# Molecular Dynamics Simulations and Modelling of the T cell receptor and the Lymphocyte-specific Protein Tyrosine Kinase

Dheeraj Prakaash

Thesis submitted in accordance with the requirements
for the degree of

*Doctor of Philosophy*


The University of Leeds

Leeds Institute of Cardiovascular and Metabolic Medicine

School of Medicine

**November 2021**

## Intellectual Property and Publication Statement

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution to the jointly authored work has been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

- Chapter 3 of my thesis titled "*The T cell receptor transmembrane region*" includes information from the following co-authored publication:

  Title: **Allosteric activation of T cell antigen receptor signaling by quaternary structure relaxation**.

  Journal: **Cell Reports** (doi:10.1016/j.celrep.2021.109375)

  Date published: 13 July 2021

  Authors: Anna-Lisa Lanz, Giulia Masi, Nicla Porciello, André Cohnen, Deborah Cipria, Dheeraj Prakaash, Štefan Bálint, Roberto Raggiaschi, Donatella Galgano, David K Cole, Marco Lepore, Omer Dushek, Michael L Dustin, Mark S P Sansom, Antreas C Kalli, Oreste Acuto.

  My contribution: Simulation methodology and data analysis, *in silico* investigation and visualisation, and writing (review, editing).

- Chapter 5 of my thesis titled "*Molecular dynamics and lipid interactions of the complete T cell receptor*" is a replica of my first-author publication:

  Title: **Multi-scale simulations of the T cell receptor reveal its lipid interactions, dynamics and the arrangement of its cytoplasmic region**.

  Journal: **PLOS Computational Biology** (doi:10.1371/journal.pcbi.1009232)

  Date published: 19 July 2021

  Authors: Dheeraj Prakaash, Graham P Cook, Oreste Acuto, Antreas C Kalli

  My contribution: Conceptualisation, data analysis, investigation, methodology, visualisation, and writing (original draft, review, editing).

# Acknowledgements

I would like to first thank Dr. Antreas Kalli for the opportunity to join his group and explore the world of molecular simulations under his supervision. This opportunity has presented me with a delightful experience as a researcher at the University of Leeds. From the early days of my doctoral research, when Antreas dedicated his time to training me, and until the end of my term as a PhD student, his support, patience, and advice have been invaluable. Being encouraged by him to participate in public engagement events and conferences, and especially being appreciated at various milestones of my research has been an uplifting experience. I am also grateful to Prof. Graham P Cook, my supervisor from the Leeds Institute of Medical Research at St James's, for his friendly and yet highly valuable discussions and advice throughout my term. Meetings with him have been something that I looked forward to, knowing that he will have an enlightening perspective to share. I would like to thank both my supervisors for providing their helpful suggestions to improve the quality of my thesis and for paving the way for my transitioning into an independent researcher.

Since the beginning of our collaboration with Prof. Oreste Acuto and his group members from Sir William Dunn School of Pathology at the University of Oxford, Prof. Oreste has played an exceptional role in developing my knowledge in my research topic, and also been a source of motivation and inspiration. I am grateful to him for reviewing my manuscripts and providing critical suggestions to improve it significantly. I would like to specially thank him for being extremely enthusiastic to address my scientific questions and for commencing interesting discussions in and outside of academics.

I wish to extend my special thanks to the Kalli group members who have always created a very pleasant environment in and out of the office. I would also like to thank a former Kalli group member, Dr. Dario De Vecchis, for his mentorship and friendly guidance in my first year. My supervisors, collaborators, and Kalli group members have all undeniably played a tremendous role in enhancing my overall PhD experience, and making these years a memorable time to cherish.

Finally, I thank my family for their extensive support in funding my studies, accelerating me in this direction of becoming a scientist, and sharing the joy of my achievements. Last but not least, I thank my partner, Dr. Sukruthi Jois, for her unparalleled support and motivation to overcome hurdles throughout these years, and also for her genuine effort toward understanding and appreciating my research.

# Abstract

The T cell antigen receptor (TCR-CD3) is an octameric protein complex located in the T cell plasma membrane. It plays a vital role in our adaptive immune system by recognising a wide variety of antigenic peptides attached to major histocompatibility complexes (pMHCs). Upon binding with pMHCs, the TCR-CD3 transmits a signal to its unstructured cytoplasmic region which undergoes phosphorylation by the lymphocyte-specific protein tyrosine kinase (LCK), further triggering a chain of events ultimately aiding in T cell-mediated immune response.

Despite studies conducted on the structure and function of the TCR-CD3 and LCK, we do not yet understand the initial phase of T cell activation in molecular detail. To achieve this, it is important to not only know their structure but also study their dynamic behaviour in their native environment on the nanosecond to microsecond time-scales. Given the lack of structural data on the TCR-CD3 cytoplasmic region and the full-length LCK, molecular modelling employed in this thesis has helped produce their complete models and enabled molecular dynamics simulation studies.

Supported by prior experimental evidence, the simulations conducted in this thesis have led to novel findings such as: (i) specific sites in the TCR-CD3 transmembrane region that potentially help transmit pMHC-induced signals into the cytoplasmic region, (ii) TCR-CD3 dynamics in a membrane and the arrangement of its cytoplasmic region, (iii) PIP lipid interactions and clustering around the TCR-CD3 and upon the association of LCK with the membrane, (iv) TCR-CD3 conformational changes, and (v) protein-protein/lipid interactions in the open and closed LCK conformations.

Overall, this thesis provides novel molecular-level insights into the dynamics of the TCR-CD3 and LCK, and also signifies the potential of molecular dynamics simulations in studying membrane-associated proteins. Further, this work encourages computational studies of other immunoreceptors to help understand various immune mechanisms and aid in clinical therapeutics.

# List of Presentations

- Poster: Annual Research Postgraduate Symposium at Leeds Institute of Medical Research, St James's Hospital, Leeds, UK (9th April 2019) – '*Molecular Dynamics of LCK SH domains and the T cell receptor transmembrane region*'
- Poster: The Faculty of Medicine and Health Postgraduate Research conference, University of Leeds (17th June 2019) – '*Molecular dynamics of LCK SH domains and the TCR ζζ dimer'*
- Poster: Astbury Centre for Structural Molecular Biology Annual Meeting, Buxton, UK (12th September 2019) – '*Molecular dynamics of LCK SH domains and the TCR ζζ dimer'*
- Image: Leeds Doctoral College Showcase 2019 – '*Experiencing the molecular realm of the T cell plasma membrane*' – ***Research Image of the Year award*** *(2nd place)*
- Poster: Leeds Institute of Cardiovascular and Metabolic Medicine (LICAMM) Annual Meeting, Ullswater, UK (16th - 17th January 2020) – '*Molecular simulations highlight PIP interactions of LCK and the T cell receptor'*
- Poster: Biophysical Society Annual Meeting (22nd - 26th Feb 2021) – '*Molecular Simulations Reveal the Dynamics of the complete T cell receptor in a T cell model membrane'* – ***Student Research Achievement Award***
- Poster: Leeds Doctoral College Showcase 2021 – '*Modelling and simulating LCK: a protein vital in T cell signalling*'
- Talk: Membrane Biology group, Astbury Centre for Structural Molecular Biology (16th Apr 2021) – *'Simulating the complete TCR-CD3 and LCK, and the clustering of TCR-CD3 in large membranes'*
- Flash talk and poster: EMBL-EBI Bioexcel Summer School (4th - 9th June 2021) – '*Simulating the T cell receptor and LCK to understand the initiation of T cell signalling*' – ***Poster Award***
- Poster: European Biophysics Conference (24th - 28th July 2021) – *'Molecular dynamics and lipid interactions of the complete T cell receptor in a plasma membrane'* – ***Poster Award***
- Talk: EMBL-EBI Bioexcel Student Webinar (21st September 2021) – *'Elucidating the dynamics and lipid interactions of the T cell receptor using molecular dynamics simulations and modelling'*

# TABLE OF CONTENTS

# List of Figures

**Chapter 1**



**Chapter 2**



**Chapter 3**

**Chapter 4**



**Chapter 5**



**Chapter 6**

**Chapter 7**

# List of Tables

# List of Abbreviations

| | | |
|---|---|---|
| **A** | APBS | Adaptive Poisson-Boltzmann Solver |
| | APC | Antigen Presenting Cell |
| | AT (MD) | Atomistic (Molecular Dynamics) |
| **B** | β2m | β2-microglobulin |
| | BRS | Basic Rich Stretch |
| **C** | Cα/Cβ | Constant domain of TCRα/β |
| | CDR | Complementarity Determining Region |
| | CHARMM (-GUI) | Chemistry at Harvard Macromolecular Mechanics (-Graphical User Interface) |
| | CG (MD) | Coarse-Grained (Molecular Dynamics) |
| | COM | Centre of Mass |
| | CP | Connecting Peptide |
| | cSMAC | Central Supramolecular Activation Cluster |
| | CYR | Cytoplasmic Region |
| **D** | DOPE | Discrete Optimised Protein Energy |
| | dSMAC | Distal Supramolecular Activation Cluster |
| **E** | ECD | Extracelllular Domain |
| | ECTM | Extracellular and Transmembrane |
| | EM | Electron Microscopy |
| | EN | Elastic Network |
| | ER | Endoplasmic Reticulum |
| | Erk | Extracellular signal Regulated Kinase |
| **F** | F-actin | Filamentous actin |
| | FL | Full / Full-length |
| | fs | Femtoseconds |
| **G** | G-actin | Globular actin |
| **H** | HU (state of MHC) | Helix Uncoiled |
| **I** | Ig | Immunoglobulin |
| | Insane | INSert membrANE |
| | ITAM | Immunoreceptor Tyrosine-based Activation Motif |
| | IS | Immunological Synapse |
| **L** | LCK | Lymphocyte-specific protein tyrosine Kinase |
| | LINCS | Linear Constraint Solver |
| | LS (state of MHC) | Linker Stretch |

| | | |
|---|---|---|
| **M** | μM | Micromolar |
| | μs | Microseconds |
| | MD | Molecular Dynamics |
| **N** | NMR | Nuclear Magnetic Resonance |
| | nm | Nanometer |
| | ns | Nanoseconds |
| **P** | PBC | Periodic Boundary Condition |
| | PC | Phosphatidylcholine |
| | PDB | Protein Data Bank |
| | PE | Phosphatidylethanolamine |
| | PI (state of MHC) | Peptide Intact |
| | PIP ($_2$ / $_3$) | Phosphatidylinositol (bi / tri) Phosphate |
| | pMHC | Peptide attached to Major Histocompatibility Complex |
| | pN | Piconewtons |
| | PRS | Proline Rich Stretch |
| | PS | Phosphatidylserine |
| | ps | Picoseconds |
| | pSMAC | Peripheral Supramolecular Activation Cluster |
| | pY | Phosphorylated Tyrosine |
| **R** | Rg | Radius of Gyration |
| | RDF | Radial Distribution Function |
| | RMSD | Root Mean Square Deviation |
| | RMSF | Root Mean Square Fluctuation |
| **S** | SH | Src Homology |
| | SH4-U | Src Homology 4 and Unique domain |
| | SM | Sphingomyelin |
| | SMD | Steered Molecular Dynamics |
| **T** | TCR-CD3 | T cell receptor-CD3 complex |
| | TMO | Transmembrane Only |
| | TMR | Transmembrane Region |
| **U** | UD | Unique Domain |
| **V** | Vα/Vβ | Variable domain of TCRα/β |
| | VdW | Van der Waal's |
| **W** | WT | Wild-Type |
| **Z** | ZAP-70 | Zeta-chain Associated Protein kinase (70 kDa) |

# Chapter 1

# Introduction

## 1.1 T cells: a key component of our adaptive immunity

Our immune system comprises of two components: the innate immune system, which provides the first layer of protection against infection non-specific to pathogens, and the adaptive immune system, which acts as the next layer of protection and responds with a high degree of specificity to various pathogens. Our adaptive immunity also involves the property of storing 'memory' of the pathogen when it is first contacted. Exposure to the same pathogen (antigen) in future results in a memory response that is more effective in neutralising infection [1].

The leading members of the adaptive immune system are T cells and B cells. In mammals, hematopoietic stem cells in the fetal liver differentiate into multipotent progenitors which then enter the thymus to mature into T cells, whereas the maturation of B cells occurs in the bone marrow [2]. Upon maturation, B cells express antibody molecules called B cell receptors in their plasma membrane. Similarly, in the thymus gland, T cells express T cell receptors (TCRs) on their plasma membrane. Unlike B cell receptors which can bind to antigens alone, TCRs bind to antigenic peptides loaded into major histocompatibility complexes (pMHCs) [3] present on the surface of antigen presenting cells (APCs) such as B cells, dendritic cells, and macrophages.

During the development of these lymphocytes, somatic recombination of gene segments responsible for variability (V), diversity (D), and joining (J) is initiated by recombination-activating gene (RAG) proteins giving rise to the extensively diverse repertoire of TCRs (potentially ~$10^{61}$ in number [4]) and BCRs [5]. The binding strength of TCRs with self-antigenic pMHC also determines the development of a diverse T cell repertoire via positive and negative selection in the thymus. This process optimises the mature TCR repertoire and contributes toward the efficacy of immune responses [6].

There are two classes of MHC molecules present on the plasma membrane of APCs: class I MHC and class II MHC. Class I MHC molecules each consist of a cytoplasmic tail, a transmembrane region, and an extracellular region divided into four domains i.e. α1, α2, α3, and β2-microglubulin (β2m), where the β2m is non-covalently bound to the rest of the protein. The antigenic peptide binding cleft is shared by α1 and α2. Peptide loaded class I MHCs bind to a sub-population of T cells called CD8+ T cells (bearing CD8 co-receptor proteins to stabilise TCR-pMHC interaction). A class II MHC is composed of two chains each consisting of a cytoplasmic tail, a transmembrane region, and an extracellular domain. The extracellular region is divided into α1, α2, β1, β2 domains where the α and β subunits are non-covalently bound. The peptide

binds to α1 and β1. Peptide-loaded class II MHCs bind to CD4+ T cells (bearing CD4 co-receptors to stabilise TCR-pMHC interaction) [7]. The presence of these CD4 glycoproteins helps distinguish T helper (CD4+) cells from T cytotoxic (CD8+) cells and bind to the corresponding MHC class (Fig 1.1A). T helper cells, upon engagement with class II MHCs, secrete cytokines which further activate B cells, T cytotoxic cells, macrophages, etc. that take part in the immune reaction (Fig 1.1B) [1].

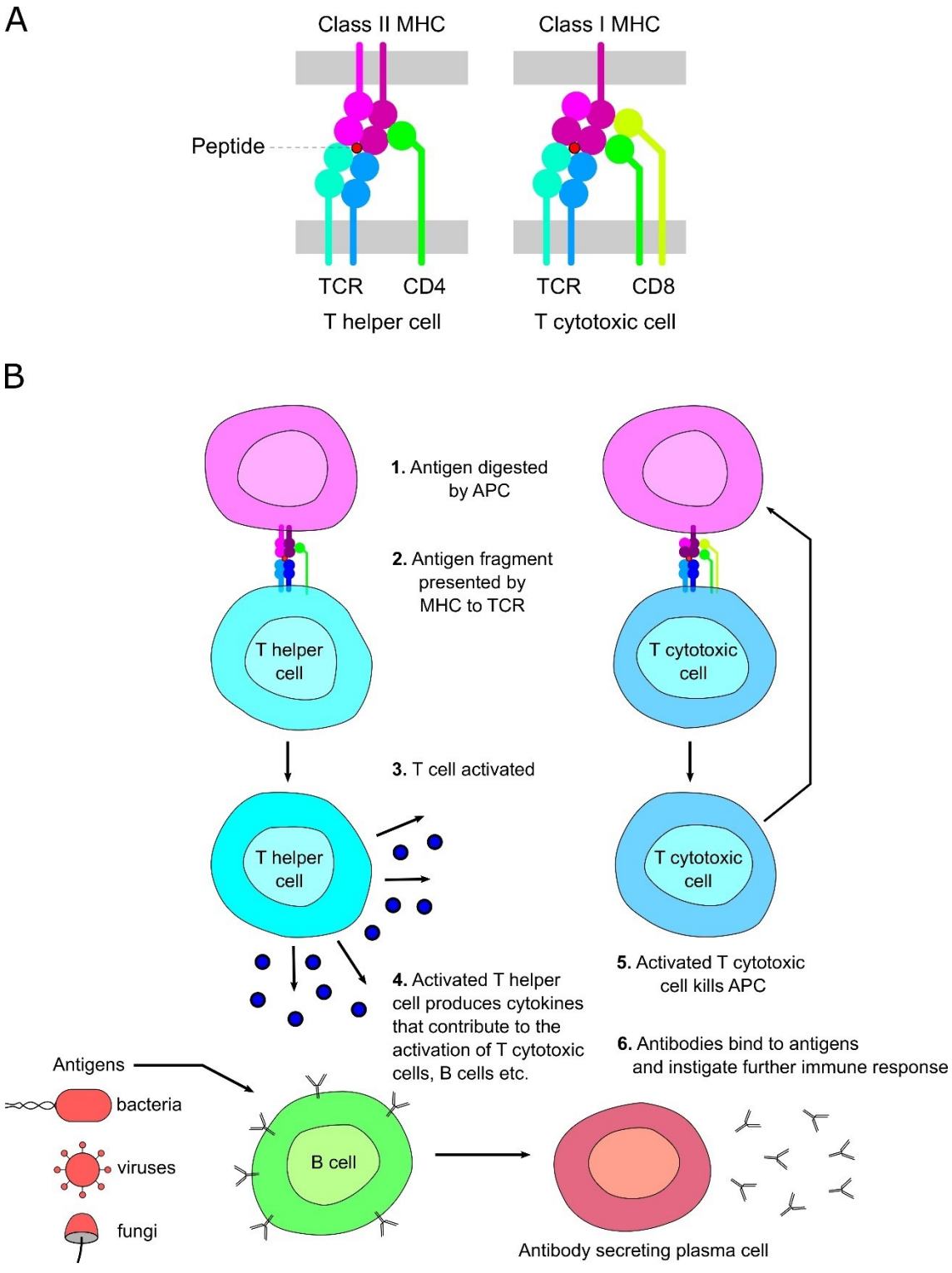

***Fig 1.1 Overview of T cell-mediated immunity.*** (A) TCR-pMHC interaction stabilised by CD4/CD8 co-receptors in T helper and T cytotoxic cells respectively, followed by (B) activation of T cells aiding elimination of infected cells via T cytotoxic cells and/or antibodies produced by B cell activation. Adapted from [1].

## 1.2 Overview of the T cell receptor and LCK

Antigen recognition and activation of T cells require TCRs to be complexed with CD3 and ζ subunits in their plasma membrane. However, the molecular mechanism by which the entire T cell receptor-CD3 (TCR-CD3) complex translates antigen recognition into intracellular signalling remains obscure. The TCR-CD3 complex is an octamer consisting of eight single-pass membrane proteins forming four dimers: TCRαβ, CD3δε, CD3γε, ζζ [8].

The activation of the T cell is initiated by the binding of the TCR alpha (α) and beta (β) subunits to pMHCs presented by APCs. The recognition of a wide variety of antigenic peptides is mediated via the highly diverse complementarity determining region (CDR) loops 1, 2, and 3 at the outermost tip of each of the TCRα and β chains. Upon TCR-pMHC ligation, signals are transmitted from the TCRαβ onto the co-stimulatory CD3δε, CD3γε, and ζζ subunits which are responsible for intracellular signalling [9]. The gamma (γ), delta (δ), and epsilon (ε) subunits are the CD3 monomers each containing extracellular, transmembrane and cytoplasmic regions. The zeta (ζ) subunit forms the ζζ homodimer playing a major role in initiating T cell signalling via its cytoplasmic region. Here, the ζ subunit is not considered as a CD3 subunit since it is not encoded by any CD3 gene. The ζ subunit exhibits low sequence similarity with the CD3 subunits and is encoded by the CD247 gene (Uniprot entry: P20963) situated on chromosome 1. All the CD3 subunits are encoded by CD3 genes located within a 300 kilobase stretch on chromosome 11, indicating that the ζ subunit is a distinct component of the TCR-CD3 complex [10]. Thus, in this thesis, ‘CD3’ only refers to the δ, γ and ε subunits. Nevertheless, when referring to the octamer as the ‘TCR-CD3 complex’, the ζζ dimer is considered inclusive.

The first and the most basic structural model of the TCR-CD3 complex based on its constituent chains and topology was proposed in 1988 [11]. Later, in 2002, a more sophisticated model of the entire complex was proposed explaining its spatial organisation and assembly in the membrane [8]. Despite available information regarding the structure and function of TCR-CD3

complexes, the detailed activation mechanism is not entirely understood. Due to the considerable controversy about its triggering mechanism, there were different models proposed to explain it. However, it was suggested in 2011 that understanding TCR-CD3 triggering still requires better information on the structure of TCR-CD3 complex and the conformational changes it undergoes on pMHC binding [12]. In addition to the availability of X-ray crystal structures of the bound state of the TCRαβ and pMHC extracellular regions [13–15], a recent cryo-electron microscopy (cryo-EM) study resolved the complete assembly of the TCR-CD3 extracellular and transmembrane region [16], revealing the quaternary arrangement and interfaces between the multiple subunits. These findings largely contributed toward understanding the TCR-CD3 structure and enabled further investigation. Although, we still lack a structural understanding of its complete cytoplasmic region (CYR). The TCR-CD3 CYR is an intrinsically disordered region that is potentially highly dynamic [17], making it challenging for experimental techniques to resolve. The TCR-CD3 CYR mediates interaction with intracellular signalling proteins and undergoes phosphorylation via its immunoreceptor tyrosine-based activation motifs (ITAMs). Therefore, it is critical to obtain structural information on the TCR-CD3 CYR to enable molecular-level studies on its signalling mechanism.

The lymphocyte-specific protein tyrosine kinase (LCK), a member of the Src family of kinases, is the key intracellular signalling protein that initiates TCR-CD3 signalling by phosphorylating its ITAMs. LCK is a single-chain protein consisting of multiple domains: Src-homology (SH)4, unique domain, SH3, SH2, and the kinase (from N to C terminus). The Y394 residue in the kinase domain mediates the phosphorylation of TCR-CD3 ITAMs upon autophosphorylation (pY394), whereas Y505 located at the C-terminus acts as an inhibitory site [18]. Upon phosphorylation of Y505 (pY505), the kinase domain binds to the SH2 domain inducing a conformational change that promotes the inactive state of LCK. This regulatory mechanism is observed among the members of the Src kinase family [19]. Upon T cell activation, this inhibitory site at Y505 in the LCK kinase domain is dephosphorylated promoting the active state of LCK [20] (Fig 1.2).

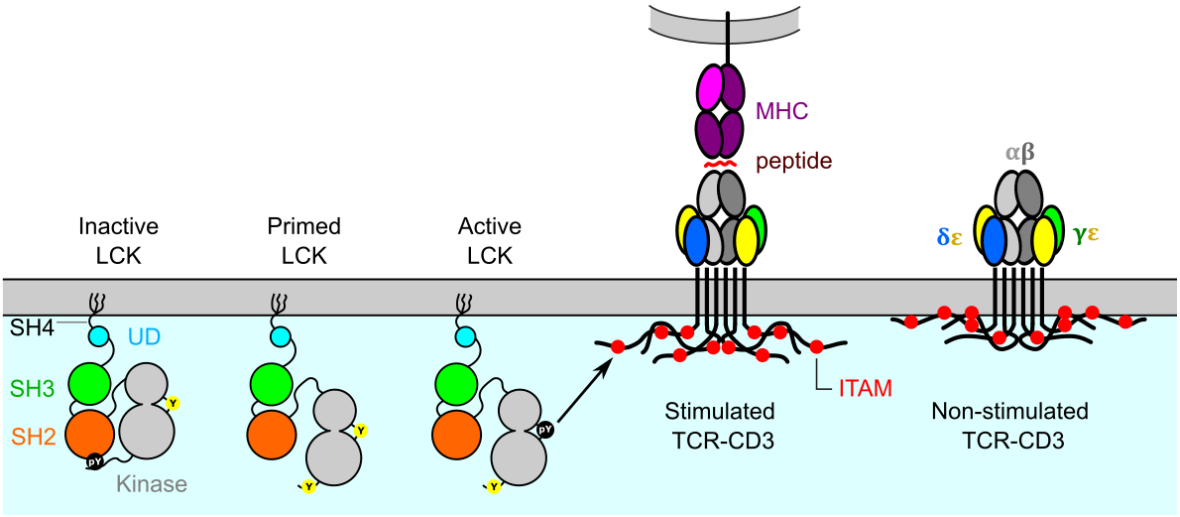


***Fig 1.2 Active and inactive states of the TCR-CD3 and LCK.*** *When Y505 located at the bottom of LCK is phosphorylated (pY shown in black) (left), LCK exhibits an inactive/closed conformation. When neither Y394 nor Y505 are phosphorylated, LCK exhibits a primed state. When only Y394 is phosphorylated, LCK can phosphorylate TCR-CD3 ITAMs upon the release of its cytoplasmic tails from the membrane.*

The SH2 and SH3 domains are known to aid in the interaction of the kinase domain to tyrosine-phosphorylated substrates [21]. Moreover, the SH2 domain is suggested to be critical for the role of LCK kinase in TCR-CD3 signalling [22]. The SH3 domain also aids in stabilising the inactive or the 'closed' conformation of LCK by binding to a proline-rich region in the linker region that connects the SH2 and the kinase domain [23]. The unique domain of LCK exhibits low sequence similarity with the unique domains of other Src family kinases. In LCK, it aids in an electrostatic interaction with CD3ε subunits of TCR-CD3 complexes which further helps in phosphorylating the ζζ and other CD3 subunits [24]. The first ~10 residues at the N-terminus of LCK constitute the SH4 domain. This includes G2, C3, C5, which are targeted for acylation as a post-translational modification. These acylated lipid tails then insert into the hydrophobic core of the membrane promoting in membrane localisation of LCK [25,26].

Previous experiments with plasma membrane-like lipid vesicles conducted with surface plasmon resonance (SPR) analysis revealed that lipids, particularly phosphatidylinositol-4,5-biphosphate i.e. PI(4,5)P2 ($PIP_2$) and phosphatidylinositol-3,4,5-triphosphate ($PIP_3$), in the plasma membrane ionically interact with the LCK-SH2 domain. Nuclear magnetic resonance (NMR), molecular docking, and electrostatic profiling indicated the PIP lipid binding site in the SH2 domain. Further, mutational studies suggested that SH2-lipid interaction sites modulated association of LCK with TCR-CD3 signalling

subunits [27]. PIP lipids have also been found to regulate TCR-CD3 activation by interacting ionically with cationic residues in the CD3ε and ζ CYRs [28–31]. Similarly, PIP interaction is found to play a functional role in other membrane proteins containing cytoplasmic regions [32]. The TCR-CD3 CYR is not only found to interact with PIPs but also penetrate into the hydrophobic core of the membrane [33,34]. This interaction between the CYR and the membrane is suggested to disengage when the local anionic charge decreases in the inner leaflet of the plasma membrane upon TCR-pMHC binding [31,35]. This demands more precise information on the lipid binding sites of the TCR-CD3 complex, and on the accessibility of their ITAMs to LCK upon pMHC binding.

Therefore, in this thesis, using multi-scale computational methods, I studied the dynamics and interactions of the TCR-CD3 and LCK with a plasma membrane-like bilayer at near-atomic resolution. As part of this study, using available experimentally resolved structures, I have obtained the first complete models of the TCR-CD3 complex and LCK and provided insights into their conformational changes and lipid interactions.

## 1.3 The structure of the TCR-CD3 complex

The TCR-CD3 is an octameric complex whose specificity of ligand binding is dictated by its disulphide-linked α and β chains [8,36]. Their immunoglobulin (Ig)-like extracellular domains (ECDs) are each composed of variable and constant regions [37], followed by a helical transmembrane region (TMR) and a short cytoplasmic tail. A relatively short and flexible stalk region, known as connecting peptides (CP), connects their TMR to their ECDs [38]. Similarly, the CD3 subunits also contain CPs that connect their helical TMRs to their ECDs (Fig 1.3). Unlike the αβ dimer which are disulphide-linked to each other, the CD3δε and CD3γε dimers are non-covalently bonded with each other but contain intra-subunit disulphide bridges in their ECDs. Apart from the αβ heterodimer, the ζζ homodimer also contains an inter-subunit disulphide-linkage in its TMR. Although the ζζ does not contain a significant extracellular region, it possesses the longest cytoplasmic tails in the complex. The CD3 and ζ subunits contain immunoreceptor tyrosine-based activation motifs (ITAMs) in their CYRs: one each in γ, δ, ε and three in each ζ monomer.

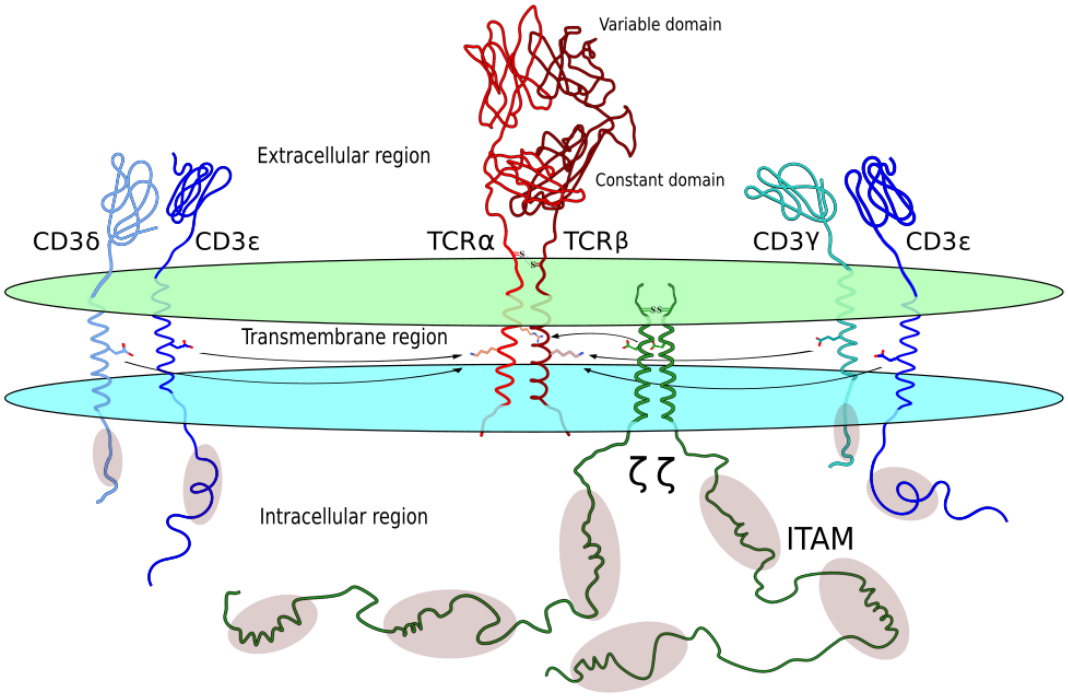


***Fig 1.3 Structural representation of the assembly of the TCR-CD3 complex.*** *The ζζ TMR and the TCRαβ ECD representations are based on structures retrieved from PDB IDs: 2HAC and 1KGC respectively.*

In the transmembrane region, as shown in Fig 1.3, the assembly of the TCR-CD3 complex occurs as follows: CD3ε chains dimerise with CD3δ and CD3γ each, and further oligomerise with the α and β chains respectively, primarily via electrostatic interactions in the core of their TMR. The lysine residue (K) of the TCRα TMR interacts with aspartic acid (D) residues of ε and δ, and the lysine residue of TCRβ with aspartic and glutamic acid (E) residues of ε respectively. The recruitment of the ζζ homodimer is suggested to be the final step to form the complete TCR-CD3 complex. Recruitment of ζζ is driven by the electrostatic contact of arginine residue (R) in the TCRα TMR with the aspartic residues of ζζ TMR [8]. This organisation of the TCR-CD3 TMR was confirmed upon the publication of the cryo-EM structure containing the ECDs and TMRs resolved at 3.7 Å (PDB:6JXR; view in 3D here: https://www.rcsb.org/3d-view/6JXR/) [16].

### 1.3.1 The extracellular region

The outermost tip of the TCRαβ dimer, containing six complementarity-determining region (CDR) loops i.e. three in each of the variable regions of the α and β subunits (VαVβ), dictates the recognition of pMHC ligands [37]. VαVβ is further connected to the constant regions of the α and β subunits (CαCβ) located closer to the surface of the outer leaflet of the membrane (Fig 1.4).

The TMR assembly of the CD3 dimers with αβ depend on electrostatic interactions. Both the α and β subunits contain a lysine at the core of their TMR that mediates interaction with anionic sidechains of CD3 subunits[8]. This raises a doubt about why the CD3δε dimer is specific to binding to TCRα and CD3γε to TCRβ. Addressing this, it was suggested that it is the complex shape of the ECDs that promotes the specificity and stabilises the organisation of CD3δε with TCRα and CD3γε with TCRβ topology after their TMR contact [39,40].

In the TCR-CD3 ECD assembly, Cβ majorly mediates the interaction with CD3γε ECDs and Cα with CD3δε ECDs. Cβ contains an FG loop bulging outward in the shape of a teacup handle that was suggested to be functionally significant not only in transmitting signals downstream [41] but also in thymic selection and T cell development [42,43]. Consistent with this, more recent experimental studies [44] combined with molecular dynamics simulations [45,46] showed that pMHC ligation induces an allosteric (conformational change-based) communication in the TCR-CD3 dependent on the FG loop in Cβ. On the other hand, Cα contains a DE loop that packs against the surface of CD3δ ECD. Cα also contains an AB loop that carries potential to undergo a conformational change upon antigen ligation irrespective of pMHC specificity and contributing allostery-mediated TCR-CD3 signalling [47].

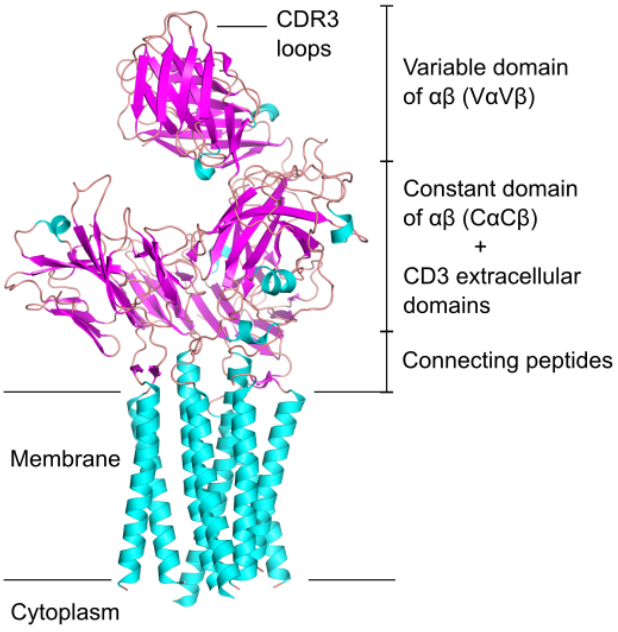


***Fig 1.4 The cryo-EM structure of the TCR-CD3 complex*** *(PDB:6JXR). The extracellular domain is mostly constituted by β sheets (magenta) and the transmembrane region by α helices (cyan). Loop regions are shown in pink.*

### 1.3.2 The connecting peptides

The extracellular CPs, or the stalk region, also contribute to this assembly of the ECDs. The TCRα CP packs against CD3δ stabilising the contacts between TCRαβ and CD3δ ECDs. This region is further stabilised by the extracellular peptide of one of the ζ subunits [16]. In the highly conserved CPs of each of the CD3 subunits (~9 amino acids), intra-subunit disulphide bonds occur within the tetracysteine (CxxC) motif [48]. However, in the CPs of the TCRαβ chains (~20 amino acids), one inter-subunit disulphide bond occurs between C227 of the α subunit and C264 of the β subunit. Similarly, the ζζ homodimer also forms an inter-subunit disulphide bond near the extracellular leaflet of the membrane [16] (Fig 1.5). Disulphide scanning experiments [49] combined with earlier studies [50] suggested that the TCRαβ inter-subunit disulphide bond is not required for membrane surface expression of the TCR-CD3 complex because it does not completely disrupt TCRαβ heterodimeric formation, rather they formed a loose contact with ζζ and were heavily dependent on the ζζ homodimer for membrane surface expression.

The CxxC motifs in each of the CD3 (δ, γ, ε) subunits are responsible for the reinforcement of their CPs as well as their TMR conformational stability [51,52]. A combination of mass spectrometry, NMR and mutation studies of these motifs also indicate that eliminating these CxxC disulphide bonds disturb the configuration of the TMR and CYR including the ITAMs, thereby suggesting their role in TCR-CD3 mechanotransduction [53].

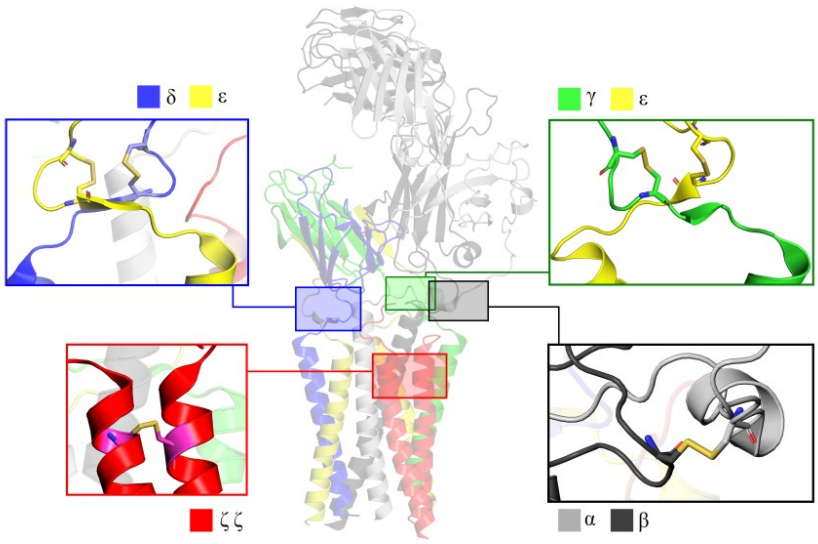


***Fig 1.5 The disulphide bonds in the TCR-CD3 connecting peptides and in the ζζ transmembrane region.***

### 1.3.3 The transmembrane region

The transmembrane assembly of the TCR-CD3 is mediated by electrostatic interactions among nine conserved amino acids [8] (Fig 1.6A, B). In the TCRαβ TMR, there are three conserved basic amino acids: one lysine on TCRα and -β chain each, and one arginine on the TCRα. The CD3δ, CD3ε and ζ TMRs each contain a conserved aspartic acid while the CD3γ TMR contains a conserved glutamic acid. The acidic residues on the CD3δε and CD3γε dimers interact with the centrally located basic residues on the TCRα and TCRβ respectively. This forms TCRα-CD3δε and TCRβ-CD3γε heterotrimers through D-K-D and E-K-D ionic interactions respectively [8]. The aspartic acids on the ζζ dimer finally pair with the conserved arginine residue of TCRα which is located toward the extracellular leaflet surface to then form the complete TCR-CD3 TMR.

Early studies have shown that TCRα-CD3δε and TCRβ-CD3γε can assemble without the requirement of ζζ indicating that ζζ is the last dimer to bind to the complex [54–56]. Nevertheless, substitutions of any of the nine ionic sidechains of the transmembrane helices with alanine disrupts the assembly of the TCR-CD3 complex [8]. Molecular dynamics simulations of the TCRα-CD3δε trimer in a membrane also showed that substituting the aspartic acids of CD3δε to asparagine or alanine destabilises their assembly with TCRα TMR [57]. This trimeric interaction of ionic sidechains between transmembrane subunits is also suggested to take place in other immune receptors such as the KIR2DS2-DAP12 (D-K-D), NKG2D-DAP10 (D-R-D) and NKG2C/CD94-DAP12 complexes (D-K-D) [58–60]. In addition, mutations in the interface between the TCRα and TCRβ TMRs, which form the core of the receptor complex, have shown to result in a defective assembly of the complex indicating that the TCRαβ heterodimer prefers a specific interface to form a closely packed structure and then enable proper assembly of the two CD3 heterodimers [49].

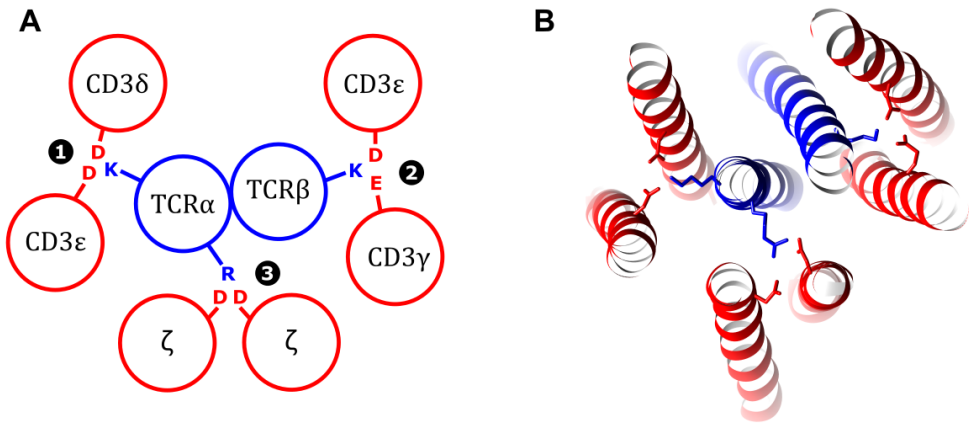

***Fig 1.6 The transmembrane assembly of the TCR-CD3 complex.*** *(A) A simplified version adapted from* [8]*. The numbers indicate the step-by-step oligomerisation process. (B) Snapshot of the TCR-CD3 transmembrane region as revealed by the recent cryo-EM structure (PDB:6JXR)* [16]*.*

The TCR-CD3 TMR interactions are reported to be highly specific during their assembly. Mutation studies showed the requirement of both aspartic acids of CD3δε TMR to interact with the lysine in the TCRα TMR and same is required for TCRβ-CD3γε trimer assembly. In addition, the TCRα strongly preferred to interact with CD3δε than CD3γε [8]. Consistent with this, it was shown that the CD3γ ECD cannot be substituted by CD3δ [61,62] while interacting with TCRβ. This suggests specific roles of TCRα-CD3δε and TCRβ-CD3γε assembly. Moreover, it was observed that TCRα-CD3δε would assemble without CD3γ and ζ, and CD3γε would not satisfactorily bind to TCRβ in the absence of CD3δ indicating the dependence of ζζ assembly on TCRβ-CD3γε, depending further on TCRα-CD3δε assembly [8].

The disulphide-linked ζζ dimer interacts with the arginine of TCRα which was shown to be free from CD3δε interaction [8]. This interaction of ζζ was primarily mediated via its aspartic acids essential for its assembly with the rest of the TCR-CD3 complex [63]. Prior to this assembly, the homodimerisation of ζζ, including the disulphide link formation, is dependent on a set of interactions i.e. D6, Y12, T17, between the two ζ chains, mutations of which cause a catastrophic decrease in dimerisation by 75% to 85% [63].

These ionic interactions responsible for the TMR assembly of the TCR-CD3 complex have been suggested as potential drug targets for the treatment of autoimmune disorders [64]. Although, no clinical studies on this have been reported lately.

### 1.3.4 The cytoplasmic region

In contrast to the long cytoplasmic tails of the CD3 and ζ chains, the cytoplasmic regions of TCRα and TCRβ are each less than 10 amino acids in length. It was speculated that the lysine residues in the TCRβ CYR might serve as a ubiquitination site for TCR-CD3 turnover or help with endoplasmic reticulum (ER) retention signalling [65]. Also, it is reported that serine residues at the C-terminus of TCRα are unorthodox sites for ubiquitination upon which the TCR-CD3 is targeted for degradation by ER [66].

The CD3ε CYRs are each a sequence of ~50 amino acids consisting of 3 major sections: a juxtamembrane polybasic region (KNRKAKAK) with a lipid binding

feature, followed by a proline-rich stretch (PRS) (PPPVPNP), and an ITAM (YxxL/Ix….YxxL/I) (Table 1.1) [33,67]. Using lipid micelles to mimic the cell membrane, NMR spectroscopy studies revealed structural configurations of an ITAM-containing peptide of the CD3ε subunit showing that it could penetrate the membrane surface via tyrosine sidechains while its protein backbone was found at the interface of the hydrophobic and hydrophilic core of the membrane [33].

The PIP lipid binding of CD3ε CYRs are mediated via their polybasic sequences, also called basic-rich stretches (BRS) [28]. Similarly, the ζ CYRs also contain BRS motifs [29,30] in addition to their ITAMs (YxxL/Ix….YxxL/I) (Table 1.1). However, the PIP lipid binding affinity of the BRS motif in ζ is not as high as that of CD3ε. Nevertheless, mutation of the basic residues in these motifs significantly reduced their lipid binding property implying the importance of ionic lipid interactions in regulating phosphorylation-mediated signalling in resting T cells [28–30]. Upon TCR-pMHC binding, these BRS-lipid interactions are somehow interrupted causing ITAMs to be exposed to intracellular space available for phosphorylation by intracellular kinases [32].

***Table 1.1 Sequences and motifs of the CD3 and ζ cytoplasmic region.*** *Adapted from* [24].

| Chain | Amino acid sequence of its cytoplasmic region |
|---|---|
| CD3δ | GHETGRLSGAADTQALLRNDQVYQPLRDRDDAQYSHLGGNWARNK |
| CD3γ | GQDGVRQSRASDKQTLLPNDQLYQPLKDREDDQYSHLQGNQLRRN |
| CD3ε | SKNRKAKAKPVTRGAGAGGRQRGQNKERPPPVPNPDYEPIRKGQRDLYSGLNQRRI |
| ζ | RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPQRRKNPQ |
| | EGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR |

ITAM BRS PRS

The CD3δ and CD3γ CYRs are devoid of BRS motifs (Table 1.1) and were not found to strongly associate with the membrane [33]. However, it is suggested that their interaction with the membrane could possibly be influenced by the CD3ε or ζ subunits [67].

## 1.4 Activation of the TCR-CD3 complex

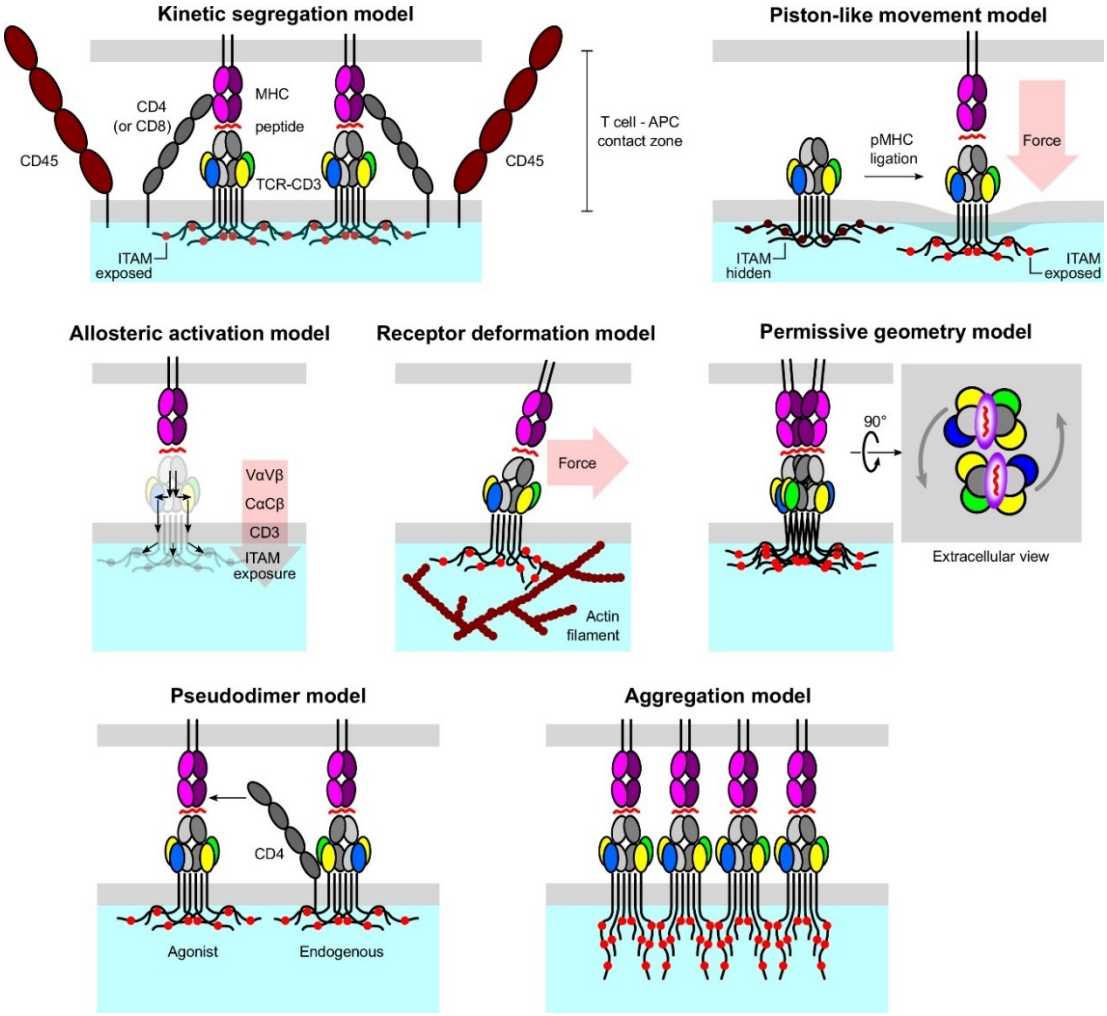


***Fig 1.7 Activation models of the TCR-CD3 complex.***

Considering numerous studies conducted in the last couple of decades, it is clear that upon pMHC ligation, the TCR-CD3 propagates signals via its CD3 and ζ ITAMs by exposing them to intracellular solvent for phosphorylation [9,68]. In its resting state, these ITAMs are attached to the lipid bilayer close to their hydrophobic core preventing access by kinases [69]. However, we still lack the knowledge of precise conformational changes that pMHC binding induces on the TCR VαVβ that eventually translates into antigen-induced signals to the CYRs. Hence, to study the possible mechanisms of TCR-CD3 triggering, there have been several models proposed (Fig 1.7):

**Kinetic segregation model**: This model suggested that inhibitory molecules like CD45, containing large ECDs, are excluded from the activation site due a close T cell-APC contact zone [70]. However, this raises a question whether it is possible for the large ECDs to bend toward the surface of the membrane and fit into the contact zone. Addressing this, a study on the dimensions and interactions of these large ECD-containing proteins suggested that this comes at a cost of high entropy and thus it is energetically favourable for them to be excluded from the contact zone [71].

**Diffusion trap model:** In addition to the kinetic segregation model as above, the formation of TCR-pMHC complexes is proposed to be dependent on the affinity or diffusion coefficient of TCR-pMHC interactions [68,72] i.e. TCR-pMHC association requires a larger number of pMHC complexes if the TCR VαVβ has lower affinity for the particular pMHC, and vice versa. In the case of high affinity, it is proposed that monomeric pMHCs strongly attach with TCRαβ and translocate to the centre of the T cell-APC contact zone i.e. immunological synapse (IS). Furthermore, as the T cell scans the surface of the APC, these high affinity TCR-pMHC interactions move along the surface possibly with the help of the actin cytoskeletal interactions [72].

**Piston-like movement model**: Mechanical force applied by pMHCs onto TCRs was proposed to instigate a perpendicular piston-like movement of the TCR-CD3 complex exposing the CD3 CYRs containing ITAMs to the aqueous environment [73].

**Allosteric activation model**: Several studies indicate that TCR-pMHC binding instigates a cascade of site-specific events from the VαVβ to CαCβ, more predominantly via the FG loop in the TCRβ subunit [44,45,74,75]. The cryo-EM structure of the TCR-CD3 ECD and TMR (PDB:6JXR) [16] also indicates that the Cβ FG loop lies just above the CD3γε ECD potentially involving it during allostery. These ECD conformational changes are transmitted to the CYR via

the TMR by quaternary structure relaxation in the TCR-CD3 complex [76] thereby exposing ITAMs for phosphorylation.

**Receptor-deformation / Mechanosensor model**: Following TCR-pMHC ligation, when the T cell continues to scan the surface of the APC, the pMHC exerts a tangential pulling force onto the TCR-CD3 ECD while the membrane-associated actin cytoskeleton interacts with its cytoplasmic tails. As a result of the pulling force driving the TCR-CD3 ECD away from its cytoplasmic tails, the TCR-CD3 ITAMs are released from their lipid-bound state in the intracellular region thereby becoming available for phosphorylation [77–80].

**Permissive geometry model**: This model gives importance to the geometry of pMHC complexes while binding to TCRs. A dimer of pMHC complexes correspondingly bind to a dimer of TCRs and form a scissor-like conformation when they cross each other at an angle. Their rotation around each other at this angle is said to cause a conformational change in the CYR, releasing them from their lipid-bound state [81].

**Pseudodimer model**: Upon ligation of a TCRαβ with an agonistic pMHC, a CD4 molecule is recruited. This CD4 molecule, previously bound to an endogenous TCR-pMHC complex, promotes dimerisation of the endogenous TCR-pMHC with the agonistic TCR-pMHC forming a TCR-pMHC heterodimer or a ‘pseudodimer’ as referred in this model [82]. This pseudodimer may contribute to inducing a TCR-CD3 conformational change, and/or bringing the CD4-bound LCK in close proximity to the ITAMs. In the same CD4-dependent manner, the pseudodimer may further form pseudomultimers promoting TCR-CD3 clusters and a cascade of phosphorylations by LCK.

**Aggregation model**: According to this model, multiple TCR-CD3 complexes bind to multiple pMHC ligands forming multimeric TCR-pMHC complexes. This further leads to contacts between the TCR-CD3 complexes allowing transphosphorylation between their cytoplasmic tails and associated LCK molecules. TCR-CD3 aggregation or clustering may also induce competition between ζ CYRs to bind to the membrane due to lack of surrounding membrane surface area resulting in solvent accessibility of their ITAMs and further signal transduction [83]. In resting TCR-CD3 complexes, the CYR is lipid-bound and inaccessible for ITAM phosphorylation [29,30,33]. A recent review also suggested that TCR-CD3 clustering can also push inhibitory phosphatases toward the periphery of the T cell-APC contact zone [84], consistent with the kinetic segregation model, thereby ensuring a wider area of inhibitor-free signalling.

The models described above are mutually non-exclusive i.e. they may occur in cooperation with each other and/or in a stepwise manner. However, it was shown that the initial triggering of TCRs are driven by monomeric complexes and not necessarily multimeric complexes but may require multimers to sustain T cell signalling [85]. Another study stated that the kinetic segregation of CD45 inhibitors is necessary for TCR-CD3 signalling. After monomeric TCR-pMHC complexes are formed, co-receptors such as CD2 on the T cell plasma membrane interact with APC co-receptors such as CD58 to further stabilise TCR-pMHC association and enhance exclusion of molecules like CD45 containing large ECDs [86]. Following this event, it is necessary that the TCR-CD3 ITAMs are exposed to intracellular solvent for ITAM phosphorylation. Consistent with the allosteric and mechanosensing models proposed above, a study was conducted in support of both models. The CDR loops located at the tip of the TCRαβ form a catch bond with the pMHC. While the allosteric activation model describes a conformational change directed from the ECD to the CYR via the Cβ FG loop, this study indicates that the TCR-CD3 employs allosteric sites in CαCβ (including the FG loop) to optimise pMHC interaction via VαVβ [46]. This potentially suggests that the mechanosensitivity and allostery of the TCR-CD3 complex complement each other in controlling extracellular binding and intracellular signalling.

## 1.5 The structure of LCK

The lymphocyte-specific protein tyrosine kinase (LCK), previously known as p56Lck due to its molecular weight of 56 kDa, is a member of the Src family of tyrosine kinases. This protein is key in propagating pMHC-stimulated TCR-CD3 signals into the cytoplasm. It is not only required to associate with the inner leaflet of the plasma membrane [27] but also spatially organise itself near the stimulated TCR-CD3 complex in an active/open configuration [87] to initiate T cell signal transduction.

Each protein member of the Src kinase family is reported to be composed of the N-terminal SH4 domain targeted for post-translational modification(s), followed by the unique domain, SH3, SH2 and a kinase domain [19] (Fig 1.8).

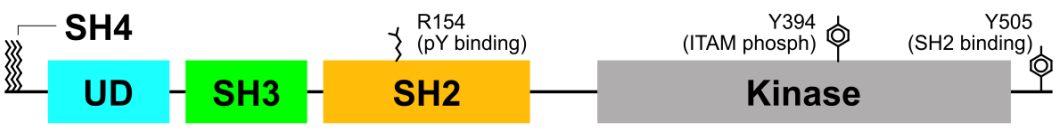


***Fig 1.8 Structural components of LCK*** *as conserved in the Src kinase family.*

### 1.5.1 The LCK-SH4 and the unique domain

The N-terminus or the LCK-SH4 domain i.e. MGCVCSSNPE is a homologous stretch that is conserved across the members of the Src family of kinases [25]. In this domain, the initiator methionine residue is removed upon post-translation modification following which G2 is covalently bonded to a myristic acid in a process called myristoylation [88,89]. Serine 6 of LCK-SH4 is shown to pose as a critical site for G2 myristoylation aiding in LCK membrane localisation [90]. C3 and C5 are covalently bonded with palmitic acids in a process called palmitoylation [25,91,92]. The myristoylated and palmitoylated residues then contain a lipid-like hydrocarbon tail preferring a membrane-inserted state [26].

The unique domain (UD) of all kinase proteins in the Src family, including LCK, are intrinsically disordered and exhibit least sequence similarity from each other [93]. The CxxC motif in the LCK-UD mediates interaction with the CxCP motif in CD4/CD8 co-receptors [94,95]. The C20 and C23 residues of the LCK-UD CxxC motif combine with C420 and C422 of CD4 co-receptor CxCP motif to form a zinc co-ordination, as reported by an NMR study. The study also suggests similar zinc co-ordination by LCK-CD8 association [96]. This metal co-ordination was reported to be critical for LCK-CD4/CD8 complex formation [97,98]. The LCK-UD also electrostatically mediates intracellular interaction with the CD3ε subunit of the TCR-CD3. Mutation studies showed that the BRS motifs of CD3ε CYRs are important to associate with D12 of LCK-UD and then conduct LCK-mediated phosphorylation. Recruitment of LCK by CD3ε CYRs was also shown to help in phosphorylating ITAMs other subunits in the TCR-CD3 [24].

### 1.5.2 The LCK-SH3 domain

The LCK-SH3 domain is a small β barrel-like structure (Fig 1.9) consisting of ~60 residues [99] playing an important role in T cell signalling pathways [100,101]. Along with the LCK-SH2 domain, they carry out critical interactions in signalling pathways.

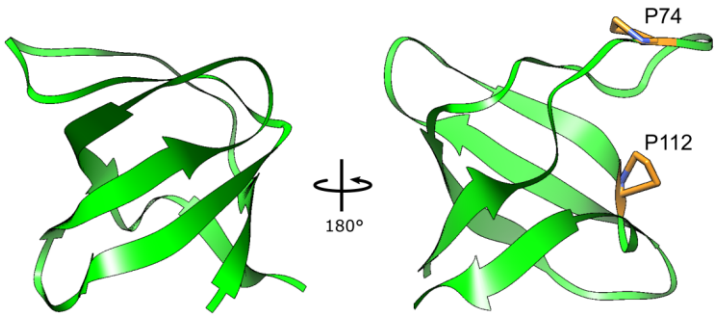


***Fig 1.9 The LCK-SH3 domain structure*** *(PDB:4D8K). The proline residues of LCK-SH3 that potentially interact with proline-rich regions of CD3ε are shown.*

As discussed above, the association of LCK with the TCR-CD3 is primarily driven by LCK-UD-CD3ε interaction. Following this, the LCK-SH3 interacts with the PRS of CD3ε, as confirmed by NMR and mutation studies [24]. More recent studies added that the TCR-CD3 CYR is required to undergo conformational changes to allow interaction of LCK-SH3 with CD3ε PRS before ITAM phosphorylation [102]. Similarly, CD3ε PRS is also reported to mediate interaction with SH3 domain of NCK upon TCR-CD3 activation [103]. Similar interactions have been reported in proline-rich regions during contact between LCK-SH3 and the LCK-kinase domain in the closed conformation of LCK. This contact is suggested to provide stability to the closed state of LCK [23].The study on NCK, a member of the SH2-SH3 containing family of proteins which forms the link between cell surface receptors and the actin cytoskeleton [104], further sheds light on the importance of SH3 domains.

### 1.5.3 The LCK-SH2 domain

The interaction of human tyrosine kinase proteins with membrane lipids via their SH2 domains (~100 residues) (Fig 1.10) is essential for their function [105]. Further, Sheng *et al.* revealed quantitative data on the PIP lipid binding properties of the LCK-SH2 domain in vesicles mimicking the plasma membrane, whereas they could not detect the same for LCK-SH3 [27].

