomena.

A net flux of charges is termed a current,  $I$ , and is measured in units of Amperes where one Ampere corresponds to one Coulomb per second. The sign of such a current depends on the direction in which the charges are moving. A positive sign indicates the flow direction of positively charged particles. The magnitude of an electric current is determined by the electric potential difference between two oppositely charged electrodes and the conductance of the solution which separates them.

The potential difference,  $E$ , describes the necessary work to move one charge unit from one point to another and is expressed in units of volts. Conductance,  $g$ , is the inverse to resistance,  $R$ , and specifies the ease with which a charge can travel from one electrode to another. Its unit of measurement is Siemens.

The correlation between current, potential difference and conductance or resistance is given by Ohm's Law:

$$I = gE \quad I = \frac{E}{R} \quad (1.1)$$

In biological systems, due to the presence of a concentration gradient, ion fluxes are also observed at a 0 mV transmembrane potential. The voltage at which net ion fluxes are equal to zero is termed the equilibrium

potential,  $E_A$ . It corresponds to the potential difference between between the intra- and the extracellular space. By convention the extracellular potential,  $E_o$ , is subtracted from the intracellular one,  $E_i$ . The relation between the concentration ratio of a given ionic species,  $A$ , and the corresponding  $E_A$  is given by Nernst's equation where  $F$  and  $R$  are the Faraday and the ideal gas constant and  $q_A$  is the ionic charge, respectively.

$$E_A = \frac{RT}{q_A F} \ln \frac{[A]_o}{[A]_i} \quad (1.2)$$

Table 1.1 shows a list of intra- ( $c_i$ ) and extracellular ( $c_o$ ) concentrations equilibrium potentials for  $\text{Na}^+$  and  $\text{K}^+$  ions in human skeletal muscle cells.

Table 1.1: Intra- and extracellular concentrations of  $\text{Na}^+$  and  $\text{K}^+$  ions in human skeletal cells and the respective equilibrium potentials at 37 °C. [3]

Ion	$c_o$ [mmol L <sup>-1</sup> ]	$c_i$ [mmol L <sup>-1</sup> ]	$E_{eq}$ [mV]
$\text{Na}^+$	145	12	+67
$\text{K}^+$	4	155	-98

To calculate the net flux or conductance of a given ion species,  $A$ , across a biological membrane, the potential difference  $E$  in equation 1.1 can, therefore, be substituted by the net potential  $E - E_A$ .

$$I_A = g_A(E - E_A) \quad (1.3)$$

#### 1.4 STRUCTURE OF HUMAN VOLTAGE-GATED SODIUM CHANNELS

The sodium channel protein consists of a large  $\alpha$  (260 kDa) and smaller  $\beta_1$  (30 to 40 kDa) subunits. The  $\alpha$  subunit is typically made up of around 2000 amino acid residues and can be divided into four homologous domains. Each domain contains six alpha-helical transmembrane segments S1-S6. Residues in segments S1-S4 constitute the channel's Voltage-Sensing Domain (VSD). The transmembrane segments S5 and S6, which are connected through a phosphate-binding loop, form the pore region through which conduction occurs. (see figure 1.1)

Prokaryotic  $\text{Na}_v$  channels are constituted by four identical domains, each of which contains a conserved Glu residue. Together they form the EEE motif which gives rise to selectivity. On the contrary, domains

in eukaryotic channels are different from each other and the selectivity filter contains a DEKA motif (figure 1.2) which is composed of an Asp, a Glu, a Lys and an Ala residue in domains I, II, III and IV, respectively. [6]

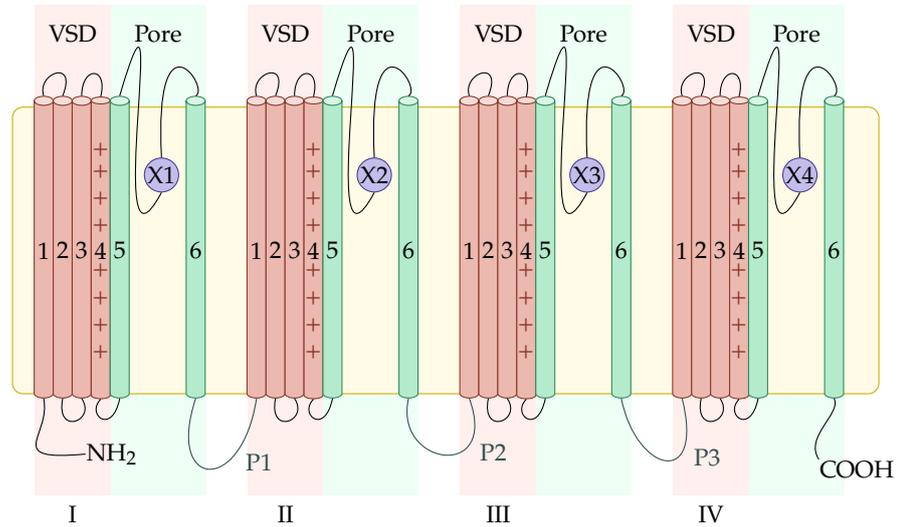


Figure 1.1: Schematic representation of the four protein domains (I-IV) making up the Na<sub>v</sub>-channel α-subunit in lipid membrane (yellow) and connected by phosphate-binding-loops P1-3 (Gray). Each domain contains a Voltage-Sensing Domain (VSD), composed of segments 1–4 (red). In the presence of an action potential, helices in segment 4 induce channel opening conformation changes. Segments 5 and 6 constitute the pore region (green). The selectivity filter residues are indicated by X1–X4 (blue). In the human Na<sub>v</sub> channel these correspond to the Asp406, Glu761, Lys1244, Ala1536 residues.

Na<sub>v</sub>-α proteins are encoded by 9 different genes and expression levels vary across different cell types. [8] The Nav1.4 channel is mostly expressed in skeletal muscle. Other types of Na<sub>v</sub> channels are primarily found in the central (Na<sub>v</sub>1.1, 1.2, 1.3 and 1.6) and the peripheral (Na<sub>v</sub>1.8 and 1.9) nervous system and in heart cells (Na<sub>v</sub>1.5). [4]

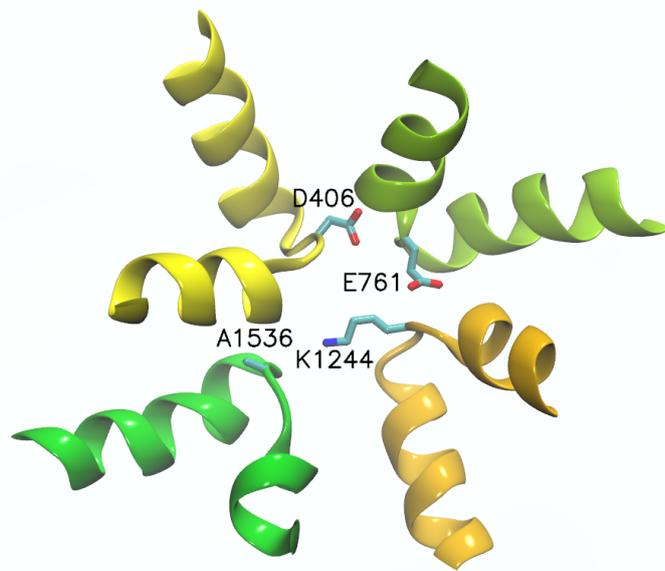


Figure 1.2: Top view on selectivity filter with DEKA residues (Asp406, Glu761, Lys1244, Ala1536) of the human Na<sub>v</sub>1.4 channel. [7]

## 1.5 MOTIVATION

Extensive research has been carried out to investigate dynamics of ion conduction in prokaryotic Na<sub>v</sub> channels. [9–11] Although, the Na<sub>v</sub> conduction mechanism is less understood than for K<sub>v</sub> channels, there exists consensus that Na<sup>+</sup> ions are traversing the channel, at least partially, hydrated and that the selectivity filter is wide enough to accommodate multiple Na<sup>+</sup> ions. Conduction is assumed to require the presence of at least two ions. The attraction of a second ion into the pore region enables a first one, which is coordinated to a binding site, at the channel entrance, to permeate further into the inner region of the protein, and consequently through the channel. Selectivity arises from more favourable interactions of Na<sup>+</sup>, compared to K<sup>+</sup>, with the Glu residues of the planar EEEE motif. [12]

In eukaryotic channels, an interaction mechanism has yet to be elucidated. Ion permeation is assumed to also proceed via a multi-ion pathway and the preference for Na<sup>+</sup> is likely related to the channel's rigid structure at the narrow constriction site between the DEKA motif's Glu, Asp, Lys and Ala residue. [12] Mutation of the Lys group were shown to significantly decrease or even eliminate selectivity. [13] Additionally, electronic effects and the extent of hydration is likely also an important factor for the discrimination between Na<sup>+</sup> and other cations. [14]

The present works aimed to use classical Molecular Dynamics (MD) simulations to shed light on the conduction mechanism in the human  $\text{Na}_v1.4$  channel. Previous research in the group which investigated ion transport across another eukaryotic  $\text{Na}_v$  channel found higher conductance for  $\text{K}^+$  than  $\text{Na}^+$ . In these simulations, interaction potentials were obtained using the classical Chemistry at Harvard Macromolecular Mechanics (CHARMM) force field parametrisation. [15] Therefore, another goal of the project was to determine whether the contradictory findings could be linked to the specific channel model or if the employed force field parametrisation is generally unsuitable for representing atomic interactions involved in ion conduction and selectivity in  $\text{Na}_v$  channels.

### 1.6 STATISTICAL AVERAGES AND THE ERGODIC HYPOTHESIS

MD simulations can be used to describe time evolution at an atomic level. Statistical averaging establishes the connection between the observations which are made at such microscopic scales and corresponding properties in macroscopic systems.

In a macroscopic system, composed of  $W$  states, in equilibrium and at constant temperature  $T$ , the average  $\langle A \rangle$  of a quantity of interest is given by the sum over all of its possible values,  $A_n$ , weighted by their corresponding probabilities,  $\rho_n$ :

$$\langle A \rangle = \sum_{n=1}^W \rho_n A_n \quad (1.4)$$

As  $n \rightarrow \infty$ , this distribution approaches a continuum and can be expressed as an integral in which  $\rho(n)$  is the partition function, given by a Boltzmann distribution which contains the Boltzmann constant,  $k_B$ , and the potential energy of state  $n$ ,  $U(n)$ .

$$\langle A \rangle = \int_0^\infty A(n) \rho(n) dn \quad \rho(n) = \frac{e^{-\frac{U(n)}{k_B T}}}{\int e^{-\frac{U(n)}{k_B T}} dn} \quad (1.5)$$

In statistical mechanics, the here defined quantity,  $\langle A \rangle$ , is termed an ensemble average. In a MD simulation, the time average of observable  $A$  can be approximated as the average of its values at each (equally spaced) time instant over a number of  $M$  time steps.

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M A(t_n) \quad (1.6)$$

As  $M \rightarrow \infty$ , this expression becomes equivalent to taking an integral over all simulation time  $t_{tot}$ :

$$\lim_{M \rightarrow \infty} \bar{A} \equiv \frac{1}{t_{tot}} \int_0^{t_{tot}} A(t') dt' \quad (1.7)$$

The ergodic hypothesis states that, if sampling is performed on a large number of states,  $W$ , and at a sufficiently large number of timesteps,  $M$ , ensemble averaging and time averaging become equivalent methods for computing the average of an observable  $A$  which depends on the coordinates and momenta of a many-particle system and is a function of the system's phase space.

In MD simulations, repeated propagation of positions and velocities of all atoms yields a trajectory which represents the time evolution of the system. The ensemble of all these points,  $i$ , in phase space, can be used to accurately predict a wide range of properties of the system, where each  $i$ , represents a possible state of the system.

## 1.7 EQUATIONS OF MOTION

Classical MD simulations address the displacement of particles due to the sum of all forces,  $\mathbf{F}_i$ , which act on them. Integrating Newton's equations of motion yields a trajectory which describes the time-dependent changes in positions, velocities and accelerations. The connection between acceleration and the force acting on each atom is given by Newton's second law.

$$\mathbf{F}_i = m_i \mathbf{a}_i \quad (1.8)$$

Newton's equations of motion can also be expressed in the more general Hamiltonian formulation which uses generalised coordinates,  $\mathbf{q}_i$ , and momenta,  $\mathbf{p}_i$ , instead of Cartesian coordinates,  $\mathbf{r}_i$ , and velocities,  $\mathbf{v}_i$ .

$$\mathbf{p}_i = m_i \mathbf{v}_i \quad (1.9)$$

Definitions for a particle's kinetic energy,  $K_i$ , and force  $\mathbf{F}_i$  can be rewritten in terms of momentum,  $\mathbf{p}_i$ .

$$K_i = \frac{1}{2} m_i \mathbf{v}_i^2 = \frac{1}{2} \frac{\mathbf{p}_i^2}{m_i} \quad (1.10)$$

$$\mathbf{F}_i = \frac{d\mathbf{p}_i}{dt} \quad (1.11)$$

The first equation of motion is given by the derivative of  $K_i$  with respect to  $\mathbf{p}_i$  which equals the change in position  $\mathbf{q}_i$  in  $dt$ .

$$\frac{d\mathbf{q}_i}{dt} = \frac{dK_i}{d\mathbf{p}_i} \quad (1.12)$$

In an isolated system, force,  $\mathbf{F}_i$ , equals the negative gradient of the potential energy surface.

$$\mathbf{F}_i = -\nabla U \equiv -\frac{\partial U}{\partial \mathbf{q}_i} \quad (1.13)$$

The second Hamilton equation for an isolated system is obtained by combining equations 1.11 and 1.13.

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial U}{\partial \mathbf{q}_i} \quad (1.14)$$

To integrate Newton's equations of motion of an atomic ensemble in cartesian coordinate space, each atom,  $i$ , can be assigned a position,  $\mathbf{r}_i$ , a velocity,  $\mathbf{v}_i$ , and an acceleration vector,  $\mathbf{a}_i$ . The three quantities are related in the following way:

$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt} \quad \text{and} \quad \mathbf{a}_i = \frac{d\mathbf{v}_i}{dt} = \frac{d^2\mathbf{r}_i}{dt^2} \quad (1.15)$$

The time evolution of  $\mathbf{r}_i$  and  $\mathbf{v}_i$  is obtained by integration from starting time  $t = 0$  to end time  $t$ :

$$\mathbf{r}_i(t) = \int_0^t \mathbf{v}_i(t') dt' + \mathbf{r}_i(0) \quad \text{and} \quad \mathbf{v}_i(t) = \int_0^t \mathbf{a}_i(t') dt' + \mathbf{v}_i(0) \quad (1.16)$$

## 1.8 THE VIRIAL EQUATION OF STATE, TEMPERATURE AND PRESSURE

The equation of state establishes the relationship between all thermodynamic properties of a system. The force acting on an atom,  $i$ , is the sum of all forces,  $\mathbf{F}_{ij}$ , arising from the interactions between the atoms present in the system, a force related to pressure,  $\mathbf{F}_p$ , and other external contributions,  $\mathbf{F}_{ext}$ , such as applied electric fields.

$$\mathbf{F}_i = \sum_{i \neq j}^N \mathbf{F}_{ij} + \mathbf{F}_p + \mathbf{F}_{ext} \quad (1.17)$$

The Virial equation of state can be derived starting from the time derivative of  $\mathbf{r}_i \cdot \mathbf{p}_i$ . Using the relations in equation 1.15 and 1.9, the following expression can be obtained:

$$\frac{d}{dt} (\mathbf{r}_i \cdot \mathbf{p}_i) = m_i \frac{d}{dt} \left( \mathbf{r}_i \frac{d\mathbf{r}_i}{dt} \right) = \underbrace{\mathbf{r}_i m_i \frac{d^2\mathbf{r}_i}{dt^2}}_{m_i \mathbf{a}_i = \mathbf{F}_i} + \underbrace{m_i \left( \frac{d\mathbf{r}_i}{dt} \right)^2}_{m_i \mathbf{v}_i^2} \quad (1.18)$$

The second term contains a force,  $\mathbf{F}_i$ , and the last term is equal to a kinetic energy contribution,  $K$ .

$$2 \left( \frac{1}{2} m_i \mathbf{v}_i^2 \right) = 2K \quad (1.19)$$

Equation 1.18 can be extended to a sum over all particles in a system.

$$\sum_{i=1}^N m_i \frac{d}{dt} \left( \mathbf{r}_i \frac{d\mathbf{r}_i}{dt} \right) = \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_i + 2 \sum_{i=1}^N K_i \quad (1.20)$$

In an ergodic system, the ensemble average term is equal to the sum of the averages of each of the individual sums. In an isolated system, external forces are absent which means that the only forces to consider are those originating from atomic interactions and pressure. Fluctuations in particle velocities and positions are random and, therefore, symmetrically distributed around the respective average value. Likewise, the time derivative of their product follows a random distribution around an average value which equals zero.

$$\underbrace{\left\langle \sum_{i=1}^N m_i \frac{d}{dt} \left( \mathbf{r}_i \frac{d\mathbf{r}_i}{dt} \right) \right\rangle}_{=0} = \underbrace{\left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_i \right\rangle}_W + 2 \langle K \rangle \quad (1.21)$$

The first term on the right side of the equation is termed the virial,  $W$ , of the system.

$$0 = \langle W \rangle + 2 \langle K \rangle \quad (1.22)$$

In a fully isolated system, the total energy is constant whereas quantities such as pressure and temperature may fluctuate.

**KINETIC ENERGY AND TEMPERATURE** The temperature,  $T$ , of a particle system, is defined by its average kinetic energy,  $\langle K \rangle$ , which is in turn determined by the average particle velocities,  $\mathbf{v}_i$ :

$$\langle K \rangle = \left\langle \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right\rangle = \frac{3}{2} N k_B T \quad (1.23)$$

The above equation states, at equilibrium, in an  $N$  atomic system, where each of the atoms has three degrees of freedom, each degree of freedom contributes  $1/2 k_B T$  to the kinetic energy ( $k_B$  is the Boltzmann constant). Equation 1.23 is the mathematical expression of the equipartition principle in statistical mechanics.

**THE VIRIAL AND PRESSURE** As given in equation 1.17, for an isolated system, the total force acting on its particles is constituted by the sum of a pressure force,  $\mathbf{F}_p$ , and the sum of all atom-atom interaction contributions,  $\mathbf{F}_{int}$ :

$$\langle W \rangle = \left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_i \right\rangle = \left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_p \right\rangle + \left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_{int} \right\rangle \quad (1.24)$$

Pressure is applied at the surface of a system, this means that for bulk atoms this force contribution term will be equal to zero. In the case of a cubic simulation box, the force acting on each face of the cube is proportional to the pressure,  $p$ , and the cube's surface area,  $L^2$ . The total pressure force is proportional to the cubic volume  $V$ .

$$F = pL^2 \quad \text{and} \quad \left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_p \right\rangle = -3pL^3 = -3pV \quad (1.25)$$

With these relations established, the equation of state (equation 1.23) can be reformulated such that it expresses the connection between a system's number of particles,  $N$ , temperature,  $T$ , pressure,  $p$ , and volume,  $V$ , as well its internal forces.

$$3Nk_B T + \left\langle \sum_{i=1}^N \mathbf{r}_i \mathbf{F}_{int} \right\rangle + 3pV = 0 \quad (1.26)$$

Using the relation between force  $\mathbf{F}$  and potential energy  $U$ , given in equation 1.13, the virial equation of state can also be expressed in

terms of potential energy attributed to particle interactions. In a non-interacting system, this contribution is equal to zero and equation 1.27 becomes to the ideal gas law.

$$\frac{pV}{k_B NT} = -\frac{1}{3k_B NT} \left[ 1 + \left\langle \sum_{i=1}^N \mathbf{r}_i \left( \frac{-\partial U_{int}}{\partial \mathbf{r}_i} \right) \right\rangle \right] \quad (1.27)$$

In a  $N$ -particle system, the interaction potential  $U_{int}$ , which depends on all  $\{\mathbf{r}_i\}$  can be split expanded into a series of  $x$ -body interaction contributions, where  $2 \leq x \leq N$ .

$$U_{int}(r_1, \dots, r_N) = \sum_{i<j}^N U_{ij}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i<j<k}^N U_{ijk}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (1.28)$$

Inserting equation 1.28 into 1.27, and separating the sum inside the average by interaction order, yields the virial equation with corresponding (volume) virial coefficients ( $B, C, \dots$ ) which specify the connection between the macroscopic properties of pressure, volume and temperature and interaction potentials arising from  $x$ -body interactions.

$$\frac{pV}{Nk_B T} = 1 + \frac{NB}{V} + \frac{N^2 C}{V^2} + \dots \quad (1.29)$$

At increasing order of  $x$ , the relative contributions of the interaction terms to the mean virial become increasingly negligible and it is common practise to truncate the expansion (usually at  $x = 2$ ) to approximate to the total force related to all particle interactions.

## 1.9 STEPS IN AN MOLECULAR DYNAMICS SIMULATION CYCLE

The aim of an MD simulation is the study of dynamic behaviour at the molecular level. More specifically, atomic positions,  $\{\mathbf{r}(t)\}$ , are repeatedly propagated from an initial time  $t$  to time  $t + \Delta t$ . Here,  $\Delta t$  corresponds to a very short time interval with respect to the timescale of motion inside the system under investigation.

This procedure requires the integration of Newton's equations of motion which, in practise, is done using a numerical method, a so-called integrator. Examples of integrators are the Verlet algorithms such as the Simple, Leap-Frog or Velocity Verlet or Runge-Kutta and Gear-Predictor correctors. The Velocity Verlet algorithm provides good stability and time-reversal properties and is also the integrator implemented in the Molecular Dynamics program NAMD which was used in for master thesis. [16]

**VELOCITY-VERLET ALGORITHM** To propagate atomic velocities,  $\mathbf{v}_i(t)$ , and positions,  $\mathbf{r}_i(t)$ , Nanoscale Molecular Dynamics (NAMD) uses the following implementation of the Velocity-Verlet integration method. The algorithm first computes  $\mathbf{v}_i$  at time  $t + \Delta t/2$  through a first order expansion of  $\mathbf{v}_i(t)$  around the initial time  $t$ . According to equation 1.17, accelerations at timestep  $t$  are evaluated from the corresponding forces  $F(\mathbf{r}_i(t))$ .

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}_i(t) + \frac{\Delta t \mathbf{a}_i(t)}{2} \quad (1.30)$$

The resulting velocities are then used to calculate the positions  $\mathbf{r}_i$  at  $t + \Delta t$ .

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) \quad (1.31)$$

$$= \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2 \mathbf{a}_i}{2} \quad (1.32)$$

These new positions are consecutively used to update  $\mathbf{a}_i(t + \Delta t)$  and propagate  $\mathbf{v}_i$ .

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) + \frac{\Delta t \mathbf{a}_i(t + \Delta t)}{2} \quad (1.33)$$

$$= \mathbf{v}_i(t) + \frac{\Delta t [\mathbf{a}_i(t) + \mathbf{a}_i(t + \Delta t)]}{2} \quad (1.34)$$

Figure 1.3 illustrates the five main steps involved in one molecular dynamics iteration which propagates  $\{\mathbf{r}_i(t)\}$  and  $\{\mathbf{v}_i(t)\}$  from  $t$  to  $t + \Delta t$ . In a MD simulation, this cycle is performed  $M$  times to yield a trajectory which describes a system's time evolution from  $t = 0$  to  $t = t_{tot}$ .

