Computational studies of talin-mediated integrin activation

Antreas C. Kalli

Structural Bioinformatics and Computational Biochemistry Unit
Department of Biochemistry
St Cross College, Oxford
Trinity Term 2012

Thesis submitted in partial fulfilment of the
requirements for the degree of Doctor of Philosophy
at the University of Oxford

ABSTRACT
Computational studies of talin-mediated integrin activation
Antreas C. Kalli
Structural Bioinformatics and Computational Biochemistry unit and St. Cross College
Submitted for the degree of Doctor of Philosophy, Trinity 2012

Integrins are large heterodimeric (αβ) cell surface receptors that play a key role in the formation of focal adhesion complexes and are involved in various signal transduction pathways. They are 'activated' to a high affinity state by the formation of an intracellular complex between the membrane, the integrin β-subunit tail and talin, a process known as 'inside-out activation'. The head domain of talin, a FERM domain homologue, plays a vital role in the formation of this complex. Recent studies also suggest that kindlins act in synergy with talin to induce integrin activation. Despite much available structural and functional data, details of how talin activates integrins remain elusive.

In this thesis a multiscale simulation approach (using a combination of coarse-grained and atomistic molecular dynamics simulations) together with NMR experiments were employed to study talin-mediated integrin inside-out activation. A number of novel insights emerged from these studies including: (i) the crucial role of negatively charged lipids in talin/membrane association; (ii) a novel V-shape conformation of the talin head domain which optimizes its interactions with negatively charged lipids; (iii) that interactions of talin with negatively charged moieties in the membrane orient talin to optimize interactions with the β cytoplasmic tail; (iv) that binding of talin to the β cytoplasmic tail promotes rearrangement of the integrin TM helices and weakens the integrin α/β association; and (v) that an increase in the tilt angle of the β integrin TM helix initiates a scissoring movement of the two integrin TM helices. These results, combined with experimental data, provide new insights into the mechanism of integrin inside-out activation.

The same multiscale approach was used to demonstrate the crucial role of the Gx_3G motif in the packing of the integrin transmembrane helices. Using recent structural data the integrin/talin complex was modelled and inserted in bilayers which resemble the biological plasma membrane. The results demonstrate the dynamic nature of the integrin receptor and suggest that the integrin/talin complex alters the lipid organization and motion in the outer and inner bilayer leaflets in an asymmetric way and that diffusion of lipids in the vicinity of the protein is slowed down.

The work in this thesis demonstrates that multiscale simulations have considerable strengths when applied to investigations of structure/function relationships in membrane proteins.

Acknowledgments

I would first wish to express my warmest thanks to Professor Mark Sansom as his advice and supervision have been truly invaluable. His support, attention to detail, hard work, and scholarship has been a source of motivation and inspiration for me throughout these four years. I would also need to extend my kindest thanks to Professor Iain Campbell who allowed me to work in his lab during my second rotation. Iain's expertise and understanding added considerably to my DPhil experience. I am grateful to both Mark and Iain who gave me the opportunity to work on a cutting-edge project and supported me with insightful ideas, critical suggestions and thorough reviews of my work and of my thesis along the way. I also wish to express my gratitude to the Wellcome Trust for its generous funding.

I also thank Professor Ilpo Vattulainen (TUT Finland) for giving me the opportunity to work in his lab during my HPC-Europa2 fellowship and offered me additional computational resources to complete my research projects. Many other people also offered me great help during this DPhil. Special thanks to Dr Kate Wegener for supervising me during my second rotation project and to Dr Nick Anthis for all the help he has provided me while working in the Campbell's lab. I also wish to thank all the members of the SBCB group for their invaluable help and support.

Last but not least, I wish to thank Katerina Tertytchnaya for her continuous encouragement and unconditional support during these four years.

Contents

Abstract ii

Acknowledgments iii

List of Publications vii

List of Abbreviations viii

1 Introduction 1
1.1 Membrane proteins and transmembrane signalling: a brief introduction 1
1.2 Integrins 3
1.2.1 Structure and biology 3
1.2.2 The integrin ectodomain and activation 6
1.2.3 Integrin TM and CT regions 9
1.2.3.1 TM and CT region topology and structure 9
1.2.3.2 Interactions in the TM and CT regions and activation 15
1.2.4 Talin and integrin inside-out activation 17
1.2.5 Other proteins involved in integrin inside-out activation 22
1.3 Molecular dynamics (MD) simulations of membrane proteins 24
1.4 Thesis aims and summary 26

2 Theory and methods 28
2.1 Introduction 28
2.2 Classical MD simulations 28
2.2.1 Bonded terms 29
2.2.2 Non-bonded terms 30
2.2.3 Integration algorithms 31
2.2.4 Force fields 32
2.2.5 Timestep-bond constrains 33
2.2.6 Periodic boundary conditions 33
2.2.7 Ensembles 33
2.2.8 Temperature and pressure coupling 34
2.2.9 Position restrains 36
2.3 CG approach 37
2.3.1 CG models 37
2.4 Multiscale approaches 39

3 Talin head domain interactions with the membrane 41
3.1 Introduction 41
3.2 Methods 43
3.2.1 NMR spectroscopy experiments 43

3.2.1.1 Protein expression and purification for NMR 43
3.2.1.2 Liposomes 44
3.2.1.3 NMR spectroscopy 44
3.2.2 Structures used in the simulations 45
3.2.3 Bilayer set-up 46
3.2.4 Simulation set-up 47
3.3 Results 49
3.3.1 NMR studies of talin/bilayer interactions 49
3.3.1.1 Interactions of the F2 and the F2-F3 subdomains with liposomes 49
3.3.2 MD simulations 52
3.3.2.1 F2 and F2-F3 interactions with the membrane 53
3.3.2.2 F0-F1 interactions with the membrane 56
3.3.2.3 Complete talin head domain (F0 to F3) interactions with the membrane 57
3.3.2.4 Refinement of the model by AT-MD simulations 65
3.3.3 High-throughput method for studying protein/lipid interactions 66
3.4 Discussion and conclusions 69

4 A helix heterodimer in a lipid bilayer: prediction of the structure of an integrin transmembrane domain via multiscale simulations 73
4.1 Introduction 73
4.2 Methods 76
4.2.1 Simulation set-up 76
4.3 Results 78
4.3.1 Wild type αIIb/β3 78
4.3.1.1 Dimerization by CG-MD simulations 78
4.3.1.2 Refinement and model assessment by AT-MD simulations 81
4.3.2 αIIb/β3 mutant forms 85
4.4 Discussion and conclusions 87

5 Conformation and dynamics of the integrin αIIbβ3 transmembrane helix dimer and the role of talin 90
5.1 Introduction 90
5.2 Structures used in the simulations 91
5.3 Results 93
5.3.1 Conformation and dynamics of the integrin αIIb/β3 helix dimer 93
5.3.2 Interactions with the talin F2-F3 domain pair modulate the integrin β subunit orientation 98
5.3.3 Changes in integrin TM helix packing induced by the talin F2-F3 domain pair 102
5.3.4 The talin head domain perturbs the packing of the integrin TM helices 109
5.4 Discussion and conclusions 111

6 The integrin receptor in more biologically relevant bilayers 116
6.1 Introduction..... 116
6.2 Methods 119
6.2.1 Integrin/talin F2-F3 complex modelling..... 119
6.2.2 Bilayer and system preparation..... 122
6.3 Results..... 123
6.3.1 Integrin/talin complex dynamics 123
6.3.2 Lipid organization around the integrin/talin complex 124
6.3.3 The integrin/talin complex alters lipid mobility 129
6.4 Discussion and conclusions 134

7 Conclusions and future directions..... 138
7.1 Conclusions..... 138
7.2 Future directions 141

APPENDIX A..... 146
A.1 Talin F0-F1 domain pair interactions with the membrane..... 146
A.2 Complete talin head domain (F0 to F3) interactions with the membrane..... 148
A.2.1 The tal-lF0F1-CG, tal-h2F1-CG and tal-h2F0-CG simulations 148
A.2.2 The role of the negatively charged lipids in the talin/membrane interactions. 150

APPENDIX B 151
B.1 High-throughput method for studying protein/lipid interactions..... 151

APPENDIX C..... 153
C.1 F2-F3/αβ, F3/β and F2-F3/β complexes modelling 153
C.2 Talin head domain (Δloop-res: 134-172) in complex with the integrin α/β integrin TM region 155

APPENDIX D..... 157
D.1 Preparation of bilayers 157
D.2 Convergence analysis..... 158
D.3 Other properties of the bilayer 159

Bibliography **160**

Publications **191**

List of Publications

1. Kalli AC, et al. (2010) The structure of the talin/integrin complex at a lipid bilayer: an NMR and MD simulation study. Structure 18(10):1280-1288.

2. Kalli AC, Hall BA, Campbell ID, & Sansom MSP (2011) A helix heterodimer in a lipid bilayer: prediction of the structure of an integrin transmembrane domain via multiscale simulations. Structure 19(10):1477-1484.

3. Kalli AC, Campbell ID, & Sansom MSP (2011) Multiscale simulations suggest a mechanism for integrin inside-out activation. Proc. Natl. Acad. Sci. USA 108(29):11890-11895.

List of Abbreviations

amu atomic mass unit
AT atomistic
AT-MD atomistic molecular dynamics simulations
CD circular dichroism
CG coarse-grained
CG2AT coarse-grained to atomistic representation
CG-MD coarse-grained molecular dynamics simulations
CHOL cholesterol
CT cytoplasmic
DMPS L-α-phosphatidylcholine/1,2-dimyristoyl-sn-glycero-3-phospho-L-serine
Dok1 docking protein 1, downstream of tyrosine kinase 1
DPPC 1,2-di-palmitoyl-sn-glycero-3-phosphatidyl-choline
DTT dithiothreitol
ECM extracellular matrix
EGF epidermal growth factor
EGFR epidermal growth factor receptor
EM electron microscopy
ENM elastic network model
ER endoplasmic reticulum
FAK focal adhesion kinase
FERM band four-point-one, ezrin, radixin, moesin (domain)
FRET fluorescence resonance energy transfer
GPCR G protein-coupled receptor
GROMACS GROningen MAchine for Chemical Simulations
GST glutathione S-transferase (protein)
H-bonds hydrogen bonds
HPLC high performance liquid chromatography
HSQC heteronuclear single quantum coherence (spectrum)
I-CAM intercellular adhesion molecule
IMC inner membrane clasp
LH left-handed
MC monte carlo
MD molecular dynamics
MDist membrane-distal
MM molecular mechanics
MProx membrane-proximal
NMR nuclear magnetic resonance
NOE nuclear Overhauser effect
NOESY nuclear Overhauser enhancement spectroscopy
PBS phosphate buffer saline
PC L-α-phosphatidylcholine
PDB protein data bank
PH pleckstrin homology (domain)

PIP_2 phosphatidylinositol-4,5-bisphosphate
PIPK1γ phosphatidylinositol phosphate kinase type 1γ
PME particle mesh Ewald
POPC 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidyl-choline
POPE 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidyl-ethanolamine
POPG 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidyl-glycerol
POPS 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidyl-serine
PPCS palmitoyl sphingomylien phosphatidyl-choline
PSI plexin, semaphoring integrin (domain)
PTB phosphotyrosine-binding (domain)
OMC outer membrane clasp
QM quantum mechanics
RH right-handed
RIAM Rap1-interacting adaptor molecule
RMSD root mean squared deviation
RMSF root mean squared fluctuation
SDS-PAGE sodium dodecyl sulphate polyacrylamide gel electrophoresis
SM sphingomyelin
SMD steered molecular dynamics simulations
SPR surface plasmon resonance
TM transmembrane
Tris tris (hydroxymethyl) aminomethane
VDW van der Waals
WT wild type
Δ(HN,N) weighted combined change in ^{1}H and ^{15}N amide chemical shifts

Chapter 1

Introduction

1.1 Membrane proteins and transmembrane signalling: a brief introduction

Membrane proteins account for a considerable fraction of the human genome (~30%) (1) and they are a major drug target (2), highlighting their importance in health and disease. In the same context, membrane proteins play an important role in many cellular processes such as signal transduction, immune recognition and cell adhesion (3, 4). Membrane proteins can be broadly divided into two categories; integral membrane proteins which span the cell membrane (e.g. transporters, channels, receptors) and peripheral proteins which interact with the cell membrane surface (e.g. FERM (F=Four point one proteins, E=Ezrin, R=Radixin, M=Moesin) domains or pleckstrin homology (PH) domains; see Fig. 1.1).

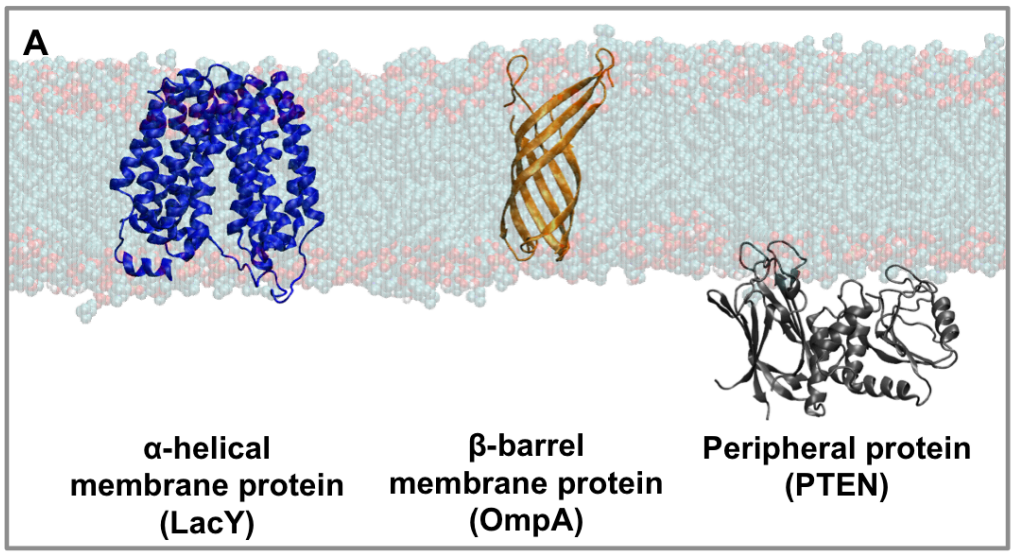


Figure 1.1: Examples of the categories of membrane proteins. The first category is the integral membrane proteins which have two main conformations – α-helical (PDB:1PV7; left) and β-barrel (PDB:1BXW; centre). The second category is the peripheral proteins (PDB:1D5R; right).

The plasma membrane is the barrier which separates the cell from its extracellular environment and regulates its communication with its surroundings. The plasma membrane is formed from a mixture of different lipid species and integral and peripheral proteins that allow information or material to be transmitted to and from the cell interior. Recent

advances in our understanding of membrane structure/dynamics changed the "fluid mosaic model" (5) of membranes, with random distributions of components moving in a Brownian fashion (5), to a more organized and structured view with lateral organization of lipids and proteins in ordered and disordered domains (6). The size and the physical properties of the two states depend on the lipid–lipid, protein–lipid or protein–protein interactions. This provides additional evidence that the functions of lipids and proteins are coupled in the events occurring at the cell membrane (e.g. transmembrane (TM) signalling (4)). Indeed, in many cases lipid organization or structure is modified in response to intracellular or extracellular stimuli. It should be noted, however, that large-scale phase separation is observed in model membranes but not in living cells (7, 8) and therefore there is ongoing debate concerning the details of the lipid phase separation process (9, 10). The high density of membrane proteins, their specificity for different lipid types and proteins and the concerted movement of protein and lipids (11-14) adds another level of complexity (15). Therefore, it is possible that, smaller microdomains or nanodomains of lipids and protein exist to regulate membrane functions (7, 8).

Some lipids in the cell membrane even modify their chemical behaviour while participating in signal transduction pathways; e.g. the phosphoinositides phosphates (PIP) which serve as markers for recruiting specific proteins despite their low concentration (~5%). During the signalling process PIP lipids can modify their headgroups via phosphorylation/dephosphorylation by different enzymes (e.g. PTEN (16)) to regulate signalling activity (17). Other classes of lipids, which can be found in the cell membrane, are phospholipids (e.g. POPC, POPE, POPS or POPG), sphingolipids (e.g. sphingomyelin), glycolipids (e.g. GM3) and sterols (e.g. cholesterol). A more detailed discussion of membrane composition and dynamics is given in the introduction to Chapter 6.

The transmission of signals across the cell membrane usually involves the binding of extracellular ligands to the ectodomains of cell surface receptors. Binding induces a stimulus which causes conformational changes in the receptor and transmission of a signal to the cell. A wide variety of proteins act in this way, including G protein-coupled receptors (GPCR), receptor kinases and ion channels (18). In the case of GPCRs, extracellular stimuli causes conformational changes which activate a trimeric G-protein (19). Extracellular ligands that bind to GPCR receptors include hormones, ions and neurotransmitters (19). Members of the large GPCR family have a seven TM helix topology and are involved in many processes, such as neurotransmission and hormone

effects (20, 21). GPCR activation also regulates secondary messengers pathways (Ca^{2+}, cAMP), which amplify the signal.

In other cases, additional events such as oligomerization may occur during signalling. In this case the lipid membrane acts as a two-dimensional platform for membrane protein diffusion. One example is the epidermal growth factor receptor (EGFR), which is a major drug target for cancer (22). The binding of EGF to the receptor promotes dimerization, which in turn results in phosphorylation of tyrosine residues in the C-terminal via a tyrosine kinase. The activated tyrosine kinase then promotes further downstream signalling when its phosphorylated regions bind to protein domains (e.g. SH2, PTB etc.) in adaptor proteins (23). Oncogenic mutations (e.g. L384R) or phosphorylation (e.g. of Tyr845) can promote EGFR dimerization and TM signalling (24).

The formation of complex protein–protein interactions during signal transmission may be also driven by lipid organization/dynamics (25). It has been shown that the formation of more ordered domains in some membranes promotes the formation of protein complexes to provide a platform where receptors can be activated by recruitment of proteins or ligands (25). One example is the Ras GTPase, which is involved in signal transmission; it was shown recently to form nanoclusters at lipid ordered and disordered interfaces (26).

Integrins are another family of cell surface receptors involved in various signal transduction pathways. Integrins (and their activation by an intercellular trigger) are the main focus of this thesis, as will be discussed in detail below.

1.2 Integrins

1.2.1 Structure and biology

Integrins are heterodimeric (αβ) adhesion cell surface receptors and can be found in all metazoans (27-43). Simpler metazoans usually have a limited number of integrins. For example, *Caenorhabditis elegans* has only two different types of integrins (44). In mammals there are 18 integrin α subunits and eight integrin β subunits, which form 24 different non-covalently linked integrins (αβ) with distinct but sometimes overlapping functions (45, 46). Each class of integrin receptor usually has different specificity for extracellular ligands which define its function, and thus specific integrins regulate different cell events (46). Each of the two integrin receptor subunits consists of a large extracellular domain with several globular domains linked via flexible linkers, a helical TM domain which spans the cell membrane and a C-terminal cytoplasmic (CT) tail (Fig. 1.2). Integrins

bind to the extracellular matrix (ECM) via their extracellular head domain and to the actin cytoskeleton via CT adaptor proteins, such as talin and filamin. They thus serve as a structural link between the cytoskeleton and the ECM (47). Some of the ECM proteins which interact with integrin extracellular regions include laminin, fibronectin, collagen and vitronectin (48, 49). Despite some similarities in structure and in activation mechanisms among the 24 different integrins, they can be further subdivided into integrins which are activated at specific times (such as αIIbβ3 in circulating blood platelets to cause clotting) and integrins with longer periods of activity where structural integrity is desired (such as β1 integrins in striated muscles) (29, 50).

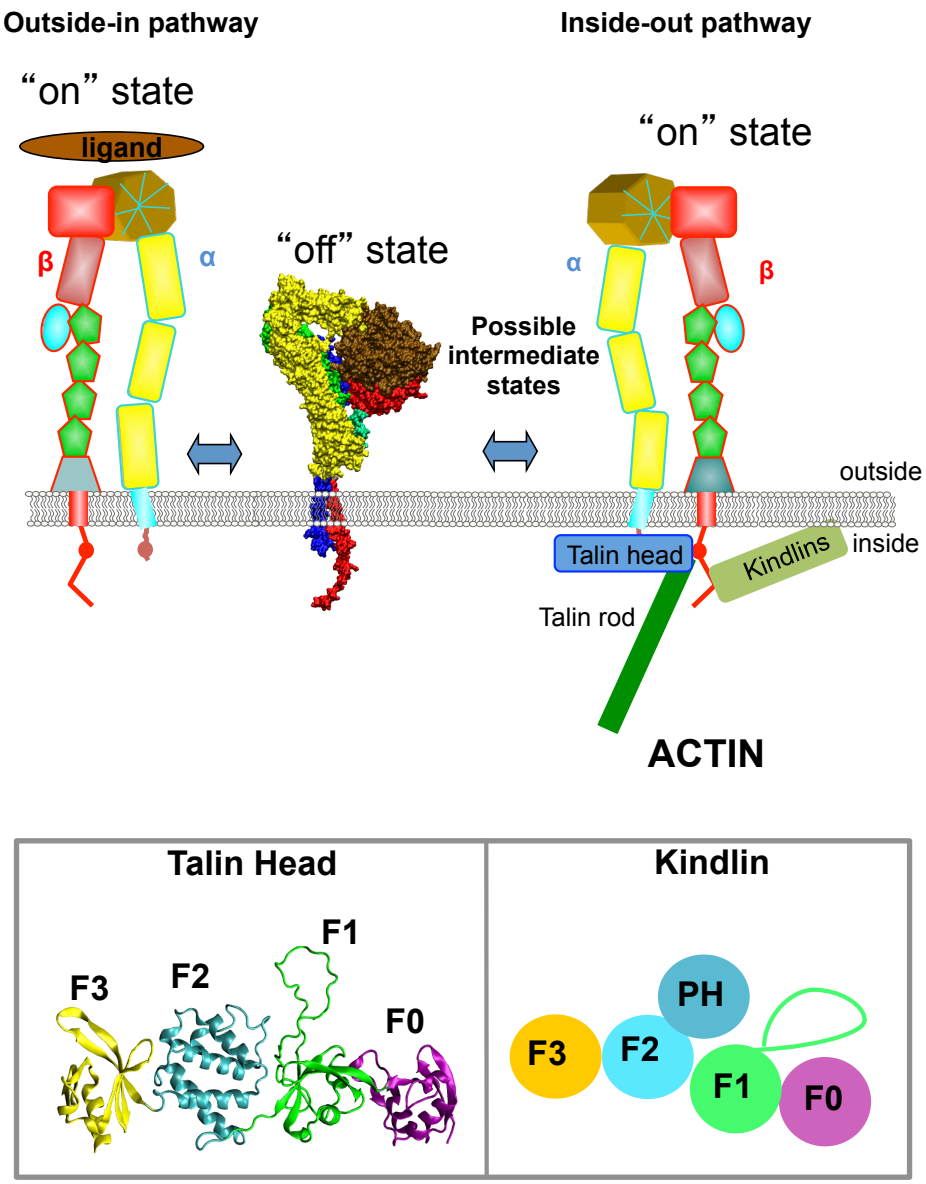


Figure 1.2: Schematic view of integrin activation. The inside-out and the outside-in activation mechanisms are presented with possible active conformation of integrins. In the picture below, the structures of the talin and kindlin FERM domains are shown.

Integrins uniquely form a bidirectional signalling pathway across the cell membrane (i.e. both inside-out and outside-in; see Fig. 1.2). In the inside-out activation pathway, CT activating proteins (e.g. talin and kindlins) bind to the CT part of integrins triggering conformational changes in the receptor (Fig. 1.2). In the active state the adhesiveness of the integrin ectodomain and its affinity for extracellular ligands increases (51-57). The inside-out activation mechanism by talin is the main focus of this thesis and is discussed below in detail. There is also evidence that another family of proteins, the kindlins (58), act in synergy with talin to induce integrin activation (59). Kindlins and their role in the integrin 'inside-out activation' are discussed in Section 1.2.5. In the outside-in activation pathway, binding of extracellular ligands induces allosteric conformational changes in the receptor and further downstream signalling (Fig. 1.2) (46, 60).

Integrins play a crucial role in many cellular processes such as cell differentiation, proliferation, survival, migration and apoptosis (61, 62). They also play a central role in cell adhesion (63, 64), which is an essential process in all living organisms that facilitates the interactions of the cells with their environment (i.e. the ECM or adjacent cells). After activation of individual integrins, clustering is believed to occur, creating focal adhesion complexes and increasing the integrin ligand binding surface (61, 65, 66). The dynamic nature of integrin activation and deactivation for most integrins (except myeloid or lymphoid integrins) induces different integrin adhesion morphologies and shapes which range from weak to more stable focal adhesion complexes (63). The formation of these assemblies is regulated by additional factors such as tension (67) and involves hundreds of membrane-spanning proteins, adapter proteins and enzymes which form the adhesion complexes via a large number of protein–protein interactions (68).

Integrins are crucial in the development of multicellular organisms and incorrect expression is associated with many defects like thrombosis, cancer and the immune response. For example, overactive cell adhesion promotes platelet aggregation and thrombosis (69), while inefficient adhesion promotes cancer metastasis (70, 71). Other associated defects include leukocyte functions, inflammation, osteoporosis, diabetes, hepatitis, haemostasis and angiogenesis (45, 72-74). The αIIbβ3 integrin, which is extensively studied in this thesis, is found in platelets and is involved in arterial thrombosis (69). Thrombus formation is initiated either by collagen binding to glycoproteins in the platelet surface or by thrombin. Binding of these proteins triggers αIIbβ3 integrin binding to fibrinogen, thus resulting in the spontaneous assembly of platelets in arteries (69). If thrombus formation occurs in wounded tissues then it prevents further bleeding while

inappropriate activation can cause clotting and prevention of blood circulation. Recent studies have also shown that GPCRs are involved in cross-talk to integrins (75). Additionally, αIIbβ3 and αVβ3 integrins are expressed in tumour cells and promote tumour progression and metastasis (70). Other integrins such as β2 and β7 integrins are expressed in leukocytes. The β2 integrin activation via cross-talk with the β7 integrin is part of the immune response (76). Defects in β2 integrins can lead to leucocytosis or inflammation (77). These examples demonstrate that tight regulation of integrin activation is crucial in preventing pathological and adhesive defects.

1.2.2 The integrin ectodomain and activation

The integrin receptor can be considered as three main regions: the ectodomain, the TM domain and the CT region. In the ectodomain, the β subunit is comprised of ~700 residues whereas the α subunit consists of ~1000 residues. The integrin ectodomain region is formed from the "head domain" (which consists of the β-propeller, the Hybrid domain, the β-I domain and, in some integrins, its homologous α-I domain) and two integrin legs (78-83). The α subunit leg is formed by a Thigh domain, a Calf-1 domain and a Calf-2 domain, whereas the β subunit leg is formed by a PSI domain (84), four EGF domains and the β-tail domain (Fig. 1.3). In eight integrin α subunits, an additional domain (α-I) exists which is homologous to the β-I domain (30). It has been shown that if the integrins have the α-I domain then it is the main ligand binding site, otherwise the β-propeller and the β-I domains provide the main ligand and metal ions binding sites (29). Structural and functional studies of the I domains show that they contain three metal ion (e.g. Mn^{2+}, Ca^{2+}, Mg^{2+}) binding sites (called MIDAS, ADMIDAS and SyMBS), which coordinate the binding of ligands to the integrin ectodomain (85-89). Some of the ligands which bind in a way that involves the metal ion binding sites are collagen (90), ICAM-1 (91), iC3b and RGD (Arg-Gly-Asp) peptides (92). For example, in the case of collagen (90) it was suggested that prior to collagen binding, rearrangement of the ion binding site occurs to create a new surface to accommodate the collagen.

The first high-resolution structure of the complete integrin ectodomain was published in 2001 for the αVβ3 integrin (83). This structure was in a bent conformation with a knee in the α integrin leg between the Calf-1 and Thigh domains and in the β integrin leg between the EGF1 and Hybrid domains (Fig 1.3). After this structure appeared, a series of ectodomain structures followed, including the human platelet αIIbβ3 integrin structure using cryo-electron microscopy and X-ray crystallography (82), the

αIIbβ3 integrin (78), the αIIbβ3 in solution using small angle neuron scattering (81), the αVβ3 integrin structure in complex with an αIIbβ3 TM dimer (79) and the αXβ2 integrin structure (93). All these structures were in a bent conformation similar to the αVβ3 structure (83). Nowadays it is generally accepted that this is the inactive conformation of integrins, although this form can still bind some ligands (see further discussion below).

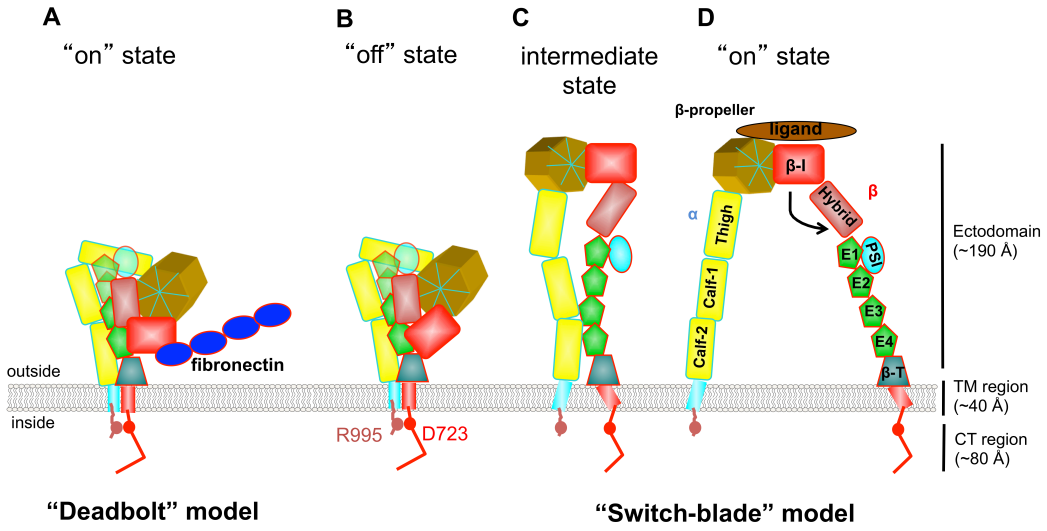


Figure 1.3: View of integrin conformations. A. The "deadbolt" model. Possible active conformation of integrins where the affinity of ligands is increased because of some moderate changes in the orientation/interactions between the β-I and the β-propeller domains. B. Bent conformation of integrin extracellular region. This conformation is considered as the inactive conformation of integrins. C,D. The "switch-blade model". C. Possible intermediate conformation of integrins where the ectodomain is extended with the headpiece in a closed state. D. Integrin active conformation with the ectodomain in an extended conformation with the β-I/Hybrid hinge open.

Upon activation, conformational changes are expected within the ectodomain (94), but there is still debate about the nature of these changes (37, 39, 95). Currently, two main functional models are considered: the "deadbolt" (96) and the "switch-blade" (43) models (see Fig. 1.3). The "deadbolt" model suggests that during the integrin activation the two integrin legs remain bent and the affinity for ligands is increased as a result of some moderate changes in the orientation/interactions between the β-I and the β-propeller domains (Fig 1.3A). In contrast, the "switch-blade" model suggests extension of the ectodomain involving changes in the β-I/Hybrid domain interface (Fig. 1.3C,D). As a result, the β-I/Hybrid hinge opens leading to separation of the two legs. Some studies also suggest a third (intermediate) state (in the "switch-blade" (43) model) where the

ectodomain is extended with the headpiece in a closed state (Fig. 1.3C) (78, 97). Evidence for both models has been found.

Evidence for the "deadbolt" model came from the Xiong *et al.* structure of the αVβ3 integrin in complex with an Arg-Gly-Asp (RGD) peptide (92). The RGD ligand was situated between the β-propeller and the βA domains and despite some changes induced in the orientation and the interactions between the two domains, the liganded ectodomain remained in a bent conformation. Later, EM and FRET studies gave more support to the "deadbolt" model by demonstrating that no large change in integrin height (~110 Å in active and inactive states) was observed upon integrin activation by Mn^{2+} (80, 98, 99). Additionally, a crystal structure where the αVβ3 integrin was in complex with fibronectin (80) showed that the integrin head was still in a bent conformation, despite the binding of fibronectin.

The "switch-blade" model is supported by EM, crystallographic and electron tomography studies, which revealed extension of the ectodomain upon activation by RGD (Arg-Gly-Asp) peptides, ions or fibrinogen (43, 78, 97, 100-103). In particular, lower resolution structures, mainly for αXβ2, αLβ2, αVβ3 and αIIbβ3 integrins, gave evidence for an extended conformation upon integrin activation with possible intermediate states, although many of the structural details of these conformations were not clear (78, 94, 104-106). Structures with smaller fractions of integrins also favoured the "switch-blade" model during activation. In particular, Xiao *et al.* published an αIIbβ3 integrin structure which contained the β-propeller, β-I, Hybrid and PSI domains, suggesting that the binding of therapeutic mimetics induced a 70 Å separation of the two integrin legs after the reorientation of the β3 I and Hybrid domains (102). Furthermore, structures of open conformations of the αIIbβ3 headpiece were published by Springer *et al.* (in complex with a fibrinogen gammaC peptide) (103) and by Zhu *et al.* (78). A study by Ye *et al*. also suggested an extended active conformation for the full-length αIIbβ3 integrin when it was inserted in nanodiscs (106). Additionally, Zhu *et al.* showed that deletion of residues which participate in the "deadbolt" region inhibited integrin activation (107). Therefore, it seems that the majority of the data favour movement of the ectodomain to an extended state upon activation.

Some of the controversy about the integrin activation models arises from the fact that the majority of the studies used cations (e.g. Mn^{2+}) to study the activation state of integrins. $C\alpha^{2+}$ and Mg^{2+} cations are present in the cell in mM concentrations whereas the Mn^{2+} is present in μM (108). Various studies showed that integrin activation can alter the

cation concentration (109). For example, activation of the α5β1 integrin in neurons increaces intracellular calcium concentration (110). Structural and functional studies demonstrated that some ligands (e.g. collagen (90), ICAM-1 (91) and RGD (Arg-Gly-Asp) peptides (92)) bind to the integrin ectodomain in a way that involves the metal ion binding sites (85-89). It should be noted, however, that it was also demonstrated that different cations have different effects on the integrin activation state (108) (e.g. Mn^{2+} was shown to increase the affinity of the ectodomain for ligands (86, 97) whereas $C\alpha^{2+}$ was shown to inhibit fibronectin binding to the integrin ectodomain (111)). Additionally, the inside-out and the outside-in activation pathways have very distinct differences (112, 113). Thus, technically it is debatable if the activation induced by cations can be considered as a good mimic of the activation pathways. Also, in many studies of integrin activation a membrane-like environment is not present or the integrins lack the TM and CT region; these may be factors which lead to artifacts (30).

Experimental data also suggested that mechanical force can activate α5β1 (114) and leukocyte integrins (115). To this end, steered molecular dynamics (SMD) simulations (78, 116) were used to explore the changes in the integrin ectodomain during force-induced integrin activation. The SMD simulations allow one to mimic a force or a combination of forces which acts on a specific region. Zhu *et al.* (78) used the ectodomain structure of the αIIbβ3 integrin for their SMD simulations and suggested that a lateral force which mimics attachment to the actin cytoskeleton via talin can regulate the transition from the bent to an extended state. Puklin-Faucher *et al.* (116) used only the integrin head piece in their SMD simulations, suggesting that a hinge opening in the β-I/Hybrid domain arises as a consequence of a small movement of a helix in the β-I domain. It should be noted, however, that both studies lacked the TM and CT regions of integrins, and these have been shown to maintain the integrin in an inactive state (see Section 1.2.3) (82, 112, 113, 117-122).

1.2.3 Integrin TM and CT regions

1.2.3.1 TM and CT region topology and structure

Interactions in the TM and CT regions are believed to lock integrins in their inactive low-affinity state (28). The integrin CT region is considerably smaller than the ectodomain in all integrin α and β subunits except β4 (124). The β integrin CT tail is typically unstructured and consists of 45 to 70 residues (119, 121, 122, 125-127), whereas the β4 CT region consists of multiple globular domains and contains ~1100 residues (46).

All β integrin CT regions contain a membrane proximal (MProx) region and a membrane distal (MDist) region (Fig. 1.4). Six out of eight β integrin CT regions (β4 and β8 are the exceptions) share sequence similarity, both in the MProx region (HDR[K/R]E sequence) and in the MDist region (two NPxY or NPxY-like motifs; Fig 1.4B). It was also shown that the β integrin CT region is the main interaction partner for many CT proteins (e.g. FAK, IPAC, Shc, talin, kindlin and Dok PTB), which are involved in both integrin inside-out (29, 128) and outside-in activation pathways (129). The α integrin CT region consists of 15 to 78 residues and all α subunits share a conserved GFFKR motif in the MProx region (Fig. 1.4A) (28, 45, 118, 119, 130-132). In the α integrin MDist region there is significant sequence variation among the different integrins and fewer interaction partners are known for α than for the β CT region. The unstructured CT region for both integrin subunits is followed by a hydrophobic α-helical region (~29 residues in the β subunit (132) and ~23 residues in the α subunit (131); see Fig. 1.4), which is embedded in the membrane. Note that proteins that associate with the α integrin CT region, for example CIB1, can have an inhibitory role in talin-mediated integrin activation (133).

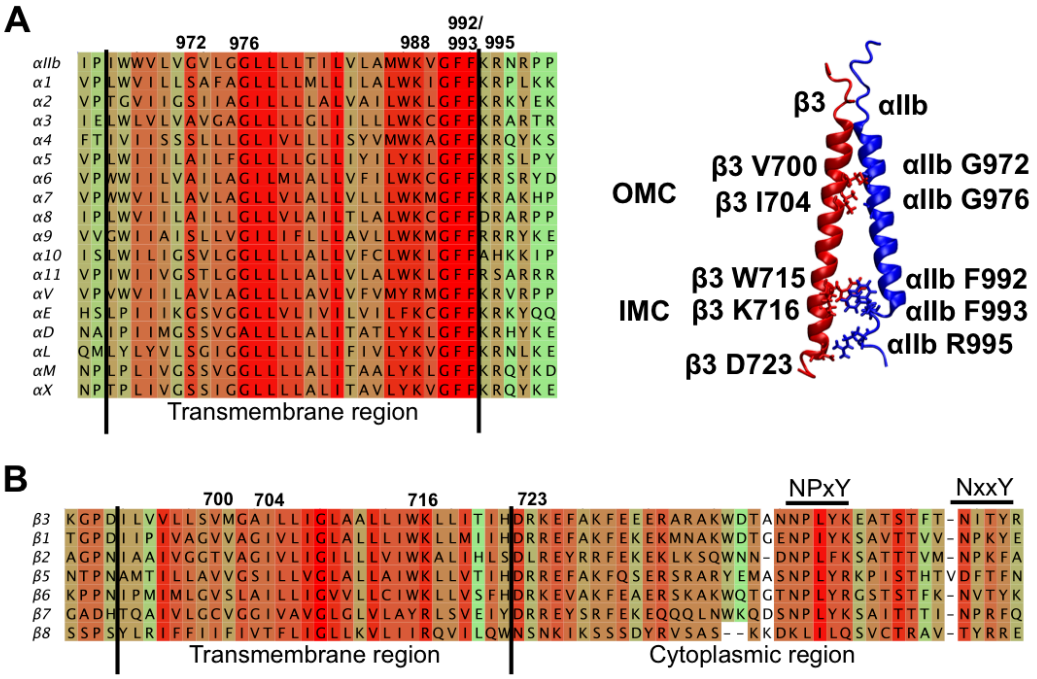


Figure 1.4: A,B. Sequence alignment of all the human integrin α (A) and β (B) TM and CT regions. The β4 sequence which has a long CT domain is omitted for clarity. Conserved amino acids in sequences are presented on a colour scale from green (low conservation) to red (high conservation). The colouring was created using Jalview (123). The key interaction regions in the TM region are shown in the αIIbβ3 structure (2K9J). The membrane proximal (MProx) and membrane distal (MDist) NxxY motifs are also indicated.

A very interesting feature of the integrin TM and CT regions is that both α and β subunits have a positively charged amino acid in the region which is expected to be located at the lipid headgroup region, followed by four or five hydrophobic residues and another charged amino acid (Fig 1.4). The presence of this positively charged residue (K716 for the β3 and K988 for the αIIb) in combination with an adjacent tryptophan or tyrosine residue in most integrins (i.e. W[K/R]motif) is believed to provide an anchor to the lipid headgroups, which regulates the tilt angle of the two helices (41). Recently, Kim *et al.* identified interactions between the β3 K716 residue (K716 (ε-NH_3)) and the lipid headgroups (PO_4^-) and suggested that this interaction controls the tilt angle of the β3 integrin tail. Thus, this motif plays a central role in the formation of both the outer membrane clasp (OMC) and the inner membrane clasp (IMC) interactions (by regulating the helix–helix crossing angle; Fig 1.4) (134). The authors also showed that mutation of this residue to amino acids other than positively charged ones moved the integrin equilibrium towards an active state, possibly by perturbing the β subunit tilt angle. In a subsequent study, the authors showed that embedding of the β tail in the membrane was increased on binding of the talin head domain (135). In contrast, Zhu *et al.* (120) suggested that a mutation in the β3K716 residue leads to integrin activation because it disrupts the H-bonds between the lysine sidechain and the αIIb Cα atoms and not because of its interactions with the lipid environment.

Earlier studies by Lau *et al.* provided structural data about the isolated αIIb (131) and β3 (132) TM regions in phospholipid bicelles and highlighted the role of the WK motif (see above) in the tilt angle of the two helices. These studies demonstrated that, when the β3 helical region is inserted in a bilayer, it needs to adopt a ~30^o tilt which positions the β3 K716 sidechain in the lipid headgroups, where it interacts with negatively charged groups (e.g. PO_4^-). In this position the five hydrophobic residues which follow the K716 residue (L717-I721) are also situated in the lipid headgroup region and the aromatic sidechain of the adjacent β3W715 residue lies in the lipid/water interface, a place tryptophans prefer (136). These residues remained within the bicelles when the β3 helix was inserted in bicelles containing lipids with shorter tails. This suggests that it is more energetically favourable for the helix to tilt further than to place these residues in the aqueous environment. These results are consistent with early glycosylation mapping experiments, which demonstrated that the conserved lysine residues (e.g. 752 in β1 and 724 in β2) and the adjacent hydrophobic region in the β integrins (i.e. β1, β2 and β8 subunits) are located in the lipid/water region, in a similar position to the β3 K716 residue described above (137,

138). In the same studies, the lysine residue (K1125 in α2 and K1022 in α5) prior to the conserved GFFKR motif were also found to be embedded in lipid headgroups, suggesting a role for this lysine in fixing the tilt angle of the α integrin subunit. The shorter hydrophobic length of the αIIb helix, in combination with the conserved lysine (K989 in αIIb), positions the αIIb helix in a vertical position relative to the bilayer normal as suggested by nuclear magnetic resonance (NMR) structures (118, 131). Interestingly, this vertical orientation positions tryptophans (W967, W968 and W988) at each side of the helix in the lipid/water interface and the two phenylalanines in the GFFKR motif in the C-terminus against the β3 helix. An NMR study using DPC micelles also showed that β3 residues K716-I721 and the αIIb K898-F993 residues are located in the lipid headgroup region (139).

NMR data also revealed two different structures for the TM region of the αIIbβ3 integrin (118, 119). The first structures obtained were of the TM regions in phospholipid bicelles (118, 140). The αIIb and β3 helices were packed in a right-handed fashion with a helix–helix crossing angle of ~-25^{o}. Inspection of the structure suggests two main areas of interaction between the two helices. In the OMC region the presence of a Gx_3G motif in the αIIb helix and glycine residues in the β3 domain (G708) allow close packing of the two helices, while in the IMC region a reverse turn allows the position of the F992-F993 towards the β3 tail. A salt bridge between the αIIbR995 and the β3723 is formed in the IMC region (118). The second NMR structure (119), obtained using a CD_3CN/H_2O (1:1) solution mixture, revealed similar packing of the two helices in the OMC region but the interactions in the IMC region differ significantly as this region adopted a helical conformation with the F992-F993 sidechains facing away from the β3 helix towards the lipid headgroups. The αIIbR995-β3723 salt bridge was also present in this structure. Note that in both cases no significant variation in the αIIbβ3 TM region within the NMR ensembles was observed. In 2009, Zhu *et al.* used the structural restraints from their disulphide linking experiments to model the TM region of the αIIbβ3 integrin. Some of the predicted structures overlaid very well with the Lau *et al.* dimer despite some small differences. Surprisingly, the unusual turn in the αIIb IMC region, which orients the two phenylalanines towards the β integrin tail, was also present in the model (32, 120). Therefore, the Lau *et al.* structure seems to be more biologically relevant and is the one given most attention in this thesis.

Prior to the above structures of the αIIbβ3 dimer, the structure of the CT and TM regions of the isolated β3 tail were studied by NMR in DPC micelles; α-helical structures

were observed for the β3 tail TM region and for CT regions which include residues K725-A735 and the NPxY motif (Fig. 1.4) (126). However, the formation of a helical conformation in the NPxY motif in solution is rather improbable since it was not observed in earlier (121, 127) or later studies (118, 119, 125, 131, 132). Ulmer *et al.* (127) used a triple-glycine linker to covalently link integrin αIIb and β3 CT regions to study the α/β complex formation. This study did not reveal any structural data for the α/β dimer and no α/β association was detected in the CT region. These results contradicted earlier studies by Haas *et al.* and by Vallar *et al.* which found interactions between the two CT regions of the αIIbβ3 integrin, possibly stabilized by the presence of cations (141, 142).

Vinogradova *et al.* also used NMR to obtain a structure for the integrin α/β CT region in solution (121). In this structure, the α subunit up to residue P998 and the β subunit up to residue K716 adopted a helical conformation whereas the rest of the β subunit CT region was unstructured (Fig. 1.4). The αIIbR995-β3723 salt bridge was assumed in the α/β CT structure (121) but other hydrophobic and electrostatic interactions between the α and β integrin tails were found. Structures by Weljie *et al.* were also published in the same year suggesting two distinct α/β conformers (143). The αIIbR995-β3723 interaction was found in only one of them and it was suggested that the two conformers represented active and inactive integrin states. Comparison of the structures by Vinogradova *et al.* and Weljie *et al.* revealed very different atomistic details for similar integrin regions suggesting that the quality of their data was not sufficient to determine the α/β structure. Comparison of both structures with the structures of αIIbβ3 integrin TM region by Lau *et al.* (118) reveals a different arrangement of the two integrin tails, suggesting that the presence of the TM region is crucial for obtaining a 'correct' αIIbβ3 structure.

NMR structures for the αIIb (144) and the αIIbβ3 CT regions (139) in DPC micelles were also determined by Vinogradova *et al.* Additional interactions of the β3 tail NPxY motif with the lipids were found and a model was presented wherein talin binding disrupted the NPxY/lipid interactions and led to CT region dissociation from the membrane and an upward movement of the β integrin subunit in the membrane (139). The idea that the "off" and "on" states of the integrin β tail may regulate integrin activation was further supported by Metcalf *et al.*, who used NMR to propose a model where the integrin β tail interacts with the membrane (via two additional transient amphipathic helices) with the same residues required for talin binding. These interactions lock the integrin in a "fully

inactive conformation". The dynamic nature of the β CT region would allow talin and kindlin binding as well as integrin activation (145).

In addition to these αIIbβ3 structures, the structure of the CT region of another integrin type (i.e. αLβ2) was obtained using a combination of NMR and modelling (122). The interface of this structure was very different from the αIIbβ3 structure, with the presence of three helices in a compact form in the αL CT region. This result was unexpected since the two integrin types share a high degree of sequence similarity. One possible reason for this result may be the fact that the structural model was based on chemical shift perturbations. Later studies of the complete TM domain of the αLβ2 integrin using either cysteine-scanning (146) experiments or coarse-grained MD simulations (147) revealed a Gx_3G-type interaction similar to the αIIbβ3 OMC region. Both studies also identified a role for β2W686 in the dimerization of the helices. The key role of the β2W686 in αLβ2 association was later confirmed by SMD simulations (148).

Besides experimental efforts to identify the interactions in the integrin TM region, several computational models have been constructed. The Zhu *et al.* study (120) was discussed above. Gottschalk *et al.* combined computational approaches with experimental restraints for the α/β interface and for the α/β salt bridge. The results suggested a right-handed coil–coil packing of the integrin TM region with the $972Gx_3G976$ region at the interface of the α/β dimer in a similar fashion to glycophorin A (GpA) (149-157). Partridge *et al.*, using a combination of experimental mutations and Monte Carlo (MC) simulations, predicted two left-handed packing dimers with a crossing angle of 30^o and 40^o respectively (158). In the same year, Li *et al.*, using experimental data and MC methods, predicted right-handed packing with the 'correct' interaction in the Gx_3G motif but not for the β3 G708 residue (159). Lin *et al.,* using all 144 possible dimers which can be derived from dimerization of the integrin subunits in mammals, performed a grid conformational space search and identified two models (inactive and intermediate) that were evolutionarily conserved (160). Other models of the integrin TM region were also constructed using coarse-grained MD simulations (147, 161). This method is described in detail in Chapter 4 of this thesis. Finally, Metcalf *et al.,* using MC algorithms, experimental constraints and/or a protein database search, produced structures of the αIIbβ3 OMC region which were in good agreement with the Lau *et al.* NMR structure (118, 162).

1.2.3.2 Interactions in the TM and CT regions and activation

In addition to the structural efforts described above, several experiments have been performed to identify the key interactions between the integrin α and β TM and CT regions. Cysteine-scanning experiments, using an intact αIIbβ3 integrin, revealed specific interactions between the two integrin helices with a preferred surface (113). Addition of an extracellular ligand to the sample resulted in no binding to the integrin receptor and therefore this state of the TM region was considered to be the inactive, low-affinity state. In contrast, experiments with activating mutations in the IMC region revealed non-specific interactions. The mutated integrins were now able to bind ligands and therefore this state was considered to be an active state. Additionally, mutation of αIIbG991, αIIb992FF993 or αIIbR995 residues to alanine (A) or complete truncation of the αIIb991GFFKR995 region resulted in integrins with increased affinity for extracellular ligands (142, 163-165). Hughes *et al.* demonstrated that deletion of the 717LLITIHD723 region (i.e. IMC) of the β3 subunit, which was shown to interact with the αIIb991GFFKR995 region, also resulted in active integrins (166).

In 2005, Partridge *et al.,* using random mutagenesis, identified mutations which resulted in integrin activation only in the OMC and IMC regions, suggesting that these regions rather than the CT region are mainly responsible for maintaining the integrin inactive state (158). What is more, GALLEX assays which included only the integrin OMC region of α2, α4, αIIβ, β1, β3 and β7 subunits (i.e. 16 or 17 residues long with truncated integrin CT region-Δ989GFFKR995) showed weak homodimerization and heterodimerization of the two integrin subunits (except for the α4β7 integrin), which in all cases was weaker than the GpA homodimerization that also contains a Gx_3G-like motif (117). This weak dimerization may be due to the lack of the IMC region, which was shown to be critical for maintaining the integrin inactive state. Li *et al.* used NMR to study the αIIb and β3 integrin tail dimerization in phospholipid micelles. In these experiments, only homo-oligomerization of the α or β subunits (i.e. mainly homodimers for the α subunit and homotrimers for the β subunit) was detected, possibly due to weak helix–helix interactions (167). This oligomerization may also be connected to integrin clustering. Using αLβ2 and αMβ2 integrins, Lu *et al.* showed that complete deletion of the β2 TM region or deletion of the MProx region resulted in activated integrins (168). In the same study, replacement of the integrin TM region with acidic and basic peptides which formed stable interactions constrained integrins in their inactive state. In contrast, replacement with peptides which did not form stable TM interactions resulted in active integrins, demonstrating the crucial

role of TM interactions in regulating activation. Additionally, Kim *et al.* used FRET to study the αLβ2 integrin TM region orientation upon activation (112). Activating mutations in the αL GFFKR motif, addition of talin or triggering outside-in signalling with ICAM or Mn^{2+} led to significant changes in the FRET signal. The authors correlated this change with integrin tail separation but the large size of the fluorescence probes and the fact that they were attached at the C-termini could indicate either separation or a more complex movement of the helices upon activation. A more extensive discussion of this aspect is presented in Chapter 5.

Of particular importance is the salt bridge in the IMC region between αIIbR995 and β3D723 residues. Mutation of either the αIIbR995 residue to alanine or aspartic acid or the β3D723 residue to alanine or arginine moves the integrin equilibrium towards an active state (164). Interestingly, simultaneous mutation of both residues in the salt bridge to opposite charged residues retained integrins in their "off" state (164). Ma *et al.* demonstrated that in αIIbβ3 integrins disruption of the αIIbR995-β3D723 salt bridge was not sufficient to produce fully activated integrin (169) and only complete disruption of the IMC interactions via the triple mutation V990A/F992A/R995D or complete truncation of this region produced fully activated integrins. In the same study, the authors demonstrated that the S752P mutation which disrupts the helical conformation in that region reduced the activation by the αIIbR995D mutation or talin. It did not, however, affect integrin activation when all the interactions in the IMC region were disrupted. What is more, mutations of the residues forming a similar salt bridge in other types of integrins such as αLβ2, α5β1, α4β1 and α4β7 resulted in intermediate or high-affinity integrins (170-172).

Mutations in the OMC region also had a strong effect on the integrin activation state. Substitution of the two glycines in the αIIb972Gx_3G976 motif, as well as the β3G708 residue, by leucine or isoleucine disrupted the close packing of the two helices and led to integrin activation (158, 159, 173). Mutation of other residues in the helix interface or at the interface edges (αIIbL980A, αIIbL981L αIIbL983A, β3I693L, β3I704A, β3I719A) also resulted in activated integrins (159, 173, 174). Hato *et al.* also demonstrated that replacement of the β3 subunit of the αIIbβ3 integrin with a β1 integrin subunit resulted in activated integrins. In the same study, the authors showed that a double β3I719M/E749S mutation, where the first residue was located in the MProx region and the second in the MDist region, greatly enhanced the integrin affinity for ligands (174). Finally, Li *et al.* suggested that mutation of β3 G708 residue to asparagine produced active integrins (175) and promoted integrin clustering.

As mentioned above, the integrin α/β CT tails provide interaction partners for a variety of proteins involved in the integrin inside-out activation mechanism (128). The primary molecule which binds to the CT region is talin (176). An extensive introduction to talin is presented below; its interactions with the CT region are also described in detail in later chapters. The binding of multiple proteins in the same region suggests competition between the proteins, which in turn may control integrin function. For example, studies in Campbell's lab showed that phosphorylation of the Y747 residue in the integrin MProx NPxY motif of β3, β1A and β7 integrin tails resulted in stronger binding of Dok1 and reduced association of talin with the β CT region (177, 178). Furthermore, phosphorylation of the tryptophan W759 on the second NxxY motif of the β3 integrin tail was shown to inhibit kindlin-2 association with the β3 integrin tail and reduce its ability to participate in talin-mediated integrin activation (179). Similarly, the phosphorylation state of the tyrosine in the NPxY motif regulates the binding of Shc to the β CT region, which was shown to initiate integrin outside-in signalling (180). Therefore, phosphorylation provides a switch for controlling integrin activation.

The interactions with various proteins can also explain the unstructured nature of the integrin CT region. An unstructured region allows weaker and more specific interactions as well as more rapid exchange of interaction partners compared to a structured region, which forms more stable interactions (181). This phenomenon can be explained in terms of the entropic cost of ordering a disordered domain on complex formation. A recent study by Anthis *et al.* examined the binding of the two talin isoforms (talin-1 and talin-2) to their β integrin partners (i.e. β1A, β1D, β3) (50). This study revealed that the affinity of talin to the MProx and MDist regions of the β tails is different depending on their structural characteristics and topology. For example, the affinity of talin to the β3 MProx region was higher than the β1 helix due to the presence of a longer region with helical propensity in the β3 MProx region (thus a lower entropic cost). The same study showed that a high-affinity interaction (1,000-fold stronger) could be engineered by replacing β1D residues D776/T777/Q778 with a valine. This was achieved using knowledge of other peptide talin complexes.

1.2.4 Talin and integrin inside-out activation

Talin is an intracellular protein that binds to the integrin CT region and triggers the inside-out activation pathway (57). It consists of a FERM head domain (~50 kDa) and a long rod tail domain (~200 kDa). In mammals, about 50 different FERM domains have

been identified and their role varies significantly depending on their interactions with protein partners and with membranes. One of their major roles is to provide a structural link between proteins in the membrane and the cytoskeleton (182). The majority of FERM domains also interact with specific proteins and lipids to initiate signalling pathways. The talin head domain consists of ~400 residues; it has some unique characteristics that are different from other FERM domains (183). In particular, canonical FERM domains usually consist of three subdomains, called F1, F2 and F3 (182, 184), in a triangular arrangement (185). On the other hand, the talin head domain has its subdomains in a linear arrangement with an extra F0 domain before the F1 domain and a large negatively charged loop insertion in the F1 domain (Fig. 1.5) (186).

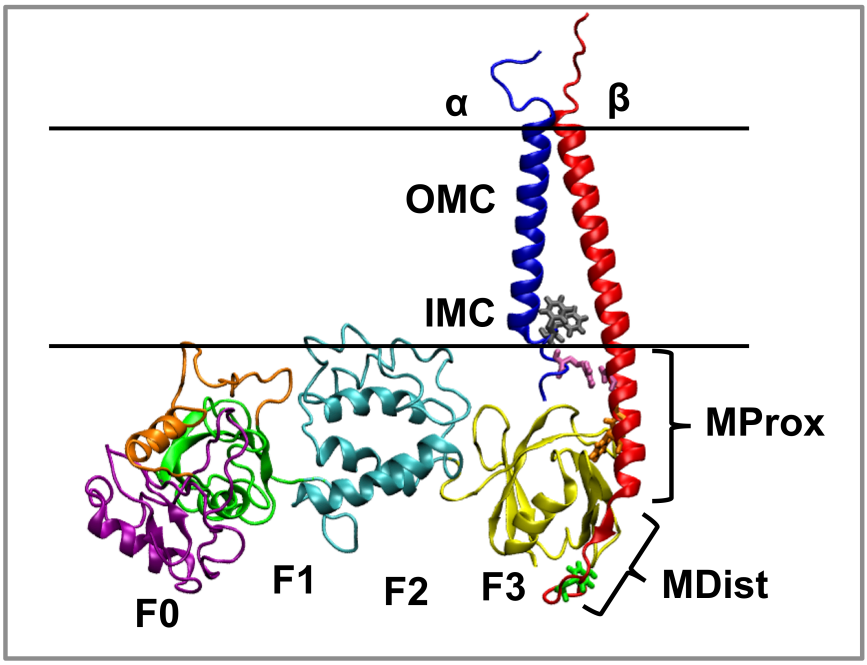


Figure 1.5: The talin head FERM domain in complex with the integrin TM region. The model was constructed as described in Chapter 5 (Section 5.2). The integrin α subunit is shown in blue, the integrin β subunit in red, the F0 in purple, the F1 in green, the F2 in cyan and the F3 in yellow. The loop in the F1 domain is shown in orange. The black line indicates the phosphate atom positions in the bilayer. The interactions in the IMC region are shown in grey (interactions of two phenylalanines with the β3 tail) and in mauve (α/β salt bridge). The MProx and MDist regions where key interactions between the integrin β subunit and the talin F3 domain occur are also shown. In the MProx region, the F3 domain interacts with two phenylalanines in the β3 tail (orange) and in the MDist region the F3 domain is associated with the integrin β tail NPxY (or NPxY-like) motif (green).

In vertebrates, talin has two isoforms: talin-1, which is expressed throughout the body, and talin-2, which is expressed mainly in muscles and the brain (187). The two talin isoforms exhibit high sequence identity (74 %) and similarity (86%) and share some

overlap in function (188). Experimental results using mice have shown that talin-1 is crucial for integrin activation in platelets and leucocytes (54). Talin-1 inactivation/depletion results in defects such as embryonic angiogenesis in early embryos (189), thrombosis or haemostasis (54). The talin-2 isoform functions at striated muscle cells where it binds relatively tightly to β1D integrin tails and it is important in the formation of stable focal adhesion complexes (188).

Talin triggers the integrin inside-out activation pathway via direct interaction of its F3 subdomain with the β MProx and MDist regions of integrins (Fig. 1.5) (52, 53, 57, 125). The F3 subdomain is a typical phosphotyrosine binding domain (PTB) and its association with integrins is sufficient to activate certain integrins, such as αIIbβ3 (51-54, 125, 190, 191). Activation of other integrins such as those containing a β1 subunit requires the presence of the F1 and F2 domains or the full-length talin (56). Full-length talin is also required to facilitate focal adhesion formation and cell spreading (192). The F3 PTB domain is associated with the integrin β tail NPxY or NPxY-like motifs and phosphorylation or substitution of the tyrosine in the motif with alanine inhibited integrin activation (Fig. 1.5) (178, 191). Other PTB-containing proteins such as ICAP-1 and Dok-1 bind to the same NPxY-like motifs, but only talin activates integrins (190).