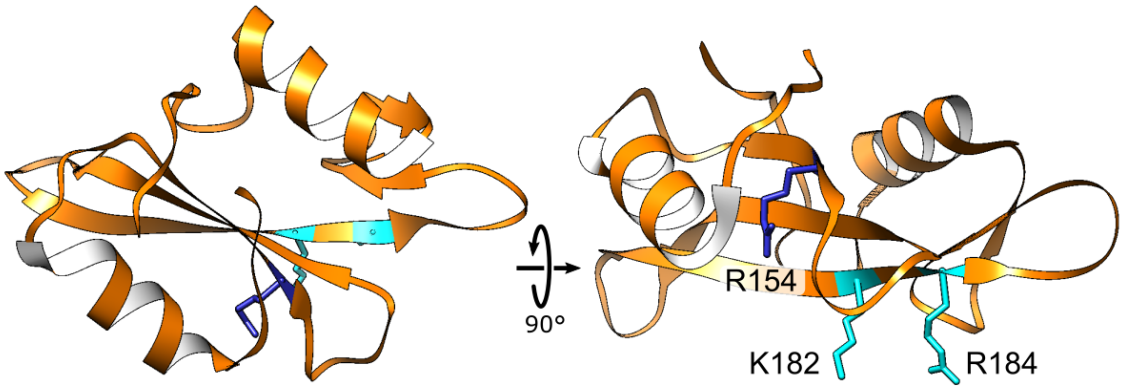


***Fig 1.10 The LCK-SH2 domain structure*** *(PDB:4D8K). The phospho-tyrosine binding site (R154) and PIP lipid binding sites (K182 and R184 in cyan) are shown as suggested by Sheng et al.* [27]

In the SH2 domain, Sheng *et al.* determined an alternative phospho-tyrosine (pY) binding pocket (K182 and R184) to the well-known R154 pocket but was not as electropositive as R154. Previous research indicated that these pockets could bind to lipids. Therefore, they conducted mutation studies on both pockets to validate this. They observed that there was no decrease in lipid binding affinity when R154 was mutated to alanine, but lipid affinity reduced 3-fold when the alternative pocket was mutated. This showed that K182 and R184 form the primary lipid binding site for LCK-SH2 via non-specific electrostatic interactions and was supported by experiments using HeLa cells. They performed NMR analyses suggesting that aromatic and hydrophobic amino acids (A160, F163, I183) also contribute to the lipid binding property of LCK-SH2. This was consistent with a A160K mutation which led to ~30% higher affinity indicating that A160 also forms an important lipid binding site [27]. More importantly, using SPR analysis, they determined the lipid selectivity/preference order of LCK-SH2 toward PIP lipids, as shown in Fig 1.11.

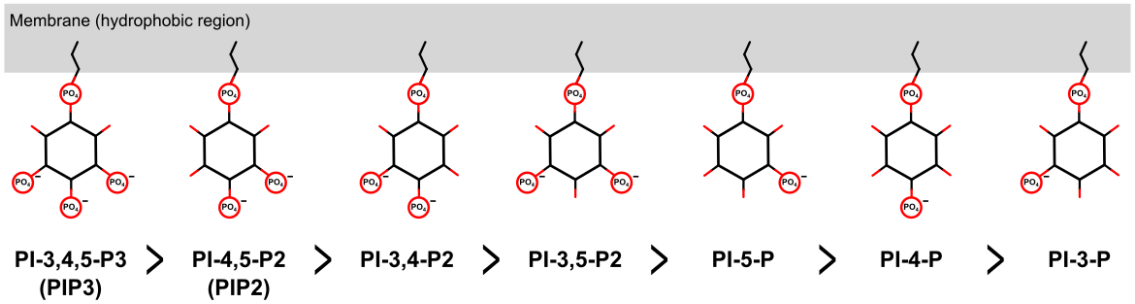


***Fig 1.11 LCK-SH2 selectivity for PIP lipids*** *as suggested by Sheng et al* [27].

They showed that $PIP_3$ (PI-3,4,5-P3) and $PIP_2$ (PI-4,5-P2) lipids play lead roles in positively influencing localisation of LCK-SH2 to the membrane. This aids in its function in the TCR-CD3 complex. Mutation of the primary lipid binding site (K182 and R184) of LCK-SH2 clearly inhibited its membrane localisation. They also performed single molecule tracking of LCK and the ζ chain of the signalling complex in a stimulated cell and observed an enhanced LCK-ζ colocalisation. This strongly suggested that anionic lipids such as PIP lipids play a critical role in recruiting LCK to the TCR-CD3 activation site and further aids in its signalling capabilities [27].

Another important functional feature of the SH2 domain is its binding potential toward phospho-tyrosines and its regulation of kinase domain activity by binding with its C-terminal tyrosine Y505 [106]. Early clinical studies also reported that phosphorylation of Y505 reduces the oncogenic potential of LCK and Y505F mutation increases its oncogenic activity [107].

In addition, Y192 phosphorylation or its substitution by an anionic amino acid (Y192E) significantly diminishes LCK-SH2 binding with pY-containing polypeptides [106]. Other mutation studies strongly suggested that Y192 substitution leads to inability of CD45 associating with LCK *in vivo* and thus preventing dephosphorylation of Y505 at the C-terminus of LCK. This restricts LCK from embracing the active conformation and hinders inducible TCR-CD3 signalling. However, different point mutations of Y192 destabilised the active conformation at different levels. On the whole, Y192 orchestrates recognition by CD45 leading to LCK-CD45 association aiding in activation of LCK and thymocyte development [108]. On the other hand, *in vitro* studies indicate that CD45 can both positively and negatively regulate TCR-CD3 ITAM phosphorylation in TCR-LCK-CD45 protein clusters [109].

### 1.5.4 The LCK kinase domain

The LCK-kinase domain consists of residues starting from ~225 until the last ($509^{th}$) residue of LCK. Homologous to the kinase structures of other Src family members, the LCK kinase contains an N-terminal lobe that connects to the LCK-SH2 via an elongated loop and a C-terminal lobe that is larger in size (Fig 1.12).

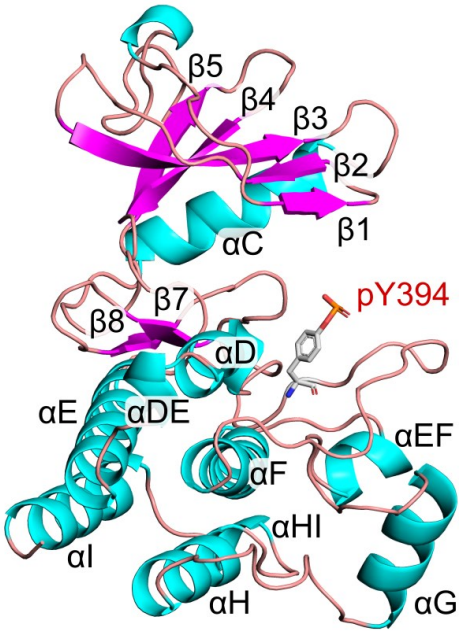


***Fig 1.12 The LCK kinase domain structure*** *in its active form where Y394 is phosphorylated, as revealed by X-ray crystallography (PDB:3LCK)* [18].

According to the crystal structure of the LCK kinase obtained from its open conformation (PDB:3LCK) [18], the N-terminal lobe is constituted mostly by β-strands (β1 to β5) and a single α-helix with 3 to 4 turns. On the other hand, the C-terminal lobe is mostly made of α-helices namely αD, αE, αDE, αF, αG, αH, αI. The structure also indicates short α-helical turns namely αDE (bridging αD and αE), αEF (bridging αE and αF), αHI (bridging αH and αI). The C-terminal lobe also contains a small β-sheet (β7, β8 strands) at the interface of the N- and C-terminal lobes.

The activity of the kinase is mediated by its catalytic loop (residues ~382 to 402) which contains the phosphorylation site at Y394. In the active form of LCK, pY394 is stabilised by R363 and R387 via ionic interactions [18]. At the C-terminal loop of the kinase domain, the Y505 phosphorylation site suppresses LCK activity by binding to the R154 of LCK-SH2 and adopting a closed LCK conformation, and this mechanism is general among the members of the Src kinase family [110].

## 1.6 Activation of LCK and its conformational states

The CD45 phosphatase and intracellular kinase protein Csk compete in order to regulate LCK in resting cells. In the ‘closed’ or ‘inactive’ form of LCK, (i) Y505 is phosphorylated by Csk [111], and/or (ii) Y394 is dephosphorylated by CD45 in the presence of pY505 [112]. When neither Y394 nor Y505 are phosphorylated,

LCK exhibits a 'primed conformation'. The 'open' or 'active' form of LCK is typically exhibited when Y394 is phosphorylated and Y505 is not. Other phosphatases such as the Src-homology domain containing Phosphatase 1 (SHP-1), Protein Tyrosine Phosphatase Non-receptor Type 2 (PTPN2) and 22 (PTPN22) also contribute to the dephosphorylation of pY residues in LCK [113].

A pool of constitutively active LCK not bound to CD4/CD8 co-receptors is suggested to initiate rapid T cell signal transduction [114,115]. It is also suggested that co-receptor-bound LCK is less active compared to free LCK [116]. In addition to this, Stirnweiss *et al.* found that a fraction of LCK undergo a conformational change into their open state after TCR-pMHC engagement, thus enhancing LCK-mediated signalling. Consistent with this, Rossy *et al.* conducted experiments showing that conformations of LCK further dictate phosphorylation events that are responsible for triggering T cell signalling [117].

The clustering radius of LCK is higher in activated T cells than in resting T cells [117]. Together, this indicates that LCK reorganisation along with its kinase activity contributes to T cell signal initiation. However, the clustering alone is independent of LCK kinase activity and driven by its conformation where its open state is correlated with relatively high-level clustering, and closed state with low-level clustering. In addition, LCK with R154K mutation induced in the SH2 domain pY binding pocket, or Y505F at the C-terminus (forced open / primed state) showed better clustering compared to its wild-type. This verified that potential closed conformational states of LCK prevented clustering, and open conformations increased it. After disabling the SH3 domain (W97A), deactivating the kinase domain (K273R) and forcing an open state of LCK (Y505F), clustering was again observed to be similar to that of the open state and greater than that of the wild-type. Thus, it was concluded that LCK clustering occurred due to intermolecular associations with other LCK molecules and not with other signaling proteins [117].

## 1.7 The TCR-CD3 signalling pathway and its regulation

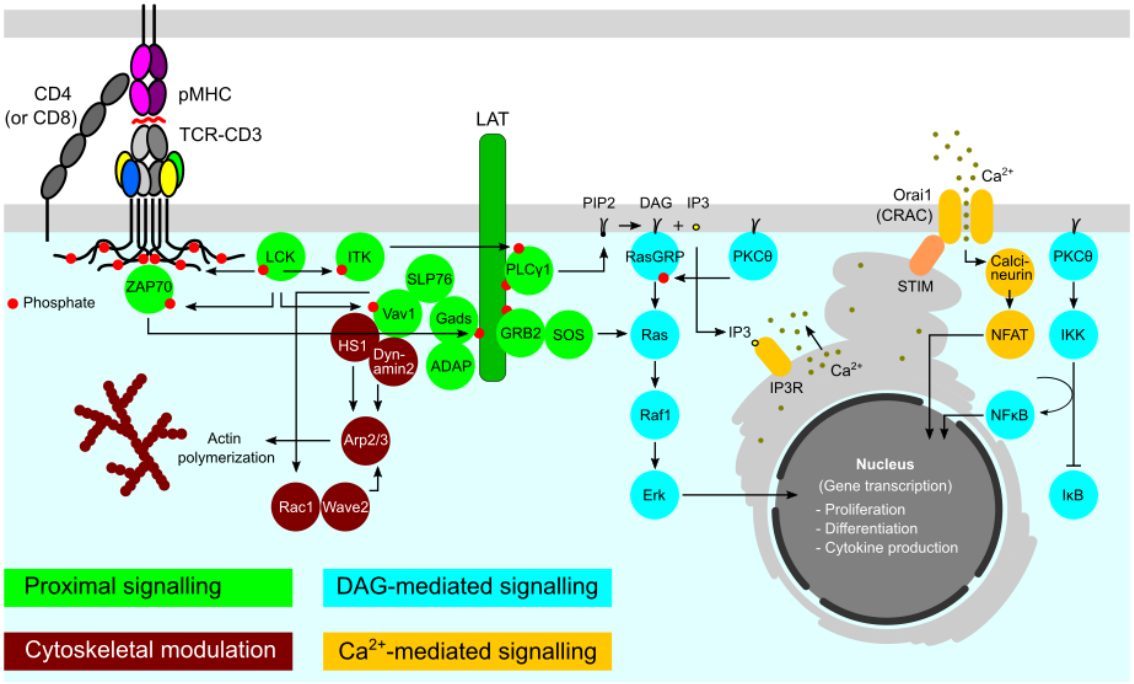


***Fig 1.13 The TCR-CD3 signalling pathway.***

### 1.7.1 Proximal signalling

The first step in TCR-CD3 intracellular signalling after TCR-pMHC ligation is the activation of LCK, an Src-family protein tyrosine kinase, which leads to the phosphorylation of TCR-CD3 ITAMs. Upon antigen recognition, the flipping of anionic lipids such as phosphatidylserine (PS) from the inner to the outer leaflet of the plasma membrane [118] aids in sustaining the exposure of ITAMs for phosphorylation. ZAP-70 protein is then recruited to the phosphorylated ITAMs paving way to phosphorylation of ZAP-70 itself by LCK [68].

The key targets of ZAP-70 are (i) Linker for Activation of T cells (LAT) of mass 36-38 kDa, and (ii) Src homology 2 (SH2) containing Leukocyte Phosphoprotein of mass 76 kDa (SLP-76), which is an intracellular adapter protein [119]. These adapter proteins are important adapter proteins in TCR-CD3 signalling since they organise effector molecules in the right spatial order to ensure activation of many downstream signalling pathways. Studies were even conducted to show that loss of either SLP-76 or LAT destructs almost the whole TCR-CD3 signalling pathway [119–121].

There are 9 tyrosine residues phosphorylated on LAT following TCR-pMHC ligation and these bind to the C-terminal of the Phospholipase-Cγ1 (PLCγ1)-SH2 domain, p85 subunit-containing phosphoinositide-3-kinase (PI3K), growth factor receptor-bound protein 2 (GRB2) and GRB2-associated adapter

downstream of Shc (Gads). Gads is a mutual partner that then helps SLP-76 bind to phosphorylated LAT [122].

Soon after TCR-pMHC ligation, Interleukin-2 inducible T cell kinase (ITK) is brought to the membrane via interactions of its PH domain with $PIP_3$ [123]. LCK then phosphorylates ITK, after which ITK-SH2 and SH3 domains interact with pY145 and the proline-rich region (PRR) of SLP-76, respectively [124–126]. While LAT and SLP-76 carry out their respective functions, there are important effector molecules that promote stability of the complex: (i) the Tec-family protein kinase ITK recruits guanine nucleotide exchange factor (GEF) Vav1 to the T cell-APC contact region, (ii) Vav1 phosphorylates SLP-76 and carries it to LAT. Vav1 also activates ITK [127–129].

The PLCγ1 isoform interacting with SLP-76, LAT, Vav1 and ITK is required provide stability to PLCγ1 in its appropriate conformation within the complex to optimise its activity [130,131]. Studies suggest that Y128F or Y145F, along with Y112F mutations in SLP-76 in primary T cells do not reduce SLP-76 / Vav1 / NCK / ITK interactions [132] but these mutations do result in severe defects in downstream signalling. The same applies to the tyrosines of Vav1 as well, with relevance to their binding partners. These results provided by immunoprecipitation suggest that interactions between SH2 and pY residues may contribute to regulation of activation of effector molecules [133]. Although, structural studies show that ITK-SH2 domain interacting with pY may allow kinase activity through conformational change [134].

With the end of the proximal signalling pathway, PLCγ1 is activated by ITK bound to SLP-76. Phosphatidylinositol-4,5-biphosphate ($PI(4,5)P_2$ / $PIP_2$) in the membrane is then hydrolysed by activated PLCγ1 giving rise to inositol triphosphate (IP3) and diacylglycerol (DAG) [68,135]. Other PLCγ1-dependent pathways are $Ca^{2+}$ signalling [136], rearrangements in the cytoskeleton and pathways to activate integrins. CD28 is a co-receptor which, on binding, optimises these pathways [68].

### 1.7.2 DAG-mediated signalling

DAG, when produced as a result of $PIP_2$ lipid hydrolysation by PLCγ1, activates 2 major pathways involving Ras and PKCθ.

**Ras pathway:** Ras is an intracellular protein activated by GTP binding and is responsible for activating serine-threonine kinase Raf1. Raf1 is further known to activate mitogen-associated protein kinases (MAPKs) which include MEK and the extracellular signal-regulated kinase (Erk). Ras activation is induced majorly

by guanine nucleotide exchange factors (GEFs) and suppressed by GTPase activating proteins (GAPs). There are 2 Ras GEFs in T cells: son of sevenless (SOS) and Ras guanyl nucleotide-releasing protein (RasGRP). Recruitment of RasGRP to the membrane is facilitated by a DAG-binding domain where it is phosphorylated by PKCθ [137,138]**.** The GRB2-SH2 domain is recruited to LAT and binds to its pY residues, thus incorporating SOS into the proximal signalling complex where it can orchestrate localised activation of Ras [120]. This model explains RasGRP dominance by its production of RasGTP. Sufficient levels of RasGTP can boost SOS activity and further amplify signal transduction.

**PKCθ pathway:** Protein Kinase C Theta (PKCθ) is a PKC-family protein containing a lipid binding domain specific for DAG [139]. This eases the recruitment of PKCθ to the membrane after the T cell is activated. Although, it is speculated that it may need to be phosphorylated by LCK to undergo a structural change enabling it to bind to phosphatidyl serine (PS) lipids. This would increase its probability of binding to DAG further leading to the activation of PKCθ [140].

Importantly, PKCθ regulates the NFκB activation pathway. NFκB is an intracellular protein and a transcription factor. Associating with its inhibitor IκB prevents NFκB from being transportated into the nucleus. Following T cell activation, IκB kinase (IKK) complex phosphorylates, ubiquitinates and degrades IκB, permitting NFκB to enter the nucleus. Here, NFκB activates functional genes responsible for the survival and homeostasis of T cells [141]. Mucosa-associated lymphoid tissue lymphoma translocation gene-1 (MALT1) enhances NFκB signalling by cleaving A20, a de-ubiquitinating enzyme and a suppressor of NFκB activation [142,143].

### 1.7.3 $Ca^{2+}$-mediated signalling

$Ca^{2+}$ ions act as messengers in living eukaryotic cells. The IP3 produced by PLCγ1 after T cell activation activates $Ca^{2+}$-permeable ion channel receptors such as IP3 receptors (IP3R) on the endoplasmic reticulum (ER) membrane, which releases ER $Ca^{2+}$ into the cytoplasm. Depletion of ER $Ca^{2+}$ provokes influx of $Ca^{2+}$ from the extracellular space. Activation of the $Ca^{2+}$ release-activated $Ca^{2+}$ (CRAC) channels such as Orai1 in the membrane help with the influx of $Ca^{2+}$ ions. This process is called store-operated $Ca^{2+}$ entry (SOCE) [136,144].

TCR-CD3 signalling also induces an increase in intracellular $Ca^{2+}$ levels which result in the activation of $Ca^{2+}$ and calmodulin-dependent transcription factors

and signalling proteins like the phosphatase calcineurin and the $Ca^{2+}$-calmodulin-dependent kinase (CaMK). These signalling proteins can then activate a variety of transcription events [145]. Calcineurin, when activated, dephosphorylates the Nuclear Factor of Activated T cells (NFAT) family members, helping them translocate to the nucleus. Inside the nucleus, NFATs can form transcriptional complexes with many transcription factors. Some pathways including NFκB and JNK are said to be $Ca^{2+}$ dependent and thus require signals emphasising interaction between other T cell signalling pathways and $Ca^{2+}$ [146]. Additionally, $Ca^{2+}$ influx into the cytoplasm is suggested to neutralise anionic lipid headgroups near TCR-CD3 complexes preventing lipid binding of their cytoplasmic region [32] and further sustaining T cell signalling.

### 1.7.4 Cytoskeletal modulation and the immunological synapse

TCR-pMHC binding also gives rise to changes in cytoskeletal morphology as the activated T cell accumulates filamentous actin (F-actin) at the site of contact [147]. TCR-CD3-induced activation of actin-related protein 2/3 complex (Arp2/3) plays a role in actin polarisation, but pathways that are independent of this complex also contribute to the accumulation [148].

Arp2/3 activation is dependent on its interaction with nucleation-promoting factors (NPF). These NPFs include Wiskott-Aldrich Syndrome protein (WASp), WASp family verprolin homologous protein (Wave2), and hematopoietic cell lineage specific protein-1 (HS1). WASp interacts with NCK, a protein associated with SLP-76, which is then brought to the TCR-CD3 activation site and stimulated depending on the activation of Rac1, Rho-family GTPase, by Vav1 [149]. Similarly, Wave2 is stimulated depending on the Vav1 induced activation of another Rac1 [147]. Henceforth, it is suggested that other Vav1 functions are important in causing TCR-CD3-induced morphological changes in actin filaments. Additionally, it was also observed that Vav1 can help in activating WASp and Wave2 by recruiting Dynamin2, a GTPase responsible for TCR-CD3-induced actin dynamics [150].

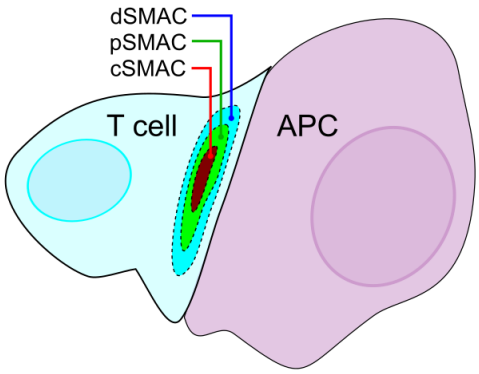


***Fig 1.14 The immunological synapse.***

The accumulation of actin filaments at the T cell-APC contact site is caused by the microtubular organising centre (MTOC) moving toward the contact site [151]. The polarisation of the MTOC toward the T cell-APC contact site is linked with the formation of 'Immunological Synapse' (IS) (Fig 1.14), an organisation of key membrane proteins, at the TCR-pMHC contact site [152,153]. The IS consists of three concentric regions: the TCR-CD3-rich central supramolecular activation cluster (cSMAC), an integrin-rich peripheral SMAC (pSMAC) surrounding cSMAC [154], and a distal SMAC (dSMAC) on the outer edge containing membrane proteins with large extracellular domains [152,155]. Since the components of the IS spatially organise to aid in T cell activation, these membrane proteins have been treated as pharmacological targets to modulate T cell mediated immune responses [156].

### 1.7.5 Negative regulation of the T cell signalling pathway

TCR-CD3 signalling brings about a series of downstream signalling events that involve activation of many effector molecules and their pathways. Similar to positive regulation of T cell function and signalling, negative regulation also occurs via TCR-CD3-mediated signals and those generated from other membrane receptors. For instance, the CD45 phosphatase can act as a positive regulator by dephosphorylating LCK pY505, as well as a negative regulator by dephosphorylating LCK pY394 [109]. Although, its positive or negative regulation effect on LCK depends on (i) its proximity to TCR-CD3-activated effector molecules, and (ii) if CD45 itself is in a favourable conformation to dephosphorylate [68].

Negative regulation of TCR-CD3 signals can also be carried out by adapter proteins like the Downstream of Kinase (Dok) adapter proteins, Dok-1 and 2, which are expressed in T cells [157]. Dok can partner with many negative

regulators such as C-terminal Src kinase (Csk), GRB2 and SH2 containing inositol polyphosphate 5-phosphatase 1 (SHIP-1) [158]. Studies have shown that pY binding domain of both Dok adapter proteins can associate with the ITAM motif of ζ chain of the TCR-CD3 complex. This instigates a speculation that Dok may compete in binding to the ITAMs with ZAP-70 during TCR-CD3 activation.

The hematopoietic progenitor (serine/ threonine) kinase (HPK1), also called MAP4K1, is found to negatively regulate TCR-CD3 proximal by binding to the SLP-76 complex, suggesting that SLP-76 may serve as an important target for negative regulation [159]. The members of the Cbl family, c-Cbl and Cbl-b, are another set of proteins that negatively regulate T cell signalling [160,161]. These proteins interact with Src family and ZAP-70 kinases and function as ubiquitin ligases to mark them for degradation impairing downstream signalling [162].

More cases of negative regulation of T cell activity upon TCR-CD3 triggering include Cytotoxic T lymphocyte Antigen-4 (CTLA-4) and Programmed Death-1 (PD-1) proteins [163]. The CD28 co-receptor being homologous to its inhibitory competitor CTLA-4, binds to B7 molecules i.e. CD80 and CD86 on the APC surface and competes in the regulation of T cell signalling [164]. However, CTLA-4 binds to the CD80 and CD86 with higher affinity than CD28 [165]. Given that CTLA-4 and PD-1 inhibitors are known to prevent anti-tumour T cell responses, they have played a significant role in cancer immunotherapy [166,167]. Antibodies were targeted against CTLA-4 and PD-1 and human clinical trials were conducted. Following their success, the anti-CTLA-4 and anti-PD-1 antibodies sought FDA approval for cancer therapy in 2011 and 2014 respectively, thereby taking immune checkpoint therapy a step forward [165].

## 1.8 Computational methods in studying membranes and membrane proteins

Membrane Proteins constitute roughly 20 to 30% of the proteome in most prokaryotic and eukaryotic organisms [168,169]. Moreover, greater than 50% of drug targets were shown to be membrane proteins [170]. Following the 50$^{th}$ anniversary of the Protein Data Bank (PDB) database [171] in 2021 and as of 31$^{st}$ August 2021, there were 181535 available biomolecular structures out of which ~5000 were membrane proteins and ~1310 were unique structures. The low outcome of membrane protein structures, unlike soluble proteins, is due to complications during expression, purification, and extraction while being

reconstituted in a membrane environment. Nevertheless, X-ray crystallography [172], electron crystallography [173], fluorescence techniques [174], and lipid nanodiscs [175] have been used to identify protein-lipid interactions. Although, these approaches are often limited to identifying strong lipid interactions and do not guarantee high spatial resolution. Given the increasing number of membrane protein structures being solved experimentally (as shown and updated here: https://blanco.biomol.uci.edu/mpstruc/), there is a need to decipher their dynamics, lipid interactions and potential activation mechanisms when inserted in a membrane. Therefore, to assist experimental efforts in the structural and functional characterisation of membrane proteins, computational modelling and molecular dynamics (MD) simulation studies have been carried out extensively [176–182].

One of the computational approaches established to address the protein-lipid interplay of static membrane protein snapshots is the automated pipeline and database called MemProtMD [183,184]. Here, membrane protein structures that are released on PDB are extracted and simulated in an explicit membrane-solvent environment using the multi-scale approach i.e. involving coarse-grained and atomistic resolutions [185]. The simulations are further analysed and their results are maintained online where they can be visualised interactively (http://memprotmd.bioch.ox.ac.uk/).

In addition, the recent advent of artificial intelligence (AI) in protein structure prediction and protein folding, such as AlphaFold2 and RoseTTaFold, has shown promising results [186–188] and may help fast-forward structural biology studies in the coming years. In addition, AI-assisted MD simulation studies have helped increase sampling and study the conformations of loop regions in protein structures in atomistic resolution. This study further enables ranking free energies of known loop conformations and discovery of new ones [189]. Parameterisation of forcefields is also a demanding area of research where machine learning techniques are starting to contribute [190–192].

Speaking of the Covid-19 pandemic, the Gordon Bell 2020 prize was awarded for high performance computing-based Covid-19 research based on innovation in computational methods. Exploiting the limits of computational power, the prize-winning authors used atomistic MD simulations combined with AI to simulate and study the viral spike proteins embedded in a membrane envelope [193]. The simulated system in this study contained a ground-breaking 305 million atoms. They employed a weighted ensemble MD method which enabled them to specifically observe the conformational change of the spike protein upon binding to human ACE2 receptors, which is suggested to lead to viral entry into the host cell thereby causing Covid-19 infection. The AI algorithm

called DeepDriveMD [194] was developed to analyse the large dataset obtained from simulations and interpret the dynamics of the SARS-Cov-2 spike protein. Overall, given a static resolved structure of proteins, these studies indicate the capabilities of MD simulations in providing greater insights into protein dynamics and provide critical information for drug designing.

In addition to the applications of AI in MD simulations, the Folding@home project (foldingathome.org) has shown immense potential of computing protein dynamics. Folding@home is a crowd-sourced computing project which is run by a community of scientists and organisations across the world who dedicate their computational resources toward running small simulations that are then used to build a larger map of protein dynamics. This project shifted its focus to the SARS-Cov-2 virus since the pandemic. Given an internet connection and a standard computer today, anybody may volunteer to contribute to the Folding@home project, thereby giving the common man an opportunity to be part of a solution. As a result, the project grew from around 30,000 active computers to more than a million (as of May 2021) collaboratively creating the first exascale computer by simulating 100 milliseconds of the SARS-Cov-2 viral proteome [195]. This expanded the options for drug designing. More importantly, it established that computational approaches are key to achieving high performance studies of protein dynamics.

In the light of lipid composition in mammalian plasma membranes, MD simulations have typically involved a handful number of lipid types. As an advancement into using complex membranes, Ingólfsson *et al.* modelled a more realistic cell membrane consisting of 63 lipid types asymmetrically arranged [176]. This led to more studies involving complex lipid bilayer environments [178,196–198]. Furthermore, a recent review reveals a glimpse of an even more complex plasma membrane environment fused with proteins, associated with actin cytoskeleton, and containing physiological metabolite gradients, pH, ions etc. [199]. However, in the process of simulating such large and complex biomolecular systems, barriers in terms of computational cost still exist.

To reduce computing time and the cost of MD simulations, a number of methods such as biased MD simulations and coarse-graining techniques have been developed. The most widely used coarse-grained (CG) forcefield is Martini where a group of four heavy atoms is replaced by a single pseudo-atom in order to reduce the number of calculations required to be performed by the computer. Although this facilitates simulations at larger time scales, it comes at a cost of lower resolution i.e. approximation of atomic coordinates [200]. Furthermore, ultra-coarse-graining techniques have also been developed where a group of amino acids are substituted by a single pseudo-atom to enable even larger MD

simulations [201,202]. Following the studies on the SARS-Cov-2 viral envelope and the spike protein, this ultra-CG technique was used to obtain the first complete model of the coronavirus in 2021 [203].

Further, a 'hybrid' scale of simulations has also been introduced where atomistic resolution (AT) and Martini forcefield-based CG structures can be simulated together [204]. Soon, another strategy was developed to convert or 'backmap' CG models to AT [205]. Multiscaling and hybrid multiscaling have thereafter been considered an integral part in the circle of computational methodologies [181,206]. However, some challenges may be met while developing hybrid models [207]. In terms of 'backmapping' complex structures, recent work proposed a new approach i.e. the TS2CG algorithm, which involves creating a triangulated mesh surface in order to convert large structures into near-atomic coarse-grained models based on the Martini forcefield. Using this approach, this study successfully modelled and simulated the double membrane of a whole mitochondrion consisting of various lipid species [208]. This study paves a way to study molecular events at sub-cellular and whole cell levels.

Overall, although MD simulation forcefields may undergo further optimisation and require experimental validation to study molecular behaviour, they are currently playing an enormous role in bringing static protein structures to life, often with results that are comparable with experiments. Computational modelling and simulations are also key in enabling visualisation and analyses at near-atomic scales, and building hypotheses to guide further studies. Given the expanding and combined power of molecular modelling, coarse-graining techniques, MD simulation methods, and the assistance of AI, computational approaches are likely to grow and provide a theoretical basis to understand life at the molecular level.

## 1.9 The aim and summary of this research

As described in this thesis chapter, there is a significant amount of evidence on the structure and function of the TCR-CD3 complex and LCK. However, due to the absence of structural data on the complete TCR-CD3 complex and the full-length LCK, and their association with the plasma membrane, our molecular understanding of T cell signal initiation is limited. While it is challenging for experiments to structurally resolve the intrinsically disordered regions of the TCR-CD3 and LCK, and provide dynamic information at the molecular level, molecular modelling combined with MD simulations serve as a feasible approach to address this question. Therefore, in this thesis, I employ these

techniques to provide complete structural models and study various aspects of the dynamics of the TCR-CD3 and LCK proteins in a membrane environment.

In chapter 2, I describe the theoretical foundation of molecular dynamics simulations in both coarse-grained and atomistic resolutions. I also discuss multi-scale approaches where the same molecular system can be studied in multiple resolutions to enable both larger time-scales and higher resolution.

In chapter 3, I study the assembly, dynamics, and allosteric effects of mutations in the TCR-CD3 transmembrane region (TMR). Due to the lack of data on the protein-protein interactions within the CD3δε and CD3γε dimers, I simulated the assembly of the CD3 dimers followed by their assembly with the TCRα/β subunits respectively. To study each of these assembly processes and yield reliable protein-protein interaction data, many individual simulations were conducted in both coarse-grained and atomistic resolutions. Subsequently, upon the release of the cryo-EM structure (PDB:6JXR), these data were compared to the static structure, indicating fairly consistent results. Further, using the coordinates of the complete TCR-CD3 TMR from the cryo-EM structure, and in collaboration with Prof. Oreste Acuto and group at the University of Oxford, I studied the effects of certain mutations i.e. βY291, ζI38A, ζI41A, that were shown to mimic an allosteric activation mechanism upon pMHC binding *in vivo* [76].

In chapter 4, using steered MD simulations, I study the interactions of certain variants of peptides attached to a class I MHC with the TCRαβ extracellular domain (ECD). This research was initiated in collaboration with Dr. Omer Dushek and group at the University of Oxford. My results indicate that the CDR3 loops of the TCRαβ ECD interact significantly with specific residues i.e. M4 and W5 of the peptide SLLMWITQV (obtained from PDB:2BNQ [14]; view in 3D here: https://www.rcsb.org/3d-view/2BNQ), and thus substitution of these residues diminishes TCR-peptide interaction and bond lifetime. More interestingly, I identify three distinct structural changes occurring within the class I MHC as a force is applied to pull it away from the TCR. The most common structural change observed in my simulations, referred in this thesis as ‘linker stretch’, is consistent with previous experimental reports [209], while the rest have not been reported before. Nevertheless, I discuss their probabilities of occurrence and potential impact on TCR-CD3 activation.

In chapter 5, I use the coordinates of the TCR-CD3 ECD and TMR from the available cryo-EM structure (PDB:6JXR) and model its cytoplasmic region (CYR) in order to enable MD simulations of the complete TCR-CD3 complex. This chapter provides insights into the (i) membrane association and dynamic nature of the cytoplasmic tails, (ii) the lipid interactions of the complete TCR-

CD3 in a complex asymmetric bilayer containing predominant lipids in the native membrane activation domain, as suggested by lipidomics studies [210], and (iii) the conformational changes in the ECD and in the TMR.

In chapter 6, I use the coordinates of the available crystal structures of the LCK-SH3 and SH2 domains (PDB:4D8K; view in 3D here: https://www.rcsb.org/3d-view/4D8K), and its kinase domain (PDB:3LCK; view in 3D here: https://www.rcsb.org/3d-view/3LCK), and further model the unique domain and the SH4 domain along with its post-translational modifications i.e. myristoylation and palmitoylation. This enabled CGMD simulations of the full-length LCK in both its open (active) and closed (inactive) conformations. These simulations shed light into the membrane association and lipid interaction sites throughout the protein. Interestingly, my results provide novel insights into a cationic patch in the kinase domain that aids in its interaction with PIP lipids. In addition, experiments have shown that PIP lipids regulate LCK activity via its SH2 domain [27]. Further to this, my simulations indicate a shift in the PIP lipid interaction site of the LCK-SH2 when comparing the membrane-bound state of the open (active) and closed (inactive) forms of LCK. This potentially suggests a correlation between lipid interaction of LCK and its conformational state. Simulations were also performed in AT resolution to study the protein-protein interactions of the open and closed states of LCK. Comparing the open conformation of LCK between its membrane-bound state (CGMD) and in solution (ATMD), my simulations show that the kinase domain possesses an ability to wobble and rotate, potentially indicating its flexibility prior to phosphorylating the ITAMs of the TCR-CD3.

Finally, in chapter 7, I discuss how my research and the data provided in this thesis inform future experimental and MD simulation studies on various aspects of TCR-CD3 activation, and help in answering additional questions to gain a deeper molecular understanding of T cell signal initiation.

# Chapter 2

# Theory and Methods

This chapter provides an explanation to the theory that lays the foundation to the molecular dynamics simulations conducted in this thesis.

## 2.1 Molecular dynamics simulations

Molecular dynamics (MD) simulations act as an important tool to study the dynamics of biological systems at the molecular level. Moreover, these biochemical systems can be studied in a range of time-scales and resolutions i.e. quantum, atomistic, coarse-grained, ultra coarse-grained (Fig 2.1).

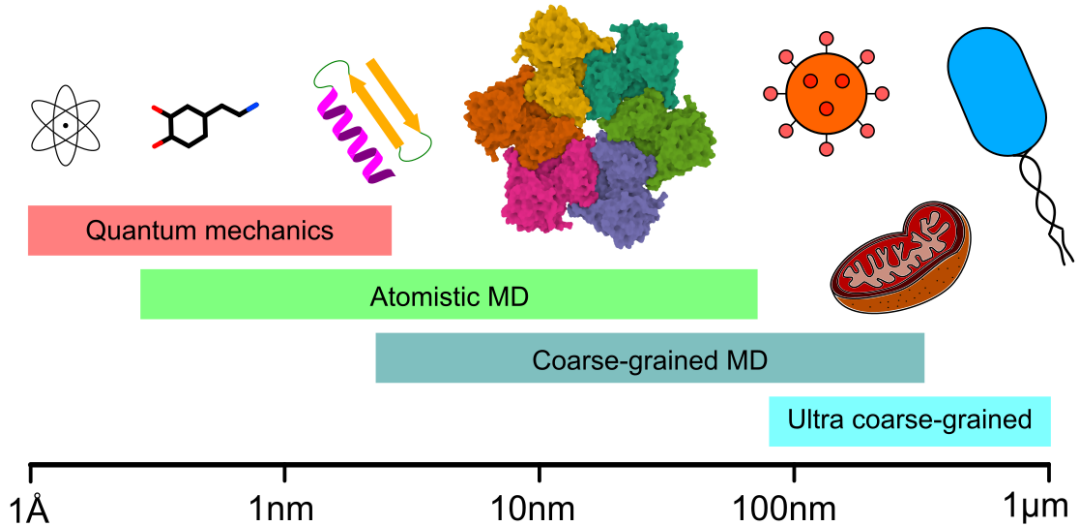


***Fig 2.1 Quantum, atomistic, coarse-grained and ultra coarse-grained scales used in simulations.***

Quantum mechanical systems, which are often used to study chemical reactions such as electron transfer at the scale of Angstroms, are described using Schrodinger's equation. However, simulating such a system containing atoms with explicit electrons can be computationally expensive and potentially simulated for a limited period of time. On the contrary, interesting biological phenomena occur at larger scales of size and time.

Classical molecular dynamics simulations based on molecular mechanics are typically conducted in atomistic resolution following the laws of Newtonian physics and the Born-Oppenheimer approximation i.e. where atoms are considered as single particles without containing explicit electrons. Although classical MD simulations do not account for quantum effects such as electron distribution and transfer, they do enable simulations of larger biomolecular

systems and/or over larger time-scales when compared to quantum mechanical simulations. While a hybrid of quantum and molecular mechanical systems (QM/MM) can be simulated [211] where a set of atoms and associated chemical bonds are treated as a quantum sub-system inside a molecular system, this technique is yet again computationally expensive compared to classical MD simulations. To enable simulations of even larger sizes and time-scales, the coarse-graining technique can be employed at the cost of reducing the resolution of the molecular system even further (explained in section 2.3). In this thesis, all coarse-grained (CG) MD simulations have been conducted using Gromacs 5.0 [212], whereas all atomistic (AT) MD simulations are done using Gromacs 2016.4 [213].

### 2.1.1 Forcefield

In this thesis, all ATMD simulations are based on the CHARMM36 forcefield [214]. A 'forcefield' refers to a set of parameters governing the energy and interactions between particles of the system. The total potential energy (U) of the system is given as the sum of the potential energy of all bonded interactions and that of all non-bonded interactions between particles in it:

$$U = \sum U_{bonded} + \sum U_{non-bonded}$$

The bonded and non-bonded components defining the potential energy of the system are further a sum of underlying components as described below.

#### 2.1.1.1 Bonded interactions

The bonded interactions component is the potential energy between particles that are covalently bonded. The covalent bond shared between two particles can give rise to potential energies during bond stretching, bond bending at an angle, and rotation of bonds (dihedral angle), which together constitute the total potential energy of the bonded interactions.

$$U_{bonded} = \sum U_{bonds} + \sum U_{angles} + \sum U_{dihedrals}$$

The stretching of a covalent bond linking two particles, *a* and *b*, of length $l_{ab}$, is given by a potential energy function taken from Hooke's law i.e. $F_s = -kx$, where $F_s$ = spring force, k = spring constant, x = stretching /compressing of the spring.

$$U_{bonds}(l_{ab}) = \frac{k_{ab}^{l}}{2}(l_{ab} - l_{ab}^{0})^2$$

*where $l_{ab}$ is the bond length, $l_{ab}^{0}$ is the reference length, and $k_{ab}^{l}$ is the spring force constant.*

Similarly, the potential of the bending of the covalent bond at an angle 'θ' is given by three particles i.e. *a, b, c*.

$$U_{angles}(\theta_{abc}) = \frac{k_{abc}^{\theta}}{2}(\theta_{abc} - \theta_{abc}^{0})^2$$

*where $\theta_{abc}$ is the bond angle between three particles, $\theta_{abc}^{0}$ is the reference angle, and $k_{abc}^{\theta}$ is the spring force constant.*

Dihedral or torsion angles are those angles calculated between two planes, one of them constituted by particles *a, b, c*, and the other constituted by *b, c, d*.

$$U_{dihedrals}(\emptyset_{abcd}) = k_{\emptyset}(1 + cos(n\emptyset_{abcd} - \emptyset_{abcd}^{0}))$$

*where $\theta_{abcd}$ is the dihedral angle between the two planes, $k_{\emptyset}$ is the spring force constant, n is the multiplicity of the dihedral term $\emptyset_{abcd}$ as the bond rotates 360°, and $\theta_{abcd}^{0}$ is the reference angle.*

There are two types of dihedral angles: proper and improper. Proper dihedral potentials are periodic, with their maximum potential energy exhibited at an unfavourable '*cis*' or 'eclipsed' conformation and a minimal potential energy exhibited at a favourable '*trans*' or 'staggered' conformation, considering butane for example as shown in Fig 2.2. Improper dihedral potentials are similar but also maintain chirality of tetrahedral groups.

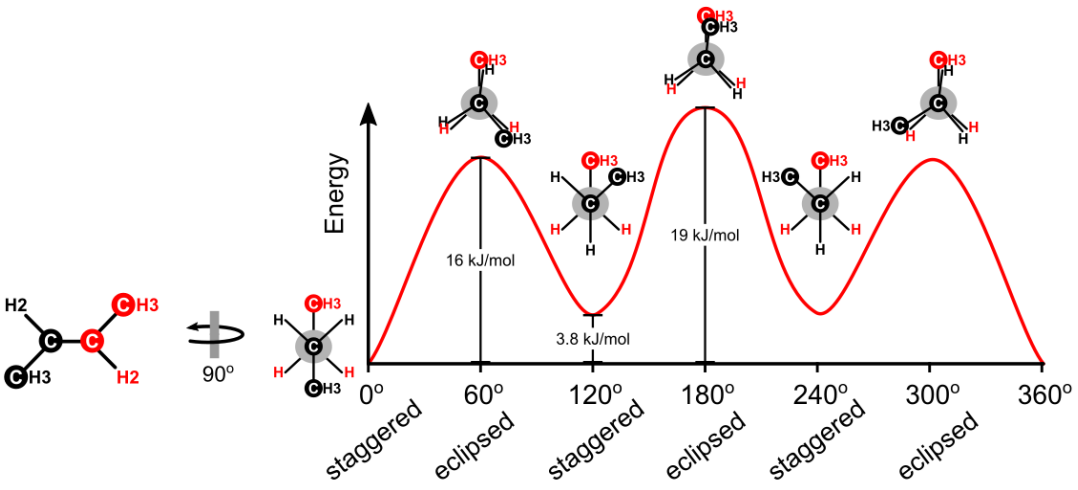


***Fig 2.2 Proper dihedral potentials** of butane* [215]*.*

#### 2.1.1.2 Non-bonded interactions

The non-bonded interactions component in the molecular system is the sum of the potential energies of Van der Waal's (VdW) (including steric repulsion) forces and electrostatic (attractive and repulsive) forces. The former is defined by the Lennard-Jones potential while the latter is defined by the Coulombic potential (or Coulomb's law).

$$U_{non-bonded} = \sum U_{VdW} + \sum U_{Electro}$$

The Lennard-Jones potential, which is applied to all atoms or particles in the simulation, follows a 12-6 rule where the repulsive forces occurring at short distances represent the $\left(\frac{1}{r_{ab}}\right)^{12}$ term and the attractive forces occurring at longer distances represent the $\left(\frac{1}{r_{ab}}\right)^{6}$ term.

$$U_{VdW}(r_{ab}) = 4\varepsilon_{ab}\left\{\left(\frac{\sigma_{ab}}{r_{ab}}\right)^{12} - \left(\frac{\sigma_{ab}}{r_{ab}}\right)^{6}\right\}$$

*where $r_{ab}$ is the distance between the two particles 'a' and 'b', $\varepsilon_{ab}$ is the magnitude of attraction, and $\sigma_{ab}$ is the distance between the particles when the potential energy is zero.*

The potential energy of the electrostatic forces arises from the interaction of charged particles based on Coulomb's law. This is modelled by the Particle Mesh Ewald (PME) method [216] used in ATMD simulations with Gromacs software. This Coulombic potential between particles 'a' and 'b' containing partial charges $q_a$ and $q_b$ respectively is given as:

$$U_{Electro}(r_{ab}) = \frac{q_a q_b}{4\pi\varepsilon_0 r_{ab}}$$

where $r_{ab}$ *is the distance between the two particles 'a' and 'b'* and $\varepsilon_0$ is the permittivity of vacuum.

The partial charges $q_a$ and $q_b$ are assigned to the particles in order to mimic the distribution of electrons in atoms in MD simulations. In addition, a certain distance cut-off for both the VdW and coulombic interactions needs to be set during the simulation set-up to mimic reality and also to minimise the number of interactions to be calculated. In this thesis, a distance cut-off of 1.2 nm was used for the radius of both VdW ('*rvdw*') and coulombic ('*rcoulomb*') interactions

in all simulations. To avoid an abrupt appearance and disappearance of these interaction potentials near the distance cut-off, the 'shift' algorithm was used for their smoothening in CGMD simulations (*'rvdw_type = Shift'; 'rcoulomb_type = Shift*'). VdW interactions were smoothened between 0.9 and 1.2 nm in CGMD simulations, and between 1.0 and 1.2 nm in ATMD simulations (upper and lower limits specified by '*rvdw*' and '*rvdw_switch*' respectively). There was no smoothening for coulombic interactions, the distance cut-offs were maintained between 0 and 1.2 nm in all simulations. In ATMD simulations, the 'switch' algorithm was replaced by the 'cut-off' and 'force-switch' algorithms for VdW interactions (*'vdwtype = Cut-off' ; 'vdw-modifier = Force-switch'*), and by the PME algorithm for coulombic interactions (*'coulombtype = pme'*).

### 2.1.2 Integration method

Integration algorithms follow Newton's laws of motion ($F = m.a$) to determine the positions and velocities of particles in MD simulations. Calculation of the position and velocity of each particle in the simulation is done at periodic intervals of time called 'time-steps'. At each time-step ($\Delta t$), the force on each particle is calculated as the vector sum of its interactions with its neighbouring particles. Using this, the acceleration of the particle is calculated which further reveals its position and velocity in the next time-step. This integration calculation is repetitively done until the end of the simulation time which is pre-defined by the user.

In this thesis, I used the leap-frog integration algorithm [217] for all simulations, as implemented by default in Gromacs. It is named 'leap-frog' because, as the positions (r) of particles change at every integer time-step, the velocities (v) are updated at half the time-step, which when plotted looks like two frogs leaping over each other (Fig 2.3).

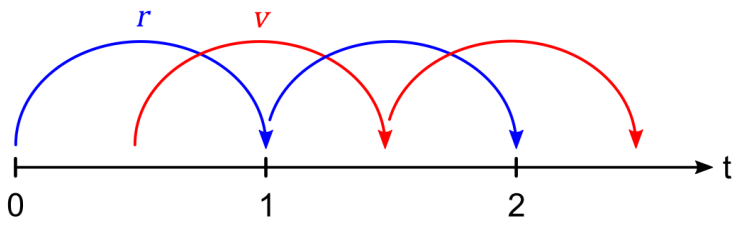


***Fig 2.3 The leap-frog integration algorithm.*** *Adapted from the Gromacs reference manual (https://manual.gromacs.org/documentation/2016.3).*

This algorithm is an extended version of the Verlet algorithm [218], but is more advantageous due to low error in velocities. The relation between the calculated positions (r) at a certain time (t), velocities (v) at time $t - \frac{1}{2}\Delta t$ is given as:

$$r(t + \Delta t) = r(t) + v(t + \frac{1}{2}\Delta t)\Delta t$$

$$v(t + \frac{1}{2}\Delta t) = v(t - \frac{1}{2}\Delta t) + \frac{\Delta t}{m} F(t)$$

*where 'm' is mass and, 'r' and 'v' are updated by force (F) at time 't'.*

### 2.1.3 Time-step and bond constraints

The time-step $\Delta t$, which is referred to as 'dt' in the Gromacs molecular dynamics parameter (mdp) file, is pre-defined by the user. While defining the time-step, an optimum value must be chosen. If the value is very small, the number of positions and velocities of each particle to be calculated will increase leading to a large amount of time and computational resource spent on the simulation. The fastest motion in an MD simulation is the vibration of bonds. Therefore, in contrast to small time-steps, if the time-step value is greater than the time required for bond vibrations, it will cause a shift in the equilibrium potential at each time-step eventually resulting in an accumulation of high energy artifacts and termination of the simulation. In order to lower the sensitivity toward the choice of time-step, algorithms such as LINCS [219] and SHAKE [220] have been developed. These algorithms apply constraints to bond lengths forcing bond lengths to return to a particular optimum length at each time-step, thereby allowing some flexibility in increasing the time-step and reducing the computational cost of performing simulations. In this thesis, the LINCS algorithm was used to constrain bond lengths. All ATMD and CGMD simulations were conducted using a time-step of 2 fs and 20 fs respectively.

### 2.1.4 Periodic boundary condition

MD simulations attempt to mimic a physiological molecular environment with a finite number of particles within a box. In contrast, there are no such physical boundaries to the existence of molecules in reality. This led to the development of the periodic boundary condition (PBC), where particles do not interact with the edges of the simulation box or stopped from moving any further. Rather, they pass through the edge of the box and appear on the opposite edge. Therefore, this simulation box referred to as a single 'unit cell' is surrounded by imaginary unit cells which are replicas of its own. In other words, at any given

time of the simulation, molecules at the edge of the box can partially be surrounded by imaginary replicas of those molecules present on the opposite edge of the simulation box (Fig 2.4).

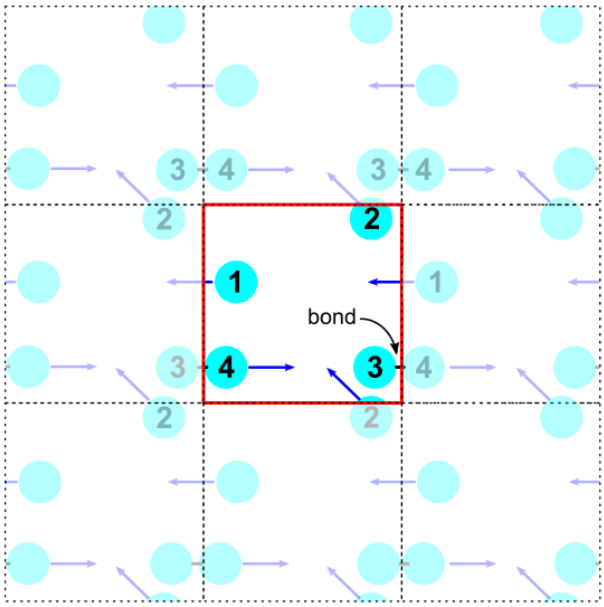


***Fig 2.4 Periodic boundary condition in two dimensions.*** *The red box represents the unit cell surrounded by replicas of its own enclosed in dotted lines. Both bonded and non-bonded interactions between particles can occur through the boundaries of the unit cell.*

While all the particles are free to leave unit cell and enter from the opposite end, the periodic boundary condition also helps in keep the number of particles as constant throughout the simulation, thereby contributing to the maintenance of potential forces. However, it is important to set up a unit cell that is large enough to ensure that a protein does not interact with itself across the boundary during the simulation. When simulating systems containing a membrane with a complex lipid composition, it is also important to ensure sufficient availability of lipid molecules that may interact with the protein. If the number of molecules of a certain lipid species is too limited in a unit cell, this may not provide enough opportunity for all the residues in the protein to explore their interaction with that lipid species, thereby causing a bias in protein-lipid interactions. In contrast, if membrane-protein simulations are performed using very large PBC dimensions, a high degree of curvature in the membrane can cause (the extracellular and intracellular regions of) membrane-embedded proteins to interact with each other across the boundary, which can be meaningless. In addition, large PBC boxes also increase the number of particles to be simulated, thereby increasing computational cost. Therefore, it is important to setup a PBC box with

appropriate dimensions to avoid meaningless protein-protein, protein-lipid interactions and to ensure efficient use computational resources.

### 2.1.5 Ensembles

Ensembles in MD simulations help in calculating the thermodynamic quantities such as the volume, energy, temperature, pressure etc., of the system. The following are popular ensembles used in MD simulations:

- NPT ensemble: maintains number of particles (N), pressure (P) and temperature (T) during the simulation.
- NVT ensemble: maintains number of particles (N), volume (V) and temperature (T) during the simulation.
- NVE ensemble: maintains number of particles (N), volume (V) and energy (E) during the simulation.

The NVT and NPT ensembles are oftentimes used together to equilibrate a molecular system, while the latter is typically used to conduct MD simulations post-equilibration. Perturbations in energy are an inherent property of biochemical processes. In addition, maintaining the pressure of the system allows proper packing of particles avoiding vacuum spaces. Therefore, in this thesis, all equilibration and production simulations were conducted using the NPT ensemble.

### 2.1.6 Temperature and pressure coupling

In order to use the NPT ensemble in MD simulations, algorithms that act as barostats and thermostats are used to maintain the pressure and temperature respectively. The Berendsen [221], Nosé-Hoover [222,223], velocity-rescale [224] algorithms provide as thermostats implemented in Gromacs, whereas the Berendsen and Parrinello-Rahman [225] algorithms provide as barostats. The Berendsen thermostat works by coupling the temperature of the system ($T$) to an external bath which is maintained at a reference temperature ($T_{ref}$) that is user-defined. Further, a temperature coupling constant ($\tau_t$) is used to specify the strength of coupling thereby potentially controlling the extent of fluctuations in temperature during the simulation. The velocity-rescale and the Nosé-Hoover algorithms are extensions of this approach that provide a more accurate and efficient method to perform constant temperature MD simulations.

In this thesis, irrespective of CGMD or ATMD simulations, I used the Berendsen thermostat during equilibration and the velocity-rescale / Nosé-Hoover

thermostat for production. In the ATMD simulations set up using CHARMM-GUI [226], I used the Nosé-Hoover thermostat implemented by default. However, for ATMD simulations conducted following 'backmapping' of CG structures to AT resolution (explained in section 2.4), I used the velocity-rescale thermostat. For all simulations conducted in solution, I maintained a temperature of 310 K, whereas I used 323 K for simulations containing a membrane to allow for lipid fluidity considering the saturation of their acyl chains. I used a temperature coupling constant of 1 ps for CGMD simulations and 0.1 ps for ATMD simulations.

The Berendsen barostat algorithm works in a way similar to the Berendsen thermostat i.e. by coupling the pressure of the system ($P$) to a user-defined reference pressure ($P_{ref}$) using a pressure coupling constant ($\tau_p$). This helps maintain the pressure of the system around the reference while the volume is allowed to change, which often leads to an alteration in PBC box dimensions relative to the initial setup.

All the CGMD and ATMD simulations conducted in this thesis were equilibrated using the Berendsen barostat while their production was run using the Parrinello-Rahman barostat which provides a more accurate method for the NPT ensemble. A pressure of 1 bar with a compressibility of $3\times10^{-4}$/bar and $4.5\times10^{-5}$/bar was used for CGMD and ATMD simulations respectively. Isotropic and semi-isotropic pressure coupling were applied to all simulations performed in solution and in a membrane environment respectively. The semi-isotropic pressure coupling ensured uniform pressure and compressibility applied on the X-Y axes along the plane of the membrane, while the isotropic type maintained uniform pressure and compressibility along all three dimensions (X,Y,Z axes) of the simulation box.

### 2.1.7 Energy minimisation and equilibration

The setup of a molecular system i.e. addition of proteins, lipids, solvent, ions etc., either in CG or AT resolution, typically results in random positioning of particles in the system generating steric clashes and potential energy forces. These lead to an accumulation of forces in the system and an unstable starting point for a molecular dynamics simulation. Therefore, the system must undergo an 'energy minimisation'. This process reduces the potential energies and steric clashes between particles, thereby relaxing the entire system over a certain number of time-steps. In addition, this process may be repeated multiple times to ensure that a minimum energy is reached. In this thesis, the steepest descent algorithm was used to perform energy minimisation for both AT and CG

systems over a maximum of 500,000 steps or until forces converged to a value less than 1000 kJ/mol/nm. In systems containing $PIP_3$ lipids, more flexible bond parameters were selected by specifying '*define = -DFLEXIBLE*' during the energy minimisation.

Following the energy minimisation process, in this thesis, all systems underwent a short equilibration between 2 and 5 ns where the protein backbone was position-restrained and the rest of the system was allowed to find its equilibrium prior to the actual 'production' simulation. Unlike the energy minimisation process, the equilibration allows spatial re-arrangement of molecules and packing of particles to remove void spaces. In simulations containing a protein embedded in a lipid membrane, the equilibration phase allowed re-adjustment of lipid molecules relative to the protein. Since equilibration simulations are not considered for analysis but rather to ensure proper packing of atoms before the production simulation, it is important not to lose any information on the dynamics of the protein during this process. Therefore, the atoms constituting the protein backbone were position-restrained during equilibration. To perform the equilibration and production simulations in this thesis, I used the NPT ensemble. The final snapshot from the short equilibration simulation is then used for generating multiple independent seeds for production simulations. All analyses performed in this thesis are a result of the multiple production simulations, unless otherwise mentioned.

## 2.2 Steered molecular dynamics simulations

Classical MD simulations or 'unbiased' MD simulations can pose a limitation in terms of obtaining samples of interesting molecular phenomena such as transitions in protein conformational states. This is due to the presence of a potentially high energy barrier between the two conformational states. In an unbiased MD simulation, depending on the magnitude of the energy barrier, a longer time-scale may be required to increase the probability of capturing the particular conformational state transition. To address this issue, 'biased' or 'non-equilibrium' methods are used where external forces can be applied to aid in a particular event within a shorter period of time. Steered MD simulations are one among these methods involving external force applied on a specified group of particles in a specified direction called a reaction coordinate.

In chapter 4 of this thesis, I have performed steered ATMD simulations using a centre-of-mass method to pull a group of particles away from a reference group. Here, the group of particles that are being pulled (pull group) are attached to an imaginary particle via an imaginary harmonic spring. Starting at the beginning of

my simulations, the imaginary particle starts to move in the specified direction (away from the reference) at a constant force thereby increasing the pulling force on the pull group with time, along the direction of pulling (reaction coordinate). The potential energy of this spring force can be given as:

$$U_{spring} = \frac{k_0}{2}(x(t) - x_0 - vt)^2$$

*where $k_0$ is the harmonic spring constant applied to the pull group, $x(t)$ is the spatial location of the pull group at a given time, t is the time of simulation, and v is the velocity of the pull group along the reaction coordinate.*

## 2.3 Coarse-grained molecular dynamics simulations

In ATMD simulations, the positions and potential forces are calculated for every atom in the system. In addition, the large number of atoms typically lead to a large number of interactions and further greater potential energy. This overall potential energy can escalate to higher values within a short period of time. Therefore, the accumulation of energy in the system needs to be neutralised over short time-steps (2 fs used in this thesis). However, these calculations performed over such short time-steps result in a high computational cost. For this reason, ATMD simulations are conducted over short time-scales (typically in ns) and may pose as a limitation in terms of the time-scale of the simulation and/or the size of the biomolecular system that needs to be simulated. Therefore, the coarse-graining approach was developed where a certain number of atoms in the AT system can be substituted by one coarse-grained bead. As a result, this technique reduces the number of particles, the number of potential interactions, and further the calculations to be computed at the cost of reducing the resolution of the biomolecular system. This simplified version of the system also helps smoothen its energetic landscape thereby allowing us to increase the time-step. Further, this allows simulations of larger biomolecular systems and/or simulations over longer time-scales (μs).

### 2.3.1 The Martini forcefield

In this thesis, I used the Martini 2.2 forcefield to conduct CGMD simulations using a 20 fs time-step. This forcefield follows a four-to-one rule, where four heavy (non-hydrogen) atoms are substituted by one coarse-grained bead on an average [200]. While this forcefield was initially developed to simulate lipids, it was later extended to enable CGMD simulations of proteins [227], carbohydrates [228], DNA [229], RNA [230], N-glycans [231] etc. The types of

coarse-grained beads / particles in this forcefield are of four standard types: non-polar (N), polar (P), apolar (C), and charged (Q). In addition, there are sub-particles (0=none, d=donor, a=acceptor, da=donor/acceptor) assigned to the N and Q particles to mimic hydrogen-bonding properties. On the other hand, the C and P particle types are sub-categorised into 5 types depending on their polarity where '1' denotes low polarity and '5' denotes high polarity. In Martini 2.2, sidechain particles that denote aromatic rings are prefixed with 'S', and those that denote apolar sidechains are prefixed with 'A' (see Table 2.1). Further, leucine, isoleucine and valine residues each comprise of an AC sidechain particle type which is non-polar, similar to the C type. However, the Martini 2.2 forcefield states that these AC type particles are required for Q-C interactions within the protein.