1. The calculation of  $\mathbf{r}(t + \Delta t)$  requires a set of initial positions,  $\{\mathbf{r}(t)\}$ , and velocities,  $\{\mathbf{v}(t)\}$ . From those, internal coordinates,  $\{r_j\}$ , such as bond lengths, bond angles or dihedral angles can be determined.

At the beginning of a simulation where  $t = 0$ , an input of starting coordinates must be supplied. These can be obtained from databases of experimental data or from quantum mechanical calculations. If not specified, initial velocities are chosen using probabilities,  $\rho(\mathbf{v}_i)$ , from a Boltzmann distribution of velocities,  $\mathbf{v}_i$ , at the chosen simulation temperature,  $T$ . Values are set such that the total momentum,  $\mathbf{p}$ , equals zero.

$$\rho(\mathbf{v}_i) = \left(\frac{m_i}{2\pi k_B T}\right) e^{-\frac{m_i \mathbf{v}_i^2}{2k_B T}} \quad (1.35)$$

$$\mathbf{p} = \sum_i^N m_i \mathbf{v}_i^2 \quad (1.36)$$

2. Classical MD simulations use so-called force field parameters to compute as a set of interaction potential terms whose sum yields the system's potential energy,  $U$ . Once the functional form of a potential energy surface is known, a corresponding force  $F$  can be calculated.
3. The magnitude of such a force is given by the magnitude of the potential energy surface gradient at a given  $r$  (eg. 1.13). If  $U$  only depends on one given internal coordinate,  $r$ , the resulting force can be calculated as the derivative with respect to this  $r$ .

$$F = -\frac{dU(r)}{dr} \quad (1.37)$$

4. If the forces,  $\{F_i\}$ , acting on each particle are known, the corresponding accelerations,  $\{\mathbf{a}(t)\}$ , can be computed (see equation 1.8).
5. Once  $\{\mathbf{a}(t)\}$  have been evaluated, an integrator is used to update atomic positions and velocities.

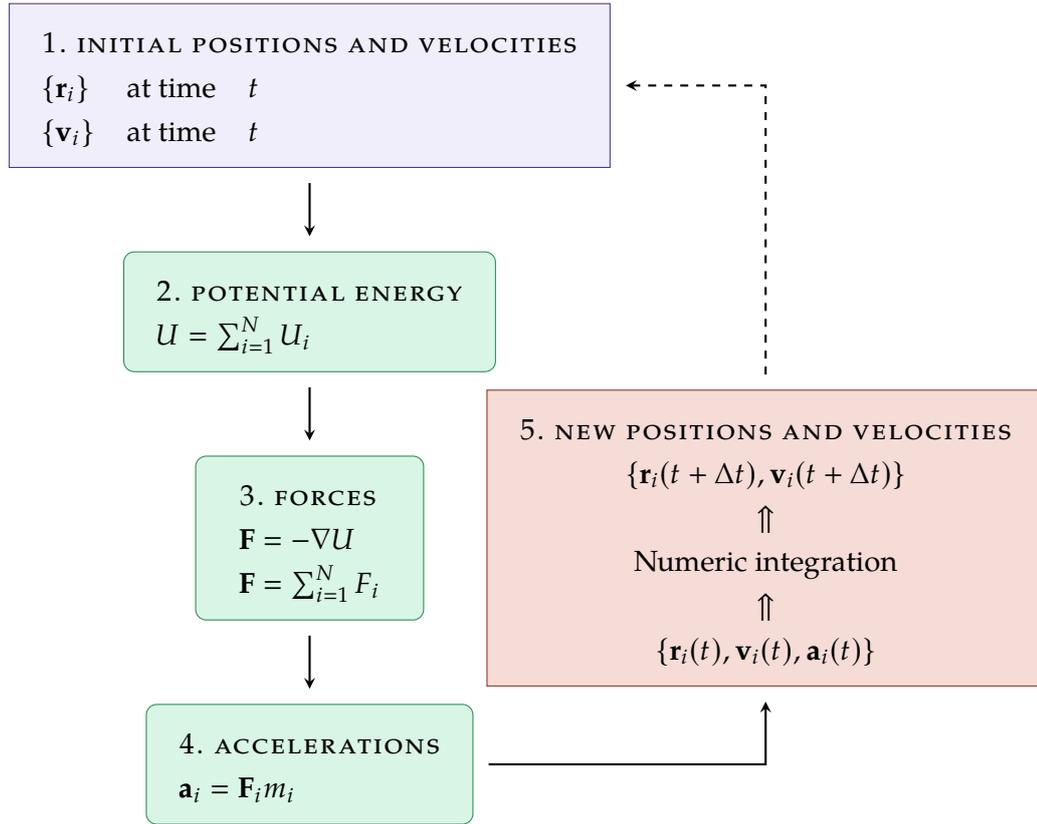


Figure 1.3: Steps in a Molecular Dynamics simulation.

## 1.10 PERIODIC BOUNDARY CONDITIONS

The upper limit for the maximum number of atoms which can be treated in an MD simulation lies between  $10^5$  and  $10^6$ . In a cubic simulation box, the ratio between surface atoms  $N_s$  and the total atom number,  $N$ , is determined by the box length,  $L$ .

$$N \propto L^3 \quad N_s \propto 6L^2 \quad \Rightarrow \quad N_s \propto 6N^{\frac{2}{3}} \quad (1.38)$$

In a system composed of  $10^6$  atoms, 6% of these atoms are at the surface. Real physical systems are made up of far more particles. For example, one mole of atoms contained in a cubic box contains only around  $7 \times 10^{-6}$  % of surface atoms. In the first case, the contribution to bulk properties originating from the interactions occurring at the simulation cell's surface is not negligible. To simulate bulk behaviour, an infinite system can be constructed by applying periodic boundary conditions. In this case, the original simulation box is reproduced in an infinite three-dimensional array. The behaviour of each image atom corresponds to that in the original box, identified by the cell vector

$\mathbf{n}(0,0,0)$ . Interactions are not only considered between all particles within the primary cell but also between periodic images separated by an integer multiple, specified by  $\mathbf{n}$ , of the box length  $L$ . The prime in the first sum in equation 1.39 indicates that interactions between atom  $i$  with itself are excluded in the original cell.

$$U = \sum_{\mathbf{n}}' \sum_{i=1}^N \sum_{j=1}^N U(|\mathbf{r}_{ij} + \mathbf{n}L|) \quad (1.39)$$

To avoid the computation of infinite sums, in practical applications, short range-interactions are only computed within a specified cutoff radius  $r_c$  beyond which their contributions are assumed to become negligible. Long range interactions cannot be treated in the same way. Their computation is therefore relies on methods which enable the approximate calculation of infinite sums. The evaluation of the long-range electrostatic potentials is mostly performed using the Ewald summation approach which is explained below. [17]

### 1.11 TEMPERATURE CONTROL IN THE ISOTHERMAL ENSEMBLE

Sections 1.7 and 1.8 considered an isoenergetic ensemble with constant particle number,  $N$ , volume,  $V$  and total energy  $E$ . At constant  $E$ , the probabilities,  $\rho_i$  of all states,  $i = \{1, \dots, N\}$ , are equal. [18]

$$\rho(i) = \frac{1}{N} \quad (1.40)$$

To represent realistic experimental conditions it is often more convenient to perform simulations at constant temperature,  $T$ , rather than total energy. In a closed system with constant volume,  $V$ , and temperature  $T$ , according to the canonical distribution function, lower energy states are more likely than high energy ones.

$$\rho(i) = \frac{e^{-\frac{E_i}{k_B T}}}{\sum_{i=1}^N e^{-\frac{E_i}{k_B T}}} \quad (1.41)$$

In an  $NVT$  simulation, temperature is readjusted in chosen intervals of timesteps  $\Delta t$ .

A constant temperature can be maintained by rescaling particle velocities inside the system (equation 1.23) or by allowing an energy exchange between the system and an external thermostat or heat bath which leads to a fluctuating total energy.

The Andersen method, for example, simulates random collisions with particle walls reassigns the velocities of a fraction of particles to

values chosen from a Boltzmann distribution of velocities for the target temperature. The number of particles to which rescaling is applied is governed by the collision frequency parameter  $f$ . [19]

Another approach to temperature regulation proceeds via adjustment of of particle velocities by a scaling factor  $\lambda$ . The velocity rescaling method [20, 21] and the Berendsen barostat [22] are based on this principle. In the simpler velocity rescaling method,  $\lambda_{VR}$  is determined by the deviation of the actual temperature  $T(t)$  from the target temperature  $T_{set}$ . [20, 21] The Berendsen Algorithm,  $\lambda_B$  contains an additional coupling parameter  $\Delta t/\tau$ .  $\tau$  which is termed 'rise time' and governs the coupling strength between the heat bath and the system. As  $\tau \rightarrow \infty$  the coupling is removed. [22]

$$\mathbf{v}_i^{new} = \lambda \mathbf{v}_i^{old} \quad (1.42)$$

$$\lambda_{VR} = \left( \frac{T_{set}}{T(t)} \right)^{\frac{1}{2}} \quad \lambda_B = \left[ 1 + \frac{\Delta t}{\tau} \left( \frac{T_{set}}{T(t)} - 1 \right) \right]^{\frac{1}{2}} \quad (1.43)$$

In the present simulations temperature was maintained using the Langevin thermostat which proceeds by extending the equation of motion for momenta (1.14) by two additional force contributions, a friction coefficient,  $\gamma$ , and a random force,  $R(t)$ . These terms are representative of the system's viscosity and random collisions between particles. The random force is selected from a Gaussian distribution with mean value and average equal to zero.[23]

$$\frac{d\mathbf{p}_i}{dt} = \frac{\partial U}{\partial \mathbf{q}_i} - \gamma \mathbf{p}_i + R(t) \quad (1.44)$$

### 1.12 PRESSURE CONTROL AND THE ISOTHERMAL-ISOBARIC ENSEMBLE

As mentioned in the previous section, in order to represent experimental conditions, it is often preferable to choose a simulation ensemble at constant temperature and pressure, rather than total energy and volume. In the isothermal-isobaric ensemble, the probabilities for finding a state with set  $p$  and  $T$  are given by:

$$\rho(i) = \frac{e^{-\frac{E_i + pV_i}{k_B T}}}{\sum_{i=1}^N e^{-\frac{E_i + pV_i}{k_B T}}} \quad (1.45)$$

This means that the probability for finding the system in a state at pressure,  $p$ , depends on the simulation cell volume. The relation between

a system's pressure,  $p$ , and its volume.  $V$ , is given by the virial equation of state (equation 1.29).

In addition to a barostat which adjusts the cell volume, temperature can be modulated by a thermostat. [18] In our simulations pressure was controlled with a Langevin Nosé-Hoover algorithm [24, 25]. This method can be viewed as an extension of the Andersen barostat [19] in which constant pressure of a closed system is maintained, by connecting it to a piston of mass,  $M$ , with an external pressure,  $\alpha$ , acting on it. This leads to uniform expansion or contraction of the simulation box volume. As the piston's mass,  $M \rightarrow \infty$ ,  $dV/dt \rightarrow 0$ .  $\alpha$  corresponds to the target pressure of the system. For this method, the equations of motion are formulated as follows:

$$\frac{d\mathbf{q}_i}{dt} = \frac{\partial K}{\partial \mathbf{p}_i} + \frac{1}{3}\mathbf{q}_i \frac{d \ln V}{dt} \quad (1.46)$$

$$\frac{d\mathbf{p}_i}{dt} = \frac{\partial U}{\partial \mathbf{q}_i} - \frac{1}{3}\mathbf{p}_i \frac{d \ln V}{dt} \quad (1.47)$$

$$\frac{d^2V}{dt^2} = \frac{1}{M} \left[ \underbrace{\frac{\left( \frac{2}{3} \sum_{i=1}^N K_i - \frac{1}{3} \sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{r}_i \right)}{V}}_{=p(t)} - \underbrace{\alpha}_{=p_{set}} \right] \quad (1.48)$$

$$\frac{d^2V}{dt^2} = \frac{1}{M} (p(t) - p_{set}) \quad (1.49)$$

The Langevin Barostat method adds a collision frequency  $\gamma$ , and a random force  $R(t)$ , to equation 1.48 which determines fluctuations in the simulation box volume. Like in the Langevin thermostat, the random force is selected from a Gaussian distribution with mean value and average equal to zero. [25]

$$\frac{d^2V}{dt^2} = \frac{1}{M} \left( p(t) - p_{set} + \gamma \frac{dV}{dt} + R(t) \right) \quad (1.50)$$

### 1.13 FORCE FIELDS

The Born-Oppenheimer approximation allows the separation of nuclear and electronic degrees of freedom. Classical potential models are only explicitly taking into account the nuclear term and handling electronic contributions in an approximate fashion. A molecular mechanics force field contains the necessary parameters to calculate interatomic interaction energies. Their values are either fitted to experimental data or quantum mechanical calculations. Each atom is assigned an atom type which does not only depend on the atomic number but also on its chemical environment. Therefore, a carbon atom inside an aromatic

ring is of a different atom type than a carbonyl carbon. In simulations of biological systems, the CHARMM or the Assisted Model Binding with Energy Refinement (AMBER) force field are commonly used. The total potential energy  $U$  is usually decomposed into bonded  $U_b$  and non-bonded  $U_{nb}$  contributions.

$$U = \sum_{i=1}^N U_{i,b} + \sum_{i=1}^N U_{i,nb}$$

### 1.13.1 Bonded Interactions

The sum of interactions between covalently bonded atoms can be further split into a sum of potentials representing 2-, 3-, and 4-body interactions, specifically bond  $U_{bond}$ , angle  $U_{angle}$  and dihedral angle  $U_{dihedral}$  terms.

$$\sum_{i=1}^N U_{i,b} = \sum_{i>j}^N U_{ij,bond} + \sum_{i>j>k}^N U_{ijk,angle} + \sum_{i>j>k>l}^N U_{ijkl,dihedral} \quad (1.51)$$

**BOND TERMS** The bond stretching contribution is approximated as a harmonic oscillator.  $k_b$  is the respective spring constant,  $R_{ij}$  is the distance between a pair of atoms  $i$  and  $j$ ,  $R_0$  corresponds to the respective equilibrium distance.

$$U_{ij,bond} = k_b(R_{ij} - R_0)^2 \quad (1.52)$$

**ANGLE TERMS** The angular bending term  $U_{angle}$  between three atoms  $i$ ,  $j$  and  $k$ , is likewise assumed to have harmonic character.  $\theta$  gives the angle between vectors  $\mathbf{r}_{ij}$  and  $\mathbf{r}_{jk}$ ,  $\theta_0$  is the respective equilibrium angle. The second Urey-Bradley term describes a spring between the two not directly bonded atoms  $i$  and  $k$  at a separation distance of  $R_{ik}$  and the equilibrium distance  $R_{ub}$  and spring constant  $k_{ub}$ .

$$U_{ijk,angle} = k_\theta(\theta - \theta_0)^2 + k_{ub}(R_{ik} - R_{ub})^2 \quad (1.53)$$

**DIHEDRAL TERMS** The dihedral or torsional interactions are calculated over all atom pairs belonging to three consecutive bonds which form a dihedral angle. The interactions are calculated using an angular harmonic potential between the two planes formed by atoms  $i, j, k$  and  $j, k, l$  where  $\psi$  the angle enclosed by the planes. In case of periodic dihedral potentials, the integer  $a > 0$  describes the order of periodicity. In this case,  $\varphi$  stands for the phase shift

angle and  $k_d$  serves as a multiplicative constant. If  $a = 0$ ,  $\varphi$  represents the equilibrium angle and  $k_d$  the corresponding spring constant.

If  $a \geq 0$ :

$$U_{ijkl, \text{dihedral}} = k_d (1 - \cos(a \psi + \varphi)) \quad (1.54)$$

If  $a = 0$ :

$$U_{ijkl, \text{dihedral}} = k_d (\psi + \varphi)^2 \quad (1.55)$$

### 1.13.2 Non-Bonded Interactions

In a simple parametrisation of interaction parameters, the sum of non-bonded potential can be split into long-range electrostatic  $U_{Coulomb}$  interactions and short-range van-der-Waals'  $U_{vdW}$  and repulsive  $U_{rep}$  interactions.

$$\sum_{i=1}^N U_{i, nb} = \sum_{i>j}^N U_{ij, Coulomb} + \sum_{i>j}^N U_{ij, vdW} + \sum_{i>j}^N U_{ij, rep} \quad (1.56)$$

**ELECTROSTATIC INTERACTIONS** Coulomb interactions are defined as the electrostatic interactions between the point charges  $q$  associated to two atoms  $i$  and  $j$ .  $\epsilon$  is the dielectric constant which accounts for screening by the solvent.

$$U_{ij, Coulomb} = \frac{1}{4\pi\epsilon} \frac{q_i q_j}{R_{ij}} \quad (1.57)$$

This energy term is attractive for opposite and negative for equal charges. The electrostatic interaction contribution decreases at larger separations  $R_{ij}$  but, being a long-range interaction, its contribution does not become negligible even for atoms which are far apart. This means that in a periodic system an atom is also interacting with its own replicate images in the adjacent simulation boxes. To account for this additional interaction term, an additional summation must be introduced for obtaining the total electrostatic energy of the system. Interactions between atom  $i$  with itself are excluded in the original cell  $\mathbf{n}(0, 0, 0)$ .

$$U_{Coulomb} = \frac{1}{2} \sum_{\mathbf{n}}' \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{R_{ij} + \mathbf{n}} \quad (1.58)$$

**EWALD SUMMATION** The result of the above expression depends on summation order and displays slow convergence. Therefore, instead of directly implementing equation 1.58, electrostatic interactions are calculated via the Ewald summation method. In this procedure, a Gaussian charge distribution of width  $\alpha$  and opposite sign but same magnitude is placed on each point charge  $q_i$ . Electrostatic energies can be computed from the interactions of  $q_i$  with the Gaussian. In this way, the long-range interactions assume a short-range character which facilitates their computation. To compensate for the introduced charge distributions, another set of Gaussian functions with opposite sign to the first ones is added to the calculation. This way,  $U_{Coulomb}$  can be decomposed into three terms.

$$U_{Coulomb} = U_{direct} + U_{reciprocal} + U_{correction} \quad (1.59)$$

where

$$U_{direct} = \frac{1}{2} \sum_{\mathbf{n}}' \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{R_{ij} + |\mathbf{n}|} \operatorname{erfc}(\alpha R_{ij} + |\mathbf{n}|) \quad (1.60)$$

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (1.61)$$

$$U_{reciprocal} = \frac{1}{2\pi U} \sum_{i=1}^N \sum_{j=1}^N q_i q_j \sum_{\mathbf{m} \neq 0} \frac{e^{-(\pi \mathbf{m}/\alpha)^2 + 2\pi \mathbf{m} R_{ij}}}{\mathbf{m}^2} \quad (1.62)$$

$$U_{correction} = -\frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^N q_i^2 \quad (1.63)$$

The additional Gaussian error function term 1.61 makes equation 1.60 decay rapidly as  $\mathbf{n} \rightarrow \infty$ . The reciprocal space vector  $\mathbf{m}$  is the Fourier transform of the real space cell vector  $\mathbf{n}$ . The correction term  $U_{correction}$  accounts for the width of the Gaussian functions distributions which are employed. [17]

**VAN-DER-WAALS' AND REPULSIVE INTERACTIONS** The attractive van der Waals' potentials  $U_{vdW}$  originates from the interaction of fluctuating instantaneous instantaneous and permanent dipole moments.

$$U_{vdW} \propto -R_{ij}^{-6} \quad (1.64)$$

A positive repulsion energy term  $U_{rep}$  arises from the repulsion between overlapping electronic orbitals at very short separations and approaches

$\infty$  as  $d \rightarrow 0$ . This contribution cannot be described exactly without the use of quantum mechanics.

$$U_{rep} \propto e^{R_{ij}} \approx R_{ij}^{-12} \quad (1.65)$$

Unlike the long-range electrostatic interactions, van-der-Waals' and repulsion energy contributions quickly become negligible as interatomic distances increase. In practise, interaction energies for atoms separated by more than a chosen cut-off radius  $R_{cut}$  are assumed to be zero and, therefore, not computed anymore. To avoid discontinuities in the potential energy surface, so called switching functions can be employed for smoothing the transition. [17]

In MD simulations, short-range interactions are usually obtained using model potentials which combine the above two energy terms. A widely-employed model is the Lennard-Jones potential  $U_{LJ}$  illustrated in figure 1.4. At short distances interaction energies are dominated by the repulsive term and as distances increase  $U_{LJ} \rightarrow 0$ . At intermediate separation the attractive van-der-Waals' term dominates yielding a minimum in the potential energy curve. Parameters  $\varepsilon$  and  $\sigma$  describe the energy of this minimum and the distance at which the two terms cancel out. For each pair of atoms  $\varepsilon$  and  $\sigma$  can be calculated as a combination of both atoms' respective parameters. [17]

$$U_{LJ} = \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] \quad (1.66)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \cdot \varepsilon_j} \quad \sigma_{ij} = \frac{\sigma_i}{2} + \frac{\sigma_j}{2}$$

#### 1.14 POLARISABILITY AND HIGHER ORDER ELECTROSTATIC INTERACTIONS

With traditional pairwise-additive potentials, electrostatic interactions are usually computed using the Coulomb-potential model (equation 1.57) which approximates electron clouds as fixed point charges centered in the atomic nuclei. All other non-bonded effects are implicitly included though an empirical potential model such as the Lennard-Jones potential (equation 1.66).

A real-case charge distribution, however, is neither spherical nor constrained to the atomic centers. The presence of increasingly complex interactions, as they are often present in biological systems, quickly make this approach become too simple to accurately depict the dynamics inside the system under investigation. A more realistic description of

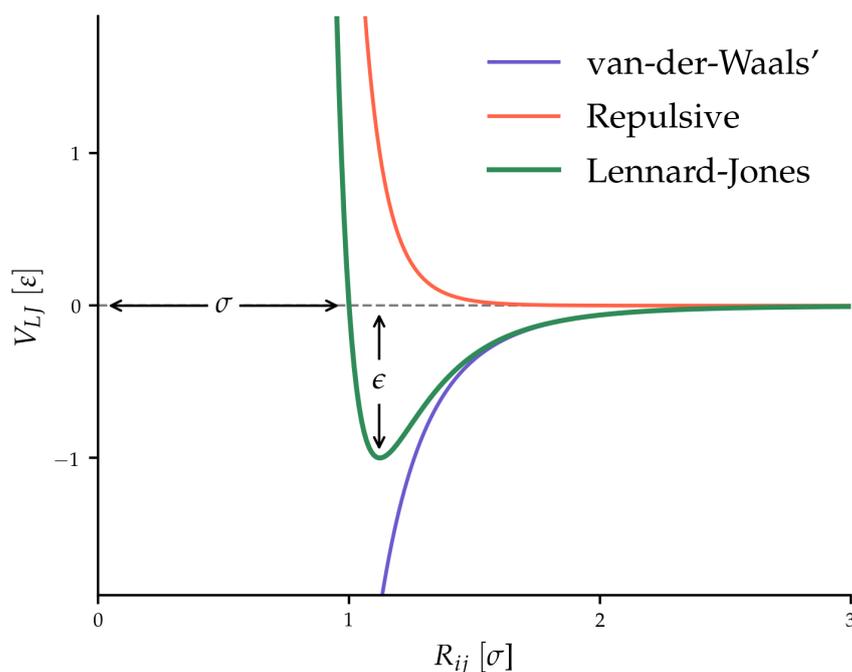


Figure 1.4: Van-der-Waals', repulsive, and Lennard-Jones potential energy functions expressed in units of minimum energy,  $\epsilon$ , and equilibrium distance,  $\sigma$ .

intermolecular interactions can be achieved through explicit account for multipole moments and polarisability effects. The high computational cost of their implementation limits the scope of such methods.