The isolated F2 domain is unable to activate integrins on its own but the F2-F3 domain pair enhances integrin activation compared to F3 alone (191). The F2 domain is a bundle of α-helices with a large positively charged patch on one side which interacts with negatively charged moieties in the membrane (29, 125). Disruption of this electrostatic interaction significantly reduces integrin activation (29), demonstrating the crucial role of negatively charged lipids. Binding studies using the complete head domain suggested that talin/lipid electrostatic interactions are regulated by the net negative charge of the membrane and not the lipid type (i.e. POPS, PIP_2) involved (186). It was shown, however, that interactions of the talin head domain with PIP_2 lipids promote integrin clustering (193). The key role of the talin/lipid interactions in inside-out activation was highlighted by recent studies using full-length αIIbβ3 integrins in phospholipid nanodiscs, which showed that talin/phospholipid association was necessary for talin-mediated activation of unclustered integrins (106).

The talin F0 and F1 domains are connected to the F2-F3 pair via a flexible linker. F0 and F1 domains have a similar ubiquitin-like fold with a large 30-residue insertion (186, 194) in the F1 domain. A recent study suggested that their contribution to the

association of talin with the membrane is minimal (195) and many of the details for their role/interactions in the talin-mediated integrin activation remain elusive.

The talin C-terminal rod domain consists of a series of four or five α-helical bundles (196-198) with multiple vinculin and actin binding sites (197, 199-202). The rod domain also interacts directly with the regions of the talin F3 domain which associate with the β integrin tail (203, 204). This induces an auto-inhibitory mechanism which regulates talin function. Studies by Yan *et al.* which compared the apparent affinity of the talin head domain and the intact talin with a β3 integrin suggested a 6-fold higher affinity in the absence of the talin rod (205). Studies using full-length αIIbβ3 integrins also demonstrated that integrin activation was enhanced when talin lacked the rod domain (204). Different models have been proposed to explain how talin is released from its auto-inhibitory state. One suggests that calpain (i.e. proteolysis) cleaves a flexible linker between the talin head domain and the talin rod domain, thus releasing the head domain (205-207). A second model suggests that interactions of talin with other proteins involved in the inside-out activation pathway, such as phosphatidylinositol phosphate kinase type 1γ (PIPK1γ) (208) or Rap 1 (209, 210) may activate talin. A third model suggests that talin is released from its auto-inhibitory mechanism via direct interactions with negatively charged moieties (e.g. PIP_2 or POPS) in the membrane. Indeed, studies by Goksoy *et al.* and by Martel *et al.* suggested that the presence of PIP lipids enhances integrin activation by intact talin (204, 211). Several studies suggest that the regions shown to interact with the talin rod domain also interact with the membrane directly, thus unmasking talin (193, 211). A recent study also suggested that unmasking of talin may include a "push-pull" mechanism which leads to an interface mismatch between F3 and the rod domain, which releases the auto-inhibitory mechanism (212). Some studies also suggest that the rod domain has an integrin-binding site which can interact directly with the β integrin tail (213-217).

Upon talin head domain binding to integrins, interactions of the talin rod domain with actin provide a direct link between integrins and the actin cytoskeleton (218). Association of talin with the actin cytoskeleton may also induce a lateral force to the integrin receptor which could play some role in activation (219, 220). It should be noted, however, that the concept of force-induced activation in talin-mediated integrin activation is still controversial because talin can also activate unclustered integrins without an applied force (106). Nevertheless, the presence of force is believed to be important in the formation of focal adhesion complexes (78). Indeed, recent studies showed that the binding of intact talin to integrins is needed for focal adhesion formation (221).

The first structural data for a talin/β CT tail complex were reported by Garcia-Alvarez *et al.* in 2003 when they published the talin-1 F2-F3 domain pair structure in complex with the CT part of a β3 integrin tail (i.e. 739WDTANNPLYDEA750) (222). This structure revealed that the Y747 sidechain in the 744NPLY747 motif is situated in an acidic pocket in F3 domain, whereas the β3 W739 residue is placed in a hydrophobic pocket. Later studies demonstrated that substitution of one of the residues which form the hydrophobic pocket in F3 with alanine (F3R358A) significantly reduced integrin activation (Fig 1.5) (42). The talin F3 domain was also shown to interact with the layilin receptor CT domain (223) and the PIPK1γ (208, 224) in a similar manner, although there were minor differences. Affinity measurements by Wegener *et al.* revealed that PIPK1γ has a higher affinity for the F3 domain compared to layilin (223). In 2007, Wegener *et al.* used a chimeric peptide in which the β3 peptide was attached to a PIPK1γ90 peptide to solve the structure of the F3 domain in complex with a β integrin tail using NMR (52). This structure suggested that an acidic loop in the F3 subdomain interacts with two phenylalanines (F727 and F730) in the β3 CT region and that mutations of these two phenylalanines or the F3L325 residue in the F3 acidic loop reduced the αIIbβ3 activation (52, 225). Prior to Wegener *et al.*, Ulmer *et al.* also identified interactions between talin and the same regions of the β3 integrin tail (55).

In 2009, crystallographic studies revealed the structure of the talin-2 F2-F3/β1D integrin tail complex (125). In this structure, the F3/β1D interface was similar to the one described above (52). Interestingly, in this study a positively charged region on F2 was also identified and shown to be crucial for inside-out activation. The authors also suggested that after talin binding, a talin/β tail salt bridge is formed which disrupts an α/β salt bridge. After the disruption of the α/β salt bridge, talin induces a 20^{o} reorientation of the β integrin tail in the membrane, thus weakening the α/β TM association. It should be noted, however, that no direct evidence for the suggested reorientation of the β integrin tail by talin was presented and further investigation is therefore required to understand how talin regulates rearrangement of the integrin TM region. Recently, Kim *et al.* also provided evidence that the binding of talin head domain increases the β integrin TM domain tilt angle, although the direction of the change in the β integrin TM domain orientation is different compared to the Anthis *et al.* model (135). In Chapter 5, a detailed investigation of the effect of talin binding to the integrin TM region is made and a mechanism for talin-mediated integrin activation is proposed.

In 2010, Goult *et al.* have published the NMR structure of the talin F0-F1 domain pair (194). In their study, the authors suggested that some regions of the dynamic loop in the F1 domain form a helical structure (i.e. regions 144-152 and 154-167). CD spectrometry measurements showed that the interactions with acidic lipids stabilized the helical region in the loop. Based on their results, the authors suggested that the role of the loop is to interact with the membrane and to locate the F0-F1 closer to the membrane (194). Extrapolation of these results to a complex formed from membrane, intact talin and integrin tails is uncertain since these studies were performed either with the isolated loop or with the isolated F1 domain. Indeed, a recent study by Moore *et al.* which quantified the binding of the talin head domain to acidic lipids and the β3 integrin tail suggested that the affinity of the talin head domain to lipids is similar to the F2-F3 binding energy and thus the F2-F3 domain pair dictates the talin binding to the membrane (195). Therefore, further analysis is required to understand the role of the F0-F1 domain pair in integrin activation. A full investigation is performed in chapters 3 and 5 of this thesis. Finally, Elliot *et al.* put forward the structure of the talin FERM head domain (186). This structure was obtained in the absence of the F1 loop insertion. A detailed discussion and investigation of this structure and its role in the integrin inside-out activation mechanism is given in Chapter 3.

The data concerning integrin inside-out activation summarised in this chapter have refined models of the integrin inside-out activation significantly. Details of the talin/integrin/membrane interactions are now clearer and studies on the TM and CT regions of integrins (Section 1.2.3) have defined the interactions in the OMC and IMC regions that are crucial for integrin inside-out activation. These studies also demonstrated the crucial role of the α/β salt bridge in the IMC region. However, controversy remains about the changes within the membrane (i.e. in the TM region of integrins) and within the integrin ectodomain that occur during inside-out activation.

1.2.5 Other proteins involved in integrin inside-out activation

Integrin inside-out activation is a dynamic process in which talin binding can be described as the “final step” (128). Before talin binding, several other proteins/enzymes are involved in signalling and focal adhesion formation. Following an initial trigger via GPCR receptors (e.g. the protease activated receptor in platelets (226)), signalling cascades are initiated which result in the activation of Rap1 (227). Rap1 belongs to the Ras GTPases and has two isoforms (Rap1A and Rap1B), which share significant sequence similarity and possibly overlapping functions (209). Rap1B deficiency in mice reduces αIIbβ3 activation

in platelets, resulting in the protection of mice from thrombosis (209). Similarly, in megakaryocytes Rap1B was shown to increase the αIIbβ3 integrin affinity for fibrinogen (228). Rap1 activation requires the presence of GTP, which is produced by a guanine nucleotide exchange factor (GEF) in the presence of Ca^{2+} and diacyl glycerol (DAG) (29). Subsequently, the activated Rap1 associates with a Rap1-GTP-interacting adapter molecule, RIAM. RIAM interacts directly with talin, thus providing a connection between Rap1 and talin (229). RIAM overexpression enhances talin-1/αIIbβ3 association in living cells and a RIAM knockout inhibits talin-1/αIIbβ3 association (230). Rap1/lipid association results in the recruitment of talin to the membrane, thus promoting talin/integrin association (29, 30). Other proteins such as lamellipodin promote talin-1/αIIbβ3 association in living cells, suggesting that they can replace RIAM function (230).

An alternative activation pathway involves direct interaction of talin with PIPK1γ. De Pereda *et al.* showed that PIPK1γ interacts with the talin F3 PTB domain via the C-terminus SPLH motif (208). Binding to PIPK1γ increases the local PIP_2 concentration, thus promoting talin/integrin association (200). In this pathway, the talin head domain is cleaved from the rod domain (auto-inhibitory mechanism) by calpain protease (205-207). It should be noted that before binding of integrin activators, integrin inhibitors such as ICAP-1 must be removed (231).

Another family of FERM domains which is believed to participate in the inside-out activation pathway is the kindlin family that includes three isoforms (kindlin-1, kindlin-2 and kindlin-3) (33, 58, 232-234). Kindlins have a FERM head domain with similar characteristics to talin and an additional PH domain which is located in the middle of the F2 domain (33). The talin and kindlin head domains share an overall 36%–50% sequence similarity between the individual domains when the characteristics of kindlin which are different from those of talin (i.e. the PH domain and the F1 loop) are omitted (235). It was suggested that when the kindlin F3 domain associates with the integrin β CT tail it forms strong interactions with the integrin MDist NPxY motif (NxxY-795 for β1/β3 integrins and NxxF for β2 integrins) and with a preceding threonine-rich region (788TT789 for β1A and 751TST753 for β3) (232, 233, 236, 237). In particular, Harburger *et al.* showed that mutation of Tyr795 (i.e. in the MDist NxxY motif), Thr788 (β1A) or Thr789 (β1A) residues to alanine inhibited the association of kindlin-1 and kindlin-2 with β1A and β3 integrin tails (236). Similarly, studies by Ma *et al.* showed that mutation in the same regions (i.e. β3 751T[S to P]T753 and Nxx[Y to A]795 regions) again reduced the kindlin-2/β3 association (232). Besides the direct interactions of kindlins with the β integrin CT

region, kindlins were also shown to interact with other proteins involved in focal adhesion. For example, kindlin-2 interacts with integrin-linked kinase (238).

Hypotheses for the role of kindlins include co-activation of integrins with talin. Thus, kindlin binding to the β integrin tail may orient the integrin CT tails for talin binding or inhibit binding of other proteins. Recently, several studies provided evidence for the role of kindlins in talin-mediated integrin activation. For example, kindlin-3 was shown to be essential for platelet αIIbβ3 and α2β1 integrin activation. Removal of kindlin-3 reduced platelet aggregation in mice (233) and elimination of kindlin-3 from leukocytes resulted in reduced β2 integrin activation and leukocyte adhesion deficiency (239). Lack of kindlin-2 or kindlin-1 also inhibit β1 integrin activation, causing peri-implantation lethality (238) and epithelial cell dysfunction respectively (240). Controversy remains, however, about the role of kindlin in talin-mediated integrin activation. For example, overexpression of kindlin-1 and kindlin-2, alongside sufficient talin head domain, activated αIIbβ3 integrins but inhibited α5β1 activation (236). In endothelial cells, kindlin-2 and kindlin-3 were also reported to have distinct functions, regulating different integrins (i.e. αVβ3 and α1β5 respectively) (241). While the participation of kindlins in talin-mediated integrin activation is apparent (237), further studies are required to understand their specific roles.

1.3 Molecular dynamics (MD) simulations of membrane proteins

In recent years, the number of known membrane protein structures has increased rapidly as techniques for obtaining protein structures (e.g. crystallography) have improved (see http://blanco.biomol.uci.edu/membrane_proteins_xtal.html). However, static structures yield limited insights into the dynamic properties of proteins (e.g. information from B-factors or changes in residues upon ligand binding), which are crucial for understanding functional mechanisms. Additionally, membrane protein structures generally lack a realistic membrane environment (except for a few cases with isolated lipids (242, 243)), although this environment is known to be crucial for proper membrane protein function (6). Among the techniques capable of giving information about protein dynamics and functionality, computational approaches are often the method of choice. MD simulations, for example, allow the motion and dynamics of the protein in the membrane to be assessed, thus promoting understanding of the relationship between structure and function (244). The downside is that the prediction of large conformational changes in proteins usually requires long simulation times that are not routinely achievable by classical MD atomistic simulations (AT-MD) despite recent advances in computational

resources and software. Therefore, alternative methodologies that enhance energy landscape exploration, such as metadynamics (245), dynamic importance sampling (DIMS) (246), biphasic mimetic models (247) or coarse-grained models (CG-MD; see Section 2.3) have been developed (248). The CG-MD approach has been particularly successful in studying the orientation and interactions of different types of membrane proteins within the membrane. For example, Scott *et al.* used self-assembly CG-MD simulations to study the localization of a variety of β-barrel and α-helical TM domains in the lipid bilayer and their interactions with lipids (249). Subsequent conversion of coarse-grained (CG) representations to atomistic (AT) resolution (i.e. a sequential multiscale approach; Section 2.4) allows more detailed analysis of the properties of the systems (e.g. the protein–lipid interactions) to be obtained. Recently, this approach was used to identify the interactions between specific lipids (e.g. PIP_2) and Kir channels (250).

MD simulations have also been used to gain insights into the dynamic properties of a variety of membrane protein families e.g. transporters (246), voltage-gated potassium channels (251), ligand-gated ions channels (252) and receptors which are involved in TM signalling (e.g. GPCRs and receptor tyrosine kinases). In the case of GPCRs, simulations of the β1 and β2 adrenergic receptors were used to examine the "ionic lock" (i.e. a salt bridge between two TM helices), which is believed to regulate the receptor activation state (253, 254). Additionally, MD simulations have been used to study the binding and interactions of different ligands or proteins with the same receptors (255-257). MD simulations also identified key interactions which regulate the activation of another signalling receptor (the EGFR receptor) (24, 258, 259) and demonstrated that the interactions of its ectodomain with the membrane promote EGFR activation (260).

MD simulations were also used to study the insertion/orientation of single TM helices in different bilayers (261-263) as well as the self-assembly of inserted TM helices. For example, CG-MD simulations were able to reproduce the packing of the GpA (152, 264, 265) homodimer and to reveal the crucial role of a common Gx_3G motif in the packing of the helices. Dimerization of TM domains in syndecan (161), ErbB (266) and integrin (147) was also studied using a similar approach. The self-assembly of more complex structures such as the BM2 proton channel TM domain and the influenza M2 tetramer (267, 268) were also studied. CG-MD simulations have also been used for 'potential of mean force' (PMF) calculations which give an estimate of the free energy of association/dissociation of GpA (269, 270) and/or the free energy of insertion of a helix in the membrane (271).

In addition to the study of the TM regions of proteins, MD simulations can be used to predict the interaction, orientation and depth of penetration of peripheral membrane proteins or enzymes in lipid bilayers (272-278). For example, simulations were used to predict the orientation of the phospholipase A2 (274) domain in the bilayer and the specific interactions of the GRP1-PH domain (278) or C2 domains (272) with PIP lipids. Both AT-MD and CG-MD simulations were used to study more complex proteins such as BAR domains, demonstrating how these domains change the lipid bilayer curvature (279, 280). MD simulations have also been used extensively to study the organization and dynamics of lipids in membranes. All-atom simulations, which are computationally expensive, were mainly used to study the details of the interactions between different lipids (281-284) or between protein and specific lipids (e.g. cholesterol (283)). Lower-resolution techniques such as CG-MD simulations or dissipative particle dynamics (DPD) simulations (14) have been used extensively to elucidate the interplay between lipid ordered and lipid disordered states in bilayers or liposomes without a protein component (14, 285, 286) or the localization of simple TM or peripheral constructs such as the WALP peptide or the Ras protein (287-289) in complex membranes. Recent studies have also used CG models to study DNA-related phenomena (290, 291).

1.4 Thesis aims and summary

As mentioned above, various structural and functional data have provided important insights into how integrins, talin and the membrane interact during the inside-out activation pathway. This pathway is, however, highly dynamic and despite the many details currently available, the precise mechanism of inside-out activation remains elusive. For example, the mechanism by which talin associates with the membrane or the changes in the TM or the extracellular regions during the inside-out activation are questions not yet fully answered. Therefore, the overall aim of this thesis is to address some of these issues by providing quantitative insights into the dynamics of the integrin receptor and its activating protein talin. The large amount of available structural information about integrins makes them an excellent target for MD simulations, which are used extensively in this thesis to give new insights into the dynamics of the integrin receptor in its biological environment.

In Chapter 2, the theory behind the MD simulation techniques (i.e. CG-MD and AT-MD simulations) is discussed. In the subsequent chapter, NMR and simulation approaches are used to elucidate the interactions and orientation of the talin head domain

relative to a POPC/POPG bilayer. The outcome of the work in Chapter 3 demonstrates how interactions between talin and the negatively charged lipids in the membrane orient talin in a plausible orientation to interact productively with the integrin CT region. Additionally, Chapter 3 also presents an automated high-throughput simulation approach for studying peripheral membrane protein/lipid interactions.

In Chapter 4, a multiscale approach with a high-throughput CG pipeline is used to identify how key interactions and a common Gx_3G motif regulates the packing of the two helices in the integrin TM region. The outcome of this chapter shows that this multiscale approach can produce "NMR accuracy" structures of TM dimers. In Chapter 5, a complex comprised of the integrin TM region and the talin head domain is simulated in order to understand how talin modulates integrin TM region orientation. The results of this chapter suggest a mechanism for the integrin inside-out activation pathway where talin reorients and dissociates the integrin TM helices to promote integrin activation.

In Chapter 6, a model of the complete integrin receptor in complex with the talin F2-F3 domain pair is constructed using recent structural data. The complex was inserted in more biologically relevant symmetric and asymmetric bilayers containing phospholipids, cholesterol and sphingolipids to study the dynamics of the integrin receptor and its effect on bilayer structure and dynamics. This work is a collaboration with Professor Ilpo Vattullainen (Tampere University of Technology). The results demonstrate the dynamic nature of the integrin receptor and suggest that the receptor induces an asymmetric distribution of lipids between the two bilayer leaflets and a slowing-down of lipid diffusion in an annulus of ~30 Å around the protein.

The outcome of this thesis is also important from a more methodological perspective since it demonstrates how a computational serial multiscale approach can be used to study various aspects of membranes and membrane-related proteins. Overall, the work presented in this thesis provides novel information about a biologically important receptor, therefore significantly enhancing our understanding of talin-mediated integrin activation.

Chapter 2

Theory and Methods

2.1 Introduction

In this chapter, the theory behind MD simulations is first discussed followed by a description of the multiscale simulation approach which is used in chapters 3 to 6.

2.2 Classical MD simulations

In classical MD simulations, the Born-Oppenheimer approximation (292) allows the motion of individual atoms to be calculated using Newton's equations of motion. Briefly, the Born-Oppenheimer approximation (292) suggests that the motion of electrons and nuclei within an atom can be treated separately. Due to the large difference between the masses of the electron and nucleus, the motion of the electrons can be adjusted to the motion of the nuclei and therefore the electronic degrees of freedom can be omitted and incorporated in one effective potential. This means that the dynamics of atoms can be described by the nuclei in this electronic potential. Thus, in classical MD simulations, atoms are treated as van der Waals (VDW) spheres with a point charge at the centre of the sphere. This approach allows much larger timescales to be addressed and larger systems to be studied but it is unable to model events where changes in the distribution of electrons are important. In such cases, quantum mechanics (QM) techniques can be used, although they are limited to small systems since they are computationally expensive (293). Alternative hybrid approaches (QM/MM) have been developed where a small part of the protein is modelled at the QM level and the rest of the protein is modelled at the molecular mechanics (MM) level (294, 295). This hybrid QM/MM approach is widely used to study protein-ligand interactions because it allows incorporation of quantum phenomena such as ligand polarization and additionally no force field parameterization is required because the ligand is represented using a QM model (296). Protein conformational flexibility upon ligand binding is also taken into consideration because of the MM representation of the largest part of the system. The major limitation of the QM/MM approach is that considerable computational resources are required and therefore it is limited to small systems. On the other hand, MD simulations do not take into consideration QM effects however much larger and thus more biologically relevant systems that include multiple

protein domains and a realistic environment can be studied. A detailed discussion about the applications of the MD simulations is made in Section 1.3.

Using the Born-Oppenheimer approximation (292) the Hamiltonian, which describes the positions (x) and momenta (p) of the particles of a system, can be simplified to:

$$H=\sum_{i=1}^{N}\frac{p_i^2}{2m_i}+U(x_1,x_2\ldots,x_N) \tag{2.1}$$

where m_i is the mass and U is the Born-Oppenheimer energy surface. In the Newtonian formulation, the equation can be written as follows:

$$F_i=-m_i\,\alpha_i \quad \text{and} \quad \alpha_i=\ddot{x}_i \tag{2.2}$$

where α_i is the acceleration of the particles. Alternatively, the force on each particle can be obtained from the potential U:

$$F=-m\,\frac{dU(x_i)}{dt} \tag{2.3}$$

The potential U in the Eq. 2.3 depends on the positions (x_i) of the particles in the system (i=1, 2, 3,.., N). The evolution of the coordinates (x) of the particles and thus the velocity and the acceleration is obtained by integrating equations 2.2 and 2.3 for all particles. The potential U in classical MD simulations is semi-empirical and is calculated as the sum of bonded and non-bonded (Eq. 2.4) interactions in the system neglecting the quantum mechanical wave functions:

$$U=U_{bonded}+U_{non\text{-}bonded} \tag{2.4}$$

2.2.1 Bonded terms

The potential for the bonded term in Eq. 2.4 is represented by the following equation:

$$U_{bonded}=\sum_{bonds}\frac{k_x}{2}\left(x-x_0\right)^2+\sum_{angles}\frac{k_\varphi}{2}\left(\varphi-\varphi_0\right)^2+ \sum_{impropers}\frac{k_\omega}{2}\left(\omega-\omega_0\right)^2+\sum_{torsions}A_n\left[1+\cos(n\theta-\theta_0)\right] \tag{2.5}$$

The first term in Eq. 2.5 uses a harmonic potential with a force constant k_x to describe the oscillation of a chemical bond (connecting two atoms) around a reference bond length (x_0). In the second term, the bond angle between three atoms (*i,j,k*) is modelled as an angle harmonic potential with a force constant k_φ which represents the angular vibration between three bonded atoms around a reference angle value (φ_0). The force constants k are different for different types of bonds or angles and they determine the

magnitude of the cost in potential energy when the angles or bonds deviate from the reference length/angle. Alternative potentials, albeit sometimes more computationally expensive, exist for modelling the bond and angle potentials (297). For example, in the GROMOS-96 force field, the angle potential is modelled as a cosine-based potential:

$$U_{angles} = \sum_{angles} \frac{k_\varphi}{2} \left(\cos(\varphi) - \cos(\varphi_0)\right)^2 \quad (2.6)$$

The third term in the Eq. 2.5 represents the improper dihedral potential, where ω represents the dihedral angle between four bonded atoms. This potential is essential for maintaining the planarity in the structure (e.g. maintaining ring sidechain planarity); k_ω is the force constant and ω_0 is the reference angle. The last term is the dihedral potential which describes the dihedral angle between two planes of atoms. The planes are defined by atoms *i,j,k* and j,*k,l*. When a group of atoms is in a *cis* conformation this potential is zero. In this term, θ represents the dihedral angle, n is the multiplicity of the dihedral term, A_n is the force constant and θ_0 is the reference dihedral angle. It should be noted that alternative forms of the bonded term also exist.

2.2.2 Non-bonded terms

The non-bonded terms include steric repulsions, attractive VDW forces and electrostatic (repulsion/attraction) interactions. The first two types are represented by a Lennard-Jones (12-6) potential and the electrostatic interactions are described by the Coulomb potential.

$$U_{non\text{-}bonded} = \sum_{i<j} \left\{ \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_r\varepsilon_0 r_{ij}} \right\} \quad (2.7)$$

In Eq 2.7, the first term is the Lennard-Jones (12-6) potential where the attractive term ($\sim r^{-6}$) arises from the dipole–dipole attraction of two neutral atoms and the repulsive term ($\sim r^{-12}$) is due to the Pauli exclusion principle. The σ_{ij} represent the distances between two atoms in which the potential energy is zero (at the position $r^* = 2^{1/6}\sigma$) and ε_{ij} is the depth of the attractive well. The second term in Eq. 2.7 is the electrostatic term which is based on Coulomb's law. The terms q_i and q_j represent the partial charges of two atoms, r_{ij} represents the separation between two atoms (i.e. *i* and *j*), the ε_r term is the relative permittivity and the ε_0 term is the vacuum permittivity. In MD simulations, partial charges are assigned to atoms to model the uniform electron distribution over the atoms. To accelerate the calculations, the non-bonded interactions are ignored above a certain cut-off distance (r_{cutoff}). Of course, the non-bonded interactions cannot be sharply stopped at the

cut-off distance and therefore methods (i.e. shift or switch) are used to smoothen the non-bonded interactions to zero. The use of a cut-off distance for the electrostatic interactions in some cases can lead to artifacts such as ordering in the simulation box and lack of long-range electrostatic interactions and thus in GROMACS one can use the alternative particle mesh Ewald (PME) approach to model the electrostatic interactions (298). The Ewald method assumes that the equation for calculating the electrostatic interactions can be calculated as the sum of terms representing direct (real) space and reciprocal space plus a constant term. The calculation in real space is made explicit using small cut-off distances and in reciprocal space using a Fourier transformation. The main problem with this approach is that the computational cost of Fourier transformation summation in reciprocal space increases as N^2 or $N^{3/2}$ (N is the number of particles), which is very slow for a larger system. The PME approach is based on Ewald's methods but it uses a Fast Fourier transformation, such that both terms are converging much faster compared to the direct calculation of the electrostatic term. In this case, the algorithm scales as Nlog(N) (298, 299). In this thesis the PME approach was used for the AT-MD simulations and the cut-off method in the CG-MD simulations.

2.2.3 Integration algorithms

In MM, different integration algorithms are used to calculate the positions and velocities of the particles using Newton's equations of motion (Eq. 2.2). These algorithms usually use analytical solutions to calculate the derivatives of the equation of motion. Therefore, the trajectory (positions, velocities) is calculated in subsequent conserved time steps (dt). In each time step, the force on all particles is calculated by considering their interactions with neighbouring particles. The acceleration is calculated from the force using Newton's equations of motion, which in turn are used to calculate the positions and velocities. This integration is repeated until a predefined time is reached.

Newton's equations of motion can be also written as quotients of finite time differences:

$$\dot{x}_i \simeq \frac{x_i(t+\Delta t)-x_i(t)}{\Delta t} = u_i(t) \qquad (2.8)$$

$$\ddot{x}_i \simeq \frac{u_i(t+\Delta t)-u_i(t)}{\Delta t} = \alpha_i(t)$$

where $\alpha_i(t)$ is the acceleration of particle *i* at a given time t. By using the Taylor expansion (300, 301), the Euler algorithm can be derived from Eq. 2.8:

$$x(t+\Delta t)=x(t)+u(t)\Delta t \qquad (2.9)$$
$$u(t+\Delta t)=u(t)+\alpha(t)\Delta t$$

The error due to limiting the terms in the Taylor expansion is $(\Delta t)^4$. Algorithms with improved errors in velocities such as the Verlet integrator algorithm are also available. GROMACS uses a variation of the Verlet algorithm, the leap-frog algorithm (302):

$$u\left(t+\tfrac{1}{2}\Delta t\right)=u\left(t-\tfrac{1}{2}\Delta t\right)\Delta t+\alpha(t)\Delta t \qquad (2.10)$$
$$x(t+\Delta t)=x(t)+u\left(t+\tfrac{1}{2}\Delta t\right)\Delta t$$

The leap-frog algorithm has two advantages compared to the Verlet algorithm; the error in velocities is lower and it introduces fewer numerical imprecisions.

2.2.4 Force fields

During recent years several force fields have been introduced such as GROMOS (303-305), OPLS (306, 307), CHARMM (308, 309) and AMBER (310). All these force fields use empirical or semi-empirical equations and their parameterization is a compromise between accuracy and efficiency. The parameterization of these force fields is based on quantum calculations and/or experimental/chemical data. One major difference is the way they treat the hydrogen atoms and calculate the non-bonded interactions. The non-bonded interactions are the most computationally demanding part of the energy function since the number of non-bonded interactions increases as the square of the number of interacting particles. Some force fields therefore use an approach where the bonded atoms are grouped into a single united atom. For example, in GROMOS (303-305) the non-polar hydrogen atoms are considered to be part of a larger united atom that retains its chemical properties. This approach reduces significantly the number of atoms in the system and therefore the computational cost. Other force fields such as OPLS-AA (306, 307) treat the hydrogen atoms explicitly. The choice of the correct force field for each system is a crucial step in obtaining the correct results. In this study all the simulations were performed using the GROMOS 43a1 (303-305) force field with a SPC water model, with the exception of a control simulation in Chapter 4 which was performed using the OPLS force field. Both forcefields are part of the GROMACS package that is a linux-based package designed for biomolecular simulations. GROMACS can be run in parallel and is a user-friendly software which uses command line options with error massages when something is wrong. The lipids parameters were obtained from Berger and co-workers (311).

2.2.5 Timestep-bond constraints

The fastest motion in a biomolecular system is the bond vibration and therefore the timestep for a simulation system (which depends on the fastest motion) should accommodate bond vibration rates. In most biological systems, however, higher frequency motions such as bond vibrations have been shown to make little contribution to the changes occurring in the system. Therefore, algorithms like LINCS (312) and SHAKE (313) have been introduced which force bond lengths to return to a reference length after each timestep. This allows the simulation step to be increased, resulting in significant simulation speed up. The LINCS (312) algorithm is employed in the simulations discussed in this thesis to constrain the bond length and a timestep of 2 fs was used for all the AT-MD simulations (see Section 2.3 for more information on the CG-MD simulations timestep).

2.2.6 Periodic boundary conditions

MD simulations calculate the properties of a macroscopic system ($\sim 10^{23}$ particles) by using a finite number of atom particles ($\sim 10^{5\text{-}7}$) in a defined box. A problem which arises is how one treats the edges of the box. If the box edges are considered as solid walls then the particles will interact with the walls and experience additional forces which are irrelevant for bulk phenomena. On the other hand, if the edges of the box are treated as a vacuum, discontinuities in the system will again affect the calculation of macroscopic quantities. Therefore, in order to reduce the errors occurring from the finite size of the system, periodic boundary conditions are used. The system is treated as a "unit cell" of periodic repeats of identical cells. Thus, a particle in the central box has "imaginary" neighbours and during the simulation it interacts with its closest neighbours (real or imaginary). When a particle leaves the simulation box it is replaced by its image particle entering the box from the opposite side. Therefore, the number of particles is kept constant without the particles experiencing any additional forces from the edges of the box. For the calculation of short-range interactions such as Lennard-Jones interactions, a cut-off distance of $l/2$ is used where l is related to box dimensions. The PME method is used for the calculation of the long-range electrostatic interactions.

2.2.7 Ensembles

One of the main goals of MM simulations is to calculate the values of macroscopic (thermodynamics) quantities such as mean energy, mean volume, mean pressure etc. This

calculation is done using statistical mechanics, correlating the microscopic states of a system with its macroscopic properties. In a system with n particles, a 6n phase space can be identified which includes 3n position coordinates and 3n momentum coordinates. Each point in this space represents one microscopic state of the system. Therefore, by knowing all the positions and momenta at a specific time one can define the exact microscopic state of the system. The macroscopic state of the system is an ensemble of multiple microscopic states. Depending on the thermodynamic quantities which are conserved we can discriminate among different ensembles:

- Microcanonical ensemble (NVE): this describes isolated systems; the volume V, number of atoms N and energy E are fixed.
- Canonical ensemble (NVT): this describes a closed system in thermal balance; the temperature T, number of atoms N and volume V are fixed.
- Isobaric-isothermal ensemble (NPT): these systems have fixed temperature T, number of atoms N and pressure P.
- Grand canonical ensemble: this describes open systems where the number of atoms varies but the temperature T, chemical potential μ and volume V are fixed.

In classical MD simulations, NVE or NPT ensembles are usually used as these resemble the experimental conditions to some degree. The NPT ensemble is used in the simulations described in this thesis. The pressure and temperature are kept constant by using a barostat and thermostat respectively (see below).

2.2.8 Temperature and pressure coupling

To use an NPT ensemble in MD simulations the temperature and the pressure must be kept constant and thus, barostatic and thermostatic algorithms are used. In GROMACS three thermostats are available: the Berendsen (314), the V-rescale and the Nose-Hoover (315, 316). In this thesis, the Berendsen thermostat was used for temperature coupling. In this thermostat the system is coupled to an external bath which is kept at the desired temperature (Θ_0). Deviations in temperature during the simulation are corrected gradually using the following equation:

$$\frac{d\Theta}{dt} = \Theta_0 - \frac{\Theta}{c} \tag{2.11}$$

where c is a time constant. The velocities for each time step are scaled by a factor μ, which is given by:

$$\mu=\sqrt{1+\frac{n_{\Theta C}\Delta t}{c_{\Theta}}\left(\frac{\Theta_0}{\Theta(t-\frac{\Delta t}{2})}-1\right)} \tag{2.12}$$

where c_Θ is given by the equation:

$$c=\frac{2C_V c_\Theta}{N_{df}k} \tag{2.13}$$

C_V is the total heat capacity of the system, k is the Boltzmann constant and N_{df} is the number of degrees of freedom. The change in kinetic energy caused by velocity scaling is distributed between the kinetic and potential energy (302). A reference temperature of 310 K was used in chapters 3 and 5, whereas in chapters 4 and 6 a reference temperature of 323 K was used. A temperature coupling constant of 1.0 ps was used for the CG-MD simulations and of 0.1 ps for the AT-MD simulations.

For pressure coupling the Berendsen barostat was used in CG-MD simulations and the Parrinello-Rahman barostat (317) was used in AT-MD simulations. The Berendsen thermostat is used to rescale the coordinates and the box vector in every time step to a reference pressure P_0 using a scaling matrix ν. The change in pressure is given by:

$$\frac{dP}{dt}=P_0-\frac{P}{\tau_P} \tag{2.14}$$

and the scaling matrix is given by:

$$\nu_{ij}=\delta_{ij}-\frac{n_{PC}\Delta t}{3\tau_P}\beta_{ij}\left(P_{0ij}-P_{ij}(t)\right) \tag{2.15}$$

where β is the isothermal compressibility and τ_P is the coupling constant.

The Parrinello-Rahman barostat gives a more precise conservation of the NPT ensemble and was thus used in the AT-MD simulations. In this case, the box vectors are represented by a matrix (b) and the fluctuation in pressure is calculated using:

$$\frac{db^2}{dt^2}=VW^{-1}b'^{-1}\left(P-P_{ref}\right) \tag{2.16}$$

where V is the volume of the box, W determines the strength of the coupling and P and P_{ref} represent the current pressure and the reference pressure respectively. The W^{-1} parameter determines how the box can be deformed. With the Parrinello-Rahman barostat, the particle's equation of motion is also modified to:

$$\frac{d^2r_i}{dt^2}=\frac{F_i}{m_i}-M\frac{dr_i}{dt} \tag{2.17}$$

where M is a function of the box vector matrix b (302). In this thesis, a reference pressure of 1 bar and a coupling constant of 1 ps were used in both AT-MD and CG-MD

simulations. The compressibility was $4x10^{-6}$ bar^{-1} for the AT-MD simulations and $5x10^{-6}$ bar^{-1} for the CG-MD simulations.

2.2.9 Position restraints

In this thesis, prior to all MD simulations the systems were equilibrated with the protein Cα atoms restrained. This restricted any changes in the protein parts while the rest of the system was equilibrated. The position restraints are applied to an atom or to a group of atoms by an additional term in the potential energy function which fixes the particles to a reference position (r_{ref}). In GROMACS this term is described by Eq. 2.18. The force constant (k_{res}) defines the strength of the position restraints. For the simulations in this thesis, a force constant of 10 kJ mol^{-1} $Å^{-2}$ was used.

$$V_{res}(r_i)=\frac{k_{res}}{2}(r_i-r_{ref})^2 \qquad (2.18)$$

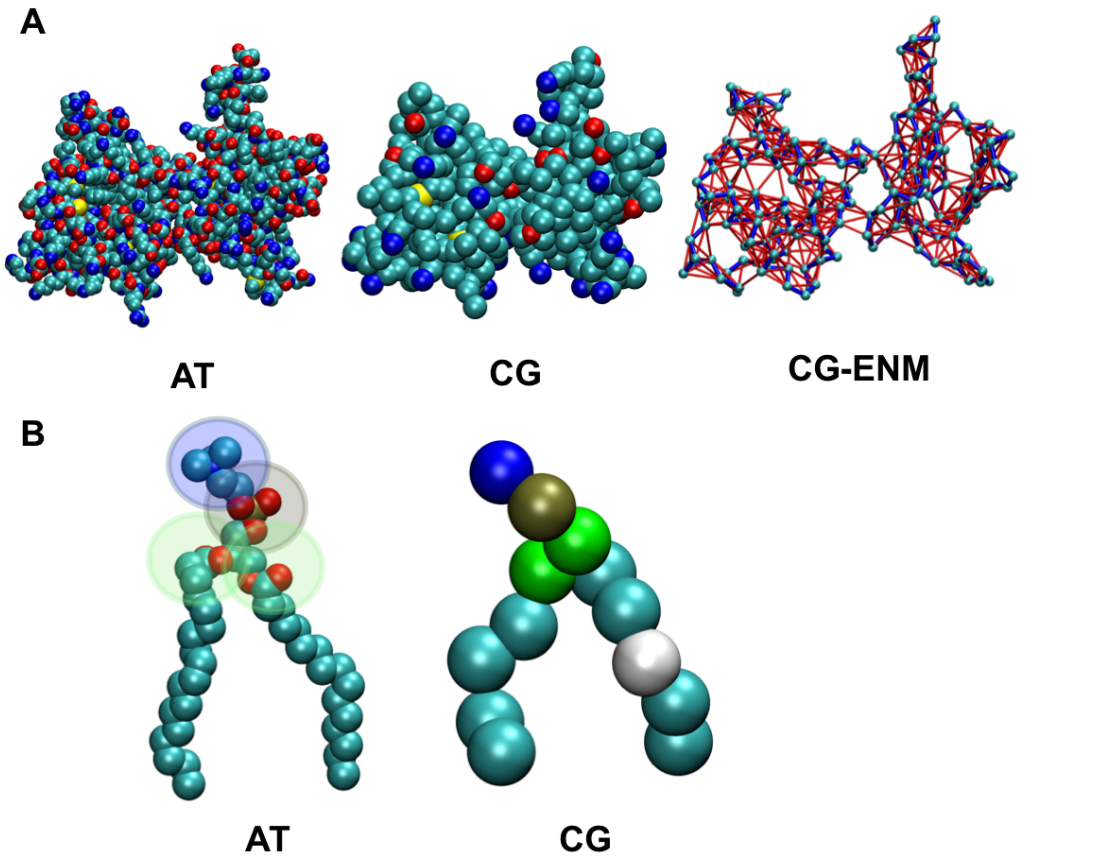


Figure 2.1: A. Atomistic (AT) representation, coarse-grained (CG) representation and the elastic network model (ENM) of the CG model illustrated for the talin F2-F3 domain pair (125). B. AT and CG representations of a POPC lipid. In the CG model a 4:1 mapping of heavy atoms is made. In the AT model, atoms are represented as: blue for nitrogen, cyan for carbon, yellow for sulphur and red for oxygen. In the CG model the CG particles are coloured as: green for glycerol, cyan for acyl chain, bronze for phosphate and blue for choline.

2.3 CG approach

All-atom simulations usually contain a large number of atoms, especially in the case of membrane proteins, and thus it is difficult to simulate large, complex and thus more biologically relevant systems. Therefore, alternative approaches such as CG-MD methodologies have been developed in which a reduced representation of the system is used to reduce the number of degrees of freedom. In CG models a group of atoms (usually consisting of two to five heavy atoms) is grouped together in a CG bit (318). The bonded and non-bonded terms of the potential equation are modified accordingly to reproduce all the thermodynamic properties. This is achieved by extensive calibration against thermodynamic data derived from experiments and/or AT simulations. The simplified representation smoothens the energy surface of the system and drastically reduces the interactions by an order of magnitude. Therefore, a time step of the order of 10 to 40 fs can be used. This gives two to three orders of magnitude speedup to the simulations, allowing access to much longer timescales (μs/ms) (319). Note that despite the speedup in the CG-MD simulations and the different timestep that is used, no normalization of the timescales was made in the CG simulations discussed in this thesis. CG approaches can reproduce properties of the systems such as area per lipid and phase transitions and thus they provide a valuable tool for studying a variety of biomolecular phenomena (see Section 1.3).

2.3.1 CG models

In this thesis, the MARTINI CG force field was used for the CG-MD simulations in Chapter 6 (320) and the Bond et al. CG force field, which is a variant of the MARTINI force field, was used for the CG-MD simulations in chapters 3, 4 and 5 (321, 322). The MARTINI force field has been released in two versions. In this chapter, the latest (v2) version will be discussed. In the MARTINI force field, four heavy atoms (e.g. carbon, oxygen) are modelled as one CG bit. Four CG particle types exist: charged (Q), non-polar (N), apolar (I) and polar (P). Polar (P) particles represent water-soluble neutral groups, apolar (C) represent hydrophobic atom groups, charged (Q) represent charged atom groups and non-polar (N) represent mixed (polar/apolar) group of atoms. Additional particle subtypes (0, d, α and dα for N and Q particles) exist for assigning H-bond capabilities. In particular, the 0 particle is used for groups with no H-bond capabilities, d is used for H-bond donors, α for H-bond acceptors and dα for H-bond donors and acceptors. Overall, MARTINI has 18 particle types, each with a particle mass of 72 amu and a VDW radius of 2.35 Å.

The protein secondary structure is imposed by restraining the proper dihedrals of the backbone particles according to the following equation:

$$F_{dih}=K_{dih}[1+\cos(n\psi_{ijkl}-\psi_{dih})] \quad (2.19)$$

For α-helices, a 60^{o} angle is defined as the equilibrium angle. A recent development of the MARTINI force field also allows the use of an elastic network to model the secondary and tertiary structure (323). The bonded interactions are described by a harmonic potential with an equilibrium value of 3.5 Å. The angle potential is also harmonic (cosine-based). Similar to AT force fields, a Lennard-Jones potential with different discrete levels ranging from attractive to repulsive interactions is used to describe the VDW interactions. The Lennard-Jones potential is shifted to zero between 9 to 12 Å. The interactions between the charged particles are modelled using the Coulomb electrostatic potential with a cut-off distance of 12 Å. The parameterization of the force field was made using the known partitioning of free energy of chemical compounds between the polar and apolar phases.

Table 2.1: CG representation of amino acids

Amino acid	CG representation
Ala	C
Arg	Nd+Qd
Asn	P
Asp	Qα
Cys	Nd
Glu	Qα
Gln	P
Gly	n/a
His	Nd+Ndα
Ile	C
Leu	C
Lys	C+Qd
Met	Nd
Phe	C+C
Pro	C
Ser	Ndα
Thr	Ndα
Trp	C+Nd
Tyr	C+P
Val	C

The Bond *et al.* CG force field has also been released in two versions. The second version, which was used for the simulations in this thesis, was modelled using solvation

free energies of neutral amino acids and transfer free energies of all amino acids from water to a bilayer environment (322). In this model, the same 4:1 grouping for non-hydrogen atoms as used in the MARTINI force field (above) is made and the same particle types are used. Each amino acid consists of a backbone CG particle plus zero to two extra particles to model the sidechains. The backbone particle is represented by a polar/apolar (N) particle and the choice of sidechain atoms depends on the partial charges and the H-bond capabilities of each amino acid. The backbone particle subtype for a certain peptide bond group is assigned based on the H-bonds within the backbone particle of the atomistic structure. For example, small hydrophobic residues (A, I, L, P, V) are modelled with a single apolar (C) particle and negatively charged amino acids (D, E) are represented by charged (Q) particles with H-bond capabilities (Qα). All the amino acid CG representations are shown in Table 2.1 (324).

The bond length is harmonically constrained with an equilibrium force constant of 50 kJ mol^{-1} $Å^{-2}$ and an equilibrium bond length (between backbone particles) of 3.6 Å. The bond angle potential for residues in a random coil is assigned a force constant of 250 kJ mol^{-1} rad^{-2}, while for residues in secondary structure elements it is assigned a force constant of 350 kJ mol^{-1} rad^{-2}. Based on high-resolution crystal structures, the equilibrium bond angle was set to 90^{o}, 130^{o} and 120^{o} for α-helical sections, β strand sections and random coil sections respectively. Secondary and tertiary structure modelling is carried out with two different types of constraint networks. In the first, harmonic distance restraints between i and i+4 backbone particles (in a helical region) are applied. The force constant is 10 kJ mol^{-1} $Å^{-2}$ and the distance equilibrium length is 6 Å. This mimics the H-bond network in isolated α-helical structures and was used in Chapter 4 of this thesis. In all the other chapters, an ENM (325) was used to model the secondary and tertiary structure. For the network, a harmonic potential was applied to all backbone particles with a cut-off distance of 7 Å. It should be noted that both types of constraint have been shown to reproduce the conformational flexibility of a variety of membrane proteins.

2.4 Multiscale approaches

Despite the simplifications, lower-resolution models (such as CG models) can give useful information about biomolecular systems. Nonetheless, coupling with a high-resolution AT model is still required to gain access to all the details of a specific system. Coupling between different resolution levels is a challenging task but it offers a valuable computational tool for studying protein dynamics in a timescale comparable to the

experimental data (326, 327). Ideally, coupling between different resolution levels should be part of the same "mixed" simulation system, where the exchange of information occurs directly with simultaneous calculation of multiple resolution components. An analogue is the hybrid QM/MM approach (328), where a small part of the system is represented by quantum mechanics and the rest by molecular mechanics. In the QM/MM approach, the two parts share the same level of AT detail and it is easier to use spatial decomposition theory to correlate the QM and MM descriptions of the system. This becomes much harder when using CG and AT representations simultaneously. Although there has been recent progress (329-333) the development of mixed multiscale force fields is still in progress.

In contrast, a serial multiscale approach has been extensively used in recent years to study many membrane proteins (161, 250, 274, 326, 327, 334-337). In the serial multiscale approach there is no mixing of the different resolution levels since the simulation for each level is performed separately, meaning no direct exchange of information between systems is required. This approach allows long timescales to be addressed by using CG-MD simulations (up to μs/ms) followed by shorter albeit more detailed AT-MD simulations, which allow the system to be refined.

During the conversion from CG to AT representations the key protein–lipid and protein–protein interactions must be retained to allow smooth continuation of the simulations. Therefore, a critical step in the serial multiscale approach is conversion between CG and AT representations. Rzepiela *et al.* (338) developed a method in which the dihedral restraints from the CG model are used to reconstruct the AT structure and a simulated annealing simulation is performed to optimize the AT structure. Stansfeld and Sansom (339) developed a fragment-based approach in which the lipids are reconstructed by aligning AT lipid fragments with the corresponding CG fragments from an energy minimized library and the protein is reconstructed by the backbone particle trace using PULCHRA (340). Aromatic sidechains in the protein are reconstructed by alignment of atomistic ring structure fragments with the CG aromatic chains. All other sidechains are reconstructed using the complete _pdb function in Modeller (341, 342). The latter methodology was used in this thesis for all the CG to AT conversions.

Chapter 3

Talin Head Domain Interactions with the Membrane

3.1 Introduction

As mentioned previously, integrin inside-out activation is mediated via the formation of a complex between the talin head domain, the membrane and the CT region of integrins. The talin head domain consists of four subdomains (F0 to F3), but it is the F3 subdomain which binds to the CT tail of integrins and triggers activation (30, 125) (Fig. 3.1). Over the last few years, an increasing amount of structural and biophysical data for the talin head domain has emerged (see Chapter 1) (52, 125, 186, 194) and this has significantly increased our understanding of the interactions of talin with the membrane and the integrin receptor. However, despite these data, many of the details of the mechanisms of talin activation are still relatively unknown. The head domain is also dynamic due to flexible regions within the protein and this induces further difficulties for structural studies. As a result, investigations have not yet been able to provide a complete picture of the talin/membrane interactions explaining in detail the role of the individual talin domains in binding and orientating the talin head domain relative to the membrane.

The talin head domain has homology with FERM domains but it differs significantly from other canonical FERM domains such as that found in focal adhesion kinase (FAK) (Fig. 3.1D) (182, 185). Canonical FERM domains have three subdomains (F1, F2 and F3) in a triagonal arrangement. In contrast, the talin head FERM domain has a linear arrangement of these three domains (Fig. 3.1B), with an extra F0 domain attached to the F1 domain as well as an additional long insertion (~30 residues) in the F1 domain. The F0-F1 and the F2-F3 domain pairs are connected with a long flexible linker. The F1 insertion is highly positively charged (60%) and is believed to be important in integrin activation, although its exact role is still unknown (186, 194). Recent NMR studies suggested that residue regions 144-152 and 154-167 of the F1 loop have a propensity to form helical structures that are stabilized by interactions with the lipids in the membrane (Fig. 3.1C) (194).

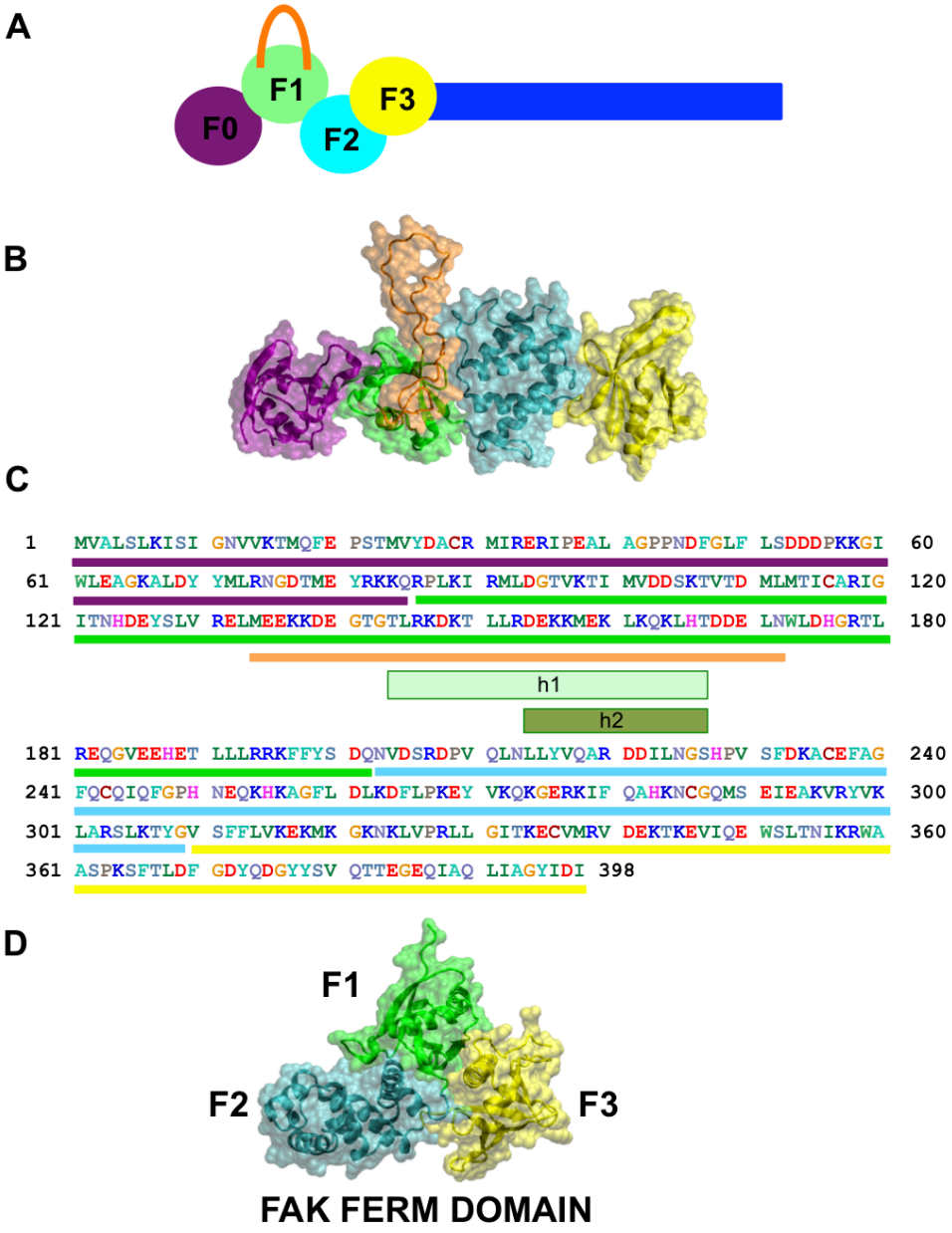


Figure 3.1: A. Schematic representation of talin. The talin rod domain is shown in blue, the F3 domain in yellow, the F2 domain in cyan, the F1 domain in green, the F1 insertion in orange and the F0 domain in purple. B. Structure of the complete talin head domain (PDB:3IVF (186)). The loop in the F1 domain was modelled using Modeller. C. Sequence of the talin head domain. The four domains are highlighted using the same colours as in A. The F1 loop region and the proposed helical regions in the loop are also shown (h1 and h2). D. Structure of the FAK FERM domains (PDB:2AL6 (185)).

Recent studies examining the interactions between the talin F2-F3 pair and/or the complete talin head domain (F0 to F3) with lipids have highlighted the crucial role of negatively charged lipids (i.e. PIP_2 and POPS) in talin-mediated integrin activation (125, 186). These studies showed that mutation of either the F2 positively charged patch or the F3 acidic loop significantly reduce integrin activation (125). In particular, mutation of four

positively charged residues (i.e. K258, K274, R276, and K280 in talin-2 F2 domain and K256, K272, K274, and R277 in talin-1 F2 domain) to glutamic acid (E) completely inhibits integrin activation. Furthermore, mutations in the F3 positively charged loop, for example the K324D mutation (125, 203), significantly altered integrin activation. These studies also suggested that the strength of the talin/lipid interactions depends on the net negative charge of the membrane and the type of negatively charged lipids involved in these interactions is not critical (186).

In this chapter, I describe the use of both NMR and a multiscale simulation approach (see Chapter 2) to study the interactions of the complete talin head domain with the membrane. In particular, the role of the different talin domains in integrin activation and the dynamic nature of the talin head domain that regulates its binding to the membrane are elucidated. In addition, by using negatively charged lipids in both NMR and simulations, the role of negatively charged moieties in the membrane on talin orientation and interactions with the membrane are explored. On the basis of these results, a mechanism for the binding and orientation of the talin head domain relative to the bilayer is proposed.

3.2 Methods

3.2.1 NMR spectroscopy experiments

3.2.1.1 Protein expression and purification for NMR

The talin wild type (WT) F2 and F2-F3 proteins and the mutated F2-4E protein (U-^{15}N-labeled) were expressed as GST-fusion proteins in M9 minimal media but with ^{15}N salts. To this end, cells were harvested by centrifugation at 7000 G for 10 minutes at 4 ^{o}C and re-suspended in phosphate-buffered saline. Subsequently, lysozyme (1 mg/mL), 1M $MgSO_4$ (10mL/mL) and 1 mg/mL DNAsel solution (10uL/mL) were added. The cells were placed on ice for 20 minutes and subsequently three freeze/thaw cycles were performed. Afterwards, 10% Triton was added and the sample was vortexed and left on ice for 5 minutes before centrifuging at 40000 G for 10 minutes at 4 ^{o}C. A glutathione-sepharose 4B (Roche Applied Science) column was set up for purification. The column was rinsed with PBS a few times and the supernatant was loaded onto the column. The column and the sample were mixed and transferred to a 50 mL tube. The protein was eluted in fractions of 15 mM glutathione in 50 mM Tris (~pH 5.8). Glutathione S-transferase fusion protease $3C^{pro}$ was added to cleave the purified fusion protein. The protein was further purified by gel filtration chromatography into NMR buffer (50 mM sodium phosphate, 100 mM NaCl,

1 mM dithiothreitol (DTT), pH 7.0). The F2-F3-4E domain was expressed as a His-tag fusion protein and purified as mentioned above, but with Talon resin (Takara Bio) in 50 mM sodium phosphate, 300 mM NaCl, 0.035% β-mercaptoethanol, pH 7.0 buffer. TEV protease was used to cleave the polyhistidine tags and gel filtration chromatography into pH 7.0 NMR buffer was used for further purification.

3.2.1.2 Liposomes

Two different types of liposomes were constructed. The first was comprised of DMPS/PC 1:4 mol % and the second was pure DMPS liposomes, where DMPS is L-α-phosphatidylcholine/1,2-dimyristoyl-*sn*-glycero-3-phospho-L-serine and PC is L-α-phosphatidylcholine (egg, chicken) with predominant species 16:0/18:1. *In vivo,* the anionic PS lipid concentration in the inner leaflet is 20 mol % (343). To construct the DMPS/PC 1:4 mol % liposomes, DMPS and PC lipids (in chloroform) were mixed. A stream of argon gas was used to remove the chloroform and subsequently the sample was attached to a vacuum pump overnight. The sample was agitated after the addition of 1 mL of 50mM phosphate, 100 mM NaCl and 1 mM DTT to the liposomes. Subsequently, the sample was rotated for 1 hour at 40 ^{o}C on a rotary evaporator with no vacuum applied. The large multilamellar vesicles (LMVs) were disrupted by five freeze-thaw cycles using liquid nitrogen. An Avanti mini-extruder kit was used to size the liposomes by following the protocol specified on the Avanti website (http://avantilipids.com/) and by using a 1000 Å polycarbonate membrane. During the extrusion state, the sample was maintained at a temperature of 40 ^{o}C (i.e. above the phase transition temperature of the lipids) and it was extruded back and forth (i.e. through the membrane) 20 times. The final liposome diameter based on the polycarbonate membrane was expected to be ~1000–1200 Å. Liposomes were stored for a maximum of 3–4 days at 4 °C in order to avoid any significant deformation of the vesicles. The DMPS liposomes were constructed using the same protocol as described above. As mentioned above the anionic PS lipid concentration in the inner leaflet is 20 mol % (343). Therefore the pure DMPS liposomes may not be physiological however they will allow the study of the talin/lipid interactions in an environment with increased negative charge.