***Table 2.1 Coarse-grained representations of amino acids in the Martini 2.2 forcefield.*** *The red and blue text represent negative and positive charges on the sidechain respectively. The alternative protonation states as shown here were not used in simulations in this thesis.*

| Amino acid | | Representative CG particles | |
|---|---|---|---|
| **1-letter** | **3-letter** | **Backbone** | **Sidechain** |
| A | ALA | P4 | - |
| G | GLY | P5 | - |
| P | PRO | P5 | C3 |
| C | CYS | P5 | C5 |
| M | MET | P5 | C5 |
| L | LEU | P5 | AC1 |
| I | ILE | P5 | AC1 |
| V | VAL | P5 | AC2 |
| T | THR | P5 | P1 |
| S | SER | P5 | P1 |
| N | ASN | P5 | P5 |
| Q | GLN | P5 | P4 |
| D | ASP | P5 | Qa |
| E | GLU | P5 | Qa |
| K | LYS | P5 | C3-Qd |
| R | ARG | P5 | N0-Qd |
| H | HIS | P5 | SC4-SP1-SP1 |
| F | PHE | P5 | SC5-SC5-SC5 |
| Y | TYR | P5 | SC4-SC4-SP1 |
| W | TRP | P5 | SC4-SNd-SC5-SC5 |
| | **Alternative protonation state** | | |
| D | ASP0 | P5 | P3 |
| E | GLU0 | P5 | P1 |
| K | LYS0 | P5 | C3-P1 |
| R | ARG0 | P5 | N0-P4 |
| H | HISH | P5 | SC4-SP1-SQd |

In chapter 6 of this thesis, available parameters for myristoylation and palmitoylation post-translational modifications for Martini 2.2 [232] were incorporated while studying LCK. In addition, as part of chapter 7, I have incorporated available parameters for N-acetylglucosamine-β(1,)-asparagine glycosylation for Martini 2.2 [231] to propose a coarse-grained model of the T cell receptor for future studies. The coarse-grained representations for these post-translational modifications are shown in Table 2.2.

***Table 2.2 Coarse-grained representations of post-translational modifications.***

| Post-translational modifications | | |
|---|---|---|
| **Name** | **Backbone** | **Sidechain** |
| N-terminal GLY Myristoyl | Nda | C1-C1-C1-C1 |
| CYS Palmitoyl | P5 | C5-C1-C1-C1-C1 |
| CYS Farnesyl | P5 | C5-C3-C3-C3-C3 |
| CYS Geranylgeranyl | P5 | C5-C3-C3-C3-C3-C3 |
| N-acetylglucosamine-β(1,)-ASN | P5 | P5-SP2-P1-P3-SP1 |

The result of coarse-graining or reducing the resolution of the biomolecular system using the Martini forcefield leads to the absence of explicitly defined non-heavy atoms such as hydrogen atoms. Given the absence of hydrogen atoms and the crucial role that hydrogen bonding plays in secondary and tertiary structure formation, it is important to mimic this in CGMD simulations. Therefore, the Martini forcefield employs an elastic network (EN) [233] to apply harmonic restraints between backbone particles occurring within a certain distance pre-defined by the user. This gives rise to a three-dimensional network of bonds that help maintain the initial secondary and tertiary structure while allowing for some backbone dynamics. The force applied to maintain the EN can also be pre-defined. In this thesis, CGMD simulations apply the EN model to backbone particles up to a 0.7 nm distance cut-off using a force constant of 1000 kJ/mol/nm. For example, in chapter 5, to mimic the strong association within each dimer and allow for the dynamics between these dimers in the octameric structure of the TCR-CD3 complex, I coarse-grained each dimer separately and merged their coordinates. As a result, the EN was applied to backbone particles (≤ 0.7 nm) only within the dimers, while inter-dimeric subunits were unrestrained, as illustrated in Fig 2.5A. I also illustrate the Martini representations of a few amino acids (Fig 2.5B), polarisable water (P4 particle type) (Fig 2.5C) that can ionically interact with charged (Q) particles, and that of a POPC lipid (Fig 2.5D) as used in this thesis.

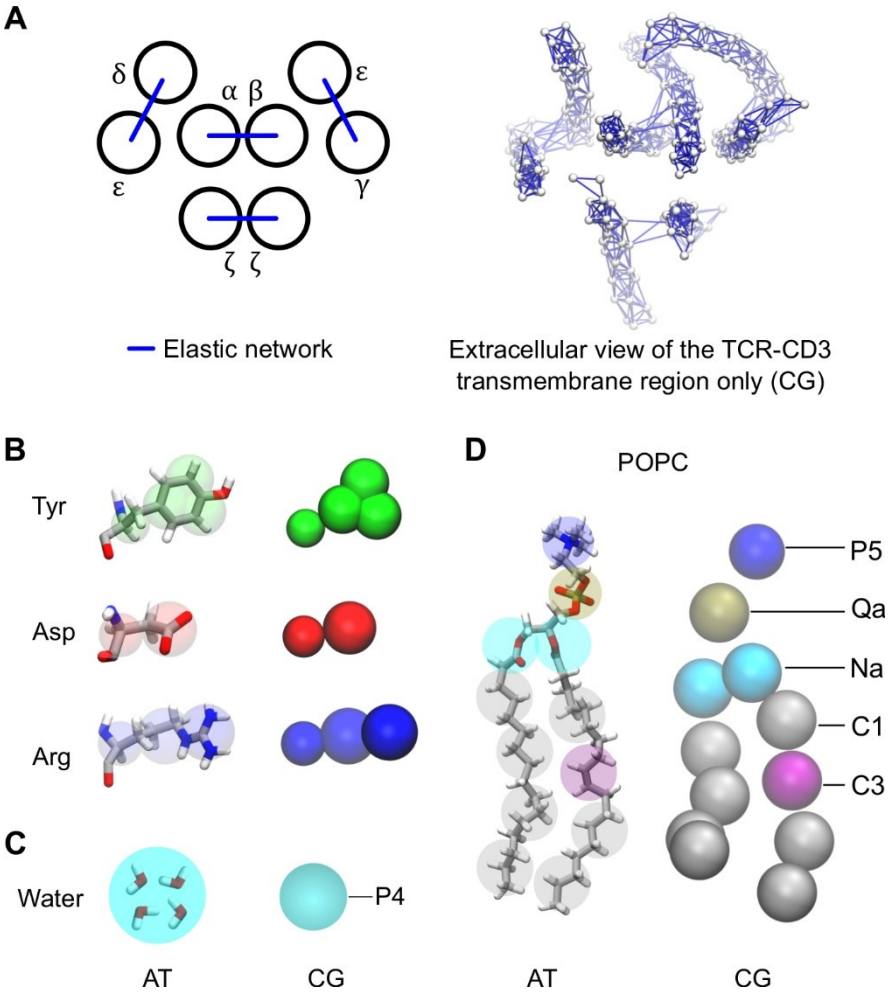


***Fig 2.5 Elastic network and coarse-grained Martini representations.*** *(A) A simplified view (left) and a snapshot of the coarse-grained TCR-CD3 transmembrane backbone connected by elastic bonds. (B) Atomistic and coarse-grained representations of some amino acids, (C) water, and (D) POPC lipid. The coarse-grained particle types of water and POPC lipid are labelled. Particle types of amino acids are shown in Table 2.1.*

## 2.4 Multi-scale molecular dynamics simulations

Molecular dynamics simulations of biomolecules in lower resolutions aid in achieving larger time-scales in comparison to those conducted in higher resolutions. However, sometimes, it may be important to achieve large time-scales while simultaneously studying a part of the system in higher resolution. This requires adopting a multi-scale resolution during simulations which can be a valuable technique. One such technique is the quantum mechanics/ molecular mechanics (QM/MM) simulation approach where the biomolecular system is considered as a hybrid model following the laws of quantum mechanics and molecular mechanics combined. Here, a part of the system is typically defined as the quantum mechanical region and linked to the rest of the system that follows molecular mechanics. In comparison with the molecular mechanical

region whose covalent bond interactions are pre-defined and conserved throughout the simulation, the quantum mechanical region can potentially transfer these bonded interactions from one atom to another. Additionally, in this approach, since both parts of the system are exhibit AT resolution, they can be simulated together. However, this is challenging to do if the different parts of the system differ in resolution such as CG and AT. For this reason, a method was developed to couple these resolutions using 'virtual sites' where the interactions between the CG and AT particles are treated similar to CG-CG interactions [204].

While multi-scale simulations can be conducted in a parallel manner as described above, a 'serial' approach can also be undertaken where a biomolecular system is completely simulated in a single resolution and then converted to another resolution [234]. This approach has been extensively employed in studying membrane proteins [235,236] and peripheral proteins [237–239] where CGMD simulations are conducted followed by ATMD simulations. Therefore, in chapter 3 and 5 of this thesis, I used this serial multi-scale approach to study the transmembrane assembly, and the complete structure of the TCR-CD3 complex (Fig 2.6) respectively. Multi-scale simulations of membrane-protein systems in this thesis were initially set up using the *Insane* tool [240], simulated in CG resolution Martini 2.2 forcefield [200], and later converted to AT resolution (CHARMM36 forcefield [214]) using the 'backmapping' technique [205]. Employing this technique in this thesis, I used final snapshots from CGMD simulations to add back the AT particles into the corresponding CG beads followed by step-wise relaxation of the resultant AT model and a short MD simulation with position restraints. This allowed me to study membrane-protein dynamics in AT detail after having simulated it on the microsecond time-scales in CG resolution.

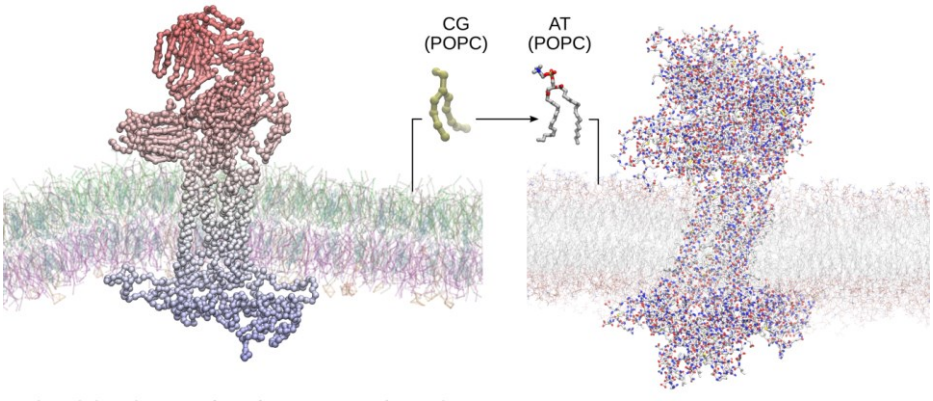


***Fig 2.6 Backmapping the T cell receptor and the membrane*** *from coarse-grained to atomistic resolution.*

# Chapter 3

# The T cell receptor transmembrane region

*Atomistic molecular dynamics simulation studies conducted on the complete (octameric) assembly of the T cell receptor transmembrane region in this thesis chapter is a derivative of my contribution to the following published paper referenced here as* [76]*:*

*Lanz et al. Allosteric activation of T cell antigen receptor signaling by quaternary structure relaxation. Cell Reports. 2021. doi: 10.1016/j.celrep.2021.109375*

## 3.1 Introduction

In this chapter, the primary focus of study is on the transmembrane region (TMR) of the TCR-CD3 complex. The key factor that drives the assembly of the TCR-CD3 complex is the electrostatic interactions between ionic residues located within its TMR [8]. This has also been observed as a common phenomenon in other immunoreceptor signalling complexes [58]. In addition, polar residues had earlier been found to drive the association of transmembrane helices that largely contained hydrophobic residues [241].

The stoichiometry of the TCR-CD3 complex is a 1:1:1:1 ratio of the TCRαβ, CD3δε, CD3γε, ζζ dimeric components [8,36]. The complete octameric assembly of the TCR-CD3 complex is suggested to be preceded by the assembly of the TCRαβ-CD3δε-CD3γε hexamer [8]. The electrostatic interactions that lead to the hexameric assembly is mediated by acidic residues situated in the middle of the CD3 TMRs (glutamic acid of CD3γ; aspartic acid of CD3δ and both CD3ε), and the lysine residues of TCRα [242] and TCRβ TMRs [243]. The hexameric assembly of TCRαβ-CD3δε-CD3γε driven by its TMR is further suggested to efficiently take place in a step-wise manner where the CD3δε heterodimer binds to TCRα followed by the CD3γε heterodimer binding to TCRβ [8]. Finally, the ζζ homodimer, via its aspartic acid residues, binds to the arginine residue of the TCRα TMR to form the complete TCR-CD3 complex (as shown in Fig 1.6 in chapter 1). Experimental data showed that mutating any one of these nine ionic residues has a disruptive effect on the assembly of the complex [8].

Although it is the electrostatic interactions in the TMR that drives this association, it cannot alone define the precise location of each CD3 component. This specificity in their location relative to the rest of the complex is driven by the protein-protein interfaces in the extracellular domains (ECDs). This

specificity in the ECD interfaces aids in TCRα-CD3δε and TCRβ-CD3γε oligomerisation as shown by experiments using chimeric CD3γ proteins [61]. The experiments first involved the TMR of CD3γ being replaced by the TMR of CD3δ. In this case, since the CD3γ extracellular domain was maintained and the CD3δ TMR also contains an anionic residue, the chimeric CD3γ was able to bind to the TCRβ chain. On the other hand, when the ECD of CD3γ was substituted with that of CD3δ, it could not bind to the TCRβ chain. Therefore, given that (i) mutations in the TMR disrupt TCR-CD3 assembly [8], and (ii) mutations in the ECD disrupt TCR-CD3 stability and signalling [244], we can infer that the TMR-mediated TCR-CD3 assembly is the primary step. Thus, with the goal of eventually building a model of the complete TCR-CD3 complex and simulating its dynamics, I aimed to initially study the dynamic aspects of the oligomerisation process of the TCR-CD3 TMR.

Following experimental studies conducted to resolve the protein-protein interfaces within the ζζ [63] and αβ TMR dimers [49], the structural basis of interactions within the CD3δε and CD3γε TMRs remained uncertain. In addition, apart from the ionic interactions that drive TCR-CD3 TMR assembly [8], we lack knowledge of the dynamics and hydrophobic interactions that contribute to its assembly. Therefore, this thesis chapter addresses this question and provides insights into the TCR-CD3 TMR oligomerisation process in a dynamic membrane environment on the microsecond time-scale. Note that this work was conducted to build a model of the complete TCR-CD3 TMR prior to the publication of its cryo-EM structure resolved at 3.7 Å resolution [16] which revealed the quaternary arrangement of its complete TMR and ECD (PDB:6JXR). Upon the release of the cryo-EM structure, I extracted the coordinates of the TCR-CD3 TMR, induced *in silico* mutations, and performed simulations to complement experiments conducted by Prof. Oreste Acuto's group at the University of Oxford to study an allosteric activation mechanism of the TCR-CD3 [76].

### 3.1.1 Key research aims

- Obtain a model of the structure and organisation of the transmembrane subunits of the TCR-CD3 complex.
- Study the protein-protein interactions of the TCR-CD3 transmembrane subunits and their dynamics in a membrane environment, further enabling investigation of the effect of mutations in the TMR.

## 3.2 Methods

### Coarse-grained molecular dynamics simulations

In this chapter, only the studies conducted on the dimerisation and trimerisation of TCR-CD3 TMR helices involved CGMD simulations. The transmembrane helices were modelled in atomistic resolution using Chimera [245] and then coarse-grained using the *martinize* script and the Martini 2.2 forcefield. An elastic network (EN) model was applied within a 0.7 nm distance cut-off using a force constant of 1000 kJ/mol/nm. Following coarse-graining, the helical peptides were inserted into a POPC membrane such they were ~6 nm apart using the *Insane* tool [240]. The system was neutralised by Martini water particles and $Na^+$ and $Cl^-$ ions to a 0.15M concentration. The entire system was then energy minimised using the steepest descent algorithm and a short MD simulation (5 ns) with the helices restrained to equilibrate the system. Production simulations were run without any restraints for 2 µs in 20 replicates and coordinates were saved every 200 ps. Other details are mentioned in chapter 2.

### Atomistic molecular dynamics simulations

The CGMD simulations conducted on the dimerisation and trimerisation of TCR-CD3 TMR helices were converted to atomistic resolution using the 'backmapping' protocol [205]. On the other hand, simulations conducted to study the entire TCR-CD3 TMR were setup using the CHARMM-GUI [226] platform. The TIP3P water model along with atomistic $Na^+$ and $Cl^-$ ions were re-added to neutralise the system and to maintain a 0.15M concentration. The systems were energy minimised using the steepest descent algorithm and equilibrated with the protein backbone restrained. The backmapped systems were equilibrated for 2 ns while the systems setup with CHARMM-GUI underwent a 6-step equilibration with the position restraints gradually released. The production simulations of backmapped systems were further conducted for 100 ns in 3 replicates, whereas those set up using CHARMM-GUI were simulated for 1250 ns in 3 replicates. Coordinates of all systems were saved at 40 ps intervals. Other details are mentioned in chapter 2.

## 3.3 Results

### 3.3.1 Predicting the TCR-CD3 transmembrane region

To enable modelling and MD simulations of the CD3δε/γε TMR dimers and their oligomerisation with TCRα/β TMRs respectively, each of their sequences and secondary structures were obtained from Uniprot (https://www.uniprot.org/) and also predicted using other online servers such as MEMSAT [246] (part of the PSIPRED 4.0 server [247]), TMHMM 2.0 (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0), and SPLIT 4.0 [248]. The amino acid sequences that were predicted to constitute the TMR are shown in Table 3.1.

All TMRs were predicted to form α-helices. To validate the accuracy of these predictions and given that the ζζ TMR is resolved by NMR in detergent micelles [63], the ζζ TMR was also predicted and compared to the TMR sequence suggested by the NMR experiments. As a result, they were observed to be quite consistent. Moreover, ATMD simulations of the NMR structure of the ζζ dimer have shown that it maintains its protein-protein interactions and tertiary structure in a POPC and in a POPC/SM/Chol membrane environment [249].

***Table 3.1 Prediction of the transmembrane region*** of the CD3δ, -γ, -ε and TCRα, -β subunits. The residues commonly predicted by all servers to lie in the transmembrane region are shown in red. The ionic residues that are key for the assembly of the TCR-CD3 are indicated by ' **^** '.

| Subunits (Uniprot entries) | Prediction server | Sequences |
|---|---|---|
| TCRα (Uniprot: P0DSE1) | Uniprot | VIGFRILLLKVAGFNLLMTLRLW |
| | MEMSAT3 | LSVIGFRILLLKVAGFNLLMTLRL |
| | TMHMM2.0 | VIGFRILLLKVAGFNLLMTLRLW |
| | SPLIT4.0 | LSVIGFRILLLKVAGFNLLMTL |
| | | ^ |
| TCRβ (Uniprot: P0DSE2) | Uniprot | TILYEILLGKATLYAVLVSALVL |
| | MEMSAT3 | ATILYEILLGKATLYAVLVSALVLM |
| | TMHMM2.0 | TILYEILLGKATLYAVLVSALVL |
| | SPLIT4.0 | SATILYEILLGKATLYAVLVSALVLM |
| | | ^ |
| CD3δ (Uniprot: P04234) | Uniprot | GIIVTDVIATLLLALGVFCFA |
| | MEMSAT3 | PATVAGIIVTDVIATLLLALGVFCFA |
| | TMHMM2.0 | AGIIVTDVIATLLLALGVFCF |
| | SPLIT4.0 | PATVAGIIVTDVIATLLLALGVFCF |
| | | ^ |
| CD3ε (Uniprot: P07766) | Uniprot | VMSVATIVIVDICITGGLLLLVYYWS |
| | MEMSAT3 | EMDVMSVATIVIVDICITGGLLLLVYYW |
| | TMHMM2.0 | VATIVIVDICITGGLLLLVYYWS |
| | SPLIT4.0 | CMEMDVMSVATIVIVDICITGGLLLLVY |
| | | ^ |
| CD3γ (Uniprot: P09693) | Uniprot | GFLFAEIVSIFVLAVGVYFIA |
| | MEMSAT3 | AATISGFLFAEIVSIFVLAVGVYFIA |
| | TMHMM2.0 | ISGFLFAEIVSIFVLAVGVYFIA |
| | SPLIT4.0 | AATISGFLFAEIVSIFVLAVGVYFI |
| | | ^ |
| ζ (Uniprot: P20963) | Uniprot | LCYLLDGILFIYGVILTALFL |
| | MEMSAT3 | LCYLLDGILFIYGVILTALFL |
| | TMHMM2.0 | LCYLLDGILFIYGVILTALFLRV |
| | SPLIT4.0 | CYLLDGILFIYGVILTALFLRVK |
| | Call *et al*, 2006 [63] | LCYLLDGILFIYGVILTALFLRV |
| | | ^ |

## 3.3.2 Dimerisation of CD3δε and CD3γε

### 3.3.2.1 Coarse-grained molecular dynamics simulations

From the sequence predictions, there were common residues that were predicted by all four online platforms to lie in the transmembrane region, as shown in red in Table 3.1. So, I first modelled only these residues and then modelled all the residues combined (sequences in black and red combined in

Table 3.1). As expected, the former modelling process gave rise to shorter helices compared to the latter. Therefore, I referred to them as 'short TM' helices and 'long TM' helices. Further, I coarse-grained both the 'short TM' and 'long TM' helix models of CD3δ, -γ, -ε using the Martini 2.2 forcefield. Further, the 'short TM' and 'long TM' models of CD3δε and CD3γε were each inserted into POPC bilayers and neutralised with Martini solvent particles using the *Insane* tool [240] (see Table 3.2).

***Table 3.2 CGMD simulation details to study CD3δε and CD3γε dimerisation***

| Protein | Lipids + solvent | PBC box size X x Y x Z (nm) | Time x replicates |
|---|---|---|---|
| CD3δε short TM | POPC (100%), 0.15M $Na^+$ & $Cl^-$, Martini water | 10.5 x 10.5 x 10 | 2 μs x 20 |
| CD3δε long TM | | 10.5 x 10.5 x 10 | 2 μs x 20 |
| CD3γε short TM | | 10.5 x 10.5 x 10 | 2 μs x 20 |
| CD3γε long TM | | 10.5 x 10.5 x 10 | 2 μs x 20 |

In each of these simulations, single helical subunits were initially placed approximately 6 nm away from each other in the membrane (Fig 3.1A) to allow them to explore their environment and different tilt angles before binding to the other subunit. The CGMD simulations were conducted for 2 μs and in 20 replicates to ensure sufficient sampling. At the end of the simulations, protein-protein contacts were analysed for both the 'long TM' and 'short TM' models for each of the CD3δε and CD3γε dimers (Fig 3.1B). However, in some of the 'short TM' simulations, I observed that the helical length did not span the membrane and therefore discontinued using them for further studies. Nonetheless, their protein-protein contact profiles were in good agreement with the 'long TM' simulations for both CD3 dimers (Fig 3.1B).

The contacts between subunits in each dimer from my 'long TM' simulations were also compared with those observed in the cryo-EM structure (using < 4 Å cut-off after adding hydrogen atoms) and were found to be consistent. However, unlike my simulations, the presence of ECDs in the cryo-EM structure seems to allow the TMR helices within CD3δε/γε dimers to interact only toward their C-terminus (the cytoplasmic side). Therefore, the similarity in protein-protein interfaces in simulations and cryo-EM structure can be largely observed toward the C-terminal half of the TMR (highlighted one-letter amino acid codes in Fig

3.1B). Snapshots of the protein-protein interfaces for both CD3 dimers from their 'long TM' simulations are shown in Fig 3.1C.

The analysis of distances between the centre of masses of the helical subunits versus time shows that dimerisation occurred in almost all simulations within the 2 μs time frame (Fig 3.1D). The simulations also suggested that both CD3 TMR dimers are likely to exhibit similar crossing angles (Fig 3.1E) which was later observed to be consistent with the cryo-EM structure (PDB:6JXR). Additionally, there were a greater number of residues forming the protein-protein interface in CD3γε TMR compared to that in CD3δε. This resulted in the CD3γε TMR often exhibiting a particular crossing angle of ~30° compared to that of CD3δε which indicated a wider distribution of crossing angles.

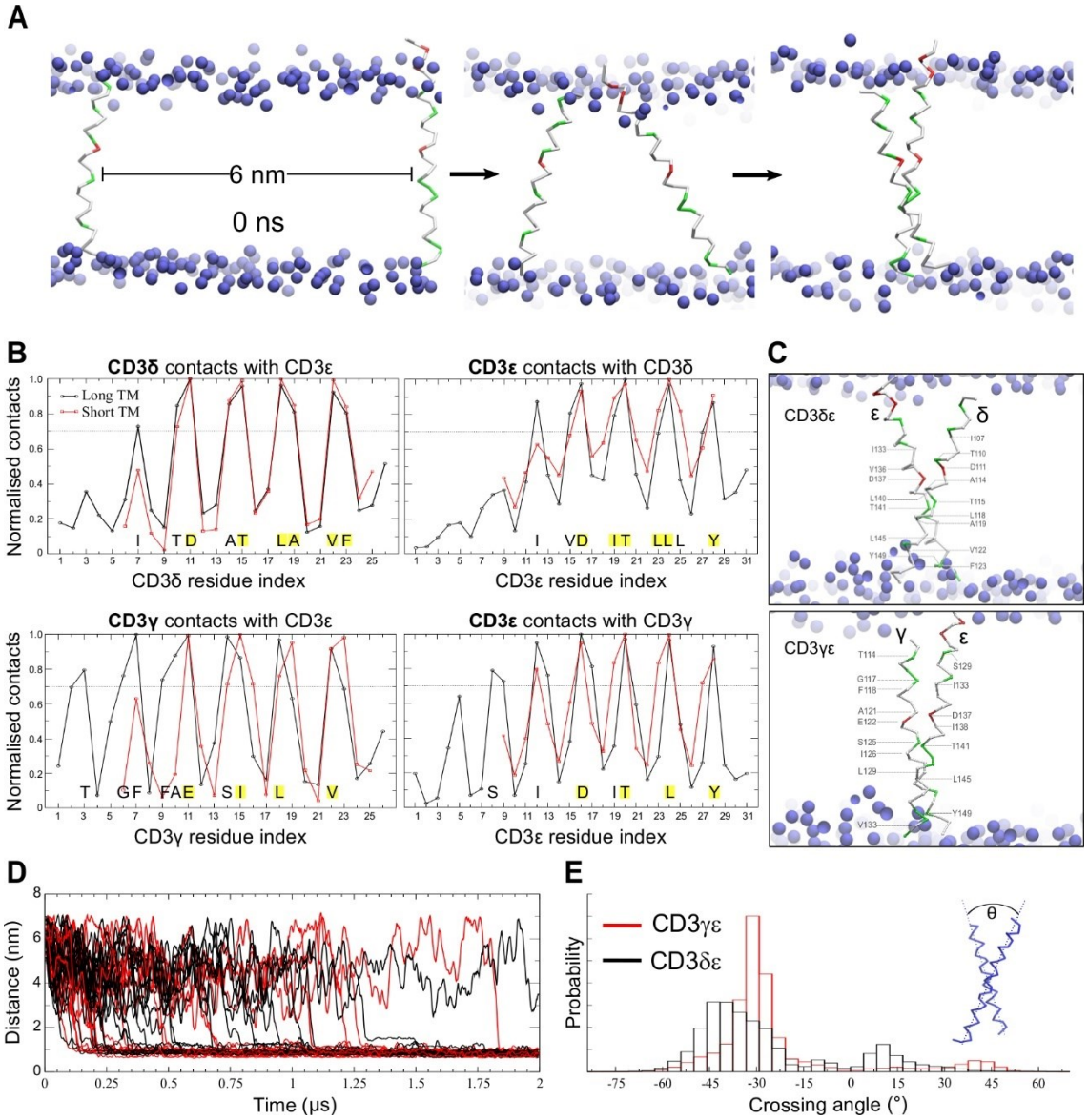

***Fig 3.1 CGMD simulations CD3δε and CD3γε TMR.*** *(A) Snapshots taken from one of the CGMD simulations show a time-lapse of transmembrane helices binding when initially separated by 6 nm distance in the membrane. (B) Normalised contacts between CD3 subunits after simulating their 'long TM' (black) and 'short TM' (red) helix models. Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). Residues exhibiting significant interactions (≥0.7 on normalised scale) in the 'long TM' simulations are labelled. Interactions consistent with those in the cryo-EM structure (PDB:6JXR) are highlighted. (C) Snapshots from a 'long TM' simulation indicating significant interactions from the 'long TM' simulations. Lipid phosphate groups are shown as purple spheres representing the membrane surface. Labelling of residue numbers are according to Uniprot entries (given in Table 3.1) and PDB (https://www.rcsb.org/sequence/6jxr). (D) Distance versus time, and (E) crossing angle distributions of CD3δε (black) and CD3γε (red) calculated from their 'long TM' simulations.*

#### 3.3.2.2 Atomistic molecular dynamics simulations

The resolution of CGMD simulations is lower than that of ATMD simulations and therefore does not provide detailed information on the interactions of atoms. For this reason, I extracted one of the final snapshots of the CD3δε and CD3γε 'long TM' dimers along with the membrane from CGMD simulations and 'backmapped' [205] them to atomistic resolution. The atomistic system was then neutralised with 0.15M $Na^+$ $Cl^-$ ions and the SPC water model, and further used to conduct ATMD simulations for 100 ns in 3 replicates each. This provided atomistic-level data on the protein-protein interfaces of both CD3 dimers.

The backmapped ATMD simulations suggested that the protein-protein contacts of the CD3δε dimer occurred mostly toward the middle of the helices (Fig 3.2A). As a result, the CD3δε dimer exhibited a scissor-like movement and thereby showing a wider distribution of crossing angles compared to CD3γε (Fig 3.2B). This observation made in ATMD simulations was consistent with CGMD simulations. In addition, RMSD analyses of the backbones of these dimers also showed that CD3δε was less stable than CD3γε (Fig 3.2C). Further, I wanted to compare this data with CD3δε/γε interfaces and crossing angles when they are oligomerised with TCRα/β TMRs respectively.

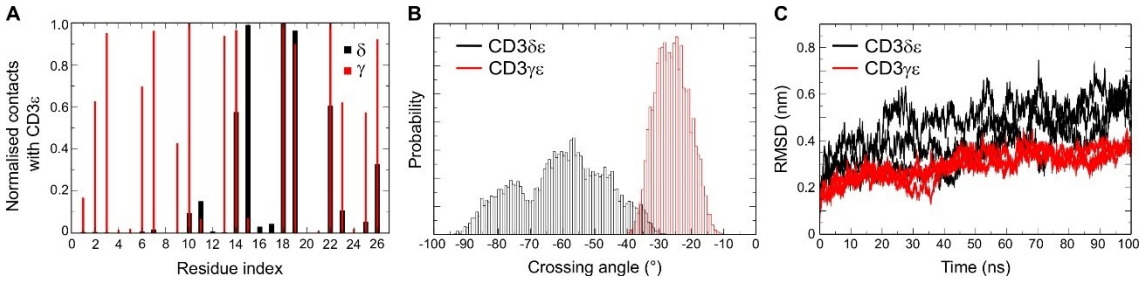


***Fig 3.2 ATMD simulations of CD3δε and CD3γε TMR.*** *(A) Normalised contacts of CD3δ and of CD3γ with CD3ε. Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). (B) Crossing angle distributions of CD3δε and CD3γε TMR. Negative values on X axis indicate right-handed crossing angles. (C) RMSD of CD3δε and CD3γε TMR backbones across time. The three lines for each CD3 dimer represent simulation replicates.*

## 3.3.3 Trimerisation of TCRα-CD3δε and TCRβ-CD3γε

### 3.3.3.1 Coarse-grained molecular dynamics simulations

Following the analyses of CD3δε/γε dimers, the TCRα-CD3δε and TCRβ-CD3γε trimeric subunits were simulated in a POPC bilayer for 5 μs in 20 replicates: (i) TCRα in the presence of CD3δ and CD3ε monomers, (ii) TCRβ in the presence of CD3γ and CD3ε monomers. In each of these simulations, the TCR/CD3 monomers were initially positioned 6 nm away from each other.

Simulations revealed that the protein-protein interfaces between the TCRα-CD3δε and TCRβ-CD3γε trimers formed significant contacts in the middle of the helices where their ionic residues are situated, consistent with experimental evidence [8]. More importantly here, combining all the 20 simulation replicates, I highlight other significant interaction sites (≥ 0.7 on normalised scale) on each of the TMR subunits (Fig 3.3A). In addition to the interaction of the ionic residues at the centre of the TMR helices, it can be observed that hydrophobic interactions are also major contributors to the oligomerisation of subunits. When comparing the protein-protein contacts from my simulations to those observed in the cryo-EM structure (using < 4 Å cut-off after adding hydrogen atoms), I found that most interface-forming residues suggested by simulations were consistent with experimental evidence [16] (highlighted one-letter amino acid codes in Fig 3.3A). Note that, in some of both TCRα-CD3δε and TCRβ-CD3γε simulations (4/20 replicates each), trimer formation did not take place.

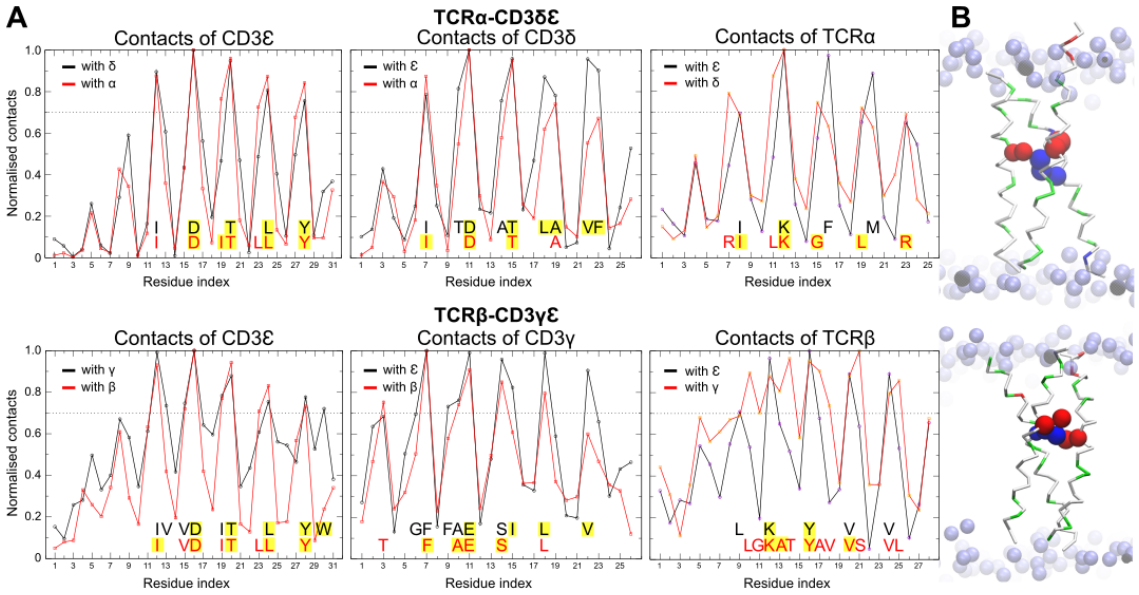


***Fig 3.3 CGMD simulations of TCRα-CD3δε and TCRβ-CD3γε TMR.*** *(A) Normalised contacts of the TCR/CD3 TMR subunits with each other. Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). Significant interactions (≥0.7 on normalised scale) that are consistent with those in the cryo-EM structure (PDB:6JXR) are highlighted. (B) Snapshots from one of my CGMD simulations of TCRα-CD3δε (above) and TCRβ-CD3γε (below) TMR trimerisation. Lipid phosphate headgroups: purple spheres, TMR helices: bonds, basic and acidic residues responsible for oligomerisation: blue and red spheres respectively.*

### 3.3.3.2 Atomistic molecular dynamics simulations

The same approach taken in studying the CD3 dimers was employed here i.e. one of the final snapshots obtained from CGMD simulations was considered for 'backmapping' to atomistic resolution. Following the backmapping process, the atomistic system containing the protein and the membrane was neutralised with solvent and prepared for ATMD simulations over 100 ns in 3 replicates.

Consistent with the contacts observed in the CD3δε dimer simulations (Fig 3.2A), I observed that the CD3δ and CD3ε subunits interact lesser at their N-terminal end (Fig 3.4A). However, in these simulations conducted in presence of the TCRα/β subunits, the same was observed in the CD3γε dimer i.e. CD3γ and CD3ε also interact less at their N-terminal end. These observations from my simulations were later found to be consistent with the TCR-CD3 cryo-EM structure (PDB:6JXR) [16]. Previously, in the simulations of the CD3 dimers, it was also observed that CD3γε exhibited a stable crossing angle value of 30° while CD3δε explored a wider variety of crossing angles (Fig 3.1E). Here, in the simulations of the trimers, I observed that the crossing angle of CD3δε was stabilised by its contacts with TCRα. As a result, both CD3 dimers indicated a

high probability of exhibiting similar crossing angles of approximately 20° (Fig 3.4B). The similar crossing angles of these CD3 dimers were also observed in the TMR of the TCR-CD3 cryo-EM structure.

Another feature observed from my ATMD simulations was a slight deformation in the membrane where POPC lipid headgroups bent inward toward the hydrophobic core of the membrane in order to interact with the ionic sidechains in the middle of the TMR subunits. As a result, I observed entry of water molecules into the hydrophobic core of the membrane. However, the bilayer remained intact. Interestingly, this occurred largely only in the outer leaflet of the membrane (Fig 3.4C). This phenomenon was also observed in the simulations of the CD3 dimers, and was further studied after modelling and simulating the complete TCR-CD3 complex in chapter 5.

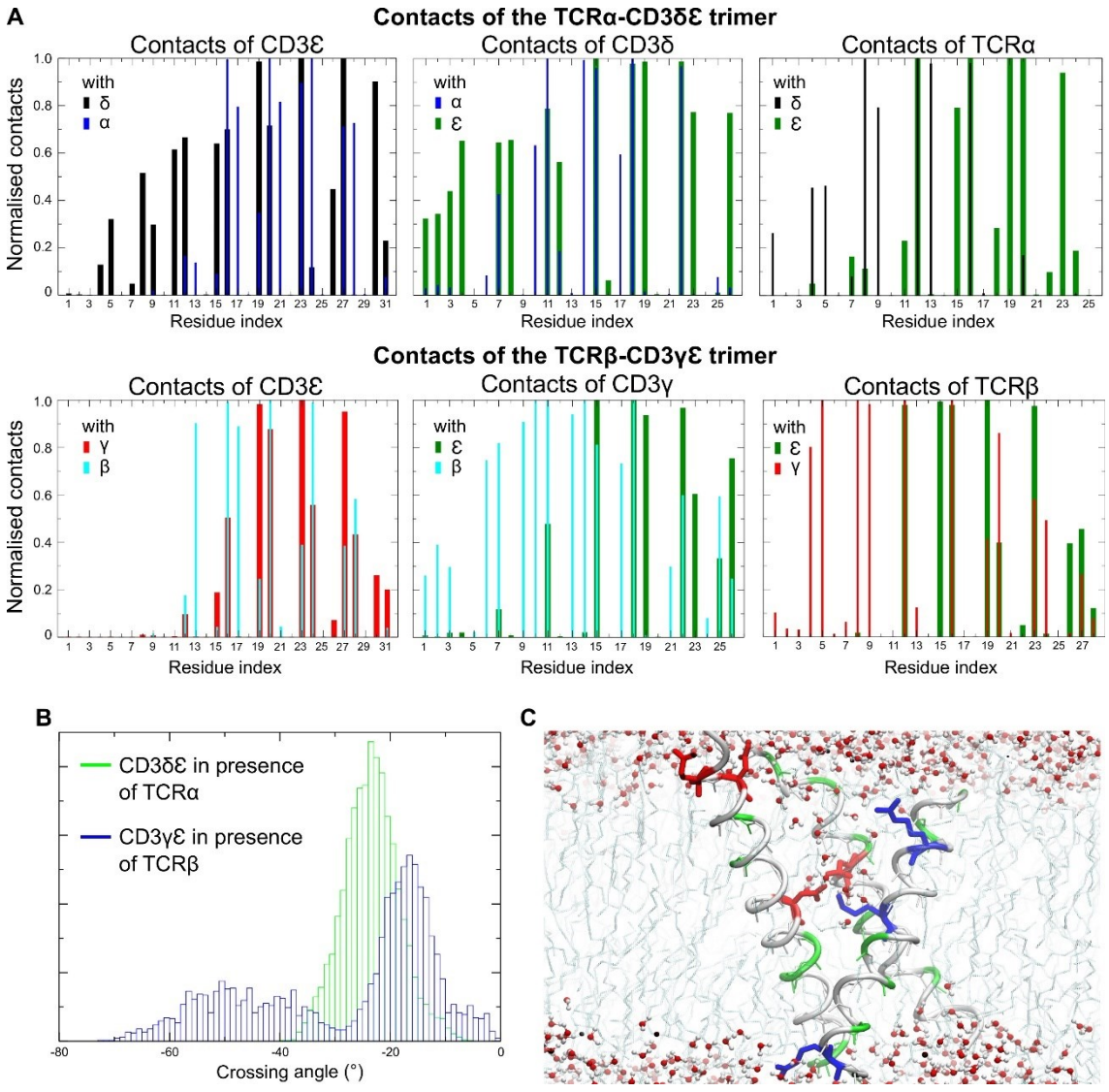

***Fig 3.4 ATMD simulations of the TCRα-CD3δε and TCRβ-CD3γε TMR.*** *(A) Normalised contacts of the TCR/CD3 TMR subunits with each other. Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). (B) Crossing angle distributions of CD3δε and CD3γε TMR in presence of TCRα/β TMRs respectively. Negative values on X axis indicate right-handed crossing angles. (C) Snapshot of TCRα-CD3δε trimerisation, and the entry of water molecules into the membrane.*

### 3.3.4 The complete TCR-CD3 transmembrane region

The publication of the cryo-EM structure of the TCR-CD3 complex [16] confirmed the organisation of its TMR as suggested by earlier studies [8]. At the time, I began studying the complete TCR-CD3 TMR in collaboration with Prof. Oreste Acuto and his group at the University of Oxford. Our collaborators identified residues in the TCRβ TMR whose mutations gave rise to the same structural changes which the TCR-CD3 undergoes upon activation. Among other residues, our collaborators were interested to obtain dynamic insights into the Y291A mutation in the TCRβ TMR (βY291A). Note that, in this chapter, I refer to the residue number as βY291 (as mentioned in our published work [76]) in contrast to βY292 in the cryo-EM structure [16]. The difference in residue numbers is due to the usage of a different isoform of the β subunit during experiments [76]. Nonetheless, the TMR sequence of the β subunit remained identical between their experiments and my simulations. In their experiments, n-dodecyl-β-D-maltopyranoside (DDM) detergent was used to extract the intact TCR-CD3 complex. Their experimental setup of probing TCR-CD3 structural stability or 'cohesion' using DDM was referred to as the DDM stability assay (DSA). This was followed by quantitative immunoblotting of the β and the CD3 δ, γ, ε subunits which helped evaluate the effect of βY291A mutation on TCR-CD3 cohesion. Further, they also explored mutations of I38A and I41A in the ζζ TMR that potentially led to loosening of the ζζ dimer from the rest of the complex thereby promoting signalling.

In parallel, to study the effects of the βY291A, ζζI38A and ζζI41A mutation on the rest of the TCR-CD3 TMR in atomistic detail, I extracted the transmembrane helices from the cryo-EM structure (PDB:6JXR) and setup ATMD simulations using the CHARMM-GUI platform [226]. The sequences of each wild-type (WT) subunit used in my simulations are shown in Table 3.3. Simulations of the WT were performed as a control to the mutation simulations. All simulations here onward were conducted over a time-scale of 1250 ns in a complex asymmetric membrane environment. The composition of lipid headgroups used (Table 3.4)

was derived from a lipidomics study of the TCR-CD3 membrane activation domain [210].

***Table 3.3 Amino acid sequences used in ATMD simulations of the complete TCR-CD3 TMR.*** *The residues that were substituted by alanine in the mutation simulations in this chapter are shown in red and underlined. The residues shown in green are those that were used to obtain helical models for the previous dimer and trimer simulations (sections 3.3.2, 3.3.3 respectively).*

| Subunits | TMR sequences used |
|---|---|
| **$\zeta_1$** | LDPKLCYLLDGILFIYGVILTALFLRVKFSR |
| **$\zeta_2$** | LDPKLCYLLDGILFIYGVILTALFLRVKF |
| **TCRα** | DTNLNFQNLSVIGFRILLLKVAGFNLLMTLRLWSS |
| **TCRβ** | TSESYQQGVLSATILYEILLGKATLYAVLVSALVLMAMVKRK |
| **CD3δ** | ELDPATVAGIIVTDVIATLLLALGVFCFAGHE |
| **CD3ε** (of δε dimer) | MDVMSVATIVIVDICITGGLLLLVYYWSKNR |
| **CD3γ** | ELNAATISGFLFAEIVSIFVLAVGVYFIAG |
| **CD3ε** (of γε dimer) | MDVMSVATIVIVDICITGGLLLLVYYWSKNRK |

***Table 3.4 Membrane lipid composition used in ATMD simulations of the TCR-CD3 TMR***

| Lipid concentration (%) | POPC | POPS | POPE | DPSM | Chol | $PIP_2$ | $PIP_3$ |
|---|---|---|---|---|---|---|---|
| **Inner leaflet** | 10 | 20 | 40 | - | 20 | 8 | 2 |
| **Outer leaflet** | 50 | - | 10 | 20 | 20 | - | - |

#### 3.3.4.1 Y291A mutation in the TCRβ subunit

Experiments conducted by our collaborators suggested that the βY291A mutation in the TCR-CD3 showed gain of function upon ligand binding. This was observed via an increase in phosphorylation of the ζ subunit and further activation of Erk (serine/threonine protein kinase) downstream. Paradoxically, βY291A also led to reduced TCR-CD3 surface expression [76]. This led to the hypothesis that βY291A altered stability of the TCR-CD3 TMR which reduced its surface expression, but nevertheless increased signalling via an allosteric

mechanism linking the TMR to the cytoplasmic region (CYR). My ATMD simulations performed using the WT and βY291A mutant provided a dynamic insight into the effect of the mutation on the rest of the TCR-CD3 TMR. The complete TCR-CD3 complex was not modelled and used in these simulations given the high expense of performing simulations of the whole receptor in the microsecond (1250 ns) time-scale in atomistic resolution. Moreover, even if the complete TCR-CD3 complex was used, it would have been unlikely to observe an allosteric reaction linking the TMR to the CYR within this time frame.

From my ATMD simulations of the WT and βY291A, I first assessed the interactions of βY291 with all other residues in the TCR-CD3 TMR. Consistent with the cryo-EM structure, my simulations suggested that βY291 did not interact with ζζ since it is located in the C-terminal half of the TCRβ TMR that protrudes in the opposite direction (Fig 3.5A). I also compared the protein-protein interactions of βY291 in the WT simulations to those identified in the cryo-EM structure (< 4 Å cut-off after adding hydrogen atoms). This showed that my simulations agreed well with the cryo-EM structure. The most significant interactions of βY291 in simulations (considering > 70% contacts on average) were εL145, γL129, γG132, αN263, and αT267 (Fig 3.5B). There were no interactions of βY291 with CD3δε and ζζ. Upon βY291A mutation, it maintained contact with γL129 and αN263 while the rest of the contacts diminished (Fig 3.5C). In addition, I assessed the overall interactions of TCRβ TMR with CD3γε (Fig 3.5D) comparing the WT and the βY291A mutant. The mutant resulted in an increase in the overall interaction of the CD3γε TMR with TCRβ despite loss of interaction specifically at the β291 site. The increase in TCRβ-CD3γε contacts due to the removal of the bulky tyrosine sidechain thereby showed a small stabilisation effect on CD3γε. Analysis of the RMSD of Cα atoms of the CD3γε dimer supported this finding (Fig 3.5E). On the other hand, immunoblotting experiments showed that βY291A affected the interaction of both CD3γε and CD3δε dimers with the rest of the complex [76]. However, note that the cryo-EM structure, upon which my simulations depended on, suggests that βY291 mostly contacts CD3γε. Moreover, given the time-scale limitation of ATMD simulations, βY291A was expected to primarily affect TCRβ-CD3γε interactions. Nevertheless, my ATMD simulations provided partial evidence that βY291A induces local rearrangements in the TMR by revealing the very beginning of the effect of the mutation.

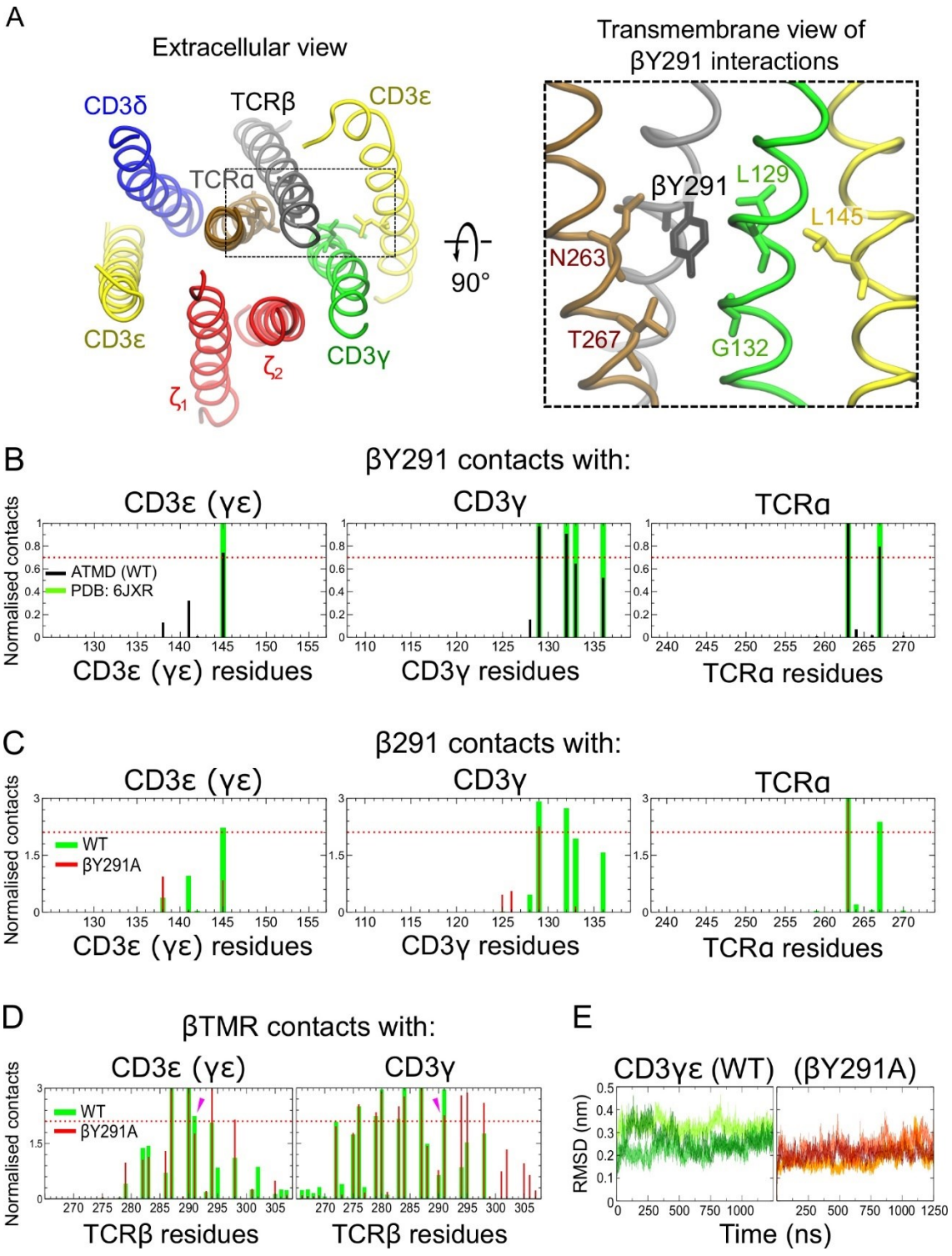
A
Extracellular view
CD3δ
TCRβ
CD3ε
TCRα
CD3ε
CD3γ
ζ2
ζ1
90°
Transmembrane view of
βY291 interactions
βY291
L129
L145
N263
T267
G132
B
βY291 contacts with:
CD3ε (γε)
CD3γ
TCRα
Normalised contacts
ATMD (WT)
PDB: 6JXR
CD3ε (γε) residues
CD3γ residues
TCRα residues
C
β291 contacts with:
CD3ε (γε)
CD3γ
TCRα
Normalised contacts
WT
βY291A
CD3ε (γε) residues
CD3γ residues
TCRα residues
D
βTMR contacts with:
CD3ε (γε)
CD3γ
Normalised contacts
WT
βY291A
TCRβ residues
TCRβ residues
E
CD3γε (WT)
(βY291A)
RMSD (nm)
Time (ns)

***Fig 3.5 Effect of βY291A on protein-protein interactions in the TCR-CD3 TMR.*** *(A) Snapshot of the TCR-CD3 TMR at the end of a 1250 ns ATMD simulation as viewed from the extracellular region. A transmembrane (side) view shows residues that significantly interact with βY291 on the right. (B) Normalised contacts of βY291 with the ε (of γε dimer), γ, and α subunit in simulations are compared with contacts seen in the cryo-EM structure (PDB:6JXR). Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). (C) Normalised contacts of βY291 with the ε (of γε dimer), γ, and α subunit are compared with contacts of βY291A mutant. (D) Normalised contacts of the complete β TMR with ε (of γε) and γ subunits when the WT and βY291A are simulated. In (C) and (D), normalisation was done by diving all contacts by the number of simulation frames. (E) RMSD versus time of the Cα atoms of the CD3γε dimer when the WT and βY291A are simulated.*

### 3.3.4.2 I38A and I41A mutations in the ζζ dimer

To correlate loosening of the TCR-CD3 TMR to the activation of the entire receptor, our collaborators were also interested in targeting the ζζ dimer and performing mutations in its TMR that could potentially lead to enhanced TCR-CD3 signalling. The ζζ dimer is the major signalling component in the TCR-CD3 complex since it alone contains 6 out of 10 ITAMs in the whole receptor [250], making it a potential target to control TCR-CD3-mediated signalling upon ligand binding.

Using the cryo-EM structure (PDB:6JXR), I first analysed the interactions of the $\zeta_1$ and $\zeta_2$ with their corresponding neighbouring subunits. It showed that the $\zeta_1$ TMR interacts with CD3ε (of δε) and $\zeta_2$ TMR interacts with CD3γ, while both ζ TMRs also interact with TCRα and TCRβ. ATMD simulations (3 x 1250 ns) of the WT that showed that ζζ-α and ζζ-β interactions were largely consistent with those identified in the cryo-EM structure (Fig 3.6A-D). ζζ-α contacts indicated that $\zeta_1$ mostly contacted α via the N-terminal half of its TMR, and $\zeta_2$ shared contacts with α throughout its TMR (Fig 3.6A, C). On the other hand, ζζ-β contacts indicated that $\zeta_1$ did not make significant contacts with β, while $\zeta_2$ contacted β via the N-terminal half of its TMR (Fig 3.6B, D).

Further, I identified the interactions between $\zeta_1$ and ε (δε) (Fig 3.6E), and between $\zeta_2$ and γ (Fig 3.6F). These contacts were also compared with those identified from the cryo-EM structure (Fig 3.6G, H). The comparison revealed that, both ζ subunits in the simulations had reduced interactions with ε (δε) and γ correspondingly toward the N-terminal edge, potentially due to the lack of secondary structure in the short extracellular peptide region of ζζ. However, most interactions toward the middle of the TMR were conserved. My

simulations also indicated a few contacts between $\zeta_2$ and γ that were not observed in the cryo-EM structure i.e. $\zeta_2$-V44, I45, A48, and γ-F127, V131. Further, it was observed from my simulations that $\zeta_1$I38 and $\zeta_2$I41 were some of the most interactive residues with the CD3ε (δε) and CD3γ subunits respectively (Fig 3.6G, H), specifically with εI138 (Fig 3.6I) and γV124 (Fig 3.6J). To check whether these site-specific interactions of ζζ with the CD3 TMRs could lead to loosening with the rest of the complex, I induced *in silico* mutations substituting I38 of both ζ subunits to alanine (ζζI38A) in one set of simulations (3 x 1250 ns) and I41 of both ζ subunits to alanine (ζζI41A) in another set of simulations (3 x 1250 ns).

ATMD simulations revealed that $\zeta_1$ at I38A lost contact with ε (δε). More importantly, the ζζI38A TMR mutant underwent loosening relative to αβ, although it was $\zeta_1$ that was more affected than $\zeta_2$. This is because $\zeta_1$I38 interacts with ε (δε) while $\zeta_2$I38 faces away from the protein complex. Alignment of snapshots of the mutated TMR complex to the WT at the end of the 1250 ns ATMD simulations showed the distortion in ζζ-δε contact (Fig 3.6K left). Additionally, the spatial distribution plot of the ζζ TMR Cα atoms relative to the αβ TMR Cα atoms (Fig 3.6K right) showed a broader density of the ζζI38A mutant demonstrating its flexibility compared to the WT.

ATMD simulations of the ζζI41A TMR mutant, compared to ζζI38A, underwent a minor fluctuation relative to αβ. This is shown by the final snapshots of the ζζI41A mutant aligned to the WT at the end of the 1250 ns simulations (Fig 3.6L left) and the average spatial distribution plot of the ζζ TMR Cα atoms relative to the αβ TMR Cα atoms (Fig 3.6L right). The spatial distribution plot showed a broader density of the ζζI41A mutant compared to the WT, but not as much as the ζζI38A mutant suggesting that ζζI41A had lesser impact compared to ζζI38A on the association of ζζ with the rest of the TCR-CD3 TMR. This was consistent with experimental results which showed that ζζI38A led to 95% reduction in ζζ surface expression compared to ζζI41A which led to 75% reduction [76]. Nonetheless, downstream phosphorylation of Erk suggested that both these mutations in the ζζ TMR increased signalling to a similar degree. Finally, with the aid of my ATMD simulations providing a dynamic insight into specific interactions in the TCR-CD3 TMR, experiments concluded that pMHC-induced TCR-CD3 complexes undergo an allosteric reaction from the extracellular region to specific sites in the TMR further leading to exposure of ITAMs in the intracellular region [76].