Higher order electrostatic moments can be calculated by extending the previously mentioned Ewald method (p. 20). This can be done in the fashion of a Taylor series. The multipole expansion involves different order electric moments where the zeroth, first, second and third correspond to the monopolar point charge, dipole, quadrupole and octopole moments, respectively. The distributed multipole description yields an accurate description at interatomic distances as long as there is no overlap in charge density. At small separations, a correction for penetration needs to be applied. [26]

Another phenomenon which is often not sufficiently well-represented in traditional pairwise-additive potential models is polarisation. Polarisation effects refer to an anisotropic redistribution of electron density in response to external electric fields. A thereby induced dipole moment creates its own, additional electric field which interacts with the environment. Therefore, despite being fully expressible in classical terms, induction effects present a non-additive term associated to many-body interactions, which makes them complicated to evaluate. Classical

force fields which do not explicitly account for polarisability often overestimate molecular dipole moments.

#### 1.14.1 *Polarisable Force Fields*

If polarisation effects need to be accounted for, a polarisable force field model can be employed. This adds an additional term to the non-bonded interactions,  $U_{nb}$  (equation 1.56).

$$\sum_{i=1}^N U_{i, nb} = \sum_{i>j}^N U_{ij, elst} + \sum_{i>j}^N U_{ij, ind} + \sum_{i>j}^N U_{ij, vdW} + \sum_{i>j}^N U_{ij, rep} \quad (1.67)$$

Examples for polarisable interaction models are the induced dipole model and the Drude oscillator.

In both cases, the total electrostatic energy  $U_{elst}$  is obtained as a sum of the Coulomb energy  $U_{Coulomb}$  and an energy term  $U_{disp}$  which represents the work required for inducing a displacement in the charge distribution.

The total electrostatic energy  $U_{elst}$  is obtained as a sum of the Coulomb energy  $U_{Coulomb}$  and an energy term  $U_{disp}$  which represents the work required for inducing a displacement in the charge distribution.

$$U_{elst} = U_{disp} + U_{Coulomb} \quad (1.68)$$

In NAMD [16], polarisable interactions can be included using the Drude polarisable CHARMM force field [27]. The Drude Oscillator model considers each charge as a pair of point charges which are connected by a harmonic spring with spring constant  $k_i$ . One of the charges is centered at the atom and represents the nuclear charge. The movement of the second so-called Drude-particle, situated at the relative location  $\mathbf{R}_i$ , is unconstrained in all directions and accounts for the electronic charge. The total atomic charge is given by the sum of the two partial charges.

$$U_{disp}^D = \sum_i^N \frac{1}{2} \frac{\mathbf{k}_i^2}{\mathbf{R}_i} \quad (1.69)$$



## COMPUTATIONAL DETAILS

---

This chapter will present how the previously explained methods of computational and theoretical chemistry were used to study and characterise ion transport in the human  $\text{Na}_v1.4$  channel. Details will be given on the performed classical MD simulations and trajectory processing, as well as, strategies and computational approaches used in constructing pore models.

### 2.1 MODEL CONSTRUCTION

A large pore model of the  $\text{Na}_v1.4$  voltage-gated sodium channel's pore region, which will be called A, was constructed in Visual Molecular Dynamics (VMD) [28] based on the cryo-electron microscopy structure file of the channel in complex with its  $\beta 1$  subunit (figure 2.1; PDB ID: 6AGF) [7]. The system comprised subunits S5 and S6 of all four domains (figure 1.1).

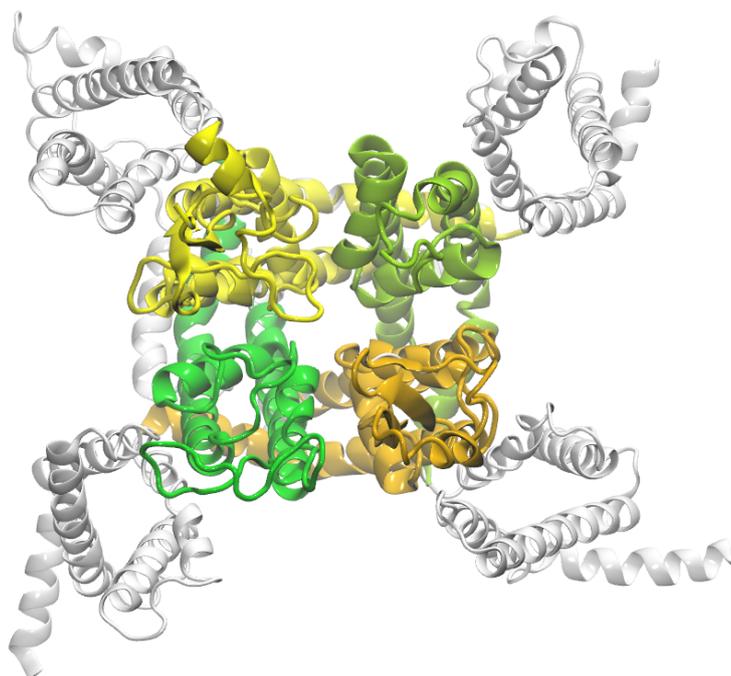


Figure 2.1: Top view on cryo-electron microscopy structure of the human  $\text{Na}_v1.4$  channel's  $\alpha$  subunit. Chains which were truncated in the construction of model a are shown in white. [7]

The protein was first correctly aligned along the z-axis using the Operations of Protein in Membranes (OPM) database and Positioning of Proteins in Membranes (PPM) web server [29]. After amidating and acetylating N- and C- end groups, the structure was placed inside a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipid bilayer and solvated in a rectangular box with aqueous  $0.75 \text{ mol L}^{-1}$  NaCl and KCl (figure 2.2A). This was done, using the CHARMM-GUI Bilayer Builder [30–33].

In addition, three different truncated models (B, C, D) of the pore region were created in VMD (see appendix A). In all cases, geometries were defined based on the coordinates of the large model (A;  $10 \times 10^5$  atoms) after equilibration (see below). N- and C- end groups were amidated and acetylated and systems were solvated in aqueous NaCl and KCl ( $0.75 \text{ mol L}^{-1}$  each). The first model (B;  $10 \times 10^4$  atoms, figure 2.2B) included the selectivity filter loop as well as the adjacent  $\alpha$ -helices on both sides. The second one (C;  $10 \times 10^4$  atoms, figure 2.2C) additionally contained parts of the two helices lying directly below those adjacent to the selectivity filter. A third model (D;  $10 \times 10^4$  atoms, figure 2.2D) consisted of the same amino acid residues as model A but did not contain a lipid membrane.

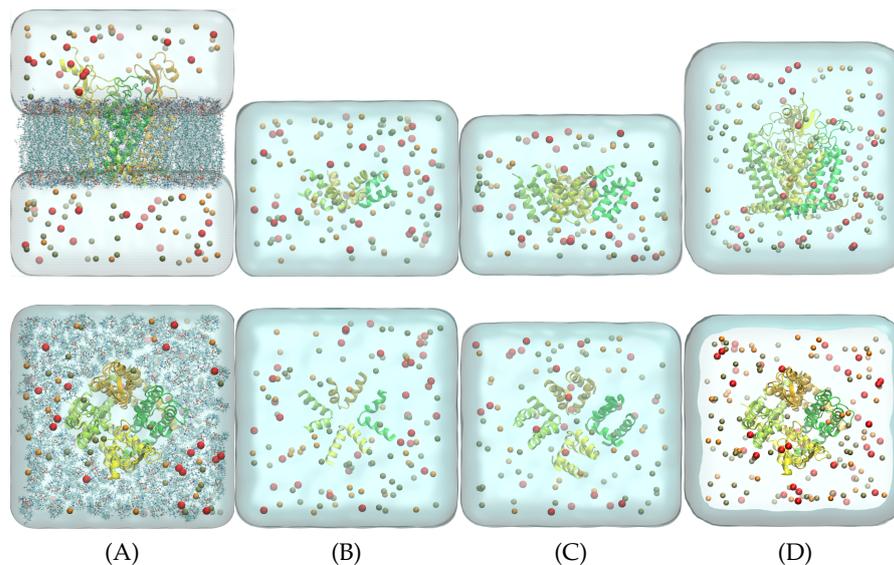


Figure 2.2: Pore model systems: Large model with membrane (A), Small model (B), Small model with additional helix segments below selectivity filter (C), Large model without membrane (D).

## 2.2 MOLECULAR DYNAMICS SIMULATIONS

Molecular dynamics simulations were carried out with NAMD. [16] At first, an energy minimisation was performed for 10000 steps. Simu-

lations were performed in the isothermal-isobaric ensemble (NPT) at 303.15 K and 1.013 25 bar and a timestep of 2 fs. Hydrogen bond lengths were kept fixed using the SHAKE algorithm. [34] The van-der-Waals' cutoff radius, switching distance and pairlist distance were set at 12.0 Å, 10 Å and 16.0 Å. Electrostatic interactions were calculated using the particle-mesh Ewald method with a grid spacing of 1.0 Å.[35] Force calculations for protein and lipids were based on the CHARMM36 force field and those of water on the TIP3P model. [36–38] Lennard-Jones parameters for ions were adopted from Joung and Cheatham. [39] Data of the first 20 ns was not included in analyses. Trajectories were processed and analysed using the MDAnalysis Python package. [40, 41]. In order to minimize the Root Mean Square Deviation (RMSD), translations and rotations in the  $xy$  plane were removed and the origin of the system was defined as the Center of Geometry of the DEKA residues'  $C_\alpha$  atoms.

### 2.2.1 *Equilibration Runs*

Pore models A and D were equilibrated in six decreasingly restrained simulations of  $2.5 \times 10^6$  timesteps (5 ns) each. Temperature and pressure were controlled with a Langevin thermostat (damping coefficient  $1 \text{ ps}^{-1}$ ) and barostat (piston period 100 fs in run 1-2 and 50 fs in run 3-6, piston decay 50 fs in run 1-2 and 25 fs in run 3-6). Harmonic constraints were imposed on the protein backbone (scaling factor = 1) and on side chain carbon atoms (scaling factor = 0.5). In simulations with model A, additional dihedral and improper restraints were placed on selected atoms inside the lipid layer and lipid heads in the membrane were harmonically restrained to the  $xy$ -plane. Table B.1 lists the values of the force constants of each run.

### 2.2.2 *Production Runs*

Temperature and pressure were controlled with a Langevin thermostat (damping coefficient  $1 \text{ ps}^{-1}$ ) and barostat (piston period 50 fs, piston decay 25 fs). Production runs with systems A, B and D were conducted for a total of 150 ns and for 500 ns with system C.

**PORE MODELS B AND C** The backbone atoms of the terminal residues was kept fixed in both cases. For the second model, also the additional helices below the selectivity filter were constrained.

**PORE MODELS A AND D** During production runs, no constraints were placed on models A and D.

### 2.2.3 Constant Electric Field Simulations and Decreased Lys1244 Charge

Two additional simulation series were performed with pore model C. In the first one, an external electric field was applied along the z-axis of ion flow (see figure 2.3a). Applied potentials were set to 100, 300 and 500 mV.

In the second one, the charge on selectivity filter's Lys1244 sidechain (see figure 2.3b) was reduced by  $0.5 e$ . The charge was fragmented into 0.17, 0.17 and 0.16  $e$  which were removed from the three  $\epsilon$ -ammonium H-atomic charges. To compensate for the reduction in positive charge,  $-0.5 e$  were also removed from one of the  $\text{Cl}^-$  anions. Simulations were carried out in the absence of an external field as well as in presence of the given external potentials.

In both cases, simulation runs were performed for a total of 500 ns.

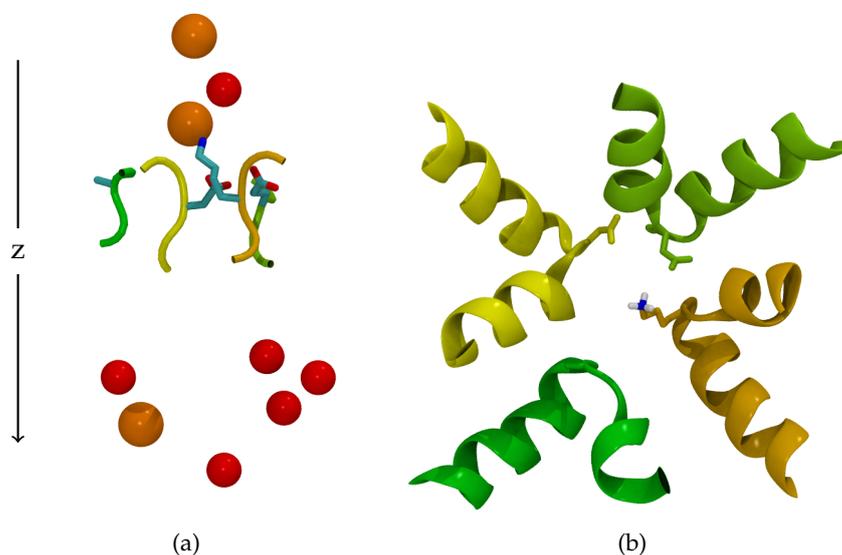


Figure 2.3: Side view on selectivity filter with indicated direction of external field (a) and top view of the Lys1244  $\epsilon$ -ammonium group (b) in pore model C.

## RESULTS

---

This chapter is dedicated to strategies and results of data analyses which were performed to select a suitable truncated pore model, quantify ion flux and evaluate the influence of external electric field strength and charge of the selectivity filter Lys1244 residue on selectivity and protein geometries. Furthermore, we will characterise four binding sites and explain approaches which led to their identification.

### 3.1 PORE MODELS

As described in section 1.2, selectivity in eukaryotic channels arises from the conserved DEKA motif. In the human  $\text{Na}_v1.4$  channel, it is constituted of Asp406, Glu761, Lys1244, Ala1536 (see figure 1.2).

Ion conduction through  $\text{Na}_v$  channels happens in response to an action potential. This depolarisation of the cell membrane is detected by the voltage-sensing domain and induces conformational changes which open the channel. [10] To investigate ion fluxes, it is therefore important to use a model which represents the pore in its conductive state.

In the present work, this condition was ensured by employing a pore model which only contained the selectivity filter region but not the gating helices. A similar strategy was adopted by Callahan and Roux. [10] In their simulations of ion conduction through the prokaryotic  $\text{Na}_v\text{Ab}$  channel they embedded the bare selectivity filter region in a pseudo-membrane constituted of a slab of Lennard-Jones particles. Here, instead of introducing a membrane, terminal protein-backbone atoms were constrained. In order to evaluate whether such a simplified pore construct was representative of the large protein, simulations were carried out with the large protein model A as well as with the three truncated pore models (B, C, D). The simplest of the small models, model B, consisted only of the selectivity filter and adjacent helices. Truncating the protein exposes hydrophobic core residues to the polar solvent and can therefore lead to altered dynamics. [7] To reduce water penetration and reproduce a more native environment, system C contained eight additional, immobile helix fragments located below the selectivity filter region. The importance of the presence of a membrane was investigated with a third model, D, where protein segments were identical to those in model A, however, they were not placed in a lipid membrane.

The quality of the selectivity filter models was evaluated by confronting the density profiles in the simplified systems (B, C, D) with data obtained in the simulations with the larger reference pore model (A). The relative positions were defined from the center of geometry of each individual residue's terminal sidechain functionality with respect to those of the common center of geometry formed by the four  $C_\alpha$  atoms.

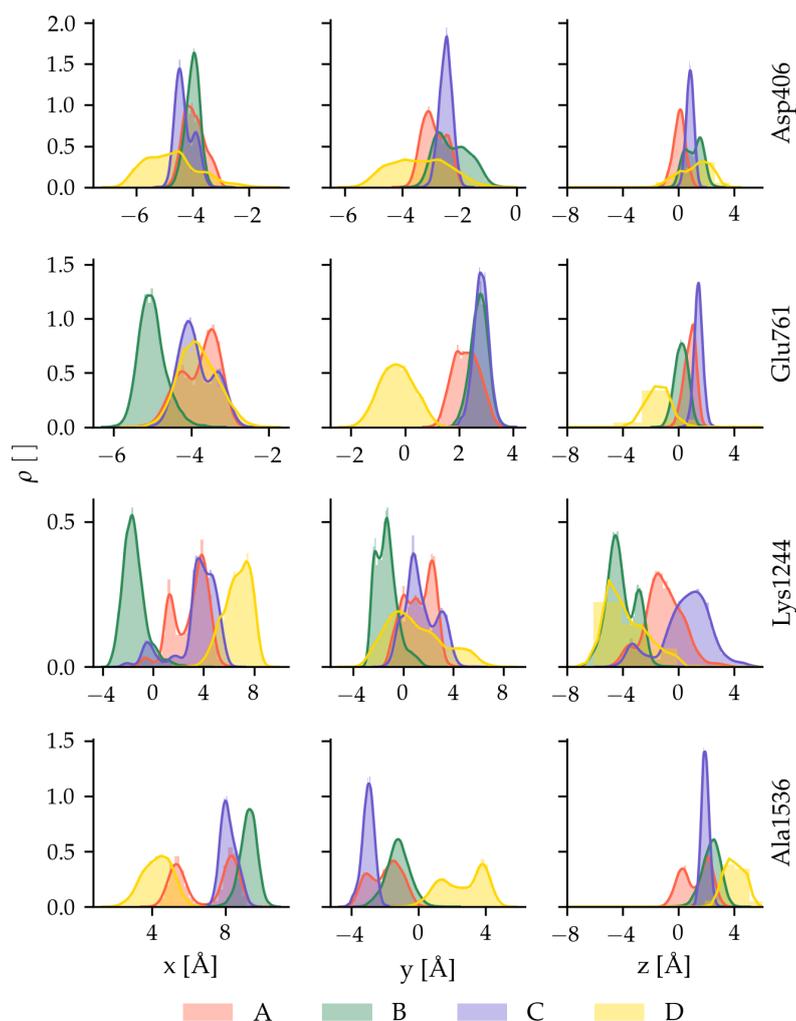


Figure 3.1: Probability distributions of the DEKA residues in pore models A, B, C, D along the  $x$ ,  $y$ ,  $z$  axes.

Figure 3.1 shows that the DEKA residue positions of model D deviated most from reference model A. The atoms in D were neither subject to constraints nor were they restrained by the presence of a lipid membrane. The broad distribution of the selectivity filter residues is representative for the increased flexibility in the protein structure.

Differences between model A and the two small pore models B and C were less pronounced. The width of their distributions were generally narrower which can be connected to the constraints placed close to the DEKA residues which are likely to impact mobility. In the  $x$  dimension, the Glu761, Lys1244 and Ala1536 residue positions in the large model showed two maxima in their distribution. In model C, the DEKA residues fluctuated the least which can be connected to the stabilisation by the additional protein chains.

In A, B, C and D, Lys1244 was the most mobile residue. On all three coordinate axes its densities spread over a range of at least 4 Å. The two acidic sidechains of Asp406 and Glu761 exhibited narrow profiles of width  $\approx 2$  Å on all axes and systems except for model D. Pan et al. [7] found that the mobility of these residues was limited mobility due to coordination with Arg756.

On the  $z$  axis, in systems A and C, Asp406, Glu761 and Lys1244 had high probabilities to be found in the same  $xy$  plane situated at around  $z = 2$  Å. In both models, the Lys1244 distribution showed more than one maximum, the smaller of which was located at around 3 Å. The larger peak was found at  $-2$  Å in model A, whereas in C Lys1244 positions were closer to those of the Asp406, Glu761 and Ala1536 residues (1 to 2 Å).

In model B, Ala1536 was found slightly above and Glu761 slightly below Asp406. Also here, Lys1244 density displayed two peaks, the first one lying around 1 Å above the common maximum of A and C (3 Å) and the second one below at approximately  $-5$  Å. The resulting constriction point in the  $z$ -direction between Lys1244 and the acidic residues presents a potential key factor for selectivity.

In the  $x$  direction, in models A and C, Lys1244 was situated up to 8 Å apart from Asp406 and Glu761 which were both found at  $-4$  Å.

In model B, Lys1244 density was highest the same  $x$  coordinate as the acidic residues.

Considering also the  $y$  dimension, in system B, Lys1244 was situated almost directly below Asp406 ( $-1$  to  $-3$  Å). This was not the case in models A and C where Lys1244 positions were more widespread across the  $xy$  plane, with  $x$  coordinates ranging from approximately  $-1$  to 5 Å and  $y$  around  $-1$  to 4 Å. Based on these observations, model C was selected to be most representative of the large pore model.

Taking the positional fluctuations of model D into account, it is likely that, unless additional constraints or restraints are placed on a fraction of atoms, the Na<sub>v</sub> channel protein model without a membrane is not able to reflect dynamics of the physical system.

### 3.2 QUANTIFICATION OF ION FLUX AND CONDUCTANCE

To quantitatively evaluate selectivity and conductance characteristics in each of the simulations, normalised  $\text{Na}^+$  and  $\text{K}^+$  fluxes along the axis of ion transport ( $z$ ) were calculated. This also corresponded to the direction in which external fields were applied (figure 2.3a). The ion current,  $I_z$ , within the entire simulation box can be obtained as the sum of all ionic displacements with respect to their positions at the previous timestep and relative to the height of the cell,  $L_z$ . [42] The mean flux,  $\langle I_z \rangle$  can be computed by averaging instantaneous data,  $I_z(t)$ , over the whole trajectory. For the computation of conductance with different Lys1244 charges, the simulation conditions could be considered equivalent because the reduction of the positive Lys1244- $\epsilon$ -ammonium charge by  $0.5 e$  was compensated by removing  $-0.5 e$  from one of the  $\text{Cl}^-$  anions.

$$I_z(t) = \frac{1}{L_z} \sum_i^N q_i \underbrace{\frac{(r_z(t) - r_z(t - \Delta t))}{\Delta t}}_{v_z(t)} \quad (3.1)$$

$$\langle I_z \rangle = \frac{\sum_0^t I_z(t)}{t} \quad (3.2)$$

In the present work, the ion channel protein was not embedded in a lipid membrane. The ions in the bulk solvent were therefore free to migrate in the direction of the external field. In the case where ionic movements in the entire simulation box were considered for calculating  $I_z$ , currents were found to increase linearly with the electric potential difference,  $U$ . As depicted in figure 3.2, the relative conductance increase of  $\text{K}^+$  exceeded that of  $\text{Na}^+$  which is in accordance with the higher electrical mobility of the ion species.

The employed pore model C occupied a relatively small fraction of the simulation boxes  $xy$ -plane (figure 2.2C) and ion fluxes across the large barrier-free area which was surrounding the protein were masking the dynamics within the pore region. Therefore, conduction analyses without considering bulk solvation were necessary to obtain information about ion transport along the selectivity filter.