3.2.1.3 NMR spectroscopy

NMR experiments were performed on 600 MHz NMR spectrometers with Oxford Instruments superconducting magnets and GE Omega consoles. The temperature was kept

at 40 °C. Note that this temperature was above the transition temperature of the lipids used in the NMR experiments. Additionally, it was confirmed that the folding of the protein was retained in temperatures above 35 ^{o}C (see Section 3.3.1.1 for more information). For all the experiments, a protein sample of 75, 100 and 150 μM of U-^{15}N-labelled protein in 5% D_2O / 95% H_2O solution was used. The buffer used for the NMR experiments was 50mM phosphate, 100 mM NaCl and 5mM DTT to prevent oxidation. In the samples with liposomes, 10 mM lipids were added. All samples were set to pH 6.5. For each sample, a two-dimensional heteronuclear single quantum coherence (HSQC) spectrum was acquired. The F2 resonance assignment was made using three dimensional (3D) gradient-enhanced [^{1}H-^{15}N]-NOESY-HSQC (τ_m= 100 ms) and [^{1}H-^{15}N]-TOCSY-HSQC (τ_m= 70 ms) spectra. The equation $\Delta\delta[HN,N] = ([\Delta\delta_{HN}W_{HN}]^2+[\Delta\delta_N W_N]^2)^{1/2}$ was used to calculate the weighted secondary shifts ($\Delta\delta=\delta_{bound}-\delta_{free}$) between bound and unbound states. W_{HN} and W_N are weighting factors for the HN and N shifts, respectively (W_{HN} = 1, W_N = 0.154) (344) and $\Delta\delta=\delta_{bound}-\delta_{free}$. The NMR assignments of the F2-F3 domain pair and the F2 domain used in this study have been submitted to the BiomagresBank: F2 (196-309), BMRB ID=16930 F2F3 (196-405), BMRB ID=169n32. NMR experiments were processed using NMRPipe (345) and visualized using SPARKY (www.cgl.ucsf.edu/home/sparky). The ratio R_{SN} was used to characterize the strength of the protein/lipid interactions and it was calculated using the following equation:

$$R_{SN} = \frac{S_{protein\ with\ liposome}}{S_{protein\ without\ liposome}} \quad (3.1)$$

In the Eq. 3.1, S is the signal-to-noise ratio for each spectrum and its calculation was made using NMRPipe and Sparky (www.cgl.ucsf.edu/home/sparky). For this calculation, first the R_{SN} values for individual peaks were calculated by dividing the signal-to-noise ratio of the peak concerned, with and without liposomes. Subsequently, the overall R_{SN} value for a particular condition was calculated as the mean of the individual peak values ± the standard deviation. For the R_{SN} calculations, 106 peaks for F2 and 201 peaks for F2-F3 were used.

3.2.2 Structures used in the simulations

A summary of all simulations is presented in tables 3.1, 3.2 and 3.3. The simulations which included the F2-F3 and the F2 talin subdomains were performed using protein coordinates from the crystal structure of the F2-F3/β1D complex (PDB:3G9W) (125). From this structure, residues 205 to 408 (i.e. F2-F3) were used for the simulations

with the isolated F2-F3 and residues 205 to 310 (i.e. F2 domain) were obtained for the simulations with the isolated F2 domain. Modeller (341, 342) was used to perform the mutations of residues K258, K274, R276, and K280 to glutamic acid (E), starting with coordinates PDB:3G9W. The simulations which included the intact talin head domain (F0 to F3) were initiated using the crystal structure PDB:3IVF (186). This structure contained the four subdomains of talin (F0-res: 1-85, F1-res: 86-202, F2-res: 203-309 and F3-res: 310-405), but it lacked a large flexible insertion in the F1 domain (res: 134-172). This loop was modelled in a random conformation using Modeller (341, 342). Since NMR studies suggested a possible helical conformation for residues 154-167 or residues 144-167 (194), these two regions were also modelled as helix (Fig. 3.1C). Subsequently, the loop in the F1 domain (with or without helical regions) was modelled in various conformations relative to the talin head domain (as shown in Fig. 3.2A). After the construction of all the models, the structure of the isolated F0-F1 domain pair was extracted (res: 1-202) and used in the simulations to investigate the F0-F1/lipid interactions.

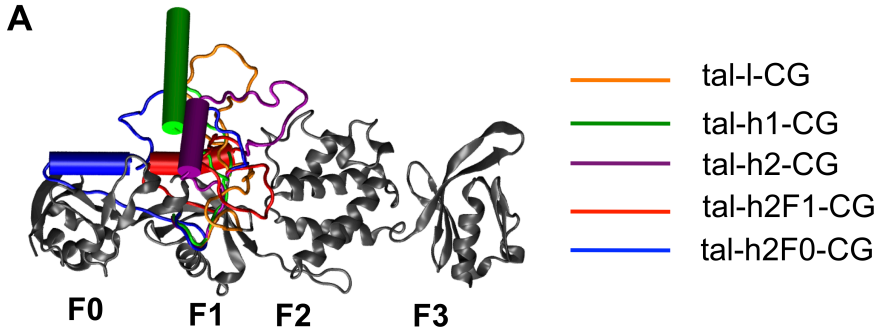


Figure 3.2: A. Positions of the F1 loop (with or without helical conformations) relative to the rest of the talin head domain at the beginning of the CG-MD simulations. See Table 3.2 and Table A.2 in Appendix A for details.

3.2.3 Bilayer set-up

The bilayers used in all the simulations (see tables 3.1, 3.2 and 3.3) were formed by performing self-assembly CG-MD simulations (249). In these simulations the lipids were randomly placed in a simulation box and solvated with CG water molecules and ions to neutralize the system. Subsequently, a production simulation was performed for 200 ns. After the first 10-15 ns of simulation the bilayer was formed with equal distribution of lipids in the two leaflets (Fig. 3.3A,B). For the simulation systems discussed in this chapter, three different bilayers were constructed. The first bilayer was comprised of 832 zwitterionic POPC lipids, the second contained 512 POPC and 320 POPG lipids (ratio of 3:2) and the third contained 154 POPC and 102 POPG lipids (ratio 3:2). The latter bilayer

was used in the simulations with the F2-F3 or the F0-F1 domains and the other two for the simulations with the complete talin head domain. In the simulations, a different type of negatively charged lipids was used (POPG instead of POPS) however this is not expected to alter the talin/lipid association since binding studies showed that talin/lipid electrostatic interactions are regulated by the net negative charge of the membrane and not the lipid type (i.e. POPG, POPS, PIP_2) involved (186).

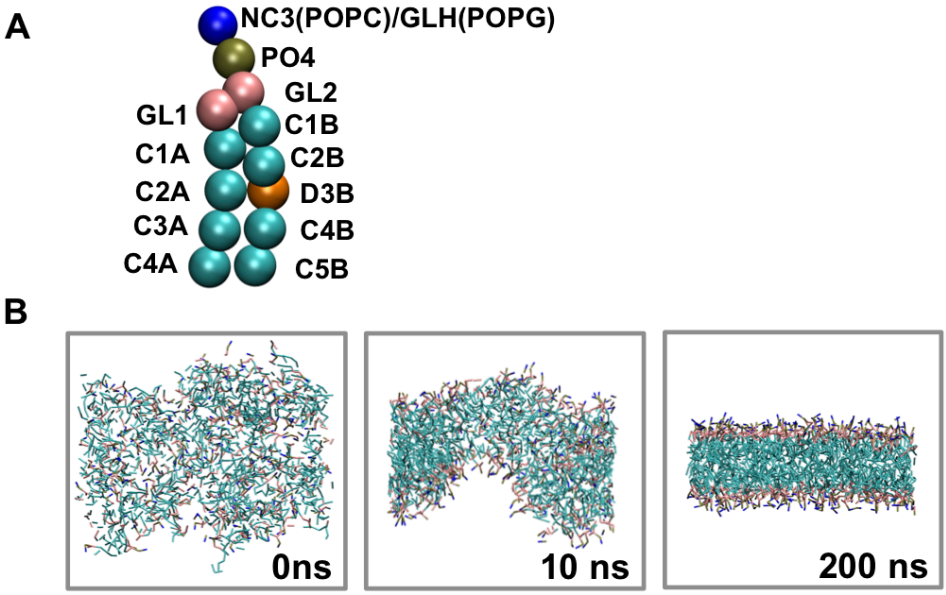


Figure 3.3: A. CG model of the POPC and POPG lipids used in the CG-MD simulations. B. Progress of the bilayer self-assembly CG-MD simulations shown for 0 ns, 10 ns and 200 ns. The lipids are shown in bonds formats with the hydrophobic tails in cyan, glycerol in pink, phosphate in brown and choline in blue. Waters and counter ions are omitted for clarity.

3.2.4 Simulation set-up

After the conversion to CG representation (discussed in Chapter 2; Section 2.3), the protein was displaced 110 Å (systems with F2-F3 and F0-F1 domains), 80 Å (systems with F2 domain) or 130 Å (systems with the talin head domain) away from the centre of mass of the preformed bilayer (see tables 3.1 and 3.2; Figs. 3.6A and 3.10A for more information). The protein/bilayer distance varied because of the different size of the proteins. In all cases, this distance was much larger than the cut-off distance used for the electrostatic and the VDW interactions in the force field and therefore the initial position of the protein did not induce any initial enforcement of the protein. The system was subsequently solvated with CG water molecules and neutralized with CG sodium particles. Then, the system was energy minimized for 250 steps and equilibrated for 5 ns with the protein Cα atoms restricted (force constant 10 kJ/mol/$Å^2$). Finally, production CG-MD simulations were

performed, as described in tables 3.1 and 3.2. The final snapshot of the CG-MD simulations was converted to an AT representation for refinement (see Table 3.3). The conversion to AT representation was made using a fragment-based approach (339) (discussed in Chapter 2; Section 2.4).

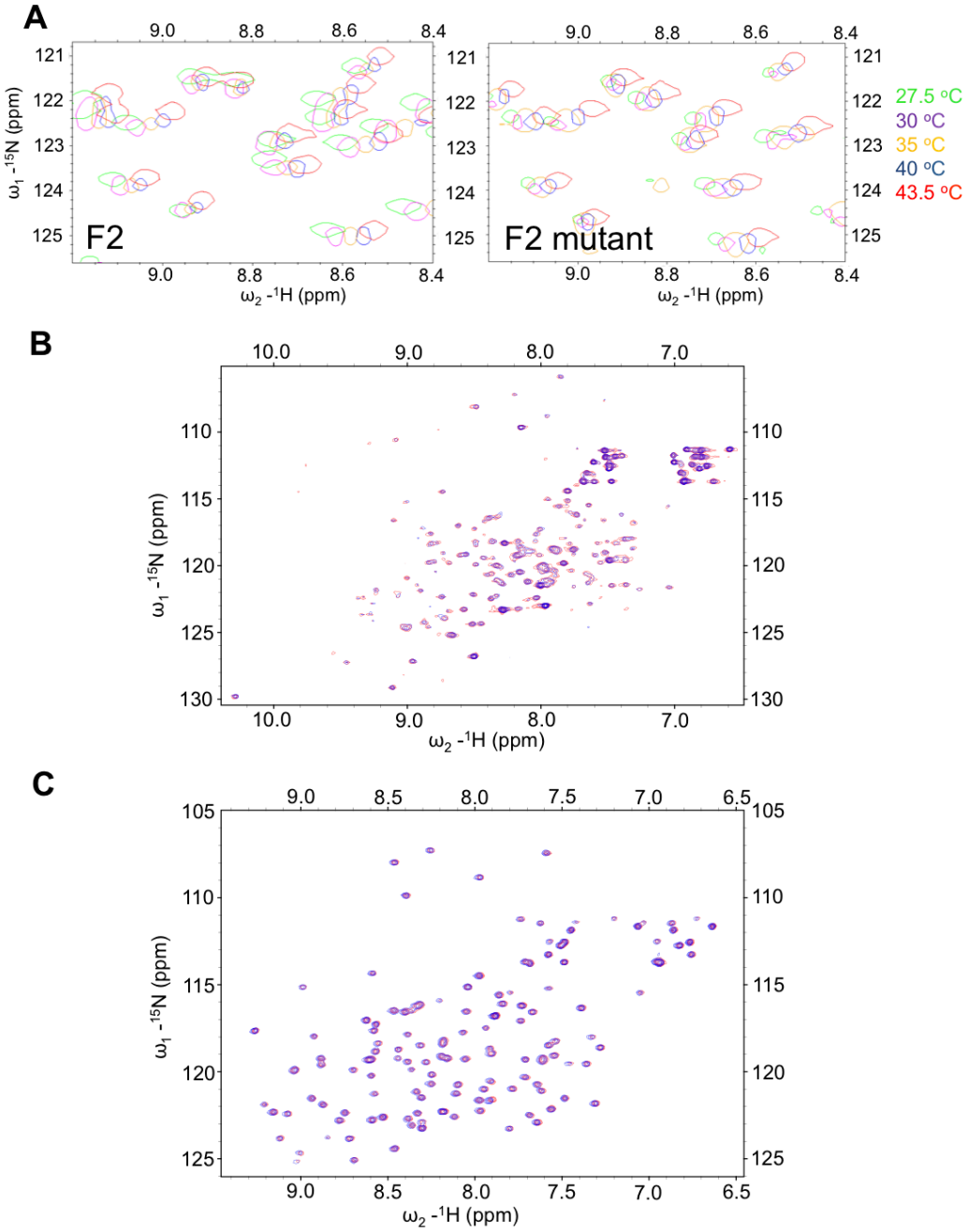


Figure 3.4: A. Temperature titration experiments using the wild type and the mutant form of the talin F2 domain. A region of the ^{1}H-^{15}N-HSQC NMR spectra is shown with the peaks observed at different temperatures (green for 27.5 ^{o}C, purple for 30 ^{o}C, yellow for 35 ^{o}C, blue for 40 ^{o}C and red for 43.5 ^{o}C). B,C. Overlaid ^{1}H-^{15}N-HSQC NMR spectra of the F2-F3-4E (B) and the F2-4E (C) domains in the absence (blue) and presence (red) of 20% DMPS/PC liposomes at a protein concentration of 150 µM. The figure demonstrates that the shifts induced to the peaks by the addition of liposomes were very small.

3.3 Results

3.3.1 NMR studies of talin/bilayer interactions

3.3.1.1 Interactions of the F2 and the F2-F3 subdomains with liposomes

As a first step towards studying the effect of negatively charged lipids in the binding of the talin F2-F3 domain pair to the membrane, NMR experiments were set up where the F2-F3 domain pair in the wild type (F2-F3-WT) and in a mutated form (K258E, K274E, R276E and K280E; F2-F3-4E) in three different concentrations (75 μM, 100μM and 150 μM) were added to 20% DMPS/PC liposomes. For each protein concentration, the chemical shift perturbations were monitored in two ^{1}H-^{15}N-HSQC NMR spectra; a spectrum with the protein alone in solution and a spectrum with the protein in the presence of the liposomes (Fig. 3.4B,C). It must be noted that the transition temperature of the DMPS lipids is 35 ^{o}C and thus, before the NMR experiments, temperature titration experiments were performed using the wild type and the mutated F2 domain to confirm that the folding of the protein is not altered in temperatures above 35 ^{o}C. Five NMR experiments were performed for each sample at 27.5 ^{o}C, 30 ^{o}C, 35 ^{o}C, 40 ^{o}C and 43.5 ^{o}C respectively and a ^{1}H-^{15}N-HSQC spectrum was obtained for each temperature. Overlapping of the spectra revealed that both structures retained their folding at all temperatures (Fig 3.4A). Therefore, all the subsequent experiments with the protein and liposomes were performed at 40 ^{o}C. Comparison of the induced chemical shifts upon the addition of liposomes revealed very small perturbation and therefore intensity changes were used to detect binding (Fig. 3.4B,C and Fig. 3.5C). Thus, a parameter R_{SN} which was a function of the signal-to-noise ratio for a given spectrum was calculated to characterize the strength of each interaction (see methods for more details). Accordingly, when the protein bound more tightly to the liposomes, the molecular weight of the complex and thus the correlation time greatly increased, resulting in the broadening of the spectral peaks and the reduction of the R_{SN} parameter. Thus, if the ratio R_{SN} was 1, then all the protein in the sample was unbound and if it was 0, then all the protein was assumed to be bound to liposomes.

For the experiments where F2-F3-WT was added to 20% DMPS/PC liposomes, the ratio R_{SN} was 0. The ratio R_{SN} increased ($R_{SN}\approx0.8$) upon the addition of the F2-F3-4E to the same liposomes. The mutations on the surface of the talin F2 domain weakened binding to liposomes and therefore some protein remained unbound and was detected by the ^{1}H-^{15}N-HSQC NMR spectra. In similar systems where the F2-F3-WT and the F2-F3-4E were

added to pure DMPS liposomes, no spectra could be obtained (R_{SN}≈0). This suggests that the talin F2-F3 domain pair bound tightly and that the liposomes were too "negative" for the mutations to disrupt the electrostatic interactions. This indicates that enrichment of the F2-F3 binding surface with negatively charged lipids increased the F2-F3 affinity to the membrane.

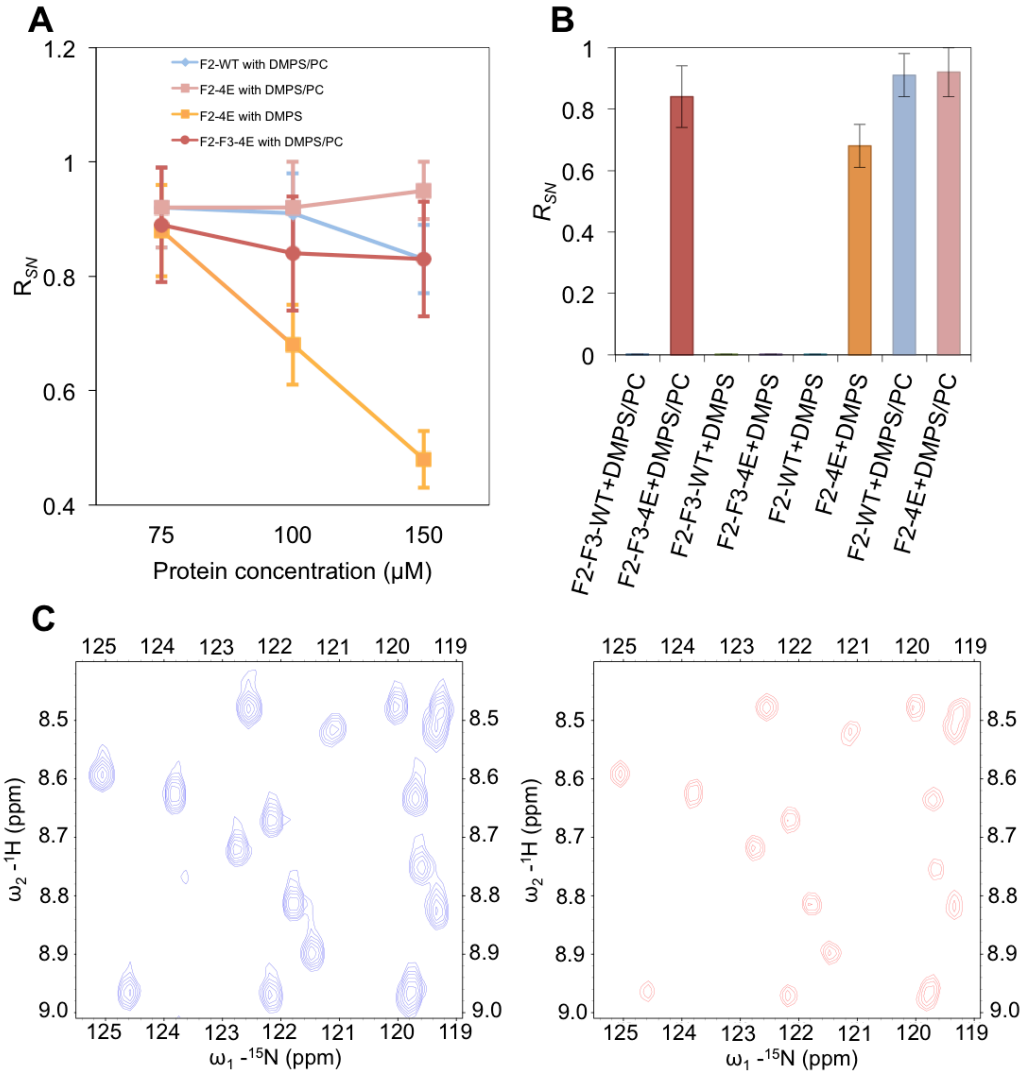


Figure 3.5: A,B. NMR protein/liposome titration experiments. A. The ratio R_{SN} is shown as a function of protein concentration for the different liposomes where an NMR spectrum could be obtained. B. The R_{SN} value is shown for all systems at a protein concentration of 100 μM. The same colour code as in A is used. C. ^{1}H-^{15}N-HSQC NMR spectra of the F2-4E (150 μM) in the absence (blue) and presence (red) of DMPS liposomes showing the reduction of intensities upon the addition of liposomes.

Similar experiments with the wild type (F2-WT) and a mutated form of the isolated F2 (F2-4E) with 20% DMPS/PC or pure DMPS liposomes were also performed. Equally, the weighted chemical shifts induced in the ^{1}H-^{15}N-HSQC NMR spectra were

very small upon the addition of the liposomes (Fig. 3.4C) and, therefore, the same R_{SN} parameter (see methods) was used to characterize binding. When F2-WT was added to 20% DMPS/PC liposomes the R_{SN} value was very high (0.85), suggesting a very weak interaction with liposomes in the absence of the F3 domain. The presence of the 4E mutation on the F2 surface weakened this interaction even further (R_{SN}≈0.92). In the presence of pure DMPS liposomes, no spectrum could be obtained for the F2-WT domain at any protein concentration. This suggests a very strong interaction of the F2-WT with the negatively charged lipids. The strength of this interaction was significantly reduced when F2-4E was added to the DMPS liposomes in good agreement with a weakened F2/lipid interaction in the presence of the 4E mutation (Fig. 3.5). It must be noted, however, that the R_{SN} value of the F2-4E with DMPS liposomes (0.68) was still lower compared to the R_{SN} value for the F2-4E (0.92) and the F2-F3-4E (0.84) with 20% DMPS/PC liposomes. Therefore, the interaction of the F2 with DMPS liposomes can be classified as "strong". Strong interactions of the F2 with negatively charged lipids were also identified by binding studies using cosedimentation assays (125). This strong interaction is expected to facilitate productive association of talin with the membrane. Using similar criteria, the ratio R_{SN} can provide a measure of the strength of these interactions. The ratio R_{SN} for the interactions of the F2-WT with 20% DMPS/PC liposomes was larger compared to all other systems, even those which contained mutated F2 and F2-F3 domains, and therefore this interaction can be characterized as "weak". Similarly, the ratio R_{SN} suggests that the interaction of F2-F3 with the negatively charged DMPS lipids can be characterized as "very strong", since all the protein in the sample in both forms, wild type and mutated, was bound in the DMPS liposomes (R_{SN}=0). Finally, the interaction of the F2-F3 with 20% DMPS/PC liposomes can be classified as "intermediate".

Table 3.1: Summary of simulations

Simulation	Protein	Duration (ns)
F2-CG	F2-WT	5x500
F2-4E-CG	F2-4E	5x500
F2-F3-CG	F2-F3-WT	5x1000
F2-F3-4E-CG	F2-F3-4E	5x1000
F2-q0-CG	F2-WT (charge of K258, K274, R276 and K280 was set to zero)	5x500
F2-F3-q0-CG	F2-F3-WT (charge of K258, K274, R276 and K280 was set to zero)	5x1000
F2-AT	F2-WT	1x30
F2-F3-AT	F2-F3-WT	1x30

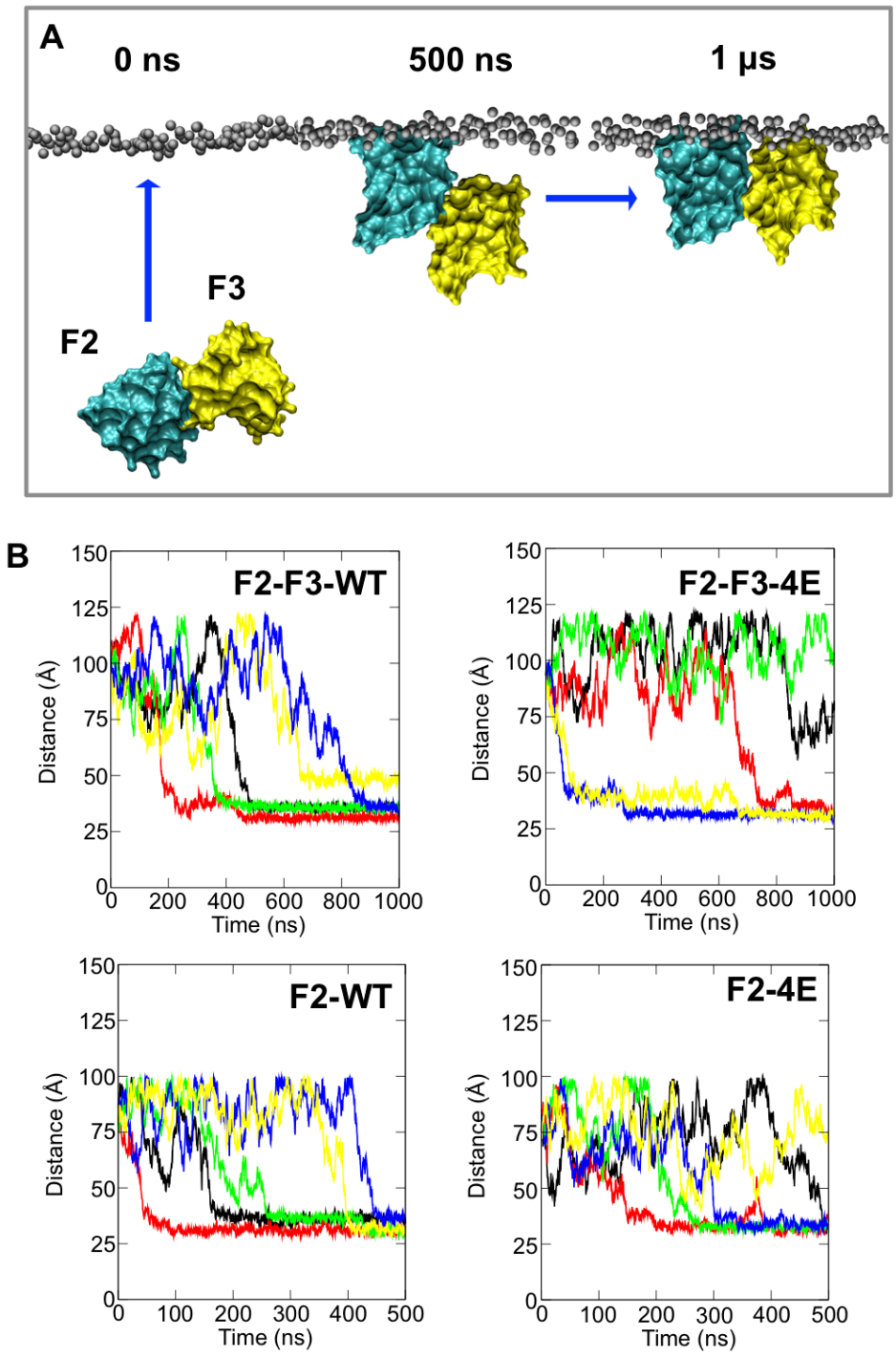


Figure 3.6: A. Progress of the *F2-F3-CG* simulation. Snapshots of the simulation are shown at 0 ns, 500 ns and 1 µs. The F2 is shown in cyan, the F3 in yellow and the bilayer phosphate atoms in grey. B. Progress of the simulations with the F2-F3-WT, the F2-F3-4E, the F2-WT and the F2-4E are shown as the distance of the protein centre of mass to the centre of mass of the bilayer as a function of time. The different coloured lines indicate the five different simulations performed for each system.

3.3.2 MD simulations

Having shown the crucial role of negatively charged lipids in the binding and

interactions of the talin F2-F3 domain pair with the membrane, a multiscale simulation approach was employed to study these interactions in more detail and to construct a molecular model which describes the talin/membrane association. All the MD simulations were performed as described in the methods section. A summary of the simulations is presented in tables 3.1, 3.2 and 3.3.

3.3.2.1 F2 and F2-F3 interactions with the membrane

In the first set of simulations, the wild type F2-F3 domain pair (F2-F3-WT) was displaced ~110 Å away from a preformed POPC/POPG bilayer and five simulations were performed for 1 μs. In all simulations, the F2-F3-WT initially diffused in the aqueous environment before anchoring to the bilayer at the lipid/water interface (Fig. 3.6A). Calculation of the separation between the centres of mass of the protein and the bilayer revealed that the F2-F3-WT bound to the bilayer in all five simulations (Fig. 3.6B). Interestingly, in three out of five simulations the F2-F3 domain pair bound to the bilayer with the same binding surface on the protein (Fig. 3.7A). Binding is assumed to be significant if the majority of the repeat simulations in a system resulted in the same binding mode. This binding surface included the positive charged patch on the F2 domain (res: 255-285) and the F3 acidic loop (res: 318-330). In particular, calculation of the contacts between the F2-F3 domains and the bilayer (contact was defined using a cut-off distance of 8 Å) revealed that the F2 residues 262, 263, 266, 275, 277 and 281 and the F3 residues 312-315 made the largest number of contacts with the lipids. The lysine and arginine residues at the interface, those mutated in the NMR experiments, also made a large number of contacts with the lipids. Calculation of the number of contacts of the F2-F3 with the POPC and POPG lipids separately revealed a larger number of contacts between the F2-F3 and the negatively charged POPG lipids (i.e. threefold) in good agreement with the NMR studies described above and other experimental studies (125, 186).

These five simulations were repeated with the F2-F3-4E variant. In this set of simulations, only three out of five simulations yielded a F2-F3/bilayer complex compared to five out of five in the *F2-F3-CG* simulation. Interestingly, in the simulations where F2-F3-4E yielded a complex with the bilayer, the pattern of interacting residues was different from the preferred binding surface of the F2-F3 domain pair in the *F2-F3-CG* system (Fig. 3.7A). Therefore, mutations on the F2 positively charged patch not only reduced the apparent affinity of F2-F3 pair to the bilayer but also perturbed its orientation relative to

the bilayer. This perturbed orientation is in turn expected to alter the binding of the F2-F3 to the integrin β CT tail.

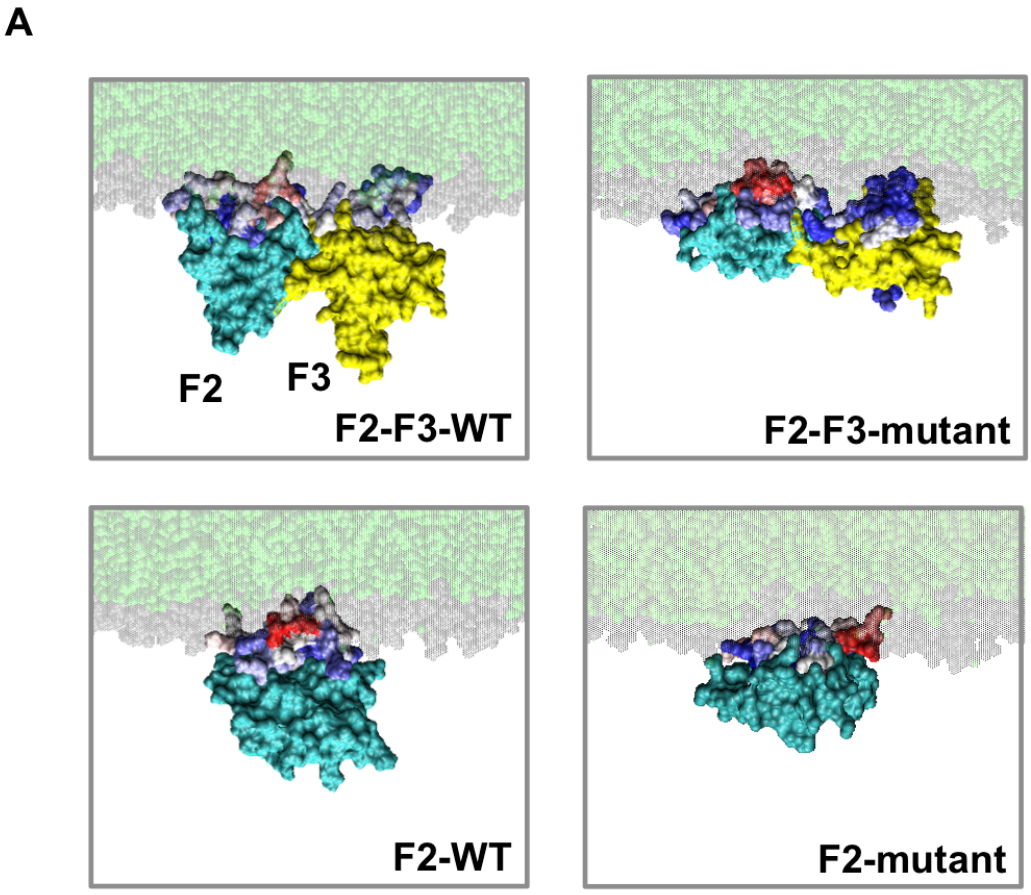


Figure 3.7: A. Contacts of the F2-F3 and the F2 domains during the CG-MD simulations. The residues which formed the largest number of contacts are mapped onto the preferred orientation. Red means a high number, white a medium number and blue a low number of contacts. The F2 domain is shown in cyan and the F3 in yellow. The normalization of contacts was made separately for each simulation system.

Similar simulations to explore the interaction of the isolated F2-WT and F2-4E domains with the same POPC/POPG bilayer were also set up (see Table 3.1). In this set of simulations the F2-WT bound to the bilayer in the lipid/water interface in all five simulations performed. In three out of five simulations which yielded a complex, the preferred binding surface was consistent with the published F2-F3/β1D complex structure (125) and the simulations with the F2-F3-WT domain (see above; Fig 3.7). Therefore, the F2 domain is capable of binding to the anionic lipids in a productive manner without the presence of the β tail and the F3 domain. In contrast, the simulations with the isolated F2-4E domain resulted in a non-productive F2-4E/bilayer complex in all four simulations in which the F2-4E was associated with the bilayer (Fig. 3.7).

To further evaluate the model for the F2-F3 pair binding to the bilayer, control simulations were run with the wild type F2 and F2-F3 domains where the charge of the residues K258, K274, R276 and K280 (those mutated above) was set to zero. Both proteins bound to the bilayer with an orientation that was different from the preferred binding mode of the F2-WT and F2-F3-WT in the above simulations. Closer examination of the interface between the F2 and the F2-F3 domains with the bilayer revealed a pattern of interacting residues similar to those in the simulations with the mutated form of the F2 and the F2-F3 domains. Therefore, we can conclude that the K/R motif on the F2 subdomain is required for productive binding of F2-F3 to the membrane.

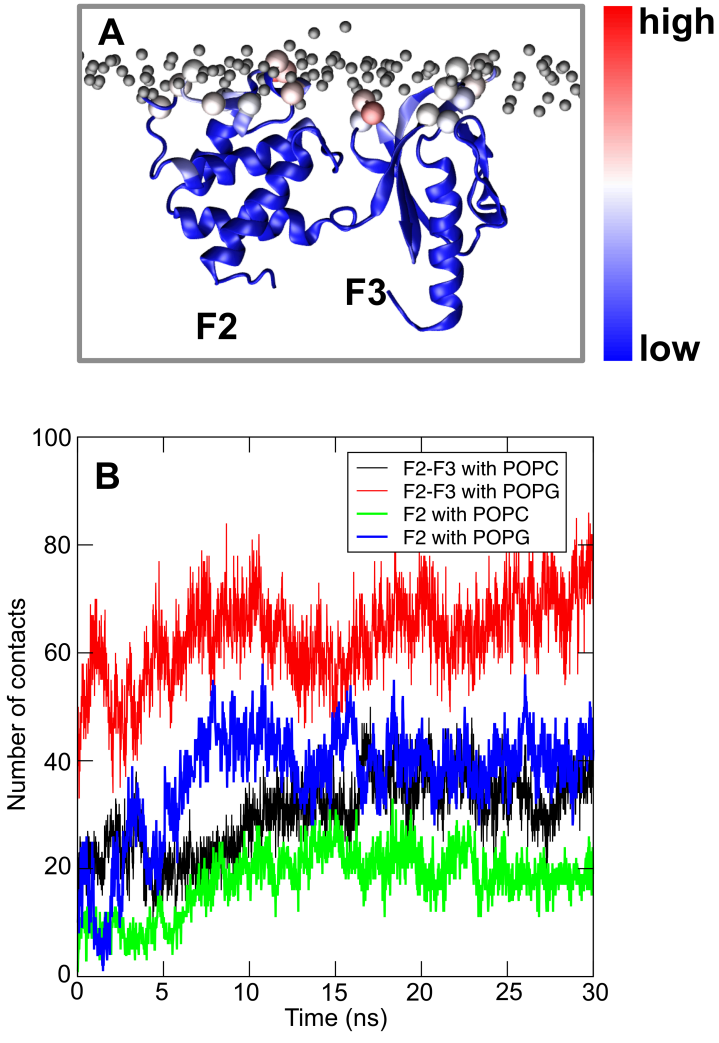


Figure 3.8: A. Number of contacts in the *F2-F3-AT* simulation mapped onto the last frame of the simulation. The residues which interacted with the lipids are colour coded: blue for infrequent/no interactions through to red for frequent interactions. The backbone particles of residues which formed the largest number of contacts are shown as spheres. B. Number of contacts of the F2-F3 and the F2 domains with the POPC and POPG lipids shown as a function of simulation time.

In order to refine the model for the F2 and F2-F3 interactions with the lipids, a CG system with the talin F2-WT or F2-F3-WT in the preferred orientation (i.e. with a similar binding surface to the one proposed by Anthis *et al.* (125)) was converted to an AT simulation (see methods). In both simulation systems, the protein maintained its orientation relative to the bilayer for the whole period of the simulation, without major conformational changes (root mean squared deviation (RMSD) relative to the initial configuration: 2 Å for F2 and 2.3 Å for F2-F3). The residue-by-residue fluctuation (RMSF) showed that, as anticipated, the core secondary structure elements have a lower RMSF than the loops. Comparison of the F2-F3 configurations in the last 10 ns of the simulation with the crystal structure revealed a small relative movement of 6±2^{o} between F2 and F3, suggesting a small reorientation of the F2 relative to the F3 domain in a plane horizontal to the bilayer xy plane during the simulation. This movement was possible due to the flexible linkage between the two subdomains and presumably optimizes contacts with the bilayer.

AT-MD simulations were also used to evaluate the F2 and F2-F3/bilayer interactions in greater detail. The talin/lipid contacts (as defined by a cut-off distance of 3.5 Å) formed in the last 5 ns of the simulation were used to analyse the F2-F3/bilayer surface (Fig. 3.8A). This analysis suggested that the binding surface found in the previous CG-MD simulations was retained. H-bonds formed between F2 or F2-F3 sidechains and the lipid headgroups suggested that residues E254, K258, K270, R276 and K280 on the binding surface of the F2 and residues K319, K321, K327, R331 and K346 on the binding surface of F3 made significant H-bonds with the bilayer. This correlates well with the residues identified in the CG-MD simulations (273). Interestingly, the same residues were also identified as making a large number of contacts and H-bonds with the isolated F2 domain in the simulation. Additionally, calculation of the number of contacts of the talin domains with the POPC and POPG lipids as a function of time revealed that the ratio of PC:PG contacts is 1:2. Taking into consideration that the overall ratio in the bilayer is PC:PG 3:2, this suggests that there is a threefold preference of the F2-F3 or F2 to interact with negatively charged POPG lipids (Fig. 3.8B). This is in good agreement with the NMR results and again demonstrates that negatively charged groups in the membrane play a crucial role in these interactions (193).

3.3.2.2 F0-F1 interactions with the membrane

The next step, after investigating the interactions of the isolated F2-F3 pair with the lipids, was to study the interactions of the F0-F1 pair with the same bilayer. In this set of

simulations, the F0-F1 pair without the F1 loop (Δloop-res: 134-172) or with the large loop in F1 with different preformed helical regions (h1 and h2) was displaced 110 Å away from the same preformed POPC/POPG (i.e. mixed zwitterionic/anionic) bilayer. Five simulations were performed for each system. Inspection of the binding interface of the F0-F1 pair relative to the bilayer revealed non-specific and non-productive binding in all systems (see Appendix A.1 for a more detailed description of the simulations). Therefore, it appears that the presence of the F1 acidic loop on the F0-F1 pair (with or without preformed helical regions) is not sufficient to result in consistent binding of the F0-F1 to the bilayer. This is in contrast to the studies described above for the interactions of the isolated F2-F3 pair with the bilayer (273).

Table 3.2: Summary of simulations

Simulation	Proteins	Duration (ns)
tal-CG	F0-F1-F2-F3 (Δloop-res: 134-172)	5x1000
tal-l-CG	F0-F1-F2-F3 (loop-res: 134-172)	5x1000
tal-h2-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167)	5x1000
tal-h1-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 144-167)	5x1000
tal-h1q0-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 144-167 [helix charge set to zero])	5x1000
tal-h2F0-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167 [helix restricted close to F0])	5x1500
tal-h2F1-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167 [helix restricted close to F1])	5x1000
tal-lF0F1-CG	F0-F1-F2-F3 (loop-res: 134-172 [loop restricted close to F0-F1])	5x1500
tal-l25-CG	F0-F1-F2-F3 (loop-res: 134-172 [harmonic constrained 25 kJ/mol/ $Å^2$])	5x1000
tal-l50-CG	F0-F1-F2-F3 (loop-res: 134-172 [harmonic constrained 50 kJ/mol/ $Å^2$])	5x1000
tal-sol-AT	F0-F1-F2-F3 (loop-res: 134-172/[solution])	2x100

3.3.2.3 Complete talin head domain (F0 to F3) interactions with the membrane

Having studied the interactions of the isolated F2-F3 and the isolated F0-F1 pairs with the bilayer (above), the studies were extended to the interaction of the complete talin head domain (PDB:3IVF) with a similar POPC/POPG bilayer (see Table 3.2). Before studying the interactions of the talin head domain with the lipids, two AT-MD simulations

with the talin head domain in solution (i.e. in the absence of a bilayer) were run for 100 ns (*tal-sol-AT* in Table 3.2). During the simulations, the flexible linker between the F2-F3 and F0-F1 pairs allowed transient displacement of the F0-F1 domain pair relative to the F2-F3 domain pair with the angle between the domains ranging from 10° to 75° (Fig 3.9A). At the initial configuration, the long loop in the F1 domain was modelled in a random configuration away from the F0-F1 pair (Fig. 3.9B). After 25 ns of simulation this loop approached the F0-F1 pair and was positioned close to the F0-F1 domains in an extended conformation on the same side of the protein as the F2 positively charged patch (Fig 3.9B). The loop then remained in this position for the remainder of the simulation. Calculation of the electrostatic profile of the talin head domain from the end of the simulation suggested that the placement of this highly positively charged loop close to the F0-F1 domains created an extensive positively charged surface on one side of protein, which is expected to regulate the talin/lipid interactions. It should be noted that, during the simulations, the loop did not form a helical conformation as suggested by recent experimental data (194). This is because the simulation timescale was not enough for the formation of a helical region. For the formation of a helical conformation a microsecond scale is required (346). Although there was a relatively large displacement between the F0-F1 and F2-F3 pairs, no significant displacement was observed within the F0-F1 or F2-F3 domain pairs.

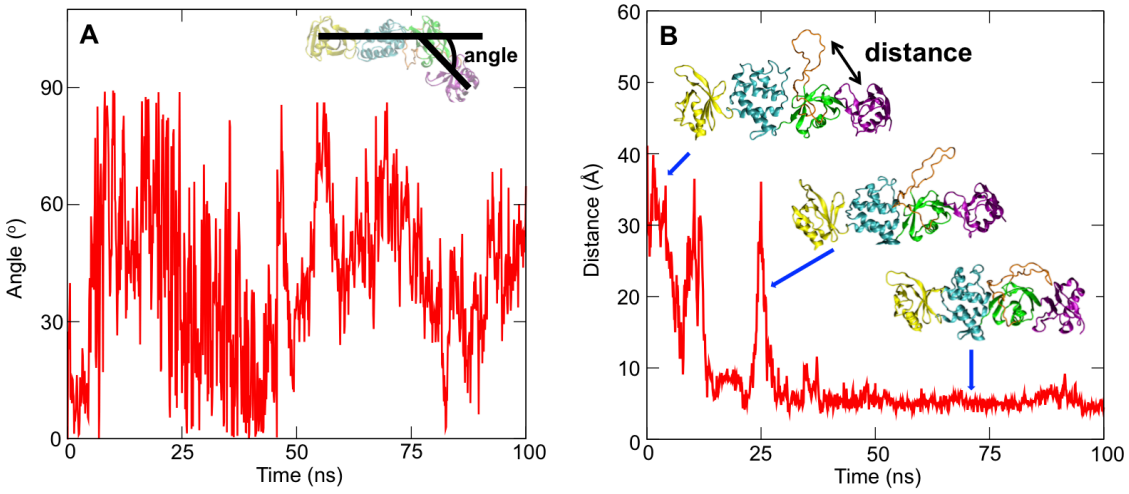


Figure 3.9: A. Change in the angle of the F0-F1 pair relative to the F2-F3 pair in one of the simulations with the talin head domain in solution as a function of time. The definition of the angle is shown in the inset picture. B. Distance between the F1 loop (res: L145) and the talin F0 domain (res: G11) as a function of time for the same simulation. The position of the loop is shown in the inset pictures at different time points.

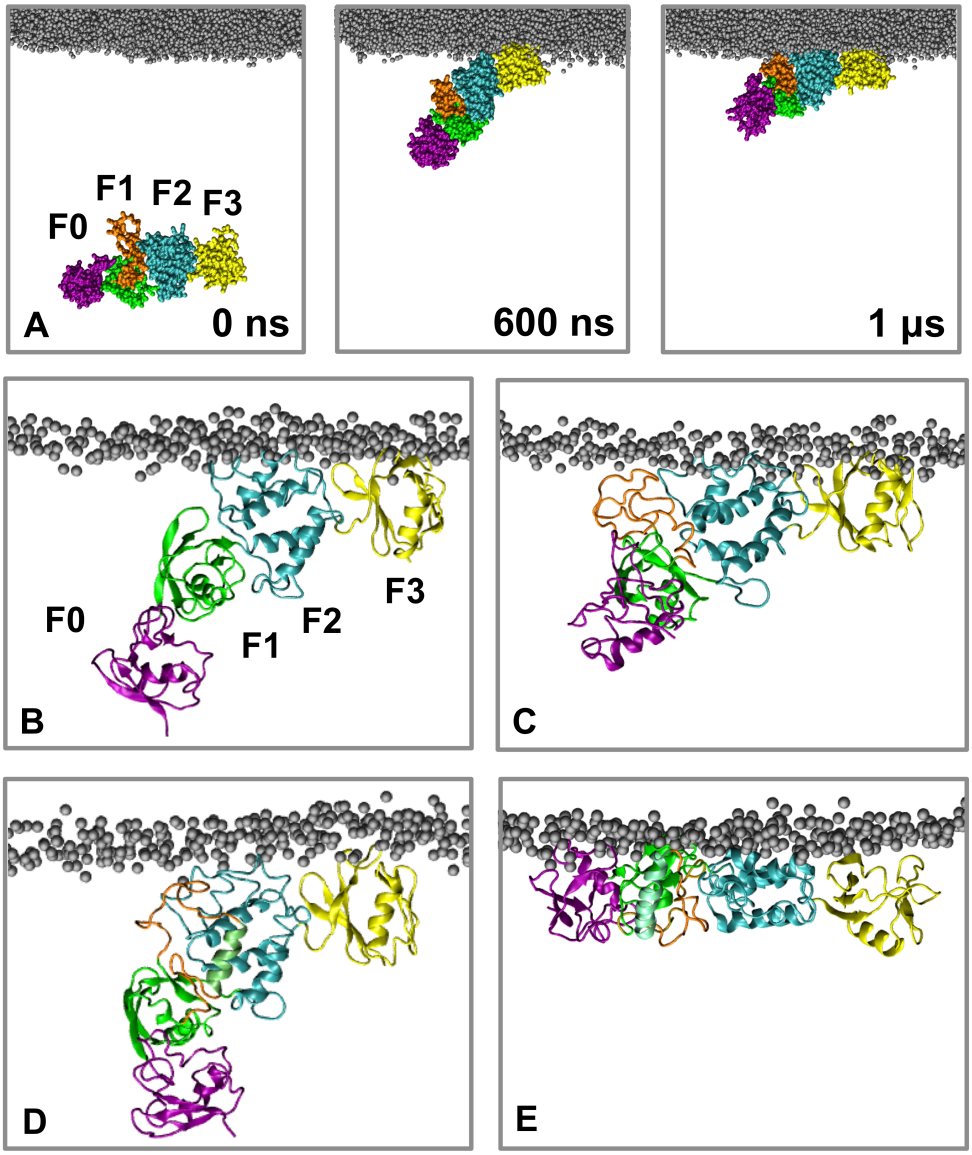


Figure 3.10: A. Progress of the simulations with the complete talin head domain. Snapshots from the simulation of the talin head domain with the F1 loop modelled are shown at 0 ns, 600 ns and 1 µs. The F0 domain is shown in purple, the F1 in green, the F1 loop in orange, the F2 in cyan and the F3 in yellow. The bilayer lipids are shown in grey. The waters in the simulation are omitted for clarity. B. Final orientation of the talin head domain in the *tal-CG* (B), *tal-l-CG* (C), *tal-h2-CG* (D) and *tal-h1-CG* (E) simulations; see Table 3.2 for more information. The same colour coding as in A is used. The lipid phosphate groups are shown as grey spheres.

Subsequently, five simulations of the talin head domain crystal structure (i.e. without the F1 loop) with a preformed POPC/POPG bilayer were run. Calculation of the separation between the centres of mass of the protein and the bilayer revealed that in three out of five simulations the talin head domain formed a complex with the bilayer. In all the simulations which yielded a talin/bilayer complex, talin bound to the bilayer initially via

the basic loop (res: 318-330) in the F3 domain and subsequently via the positively charged patch in the F2 domain (res: 255-285) (Fig. 3.10A). These two regions were identified above as promoting the productive binding of the isolated F2-F3 pair to the bilayer. The F0-F1 pair remained in the aqueous environment without forming any contacts with the bilayer even after F2-F3 binding (Fig. 3.10B). Calculation of the electrostatic profile of the talin head domain in the absence of the F1 loop revealed large positively charged patches in the F3 and F2 domains but not in the F0 or F1 domains. Therefore, the absence of a clear positively charged region on F0-F1 in the absence of the F1 loop may disfavour direct F0-F1/bilayer interactions. In the presence of the F1 loop (*tal-l-CG*), the talin head domain formed a complex with the bilayer in four out of five simulations, revealing a small increase in the docking of the talin head domain because of the addition of a highly positively charged loop. Interestingly, the talin head domain was associated with the bilayer in a similar fashion to that described above for the *tal-CG* simulation. In this case, however, after the association of F3 and F2 with the bilayer the F1 loop was positioned between the F1 subdomain and the bilayer in a compact form, as shown in Fig. 3.10C. Other regions of the F0-F1 domain pair remained in the aqueous environment and did not interact directly with the bilayer (Fig. 3.10C). Addition of a small helical region in the F1 loop (res: 154-167, h2 in Fig. 3.1C) resulted in talin/bilayer association in all five simulations performed. In this case, the talin head domain again adopted a similar orientation relative to the bilayer as described above; however, after the productive binding of F2-F3 to the bilayer, the helical region of the F1 loop (res: 154-167, h2 in Fig. 3.1C) was not positioned between the F1 domain and the bilayer (Fig. 3.10D). This may be due to the small space between the F1 and the bilayer after the F2-F3 interactions, which cannot accommodate the rigid helical region. The helical region fluctuated in the aqueous environment during the simulations (Fig. 3.10 D).

Simulations were set up to examine the suggestion that a longer helical region may be formed in the F1 loop (*tal-h1-CG*). In the presence of the longer helical region in the loop, the talin head domain bound to the bilayer in three out of five simulations, although in all three simulations talin was associated with the bilayer in a manner that was different from the above simulations (Fig. 3.10E). In this case, the presence of a large positively charged rigid domain (i.e. helix) in the F1 domain promoted the interaction of the F0-F1 pair with the bilayer before the F2-F3 interaction, resulting in non-productive binding to the bilayer (Fig. 310E). Intriguingly, the perturbed binding surface of the talin head domain observed in these simulations was similar to the perturbed binding surface

identified for the mutated form of the isolated F2-F3 and F2 domains and the isolated F0-F1 domain pair. When similar simulations were run with the charge of this helical region set to 0, the talin head domain associated with the bilayer in what is assumed to be the productive manner. It must be noted that, in this case, talin bound to the bilayer in only two out of five simulations. Interestingly, after the F2-F3 productive binding, the helical region remained in the aqueous environment in a similar form to that found in the *tal-h2-CG* simulation (i.e. with the small helical region; Fig. 3.10D). Therefore, the formation of a large helical region in the F1 loop prior to bilayer binding is rather improbable since it results in a perturbed talin/membrane complex. For the remainder of this study, only talin conformations with the h2 or no helical constraints (res: 154-167; h2 in Fig. 3.1C) were therefore used.

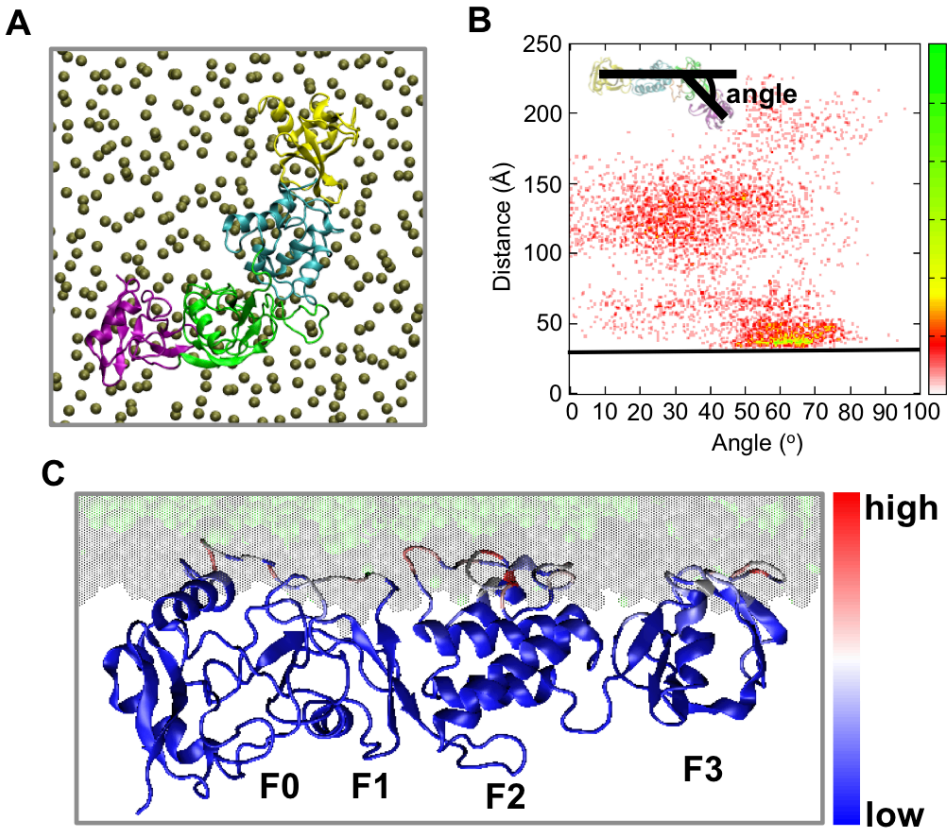


Figure 3.11: A. Final snapshot from the *tal-h2F0-CG* system. The figure demonstrates the displacement of the F0-F1 pair (purple-green) relative to the F2-F3 pair (cyan-yellow), which resulted in a novel V-shape conformation. B. Displacement of the F0-F1 pair relative it its initial position (shown as the change in the angle) as a function of distance from the bilayer for the *tal-h2F0-CG* simulation. The black line indicates the phosphate atom positions in the bilayer. C. Normalized number of contacts between the talin and the lipids mapped onto the final snapshot of the *tal-h2F0-CG* simulation. Blue indicates a low number, white indicates a medium number and red a large number of contacts. The bilayer headgroups are shown as grey spheres and the lipid tails as green spheres.

In all the above simulations, the loop (with or without any helical conformation) was modelled to be in random configuration in solution away from the four domains. However, the simulations with the talin head domain in solution suggested a tendency for the loop to be situated close to the F0-F1 on the same side as the F2 and F3 positively charged surfaces (Fig. 3.9B). To this end, simulations were set up where the loop (with or without any helical conformation) was restricted close to the F0 or the F1 domain in a similar form to that observed in the simulation with the talin in solution (*tal-lF0F1-CG, tal-h2F1-CG* and *tal-h2F0-CG*; Fig. 3.9B and Appendix A.2.1).

In the simulations where the helical region was constrained to be close to the F0 domain, talin bound to the bilayer in a productive orientation, similar to the *tal-h2-CG* and *tal-l-CG* simulations (above), albeit with the difference that the additional surface created by the loop between the F0-F1 pair and the bilayer now also interacted with the bilayer after the association of the F3 and F2 domains with the bilayer (Fig. A.2A). Interestingly, when this loop surface was associated with the bilayer the talin head domain adopted a V-shaped configuration due to rotation of the F0-F1 pair relative to the F2-F3 pair in the xy bilayer plane (Fig. 3.11A). Similar transient displacement between the F0-F1 and F2-F3 domains was also observed when talin was in solution and it was stabilized after the association with the bilayer (Fig 3.11B). The novel V-shape configuration allows the optimization of the talin/lipid interactions and induces a more compact packing of the domains, which is, however, still different to the canonical packing of other FERM domains (182). This optimized binding mode was also obtained in the *tal-lF0F1-CG* simulations where the F1 loop was restricted in a similar position close to the F0-F1 with no helical conformation. In contrast, in the simulations where the small helical region was restricted close to F1 domain (i.e. the new binding surface did not include the F0 domain; *tal-h2F1-CG* in Fig. 3.2), the talin head domain was no longer able to productively bind the bilayer. See Appendix A.2.1 for detailed discussion of the *tal-lF0F1-CG*, *tal-h2F1-CG* and *tal-h2F0-CG* simulations discussed above.

All the above simulations highlighted the crucial role of negatively charged lipids in the binding and interactions of the talin head domain with the membrane. To test this observation further, similar simulations were set up where the electrostatic interactions between talin and the bilayer were disrupted (i.e. either by mutations in the F3 or the F2 binding surfaces or by using a pure POPC bilayer; see Appendix A.2.2 for detailed discussion). In good agreement with previous observations, disruption of the interactions of talin with negatively charged lipids significantly decreased the talin/bilayer association.

Furthermore, in the simulations in which the talin/bilayer complex was formed, talin bound to the bilayer in a perturbed orientation (Fig. 3.12).

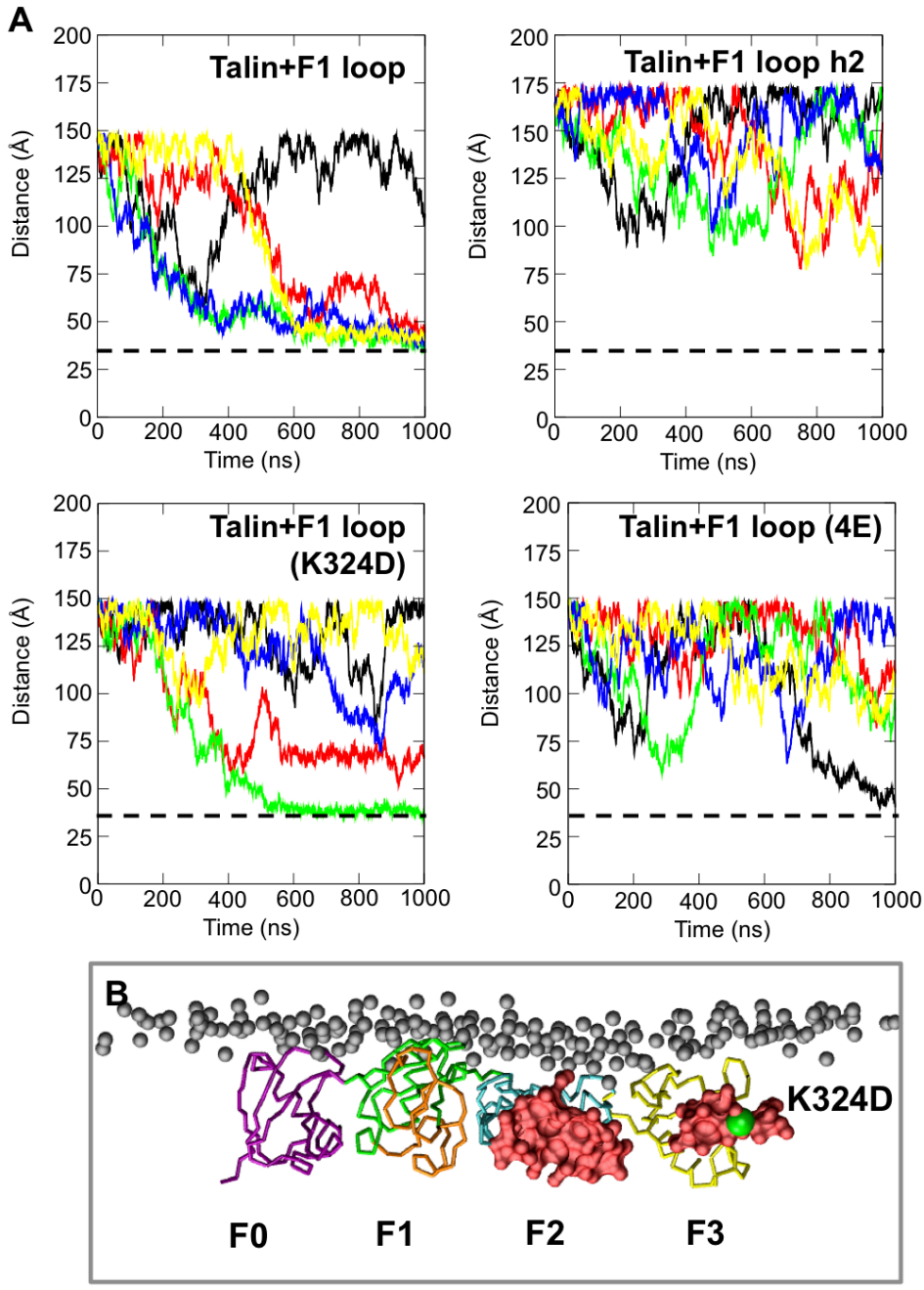


Figure 3.12: A. Separation between the centres of mass of talin and the bilayer as a function of time for the simulations with the pure POPC bilayer or the mutated form of talin. The same figure is shown for the simulation of the talin with the loop modelled in the F1 domain for comparison. B. Final snapshot demonstrating the final orientation of the talin head domain in the system with the K324D mutation in the F3 positively charged loop. The F0 domain is shown in purple, the F1 in green, the F1 loop in orange, the F2 in cyan and the F3 in yellow. The bilayer lipid headgroups are shown in grey and the backbone particle of the mutated residue is shown as a green sphere. The "correct" binding surface of the F2 and F3 domains, identified earlier, is shown as a light red surface.