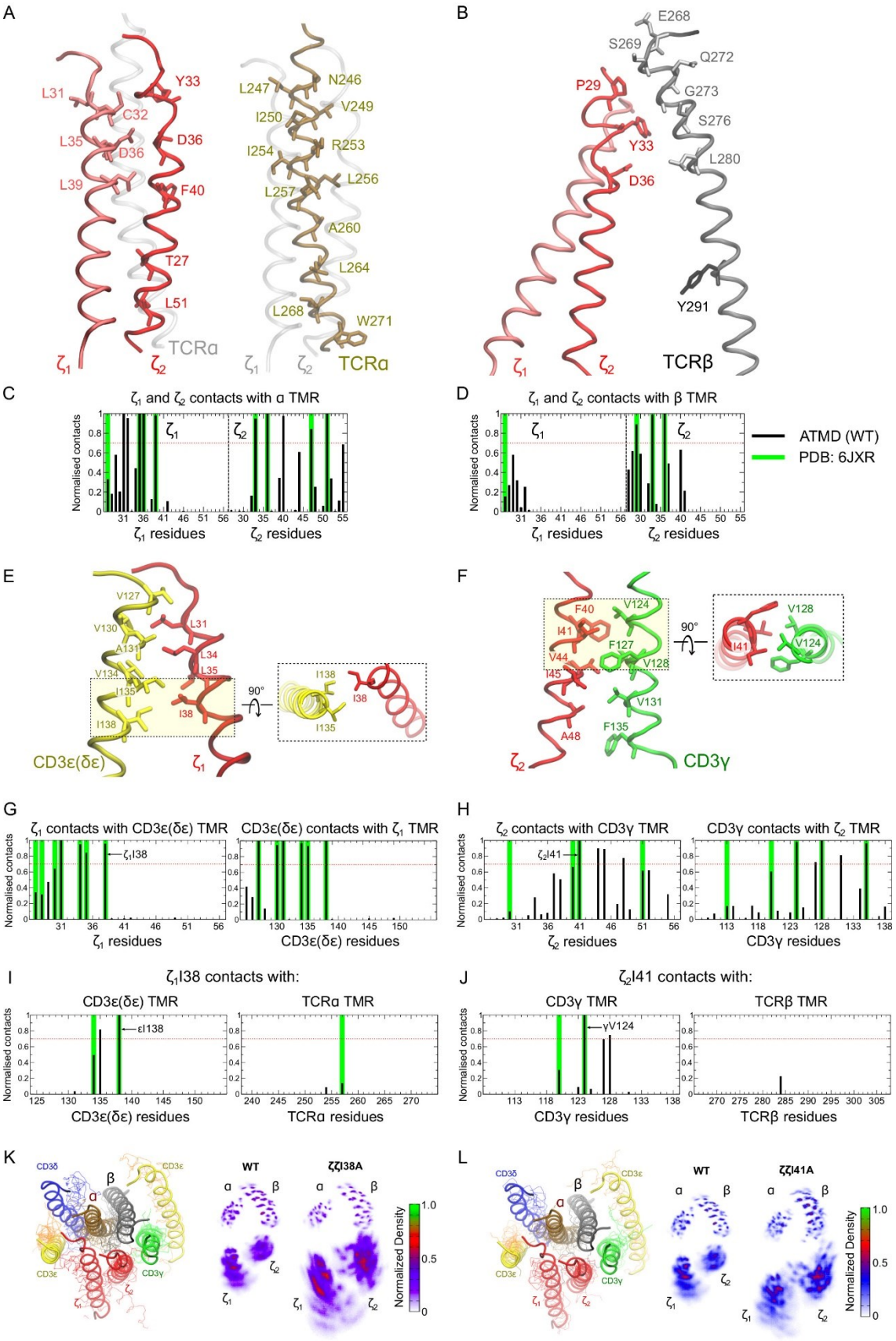
A
L31
Y33
C32
L35
D36
D36
L39
F40
T27
L51
TCRα
ζ1
ζ2
L247
N246
V249
I250
I254
R253
L256
L257
A260
L264
L268
W271
ζ1
ζ2
TCRα
B
E268
S269
Q272
P29
G273
S276
Y33
L280
D36
Y291
ζ1
ζ2
TCRβ
C
ζ1 and ζ2 contacts with α TMR
Normalised contacts
ζ1
ζ2
ζ1 residues
ζ2 residues
D
ζ1 and ζ2 contacts with β TMR
Normalised contacts
ζ1
ζ2
ζ1 residues
ζ2 residues
ATMD (WT)
PDB: 6JXR
E
V127
V130
L31
A131
L34
V134
L35
I135
I38
I138
CD3ε(δε)
ζ1
90°
I138
I38
I135
F
F40
V124
I41
F127
V44
V128
I45
V131
A48
F135
ζ2
CD3γ
90°
V128
I41
V124
G
ζ1 contacts with CD3ε(δε) TMR
CD3ε(δε) contacts with ζ1 TMR
Normalised contacts
ζ1I38
ζ1 residues
CD3ε(δε) residues
H
ζ2 contacts with CD3γ TMR
CD3γ contacts with ζ2 TMR
Normalised contacts
ζ2I41
ζ2 residues
CD3γ residues
I
ζ1I38 contacts with:
CD3ε(δε) TMR
TCRα TMR
Normalised contacts
εI138
CD3ε(δε) residues
TCRα residues
J
ζ2I41 contacts with:
CD3γ TMR
TCRβ TMR
Normalised contacts
γV124
CD3γ residues
TCRβ residues
K
CD3δ
CD3ε
β
α
CD3ε
CD3γ
ζ1
ζ2
WT
ζζI38A
α
β
ζ2
ζ1
Normalized Density
1.0
0.5
0
L
CD3δ
CD3ε
β
α
CD3ε
CD3γ
ζ1
ζ2
WT
ζζI41A
α
β
ζ2
ζ1
Normalized Density
1.0
0.5
0

***Fig 3.6 Interactions of ζζ TMR in the TCR-CD3.*** *(A) Snapshots of significant contacts between ζζ TMR and TCRα, and (B) ζζ TMR and TCRβ. βY291 is only shown (in black) for reference. (C) Normalised contacts between ζζ TMR and TCRα, and (D) ζζ TMR and TCRβ. (E) Snapshots of significant contacts between $\zeta_1$ TMR and CD3ε (δε), and (F) $\zeta_2$ TMR and CD3γ. Normalised contacts (G) between $\zeta_1$ TMR and CD3ε (δε), (H) between $\zeta_2$ TMR and CD3γ, (I) of $\zeta_1$I38 with CD3ε (δε) and TCRα, (J) of $\zeta_2$I41 with CD3γ and TCRβ. All normalised contacts (black bars) are compared with those identified from the cryo-EM structure (PDB:6JXR) using < 4 Å cut-off after adding hydrogen atoms (green bars). Normalisation is done by dividing all contacts by the highest number of contacts (0 to 1 scale). (K) Comparing the structural alignments (left) and spatial distribution of the Cα atoms of WT ζζ TMR (right) with ζζI38A and (L) with ζζI41A mutants, relative to the αβ TMR Cα atoms. In the structural alignments shown, the WT is shown in cartoon representation and the snapshots of the mutants at the end of 1250 ns are shown as lines.*

## 3.4 Discussion

In this chapter, I initially modelled the transmembrane helices of the α, β, δ, γ, ε subunits of the TCR-CD3 complex and simulated them in a POPC bilayer using classical / unbiased multi-scale (CG/AT) MD simulations to study their protein-protein interfaces and crossing angles. The aim of this work was to begin with answering questions concerning the orientation of the TMR helices and further model the rest of the protein complex based on experimental evidence. The simulations conducted in coarse-grained and then in atomistic resolution collectively revealed that the CD3δε TMR exhibits a more dynamic behaviour exhibiting a wider distribution of crossing angles since it shares a shorter protein-protein interface compared to the CD3γε TMR. However, when TCRα/β TMRs are bound to CD3δε/γε TMRs respectively, their crossing angles converged such to those observed in the cryo-EM structure (PDB:6JXR) [16]. More importantly, in addition to the crossing angles, the protein-protein interfaces calculated from my CGMD simulations of the dimers and trimers were similar to those observed in the cryo-EM structure. This provides more evidence that CGMD simulations using the Martini 2.2 forcefield [251] are capable of reproducing experimental results and thus lay a strong foundation in studying membrane proteins at near-atomic resolution *in silico*.

In addition, an interesting phenomenon that occurred during simulations of the dimers and trimers was the entry of waters into the hydrophobic core of the membrane via the outer leaflet. The local deformation in the bilayer near the core of the protein was caused by bending of lipids toward the core of the membrane so that their headgroups interact with the ionic sidechains of the protein TMR helices. Earlier MD simulation studies [57] reported this occurrence

but only upon dimerisation of the CD3 subunits. Here, I observe that it holds true upon TCRα-CD3δε and TCRβ-γε trimerisation as well.

The presence of ionic sidechains at the core of the TMR is primarily responsible for the assembly of the TCR-CD3 complex. They result in electrostatic interactions driving the oligomerisation of the subunits and this pattern in the assembly mechanism was found consistent in other immunoreceptor complexes such as the natural killer (NK) cell-activating NKG2D-DAP10 complex [59]. The release of the cryo-EM structure of the human TCR-CD3 complex [16] revealed the organisation of the TMR including these electrostatic interactions confirming earlier studies [8]. Using the atomic coordinates revealed by the cryo-EM study, and in collaboration with Prof. Oreste Acuto and his group, I studied the entire TCR-CD3 TMR and also induced *in silico* mutations. These mutations were performed to identify potential allosteric sites that could lead to loosening of the protein complex.

In this chapter, employing ATMD simulations to study the entire TCR-CD3 TMR, I first explored the effect of substituting βY291 to alanine. This mutant showed that it had a direct influence on the interactions between CD3γε and TCRβ TMR. This was expected since these subunits are immediate neighbours to each other sharing significant protein-protein interfaces. However, experimental results indicated that this mutation also had an indirect effect on the rest of the complex including the CD3δε TMR located on the opposite end. More specifically, they showed that it reduced association of the β subunit with CD3ε and ζ. This reduction in association with CD3ε and ζ subunits, which the TCR-CD3 depends on for signalling, led to an increase in phosphorylation of downstream kinase protein, Erk. Further, this indicated that stimulated TCR-CD3 complexes undergo activation as a result of quaternary structure relaxation in the TMR [76]. However, due to the limitation of time-scale in ATMD simulations, the effect of βY291A could probably be observed only in the region surrounding the mutated residue and not on the rest of the complex. Nevertheless, I also explored mutations in the ζζTMR i.e. ζζI38A and ζζI41A, which strongly suggested that they lead to a loosening in ζζ-αβ association. When replicated in experiments, these mutations also revealed TCR-CD3 activation similar to βY291A [76].

On the whole, my studies contributed to experimental findings [76] that the TCR-CD3 TMR plays a key role in transmitting mechanical signals from the pMHC binding site to the intracellular region by exposing its ITAMs. Previous studies have showed that allostery is a plausible mechanism of signal transduction due to the interactions between TCRαβ and the CD3 ECDs

[45,47,74,75]. Further, the connecting peptides that link the ECDs to their TMRs have also been shown to influence signalling and development of TCR-CD3 complexes [252]. The TCR-CD3 TMR however, apart from its contribution toward TCR-CD3 assembly, has not been extensively studied as a medium to propagate signals from the ECD to the CYR. The collaborative work, as explained in this chapter, addresses this void in our knowledge and suggests that modifications taking place in the TMR of stimulated TCR-CD3 potentially results in rearrangements in its CYR favouring ITAM exposure. Anionic lipids in the inner leaflet of the plasma membrane are known to interact strongly with basic-rich stretch (BRS) motifs in the cytoplasmic tails of CD3ε and ζ, thereby maintaining a hidden state of ITAMs in the membrane [33]. There are different mechanisms that may aid in dissociation of ITAMs from the membrane. For example, (i) interactions of TCR-CD3 CYR with actin cytoskeleton – but actin polymerisation and polarisation near the activation site occurs as a result of TCR-CD3-mediated signalling downstream and thus cannot be responsible for signal initiation, (ii) clustering of TCR-CD3 complexes or the 'aggregation model' – in this model, the cytoplasmic tails are released from the plasma membrane due to over-crowding of proteins and lack of space for the tails to interact with the membrane. However, it has been shown that monomeric TCR-CD3 complexes are sufficient to trigger the initiation of signalling [85]. Therefore, upon pMHC stimulation, the untying of TCR-CD3 cytoplasmic tails and exposure of ITAMs is potentially a result of alterations or more specifically 'loosening' in the TMR.

To observe alterations in the TCR-CD3 TMR upon stimulation using MD simulations, the complete TCR-CD3 will have to be modelled and stimulated with a pMHC molecule, potentially using the steered MD simulation approach. However, given the limitation of time-scale in ATMD simulations and of the elastic network restraints in CGMD simulations, providing dynamic insights into the TCR-CD3 TMR alterations including potential changes in the CYR demands a simulation of much larger time-scale. To achieve this, a different approach such as using implicit solvent models [253,254] may be fruitful. Simulating explicit water molecules in addition to protein / lipids can be very time-consuming and/or computationally expensive. Moreover, solvent accessibility of TCR-CD3 cytoplasmic tails can still be calculated without explicit waters. Therefore, future *in silico* studies can apply this approach to minimise computational cost, achieve larger simulation time-scales, and push the limits to understand TCR-CD3 activation in atomistic detail.

### 3.4.1 Key research findings

- Molecular modelling combined with multi-scale MD simulations (using the Martini 2.2 and CHARMM36 forcefields) in this chapter provided an insight into the protein-protein interfaces of the CD3 TMR dimers and TCR-CD3 TMR trimers. These resembled the interfaces observed in the cryo-EM structure that was later published [16]. This potentially indicates that this methodology can be used to predict inter-subunit interactions of more such membrane proteins in a dynamic environment.
- ATMD simulations showed that the ζζI38A and ζζI41A mutants in the complete TCR-CD3 TMR led to the loosening of ζζ relative to TCRαβ. This conformational change in the TMR potentially leads to ITAM exposure in the cytoplasm. Consistently, experiments showed that these mutants mimicked the effect of pMHC binding and enhanced downstream T cell signalling [76].

# Chapter 4

# Steered molecular dynamics of the peptide-MHC and its interactions with the T cell receptor extracellular domain

## 4.1 Introduction

In the previous chapter, I simulated the TCR-CD3 transmembrane region and provided insights into residues that potentially aid in signalling from the extracellular to the intracellular region supporting experimental findings. In this chapter, using steered molecular dynamics (SMD) simulations, I study the first and foremost event of TCR-CD3 activation i.e. the interaction between the extracellular domain (ECD) of the TCRαβ dimer and antigenic peptides attached to major histocompatibility complexes (pMHCs).

Antigen recognition carried out by the TCRαβ immunoglobulin-like ECD (containing constant and variable domains) relies on interactions of antigenic peptides with the complementarity determining region (CDR) loops 1, 2, and 3, that are situated on the outer edge of the TCRαβ variable domain [255,256]. Experimental evidence has also suggested that the TCRαβ and/or the pMHC contain flexible binding surfaces which undergo stabilisation upon binding, potentially contributing to cross-reactivity [257] i.e. the ability of an antibody to bind to different antigens. Upon pMHC ligation, the TCRαβ constant domain transmits allosteric conformational changes toward the intracellular side, with the protruding FG loop of the β constant domain being the key intermediate [44,45,47] (indicated in Fig 4.1 left). This allosteric reaction initiated in the ECD not only transmits the signal toward the intracellular region but was also found to help optimise pMHC binding [46].

How a single TCR-CD3 discriminates between a plethora of antigenic peptides has remained elusive. Addressing this, single molecule fluorescence resonance energy transfer (smFRET) experiments were conducted to measure the bond distances between TCR-CD3 and pMHCs [258]. This study revealed that the TCR-CD3 formed closer interactions when associated with high affinity antigens and farther interactions with lower affinity antigens. The binding of CD4 co-receptor with the pMHC was found to have much lower affinity [258,259] compared to TCR-pMHC interactions and had a minimal effect on them [258,260]. However, CD8 co-receptors are suggested to exhibit greater affinity toward class I MHC than CD4 does toward class II MHC, thereby influencing antigenic specificity during TCR-pMHC interaction [261]. Interestingly, the smFRET study also showed that $Ca^{2+}$ signalling along with the dissociation of the ζ cytoplasmic region (CYR) from the membrane were strongly (negatively) correlated to the TCR-pMHC bond distance, meaning that high affinity antigens forming closer contacts with TCRαβ induced a higher degree of T cell activation [258]. In addition to the affinity or ‘strength’ of TCR-pMHC binding, it is also proposed that the triggering of the TCR-CD3 complex is driven by the

confinement time of TCR-pMHC contact [262], and mechanical force exerted onto the TCRαβ ECD that potentially induces conformational changes in the ECD, followed by the TMR and the CYR [77].

The affinity of antigenic peptides toward TCRαβ has immense implications in designing vaccines [14,263]. Immunisation using peptides that are altered has helped enhance T cell responses in cancer patients and reported to be a promising immunotherapy [264]. Further, cysteinylation is suggested to decrease the antigenicity of peptides attached to class I MHC molecules [265]. Consistent with this, an experimental study showed that the substitution of cysteine to valine in the 9$^{th}$ position (9V) in a native SLLMWITQC peptide enhanced interactions with the TCRαβ ECD in solution [14]. The study also provided insights into the potential of altered peptides in improving immunogenicity and highlighted the importance of structural studies in optimising peptides for clinical trials. In addition, it resolved the structure of the extracellular region of the TCRαβ-pMHC complex containing the 9V peptide using X-ray crystallography at 1.7 Å [14] (PDB:2BNQ; view in 3D here: https://www.rcsb.org/3d-view/2BNQ/).

In this chapter, using the atomic coordinates of the crystal structure (PDB:2BNQ), I induced single point *in silico* mutations, simulated them in atomistic resolution, and compared them with the 9V peptide-MHC complex in terms of interactions and dissociation from the TCRαβ ECD. The simulation study conducted in this chapter was initiated in collaboration with Dr. Omer Dushek and his group at the University of Oxford but was not published. The selection of the specific mutations was made depending on their experimental results when studying the dissociation rates of peptide variants and the ability of the TCRαβ to discriminate between them [266]. In addition, my simulations reveal some major conformational changes in the MHC upon its dissociation from the TCRαβ ECD that have not been previously studied.

### 4.1.1 Key research aims

- Study the bond lifetime of the TCRαβ with different variants of the peptide (obtained by single point mutations) loaded into a class I MHC.
- Study the cause of a ‘catch bond’ i.e. increase in TCR-pMHC bond lifetime upon applied force.
- Shed light on potential structural changes upon TCR-pMHC dissociation.

## 4.2 Methods

### Steered molecular dynamics simulations

The crystal structure of the extracellular region of the TCR-pMHC complex (PDB:2BNQ) was used as the starting structure for conducting steered molecular dynamics (SMD) simulations. This protein complex was simulated in solution (SPC water model) neutralised by a 0.15M concentration of $Na^+$ and $Cl^-$ ions. The system was energy minimised using the steepest descent algorithm and equilibrated for 2 ns using the NPT ensemble with the protein backbone restrained. Production SMD simulations (5 replicates x 100 ns) for each pMHC variant (9V peptide (WT; PDB:2BNQ), 4A, 5F, 6T mutants) were carried out by position-restraining (using 1000 kJ/mol/nm) the backbone of TCRαβ ECD except their CDR loops, and a constant pulling force was applied on the MHC backbone except on the peptide and the long helices that hold the peptide (Fig 4.1 middle). For each pMHC variant, a number of different force constants used for pulling (see results). The direction used to pull the MHC away from the TCR was along the +Z axis. Sufficient space was provided between the protein and the edge of the PBC box (X = Y = 13 nm; Z = 18 nm) in the direction of pulling. The temperature was maintained at 310 K using V-rescale thermostat [224] and an isotropic pressure of 1 bar was used with the Parrinello-Rahman barostat [225].

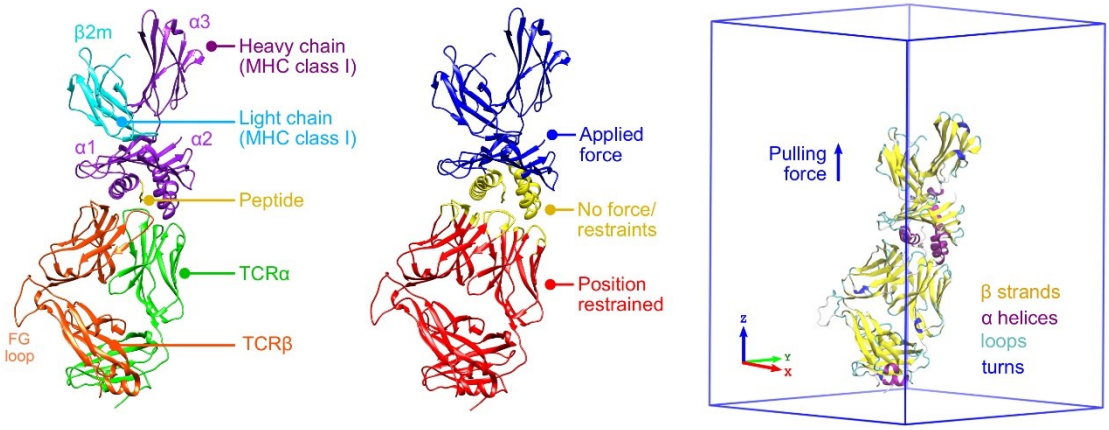


***Fig 4.1 Snapshot of SMD simulation setup.*** *Labelling of the chains comprising the extracellular domains of the TCRαβ and pMHC complex obtained from PDB:2BNQ (left). Labelling of the backbone region of the TCRαβ and pMHC ECDs indicating where an applied pulling force is exerted, where no force is exerted, and where position restraints are applied (middle). The protein placed in the PBC box before the simulations (right).*

## 4.3 Results

### 4.3.1 SMD simulations of TCRαβ-pMHC ECD indicate relative time of dissociation between peptide variants

Using the crystal structure of the extracellular region of the TCRαβ-pMHC complex containing the high affinity 9V peptide (PDB:2BNQ) [14], I setup SMD simulations where external forces were applied on the MHC to be pulled away from TCRαβ (Fig 4.1 right). This experimentally altered (C9V substituted) peptide was considered as a WT in my simulations. Experiments have shown that TCR-pMHC dissociation can occur at forces as low as ~10 to 20 pN, but occur on the order of milliseconds to seconds [209]. To tackle the cost of computation, a fixed time-scale (100 ns) was simulated with varied force constants. The forces applied were 12, 30, 60, 150 kJ/mol/nm i.e. 20, 50, 100, 250 pN (1 kJ/mol/nm = 1.66 pN). However, in the WT, these forces did not result in TCRαβ-pMHC dissociation as they were probably lesser than the required force for dissociation within 100 ns. I then used a force constant of 181 kJ/mol/nm i.e. 300 pN and simulated the same system. I observed that some simulations (3/5) showed dissociation, and thus used this force constant to simulate the mutants i.e. 4A, 5F, 6T in addition to the WT (5 x 100 ns). However, only the simulations resulting in TCRαβ-pMHC dissociation at the end of the time-scale were analysed.

I was interested in studying the influence of the four sequential variations in the peptide on TCR-pMHC interactions and dissociation. Following visual inspection, in some simulations, I observed that R65 in the α1 helical region of the MHC made ionic interactions with D54 of the variable domain of TCRβ (Vβ) (Fig 4.2A, B) which delayed the time of TCR-peptide dissociation, potentially biasing the data. Therefore, simulations indicating delay in TCR-peptide dissociation due to R65 were not considered for analyses. Finally, from the rest of the simulations, I compared the time of dissociation between each peptide variant by calculating the distance versus time (Fig 4.2C), contacts (< 4 Å) versus time (Fig 4.2D), and hydrogen bonds versus time (Fig 4.2E). These data showed that the time of dissociation (WT>6T>4A>5F = 88>72>20>10 ns) agreed with the inverse order of their dissociation constants ($K_D$) (9V<6T<4A<5F = 8<162<299<1309 μM) revealed by experimental findings [266] but not with the proportion of the experimental dissociation constants. Nonetheless, according to the contacts and hydrogen bonds versus distance analyses, the WT was observed to make more contact with TCRαβ at greater

distances compared to the mutants (Fig 4.2F-G), consistent with experimental evidence that it exhibits least dissociation ($K_D$= 8 μM) [266].

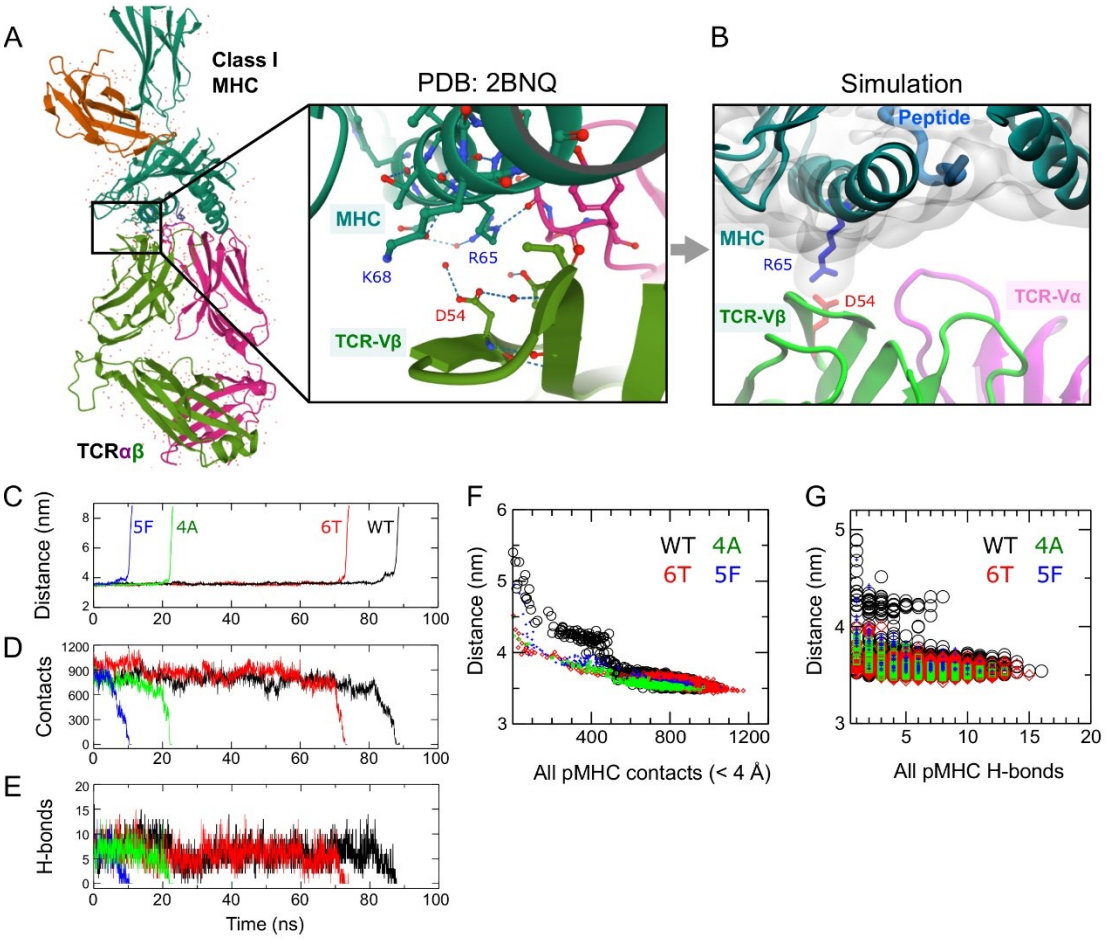


***Fig 4.2 TCR-pMHC dissociation and contacts.*** *(A) Potential ionic interactions between the MHC and TCR-Vβ in the crystal structure (PDB:2BNQ). (B) In some simulations, R65 of the MHC heavy chain maintained contact with D54 of TCR-Vβ after the peptide and the rest of the MHC had dissociated from the TCR. The surface of the pMHC is shown in transparent to indicate that R65 was the last residue to dissociate after delaying TCR-pMHC interaction. (C) TCR-pMHC distance versus time. (D) All TCR-pMHC contacts (< 4 Å) versus time. (E) TCR-pMHC H-bonds versus time. (F) Distance versus all TCR-pMHC contacts. (H) Distance versus TCR-pMHC H-bonds.*

### 4.3.2 Comparing TCRαβ-pMHC sidechain contacts at specific distances along the SMD trajectory

To study the TCRαβ-pMHC interactions of each peptide variant in more detail, I extracted snapshots of the system along the SMD trajectory where each snapshot represented the position of the pMHC at certain distances (35, 37.6,

40, 41.8, 45, 50, 56 Å) from the TCRαβ. The distances were measured from the centre of mass of the TCRαβ ECD backbone to that of the peptide and the adjacent helices, as shown in Fig 4.3A. Using these extracted snapshots as starting structures, classical MD simulations were further performed for 20 ns time for each distance and each peptide variant. In these simulations, the backbones of the TCRαβ and MHC were restrained except their interacting regions (highlighted in yellow in Fig 4.3A) to allow their sidechains to make as many contacts as possible within the time frame.

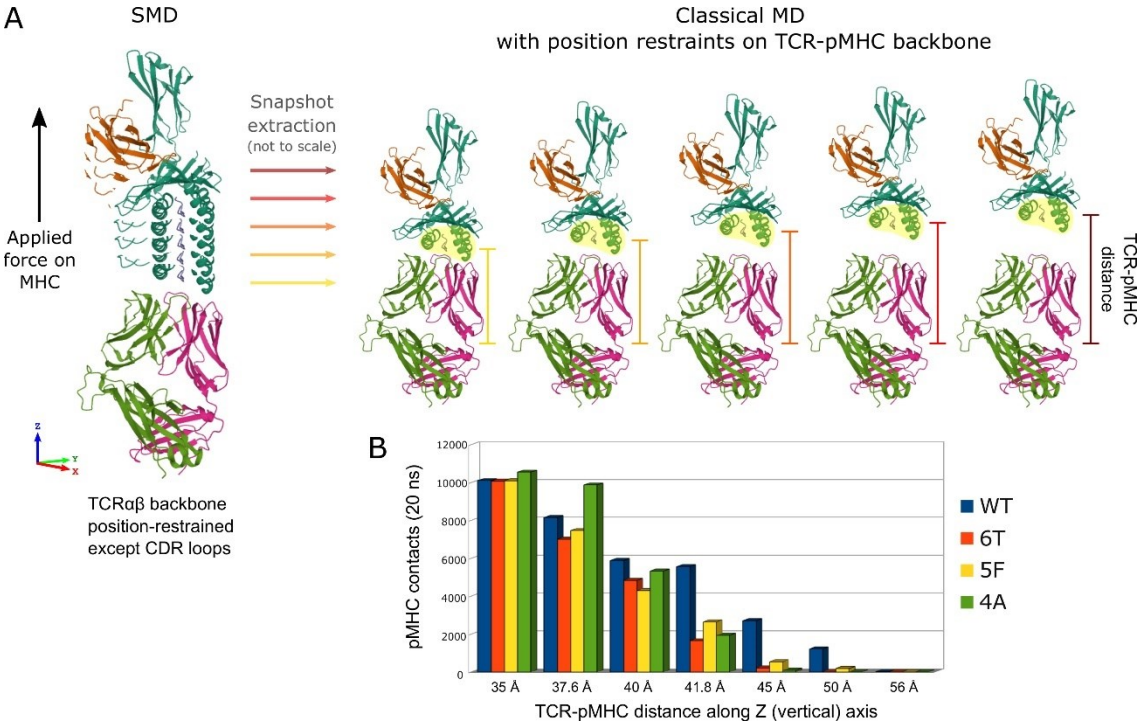


***Fig 4.3 Classical MD simulations upon extracting snapshots from SMD simulations.*** *(A) The snapshots were extracted from specific distances measured from the centre of mass of the TCRαβ backbone to the highlighted region (yellow) of the pMHC. The highlighted region corresponds to the backbone of the peptide and of the helices located on either side of the peptide. (B) Contacts of the pMHC with the TCRαβ when its backbone is position-restrained (fixed) at specific points along the SMD trajectory.*

These classical MD simulations, which provided additional sampling for calculating TCRαβ-pMHC interactions, confirmed that their overall interactions reduced with increase in the distance between them in all systems as expected. More importantly, they confirmed my previous observation (Fig 4.2F-G) that the WT peptide-MHC made more contacts with the TCRαβ at greater distances (≥ 40 Å) compared to the mutants (Fig 4.3B). Further, I also calculated the contacts of only the peptide with the TCRαβ, and made a comparison of

contacts between peptide variants and between the different distances along the trajectory at which the pMHC was fixed (Fig 4.4).

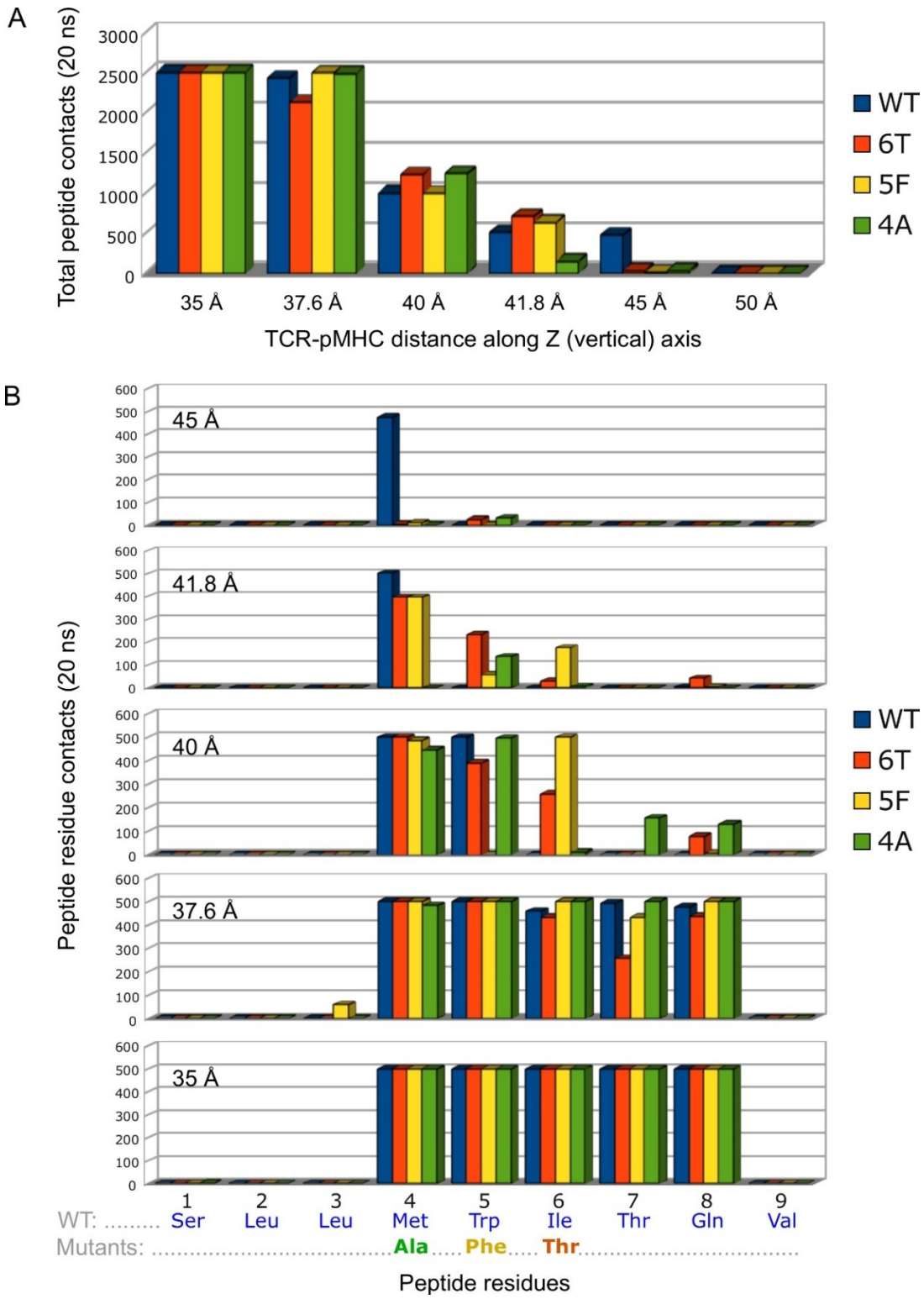


***Fig 4.4 Peptide interactions with the TCRαβ.*** *(A) Total peptide-TCRαβ contacts across peptide variants and distances along the SMD trajectory. (B) Number of contacts of each residue in each peptide variant at specified distances along the SMD trajectory.*

Further, these residual contacts between the pMHC and TCRαβ are illustrated after mapping them onto the structure (Fig 4.5).

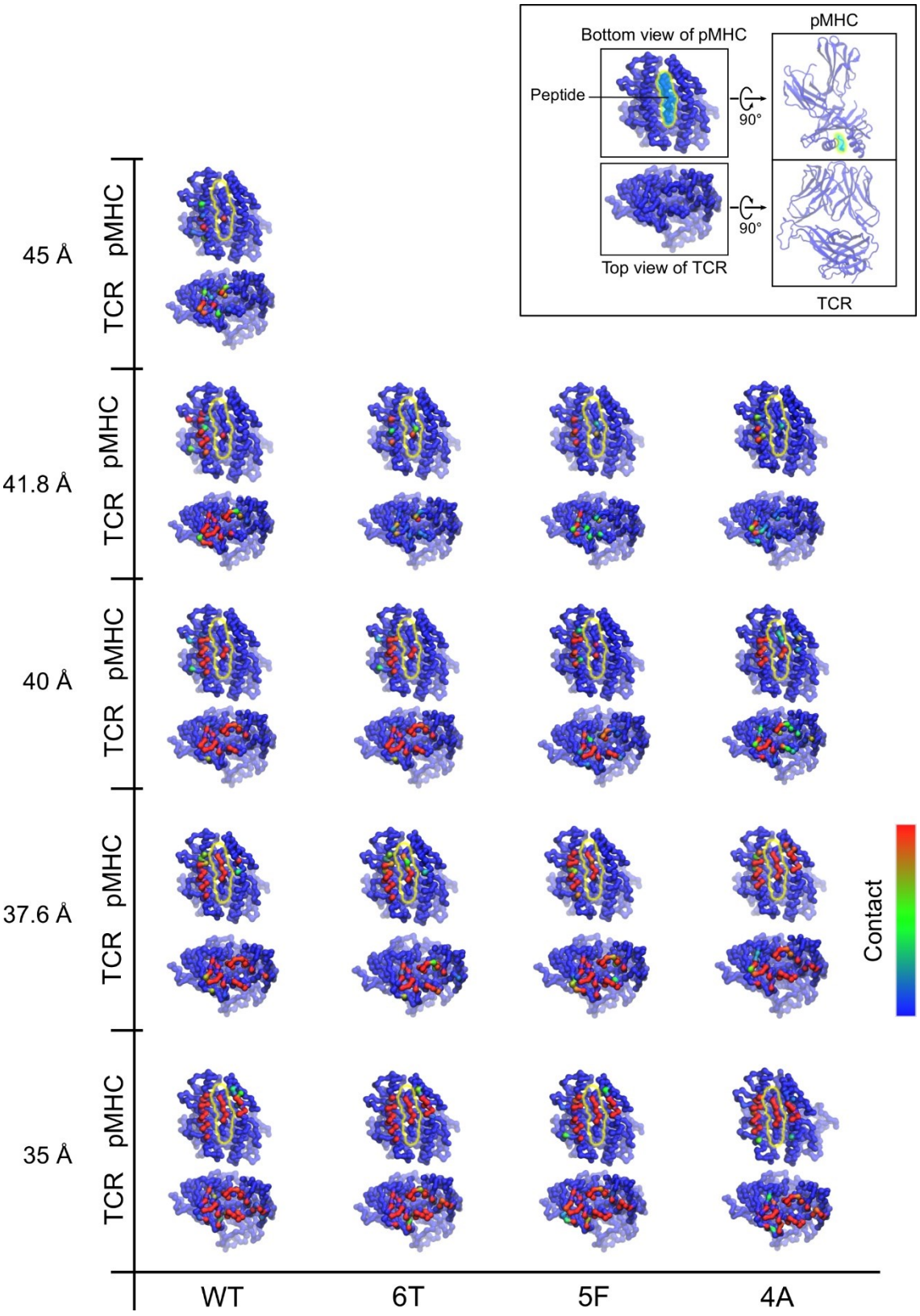

***Fig 4.5 Peptide-MHC and TCRαβ ECD interactions mapped onto 3D structure.*** *The peptide is highlighted within a fuzzy yellow line for clarity. 3D structural maps that did not indicate TCR-peptide contacts are not shown. The box above indicates the rotated representation of the TCR/pMHC.*

The above analysis (Fig 4.4 and Fig 4.5) indicated that the WT peptide retained contacts until a distance of ~45 Å specifically via the 4$^{th}$ residue (M4) while the contacts of other peptide variants had diminished when they reached a distance of ~42 Å. The fact that the WT peptide maintained contact with the TCRαβ CDR3 compared to the other peptide variants despite the applied force potentially indicates the contribution of specific peptide residues toward a 'catch bond' formation (meaning an increase in TCR-pMHC bond lifetime when a pulling force is applied) as reported by experiments [209]. On the other hand, the protein-protein contact analyses from my simulations could not distinguish very well between the mutants, potentially due to the limited time-scale and higher force applied compared to those used in experiments (~10 to 20 pN) on the order of hundreds of milliseconds [209].

### 4.3.3 Conformational changes in the class I MHC upon dissociation from the TCR

Significant conformational changes were observed in the class I MHC ECD upon dissociation from the TCRαβ in my SMD simulations (using 181 kJ/mol/nm or 300 pN force) simulations. These included the 'linker stretch' (LS), the 'peptide-intact' (PI), and the 'helix-uncoiled' (HU) conformational states (Fig 4.6) as referred to here. Simulations suggested that the LS was a more common conformational change where the β2-microglobulin (β2m) and the α3 region of the class I MHC (labelled in Fig 4.1 left) first dissociated under force, and thus the loop connecting the α3 and α2 regions of the heavy chain underwent extension before the entire pMHC dissociated from TCRαβ (Fig 4.6A). This was consistent with previous computational and experimental reports [209] and was thus considered during analyses. In the PI state, the peptide stayed intact with the TCRαβ after the MHC dissociated (Fig 4.6B). The HU state was an event that accompanied the PI state (Fig 4.6C). Here, one of the helices that holds the peptide in the MHC was observed to uncoil as a result of a strong ionic interaction (between R65 and D54, as shown in Fig 4.2B) that seemed to resist the pulling force. However, the PI and HU states were not considered for analyses since the number of simulations that resulted in these were negligible, indicating that they may be an unlikely event.

To further test the probability of these changes in a more rigorous manner, I conducted more simulations using varying force constants to pull the MHC i.e. 181, 196, 211, 226, 241, 256, 271 kJ/mol/nm or 300, 325, 350, 375, 400, 425 pN (1 kJ/mol/nm = 1.66 pN). For each of these forces and each of the peptide variants, 5 individual simulations of 100 ns each were conducted, totally amounting to 140 simulations and 14 μs simulation time. The LS, PI, and HU conformational states were observed in 30%, 15%, and 5% of all the simulations respectively. Note that the PI (Fig 4.6B) and HU (Fig 4.6C) states mostly occurred when higher TCR-peptide affinity was involved i.e. in the WT followed by 6T. This is because the wild-type residues in the 4$^{th}$ and 5$^{th}$ positions i.e. M4 and W5 respectively, were majorly responsible for the heightened TCR-peptide interactions (Fig 4.6D). When these residues were mutated in the 4A and 5F variants, the peptide dissociated more easily along with the MHC.

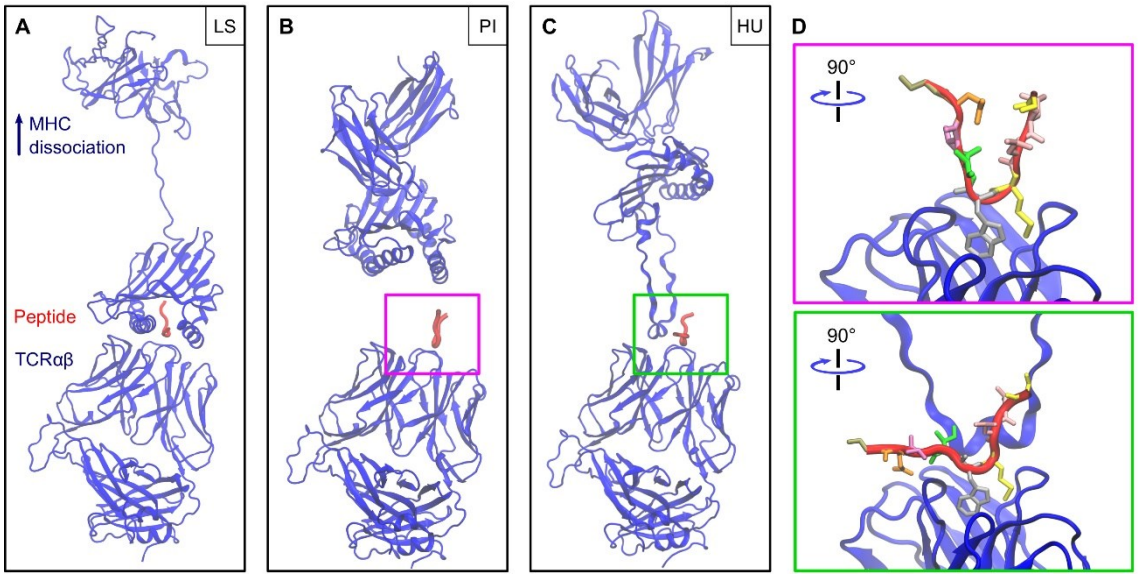


***Fig 4.6 Class I MHC conformational changes upon dissociation from TCRαβ****. (A) The 'linker stretch' (LS), (B) the 'peptide-intact' (PI), and (C) PI coupled with the 'helix-uncoiled' (HU) state. (D) A zoomed-in view of the TCR-peptide interactions in the PI and HU states. The binding of the WT peptide residues at the 4$^{th}$ (Met; yellow sticks) and 5$^{th}$ (Trp; silver sticks) positions are also shown.*

Consistent with the observation of the LS in my simulations, experiments showed that this event not only holds true but is also the reason for the catch bond formation [209]. Similarly, given that the PI state was observed in 15% of the simulations, compared to 30% of the LS, it is possible that this may also

hold true. Note that PI has not been reported before, and thus requires further investigation for validation and to examine its role in TCR-CD3 activation. Although it does not necessarily change our current understanding of TCR-CD3 activation, it does lead to interesting questions such as – (i) Upon MHC dissociation, are TCR-CD3 complexes with intact peptides (PI-TCRs) incapable of signalling thereafter?, (ii) can PI-TCRs be part of signalling TCR-CD3 clusters at the immunological synapse and minimise signalling capacity?, (iii) when another pMHC collides with a PI-TCR with a certain mechanical force, can it lead to signalling against the wrong antigen?

## 4.4 Discussion

In this chapter, I studied the dissociation of the class I MHC containing a peptide (human histocompatibility leukocyte antigen (HLA)-A2 peptide) that was altered (C9V) to enhance its affinity toward the (1G4) TCR [14]. The experimental study also employed X-ray crystallography to resolve the structure of the altered peptide at 1.7 Å (PDB:2BNQ). This altered peptide obtained from the crystal structure was considered as a WT in my simulations. The crystal structure, that my simulations started from, shows that the V9 residue faces toward the MHC, distant from the CDR loops of TCRαβ. Consistently, the C9V residue in my simulations did not interact with the TCRαβ and would potentially yield similar results with the native C9 residue. More importantly, I observed that the M4 residue interacted the most in the WT simulations.

In addition to the C9V peptide alteration, I induced other *in silico* mutations (M4A, W5F, I6T) in the interest of our collaborators, Dr. Omer Dushek and group at the University of Oxford. My simulations revealed that the WT peptide delayed the dissociation of the MHC from the TCR and exhibited significant number of interactions at greater TCR-pMHC distances compared to the mutant peptides. Moreover, these simulations enabled me to study the contributions of certain wild-type residues, especially the M4 and W5 residues toward increased TCR-pMHC interactions under applied force.

About a third of my simulations (100 ns each) that aggregated to 14 μs time showed that the MHC underwent a major conformational change termed here as the ‘linker stretch’. This is consistent with previous experiments indicating that the stretching of the loop region linking α3 to α2 in the class I MHC heavy chain increases the TCR-pMHC bond lifetime giving rise to a catch bond. When this stretching was prevented by inducing a disulphide bond between the heavy and light chains, the bond lifetime was observed to reduce by ~2-fold [209].

Further, I also observed other MHC conformational changes taking place in my simulations upon dissociation from TCRαβ that have not been reported before. These are the 'peptide-intact' (PI) and 'helix-uncoiled' (HU) states of the class I MHC upon dissociation from the TCR, occurring due to strong peptide-TCR interactions and ionic MHC-TCR-Vβ interactions respectively. The ensemble of my simulations also showed that the MHC exhibited PI and HU states in 15% and 5% of all simulations. These conformational states were observed mostly in presence of the wild-type residues at $4^{th}$ and $5^{th}$ positions (M4 and W5) since these residues strongly interacted with the TCR and seemed to resist the pulling force to some extent.

The HU state of the MHC was observed to occur alongside the PI state in my simulations. This was expected because the loss of the helical structure in either α1 or α2 of the class I MHC would disable it from holding the peptide thereafter. Moreover, even if we speculate the possibility of the HU state *in vivo* despite its low occurrence in my simulations, it may not significantly aid in our understanding of TCR-CD3 activation. This is because a non-structured MHC devoid of a peptide present on the surface of an antigen presenting cell (APC) cannot trigger TCR-CD3 complexes, and is likely to be internalised. On the other hand, the PI state may have implications during T cell activation. TCR-CD3 complexes that are already attached to peptides previously transferred from MHCs, similar to the event of 'trogocytosis' [267–270], may reduce the overall magnitude of antigen recognition in a following signalling event when clustered at the T cell-APC interface. Experimental approaches such as mass spectrometry may be employed to detect peptide bound to TCRs on T cell surfaces, similar to experiments that were conducted to isolate and detect peptides bound to MHCs [271]. Moreover, structural mass spectrometry has been used to determine populations of protein complexes bound to small-sized ligands such as ATP, and the same approach is suggested to be applicable in detecting a wider range of protein-ligand complexes [272]. The probability of peptides dissociating from MHCs may also be computationally predicted by measuring the affinity of peptides to MHCs [273–275] compared to their affinities with associated TCRs. However, my simulations indicating the PI state do not account for glycosylated peptides. Additionally, the dissociation of peptides from MHCs is highly unlikely if they are disulphide-linked [276].

### 4.4.1 Limitations

The time-scale of my SMD simulations was limited to 100 ns each. Therefore, a larger force was used to aid in TCR-pMHC dissociation within the time-frame.

This is likely to have reduced the accuracy in differentiating between the binding affinities of peptide variants. For protein-protein interaction analyses, only the simulations that did not exhibit the PI and HU states were considered. As a result, it limited the number of individual simulations sampled. Given the computational cost and the large dataset involving multiple peptide variants and force constants, a single direction of pulling the MHC away from the TCR was considered. Although, previous simulations have also employed a single direction of pulling and reproduced experimental observations of TCR-pMHC bond lifetimes [209]. However in future, multiple pulling directions can be explored to mimic multiple possibilities and study their influence on TCR-pMHC interactions.

### 4.4.2 Key research findings

- Steered ATMD simulations suggest that the residues M4 followed by W5 in the HLA-A2 peptide complexed with a class I MHC (PDB:2BNQ) play a significant role in binding with the TCR. Mutation of these residues led to TCR-pMHC dissociation within a shorter time-scale indicating a higher dissociation rate. Therefore, in contrast to enhancing TCR-pMHC binding affinity to induce stronger T cell responses in cancer patients [14], this chapter reveals specific residues that can possibly be altered to regulate T cell responses.
- In addition to a previously reported conformational change (referred to in this chapter as ‘linker stretch’) in the class I MHC that gives rise to a catch bond upon TCR-pMHC dissociation [209], this chapter highlights other types of potential structural changes referred to as ‘peptide-intact’ and ‘helix-uncoiled’.

# Chapter 5

# Molecular dynamics and lipid interactions of the complete T cell receptor

This chapter is based on the following publication cited in this thesis as [277]:

*Prakaash et al. Multi-scale simulations of the T cell receptor reveals its lipid interactions, dynamics and the arrangement of its cytoplasmic region. PLOS Comp Biol. 2021. doi: 10.1371/journal.pcbi.1009232*

## 5.1 Introduction

In the previous chapters, I conducted studies explicitly on the T cell receptor (TCR-CD3) transmembrane region revealing potential sites that could transmit extracellular signals into the intracellular region, and on its extracellular region that interacts with peptides attached to major histocompatibility complexes (pMHCs) showing its sensitivity toward antigenic peptide sequences. In this chapter, I modelled its cytoplasmic region to provide a model of the complete TCR-CD3 complex, and studied its dynamics and lipid interactions in its resting state.

T lymphocytes express a diverse repertoire of TCR-CD3 complexes on their plasma membrane which initiate T cell signalling upon identifying antigenic peptides displayed in the jaws of major histocompatibility complexes [9,278]. The TCR-CD3 complex consists of four non-covalently assembled dimers: TCRαβ, CD3δε, CD3γε heterodimers and the ζζ homodimer [8]. Disulphide bridges aid in linking the subunits within the αβ and ζζ dimers in the extracellular region. Inter-subunit interactions are mediated by the ectodomains (ECDs) and the transmembrane regions (TMRs) which contribute to the stability of the complex and determine its precise topology [8,16,244]. The α and β subunits each display variable domains, Vα and Vβ, featuring three variable loops of complementarity determining regions (CDRs) 1, 2 and 3 that together form the VαVβ binding site for peptide-MHC (pMHC) ligands [279]. The cytoplasmic regions (CYRs) of the α and β subunits each contain short peptides of less than ten amino acids long and do not transmit signals to the intracellular region.

TCR-CD3 signal transduction, which is initiated by pMHC binding to TCRαβ, is governed by the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) in the intracellular region of CD3 and ζ subunits [280]. Circular dichroism (CD) spectroscopy experiments suggested that the CYRs of each ITAM-containing subunit are intrinsically disordered in both monomeric and

oligomeric states, but exhibit lipid binding with acidic phospholipid-containing vesicles [17]. CD3ε and ζ CYRs contain basic-rich stretches (BRS) that have been suggested to mediate robust ionic interactions with negatively charged headgroups of phosphatidylinositols [28,29] and phosphatidylserine [33] in the inner leaflet of the membrane. Moreover, the tyrosine sidechains in the ITAM-containing segments of both CD3ε and ζ CYRs were found to penetrate into the hydrophobic core of the membrane [33,34]. However, it remains unclear whether this configuration applies to all cytoplasmic tyrosines in an entire TCR-CD3 complex. A 'stand-by' model of TCR-CD3 signalling was proposed based on the evidence that a pool of constitutively active LCK at the T cell plasma membrane phosphorylates the ITAMs of TCR-CD3 CYRs upon their disengagement from the membrane [114]. However, the molecular mechanism of pMHC binding that induces a change in the ITAM configuration which favours accessibility by LCK remains unclear. The key to these questions potentially lies in the interactions of the TCR-CD3 complex with its local membrane environment which is also currently poorly understood.

A recent cryo-electron microscopy (cryo-EM) study [16] revealed the 3D structure of the human TCR-CD3 complex at a resolution of 3.7 Å (PDB:6JXR) and supports previous findings that TCRαβ maintains critical ionic contacts with CD3δε, CD3γε and ζζ in the TMR [8]. This study shed light on the quaternary structure arrangement featuring highly interlaced contacts among subunits' ECDs and TMRs, suggesting a dense connectivity maintaining the topology of the entire complex. However, the cryo-EM structure could not identify the arrangement of the CYRs presumably due to their disordered state. Moreover, the dynamic behaviour of the TCR-CD3 ECDs, TMRs and CYRs when embedded in their native membrane environment has not been studied. This information may provide clues to the signal transduction mechanism.

Here, the cryo-EM structure was used to generate the first molecular model of the entire TCR-CD3 embedded in a complex asymmetric bilayer containing the predominant lipid headgroup species found in its native environment upon receptor activation [210]. The multi-scale molecular dynamics (MD) simulation approach, in coarse-grained and atomistic resolutions, provided insights into the conformational flexibility of the TCR-CD3 and its interactions with membrane lipids in the microseconds timescale. The CYRs is observed to assemble into a coiled conformation and interact with the inner membrane leaflet, while anionic headgroups of phosphatidylinositol phosphate (PIP) lipids interact selectively with the TCR-CD3 CYR. The simulations also reveal that the ECDs, TMRs, and CYRs each exhibit conformational changes when the TCR-CD3 is embedded

within the lipid bilayer, potentially supporting a model of allosteric activation of the TCR-CD3 complex.

### 5.1.1 Key research aims

- Obtain a structural model of the complete TCR-CD3 complex, and elucidate the arrangement and dynamics of its cytoplasmic region.
- Study the lipid interactions of the complete TCR-CD3 complex and the contribution of the cytoplasmic region towards it.
- Study protein-protein interactions and conformational changes occurring in the TCR-CD3 ECD and TMR compared to its cryo-EM structure.

## 5.2 Methods

### 5.2.1 Molecular Modelling

The cryo-EM structure (PDB:6JXR) [16] was used as a template to obtain the complete TCR-CD3 model. Sequences of each subunit were obtained from UniProt (uniprot.org): ζ:P20963, δ:P04234, ε:P07766, γ:P09693, α:A0A0B4J271, β:P0DSE2. Modeller 9.2 was used to model the entire TCR-CD3 complex [281,282] (Fig 5.1) along with UCSF Chimera [245]. Hydrogen atoms were added to the model and topologies were generated using the CHARMM36 forcefield [214] and Gromacs 2016 [213].


***Fig 5.1 Sequences of the full-length TCR-CD3 subunits used for modelling.*** *Their secondary structure was predicted by the PSIPRED 4.0 server. The extracellular and intracellular residues modelled in this chapter are shown in boxes and underlined respectively.*

### 5.2.2 Coarse-grained molecular dynamics simulations

The structural models were coarse-grained using the *martinize* script. CGMD simulations were conducted using Gromacs 5.0 [212] with the Martini 2.2 forcefield [251]. The TCR-CD3 is comprised of the αβ, δε, γε, ζζ dimers non-covalently bonded with each other. To replicate this and to avoid restraints between each dimer, an elastic network model [233] with a lower cut-off distance of 0 nm, an upper cut-off distance of 0.7 nm, and a force constant of 1000 kJ/mol/nm was applied only between subunits within each dimer to maintain their dimeric tertiary/quaternary structures. This was done by martinizing each dimer separately and then concatenating their coordinates. Note also that there were no elastic network restraints between the unstructured regions of the tails, but only within the small helical regions in the ζζ tails. The tails of each subunit were modelled sufficiently far away from each other to observe intracellular dynamics without any restraints. The protein complex was then placed in a simulation box and inserted into a complex asymmetric bilayer

using the *Insane* tool [240]. The concentration of lipid headgroups in both membrane 1 and 2 was the same (Table 5.1) and is based on lipidomics studies of TCR-CD3 activation domains [210].

***Table 5.1 Composition of lipid headgroup types (%) in the membrane***

| **Lipid concentration (%)** | PC | PE | SM | Chol | PS | $PIP_2$ | $PIP_3$ |
|---|---|---|---|---|---|---|---|
| **Outer leaflet** | 50 | 10 | 20 | 20 | - | - | - |
| **Inner leaflet** | 10 | 40 | - | 20 | 20 | 8 | 2 |

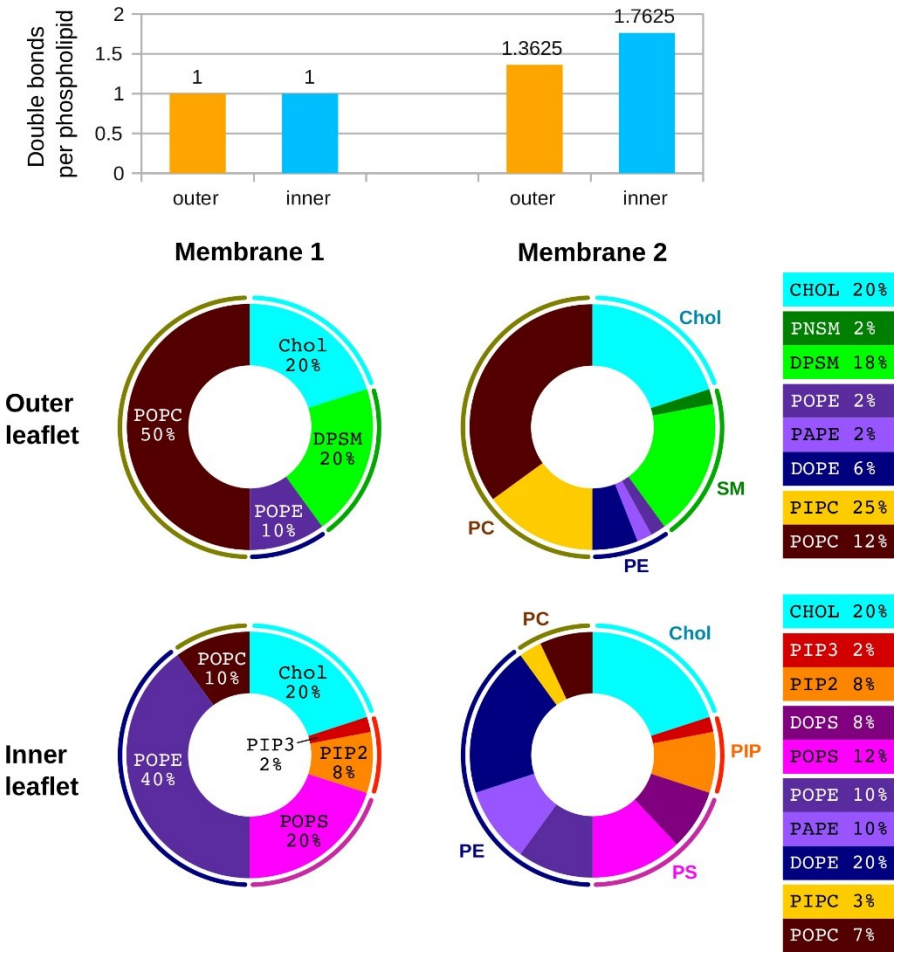


***Fig 5.2 Unsaturation levels of membranes 1 and 2*** *(top; histogram) and their lipid compositions (below; pie charts). Numbers of each lipid type in membrane 1 and 2 are given in Table 5.3.*

Membrane 1 contained mono-unsaturated lipid acyl chains or saturated acyl chains whereas membrane 2 contained a wider variety and a higher percentage of poly-unsaturated lipid acyl chains (Fig 5.2) which was also derived from the

same lipidomics study. CG waters were added, and the system was neutralised with 0.15M of Na+ and Cl- ions. They were energy minimised using the steepest descent algorithm followed by equilibration with the protein backbone particles position-restrained for 2.5 ns. The equilibrated system was used to generate systems with differing initial velocities for five production simulations run for 5 μs each with 20 fs time-step. Equilibration and production simulations were conducted using the NPT ensemble. Every frame in each production simulation was generated at 200 ps intervals. The semi-isotropic barostat and thermostat used for CGMD production simulations were Parrinello-Rahman (1 bar) [225] and V-rescale (323 K) [224], respectively.

### 5.2.3 Atomistic molecular dynamics simulations

ATMD simulations (100 ns x 3 replicates) were initially performed starting from the same initial model as CGMD simulations to confirm the coiled conformation of the TCR-CD3 CYR. This was set up using CHARMM-GUI [226] using the membrane 1 composition, TIP3P water model, and neutralised with 0.15M Na+ and Cl- ions. Energy minimisation using the steepest descent algorithm was conducted until forces converged to 1000 kJ/mol/nm followed by a 6-step semi-isotropic equilibration at 323 K where position restrains in the system were gradually released. Production simulations were run for 100 ns x 3 replicates using the Nose-Hoover thermostat (323 K) and Parrinello-Rahman semi-isotropic barostat (1 bar) with a compressibility of $4.5\times10^{-5}$ $bar^{-1}$ and frames were generated every 100 ps. The LINCS algorithm was used to constrain hydrogen bond lengths and the Particle Mesh Ewald algorithm defined long-range electrostatics.