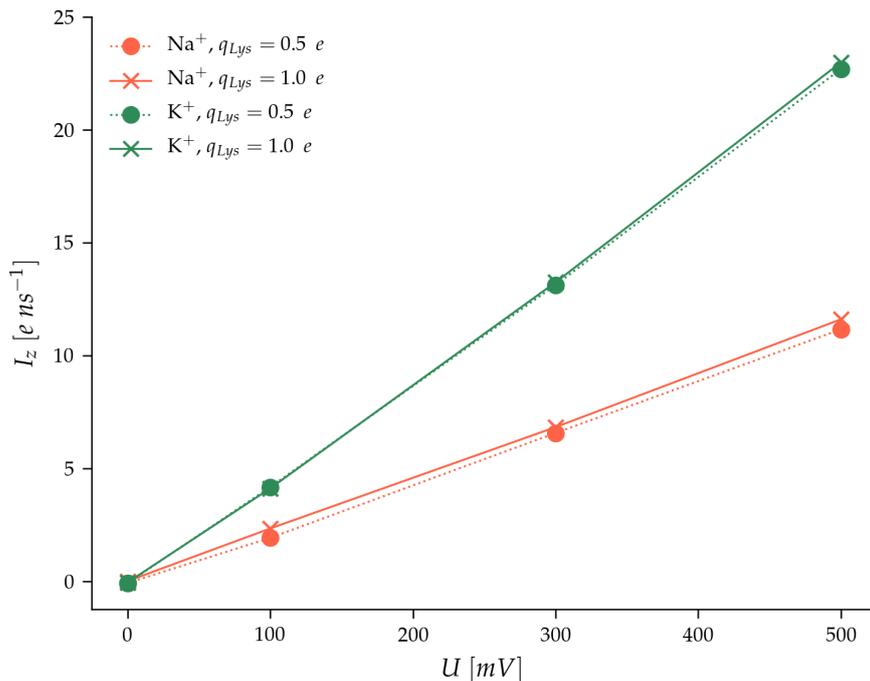


Figure 3.2: Average Na<sup>+</sup> and K<sup>+</sup> current, in the simulation box in function of different external electric potentials,  $U$ .

To evaluate conductance in the selectivity filter region, ion fluxes were calculated within a representative cylindrical segment along the  $z$ -axis. The required diameter of this cylinder was determined based on distances between the C $_{\alpha}$ -atoms of the DEKA motif's Asp406 and Lys1244, and Glu761 and Ala1536 residues as given in figure 3.3.

Largest mean and maximum separations ( $d_{avg} = 14.2 \text{ \AA}$ ,  $d_{max} = 15.5 \text{ \AA}$ ) were found between Glu761 and Lys1244 at  $U_z = 500 \text{ mV}$  and  $q_{Lys1244} = 1.0 e$  (see appendix C table C.1). In order to assure that ion transport within the entire selectivity filter region was captured, conductance analysis was carried out within a cylindrical volume element of radius  $9 \text{ \AA}$  around the COG of the DEKA C $_{\alpha}$  atoms, COG<sub>DEKA</sub> (see figure 3.4).

Considering this cylinder, Z1, which stretched over the whole  $z$ -axis, relative changes in conduction flux were comparable to those in the full box (figure 3.2 and 3.5). The barrier posed by the protein spanning the entire  $xy$ -dimension of the analysis region led to a smaller absolute conductance.

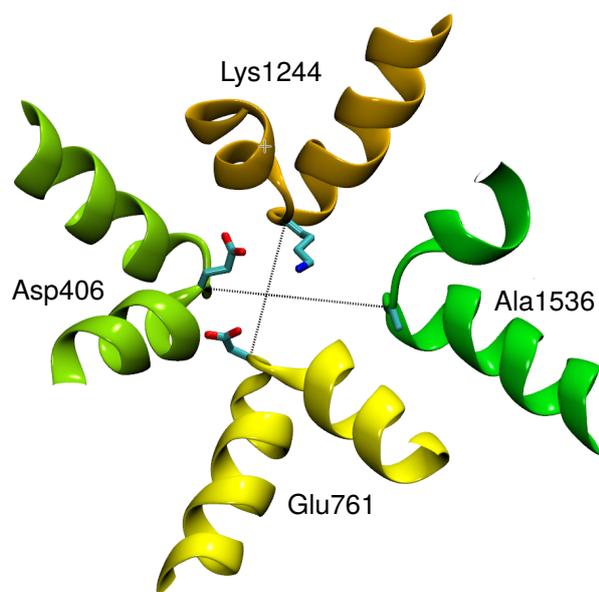


Figure 3.3: Distances used for determining the size of the pore region.

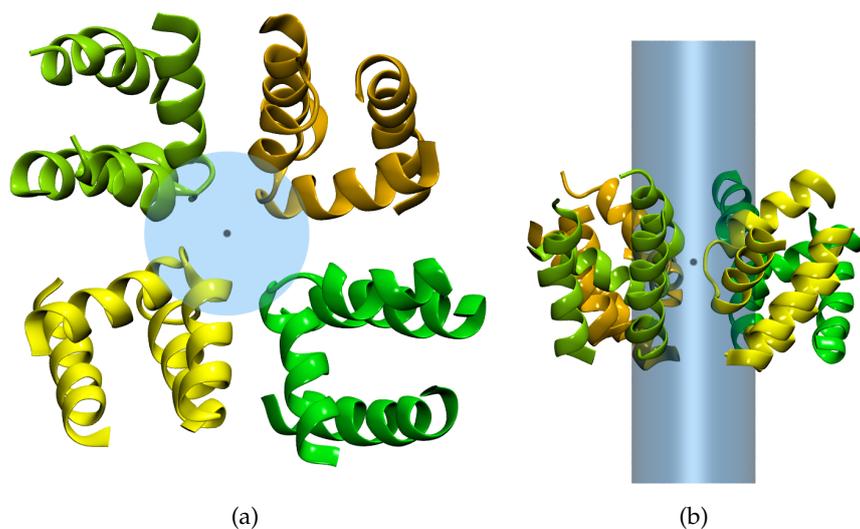


Figure 3.4: Top (a) and side (b) view on Z1 volume element,  $\text{COG}_{\text{DEKA}}$  indicated by black sphere.

The here performed analysis only targeted displacements along the  $z$ -axis. However, even though the applied field exerts a force on charged particles, ionic movement does not become constrained to one dimension. Including a large fraction of the bulk solvent region above the cylinder still captures the flux of ions which are not crossing the  $xy$

plane through the channel but are bypassing the protein barrier in regions beyond the area of protein extension.

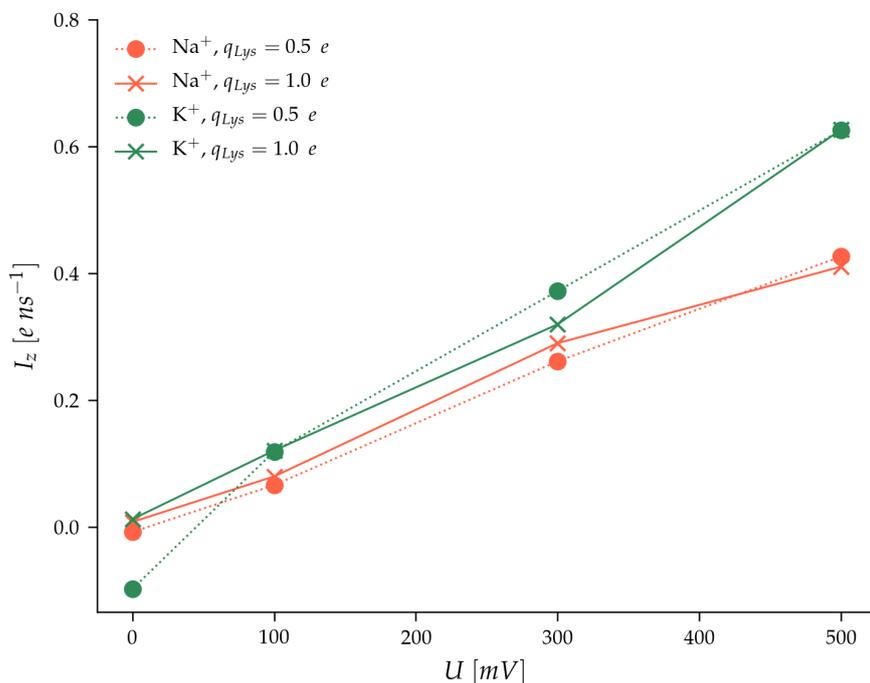


Figure 3.5: Average Na<sup>+</sup> and K<sup>+</sup> current,  $I_z$  in cylindrical volume element Z1 in function of different external electric potentials,  $U$ .

As a consequence, to exclusively investigate conductance associable to ions which are crossing the  $xy$ -plane through the channel pore, the  $z$  dimension of the sampling region was further reduced to the volume elements enclosed by the protein (Z2) or in close proximity to the selectivity filter (Z3). Representative fluxes were calculated within the two smaller cylindrical volume elements which are shown in figure 3.6a and 3.6b. Their upper and lower  $z$ -limits were determined by analysing mean and maximum absolute  $z$ -dimensions of the entire protein (Z2; figure 3.6a; appendix C, table C.4) and a fraction of it which included only the DEKA motif and 20 adjacent residues (Z3; figure 3.6b; appendix C, table C.3).

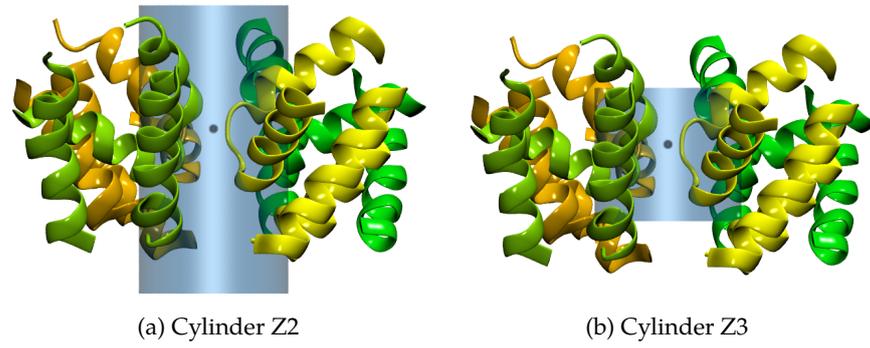


Figure 3.6: Truncated cylindrical volume elements representing protein (a) and selectivity filter volume elements (b).

In both analysis regions, absolute conductance continued to increase in function of applied field strength,  $U_z$  and also  $K^+$  conductance exceeded that of  $Na^+$  (figure 3.7, 3.8) which is contradictory to experimental results where  $Na^+/K^+$  permeation ratios in mammalian sodium channels were found to lie between 10 and 30. [43]

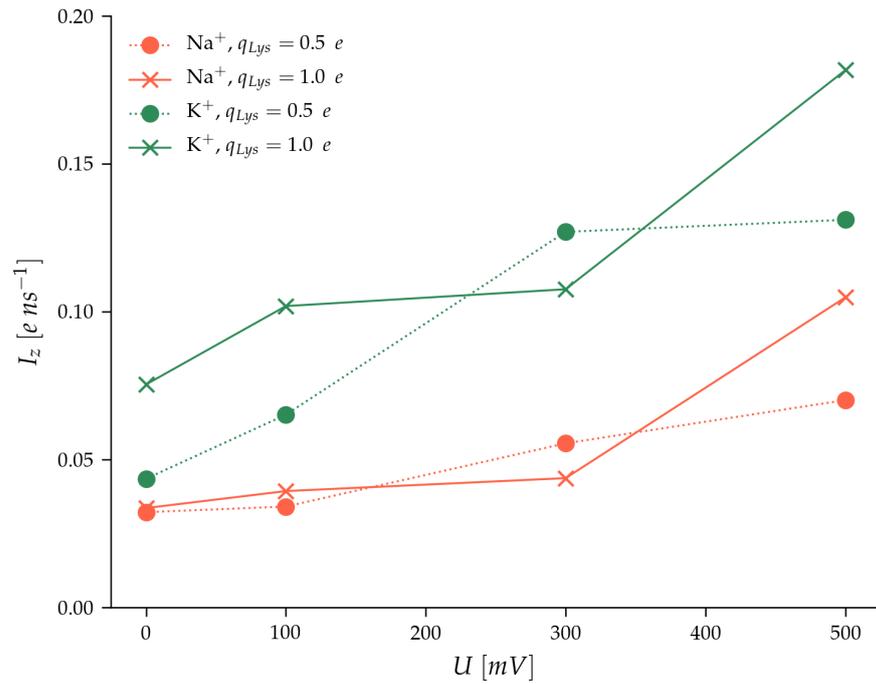


Figure 3.7: Average  $Na^+$  and  $K^+$  current,  $I_z$  in cylindrical volume element Z2 in function of different external electric potentials,  $U$ .

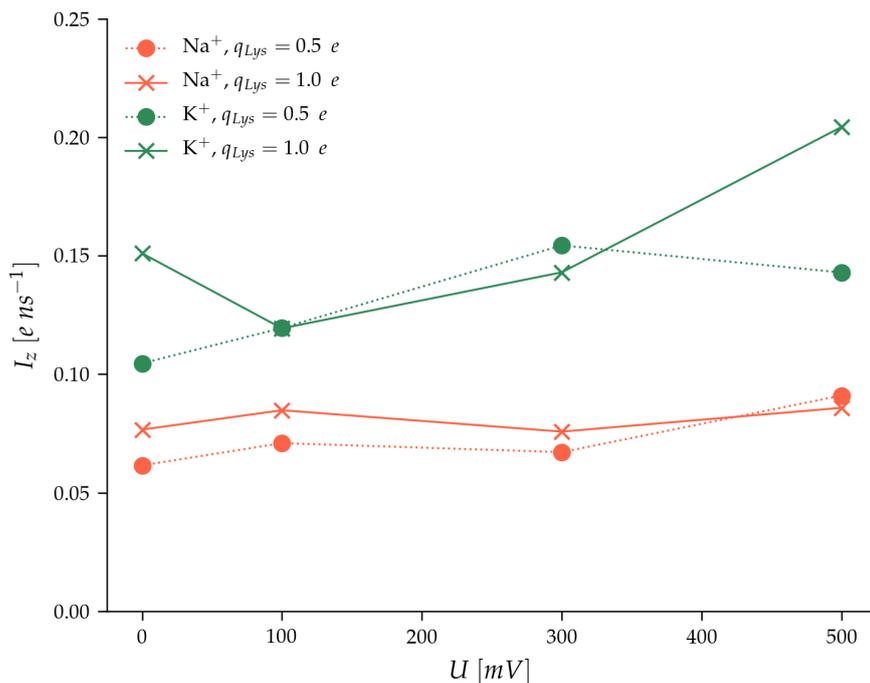


Figure 3.8: Average Na<sup>+</sup> and K<sup>+</sup> current,  $I_z$  in cylindrical volume element Z3 in function of different external electric potentials,  $U$ .

One explanation for this misrepresentation could be the approximating description of ion-ligand interactions in the selectivity filter region. The employed classical CHARMM36 force field [15] separates the non-bonded intermolecular potential into two terms. The first one describes long-range electrostatic interactions between point charges and is evaluated as a Coulomb potential (equation 1.57). The second term accounts for all other non-bonded, short-range contributions in an average fashion and is represented in a Lennard-Jones potential.

In this case, since their charge is equal, the differences in ligand-ion interactions between Na<sup>+</sup> and K<sup>+</sup> are governed by the Lennard-Jones  $\sigma$  and  $\epsilon$  parameters.

However, Na<sup>+</sup> has a smaller ionic radius than K<sup>+</sup>, which results in a higher charge-to-volume ratio with a stronger polarising effect on its environment. [44] Especially, an oversimplified accounting for polarisation interactions is, therefore, likely to explain the failure of an accurate description of the ion-protein, ion-water and protein-water interactions which are key to Na<sup>+</sup> selectivity.

It has been shown that the positively charged Lys1244 residue in the DEKA motif is essential for selectivity in mammalian Na<sub>v</sub> channels. [45] Polarising effects of a metal cation would reduce the positive charge concentration on the Lys1244  $\epsilon$ -ammonium group (see figure 2.3b).

To model polarisation interactions with this residue, simulations were also carried out with a reduced charge on the Lys1244  $\epsilon$ -ammonium group's H-atoms.

In this case, the 'charge redistribution' was not determined by the nature of a present ion species but an increased conductance could, nonetheless, underline the importance of polarisation interactions for selective permeability.

As illustrated in figure 3.9, in Z2, at  $q_{Lys1244} = 1.0 e$  and field strengths ranging from 0 to 300 mV, the ratio between  $Na^+$  and  $K^+$  conductance remained constant at 0.4. At 500 mV, it increased to 0.6 which means that here the voltage-dependence of  $I_z$  was slightly higher for  $Na^+$  conduction.

With  $q_{Lys1244} = 0.5 e$ , at 100 and 500 mV, conductance was found 50% lower for  $Na^+$  than for  $K^+$ , and 40% lower at 300 mV. With the full lysine charge, at 100 mV, a charge reduction increased relative  $Na^+$  permeation from 0.4 to 0.5 whereas it had no effect at 300 mV and an opposite influence at 500 mV.

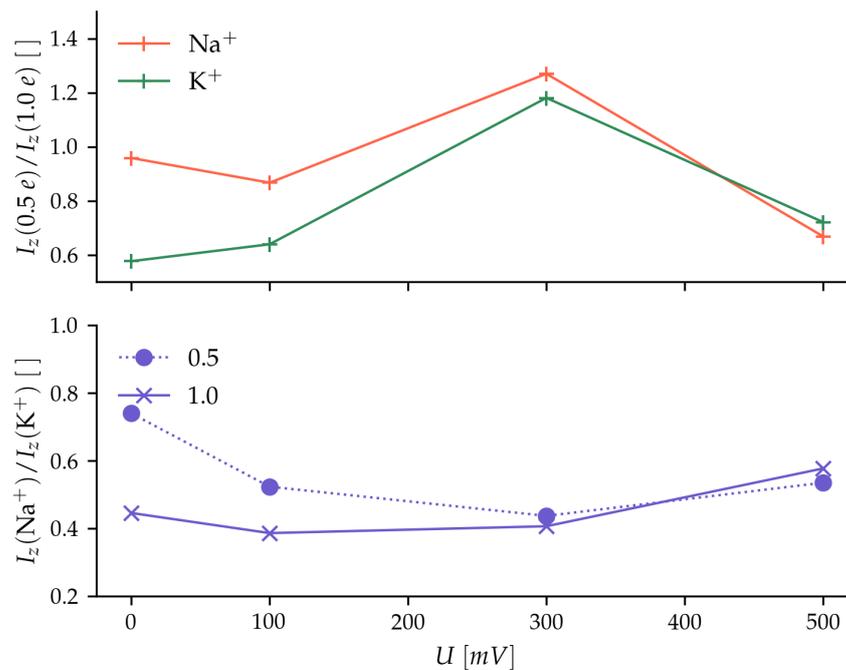


Figure 3.9: Voltage dependence of  $q_{Lys1244} = 0.5 e / q_{Lys1244} = 1.0 e$  conductance ratios for  $Na^+$  and  $K^+$  and  $Na^+ / K^+$  conductance ratios for  $q_{Lys1244} = 0.5 e$  and  $q_{Lys1244} = 1.0 e$  in Z2.

Comparing the effects of lowering  $q_{Lys1244}$  on the fluxes of each ion species separately, at 100 and 500 mV, decreasing  $q_{Lys1244}$  cut down  $Na^+$  and  $K^+$  conduction. With  $U_z = 500$  mV, reducing  $q_{Lys1244}$  lowered both

ionic fluxes by 30 %. Only at 300 mV, did conductance ratios increase to 1.2 to 1.3 for  $\text{Na}^+$  and  $\text{K}^+$ , respectively.

Interestingly, at the physical field strength of 100 mV,  $\text{K}^+$  permeation declined by 40 % whereas 90 % of  $\text{Na}^+$  fluxes were found to persist. Overall, in cylinder Z2, the reduced charge on the Lys1244 sidechain yielded an better representation of selectivity and employing a less crude approximation for modelling polarisation effects could thus further improve the description of ion discrimination.

In the smallest cylindrical selection Z3, trends in the dependence of conductance of the individual ion species on  $q_{\text{Lys1244}}$  were similar to those found for Z2 (figure 3.10). Here, however,  $\text{Na}^+$  permeation only decreased at the highest field strength of 500 mV while at 100 and 300 mV it was found equal or even higher with the reduced Lys1244 charge. Differently from the observations in Z2, in this case,  $\text{K}^+$  conduction only increased by 10 % at 500 mV whilst at 100 and 300 mV it was found 20 and 10 % below the respective value at full Lys1244 charge.

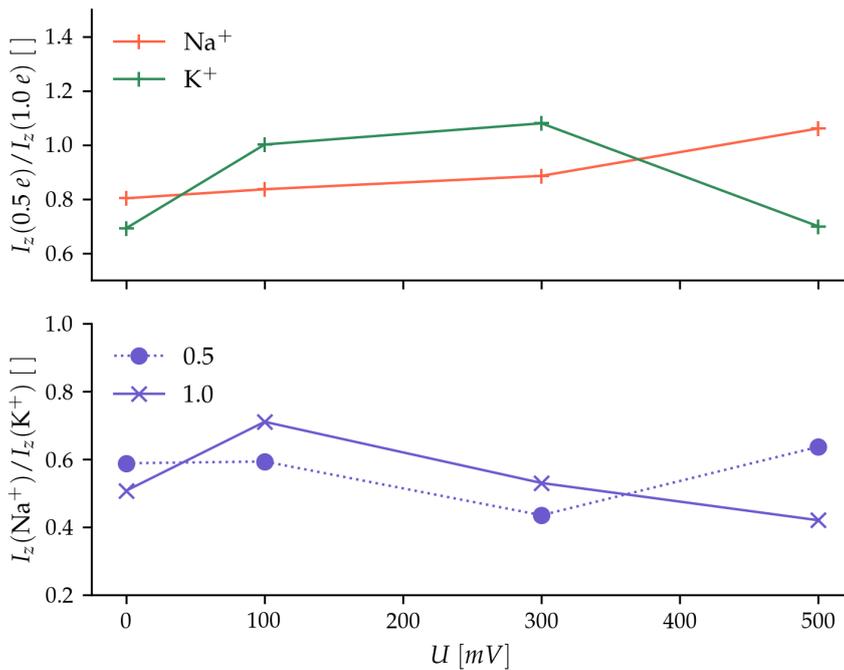


Figure 3.10: Voltage dependence of  $q_{\text{Lys1244}} = 0.5e/q_{\text{Lys1244}} = 1.0e$  conductance ratios for  $\text{Na}^+$  and  $\text{K}^+$  and  $\text{Na}^+/\text{K}^+$  conductance ratios for  $q_{\text{Lys1244}} = 0.5e$  and  $q_{\text{Lys1244}} = 1.0e$  in Z3.

With a  $\text{Na}^+/\text{K}^+$  conductance ratio of 0.7, selectivity was highest at 100 mV and  $q_{\text{Lys1244}} = 1.0e$ . The decreases in relative  $\text{Na}^+$  currents in simulations with the full lysine charge, were identical to those calculated

in the larger cylinder, Z2, with  $q_{Lys1244} = 0.5 e$ . Also  $Na^+/K^+$  selectivity was found to worsen by up to 20% with reduced charge at 500 mV.