The above simulations also suggest that the optimum association of talin with the membrane is enhanced by flexibility within the molecule, especially between the F2-F3 and F0-F1 domain pairs (Fig 3.1B). To this end, CG models of talin were prepared where the force constant and the cut-off distance of the ENM network were increased. This essentially means that no flexibility between the domains was allowed during these simulations. Interestingly, in none of the simulations performed did talin bind to the bilayer in the "correct" binding mode – a lack of flexibility prevented the association of the F2-F3 pair with the bilayer before the F0-F1/bilayer association.

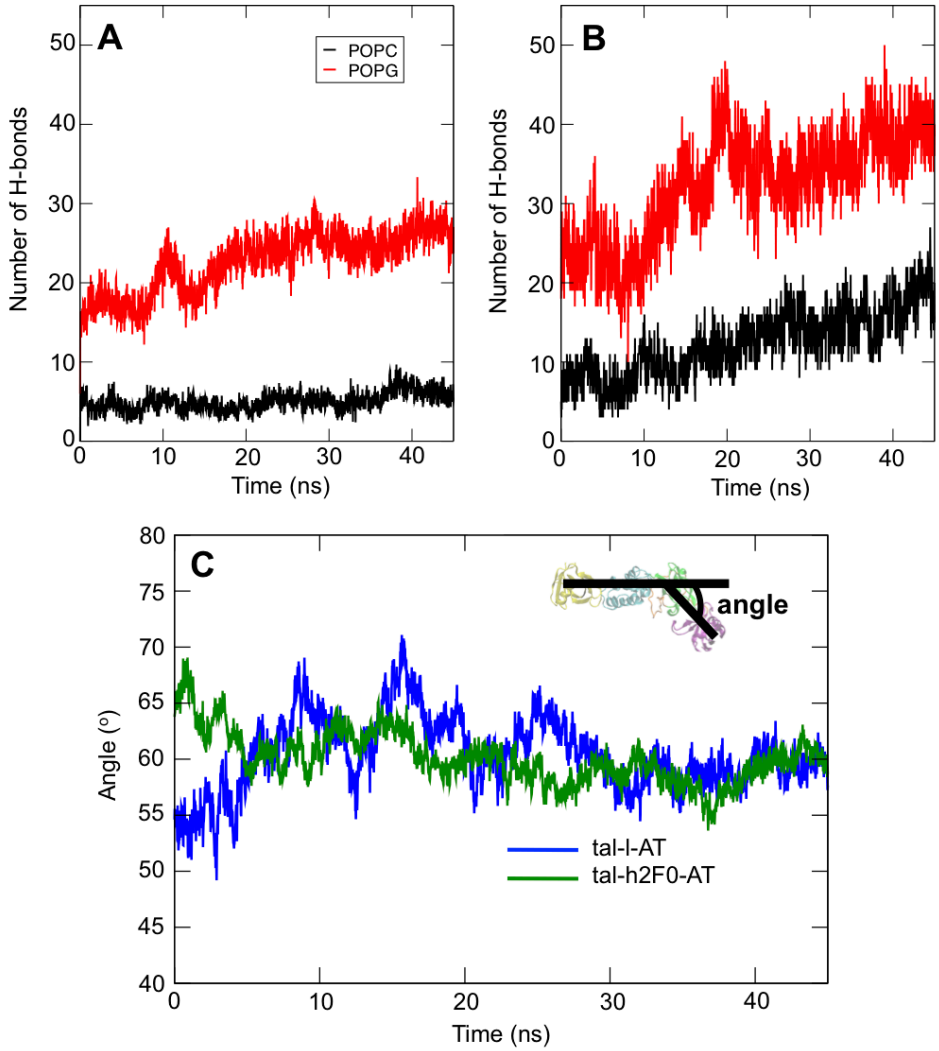


Figure 3.13: A,B. Number of H-bonds between the protein and the POPC (black) and POPG (red) lipids separately for the *tal-AT* (A) and the *tal-l-AT* (B) simulations. C. Mean value of the angle (over three AT-MD simulations) between the F0-F1 and the F2-F3 talin domains as a function of time for the *tal-l-AT* (blue) and the *tal-h2F0-AT* (green) simulations. See Table 3.3 for more information about the simulations.

Table 3.3: Summary of simulations

Simulation	Proteins	Duration (ns)
tal-AT	F0-F1-F2-F3 (Δloop-res: 134-172)	3x45
tal-l-AT	F0-F1-F2-F3 (loop-res: 134-172)	3x45
tal-h2-AT	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167)	3x45
tal-h2F0-AT	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167)	3x45

3.3.2.4 Refinement of the model by AT-MD simulations

In order to refine the binding modes derived from the CG-MD simulations, the last snapshot from one of the *tal-CG*, *tal-l-CG*, *tal-h2-CG* and *tal-h2F0-CG* simulations was converted to an AT representation (see Chapter 2 for the conversion process). In all these simulations, the talin head domain was associated with the bilayer in an orientation which allowed productive association of talin with the β integrin tail and was consistent with all the available experimental data (125, 186, 193, 195). In the *tal-AT* system, where the talin head domain lacked the F1 loop, the F0-F1 pair remained in the aqueous environment in a similar form to that observed in the CG-MD simulations above, without interacting with the bilayer (Fig. 3.10B). The F2-F3 domain pair orientation was also similar to the CG-MD simulations with the domains interacting preferentially with negatively charged lipids (i.e. POPG) via the positively charged regions on F2 and F3. In the simulations where a helical region was added to F1 loop (i.e. not restricted between F0 and the bilayer; *tal-h2-AT)*, talin also retained the orientation and interactions with the bilayer observed in the CG-MD simulations above (Fig. 3.10D). In contrast, in the *tal-l-AT* simulation where the F1 loop was added, additional interactions between the F0 domain and the lipids were observed in two out of three simulations. In this case, the pattern of interacting residues between the F2-F3 domain pair and the bilayer and the loop position between the F1 domain and the lipids were retained during the simulations, thus allowing F0 to interact directly with the bilayer. Interestingly, when the F0 domain was associated with the bilayer, the F0-F1 pair was displaced by ~65^{o} relative to the F2-F3 pair in the bilayer xy plane, allowing talin to adopt a similar V-shape configuration to the one observed in the *tal-h2F0-CG* simulation (Fig. 3.11A). The novel V-shape conformation of talin was also retained in the *tal-h2F0-AT* simulations, where the small helical region was restricted between the F0 and the bilayer (Fig. 3.13C). In this system, the loop region (which was situated between the F0-F1 pair and the bilayer) but no other regions of the F0-F1 domains interacted with the bilayer

together with the positively charged regions on F2 and F3 identified above. Calculation of the contacts and the H-bonds of the protein with the POPC and POPG lipids revealed a clear preference for talin to interact with the acidic POPG lipids in all simulation systems (Fig. 3.13A,B). Based on the work described in Chapter 6 of this thesis, where the talin F2-F3 pair was simulated for much longer timescales, local clustering of negatively charged lipids within the plane of the bilayer whilst forming contacts with talin is expected. Furthermore, in all simulations no significant displacement was observed between the F0 and F1 or F2 and F3 domains.

3.3.3 High-throughput method for studying protein/lipid interactions

The previous simulations suggested that, despite their simplifications, CG-MD simulations can predict the orientation of peripheral proteins relative to a bilayer and capture the main protein/lipid interactions. The simplification of the CG model (see Chapter 2), however, often raises some concerns about the accuracy of the predictions and the statistical significance of the observations (347). One direct way to check the accuracy of CG-MD simulations is to compare the results directly with experimental data. However, experimental results are not always available, especially for more complex proteins, and thus running a larger ensemble of simulations, where the initial protein configuration and position varies, can increase the accuracy and the statistical significance of the results. To this end, a fully automated high-throughput pipeline was developed. This pipeline is designed to run a large ensemble of CG-MD simulations to investigate the peripheral protein/lipid interactions in an automated way. Such automated high-throughput pipelines using CG-MD simulations have also been developed for studying single peptide insertions and orientations in a bilayer (348) and helix–helix dimerization (see Chapter 4). The pipeline is completely automated and therefore the user only needs to provide an AT structure of the protein of interest, along with some information on the lipid bilayer type and composition. Then, the pipeline prepares the protein, the bilayer, sets up the individual simulations, distributes the simulations to local clusters, retrieves the data and finally does some basic analysis of the system. By default, the pipeline produces an ensemble of 100 simulations of 1μs each but the user can change this number.

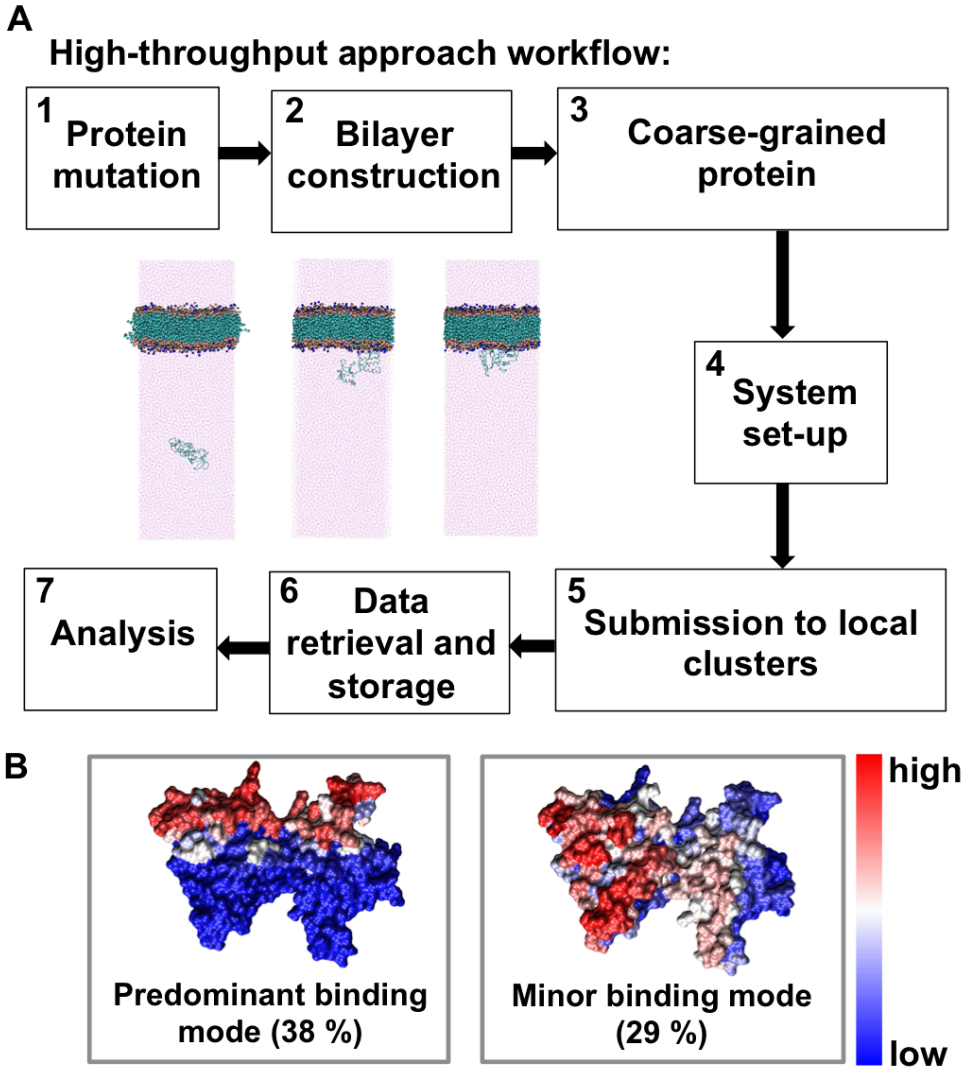


Figure 3.14: A. Workflow of the automated high-throughput pipeline. *Stage 1*: The pipeline uses Modeller (341, 342) to perform any mutations on the protein. *Stage 2*: The bilayer is constructed using self-assembly CG-MD simulations. The user can choose from a variety of phospholipids, phosphoinisitol, sphingolipids and cholesterol when constructing the bilayer. *Stage 3*: The protein is coarse-grained either using the MARTINI representation (320) or a local variant of MARTINI (321, 324). *Stage 4:* The individual simulation systems are built. In brief, the protein and the bilayer are inserted in a simulation box and the protein centre of mass is displaced 100 Å away from the bilayer. In each individual system, the protein is rotated around the x, y and z axes to form a new initial configuration. The system is solvated with CG water molecules, counter ions are added to neutralize the system and the system is energy minimized, equilibrated and production simulations are set up. *Stage 5:* The script connects to local (SBCB) clusters and runs the individual simulations. *Stage 6:* As soon as a batch of ten simulations finishes it moves all the data back to the local storage disks. *Stage 7:* After all the data are retrieved, the pipeline performs a basic analysis by calculating the protein Cα RMSD and RMSF, the protein binding surface and the protein contacts with each individual lipid type. B. The predominant (left) and minor (right) binding modes of the F2-F3 domain pair derived from an ensemble of ~100 simulations. The blue colour represents a small, white medium and red a high number of contacts.

Figure 3.14 displays the workflow of the pipeline. From a more methodological perspective, the user runs a central script which, for each stage, calls a second python script responsible for the specific stage. The pipeline is separated into seven stages, which are described in detail in Appendix B.

In order to evaluate this approach the F2-F3 domain pair which was studied above and a POPC/POPG bilayer at a ratio of 3:2 were used. For this system, ~100 individual simulations were set up using the automated pipeline. Each simulation was run for 1 μs. The simulation parameters were identical to the parameters used for the previous simulations of the F2-F3 pair (see methods). Analysis of the pattern of interacting residues in the simulation ensemble revealed that in 36 out of the 100 simulations F2-F3 bound to the bilayer with the same pattern of interacting residues as the previous CG-MD simulations (Fig. 3.14B). In 30 simulations, the F2-F3 domain pair did not form a complex with the bilayer and in the rest of the simulations it was associated with the bilayer with a perturbed orientation via two different binding modes (21 in one mode; Fig. 3.14B and eight in the second mode). These orientations were different from the predominant orientation (Fig. 3.14B) and they are expected to result in a perturbed, or unperturbed, association of talin with the β integrin tail. The F2-F3 Cα RMSD reached a plateau below 3 Å, suggesting a stable conformation in all simulations. The RMSF was as expected, with higher values for loop regions and lower for core structured regions. Calculation of the F2-F3/lipid contacts for different fractions of the simulation (i.e. 500 ns, 750 ns and 1μ) suggested that the F2-F3 bound to the bilayer directly with its preferred binding surface and then remained in the same orientation for the remainder of the simulation. Calculation of the number of interactions with individual lipid species separately revealed a threefold higher number of interactions with negatively charged POPG lipids compared to neutral POPC lipids.

Overall, these results are in good agreement with the previous studies and suggest that the talin F2-F3 domain pair binds to the bilayer in a preferred orientation via a positively charged patch on the F2 domain and an acidic loop on the F3 domain, thus orienting itself to interact productively with the integrin receptor. From a more methodical perspective, this “test” case suggests that this automated high-throughput approach can unravel the interaction details of peripheral membrane proteins with the membrane. Indeed, this pipeline provides an automated and efficient way of studying protein/lipid interactions and can be a valuable addition to ongoing efforts to study in detail the orientation and interactions of other peripheral membrane proteins or enzymes. Other types

of protein which this method could study include FERM domains (182, 184, 185) or C2 domains (349, 350).

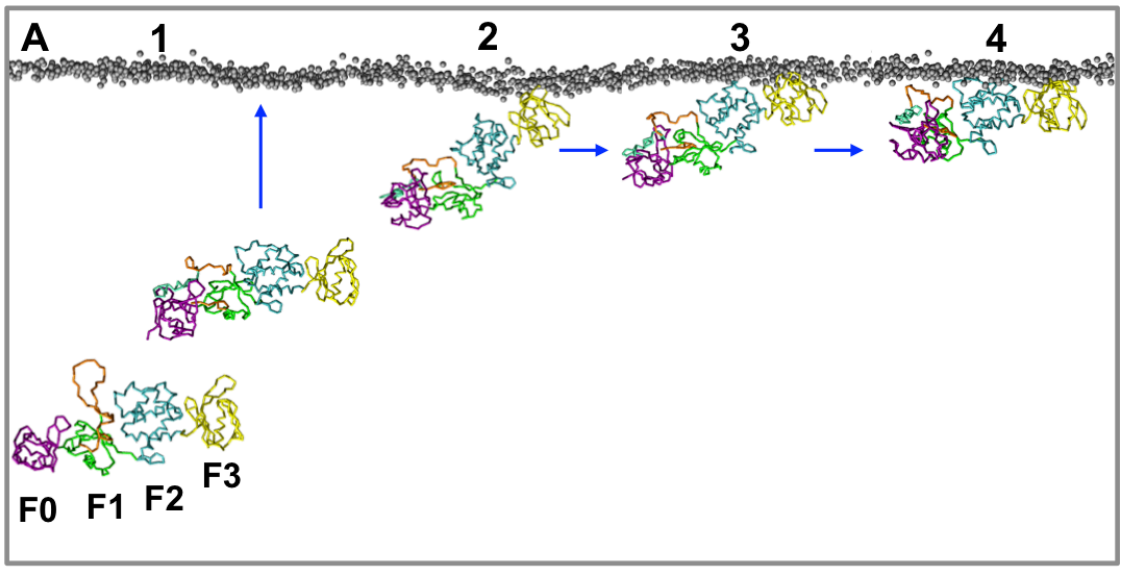


Figure 3.15: Summary of the mechanism of the talin head domain binding to a negatively charged membrane bilayer. 1. The long loop in the F1 domain is situated close to F0-F1 and next to the F2 positively charged patch. A small transient helix may be formed close to the F0 subdomain prior to binding to the membrane (194). The formation of a small helical region close to F0 may provide an additional solid positively charged surface which will facilitate the interactions of the F0-F1 pair with the membrane. It is also possible that the formation of the small helical region is related to the binding of other proteins (e.g. kindlins) which act in synergy with talin to induce integrin activation. 2. The talin head domain binds to the membrane via the F3 positively charged loop. 3. Subsequently, the F2 positive patch interacts with the membrane, orienting talin in a way that would facilitate binding to the β CT integrin tail. 4. After F2-F3 binding, the F0-F1 pair interacts with the membrane via the additional surface created by the F1 loop. Upon binding, the F0-F1 pair is displaced by ~65^{o} relative to the F2-F3 pair on a plane perpendicular to the bilayer normal, allowing talin to adopt a compact V-shape packing. The F0 domain is shown in purple, the F1 in green, the F1 loop in orange, the F2 in cyan and the F3 in yellow. The bilayer lipid headgroups are shown as grey spheres.

3.4 Discussion and conclusions

In this chapter, NMR and simulation techniques were combined to probe the nature of the talin head domain interactions with the membrane. A talin/membrane complex is a necessary step in integrin inside-out activation (29) and understanding the details of these interactions is a valuable addition to the integrin field. In this study, NMR and a multiscale simulation approach were initially used to demonstrate the crucial role of negatively charged lipids in F2-F3/membrane interactions, as well as to suggest key residues on the F2-F3 domain pair for these interactions. Subsequently, studies of the interactions of the

isolated F0-F1 pair suggested a non-specific interaction of this domain pair with the membrane. Finally, the interactions of the complete talin head domain with the lipids were explored, demonstrating how the F2-F3 pair facilitates productive binding of the complete talin head domain to the bilayer and how talin adopts a novel V-shape conformation which optimizes its interactions with the membrane.

In the NMR studies, by using two different types of charged liposomes it was demonstrated that enrichment of negatively charged headgroups in the bilayer promotes talin binding to the membrane. This correlates well with the observation from the simulations that the number of contacts between either the F2-F3 pair or the F2 domain and the anionic headgroups of POPG lipids were consistently threefold higher than with the zwitterionic headgroups of POPC lipids. In the NMR studies, the presence of F3 in combination with F2 seems to enhance the binding of talin to the membrane. Thus, in the presence of the F3 domain the protein is capable of binding strongly to 20% anionic liposomes, unlike the F2 subdomain alone. Therefore, anionic lipids in the bilayer and the presence of the F3 domain alongside the F2 domain are two key factors that enhance the interaction of talin with the membrane. The simulations with the complete talin head domain (F0 to F3) also revealed that electrostatic interactions of positively charged parts of the complete head domain orient talin productively, optimizing its interactions with the membrane. The surface of the talin head domain which binds to the membrane consists of a stretch of basic and acidic residues on the F3 domain, the F2 domain and the F1 acidic loop. The binding of these three areas to the membrane and their interactions with negatively charged lipids is key to the optimum binding of talin to the membrane and to the integrin receptor. Mutations in this interface resulted in non-productive binding of talin to the membrane, a finding in good agreement with other experimental studies (125, 186, 193, 194). In particular, residues K258, K274, R276, and K280 on the F2 domain and residue K324 on the F3 domain appear to make a substantial number of contacts with anionic lipids in the simulations, explaining why mutations in these residues disrupt interactions with the membrane. This binding surface is consistent with NMR (52), crystallographic (125) and recent TIRF microscopy studies (193).

Another significant outcome of this chapter relates to the role of the large insertion in the F1 domain. Experimental studies gave limited information on the role of this loop, mostly because of its dynamic nature (186, 194). The simulations presented in this chapter allowed a detailed investigation of the “conformational dynamics” of this loop, suggesting that it provides an additional positively charged binding surface close to the F0-F1 domain

pair and next to the F2-F3 positive patches. The loop is situated close to the F0-F1 pair while the talin is still away from the bilayer, possibly forming a small helical region close to the F0 domain. In contrast, the formation of a larger helical region on the F1 loop region while talin is still in the aqueous environment is rather improbable since it appears to perturb the talin/membrane interactions (194). The formation of this extra surface between the F0-F1 pair and the membrane is necessary for talin binding in order to avoid non-productive and non-specific association of F0-F1 to the membrane, as demonstrated by the simulations with the isolated F0-F1 pair. Indeed, in all simulations that yielded productive binding of the talin head domain to the membrane, the association of the F3 and F2 domains with the membrane always occurred prior to the F0-F1 association. In contrast, in all cases where structural characteristics of talin promoted binding of the F0-F1 pair to the membrane prior to the F2-F3 pair, the talin head domain bound in a perturbed, non-productive orientation. This result is in good agreement with recent quantitative data suggesting that the association constant of the complete talin head domain with lipids is similar to that of the F2-F3 to lipids, implying that the role of F0-F1 in binding is minimal (195). Therefore, it is likely that it is F2-F3 which regulates the talin/membrane association and that the F0-F1 pair is not required for productive binding of talin to the membrane, although it is a part of a larger binding surface that optimizes the talin/lipid interactions.

It is also important that the productive orientation of both F2-F3 and the complete talin head domain were observed in the simulations without the presence of the β tail. This supports the suggestion that the talin/bilayer interaction positions F3 for productive association with the β integrin CT tail. A perturbed orientation of F2-F3 relative to the membrane would be expected to result in a less productive interaction of F3 with the β integrin tail. This could also explain how talin can be released from its auto-inhibitory mechanism (203, 204). Upon binding to the membrane, the F3/lipid interactions will disrupt the F3/talin rod domain interactions, allowing F3 to interact with the β integrin tail. Indeed, the regions of the F3 which were identified in this chapter to interact with the membrane were also shown to be involved in the F3/rod domain interactions (203, 204).

The current study also explains how the dynamic nature and in particular a flexible linker between the F2-F3 and F0-F1 pairs allows talin to adopt a novel V-shape conformation, optimizing its contacts with the lipids (Fig. 3.11A). This V-shape conformation was also observed while talin was in solution and it is stabilized when talin binds to the membrane, possibly due to the interactions between talin and negatively charged lipids in the membrane. Studies by Bouaouina *et al.* suggested that the activation

of β3 integrins is enhanced in the presence of the complete talin head domain compared to the F3 alone. In addition, in the same study the authors suggested that the F1 domain is required for the activation of the α5β1 integrins (56). Therefore, the formation of the V-shape conformation is likely to be of mechanistic importance since it promotes a more stable talin/membrane complex, which may be important for how talin regulates the conformational changes on different integrins.

Chapter 4

A Helix Heterodimer in a Lipid Bilayer: Prediction of the Structure of an Integrin Transmembrane Domain via Multiscale Simulations

4.1 Introduction

Lateral association of TM α-helices is a key stage in the assembly and function of a number of α-helical bundle proteins, including cell surface receptors, channels and transporters. The membrane protein assembly process can be explained using simple models, for example a two-stage model (264, 351) which involves insertion and folding of α-helices in the membrane (352) followed by self-assembly of the helices into helical bundles (353). Helix–helix dimerization provides a model for the second stage of this procedure which, despite its simplicity, can be used to explain some of the main principles of bundle assembly. Whilst the TM association can be driven by different factors such as VDW interactions, hydrogen bonding interactions, ionic interactions, polar interactions and the lipid environment (354-357), several studies have suggested that certain motifs in the TM region have an important role in the helix–helix association (353, 358-360). Some of these motifs include prolines, polar residues such as Asn, Gln, Asp, His (361), and SxxG (362) or QxxS motifs (363) and these are found in many known structures of α-helical bundle proteins. This chapter will focus on the Gx_3G or Gx_3G-like motifs (364-366), since they seem to be particularly important in TM signalling (367) and are present in the integrin TM region (118).

Senes *et al.* performed a statistical analysis on sequences of the TM regions of membrane proteins and suggested that a small-x_3-small amino acid (small might be alanine, glycine or serine) motif was frequently observed in TM helices, especially in TM regions of proteins involved in TM signalling (368). This motif is found in integrins, ErbB (369), GPCR, BNIP (366), syndecans (370) and GpA. Small apolar residues (i.e. glycine or alanine) allow close packing of the helices, stabilize the TM interactions and regulate the helix–helix crossing angle (152, 355, 358, 359, 371). TM dimerization often regulates the affinity of ectodomains for ligands (372) and mutations in these motifs can alter signalling and result in severe defects (23). It must be noted, however, that even in the simple case of

dimer formation by two helices there is potential complexity, since helices might interact with multiple interfaces (373, 374) or specific residues, e.g. proline, can perturb helix–helix packing (375). Moreover, helix–helix interactions can be enhanced by H-bond networks or other specific interactions (376-378).

One of the best studied receptors where a Gx_3G motif regulates the packing of its TM region is GpA (152-157). In GpA, the Gx_3G motif is located in the interface of the TM dimer, allowing the formation of VDW interactions and Cα-carbonyl H-bonds (376). Various studies have shown that mutations in the two glycines in the Gx_3G motif (155, 379) or in residues L75, I76, V84 and T87 which are associated with the Gx_3G motif (153, 364) significantly reduce or alter TM dimerization. Another family of proteins which contains the Gx_3G-like motif is the ErbB receptor (380). ErbB has a single TM helix with two GxxxG-like motifs, which are important both for receptor homodimerization (369) and heterodimerizaton between ErbB2 and ErbB1 (381). Similarly, a Gx_3G motif in TM helix 1 of the G protein α-factor was also shown to be important for oligomerization of this type of receptor (367). In some receptors, the Gx_3G motif is part of longer motifs. For example, the TM region of the BNIP receptor has a Hx_6Gx_3G (366, 382) motif and the SRBI receptor contains a $Gx_2Gx_3Ax_2G$ motif. Studies by Unterreitmeier *et al.* demonstrated that certain amino acids such as phenylalanines combine with the Gx_3G motif to promote helix–helix association (383). It must be noted, however, that there are proteins, such as SNARE proteins, which have a Gx_3G motif not involved directly in helix–helix interactions (384).

Integrins are another example of cell surface receptors which are involved in TM signalling and have a Gx_3G motif in their TM region (Fig. 4.1A) (34, 41, 45). It was shown by recent structural (118, 120) and functional data (159, 173, 385) (reviewed in Chapter 1) that this Gx_3G sequence, located in the OMC region of the integrin TM region, regulates the packing of the integrin TM helices. The integrin TM heterodimer is a more complex system than GpA and thus provides a valuable test of computational approaches which model the TM dimerization of membrane receptors. In this chapter, a multiscale MD approach (combining CG-MD and AT-MD simulations) is used to explore αIIb/β3 integrin TM heterodimerization. The results of this chapter demonstrate that simulation methods predict a structure of the αIIb/β3 integrin TM region that is similar to that observed in NMR studies. The chapter also shows that mutations of the Gx_3G motif perturb the packing of the TM helices.

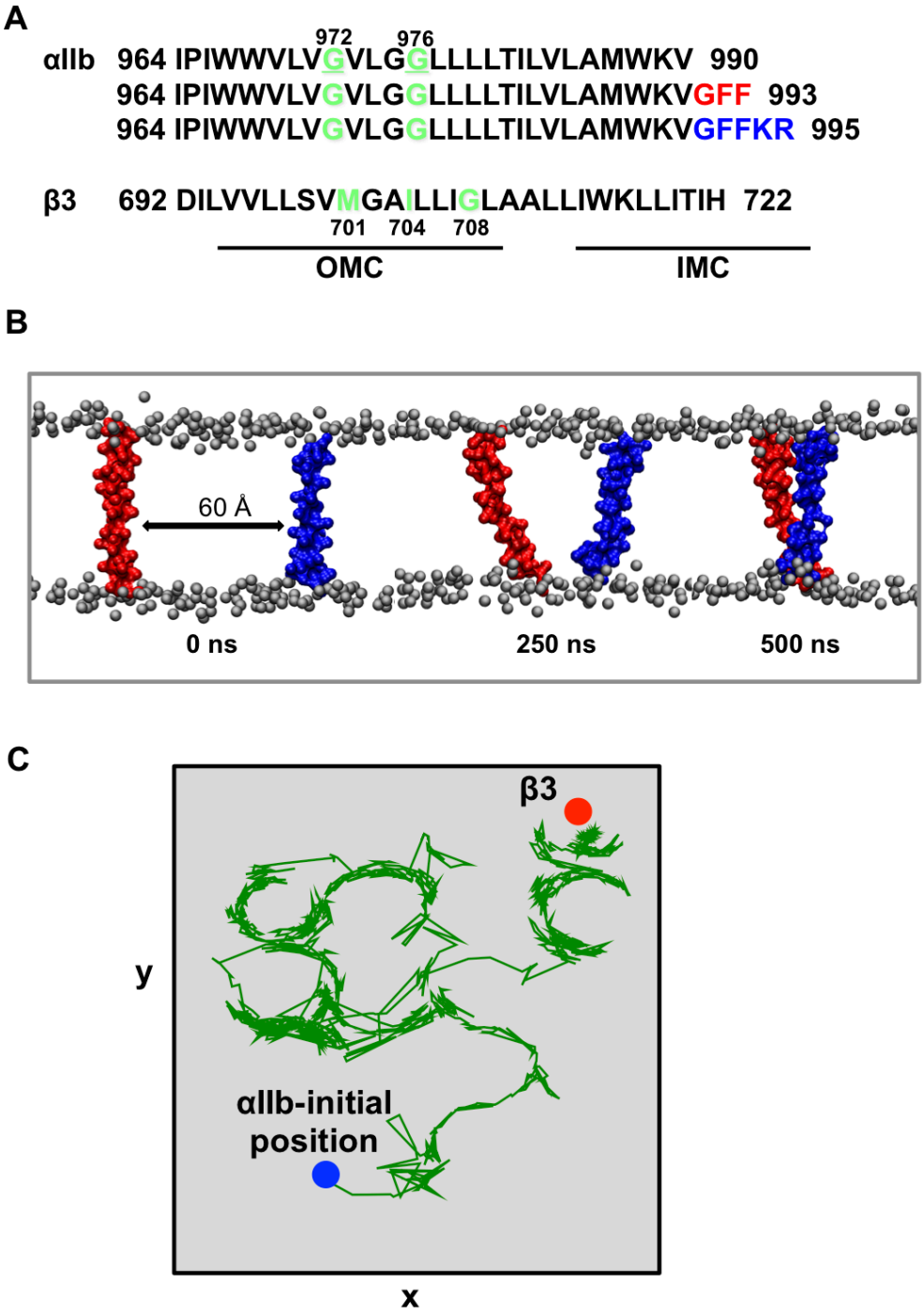


Figure 4.1: A. Sequences of the αIIb and β3 integrin TM helices used in the simulations. Key residues which are involved in the packing of the two helices (and mutated in the simulations) are coloured in green. The residues in the IMC region which were added to the second set of simulations are coloured in red (*CG-IMC-1*) and blue (*CG-IMC-2*). B. Progress of the CG-MD simulations. Snapshots of the system are shown at 0 ns, 250 ns and 500 ns. The αIIb subunit is shown in blue, the β3 subunit in red and the lipid phosphate groups are shown as grey spheres. C. Motion of the centre of mass of the αIIb helix relative to the β3 helix in the plane of the lipid bilayer. The coordinates (x,y) of the mobile αIIb helix (green) are shown relative to the position of the fixed β3 helix to show the random movement of the αIIb helix towards the β3 helix as the dimer is formed.

Table 4.1: Summary of simulations

Simulation	System	Duration (ns)
CG-WT	αIIb/β3 WT	~100x500
CG-WT-rot	αIIb/β3 WT	~100x500
CG-αG972L	αIIb (G972L)/β3	~100x500
CG-αG976L	αIIb (G976L)/β3	~100x500
CG-αG972L,G976L	αIIb (G972L,G976L)/β3	~100x500
CG-βG708I	αIIb/β3 (G708I)	~100x500
CG-IMC-1	αIIb (F993)/β3 (H772)	~100x500
CG-IMC-2	αIIb (R995)/β3 (H772)	~100x500
AT-RH1	αIIb/β3 WT, RH1 dimer	3x50
AT-RH2	αIIb/β3 WT, RH2 dimer	3x30
AT-LH	αIIb/β3 WT, LH dimer	3x30
AT-αG972L	αIIb (G972L)/β3	3x30
AT-αG976L	αIIb (G976L)/β3	3x30
AT-αG972L,G976L	αIIb (G972L,G976L)/β3	3x30
AT-βG708I	αIIb/β3 (G708I)	3x30
AT-NMR	αIIb/β3 (PDB:2K9J)	3x50
AT-RH1-OPLS	αIIb/β3 (RH1 dimer)	3x50

4.2 Methods

4.2.1 Simulation set-up

In order to study integrin TM region dimerization, the sequences of αIIb (res: 964-995) and β3 (res: 692-722) were employed (Fig 4.1A). In the first set of simulations only the helical regions (as defined by structural studies (118, 119)) of the αIIb (res: 964-990) and β3 tails (res: 692-722) were used (Fig 4.1A). The non-helical IMC region, which includes the αIIb 991GFFKR995 motif, was added in a second set of simulations where the C-terminus of the αIIb helix was extended initially by three residues to include the aromatic phenylalanine residues and subsequently by five residues to study how two basic residues (K994 and R995) affect the packing of the two helices (Fig 4.1A). In all CG-MD simulations, the sequences were modelled as α-helices. After conversion to CG representation (see Chapter 2), the two helices were positioned ~60 Å away from each other (in parallel orientation) in a ~128 DPPC bilayer (Fig. 4.1B). This separation distance between the helices was significantly larger than the cut-off distance for electrostatic and VDW interactions in the force field, and therefore no inter-helix interactions occurred at the beginning of the simulation. This ensured that the initial position did not favour dimer formation. The CG-MD simulations were performed as described in Table 4.1 using a

high-throughput framework (348). For all systems (see Table 4.1), representative structures were selected which were then converted to AT representation using a fragment-based approach as described in Chapter 2 (339). The AT-MD simulations were used to check the stability of the dimers obtained by the CG-MD simulations and to refine the helix–helix packing.

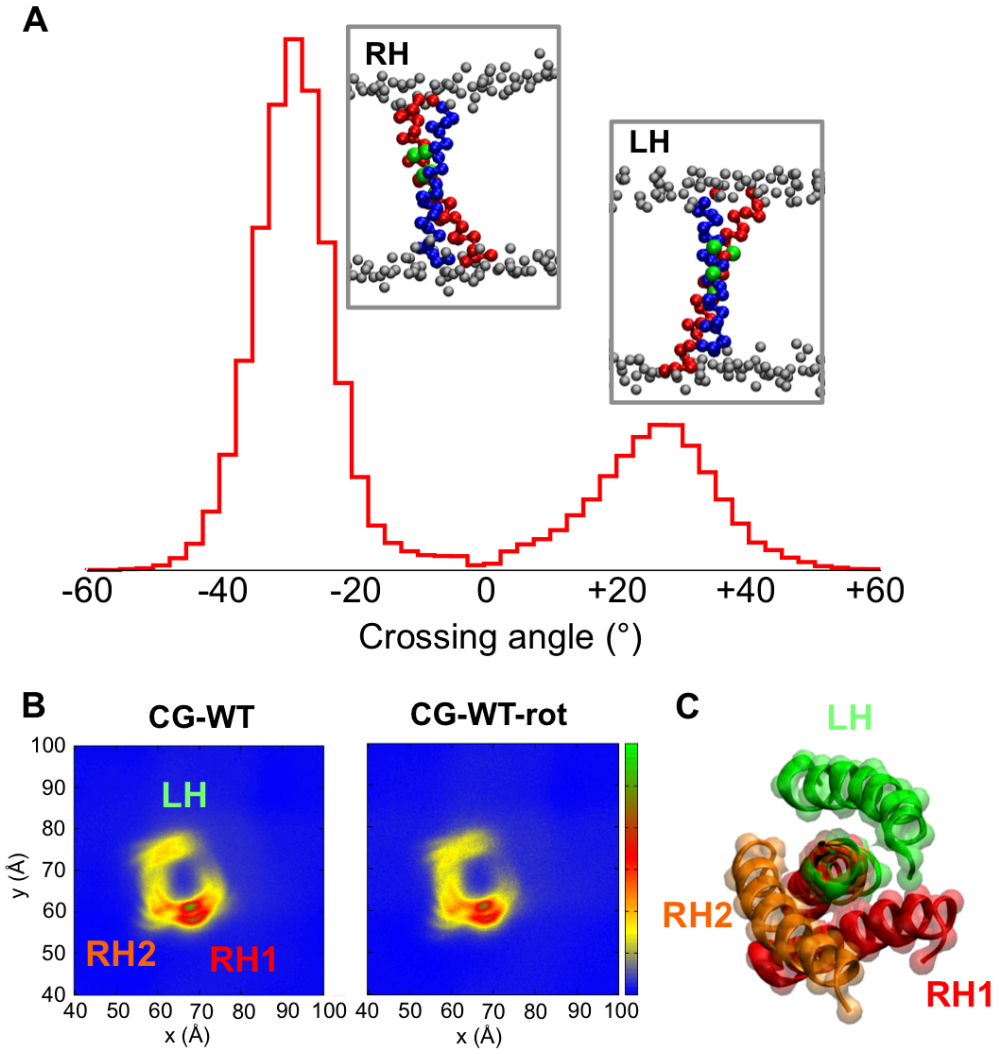


Figure 4.2: A. Helix–helix crossing angle distribution (for the CG-MD simulations) observed for WT sequences. The negative values of the crossing angle represent right-handed (RH) packing of the dimer and the positive values represent left-handed (LH) packing. The inset pictures show representative structures for the RH and the LH dimers derived from the *CG-WT* simulation. The αIIb helix is shown in blue, the β3 helix in red and residues which were involved in the packing of the two helices (and mutated in the simulations) are coloured in green. The lipid phosphate atoms are coloured in grey. B. Spatial distribution maps of the αIIb helix relative to the β3 helix for the *CG-WT* (left) and the *CG-WT-rot* (right) simulations. In this analysis, the β3 helix was fitted to a reference structure (see text for more information). The probability density of finding the αIIb helix at any point of the bilayer plane is shown as blue for low probability through to red and then green for high probability. C. Representative structures from the three maxima (i.e. RH1, RH2 and LH) found in the spatial distribution analysis.

4.3 Results

4.3.1 Wild type αIIb/β3

4.3.1.1 Dimerization by CG-MD simulations

During the CG-MD simulations of the WT sequences, the two helices diffused randomly in the bilayer until a stable dimer was formed, usually within a few hundred nanoseconds (Fig. 4.1B,C). Visual inspection of the dimers revealed a preference for the helices to pack in a right-handed (RH) manner as expected for packing of helices involving a Gx_3G motif (152). It must be noted, however, that there was a degree of variation both in regard to the helix packing and the time which the packing occurred between the individual simulations. This dynamic aspect of the helix packing can be seen by analysing the crossing angle distribution of the whole ensemble. The crossing angle distribution revealed a bimodal distribution (Fig 4.2), with the majority corresponding to RH packing of the helices (approximately 70% of the structures). A similar bimodal distribution for the integrin TM sequences was also suggested by preliminary studies by Psachoulia *et al.* (161). The mean crossing angle in the RH packing mode was ~-30^o, which is in good agreement with the crossing angle proposed by the αIIb/β3 TM NMR structures in bicelles ($-25^o \pm 3^o$) and in a CD_3CN/H_2O solvent system (-30^o) (118, 119). In approximately 30% of the simulations, the helices were packed in a left-handed (LH) packing (minor) mode with a crossing angle of $+30^o$.

Inspection of the selected structures corresponding to RH and LH packing revealed significant differences in the helix–helix interface. The major difference was found in the position of the αIIb Gx_3G (G972 to G976) motif, which in the LH packing mode was directed away from the β3 helix (not in the helix–helix interface), whereas in the RH packing mode the same Gx_3G motif was directed toward the β3 helix (Fig. 4.2A). For more detailed investigation of the packing of the two helices, the trajectories from all CG-MD simulations were merged and the backbone particles of the β3 helix were fitted to a reference structure. This allowed the calculation of the spatial distribution of a mobile αIIb around a static β3 helix (Fig. 4.2B). This analysis suggested a major binding mode (position RH1 in Fig. 4.2B) and two minor binding modes (RH2 and LH in Fig. 4.2B). Examination of the helix–helix contacts (using a cut-off distance of 8 Å) in the major RH1 mode revealed that αIIb residues W968, G972, G976 and L980 and β3 residues I693, L967, V700, M701, I704, L705 and L712 were in the helix–helix interface, in good agreement with the interface residues suggested by the NMR structures (118, 119). Calculation of the Cα RMSD between the RH1 dimer and the same part of the NMR

structure (PDB:2K9J) revealed a value of 2.3 Å, suggesting that the predominant structure found in simulations resembled the experimental structure. This RMSD was improved after the AT-MD simulations, which refined the structure (see below). Similar calculations for the RH2 and LH modes revealed very large values for the Cα RMSD, suggesting a very different packing of the helices compared to the NMR structure. Finally, analysis of the position of the two helices in the bilayer revealed a tilt angle relative to the bilayer normal of ~15^{o} for the αIIb and of ~27^{o} for the β3. Convergence analysis of this simulation ensemble by randomly using 25%, 50%, 75% and 100% of the simulation ensemble revealed that the system converged after using 50% of the simulations.

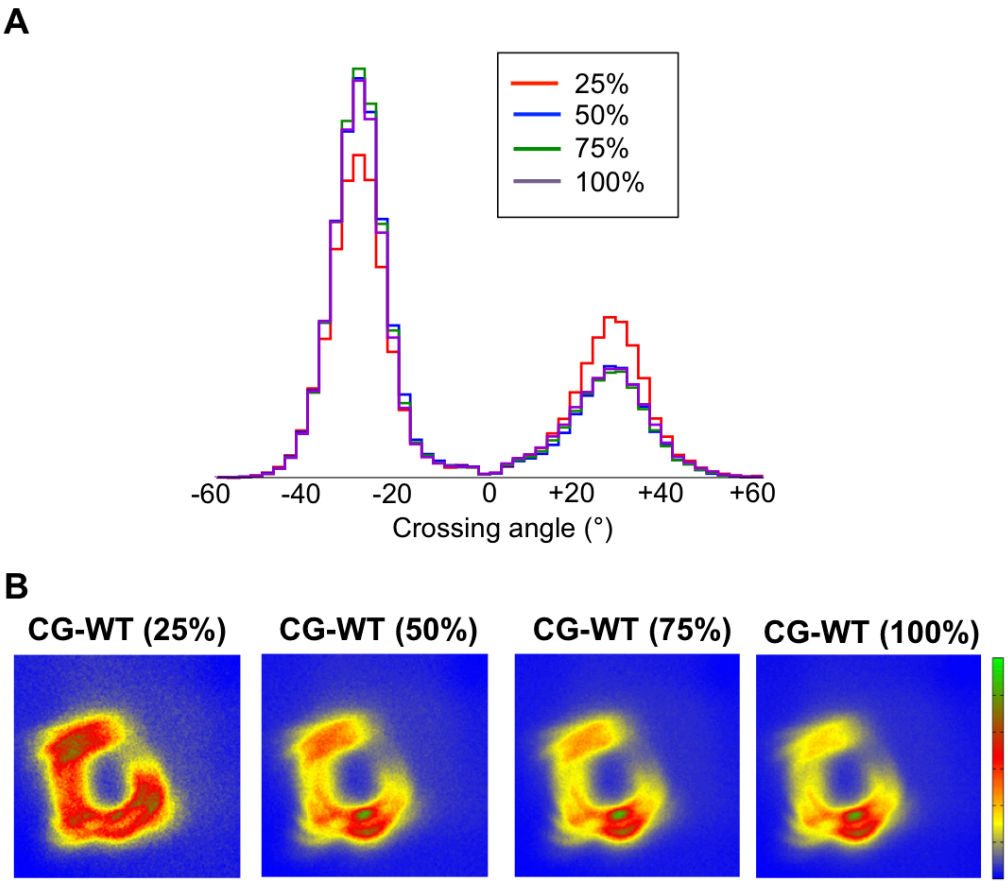


Figure 4.3: A,B. Convergence analysis of the crossing angle distribution (A) and the spatial distribution of the αIIb helix relative to the β3 helix (B) from the *CG-WT* simulations is shown for 25%, 50%, 75% and 100% of the simulation ensemble. The individual simulations from the ensemble used for the 25%, 50% and 75% subsets were chosen randomly.

To test if the initial position of the helices affected helix–helix dimerization, the same protocol was used as above with the two helices in the same initial positions but with the β3 helix rotated by a 15^{o} increment in each subsequent individual simulation (*CG-WT-*

rot). Approximately ~100 simulations were again performed. The results (i.e. binding modes, crossing angle distribution and helix–helix interface) were essentially the same (Fig. 4.2B) as in the above simulation (i.e. *CG-WT*). Therefore, for the remainder of this study the protocol described in the methods section is used.

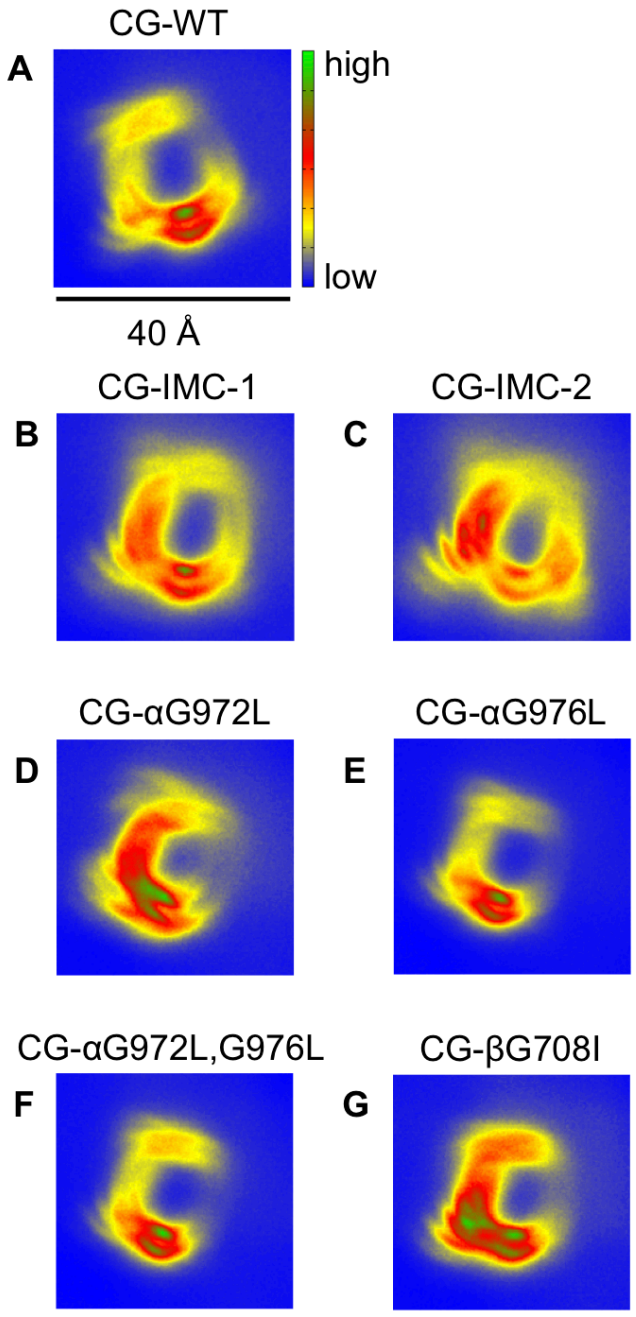


Figure 4.4: Spatial distribution maps of the αIIb helix relative to the fitted β3 helix when the IMC region was added (B, C) or when mutations of glycines (αIIb G972, G976 and β3 G708) in the OMC region were performed (D, E, F, G). The same distribution is displayed for the simulation with the WT dimer for comparison (A). The probability of finding the αIIb helix at any point of the bilayer plane is shown as blue for low probability through to green for high probability.

In addition to the studies above, which included only the OMC region of the αIIb/β3 dimer (118), simulations were set up to explore the role of residues in the IMC region (Fig 4.1A). In the first set of simulations, the C-terminus of the αIIb subunit was extended by three residues to include the aromatic cluster (i.e. F992, F993) in the IMC region (Fig. 4.1A). This extension of the αIIb helix did not greatly perturb the TM helix packing profile (Fig. 4.4B); however, a slightly modified RH1 binding mode was observed to be the major mode. In contrast, extension of the α-helix by five residues (res: 694-995) to include two basic residues in the IMC region significantly perturbed the helix packing (Fig. 4.4C). In this case, the RH1 mode observed in the simulations above vanished almost completely. However, the presence of two additional positively charged residues in this region resulted in stronger interactions of the αIIb with lipid phosphate groups compared to systems with only the OMC region (above). The loosening of the native packing in the presence of a helical conformation in the IMC region was also suggested by recent CG-MD simulations of the same integrin region using the MARTINI force field (147). Thus, the studies in the remainder of this chapter focus on the OMC region.

4.3.1.2 Refinement and model assessment by AT-MD simulations

In order to refine and further evaluate the stability of the binding modes derived from the spatial distribution analysis of the CG-MD simulations of the WT sequences (RH1, RH2 and LH), a representative structure from each binding mode was converted to an AT representation. Three AT-MD simulations for each system, starting from the same initial configuration but with different velocities, were performed (Table 4.1). Similar simulations have been used previously by Psachoulia *et al.* (161) and Volynsky *et al.* (386) for the evaluation of model structures of membrane proteins derived from simulations. Alignment of the final snapshot from the AT-MD simulations of the RH1 mode with the NMR structure (PDB:2K9J) (118) revealed very similar packing (Fig 4.5B). The conformational stability of the dimer during the simulation was assessed by calculating the Cα RMSD from the initial structure. For the RH1 dimer the Cα RMSD reached a plateau at a value of ~3 Å, suggesting a conformationally stable dimer which retained its RH packing (387) (Fig. 4.5A). The crossing angle distribution for this system was relatively narrow at around -35^{o}, which is characteristic of "stable" dimers. As a result, the interface between the two helices remained intact throughout the simulation and was identical to the interface proposed by two NMR structures and the CG-MD simulations (see above; Fig. 4.6) (118, 119).

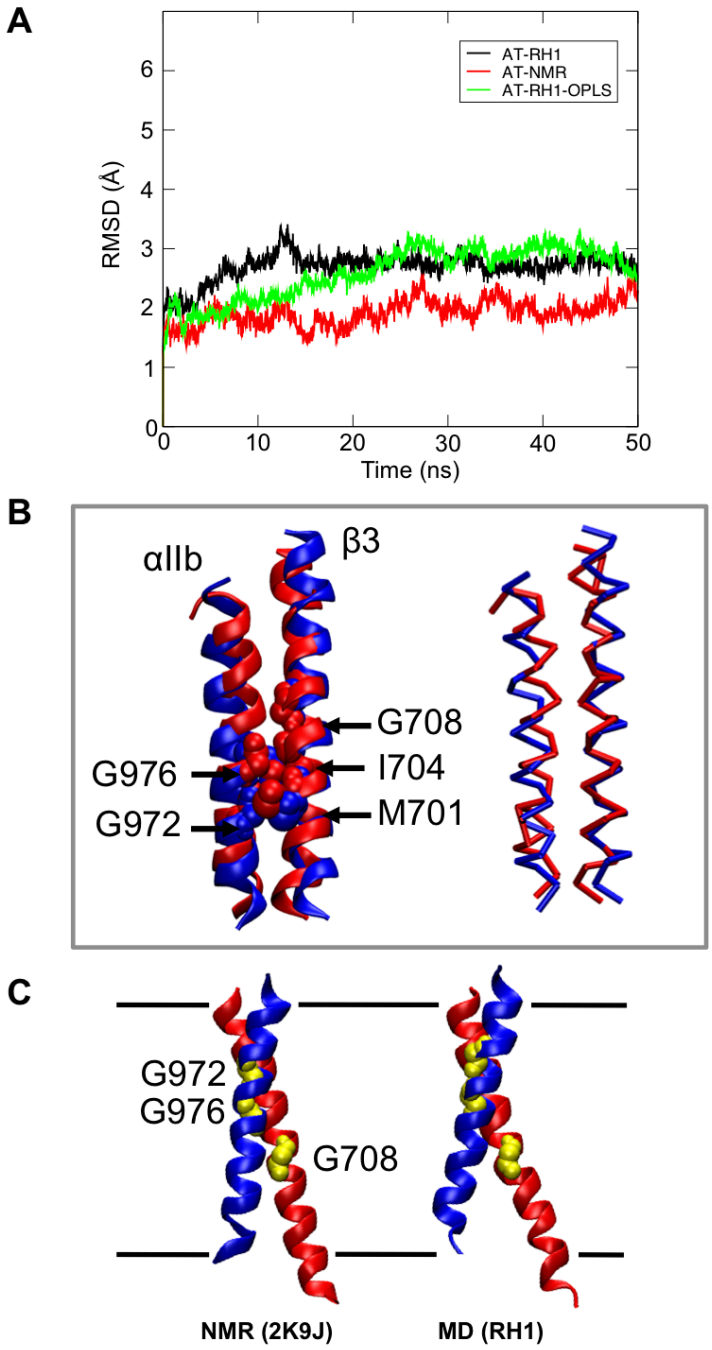


Figure 4.5: A. Cα RMSD averages across the three repeat simulations of the NMR structure (2K9J; *AT-NMR*) and the RH1 structure with GROMOS 43a1 (*AT-RH1*) and OPLS (*AT-RH1-OPLS*) force fields. B. Comparison of the NMR structure (blue) with a structure from the end of one of the *AT-RH1* simulations (red). The αIIb helix G972 and G976 residues and the β3 helix M701, I704 and G708 residues are shown in VDW format. C. The final structure derived from the simulation approach compared with the NMR structure. Glycine residues αIIbG972, αIIbG976 and β3G708 are shown in yellow.

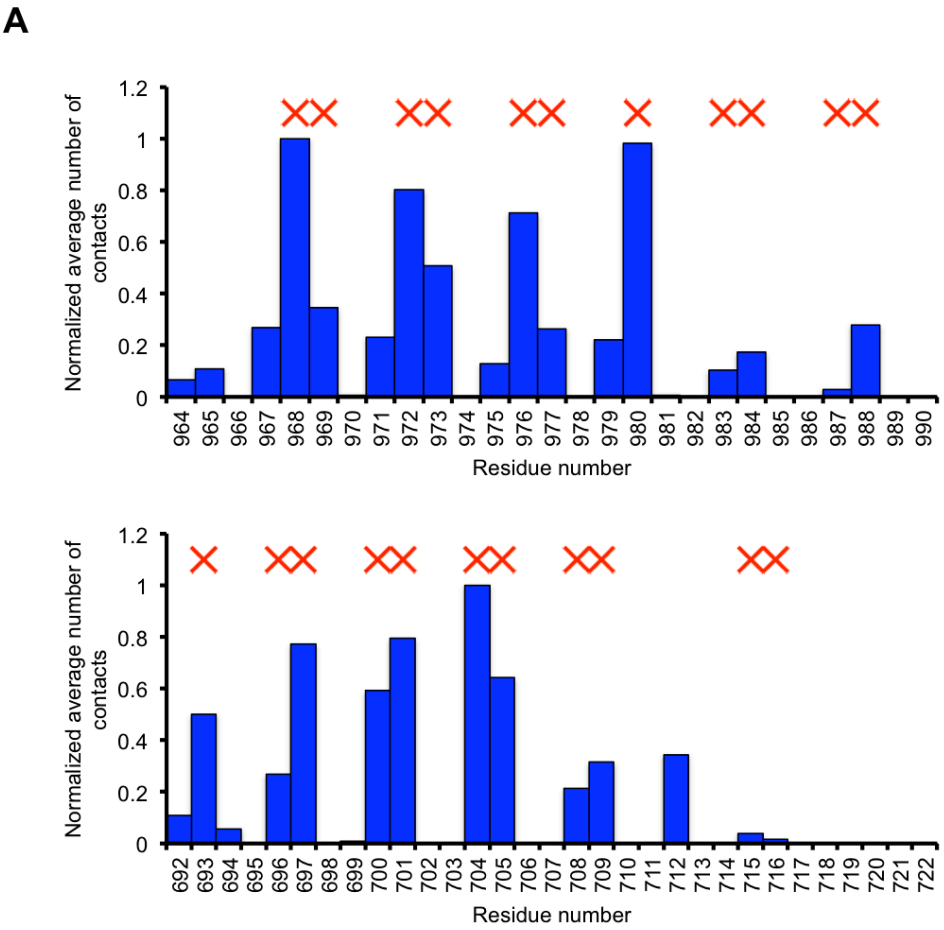


Figure 4.6: A. Normalized average number of contacts between the αIIb helix and the β3 helix in the *AT-RH1* simulation. The contacts were calculated using a cut-off distance of 3.5 Å. The residues suggested to be in the dimer interface by the NMR structure (PDB:2K9J) are indicated by red crosses.

Alongside this simulation system (i.e. *AT-RH1*), three simulations using the same DPPC bilayer with the αIIβ/β3 NMR structure obtained in bicelles were also performed (118). These simulations were used to compare the most stable predicted structure (RH1) with an NMR structure solved in a membrane-like (albeit different, with fluid bilayer as opposed to bicelles) environment (388, 389). For these simulations, the same part of the NMR structure (αIIbI964-V990/β3D692-H722) was used as in the *AT-RH1* simulation (see above). During the simulations, the dimer was conformationally stable with a crossing angle of ~-30^{o} ± 3^{o}. Calculation of the Cα RMSD after 50 ns of simulation revealed a value of ~2.5 Å, which was comparable to the Cα RMSD of the *AT-RH1* simulation. Pairwise Cα RMSD calculations between the simulations with the NMR and the RH1 structures revealed values between 1 and 3 Å for the last 15 ns of the simulations, suggesting that they converge to the same structure. Finally, to test the effect of different force fields on

the stability of the RH1 structure, additional simulations using the OPLS force field (306, 390) were performed. These simulations yielded comparable values of the Cα RMSD with the RH1 and NMR structures (Fig. 4.5A).