A frequently observed conformation in which (i) $\zeta_1$ and CD3ε(δ) TMRs had undergone loosening and (ii) ITAM Y177 of CD3ε(γ) was inserted in the membrane, was extracted from the end of a CGMD simulation and backmapped [205] to AT resolution. Gromacs 2016 with the CHARMM36 forcefield [214] was used for backmapping. Lipid parameters were obtained from CHARMM-GUI [226,283]. Before performing ATMD, the simulation box size was reduced along the vertical (Z) axis to prevent simulating excess solvent particles, thus minimising the computational cost. The protein-lipid conformation was retained during the resizing of Z axis of the simulation box. The TIP3P water model was used along with 0.15M of Na+ and Cl- ions to neutralise the system. The system was energy minimised with the steepest descent algorithm followed by a step-wise (0.25→1→1.5→1.8→1.9→2 fs time-step) NPT equilibration with the protein backbone position-restrained. The final equilibration step was conducted

for 2 ns with a 2 fs time-step. The equilibrated system was then used to generate systems with different initial velocities for three repeat production simulations each of which were run for 250 ns with a 2 fs time-step. Every frame in each simulation was generated at 40 ps intervals. The V-rescale thermostat (323 K) and Parrinello-Rahman semi-isotropic barostat (1 bar) [225] was used with a compressibility of $4.5\times10^{-5}$ $bar^{-1}$. The LINCS algorithm [219] applied constraints on all bond lengths and the Particle Mesh Ewald algorithm [216] defined long-range electrostatics. A summary of simulations is shown in Table 5.2. The number of each lipid type used in my simulations is shown in Table 5.3.

***Table 5.2 Summary of simulations conducted in this chapter.***

| **Simulations** | **Resolution** | **Particles** | **Simulation box (X × Y × Z) (nm)** | **Duration x replicas** |
|---|---|---|---|---|
| TMO (memb 1) | CG | 26728 | 17 × 17 × 10 | 5 μs × 5 |
| ECTM (memb 1) | CG | 54106 | 17 × 17 × 21 | 5 μs × 5 |
| FL (memb 1) | CG | 196303 | 25 × 25 × 38 | 5 μs × 5 |
| FL (memb 1) | AT | 1517173 | 20 × 20 × 37 | 100 ns × 3 |
| FL (memb 1) | AT (backmapped) | 1145375 | 25 × 25 × 21 | 250 ns × 3 |

***Table 5.3 Number of lipids of each type and of CG waters used in simulations.***

| | **TMO / ECTM (XY plane = 17 nm$^2$) (membrane 1)** | | **FL (XY plane = 25 nm$^2$) (membrane 1)** | | **FL (XY plane = 25 nm$^2$) (membrane 2)** | |
|---|---|---|---|---|---|---|
| | Outer | Inner | Outer | Inner | Outer | Inner |
| POPC | 254 | 50 | 502 | 99 | 351 | 69 |
| PIPC | - | - | - | - | 150 | 29 |
| POPE | 50 | 201 | 100 | 399 | 20 | 99 |
| PAPE | - | - | - | - | 20 | 99 |
| DOPE | - | - | - | - | 60 | 199 |
| DPSM | 101 | - | 201 | - | 180 | - |
| PNSM | - | - | - | - | 20 | - |
| Chol | 101 | 100 | 201 | 199 | 200 | 199 |
| POPS | - | 100 | - | 199 | - | 119 |
| DOPS | - | - | - | - | - | 79 |
| $PIP_2$ | - | 40 | - | 79 | - | 79 |
| $PIP_3$ | - | 10 | - | 19 | - | 19 |
| CG waters | TMO = 13799<br>ECTM = 38666 | | 165366 | | 166370 | |

## 5.2.4 Analysis

Calculation of all protein-lipid, protein-protein and protein-solvent contacts in the CGMD simulations used a 0.55 nm distance cut-off to define a contact. The same cut-off value was used to calculate contacts of protein residues with hydrophobic lipid acyl chains. Similarly, all contact analyses for ATMD simulations used a 0.4 nm cut-off. All interaction profiles represent merged data from all simulation repeats and were performed using *gmx mindist* command and in-house python scripts. Residence times of lipids were calculated using the PyLipID tool (https://doi.org/10.5281/zenodo.4999404) [284]. To calculate clusters of TCR-CD3 cytoplasmic conformations, and lipid densities around the protein, the trajectories of all simulation repeats were concatenated using *gmx trjcat* and the protein orientation was fixed in the centre using *gmx trjconv*. The *gmx densmap* and *gmx xpm2ps* commands were used to produce lipid density

images. The clustering analysis of the cytoplasmic conformations was performed using the *gmx cluster* command using the single linkage method. Clustering analysis was done using a 0.35 nm cut-off for the protein backbone RMSD for every 10$^{th}$ frame i.e. 400 ps. RMSDs and RMSFs were calculated with *gmx rms* and *gmx rmsf* respectively. Distance and radius of gyration analyses were performed using the *gmx distance* and *gmx gyrate* commands respectively. Radial distribution functions were calculated using *gmx rdf*. VMD was used for visualisation and rendering. The RMSD trajectory tool of VMD [285] was also used to perform alignments of TMR helices. The electrostatic potential (kT/e) of the entire TCR-CD3 was obtained using the APBS electrostatics tool [286] integrated with VMD, whose input was provided by the PDB2PQR tool [287]. The MUSCLE tool [288] was used to perform multiple sequence alignments. Calculation of the relative height of lipid nitrogen atoms was based on their positions along the vertical (Z) axis i.e. perpendicular to the membrane. For this calculation, the TCR-CD3 orientation was fixed in the centre in each ATMD simulation before concatenating all of them. The python script to perform this calculation was obtained from https://github.com/jiehanchong/membrane-depth-analysis. The *gmx do_dssp* command was used to calculate the secondary structure formations in the TCR-CD3 CYRs considering 1 ns intervals from all ATMD simulations combined. XmGrace (https://plasma-gate.weizmann.ac.il/Grace/) and Matplotlib 3.3 (https://doi.org/10.5281/zenodo.3948793) were used for plotting.

## 5.3 Results

### 5.3.1 Modelling the entire TCR-CD3 complex

The cryo-EM structure of the human TCR-CD3 complex (PDB:6JXR) was used as a template and performed secondary structure predictions with the PSIPRED 4.0 workbench [247] using sequence data from Uniprot (see Methods section 5.2.1) to model the entire complex. The predictions suggested that the cytoplasmic region (CYR) of the CD3δ, γ, ε subunits lacked secondary structure in agreement with circular dichroism experiments [17]. However, the ζ CYR was predicted to contain short α-helices consistent with NMR spectroscopy data [289]. These regions were modelled as helical in the complete TCR-CD3 structure while the rest of the ζ CYR were modelled as unstructured regions. Some extracellular residues in the CD3ε, CD3γ, TCRα subunits that were missing from the cryo-EM structure were also modelled as unstructured regions (Fig 5.1). The CYRs of the CD3 and ζ subunits were modelled in a linear extended orientation perpendicular to the membrane to avoid bias in inter-

subunit contacts at the beginning of the simulations. During modelling, the TMRs and the ECDs of the TCR-CD3 subunits were position-restrained to preserve their experimentally derived structural integrity. Finally, twenty different models including each of their discrete optimised protein energy (DOPE) scores [290] were obtained. The structural models showed minimal differences from one another due to position restrains applied on the tertiary structure of the ECDs and TMRs, and on the predicted α-helices in both the ζ CYRs. The most energetically favourable model (least DOPE score) of the entire TCR-CD3 complex was used and further energy minimised. This atomistic model (Fig 5.3) was converted to a coarse-grained (CG) resolution using the Martini 2.2 forcefield and used for CGMD simulations.

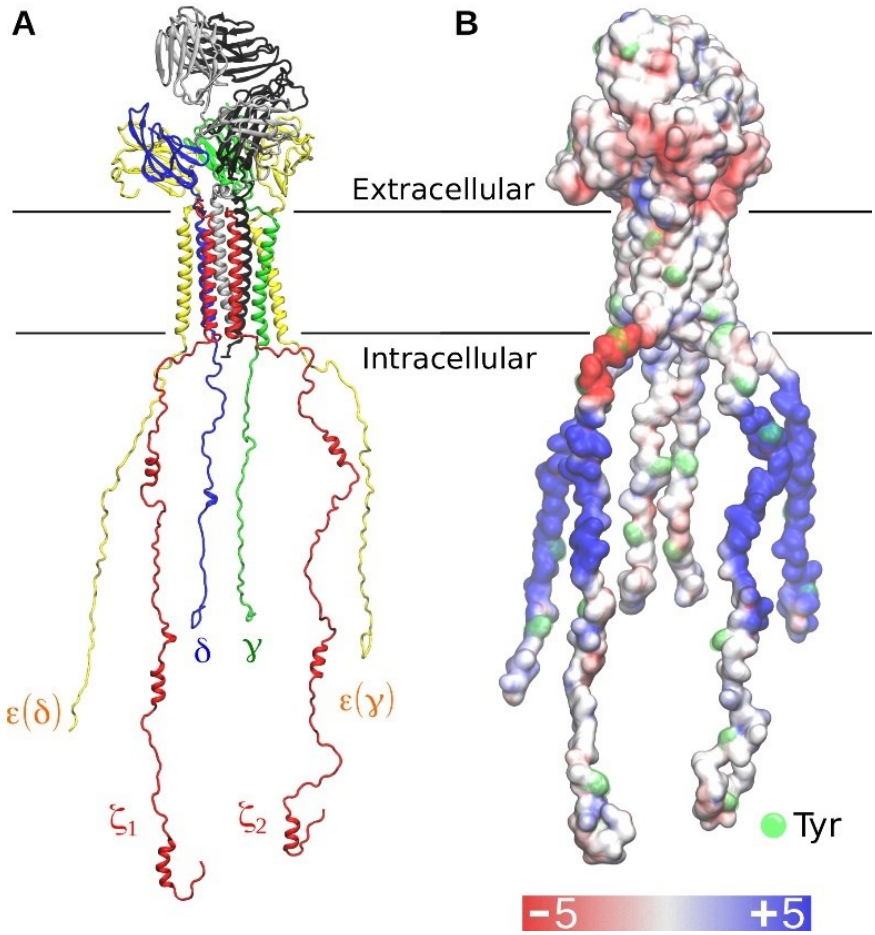


***Fig 5.3 The complete T cell receptor structure model.*** *(A) Model of the entire TCR-CD3 complex used in my simulations, and its (B) electrostatic profile. A range of ±5 kT/e was used to indicate the electronegative and electropositive regions as shown in red and blue respectively. The calculation of the electrostatic profile was done using the APBS tool* [286].

### 5.3.2 The TCR-CD3 cytoplasmic region exhibits a coiled conformation

The CGMD simulations of the entire TCR-CD3 complex were used to analyse its dynamic nature when embedded in a lipid bilayer comprised of the predominant lipid headgroup types found in the TCR-CD3 activation domain (Table 5.1) [210]. Five independent CGMD simulations of the TCR-CD3 were performed in an asymmetric complex membrane for 5 μs each. During the simulations, the CYRs of ζζ and CD3 dimers that were initially modelled in an extended configuration, rapidly coiled forming inter-chain interactions and then associated with membrane lipids of the inner leaflet (Fig 5.4A). Calculation of the distance between the centre of mass (COM) of the CYRs and the COM of the lipid bilayer as a function of time suggested that the association of the CYRs with the membrane occurred within the first 100 ns of the simulations. Although the time taken for the CYRs to coil and associate with the membrane was consistent amongst all CGMD simulations, their radius of gyration varied indicating that the coiled conformation of the CYRs is dynamic. To confirm the propensity of the CYR of TCR-CD3 to form coiled structures and to associate with the membrane, ATMD simulations were also performed for 100 ns starting from the same initial model (with cytoplasmic tails in an extended conformation perpendicular to the membrane). When compared with CGMD simulations, these simulations confirmed the observation that TCR-CD3 CYR can form coiled structures and associate with the membrane (Fig 5.4B-D).

Following the observation of the fluctuation in the CYR assembly, its most commonly occurring structural conformation was determined. Therefore, using the coiled and membrane-bound state of the CYRs (from the last 4 μs of the CGMD simulations), 10,000 cytoplasmic configurations were extracted from all simulations combined and grouped into clusters using a 3.5 Å RMSD cut-off (RMSD calculated relative to the initial model shown in Fig 5.3). The largest cluster contained 867 similar structures (Fig 5.4E) representing the most frequent structural conformation of the CYRs across all simulations (shown in a box in Fig 5.4A).

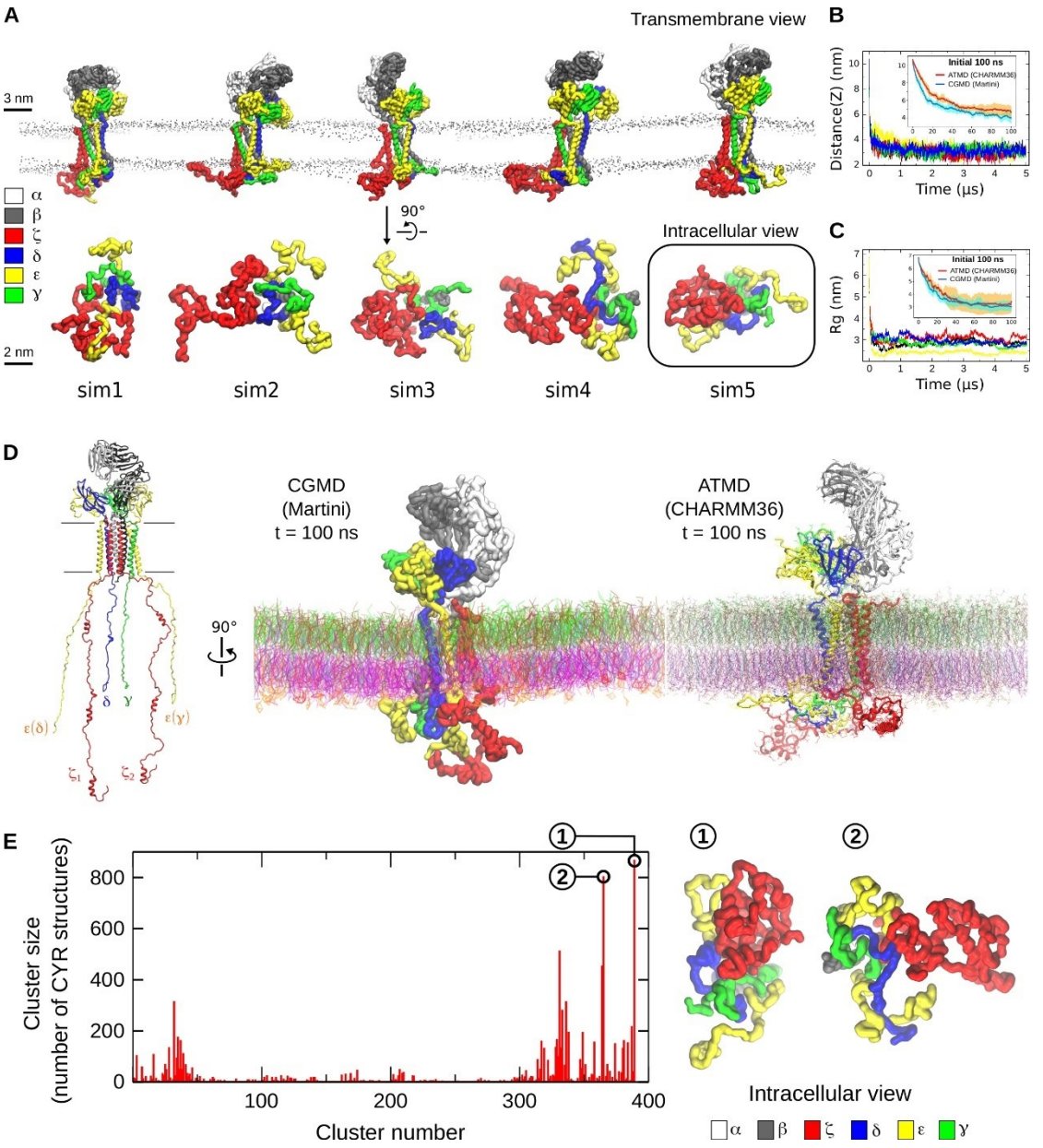


***Fig 5.4 Association of the TCR-CD3 CYR with the membrane.*** *(A) Transmembrane / side view of the entire TCR-CD3 (top) and intracellular view of its cytoplasmic region (bottom) taken from the end of the five CGMD simulations of membrane 1. The box highlights the most common CYR conformation calculated by the clustering analysis. (B) Distances between the centre of mass (COM) of the TCR-CD3 CYR and the COM of membrane 1 calculated from the five CGMD simulations along the vertical (Z) axis versus time. (C) Radius of gyration (Rg) of the TCR-CD3 CYR versus time from all CGMD simulations. In (B) and (C), distance(Z) and Rg were also calculated from ATMD simulations performed for 100 ns. Comparison between the average distance(Z) and Rg between the ATMD (red) and CGMD (blue) simulations is shown in the inset graph. Standard deviation is shown in cyan for CGMD and in orange for ATMD. (D) Snapshot from one of the CGMD and ATMD simulations at simulation time (t) = 100 ns, starting from the same initial model (shown on the left). (E) TCR-CD3 cytoplasmic conformations grouped into clusters using a 3.5 Å RMSD cut-off. The clusters containing the highest number of structures indicate the most stable conformation of the TCR-CD3 cytoplasmic region in my simulations.*

### 5.3.3 Membrane penetration by ITAM tyrosines

Previous NMR studies of the CD3ε ITAM-containing peptide interacting with a lipid micelle suggested that two ITAM tyrosines, one isoleucine and one leucine residue penetrated the hydrophobic core of the membrane [33]. To investigate whether my simulations showed similar membrane-penetrating activity of the ITAM tyrosines, their interactions with the hydrophobic acyl chains of the lipids were calculated from the CGMD simulations. It was found that membrane penetration was only achieved by some ITAM tyrosines. In all simulations combined, the tyrosines that displayed the most membrane-penetrating capabilities belonged to the CD3ε and ζ subunits only. Y177 of CD3ε pairing with γ, referred to as CD3ε(γ), made the highest number of contacts with the lipid acyl chains, followed by Y177 of CD3ε(δ). Y166 of CD3ε(γ) penetrated the membrane more than Y166 of CD3ε(δ). In addition, only one of the subunits of the ζζ dimer mostly showed ITAM tyrosine contacts with lipid acyl chains. The contacts of the entire TCR-CD3 subunits with the lipid acyl chains were also calculated (Fig 5.5A). Interestingly, the short extracellular segments of both ζ subunits contacted the acyl chains of lipids in the outer membrane leaflet suggesting their tendency to anchor onto the extracellular leaflet during the simulations. In these simulations, mono-unsaturated phospholipids were used.

Given the finding that tyrosine residues on ITAMs penetrate the membrane, it is possible that a higher unsaturation of the phospholipid acyl chains may facilitate tyrosine penetration even further. Therefore, as a control, additional simulations were conducted in which the TCR-CD3 was inserted in a bilayer containing the same concentration of lipid headgroups but the degree of lipid tail unsaturation was increased. The new membrane composition is shown in Fig 5.2. The increase in unsaturation resulted in similar tyrosine penetration with a somewhat higher penetration of some tyrosines of the $\zeta_2$ and the ε(δ) CYRs (Fig 5.5A). This observation augments the observation that tyrosine residues are able to penetrate the bilayer when the TCR-CD3 is in a steady state. The change in TMR tilt angle for each of the subunits was calculated to assess the influence of membrane unsaturation. It was observed that the range of tilt angles for all TMRs were similar but with minor differences in the peak of the distribution of the tilt angles (Fig 5.5B).

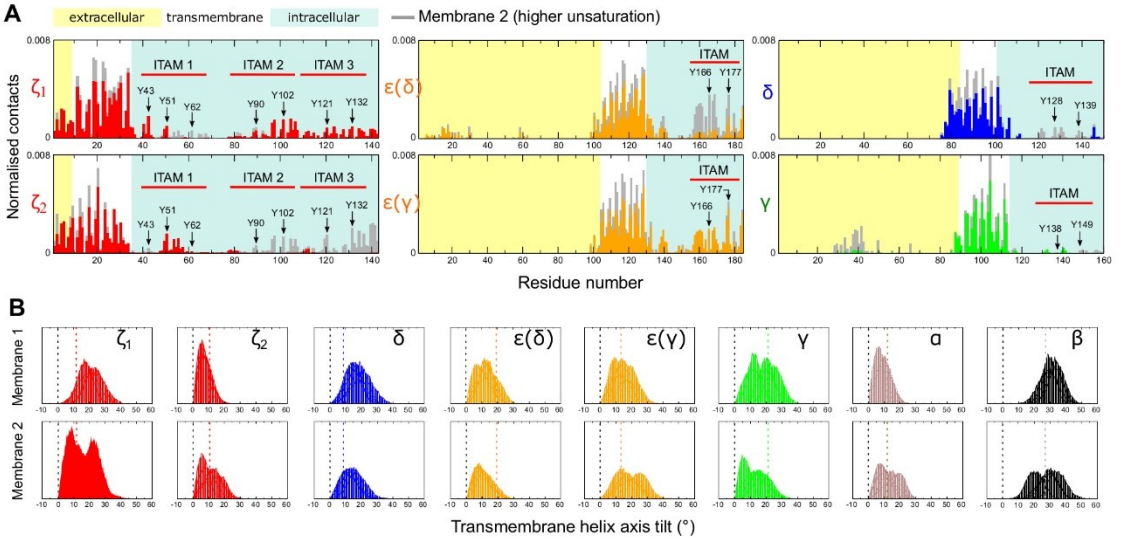


***Fig 5.5 Interactions of the ITAMs with the hydrophobic region of the membrane and transmembrane helix tilt angles.*** *(A) Normalised number of contacts of ζ and CD3 subunits with lipid acyl chains in membrane 1 (coloured by subunit) and membrane 2 (grey). The position of the ITAM tyrosines is indicated by arrows. To normalise, the number of contacts of each residue was divided by the number of lipids in the membrane and the number of simulation frames. Contacts of residues with membrane 1 are coloured by subunit whereas those with membrane 2 are shown in grey. (B) Transmembrane helix tilt angle distribution of each subunit calculated from all simulations from membrane 1 (top row) and membrane 2 (bottom row).*

### 5.3.4 The TCR-CD3 selectively interacts with lipid headgroups

It was questionable whether the association of the CYRs onto the inner leaflet of the bilayer affected the TCR-CD3 lipid environment. From all CGMD simulations combined, the contacts of the entire TCR-CD3 complex with all lipid headgroups including the sterol in both leaflets of the membrane were analysed. Further, the number of interactions of each lipid-type were normalised by their respective concentrations in membrane 1. These data showed the relative enrichment of certain lipids over others in the vicinity of the TCR-CD3 complex. The TCR-CD3 showed a high propensity to contact phosphatidylinositol 4,5-biphosphate ($PIP_2$) and phosphatidylinositol 3,4,5-triphosphate ($PIP_3$) (Fig 5.6A, B) despite their relatively low abundance in the inner leaflet (8% and 2% respectively). A closer inspection revealed that the cationic residues dispersed throughout the CYRs of CD3ε and ζ subunits dominated the interaction with $PIP_2$ and $PIP_3$ lipids. The CD3δ and CD3γ CYRs also contacted the PIP lipid headgroups, though to a lower extent. In the CD3ε and ζ subunits, the basic residue-rich stretches (BRS) interacted most with PIPs, while the poly-proline motifs (PPPVPNP) in both CD3ε subunits (labelled 'PPP' in Fig 5.6A) showed the least contact. Basic residues at the TMR-CYR

interface of the ζ subunits (R31, K33, R36) also contacted PIP lipids. Other anionic lipid headgroups i.e. POPS, and also neutral lipids made contacts with the CD3 and ζ CYRs but less frequently when compared with PIP lipids (Fig 5.6A, B).

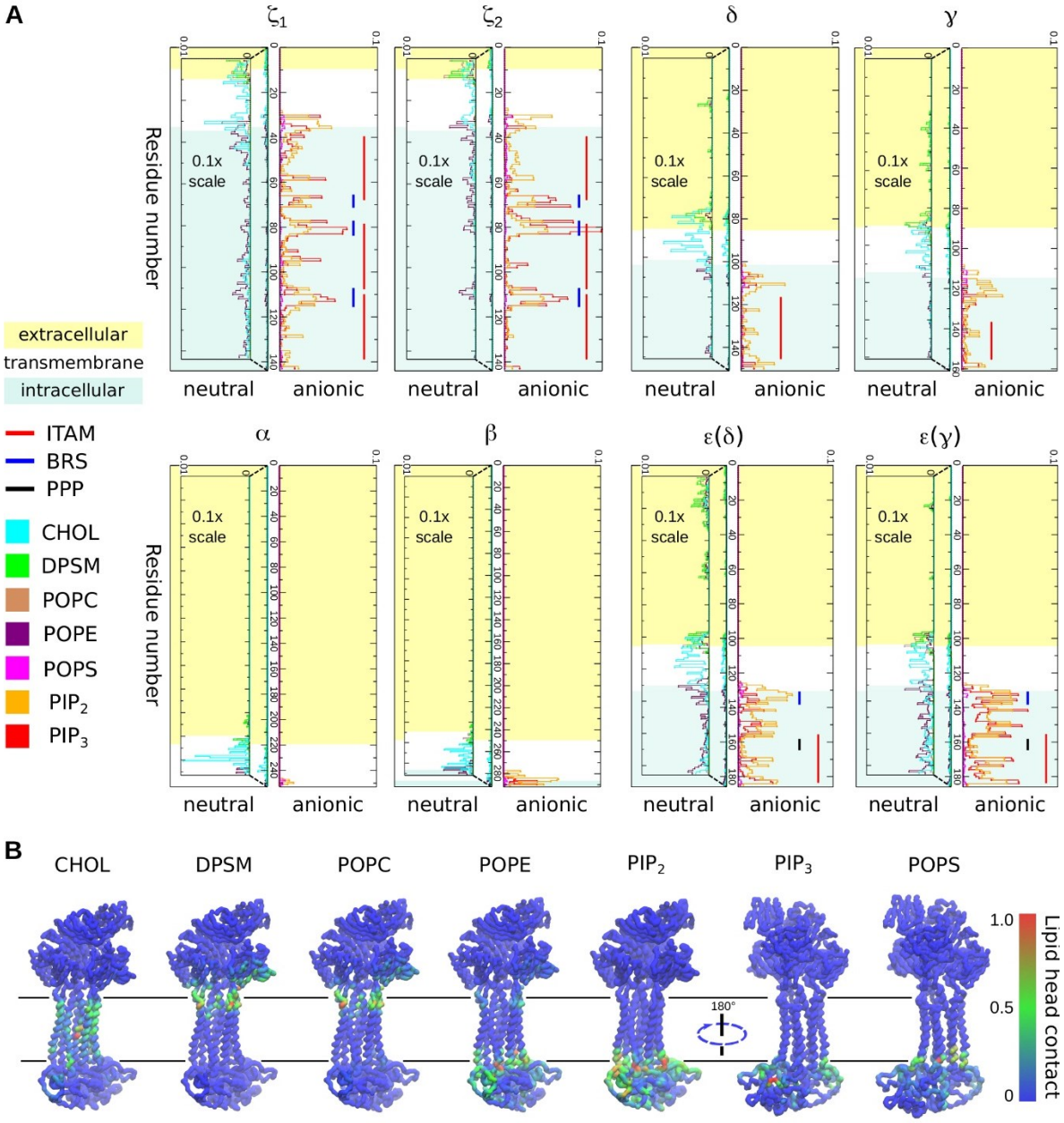

***Fig 5.6 Interactions of the complete TCR-CD3 with lipid headgroups.*** *(A) Normalised number of contacts of each full-length TCR-CD3 subunit with lipid headgroups. The normalisation (N) is done by dividing the number of lipid contacts of each residue (n1) by the number of the specific lipid in the bilayer (n2) and by the number of simulation frames (n3) i.e. N=n1/n2/n3. Interactions of each TCR-CD3 subunit with anionic lipids are shown on the right whereas those with neutral lipids are shown on the left. The scale of non-anionic interactions was magnified 10 times (scale: 0 to 0.1) for clarity. The ITAMs (red lines) of ζ and all CD3 subunits, the BRS motifs (blue lines) of CD3ε and ζ subunits, and poly-proline motifs (black lines) of the CD3ε subunits are also indicated. (B) Contacts of the TCR-CD3 with each lipid headgroup type are normalised separately on a scale of 0 to 1 and mapped as a colour gradient (blue: low, green: medium, red: high) on the TCR-CD3 structure extracted from the end of a simulation.*

Cholesterol interactions occurred across all the TMRs of the TCR-CD3, and were accompanied by minor interactions with CYRs of CD3ε and ζ (Fig 5.6A), explained by their ability to penetrate the membrane's surface. POPE headgroups in the inner membrane leaflet also interacted with the CYRs of CD3ε and ζ to an extent similar to cholesterol interactions (Fig 5.6A, B). Sphingomyelin (DPSM), present only in the outer leaflet, interacted with all subunits but mostly contacted the extracellular segment of ζ and the ECD of CD3ε(δ), followed by the other CD3 subunits. It was also observed that POPC headgroups in the outer leaflet interacted with the extracellular segment of ζ and with the connecting peptides (CPs) of CD3 subunits (Fig 5.6A, B). While in agreement with previous studies suggesting ionic interactions of CYRs with the plasma membrane [32], data shown here points to a potentially more complex scenario of contacts of the TCR-CD3 with membrane lipid headgroups than previously suggested.

### 5.3.5 PIP clustering and the significance of the cytoplasmic region

Investigation of PIP interactions with TCR-CD3 CYRs suggested clustering of PIPs around the TCR-CD3 complex. This led to the analysis of the densities of each lipid type around the TCR-CD3. In the outer leaflet, there was no clustering of POPE, POPC, and DPSM, but in the inner leaflet the densities of PIP lipids around the protein dominated that of other lipids. To better discern the origin of this effect, the contribution of the ECDs, TMRs and CYRs of the TCR-CD3 toward lipid interaction was quantified. In addition to the CGMD simulations of the full TCR-CD3 complex (referred here as FL), two sets of CGMD simulations were performed (5 simulations x 5 μs each) using the lipid composition of membrane 1: (i) excluding the ECDs and CYRs, and retained

only the TMRs (referred here as TMO), (ii) excluding only the CYRs, and retained the ECDs and TMRs (referred here as ECTM) (Fig 5.7A). The densities of lipids combining all five repeats of the TMO and ECTM simulations were calculated and compared to the densities retrieved from FL simulations. The most striking observation in all three conditions was the clustering of PIP lipids. To assess the influence of poly-unsaturated lipid acyl chains on lipid clustering, the densities of lipids around TCR-FL in membrane 2 were compared to the densities around TCR-FL in membrane 1 and found that they were similar (Fig 5.7B). The radial distribution functions of all lipid types in membrane 1 and 2 were also calculated, and it was observed that PIP lipids interacted the most with the TCR-CD3 irrespective of the difference in lipid unsaturation level (Fig 5.7C).

A distinct cholesterol annulus was clearly observed in the TMO simulations concentrated on the opposite side of the ζζ dimer, suggesting it to be a cholesterol-binding hotspot. Cholesterol bound to similar sites in the ECTM and FL simulations was also observed (Fig 5.7B). It was also found that the average number of cholesterol molecules interacting with the TCR-CD3 were consistent in all three conditions (Fig 5.7D), suggesting that cholesterol binds to the TMRs independent of the ECDs or CYRs. Similarly, the change in the average number of PIP lipids contacting the TCR-CD3 across time was quantified. This revealed a three-fold increase in the number of one or both $PIP_2$/$PIP_3$ lipids in the FL simulations compared to the TMO and ECTM simulations (Fig 5.7E), suggesting that the CYRs played a significant role in enhancing PIP clustering. The number of interacting POPE lipids with the TCR-CD3 also increased in presence of the CYRs while the number of interacting POPS lipids showed a minor increase. There were no differences in the average number of interacting POPC and DPSM lipids (Fig 5.7D) when comparing the FL simulations to TMO and ECTM simulations. Calculation of the lipid residence times shows that on average, $PIP_2$ lipids spent the most amount of time bound to the protein without detaching followed by cholesterol, $PIP_3$ and PS, and other lipids. This order was maintained in membrane 2 where the unsaturation of lipid acyl chains in the inner leaflet was higher than that of the inner leaflet in membrane 1 (Fig 5.7F).

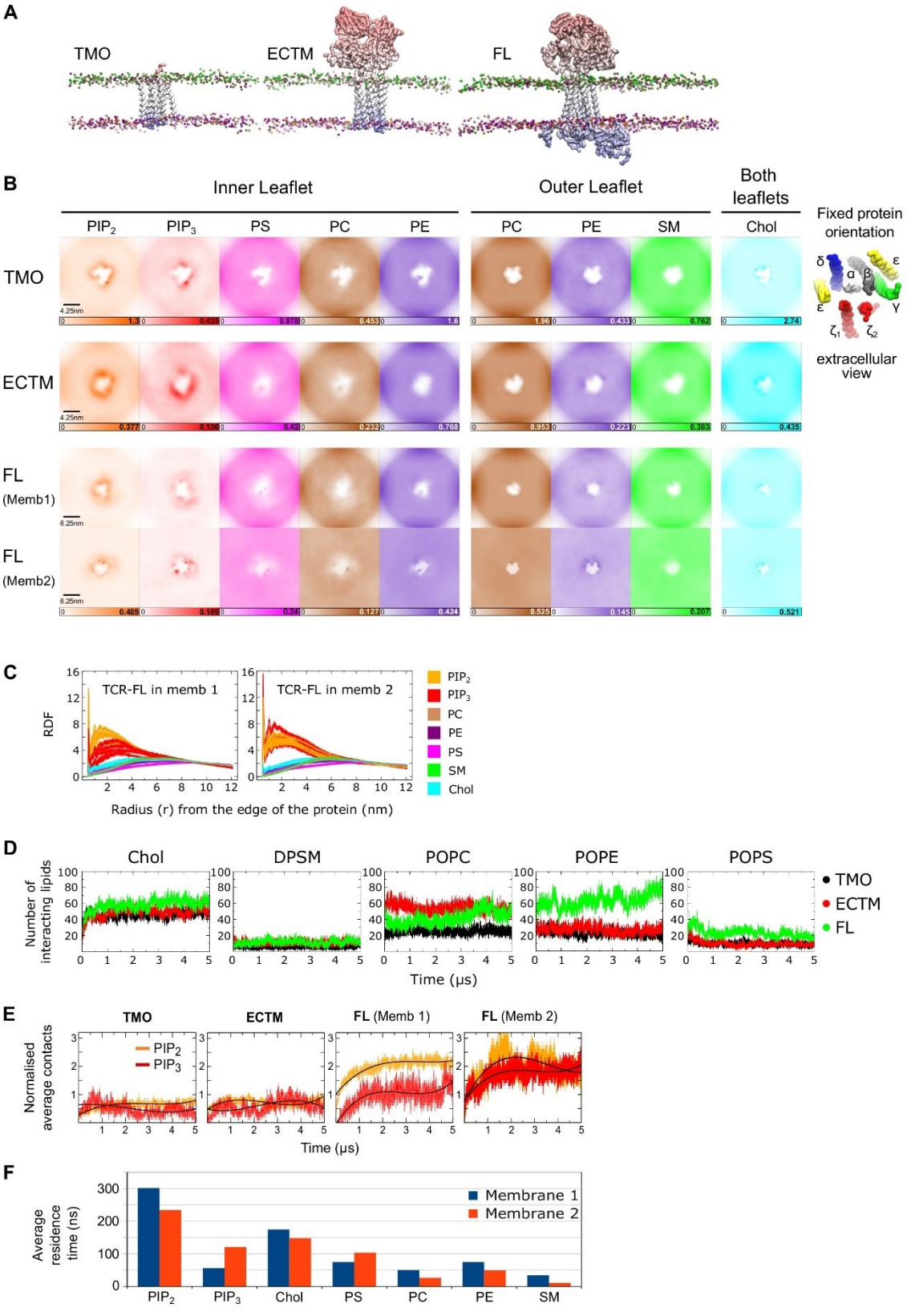
A
TMO
ECTM
FL
B
Inner Leaflet
Outer Leaflet
Both leaflets
PIP2
PIP3
PS
PC
PE
PC
PE
SM
Chol
Fixed protein orientation
extracellular view
TMO
ECTM
FL (Memb1)
FL (Memb2)
C
TCR-FL in memb 1
TCR-FL in memb 2
RDF
Radius (r) from the edge of the protein (nm)
PIP2
PIP3
PC
PE
PS
SM
Chol
D
Chol
DPSM
POPC
POPE
POPS
Number of interacting lipids
Time (μs)
TMO
ECTM
FL
E
TMO
ECTM
FL (Memb 1)
FL (Memb 2)
Normalised average contacts
PIP2
PIP3
Time (μs)
F
Average residence time (ns)
Membrane 1
Membrane 2
PIP2
PIP3
Chol
PS
PC
PE
SM

***Fig 5.7 Lipid densities and clustering of PIP lipids around the TCR-CD3.*** *(A) Snapshots from the three sets of CGMD simulations (TMO, ECTM, FL) performed in this chapter. Here, only the protein and the lipid phosphate groups are shown for clarity. (B) Extracellular view of the densities of all lipids in the XY plane of the membrane from all five CGMD simulations combined. The protein was fixed in the centre and its TMR orientation is shown for reference. (C) Radial distribution functions of all lipid types in membrane 1 and 2. The different lines for each lipid type represent the RDF for the five repeat simulations that I performed for each system. (D) The average number of cholesterol, DPSM, POPC, POPE, and POPS lipids interacting with the TCR-CD3 over 5 μs time from all CGMD simulations of TMO, ECTM, FL systems conducted in membrane 1. (E) Normalised average number of $PIP_2$ and $PIP_3$ lipids contacting the protein in all CGMD simulations (TMO, ECTM, FL) versus time. Normalisation was done by dividing the number of $PIP_2$ and $PIP_3$ lipid contacts by their respective number in the membrane. The smoothened black lines are a cubic regression of the number of PIP contacts across time. (F) Average residence time of all lipid types in membrane 1 and 2.*

It was still uncertain as to why there were some anionic lipid headgroups contacting the TCR-CD3 in the ECTM simulations despite the absence of the CYRs. This was due to the cationic residues present at the juxtamembrane region of TCR-CD3 (Fig 5.8A, B), collectively termed here as the ‘cationic anchor’. A multiple sequence alignment of the TMRs and the juxtamembrane region indicates that the cationic anchor is conserved across various species (Fig 5.8C). By comparing the ECTM to FL simulations, the interaction of the TCR-CD3 with anionic lipid headgroups in the absence of its CYRs potentially suggest that the cationic anchor can alone retain approximately one-third of anionic lipids around the TCR-CD3.

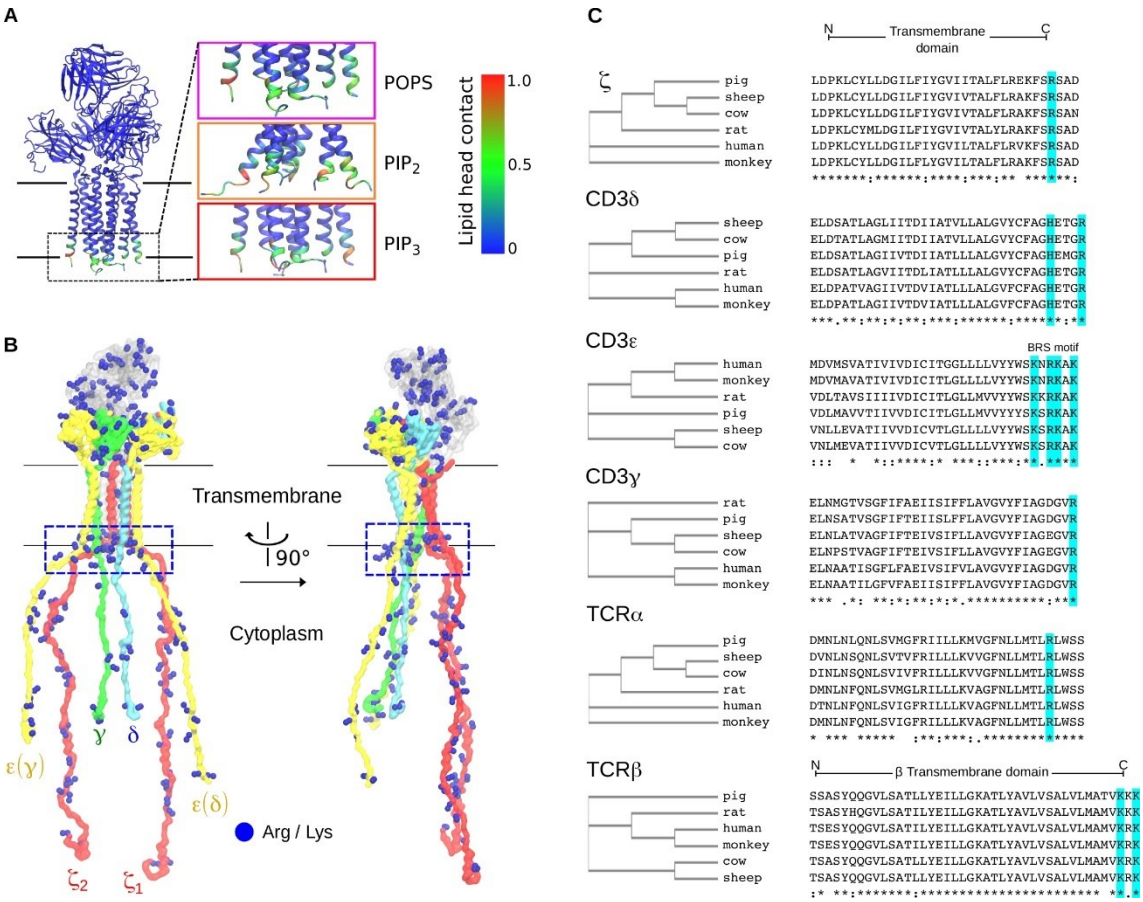


***Fig 5.8 The cationic anchor of the TCR-CD3 complex.*** *(A) Interaction of the TCR-CD3 cationic anchor with anionic lipids mapped onto the ECTM structure. (B) The cationic anchor of the TCR-CD3 located at the interface of the TMR and CYR is indicated within a box. (C) Multiple sequence alignment of the TMR and juxtamembrane residues of all TCR-CD3 subunits indicates the conservation of cationic residues at the TMR-CYR interface across different species i.e. Homo sapiens (humans), Rattus norvegicus (rat), Sus Scrofa (pig), Bos taurus (cow), Ovis aries (sheep), and Macaca mulatta (monkey).*

### 5.3.6 Conformational changes and inter-chain interactions in the TCR-CD3

To determine the conformational changes occurring within the ECD, TMR, and CYRs of the complete TCR-CD3, the backbone RMSD of each of these regions were calculated. RMSD analysis of the ECD backbone during the simulations confirmed its flexibility and also indicated some distinct conformations (Fig 5.9A). The backbone RMSF of all residues in the complete TCR-CD3 were also calculated and mapped onto the structure highlighting the large fluctuations in the CYRs due to their dynamics (Fig 5.9B).

Inspection of the TMRs in the TMO, ECTM, and FL simulations at the end of 5 μs revealed a consistent loosening of the $\zeta_1$ TMR from the CD3ε(δ) TMR (Fig 5.9C). Note that, in the cryo-EM structure (PDB:6JXR), their TMRs are in

contact only at their N-terminal ends near the surface of the outer membrane leaflet. Calculation of the distance between the COM of $\zeta_1$ and of CD3ε(δ) TMRs in the FL simulations showed an increase in the distance between their TMRs (along the XY plane i.e. parallel to the membrane) compared to their initial distance calculated from the cryo-EM structure in 7 out of 10 simulations (Fig 5.9D). In addition, a decrease in the distance between CD3δ and TCRβ TMRs was observed in 9 out of 10 FL simulations (Fig 5.9E).

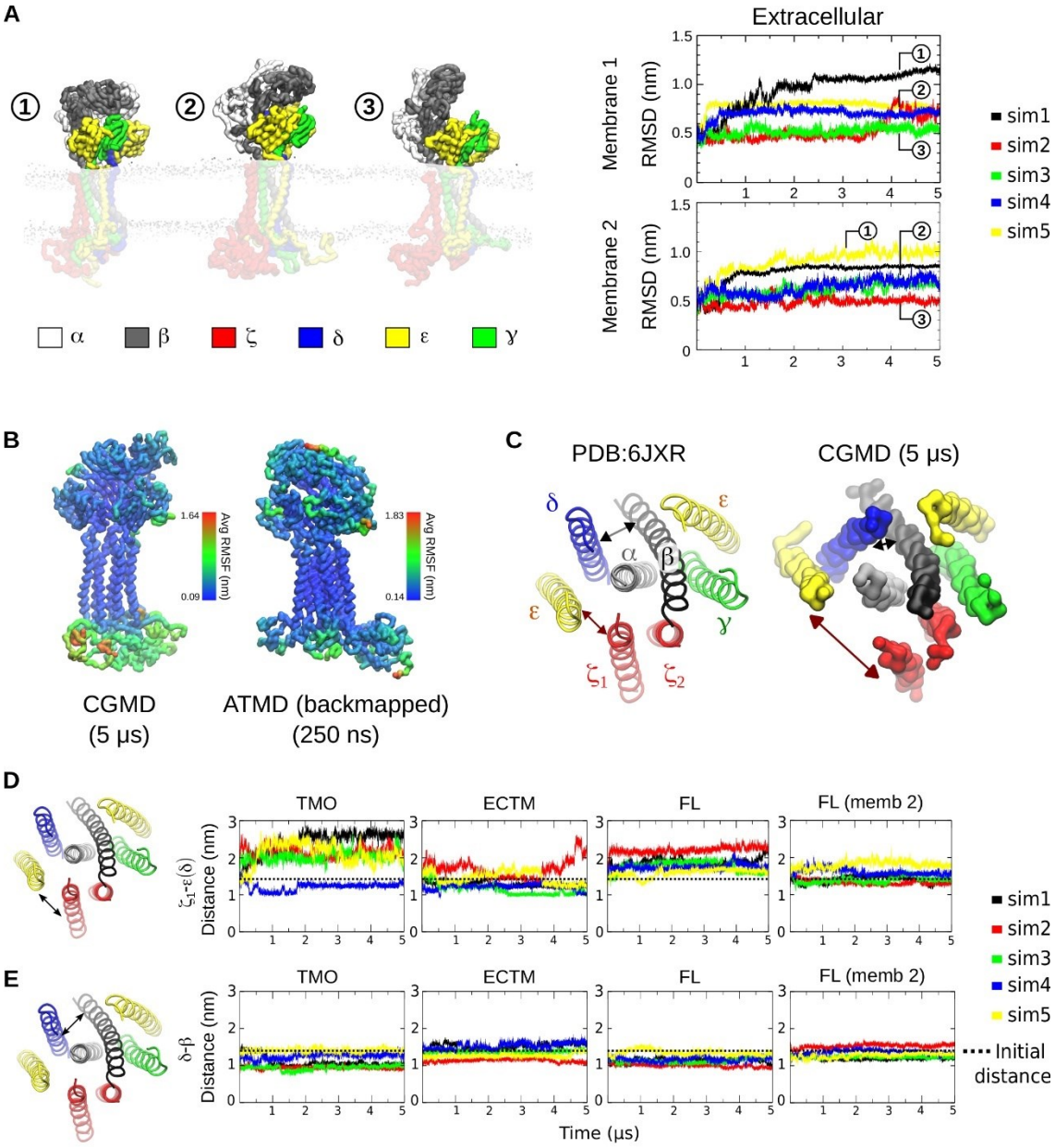

***Fig 5.9 Conformational changes within the TCR-CD3 ECD and TMR.*** *(A) Backbone RMSD of the TCR-CD3 ECD. (B) Backbone RMSF of the complete TCR-CD3 in the CGMD and ATMD simulations. A structure extracted at 1 μs was used a reference for this RMSF calculation. (C) Extracellular view of the TCR-CD3 TMR from the cryo-EM structure (PDB:6JXR) compared to the TMR conformational change seen in the CGMD simulations. (D) Distance between the centre of mass of $\zeta_1$ and of ε(δ) subunits, and (E) distance between the centre of mass of δ and of β subunits in all simulations systems over 5 μs compared to their initial distances calculated from the cryo-EM structure.*

During the simulations, the ζζ CYRs also formed contacts with CD3 CYRs (Fig 5.10). TCRαβ-ζζ interactions occurred mostly in the TMR with some interactions observed in the membrane-proximal extracellular region. In this figure, the interactions of TCRαβ with CD3δε and CD3γε highlight the contacts of the DE loop in the TCRα constant domain and FG loop in the TCRβ constant domain. Although the cryo-EM structure shows that the CD3 ECDs are only contacted by the TCRαβ constant domains, the CGMD simulations revealed additional interactions of the TCRαβ variable domain with the CD3 ECDs indicating some flexibility within the TCR-CD3 ECDs (shown in Fig 5.9A). This flexibility of VαVβ may be required for antigen recognition and initiating allosteric effects onto the rest of the complex [76].

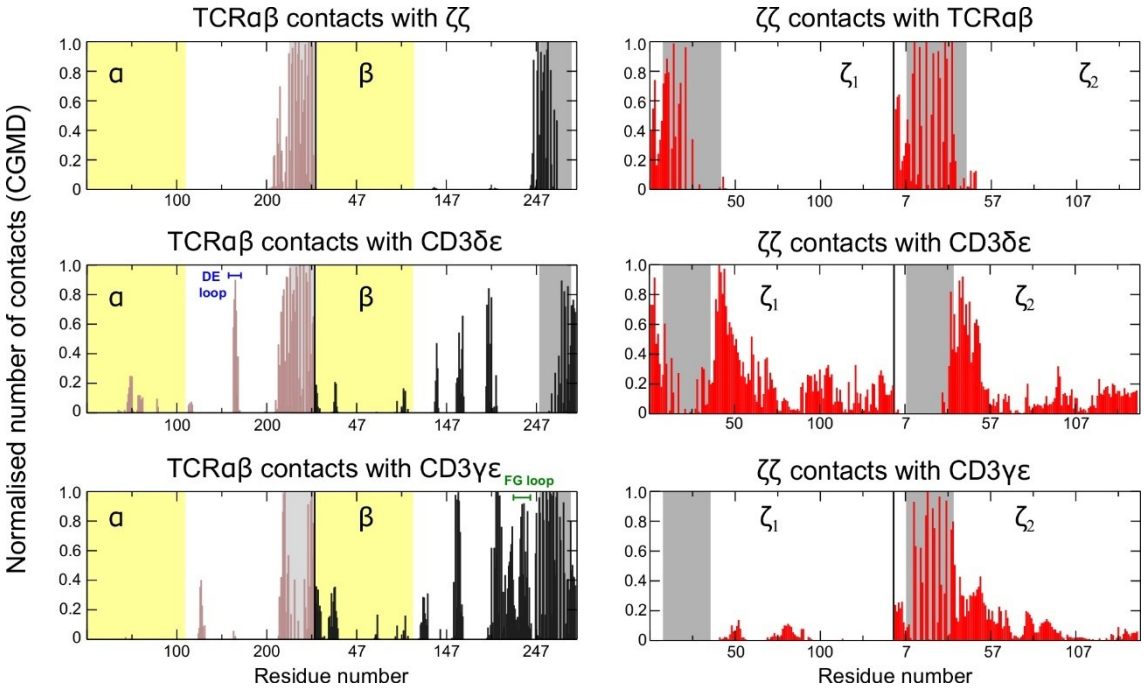


***Fig 5.10 Normalised inter-dimeric interactions*** *of the TCRαβ dimer with the ζζ, CD3δε, CD3γε dimers (left), and of the ζζ dimer with the TCRαβ, CD3δε, CD3γε dimers, in CGMD simulations (right). Normalisation is done by dividing all contacts by the highest number of contacts within the αβ / ζζ dimers.*

### 5.3.7 Atomistic molecular dynamics simulations

The CGMD simulations involved applying an elastic network [233] within each dimer (αβ, δε, γε, ζζ) to maintain their dimeric tertiary/quaternary structures during simulations, hence it restricted conformational changes within the dimers. Therefore, in addition to the ATMD simulations performed for 100 ns using the initial model to confirm the coiling of TCR-CD3 CYRs, ATMD simulations using the final snapshot from one of the five CGMD simulations were also performed where the FL protein was extracted along with membrane 1. This snapshot was converted from CG to AT resolution using the backmapping method [205], and further three replicates of ATMD simulations were performed for 250 ns each, thereby providing information on the protein-protein interactions and TCR-CD3 dynamics in atomistic (AT) detail. The following criteria were considered for backmapping: (i) the most observed TMR configurational change as shown in Fig 5.9C, i.e. where the distance between ε(δ) and $\zeta_1$ TMRs increased, and (ii) penetration of tyrosines into the bilayer, and in particular the penetration of Y177 of ε(γ) subunit that was observed most frequently.

During the ATMD simulations, the protein-lipid interactions observed in CGMD simulations were retained. The loosening between $\zeta_1$ and CD3ε(δ) TMRs, seen in CGMD simulations was also maintained throughout the ATMD simulations suggesting that this conformation may be an energetically favourable configuration for the TCR-CD3. This loosening between $\zeta_1$ and CD3ε(δ) TMRs brought CD3δ TMR closer to TCRβ TMR as observed in CGMD simulations (Fig 5.9C). Although $\zeta_1$ and CD3ε(δ) TMRs made minimal contact, they interacted via their extracellular segments and CYRs. Consistent with the CGMD simulations, $\zeta_1$ made no contact with the CD3γε dimer, while $\zeta_2$ interacted only with CD3γ via its TMR and CYR. Similarly, ATMD simulations also showed the interactions of the DE loop in the TCRα constant domain and of the FG loop in the TCRβ constant domain with CD3δε and CD3γε ECDs respectively. They also revealed new interactions of Vα and Vβ with the ECDs of CD3δε and CD3γε respectively (Fig 5.11A). Furthermore, the ATMD simulations showed that the α-helical regions in the CYRs of ζζ were retained. Additionally, despite starting as unstructured regions, the CD3δ, ε, γ CYRs tended to form short α-helices in all ATMD simulations. CD3δ CYR had the smallest helical regions compared to all other CYRs (Fig 5.11B).

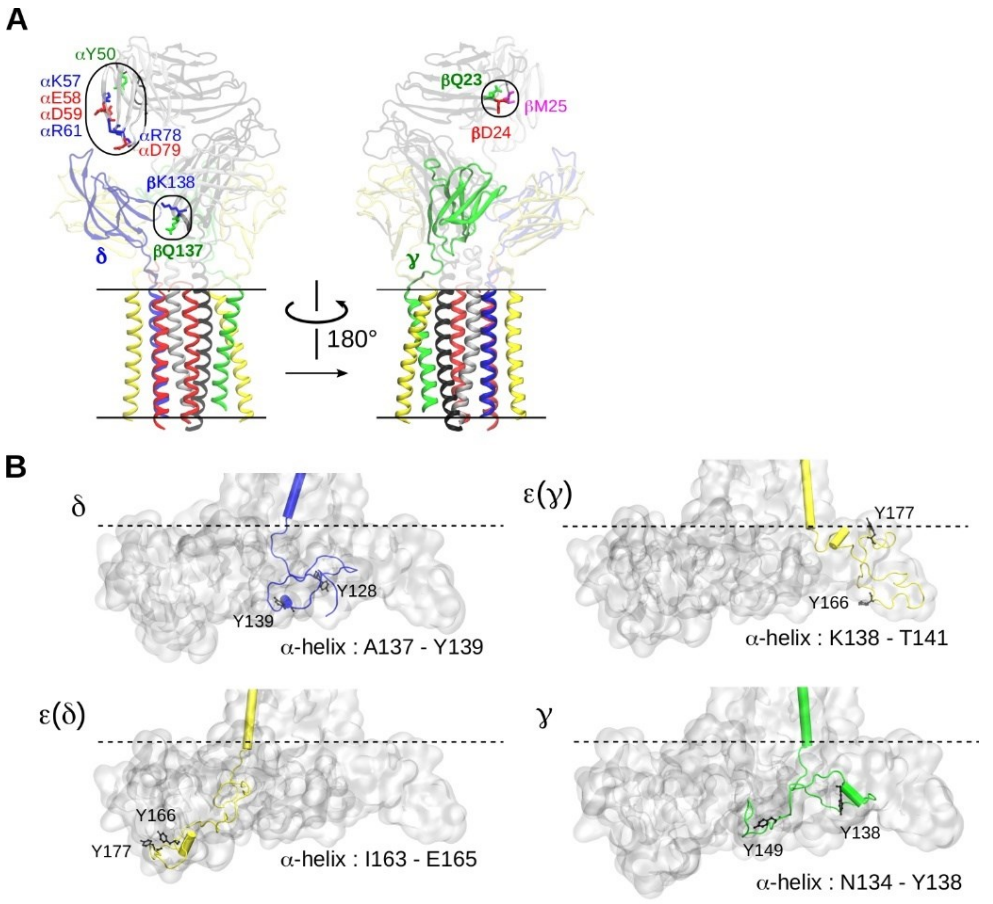


***Fig 5.11 Novel identification interactions in the TCR-CD3 ECD and helix formation in the CD3 CYRs.*** *(A) New interactions of the TCRα variable domain and of TCRβ constant domain with CD3δ (left), and of the TCRβ variable domain with CD3γ (right) seen in the ATMD simulations, compared with the cryo-EM structure (PDB:6JXR). (B) Location of cytoplasmic α-helices relative to the ITAM tyrosines of the respective subunits. The ITAM tyrosines and α-helix forming residues in the CD3 subunits observed at the end of 250 ns in one of the simulations are labelled.*

ATMD of the TCR-CD3 complex in the membrane also revealed that, in the outer membrane leaflet, POPC and POPE headgroups were pulled toward the hydrophobic region of the membrane to ionically interact with anionic residues of the CD3 and ζ TMRs (Fig 5.12A). From all ATMD simulations combined, the number of contacts were calculated between the nitrogen atoms of all outer leaflet lipid headgroups i.e. POPC, POPE, DPSM, and the anionic residues of the CD3 and ζ TMRs. The nitrogen atoms of DPSM made negligible number of contacts, POPC made some contacts, while POPE dominated interaction. The CD3γε TMR induced the most pulling of POPC/POPE headgroups into the membrane in contrast to the CD3δε and ζζ TMRs (Fig 5.12B, C). Moreover, the pulling down of lipid headgroups resulted in extracellular water solvating some charged residues of the TCR-CD3 TMR, thereby leading to a local membrane deformation in the outer membrane leaflet. Previous MD studies proposed that

this occurs prior to TCR-CD3 assembly [57] whereas this study suggests that this occurrence is also a feature of the entire TCR-CD3 complex.

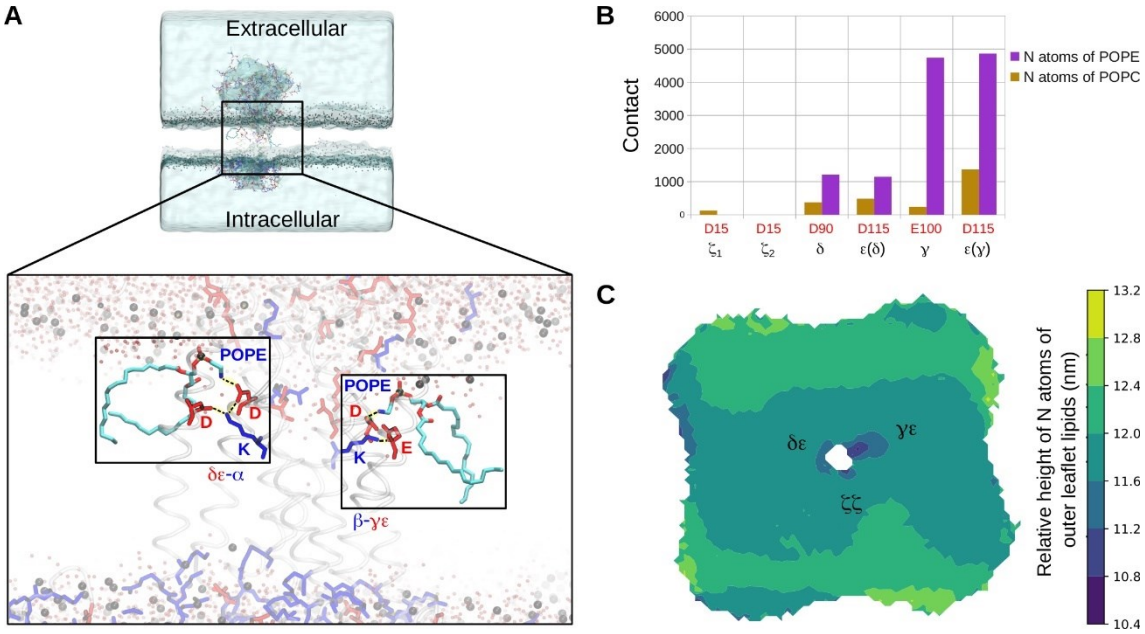


***Fig 5.12 Pulling-in of POPE headgroups toward the anionic sidechains of the TCR-CD3 TMR.*** *(A) Snapshot from the end of one of the ATMD simulations of the entire TCR-CD3. The regions highlighted in boxes show pulled-in POPE headgroups whose amine (NH3) groups interact with the anionic sidechains of CD3 TMRs. The pulled-in POPE lipids are shown in cyan with their oxygen atoms shown in red and nitrogen atoms in blue. Hydrogen atoms are not shown for clarity. Anionic and cationic protein residues are represented as licorice sticks in red and blue, respectively. The TMR helices are shown as transparent cartoons. Water molecules are shown as transparent red spheres and the phosphorus atoms of all phospholipids are shown as black spheres defining the surface of the membrane. (B) Number of contacts of anionic residues of CD3 TMRs with nitrogen (N) atoms of POPC and POPE lipids from all ATMD simulations combined. (C) Extracellular view of the relative height of nitrogen atoms of all outer leaflet lipids (POPC, POPE, DPSM) mapped as a colour gradient from dark blue (low) to yellow (high). The TCR-CD3 TMR orientation is fixed in the centre.*

## 5.4 Discussion

The CYRs of the CD3 and ζ subunits are essential to mediate T cell activation, as pMHC binding induces tyrosine phosphorylation of their ITAMs by LCK and intracellular signal propagation by the ZAP-70 tyrosine kinase [280]. Biochemical and biophysical studies have suggested that the CYRs of the CD3ε and ζ subunits of a non-stimulated TCR-CD3 complex interact with the inner leaflet of the plasma membrane [28,29,33,34]. These data may imply that TCR-CD3 ligation induces dissociation of the CYRs from the inner leaflet of the

membrane. This dissociation allows augmented ITAM access by active LCK [114] followed by sustained ITAM phosphorylation leading to activation of ZAP-70 to propagate intracellular signalling [278]. Recent data indicate that allosterically-induced conformational changes triggered by pMHC binding to TCR-CD3 facilitate ITAM exposure [76]. Upon initial phosphorylation, it is plausible that the increase in the net electronegative charge of the CYRs aid in their dissociation from the membrane promoting continuity in ITAM phosphorylation and signalling.