The discrepancy between trends observed in Z2 and Z3 indicate that, the reduced Lys1244 charge might have positively affected the description of ion fluxes in the region beyond the extension of Z3 but the interactions occurring in immediate proximity of the selectivity filter were likely too complex to be represented with the employed force field parametrisation and the very simple approach to consider polarisation.

### 3.3 EXTERNAL FIELD STRENGTH AND LYS1244 CHARGE

The two driving forces for ion transport across cell membranes are a concentration and an electric potential gradient. Due to the absence of a membrane and the use of a periodic simulation system no concentration gradient was implemented in the simulations. The membrane potential was accounted for in the form of an external electric field which was applied along the  $z$ -axis, corresponding to the direction of ionic flux. In line with conventions, coordinates were set such that the extracellular space was represented by  $z > 0$  and the intracellular box fraction was assigned  $z < 9$ . Positive ions flow along the electric potential gradient, therefore  $U < 0$  leads to negative ion fluxes along the  $z$ -axis. The direction of positive conductance (appendix D, tables D.1, D.2, D.3, D.4) is defined in direction of ion transport.

The timescales at which ion conduction takes place is in the ns range. [46] Applying electric fields is one strategy for accelerating dynamics. By increasing the amount of conduction events per time interval, required simulation times and computational cost can be reduced. [12] In this case it is, however, crucial to consider the possibility that the additional electric force can alter dynamics and make them unsuitable to accurately model processes in the system under investigation.

The DEKA motif of the eukaryotic selectivity filter contains two negatively charged (Asp406 and Glu761), one positively charged (Lys1244) and one neutral residue (Ala1536). As depicted in figure 3.11, increasing field strengths caused positions of the positively charged Lys1244 residue to shift towards the negative  $z$ -direction along which ion transport takes place.

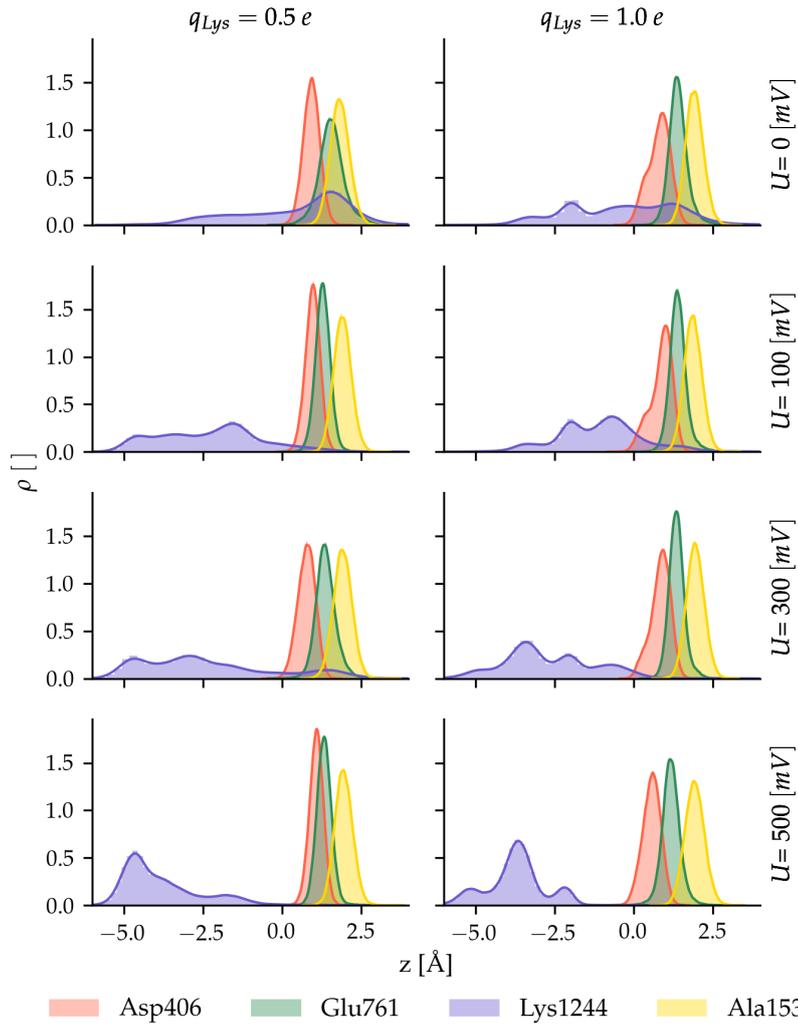


Figure 3.11: Density profiles of the COG of the terminal selectivity filter DEKA residue sidechains along the  $z$ -axis with respect to  $\text{COG}_{\text{DEKA}}$ .

Positions of the neutral Ala1536 and the negative Glu761 were neither influenced by the external field nor by  $q_{Lys1244}$ . Pan et al. [7] discovered that the acidic Glu761, together with Glu764, is stabilised by the conserved Arg756 guanidinium group. This finding might be an explanation for the narrow density distributions.

Only at 500 mV and reduced Lys1244 charge could a positive shift be observed on the Asp406 density. This could be related to electrostatic interaction between Lys1244 and Asp406 sidechains which limited their mobility in response to the external field. At  $q_{Lys1244} = 1.0 e$ , this attraction was found to drag Asp406 densities in the direction of the field. Likewise, with  $q_{Lys1244} = 0.5 e$ , already low field strength of 100 mV caused a significant fraction of Lys1244 density to shift by

approximately 5 Å along the  $z$ -axis. At its full charge, this effect was less pronounced and the maximum of the distribution at 500 mV was found at  $-3.6$  Å compared to  $-4.7$  Å at  $q_{Lys1244} = 0.5 e$ . This observation shows that intermolecular attractive forces between Lys1244 and the acidic residues is stronger than that of the external field.

Figure 3.12 confronts the DEKA  $xy$ -positions at different  $U_z$  and  $q_{Lys1244}$ . In all cases, Asp406, Glu761 and Ala1536 positions remained unchanged under the influence of an external potential or altered  $q_{Lys1244}$  whereas a high mobility was found in the Lys1244 sidechain. Lys1244 was more located at full charge but no clear correlation between position, charge and field strength could be observed.

However, the relatively wide density distribution of Lys1244 compared to those of the other DEKA residues, affirms that the dynamics of this particular sidechain should be decisive in the process of ion transport, and since these are governed by intermolecular forces, it is vital to use a force field which is able to accurately reproduce interactions with this residue.

Additionally, density profiles of the Lys1244 sidechain, were found to remain almost identical, independently of whether  $Na^+$  or  $K^+$  were observed within a 10 Å of the  $\epsilon$ -ammonium group (see figure 3.13) The same observation held true for the three other DEKA residue endgroups.

The radius of hydrated  $Na^+$  ions is smaller than that of  $K^+$ , however,  $Na_v$  channels are wider than voltage-gated potassium ( $K_v$ ) channels. [12, 47] The central cavity above the constriction site is wide enough to accommodate hydrated ions and with a radius of around 3.5 Å even the smallest diameter of the  $Na_v1.4$  channel is still large enough to allow the passage of partially hydrated ions of both species. [7, 12]

From that follows that size exclusion cannot be the only factor governing selectivity. It also implies that discrimination between  $Na^+$  and  $K^+$  should arise from a difference in interactions between residues of the pore region and the respective ion species.

Selectivity could be related to the presence of the conserved DEKA motif, and in particular the lysine residue. Differences in interaction characteristics with  $Na^+$  or  $K^+$  should, therefore, become evident from a variation in density distributions of these four residues.

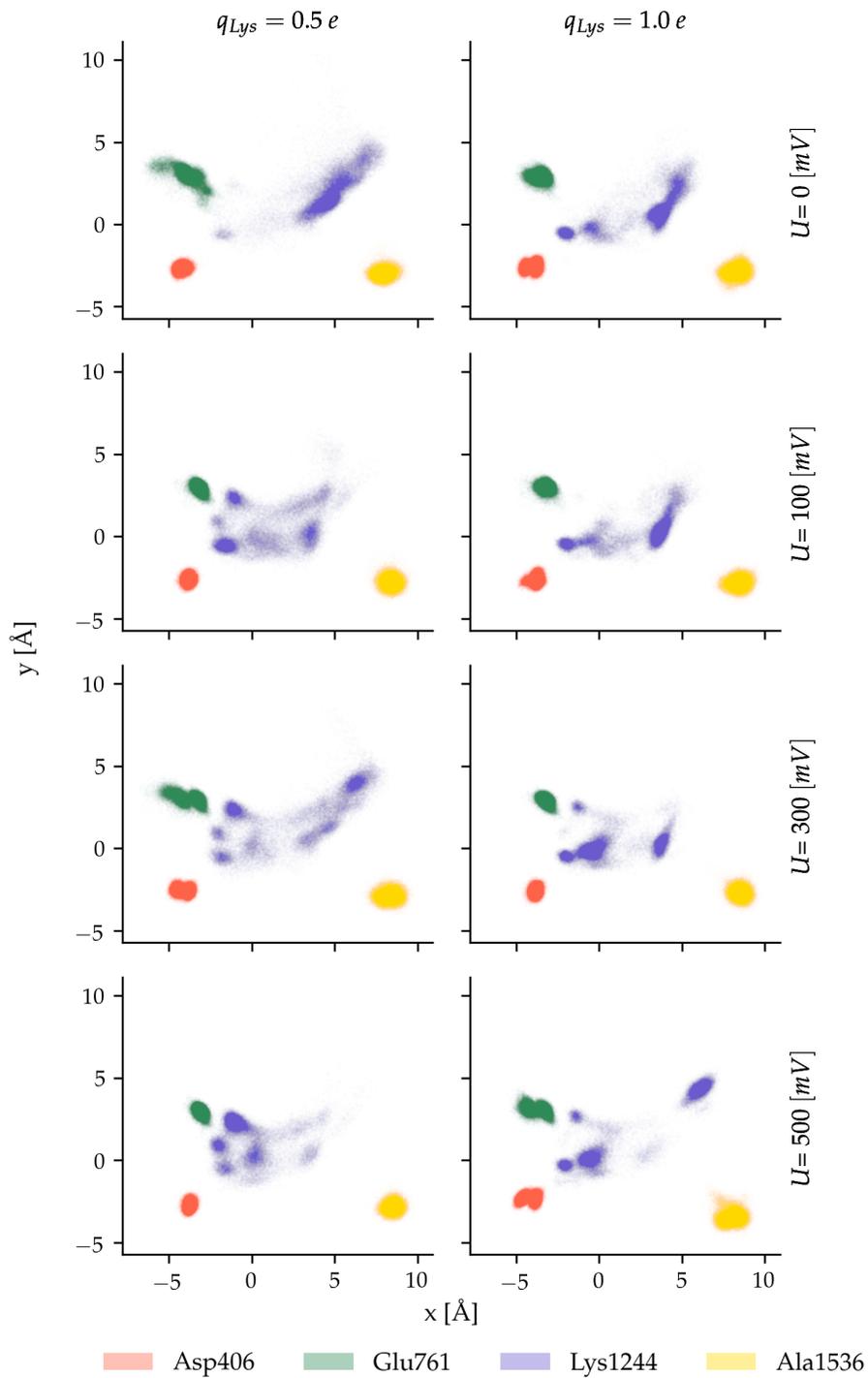


Figure 3.12: Position distribution of the COG of the terminal selectivity filter DEKA residue sidechains within the  $xy$ -plane with respect to  $\text{COG}_{\text{DEKA}}$ .

A probable explanation for the misrepresentation of interactions leading to discriminatory behaviour is, again, the oversimplification and inadequacy of the employed interaction potential parametrisation. Including an explicit potential term for describing polarisation effects would likely yield a more accurate interaction potential.

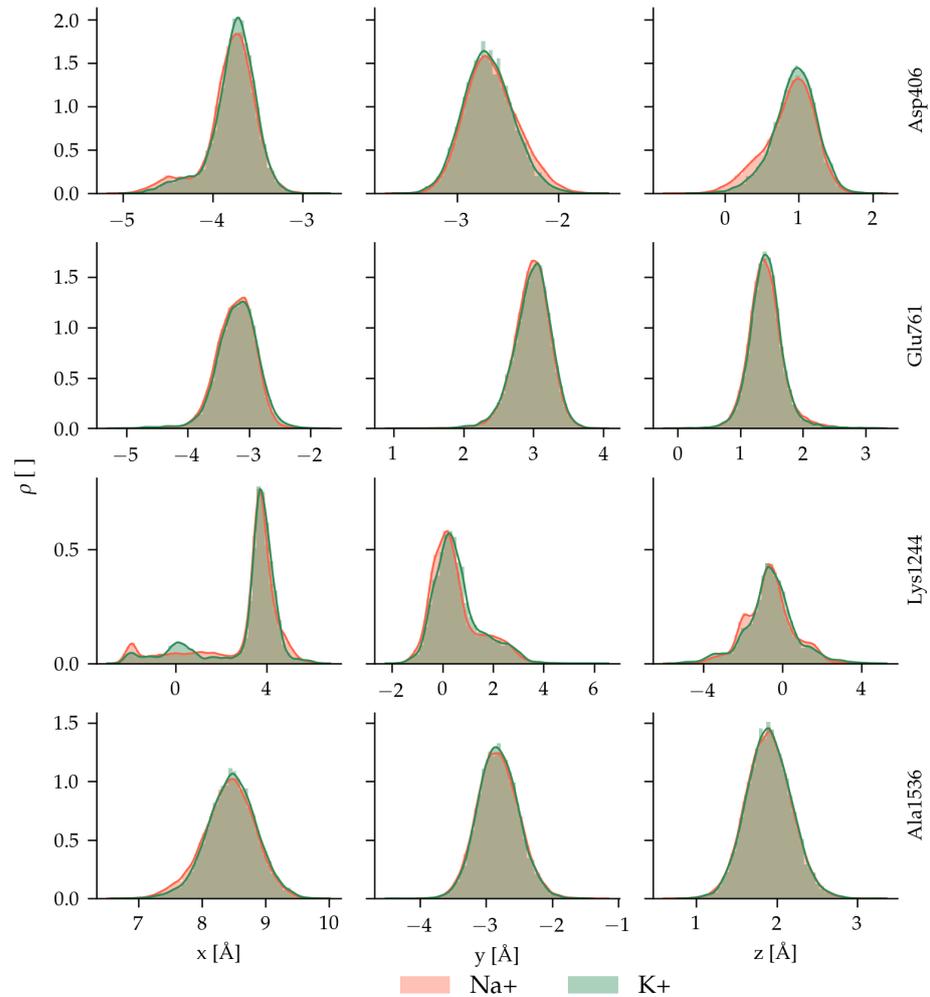


Figure 3.13: Lys1244  $x$ ,  $y$  and  $z$  densities,  $\rho$ , at  $q_{\text{Lys1244}} = 1.0e$  and  $U_z = 100$  mV and  $\text{Na}^+$  or  $\text{K}^+$  within  $10 \text{ \AA}$  of  $\text{COG}_{\text{Lys1244 } \epsilon\text{-NH}_3^+}$ .

Previous research by Callahan and Roux [10] has explored the effect of a variant electric potential on conductivity behaviour and selectivity site geometry of ion channel models. In their simulations they found that magnitudes of potential difference and ion concentrations affected the orientation of Glu residues in a prokaryotic  $\text{Na}_v$  selectivity filter. In

their work, at biological membrane potentials of  $-200$  mV, they did not observe selectivity. At  $-500$  mV, a preference for  $K^+$  conduction was found.

The present data agrees with these results. As it can be seen in figures 3.6a, 3.7 and 3.8, at  $500$  mV,  $K^+$  currents exceeded those of  $Na^+$  in both selections. The increasing  $z$ -separation between Lys1244 and the other three DEKA residues was found to increase with greater applied field strengths. This explains the generally higher permeability but not the  $K^+$  selectivity.

### 3.4 BINDING POCKETS AND DEHYDRATION

The inconsistency in selectivity and conductance trends between the Z2 and Z3 analysis regions indicate the presence of additional coordination sites, located beyond the extension of the shorter cylinder, Z3. As discussed in section 3.2, the smaller Lys1244 charge seems to have affected binding strength in the two regions differently for  $Na^+$  and  $K^+$  (see figure 3.9, 3.10). In Z3, at  $100$  mV, a reduction of  $q_{Lys1244}$  left  $Na^+$  conductance unaltered whereas in Z2 it led to a decrease by  $40\%$ . On the other hand, under the same conditions, in Z3,  $K^+$  current was reduced by  $20\%$  but only by  $10\%$  in Z2. A decrease in  $q_{Lys1244}$  appears to have increased  $Na^+$  binding strength in these regions while it had the opposite effect on  $K^+$ . Weakened binding in the entrance region could be explained by a reduction off access resistance from the positively charged Lys1244 sidechain. In this case, however, a similar effect would be expected on the conductance of both ion species.

In order to identify the potential ion binding pockets, density profiles for  $Na^+$  and  $K^+$  ions along the  $z$ -axis were determined. For both ions, these peaked at the entrance and exit of the pore which lead to the assumption that binding sites should be present in these regions.  $Na^+$  ions had a relatively higher probability to be found in these sites than  $K^+$  ions which could be related to tighter ion-protein binding and also explained the lower conductance of this ion species.

The obtained probability densities,  $\rho(z)$ , could be used to calculate relative binding free energies,  $G(z)$ .

$$G(z) \propto -\ln \rho(z) \quad (3.3)$$

Independently of applied field strengths and Lys1244 charge, the resultant relative binding free energies were consistently more negative for  $Na^+$  than  $K^+$  (see appendix E, figures E.1–E.4) which indicated more attractive protein-ion interactions with  $Na^+$ . A correlation could also be found between relative dwell-times along the  $z$ -axis and the number of coordinating water molecules.

Favourable protein-ion interactions can compensate the energetic cost associated to dehydration. In order to elucidate the connection between binding strength and the extent of hydration,  $\text{Na}^+$  and  $\text{K}^+$  coordination numbers along the  $z$ -axis were determined based on the number of water oxygen atoms within the coordination radius of the respective ion species. Values of the coordination radii were adopted from Corry and Thomas ( $\text{Na}^+$ : 3.0 Å,  $\text{K}^+$ : 3.4 Å). [9]

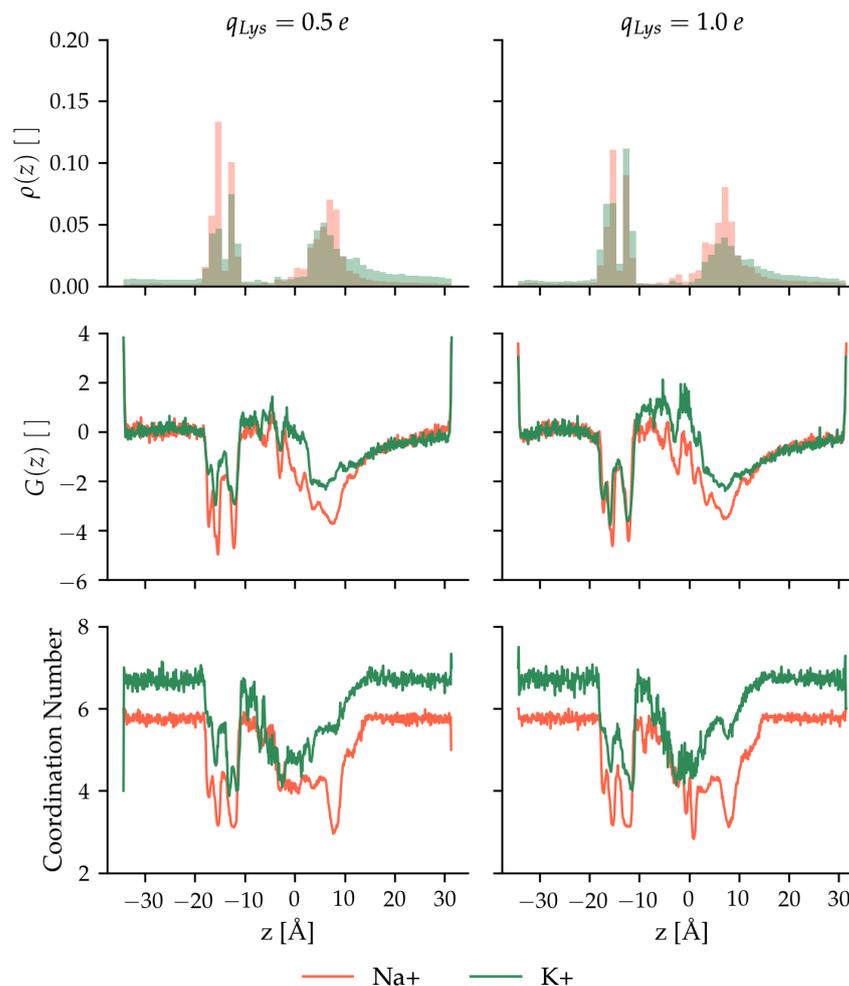


Figure 3.14: Probability densities,  $\rho(z)$ , relative binding free energies  $G(z)$  and hydration numbers for  $\text{Na}^+$  and  $\text{K}^+$  ions along the  $z$ -axis during simulations with  $U = 100$  mV

Negative relative binding free energies, corresponding to strong protein-ion interactions, coincided with low coordination numbers (figure 3.14). In the bulk solvent region,  $\text{Na}^+$  was coordinated to an average of 5.8 and  $\text{K}^+$  to a mean of 6.7 water molecules. This is consistent with data

from previous Molecular Dynamics simulations which determined coordination numbers of 5.0 to 5.8 and 6.0 to 7.0 for  $\text{Na}^+$  and  $\text{K}^+$ , respectively. The preference for a larger hydration number of the  $\text{K}^+$  ion is connected to its larger radius. [47, 48]

At 100 mV and  $q_{\text{Lys1244}} = 1.0 e$ , the  $\text{Na}^+$  coordination profile along the  $z$ -axis displayed distinct minima at 8, 1, -1, -12, -15 and -17 Å. In these sites, the coordination number dropped below 3 or 4 (figure 3.14). The greatest extent of dehydration was observed at 1 Å which marked the location of the selectivity filter residues.

In the dehydration minima around the selectivity filter, the degree of dehydration decreased by three water molecules for both ion species which means that  $\text{Na}^+$  crossed the channel with a mean of three and  $\text{K}^+$  with four coordinated water molecules.

Using Density-Functional Theory (DFT), Dudev and Lim [14] found that the three coordinating water molecules around a  $\text{Na}^+$  ion can coordinate to the Asp406 and Glu761 carboxylate O-atoms and the backbone carbonyl-O of Ala1536. In this conformation, the Lys1244  $\epsilon$ -ammonium group can form H-bonds with the second Glu761 sidechain O-atom and with the Ala1536 carbonyl-O.