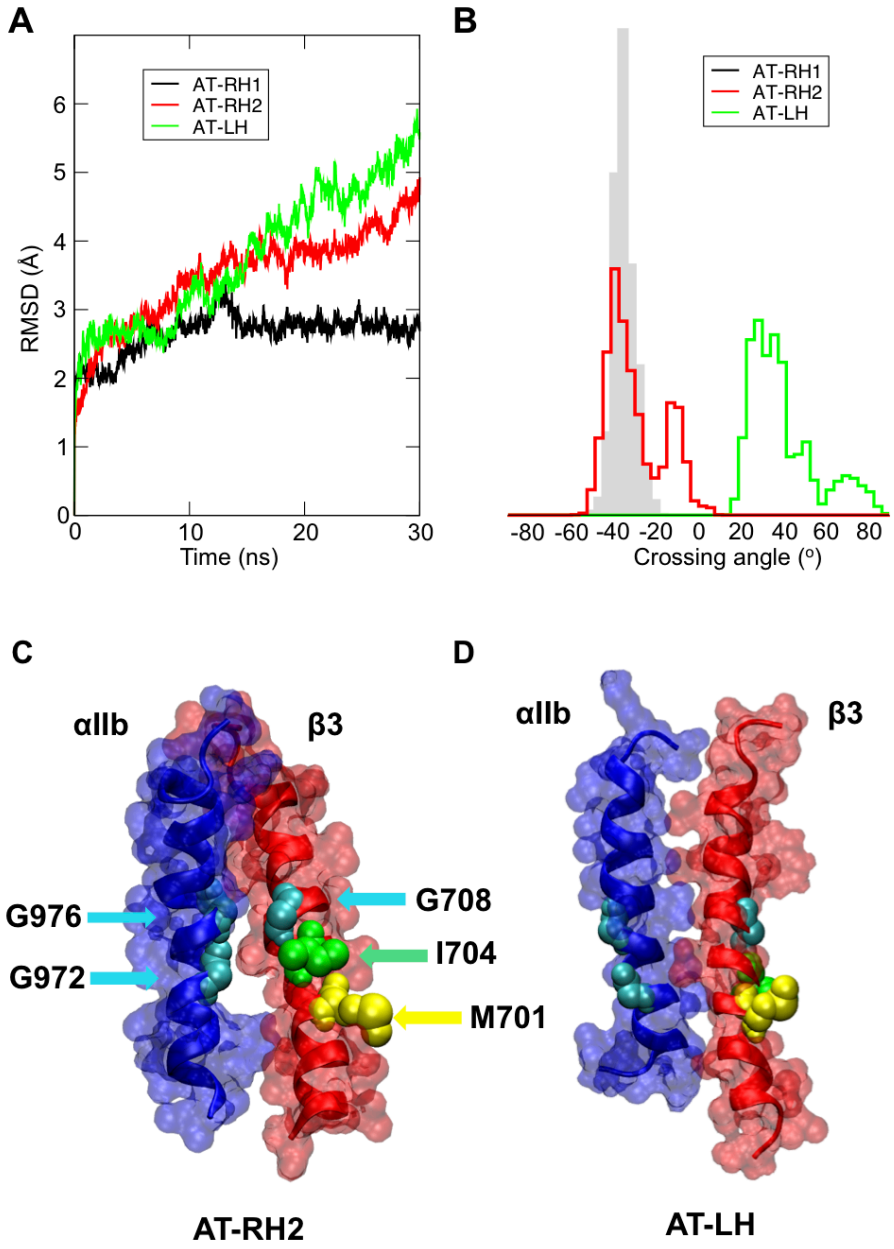


Figure 4.7: A,B. Cα RMSD averages (A) and crossing angle distributions (B) for the three repeat simulations of the RH1 mode (*AT-RH1*) and the minor binding modes (RH2 and LH). C,D. Final structures from one of the simulations of the minor binding modes (RH2 and LH). The key residues identified to be important in the packing of the helices in the NMR structure are shown in cyan for αIIbG972, αIIbG976 and β3G708, in yellow for β3M701 and in green for β3I704.

In the simulations which were initiated with the LH and the RH2 structures, the Cα RMSD rapidly increased to very high values for all simulations, suggesting a

conformationally unstable dimer (Fig 4.7A). The crossing angle distribution was considerably wider compared to the distribution of the RH1 binding mode, ranging from -60° to 0° for the RH2 position and from +10° to +60° for the LH position (Fig. 4.7B). This indicates that disruption of the close packing in the OMC region of the dimer resulted in a "looser" packing (Fig. 4.7C,D) of the helices, which is in good agreement with mutagenesis data (159, 173, 385).

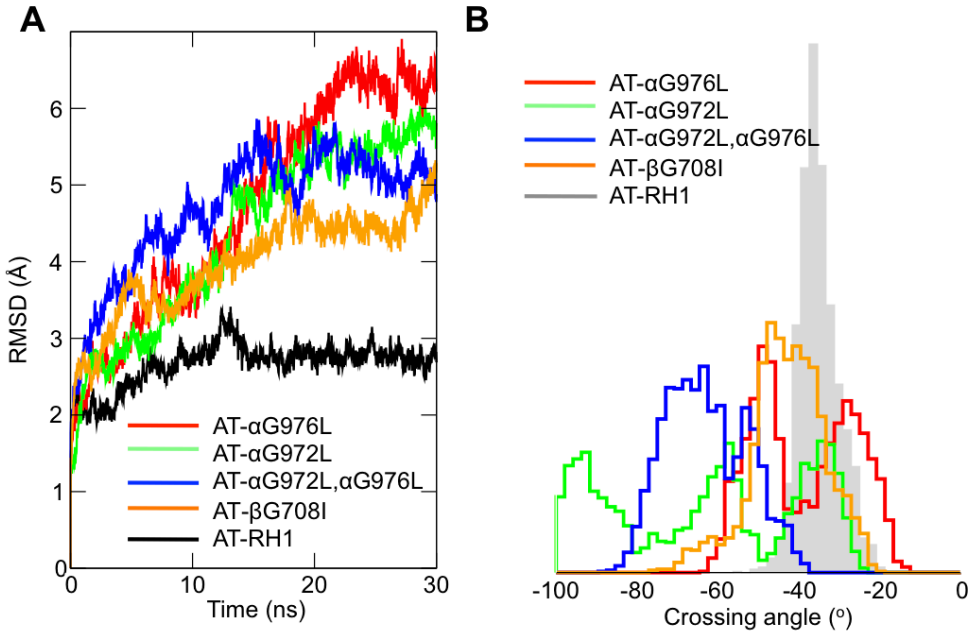


Figure 4.8: A,B. Cα RMSD averages (A) and crossing angle distributions (B) for the three repeat simulations of the RH1 mode (*AT-RH1*) and the simulations with the αIIb/β3 mutant forms (see Table 4.1 for more information).

4.3.2 αIIb/β3 mutant forms

Similar helix–helix self-assembly CG-MD simulations were performed, with mutations in the αIIb Gx_3G motif and the β3 G708 residue which had been shown experimentally to be disruptive. The aim was to investigate the effect of experimentally proposed mutations on helix packing. It also tested the ability of CG-MD simulations to detect disruptive mutations. The simulations are summarized in Table 4.1.

In all CG-MD simulations, the two helices diffused in the bilayer and after a few hundred nanoseconds formed dimers with perturbed packing modes compared to the WT helices. The relative location of the two helices in the bilayer revealed a favoured position for the αIIb helix in the *CG-αG976L* and *CG-αG972L,G976L* simulations. However, in both cases this position was different from the RH1 mode observed above. In the *CG-αG972L* and the *CG-βG708I* simulations, the mobile helix was found on different sides of

the fixed β3 helix, again showing that the RH1 binding mode was disfavoured. Closer examination of the helix–helix interactions in the different packing modes revealed that the αIIb Gx_3G motif and the β3G708 residue were no longer in the dimer interface expected from the αIIβ/β3 NMR structures (118, 119). As a result, calculation of the Cα RMSD between the representative structure for each system and the NMR structure (PDB:2K9J) revealed very high values, thus suggesting a "looser" packing of the two helices. In all simulations, the tilt angle of the αIIb helix was ~15° and of the β helix ~28°.

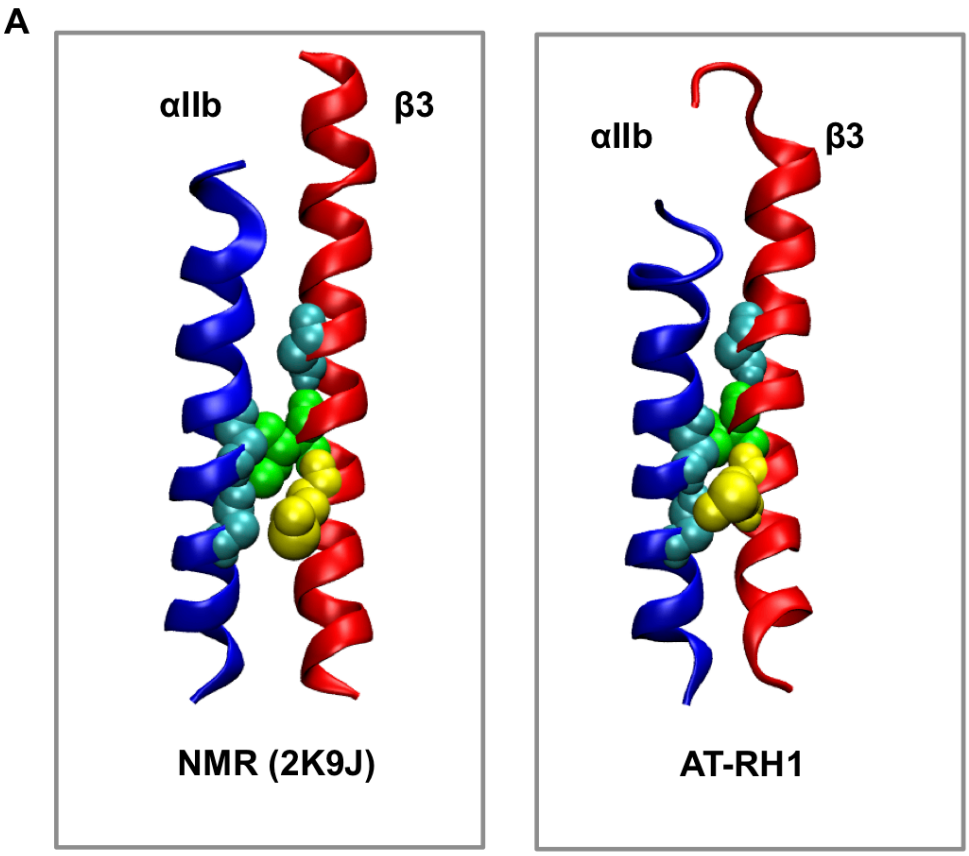


Figure 4.9: Comparison of the final structure from one of the *AT-RH1* simulations with the NMR structure. The key residues identified to be important in the packing of the helices in the NMR structure are shown in cyan for αIIbG972, αIIbG976 and β3G708, in yellow for β3M701 and in green for β3I704.

A structure from a favoured position in the spatial distribution analysis for each system was converted to AT representation and AT-MD simulations were performed for 30 ns (see Table 4.1). For all systems the Cα RMSD of the mutant dimer increased to very high values (5-7 Å) suggesting that mutations in the Gx_3G motif resulted in conformationally unstable dimers (Fig. 4.8A). The crossing angle distributions for all systems were also significantly wider compared to the simulations with the WT dimer,

ranging from -90^{o} to 0^{o} (Fig. 4.8B). This highlights the crucial role of the Gx_3G in regulating the crossing angle of a TM dimer. Finally, analysis of the residues that formed contacts between the two helices revealed a perturbed αIIb/β3 interface, with the αIIb Gx_3G motif facing away from the β3 helix.

4.4 Discussion and conclusions

This chapter focused on studying the association of the integrin αIIb and β3 TM helices and in particular on how the Gx_3G motif, a common feature in the TM region of many signalling receptors, regulates interactions between the two helices. The results demonstrate that a purely computational multiscale approach with a high-throughput framework can reproduce the packing and the interfacial contacts of the integrin αIIb/β3 dimer. This in turn demonstrates the predictive capabilities of this multiscale approach and indicates that the same method can be used to model and characterize the packing and conformational stability of other TM structures.

The current study focused on the OMC region of integrins, which includes a common Gx_3G motif (118, 119). The presence of two glycines in the interface between the helices allows the bulky hydrophobic sidechains of the β3 residues M704 and I704 to pack tightly and the αIIb Leu980 sidechain to pack against the β3 glycine G708. The simulations with the WT sequences of the two helices reproduced the packing of these sidechains (Fig. 4.9). Mutation of either of the two glycines in the Gx_3G motif or the β3 glycine G708 resulted in a loosening of the packing of the TM region. This is in good agreement with experimental studies which demonstrated that disruption of the interaction in this motif alters the packing of the helices; this may be a key step in moving the integrin equilibrium towards an active state. In particular, DN-ToxRed assays using leucine and alanine mutations in different α and β integrin TM pairs revealed the crucial role of the motif containing small amino acids in the packing of the integrin TM helices (385). Furthermore, leucine scanning of a 23-residue TM integrin domain showed that mutations in αIIbG972, αIIbG976 and β3G708 resulted in a high level of integrin activation (173). Finally, TOXCAT assays confirmed the crucial role of the Gx_3G motif in the association of the two integrin helices (159). Besides the αIIb/β3 heterodimer, recent studies (147, 148) of the TM region of the αLβ2 integrin suggested similar packing of the helices in the OMC region. The packing in that case was regulated by a Gx_3G-like motif (i.e. SxxxG). Sequence alignment of the different integrin α subunits (see Chapter 1) revealed the presence of a small-xxx-small motif in the majority of integrin α subunits, suggesting that

this is a general property of integrins (148). Therefore, the results of this chapter are of particular biological significance because they can be generalized for a variety of integrins and can explain, in combination with the existing experimental and computational data, the role of the OMC region in the regulation of the integrin activation state.

Another interesting aspect arising from this study is that the integrin TM region in the simulations is more dynamic in nature than GpA. Simulation studies by Psachoulia *et al.* (161, 391) and experimental studies using the TOXCAT assay (392) suggested an exceptionally stable dimer for GpA compared to other TM dimers such as tyrosine kinases (392). During integrin activation, large changes in the packing of the integrin helices are expected (through separation or a scissoring movement) and thus one would expect a more flexible helix–helix interface and possibly the existence of alternative packing modes. It must be noted that alternative packing modes have been reported by Bocharov *et al.* for the EphA2 receptor tyrosine kinase (393) and other signalling receptors (373). This is an aspect of TM helix dimerization which is usually ignored because many of the assays used to detect such interactions (e.g. TOXCAT (157, 370, 392, 394)) may give an average over all possible modes of interaction.

The results described in this chapter are also interesting from a more methodological perspective because they demonstrate that a computational multiscale approach predicts the interfacial contacts and the packing of TM dimers. This approach combines the predictive nature of CG-MD simulations to assess the assembly of TM helices and their crossing angle and all-atom simulations to refine the resulting structures. The integrin TM region provides a good test case because it is a more complex heterodimer compared to the GpA homodimers (152, 264, 265), which have been used extensively to study the role of Gx_3G motif. In this study, use of a high-throughput method also ensured better sampling in the CG-MD simulations. Other approaches which combine modelling and data-based filtering or experimental constraints have also been used to model the integrin TM region and these also show relatively good agreement with the NMR structures (149, 150, 162). For example, early studies by Gottschalk *et al.* predicted glycophorin-like packing in the integrin Gx_3G motif using global search methods in combination with experimental restraints (149-151). There is, however, significant variation between their model structures in three different publications. Later studies by Luo *et al.* used cysteine scanning experimental restraints together with the Rosseta algorithm to produce a structure of the integrin TM region which could reproduce both the interactions in the OMC and IMC regions of integrins (120). Recently, Metcalf *et al.* used

MC and threading approaches, with additional mutational and cysteine-scanning restraints, to successfully predict the packing of the OMC region. Not all models were successful, however; for example, using a similar approach to that of Metcalf *et al.*, Partridge *et al.* predicted a LH packing of the dimer without the involvement of the Gx_3G motif in the interaction of the helices (158). Similar approaches have been used to model parts of more complex proteins such as GPCRs and Kv channels with some success (395). These approaches, which are usually much faster than MD simulations, have been useful in some cases but the predictions rely heavily on the quality and availability of experimental data or statistical information from known experimental structures. With more known structures becoming available, the accuracy of these methods should improve. These methods also select a structure using a scoring function of questionable validity. A refinement step similar to that described in this chapter might improve the accuracy of the structure by optimizing the packing/interactions of the helices and the sidechains in a lipid environment. It should be noted, however, that the force fields used in MD simulation approaches are parameterized using a defined set of experimental data and therefore their ability to reproduce certain properties of a biomolecular system should also be tested.

In summary, the results described in this chapter demonstrate that the predictive nature of CG-MD simulations in combination with all-atom simulations gives valuable information about the packing of the integrin TM region in a membrane-like environment. It was possible to predict a structure with similar accuracy to that obtained by NMR. This purely computational method can therefore be applied to other TM dimers and should provide structural insights that complement the experimental data. Additionally, the approach can provide valuable insight into the role of structural motifs in the biological function of proteins. The extension of the multiscale approach, for example with inclusion of QM/MM methods, would enable a more detailed investigation of the role of H-bonds in helix–helix interactions (396).

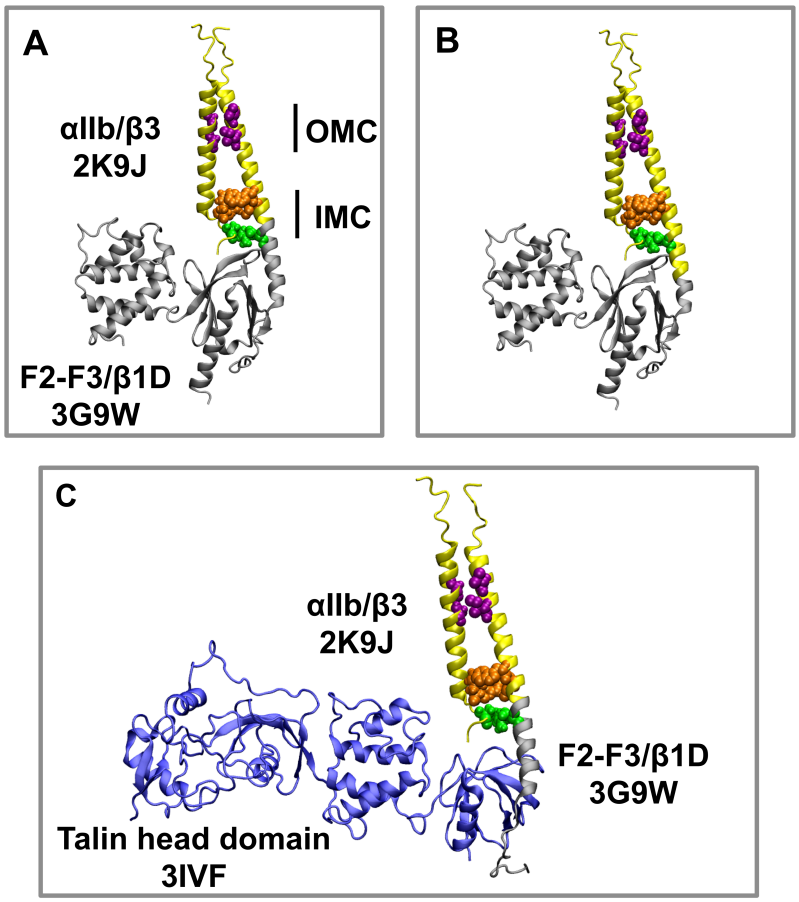
Chapter 5

Conformation and Dynamics of the Integrin αIIb/β3 Transmembrane Helix Dimer and the Role of Talin

5.1 Introduction

Having studied the interactions and the orientation of the talin head domain relative to the membrane (Chapter 3) and some aspects of integrin TM α-helices packing (Chapter 4), this chapter now focuses on understanding the effect of the talin head domain on the structure of the integrin TM region. There is good experimental evidence that talin binding to the integrin CT region moves the integrin conformational equilibrium towards an active state (see Chapter 1). It is believed that activation is associated with large conformational changes in the integrin ectodomain coupled to changes in TM helix packing. Three models for TM helix packing have been suggested to explain these changes; namely the "piston" (397), "scissor" (151) and "separation" models (37, 52, 112, 113, 158, 168)). In the "piston" model there is vertical movement of the two helices in the membrane. In the "scissor" model there is a large increase in the crossing angle of the helices in the TM region during activation. In the "separation" model there is a complete dissociation of the two helices upon talin binding, which in turn separates the integrin ectodomain legs, thus moving the integrin ectodomain towards an extended high-affinity state. The existence of three different models indicates that the details of talin-mediated changes in the TM region remain unclear. Elucidation of how talin regulates the rearrangement of integrin TM helices during activation would thus be a valuable addition to the integrin field.

In this chapter, models of the talin head domain in complex with the β and the α/β integrin TM regions were constructed using recent structural data (118, 119, 125, 186), and a similar multiscale simulation approach to that described previously was used to probe the changes which occur in the integrin TM region upon talin binding. A mechanism is proposed for the talin-modulated rearrangement of the integrin TM region that promotes integrin activation.



D

β3 685 PESPKG PDILVVLLSV MGAILLIGLA ALLIWKLLIT IHDRKEFAKF EEERA 735
β1D 721 PESPKG PDILVVLLSV MGAILLIGLA ALLIWKLLMI IHDRREFAKF EKEKM 771
αIIb 957 GALEER AIPIWWVLVG VLGGLLLLTI LVLAMWKVGF FKRNRP 998

β3 736 RAKWDTANNP LYKEATSTFT NITYRGT 762
β1D 772 NAKWDTQENPI YKSPINNFKN PNYGRKAGL 801

Figure 5.1: A,B. Models of the F2-F3/αβ complex. In A, the β chimeric peptide was comprised of β3 residues 688-719/β1D residues 753-787, whereas in B it was comprised of β3 688-727 residues/β1D 764-787 residues. C. Model of the talin/αβ complex. In this case, the β chimeric peptide was comprised of β3 residues 688-719/β1D residues 753-787. The part of the structure corresponding to the αIIb/β3 structure (PBD:2K9J) is shown in yellow, the F2-F3/β1D (PDB:3G9W) is in grey and the talin head domain (PDB:3VIF) is in blue. D. Sequences of the αIIb, β3 and β1D integrin tails.

5.2 Structures used in the simulations

In the simulations of the integrin TM region (Table 5.1), the NMR structure of the αIIb/β3 integrin TM dimer (PBD:2K9J) was used. This structure is thought to represent the integrin inactive low affinity state (118). Modeller (341, 342) was used to perform all the mutations in the αIIb helix (R995D, G972S, G976L) and in the β3 helix (F992A, F993A) starting with coordinates PBD:2K9J. The F2-F3/β, the F3/β and the F3-F3/αβ complexes were constructed using the crystal structure of the F2-F3/β1D complex (PDB:3G9W) (125)

and the αIIb/β3 NMR structure (PBD:2K9J) (Fig. 5.1A,B) (118). See Appendix C.1 for a detailed discussion on how the complex was constructed and evaluated. The talin complete head domain/αβ complex was constructed in the same fashion by using talin head domain configurations obtained from the simulations in Chapter 3 (PDB:3IVF) (Fig. 5.1C).

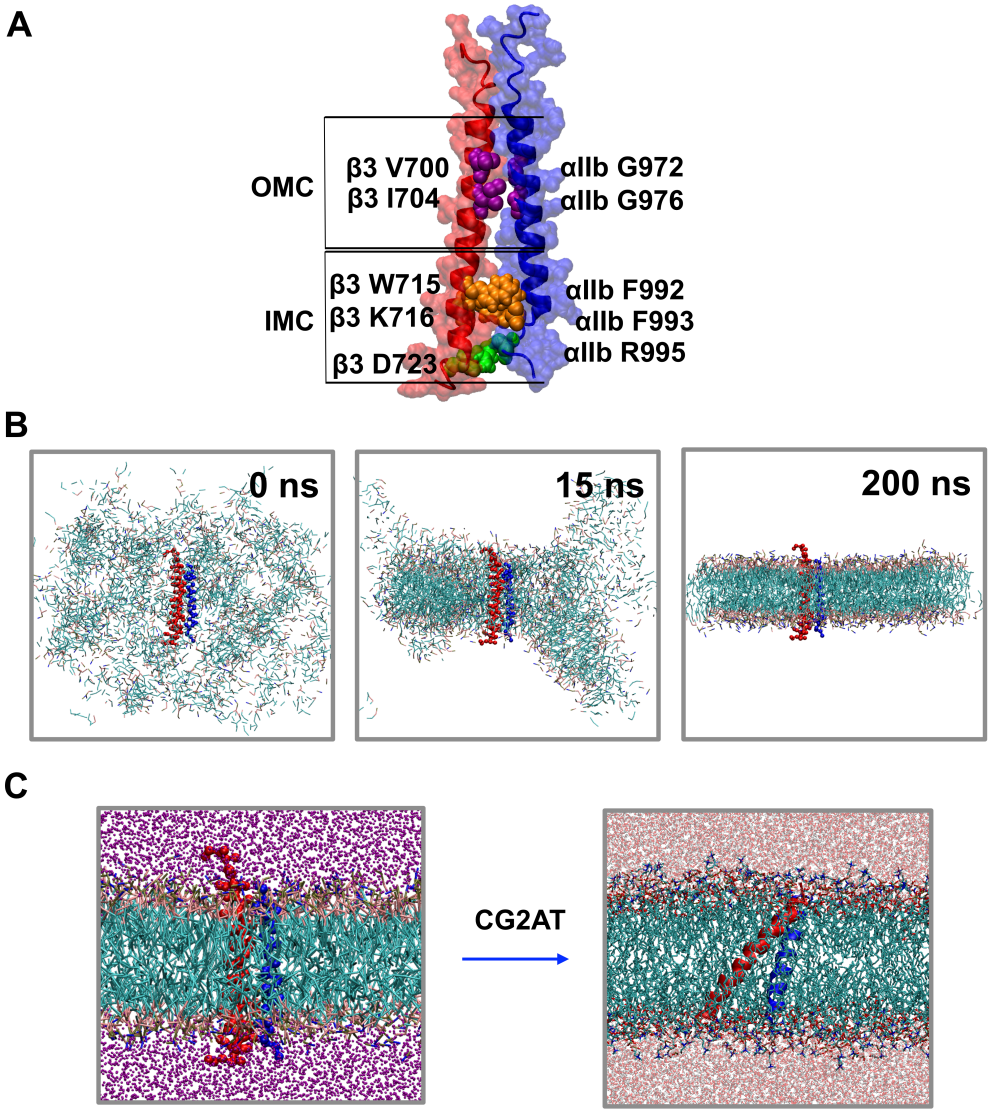


Figure 5.2: A. Structure of the αIIb/β3 dimer. The α subunit is shown in blue and the β subunit in red. The key interactions in the OMC region (i.e. Gx_3G) are shown in magenta and those in the IMC region are shown in orange (i.e. F992F993) and green (i.e. αIIbR995-β3D723 salt bridge). B. Progress of the self-assembly CG-MD simulation with the αIIb/β3 dimer. For the α and β helices, the same colour code as in A is used. The lipids are shown in bonds format with the hydrophobic tails in cyan, glycerol in pink, phosphate in brown and choline in blue. Waters and counter ions are omitted for clarity. C. AT and CG representation of the system with the αIIb/β3 dimer inserted in the POPC/POPG bilayer.

Table 5.1: Summary of simulations

Simulation	Proteins	Coordinates	Duration (ns)
αβ1-AT	αIIb/β3 TM	2K9J	3x100
αβ2-AT	αIIb(R995D)/β3 TM		3x100
αβ3-AT	αIIb(F992A,F993A)/β3 TM		3x100
αβ4-AT	αIIb(F992A,F993A,Δ995-998)/β3 TM		3x100
αβ5-AT	αIIb(G972S)/β3 TM		1x100
αβ6-AT	αIIb(G976L)/β3 TM		1x100
β-AT	β* TM	model based on	3x70
β-F3-AT	β* TM/talin F3	2K9J β3 + 3G9W	3x70
β-F2F3-AT	β* TM/talin F2-F3		3x70
β-F2F3_4E-AT	β* TM/talin F2-F3(4E)		3x70
β-F2F3-CG	β* TM/talin F2-F3		1x500

β* indicates a chimeric β chain (β3/β1D) – see methods for details).

5.3 Results

5.3.1 Conformation and dynamics of the integrin αIIb/β3 helix dimer

In the first set of simulations, the αIIb/β3 NMR structure (PDB:2K9J) (118) was converted to CG representation and a self-assembly CG-MD simulation was performed for 200 ns (Fig. 5.2). In this simulation, the αIIb TM helix was restricted to an orientation parallel to the z-axis (at the centre of the simulation box) to satisfy previous experimental observations (118, 119, 131, 132) and thus the β3 TM helix adopted a tilted position of ~30^{o}. The CG POPC and POPG lipids (at a ratio of 3:2), with CG waters and ions, were randomly placed in the box. After ~25 ns of simulation, the bilayer was formed around the αIIb/β3 dimer (Fig. 5.2B). After the self-assembly simulation, the N-terminal residues αIIb965-969/β3692-696 and the C-terminal residues αIIb988-994/β3715-723 were found close to the lipid headgroups, in good agreement with experimental data (140). The resultant CG system was then converted to an AT representation to provide the initial system for the unrestrained AT-MD simulations with the αIIb/β3 wild type dimer (249) (see Fig. 5.2C and Table 5.1 for more information).

During the AT-MD simulations the dimer was conformationally stable, retaining its RH packing with a crossing angle of ~-35^{o} (Fig. 5.3). As a result, the αIIb/β3 interface identified by the simulations was in good agreement with the interface proposed by the NMR structure (118) and previous computational work (120, 162) (Fig. 5.3A). The tilt angles for the αIIb and the β3 helices (relative to the bilayer normal) were ~10^{o} and ~30^{o} respectively (Fig 5.3D). Calculation of the interactions between the αIIb/β3 dimer and the

lipid headgroups revealed that αIIb R997, K989, W988, K994 and β3 K716, T270, H722 and R724 residues in the C-terminal and αIIb G957-R962 and β3 P685-K689 residues in the N-terminal formed the highest number of contacts with the lipid headgroups.

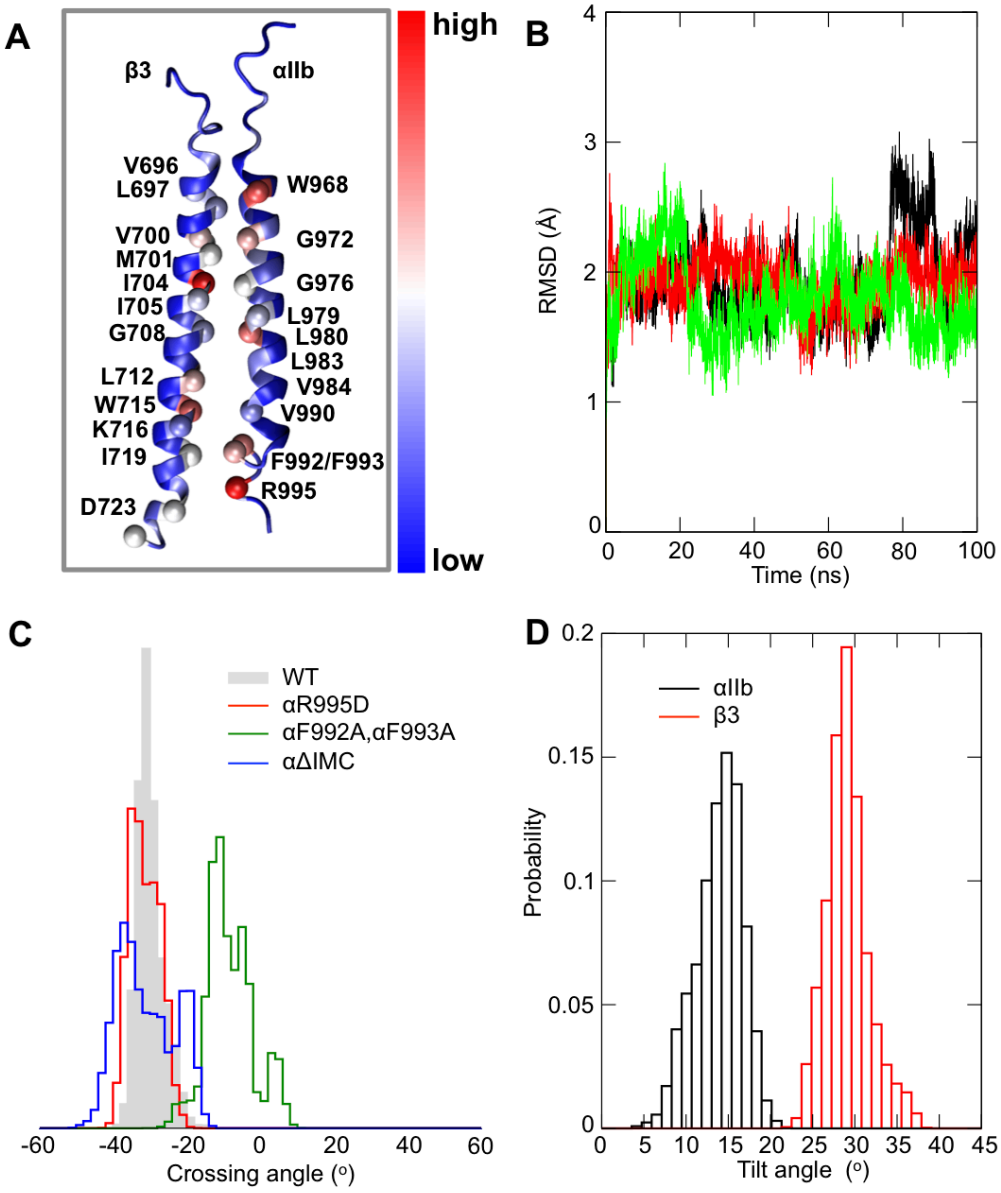


Figure 5.3: A. Normalized average number of contacts (using a cut-off distance of 4 Å) between the αIIb and β3 TM regions (over the last 10 ns of the *αβ1-AT* simulations). The normalized average number of protein–protein contacts was mapped on the αIIb/β3 NMR structure. Blue represents a low number of contacts and red represents a high number of contacts. Backbone atoms of residues which made the largest number of contacts are shown as spheres. B. Cα RMSD of the αIIb/β3 structure over time for the three *αβ1-AT* simulations (shown in different colours). C. Crossing angle distributions for the *αβ1-AT* (grey), *αβ2-AT* (red), *αβ3-AT* (green) and *αβ4-AT* (blue) simulations (see Table 5.1 for more information). The negative crossing angle corresponds to RH packing of the helices and the positive corresponds to the LH packing of the helices. D. Tilt angle distribution of the α (black) and β (red) integrin helices in the *αβ1-AT* simulation.

Similar AT-MD simulations were also set up to further explore the effect of the key interactions in the IMC region on the stability of the αIIb/β3 dimer. To this end, the αIIbR995 residue in the αIIbR995-β3D723 salt bridge was mutated to aspartic acid (D) and again three AT-MD simulations were performed for 100 ns each (*αβ2-AT* in Table 5.1). The αIIbR995D mutation resulted in the disruption of the αIIbR995-D723 salt bridge after ~20 ns of simulation; however, the rest of the interfacial contacts of the dimer remained intact, as the hydrophobic interactions of the αIIbF992F993 residue pair with the β3 helix and the packing in the OMC region were very stable (Fig. 5.4).

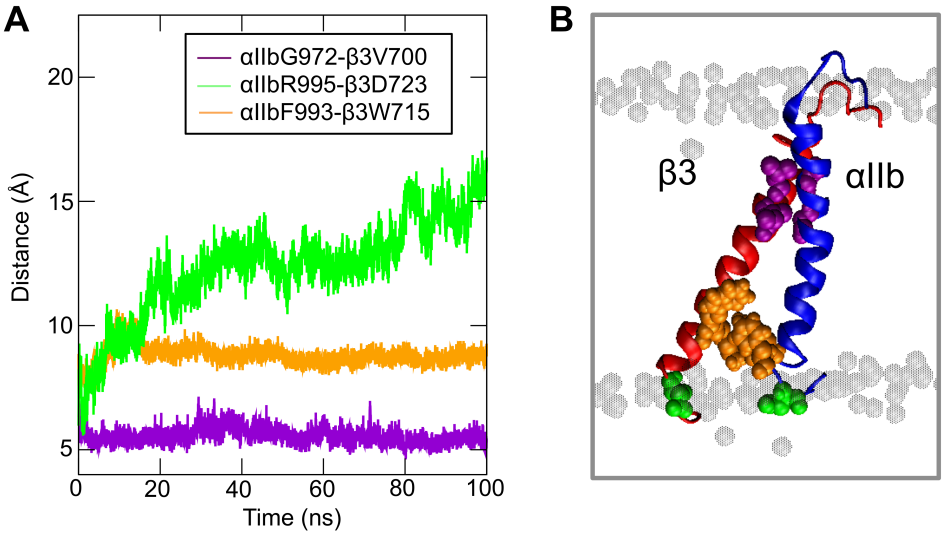


Figure 5.4: A. Separation between the αIIbG972-β3V700 pair (magenta), the αIIbR995-β3D723 pair (green) and the αIIbF993-β3W715 pair (orange) as a function of time for the *αβ2-AT* simulation. B. Final snapshot (i.e. at 100 ns) from one of the *αβ2-AT* simulations. The key interactions in the OMC region are shown in magenta and in the IMC region in orange (F992F993) and green (αIIbR995-β3D723 salt bridge). The lipid phosphate groups are shown in grey.

To test the contribution of the two aromatic residues (i.e. αIIbF992F993) on the stability of the αIIb/β3 dimer, a double mutation (to alanine) was made in this aromatic cluster (*αβ3-AT* simulation) (118). This double mutation significantly perturbed the packing of the two helices during the simulations. In particular, the β3 tilt angle was retained at ~30^{o}, which is similar to the *αβ1-AT* and *αβ2-AT* simulations but the tilt angle of the αIIb helix increased to ~25^{o} (Fig. 5.5B), resulting in a reduction of the crossing angle between the two helices to ~-10^{o} (Fig. 5.3C). The new position of the αIIb helix created a modified αIIb/β3 interface where the β3 residues M701 and L705 in the OMC region and the residues L712 and K716 in the IMC region which were in the interface of the two helices in the NMR structure no longer formed contacts with the αIIb helix. The

αIIbR995-β3D723 salt bridge remained intact during the simulations, thus preventing further separation of the two helices (Fig. 5.5A).

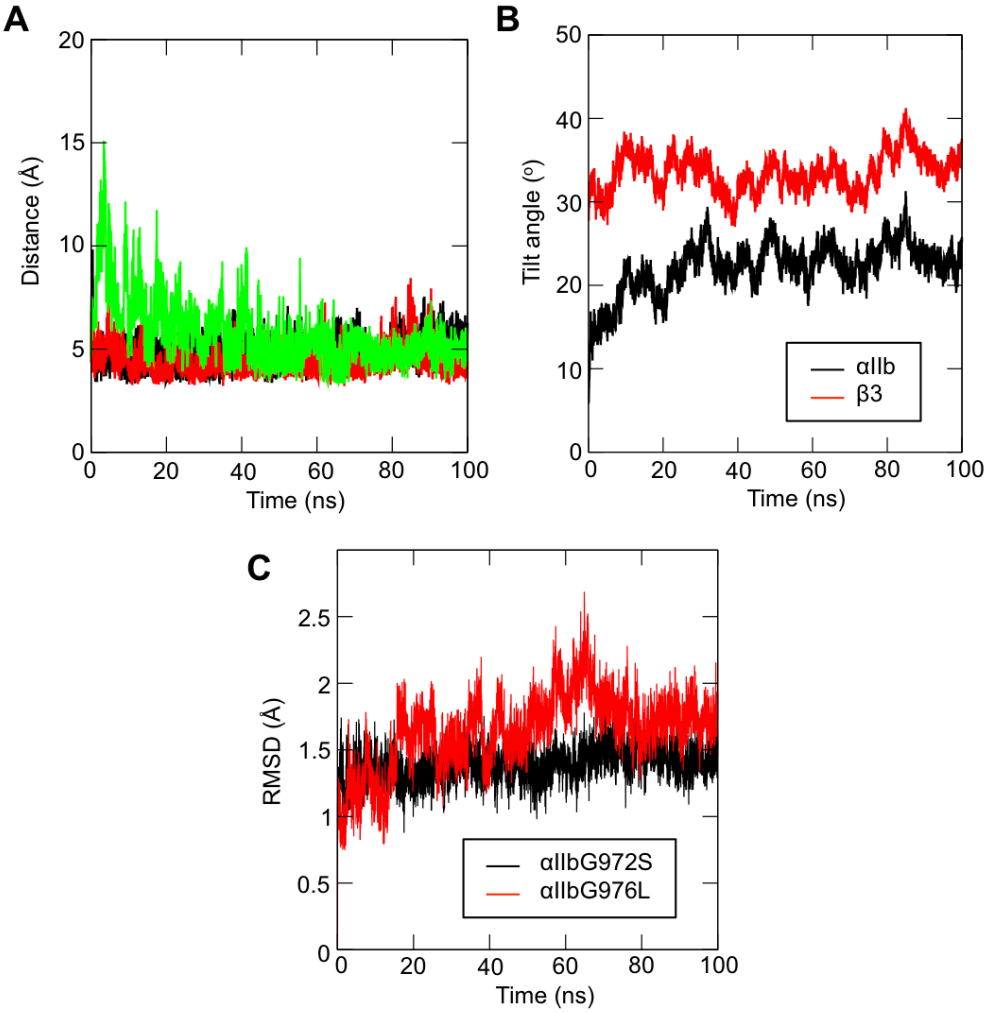


Figure 5.5: A. Separation between the αIIbR995 and β3D723 residues as a function of time for the *αβ3-AT* simulation. The three simulations are shown in different colours. B. Tilt angle relative to the bilayer normal as a function of time for the αIIb (black) and the β3 (red) subunits in the *αβ3-AT* simulation. C. Cα RMSD over time for the *αβ5-AT* (black) and *αβ6-AT* (red) simulations.

Additional simulations were performed where the αIIb residues from R995 to K998 were cleaved from the initial structure and, at the same time, the αIIbF992F993 residues were mutated to alanines (*αβ4-AT* in Table 5.1). In this case, the separation between the αIIbA992 and αIIbA993 residues and the β3 helix increased. This resulted in a wider distribution of crossing angles, which suggested looser packing of the two helices (Fig. 5.3C). However, disruption of the packing in the OMC region was observed in only one of three simulations, suggesting that disruption of the IMC region alone is not sufficient to

disrupt all the αIIb/β3 TM interactions. Note that the helical structure of the αIIb and the β3 subunits was retained during the simulation.

Experimental studies suggested that mutation of the G972 residue to serine or the G976 residue to leucine resulted in integrin activation by disrupting the close packing of the two helices (118, 173). To examine the effect of these two mutations on the αIIb/β3 structure, AT-MD simulations with the mutated dimers were also run (*αβ5-AT* and *αβ6-AT* in Table 5.1). During the simulations, both dimers were conformationally stable (Fig. 5.5C). As a result, the interface between the two dimers and the crossing angle distribution were similar to the *αβ1-AT* simulation (above). It must be noted, however, that in the simulations described in Chapter 4 mutations in the Gx_3G motif do indeed perturb the packing of the dimer, resulting in unstable dimers.

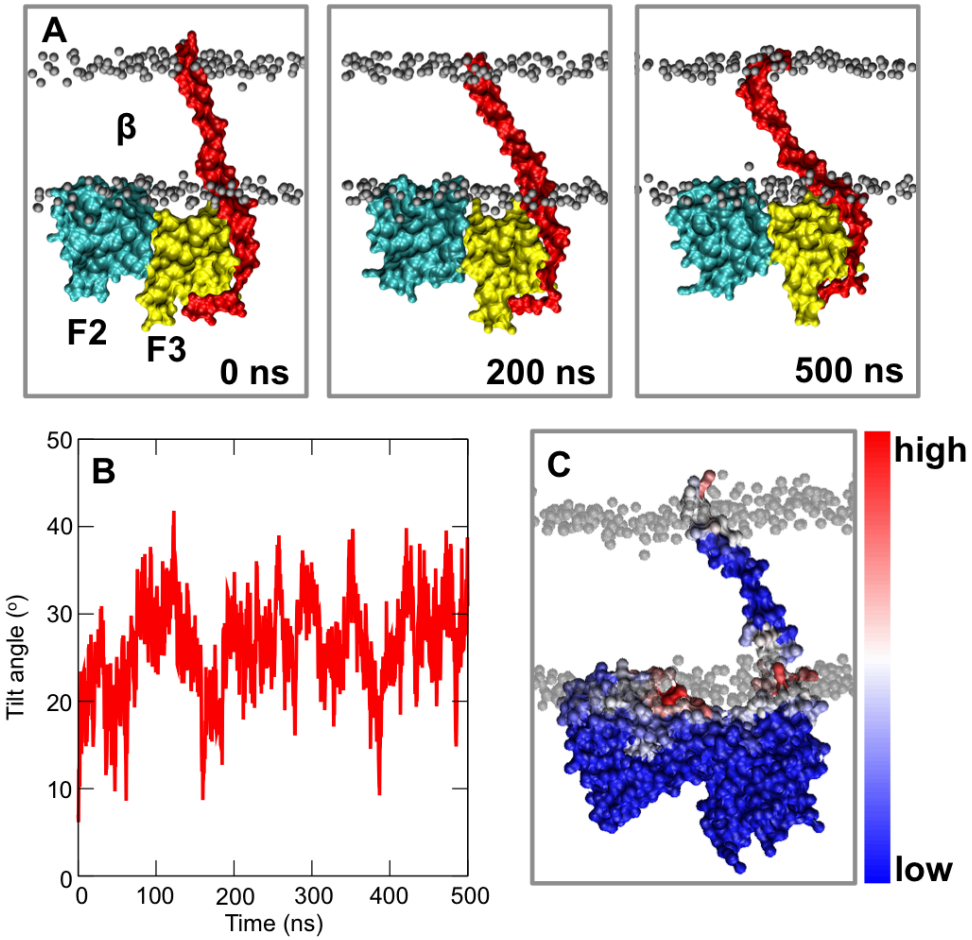


Figure 5.6: A. Snapshots (at 0 ns, 200 ns and 500 ns) from the CG-MD simulation with the F2-F3/β complex. The F2 domain is shown in cyan, the F3 domain in yellow and the β tail in red. The lipid phosphate atoms are shown in grey. B. Integrin β tail tilt angle relative to the bilayer normal as a function of time. C. Normalized average number of protein/lipid contacts for the *β-F2F3-CG* simulation mapped onto the last snapshot of the simulation. Blue represents low/no contacts, white medium contacts and red large contacts.

5.3.2 Interactions with the talin F2-F3 domain pair modulate the integrin β subunit orientation

Subsequently, modulation of the orientation of the β integrin tail in the membrane by the talin F2-F3 domain pair was investigated. For this study, the model of the F2-F3/β complex (see methods) was inserted in a preformed POPC/POPG bilayer at a ratio of 3:2 and a CG-MD simulation for 500 ns was performed (Fig. 5.6A). In the initial frame of the simulation, the complex was positioned in such a way that the F2 positively charged patch (identified in Chapter 3) interacted with the lipid headgroups. This positioned the β integrin tail with a tilt angle of ~10^{o} relative to the bilayer normal. During the simulation, there was an increase in the β tail tilt angle by ~20^{o} to a final value of ~35^{o} (Fig. 5.6A,B). Analysis of the F2-F3/lipid interactions revealed that residues L262, D263, E266, D275, R276, K277, R281 and K319-M322 and the interface lysine and arginine residues in the talin F2 domain, which were mutated in the NMR experiments (Chapter 3), formed a significant number of contacts with the lipids (Fig. 5.6C). Calculation of the number of contacts between the protein and the POPC and POPG lipids separately revealed a threefold higher number of contacts with negatively charged POPG lipids suggesting that the F2-F3 binding surface was enriched in anionic lipids.

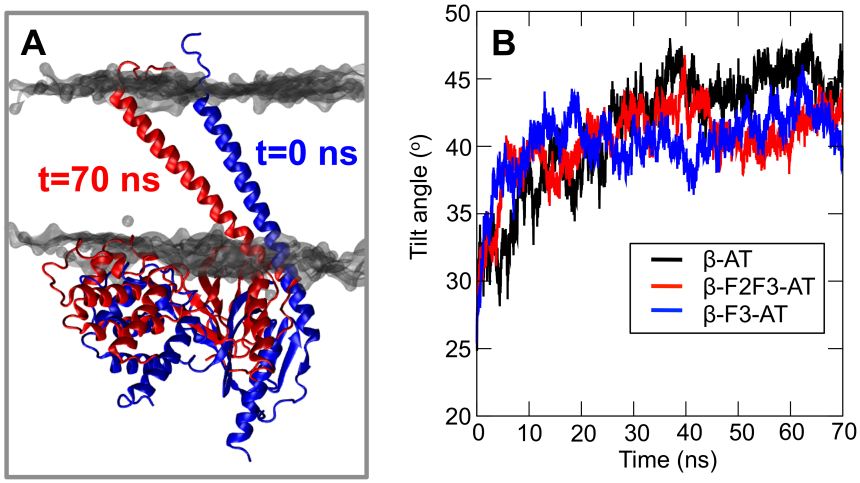


Figure 5.7: A. Initial (t=0 ns; blue) and final (t=70 ns; red) snapshots from one of the *β-F2-F3-AT* simulations. B. Average (over three simulations) β helix tilt angle relative to the bilayer normal as a function of time for the *β-AT* (black), *β-F2F3-AT* (red) and *β-F3-AT* (blue) simulations.

Subsequently, the final snapshot from the CG-MD simulation was converted to an AT representation and three simulations were performed (*β-F2F3-AT*). During the 70 ns simulation, a large increase (~15^{o}) in the tilt angle of the β TM helix relative to the bilayer

normal was observed (Fig. 5.7A,B). Calculation of the relative movements of the F2 and the F3 domains during the simulation suggested that the F2 subdomain rotated 10±4° relative to the F3 domain in a plane perpendicular to the bilayer normal. In contrast, the same analysis revealed no significant displacement between the F3 domain and the β helix IMC region. During the simulations, the F3 subdomain regions F318-P330 and T357-T385 interacted with the β tail MProx region and the irregular region from A773 to I787 respectively (Fig 5.9A). In particular, β tail residues I782, F756, A773, P781, Y783 and W775 made frequent contacts with the F3 domain. Calculation of the H-bonds and close contacts in the last 5 ns of the simulations revealed that F2-F3 interacted with the bilayer with the same pattern of interacting residues as described in the CG-MD simulation above (Fig. 5.8).

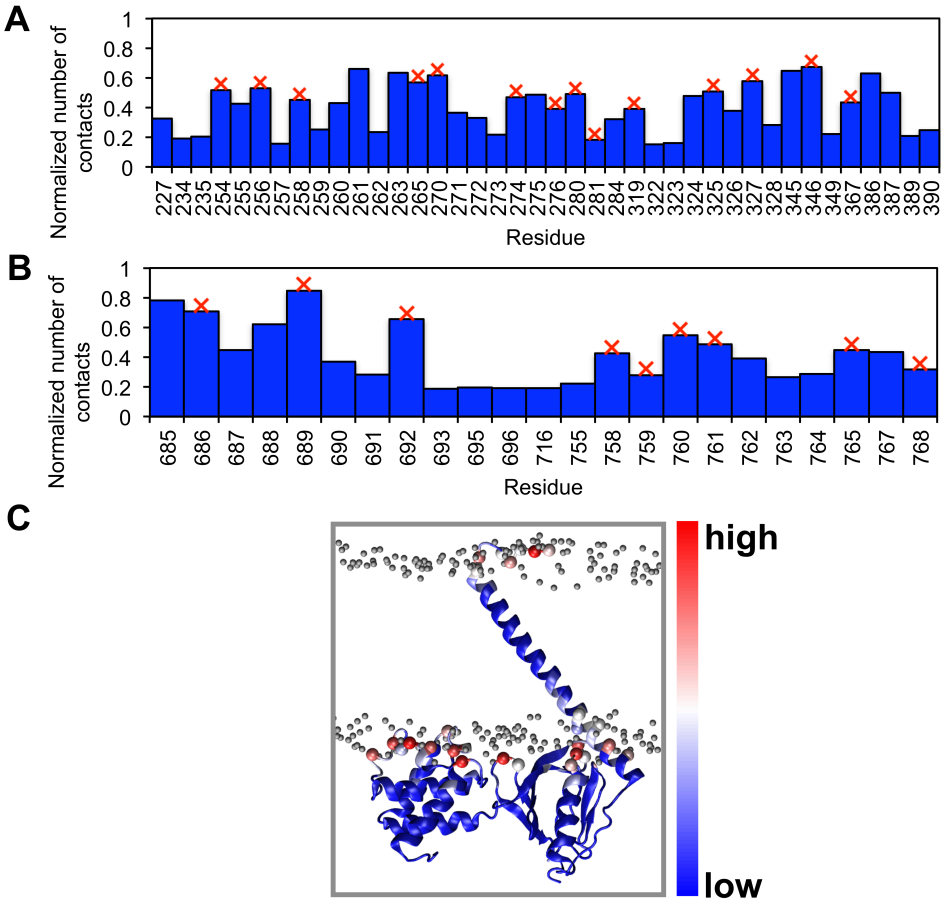


Figure 5.8: A,B. Normalized average number of contacts as a function of residue number for (A) the F2-F3 domain pair and (B) the β subunit and lipid headgroups in the *β-F2F3-AT* simulation. Residues which made less than 10% of the contacts are omitted for clarity. Residues which made more than 10% of the H-bonds are marked with crosses. C. The normalized average number of protein/lipid contacts was mapped onto the last frame of the *β-F2F3-AT* simulation and is shown on a colour scale from blue (low number of contacts) through to red (high number of contacts). The Cα atoms of the residues which made the largest number of contacts are shown as spheres.

Figure 5.8 demonstrates the interactions between the β helix and the lipid headgroups. In the outer part of the membrane, residues V696-P685 maintained their contacts with lipid headgroups throughout the simulation. In the CT region of the bilayer, the β W715 and the K716 sidechains pointed towards the phosphate groups of the lipids. H758, R760 and R761 sidechains lay in the headgroup/tail interface while E762, P763, K765 and K768 sidechains were immersed in the lipid headgroups. Again, there was a clear preference for the β tail to interact with acidic POPG lipids. This preference for the POPG lipids was a result of interactions with the positively charged regions in the talin F2-F3 pair, which were situated close to the integrin TM region (i.e. they interacted with the TM domain) and the cluster of positively charged residues in the integrin TM region.

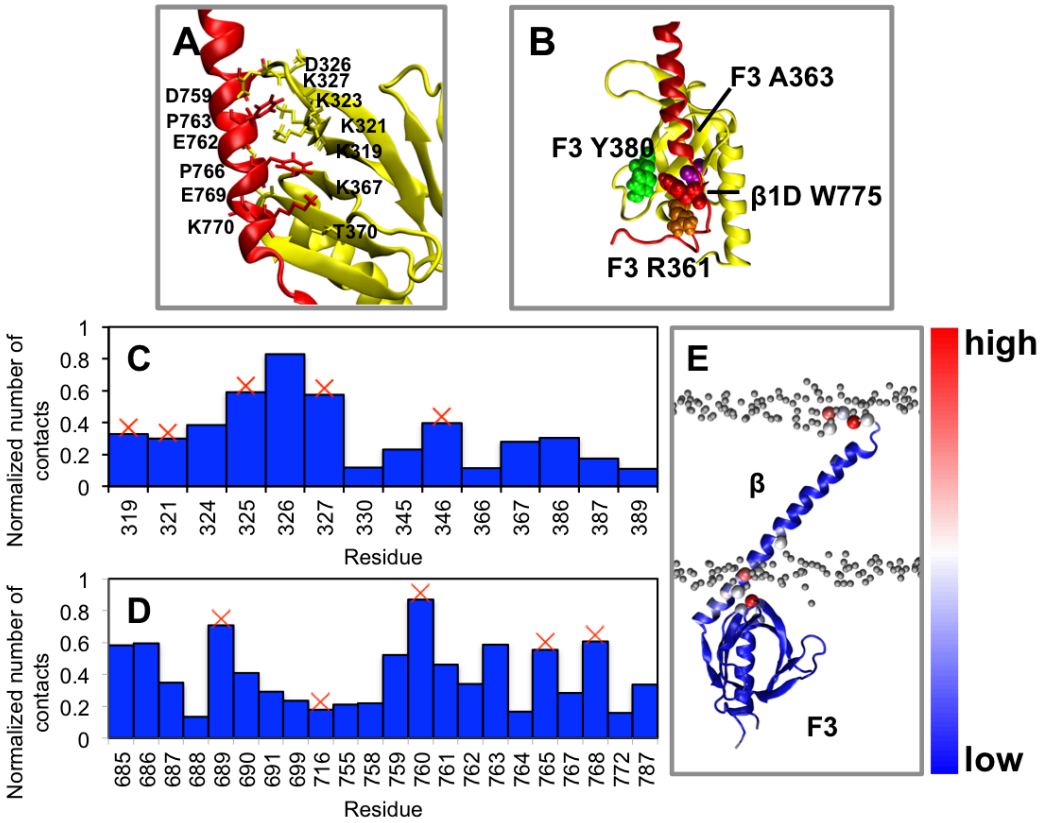


Figure 5.9: A. The F3/β tail interface in the MProx region (*β-F2F3-AT* simulation). Residues which are involved in the interaction are highlighted. B. The β1D residue W775 (β3 W739) is in a hydrophobic pocket formed by F3 residues R361, A363 and Y380 (R358, A360 and Y377 in talin-1) in the *β-F2F3-AT* simulation. C,D. Normalized average number of contacts as a function of residue number for the F3 subdomain (C) and the β subunit (D) and lipid headgroups in the *β-F3-AT* simulation. Residues which made less than 10% of the contacts are omitted for clarity. Residues which made more than 10% of the H-bonds are marked with a cross. E. Snapshot corresponding to the last frame of the *β-F3-AT* simulation. The protein/lipid contacts (normalized average number) were mapped on the structure on a colour scale from blue (low number of contacts) through to red (high number of contacts). Residues which made the largest number of contacts are shown in spheres.

Comparable simulations with a mutated form of the F2-F3 domain pair in complex with the β integrin tail (i.e. K258E, K274E, R276E, and K280E) were also performed. During these simulations, the β helix adopted a similar tilt angle (i.e. ~40^{o}) in the bilayer to the one observed in the *β-F2F3-AT* simulation described above. No significant perturbation of the mutated F2-F3/lipid interactions was observed. This stable F2-F3/lipid interface implies that the mutation of four positive residues in the F2 surface is not sufficient to disrupt the interactions of the F2-F3 with the lipids after the binding of the F2-F3 because other positive residues on the same surface retained their interactions with the lipids. Therefore, this 4E mutation is only able to alter the binding of talin to the membrane prior to the association with the β integrin tail as demonstrated in Chapter 3.

To explore how the isolated F3 domain modulates the β tail TM region orientation, the F2 domain was removed from the AT system above and three simulations with the F3/β complex were performed. An increase of ~15^{o} in the tilt angle of the β tail TM region relative to the bilayer normal was observed (Fig. 5.7B). The F3/β and the F3/bilayer interaction surfaces were similar to the *β-F2F3-AT* simulation (Fig. 5.9C,D,E). The β TM helix also adopted a similar orientation when the F2-F3 domain pair was removed and simulations with the isolated β TM helix alone in a bilayer were performed (Fig. 5.7B).

Table 5.2: Summary of simulations

Simulation	Proteins	Coordinates	Duration (ns)
αβ-F2F3-SACG	αIIb/β* TM/talin F2-F3	model based on	1x200
αβ-F2F34E-SACG	αIIb/β* TM/talin F2-F3-4E	2K9J + 3G9W	10x200
αβ-F2F3-CG	αIIb/β* TM/talin F2-F3		5x5000
αβ-F2F3c-AT	αIIb/β* TM/talin F2-F3 closed		3x100
αβ-F2F3p-AT	αIIb/β* TM/talin F2-F3 partial		3x100
αβ-F2F3o-AT	αIIb/β* TM/talin F2-F3 open		3x100
αβ-tal-CG	αIIb/β* TM/talin (ΔF1 loop)	model based on	5x4000
αβ-talh2-CG	αIIb/β* TM/talin (loop-res: 134-172/helix-res: 154-167)	2K9J + 3G9W+ 3IVF	5x4000
αβ-tal-AT	αIIb/β* TM/talin (ΔF1 loop)		1x100
αβ-talh2p-AT	αIIb/β* TM/talin (loop-res: 134-172/helix-res: 154-167) partial		1x100
αβ-talh2o-AT	αIIb/β* TM/talin (loop-res: 134-172/helix-res: 154-167) open		3x100

TM = transmembrane helix. β* indicates a chimeric β chain (β3/β1D) – see methods for details).