This study shows that the CYRs of CD3δε, CD3γε and ζζ dimers exhibit a coiled and interlaced conformation forming contacts with each other. This coiled conformation of the tails allowed some cytoplasmic tyrosines to penetrate into the membrane whilst the sidechains of some other cytoplasmic tyrosines were hidden within the CYR coiled conformation. The transient exposure of these hidden ITAM tyrosines to intracellular solvent during my simulations supports findings that resting T cells can undergo basal phosphorylation [291,292]. Moreover, these simulations suggest that the network of protein-protein interactions among the CYRs reduce the solvent accessibility surface area, thereby reducing the probability of ITAM phosphorylation in its resting state. Therefore, an allostery-based stimulation of conformational changes in the TCR-CD3 which alters the interactions of ζζ TMR with the rest of the TCR-CD3 as seen in these simulations can potentially contribute to the unbinding of ζζ CYRs from CD3 CYRs and further increase the exposure of their ITAMs for phosphorylation. This mechanism is consistent with recent data showing that inducing ζζ loosening in the TMR via mutations results in increased downstream signalling upon TCR-pMHC engagement [76]. My simulations in this chapter suggest that the TCR-CD3 CYRs undergo coiling, thereby not allowing all ITAM tyrosines to penetrate the membrane. As a result, the ζ subunits containing the lengthiest CYRs with the highest number of ITAM tyrosines are more likely to expose their ITAMs, compared to the CD3 CYRs. In addition, the total number of tyrosines in the ζζ CYRs is higher compared to the number of tyrosines present in all CD3 CYRs combined, potentially explaining why ζζ plays a major role in signal transduction. This is in agreement with a previous study that proposed that ζζ ITAM multiplicity can enhance signalling [293].

A recent cryo-EM study revealed most of the TCR-CD3 quaternary structure in a detergent environment, i.e. its TMRs and ECDs, but not its CYRs [16]. This allowed employing molecular modelling to complete the TCR-CD3 structure and simulate its dynamics in a bilayer that closely mimics the lipid headgroup composition of a TCR-CD3 activation domain [210]. In simulations performed here, the TCR-CD3 conformation was somewhat divergent from the cryo-EM

structure, indicating that its conformation may differ in a membrane environment compared to that obtained in a detergent environment. The simulations suggested an alteration in the TMR configuration where the ζζ and CD3δε dimers lost contact thereby bringing CD3δ and TCRβ TMRs closer than observed in the cryo-EM structure. Moreover, possible conformations of the TCR-CD3 ECD and interactions of VαVβ with CD3 ECDs were identified indicating flexibility of the antigen-binding domain of the TCRαβ, in line with studies that identified allosteric sites in the TCRαβ constant domains [9,45,47,74].

This work also demonstrates that the TCR-CD3 creates a unique lipid fingerprint in the membrane by forming selective interactions with anionic headgroups of PIP lipids. This is in agreement with experimental findings that have suggested that PIP lipids regulate membrane dynamics and TCR-CD3 activation [31]. Such unique membrane footprints have also been suggested for other membrane proteins [294]. The simulations show that PIP lipids interact strongly with the BRS of CD3ε and ζ subunits in the intracellular region, consistent with findings suggesting that BRS mediate interaction with PIPs and modulate signalling [28,29]. Moreover, the BRS of CD3ε is suggested to serve as a docking site for LCK [24,102] whereas that of ζ is suggested to help localise TCR-CD3s at the immunological synapse [29]. Similar cationic patches were observed in the juxtamembrane regions of receptor tyrosine kinases [295–297] and other signalling receptors such as the integrin-talin complexes [238,298] where they were shown to play critical functions in receptor activation by interacting with anionic lipids.

In this chapter, the cationic residues that are situated at the interface of the TMR and CYR of the TCR-CD3 were shown to maintain an anionic lipid environment around TCR-CD3 in the absence of the CYR. This suggests that the TCR-CD3 complex can maintain an anionic environment in its vicinity, albeit smaller, even when its cytoplasmic tails are not associated with the membrane. The positively charged regions of the cytoplasmic tails enhance the formation of a distinct annulus of PIP lipids around the TCR-CD3. The triggering of PIP clustering may further create a suitable lipid environment for recruiting peripheral proteins such as LCK. Mutation studies also showed that the LCK-SH2 domain interacts with PIP lipids via a cationic patch at K182 and R184, which is distinct from its phospho-tyrosine binding site [27]. The SH2 domain of ZAP-70, which is homologous to LCK-SH2, was also found to bind to $PIP_2$ headgroups and phospho-tyrosines independently. It was also found that these SH2 domains cannot associate with stimulated TCR-CD3s when their lipid binding site is mutated [105] suggesting that lipid interaction of SH2 domains

could be critical to their association with the TCR-CD3. Therefore, the formation of an anionic lipid environment enriched in PIPs around the TCR-CD3 shown in this chapter may be key for its interactions with LCK and other SH2 domain-containing protein kinases.

In summary, the first molecular model of the entire resting TCR-CD3 complex inserted in an asymmetric bilayer containing lipid types found in its membrane activation domain is hereby proposed. The dynamic model suggests conformational changes in the TCR-CD3, membrane penetration of some ITAM tyrosines and selective interactions of the TCR-CD3 with cholesterol and anionic headgroups, especially those of PIP lipids whose interactions were enhanced in the presence of CD3 and ζ CYRs. The clustering of PIP lipids around TCR-CD3 may facilitate the initial interaction of its CYRs with LCK, participate in TCR-CD3 clustering and aid in the spatial organisation of the immunological synapse. Therefore, these findings can lead to further studies, e.g. involving mutations in the CYRs to alter protein-lipid interactions, to better understand the molecular mechanism of TCR-CD3 activation and signalling.

### 5.4.1 Limitations

It is important to consider possible limitations of the simulations performed in this chapter. Given the lack of structural information on the arrangement of the TCR-CD3 CYR, this region was modelled in an extended conformation perpendicular to the membrane. This avoided bias in inter-chain contacts at the beginning of the simulations. For the CGMD simulations, the Martini 2.2 forcefield was used. As a result, the structured regions of the ECD and TMR were restrained to maintain their tertiary structure as observed in the cryo-EM structure (PDB:6JXR). Note that these restraints were applied only within each dimer and not between dimers to allow inter-dimeric relaxation as the TCR-CD3 is a complex of four dimers. This limitation was partly addressed by the multi-scale approach where the final structural configuration of the TCR-CD3 was backmapped along with the membrane to AT resolution and further simulated. From these ATMD simulations, the stability of the different TCR-CD3 domains was found to be maintained. It is also suggested that CGMD simulations with the Martini forcefield may exaggerate protein-protein interactions. Therefore, to confirm that the TCR-CD3 CYR forms coiled conformations, ATMD simulations were conducted starting from the same initial model used at the beginning of the CGMD simulations (with the CYRs in an extended conformation). These ATMD simulations also showed the CYRs coiling, supporting the CGMD simulations. Additionally, the ITAM tyrosine interactions with the hydrophobic

region of the membrane observed in this chapter are in agreement with experimental data showing Y177 of CD3ε penetrating the membrane [33]. The fact that the CGMD simulations agree with experimental findings provides a validation for this study on the TCR-CD3 dynamics.

### 5.4.2 Key research findings

- Multi-scale simulations highlight the dynamics of the complete TCR-CD3 complex in its resting state where it is observed that:
  (i) the variable domain in the extracellular region is flexible and can interact with the constant domain,
  (ii) the helical transmembrane subunits are dynamic and capable of undergoing slight rearrangement relative to the cryo-EM structure to relax themselves in a membrane environment,
  (iii) the disordered cytoplasmic region exhibits a random coiled conformation, and inserts some of the tyrosine residues (especially Y177 of CD3ε) into the membrane while the rest are shielded within the coiled conformation.
- Lipid interactions of the TCR-CD3:
  (i) $PI(4,5)P_2$ and $PIP_3$ lipid headgroups interact strongly with the BRS motifs of the cytoplasmic region. Further, the cytoplasmic region enhances PIP clustering and therefore aids in a negatively charged environment around the TCR-CD3.
  (ii) Cholesterol interacts strongly with the TCR-CD3 TMR, especially near the αβ dimer, opposite to the ζζ dimer.

# Chapter 6

# Molecular dynamics and lipid interactions of the full-length LCK

## 6.1 Introduction

In the previous chapters, I have provided molecular insights into the dynamics of a resting TCR-CD3 and how it maintains its ITAM tyrosines hidden within the membrane or within the coiled state of its cytoplasmic region. In addition, my simulations also indicated that specific interactions of pMHCs with the CDR loops of TCRαβ can influence their bond lifetime, and that certain residues in the transmembrane region can potentially transmit the pMHC-induced signal to the cytoplasmic region allowing their phosphorylation and the initiation of the signalling cascade. This phosphorylation event is mediated by the lymphocyte-specific protein tyrosine kinase (LCK). In this chapter, I have modelled the full-length LCK and simulated its dynamics in association with the membrane.

LCK is a member of the Src family of tyrosine kinases responsible for initiating T cell signalling upon kindled by the TCR-pMHC engagement [9,278]. The activation of the TCR-CD3 leads to solvent exposure of its immunoreceptor tyrosine-based activation motifs (ITAMs) from its membrane-bound state [28–30,33,34,69]. The phosphorylation of these ITAMs is found to be mediated by a pool of free LCK [114] and less likely by those bound to CD4/CD8 co-receptor [115,116]. In addition, LCK is found constitutively active in the plasma membrane of naive T cells and does not increase in number upon TCR-pMHC engagement but aids in phosphorylation of ITAMs [114]. This suggests that spatial organisation of LCK is key in signalling [87]. Further, studies using imaging techniques showed that the conformational states of LCK are key in their organisation and therefore in TCR-CD3 ITAM phosphorylation and regulation [117]. The phosphorylated ITAMs then act as a substrate for the stable binding of ZAP-70 [299,300]. However, ZAP-70 needs to be phosphorylated by LCK to stimulate its catalytic activity [301] and initiate the next phase of T cell signalling.

Understanding the role of LCK in molecular detail is important in deciphering the initial phases of T cell activation. To achieve this, it is key to obtain the full-length 3D structure of LCK which remains structurally unresolved until date. The full-length LCK contains the following domains (from N to C terminus): the SH4 (first ~10 residues), unique domain (UD), SH3, SH2, and the kinase domain (Fig 6.1). X-ray crystallography experiments have resolved the structures of the isolated SH2 and SH3 domains at 2.36 Å (PDB:4D8K), and the isolated kinase domain in its active state i.e. phosphorylated at Y394 at 1.7 Å (PDB:3LCK) [18]. The active (open) and inactive (closed) states of LCK are controlled by tyrosine phosphorylations in the kinase domain. Y394, present in the catalytic loop of the kinase domain, is an autophosphorylation site. Upon phosphorylation (p),

pY394 is responsible for phosphorylating the TCR-CD3 ITAMs. On the other hand, Y505 is an inhibitory site present in the C-terminal tail. When Y505 alone is phosphorylated, it binds to a cationic pocket in the SH2 domain and renders a compact and a closed LCK conformation. When both Y394 and Y505 are dephosphorylated by pY phosphatases (PTPases) such as CD45, LCK enters an intermediate state between open and closed, known as the primed state (Fig 6.1). The CD45 PTPase binds to LCK via Y192 in its SH2 domain [108]. Studies suggested that, when Y192 is mutated and thereby disabling CD45-LCK binding, both Y394 and Y505 of LCK become phosphorylated. This leads to LCK adopting its closed conformation due to pY505-SH2 binding (Fig 6.1). In resting T cells with wild-type LCK, LCK is regulated by CD45 PTPases competing with LCK's autophosphorylation at Y394 and with Csk kinases at the LCK's C-terminal Y505 [113]. Upon T cell-APC contact, the large extracellular domains of CD45 phosphatases are pushed away [70] thereby segregating CD45 from the activation site and allowing phosphorylation-mediated signalling.

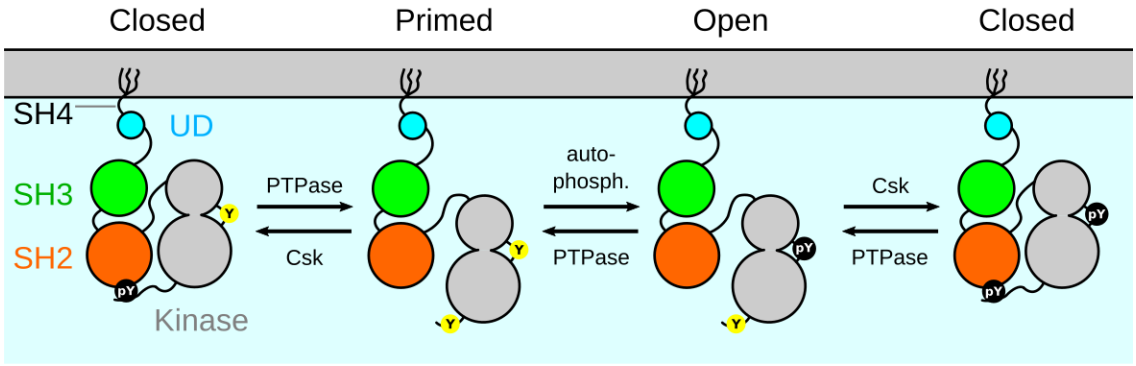


***Fig 6.1 Schematic representation of the closed/inactive, primed and open/active conformations of LCK.*** *This figure was derived from a review by Chakraborty and Weiss* [23] *and modified based on later findings* [108].

The SH4 and UD remain structurally unresolved. In the SH4 domain, G2, C3 and C5 undergo acylation as a post-translational modification i.e. myristoylation at G2 [88,89], and palmitoylation at C3 and C5 [25,91,92]. As a result, these covalently added acyl chains, also referred hereon as lipid tails, insert into the hydrophobic core of the membrane and aid in membrane localisation of LCK [26]. In addition, S6 of LCK was shown to serve as a critical site for G2 to be myristoylated [90]. The unique domain of LCK, which exhibits low sequence similarity across Src family members, lies after the acylated residues and before the SH3 domain. In this chapter, I refer to the N-terminal residues and the UD combined as the 'SH4-U' domain for simplicity.

Lipids in the plasma membrane have been reported to play an important role in interacting with LCK via its SH2 domain and in turn regulating TCR-CD3 signalling [27]. In particular, anionic lipids such as phosphatidylinositol-4,5-biphosphate i.e. PI(4,5)P2 ($PIP_2$) and phosphatidylinositol-3,4,5-triphosphate ($PIP_3$) are suggested to aid in LCK interaction with the TCR-CD3 in a spatiotemporal manner. LCK-SH2 domain is found to be key in lipid interaction by selectively contacting these PIP lipids via its cationic patch [27]. SH2 domains in other receptor tyrosine kinases (RTKs) such as ZAP-70 have also been reported to direct signalling pathways by binding to PIP lipids [105]. In the previous chapter, my simulations suggested that the TCR-CD3 maintains an anionic lipid environment enriched in PIP lipids with the help of its cytoplasmic region [277]. Since LCK is also shown to possess high affinity for PIP lipids [27] and its clustering is driven by their open conformational state [117], it is important to understand the correlation between its lipid interactions and conformational states in molecular detail. This could further aid in our understanding of its interaction with ITAMs of stimulated TCR-CD3 complexes.

In this chapter, I modelled the full-length LCK by predicting a model of the SH4-U domain and integrating it with the experimentally resolved structures of the SH2, SH3 and kinase domains. Further, I performed CGMD simulations of an accumulative time of 100 microseconds for each of the open and closed states of LCK in a complex symmetric bilayer. The lipid headgroup composition of the bilayer mimics the inner leaflet of the T cell membrane activation domain [210]. From these simulations, I study the dynamics of the open and closed conformations of LCK and their lipid interactions.

### 6.1.1 Key research aims

- Model the full-length LCK along with its post-translational modifications in its open and closed conformational states, and elucidate their structural forms in a dynamic membrane environment.
- Study the lipid interactions of the open and closed conformations of the full-length LCK.

## 6.2 Methods

### 6.2.1 Coarse-grained molecular dynamics simulations

To coarse-grain the post-translational modifications i.e. the lipid tails along with the rest of the protein, the *martinize* script, and the Martini 2.2 amino acid topology were modified to include published parameters [232], and made publicly available (https://github.com/DJ004/martini_mod). The elastic network (EN) distance cut-off of 0.7 nm created restraints only within each domain of the LCK open state. This maintained each of their tertiary structures independently and allowed unbiased inter-domain interactions. Membrane lipid compositions used to set up all CGMD simulations are shown in Table 6.1. In all CGMD simulations, each lipid contained one saturated acyl chain and one mono-unsaturated acyl chain, while their headgroup composition is based on the composition of TCR-CD3 activation domains in the T cell plasma membrane [210].

***Table 6.1 Summary of CGMD simulations conducted in this chapter.***

| Simulations | Membrane | Particles | Simulation box dimension (X x Y x Z) (nm) | Duration | Replicas |
|---|---|---|---|---|---|
| LCK-SH4-U | simple | 22134 | 12 × 12 × 17 | 1 μs | 20 |
| LCK-SH3 | simple | 25676 | 12 × 12 × 20 | 1 μs | 20 |
| LCK-SH2 | simple | 25914 | 12 × 12 × 20 | 1 μs | 20 |
| LCK-SH2,3,4-U | complex | 53576 | 16 × 16 × 23 | 5 μs | 20 |
| LCK-SH2,3,4-U mut2 | complex | 51020 | 16 × 16 × 22 | 5 μs | 20 |
| LCK-SH2,3,4-U mut3 | complex | 50692 | 16 × 16 × 22 | 5 μs | 20 |
| LCK-SH2,3,4-U mut5 | complex | 51173 | 16 × 16 × 22 | 5 μs | 20 |
| LCK-FL open mut5 | complex | 83194 | 19 × 19 × 26 | 5 μs | 20 |
| LCK-FL open | complex | 83137 | 19 × 19 × 26 | 5 μs | 20 |
| LCK-FL closed | complex | 46187 | 16 × 16 × 20 | 5 μs | 20 |

(Symmetric) membrane composition:

- simple: POPC/POPS/$PIP_2$/$PIP_3$ = 72/20/6/2
- complex: POPC/POPE/POPS/Chol/$PIP_2$/$PIP_3$ = 12/40/20/20/6/2

### 6.2.2 Atomistic molecular dynamics simulations

CHARMM-GUI [226] was used with the CHARMM36 forcefield [214] to setup ATMD simulations of the initial LCK-SH4-U model, the LCK-FL open model and closed model in solution using the TIP3 water as solvent neutralised with 0.15M Na+ and Cl- ions. All systems were energy minimised using Gromacs 2016 until the maximum force converged to 1000 kJ/mol/nm, followed by isotropic (NPT) equilibration where the protein backbone was position-restrained. The equilibrated system was used to generate differing initial velocities for three production simulations run 250 ns each. Coordinates were saved at 40 ps intervals.

## 6.3 Results

### 6.3.1 Modelling the full-length LCK in its open and closed states

To obtain a model of the 3D structure of full-length LCK, I first modelled its SH4-U domain since it is structurally unknown. To predict the structure of the SH4-U domain (Fig 6.2A), I used the PSIPRED secondary structure prediction server [247] along with 3D structure predictions by two independent servers i.e. I-Tasser [302] and Robetta [303] were used. Hydrogens were added to both the models and their topologies were generated using the CHARMM36 forcefield [214] and Gromacs 2016 [212]. Both models were taken and subjected to 250 ns ATMD simulations in solution neutralised by 0.15M Na+ and Cl- ions. These simulations allowed both models to relax their tertiary structure and embrace their preferred state of folding. At the end of the simulations, the resultant structures were once again compared to the initial secondary structure predictions. Based on their agreement with the predictions and some structural evidence from an NMR study [96], the model derived from the Robetta server was considered for further model building. The NMR study revealed a small loop region of the UD consisting of a hairpin-like bend responsible for binding CD4 and CD8 co-receptors via a coordinating $Zn^{2+}$ ion [96]. When compared with this NMR-resolved region, my SH4-U model also showed that this region is unstructured, bent in a similar way, and likely to be exposed to solvent (magnified region in Fig 6.2B). Due to the absence of a zinc ion in my model, a disulphide bond was taken up. Nonetheless, this region of my model mimics the NMR findings [96]. In addition, I identified a potential anionic binding surface involving residues E10, D11, E15, E21 apart from D12 (Fig 6.2B) whose mutation (D12N) was shown to reduce binding with CD3ε BRS by NMR experiments [24].

Following modelling of the SH4-U domain, the crystal structures of SH2, SH3 (PDB:4D8K), and kinase domains (PDB:3LCK) along with the SH4-U model were assembled in a linear manner to model the full-length LCK open conformation (Fig 6.2C) using UCSF Chimera [245]. Missing residues were predicted to be unstructured and hence modelled as loops between the domains using Modeller 9.2 [282]. The crystal structures were assembled sufficiently far from each other to avoid bias in protein-protein interactions at the beginning of the simulations. Note that the positioning of the SH3 and SH2 domains relative to each other were not altered and were used as obtained from the crystal structure (PDB:4D8K). On the other hand, a homolog of the Src family of kinases, HCK, was used to homology model the LCK closed state. During homology modelling, the crystal structure of HCK closed state resolved at 1.65 Å (PDB:5H0B) was used as a structural template to map the 3D structure of LCK closed state using Modeller 9.2 [304] (Fig 6.2D). This modelling involved a multiple sequence alignment of HCK and LCK produced by Clustal Omega [305]. The structures of the SH2, SH3, kinase domains of LCK were also individually aligned with those of HCK (Fig 6.2E) using the 'super' aligning method in PyMOL (https://pymol.org/2/). This indicated a structural similarity between their SH2, SH3, kinase domains with a root mean square deviation of 0.582, 1.135, 0.92 Å respectively. The resultant homology modelled structure of the closed state of LCK (SH2, SH3, kinase domains) was then conjoined with the previously obtained SH4-U model to obtain the closed state of the full-length LCK (LCK-FL) (Fig 6.2D).

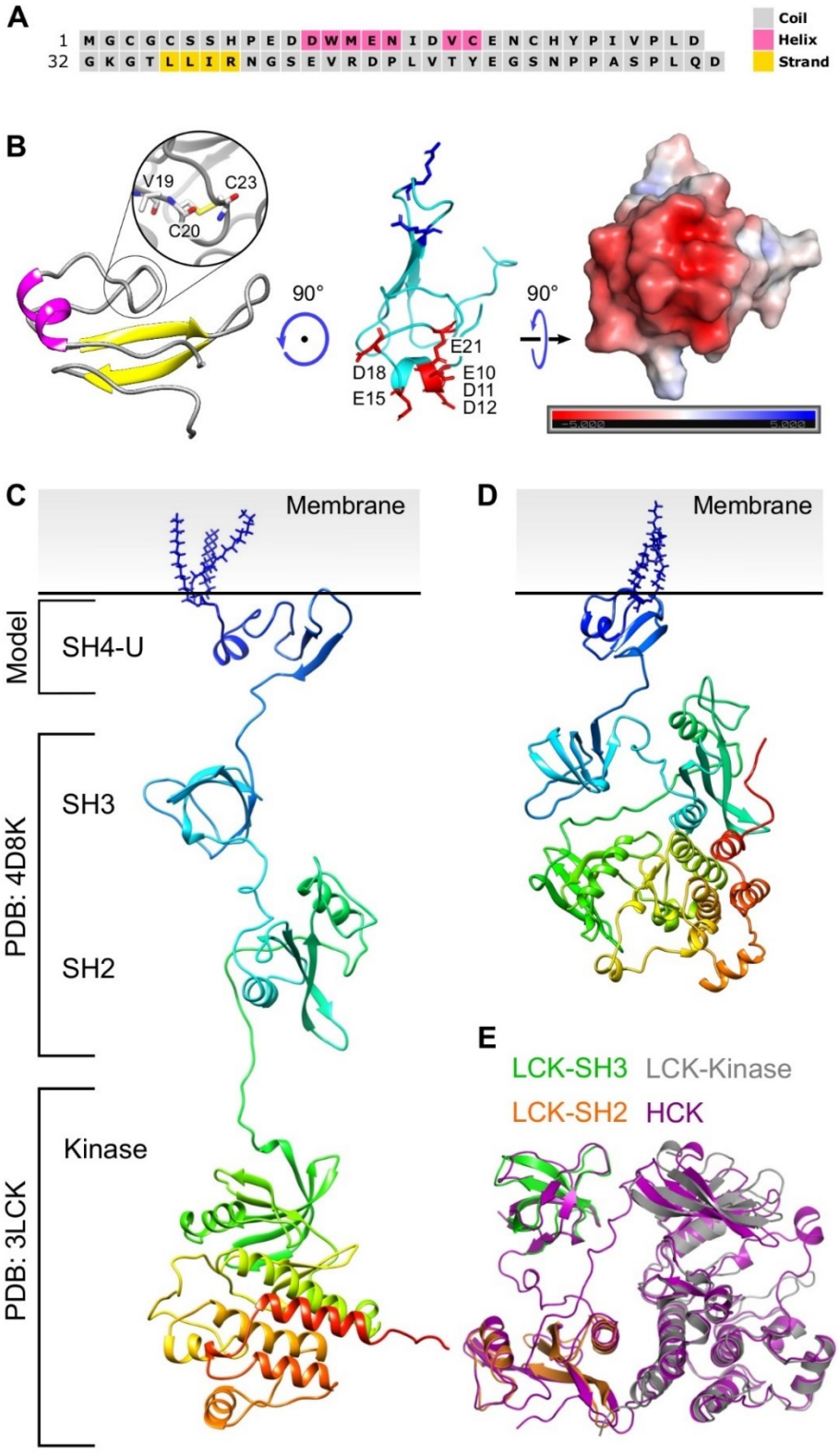
A
1 M G C G C S S H P E D D W M E N I D V C E N C H Y P I V P L D
32 G K G T L L I R N G S E V R D P L V T Y E G S N P P A S P L Q D
Coil
Helix
Strand
B
V19
C23
C20
90°
E21
D18
E10
D11
E15
D12
90°
C
Membrane
Model
SH4-U
PDB: 4D8K
SH3
SH2
PDB: 3LCK
Kinase
D
Membrane
E
LCK-SH3
LCK-Kinase
LCK-SH2
HCK

***Fig 6.2 Model of the SH4-U domain and, the open and closed full-length LCK conformations.*** *(A) Secondary structure prediction of LCK-SH4-U by PSIPRED 4.0 server. (B) The model of the SH4-U domain (residues 2 to 63) used in this chapter. The region responsible for coordinating a Zn2+ ion is magnified (left). Residues forming the anionic patch are shown as red sticks and labelled. Lipid binding cationic residues as shown in this chapter (R39, R45) are shown as blue sticks and located on the opposite side of the anionic patch (middle). The electrostatic profile of the anionic patch (right) shown was calculated in the ±5 kT/e range and at pH 7.0 using the PDB2PQR* [287] *and APBS* [286] *tools. Electronegative and electropositive regions are indicated by red and blue intensities respectively. (C) The open LCK-FL and (D) closed LCK-FL model, including the post-translational modifications at the N-terminus (blue), used in this chapter. (E) Isolated SH2, SH3 and kinase domains aligned to the closed state of HCK (PDB:5H0B; view in 3D here: https://www.rcsb.org/3d-view/5H0B).*

Finally, to both the LCK-FL open and closed models, post-translational modifications (PTMs) were made to the N-terminal residues i.e. G2 was myristoylated and, C3 and C5 were palmitoylated. The initiator M1 residue was removed during this process since they are known to be cleaved in mature eukaryotic proteins upon glycine-myristoylation at the N-terminus [89]. All atomistic protein models were energy minimised before coarse-graining. Coarse-grained protein structures were then inserted in symmetric membrane environments using the *Insane* tool [240], neutralised with solvent and ions, and equilibrated using the NPT ensemble before CGMD simulations (see Table 6.1 for membrane composition).

### 6.3.2 Membrane association and lipid interaction of the full-length LCK

To assess the association of the LCK-FL open model with the membrane, I began the CGMD simulations where its domains did not initially interact with each other as shown in Fig 6.2C. Similarly, CGMD simulations of the LCK-FL closed model began with the homology model of LCK (SH2, SH3, kinase domains) linked with SH4-U, as shown in Fig 6.2D. Note that all simulations performed here onward contained the PTMs in the SH4-U domain. In addition, simulations of the open and closed models of LCK-FL started with their PTMs (i.e. lipid tails) partially inserted into the membrane. Further, 20 replicates of each of these models were simulated individually for 5 μs leading to an accumulative 100 μs simulation time for both models. By the end of the simulations, I found that both models of LCK-FL associated with the membrane.

Calculation of the average distance versus time of the centre of mass (COM) of the initial protein model to the COM of the membrane along the vertical (Z) axis showed that both LCK-FL models closely associated with the membrane within 1 μs simulation time (Fig 6.3A). I hypothesised that the close association of these LCK-FL models to the membrane was due to strong electrostatic interactions between cationic residues of LCK-FL and anionic headgroups. For this reason, I analysed their anionic lipid headgroup interactions versus time as they contacted the membrane.

Consistent with the time of association of both LCK-FL models to the membrane over the first microsecond, I observed an increase in the number of anionic ($PIP_2$ and $PIP_3$) lipid headgroups around the protein (open and closed conformations) over the same time interval. In this calculation, lipid interactions were normalised by the number of those lipids present in the membrane, thereby highlighting the preference in lipid interactions (Fig 6.3B). For both models, it was observed that $PIP_2$ and $PIP_3$ were preferred by LCK over other lipids. The radial distribution function (RDF) of all lipids surrounding both LCK-FL models further confirmed the abundance of PIP lipids in their vicinity (Fig 6.3C). When normalised by lipid concentrations, cholesterol, POPC, POPE, and POPS were less preferred compared to PIPs in both LCK-FL models. Further, I calculated the binding sites of both LCK-FL open and closed conformations with the less preferred lipid headgroups i.e. cholesterol, POPC, POPE, and POPS (Fig 6.3D). I found that the first 50 residues at the N-terminus interacted with cholesterol the most. These interactions were largely mediated by the acylated residues myristoylated G2, palmitoylated C3 and C5, followed by H24, Y25, P26, V44, R45, D46, suggesting their ability to penetrate the membrane surface. Cholesterol interactions throughout the rest of the LCK-FL open and closed models were negligible compared to the SH4-U domain. Interactions of POPC, POPE, and POPS were also higher in the first 50 residues compared to the rest of LCK showing that this region embraces a close association with the membrane provided that the SH4-U lipid tails are inserted.

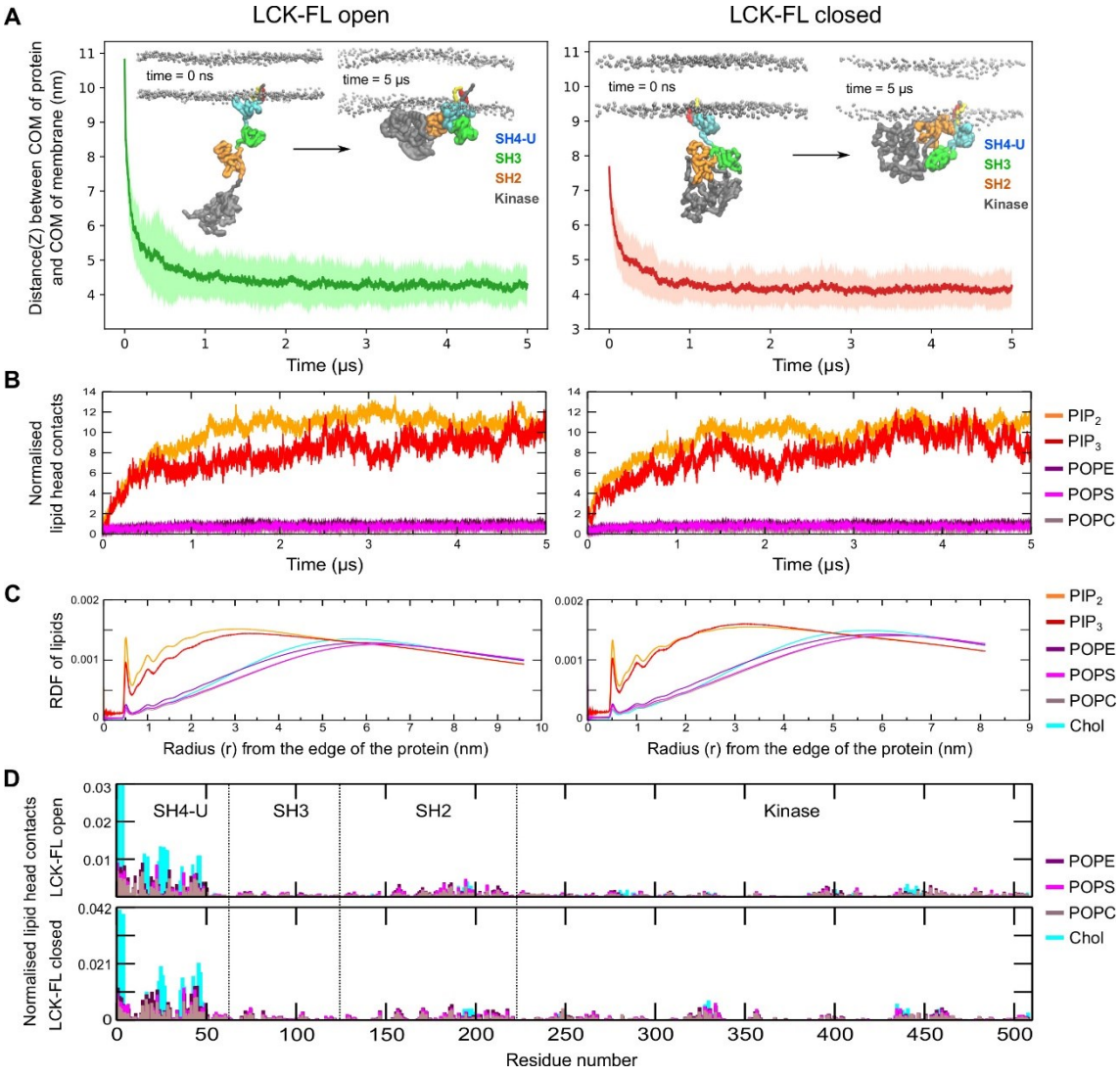


***Fig 6.3 Membrane association and lipid interactions of the open and closed full-length LCK conformations.*** *Association of open and closed conformations of LCK-FL with the membrane is indicated by the reduction in distance between the centre of mass (COM) of LCK-FL and COM of the membrane versus time. (B) The increase in the headgroup interactions of PIP lipids is observed as LCK associates with the membrane. The number of headgroup interactions of each phospholipid type is normalised by the number of lipids of the respective lipid type in the membrane. (C) The radial distribution function (RDF) of all lipid types around LCK calculated throughout the simulation time. The RDF is normalised by the total number of lipids in the membrane of that system to enable comparison between the open and closed conformations of LCK-FL. Note: Simulations of the LCK-FL open and closed systems contain different number of lipids in the membrane due to different dimensions of the XY plane (along the membrane). (D) Normalised headgroup contacts of cholesterol, POPC, POPE, POPS with LCK-FL open (top) and closed (bottom) conformations. Normalisation was done by dividing lipid contacts by the number of respective lipids present in the membrane.*

### 6.3.3 PIP lipid binding sites

#### 6.3.3.1 Simulations of individual LCK SH2, SH3 and SH4-U domains

Following my investigation of lipid interactions of the LCK-FL open and closed states which showed an overall preference for PIP lipids, I proceeded to study their PIP lipid binding sites. Prior to investigation of the PIP lipid binding sites of the open and closed states of LCK-FL, I simulated the LCK-SH2, SH3, SH4-U domains individually to be able to analyse their lipid interactions with and without the influence of the other domains. Given that the kinase domain is the largest domain constituting greater than half of the LCK-FL sequence, and to assess their lipid interactions without the influence of the kinase domain, I also simulated the LCK-SH domains combined together (LCK-SH2,3,4-U). In all the individual domain and in the LCK-SH2,3,4-U simulations which were run in 20 replicates for 1 μs each, the protein structure was placed ~6 nm away from the membrane (Fig 6.4A) to allow it to explore all possible orientations in solution before binding to the membrane, whereas simulations of the full-length LCK started with the lipid tails partially inserted into the membrane.

After simulating the individual LCK-SH domains for 1 μs in a symmetric membrane environment (see Table 6.1 for membrane composition), I found that K182 and R184 of LCK-SH2 were the most interactive residues with PIP lipids, preferring $PIP_3$ over $PIP_2$, as suggested by previous studies [27]. However, my simulations also suggested additional cationic residues interacting strongly with the anionic headgroups of PIP lipids in the order: K182 > R184 ~ R134 > K135 > K179. LCK-SH2 was also found to bind to the membrane within a shorter period of time (Fig 6.5A) compared to SH3 and SH4-U domains when simulated individually. In the individual simulations, while LCK-SH2 dominated interaction with the membrane, LCK-SH3 made much weaker contacts. R89 of LCK-SH3 was distinctly the most interactive residue with PIP lipids. Further, LCK-SH4-U made negligible contacts (Fig 6.4B). I also found a significantly larger fraction of unbound SH4-U compared to SH3 and SH2 domains (Fig 6.5B). This was due to the lack of a strong lipid binding site and because the myristoylated and palmitoylated lipid tails failed to insert into the membrane in the majority of the simulations. This suggests that membrane association of LCK-FL is likely to be primarily driven by a strong lipid binding site rather than depending on the SH4 lipid tail insertion. This is consistent with previous experimental findings which showed that mutating the LCK-SH2 lipid binding site reduces it localisation at the plasma membrane [27]. On the other hand, in my simulations where the lipid tails inserted into the membrane, SH4-U stayed membrane-bound for the rest of the simulation time.

#### 6.3.3.2 Simulations of the LCK SH2, SH3, SH4-U domains combined

In the LCK-SH2,3,4-U simulations, the LCK-SH2 dominated the interactions with PIP lipids (via R134, K135, K182, R184), while those of SH3 (peak at K84) and SH4-U (peak at R45) were not significant. This indicated that LCK-SH2 lipid interactions were not influenced by the SH3 / SH4-U domains (Fig 6.4C). These PIP interactions sites were maintained when compared to those of the individual LCK-SH domain simulations, except that the SH3 domain exhibited a different residue but nonetheless interacted electrostatically.

Due to the high affinity of LCK-SH2 for the membrane, it seemed to pull the SH3 and SH4-U domains toward the membrane but did not aid in the membrane insertion of SH4-U lipid tails. Moreover, given the hydrophobicity of the lipid tails, they were found binding to a small cavity near the SH2-SH3 linker before the protein attained its membrane-bound state (Fig 6.5C). Similarly, *in vitro* studies have shown that the myristoylated acyl chain in the c-Src protein kinase binds to the SH3 domain and modulates its insertion into the membrane [306]. In my simulations, this binding of the SH4-U lipid tails to the SH2-SH3 linker prevented its membrane insertion despite SH2-mediated lipid binding, potentially due to the energy barrier of unbinding from the SH2-SH3 linker. Nevertheless, this augments my previous observation from the individual SH domain simulations that lipid tail insertion may not be required for LCK to associate with the membrane, in agreement with experiments [117,307]. Although the lipid tails of SH4-U can promote the membrane-bound state of LCK, it is suggested that LCK-SH2 is primarily responsible for membrane association [27].

#### 6.3.3.3 Simulations of the full-length LCK in its open conformation indicate the primary lipid binding site of LCK-SH2

To study the lipid interaction profiles of LCK-FL in its open and closed conformations, the protein models were placed near the membrane such that the SH4 lipid tails would immediately insert into the membrane. The SH2, SH3 and kinase domains in the LCK-FL closed model were initially bound to each other while the LCK-FL open model exhibited a linearly stretched conformation (Fig 6.4A), The SH4-U domain was not associated with the rest of the protein in either model. The PIP lipid binding sites of the SH2 and SH3 domains in the LCK-FL open conformation were similar to those observed in the simulations of

the individual domains and of LCK-SH2,3,4-U. In comparison with all domains in the open LCK-FL, the SH2 domain made significant contacts via its primary binding site (K182 > R184 ~ R134 > K135 ~ K179) while the SH3 domain made least contact with the membrane, interacting mostly via K84. Similar to the SH2 domain, the SH4-U domain made significant contacts by preferring to bind to PIP lipids via R39 and R45, followed by cholesterol interactions (shown in cyan in Fig 6.3D) via myristoylated G2, palmitoylated C3 and C5, followed by H24, Y25, P26, V44, R45, D46. Interestingly, the kinase domain of LCK, whose lipid interactions has not been studied until date, also showed significant PIP lipid interactions. It exhibited specific binding sites for PIPs, the most interactive residues being R455, R458, R461, R474, K478, R480 (Fig 6.4D) in the open conformation. These residues are situated at the bottom surface of the C-terminal lobe of the kinase and constitute a flat cationic area acting as a suitable lipid binding site (Fig 6.5D). In the closed conformation, the kinase domain retained significant PIP interactions via R455, R458, R461 (Fig 6.4E) at the bottom of its C-terminal lobe.

#### 6.3.3.4 Simulations of the full-length LCK in its closed conformation reveal a secondary lipid binding site of LCK-SH2

Compared to the LCK-FL open simulations, the lipid interaction profile of the SH4-U, SH3 and kinase domains remained fairly similar in the closed state of LCK-FL whereas the SH2 domain exhibited a distinct lipid binding site (H148, R219, R222) along with some residues of the primary binding site i.e. K182 and R184 (Fig 6.4E). This means that K182 and R184 interacted with lipids in both the open and closed states of LCK-FL, however they dominated PIP lipid interaction in the open state while H148, R219, R222 interacted more in the closed state.

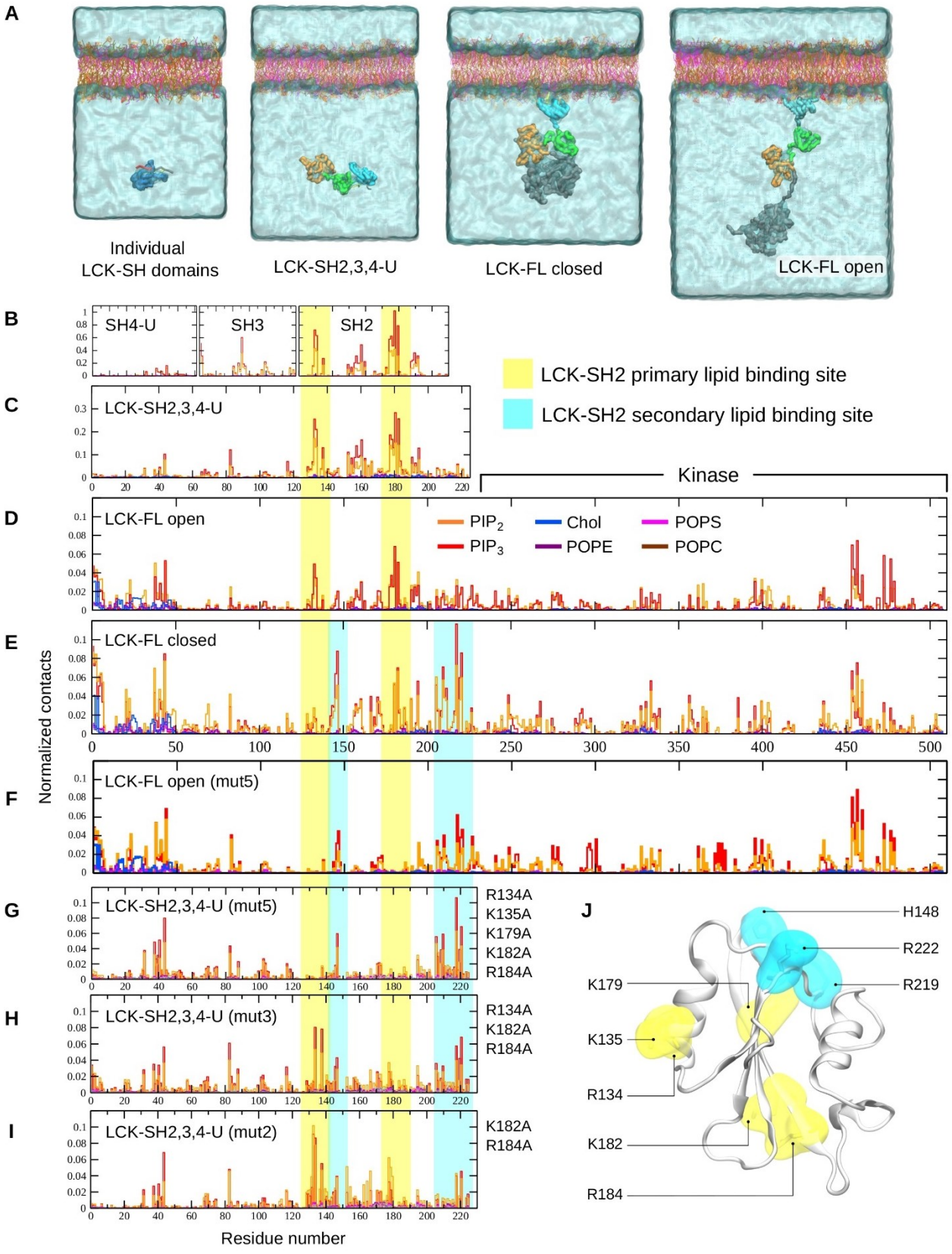
A
Individual
LCK-SH domains
LCK-SH2,3,4-U
LCK-FL closed
LCK-FL open
B
SH4-U
SH3
SH2
C
LCK-SH2,3,4-U
LCK-SH2 primary lipid binding site
LCK-SH2 secondary lipid binding site
Kinase
D
LCK-FL open
PIP2
PIP3
Chol
POPE
POPS
POPC
E
LCK-FL closed
Normalized contacts
F
LCK-FL open (mut5)
G
LCK-SH2,3,4-U (mut5)
R134A
K135A
K179A
K182A
R184A
H
LCK-SH2,3,4-U (mut3)
R134A
K182A
R184A
I
LCK-SH2,3,4-U (mut2)
K182A
R184A
Residue number
J
H148
R222
R219
K179
K135
R134
K182
R184

***Fig 6.4 Snapshots of the initial simulation frames and PIP lipid binding sites.*** *(A) The initial simulation frames of the individual LCK-SH domains (SH4-U, SH3, SH2) contained the protein placed in solution ~6 nm away from the membrane. The same applied to the LCK-SH2,3,4-U simulations whereas the closed and open conformations of LCK-FL started with the lipid tails of the SH4 domain partially inserted into the membrane. (B) Normalised lipid interactions of the SH4-U, SH3, SH2 domains when individually simulated, and (C) of the SH4-U, SH3, SH2 domains simulated in conjunction (LCK-SH2,3,4-U). (D) Normalised lipid interactions of LCK-FL open, (E) LCK-FL closed, (F) LCK-FL open when mutated (mut5 i.e. R134A, K135A, K179A, K182A, R184A), (G) LCK-SH2,3,4-U when mutated (mut5), (H) LCK-SH2,3,4-U mut3 (R134A, K182A, R184A), and (I) LCK-SH2,3,4-U mut2 (K182A, R184A). (J) Residues constituting the primary (yellow) and secondary (blue) lipid binding sites of LCK-SH2 as observed in CGMD simulations of LCK-FL open and closed.*

#### 6.3.3.5 LCK adopts the secondary lipid binding site upon mutation of the primary binding site in the SH2 domain

To investigate whether, in the absence of the primary binding site, the secondary binding site would help the LCK-FL open conformation attain a membrane-bound state, I induced *in silico* mutations in the LCK-FL open model. I substituted the residues that constituted the primary PIP lipid binding site i.e. R134A, K135A, K179A, K182A, R184A, as observed in the individual LCK-SH2 simulations, LCK-SH2,3,4-U simulations, and LCK-FL open simulations. This mutation (referred as mut5) in the open LCK-FL led to the loss of lipid interaction via the primary binding site but retained contacts with PIPs via the secondary binding site, although they were reduced (Fig 6.4F) compared to the closed LCK-FL (Fig 6.4E). The lipid interaction of other domains remained unaffected. However, performing this mutation in the LCK-FL open conformation while its SH4 lipid tails are initially associated with the membrane may impose some bias in its lipid interactions. Therefore, I also performed the same mutation in the LCK-SH2,3,4-U structure and placed the protein ~6 nm away from the membrane. I knocked out the large kinase domain to allow the LCK-SH2,3,4-U structure to freely explore orientations prior to membrane association.

The mut5 mutations in the LCK-SH2 domain in the LCK-SH2,3,4-U structure led to lipid interactions of the secondary lipid binding site via R219 and R222. These were comparable with those of the open LCK-FL model, suggesting that the secondary binding site does help in membrane association provided LCK-FL is in vicinity of the membrane. In addition, I also performed other mutations into the LCK-SH2,3,4-U structure referred here as mut3 i.e. R134A, K182A, R184A (Fig 6.4H), and mut2 i.e. R134A, K135A (Fig 6.4I) to assess the

importance of the residues involved in the primary binding site. I found that LCK-SH2,3,4-U mut2 still maintained lipid interactions in its primary lipid binding orientation via R134 and K135, and LCK-SH2,3,4-U mut3 via K135 and R139. On the whole, simulations of these mutants suggested that, as the cationic residues of the primary binding site are knocked out, LCK-SH2 gradually shifts its preference to the secondary binding site (Fig 6.4J). The fact that H148, R219 and R222 in the secondary binding site dominated PIP interaction in the closed LCK-FL suggests that the closed conformation potentially alters the preferred membrane-binding orientation of LCK-SH2 to some degree. The lipid binding orientation of the other domains remained the same.

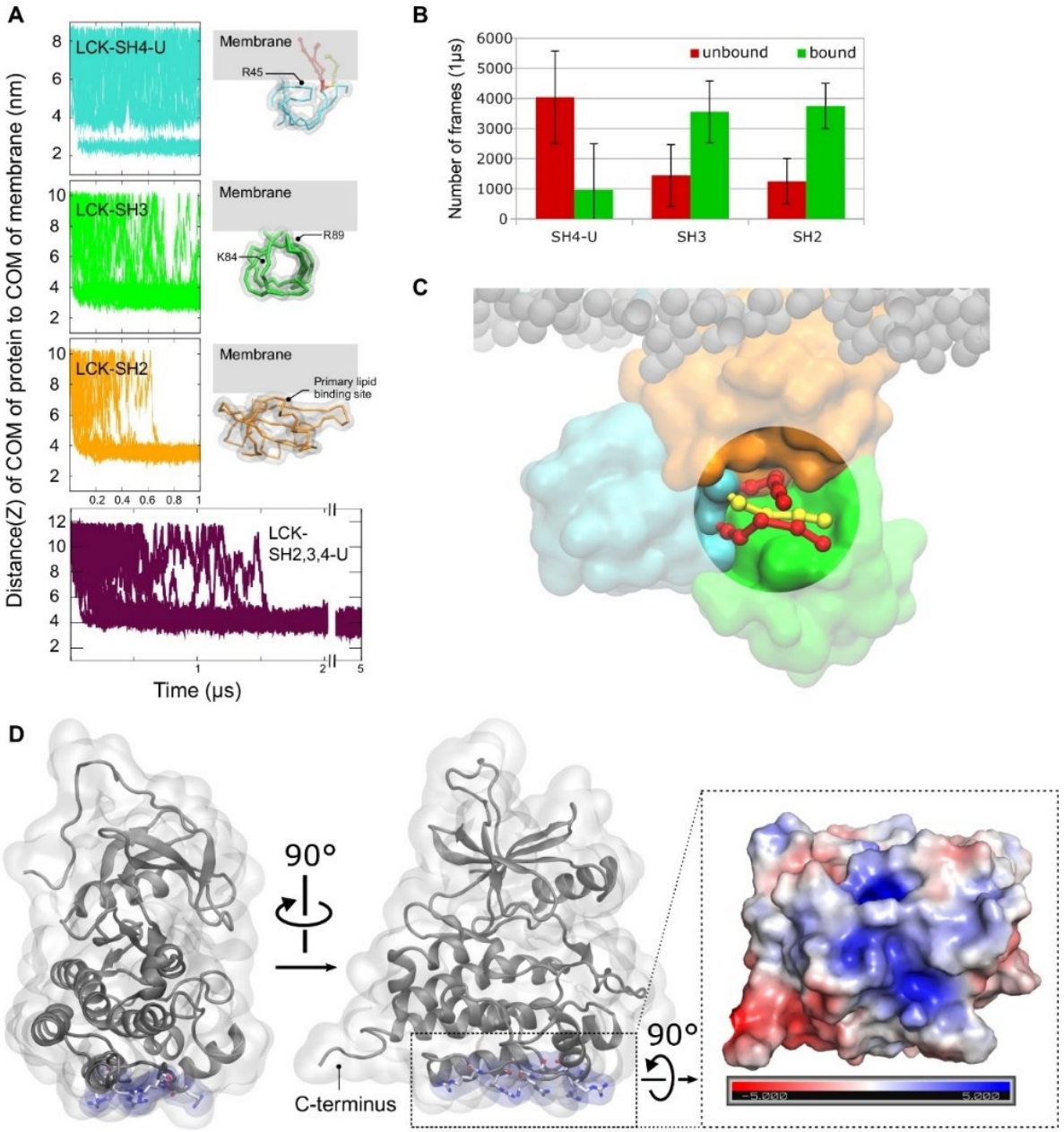

***Fig 6.5 Membrane binding and orientation of the LCK-SH2, SH3, SH4-U domains, and electrostatics of the kinase domain.*** *(A) Distance between the centre of mass (COM) of protein to COM of membrane in all 20 simulation replicates: SH4-U (cyan), SH3 (green), SH2 (orange), LCK-SH2,3,4-U (maroon). The most observed membrane-bound conformations of the individually simulated LCK-SH domains are also shown. Residues interacting most with the membrane in these simulations are labelled. (B) Average number of frames in the individual LCK-SH simulations that the protein stayed unbound (red) or bound (green) to the membrane. (C) A snapshot from one of the LCK-SH2,3,4-U simulations highlighting the binding pocket of the SH4 lipid tails (near the SH2-SH3 linker loop region) when they did not insert into the membrane. Myristoylated G2 residue is shown as yellow ball and stick. Palmitoylated C3 and C5 residues are shown as red ball and stick. SH2, SH3, SH4-U domains are shown as orange, green, cyan surfaces, while phospholipid headgroups are shown as grey spheres. (D) Residues in the kinase domain forming a flat cationic patch. Their electrostatic profile was calculated using in the ±5 kT/e range and at pH 7.0 using the PDB2PQR and APBS tools. Electronegative and electropositive regions are indicated by red and blue intensities respectively.*

### 6.3.4 Atomistic molecular dynamics simulations in solution

To study the dynamics of LCK before it binds to the membrane in atomistic resolution, I simulated both the LCK-FL open and closed models in solution using ATMD simulations in 3 replicates and for 250 ns each. In the LCK-FL open model, Y394 is phosphorylated (pY394) and in the closed model, Y505 is phosphorylated (pY505) and tethered to the SH2 domain interacting with its phospho-tyrosine binding site R154. However, my ATMD simulations also suggested that pY505 interacted with LCK-SH2 residues R134 and K182 in the closed state of LCK-FL (Fig 6.6). These residues in LCK-SH2 were observed to play significant roles as the primary lipid binding site in my CGMD simulations. Therefore, this interference of pY505 with the LCK-SH2 primary lipid binding site may be responsible for weakening its membrane interaction potentially reducing its localisation in TCR-CD3 activation membrane domains, in agreement with literature suggesting that LCK clustering at the plasma membrane is driven by its open conformation [117].

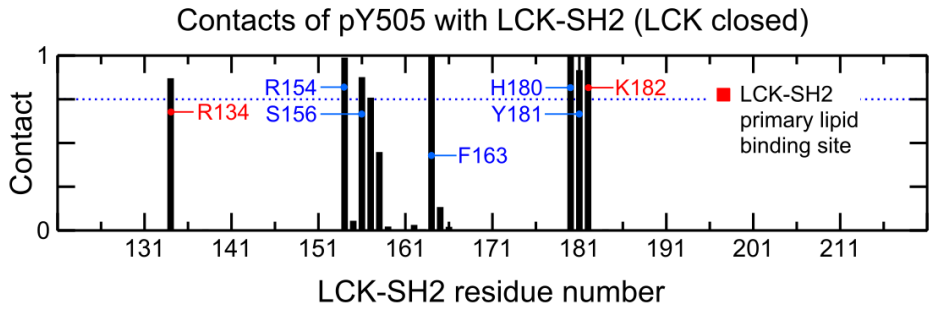


***Fig 6.6 Normalised average contacts of pY505 with LCK-SH2 domain in the LCK-FL closed state*** *in ATMD simulations. Normalisation was done by dividing the number of contacts by the number of simulation frames thereby obtaining a scale of 0 to 1.*

#### 6.3.4.1 Structural orientation of open and closed LCK in solution

To deduce the most representative structural conformation of the open and closed LCK-FL in solution, I performed a clustering analysis where protein structures were extracted from the ATMD simulations at 200 ps intervals and grouped into clusters containing similar conformations within an RMSD cut-off of 0.35 nm. The SH4-U domain was excluded from this analysis to avoid overall inaccuracy due to its unstructured and dynamic nature. The clustering analysis revealed that the SH2, SH3 and kinase domains in the LCK-FL closed model maintained its quaternary structure in solution whereas the domains in the open model, which were initially modelled in a linear manner, underwent compaction. This compaction was also observed in the LCK-FL open CGMD simulations, shown by comparing the radius of gyration of the protein backbone (Fig 6.7A). As a result of this compaction in the LCK-FL open conformation, the clustering analysis indicated the SH2, SH3, and kinase domains in the LCK-FL open simulations in solution exhibited a similar orientation as that observed in the LCK-FL closed model (Fig 6.7B). In this orientation of the open model (Fig 6.7C), the C-terminal lobe of the kinase interacts with the SH2 domain, similar to the closed state of LCK-FL (Fig 6.7B). However, since my simulations did not contain co-receptor proteins or mimic LCK clustering, this compact orientation of open LCK (Fig 6.7C) may be applicable when LCK is free and not bound to another protein at the plasma membrane.

Further, the similar orientation between the open and closed LCK-FL structures in the absence of a membrane led me to inspect the differences in the position of the kinase relative to the SH2 and SH3 domains. Interestingly, my ATMD simulations of the LCK-FL open state showed that Y505 faced away from the protein while pY394 interacted with the SH2 domain via S158 and T159. In addition, pY394 was found ionically interacting with R363, R387, R397 residues

within the kinase domain (Fig 6.7C and Fig 6.7D), consistent with crystallographic findings (PDB:3LCK) [18]. The SH2-pY394 interaction via S158 and T159 occurred only half the simulation time on average (Fig 6.7E) in ATMD simulations, and was not observed in CGMD simulations possibly due to the lack of coarse-grained particles to represent the anionic phosphate group on Y394. On the other hand, the kinase domain in the LCK-FL closed conformation did not undergo any significant conformational change. The most observed SH2-SH3-kinase orientation, considering all three simulation replicates, was consistent with the starting structure where pY505 was bound to the SH2 domain (Fig 6.6).

It was suggested that the closed state of LCK is stabilised by the interaction of the SH3 domain with proline residues located in the SH2-kinase loop region [23]. My ATMD simulations of the LCK-FL closed state are in agreement with this indicating that P229 and P232 of the SH2-linker region interacted with the SH3 domain (Fig 6.7F). P112 of the SH3 domain also interacted with the SH2-kinase linker region (Fig 6.7B and Fig 6.7G). P74 of the SH3 domain was in proximity with the SH2-kinase linker (Fig 6.7B) but it established negligible contacts (Fig 6.7G). I observed that these proline interactions also formed in the simulations of LCK-FL open conformation in solution (Fig 6.7C), although occurring less frequently compared to the closed conformation (Fig 6.7F). Analysis of contacts between the prolines of SH3 domain (P74, P112) and prolines of the SH2-kinase linker (P229, P232) suggested that neither the open nor the closed LCK-FL conformation could maintain these proline interactions throughout the simulation suggesting their tendency to disengage. However, in the most representative structure, I observed that the open conformation of LCK-FL does allow the interaction of SH3 with the proline motif to some extent despite the dynamics of the SH2-kinase linker (Fig 6.7C). This may convey that the disengagement of Y505 from the SH2 domain, or the transitioning of LCK from its closed to open conformation, is not dependent on the proline interactions between the SH3 and SH2-kinase linker. Nonetheless, the unbinding of the SH3 from the SH2-kinase linker might result in further extension of the SH2-kinase linker promoting overall flexibility of LCK.