Ionic radii of solvated  $\text{Na}^+$  and  $\text{K}^+$  have been experimentally determined to lie around 1.1 and 1.4 Å. The higher charge concentration in the smaller  $\text{Na}^+$  ion increases binding strength between the cation and its coordinating water molecules and yields more negative hydration enthalpies ( $\Delta H_{\text{hyd}} = -391 \text{ kJ mol}^{-1}$  ( $\text{Na}^+$ );  $-308 \text{ kJ mol}^{-1}$  ( $\text{K}^+$ )). Consequently, with 2.4 Å, also distances between  $\text{Na}^+$  and the O-atoms of the first hydration shell are shorter than those of  $\text{K}^+$ -O (2.8 Å). The smaller  $\text{Na}^+$  ion also prefers smaller coordination numbers, therefore, a selectivity filter with three metal-binding sites favours  $\text{Na}^+$  over  $\text{K}^+$  binding. In addition,  $\text{Na}^+$  is a better electron acceptor which can exert a greater polarising effect on its coordinating water molecules. This is reflected in a greater field strength of the  $\text{Na}^+$  hydration shell, which in turn favours binding with Lewis base functionalities of the protein. Due to the higher energetic cost associated to dehydration,  $\text{Na}^+$ -binding is facilitated by the presence of the electronegative amino acids carboxylate groups. [47, 48]

At physical field strength of 100 mV, a decrease in  $q_{\text{Lys1244}}$  diminished  $\text{Na}^+$  selectivity which can be explained by an increased coordination number arising from to reduced spatial constriction in the selectivity filter region. The selectivity region was defined to lie within  $z(\text{COG}_{\text{DEKA}} \pm 2.5 \text{ Å})$ . Due to the decreased repulsion by Lys1244,  $\text{Na}^+$  was able to retain an average of 0.3 more water molecules and cross the selectivity filter with a coordination number of 4.2. Similarly, also for  $\text{K}^+$ , the mean coordination number in the selectivity filter region increased from 4.7 to 4.8 (see appendix E table E.1).

Comparing the extents of dehydration and relative binding free energies in the central cavity to those at the channel entrance showed that for similar coordination numbers, binding free energies in the binding region at  $z = 8 \text{ \AA}$  were disproportionately more negative than those found in the selectivity filter.

Table 3.1 shows relative binding free energies and coordination numbers within  $\pm 2.5 \text{ \AA}$  of the binding site at  $z = 8 \text{ \AA}$ . Independently of external electric field strength, with a median coordination number of 5.5,  $\text{K}^+$  had around two additional coordinating water molecules compared to 3.6 around  $\text{Na}^+$ . This larger desolvation of  $\text{Na}^+$  could possibly explain the more negative relative binding free energies found for  $\text{Na}^+$  ( $-3.1$  to  $-4.0$ ) compared to  $\text{K}^+$  ( $-1.6$  to  $-2.8$ ). Considering both Lys1244 charges and all field strengths,  $G(z)$  was by  $-1.1$  to  $-1.5$  lower for  $\text{Na}^+$  than  $\text{K}^+$ . This difference in relative binding free energy represents an additional energy barrier which  $\text{Na}^+$  has to overcome in order to leave the binding site. Therefore, the more attractive interactions around  $z = 8 \text{ \AA}$  have a greater inhibitory effect on  $\text{Na}^+$  flux through the channel which is reflected in the higher  $\text{K}^+$  permeation ratios.

Table 3.1: Average relative binding free energies,  $G(z)$ , and coordination numbers for  $\text{Na}^+$  and  $\text{K}^+$  within  $z = 8 \pm 2.5 \text{ \AA}$  at different electric field strengths,  $U_z$ , and Lys1244 charge,  $q_{\text{Lys1244}}$ .

$q_{\text{Lys1244}} [e]$	$U_z [\text{mV}]$	$G(z)$		Coord. Number	
		$\text{Na}^+$	$\text{K}^+$	$\text{Na}^+$	$\text{K}^+$
0.5	0	-3.1	-1.8	3.8	5.5
	100	-3.5	-2.0	3.6	5.5
	300	-3.7	-2.6	3.5	5.4
	500	-3.9	-2.8	3.5	5.6
1.0	0	-3.1	-1.6	3.6	5.6
	100	-3.3	-2.2	3.7	5.5
	300	-3.7	-2.3	3.6	5.5
	500	-4.0	-2.5	3.6	5.6

For  $\text{K}^+$  no well-defined sites of dehydration were identified in the selectivity filter region. As for  $\text{Na}^+$ , increased dwell-times could be observed at the channel entrance and exit. However, relative binding free energies and hydration numbers were always above those of  $\text{Na}^+$ , which explains the higher  $\text{K}^+$ -current.

These results illustrate that conductance and selectivity appear to be product of various terms including the energetic cost of (de-)hydration, polarisation effects, possibly also involving charge-transfer, and finally

spatial constriction and rigidity of the selectivity filter architecture. The employed force field parametrisation appears incapable of capturing this complex combination of contributions.

A more precise description of the identified binding regions was obtained from the correlation between probability densities on all three coordinate axes.  $x$ ,  $y$  and  $z$  density profiles of  $\text{Na}^+$  and  $\text{K}^+$  of the simulations with  $U_z = 100 \text{ mV}$  and  $q_{\text{Lys1244}} = 1.0 e$  were discretised into  $1 \text{ \AA}$  sized bins. Ordering these by their common occurrence, four potential binding pockets, (P1-4; see figure 3.15) could be identified (see appendix C, table C.2 and figure 3.16). The center of each pocket was determined from a probability weighted average over the coordinates of all bins belonging to a given site. All high density regions on the individual coordinate axes could be associated to a binding site.

Except for site P4, where  $I_z < 0.001 e \text{ ns}^{-1}$  was obtained, no conductance within cylindrical regions of  $3 \text{ \AA}$  radius and  $6 \text{ \AA}$  height around the identified binding pockets.

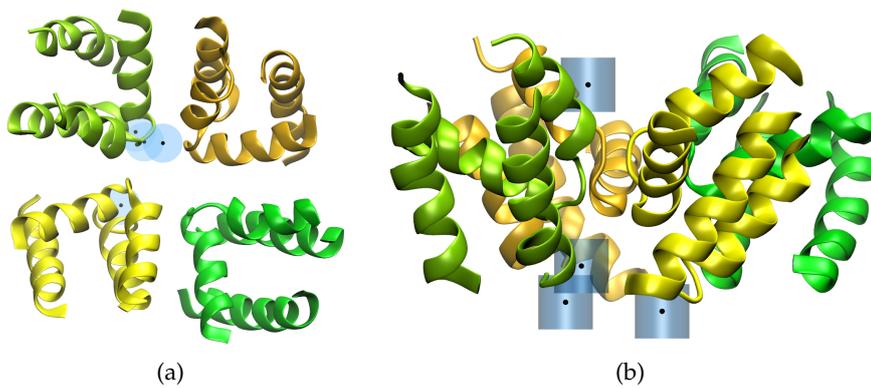


Figure 3.15: Top view (a) and side view (b) of binding pockets and selection cylinders (centers indicated by black spheres).

The first pocket, P1, was found at  $8 \text{ \AA}$  above the  $\text{COG}_{\text{DEKA}}$  (figure 3.17). It is situated between the negatively charged of the Glu764 and the Asp1248 side-chains which lie above the selectivity filter's Lys1244 and Glu761. In addition to repulsion arising from interactions with Lys1244, there is ions are likely also hindered from leaving the site by the Arg756 residue. Strong attractive  $\text{Na}^+$ -protein interactions in the P1 region are likely to limit ion passage rates.

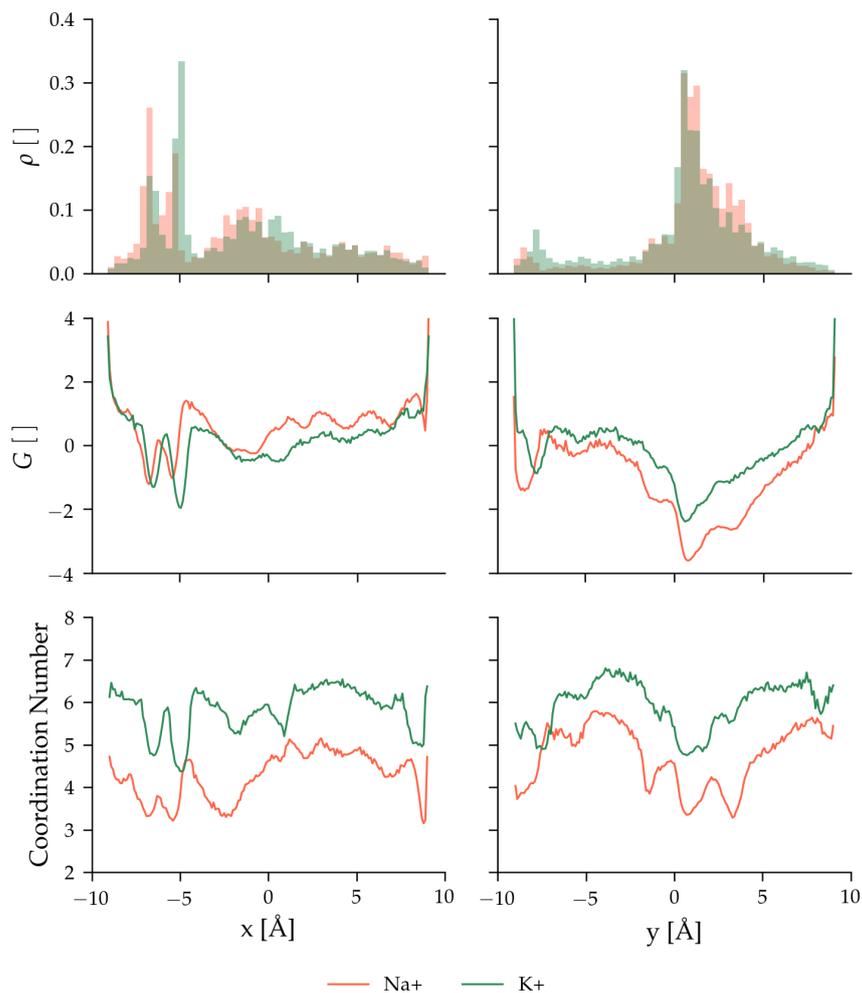


Figure 3.16: Probability densities, relative binding free energies and hydration numbers for  $\text{Na}^+$  and  $\text{K}^+$  ions along the  $x$ - and  $y$ -axes during simulations with  $U = 100 \text{ mV}$  and  $q_{\text{Lys1244}} = 1.0 e$

Figures 3.18 and 3.19 illustrate the binding regions P2-4 which are located at the exit of the pore.

P2 was located in a loop formed by residues 435 to 439. The backbone oxygen atoms were likely responsible for elevated dwell times in this region. Unlike in the real channel, in the employed pore model, Ile439 corresponded to a terminal residue. The increased binding affinity might therefore be connected to the presence of the additional electronegative carboxyl end-group.

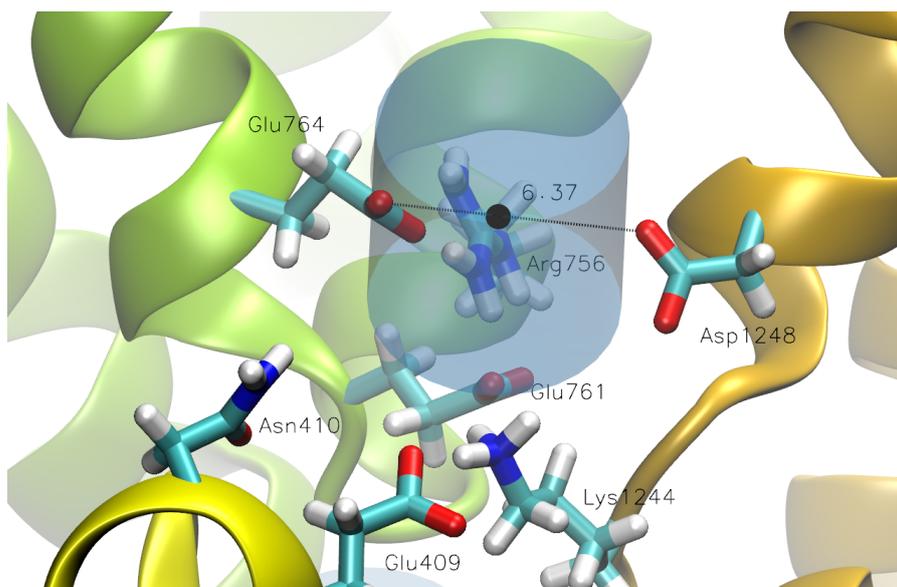


Figure 3.17: Selection cylinder and residues around binding site P1 (center indicated by black sphere).

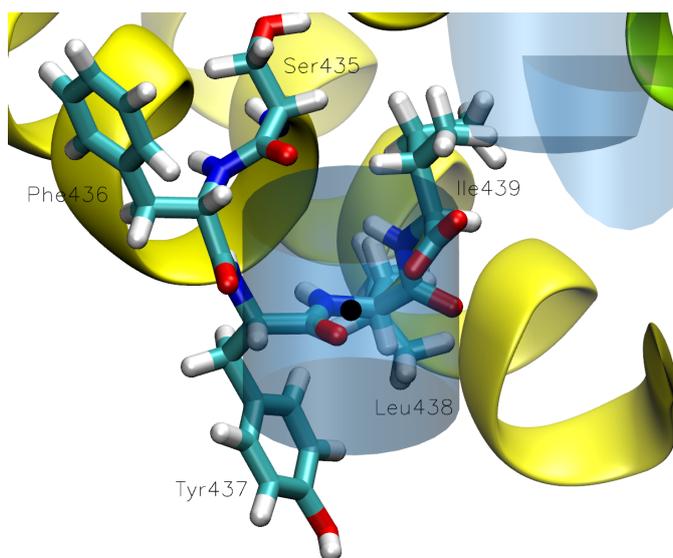


Figure 3.18: Selection cylinder and residues around binding site P2 (center indicated by black sphere).

Similarly, binding pockets P3 and P4 were found in a loop formed by residues 790 to 793 with a terminal carboxylic acid functionality on the Val793 residue. The high dwell-times found in the present simulations might therefore be artifacts of the simplified pore model and without physical significance.

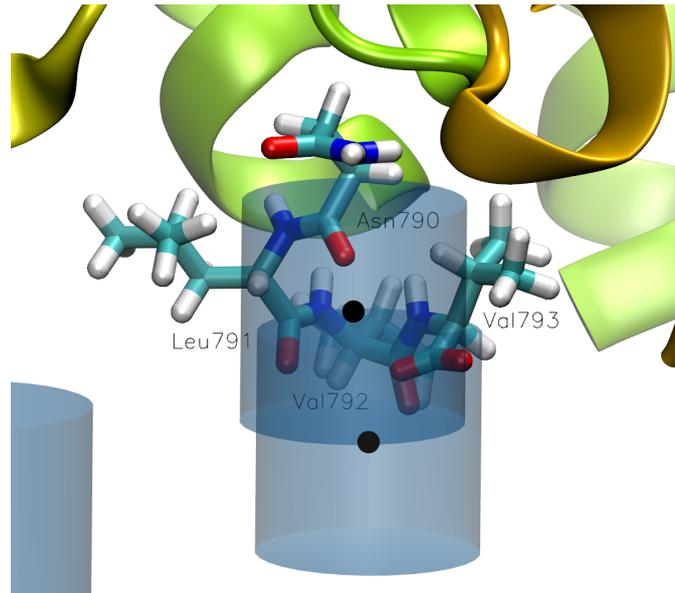


Figure 3.19: Selection cylinder and residues around binding sites P3 and P4 (centers indicated by black spheres).

The top view on the selectivity filter region with the indicated high density sites (figure 3.15) shows that the P1-P4 were located above or below the two acidic Asp406 and Glu761 residues. According to Pan et al. [7], together with Glu764, these two residues are responsible for attracting ions from the bulk solvent. Being positioned above the basic Lys1244 residue, they define a point of constriction from selectivity originates.

The  $x$  and  $y$  location of the binding sites suggests that ions are traversing the channel along the domains I and II rather than in the proximity to Ala1536. This assumption is supported by the relation between Lys1244 density and observed ion flux. As described in section 3.3, higher ionic currents were observed when the Lys1244  $\epsilon$ -ammonium group was more likely to be found at large  $xy$  separations from Asp406 and Glu761. Additionally, density profiles in sites P2-P4 exhibited peaks at different voltages which lets assume that these regions trapped ions which had crossed the channel instead of only attracting them from the bulk solvent region.

## CONCLUSIONS

---

In our investigation we used classical MD simulations to investigate the effects of variant electric field strength on ion conduction and selectivity in a truncated model of the human Na<sub>v</sub>1.4 channel. The molecular dynamics of this simplified system were validated against a larger pore model inside a POPC membrane. Additionally, to coarsely model polarisation effects on the selectivity filter's Lys1244 residue, the positive charge on the Lys1244 ε-ammonium group was reduced by 0.5 *e*.

Stronger external electric potentials were found to increase separation between the acidic and basic DEKA residues of the selectivity filter. This trend was also reflected in ion conductance which augmented in function of electric field strength. At all voltages, K<sup>+</sup> flux outnumbered that of Na<sup>+</sup>.

Decreasing  $q_{Lys1244}$  could not improve selectivity in ion permeation. Accordingly, in the selectivity filter region (cylinder III), best ratios of Na<sup>+</sup>/K<sup>+</sup> (0.7) were found at 100 mV and full Lys1244 charge.

Due to strong attractive interactions between acidic and basic DEKA residues, the voltage-dependent dilatation of the constriction site, defined by Asp406, Glu761 and Lys1244, was greater with a reduced lysine charge. In addition to the decreased spatial constriction, also the number of coordinating waters in the ions' first hydration layer was found higher when  $q_{Lys1244} = 0.5 e$ . A higher coordination number in the selectivity filter region could be associated to increased K<sup>+</sup> current.

Along the *z*-axis, which corresponded to the direction of ion flux, we could identify three regions in which Na<sup>+</sup> coordination numbers dropped from 6 to 3 (and from 7 to 4 for K<sup>+</sup>). The same sites, defined maxima in ionic density distributions, and in all of them, K<sup>+</sup> binding was weaker than for Na<sup>+</sup> and total coordination numbers of K<sup>+</sup> never dropped below those of Na<sup>+</sup>. Using these density profiles, and their relation to the relative binding free energies, four binding pockets, one at the pore entrance (P1) and three at the exit (P2–4) could be identified. In these regions, no, or almost no, ion flux was registered. Overestimated Na<sup>+</sup>-protein interactions with P1 are one possible explanation for the nonphysical conduction ratios.

The here obtained incoherence in Na<sup>+</sup>/K<sup>+</sup> conductance ratios with experimental data led to the conclusion that the employed force field parametrisation is was not suitable to reproduce experimentally observed selectivity in the human Na<sub>v</sub>1.4 channel. [46]

Selectivity in eukaryotic Na<sub>v</sub> channels seems to be tightly linked to the presence of a conserved lysine residue in the selectivity filter. [13] In the present simulations, out of the four DEKA sidechains, the basic Lys1244 functionality displayed the greatest mobility which highlights the importance of a proper description of the associated dynamic behaviour. A more accurate replication of selectivity in future simulations, should be achieved with adequate interaction parameters for this residue and an account for polarisation effects.

## BIBLIOGRAPHY

---

- (1) Oakes, V.; Furini, S.; Domene, C. *Curr. Top. Membr.* **2016**, *78*, 183–214.
- (2) Hille, B.; Catterall, W. A. In *Basic Neurochem.* Elsevier: 2012, pp 63–80.
- (3) Hille, B., *Ion Channels of Excitable Membranes*, 3rd ed.; Sinauer Associates Inc.: Sunderland, 2001.
- (4) Catterall, W. A. *J. Physiol.* **2012**, *590*, 2577–2589.
- (5) Bean, B. P. *Nat. Rev. Neurosci.* **2007**, *8*, 451–465.
- (6) Catterall, W. A. *Neurochem. Res.* **2017**, *42*, 2495–2504.
- (7) Pan, X. et al. *Science (80-. )*. **2018**, *362*, eaau2486.
- (8) Goldin, A. L. *Annual review of physiology* **2001**, *63*, 871–894.
- (9) Corry, B.; Thomas, M. J. *Am. Chem. Soc.* **2012**, *134*, 1840–1846.
- (10) Callahan, K. M.; Roux, B. J. *Phys. Chem. B* **2018**, *122*, 10126–10142.
- (11) Furini, S.; Domene, C. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 5450–5455.
- (12) Nogueira, J. J.; Corry, B. In *The Oxford Handbook of Neuronal Ion Channels*.
- (13) Favre, I.; Moczydlowski, E.; Schild, L. *Biophys. J.* **1996**, *71*, 3110–3125.
- (14) Dudev, T.; Lim, C. *Acc. Chem. Res.* **2014**, *47*, 3580–3587.
- (15) Huang, J. et al. *Nat. Methods* **2017**, *14*, 71–73.
- (16) Phillips, J. C. et al. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (17) Frenkel, D.; Smit, B. *Understanding molecular simulation: From algorithms to applications*, 2002.
- (18) Leach, A. R.; Leach, A. R., *Molecular modelling: principles and applications*; Pearson education: 2001.
- (19) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
- (20) Woodcock, L.-V. *Chemical Physics Letters* **1971**, *10*, 257–261.
- (21) Evans, D. J. *The Journal of Chemical Physics* **1983**, *78*, 3297–3302.
- (22) Berendsen, H. J. et al. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (23) Thijssen, J., *Computational physics*; Cambridge University Press: 1999.
- (24) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 8577.

- (25) Feller, S. E. et al. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (26) Cisneros, G. A. et al. *Chem. Rev.* **2014**, *114*, 779–814.
- (27) Vanommeslaeghe, K.; Mackerell, A. D. CHARMM additive and polarizable force fields for biophysics and computer-aided drug design, 2015.
- (28) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33–38.
- (29) Lomize, M. A. et al. *Nucleic Acids Res.* **2012**, *40*, D370–D376.
- (30) Jo, S. et al. *J. Comput. Chem.* **2008**, *29*, 1859–1865.
- (31) Lee, J. et al. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
- (32) Wu, E. L. et al. *J. Comput. Chem.* **2014**, *35*, 1997–2004.
- (33) Jo, S.; Kim, T.; Im, W. *PLoS One* **2007**, *2*, ed. by Yuan, A., e880.
- (34) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (35) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (36) Best, R. B. et al. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (37) Klauda, J. B. et al. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
- (38) Jorgensen, W. L. et al. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (39) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
- (40) Michaud-Agrawal, N. et al. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (41) Gowers, R. et al. In, 2016, pp 98–105.
- (42) Gumbart, J. et al. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818*, 294–302.
- (43) Hille, B. *J. Gen. Physiol.* **1975**, *66*, 535–560.
- (44) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475–1498.
- (45) Favre, I.; Moczydlowski, E.; Schild, L. *Biophysical Journal* **1996**, *71*, 3110–3125.
- (46) Hodgkin, A. L.; Huxley, A. F. *J. Physiol.* **1952**, *117*, 500–544.
- (47) Mähler, J.; Persson, I. *Inorg. Chem.* **2012**, *51*, 425–438.
- (48) Yang, Z. H. *RSC Adv.* **2015**, *5*, 1213–1219.

## APPENDICES





## PORE MODELS

---

Table A.1: Selected residues for the pore models A and D.

Domain	Chain ID	Residues
I	A; B	234 – 286; 336 – 451
II	C	683 – 805
III	D	1143 – 1298
IV	E	1464 – 1601

Table A.2: Selected residues for the pore model B.