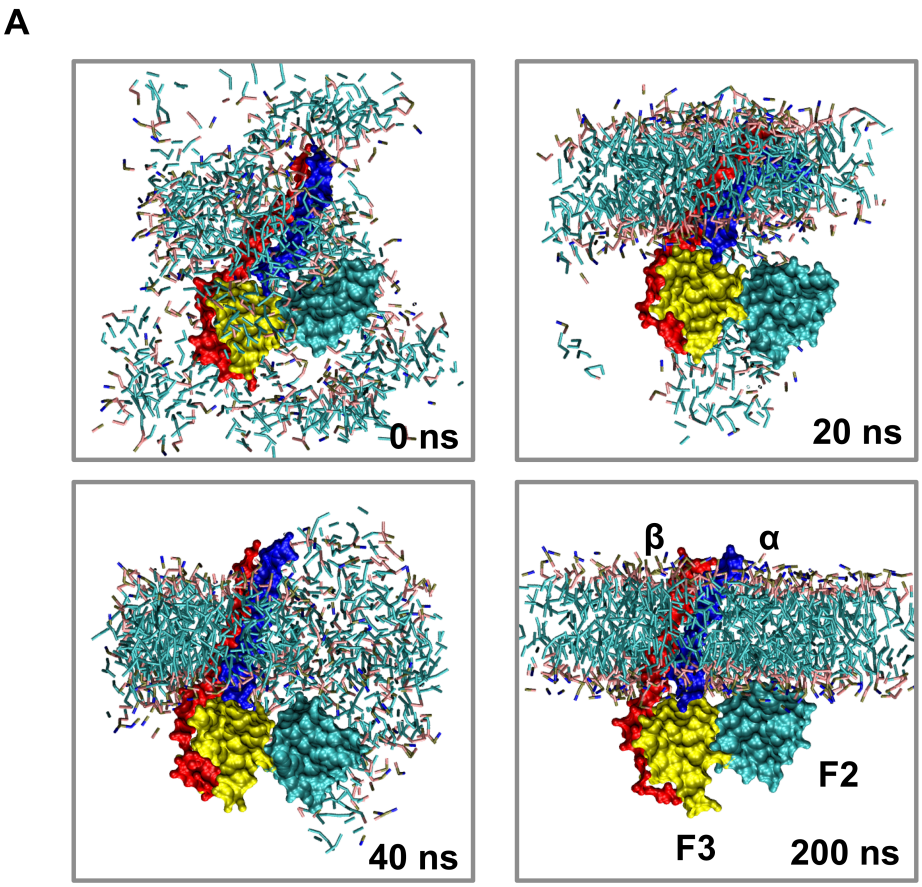


Figure 5.10: A. Progress (at 0 ns, 20 ns, 40 ns and 200 ns) of the *αβ-F2F3-SACG* simulation. The F2 domain is shown in cyan, the F3 domain in yellow, the β tail in red and the α tail in blue. The lipids are shown in bonds format with the hydrophobic tails in cyan, glycerol in pink, phosphate in brown and choline in blue. Waters and counter ions are omitted for clarity.

5.3.3 Changes in integrin TM helix packing induced by the talin F2-F3 domain pair

Having studied how the F2-F3 modulates the orientation of the integrin β tail, the F2-F3/αβ complex (see methods) was then used to explore how the F2-F3 pair facilitates the orientation of the integrin TM dimer (Table 5.2). Initially, a self-assembly CG-MD simulation of the F2-F3/αβ complex was performed to obtain a starting system for the all-atom simulations (Fig. 5.10). At the end of this simulation, the F2-F3 pair adopted a final position and orientation at the lipid/water interface which was identical to the orientation in the simulations above. The two helices were embedded in the bilayer with the αIIb helix in a perpendicular orientation and the β integrin tail in a tilted position (~35^{o} relative to the bilayer normal; see Fig. 5.10). This large tilt angle of the β integrin tail is necessary to better accommodate its long hydrophobic region (~40 Å). To test the effect of the 4E mutations on the F2-F3 binding surface, similar simulations were set up but with the mutated F2-F3 (4E) domain pair in complex with the two TM helices. Ten individual self-

assembly CG-MD simulations were performed with F2-F3-4E, although none of them reached a position similar to the productive binding surface observed in the simulations above.

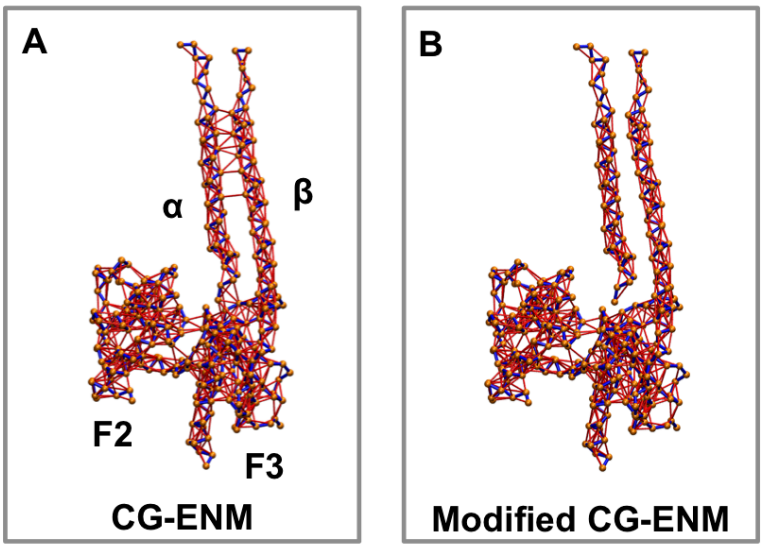


Figure 5.11: A,B. Modified CG-ENM network (B) used in the simulations with the F2-F3/αβ and the talin/αβ complexes. The canonical ENM network is presented in A for comparison.

The final snapshot from the self-assembly CG-MD simulation with the F2-F3/αβ complex (above) was used to initiate longer CG-MD simulations (5x5μs). For these simulations, a modification to the CG model was made in which all the restraints between the α subunit and the F2-F3/β complex were removed and thus the α subunit could move relative to the rest of the complex (Fig. 5.11). Inspection of the α/β dimer at the end of the simulations suggested that the presence of talin F2-F3 domain pair destabilized the dimer significantly. As a result, the interactions in both the α/β IMC and OMC regions were disrupted (Fig. 5.12A) and a new α/β interface was formed as shown in Fig. 5.12B. The αIIb helix maintained its perpendicular orientation relative to the bilayer (tilt angle of ~10^{o} relative to the bilayer normal), whereas the β helix adopted an orientation with a tilt angle of ~28^{o} (Fig. 5.12C). Analysis of the movement of the individual domains within the complex during the simulations revealed a 25±4^{o} rotation of the F2 domain in a plane horizontal to the bilayer surface alongside a 22±1^{o} rotation of the F3 domain in the same plane but in the opposite direction (Fig. 5.13A). The F3 rotation in turn induced a 23±1^{o} rotation of the β TM domain perpendicular to the bilayer in the same direction as for F3. This displacement of the β tail resulted in the disruption of the α/β interactions in both the

IMC and OMC regions. Despite the reorientation of F2-F3, the same interactions of F2-F3 with the bilayer that were identified above were observed.

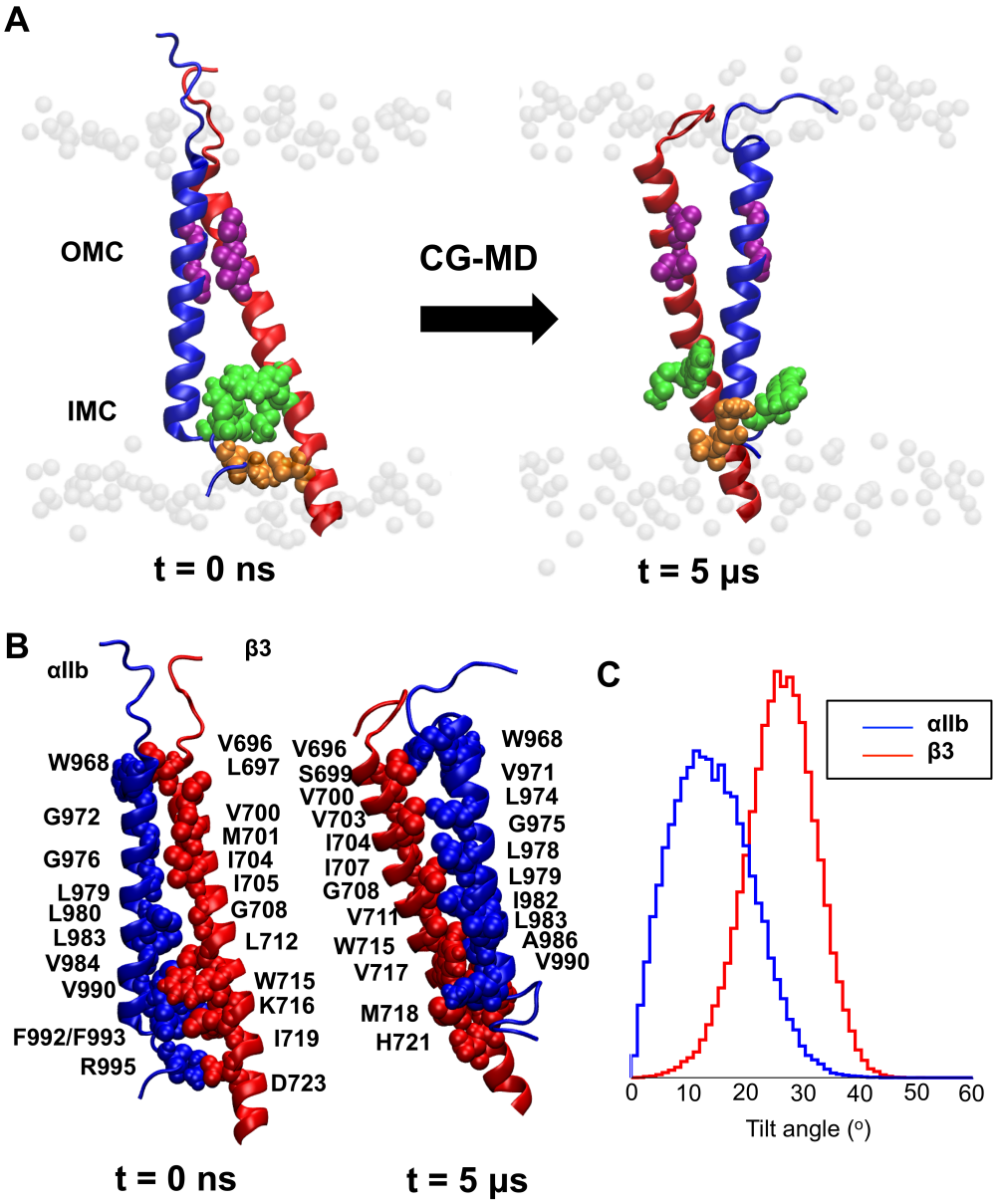


Figure 5.12: A. Movement of the integrin α and β TM regions during the CG-MD simulations with the F2-F3/αβ complex. The structures are shown at the beginning (left) and end (right) of the simulation. In the diagram, the key interactions in the OMC region (Gx_3G) are shown in magenta and in the IMC region in green (F992F993) and orange (αIIbR995-β3D723 salt bridge). The lipid phosphate atoms are shown as grey spheres. B. Interface between the integrin α (blue) and β (red) TM domains at the beginning (t=0 ns) and end (t=5 μs) of the CG-MD simulations. Residues which form key interactions are highlighted. In A and B, the F2-F3 domain pair is omitted for clarity. C Tilt angle distribution for the αIIb (blue) and the β (red) subunits in the *αβ-F2F3-CG* simulation.

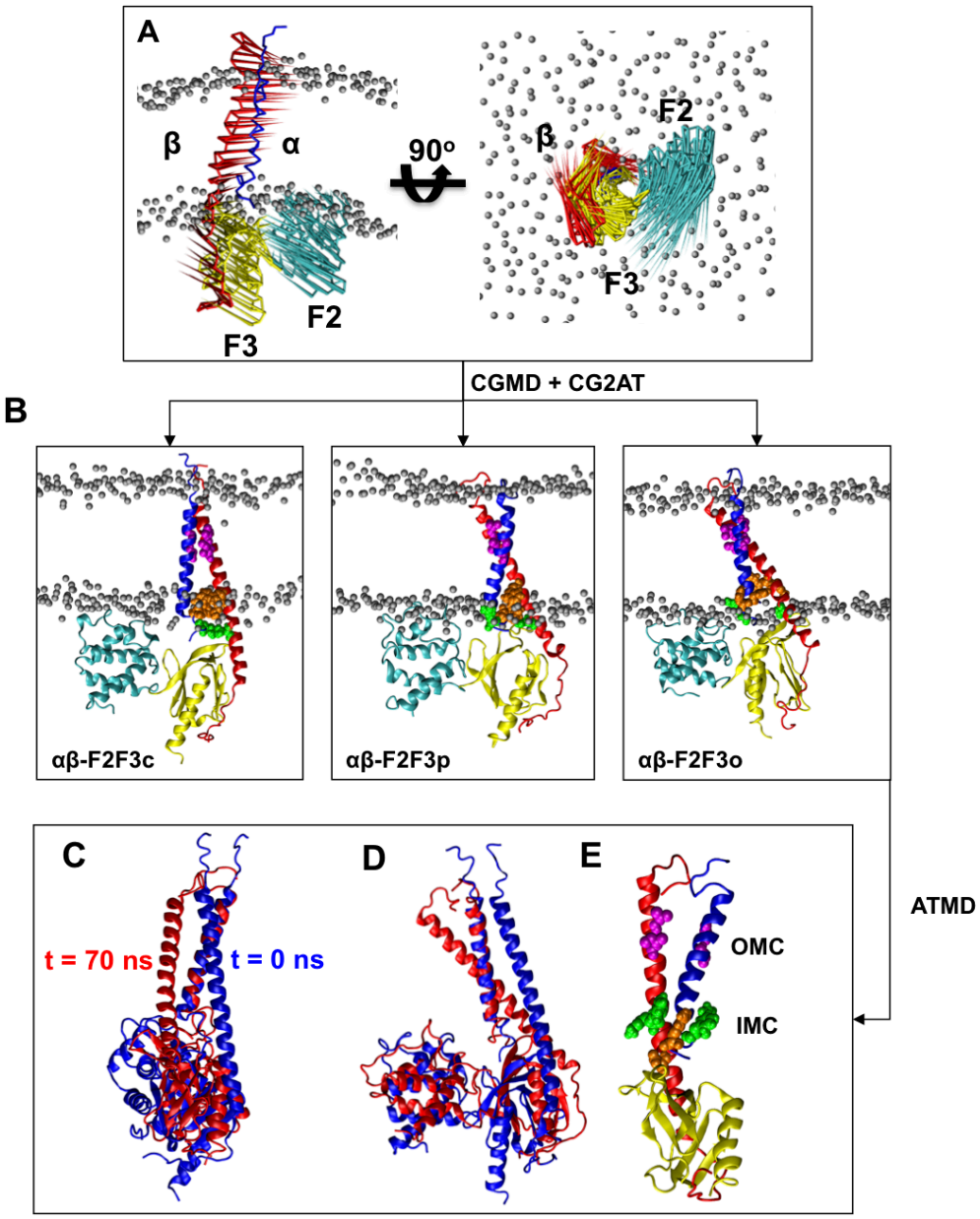


Figure 5.13: A. Relative movement of the subdomains in the F2-F3/αβ complex during the *αβ-F2F3-CG* simulation. The displacement of all the Cα atoms of the F2 (cyan), the F3 (yellow) and the β TM helix (red) through the simulation is shown as arrows. The arrows were mapped onto the initial structure and the length of each arrow represents the displacement/direction of the corresponding Cα atom. Lipid headgroups are shown as grey spheres. B. Initial configuration of the AT-MD simulations. The *αβ-F2F3o* structure represents the "open" integrin model, *αβ-F2F3p* represents the "partially" open integrin model and *αβ-F2F3c* represents the "closed" integrin model. The F2 domain is shown in cyan, the F3 domain in yellow, the β tail in red, the α tail in blue and the lipid phosphates are shown as grey spheres. C,D. β tail displacement (C) and tilt angle relative to bilayer normal (D). The initial position of the complex is shown in blue and the final position in red. E. Snapshot of the last frame of the *αβ-F2F3o-AT* simulation showing the disruption of the interactions in the integrin TM region.

Three representative structures from the *αβ-F2F3-CG* simulation (above) were converted to AT representations to allow a more detailed analysis of the system (Fig. 5.13). These structures were obtained by combining the trajectories from all the CG-MD simulations and clustering the F2-F3/αβ structures as described in Daura *et al.* (398). In the first structure, the interactions in both the IMC and OMC regions were disrupted and thus represent an "open" integrin model. In the second structure, the OMC region interactions were disrupted, although the IMC interactions remained intact. This structure was taken to represent a "partially" disrupted case. In the third structure, both the OMC and IMC interactions were preserved, thus representing a "closed" integrin model.

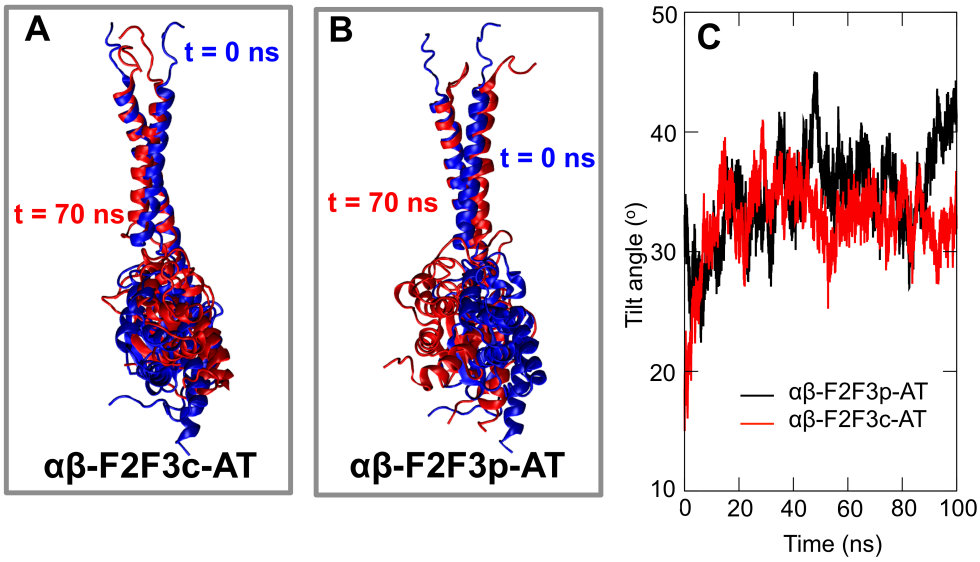


Figure 5.14: A,B. Alignment of the β integrin tail from the structure of the F2-F3/αβ complex (blue) and the final snapshot (red) of (A) the *αβ-F2F3c-AT* and (B) the *αβ-F2F3p-AT* simulations. C. Integrin β TM domain tilt angle relative to the bilayer normal for the *αβ-F2F3c-AT* (red) and the *αβ-F2F3p-AT* (black) simulations.

In the simulations with the "closed" model, the tilt angle of the β helix increased to a final value of ~35° at the end of the simulations, which is similar to the tilt angle observed in the *β-F2F3-AT* simulations (above). The αIIbR995-βD759 salt bridge was disrupted after 10 ns of simulation due to interactions between the αIIbR995 residue and the adjacent talin F3 acidic loop. Interestingly, the interactions between the αIIbF992-F993 aromatic cluster and the β tail, as well as the packing in the OMC region, were very stable and thus no significant destabilization of the α/β dimer occurred (Fig. 5.14A). The increase in the β tail tilt angle and the fact that the interactions in the TM part of the helices remained intact increased the αIIb helix tilt angle to ~20° (Fig. 5.14C). For the "partially" disrupted system, three simulations were also performed for 100 ns in the same

POPC/POPG bilayer. Alignment of the final configuration of the system with the initial structure revealed no significant destabilization of the α/β dimer due to the fact that the aromatic cluster maintained its interactions with the β subunit (Fig. 5.14B).

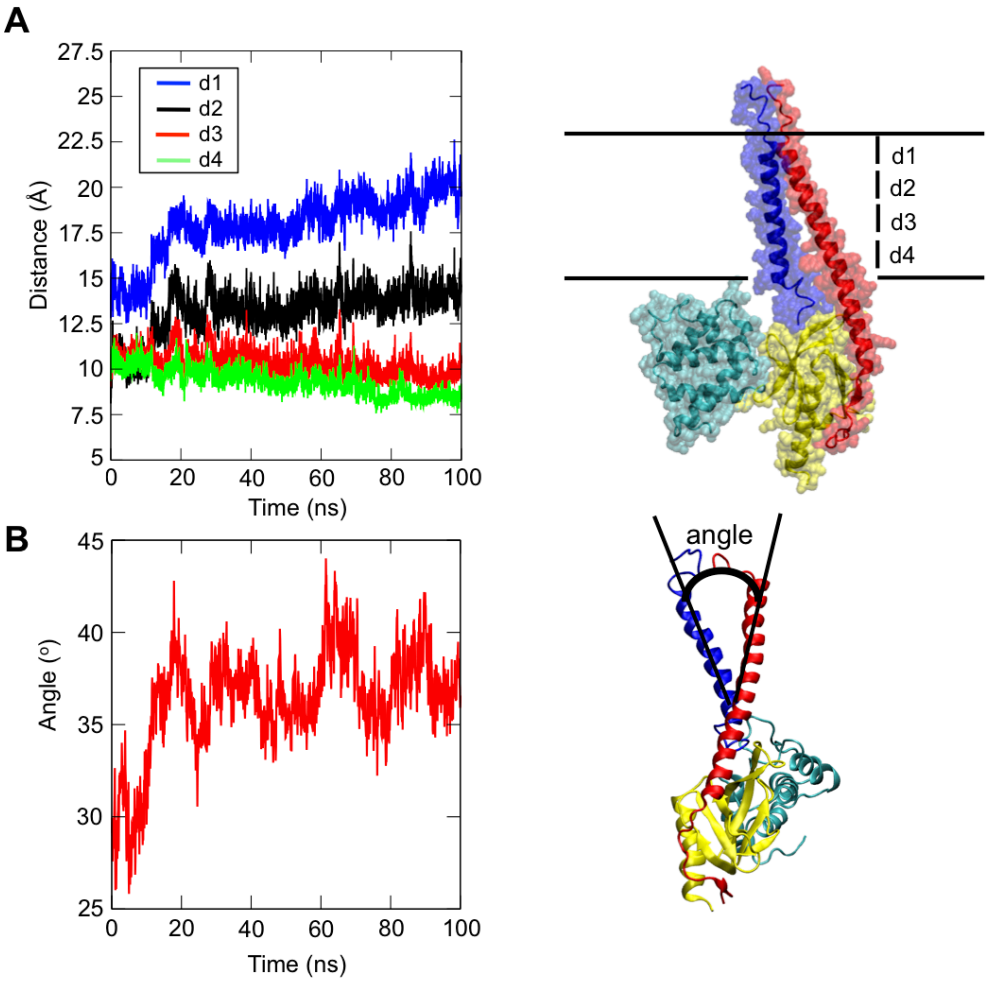


Figure 5.15: A. Interhelical distances (d1 to d4) showing the scissoring movement of the integrin TM region in the *αβ-F2F3o-AT* system. B. Angle between the α and β helices with the centre of the scissor in the IMC region.

In contrast, in the "open" model the interactions in both the OMC and IMC regions remained disrupted and thus the αIIb/β3 dimer remained destabilized throughout the simulations. Interestingly, the weak α/β TM association in combination with the large increase in the β helix tilt angle relative to the bilayer normal resulted in further dissociation of the helices in the OMC region (Fig. 5.15A). In particular, the distance between the α and the β helices in this region increased from ~5 Å to ~20 Å. In contrast, the distance between the two helices in the IMC region was stable and thus a scissoring movement of the two helices with the centre of the scissor in a modified IMC region was observed. The modified IMC region consisted of the αIIb L978, L979, I982, A986, V990

and the β I707, G708, A711, T715 and M718 residues. This scissoring movement resulted in the increase of the angle between the TM parts of the two helices to a final value of 40° (Fig. 5.15B).

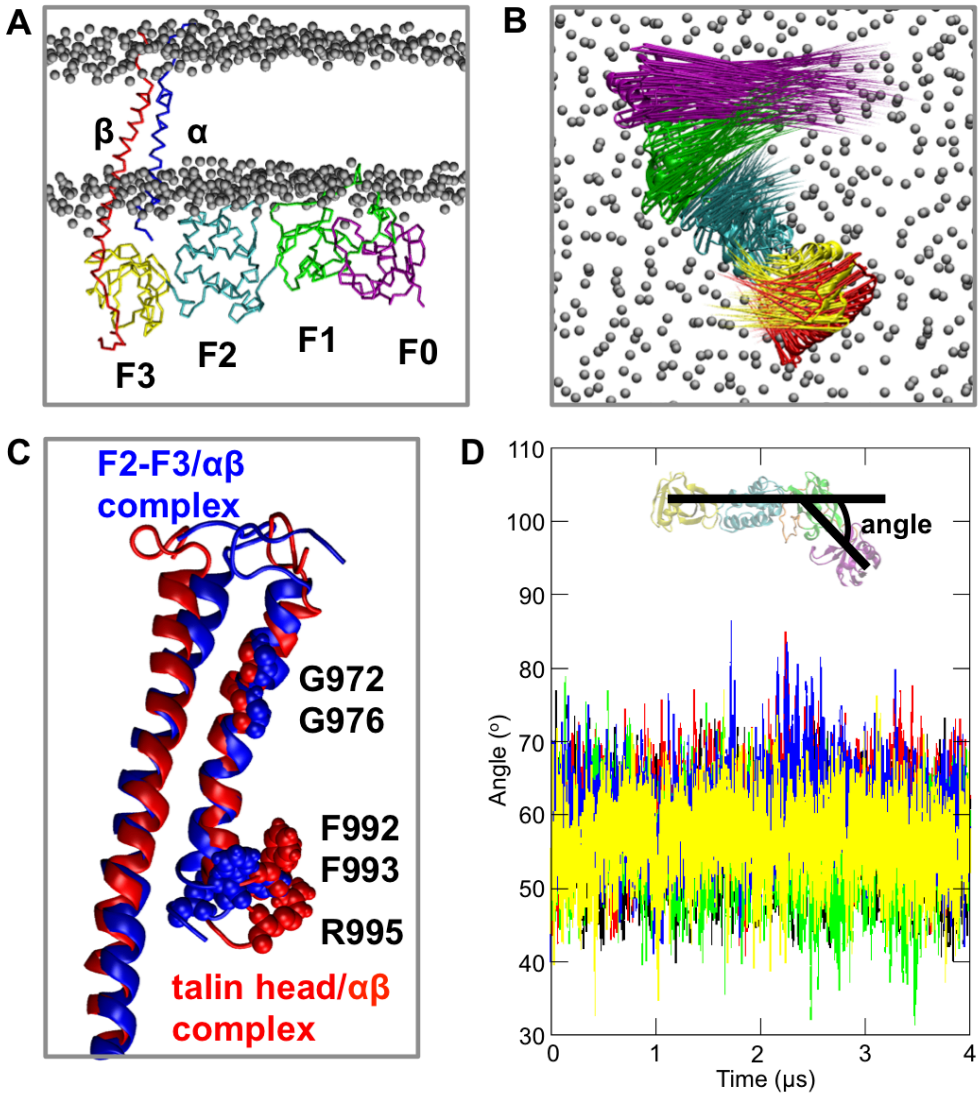


Figure 5.16: A. The *αβ-talh2-CG* simulation set-up. The integrin α subunit is shown in blue, the integrin β subunit in red, the F0 in purple, the F1 in green, the F2 in cyan and the F3 in yellow. The phosphate atoms are shown as grey spheres. B. Relative movement of the domains of the talin/αβ complex in the *αβ-talh2-CG* simulation. The displacement of all the Cα atoms of the F0 (purple), the F1 (green), the F2 (cyan), the F3 (yellow) and β subunit (red) during the simulation is shown as arrows. The arrows are mapped onto the initial structure and the length of each arrow represents the displacement/direction of the corresponding Cα atom. C. Alignment of the integrin TM region from the *αβ-talh2-CG* (red) and the *αβ-F2F3-CG* (blue) simulations. The key residues in the IMC and OMC regions are also shown. D. Angle between the F2-F3 and F0-F1 pairs for the five CG-MD simulations. 0° represents a linear arrangement of the two pairs and 90° represents a conformation where the two pairs are perpendicular. See inset picture for the angle definition.

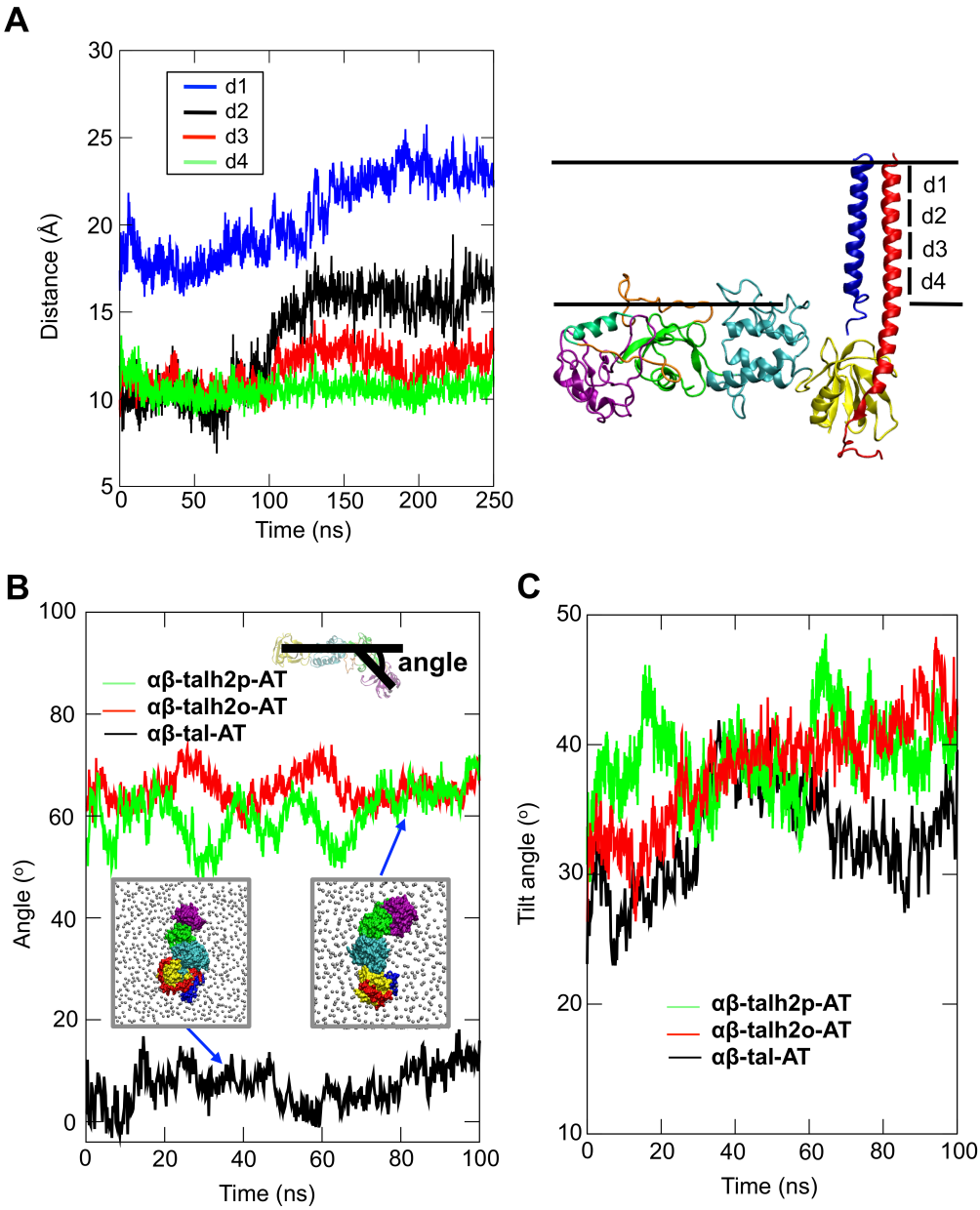


Figure 5.17: A. Interhelical distances (d1 to d4) as a function of time for the *αβ-talh2o-AT* simulation. B. Displacement of the F0-F1 pair relative to the F2-F3 pair as a function of time for the three AT systems with the talin head domain. Inset pictures show the talin conformation for each case. C. The β helix tilt angle relative to the bilayer normal as a function of time for the *αβ-talh2o-AT* (red), the *αβ-talh2p-AT* (green) and the *αβ-tal-AT* (black) simulations. See Table 5.2 for more information.

5.3.4 The talin head domain perturbs the packing of the integrin TM helices

The effect of the complete talin head domain on the structure of the integrin TM region was also studied. To this end, the talin head domain with the optimum V-shape configuration derived from the simulations in Chapter 3, in complex with the α/β integrin TM complex (see methods), was inserted in a POPC/POPG bilayer (Fig. 5.16A).

Calculation of the movement of the integrin TM region during the simulations revealed a similar displacement of the β helix perpendicular to the bilayer normal as seen in the CG-MD simulations of the F2-F3/αβ complex (Fig 5.16C). As a result, the close packing in the OMC region was perturbed and all the key interactions in the IMC region (i.e. hydrophobic and salt bridge) were disrupted. Closer examination of the movement of the individual domains in the complex revealed a reorientation of the talin head domain in the bilayer surface, comparable to the reorientation observed in the *αβ-F2F3-CG* system (Fig. 5.16B). Interestingly, calculation of the angle between the F2-F3 and the F0-F1 domain pairs during the simulations revealed a stable angle of ~60^{o}, suggesting that the novel V-shape talin configuration was retained in all the simulations (Fig. 5.16D).

In order to study in more detail the talin/αβ complex, two structures which represent the "open" and "partially" disrupted states of integrin were selected from the *αβ-talh2-CG* simulation using similar criteria to the *αβ-F2F3-CG* simulations (above) and converted to AT representation. At the end of the simulations with the "open" structure, the V-shape conformation of the talin head domain was retained (i.e. the angle between the F2-F3 and F0-F1 pairs was 60^{o} at the end of the simulations; see Fig. 5.17B). Intriguingly, calculation of the interhelical distances in the OMC and IMC regions revealed an increase in the helix–helix distance in the OMC region but not in the IMC region. This type of movement was similar to that seen in the *αβ-F2F3o-AT* simulations above and indicates a scissoring movement of the two helices with a modified IMC at the centre of the scissor (Fig. 5.17A). In addition to the scissoring movement, a large increase in the tilt angle of the β helix relative to the bilayer normal, to a final value of 40^{o}, was also observed (Fig. 5.17C). To further investigate the scissoring movement, one of the simulations was extended to 250 ns. At the end of this simulation, the helix–helix distance at the extracellular region (d1) reached a value of 23 Å, which is similar to the distance at 100 ns.

In contrast, in the simulation with the "partially" disrupted model no significant movement between the two helices was observed despite the comparable increase in the β integrin TM region tilt angle during the simulation (~40^{o}; see Fig. 5.17C). This was due to the fact that the interactions of the aromatic cluster (F992, F993) with the β integrin W715 and K716 residues were retained throughout the simulation, thus preventing further dissociation of the two helices. The talin V-shape conformation was again retained (Fig. 5.17B).

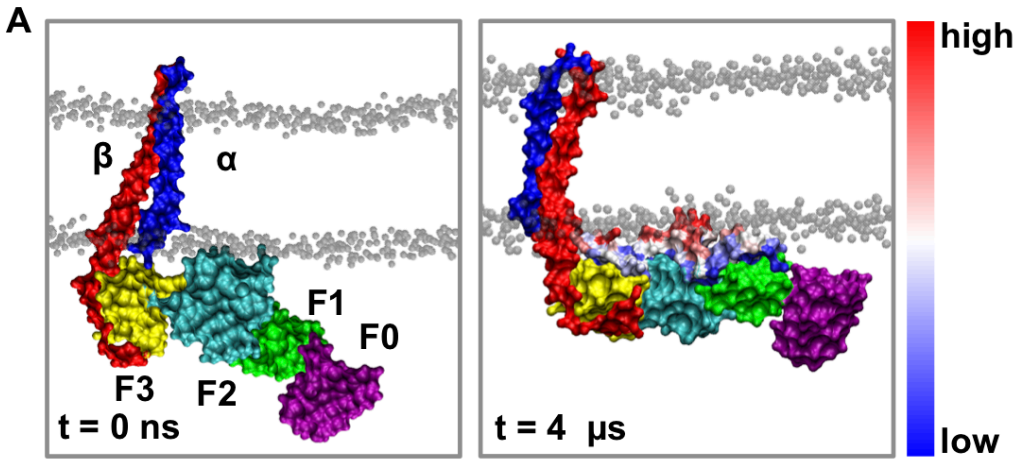


Figure 5.18: A. Progress of the *αβ-tal-CG* simulation. Snapshots from one of the simulations are shown at 0 ns and 4 μs. The integrin α subunit is shown in blue, the β subunit in red, the F0 in purple, the F1 in green, the F2 in cyan and the F3 in yellow. The phosphate atoms are shown as grey spheres. In the simulation snapshot at 4 μs, the contacts between the talin and the bilayer are mapped on the structure on a colour scale from blue (low number of contacts) through to red (high number of contacts).

A similar talin/αβ complex constructed using the talin head domain conformation obtained from the *tal-AT* simulation in Chapter 3 was also inserted in the same POPC/POPG bilayer (Fig. 5.18A). Interestingly, at the end of the CG-MD simulations, similar destabilization of the α/β integrin TM dimer was observed as in the simulations above (Fig. 5.17) despite the fact that the optimum V-shape talin configuration was not observed in any of the simulations. This indicates that the productive F2-F3 pair interaction with the bilayer was sufficient to disrupt the interactions in the integrin TM region. A snapshot representing an "open" integrin structure was selected and an all-atom simulation was performed for 100 ns. The perturbed packing of the two helices was maintained during the simulation, but no further dissociation (i.e. scissoring movement) of the two helices occurred. This suggests that a more stable association of talin with the lipids is needed for the scissoring movement of the helices. See Appendix C for detailed discussion of this system.

5.4 Discussion and conclusions

Recent structural studies of talin and integrins (52, 118, 119, 125, 186, 203) have greatly increased our understanding of integrin activation from an intracellular trigger. In this chapter, it has been demonstrated that a detailed investigation of the system using simulation methods can give novel information about the activation mechanism which

static structures fail to provide. In particular, by using long multiscale MD simulations it is shown how the talin head domain, the membrane and the α/β TM region interact to promote the integrin inside-out activation. Based on the results of this chapter, a mechanism is suggested which explains how the talin head domain facilitates rearrangement of the integrin TM region, thus stimulating integrin activation (Fig. 5.19A).

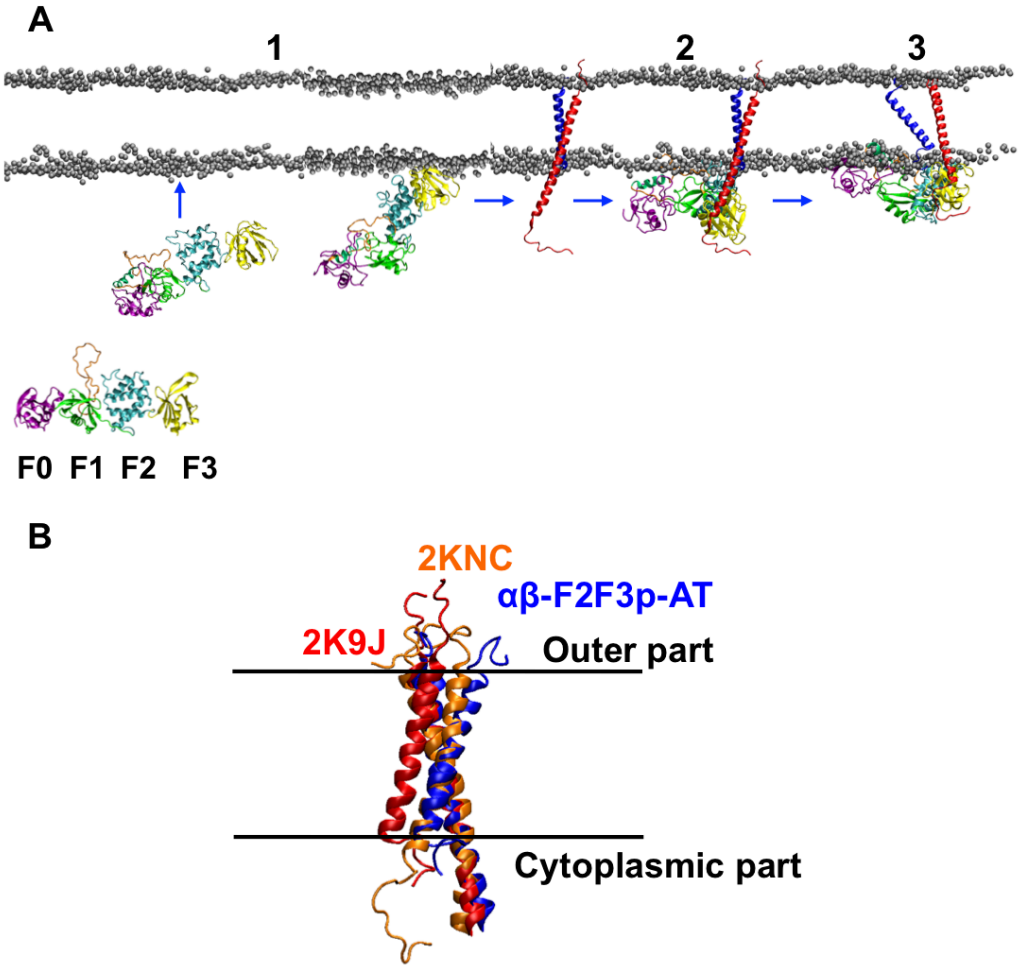


Figure 5.19: A. Proposed mechanism for the integrin inside-out activation by the talin head domain. 1. Electrostatic interactions between the talin head domain (see Chapter 3) and the negatively charged lipids in the membrane orient talin in a plausible orientation to associate with the integrin β CT region. 2. Strong talin/negatively charged lipid interaction promotes reorientation of the talin head domain in a plane perpendicular to the bilayer normal, which in turn rotates the β integrin tail (~30^{o}) perpendicular to the membrane. This disrupts the interactions in both the αIIb/β3 IMC and OMC regions. 3. The weakened αIIb/β3 interactions in combination with a ~15^{o} increase in the β tail tilt angle relative to the bilayer normal result in a scissoring movement of the TM part of the two helices with a modified IMC at the centre of the scissor. B. Superposition of the β TM region (residues 697-726) from the *αβ-F2F3p-AT* simulation (blue), the αIIb/β3 (PDB:2KNC) structure (orange) and the αIIb/β3 (PDB:2K9J) structure (red). The alignment shows that the αIIβ/β3 (PDB:2KNC) structure is similar to the partially disrupted state (*αβ-F2F3p-AT* simulation) and could therefore correspond to an intermediate state. The F2-F3 domains are omitted for clarity.

Initially, the αIIb/β3 NMR (118) structure was simulated in a membrane-like environment. These simulations are in good agreement with a RH packing of the two helices, with a crossing angle of ~-35^{o} where the αIIb helix is in a straight position in the membrane and the β3 is tilted ~30^{o} (118, 119). A double mutation in the IMC aromatic cluster or deletion of the IMC region but not disruption of the αIIbR995-β3D723 salt bridge strongly affected the crossing angle and thus the packing of the two helices (118, 164, 399). This is in good agreement with the experimental results of Lau *et al*. and Kim *et al.*, which suggested that disruption of the αIIbR995-β3D723 salt bridge weakens but does not eliminate the integrin TM region association (118, 399). Additionally, it is in good agreement with experimental results which suggested that substitution of either of the two phenylalanines to alanine moves the integrin equilibrium towards an active state (164). However, in all cases modification of the IMC region was not sufficient to disrupt all the αIIb/β3 interactions and, therefore, the additional role of talin in disrupting the interactions in the IMC and OMC regions is crucial.

Indeed, the simulations that included the entire talin head domain (or the F2-F3 domain pair) showed that talin interactions with the integrin β tail and the membrane surface disrupted the interactions in both the OMC and IMC regions, resulting in separation of the helices. Recent crystal structures of the integrin ectodomain (78, 79, 81) suggested that in the inactive state the α and β domains are in close proximity; therefore, the scissoring movement proposed by the simulations may provide a plausible model for how the ectodomain is stimulated to move to an extended active state. This mechanism also clarifies the role of the IMC and OMC regions in maintaining the integrin inactive state and explains much of the experimental data, which have indicated that disruption of the association in the OMC and IMC regions is needed for integrin activation (113, 118, 119, 158, 159, 163, 164, 166, 173, 385, 400). In particular, it was suggested that deletion or mutation of residues in the CT region of the α or the β tails promotes integrin activation (37, 163, 164, 166, 168, 400). Furthermore, it was shown that mutations of the two glycines in the OMC region cause steric clashes, which disrupt the packing of the α/β dimer (118, 159, 173). A study by Kim *et al.* using FRET (112) and an αLβ2 integrin revealed a significant reduction in FRET signal upon talin addition. The authors interpreted this reduction as integrin tail separation. The scissoring movement suggested above may separate the handles as well as the blades. Even though the distance in the IMC region (i.e. d4) does not suggest a tail separation in this region, the disruption of the αIIbR995-β3D723

salt bridge during the simulations and the flexibility of the α and β tails imply that tail separation is not inconsistent with the model suggested.

The results of this chapter also explain the experimental observations that disruption of the F3/β interface significantly reduces the ability of talin to activate integrins (51-57, 125). All simulations revealed a strong association between the F3 domain and the integrin β CT tail in the MProx region, the NPxY motif and residue W775 in β1D (β3 W739). In particular, the β residue Y783 in the NPxY motif of the model, which corresponds to the Y747 residue of the β3 subunit, exhibited a strong interaction with the F3 domain. This is in good agreement with previous studies which suggested that mutation of the tyrosine in the proximal NPxY motif to alanine inhibits talin-mediated integrin activation (52, 53, 127, 222, 401). Furthermore, in all simulations the β1D residue W775 (β3 W739) remained in a hydrophobic pocket comprised of F3 residues R361, A363 and Y380 (R358, A360 and Y377 in talin-1), in good agreement with the suggestion that the R358A mutation significantly reduces integrin activation (42). A recent NMR study (145) suggested that the activation of integrin is also regulated by the state of the β integrin CT region. In their study, the authors demonstrated that the β TM is bent at residue R724, interacting with the membrane with the same residues required for the F3/β association. Promotion of a helical structure in the R724 region would extend the β helix, exposing the β residues and allowing the productive binding of F3 domain.

Another interesting outcome of the work documented in this chapter is that the β TM helix consistently favoured a higher tilt angle in the bilayer. Such a large tilt angle is necessary to better accommodate its long hydrophobic region (~40 Å). Furthermore, this tilted position promotes the interactions of the K716 sidechain with the lipid phosphate groups and positions the W715 residue in the lipid headgroup/tail region, a position which is favoured by tryptophans (136). These interactions provide an anchor to the lipid headgroups, which in combination with the interactions between the α and β TM domains in the OMC and IMC regions regulate the β subunit tilt angle (41). In the absence of the α subunit or after the attachment of talin, the five hydrophobic amino acids which follow the K716 residue (charged) are also immersed in the lipid headgroups. This tendency of the β TM helix to adopt a tilted orientation in the membrane is also demonstrated by experimental (132) and other computational (193) results and it has been suggested that this promotes further dissociation of the integrin TM helices in the OMC region.

Besides the αIIb/β3 structure determined by NMR in bicelles (PDB:2K9J), other structures have been derived either using organic solvents (PDB:2KNC) (119, 121) or from

other computational approaches (120, 162). The computational models are in good agreement with the αIIb/β3 structure in bicelles (PDB:2K9J), largely reproducing both the IMC and OMC interactions. In contrast the 2KNC structure, which was obtained in the absence of a lipid environment, indicated different packing for the aromatic cluster in the IMC part, demonstrating that the lipid bilayer environment is important for the formation of the IMC interactions. Interestingly, alignment of the 2KNC structure with the structures derived from the simulations of the F2-F3/αβ complex suggests that it is rather similar to the partially disrupted model (*αβ-F2F3p-AT*; Fig. 5.19B). Therefore, this suggests that this structure might relate to an intermediate state between integrin "off" and "on" states. This state may be related to the intermediate state suggested by the "switch-blade' model where the ectodomain is extended with the headpiece in a closed state (78, 97).

In summary, the results presented in this chapter suggest a mechanism (Fig. 5.19A) for integrin inside-out activation in which binding of talin destabilizes the α/β interactions, thus weakening the α/β association and leading to a scissoring motion of the two helices. This model shows how protein/membrane interaction can modulate potential changes in receptor ectodomains and it is also a good demonstration of the strength of this multiscale simulation approach for obtaining information about the position, dynamics and orientation of membrane receptors, thus allowing construction of dynamic models for their function.

Chapter 6

The Integrin Receptor in more Biologically Relevant Bilayers

6.1 Introduction

In recent years, the development of new experimental techniques (see, for example, the review on 'lipidomics' (402)) has unravelled the diversity of lipids in the membranes of different organisms. It has been shown that 5% of the cell genome is used to specify lipid synthesis (343) and that there is significant variation of lipids in different membranes. This provides additional evidence that variation in lipid type is important in many cellular functions (402, 403), including TM signalling (404, 405) and protein trafficking (406). Recent studies have also suggested that the membrane is an organized and structured entity with interplay between 'lipid-ordered' (i.e. lipids rafts) and 'lipid-disordered' states (407). The two states can coexist in the same membrane with different size and with different physical properties depending on the lipid–lipid, protein–lipid or protein–protein interactions. It should be noted, however, that there is ongoing debate concerning the details of the lipid phase separation process (9, 10). This is due to the fact that large-scale phase separation is observed in model membranes but not in living cells (7, 8). Additionally, the high density of membrane proteins and their specificity for different lipid types and proteins adds another level of complexity which is still poorly understood (15). Therefore, it is possible that, instead of large-scale separation, smaller microdomains or nanodomains of lipids and protein exist to regulate membrane functions (7, 8). Fluorescence measurements are potentially a good probe of these properties (282, 408, 409) but the concerted movement of protein and lipids (11-14), combined with the scale and the lifetime of the lipid domains, makes the study of protein/lipid dynamics experimentally difficult (410-412).

Protein/lipid dynamics and organization occur on scales of nanometers and microseconds, scales which can be investigated using MD simulations (244, 318, 413). All-atom simulations have been used to gain insights into protein/lipid organization and dynamics, although they are computationally expensive mostly because of the low diffusion rate of the lipids (e.g. $\sim 10^{-12}$ for POPC lipid m^2/s (414)) (282-284). Therefore,

lower resolution techniques such as CG-MD simulations have been extensively used to study lipid dynamics and organization in bilayers or liposomes without a protein component (14, 285, 286) or with simple TM proteins such as the WALP peptide (287, 288). These CG studies have, so far, used rather simple systems and, although they have provided useful insights, studies of a more complex protein with extracellular, TM and CT regions would give novel information. One such example is the complete integrin receptor in complex with its activating protein (talin). Information about the influence of lipid structure and dynamics on this biologically relevant system will help increase our knowledge of the behaviour of a signalling receptor in a more realistic environment.

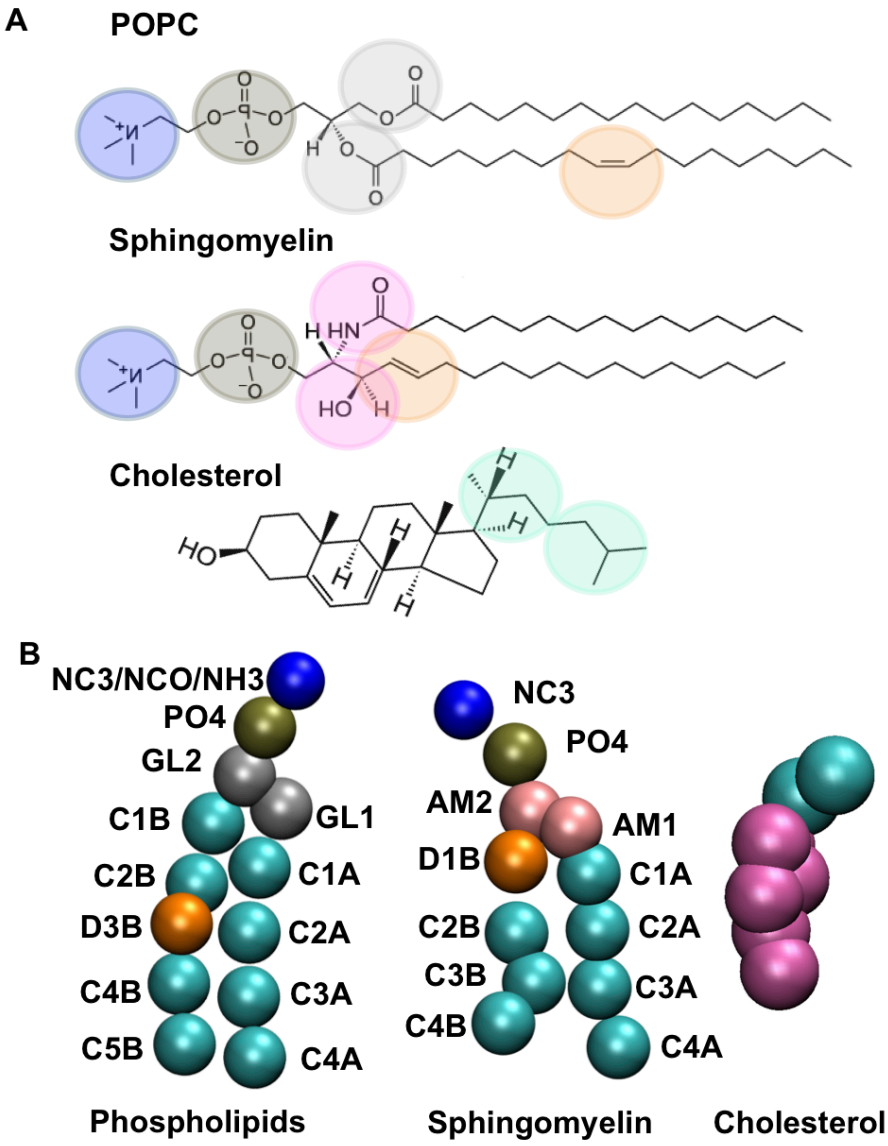


Figure 6.1: A,B. AT (A) and CG (B) structures of the phospholipids, sphingomyelin and cholesterol used in the simulations.

The available structural data for integrins (reviewed in Chapter 1) (79, 118, 125) allow the construction of models of the complete integrin receptor, which can then be inserted in bilayers resembling the mammalian plasma membrane where integrins are found (see Chapter 1) (15). The plasma membrane has an asymmetric distribution of lipids with sphingolipids (mainly sphingomyelin (SM)) and sterols (i.e. cholesterol (CHOL)) but not glycerolipids in high concentrations in the outer leaflet (343). SM has a polar headgroup with saturated lipid tails and a double bond between the C4 and C5 atoms in the acyl chain (Fig. 6.1) (415). It is believed to be involved in TM signalling (405). CHOL is believed to increase the thickness and stiffness of lipid bilayers (416) by increasing the packing density (417, 418). Furthermore, it was shown to facilitate receptor activity (283), TM peptide segregation (282) and to affect the lateral pressure of the membrane (419). The asymmetric distribution of lipids in the two leaflets is believed to be important for some functions of the plasma membrane, such as vesicle budding (420, 421). Plasma membranes in mammals are not directly involved in lipid synthesis and therefore the asymmetric distribution is achieved either by direct placement of specific lipid types in the leaflet where they function (e.g. SM in the outer leaflet) or by transport of lipids between the leaflets. For example, the P4 ATPase has been shown to transport POPC or POPE between the two leaflets (422).

Two main approaches are applied in this chapter. First, extended (10 μs) CG-MD simulations are used to examine the dynamics and interactions of an integrin/talin complex in biologically relevant, symmetric and asymmetric, bilayers containing CHOL, phospholipids and SM. Second, the same simulations are used to explore the effect of the integrin/talin complex on lipid organization and dynamics. The results demonstrate the dynamic nature of the integrin receptor and suggest that the integrin/talin complex alters the local lipid organization and motion in the outer and inner bilayer leaflets in an asymmetric fashion and that the dynamics of lipids in the vicinity of the protein are slowed down.

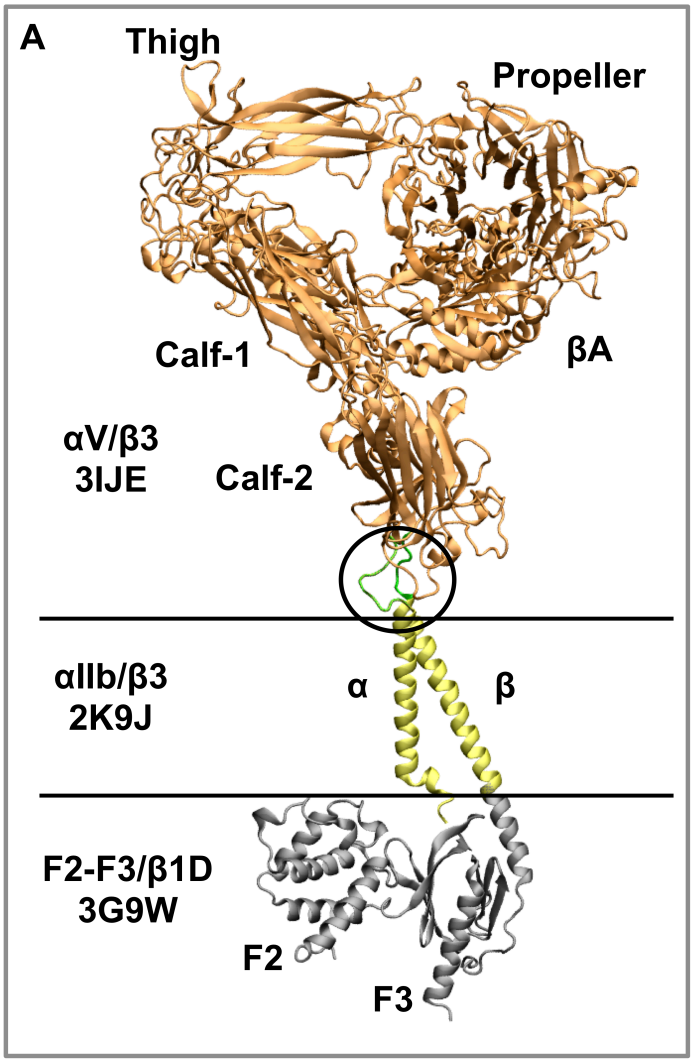


Figure 6.2: A. Model of the integrin/talin F2-F3 complex. The β subunit was comprised of β3 G1-P685 residues/β3 688-719 residues/β1D 753-787 residues, whereas the α subunit was comprised of αV F1-P963 residues/αIIb P965-P998 residues. The part of the structure corresponding to the αIIb/β3 structure (PBD:2K9J) is shown in yellow, the F2-F3/β1D (PDB:3G9W) is in grey and the integrin ectodomain (PDB:3IJE) is in orange. The area between the integrin TM and ectodomain regions which was predicted to be unstructured is shown in green (see Section 6.2.1 for more information).