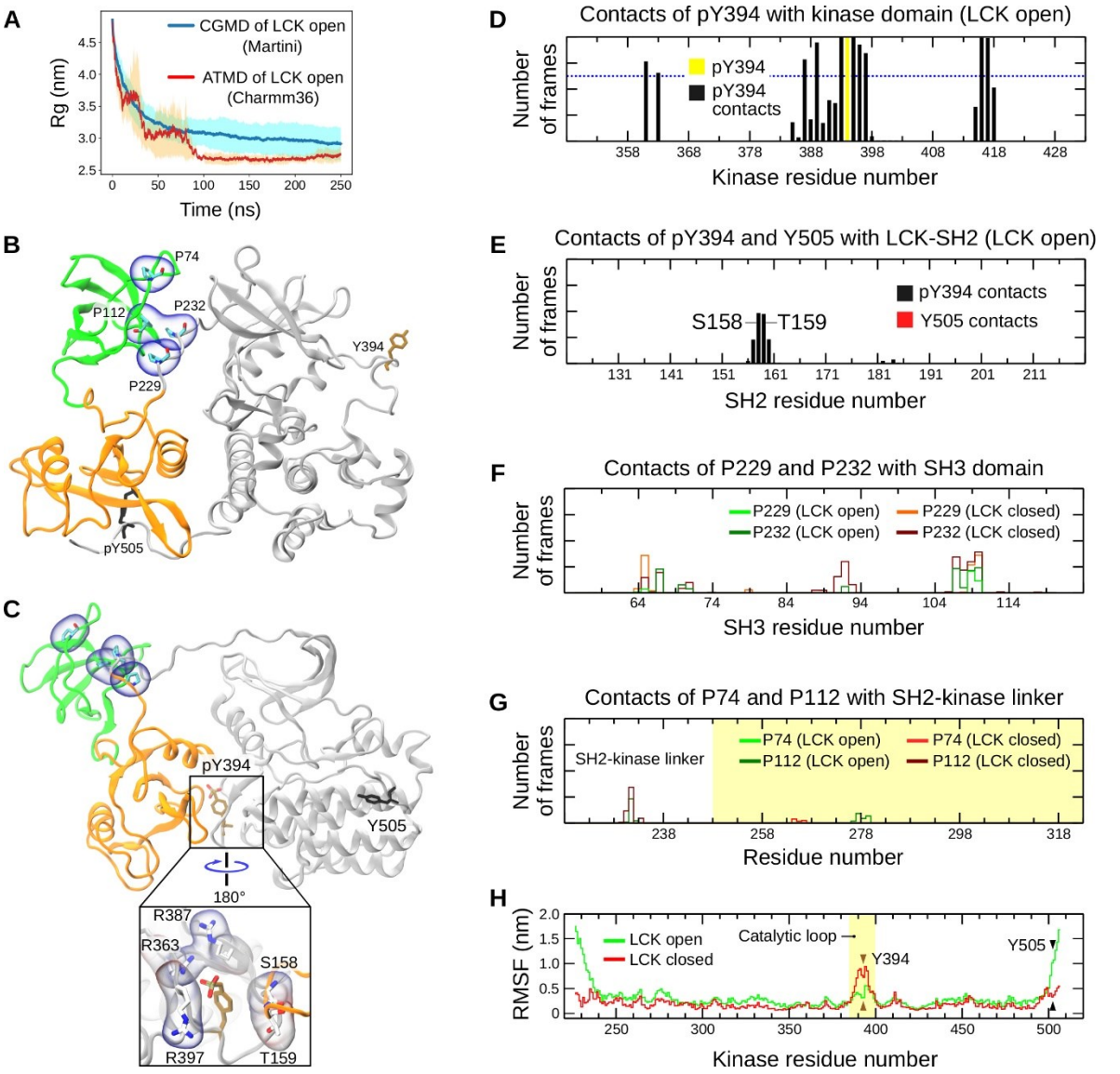


***Fig 6.7 Conformation and protein-protein interactions of the SH2, SH3, kinase domains in the open and closed states of LCK-FL*** *in ATMD simulations. (A) Average radius of gyration of the backbone of the LCK-FL in CGMD and ATMD simulations with their standard deviation shown in cyan and orange respectively. (B) The most observed conformation of the SH2, SH3, kinase domains as a result of clustering analysis of the open LCK-FL, and (C) the closed LCK-FL. The ionic contacts of pY394 in the kinase domain and its polar contacts with the SH2 domain are shown as observed in the most exhibited conformation in the ATMD simulations. In (B) and (C), the proline residues at the interface of the SH3 domain and loop region N-terminus of the kinase domain are shown as blue bubbles. Y394 and Y505 are shown as brown and black sticks respectively, and are observed to be located on opposite sides of the kinase. In the ATMD simulations, Y394 and Y505 carry a phosphate group in the open and closed states respectively. (D) Normalised average contacts of pY394 with the kinase domain in the open LCK-FL state. (E) Normalised average contacts of pY394 and Y505 with LCK-SH2 in the open LCK-FL state. (F) Normalised average contacts of P229 and P232 with LCK-SH3 in the open and closed states of LCK-FL. (G) Normalised average contacts of P74 and P112 with the SH2-kinase linker region. The white region represents the SH2-kinase linker residues while the yellow region represents the residues*

*of the N-lobe of the kinase domain. In (D) to (G), normalisation was done by dividing the number of contacts by the number of simulation frames. (H) Root mean square fluctuation (RMSF) of residues in the kinase domain, including the loop region N-terminal to the kinase (~230 to 240), in both the open and closed states of LCK-FL. The RMSF of Y394 and Y505 is indicated by brown and black arrows respectively.*

RMSF analysis of the kinase also indicated that this SH2-kinase linking loop region (residues ~230 to 240) experienced major fluctuation in the open state in contrast to the closed state where it was significantly stable. In addition, as expected, Y505 was stabilised i.e. exhibited a lower RMSF compared to Y394 in the closed LCK-FL, and vice versa in the open LCK-FL (Fig 6.7H). In the open LCK-FL, the contacts of residues in the SH2 and/or kinase domains with pY394 (Fig 6.7D and Fig 6.7E) potentially contribute to its stability and aid phosphorylation of TCR-CD3 ITAMs.

#### 6.3.4.2 Comparing the structural orientation of open LCK-FL in solution with its membrane-bound state

Following the orientation of open LCK-FL in solution, I also deduced the most representative conformation of the membrane-bound open LCK-FL from the CGMD simulations using clustering analysis using the same 0.35 nm RMSD cut-off. Here, all domains were considered including the SH4-U due to its membrane-induced stability and the presence of EN to conserve its tertiary structure. When comparing the three most exhibited conformations of the CGMD (membrane-bound LCK) with the ATMD (LCK in solution) simulations, I observed that Y394 and Y505 often switched positions i.e. both residues alternatively being proximal to the SH2 domain (Fig 6.8) indicating that the kinase domain was able to rotate and re-orient relative to the SH2 domain. This flexibility of the kinase domain in the LCK-FL open conformation is potentially key to the dynamics of its catalytic activity. Note that, despite taking up positions near the SH2 domain, neither pY394 nor Y505 contacted the LCK-SH2 lipid binding site (Fig 6.7E) suggesting that LCK-SH2 is free to bind to lipids in the open state, unlike in the closed state where pY505 interacted with some lipid binding residues of LCK-SH2 (Fig 6.6).

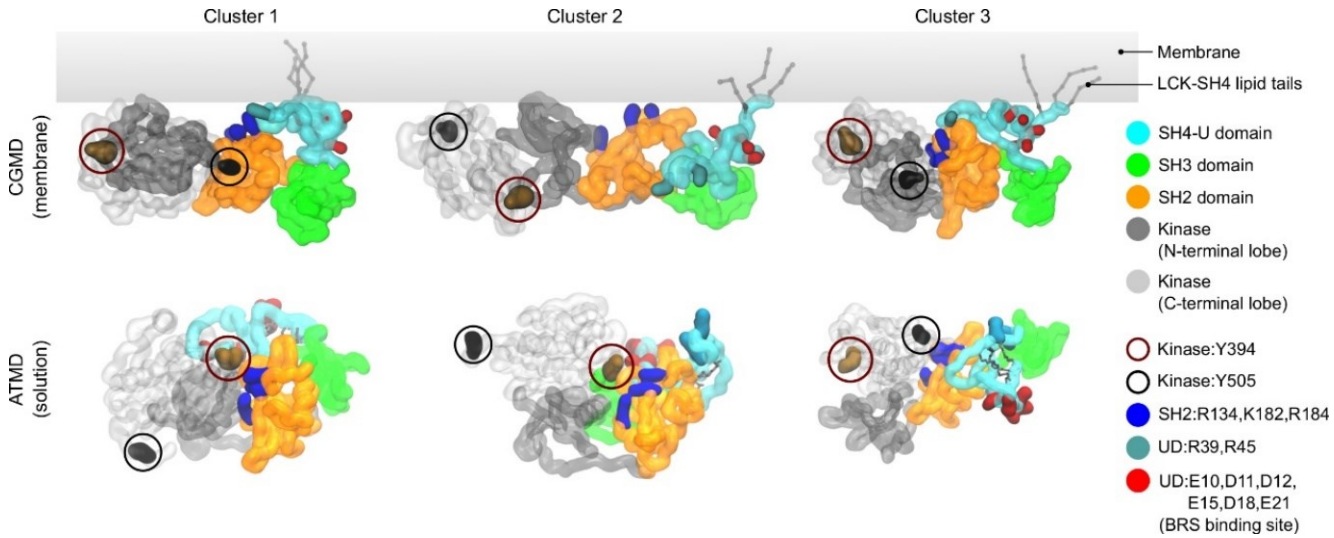


***Fig 6.8 Top three representative conformations of LCK-FL open*** *from CGMD (top) and ATMD simulations (bottom) derived from clustering analyses. Positions of Y505 and Y394 are highlighted to indicate the rotation of the kinase domain i.e. Y505 and Y394 alternatively being proximal to LCK-SH2. In the ATMD simulations, the SH4-U is hidden since it was not considered for clustering analysis due to its dynamics.*

### 6.3.4.3 The structural conformation of LCK-SH4-U domain in solution

From the ATMD simulations of the open and closed LCK-FL simulations in solution, clustering analysis was also performed on the SH4-U domain alone. Using its most representative conformation, I obtained its electrostatic profile. This revealed that the SH4-U domain, despite its dynamic nature, maintained an anionic patch (E10, D11, D12, E15, D18, E21) independent of the open and closed LCK conformations, consistent with the initial model (Fig 6.9). In addition, this anionic patch was identified on the opposite side of its lipid binding surface (R39, R45) suggesting its availability to bind to cationic residues, potentially those of the BRS motif of ζ and CD3ε subunits of the TCR-CD3.

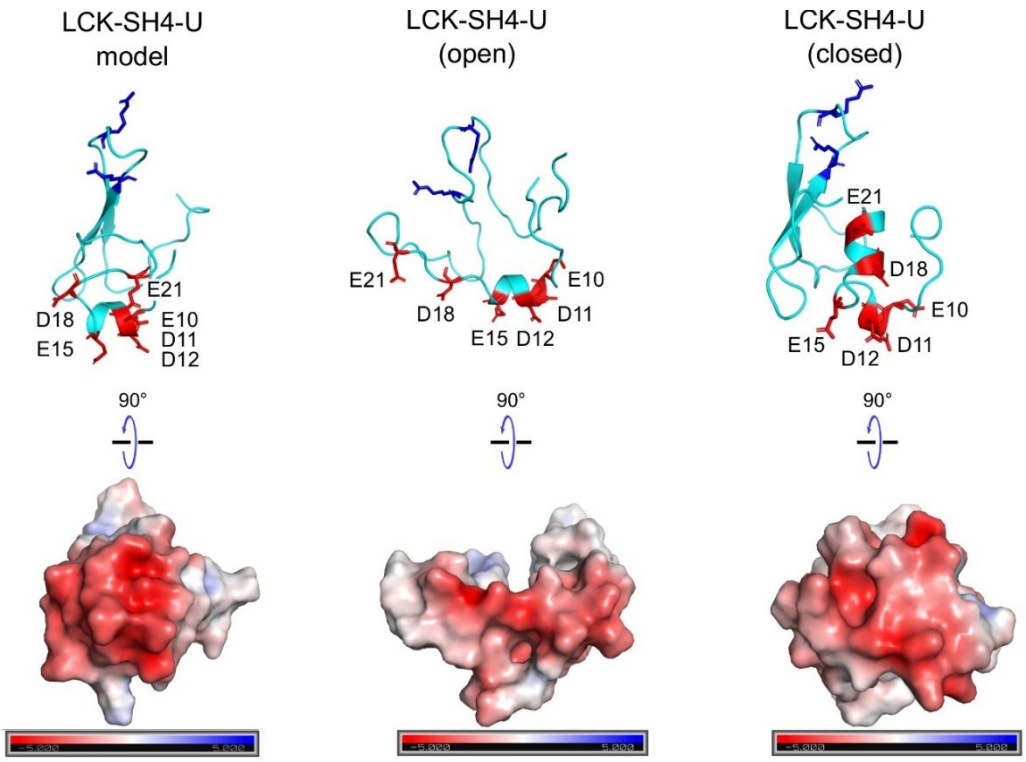


***Fig 6.9 The electrostatic profiles of LCK-SH4-U*** *as in the initial model (left), and according to clustering analyses of the LCK-FL open state (middle), and of the LCK-FL closed state (right). Their electrostatic profiles were calculated in the ±5 kT/e range and at pH 7.0 using the PDB2PQR and APBS tools. Electronegative and electropositive regions are indicated by the red and blue intensities respectively. The residues forming the anionic patch are shown as red sticks and labelled. Lipid binding residues (R39, R45) are shown as blue sticks for reference.*

## 6.4 Discussion

LCK is a member of the Src family of protein kinases that associates with the inner leaflet of the T cell plasma membrane. During antigen-induced allosteric activation of the TCR-CD3 complex [76], the phosphorylation event is first driven by a pool of constitutively active LCK [114] that is less likely to be associated with co-receptors [115,116]. In this chapter, I modelled and simulated the full-length open LCK which is not bound to co-receptors and contains phosphorylated (p)Y394 and dephosphorylated Y505. I also modelled and studied the inactive or the closed state of the full-length LCK which contains a dephosphorylated Y394, SH2-bound pY505, and whose SH3 interacts with the proline-containing motif located N-terminal to the kinase domain, as previously suggested [23]. This closed conformation of LCK is homologous to other members of the Src kinase family [19,308]. Further, the open conformation of LCK has always been schematically represented to exhibit a

linear conformation roughly perpendicular to the membrane [23]. However, in my simulations, I observed that the open state of LCK embraces a similar conformation as the closed state where the kinase domain places itself proximal to the SH2 and SH3 domains. This observation was made during simulations of LCK-FL conformations in both solution and while associated with membrane.

In addition, despite studies suggesting that the SH3 domain interacts with a proline motif in the SH2-kinase linker region and stabilises the closed state of LCK [23], my results show that the SH3 domain in open LCK also tends to bind to proline motifs. This is consistent with experimental findings that the SH3 domain in open LCK binds PxxP proline motifs in the CD3ε subunits of TCR-CD3 complexes upon activation [102]. The fact that my simulations of both the open and closed LCK conformations exhibit this SH3-proline interactions also suggests that the closed state of LCK may primarily be driven by the binding of pY505 to SH2. In the LCK open conformation, apart from its SH3 domain interacting with proline-rich motifs of the TCR-CD3, the phospho-tyrosine binding site of SH2 i.e. R154 can also attach to one of the ITAMs and facilitate phosphorylation of other ITAMs mediated by pY394 of the kinase. Consistent with and in addition to this, experimental findings indicated that R154A mutation does not reduce membrane association of LCK indicating that phospho-tyrosine binding can take place independent of lipid interaction [27].

The primary lipid binding site of the SH2 domain has been shown to be mediated largely via K182 and R184. The preference of the SH2 domain for $PIP_3$ compared to $PI(4,5)P_2$ has also been quantified [27]. In agreement, lipid interaction profiles from my simulations largely indicate interactions of LCK with $PIP_3$ despite its low concentration in the membrane. Amongst all lipids in the membrane, LCK interacted strongly with both $PIP_3$ and $PIP_2$ lipids. Interestingly, these lipid species were found to cluster around the TCR-CD3 and interact with its cytoplasmic region as seen in the previous chapter [277].

Experimental studies showed the number of activated LCK does not increase upon TCR-CD3 stimulation and thus, activation is led by constitutively active LCK [114], meaning that there is no transition from closed to open LCK conformations. Consistent with this, further studies revealed that TCR-CD3 activation is followed by LCK clustering that is driven by its open conformation [117,307]. Similar to my simulations, experiments have also suggested that TCR-CD3 complexes are likely to be activated by free LCK and not those bound to CD4/CD8 co-receptors [116]. In this open state of LCK in my simulations, in addition to previously suggested lipid binding residues i.e. K182 and R184, I identified other lipid-interacting residues such as R134 and K135 contributing to

the primary lipid binding site. Furthermore, in my simulations, residues including and neighbouring A160 was observed to interact with lipids, but not as significantly as K182 and R184. Nevertheless, this observation is consistent with mutation studies which revealed that A160K reduces dissociation of LCK-SH2 from plasma membrane-mimetic vesicles [27].

In this chapter, I also identified a secondary binding site in the SH2 domain consisting of residues H148, R219, and R222, that come into play when the primary binding site is mutated or not in contact with the membrane. This suggests that LCK-SH2 is capable of attaining a membrane-bound state by adopting a different conformation, although less preferred and potentially weaker. This secondary binding site was employed frequently in the closed state of LCK compared to the open state. However, it is shown that LCK does not bind to the ζζ component of the TCR-CD3 if its preferred binding site is altered or if its phospho-tyrosine binding site R154 is mutated [27]. Therefore, this secondary lipid binding may promote membrane association to a certain degree but may reduce colocalisation with the TCR-CD3 due to change in orientation. During spatial organisation of LCK with TCR-CD3 on activation [87], this alteration in its SH2 domain orientation and lipid binding site may also reduce its competence with the preferred open LCK conformations, as previously suggested [117,307].

In addition to experimental findings that the LCK-SH2 binds to the membrane [27] preferring its primary binding site, and the SH4 lipid tail insertion [25,26], my studies suggest that the kinase domain also tends to interact with the membrane via a considerably cationic patch irrespective of the open and closed conformations. I also highlight that an anionic patch on LCK-SH4-U is maintained despite its dynamics, and this anionic site is suggested to act as a suitable binding site for basic-rich motifs of CD3ε subunits of the TCR-CD3 by NMR experiments [24]. Given the potential of strong electrostatic interactions between LCK and the CD3 subunits, and the close association of LCK and of the TCR-CD3 cytoplasmic tails to the inner leaflet of the membrane [277], I propose that initial TCR-LCK association and ITAM phosphorylation occurs proximal to the inner leaflet of the plasma membrane. Based on our current understanding from experimental evidence, constitutively active LCK interacting with the plasma membrane via its SH2 domain are spatially reorganised in the vicinity of the TCR-CD3. Upon allosteric activation of the TCR-CD3 complex [76] when its BRS and PRS motifs are exposed along with its ITAMs, the anionic surface of the LCK-SH4-U domain binds to the BRS motif of CD3ε or ζ subunit whereas the LCK-SH3 binds to the PRS motif of CD3ε. Further, the

phospho-tyrosine binding site of LCK-SH2 R154 is capable of binding to TCR-CD3 ITAMs. These interactions may together promote TCR-LCK association independent of lipid interaction and aid in ITAM phosphorylation by the kinase domain.

In summary, this chapter sheds light on the lipid interactions of the full-length LCK at near-atomic resolution over an accumulative simulation time of 100 microseconds for its open and closed models each. In addition to providing insights into the complete structure and membrane-associated dynamics of LCK, my studies pave way to further experimental investigation on the membrane-binding of LCK via its kinase domain, and on the secondary lipid interaction sites of the SH2 domain. Using structural models of LCK from this chapter and of the TCR-CD3 from the previous chapter, this work can also be extended into a steered MD and/or other non-equilibrium MD studies of pMHC-induced TCR-CD3 stimulation in a complex membrane environment followed by LCK-ITAM association.

### 6.4.1 Limitations

In this study, I performed CGMD simulations using the MARTINI forcefield which involved elastic network (EN) restraints [233] within each domain of the open state of LCK in order to maintain their tertiary structures as suggested by experiments (PDB:4D8K, 3LCK). To make the simulations more realistic, I avoided inter-domain restraints allowing each domain to freely associate with each other. In the closed state of LCK, it is known that the kinase binds to the SH2 and SH3 domains. Therefore, such a configuration was homology modelled based on the structure of the closed state of HCK protein, a member of the Src kinase family, resolved at 1.65 Å (PDB:5H0B) [309]. This model was restrained using EN in my CGMD simulations to retain the LCK closed conformation. Finally, the SH4-U domain was modelled based on secondary structure predictions and validated using ATMD simulations combined with available experimental evidence. However, the resultant model was also restrained using EN during CGMD simulations.

### 6.4.2 Key research findings

- CGMD simulations of the full-length LCK indicate that the membrane-associated form of both the open and closed conformations are similar i.e. where all domains prefer to be proximal to each other and to the surface of the membrane.
- Lipid interactions of the full-length LCK (open/closed):
  (i) The interaction of the SH4-U domain with the membrane is observed when its myristoylated (G2) and palmitoylated (C3, C5) residues are inserted into the membrane. Once inserted, the proximity of SH4-U with the membrane leads to its interaction with ionic and non-ionic lipids including cholesterol. In addition, R39 and R45 of SH4-U are observed to significantly interact with PIP lipid headgroups.
  (ii) The SH3 domain interacts least with the membrane with only R89 consistently interacting with PIP lipid headgroups.
  (iii) Further to previous reports that the SH2 domain interacts with PIP lipids via K182 and R184 [27], this chapter suggests R134, K135, and K179 as additional PIP lipid binding sites on the SH2 domain in the open conformation of LCK. In its closed conformation, PIP interaction via K182 and R184 were retained but dominated by H148, R219, and R222. This indicates that the PIP interaction sites of the SH2 domain and its orientation relative to the membrane are altered in the closed conformation of the full-length LCK.
  (iv) The kinase domain has not been previously reported to interact with the membrane. This chapter not only suggests that it does, but also indicates a cationic patch in the C-terminal lobe which mediates the majority of its interactions with PIP lipid headgroups.
- ATMD simulations reveal that the kinase domain interacts with the SH2 domain via pY394 when LCK is open, in contrast to pY505 interacting with the SH2 when LCK exhibits a closed conformation. This rotation of the kinase between the open and closed states of LCK potentially suggests its flexibility that may be useful during phosphorylation of TCR-CD3 ITAMs.

# Chapter 7

# Future Directions

In this chapter, I discuss potential investigations that can be further conducted using the data presented in previous chapters, with the ultimate aim of gaining dynamic insights into various aspects of TCR-CD3 signal initiation.

## 7.1 Mutations in the complete TCR-CD3 complex

In chapter 5, I modelled and simulated the complete TCR-CD3 complex revealing its interactions with lipids in an asymmetric complex membrane. In addition to anionic lipids binding to the juxtamembrane region of the receptor, I found that PIP lipids made significant interactions with the cationic / basic-rich stretch (BRS) motifs of the ζ and ε cytoplasmic tails. Moreover, calculation of lipid densities over time indicated that PIP lipids formed a distinct annulus around the receptor suggesting an anionic lipid environment in its resting state. This is consistent with studies suggesting that ionic interactions between BRS motifs and acidic phospholipids regulate the activity of the TCR-CD3 complex [28–30,310] and other plasma membrane proteins [32].

Further, the dynamic insights provided by my simulations aid in identifying specific sites in the TCR-CD3 cytoplasmic region (CYR) that may be targeted to alter protein-lipid interactions to either up- or down-regulate TCR-CD3 activation. While simulations can determine the degree of dissociation of the CYR from the membrane and its solvent accessibility, experiments involving site-directed mutagenesis combined with functional assays can help reveal the magnitude of downstream signalling upon mutations in the CYR.

In chapter 3, I studied the dynamics and effects of certain *in silico* mutations that mimic an allosteric reaction in the TCR-CD3 TMR. My simulations supported experiments showing that ζI38A and ζI41A mutations lead to quaternary structure relaxation or ‘loosening’ which further enhances TCR-CD3 signalling. Although, my simulations only involved the structure of the TCR-CD3 TMR due to the lack of a complete TCR-CD3 model at the time. This gave rise to a limitation in terms of observing the dynamics of the allosteric mechanism in the whole TCR-CD3 complex. Therefore, in addition to substitutions in the CYR, the ζ TMR mutations can be conducted in comparison with the WT using the complete TCR-CD3 model. These simulation studies, when conducted for a sufficient time-scale and with a bound pMHC, can augment our molecular understanding of the allosteric pathway that leads to the exposure of ITAM tyrosines.

## 7.2 TCR-CD3 triggering upon pMHC binding

In chapter 4, I performed SMD simulations where I applied a pulling force on a class I MHC, containing a high affinity peptide, away from the TCRαβ ECD. This was primarily done to study the time of dissociation (indicative of binding affinity), and specific interactions between a set of peptide variants and the TCRαβ. The limitation in this chapter was that multiple pulling angles were not considered. The fact that other SMD simulation studies have also used a specific angle / direction of pulling [209,311] instigates this aspect to be explored. This is because the angle of TCR-pMHC contact when a T cell is scanning a target cell is likely to be random, thereby inducing a variety of tilt angles in the TCRαβ ECD. The pMHC may also turn upon binding with the TCRαβ potentially inducing a twist in the TCR-CD3 complex. This may be simulated using the enforced rotation approach in Gromacs in addition to pulling the MHC from the TCR. Moreover, simulating the twisting and pulling of the pMHC when bound with the complete TCR-CD3 model may reveal more conformational changes in the TCR-CD3 structure than previously studied. Furthermore, the contributions of high and low affinity pMHCs toward the dissociative mechanism of the complete TCR-CD3 CYR from the membrane can also be studied.

The above-mentioned plans are to be executed preferably using ATMD simulations which will yield more accurate results compared to CGMD simulations, but may demand a high cost of computation to sample sufficient time-scales and individual simulation replicas. Although CGMD simulations may be employed to tackle the computational cost, they may produce elastic network (EN)-related artifacts when trying to compute conformational changes in the TCR-CD3 structure. To overcome this, each subunit in the TCR-CD3 can be martinized individually to avoid inter-subunit restraints, and EN parameters can be fine-tuned to exert ‘softer’ restraints on secondary/tertiary structures but at the cost of losing hydrogen bonds during CGMD simulations.

## 7.3 Clustering of TCR-CD3 complexes

In addition to single TCR-CD3 complexes being capable of initiating downstream T cell signalling [85], experiments have suggested that the formation of TCR-CD3 clusters precedes ITAM phosphorylation and also that the strength of TCR-pMHC interaction is translated downstream via dense nanoclusters of TCR-CD3 [312]. Given that TCR-CD3 clustering is a key aspect of T cell signalling, simulating it using the CGMD approach can potentially

reveal intricate structural and dynamic aspects such as patterns in protein-protein interactions / binding sites.

To support this study, some preliminary simulations have been performed. Two individual simulations were run for 25 μs each where 16 TCR-CD3 complexes were simulated starting with their cytoplasmic tails in an extended conformation (Fig 7.1A; as done in chapter 5 with a single complex) allowing each TCR-CD3 to dynamically embrace a random CYR conformation. In one simulation, the proteins are embedded in a complex asymmetric bilayer (Fig 7.1B) (20812 lipids; ~77 $nm^2$) whose lipid headgroup composition was derived from a lipidomics study of the T cell activation membrane domain [210]. This composition was referred to as 'membrane 1' in chapter 5. In the other simulation, the membrane is a pure POPC bilayer (Fig 7.1C) (18500 lipids; ~75 $nm^2$). Analysing these simulations will provide insights into (i) the influence that lipid composition and membrane curvature may have on TCR-CD3 clustering and organisation, (ii) large-scale lipid organisation around TCR-CD3 clusters, (iii) possible patterns in protein-protein contacts, etc. Markov state models may also be used in analysing the dynamics of the protein oligomerisation process, as recently done for a class A G-protein-coupled receptor [313].

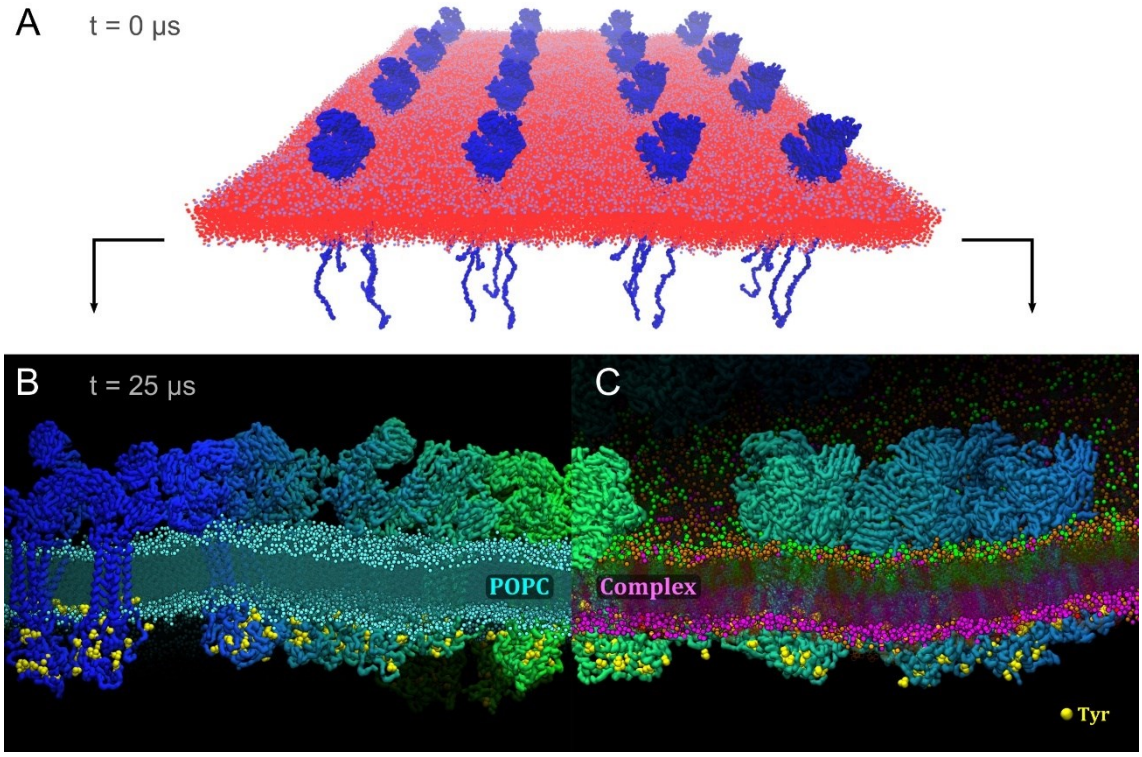


***Fig 7.1 TCR-CD3 clustering snapshots*** *(A) extracted from the beginning of the simulation, (B) from the end of the POPC bilayer simulation, and (C) the end of the complex bilayer simulation. Membrane curvature can be observed behind the cluster of TCR-CD3 complexes in (C) in contrast to (B).*

In chapter 5, I showed that PIP lipids form an anionic environment around the TCR-CD3 complex. Similarly, in the TCR-CD3 clustering simulations with the complex bilayer, I would expect to see clustering of anionic lipids on a larger scale forming an outline around groups of receptors. More importantly, the novel information that these simulations can provide is the protein-protein interactions between each TCR-CD3 complex and possibly a pattern suggesting preferential binding site(s).

## 7.4 Glycosylations in the TCR-CD3 complex

The role of glycosylations i.e. the covalent attachment of carbohydrate moieties to amino acids, have been studied with relevance to immune receptor-mediated responses and cancer immunotherapy [314–316]. Glycosylation of antigenic complexes [276] can also impact T cell activation and response [317–319]. Further, glycosylations of antigenic fragments have been reported to assist in their escape from being recognised by our immune system [318,319]. Given that it primarily plays a role in cell surface localisation and organisation [320], it is worth considering simulating clustering of glycosylated TCR-CD3 complexes in a complex bilayer in addition to previous clustering simulations. This can be compared with the non-glycosylated TCR-CD3 clustering (complex bilayer) simulation to reveal the influence of glycosylations on TCR-CD3 organisation in molecular detail.

CGMD simulations enable large-scale simulations involving sufficient time-scales. For this reason, the modelling and simulation of clustering of glycosylated TCR-CD3 complexes demand coarse-grained parameters of glycans. Recently, the Martini forcefield has been extended to include parameters for N-linked glycans [231]. Therefore, using these parameters and the information on the location of glycosylations (N-acetyl glucosamine Asn modification for all) provided by the cryo-EM study of the TCR-CD3 [16] (PDB:6JXR), I constructed a CG representation of the same cryo-EM-derived structure (Fig 7.2). This was done following the addition of parameters into the martini_v2.2_aminoacids.itp and martinize.py files (updated files publicly available here: https://github.com/DJ004/martini_mod). Similarly, future work can involve constructing a coarse-grained representation of the glycosylated form of the complete TCR-CD3 complex to use it for clustering simulations.

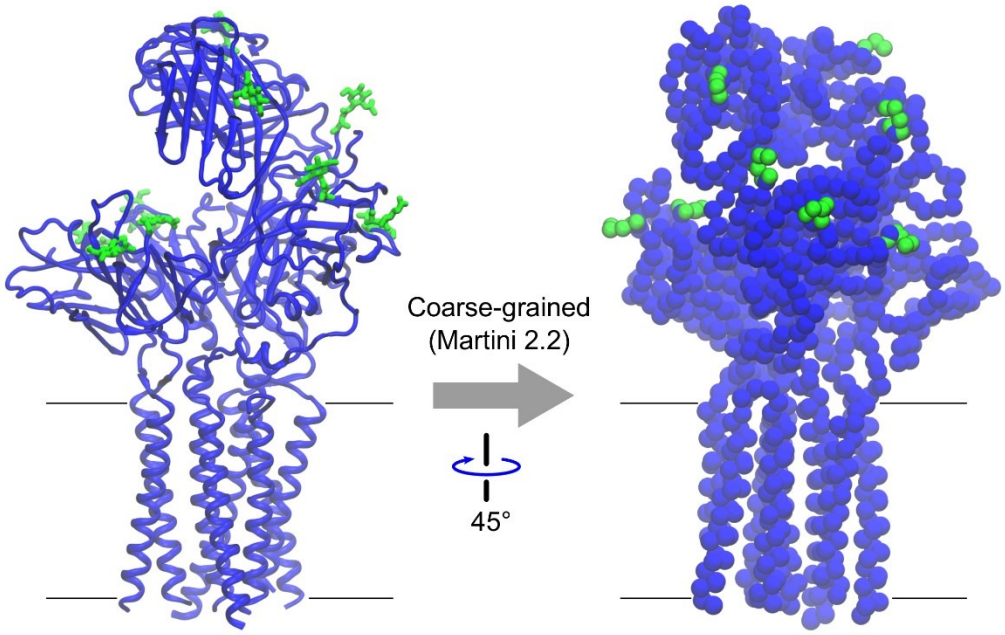


***Fig 7.2 The glycosylated form of the TCR-CD3 cryo-EM structure.*** *The backbone of the atomistic (left) and coarse-grained (right) structures is indicated in blue and N-acetyl glucosamine Asn glycosylations in green.*

In addition to the N-acetyl glucosamine Asn parameters, future clustering simulations may also involve more complex glycosylations, as parameterised by Shivgan *et al.* [231], to determine if the size of glycosylations influences TCR-CD3 binding and organisation. Following identification of potential protein-protein binding sites from CGMD simulations, they can be evaluated in atomistic resolution. This can be achieved by first backmapping [205,321] all protein-protein bound conformations that represent potential binding sites and then employing constant velocity SMD to determine the strength or the rate of dissociation of each binding site, thereby providing insights on the probability of each binding site during TCR-CD3 oligomerisation. A further step to this protocol would be to perform umbrella sampling and calculate their free energies of unbinding. It is also possible that TCR-CD3 complexes non-specifically bind to each other. In that case, the above proposed protocol can be used to determine whether glycosylations have a significant impact on TCR-CD3 interactions irrespective of binding site.

## 7.5 Testing the TCR-CD3 aggregation model

The TCR-CD3 aggregation model, as described in section 1.4 in chapter 1, suggests that TCR-CD3 clustering aids in the dissociation of ITAMs by reducing the interactions between the TCR-CD3 CYR and the plasma membrane.

However, this has not been studied in molecular detail. Using the complete TCR-CD3 model as proposed in chapter 5 of this thesis, SMD simulations may again be performed to test the aggregation model. More specifically, as shown in Fig 7.3, a pulling force may be applied on the TMR backbones of multiple TCR-CD3 complexes steering them to collide with a single TCR-CD3 toward the centre of the membrane. Such a simulation conducted in atomistic resolution is expected to show whether the forces of TCR-CD3 aggregation can override the electrostatic forces between the CYR and the membrane and aid in the solvent accessibility of ITAMs.

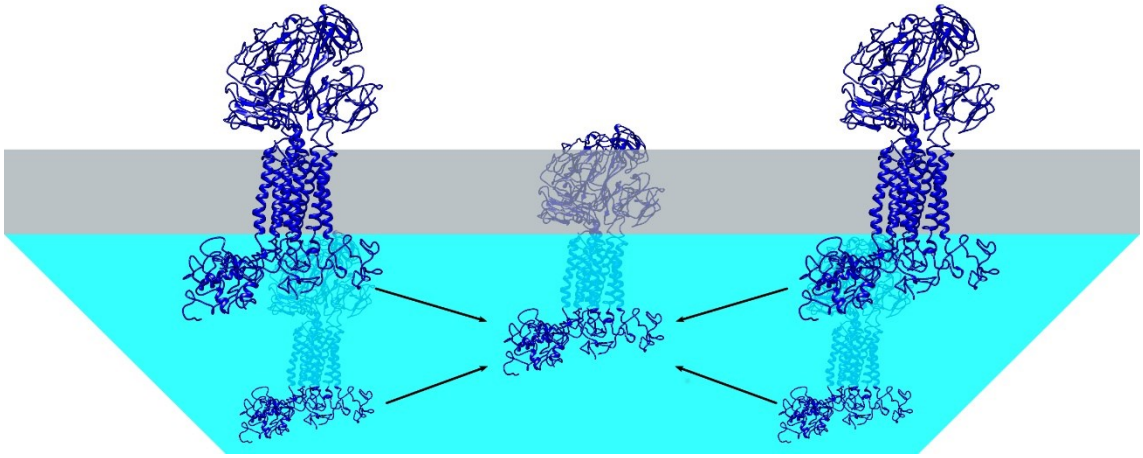

***Fig 7.3 Steered MD simulations to test the TCR-CD3 aggregation model.*** *The grey and blue areas represent the membrane and the surface of its inner leaflet respectively. The arrows indicate the direction of pulling the four TCR-CD3 complexes toward the central one.*

## 7.6 Clustering of LCK and mutations in the kinase

Following experimental evidence on the correlation between clustering of LCK and its open conformational state [117,307], multiple full-length LCKs can be simulated to determine the influence of their clustering on their conformational state in molecular detail. The quantum-level phosphorylation mechanisms defining the switch between open and closed LCK conformations cannot be simulated for reasons concerning the computational cost. Therefore, CGMD simulations can be used to simulate multiple LCKs in their active state. Moreover, LCK is known to exhibit an active state prior to TCR-pMHC ligation [114]. By simulating its clustering, it will be interesting to study LCK-LCK interactions that potentially maintain their active state and influence their organisation, and also observe potential conformational changes upon binding. Alternatively, to avoid simulating large time-scales in coarse-grained resolution,

steered MD simulations can be conducted in atomistic resolution to drive LCK clustering (similar to Fig 7.3).

In chapter 6, my CGMD simulations revealed lipid interaction sites in the unique domain and kinase domain of LCK. While the unique domain showed significant interactions with anionic PIP lipid headgroups via a couple of residues (R39, R45), the kinase domain contacted PIP lipids via a broader range of residues which constituted a distinct flat cationic patch. This cationic patch is proposed to largely drive the proximity of the LCK kinase domain to the membrane. Given that the TCR-CD3 CYR also occurs proximal to the membrane, as shown in chapter 5, ITAM phosphorylation is likely to occur close to the membrane. In contrast, due to the lack of completely resolved experimental structures of the TCR-CD3 and LCK, schematic illustrations often assume that activated TCR-CD3 complexes stretch their cytoplasmic tails into the intracellular region where it is phosphorylated by the LCK kinase. In addition, the kinase domain is often understood to float in the intracellular region while the SH2 domain strongly interacts with the membrane [27]. The idea that the kinase domain interacts with the membrane and may further aid in ITAM phosphorylation, suggested by my simulations, may be tested by neutralising the cationic patch of the kinase via experiments. Consequently, a reduction in ITAM phosphorylation may then support the hypothesis proposed here.

## 7.7 TCR-CD3 / LCK interactions with the actin filament

Upon antigen recognition, in the T cell, it is known that actin rapidly polymerises toward the activation site and thereby helps in localising signalling proteins along with TCR-CD3 complexes at the immunological synapse [322,323]. Structural data on the monomeric / globular actin (G-actin) [324] (1.54 Å, PDB:1J6Z) and multimeric / filamentous actin (F-actin) [325,326] have enabled MD simulation studies on nanosecond time-scales [327–329]. However, note that the interactions of G- and F-actin with TCR-CD3 and/or LCK have not been previously studied in molecular detail. Therefore, using the available crystal structure, F-actin can be modelled and simulated in a membrane environment containing a single / multiple TCR-CD3 complexes and/or LCK molecules. This may reveal protein-protein interactions of the TCR-CD3 / LCK with actin, and also enable molecular-level studies on the conformational impact of a moving F-actin molecule while in contact with the TCR-CD3. However, it is important to consider that, while it is feasible to perform CGMD simulations using the Martini forcefield to study F-actin in a membrane environment containing a single TCR-

CD3 / LCK, a larger system containing F-actin along with many TCR-CD3 complexes and/or LCK molecules may represent a more realistic scenario but with a substantial rise in computational cost.

## 7.8 Summary and concluding remarks

To summarise this thesis, I have employed coarse-grained and atomistic MD simulations to study and provide molecular insights into the structure and function of the complete TCR-CD3 complex and the full-length LCK in a membrane environment. Although my simulations reveal the dynamics and flexibility of the TCR-CD3 ECD (in chapter 5), the precise chain of events that transmit pMHC-induced signals to the TMR remains obscure. Nevertheless, the conformational changes that occur upon mutations in the TCR-CD3 TMR, as described in chapter 3, strongly suggest the signalling mechanism from the TMR to the CYR. Furthermore, my findings on the lipid interactions of the TCR-CD3 and LCK (chapters 5 and 6) together suggest that a pool of PIP lipids surrounding TCR-CD3 complexes may aid in recruiting SH2 domain-containing kinases to the activation site. Moreover unlike PS lipids that can flip into the outer leaflet of the plasma membrane with the help of TMEM16F scramblases [330], PIP lipids have not been reported to do so. Therefore, due to the presence of PIPs in the inner leaflet, TCR-CD3 complexes may not completely disengage from the membrane initially but instead undergo conformational changes to sufficiently expose their ITAMs upon pMHC binding. This suggests that the initial phosphorylation of ITAMs by LCK occurs proximal to the membrane, followed by spatial organisation and protein clustering which then amplify ITAM exposure and phosphorylation. Additionally, the influx of $Ca^{2+}$ cations neutralise anionic phospholipid headgroups and contribute to the dissociation of the TCR-CD3 CYR from the membrane [32,35,331,332]. However, it is unlikely that its CYR exhibits a linear configuration perpendicular to the membrane, as commonly illustrated schematically [310]. Preliminary TCR-CD3 clustering simulations indicate that their CYRs maintain a coiled state despite protein clustering and neutralisation of membrane charge (as shown in Fig 7.1B, in a POPC bilayer). This is potentially due to protein-protein interactions within the TCR-CD3 CYR. Therefore, I would expect LCK to also exhibit a membrane-proximal orientation to be able to access the ITAMs throughout the signalling process.

This work provides a methodological example to conduct computational research on other immunoreceptors such as B cell receptors and natural killer

(NK) cell receptors to explore their complete structures, dynamics, and lipid interactions in a membrane. While atomistic MD forcefields such as CHARMM36 [214], as used in this thesis, may serve as a more accurate representation of biomolecular dynamics, they are limited to simulations on shorter time-scales. On the other hand, Martini 2.2 [251] has been a promising CGMD forcefield to sample larger time-scales enabling studies of a wider range of biochemical phenomena. However, not only does it limit the resolution and not provide hydrogen bond information, but it is also suggested to exaggerate protein-protein interactions [333]. Nonetheless, it has been considerably improved in Martini 3 [334], which is recommended for future CGMD simulations of membrane proteins.

Exploring the spatial organisation of various membrane proteins with membrane-associated molecules in a larger membrane environment including actin cytoskeletal dynamics is likely to become of interest, and something that can be achievable with ultra-coarse-graining (UCG) [335–337]. However, this approach as of today probably requires more extensive studies and better parameterisation to be used as a general purpose UCG forcefield. In addition, building triangulated surfaces of organelle-sized lipid vesicles (such as ~80 million particles), backmapping them to coarse-grained resolution, and further simulating their dynamics in implicit solvent (Dry Martini) has paved way to mesoscale simulations [208,338], but yet with a constraint on time-scale. Nevertheless, with the development of such new approaches along with improvisation of existing methods, assistance of Markov state models, machine learning-based analyses, and robust computing and parallelisation techniques, MD simulations have become a strong source of providing molecular-level insights and hypotheses. Although they may cost us a certain degree of uncertainty, the range of discoveries and methodological inventions that emerge from this field is truly immense. Inspired by that, I have carried out this work as a contribution toward deciphering some structural and dynamic aspects of the initiation of T cell activation, and in a broader sense, toward understanding the nature of our immune system.

# References


1. Goldsby RA, Kindt TJ, Osborne BA, Kuby J. Immunology 5th edition. W.H. 2003.
2. Cano RLE, Lopera HDE. Introduction to T and B lymphocytes. Autoimmunity: From Bench to Bedside. El Rosario University Press; 2013. Available: https://www.ncbi.nlm.nih.gov/books/NBK459471/
3. Charles A Janeway J, Travers P, Walport M, Shlomchik MJ. Antigen Recognition by B-cell and T-cell Receptors. Immunobiology: The Immune System in Health and Disease 5th edition. Garland Science; 2001. Available: https://www.ncbi.nlm.nih.gov/books/NBK10770/
4. Mora T, Walczak AM. Quantifying lymphocyte receptor diversity. Syst Immunol. 2018; 183–198.
5. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. Nat Rev Immunol. 2011;11: 251–263. doi:10.1038/nri2941
6. Mandl JN, Monteiro JP, Vrisekoop N, Germain RN. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. Immunity. 2013;38: 263–274. doi:10.1016/J.IMMUNI.2012.09.011
7. Wieczorek M, Abualrous ET, Sticht J, Álvaro-Benito M, Stolzenberg S, Noé F, et al. Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. Front Immunol. 2017;8: 292. doi:10.3389/FIMMU.2017.00292/BIBTEX
8. Call ME, Pyrdol J, Wiedmann M, Wucherpfennig KW. The organizing principle in the formation of the T cell receptor-CD3 complex. Cell. 2002;111: 967–979. doi:10.1016/S0092-8674(02)01194-7
9. Mariuzza RA, Agnihotri P, Orban J. The structural basis of T-cell receptor (TCR) activation: An enduring enigma. Journal of Biological Chemistry. American Society for Biochemistry and Molecular Biology Inc.; 2019. pp. 914–925. doi:10.1074/jbc.REV119.009411
10. Weissman AM, Hou D, Orloff DG, Modi WS, Seuanez H, O'Brien SJ, et al. Molecular cloning and chromosomal localization of the human T-cell receptor zeta chain: distinction from the molecular CD3 complex. Proc Natl Acad Sci U S A. 1988;85: 9709–13. doi:10.1073/pnas.85.24.9709
11. Bonifacino JS, Chen C, Lippincott-Schwartz J, Ashwell JD, Klausner RD. Subunit interactions within the T-cell antigen receptor: clues from the study of partial complexes. Proc Natl Acad Sci U S A. 1988;85: 6929–33. doi:10.1073/pnas.85.18.6929
12. van der Merwe PA, Dushek O. Mechanisms for T cell receptor triggering. Nat Rev Immunol. 2011;11: 47–55. doi:10.1038/nri2887
13. YH D, BM B, DN G, WE B, DC W. Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. Immunity. 1999;11: 45–56. doi:10.1016/S1074-7613(00)80080-1
14. JL C, G S-J, G B, NM L, L W, EM C, et al. Structural and kinetic basis for heightened immunogenicity of T cell vaccines. J Exp Med. 2005;201:

1243–1255. doi:10.1084/JEM.20042323

15. ME B, JL M, DK S, S D, J G, J D, et al. Deconstructing the peptide-MHC specificity of T cell recognition. Cell. 2014;157: 1073–1087. doi:10.1016/J.CELL.2014.03.047

16. Dong D, Zheng L, Lin J, Zhang B, Zhu Y, Li N, et al. Structural basis of assembly of the human TCR–CD3 complex. Nature. 2019;573: 546–552. doi:10.1038/s41586-019-1537-0

17. Sigalov AB, Aivazian DA, Uversky VN, Stern LJ. Lipid-binding activity of intrinsically unstructured cytoplasmic domains of multichain immune recognition receptor signaling subunits. Biochemistry. 2006;45: 15731–15739. doi:10.1021/bi061108f

18. Yamaguchi H, Hendrickson WA. Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. Nature. 1996;384: 484. doi:10.1038/384484a0

19. Boggon TJ, Eck MJ. Structure and regulation of Src family kinases. Oncogene. 2004. pp. 7918–7927. doi:10.1038/sj.onc.1208081

20. Stirnweiss A, Hartig R, Gieseler S, Lindquist JA, Reichardt P, Philipsen L, et al. T cell activation results in conformational changes in the Src family kinase Lck to induce its activation. Sci Signal. 2013;6. doi:10.1126/scisignal.2003607

21. Koch CA, Anderson D, Moran MF, Ellis C, Pawson T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science. 1991;252: 668–74. doi:10.1126/science.1708916

22. Straus DB, Chan AC, Patai B, Weiss A. SH2 domain function is essential for the role of the Lck tyrosine kinase in T cell receptor signal transduction. J Biol Chem. Chicago, Illinois, USA; 1996. doi:10.1074/jbc.271.17.9976

23. Chakraborty AK, Weiss A. Insights into the initiation of TCR signaling. Nat Immunol. 2014;15: 798–807. doi:10.1038/ni.2940

24. Li L, Guo X, Shi X, Li C, Wu W, Yan C, et al. Ionic CD3−Lck interaction regulates the initiation of T-cell receptor signaling. Proc Natl Acad Sci U S A. 2017;114: E5891–E5899. doi:10.1073/pnas.1701990114

25. Koegl M, Zlatkine P, Ley SC, Courtneidge SA, Magee AI. Palmitoylation of multiple Src-family kinases at a homologous N-terminal motif. Biochem J. 1994;303: 749. doi:10.1042/BJ3030749

26. MD R. Myristylation and palmitylation of Src family members: the fats of the matter. Cell. 1994;76: 411–413. doi:10.1016/0092-8674(94)90104-X

27. Sheng R, Jung DJ, Silkov A, Kim H, Singaram I, Wang ZG, et al. Lipids regulate Lck protein activity through their interactions with the Lck Src homology 2 domain. J Biol Chem. 2016;291: 17639–17650. doi:10.1074/jbc.M116.720284

28. DeFord-Watts LM, Tassin TC, Becker AM, Medeiros JJ, Albanesi JP, Love PE, et al. The Cytoplasmic Tail of the T Cell Receptor CD3 ε Subunit Contains a Phospholipid-Binding Motif that Regulates T Cell Functions. J Immunol. 2009;183: 1055–1064.

doi:10.4049/jimmunol.0900404

29. DeFord-Watts LM, Dougall DS, Belkaya S, Johnson BA, Eitson JL, Roybal KT, et al. The CD3 zeta subunit contains a phosphoinositide-binding motif that is required for the stable accumulation of TCR-CD3 complex at the immunological synapse. J Immunol. 2011;186: 6839–47. doi:10.4049/jimmunol.1002721

30. Zhang H, Cordoba S-P, Dushek O, Anton van der Merwe P. Basic residues in the T-cell receptor ζ cytoplasmic domain mediate membrane association and modulate signaling. Proc Natl Acad Sci. 2011;108: 19323–8. doi:10.1073/pnas.1108052108

31. Chouaki-Benmansour N, Ruminski K, Sartre AM, Phelipot MC, Salles A, Bergot E, et al. Phosphoinositides regulate the TCR/CD3 complex membrane dynamics and activation. Sci Rep. 2018;8. doi:10.1038/s41598-018-23109-8

32. Li L, Shi X, Guo X, Li H, Xu C. Ionic protein–lipid interaction at the plasma membrane: what can the charge do? Trends Biochem Sci. 2014;39: 130–140. doi:10.1016/J.TIBS.2014.01.002

33. Xu C, Gagnon E, Call ME, Schnell JR, Schwieters CD, Carman C V, et al. Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. Cell. 2008;135: 702–13. doi:10.1016/j.cell.2008.09.044

34. Zimmermann K, Eells R, Heinrich F, Rintoul S, Josey B, Shekhar P, et al. The cytosolic domain of T-cell receptor associates with membranes in a dynamic equilibrium and deeply penetrates the bilayer. J Biol Chem. 2017;292: 17746–17759. doi:10.1074/jbc.M117.794370

35. Wu W, Shi X, Xu C. Regulation of T cell signalling by membrane lipids. Nature Reviews Immunology. Nature Publishing Group; 2016. pp. 690–701. doi:10.1038/nri.2016.103

36. Call ME, Pyrdol J, Wucherpfennig KW. Stoichiometry of the T-cell receptor-CD3 complex and key intermediates assembled in the endoplasmic reticulum. EMBO J. 2004;23: 2348–2357. doi:10.1038/sj.emboj.7600245

37. Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, Peterson PA, et al. An αβ T Cell Receptor Structure at 2.5 Å and Its Orientation in the TCR-MHC Complex. Science (80- ). 1996;274: 209–219. doi:10.1126/SCIENCE.274.5285.209

38. Bäckström BT, E M, Peter A, Jaureguiberry B, Baldari CT, Palmer E. A motif within the T cell receptor alpha chain constant region connecting peptide domain controls antigen responsiveness. Immunity. 1996;5: 437–447. doi:10.1016/S1074-7613(00)80500-2

39. T W, LP K, J Y, GR C, C T. Associations between subunit ectodomains promote T cell antigen receptor assembly and protect against degradation in the ER. J Cell Biol. 1993;122: 67–78. doi:10.1083/JCB.122.1.67

40. Kuhns MS, Davis MM, Garcia KC. Deconstructing the form and function of the TCR/CD3 complex. Immunity. 2006;24: 133–139. doi:10.1016/j.immuni.2006.01.006

41. Kuhns MS, Girvin AT, Klein LO, Chen R, Jensen KDC, Newell EW, et al. Evidence for a functional sidedness to the αβTCR. Proc Natl Acad Sci. 2010;107: 5094–5099. doi:10.1073/PNAS.1000925107

42. Sasada T, Touma M, Chang H-C, Clayton LK, Wang J, Reinherz EL. Involvement of the TCR Cβ FG Loop in Thymic Selection and T Cell Function. J Exp Med. 2002;195: 1419–1431. doi:10.1084/jem.20020119

43. Touma M, Chang H-C, Sasada T, Handley M, Clayton LK, Reinherz EL. The TCR C beta FG Loop Regulates alpha beta T Cell Development. J Immunol. 2006;176: 6812–6823. doi:10.4049/jimmunol.176.11.6812

44. Das DK, Feng Y, Mallis RJ, Li X, Keskin DB, Hussey RE, et al. Force-dependent transition in the T-cell receptor β-subunit allosterically regulates peptide discrimination and pMHC bond lifetime. Proc Nati Acad Sci USA. 2015;112: 1517–22. doi:10.1073/pnas.1424829112

45. Rangarajan S, He Y, Chen Y, Kerzic MC, Ma B, Gowthaman R, et al. Peptide–MHC (pMHC) binding to a human antiviral T cell receptor induces long-range allosteric communication between pMHC- and CD3-binding sites. J Biol Chem. 2018;293: 15991–16005. doi:10.1074/jbc.RA118.003832

46. Hwang W, Mallis RJ, Lang MJ, Reinherz EL. The αβTCR mechanosensor exploits dynamic ectodomain allostery to optimize its ligand recognition site. Proc Natl Acad Sci U S A. 2020;117: 21336–21345. doi:10.1073/pnas.2005899117

47. Beddoe T, Chen Z, Clements CS, Ely LK, Bushell SR, Vivian JP, et al. Antigen Ligation Triggers a Conformational Change within the Constant Domain of the αβ T Cell Receptor. Immunity. 2009;30: 777–788. doi:10.1016/j.immuni.2009.03.018

48. Xu C, Call ME, Wucherpfennig KW. A membrane-proximal tetracysteine motif contributes to assembly of CD3δε and CD3γε dimers with the T cell receptor. J Biol Chem. 2006;281: 36977–36984. doi:10.1074/jbc.M607164200

49. Krshnan L, Park S, Im W, Call MJ, Call ME. A conserved αβ transmembrane interface forms the core of a compact T-cell receptor-CD3 structure within the membrane. Proc Natl Acad Sci. 2016;113: 6649–58. doi:10.1073/pnas.1611445113

50. Arnaud J, Hucheng A, Vernhes M, Caspar-Bauguil S, Lenfant F, Sancho J, et al. The interchain disulfide bond between TCR alpha beta heterodimers on human T cells is not required for TCR-CD3 membrane expression and signal transduction. Int Immunol. 1997;9: 615–626. doi:10.1093/intimm/9.4.615

51. Iqbalsyah TM, Moutevelis E, Warwicker J, Errington N, Doig AJ. The CXXC motif at the N terminus of an α-helical peptide. Protein Sci. 2006;15: 1945. doi:10.1110/PS.062271506

52. S I, ST K, S T, M G, RM B, C R, et al. Disulfide conformation and design at helix N-termini. Proteins. 2010;78: 1228–1242. doi:10.1002/PROT.22641

53. Brazin KN, Mallis RJ, Li C, Keskin DB, Arthanari H, Gao Y, et al.

Constitutively Oxidized CxxC Motifs within the CD3 Heterodimeric Ectodomains of the T Cell Receptor Complex Enforce the Conformation of Juxtaposed Segments. J Biol Chem. 2014;289: 18880–18892. doi:10.1074/jbc.M114.574996

54. JJ S, JS B, J L-S, AM W, T S, RD K, et al. Failure to synthesize the T cell CD3-zeta chain: structure and function of a partial T cell receptor complex. Cell. 1988;52: 85–95. doi:10.1016/0092-8674(88)90533-8

55. Geisler C, Kuhlmann J, Rubin B. Assembly, intracellular processing, and expression at the cell surface of the human alpha beta T cell receptor/CD3 complex. Function of the CD3-zeta chain. J Immunol. 1989;143: 4069–77. Available: http://www.jimmunol.org/content/143/12/4069

56. Weissman AM, Frank SJ, Orloff DG, Merćep M, Ashwell JD, Klausner RD. Role of the zeta chain in the expression of the T cell antigen receptor: genetic reconstitution studies. EMBO J. 1989;8: 3651–3656. doi:10.1002/J.1460-2075.1989.TB08539.X

57. Sharma S, Juffer AH. An atomistic model for assembly of transmembrane domain of T cell receptor complex. J Am Chem Soc. 2013;135: 2188–2197. doi:10.1021/ja308413e

58. Call ME, Wucherpfennig KW. Common themes in the assembly and architecture of activating immune receptors. Nat Rev Immunol. 2007;7: 841–850. doi:10.1038/nri2186

59. Call ME, Wucherpfennig KW, Chou JJ. The structural basis for intramembrane assembly of an activating immunoreceptor complex. Nat Immunol. 2010;11: 1023–1029. doi:10.1038/ni.1943

60. Dube N, Marzinek JK, Glen RC, Bond PJ. The structural basis for membrane assembly of immunoreceptor signalling complexes. J Mol Model. 2019;25. doi:10.1007/s00894-019-4165-6

61. AM W, X H, J D, C G. Distinct domains of the CD3-gamma chain are involved in surface expression and function of the T cell antigen receptor. J Biol Chem. 1995;270: 4675–4680. doi:10.1074/JBC.270.9.4675

62. Dietrich J, Neisig A, Hou X, Wegener A-MK, Gajhede M, Geisler C. Role of CD3 gamma in T cell receptor assembly. J Cell Biol. 1996;132: 299–310. doi:10.1083/jcb.132.3.299

63. Call ME, Schnell JR, Xu C, Lutz RA, Chou JJ, Wucherpfennig KW. The Structure of the ζζ Transmembrane Dimer Reveals Features Essential for Its Assembly with the T Cell Receptor. Cell. 2006;127: 355–368. doi:10.1016/j.cell.2006.08.044

64. Nicholas Manolios, Marina Ali VB. T-cell antigen receptor (TCR) transmembrane peptides: A new paradigm for the treatment of autoimmune diseases. Cell Adh Migr. 2010;4: 273–283. doi:10.4161/cam.4.2.11909

65. Kuhns MS, Badgandi HB. Piecing together the family portrait of TCR-CD3 complexes. Immunol Rev. 2012;250: 120–143. doi:10.1111/imr.12000