Domain	Chain ID	Residues
I	A	392 – 420
II	B	747 – 772
III	C	1230 – 1255
IV	D	1522 – 1548

Table A.3: Selected residues for the pore model C.

Domain	Chain ID	Residues
I	A	392 – 420
II	B	747 – 772
III	C	1230 – 1255
IV	D	1522 – 1548
I	E; F	254 – 271; 423 – 439
II	G; H	707 – 724; 775 – 793
III	I; J	1164 – 1182; 1269 – 1284
IK; L		1486 – 1504; 1571 – 1589

Table A.4: Numbers of atoms of the different pore models.

System		Numbers of atoms						
		Protein	Membrane	Water	Ions			Total
					Na <sup>+</sup>	K <sup>+</sup>	Cl <sup>-</sup>	
Large	A	9328	26800	61890	35	34	55	98142
	B	1746	–	61908	32	32	58	63776
Small	C	4193	–	50265	26	26	47	64557
	D	9328	–	73944	42	42	70	83426

Table A.5: Simulation box dimensions of the different pore models.

System		Box lengths [Å]		
		x	y	z
Large	A	95.5	95.5	117.2
	B	96.1	95.5	72.6
Small	C	96.8	89.2	66.4
	D	94.8	90.2	101.4

# B

## FORCE CONSTANTS OF CONSTRAINTS AND RESTRAINTS DURING EQUILIBRATION RUNS

Table B.1: Constraints and restraints placed on the protein and lipid membrane in system A and on the protein in system D during the equilibration runs.

Run	Force constants [kcal mol <sup>-1</sup> Å <sup>2</sup> ]		
	Protein constraints	Lipid restraints	
		Dihedral and improper	Planar
1	10.0	500	5.0
2	5.0	200	5.0
3	5.0	100	2.0
4	2.0	100	1.0
5	1.0	50	0.2
6	0.1	0	0.1



## PROTEIN GEOMETRIES

Table C.1: Mean  $R_{mean}$  and maximum  $R_{max}$  distances between the Asp406 and Lys1244  $C_{\alpha}$  atoms.

$q_{Lys1244}$ [e]	$U_{ext}$ [mV]	$R_{mean}$ [Å]	$R_{max}$ [Å]
0.5	0	13.6	15.2
	100	13.6	15.0
	300	13.9	15.3
	500	13.6	15.0
1.0	0	13.8	15.4
	100	13.6	15.0
	300	13.7	15.1
	500	14.2	15.5

Table C.2: Coordinates of the identified binding pockets P1-P4.

Binding Pocket	$x$ [Å]	$y$ [Å]	$z$ [Å]
P1	-2.8	3.1	7.9
P2	-0.4	-8.2	-17.2
P3	-6.8	0.9	-15.6
P4	-5.4	1.1	-12.2

Table C.3: Minimum and maximum z-coordinates ( $r_{z, min}$ ,  $r_{z, max}$ ) of the selectivity filter region with respect to the center of geometry formed by the DEKA residues'  $C_{\alpha}$  atoms.

$q_{Lys1244}$ [ $e$ ]	$U_{ext}$ [mV]	$r_{z, min}$ [ $\text{\AA}$ ]	$r_{z, max}$ [ $\text{\AA}$ ]
0.5	0	5.6	-7.5
	100	5.8	-7.8
	300	6.1	-7.3
	500	5.9	-8.4
1.0	0	6.1	-7.9
	100	5.9	-8.1
	300	5.8	-8.1
	500	7.0	-7.6

Table C.4: Minimum and maximum z-coordinates ( $r_{r, min}$ ,  $r_{z, max}$ ) of protein model with respect to the center of geometry formed by the DEKA residues'  $C_{\alpha}$  atoms.

$q_{Lys1244}$ [ $e$ ]	$U_{ext}$ [mV]	$r_{z, min}$ [ $\text{\AA}$ ]	$r_{z, max}$ [ $\text{\AA}$ ]
0.5	0	14.9	-19.9
	100	14.7	-20.0
	300	14.8	-20.0
	500	14.8	-19.9
1.0	0	14.7	-20.0
	100	14.8	-20.0
	300	14.8	-20.0
	500	14.5	-20.1

## CONDUCTANCE DATA

Table D.1: Average Na<sup>+</sup> and K<sup>+</sup> currents in the entire simulation box (figure 3.2) within 500 ns.

Ion Species	$q_{Lys1244}$ [ $e$ ]	$U$ [mV]	$I_z$ [ $e \text{ ns}^{-1}$ ]
Na+	0.5	0	-0.1
		100	1.9
		300	6.6
		500	11.2
	1.0	0	0.0
		100	2.3
		300	6.8
		500	11.6
K+	0.5	0	-0.1
		100	4.2
		300	13.1
		500	22.7
	1.0	0	0.0
		100	4.1
		300	13.2
		500	23.0

Table D.2: Average Na<sup>+</sup> and K<sup>+</sup> currents in the cylinder selection Z1 (figure 3.4 and 3.5) within 500 ns.

Ion Species	$q_{Lys1244}$ [ $e$ ]	$U$ [mV]	$I_z$ [ $e \text{ ns}^{-1}$ ]
Na <sup>+</sup>	0.5	0	-0.01
		100	0.07
		300	0.26
		500	0.43
	1.0	0	0.01
		100	0.08
		300	0.29
		500	0.41
K <sup>+</sup>	0.5	0	-0.10
		100	0.12
		300	0.37
		500	0.63
	1.0	0	0.01
		100	0.12
		300	0.32
		500	0.63

Table D.3: Average Na<sup>+</sup> and K<sup>+</sup> currents in cylinder selection Z2 (figure 3.6 and 3.7) within 500 ns.

Ion Species	$q_{Lys1244}$ [e]	$U$ [mV]	$I_z$ [ $e \text{ ns}^{-1}$ ]
Na+	0.5	0	0.03
		100	0.03
		300	0.06
		500	0.07
	1.0	0	0.03
		100	0.04
		300	0.04
		500	0.11
K+	0.5	0	0.04
		100	0.07
		300	0.13
		500	0.13
	1.0	0	0.08
		100	0.10
		300	0.10
		500	0.18

Table D.4: Average Na<sup>+</sup> and K<sup>+</sup> currents in cylinder selection Z3 (figure 3.6 and 3.8) within 500 ns.

Ion Species	$q_{Lys1244}$ [ $e$ ]	$U$ [mV]	$I_z$ [ $e \text{ ns}^{-1}$ ]
Na <sup>+</sup>	0.5	0	0.06
		100	0.07
		300	0.07
		500	0.09
	1.0	0	0.08
		100	0.09
		300	0.07
		500	0.09
K <sup>+</sup>	0.5	0	0.11
		100	0.12
		300	0.15
		500	0.14
	1.0	0	0.15
		100	0.12
		300	0.14
		500	0.20

ION DENSITY PROFILES, RELATIVE BINDING FREE ENERGIES AND WATER COORDINATION NUMBERS

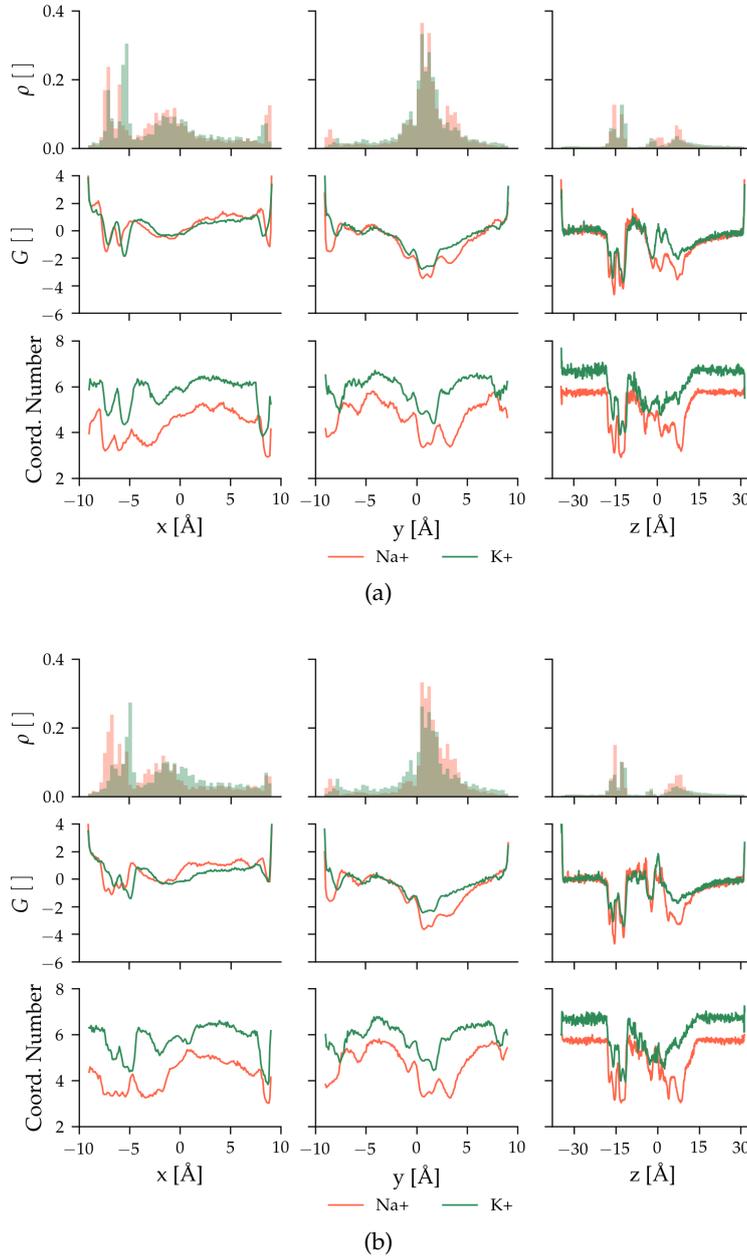


Figure E.1: Probability densities, relative binding free energies and coordination numbers for  $\text{Na}^+$  and  $\text{K}^+$  ions along the  $x$ ,  $y$  and  $z$ -axes during simulations with  $U = 0 \text{ mV}$  and (a)  $q_{Lys1244} = 0.5 e$ , (b)  $q_{Lys1244} = 1.0 e$ .

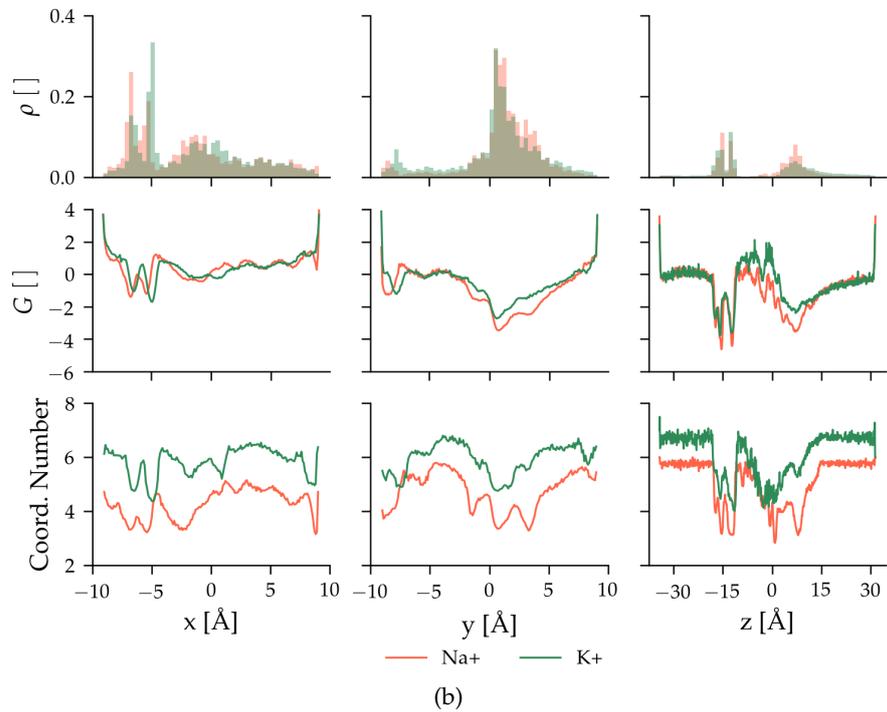
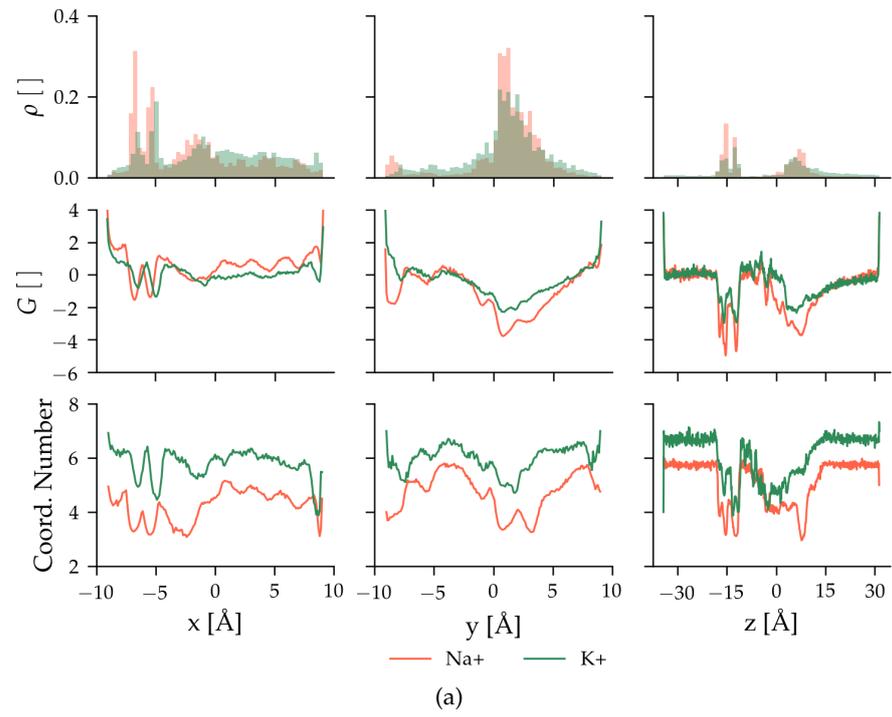


Figure E.2: Probability densities, relative binding free energies and coordination numbers for Na<sup>+</sup> and K<sup>+</sup> ions along the  $x$ ,  $y$  and  $z$ -axes during simulations with  $U = 100$  mV and (a)  $q_{Lys1244} = 0.5 e$ , (b)  $q_{Lys1244} = 1.0 e$ .

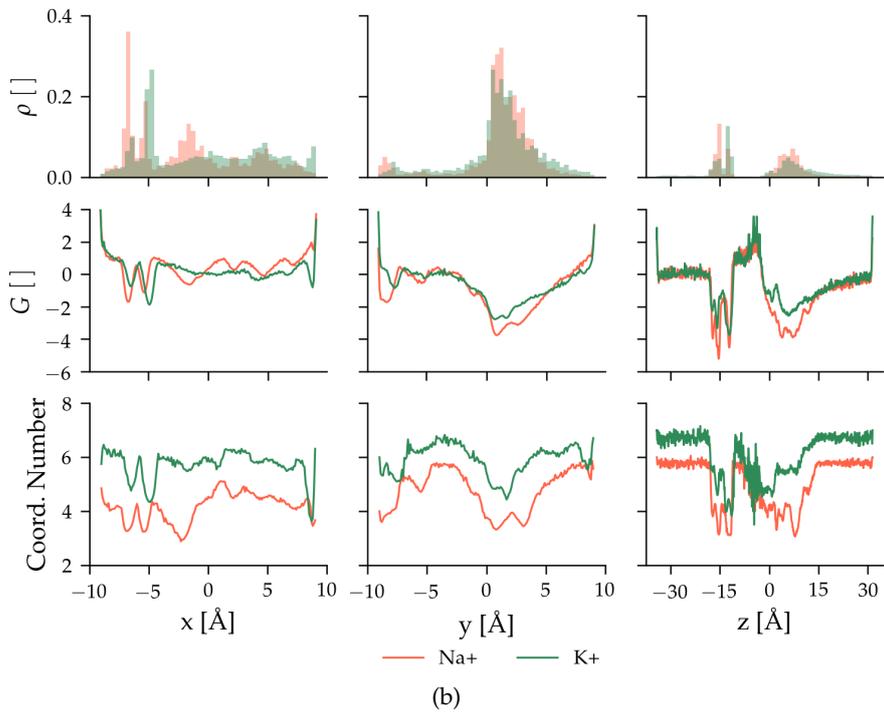
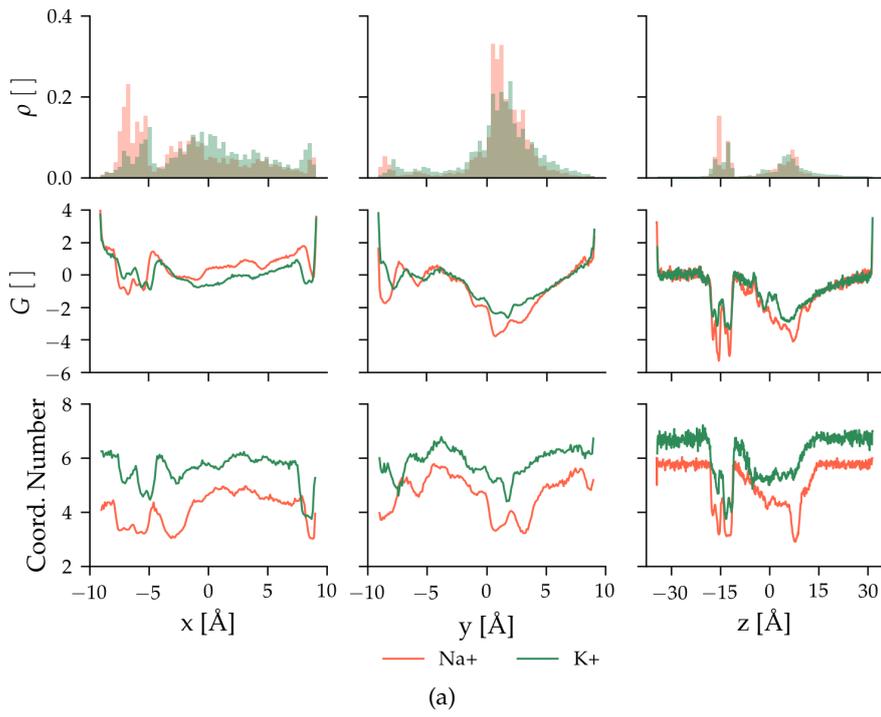


Figure E.3: Probability densities, relative binding free energies and coordination numbers for Na<sup>+</sup> and K<sup>+</sup> ions along the  $x$ ,  $y$  and  $z$ -axes during simulations with  $U = 300$  mV and (a)  $q_{Lys1244} = 0.5 e$ , (b)  $q_{Lys1244} = 1.0 e$ .

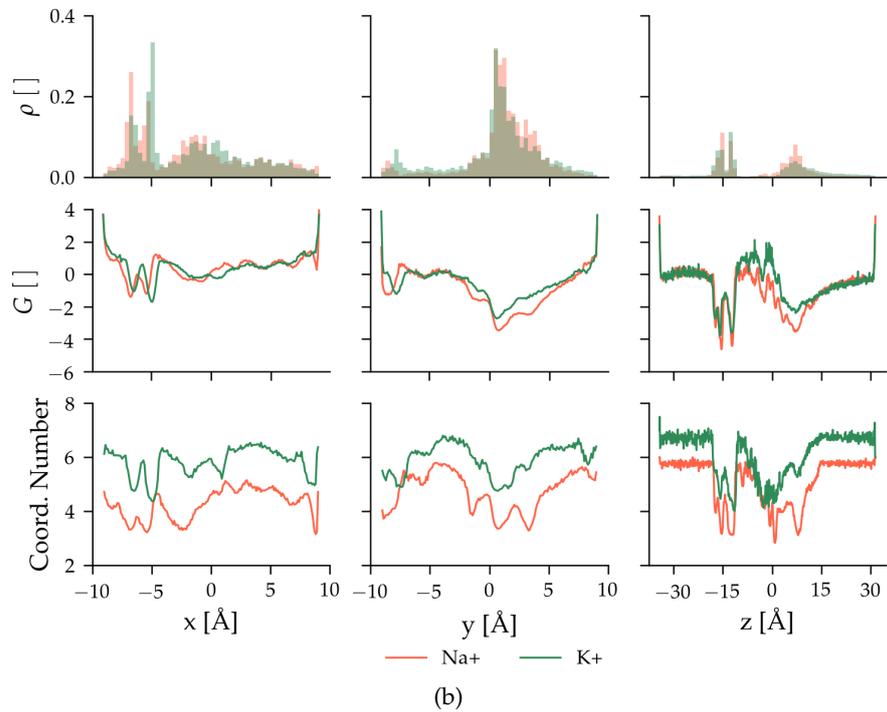
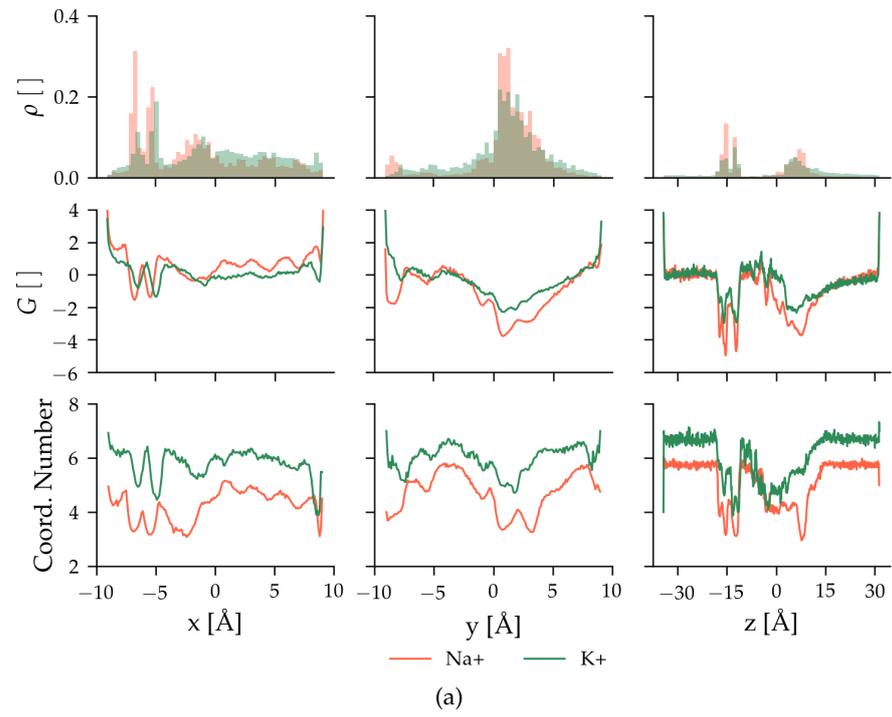


Figure E.4: Probability densities, relative binding free energies and coordination numbers for Na<sup>+</sup> and K<sup>+</sup> ions along the  $x$ ,  $y$  and  $z$ -axes during simulations with  $U = 500$  mV and (a)  $q_{Lys1244} = 0.5 e$ , (b)  $q_{Lys1244} = 1.0 e$ .