6.2 Methods

6.2.1 Integrin/talin F2-F3 complex modelling

The complete integrin receptor was modelled using the integrin TM αIIb/β3 NMR structure (PDB:2K9J) (118) and the αVβ3 ectodomain structure (PDB:3IJE) (79). During the modelling, the αIIb TM region was oriented at ~30^{o} angle relative to the axis of the α subunit ectodomain leg (i.e. Calf-1 and Calf-2 domains) to satisfy available experimental data (Fig. 6.2) (79, 81). All the modelling was performed using Modeller (342) and the relative orientation of the two structures during the modelling was kept restrained. To

construct the integrin α subunit, the two structures were connected at αIIbP965 (PDB:2K9J) and αVP963 (PDB:3IJE) residues. To construct the integrin β subunit, the structures were connected at β3P685 (PDB:2K9J) and β3P685 (PDB:3IJE) residues. This resulted in a structure with αV F1-P963/αIIb P965-P998 and β3 G1-P685/β3 P685-F727 subunits (Fig. 6.2). To model the conformation of the region between the TM region and the ectodomain, for which there is no available structural data (res: α 956-965/β 685-694; green in Fig. 6.2), a structure prediction (performed using PSIPRED (423)) suggested that both regions are unstructured. This same region has been suggested previously to be unstructured (79). It should be noted that recent studies also suggest that glycosylation is important for integrin function (424). Therefore, it is possible that the flexibility of the aforementioned region between the TM and the ectodomain regions is affected by glycosylation but in this chapter the effect of glycosylation on the integrin receptor will not be taken into consideration. The resulting model structure was energy minimized to remove any steric clashes. The integrin/talin complex was constructed by aligning overlapping residues at the TM region of the integrin receptor model (2K9J part of the structure; Fig. 6.2) with the TM region of the F2-F3/β1D (PDB:3G9W) (125) complex as described in Appendix C.

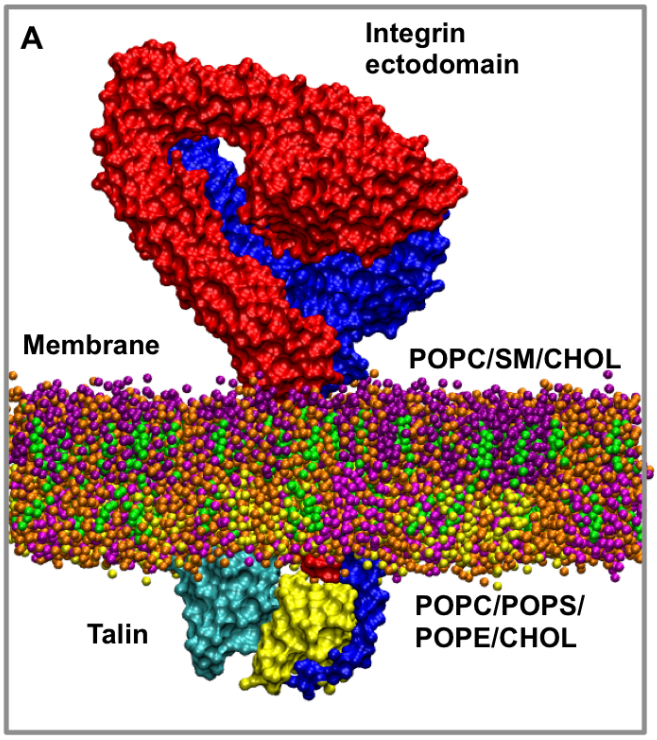


Figure 6.3: A. Illustration of the simulation system. The integrin α subunit is shown in red, the β subunit in blue, the talin F2 domain in cyan and the F3 domain in yellow. Different colours in the bilayer represent different lipid types.

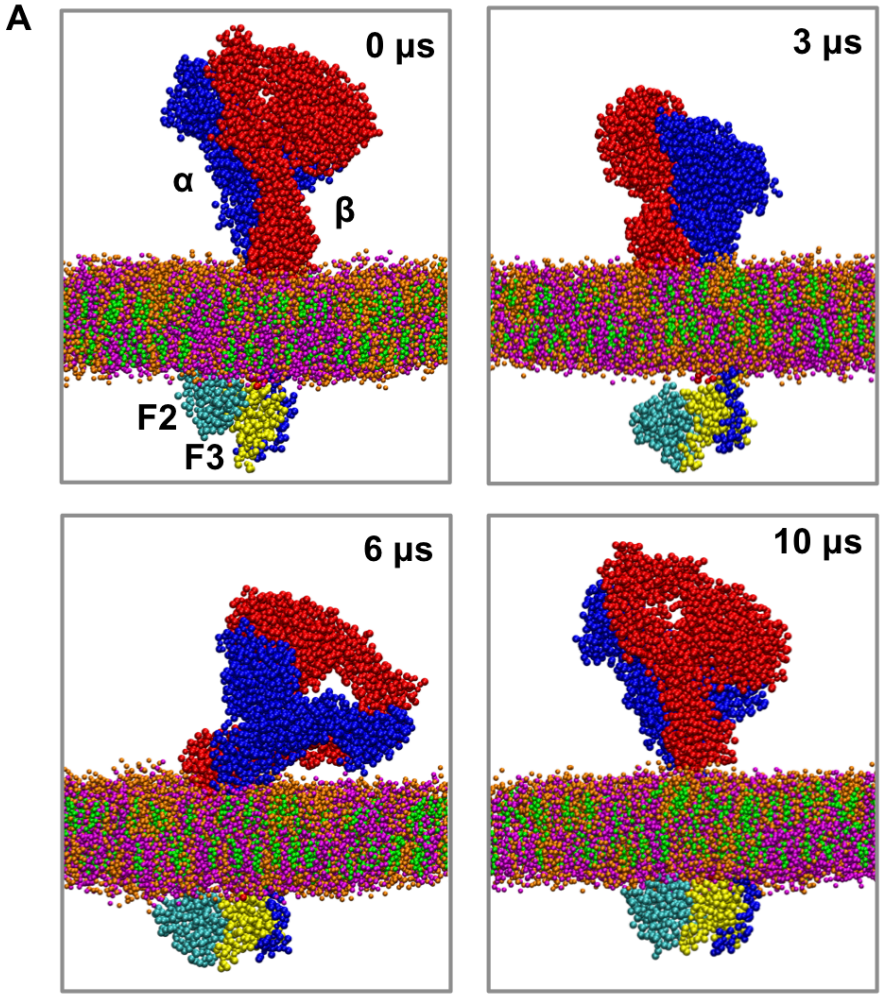


Figure 6.4: A. Progress of the *int_tal-1* simulation. Snapshots of the simulation are shown at 0 µs, 3 µs, 6 µs and 10 µs. The integrin α subunit is shown in red, the β subunit in blue, the talin F2 domain in cyan and the F3 domain in yellow. The bilayer lipids are shown in magenta for POPC, orange for SM and green for CHOL. The waters and counter ions are omitted for clarity.

Table 6.1: Bilayers used in the simulations

Bilayer	Inner Leaflet Lipid Concentration	Outer Leaflet Lipid Concentration	Size (Å)
asymmetric	POPC/POPS/POPE/CHOL (35%/20%/25%/20%)	POPC/SM/CHOL (1:1:1)	190x190
outer	POPC/SM/CHOL (1:1:1)	POPC/SM/CHOL (1:1:1)	190x190
inner	POPC/POPS/POPE/CHOL (35%/20%/25%/20%)	POPC/POPS/POPE/CHOL (35%/20%/25%/20%)	190x190
outer-2	POPC/SM/CHOL (2:1:1)	POPC/SM/CHOL (2:1:1)	200x200

* See Appendix D.1 for more information.

Table 6.2: Summary of simulations

Simulation	Protein	Bilayer
int_tal-1	F2-F3+α/β* TM+ectodomain	outer
int_tal-2	F2-F3+α/β* TM+ectodomain	inner
int_tal-3	F2-F3+α/β* TM+ectodomain	asymmetric
tal_αβ-1	F2-F3+α/β* TM	outer
tal_αβ-2	F2-F3+α/β* TM	inner
tal_αβ-3	F2-F3+α/β* TM	asymmetric
tal_αβ-4	F2-F3+α/β* TM (protein Cα atoms restrained)	outer
tal_αβ-5	F2-F3+α/β* TM	outer-2

TM = transmembrane helix; β* indicates a chimeric β chain (β3/β1D) – see main text for details).

6.2.2 Bilayer and system preparation

As mentioned above, the plasma membrane has an asymmetric distribution of lipids with the outer leaflet rich in SM, POPC and CHOL and the inner leaflet consisting mainly of phospholipids and CHOL (343). Therefore, in order to study the behaviour of the integrin/talin complex in the plasma membrane, an asymmetric bilayer (which resembles the plasma membrane; see Table 6.1) was constructed and the integrin/talin complex was inserted in the bilayer. The asymmetric nature of the bilayer (i.e. it has different concentration/types of lipids in each leaflet) induces further difficulties in constructing this type of bilayer by simply using self-assembly CG-MD simulations (described in Section 3.2.3). Thus, initially two symmetric bilayers (of the same size; Table 6.1), which had the same lipid concentration as either the inner or the outer leaflet of the asymmetric bilayer, were constructed (inner and outer bilayers in Table 6.1). For the asymmetric bilayer construction, the outer leaflet from the outer bilayer and the inner leaflet from the inner bilayer were combined (see Appendix D.1).

The complexity of the integrin/asymmetric bilayer complex may induce difficulties in explaining some of the phenomena which occur in the system and thus the integrin receptor was also inserted (and simulated) in the two symmetric bilayers (*int_tal-1* and *int_tal-2* simulations in Table 6.2). These symmetric bilayers are simpler compared to the asymmetric bilayer and some of their properties have been studied before; thus, they were used to better understand/explain the phenomena which occur at the more biologically relevant integrin/asymmetric bilayer system.

The integrin/talin complex (*int_tal* systems in Table 6.2) was inserted in the bilayer in such way that the integrin TM region (2K9J part of the model; Fig. 6.2) was embedded in the bilayer with the αIIb TM helix parallel to the bilayer normal (118, 119, 131, 132). Thus, the β TM helix adopted a tilted position of ~30^{o} and the F2 positively charged patch was situated on the bilayer surface (Fig. 6.3). In the *tal_αβ* systems in Table 6.2, the ectodomain (3IJE part of the structure; Fig. 6.2) was removed in order to examine how the dynamics of the F2-F3/αβ complex (3G9W/2K9J part of the structure; Fig. 6.2) and its effect on the bilayer structure and dynamics were affected by the integrin ectodomain. During the insertion, all lipids which overlapped with the protein were removed using GROMACS (302, 303). After the protein insertion in the bilayer, CG-MD simulations were performed for each system.

6.3 Results

6.3.1 Integrin/talin complex dynamics

Inspection of the motion of the integrin receptor during the simulations revealed large fluctuations of the ectodomain in all simulation systems due to the flexible linkers between the TM region and the ectodomain (Fig. 6.4). Calculation of the contacts between the integrin ectodomain and the lipids in the bilayer (a contact is defined by using a cut-off distance of 8 Å) demonstrated that, in the *int_tal-1* and *int_tal-3* simulations (i.e. in the absence of any negatively charged lipids in the outer leaflet), no direct ectodomain/bilayer interactions occurred. In contrast, in the *int_tal-2* simulation, analysis of the ectodomain movement revealed transient interactions between the legs of the ectodomain and the bilayer, mainly with POPS and POPE lipids. It should be noted that, due to the ENM model which was applied to model the secondary and tertiary structure in the CG model, no conformational changes were allowed within the integrin receptor and thus the integrin receptor was "locked" in its inactive state. Note that the integrin ectodomain was highly dynamic even when the two-dimensional average lipid headgroup densities were converged. Therefore, if the simulations are extended to larger timescales then the ectodomain is expected to continue to be highly dynamic. This indicates that in the plasma membrane (which is asymmetric) the integrin ectodomain is very dynamic even in the inactive state. However, it does not interact directly with the outer leaflet of the membrane. This is expected to have implications for its functions because possible stable interactions with the lipids might hide some ligand binding sides, thus inhibiting association of ligands with the receptor. The dynamic nature of the integrin ectodomain may also be important in

the interactions of the integrin ectodomain with various proteins in the ECM and the formation of focal adhesion complexes. The integrin/ECM association is followed by attachment of the integrin β subunit, via talin, to polymerising actin filaments (78). This results in a pulling force which possibly activates integrins and promotes the formation of focal adhesion complexes. These complexes either pull the matrix towards the cell or the cell towards the matrix (219). This result is also consistent with studies of integrins in nanodiscs, which demonstrated that in the active state the integrin ectodomain stands in an extended form perpendicular to the membrane (106). In contrast, experimental and computational studies on the epidermal growth factor receptor (which has some similar characteristics to the integrin receptor) demonstrated that its ectodomain interacts strongly with the membrane and that these interactions regulate its function (260, 425).

Calculation of the contacts between the F2-F3 domain pair and the lipids in the *int_tal-1* simulation suggested transient interactions between the talin F2 positively charged patch (res: 255-285) and the bilayer (Fig. 6.4). Based on previous work described in this thesis, these interactions are transient because there are no negatively charged lipids in the membrane to stabilize the talin F2/lipid interactions. In the *int_tal-2* and *int_tal-3* simulations, the presence of negatively charged lipids stabilized the interactions between the CT protein and the bilayer. The talin F3 subdomain interacts with the membrane with a positively charged loop (res: 315-328). In all the simulations with the F2-F3/αβ complex, the interactions with the bilayer were similar reflecting the fact the ectodomain does not form a stable complex with the bilayer.

6.3.2 Lipid organization around the integrin/talin complex

In order to study in more detail the organization (and the interactions) of the different lipid types around the integrin receptor, the two-dimensional average lipid headgroup densities around the integrin receptor were calculated. For the *int_tal-1* system, this analysis revealed an asymmetric distribution of lipids between the two leaflets with a high-density ring of POPC lipids around the integrin TM region in the inner leaflet (Fig. 6.5C). The same analysis revealed a homogenous distribution for SM in both leaflets and a high density of CHOL close to the TM region (Fig. 6.5A,B). The CHOL distribution was not uniformly distributed around the protein TM region although it was similar in both leaflets. Closer examination of the protein/CHOL interactions suggested that cholesterol interacted mainly with leucines, isoleucines and valines in the integrin TM region however AT simulations are required to study the details of theses interactions. Cholesterol was

shown previously to interact with TM helices (452-454) and regulate receptor activity (283, 455). The presence of a high density of cholesterol around the integrin TM region may also regulate the integrin TM region tilt angle in the membrane. A similar asymmetric distribution of lipids between the two leaflets was also observed in the *int_tal-2* system. Interestingly, the high-density ring of POPC lipids around the integrin TM region in the inner leaflet (in the *int_tal-1* simulation) was replaced by a similar high-density ring of POPS lipids. This ring was uniformly distributed around the TM region, indicating that the POPS distribution is affected not only by the F3 acidic loop which is close to the TM region but also by the positively charged residues in the integrin TM region, which were situated in that region (i.e. αIIb K994, R995, R997 and K989 and β3 K716, R723, K724, K725 and K728 residues). The stronger and more stable interactions of the F2 domain with the bilayer in this simulation (see above) resulted in a second high-density ring of POPS lipids around the F2 domain which was not present in the *int_tal-1* simulation (Fig. 6.5). The distribution of CHOL was highly correlated in the *int_tal-1* and *int_tal-2* simulations.

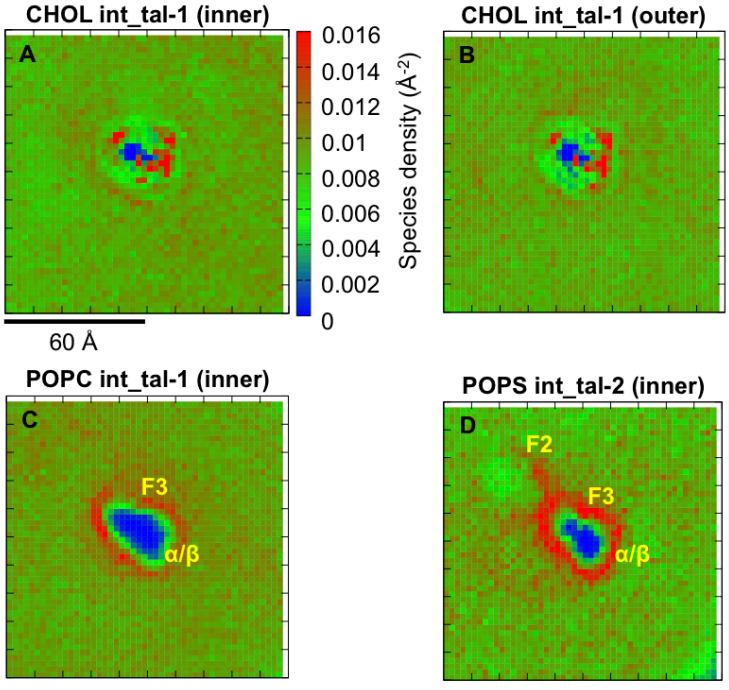


Figure 6.5: Two-dimensional average lipid headgroup densities around the integrin receptor for the CHOL molecules from the *int_tal-1* simulation (A,B), the POPC lipids from the *int_tal-1* simulation (C) and the POPS lipids from the *int_tal-2* simulation (D). The diagram shows the probability of finding the lipid at a given point in the bilayer plane. Blue represents low probability through to red for high probability. In the *int_tal-2* system (D), a second high-density region was observed, which reflects the strong interactions of the F2 domain with the lipids.

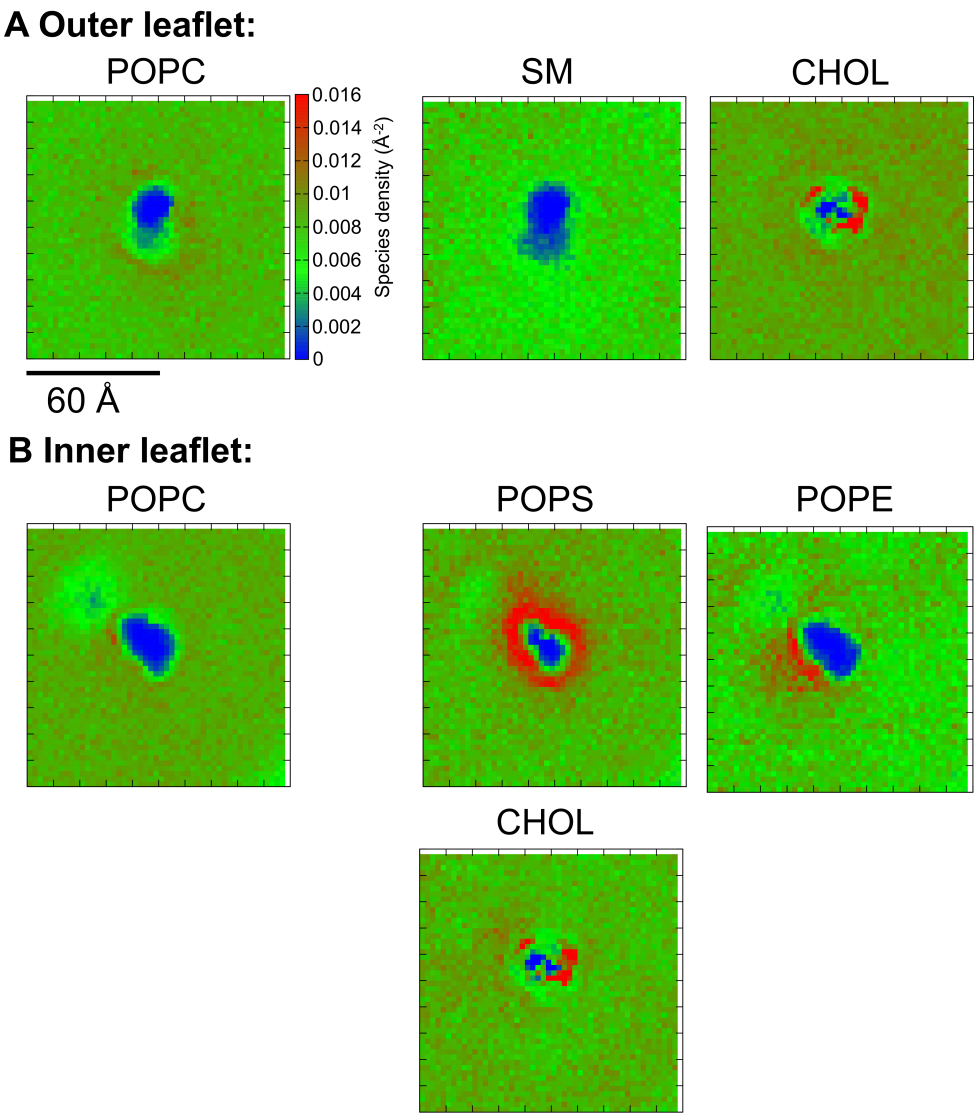


Figure 6.6: A. Two-dimensional average lipid headgroup densities around the integrin receptor for the outer (A) and inner (B) leaflets in the *int_tal-3* simulation. The diagram shows the probability of finding the lipid at a given point in the bilayer plane around the protein. Blue represents low probability through to red for a high probability.

Comparison of the two-dimensional average lipid headgroup densities of the outer and the inner leaflets of the asymmetric bilayer with the same leaflets in the *int_tal-1* and *int_tal-2* simulations (i.e. the outer leaflet of the *int_tal-3* simulation was compared with the outer leaflet of the *int_tal-1* simulation and the inner leaflet of the *int_tal-3* simulation was compared with the inner leaflet of the *int_tal-2* simulation) revealed similar distribution for all lipids (Fig. 6.6). Therefore, the organization of each leaflet in the asymmetric bilayer closely resembled that of the same leaflet in the two symmetric bilayers described above.

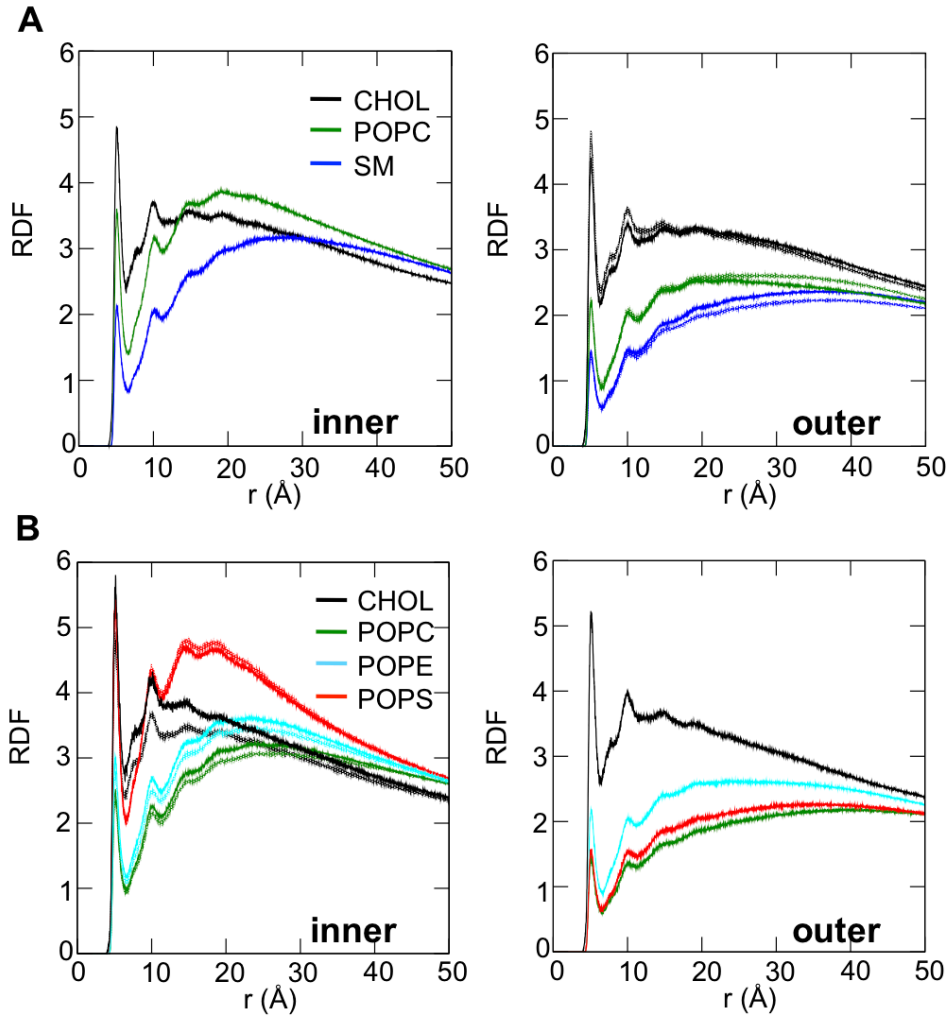


Figure 6.7: A,B. The lipid radial distribution function for the *int_tal-1* (A) and the *int_tal-2* (B) simulations. The distribution for each lipid type is shown separately for the outer (right) and the inner (left) leaflet. The radial distributions were calculated around the integrin TM region. The dotted lines in the outer leaflet of the *int_tal-1* simulation (A) and in the inner leaflet of the *int_tal-2* simulation (B) represent the radial distribution function for lipids in the outer (A) and inner (B) leaflets of the *int_tal-3* simulation.

To quantify the behaviour of the lipids around the integrin TM region, the radial distribution functions of the lipids relative to the protein TM region were calculated (Fig. 6.7). For the outer leaflet in the *int_tal-1* simulation, a large peak at ~5 Å was observed for CHOL, indicating its preference to accumulate close to the TM region. The same peaks (i.e. close to the integrin TM region) were lower for POPC and much lower for SM, suggesting a preferred pattern of lipid distribution close to the integrin TM region (i.e. CHOL>POPC>SM). The peak for CHOL had the same height for both leaflets, indicating that CHOL distribution was not affected by the presence of the CT protein or the ectodomain. In the inner leaflet, the first peak for the POPC and SM lipids was followed

by a second peak at ~20 Å, where the talin F2 domain was situated. This indicates that, despite the transient interactions between the talin F2 domain and the lipids (see above), the POPC/SM distribution was affected (Fig. 6.7A). In the outer leaflet, no second peak was observed, reflecting the fact that there is no direct interaction between the bilayer and the integrin ectodomain.

In the *int_tal-2* simulation, where a different bilayer (i.e. POPC/POPE/POPS/CHOL) was used, in the outer leaflet the high density of CHOL around the integrin TM region was followed by lower densities of all phospholipids (POPC, POPE and POPS). The density of POPE lipids was higher than POPC and POPS, suggesting the following pattern for the lipid distribution around the integrin TM region: CHOL>POPE>POPC>POPS. This lipid distribution pattern was modified in the inner leaflet since the density of negatively charged lipids (POPS) close to the TM region was similar to the CHOL density (Fig. 6.7B). The high density of POPS was a result of interactions with the positively charged loop in the talin F3 subdomain, which was situated close to the integrin TM region (i.e. it interacted with the TM domain) and the cluster of positively charged residues in the integrin TM region. Similar to the *int_tal-1* simulation, the first peak for POPS in the inner leaflet was followed by a second high peak at 20 Å (which was higher compared to POPC/SM peaks at the same area in the *int_tal-1* simulation), suggesting a strong preference for the CT protein to interact with negatively charged POPS lipids.

Comparison of the lipid radial distribution functions observed in the simulation with the asymmetric bilayer (*int_tal-3*) with those observed in the simulations with the symmetric bilayers revealed very similar distributions for all lipid types in both leaflets (Fig. 6.7). This augments the previous observation that the lipid organization of each leaflet in the asymmetric bilayer was similar to that of the same leaflet in the two symmetric bilayers described above. Therefore, each leaflet modifies its lipid organization separately to accommodate the shape and the interactions with the protein domains. Note that in the simulations discussed in this chapter all the lipid models maintain the key characteristics of their tails (e.g. saturation) and thus the association of different lipid tails is also taken into consideration in the lipid organization. For the simulations with the F2-F3/αβ complex, the distribution of lipids was similar (as above) in all cases, reflecting the fact that no direct stable interactions between the lipids and the ectodomain occurred.

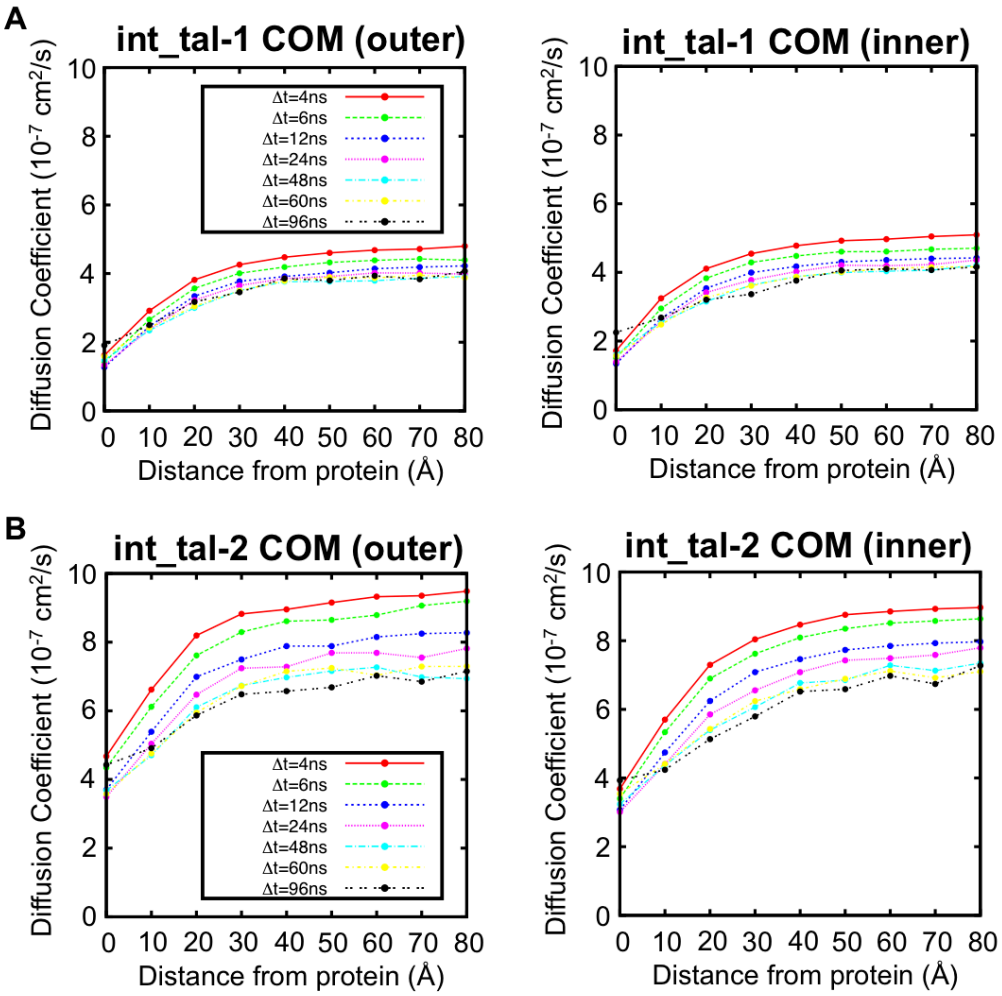


Figure 6.8: A,B. Diffusion coefficient of the centre of mass (COM) of the lipids in the outer (left) and inner (right) leaflets of the *int_tal-1* (A) and *int_tal-2* (B) simulations as a function of distance from the protein centre of mass over different timescales (Δt).

6.3.3 The integrin/talin complex alters lipid mobility

The local clustering of lipids because of their interactions with the integrin/talin complex (see above) is also anticipated to alter lipid mobility. To examine this further, the lipid lateral diffusion coefficient (*D*) in the two leaflets was calculated separately. In this calculation, the lateral diffusion coefficient was a function of different timescales (Δt) to capture all the different modes of diffusion (those dominating at short timescales and those dominating at longer timescales). At longer timescales (Δt), this calculation is expected to yield converging values for the lateral diffusion coefficient (*D*). The diffusion coefficient (*D*) can be also measured experimentally using electron paramagnetic resonance (EPR) with spin-labelled lipids (426-428) or fluorescence measurements (282, 408, 409) (see Section 6.4 for more details and examples). Overall, this calculation revealed a value of ~10^{-7} cm^2/s for all lipid components in all the simulations. It should be noted that

comparison of the diffusion coefficient between the MARTINI force field and FRET experiments for simpler lipid bilayers (319, 429) revealed a speeding up of the diffusion coefficient by a factor of ~4 when the diffusion coefficient was calculated with the MARTINI force field. This implies that this factor should be applied in these simulations; however, due to the fact that in this study mixed bilayers were used with a protein component (and thus this factor is expected to vary significantly between the different lipid types (287)), all the values of the diffusion coefficient are presented without any scaling. Interestingly, calculation of the diffusion coefficient relative to the distance from the integrin/talin complex revealed a reduced diffusion coefficient for lipids in a layer around the complex, thus suggesting a slowing-down effect for lipids. The effect was strongest for the region <10 Å around the protein and less for ~25 Å around the protein. For greater distances from the protein the diffusion coefficient reached a plateau (Figs. 6.8 and 6.9).

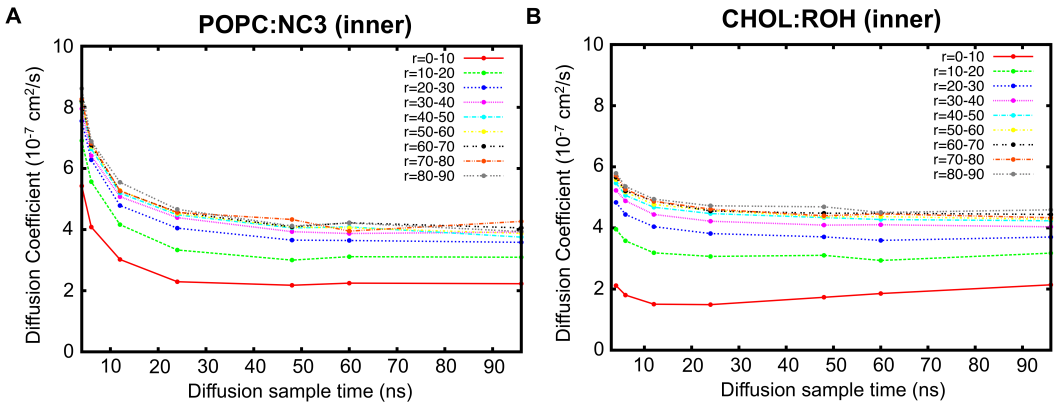


Figure 6.9: A,B. Diffusion coefficient (in layers of 10 Å radii around the protein) of the POPC (A) and CHOL (B) headgroups in the inner leaflet of the *int_tal-1* simulation as a function of observation time.

To characterize this effect more quantitatively, the space around the protein was separated in layers of 10 Å radii each and the diffusion coefficient was measured in each annulus (Fig. 6.9). For the inner leaflet of the *tal_int-1* simulation, this analysis showed that the diffusion coefficient of the POPC headgroups was ~2.5x10^{-7} cm^2/s in the first annulus (0-10 Å from the protein), ~4.25x10^{-7} cm^2/s in the second annulus (10-20 Å) and ~6x10^{-7} cm^2/s in the last annulus (80-90 Å). In the same regions, the CHOL diffusion coefficient was ~1.75x10^{-7} cm^2/s (0-10 Å), ~3.75x10^{-7} cm^2/s (10-20 Å) and ~5.75x10^{-7} cm^2/s (80-90 Å). The SM diffusion coefficient was ~3.75x10^{-7} cm^2/s (0-10 Å), ~4.75x10^{-7} cm^2/s (10-20 Å) and ~6.5x10^{-7} cm^2/s (80-90 Å), demonstrating a slowing effect for all lipid

species in the system. Equally, in the *int_tal-2* and *int_tal-3* simulations where a different type of bilayer was used, a slowing effect of lipids in the same regions was also observed.

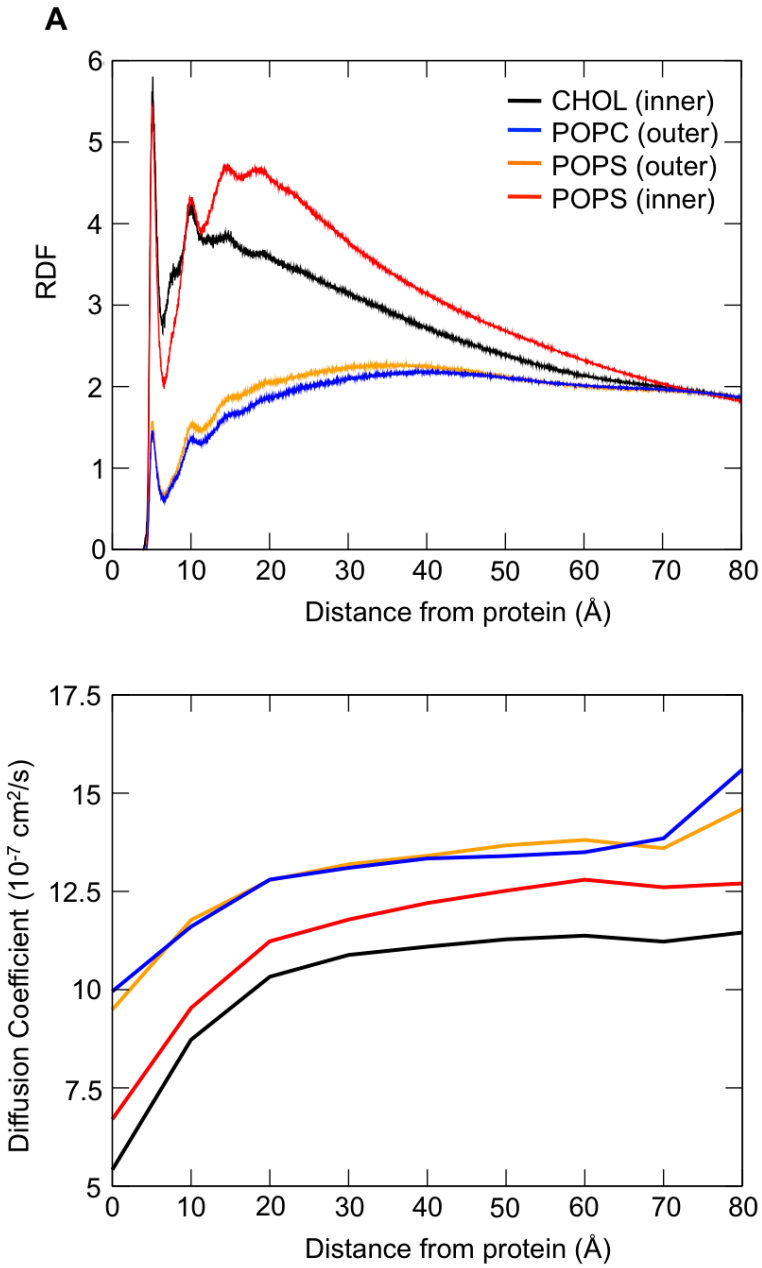


Figure 6.10: A. The lipid radial distribution function for the POPC (blue), CHOL (black) and POPS (inner leaflet-red, outer leaflet-orange) lipids from the *int_tal-2* simulation. The diffusion coefficient of the headgroups of the same lipids is shown as a function of distance from the protein COM.

Intriguingly, the lipid mobility profile in each simulation system can be explained in correlation with the lipid organization (i.e. radial distribution function; Fig. 6.10) around the integrin/talin complex. Overall, in the area around the integrin/talin complex (0-25 Å) where the lipid radial distribution function analysis (see above) revealed higher values (i.e.

peaks), indicating association of the lipids with the integrin TM region and talin, the motion of the lipids was lower compared to the regions away from the complex where lower values for the radial distribution function were calculated (i.e. no peaks observed). Above ~35 Å around the protein where the radial distribution function does not change, the lipid diffusion coefficient reaches a plateau. Lipid types whose radial distribution profile had higher value (i.e. peaks) closer to the protein and thus they were strongly associated with the integrin/talin complex, they had lower mobility compared to lipids which their radial distribution profile revealed lower values. For example, CHOL, the radial distribution profile of which had a high peak at the 0-10 Å region around the protein, had a lower diffusion coefficient compared to lipids in the same system such as POPC, whose radial distribution profile had lower peaks in the same area (Fig. 6.10).

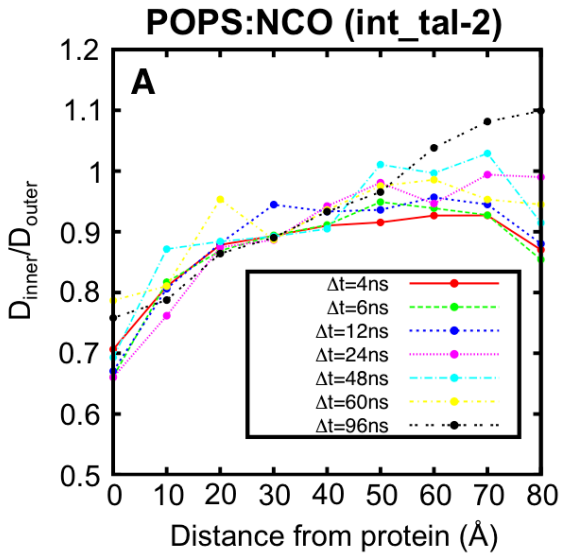


Figure 6.11: A. Diffusion coefficient ratio of the POPS headgroups for the inner and the outer leaflets in the *int_tal-2* simulation as a function of distance from the protein over different timescales (Δt).

Additionally, an asymmetry in lipid diffusion between the two leaflets was observed when lipids (e.g. the POPS lipids in the *int_tal-2* system) had very different radial distribution profiles in the inner (i.e. with two peaks) and the outer leaflet (i.e. with one peak). In particular, calculation of the diffusion coefficient of the POPS headgroups in the inner leaflet (where they form stronger interactions with the integrin/talin complex) gave a value of ~$5x10^{-7}$ cm^2/s. On the other hand, the diffusion coefficient for the POPS for the same region in the outer leaflet was ~$8x10^{-7}$ cm^2/s, suggesting that POPS mobility was slower in the inner leaflet because of its strong interactions with the F3 domain acidic

loop (Fig. 6.11). The diffusion coefficient profile for lipids with the same radial distribution profile (and thus with similar interaction strength with the protein) was similar in both leaflets (Fig. 6.10). Calculation of the mean interaction time (i.e. the mean time which a single lipid was continuously in contact with the protein during the simulation) revealed transient protein/lipid complexes with rapid exchange of lipids. In particular, calculation of the mean interaction time of the POPC/protein interactions for the *int_tal-3* revealed a value of 27 ns. The same calculation revealed 70 ns for the POPS, 25 ns for SM, 27 ns for POPE and 45 ns for CHOL. Similar fast exchange of lipids (which depends on the type of lipids and the topology of the membrane proteins (430)) was previously suggested by EPR and fluorescence data (431). See Section 6.4 for more detailed discussion.

The crucial role of the protein/lipid interactions in lipid mobility (see above) also indicates that integrin/talin mobility may also affect the motion of the lipids. To test this further, a simulation was also performed where the integrin/talin complex was restrained (*tal_αβ-4*). Since the previous results in this chapter suggested that the ectodomain does not affect the lipid profile, the F2-F3/αβ complex (i.e. without the ectodomain) was used. In this simulation, the mobility of the lipids was somewhat lower compared to the *tal_αβ-1* simulation (Fig. 6.12A). The reduction in lipid mobility was larger < 25 Å around the protein.

In addition to the interactions of the lipids with the protein, lipid mobility (especially around the integrin/talin complex) may also be regulated by the lipid–lipid interactions. The lipid–lipid interactions involve large number of H-bonds and thus AT-MD simulations are needed to study them in detail. However, from the simulation systems in this chapter it is clear that the levels of CHOL in the bilayer, and thus its interactions with the lipid tails, altered lipid mobility significantly. For example, the values of the diffusion coefficient for all lipids were higher in the *int_tal-2* system where the levels of CHOL were lower (i.e. 254 molecules) compared to the *int_tal-1* system (i.e. 474 CHOL molecules). This indicates that the reduction of CHOL in the bilayer significantly enhances lipid mobility. To test this further, another simulation was set-up where the amount of CHOL in the bilayer was reduced (*tal_αβ-5*). In this simulation, the lipid diffusion coefficient was approximately twofold higher for all lipid species in the vicinity of the protein and 1.5-fold higher further away from the protein compared to the *tal_αβ-1* simulation (Fig. 6.12B). This demonstrates that the amount of CHOL and thus the CHOL–lipid association also affected lipid mobility in the system studied.

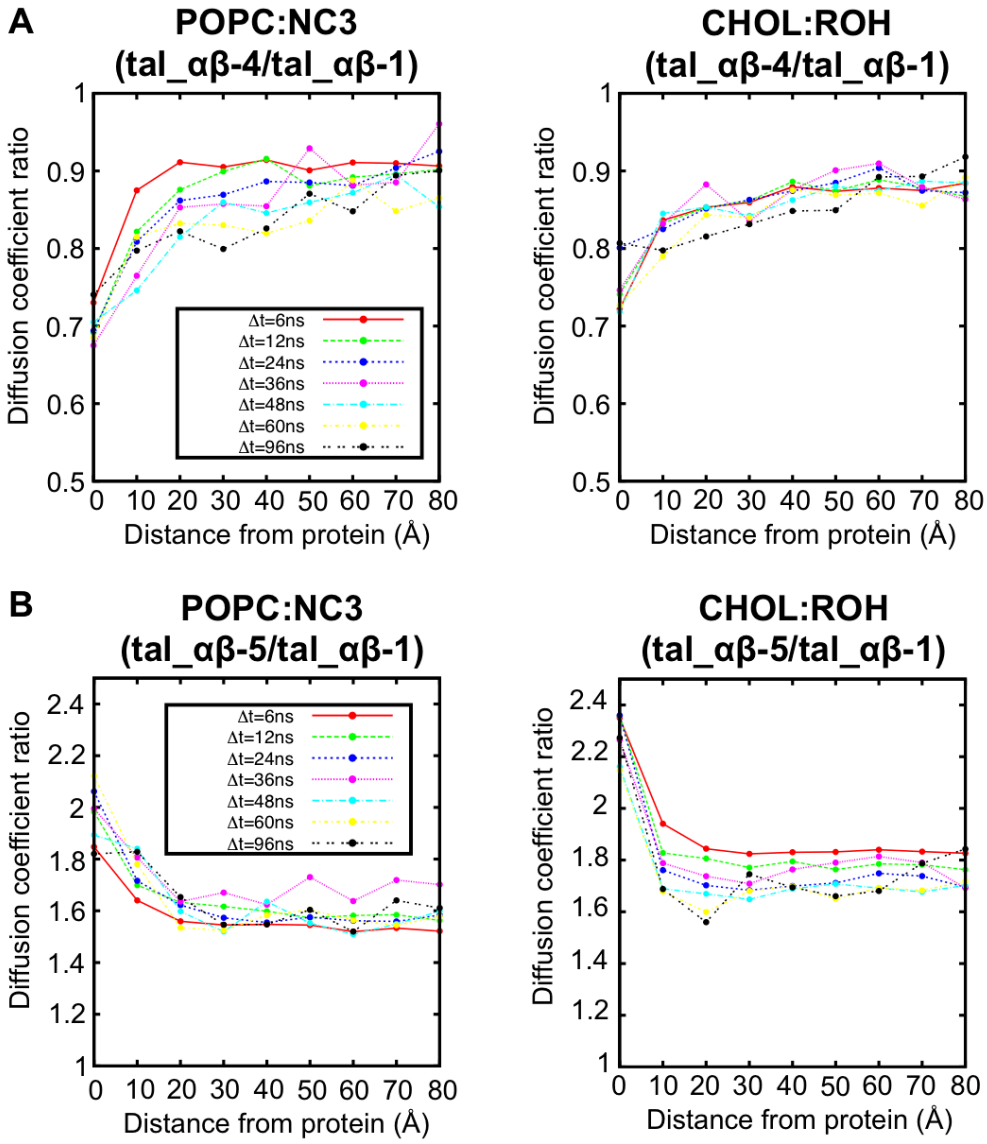


Figure 6.12: A,B. Diffusion coefficient ratio of the POPC (left) and the CHOL (right) headgroups for the *tal_αβ-4/tal_αβ-1* simulations (A) and the *tal_αβ-5/tal_αβ-1* simulations (B) as a function of the distance from the protein COM over different timescales (Δt).

6.4 Discussion and conclusions

In this chapter, extended simulations were used to probe the dynamics of the integrin/talin complex in more complex bilayers (which resemble real plasma membranes) and its effect on lipid structure and dynamics. The results illustrate the dynamic nature of the integrin ectodomain and demonstrate the prominent role of the integrin/lipid interactions in membrane leaflet structure and dynamics. In particular, the results suggest asymmetric lipid dynamics and organization between the bilayer leaflets in the presence of

the integrin/talin complex and a significant slowing effect for the lipids in the vicinity of the protein, which is due to the formation of transient protein/lipid complexes.

The first outcome of the work detailed in this chapter is that the presence of the integrin/talin complex modifies the lipid profile in all types of bilayers examined. In the simulations, depending on the nature/strength of the protein/lipid interactions, the organization of the lipids in each leaflet was altered to accommodate the interactions with the protein domains. This change in the lipid profile resulted in an asymmetric distribution of lipids between the leaflets of the bilayer. Experimental studies have suggested that other membrane proteins such as rhodopsin induced reorganization of phospholipids in rod disc membranes (432). Lipid reorganization was also observed after the binding of membrane-active peptides to the membrane (433), after clustering of proteins on the bilayer surface (434) and after binding of drugs molecules (435). Additionally, computational approaches suggested reorientation of lipids in one leaflet upon binding of nanosubstrates (436) to the lipid bilayer. The asymmetric distribution of the two leaflets is expected to have implications in biological processes (437), since previous studies have shown that leaflet asymmetry is important for the regulation of permeability in the membrane (438), for the accommodation of external moieties (436) and for membrane stability (439).

The diffusion rate in all simulations was ~10^{-7} cm^2/s, which is typical for the lipid fluid phase as suggested by computational and experimental results (14, 284, 440-442). The various protein/lipid interactions, however, limited the motion of the lipids in an annulus of ~25 Å around the protein, suggesting a slowing-down effect for lipids in the vicinity of the protein. Lipids around a protein often are referred to as "annular" lipids. The concept of a lipid annulus around membrane proteins is supported by electron paramagnetic resonance (EPR) studies with spin-labelled lipids, where it was possible to detect the slower mobility of lipids in membranes with protein domains (426-428). For example, lower lipid mobility was observed around $C\alpha^{2+}$-ATPase (443), Na^+/K^+-ATPase (444) and K^+ channel associated peptides (445). This concept is further supported by simulation studies (446) and by crystal structures obtained with lipids bound on the protein surface (242, 243). Of course, the lipids which are bound to crystal structures are usually very limited (447), suggesting that they are a small component of the protein/lipid complex. MD simulations have previously been used to study the details of the protein/lipid and lipid/lipid interactions (282-284). Therefore, these crystal structures, combined with MD simulations, can give valuable insight into how different lipid types associate with proteins and how these interactions regulate the function of the protein.

Despite the fact that the TM dimer surface is relatively smooth, the slowing effect observed in the simulations here was relatively strong. Larger proteins or proteins with well-defined binding surfaces might further enhance this phenomenon (446). Analysis of the protein/lipid interactions in the simulations suggested that the interactions between the lipids and the protein are rapid and transient. This is in good agreement with EPR and fluorescence data, which also suggested fast exchange of lipids despite the slowing effect in the annulus around the protein (431). The exchange rate depends on the type of lipid and the topology of the membrane proteins. For example, fluorescence experiments on mechanosensitive channels demonstrated that anionic lipids exchange more slowly than neutral lipids because of their interactions with a charged cluster in the channel (430). This observation raises the idea that the protein and the adjacent lipids form transient complexes and move in a concerted manner in the membrane (446). This, in combination with the fact that membranes are highly crowded (up to 30% proteins) (448, 449), suggests that the number of free lipids in cell membranes is limited due to the formation of various protein/lipid complexes.

Another interesting outcome is that CHOL concentration regulates lipid mobility in the membrane. In the simulations, reduction of CHOL concentration resulted in an approximately twofold increase in the diffusion coefficient for all lipid components in all regions of the bilayer. This result is in good agreement with experimental results, which suggest that CHOL increases membrane rigidity and reduces lipid mobility (441, 442, 450). Closer examination of the lipid distribution around the protein also revealed a high density of CHOL close to the integrin TM region in both symmetric and asymmetric bilayers. This is in good agreement with experimental results which suggested that the presence of integrins affects CHOL distribution (451). Additionally, previous data suggested that CHOL interacts with TM helices (452-454) and regulates receptor activity (283, 455).

The simulations also confirm the electrostatic nature of the talin/bilayer interactions. Indeed, in the simulations the interactions between the talin and the bilayer were stable only when negatively charged POPS lipids were in the inner leaflet. This observation is consistent with experimental results (125) and the outcome of chapters 3 and 5 of this thesis, which suggests that negatively charged moieties are crucial for the talin/membrane interactions and thus for integrin activation. The fact that in chapters 3 and 5 of this thesis a different type of negatively charged lipids was used (i.e. POPG) augments the observation from experimental studies that talin/lipid electrostatic interactions are

regulated by the net negative charge of the membrane and not the lipid type (i.e. POPS, POPG) involved (186). Moreover, the simulations also demonstrate the dynamic nature of the integrin ectodomain, even in the inactive state. The flexible linker between the TM region and the integrin ectodomain allows large fluctuations of the ectodomain. The ectodomain only interacted transiently with the bilayer in the presence of negatively charged lipids.

In summary, this study demonstrates the dynamic nature of the integrin receptor and suggests that its presence has a significant effect on the structure and dynamics of membrane lipids. Other receptors or membrane proteins (22, 456) exhibit similar topology to the integrin receptor (i.e. having a dimeric TM region and multi-domain ectodomain) and therefore the results concerning both the dynamics of the receptor and its effect on the structure and dynamics of the bilayer leaflets are of particular biological significance because they can be generalized to other receptors.

Chapter 7

Conclusions and Future Directions

7.1 Conclusions

Prior to this thesis, studies of integrin inside-out activation mainly focused on the structural aspects of talin and integrins and the identification of key residues and interactions. Structural data that emerged after 2009 included the αIIb/β3 TM dimer (118), the F2-F3/β1D complex (125), the talin F0-F1 domain pair (194) and the complete talin head domain (186). These significantly enhanced our understanding of the structural basis of inside-out activation. However, inside-out activation is a very dynamic process involving a large membrane-spanning protein, the membrane and talin; thus, many details of the mechanism remained incomplete. The available structural data for the various parts of integrins and talin did, however, allow the application of simulation techniques to unravel the dynamic details of the integrin inside-out activation process. A limitation is that the timescales which can be addressed by all-atom simulations are very limited for larger protein complexes; therefore, a multiscale approach was employed using a combination of AT-MD and CG-MD simulations. The use of CG-MD simulations allows extended timescales to be explored and the subsequent use of AT-MD simulations allows refinement of the system, thus giving a more detailed view of its properties and interactions.

Recent experimental data suggested that talin interactions with negatively charged lipids in the membrane are crucial for inside-out activation (125, 186), but a molecular model of the talin/membrane association was missing. The individual roles of the four subdomains and the negatively charged F1 loop in the 50 kDa talin head domain were also uncertain. In Chapter 3, the multiscale simulation approach was used alongside NMR experiments to probe the association and orientation of the talin head domain on the membrane. The results of Chapter 3 demonstrate that flexibility within the talin head domain and the interactions between talin and the negatively charged lipids are crucial for the productive association of talin with the membrane. Mutations which disrupt these electrostatic interactions result in perturbed binding of talin to the membrane, which in turn is expected to perturb the integrin/talin association. On the basis of these results, a molecular model for the talin/membrane association has been proposed. In this model the

talin head domain binds to the membrane initially via a positively charged loop in the F3 domain and subsequently via a positive patch in F2, orienting talin in a way that facilitates binding to the β CT integrin tail. After the F2-F3 domain pair binds, the F0-F1 domain pair interacts with the membrane via an additional interaction created by the F1 loop. Upon binding, the F0-F1 pair reorients by ~65^{o} relative to the F2-F3 pair in a plane perpendicular to the bilayer normal, allowing talin to adopt a compact V-shape. This novel V-shape conformation of talin is likely to be of mechanistic importance since it optimizes the talin/lipid association and promotes a more stable talin/membrane complex. This may be important for how talin regulates the conformational changes of different types of integrins.

Subsequently, the talin head domain (and the isolated F2-F3 pair) was used to construct a talin/αβ complex which allowed elucidation of how talin modulates the TM region of integrins. Based on this study, a model for how talin promotes integrin inside-out activation has been proposed here. According to this model, strong talin/negatively charged lipid interactions promote reorientation of the talin head domain in a plane perpendicular to the bilayer normal, which disrupts the interactions in both the αIIb/β3 IMC and OMC regions of the membrane-spanning TM helices. The weakened αIIb/β3 interaction, in combination with a large increase in the tilt of the β TM helix, results in a scissoring movement of the TM part of the two helices with a modified IMC at the centre of the scissors. This movement increases the separation of the extracellular part of the TM region of integrins, thus providing a plausible mechanism for the induction of conformational changes in the ectodomain. This study also demonstrated the stability of the 'off' state and the key roles of salt bridges and residues F992 and F993 in the integrin αIIb subunit in stabilizing this state.

The integrin TM region is a rather more complex system than is found in many TM regions of signalling receptors due to its heterodimeric nature, although it has a Gx_3G motif. Therefore, it provides a valuable test of computational approaches which model the TM dimerization of membrane receptors. In Chapter 4, the multiscale approach, as used in Chapter 3, was extended with a high-throughput pipeline. Comparison of the predominant binding mode after refinement with AT-MD simulations revealed very similar packing of the integrin helices to that found in a known NMR structure. Disruptive mutations in the Gx_3G motif perturbed the helix–helix packing, thus demonstrating the crucial role of the Gx_3G motif in the packing of the integrin TM helices.

In Chapter 6, an integrin/talin complex was constructed and used to probe the dynamics of this complex and its effect on lipid structure and dynamics in more complex bilayers closely resembling real plasma membranes. The results illustrate the dynamic nature of the integrin ectodomain even in the inactive state and demonstrate the prominent role of the integrin\lipid interactions in the structure and dynamics of membrane leaflets. In particular, the results suggest asymmetric lipid dynamics and organization between the bilayer leaflets in the presence of the integrin/talin complex and a significant slowing effect for the lipids in the vicinity of the protein due to the formation of transient protein/lipid complexes.

From a more methodological perspective, this thesis demonstrates that a computational multiscale approach can predict various aspects of membrane and peripheral proteins. Chapter 3 demonstrates that this approach can be used to predict the position, orientation and interactions of a protein domain which binds to a lipid bilayer. The combination of the predictive nature of the CG-MD simulations, which can identify the orientation of a peripheral membrane protein complex with a bilayer, and the subsequent use of AT-MD simulations, which allow the refinement of protein/lipid interactions, is a powerful way of studying the nature of the interactions. The size and characteristics of the complex (i.e. four flexibly connected domains and a large flexible loop) added an extra level of complexity to the system compared to previously studied systems (274, 457). Therefore, this study is a good demonstration that the serial multiscale approach can be used to identify the protein/lipid interactions of more complex systems, and to provide novel information about features like flexibility between the domains and their role in lipid interactions.

The development of an automated high-throughput approach offers a convenient way of running large ensembles of simulations, thus ensuring better sampling in CG-MD simulations and better statistics. Chapter 4 demonstrates that the predictive nature of CG-MD simulations in combination with all-atom simulations also gives valuable information about the packing of membrane protein TM regions in a lipid environment. By using this approach, it was possible to predict a structure with similar accuracy to that obtained by NMR. This purely computational method can therefore be applied to other TM dimers and should provide structural insights into the assembly of membrane protein TM regions. Chapter 5 is a good demonstration of the strength of the multiscale simulation approach for obtaining information about the position, dynamics and orientation of membrane receptors,

which allows construction of models that show how protein/membrane interactions can modulate potential changes in receptor ectodomains.

Overall, this thesis offers novel insights into the dynamics of talin-mediated integrin activation and proposes a novel mechanism which explains how talin, the membrane and the integrin interact to promote integrin activation. Additionally, it demonstrates that a multiscale approach can be used to give insights into various aspects of membranes and membrane-related proteins.

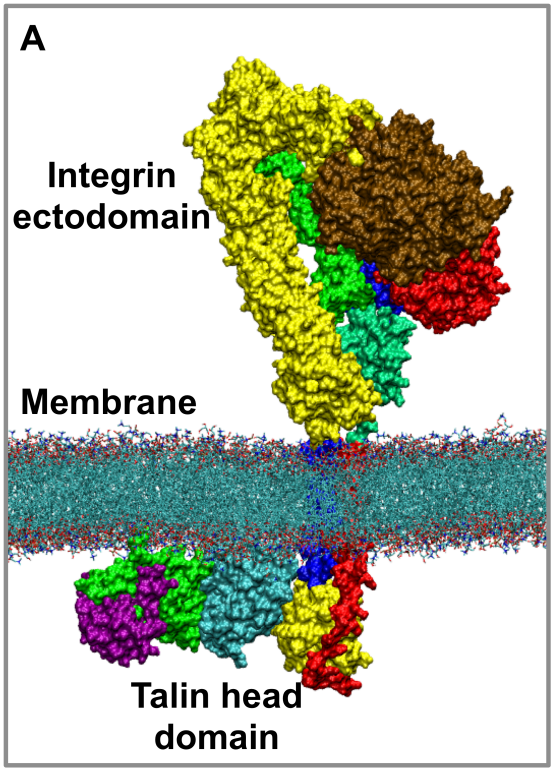


Figure 7.1: A. Model of the complete integrin receptor in complex with the complete talin head domain, inserted in a lipid bilayer. The talin F0 domain is shown in purple, F1 in green, F2 in cyan and F3 in yellow. The receptor is in a bent, inactive state.