66. Ishikura S, Weissman AM, Bonifacino JS. Serine Residues in the Cytosolic Tail of the T-cell Antigen Receptor-Chain Mediate Ubiquitination

and Endoplasmic Reticulum-associated Degradation of the Unassembled Protein. J Biol Chem. 2010;285: 23916–23924. doi:10.1074/jbc.M110.127936

67. Wucherpfennig KW, Gagnon E, Call MJ, Huseby ES, Call ME. Structural biology of the T-cell receptor: insights into receptor assembly, ligand recognition, and initiation of signaling. Cold Spring Harb Perspect Biol. 2010;2: a005140–a005140. doi:10.1101/cshperspect.a005140

68. Smith-garvin JE, Koretzky GA, Jordan MS. T Cell Activation. Immunology. 2009;27: 591–619. doi:10.1146/annurev.immunol.021908.132706.T

69. Kuhns MS, Davis MM. The Safety on the TCR Trigger. Cell. 2008;135: 594–596. doi:10.1016/J.CELL.2008.10.033

70. Davis SJ, van der Merwe A. The kinetic-segregation model: TCR triggering and beyond. Nat Immunol. 2006;7: 803–809. doi:10.1038/ni1369

71. Junghans V, Santos AM, Lui Y, Davis SJ, Jönsson P. Dimensions and Interactions of Large T-Cell Surface Proteins. Frontiers in immunology. NLM (Medline); 2018. p. 2215. doi:10.3389/fimmu.2018.02215

72. Varma R. TCR triggering by the pMHC complex: valency, affinity, and dynamics. Sci Signal. 2008;1. doi:10.1126/stke119pe21

73. Sun Z-YJ, Kim KS, Wagner G, Reinherz EL. Mechanisms Contributing to T Cell Receptor Signaling and Assembly Revealed by the Solution Structure of an Ectodomain Fragment of the CD3εγ Heterodimer. Cell. 2001;105: 913–923. doi:10.1016/s0092-8674(01)00395-6

74. Natarajan K, McShan AC, Jiang J, Kumirov VK, Wang R, Zhao H, et al. An allosteric site in the T-cell receptor Cβ domain plays a critical signalling role. Nat Commun. 2017;8. doi:10.1038/ncomms15260

75. He Y, Agnihotri P, Rangarajan S, Chen Y, Kerzic MC, Ma B, et al. Peptide–MHC Binding Reveals Conserved Allosteric Sites in MHC Class I- and Class II-Restricted T Cell Receptors (TCRs). J Mol Biol. 2020;432: 166697. doi:10.1016/j.jmb.2020.10.031

76. AL L, G M, N P, A C, D C, D P, et al. Allosteric activation of T cell antigen receptor signaling by quaternary structure relaxation. Cell Rep. 2021;36: 109375. doi:10.1016/J.CELREP.2021.109375

77. Ma Z, Janmey PA, Finkel TH. The receptor deformation model of TCR triggering. FASEB. 2008;22: 1002–1008. doi:10.1096/fj.07-9331hyp

78. Kim ST, Takeuchi K, Sun ZYJ, Touma M, Castro CE, Fahmy A, et al. The αβ T cell receptor is an anisotropic mechanosensor. J Biol Chem. 2009;284: 31028–31037. doi:10.1074/jbc.M109.052712

79. Brazin KN, Mallis RJ, Das DK, Feng Y, Hwang W, Wang J, et al. Structural Features of the αβTCR Mechanotransduction Apparatus That Promote pMHC Discrimination. Front Immunol. 2015;6: 441. doi:10.3389/FIMMU.2015.00441

80. Feng Y, Brazin KN, Kobayashi E, Mallis RJ, Reinherz EL, Lang MJ. Mechanosensing drives acuity of αβ T-cell recognition. Proc Natl Acad Sci U S A. 2017;114: E8204–E8213. doi:10.1073/pnas.1703559114

81. Minguet S, Schamel WWA. A permissive geometry model for TCR-CD3 activation. Trends Biochem Sci. 2008;33: 51–57. doi:10.1016/j.tibs.2007.10.008

82. Krogsgaard M, Li Q-J, Sumen C, Huppa JB, Huse M, Davis MM. Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. Nature. 2005;434: 238–43. doi:10.1038/nature03362

83. Aivazian D, Stern LJ. Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition. Nat Struct Biol. 2000;7: 1023–6. doi:doi.org/10.1038/80930

84. Goyette J, Nieves DJ, Ma Y, Gaus K. How does T cell receptor clustering impact on signal transduction? J Cell Sci. 2019;132. doi:10.1242/jcs.226423

85. Brameshuber M, Kellner F, Rossboth BK, Ta H, Alge K, Sevcsik E, et al. Monomeric TCRs drive T cell antigen recognition. Nat Immunol. 2018;19: 487–496. doi:10.1038/s41590-018-0092-4

86. James JR, Vale RD. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. Nature. 2012;487: 64–69. doi:10.1038/nature11220

87. Rossy J, Williamson DJ, Gaus K. How does the kinase Lck phosphorylate the T cell receptor? Spatial organization as a regulatory mechanism. Frontiers in Immunology. 2012. doi:10.3389/fimmu.2012.00167

88. Wingfield P. N-Terminal Methionine Processing. Curr Protoc protein Sci. 2017;88: 6.14.1. doi:10.1002/CPPS.29

89. Udenwobele DI, Su RC, Good S V., Ball TB, Shrivastav SV, Shrivastav A. Myristoylation: An important protein modification in the immune response. Frontiers in Immunology. Frontiers Media S.A.; 2017. p. 1. doi:10.3389/fimmu.2017.00751

90. Yasuda K, Kosugi A, Hayashi F, Saitoh S, Nagafuku M, Mori Y, et al. Serine 6 of Lck tyrosine kinase: a critical site for Lck myristoylation, membrane localization, and function in T lymphocytes. J Immunol. 2000;165: 3226–31. doi:10.4049/jimmunol.165.6.3226

91. PS K, AI M, SC L. S-acylation of LCK protein tyrosine kinase is essential for its signalling function in T lymphocytes. EMBO J. 1997;16: 4983–4998. doi:10.1093/EMBOJ/16.16.4983

92. Shenoy-Scaria AM, Gauen LK, Kwong J, Shaw AS, Lublin DM. Palmitylation of an amino-terminal cysteine motif of protein tyrosine kinases p56lck and p59fyn mediates interaction with glycosyl-phosphatidylinositol-anchored proteins. Mol Cell Biol. 1993;13: 6385. doi:10.1128/MCB.13.10.6385

93. Amata I, Maffei M, Pons M. Phosphorylation of unique domains of Src family kinases. Front Genet. 2014;5: 181. doi:10.3389/FGENE.2014.00181/BIBTEX

94. JM T, MH B, BA I, SD L, RM P, DR L. Interaction of the unique N-terminal region of tyrosine kinase p56lck with cytoplasmic domains of CD4 and CD8 is mediated by cysteine motifs. Cell. 1990;60: 755–765. doi:10.1016/0092-8674(90)90090-2

95. AS S, J C, JA W, C H, KE A, P K, et al. Short related sequences in the cytoplasmic domains of CD4 and CD8 mediate binding to the amino-terminal domain of the p56lck tyrosine protein kinase. Mol Cell Biol. 1990;10: 1853–1862. doi:10.1128/MCB.10.5.1853-1862.1990

96. Kim PW, Sun ZYJ, Blacklow SC, Wagner G, Eck MJ. A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8. Science (80- ). 2003;301: 1725–1728. doi:10.1126/science.1085643

97. RS L, C R, A V, HF L. Zinc is essential for binding of p56(lck) to CD4 and CD8alpha. J Biol Chem. 1998;273: 32878–32882. doi:10.1074/JBC.273.49.32878

98. M H, MJ E, SC H. A Zn2+ ion links the cytoplasmic tail of CD4 and the N-terminal region of Lck. J Biol Chem. 1998;273: 18729–18733. doi:10.1074/JBC.273.30.18729

99. Eck MJ, Atwell SK, Shoelson SE, Harrison SC. Structure of the regulatory domains of the Src-family tyrosine kinase Lck. Nature. 1994;368: 764–9. doi:10.1038/368764a0

100. MF D, HC K, AC C, DB S. The lck SH3 domain is required for activation of the mitogen-activated protein kinase pathway but not the initiation of T-cell antigen receptor signaling. J Biol Chem. 1999;274: 5146–5152. doi:10.1074/JBC.274.8.5146

101. Rudd ML, Tua-Smith A, Straus DB. Lck SH3 Domain Function Is Required for T-Cell Receptor Signals Regulating Thymocyte Development. Mol Cell Biol. 2006;26: 7892. doi:10.1128/MCB.00968-06

102. Hartl FA, Beck-Garcìa E, Woessner NM, Flachsmann LJ, Cárdenas RMHV, Brandl SM, et al. Noncanonical binding of Lck to CD3ε promotes TCR signaling and CAR function. Nat Immunol. 2020;21: 902–913. doi:10.1038/s41590-020-0732-3

103. Gil D, Schamel WWAA, Montoya M, Sá Nchez-Madrid F, Alarcó B, Sánchez-Madrid F, et al. Recruitment of Nck by CD3 Reveals a Ligand-Induced Conformational Change Essential for T Cell Receptor Signaling and Synapse Formation. Cell. 2002;109: 901–912. doi:10.1016/S0092-8674(02)00799-7

104. Li W, Fan J, Woodley DT. Nck/Dock: an adapter between cell surface receptors and the actin cytoskeleton. Oncogene 2001 2044. 2001;20: 6403–6417. doi:10.1038/sj.onc.1204782

105. Park M-J, Sheng R, Silkov A, Jung D-J, Wang Z-G, Xin Y, et al. SH2 Domains Serve as Lipid-Binding Modules for pTyr-Signaling Proteins. Mol Cell. 2016;62: 7–20. doi:10.1016/j.molcel.2016.01.027

106. Couture C, Songyang Z, Jascur T, Williams S, Tailor P, Cantley LC, et al. Regulation of the Lck SH2 domain by tyrosine phosphorylation. J Biol Chem. 1996;271: 24880–24884. doi:10.1074/jbc.271.40.24880

107. Amrein KE, Sefton BM. Mutation of a site of tyrosine phosphorylation in the lymphocyte-specific tyrosine protein kinase, p56lck, reveals its oncogenic potential in fibroblasts. Proc Nati Acad Sci USA. 1988;85: 4247–4251. doi:10.1073%2Fpnas.85.12.4247

108. Courtney AH, Amacher JF, Kadlecek TA, Mollenauer MN, Au-Yeung BB,

Kuriyan J, et al. A Phosphosite within the SH2 Domain of Lck Regulates Its Activation by CD45. Mol Cell. 2017;67: 498-511.e6. doi:10.1016/j.molcel.2017.06.024

109. Furlan G, Minowa T, Hanagata N, Kataoka-Hamai C, Kaizuka Y. Phosphatase CD45 both positively and negatively regulates T cell receptor phosphorylation in reconstituted membrane protein clusters. J Biol Chem. 2014;289: 28514–25. doi:10.1074/jbc.M114.574319

110. JA C, B H. The when and how of Src regulation. Cell. 1993;73: 1051–1054. doi:10.1016/0092-8674(93)90634-3

111. Chow LML, Fournel M, Davidson D, Veillette A. Negative regulation of T-cell receptor signalling by tyrosine protein kinase p50csk. Nature. 1993;365: 156–160. doi:10.1038/365156a0

112. D'Oro U, Ashwell JD. Cutting edge: the CD45 tyrosine phosphatase is an inhibitor of Lck activity in thymocytes. J Immunol. 1999;162: 1879–83. Available: http://www.ncbi.nlm.nih.gov/pubmed/9973453

113. Bozso SJ, Kang JJH, Nagendran J. The role of competing mechanisms on Lck regulation. Immunologic Research. Springer; 2020. pp. 289–295. doi:10.1007/s12026-020-09148-2

114. Nika K, Soldani C, Salek M, Paster W, Gray A, Etzensperger R, et al. Constitutively active lck kinase in T cells drives antigen receptor signal transduction. Immunity. 2010;32: 766–777. doi:10.1016/j.immuni.2010.05.011

115. Casas J, Brzostek J, Zarnitsyna VI, Hong J, Wei Q, Hoerter JAH, et al. Ligand-engaged TCR is triggered by Lck not associated with CD8 coreceptor. Nat Commun. 2014;5: 5624. doi:10.1038/ncomms6624

116. Wei Q, Brzostek J, Sankaran S, Casas J, Hew LSQ, Yap J, et al. Lck bound to coreceptor is less active than free Lck. Proc Natl Acad Sci U S A. 2020;117: 15809–15817. doi:10.1073/pnas.1913334117

117. Rossy J, Owen DM, Williamson DJ, Yang Z, Gaus K. Conformational states of the kinase Lck regulate clustering in early T cell signaling. Nat Immunol. 2013;14: 82–89. doi:10.1038/ni.2488

118. Fischer K, Voelkl S, Berger J, Andreesen R, Pomorski T, Mackensen A. Antigen recognition induces phosphatidylserine exposure on the cell surface of human CD8+ T cells. Blood. 2006;108: 4094–4101. doi:10.1182/BLOOD-2006-03-011742

119. Bubeck Wardenburg J, Fu C, Jackman JK, Flotow H, Wilkinson SE, Williams DH, et al. Phosphorylation of SLP-76 by the ZAP-70 Protein-tyrosine Kinase Is Required for T-cell Receptor Function. J Biol Chem. 1996;271: 19641–19644. doi:10.1074/jbc.271.33.19641

120. Finco TS, Kadlecek T, Zhang W, Samelson LE, Weiss A. LAT is required for TCR-mediated activation of PLCgamma1 and the Ras pathway. Immunity. 1998;9: 617–626. doi:10.1016/S1074-7613(00)80659-7

121. W Z, CL S, DN B, CC S, JB D, RP T, et al. Essential role of LAT in T cell development. Immunity. 1999;10: 323–332. doi:10.1016/S1074-7613(00)80032-1

122. Liu SK, Fang N, Koretzky GA, Mcglade CJ. The hematopoietic-specific adaptor protein Gads functions in T-cell signaling via interactions with the SLP-76 and LAT adaptors. Curr Biol. 1999;9: 67–75. doi:10.1016/s0960-9822(99)80017-7

123. Berg LJ, Finkelstein LD, Lucas JA, Schwartzberg PL. Tec Family Kinases in T Lymphocyte Development and Function. Annu Rev Immunol. 2005;23: 549–600. doi:10.1146/annurev.immunol.22.012703.104743

124. Shan X, Wange RL. Itk/Emt/Tsk Activation in Response to CD3 Cross-linking in Jurkat T Cells Requires ZAP-70 and Lat and Is Independent of Membrane Recruitment. J Biol Chem. 1999;274: 29323–30. doi:10.1074/jbc.274.41.29323

125. Bunnell SC, Diehn M, Yaffe MB, Findell PR, Cantley LC, Berg LJ. Biochemical Interactions Integrating Itk with the T Cell Receptor-initiated Signaling Cascade. J Biol Chem. 2000;275: 2219–2230. doi:10.1074/jbc.275.3.2219

126. Su Y-W, Zhang Y, Schweikert J, Koretzky GA, Reth M, Wienands J. Interaction of SLP adaptors with the SH2 domain of Tec family kinases. Eur J Immunol. 1999;29: 3702–3711. doi:10.1002/(SICI)1521-4141(199911)29:11<3702::AID-IMMU3702>3.0.CO;2-R

127. Pamela Schwartzberg Daniel D Billadeau and L, Wange RL, Burkhardt Labno JK, Precht P, Takesono A, Caplen NJ, et al. Kinase-independent functions for Itk in TCR-induced regulation of Vav and the actin cytoskeleton. J Immunol. 2005;174: 1385–92. doi:10.4049/jimmunol.174.3.1385

128. Reynolds LF, Smyth LA, Norton T, Freshney N, Downward J, Kioussis D, et al. Vav1 Transduces T Cell Receptor Signals to the Activation of Phospholipase C-1 via Phosphoinositide 3-Kinase-dependent and-independent Pathways. J Exp Med. 2002;195: 1103–1114. doi:10.1084/jem.20011663

129. Reynolds LF, De Bettignies C, Norton T, Beeser A, Chernoff J, Tybulewicz VLJ. Vav1 Transduces T Cell Receptor Signals to the Activation of the Ras/ERK Pathway via LAT, Sos, and RasGRP1*. J Biol Chem. 2004;279: 18239–46. doi:10.1074/jbc.M400257200

130. Qi Q, August A. Keeping the (Kinase) Party Going: SLP-76 and ITK Dance to the Beat. Sci STKE. 2007;2007: pe39–pe39. doi:10.1126/stke.3962007pe39

131. Beach D, Gonen R, Bogin Y, Reischl IG, Yablonski D. Dual Role of SLP-76 in Mediating T Cell Receptor-induced Activation of Phospholipase C-gamma1. J Biol Chem. 2007;282: 2937–46. doi:10.1074/jbc.M606697200

132. Jordan MS, Smith JE, Burns JC, Austin J-ET, Nichols KE, Aschenbrenner AC, et al. Complementation in Trans of Altered Thymocyte Development in Mice Expressing Mutant Forms of the Adaptor Molecule SLP76. Immunity. 2008;28: Immunity. doi:10.1016/j.immuni.2008.01.010

133. Miletic A V, Sakata-Sogawa K, Hiroshima M, Hamann MJ, Gomez TS, Ota N, et al. Vav1 Acidic Region Tyrosine 174 Is Required for the Formation of T Cell Receptor-induced Microclusters and Is Essential in T

Cell Development and Activation. J Biol Chem. 2006;281: 38257–65. doi:10.1074/jbc.M608913200

134. Pletneva E V, Sundd M, Fulton DB, Andreotti AH. Molecular Details of Itk Activation by Prolyl Isomerization and Phospholigand Binding: The NMR Structure of the Itk SH2 Domain Bound to a Phosphopeptide. J Mol Biol. 2006;357: 550–61. doi:10.1016/j.jmb.2005.12.073

135. Huse M. The T-cell-receptor signaling network. J Cell Sci. 2009;122: 1269–1273. doi:10.1242/jcs.042762

136. Bootman MD. Calcium signaling. Cold Spring Harb Perspect Biol. 2012;4: a011171. doi:10.1101/cshperspect.a011171

137. Ebinu JO, Stang SL, Teixeira C, Bottorff DA, Hooton J, Blumberg PM, et al. RasGRP links T-cell receptor signaling to Ras. Blood. 2000;95: 3199–203. doi:10.1182/blood.V95.10.3199

138. Roose JP, Mollenauer M, Gupta VA, Stone J, Weiss A. A Diacylglycerol-Protein Kinase C-RasGRP1 Pathway Directs Ras Activation upon Antigen Receptor Stimulation of T Cells. Mol Cell Biol. 2005;25: 4426–4441. doi:10.1128/MCB.25.11.4426-4441.2005

139. Hayashi K, Altman A. Protein Kinase C Theta (PKCθ): A Key Player in T Cell Life and Death. Pharmacol Res. 2007;55: 537–44. doi:10.1016%2Fj.phrs.2007.04.009

140. Melowic HR, Stahelin R V, Blatner NR, Tian W, Hayashi K, Altman A, et al. Mechanism of Diacylglycerol-induced Membrane Targeting and Activation of Protein Kinase Cθ. J Biol Chem. 2007;282: 21467–76. doi:10.1074/jbc.M700119200

141. Schulze-Luehrmann J, Ghosh S. Antigen-Receptor Signaling to Nuclear Factor kappa B. Immunity. 2006;25: 701–15. doi:10.1016/j.immuni.2006.10.010

142. Coornaert B, Baens M, Heyninck K, Bekaert T, Haegman M, Staal J, et al. T cell antigen receptor stimulation induces MALT1 paracaspase-mediated cleavage of the NF-jB inhibitor A20. Nat Immunol. 2008;9: 263–71. doi:10.1038/ni1561

143. Rebeaud F, Hailfinger S, Posevitz-Fejfar A, Tapernoux M, Moser R, Rueda D, et al. The proteolytic activity of the paracaspase MALT1 is key in T cell activation. Nat Immunol. 2008;9: 272–81. doi:10.1038/ni1568

144. Oh-Hora M, Rao A. Calcium signaling in lymphocytes. Curr Opin Immunol. 2008;20: 250–8. doi:10.1016/j.coi.2008.04.004

145. Savignac M, Mellström B, Naranjo JR, Savignac M, Mellström B, Naranjo JR. Calcium-dependent transcription of cytokine genes in T lymphocytes. Pflugers Arch-Eur J Physiol. 2007;454: 523–533. doi:10.1007/s00424-007-0238-y

146. Dolmetsch R, Lewis R, Goodnow C, Healy JI. Differential activation of transcription factors induced by Ca2+ response amplitude and duration. Nature. 1997;386: 855–8. doi:10.1038/386855a0

147. Burkhardt JK, Carrizosa E, Shaffer MH. The Actin Cytoskeleton in T Cell Activation. Annu Rev Immunol. 2008;26: 233–59.

doi:10.1146/annurev.immunol.26.021607.090347

148. Gomez TS, Kumar K, Medeiros RB, Shimizu Y, Leibson PJ, Billadeau DD. Formins regulate the Arp2/3-independent polarization of the MTOC to the Immunological Synapse. Immunity. 2007;26: 177–90. doi:10.1016/j.immuni.2007.01.008

149. Zeng R, Cannon JL, Abraham RT, Way M, Billadeau DD, Bubeck-Wardenberg J, et al. SLP-76 Coordinates Nck-Dependent Wiskott-Aldrich Syndrome Protein Recruitment with Vav-1/Cdc42-Dependent Wiskott-Aldrich Syndrome Protein Activation at the T Cell-APC Contact Site. J Immunol. 2014;171: 1360–1368. doi:10.4049/jimmunol.171.3.1360

150. Gomez TS, Hamann MJ, Mccarney S, Savoy DN, Lubking CM, Heldebrant MP, et al. Dynamin 2 regulates T cell activation by controlling actin polymerization at the immunological synapse. Nat Immunol. 2005;6: 261–70. doi:10.1038/ni1168

151. Kupfer A, Swain SL, Singer SJ. The specific direct interaction of helper T cells and antigen-presenting B cells. II. Reorientation of the microtubule organizing center and reorganization of the membrane-associated cytoskeleton inside the bound helper T cells. J Exp Med. 1987;165: 1565–80. doi:10.1084/jem.165.6.1565

152. Dustin ML. The immunological synapse. Cancer Immunol Res. 2012;2: 1023–33. doi:10.1158/2326-6066.CIR-14-0161

153. Ritter AT, Angus KL, Griffiths GM. The role of the cytoskeleton at the immunological synapse. Immunol Rev. 2013;256: 107. doi:10.1111/IMR.12117

154. Monks C, Freiberg B, Kupfer H, Sciaky N, Kupfer A. Three-dimensional segregation of supramolecular activation clusters in T cells. Nature. 1998;395: 82–6. doi:10.1038/25764

155. Alarcón B, Mestre D, Martínez-Martín N. The immunological synapse: A cause or consequence of T-cell receptor triggering? Immunology. Wiley-Blackwell; 2011. pp. 420–425. doi:10.1111/j.1365-2567.2011.03458.x

156. F F, CT B. The immunological synapse as a pharmacological target. Pharmacol Res. 2018;134: 118–133. doi:10.1016/J.PHRS.2018.06.009

157. T Y, K B, A H, K H, A I, M S, et al. Dok-1 and Dok-2 are negative regulators of T cell receptor signaling. Int Immunol. 2007;19: 487–495. doi:10.1093/INTIMM/DXM015

158. Dong S, Corre B, Foulon E, Dufour E, Veillette A, Acuto O, et al. T cell receptor for antigen induces linker for activation of T cell–dependent activation of a negative signaling complex involving Dok-2, SHIP-1, and Grb-2. J Exp Med. 2006;203: 2509. doi:10.1084/JEM.20060650

159. Shui J-W, Boomer JS, Han J, Xu J, Dement GA, Zhou G, et al. Hematopoietic progenitor kinase 1 negatively regulates T cell receptor signaling and T cell–mediated immune responses. Nat Immunol 2006 81. 2006;8: 84–91. doi:10.1038/ni1416

160. Naramura M, Jang IK, Kole H, Huang F, Haines D, Gu H. C-Cbl and Cbl-b regulate T cell responsiveness by promoting ligand-induced TCR down-modulation. Nat Immunol. 2002;3: 1192–1199. doi:10.1038/ni855

161. Balagopalan L, Barr VA, Sommers CL, Barda-Saad M, Goyal A, Isakowitz MS, et al. c-Cbl-Mediated Regulation of LAT-Nucleated Signaling Complexes. Mol Cell Biol. 2007;27: 8622–8636. doi:10.1128/mcb.00467-07

162. Rao N, Miyake S, Reddi AL, Douillard P, Ghosh AK, Dodge IL, et al. Negative regulation of Lck by Cbl ubiquitin ligase. Proc Natl Acad Sci U S A. 2002;99: 3794–3799. doi:10.1073/PNAS.062055999

163. Parry R V., Chemnitz JM, Frauwirth KA, Lanfranco AR, Braunstein I, Kobayashi S V., et al. CTLA-4 and PD-1 Receptors Inhibit T-Cell Activation by Distinct Mechanisms. Mol Cell Biol. 2005;25: 9543–9553. doi:10.1128/MCB.25.21.9543-9553.2005

164. Buchbinder EI, Desai A. CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition. Am J Clin Oncol. 2016;39: 98. doi:10.1097/COC.0000000000000239

165. P S, JP A. The future of immune checkpoint therapy. Science (80- ). 2015;348: 56–61. doi:10.1126/SCIENCE.AAA8172

166. MK C, JD W. At the bedside: CTLA-4- and PD-1-blocking antibodies in cancer immunotherapy. J Leukoc Biol. 2013;94: 41–53. doi:10.1189/JLB.1212631

167. Rotte A. Combination of CTLA-4 and PD-1 blockers for treatment of cancer. J Exp Clin Cancer Res 2019 381. 2019;38: 1–12. doi:10.1186/S13046-019-1259-Z

168. E W, G von H. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. 1998;7: 1029–1038. doi:10.1002/PRO.5560070420

169. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J Mol Biol. 2001;305: 567–580. doi:10.1006/JMBI.2000.4315

170. TM B, AJ D. Properties and identification of human protein drug targets. Bioinformatics. 2009;25: 451–457. doi:10.1093/BIOINFORMATICS/BTP002

171. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28: 235–242. doi:10.1093/nar/28.1.235

172. Loll PJ. Membrane proteins, detergents and crystals: what is the state of the art? Acta Crystallogr Sect F, Struct Biol Commun. 2014;70: 1576. doi:10.1107/S2053230X14025035

173. S R, T W. Electron crystallography as a technique to study the structure on membrane proteins in a lipidic environment. Annu Rev Biophys. 2009;38: 89–105. doi:10.1146/ANNUREV.BIOPHYS.050708.133649

174. Loura LMS, Prieto M, Fernandes F. Quantification of protein–lipid selectivity using FRET. Eur Biophys J. 2010;39: 565. doi:10.1007/S00249-009-0532-Z

175. MA S, IG D, SG S. Nanodiscs as a new tool to examine lipid-protein

interactions. Methods Mol Biol. 2013;974: 415–433. doi:10.1007/978-1-62703-275-9_18

176. Ingólfsson HI, Melo MN, van Eerden FJ, Arnarez C, Lopez CA, Wassenaar TA, et al. Lipid Organization of the Plasma Membrane. J Am Chem Soc. 2014;136: 14554–14559. doi:10.1021/ja507832e

177. Chavent M, Duncan AL, Sansom MS. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. Curr Opin Struct Biol. 2016;40: 8–16. doi:10.1016/j.sbi.2016.06.007

178. Corradi V, Mendez-Villuendas E, Ingoífsson HI, Gu R-X, Siuda I, Melo MN, et al. Lipid−Protein Interactions Are Unique Fingerprints for Membrane Proteins. ACS Cent Sci. 2018;4: 709–17. doi:10.1021/acscentsci.8b00143

179. W I, S K. Molecular Simulations of Gram-Negative Bacterial Membranes Come of Age. Annu Rev Phys Chem. 2020;71: 171–188. doi:10.1146/ANNUREV-PHYSCHEM-103019-033434

180. Koehler Leman J, Ulmschneider MB, Gray JJ. Computational modeling of membrane proteins. Proteins. 2015;83: 1–24. doi:10.1002/prot.24703

181. Pluhackova K, Böckmann RA. Biomembranes in atomistic and coarse-grained simulations. J Phys Condens Matter. 2015;27: 323103. doi:10.1088/0953-8984/27/32/323103

182. Almeida JG, Preto AJ, Koukos PI, Bonvin AMJJ, Moreira IS. Membrane proteins structures: A review on computational modeling tools. Biochim Biophys Acta - Biomembr. 2017;1859: 2021–2039. doi:10.1016/J.BBAMEM.2017.07.008

183. Stansfeld PJ, Goose JE, Caffrey M, Carpenter EP, Parker JL, Newstead S, et al. MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. Structure. 2015;23: 1350. doi:10.1016/J.STR.2015.05.006

184. TD N, MSP S, PJ S. The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. Nucleic Acids Res. 2019;47: D390–D397. doi:10.1093/NAR/GKY1047

185. Stansfeld PJ, Sansom MSP. From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. J Chem Theory Comput. 2011;7: 1157–1166. doi:10.1021/ct100569y

186. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 1–11. doi:10.1038/s41586-021-03819-2

187. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science (80- ). 2021; eabj8754. doi:10.1126/SCIENCE.ABJ8754

188. Noé F, De Fabritiis G, Clementi C. Machine learning for protein folding and dynamics. Curr Opin Struct Biol. 2020;60: 77–84. doi:10.1016/J.SBI.2019.12.005

189. Smith Z, Ravindra P, Wang Y, Cooley R, Tiwary P. Discovering Protein

Conformational Flexibility through Artificial-Intelligence-Aided Molecular Dynamics. J Phys Chem B. 2020;124: 8221–8229. doi:10.1021/ACS.JPCB.0C03985

190. Wang J, Olsson S, Wehmeyer C, Pérez A, Charron NE, de Fabritiis G, et al. Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. ACS Cent Sci. 2019; acscentsci.8b00913. doi:10.1021/acscentsci.8b00913

191. Bejagam KK, Singh S, An Y, Deshmukh SA. Machine-Learned Coarse-Grained Models. J Phys Chem Lett. 2018;9: 4667–4672. doi:10.1021/acs.jpclett.8b01416

192. Bereau T, DiStasio RA, Tkatchenko A, von Lilienfeld OA. Non-covalent interactions across organic and biological subsets of chemical space: Physics-based potentials parametrized from machine learning. J Chem Phys. 2018;148: 241706. doi:10.1063/1.5009502

193. Casalino L, Dommer AC, Gaieb Z, Barros EP, Sztain T, Ahn S-H, et al. AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics: Int J High Perform Comput Appl. 2021 [cited 2 Sep 2021]. doi:10.1177/10943420211006452

194. Lee H, Ma H, Turilli M, Bhowmik D, Jha S, Ramanathan A. DeepDriveMD: Deep-Learning Driven Adaptive Molecular Simulations for Protein Folding. Proc DLS 2019 Deep Learn Supercomput - Held conjunction with SC 2019 Int Conf High Perform Comput Networking, Storage Anal. 2019; 12–19. Available: https://arxiv.org/abs/1909.07817v1

195. Zimmerman MI, Porter JR, Ward MD, Singh S, Vithani N, Meller A, et al. SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. Nat Chem. 2021;13: 651–659. doi:10.1038/s41557-021-00707-0

196. Koldsø H, Sansom MSP. Organization and Dynamics of Receptor Proteins in a Plasma Membrane. J Am Chem Soc. 2015;137: 14694–14704. doi:10.1021/jacs.5b08048

197. Koldsø H, Shorthouse D, Hélie J, Sansom MSP. Lipid Clustering Correlates with Membrane Curvature as Revealed by Molecular Simulations of Complex Lipid Bilayers. Fradin C, editor. PLoS Comput Biol. 2014;10: e1003911. doi:10.1371/journal.pcbi.1003911

198. Duncan AL, Reddy T, Koldsø H, Hélie J, Fowler PW, Chavent M, et al. Protein crowding and lipid complexity influence the nanoscale dynamic organization of ion channels in cell membranes. Sci Rep. 2017;7: 16647. doi:10.1038/s41598-017-16865-6

199. Marrink SJ, Corradi V, Ingólfsson HI, Souza PCT, Tieleman DP, Sansom MSP. Computational Modeling of Realistic Cell Membranes. Chem Rev. 2019;119: 6184–6226. doi:10.1021/acs.chemrev.8b00460

200. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. J Phys Chem B. 2007;111: 7812–7824. doi:10.1021/jp071097f

201. Saunders MG, Voth GA. Coarse-Graining Methods for Computational

Biology. Annu Rev Biophys. 2013;42: 73–93. doi:10.1146/ANNUREV-BIOPHYS-083012-130348

202. Dama JF, Sinitskiy A V., McCullagh M, Weare J, Roux B, Dinner AR, et al. The Theory of Ultra-Coarse-Graining. 1. General Principles. J Chem Theory Comput. 2013;9: 2466–2480. doi:10.1021/CT4000444

203. A Y, AJ P, P H, V M-G, L C, Z G, et al. A multiscale coarse-grained model of the SARS-CoV-2 virion. Biophys J. 2021;120: 1097–1104. doi:10.1016/J.BPJ.2020.10.048

204. Rzepiela AJ, Louhivuori M, Peter C, Marrink SJ. Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. Phys Chem Chem Phys. 2011;13: 10437. doi:10.1039/c0cp02981e

205. Wassenaar TA, Pluhackova K, Böckmann RA, Marrink SJ, Tieleman DP. Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. J Chem Theory Comput. 2014;10: 676–690. doi:10.1021/ct400617g

206. Zavadlav J, Bevc S, Praprotnik M. Adaptive resolution simulations of biomolecular systems. Eur Biophys J. 2017;46: 821–835. doi:10.1007/s00249-017-1248-0

207. Wassenaar TA, Ingólfsson HI, Prieß M, Marrink SJ, Schäfer L V. Mixing MARTINI: Electrostatic Coupling in Hybrid Atomistic–Coarse-Grained Biomolecular Simulations. J Phys Chem B. 2013;117: 3516–3530. doi:10.1021/jp311533p

208. Pezeshkian W, König M, Wassenaar TA, Marrink SJ. Backmapping triangulated surfaces to coarse-grained membrane models. Nat Commun. 2020;11: 1–9. doi:10.1038/s41467-020-16094-y

209. Wu P, Zhang T, Liu B, Fei P, Cui L, Qin R, et al. Mechano-regulation of Peptide-MHC Class I Conformations Determines TCR Antigen Recognition. Mol Cell. 2019;73: 1015-1027.e7. doi:10.1016/j.molcel.2018.12.018

210. Zech T, Ejsing CS, Gaus K, Wet B de, Shevchenko A, Simons K, et al. Accumulation of raft lipids in T-cell plasma membrane domains engaged in TCR signalling. EMBO J. 2009;28: 466. doi:10.1038/EMBOJ.2009.6

211. Senn HM, Thiel W. QM/MM studies of enzymes. Curr Opin Chem Biol. 2007;11: 182–187. doi:10.1016/J.CBPA.2007.01.684

212. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem. 2005;26: 1701–1718. doi:10.1002/jcc.20291

213. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1–2: 19–25. doi:10.1016/J.SOFTX.2015.06.001

214. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. J Comput Chem. 2013;34: 2135–2145. doi:10.1002/jcc.23354

215. Mo Y. A Critical Analysis on the Rotation Barriers in Butane. J Org Chem.

2010;75: 2733–2736. doi:10.1021/JO1001164

216. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys. 1995;103: 8577–8593. doi:10.1063/1.470117

217. Hockney RW, Goel SP, Eastwood JW. Quiet high-resolution computer models of a plasma. J Comput Phys. 1974;14: 148–158. doi:10.1016/0021-9991(74)90010-2

218. Verlet L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. Phys Rev. 1967;159: 98. doi:10.1103/PhysRev.159.98

219. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem. 1997;18: 1463–1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

220. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys. 1977;23: 327–341. doi:10.1016/0021-9991(77)90098-5

221. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys. 1984;81: 3684–3690. doi:10.1063/1.448118

222. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Mol Phys. 2006;52: 255–268. doi:10.1080/00268978400101201

223. Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Phys Rev A. 1985;31: 1695. doi:10.1103/PhysRevA.31.1695

224. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys. 2007;126: 014101. doi:10.1063/1.2408420

225. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys. 1981;52: 7182–7190. doi:10.1063/1.328693

226. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. J Comput Chem. 2008;29: 1859–1865. doi:10.1002/JCC.20945

227. Monticelli L, K. Kandasamy S, Periole X, G. Larson R, Peter Tieleman D, Marrink S-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. J Chem Theory Comput. 2008;4: 819–834. doi:10.1021/ct700324x

228. López CA, Rzepiela AJ, Vries AH de, Dijkhuizen L, Hünenberger PH, Marrink SJ. Martini Coarse-Grained Force Field: Extension to Carbohydrates. J Chem Theory Comput. 2009;5: 3195–3210. doi:10.1021/CT900313W

229. Uusitalo JJ, Ingólfsson HI, Akhshi P, Tieleman DP, Marrink SJ. Martini Coarse-Grained Force Field: Extension to DNA. J Chem Theory Comput. 2015;11: 3932–3945. doi:10.1021/ACS.JCTC.5B00286

230. Uusitalo JJ, Ingólfsson HI, Marrink SJ, Faustino I. Martini Coarse-Grained Force Field: Extension to RNA. Biophys J. 2017;113: 246–256. doi:10.1016/j.bpj.2017.05.043

231. Shivgan AT, Marzinek JK, Huber RG, Krah A, Henchman RH, Matsudaira P, et al. Extending the Martini Coarse-Grained Force Field to N-Glycans. J Chem Inf Model. 2020;60: 3864–3883. doi:10.1021/ACS.JCIM.0C00495

232. Atsmon-Raz Y, Tieleman DP. Parameterization of Palmitoylated Cysteine, Farnesylated Cysteine, Geranylgeranylated Cysteine, and Myristoylated Glycine for the Martini Force Field. J Phys Chem B. 2017;121: 11132–11143. doi:10.1021/acs.jpcb.7b10175

233. Periole X, Cavalli M, Marrink S-J, Ceruso MA. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. J Chem Theory Comput. 2009;5: 2531–2543. doi:10.1021/ct9002114

234. GS A, WG N, GA V. Multiscale modeling of biomolecular systems: in serial and in parallel. Curr Opin Struct Biol. 2007;17: 192–198. doi:10.1016/J.SBI.2007.03.004

235. Stansfeld PJ, Hopkinson R, Ashcroft FM, Sansom MSP. PIP2-Binding Site in Kir Channels: Definition by Multiscale Biomolecular Simulations. Biochemistry. 2009;48: 10926–10933. doi:10.1021/BI9013193

236. Kalli AC, Campbell ID, Sansom MSP. Multiscale simulations suggest a mechanism for integrin inside-out activation. Proc Natl Acad Sci U S A. 2011;108: 11890–5. doi:10.1073/pnas.1104505108

237. Wee CL, Balali-Mood K, Gavaghan D, Sansom MSP. The Interaction of Phospholipase A2 with a Phospholipid Bilayer: Coarse-Grained Molecular Dynamics Simulations. Biophys J. 2008;95: 1649. doi:10.1529/BIOPHYSJ.107.123190

238. Kalli AC, Campbell ID, Sansom MSP. Conformational Changes in Talin on Binding to Anionic Phospholipid Membranes Facilitate Signaling by Integrin Transmembrane Helices. Ben-Tal N, editor. PLoS Comput Biol. 2013;9: e1003316. doi:10.1371/journal.pcbi.1003316

239. Wee CL, Gavaghan D, Sansom MSP. Interactions Between a Voltage Sensor and a Toxin via Multiscale Simulations. Biophys J. 2010;98: 1558–1565. doi:10.1016/J.BPJ.2009.12.4321

240. Wassenaar TA, Ingólfsson HI, Böckmann RA, Tieleman DP, Marrink SJ. Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. J Chem Theory Comput. 2015;11: 2144–2155. doi:10.1021/acs.jctc.5b00209

241. Zhou FX, Merianos HJ, Brunger AT, Engelman DM. Polar residues drive association of polyleucine transmembrane helices. Proc Natl Acad Sci U S A. 2001;98: 2250–5. doi:10.1073/pnas.041593698

242. Blumberg RS, Alarcon B, Sancho J, McDermott F V, Lopez P, Breitmeyer J, et al. Assembly and function of the T cell antigen receptor. Requirement of either the lysine or arginine residues in the transmembrane region of the alpha chain. J Biol Chem. 1990;265: 14036–43. Available: https://pubmed.ncbi.nlm.nih.gov/2143190/

243. Alcover A, Mariuzza RA, Ermonval M, Acuto O. Lysine 271 in the transmembrane domain of the T-cell antigen receptor beta chain is necessary for its assembly with the CD3 complex but not for alpha/beta dimerization. J Biol Chem. 1990;265: 4131–5. Available: http://www.ncbi.nlm.nih.gov/pubmed/2137462

244. Kuhns MS, Davis MM. Disruption of Extracellular Interactions Impairs T Cell Receptor-CD3 Complex Stability and Signaling. Immunity. 2007;26: 357–369. doi:10.1016/j.immuni.2007.01.015

245. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25: 1605–1612. doi:10.1002/jcc.20084

246. DT J. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics. 2007;23: 538–544. doi:10.1093/BIOINFORMATICS/BTL677

247. Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 Years On. Nucleic Acids Res. 2019;47: W402–W407. Available: https://pubmed.ncbi.nlm.nih.gov/31251384/

248. D J, L Z, D Z. Basic charge clusters and predictions of membrane protein topology. J Chem Inf Comput Sci. 2002;42: 620–632. doi:10.1021/CI010263S

249. Petruk AA, Varriale S, Coscia MR, Mazzarella L, Merlino A, Oreste U. The structure of the CD3 ζζ transmembrane dimer in POPC and raft-like lipid bilayer: A molecular dynamics study. Biochim Biophys Acta - Biomembr. 2013;1828: 2637–2645. doi:10.1016/j.bbamem.2013.07.019

250. Love PE, Shores EW. ITAM multiplicity and thymocyte selection: How low can you go? Immunity. Cell Press; 2000. pp. 591–597. doi:10.1016/S1074-7613(00)80210-1

251. de Jong DH, Singh G, Bennett WFD, Arnarez C, Wassenaar TA, Schäfer L V., et al. Improved Parameters for the Martini Coarse-Grained Protein Force Field. J Chem Theory Comput. 2013;9: 687–697. doi:10.1021/ct300646g

252. Wang Y, Becker D, Vass T, White J, Marrack P, Kappler JW. A Conserved CXXC Motif in CD3ε Is Critical for T Cell Development and TCR Signaling. PLOS Biol. 2009;7: e1000253. doi:10.1371/JOURNAL.PBIO.1000253

253. Onufriev A. Implicit Solvent Models in Molecular Dynamics Simulations: A Brief Overview. Annual Reports in Computational Chemistry. Elsevier; 2008. pp. 125–137. doi:10.1016/S1574-1400(08)00007-8

254. Kleinjung J, Fraternali F. Design and application of implicit solvent models in biomolecular simulations. Curr Opin Struct Biol. 2014;25: 126. doi:10.1016/J.SBI.2014.04.003

255. Chien YH, Davis MM. How alpha beta T-cell receptors “see” peptide/MHC complexes. Immunol Today. 1993;14: 597–602. doi:10.1016/0167-5699(93)90199-u

256. Rock EP, Sibbald PR, Davis MM, Chien YH. CDR3 length in antigen-

specific immune receptors. J Exp Med. 1994;179: 323–328. doi:10.1084/jem.179.1.323

257. Willcox BE, Gao GF, Wyer JR, Ladbury JE, Bell JI, Jakobsen BK, et al. TCR binding to peptide-MHC stabilizes a flexible recognition interface. Immunity. 1999;10: 357–65. Available: http://www.ncbi.nlm.nih.gov/pubmed/10204491

258. Sasmal DK, Feng W, Roy S, Leung P, He Y, Cai C, et al. TCR–pMHC bond conformation controls TCR ligand discrimination. Cell Mol Immunol. 2020;17: 203–217. doi:10.1038/s41423-019-0273-6

259. P J, JH S, AM S, J H, RA F, J M, et al. Remarkably low affinity of CD4/peptide-major histocompatibility complex class II protein interactions. Proc Natl Acad Sci U S A. 2016;113: 5682–5687. doi:10.1073/PNAS.1513918113

260. Y X, P K, H C, E R. T Cell Receptor Binding to a pMHCII Ligand Is Kinetically Distinct from and Independent of CD4. J Biol Chem. 2001;276: 5659–5667. doi:10.1074/JBC.M009580200

261. Hoerter JAH, Brzostek J, Artyomov MN, Abel SM, Casas J, Rybakin V, et al. Coreceptor affinity for MHC defines peptide specificity requirements for TCR interaction with coagonist peptide–MHC. J Exp Med. 2013;210: 1807–1821. doi:10.1084/JEM.20122528

262. Aleksic M, Dushek O, Zhang H, Shenderov E, Chen J-L, Cerundolo V, et al. Dependence of T Cell Antigen Recognition on T Cell Receptor-Peptide MHC Confinement Time. Immunity. 2010;32: 163. doi:10.1016/J.IMMUNI.2009.11.013

263. Apostolopoulos V, Yu M, Corper AL, Teyton L, Pietersz GA, McKenzie IFC, et al. Crystal Structure of a Non-canonical Low-affinity Peptide Complexed with MHC Class I: A New Approach For Vaccine Design. J Mol Biol. 2002;318: 1293–1305. doi:10.1016/S0022-2836(02)00196-1

264. Rosenberg SA, Yang JC, Schwartzentruber DJ, Hwu P, Marincola FM, Topalian SL, et al. Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. Nat Med. 1998;4: 321–327. doi:10.1038/nm0398-321

265. W C, JW Y, RL L, JR B. Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants. J Exp Med. 1999;189: 1757–1764. doi:10.1084/JEM.189.11.1757

266. Pettmann J, Huhn A, Shah EA, Kutuzov MA, Wilson DB, Dustin ML, et al. The discriminatory power of the T cell receptor. Elife. 2021;10: e67092. doi:10.7554/ELIFE.67092

267. Rosenits K, Keppler SJ, Vucikuja S, Aichele P. T cells acquire cell surface determinants of APC via in vivo trogocytosis during viral infections. Eur J Immunol. 2010;40: 3450–3457. doi:10.1002/eji.201040743

268. DG O, SA W. Trogocytosis results in sustained intracellular signaling in CD4(+) T cells. J Immunol. 2012;189: 4728–4739. doi:10.4049/JIMMUNOL.1201507

269. Kim HR, Mun YV, Lee KS, Park YJ, Park JS, Park JH, et al. T cell

microvilli constitute immunological synaptosomes that carry messages to antigen-presenting cells. Nat Commun. 2018;9. doi:10.1038/s41467-018-06090-8

270. Nakayama M, Hori A, Toyoura S, Yamaguchi S-I. Shaping of T Cell Functions by Trogocytosis. Cells. 2021;10: 1155. doi:10.3390/CELLS10051155

271. S K, A P, G B, J A, I D. Mass Spectrometry-Based Identification of MHC-Associated Peptides. Cancers (Basel). 2020;12. doi:10.3390/CANCERS12030535

272. Dyachenko A, Gruber R, Shimon L, Horovitz A, Sharon M. Allosteric mechanisms can be distinguished using structural mass spectrometry. Proc Natl Acad Sci. 2013;110: 7235–7239. doi:10.1073/PNAS.1302395110

273. Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology. 2018;154: 394–406. doi:10.1111/IMM.12889

274. Ochoa R, Laio A, Cossio P. Predicting the Affinity of Peptides to Major Histocompatibility Complex Class II by Scoring Molecular Dynamics Simulations. J Chem Inf Model. 2019;59: 3464–3473. doi:10.1021/ACS.JCIM.9B00403

275. Xiao Z, Zhang Y, Yu R, Chen Y, Jiang X, Wang Z, et al. In silico design of MHC class I high binding affinity peptides through motifs activation map. BMC Bioinforma 2018 1919. 2018;19: 13–24. doi:10.1186/S12859-018-2517-3

276. Malaker SA, Ferracane MJ, Depontieu FR, Zarling AL, Shabanowitz J, Bai DL, et al. Identification and Characterization of Complex Glycosylated Peptides Presented by the MHC Class II Processing Pathway in Melanoma. J Proteome Res. 2017;16: 228. doi:10.1021/ACS.JPROTEOME.6B00496

277. Prakaash D, Cook GP, Acuto O, Kalli AC. Multi-scale simulations of the T cell receptor reveal its lipid interactions, dynamics and the arrangement of its cytoplasmic region. Wallqvist A, editor. PLOS Comput Biol. 2021;17: e1009232. doi:10.1371/JOURNAL.PCBI.1009232

278. Courtney AH, Lo WL, Weiss A. TCR Signaling: Mechanisms of Initiation and Propagation. Trends in Biochemical Sciences. Elsevier Ltd; 2018. pp. 108–123. doi:10.1016/j.tibs.2017.11.008

279. Garcia KC, Degano M, Speir JA, Wilson IA. Emerging principles for T cell receptor recognition of antigen in cellular immunity. Rev Immunogenet. 1999;1: 75–90. Available: https://pubmed.ncbi.nlm.nih.gov/11256574/

280. Acuto O, Bartolo V Di, Michel F. Tailoring T-cell receptor signals by proximal negative feedback mechanisms. Nature Reviews Immunology. Nature Publishing Group; 2008. pp. 699–712. doi:10.1038/nri2397

281. Fiser A, Šali A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. Methods Enzymol. 2003;374: 461–491. doi:10.1016/S0076-6879(03)74020-8

282. Webb B, Sali A. Protein structure modeling with MODELLER. Methods

Mol Biol. 2014;1137. doi:10.1007/978-1-4939-0366-5_1

283. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J Phys Chem B. 2010;114: 7830–7843. doi:10.1021/JP101759Q

284. Song W, Corey RA, Ansell TB, Cassidy CK, Horrell MR, Duncan AL, et al. PyLipID: A Python Package for Analysis of Protein–Lipid Interactions from Molecular Dynamics Simulations. J Chem Theory Comput. 2022;18: 1188–1201. doi:10.1021/ACS.JCTC.1C00708

285. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996;14: 33–38. doi:10.1016/0263-7855(96)00018-5

286. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001;98: 10037–10041. doi:10.1073/pnas.181342398

287. TJ D, JE N, JA M, NA B. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 2004;32. doi:10.1093/NAR/GKH381

288. Edgar RC. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5. doi:10.1186/1471-2105-5-113

289. Duchardt E, Sigalov AB, Aivazian D, Stern LJ, Schwalbe H. Structure Induction of the T-Cell Receptor ζ-Chain upon Lipid Binding Investigated by NMR Spectroscopy. ChemBioChem. 2007;8: 820–827. doi:10.1002/cbic.200600413

290. Shen M, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci. 2006;15: 2507–2524. doi:10.1110/ps.062416606

291. Schoenborn JR, Tan YX, Zhang C, Shokat KM, Weiss A. Feedback circuits monitor and adjust basal Lck-dependent events in T cell receptor signaling. Sci Signal. 2011;4: ra59–ra59. doi:10.1126/scisignal.2001893

292. Sjölin-Goodfellow H, Frushicheva MP, Ji Q, Cheng DA, Kadlecek TA, Cantor AJ, et al. The catalytic activity of the kinase ZAP-70 mediates basal signaling and negative feedback of the T cell receptor pathway. Sci Signal. 2015;8: ra49. doi:10.1126/scisignal.2005596

293. Love PE, Hayes SM. ITAM-mediated signaling by the T-cell antigen receptor. Cold Spring Harb Perspect Biol. 2010;2. doi:10.1101/cshperspect.a002485

294. Corradi V, Sejdiu BI, Mesa-Galloso H, Abdizadeh H, Noskov SY, Marrink SJ, et al. Emerging Diversity in Lipid-Protein Interactions. Chemical Reviews. American Chemical Society; 2019. pp. 5775–5848. doi:10.1021/acs.chemrev.8b00451

295. Hedger G, Sansom MSP, Koldsø H. The juxtamembrane regions of human receptor tyrosine kinases exhibit conserved interaction sites with anionic lipids. Sci Rep. 2015;5: 1–10. doi:10.1038/srep09198

296. Chavent M, Karia D, Kalli AC, Seiradake E, Jones EY, Sansom MSP.

Interactions of the EphA2 Kinase Domain with PIPs in Membranes: Implications for Receptor Function. Structure. 2018;26: 1025–34. doi:10.1016/j.str.2018.05.003

297. Michailidis IE, Rusinova R, Georgakopoulos A, Chen Y, Iyengar R, Robakis NK, et al. Phosphatidylinositol-4,5-bisphosphate regulates epidermal growth factor receptor activation. Pflügers Arch - Eur J Physiol. 2011;461: 387–397. doi:10.1007/s00424-010-0904-3

298. Saltel F, Mortier E, Hytönen VP, Jacquier MC, Zimmermann P, Vogel V, et al. New PI(4,5)P2- and membrane proximal integrin-binding motifs in the talin head control β3-integrin clustering. J Cell Biol. 2009;187: 715–731. doi:10.1083/jcb.200908134

299. Hatada MH, Lu X, Laird ER, Green J, Morgenstern JP, Lou M, et al. Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. Nature. 1995;377: 32–38. doi:10.1038/377032a0

300. Katz ZB, Novotná L, Blount A, Lillemeier BF. A cycle of Zap70 kinase activation and release from the TCR amplifies and disperses antigenic stimuli. Nat Immunol. 2017;18: 86–95. doi:10.1038/ni.3631

301. Palacios EH, Weiss A. Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. Oncogene. 2004;23: 7990–8000. doi:10.1038/sj.onc.1208074

302. J Y, R Y, A R, D X, J P, Y Z. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12: 7–8. doi:10.1038/NMETH.3213

303. DE K, D C, D B. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 2004;32. doi:10.1093/NAR/GKH468

304. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, et al. Comparative protein structure modeling using Modeller. Curr Protoc Bioinforma. 2006;Chapter 5: Unit-5.6. doi:10.1002/0471250953.bi0506s15

305. F S, DG H. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 2018;27: 135–145. doi:10.1002/PRO.3290

306. Le Roux AL, Mohammad IL, Mateos B, Arbesú M, Gairí M, Khan FA, et al. A Myristoyl-Binding Site in the SH3 Domain Modulates c-Src Membrane Anchoring. iScience. 2019;12: 194–203. doi:10.1016/J.ISCI.2019.01.010

307. Hilzenrat G, Pandžić E, Yang Z, Nieves DJ, Goyette J, Rossy J, et al. Conformational States Control Lck Switching between Free and Confined Diffusion Modes in T Cells. Biophys J. 2020;118: 1489–1501. doi:10.1016/j.bpj.2020.01.041

308. Xu W, Doshi A, Lei M, Eck MJ, Harrison SC. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol Cell. 1999;3: 629–638. doi:10.1016/S1097-2765(00)80356-1

309. H Y, K K, Y K, J M, Y T, M K-N, et al. Activity cliff for 7-substituted pyrrolo-pyrimidine inhibitors of HCK explained in terms of predicted basicity of the amine nitrogen. Bioorg Med Chem. 2017;25: 4259–4264.

doi:10.1016/J.BMC.2017.05.053

310. Xu X, Li H, Xu C. Structural understanding of T cell receptor triggering. Cellular and Molecular Immunology. Springer Nature; 2020. pp. 193–202. doi:10.1038/s41423-020-0367-1

311. Sibener L V., Fernandes RA, Kolawole EM, Carbone CB, Liu F, McAffee D, et al. Isolation of a Structural Mechanism for Uncoupling T Cell Receptor Signaling from Peptide-MHC Binding. Cell. 2018;174: 672-687.e27. doi:10.1016/J.CELL.2018.06.017

312. Pageon S V, Tabarin T, Yamamoto Y, Ma Y, Nicovich PR, Bridgeman JS, et al. Functional role of T-cell receptor nanoclusters in signal initiation and antigen discrimination. Proc Natl Acad Sci U S A. 2016;113: E5454-63. doi:10.1073/pnas.1607436113

313. Song W, Duncan AL, Sansom MSP. Modulation of adenosine A2a receptor oligomerization by receptor activation and PIP2 interactions. Structure. 2021;29: 1312-1325.e3. doi:10.1016/J.STR.2021.06.015

314. Costa AF, Campos D, Reis CA, Gomes C. Targeting Glycosylation: A New Road for Cancer Drug Discovery. Trends in Cancer. 2020;6: 757–766. doi:10.1016/J.TRECAN.2020.04.002

315. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 2015;15: 540–555. doi:10.1038/nrc3982

316. Sun R, Kim AMJ, Lim S-O. Glycosylation of Immune Receptors in Cancer. Cells. 2021;10: 1100. doi:10.3390/CELLS10051100

317. Streng-Ouwehand I, Ho NI, Litjens M, Kalay H, Boks MA, Cornelissen LAM, et al. Glycan modification of antigen alters its intracellular routing in dendritic cells, promoting priming of T cells. Elife. 2016;5: e11765. doi:10.7554/ELIFE.11765

318. Wolfert MA, Boons G-J. Adaptive immune activation: glycosylation does matter. Nat Chem Biol. 2013;9: 776. doi:10.1038/NCHEMBIO.1403

319. Olvera A, Cedeño S, Llano A, Mothe B, Sanchez J, Arsequell G, et al. Does Antigen Glycosylation Impact the HIV-Specific T Cell Immunity? Front Immunol. 2020;11: 573928. doi:10.3389/FIMMU.2020.573928

320. Bousser E De, Meuris L, Callewaert N, Festjens N. Human T cell glycosylation and implications on immune therapy for cancer. Hum Vaccin Immunother. 2020;16: 2374–2388. doi:10.1080/21645515.2020.1730658

321. Vickery ON, Stansfeld PJ. CG2AT2: an Enhanced Fragment-Based Approach for Serial Multi-scale Molecular Dynamics Simulations. J Chem Theory Comput. 2021;17: 6472–6482. doi:10.1021/ACS.JCTC.1C00295

322. Beemiller P, Krummel MF. Mediation of T-cell activation by actin meshworks. Cold Spring Harbor perspectives in biology. Cold Spring Harbor Laboratory Press; 2010. p. a002444. doi:10.1101/cshperspect.a002444

323. Barda-Saad M, Braiman A, Titerence R, Bunnell SC, Barr VA, Samelson LE. Dynamic molecular interactions linking the T cell antigen receptor to the actin cytoskeleton. Nat Immunol. 2005;6: 80–89. doi:10.1038/ni1143

324. LR O, P G, R D. The crystal structure of uncomplexed actin in the ADP

state. Science (80- ). 2001;293: 708–711. doi:10.1126/SCIENCE.1059700

325. Holmes KC, Popp D, Gebhard W, Kabsch W. Atomic model of the actin filament. Nature. 1990;347: 44–49. doi:10.1038/347044a0

326. T O, M I, T A, Y M, A N. The nature of the globular- to fibrous-actin transition. Nature. 2009;457: 441–445. doi:10.1038/NATURE07685

327. Chu J-W, Voth GA. Allostery of actin filaments: Molecular dynamics simulations and coarse-grained analysis. Proc Natl Acad Sci U S A. 2005;102: 13111. doi:10.1073/PNAS.0503732102

328. Pfaendtner J, Lyman E, Pollard TD, Voth GA. Structure and Dynamics of the Actin Filament. J Mol Biol. 2010;396: 252–263. doi:10.1016/J.JMB.2009.11.034

329. Splettstoesser T, Holmes KC, Noé F, Smith JC. Structural modeling and molecular dynamics simulation of the actin filament. Proteins Struct Funct Bioinforma. 2011;79: 2033–2043. doi:10.1002/PROT.23017

330. Connolly A, Panes R, Tual M, Lafortune R, Bellemare-Pelletier A, Gagnon E. TMEM16F mediates bystander TCR-CD3 membrane dissociation at the immunological synapse and potentiates T cell activation. Sci Signal. 2021;14. doi:10.1126/SCISIGNAL.ABB5146

331. Ma Y, Poole K, Goyette J, Gaus K. Introducing Membrane Charge and Membrane Potential to T Cell Signaling. Front Immunol. 2017;8: 1513. doi:10.3389/fimmu.2017.01513

332. Shi X, Bi Y, Yang W, Guo X, Jiang Y, Wan C, et al. Ca2+ regulates T-cell receptor activation by modulating the charge property of lipids. Nature. 2013;493: 111–115. doi:10.1038/NATURE11699

333. Javanainen M, Martinez-Seara H, Vattulainen I. Excessive aggregation of membrane proteins in the Martini model. PLoS One. 2017;12: 1–20. doi:10.1371/journal.pone.0187936

334. Souza PCT, Alessandri R, Barnoud J, Thallmair S, Faustino I, Grünewald F, et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. Nat Methods. 2021;18: 382–388. doi:10.1038/s41592-021-01098-3

335. Z Z, L L, WG N, V K, J P, GA V. A systematic methodology for defining coarse-grained sites in large biomolecules. Biophys J. 2008;95: 5073–5083. doi:10.1529/BIOPHYSJ.108.139626

336. Z Z, J P, A G, GA V. Defining coarse-grained representations of large biomolecules and biomolecular complexes from elastic network models. Biophys J. 2009;97: 2327–2337. doi:10.1016/J.BPJ.2009.08.007

337. Li M, Zhang JZH, Xia F. A new algorithm for construction of coarse-grained sites of large biomolecules. J Comput Chem. 2016;37: 795–804. doi:10.1002/JCC.24265

338. Pezeshkian W, Marrink SJ. Simulating realistic membrane shapes. Curr Opin Cell Biol. 2021;71: 103–111. doi:10.1016/J.CEB.2021.02.009