Table E.1: Average relative binding free energies,  $G(z)$ , and coordination numbers for  $\text{Na}^+$  and  $\text{K}^+$  within  $z(\text{COG}_{DEKA}) \pm 2.5 \text{ \AA}$ .

$q_{\text{Lys1244}} [e]$	$U_z [\text{mV}]$	$G(z)$		Coord. Number	
		$\text{Na}^+$	$\text{K}^+$	$\text{Na}^+$	$\text{K}^+$
0.5	0	-2.5	-0.9	4.5	5.2
	100	-1.5	-0.1	4.2	4.8
	300	-2.5	-1.6	4.4	5.3
	500	-3.2	-2.5	4.1	5.0
1.0	0	-0.3	0.1	4.6	5.0
	100	-1.2	0.5	3.9	4.7
	300	-1.8	-0.9	4.1	4.9
	500	-3.0	-2.1	4.1	5.1



## DENSITY PROFILES OF THE SELECTIVITY FILTER RESIDUES

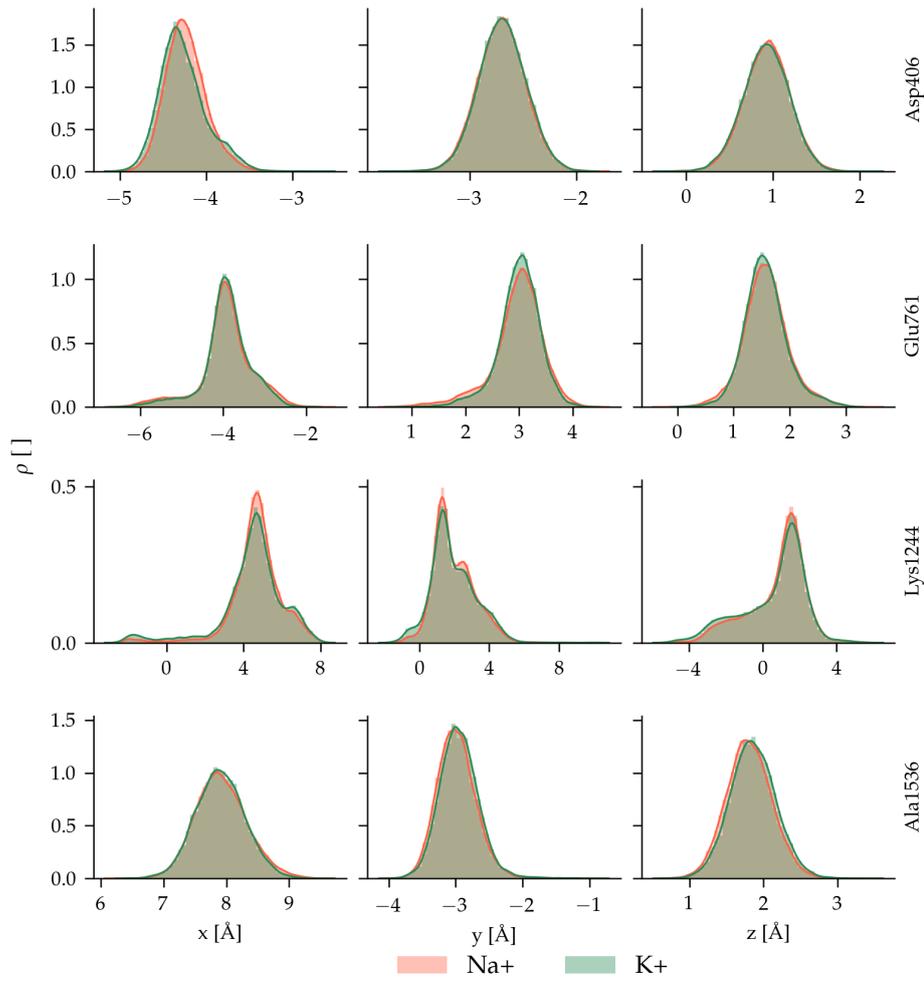


Figure F.1: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 0.5 e$  and  $U_z = 0$  mV and  $Na^+$  or  $K^+$  within  $10 \text{ \AA}$  of  $COG_{Lys1244 \epsilon-NH_3^+}$ .

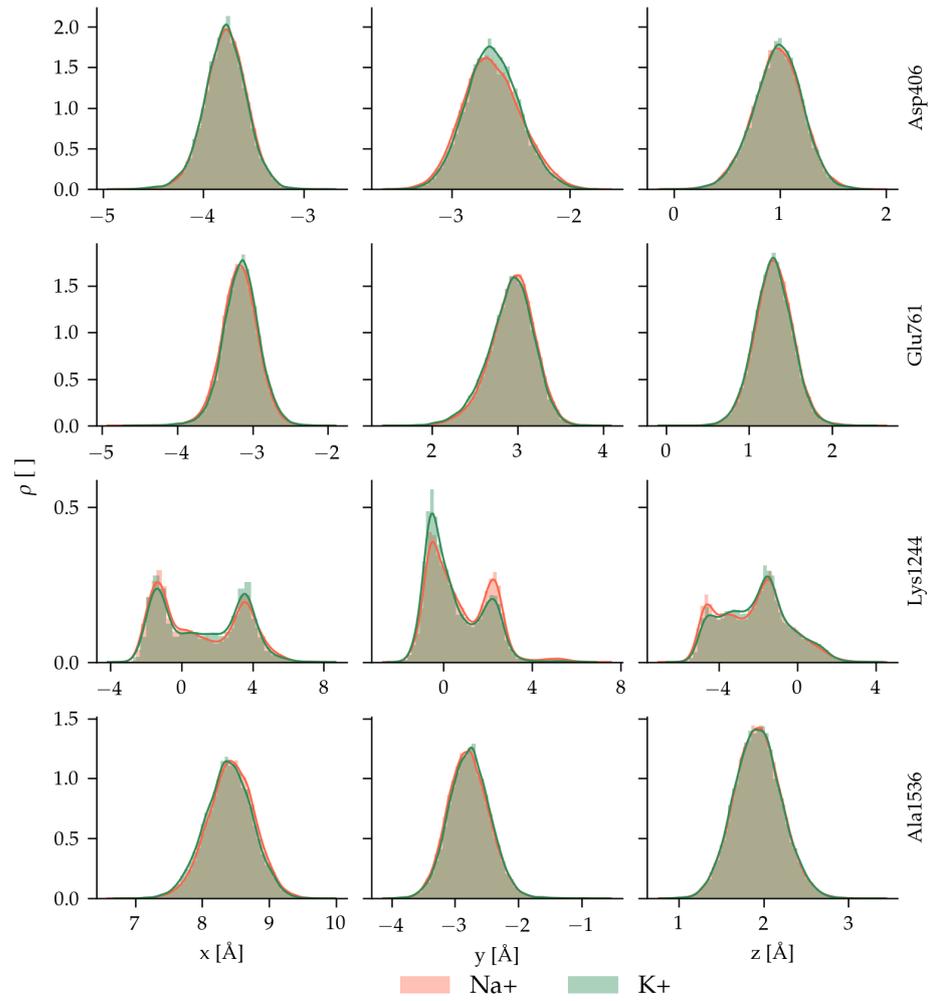


Figure F.2: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 0.5e$  and  $U_z = 100$  mV and Na<sup>+</sup> or K<sup>+</sup> within 10 Å of COG<sub>Lys1244  $\epsilon$ -NH<sub>3</sub><sup>+</sup></sub>.

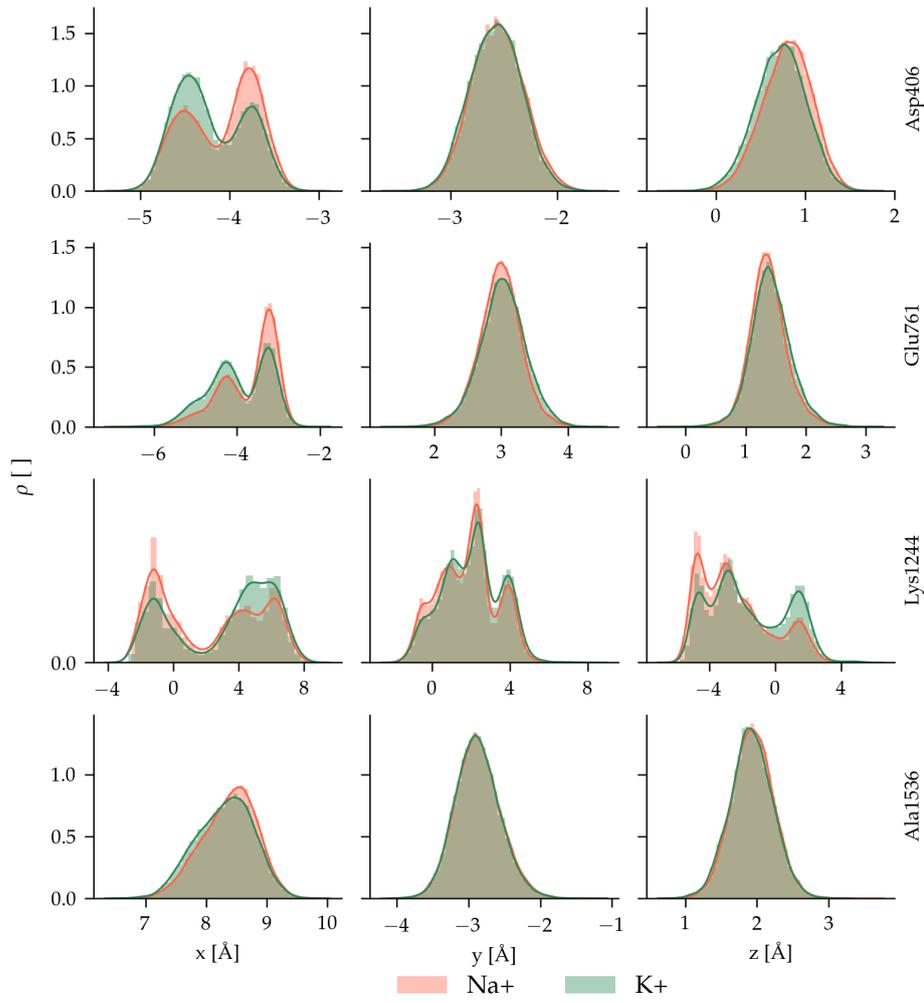


Figure F.3: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 0.5 e$  and  $U_z = 300$  mV and Na<sup>+</sup> or K<sup>+</sup> within 10 Å of COG<sub>Lys1244  $\epsilon$ -NH<sub>3</sub><sup>+</sup></sub>.

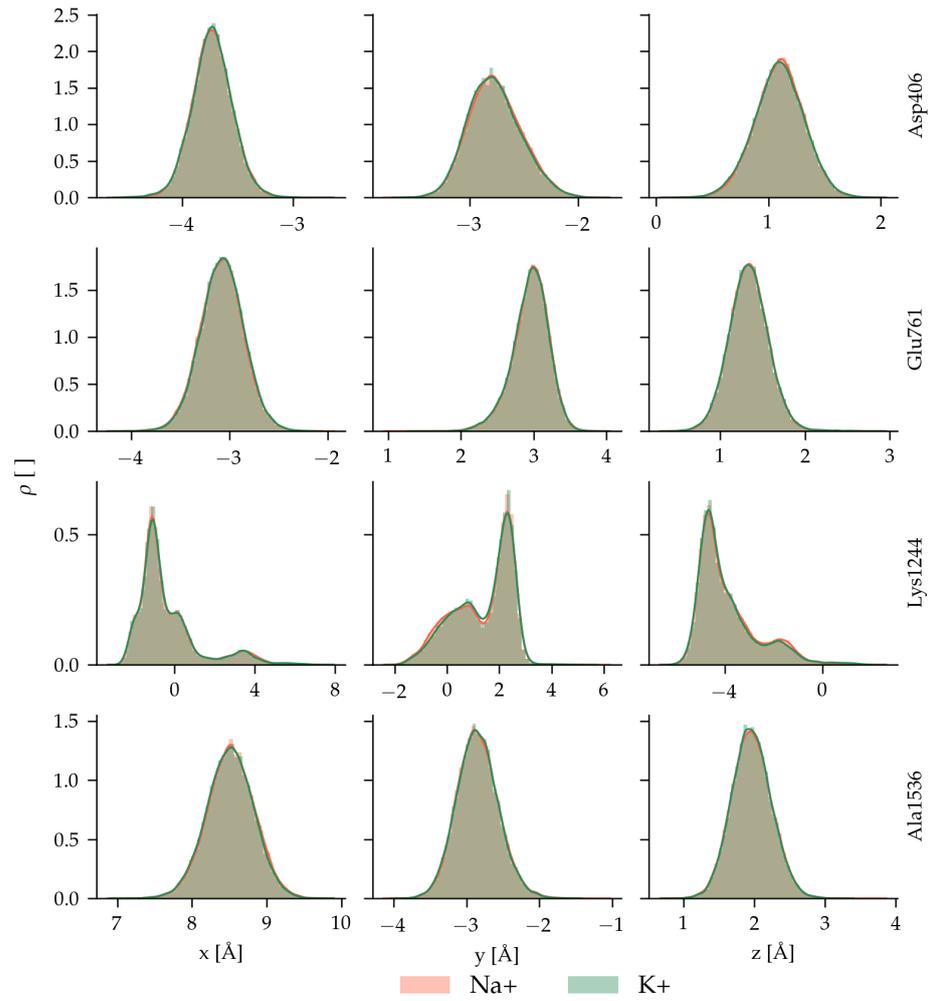


Figure F.4: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 0.5e$  and  $U_z = 500$  mV and Na<sup>+</sup> or K<sup>+</sup> within 10 Å of COG<sub>Lys1244  $\epsilon$ -NH<sub>3</sub><sup>+</sup></sub>.

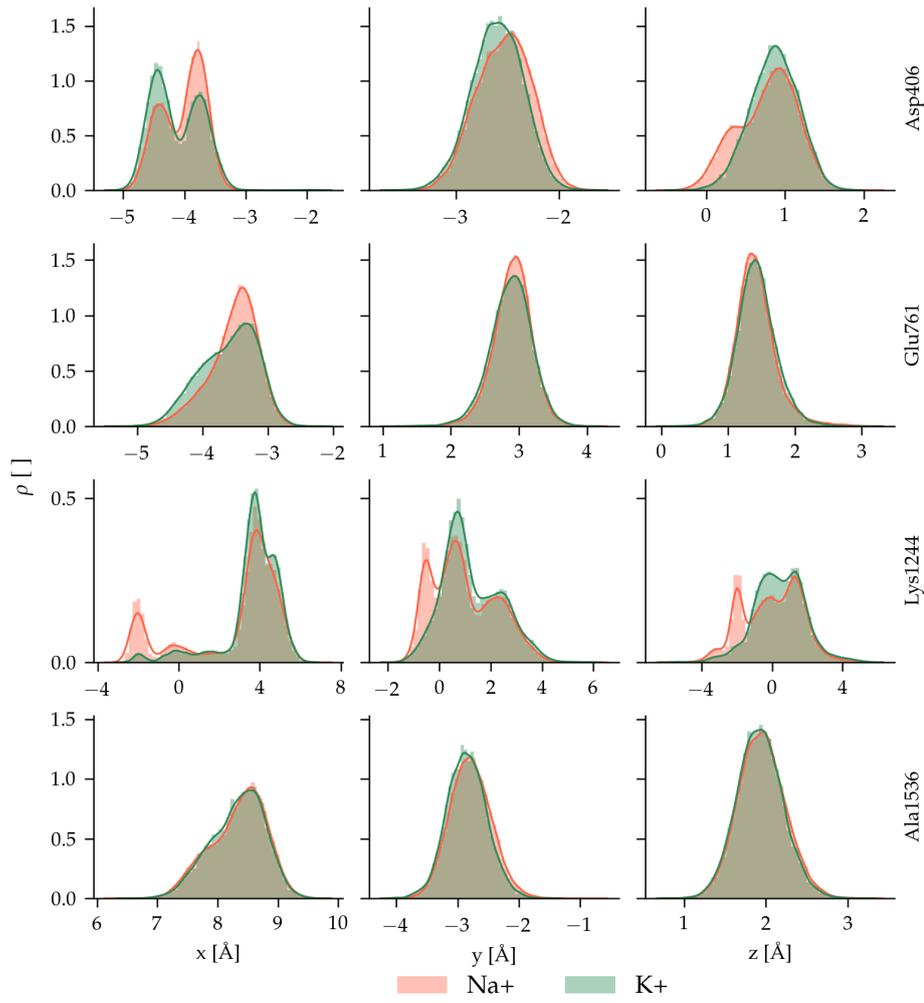


Figure F.5: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 1.0 e$  and  $U_z = 0$  mV and Na<sup>+</sup> or K<sup>+</sup> within 10 Å of COG<sub>Lys1244</sub> ε-NH<sub>3</sub><sup>+</sup>.

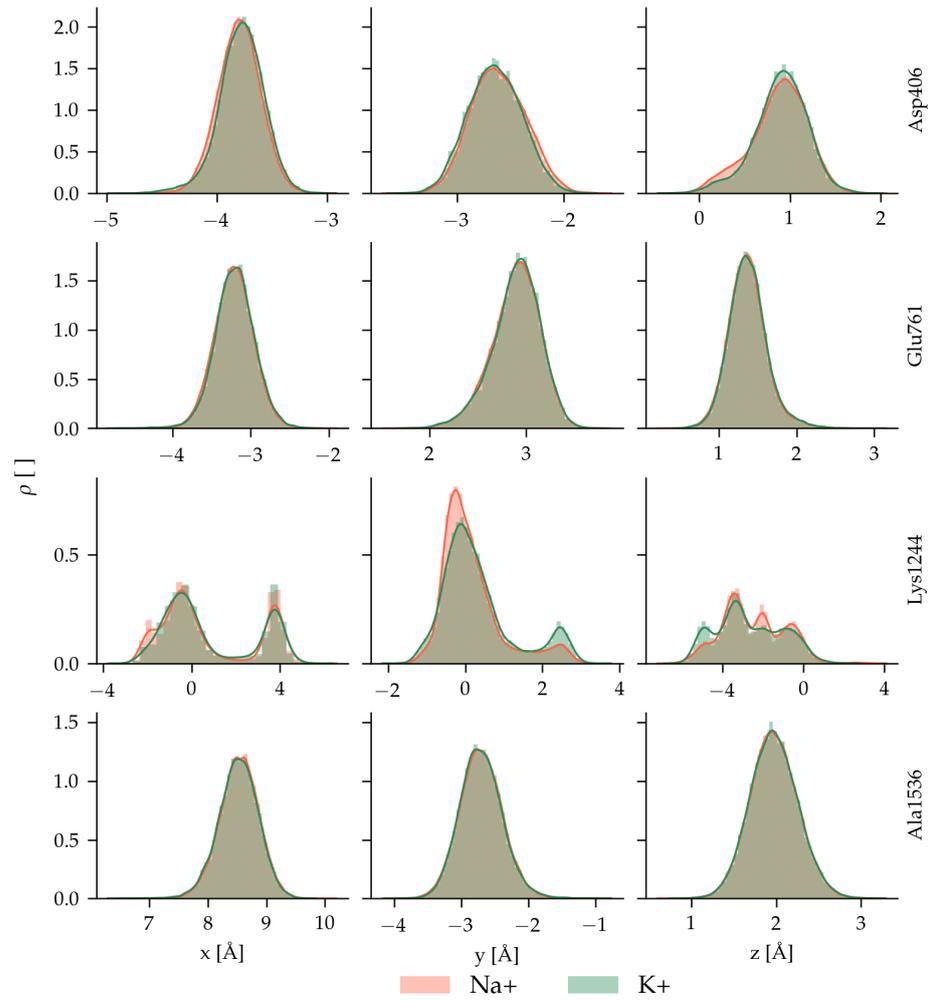


Figure F.6: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{\text{Lys1244}} = 1.0e$  and  $U_z = 300 \text{ mV}$  and  $\text{Na}^+$  or  $\text{K}^+$  within  $10 \text{ \AA}$  of  $\text{COG}_{\text{Lys1244 } \epsilon\text{-NH}_3^+}$ .

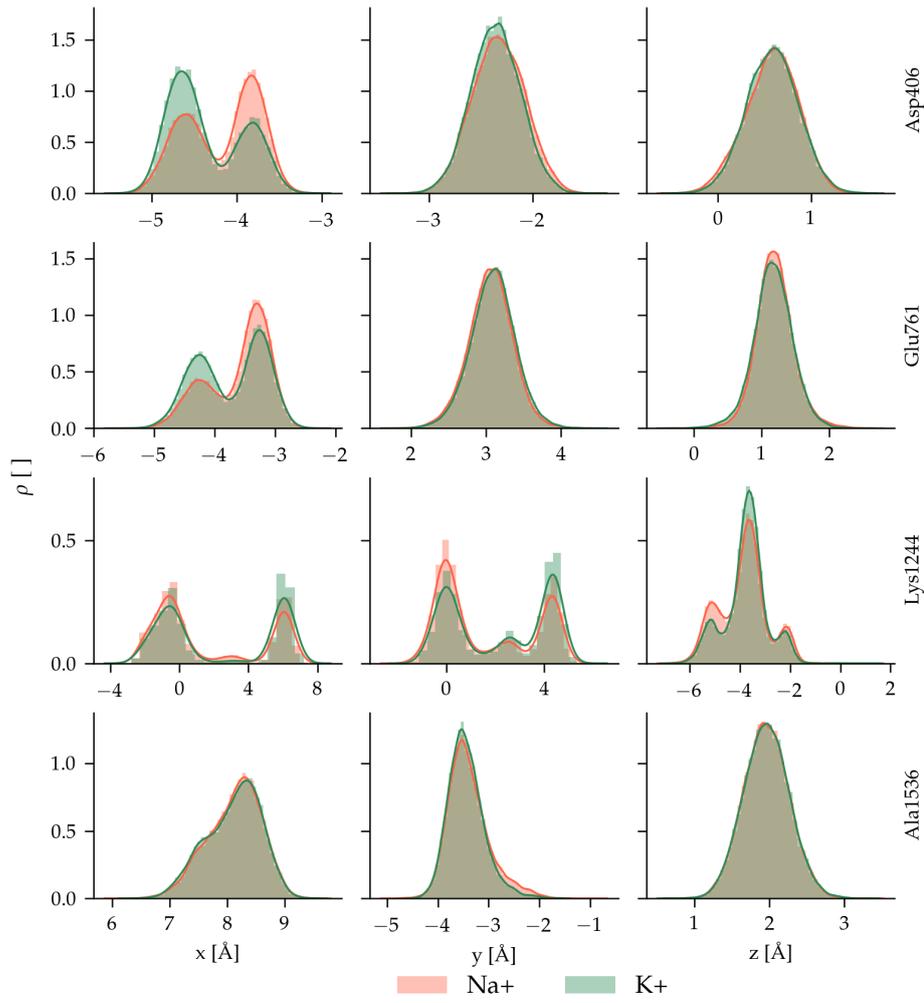


Figure F.7: Lys1244  $x$ ,  $y$  and  $z$  densities at  $q_{Lys1244} = 1.0 e$  and  $U_z = 500$  mV and Na<sup>+</sup> or K<sup>+</sup> within 10 Å of COG<sub>Lys1244  $\epsilon$ -NH<sub>3</sub><sup>+</sup></sub>.