7.2 Future directions

Despite all the available structural data for the component parts of integrins (52, 78, 79, 118, 119, 125), the structure of the complete integrin receptor is still not available. Therefore, in this thesis, component parts of the integrin system were used to study different aspects of the activation mechanism. However, the use of relatively small systems cannot give information about all stages of the activation mechanism and thus answer important questions such as the activation states of the complete integrin receptor. Therefore, a simulation of a larger system which contains the complete integrin receptor in

the presence of its activating protein (i.e. talin) is needed to obtain this information (see Fig. 7.1). In Chapter 6, CG-MD simulations of the integrin/talin complex were performed but the elastic network used to model the secondary and tertiary structure "locked" the receptor in its inactive conformation and therefore no information about the active conformations of integrins could be obtained. However, the model in Chapter 6 could be used to convert the system back to an all-atom representation. This AT system would allow the study of all the changes within the integrin during the inside-out activation mechanism. Of course, a major limitation will then be its size since it consists of ~1,500,000 particles with large demands on computational resources.

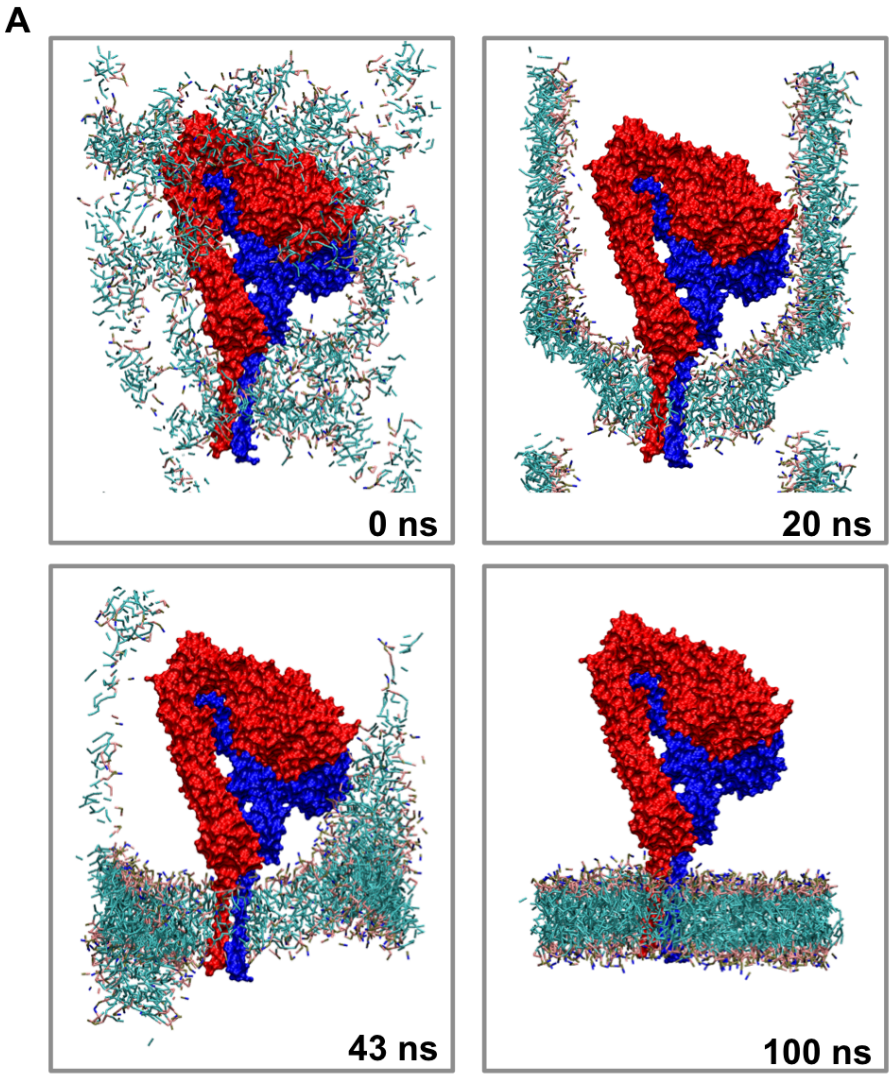


Figure 7.2: A. Progress of a simulation showing bilayer self-assembly around the integrin receptor. The integrin α subunit is shown in red, the integrin β subunit is shown in blue and the lipids (POPC and POPG) are shown in bonds format and are coloured as follows: cyan for the hydrophobic tails, pink for glycerol, brown for phosphate and blue for choline. Water and counter ions are omitted for clarity. The lipid layers at 20 ns are due to the periodic boundary conditions used in the simulations.

An initial investigation of a similar, albeit less complex, system has already been performed using the integrin receptor model (see Chapter 6) in complex with the talin F2-F3 domain pair. In this case, the complex was inserted in a POPC/POPG (ratio 3:2) bilayer using self-assembly CG-MD simulations (Fig. 7.2). After the insertion in the bilayer, the system was converted back to an AT representation. Using this model a ~500 ns simulation was performed on the Louhi supercomputer at CSC Finland. During the simulation, the talin F2-F3 domain pair interacted preferentially with negatively charged lipids, although no significant destabilization was observed in the integrin TM region. The ectodomain fluctuated in the solvent but again no major conformational changes occurred. This simulation forms the basis for further simulations to gain access to longer timescales and deeper insight into the integrin activation mechanism.

Recent studies suggest that kindlins also bind to the integrin β CT region and cooperate with talin to induce integrin activation. There is still relatively little structural information about kindlins but they are expected, from sequence comparisons and structural studies of some subdomains, to have similar characteristics to the talin head domain, although there are differences, such as a larger loop insertion in the F1 domain and an extra PH domain. However, using the talin head domain structure and knowledge of the PH domain structure it should be possible to construct initial models of the kindlin head domain and to study its interactions with integrins and the membrane. This model should also allow examination of how kindlins and talin bind to the CT region of integrins and cooperate together to regulate integrin affinity.

Recent studies also suggest that mechanical forces may regulate the activation state of integrins (78). Mechanical force may be induced extracellularly by ligand binding to the ectodomain or intracellularly by talin and/or kindlin attachment. Binding of talin to integrins provides a molecular linkage to the actin cytoskeleton, which could induce a lateral force to the integrin β subunit (78). The available structural information about integrins (79) and their importance make them key targets for state-of-the-art MD simulations, which can give new insights into possible conformational states and the effect of applied forces.

This thesis has focused on a limited number of integrins (mainly the αIIbβ3). There is considerable diversity within the integrin family and it will be particularly interesting to extend these studies to other types of integrins. Of particular interest is how talin regulates the tilt angle of other types of integrin β TM regions and what the molecular details of the TM region dimerization are in integrins which also contain the small-xxx-small structural

motif in their TM region. Finally, by using recent advances in experimental and simulation techniques, which allow the construction of more biologically relevant bilayers (see Chapter 6), the model for integrin inside-out activation proposed in this thesis could be further tested.

APPENDICES

APPENDIX A

A.1 Talin F0-F1 domain pair interactions with the membrane

In the first set of simulations, the F0-F1 pair without the F1 loop (Δloop-res: 134-172) was displaced 110 Å away from a preformed POPC/POPG (i.e. mixed zwitterionic/anionic) bilayer and five simulations were performed. Similar to F2-F3 (see main text), the F0-F1 pair diffused in the aqueous environment before associating with the bilayer in four out of five simulations. After the association, the F0-F1 remained bound in the same orientation until the end of the simulation. Inspection of the binding interface of the F0-F1 pair relative to the bilayer revealed non-specific binding, since in all four simulations in which the F0-F1/bilayer complex was formed the F0-F1 bound with a different pattern of interacting residues. Interestingly, mapping of the interacting residues on the complete talin head domain structure showed that these interfaces were situated on a different side compared to the productive F2-F3 binding surface (125, 186). The addition of the large loop in F1 with different preformed helical regions (h1 and h2) as described in Table A.1 (*F0-F1-l-CG*, *F0-F1-h1-CG* and *F0-F1-h2-CG*) again resulted in the non-specific and non-productive binding of the F0-F1 pair (Fig. A.1). In the presence of the F1 loop four out of five simulations, similar to the *F0-F1-CG* simulation above, yielded a F0-F1/bilayer complex, thus suggesting that the presence of the loop did not increase the F0-F1 affinity for the bilayer (Fig. A.1).

Table A.1: Summary of simulations

Simulation	Protein	Duration (ns)
F0-F1-CG	F0-F1(Δloop-res: 134-172)	5x1000
F0-F1-l-CG	F0-F1(loop-res: 134-172)	5x1000
F0-F1-h1-CG	F0-F1(loop-res: 134-172/helix-res: 144-167)	5x1000
F0-F1-h2-CG	F0-F1(loop-res: 134-172/helix-res: 154-167)	5x1000

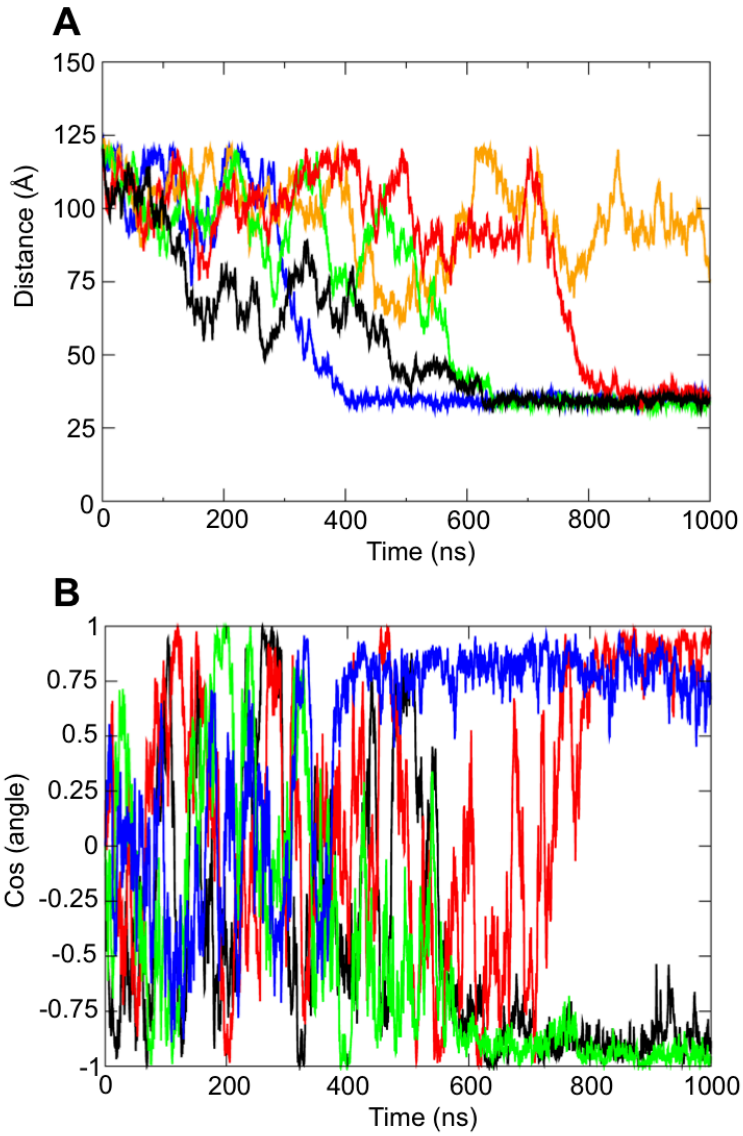


Figure A.1: A. Progress of the simulations with the F0-F1 domain pair with the F1 modelled insertion. The progress of the simulations is shown as the separation between the centres of mass of the protein and the bilayer as a function of time. The five different colours represent the five different simulations performed. B. The cosine of the angle between the plane of the protein and the bilayer plane is shown as a function of time for the four simulations that yielded a F0-F1/bilayer complex (shown using the same colour coding as in A). This analysis revealed that two different binding modes of the F0-F1 were observed in the simulations.

A.2 Complete talin head domain (F0 to F3) interactions with the membrane

A.2.1 The *tal-lF0F1-CG*, *tal-h2F1-CG* and *tal-h2F0-CG* simulations

The simulations with the talin head domain in solution (see *tal-sol-AT* in main text) suggested a tendency for the loop to be situated close to the F0-F1 on the same side as the F2 and F3 positively charged surfaces (Fig. 3.9B). To this end, simulations were set up where the loop (with or without any helical conformation) was restricted close to the F0 or the F1 domain in a similar form to that observed in the simulations with the talin in solution (*tal-lF0F1-CG, tal-h2F1-CG* and *tal-h2F0-CG*; Fig. 3.9B). To decide the orientation in which the helical region (i.e. h2) would be restricted close to the F0 or the F1 domain, simulations were set up with the isolated helical region displaced 80 Å away from the POPC/POPG bilayer. Five simulations were performed, with the helix starting from the same initial position but with different initial velocities. In all five simulations, the helix associated with the bilayer with the same pattern of interacting residues. In the models where the helical region was restricted close to the F0 and F1 talin domains, this surface was modelled to face towards the bilayer (Fig. 3.2).

In the simulations where the helical region was constrained to be close to the F0 domain, talin yielded a complex with the bilayer in four out of five simulations, similar to the simulation system where the loop with the helical region was modelled in a random configuration away from talin. In all four of these simulations, talin bound to the bilayer in a productive orientation, similar to the *tal-h2-CG* and *tal-l-CG* simulations (see main text), with the difference that the additional surface created by the loop between the F0-F1 pair and the bilayer now also interacted with the bilayer after the association of the F3 and F2 domains with the bilayer (Fig. 3.11). In particular, talin was associated with the bilayer initially via the F3 positively charged loop (res: 318-330), subsequently with the F2 positively charged patch (res: 255-285), and finally with the additional positively charged surface created by the loop between the F0-F1 pair and the bilayer. Interestingly, when this loop surface was associated with the bilayer the talin head domain adopted a V-shape configuration due to rotation of the F0-F1 pair relative to the F2-F3 pair in the xy bilayer plane. This F0-F1 displacement was approximately 65^{o} from its initial position (i.e. the crystal structure) (Fig. 3.1A). Calculation of the same angle relative to the distance from the centre of mass of the bilayer revealed that the angle between the F0-F1 and F2-F3 domains fluctuated between 0^{o} and 80^{o} when talin was in solution and it was stabilized to approximately 65^{o} after the association with the bilayer (Fig. 3.11B). The presence of the F1 loop between the F1 and the bilayer did not allow any direct interactions between the

rest of the F1 domain and the lipids. It must be noted, however, that some transient minor direct interactions of the F0 domain with the bilayer were observed (Fig 3.11C). This optimized binding mode was also obtained in the *tal-lF0F1-CG* simulations where the F1 loop was restricted in a similar position close to the F0-F1 with no helical conformation. In this case, however, only three out of five simulations resulted in a talin/bilayer complex compared to four out of five in the *tal-h2F0-CG* (above), suggesting that the formation of a small helical region close to the F0 domain prior to binding to the bilayer slightly increased the talin/lipid association. In contrast, in the simulations where the small helical region was restricted close to the F1 domain (i.e. the new binding surface did not include the F0 domain; *tal-h2F1-CG* in Fig. 3.2), the talin head domain was associated with the bilayer in four out of five simulations, although it was no longer able to productively bind the bilayer. In this case, the presence of an extra rigid positive binding surface (formed by the helix) next to the F2 patch promoted the binding of domains other than F3, which was shown above to be needed for the productive orientation of talin head domain (Fig. A.2).

Table A.2: Summary of simulations

Simulation	Proteins	Duration (ns)
tal-h2F0-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167 [helix restricted close to F0])	5x1500
tal-h2F1-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167[helix restricted close to F1])	5x1000
tal-lF0F1-CG	F0-F1-F2-F3 (loop-res: 134-172 [loop restricted close to F0-F1])	5x1500
tal-l4E-CG	F0-F1-F2-F3 (loop-res: 134-172 [4E])	5x1000
tal-lK324D-CG	F0-F1-F2-F3 (loop-res: 134-172 [K324D])	5x1000
tal-h2pc-CG	F0-F1-F2-F3 (loop-res: 134-172/helix-res: 154-167 [POPC bilayer])	5x1000

All simulations were performed using a preformed POPC/POPG bilayer except *tal-h2pc-CG*, where a preformed POPC bilayer was used.

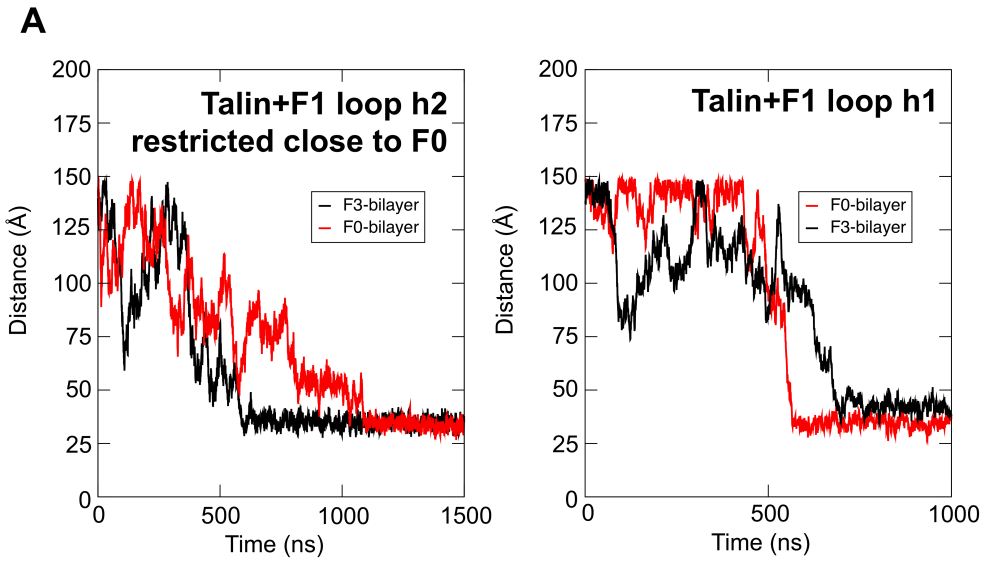


Figure A.2: A. Distance of the centres of mass of F3 (black) or F0 (red) from the bilayer centre of mass shown as a function of time.

A.2.2 The role of the negatively charged lipids in the talin/membrane interactions

All the simulations in the main text highlighted the crucial role of negatively charged lipids in the binding and interactions of the talin head domain with the membrane. To test this observation further, simulations were set up where the POPC/POPG bilayer used earlier was replaced by a pure POPC bilayer (i.e. with a net zero charge). In good agreement with previous observations, talin remained in the aqueous environment in all five simulations without forming a complex with the bilayer. Furthermore, additional simulations using the same POPC/POPG bilayer as above (see main text) were set up with mutations in the F3 acidic loop (*tal-K324D-CG)* or the F2 positively charged patch (*tal-l4E-CG*). Experiential studies have suggested that these mutations reduce the affinity of talin head domain to the membrane and reduce integrin activation (125, 186, 193). The simulations with the mutated form of talin are in good agreement with these studies, since in both simulation systems the talin head domain bound to the bilayer with reduced affinity (one out of five simulations in the *tal-l4E-CG* and two out of five simulations in the *tal-K324D-CG*; Fig. 3.12A). Furthermore, in all simulations in which the talin/bilayer complex was formed, talin bound to the bilayer in a perturbed orientation as shown in Fig. 3.12B (main text).

APPENDIX B

B.1 High-throughput method for studying protein/lipid interactions

This high-throughput pipeline was written in the Python programming language (http://www.python.org/) using GROMACS (303, 304) for setting up, running and analysing the simulations and Modeller (341, 342) for performing mutations on the protein of interest. The pipeline is separated into seven stages:

- *Stage 1*: The pipeline uses Modeller (341, 342) to perform any mutations on the protein the user has requested. Multiple mutations on the same protein can also been made.
- *Stage 2*: The protein is coarse-grained either using the MARTINI representation (320) or a local variant of MARTINI (321, 324), depending on the user's choice.
- *Stage 3*: The bilayer is constructed by using self-assembly CG-MD simulations. The user can choose from a variety of phospholipids, phosphoinisitol, sphingolipids and cholesterol when constructing the bilayer. In brief, the distance between all atoms in the protein is calculated and subsequently the largest distance is used to calculate the bilayer size in order to be large enough to accommodate the protein. The bilayer size is the protein largest length (Å) + 30 Å (in each side). When the bilayer size is calculated, lipids are randomly added to a simulation box, solvated, ionized and a simulation of a small portion of the bilayer is performed. After the end of the simulation the script checks if the bilayer was properly formed and replicates the bilayer patches to reach the final bilayer size. If the bilayer is not properly formed the whole procedure is repeated automatically up to five times. The user also has the option to add a preformed bilayer.
- Stage 4: The individual simulation systems are built depending on the number of simulations the user desires. In brief, the protein and the bilayer are inserted in a simulation box and the protein centre of mass is displaced 100 Å away from the bilayer. In each individual system, the protein is rotated around the x, y and z axes to form a new initial configuration. The system is solvated with CG water molecules, counter ions are added to neutralize the system and the system is energy minimized for 300 steps. After the energy minimization all the individual systems are equilibrated for 5 ns and production simulations are set up. The default timestep

for equilibration/production is 20 fs, as this helps maintain the temperature and energy distribution. Other time steps (from 10 to 40 fs) can also be used.

- Stage 5: Submit simulations. After constructing and equilibrating a large number of individual simulations the script connects to local (SBCB) clusters and runs the simulations in four or eight processor nodes. The simulations are submitted in batches of ten simulations. When running the simulations it checks whether the simulations have crashed and notifies the user if more than 50% of the simulations are incomplete.
- Stage 6: Data retrieval and storage. As soon as a batch of ten simulations finishes it moves all the data back to the local storage disks. This procedure is repeated until all the simulations are completed.
- Stage 7: Basic analysis. After all the data are retrieved, the pipeline performs a basic analysis by calculating the protein Cα RMSD and RMSF, the protein binding surface for different portions of the simulations and the protein contacts with each individual lipid type. All the data are available to the user, allowing them to perform additional analysis.

APPENDIX C

C.1 F2-F3/αβ, F3/β and F2-F3/β complexes modelling

For the construction of the F2-F3/αβ, F3/β and F2-F3/β complexes, the helical part of the β1D (PDB:3G9W) (125) and of the β3 subunits (PBD:2K9J) (118) were aligned using the backbone particles of the residues that overlap (based on sequence alignment). After the alignment, the β3 residues from T720 to F727 in the C-terminus were removed, thus keeping the equivalent β1D residues at the same positions. This resulted in a chimeric structure (β3 residues 688-719/β1D residues 753-787) for the β subunit (Fig. 5.1A). It should be noted that the TM regions of the β3 and β1D subunits have 90% sequence similarity and 52% sequence identity. Subsequently, the resulting model was energy minimized to remove any steric clashes. For the F2-F3/β and F3/β complexes, the integrin α subunit was removed after the modelling.

During the modelling no structural changes were allowed and thus all the F3/β and α/β interactions were retained, with the exception of the β3D723-αIIbR995 salt bridge. This salt bridge was situated at the region where the two structures were connected and therefore further investigation was required to evaluate its stability in the resulting structure (Fig. 5.1A,B). In the current model, the salt bridge was formed by the β1D D759 and the αIIbR995 residues (Fig. C.1A). To test its stability, four AT-MD simulations were performed (20 ns each) with the structure inserted in a POPC/POPG bilayer. In all four simulations, rapid disruption of the salt bridge was observed due to electrostatic interactions of αIIbR995 with residues M322-L328 in the positively charged loop of the talin F3 domain, which was situated close to the salt bridge (Fig. C.1C). An additional model of the F2-F3/αβ complex was prepared as described above, where the β chimeric peptide consisted of β3 688-727 residues and β1D 764-787 residues (Fig. 5.1B); in this case, the homologous salt bridge was formed between β3D723 and αIIbR995 residues (Fig. C.1B). The same rapid disruption of the salt bridge was observed in the simulations because of the interactions of the αIIbR995 sidechain with the adjacent charged loop in the talin F3 domain (res: M322-L328) (Fig. C.1D). This indicated that the salt bridge in the two models had similar behaviour and therefore the first model was used for the simulations.

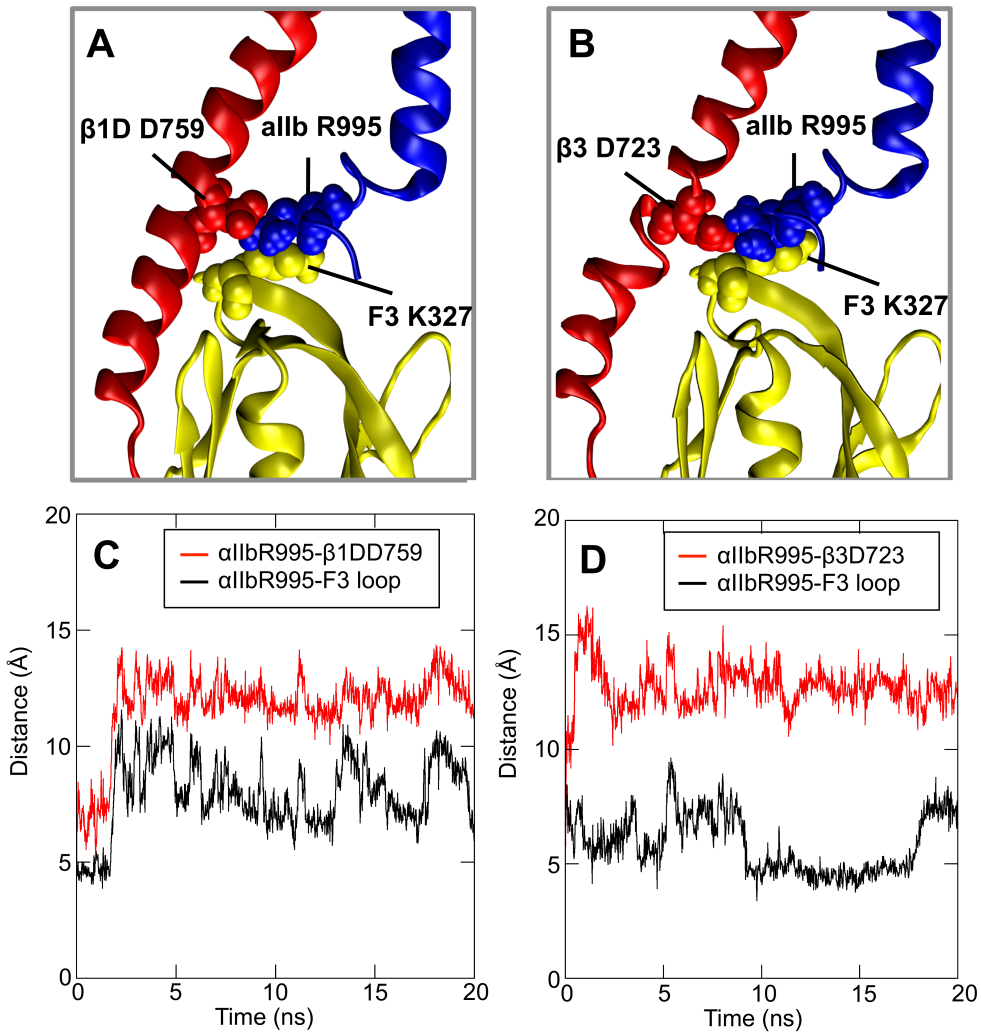


Figure C.1: A,B. Packing of the αIIbR995 (blue), β1DD759 or β3D723 (red) and F3K327 (yellow) sidechains in the two models constructed for the F2-F3/αβ complex. In A the salt bridge consisted of residues β1DD759 and αIIbR995, while in B it consisted of residues β3D723 and αIIbR995. C,D. Distance between the αIIbR995 and the β1DD759 residues (red/C) or between the αIIbR995 and the β3D723 residues (red/D) as a function of time. In both figures, the distance between the αIIbR995 residue and the F3 loop (i.e. residues K327-P330) is shown in black.

C.2 Talin head domain (Δloop-res: 134-172) in complex with the integrin α/β integrin TM region

In this set of simulations, a similar multiscale approach (see main text) was used to study the talin/αβ complex which was constructed using the talin head domain conformation obtained from the *tal-AT* simulation in Chapter 3 in the same POPC/POPG bilayer. In this configuration, the F2 and F3 subdomains interacted with the bilayer via their positively charged patches identified earlier, whereas the F0-F1 pair did not interact with the bilayer (Fig. 5.18A). Interestingly, at the end of the CG-MD simulations a similar displacement of the β integrin helix in the bilayer was observed as in the simulations in the main text (Fig. C.2B). Alignment of the talin F3 domain from the initial and the final snapshots of the five individual simulations suggest a conserved movement of the F2-F3 pair and a very dynamic movement of the F0-F1 pair during the simulations (Fig. C.2A). This dynamic movement may be due to the absence of the F1 flexible loop, which was shown to stabilize and optimize the F0-F1/bilayer interactions. While the F2-F3 pair retained its orientation relative to the bilayer during the simulations, the F1 domain formed direct interactions with the bilayer via the residue regions K89-I100, R118-I121, D125-L133 and H188-R196. Interestingly, positively charge residues in these regions, K89, R91, K98, R118, R131, H188, R194 and R195 formed the largest number of contacts with the bilayer (Fig. 5.18A). The F0 domain remained in the aqueous environment with rare transient interactions with the lipids. The optimum V-shape talin configuration was not observed in any of these simulations, indicating that the presence of the F1 loop is necessary for optimization of talin/lipid interactions. Nevertheless, productive F2-F3 pair interactions with the bilayer were sufficient to disrupt the interactions in the integrin TM region. This is in good agreement with experimental data, which suggest that the F3 domain alone is sufficient for activating the αIIbβ3 integrin (191).

Subsequently, an AT-MD simulation (100 ns) was performed with the final snapshot from one of the *αβ-tal-CG* simulations. In this system, the interactions in both the IMC and OMC regions were disrupted at the beginning of the simulation, thus representing an "open" integrin structure. Interestingly, the perturbed packing of the two helices was maintained during the simulation but no further dissociation of the two helices occurred. The talin interacted with the lipids via the interaction surface described in the *αβ-tal-CG* simulations, with the difference that after 20 ns of simulation the F0 domain also associated with the bilayer. Despite the direct interactions between the F1-F0 pair and the bilayer, the absence of the loop between the F0-F1 and the bilayer resulted in a linear

arrangement of the domains (Fig. 5.17B). This suggests that the optimum and thus more stable association of talin with the lipids may be needed for the scissoring movement of the helices. In all cases, the F3/β integrin tail interactions were as identified in the main text for the F2-F3/αβ complex (Fig. 5.9).

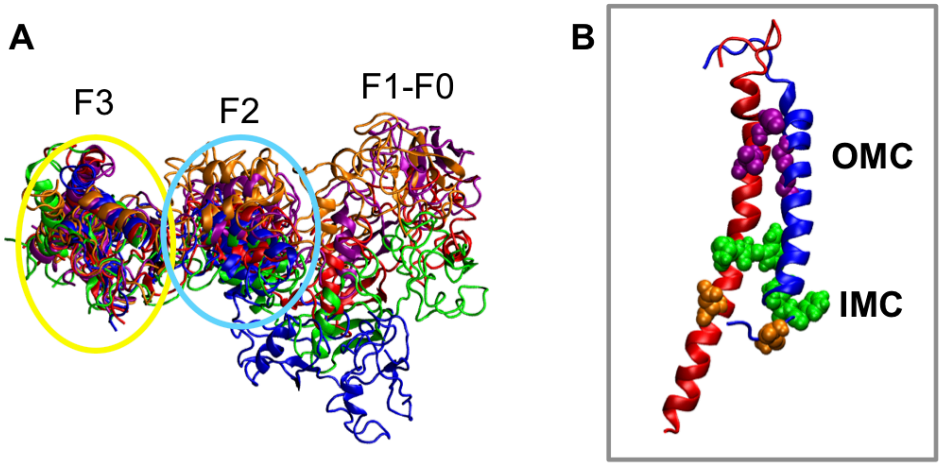


Figure C.2: A. Alignment of the F3 domain of the final snapshots from the *αβ-tal-CG* simulation. Each colour represents a different simulation. B. Snapshot of the α/β integrin TM region at the end of the *αβ-tal-CG* simulation showing the disruption of the interactions in the IMC and OMC regions.

APPENDIX D

D.1 Preparation of bilayers

All symmetric bilayers (see Table 6.1) were prepared using the approach described in the methods section of Chapter 3 (Section 3.2.3). Initially, only a small portion of the final lipid bilayer was prepared and, after the end of the simulation, this portion was replicated in the x and y axes to reach the final bilayer size. A further 200 ns of simulation was performed to remove any discontinuities in the newly formed larger bilayer. The outer bilayer contained 452 POPC lipids, 459 SM lipids and 474 CHOL molecules, the outer-2 bilayer contained 633 POPC lipids, 305 SM lipids and 332 CHOL and the inner bilayer contained 379 POPC lipids, 248 POPS lipids, 313 POPE lipids and 254 CHOL molecules. The asymmetric bilayer contained 240 POPC lipids, 261 SM lipids and 220 CHOL molecules in the outer leaflet and 193 POPC lipids, 123 POPS lipids, 139 POPE lipids and 121 CHOL molecules in the inner leaflet.

D.2 Convergence analysis

In order to study the convergence of the simulations, the two-dimensional average lipid headgroup densities were calculated for the POPC lipids in the *tal_αβ-1* simulation using 2.5 µs, 5 µs, 7.5 µs and 10 µs of the simulation. This analysis showed that the simulation converged after 7.5 µs.

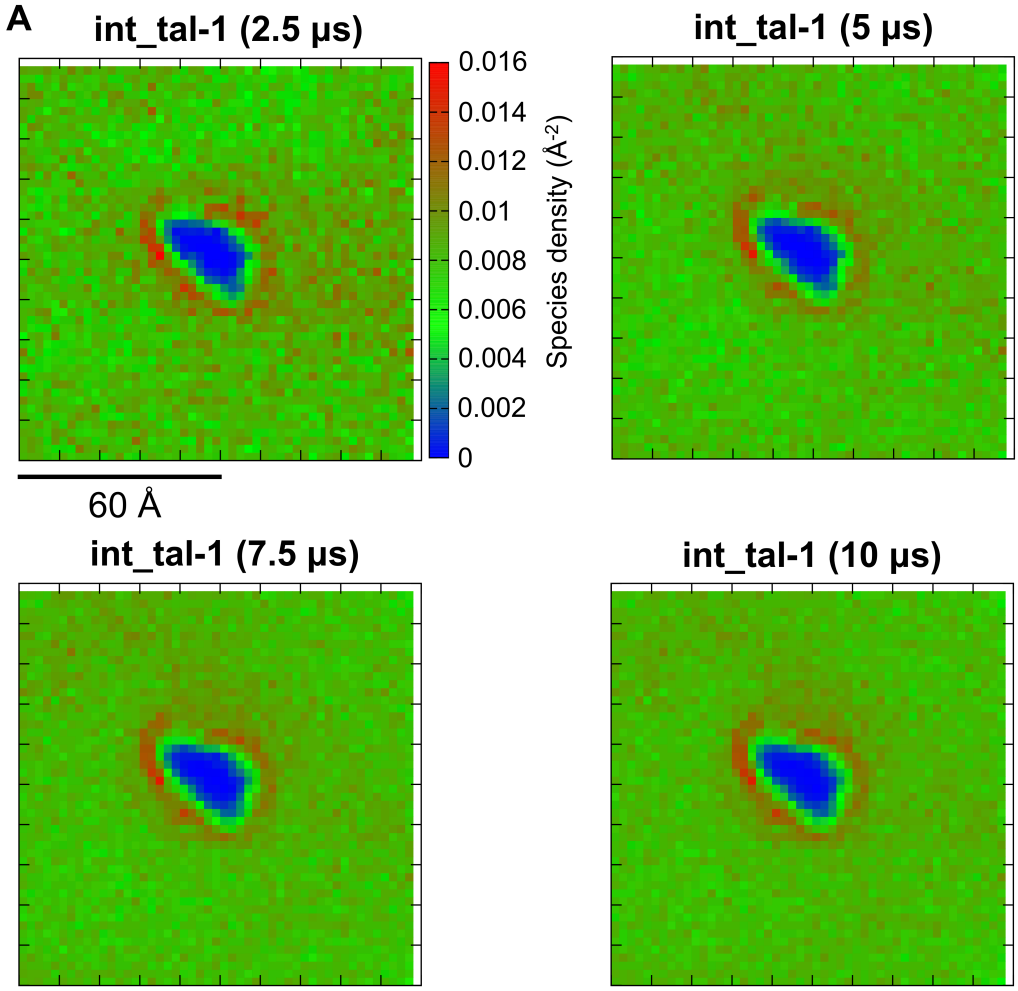


Figure D.2: A. Convergence analysis. Two-dimensional average lipid headgroup densities of POPC lipids in the *tal_αβ-1* simulation are shown at 2.5 µs, 5 µs, 7.5 µs and 10 µs of the simulation.

D.3 Other properties of the bilayer

Calculation of other bilayer properties such as the bilayer thickness revealed a ~2.5 Å perturbation of bilayer thickness in the vicinity of the protein in all systems (Fig. D.3.1). Monitoring ordering of the lipid tails revealed increased ordering, possibly due to the presence of CHOL (Fig. D.3.2).

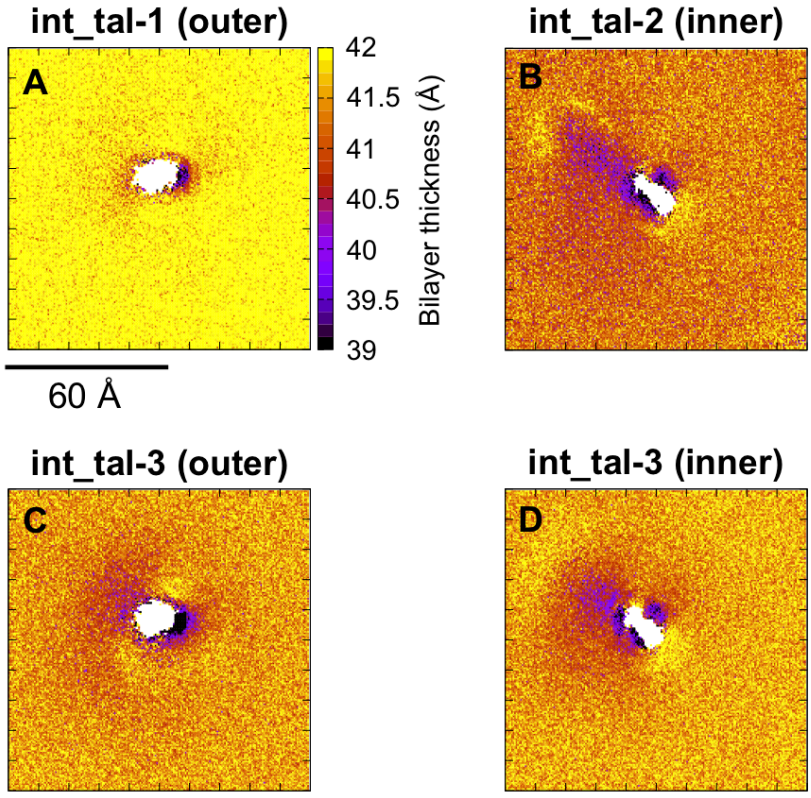


Figure D.3.1: Bilayer thickness for the outer leaflet of the *int_tal-1* simulation (A), the inner leaflet of the *int_tal-2* simulation (B) and the outer and the inner leaflets of the *int_tal-3* (C,D) simulation.

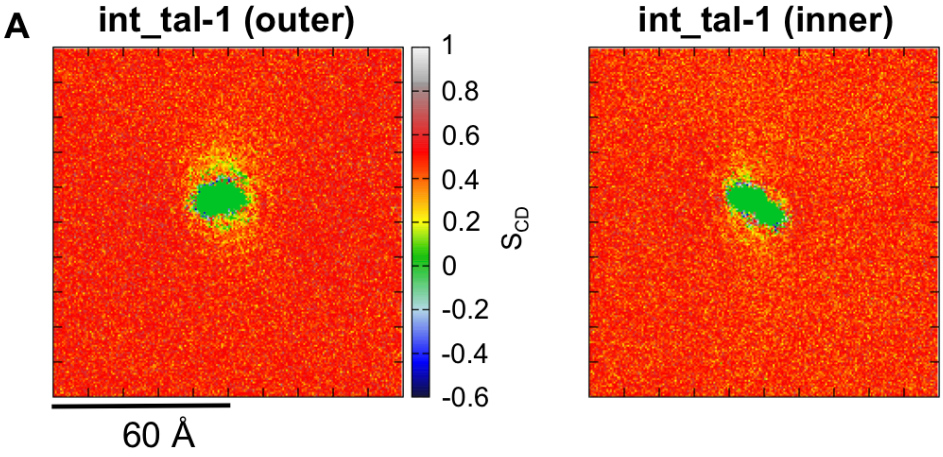


Figure D.3.2: Two-dimensional deuterium order parameters of the lipid tails around the protein for the outer (left) and inner (right) leaflets. Green (0) represents completely disordered lipid tails and grey (1) represents completely ordered tails.

Bibliography

1. Almen M, Nordstrom K, Fredriksson R, & Schioth H (2009) Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* 7(1):50.

2. Arinaminpathy Y, Khurana E, Engelman DM, & Gerstein MB (2009) Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov. Today* 14(23–24):1130-1135.

3. Grigoryan G, Moore DT, & DeGrado WF (2011) Transmembrane communication: general principles and lessons from the structure and function of the M2 proton channel, K+ channels, and integrin receptors. *Ann. Rev. Biochem.* 80(1):211-237.

4. Grecco HE, Schmick M, & Bastiaens PIH (2011) Signaling from the living plasma membrane. *Cell* 144(6):897-909.

5. Singer SJ & Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes. *Science* 175(4023):720-731.

6. Coskun U & Simons K (2011) Cell membranes: the lipid perspective. *Structure* 19(11):1543-1548.

7. Kusumi A*, et al.* (2005) Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu. Rev. Biophys. Biomol. Struct.* 34(1):351-378.

8. Lagerholm BC, Weinreb GE, Jacobson K, & Thompson NL (2004) Detecting microdomains in intact cell membranes. *Annu. Rev. Phys. Chem.* 56(1):309-336.

9. Shaw AS (2006) Lipid rafts: now you see them, now you don't. *Nat. Immunol.* 7(11):1139-1142.

10. Sean M (2003) Lipid rafts: elusive or illusive? *Cell* 115(4):377-388.

11. Anderson RGW & Jacobson K (2002) A role for lipid shells in targeting proteins to caveolae, rafts, and other lipid domains. *Science* 296(5574):1821-1825.

12. Gambin Y*, et al.* (2006) Lateral mobility of proteins in liquid membranes revisited. *Proc. Natl. Acad. Sci. USA* 103(7):2098-2102.

13. Falck E, Rog T, Karttunen M, & Vattulainen I (2007) Lateral diffusion in lipid membranes through collective flows. *J. Am. Chem. Soc.* 130(1):44-45.

14. Apajalahti T*, et al.* (2010) Concerted diffusion of lipids in raft-like membranes. *Faraday Discuss.* 144:411-430.

15. Simons K & Sampaio JL (2011) Membrane organization and lipid rafts. *Cold Spring Harbor Perspect. Biol.*

16. Rahdar M, *et al.* (2009) A phosphorylation-dependent intramolecular interaction regulates the membrane association and activity of the tumor suppressor PTEN. *Proc. Natl. Acad. Sci. USA* 106(2):480-485.

17. Di Paolo G & De Camilli P (2006) Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443(7112):651-657.

18. Holland KA & Holland IB (2001) Transmembrane signalling. *eLS*, (John Wiley & Sons, Ltd).

19. Choe H-W, Park JH, Kim YJ, & Ernst OP (2011) Transmembrane signaling by GPCRs: insight from rhodopsin and opsin structures. *Neuropharmacology* 60(1):52-57.

20. Guirland C, Buck KB, Gibney JA, DiCicco-Bloom E, & Zheng JQ (2003) Direct cAMP signaling through G-protein-coupled receptors mediates growth cone attraction induced by pituitary adenylate cyclase-activating polypeptide. *J. Neurosci.* 23(6):2274-2283.

21. Franco R (2009) G-protein-coupled receptor heteromers or how neurons can display differently flavoured patterns in response to the same neurotransmitter. *Brit. J. Pharm.* 158(1):23-31.

22. Lemmon MA & Schlessinger J (2010) Cell signaling by receptor tyrosine kinases. *Cell* 141(7):1117-1134.

23. Li E & Hristova K (2006) Role of receptor tyrosine kinase transmembrane domains in cell signaling and human pathologies. *Biochemistry* 45(20):6241-6251.

24. Shan Y, *et al.* (2012) Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* 149(4):860-870.

25. Simons K & Toomre D (2000) Lipid rafts and signal transduction. *Nat. Rev. Mol. Cell Biol.* 1(1):31-39.

26. Janosi L, Li Z, Hancock JF, & Gorfe AA (2012) Organization, dynamics, and segregation of Ras nanoclusters in membrane domains. *Proc. Natl. Acad. Sci. USA* 109(21):8097-8102.

27. Ye F, Kim C, & Ginsberg MH (2011) Molecular mechanism of inside-out integrin regulation. *J. Thromb. Haemost.* 9:20-25.

28. Kim C, Ye F, & Ginsberg MH (2011) Regulation of integrin activation. *Annu. Rev. Cell Dev. Biol.* 27(1):321-345.

29. Anthis NJ & Campbell ID (2011) The tail of integrin activation. *Trends Biochem. Sci.* 36(4):191-198.

30. Shattil SJ, Kim C, & Ginsberg MH (2010) The final steps of integrin activation: the end game. *Nat. Rev. Mol. Cell Biol.* 11(4):288-300.

31. Baker EL & Zaman MH (2010) The biomechanical integrin. *J. Biomech.* 43(1):38-44.

32. Wei W & Bing-Hao L (2009) Structural basis of integrin transmembrane activation. *J. Cellular Biochem.* 9999(9999):n/a.

33. Moser M, Legate KR, Zent R, & Fassler R (2009) The tail of integrins, talin, and kindlins. *Science* 324(5929):895-899.

34. Harburger DS & Calderwood DA (2009) Integrin signalling at a glance. *J. Cell Sci.* 122(2):159-163.

35. Bennett JS, Berger BW, & Billings PC (2009) The structure and function of platelet integrins. *J. Thromb. Haemost.* 7(s1):200-205.

36. Askari JA, Buckley PA, Mould AP, & Humphries MJ (2009) Linking integrin conformation to function. *J. Cell Sci.* 122(2):165-170.

37. Luo B-H, Carman CV, & Springer TA (2007) Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.* 25(1):619-647.

38. Lad Y, Harburger DS, Calderwood DA, & David AC (2007) Integrin cytoskeletal interactions. *Methods Enzymol.*, (Academic Press), Vol Volume 426, pp 69-84.

39. Arnaout MA, Goodman SL, & Xiong J-P (2007) Structure and mechanics of integrin-based cell adhesion. *Curr. Opin. Cell Biol.* 19(5):495-507.

40. Luo B-H & Springer TA (2006) Integrin structures and conformational signaling. *Curr. Opin. Cell Biol.* 18(5):579-586.

41. Ginsberg MH, Partridge A, & Shattil SJ (2005) Integrin regulation. *Curr. Opin. Cell Biol.* 17(5):509-516.

42. Campbell ID & Ginsberg MH (2004) The talin-tail interaction places integrin activation on FERM ground. *Trends Biochem. Sci.* 29(8):429-435.

43. Takagi J & Springer TA (2002) Integrin activation and structural rearrangement. *Immunol. Rev.* 186(1):141-163.

44. Brown NH (2000) Cell–cell adhesion via the ECM: integrin genetics in fly and worm. *Matrix Biol.* 19(3):191-201.

45. Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* 110(6):673-687.

46. Hynes RO (1992) Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 69(1):11-25.

47. Delon I & Brown NH (2007) Integrins and the actin cytoskeleton. *Curr. Opin. Cell Biol.* 19(1):43-50.

48. Takagi J, Strokovich K, Springer TA, & Walz T (2003) Structure of integrin α5β1 in complex with fibronectin. *EMBO J.* 22(18):4607-4615.

49. Arnaout MA, Goodman SL, & Xiong J-P (2002) Coming to grips with integrin binding to ligands. *Curr. Opin. Cell Biol.* 14(5):641-652.

50. Anthis NJ, Wegener KL, Critchley DR, & Campbell ID (2010) Structural diversity in integrin/talin interactions. *Structure* 18(12):1654-1666.

51. Calderwood DA*, et al.* (1999) The talin head domain binds to integrin β subunit cytoplasmic tails and regulates integrin activation. *J. Biol. Chem.* 274(40):28071-28074.

52. Wegener KL*, et al.* (2007) Structural basis of integrin activation by talin. *Cell* 128(1):171-182.

53. Tadokoro S*, et al.* (2003) Talin binding to integrin β tails: a final common step in integrin activation. *Science* 302(5642):103-106.

54. Petrich BG*, et al.* (2007) Talin is required for integrin-mediated platelet function in hemostasis and thrombosis. *J. Exp. Med.*:jem.20071800.

55. Ulmer TS, Calderwood DA, Ginsberg MH, & Campbell ID (2003) Domain-specific interactions of talin with the membrane-proximal region of the integrin β3 subunit. *Biochemistry* 42(27):8307-8312.

56. Bouaouina M, Lad Y, & Calderwood DA (2008) The N-terminal domains of talin cooperate with the phosphotyrosine binding-like domain to activate β1 and β3 integrins. *J. Biol. Chem.* 283(10):6118-6125.

57. Calderwood DA (2004) Talin controls integrin activation. *Biochem. Soc. Trans.* 32(Pt3):434-437.

58. Larjava H, Plow EF, & Wu C (2008) Kindlins: essential regulators of integrin signalling and cell-matrix adhesion. *EMBO Rep.* 9(12):1203-1208.

59. Yates LA, Fuzery AK, Bonet R, Campbell ID, & Gilbert RJC (2012) Biophysical analysis of kindlin-3 reveals an elongated conformation and maps integrin binding to the membrane-distal β-subunit NPxY motif. *J. Biol. Chem.*

60. Schwartz MA & Ginsberg MH (2002) Networks and crosstalk: integrin signalling spreads. *Nat. Cell Biol.* 4(4):E65-68.

61. Cluzel C*, et al.* (2005) The mechanisms and dynamics of αvβ3 integrin clustering in living cells. *J. Cell Biol.* 171(2):383-392.

62. Danen EHJ & Sonnenberg A (2003) Erratum: Integrins in regulation of tissue development and function. *J. Pathol.* 201(4):632-641.

63. Burridge K & Chrzanowska-Wodnicka M (1996) Focal adhesions, contractility, and signaling. *Ann. Rev. Cell Dev. Biol.* 12(1):463-519.

64. Schoenwaelder SM & Burridge K (1999) Bidirectional signaling between the cytoskeleton and integrins. *Curr. Opin. Cell Biol.* 11(2):274-286.

65. Roca-Cusachs P, Gauthier NC, del Rio A, & Sheetz MP (2009) Clustering of α5β1 integrins determines adhesion strength whereas αvβ3 and talin enable mechanotransduction. *Proc. Natl. Acad. Sci. USA* 106(38):16245-16250.

66. Carman CV & Springer TA (2003) Integrin avidity regulation: are changes in affinity and conformation underemphasized? *Curr. Opin. Cell Biol.* 15(5):547-556.

67. Wolfenson H, Henis YI, Geiger B, & Bershadsky AD (2009) The heel and toe of the cell's foot: a multifaceted approach for understanding the structure and dynamics of focal adhesions. *Cell Motil. Cytoskeleton* 66(11):1017-1029.

68. Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, & Geiger B (2007) Functional atlas of the integrin adhesome. *Nat. Cell Biol.* 9(8):858-867.

69. Furie B & Furie BC (2008) Mechanisms of thrombus formation. *N. Engl. J. Med.* 359(9):938-949.

70. Desgrosellier JS & Cheresh DA (2010) Integrins in cancer: biological implications and therapeutic opportunities. *Nat. Rev. Cancer* 10(1):9-22.

71. Assoian RK (1997) Anchorage-dependent cell cycle progression. *J. Cell Biol.* 136(1):1-4.

72. Sheppard D (2000) In vivo functions of integrins: lessons from null mutations in mice. *Matrix Biol.* 19(3):203-209.

73. De Arcangelis A & Georges-Labouesse E (2000) Integrin and ECM functions: roles in vertebrate development. *Trends Genet.* 16(9):389-395.

74. Krissansen GW & Danen EHJ (2001) Integrins: signalling and disease. *eLS*, (John Wiley & Sons, Ltd).

75. Gong H, *et al.* (2010) G protein subunit Gα13 binds to integrin αIIbβ3 and mediates integrin "outside-in" signaling. *Science* 327(5963):340-343.

76. Rose DM, Alon R, & Ginsberg MH (2007) Integrin modulation and signaling in leukocyte adhesion and migration. *Immunol. Rev.* 218(1):126-134.

77. Cantor JM, Ginsberg MH, & Rose DM (2008) Integrin-associated proteins as potential therapeutic targets. *Immunol. Rev.* 223(1):236-251.

78. Zhu J, *et al.* (2008) Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces. *Mol. Cell* 32(6):849-861.

79. Xiong J-P, *et al.* (2009) Crystal structure of the complete integrin αVβ3 ectodomain plus an α/β transmembrane fragment. *J. Cell Biol.* 186(4):589-600.

80. Adair BD, *et al.* (2005) Three-dimensional EM structure of the ectodomain of integrin αVβ3 in a complex with fibronectin. *J. Cell Biol.* 168(7):1109-1118.

81. Nogales A, *et al.* (2009) Three-dimensional model of human platelet integrin αIIbβ3 in solution obtained by small angle neutron scattering. *J. Biol. Chem.* 285(2):1023-1031.

82. Adair BD & Yeager M (2002) Three-dimensional model of the human platelet integrin αIIbβ3 based on electron cryomicroscopy and x-ray crystallography. *Proc. Natl. Acad. Sci. USA* 99(22):14059-14064.

83. Xiong J-P, *et al.* (2001) Crystal structure of the extracellular segment of integrin αVβ3. *Science* 294(5541):339-345.

84. Shi M, *et al.* (2005) The crystal structure of the plexin-semaphorin-integrin domain/hybrid domain/I-EGF1 segment from the human integrin β2 subunit at 1.8-Å resolution. *J. Biol. Chem.* 280(34):30586-30593.

85. Lee J-O, Bankston LA, & Robert C Liddington MAAa (1995) Two conformations of the integrin A-domain (I-domain): a pathway for activation? *Structure* 3(12):1333-1340.

86. Raborn J, Wang W, & Luo B-H (2011) Regulation of Integrin αIIbβ3 Ligand Binding and Signaling by the Metal Ion Binding Sites in the β I Domain. *Biochemistry* 50(12):2084-2091.

87. Qu A & Leahy DJ (1996) The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. *Structure* 4(8):931-942.

88. Nolte M, *et al.* (1999) Crystal structure of the α1β1 integrin I-domain: insights into integrin I-domain function. *FEBS Lett.* 452(3):379-385.

89. Valdramidou D, Humphries MJ, & Mould AP (2008) Distinct Roles of β1 Metal Ion-dependent Adhesion Site (MIDAS), Adjacent to MIDAS (ADMIDAS), and Ligand-associated Metal-binding Site (LIMBS) Cation-binding Sites in Ligand Recognition by Integrin α2β1. *J. Biol. Chem.* 283(47):32704-32714.

90. Emsley J, Knight CG, Farndale RW, Barnes MJ, & Liddington RC (2000) Structural basis of collagen recognition by integrin α2β1. *Cell* 101(1):47-56.

91. Shimaoka M, *et al.* (2003) Structures of the αL I Domain and Its Complex with ICAM-1 Reveal a Shape-Shifting Pathway for Integrin Regulation. *Cell* 112(1):99-111.

92. Xiong J-P, *et al.* (2002) Crystal structure of the extracellular segment of integrin αVβ3 in complex with an Arg-Gly-Asp ligand. *Science* 296(5565):151-155.

93. Xie C, *et al.* (2010) Structure of an integrin with an αI domain, complement receptor type 4. *EMBO J.* 29(3):666-679.

94. Nishida N, *et al.* (2006) Activation of leukocyte β2 integrins by conversion from bent to extended conformations. *Immunity* 25(4):583-594.

95. Mould AP & Humphries MJ (2004) Regulation of integrin function through conformational complexity: not simply a knee-jerk reaction? *Curr. Opin. Cell Biol.* 16(5):544-551.

96. Xiong J-P, Stehle T, Goodman SL, & Arnaout MA (2003) New insights into the structural basis of integrin activation. *Blood* 102(4):1155-1159.

97. Springer TA & Dustin ML (2012) Integrin inside-out signaling and the immunological synapse. *Curr. Opin. Cell Biol.* 24(1):107-115.

98. Coutinho A, Garcia C, Gonzalez-Rodriguez J, & Lillo MP (2007) Conformational changes in human integrin αIIbβ3 after platelet activation, monitored by FRET. *Biophys. Chem.* 130(1,Äì2):76-87.

99. Ye F, Liu J, Winkler H, & Taylor KA (2008) Integrin αIIbβ3 in a membrane environment remains the same height after Mn2+ activation when observed by cryoelectron tomography. *J. Mol. Biol.* 378(5):976-986.

100. Shi X*, et al.* (2007) The MIG-2/integrin interaction strengthens cell-matrix adhesion and modulates cell motility. *J. Biol. Chem.* 282(28):20455-20466.

101. Takagi J, Petre BM, Walz T, & Springer TA (2002) Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* 110(5):599-611.

102. Xiao T, Takagi J, Coller BS, Wang J-H, & Springer TA (2004) Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* 432(7013):59-67.

103. Springer TA, Zhu J, & Xiao T (2008) Structural basis for distinctive recognition of fibrinogen γC peptide by the platelet integrin αIIbβ3. *J. Cell Biol.* 182(4):791-800.

104. Hantgan RR, Paumi C, Rocco M, & Weisel JW (1999) Effects of ligand-mimetic peptides Arg-Gly-Asp-X (X = Phe, Trp, Ser) on αIIbβ3 integrin conformation and oligomerization. *Biochemistry* 38(44):14461-14474.

105. Erb E-M, Tangemann K, Bohrmann B, Müller B, & Engel J (1997) Integrin αIIbβ3 reconstituted into lipid bilayers is nonclustered in Its activated state but clusters after fibrinogen binding. *Biochemistry* 36(24):7395-7402.

106. Ye F*, et al.* (2010) Recreation of the terminal events in physiological integrin activation. *J. Cell Biol.* 188(1):157-173.

107. Zhu J*, et al.* (2007) Requirement of α and β subunit transmembrane helix separation for integrin outside-in signaling. *Blood* 110(7):2475-2483.

108. Tiwari S, Askari JA, Humphries MJ, & Bulleid NJ (2011) Divalent cations regulate the folding and activation status of integrins during their intracellular trafficking. *Journal of Cell Science* 124(10):1672-1680.

109. Dabiri BE, Lee H, & Parker KK (2012) A potential role for integrin signaling in mechanoelectrical feedback. *Progr. Bioph. Mol. Biol.* 110(2–3):196-203.

110. Lin CY, Hilgenberg LGW, Smith MA, Lynch G, & Gall CM (2008) Integrin regulation of cytoplasmic calcium in excitatory neurons depends upon glutamate receptors and release from intracellular stores. *Mol. Cell. Neurosci.* 37(4):770-780.

111. Gailit J & Ruoslahti E (1988) Regulation of the fibronectin receptor affinity by divalent cations. *J. Biol. Chem.* 263(26):12927-12932.

112. Kim M, Carman CV, & Springer TA (2003) Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science* 301(5640):1720-1725.

113. Luo B-H, Springer TA, & Takagi J (2004) A specific interface between integrin transmembrane helices and affinity for ligand. *PLoS Biol.* 2(6):e153.

114. Friedland JC, Lee MH, & Boettiger D (2009) Mechanically activated integrin switch controls α5β1 function. *Science* 323(5914):642-644.

115. Alon R & Ley K (2008) Cells on the run: shear-regulated integrin activation in leukocyte rolling and arrest on endothelial cells. *Curr. Opin. Cell Biol.* 20(5):525-532.

116. Puklin-Faucher E, Gao M, Schulten K, & Vogel V (2006) How the headpiece hinge angle is opened: new insights into the dynamics of integrin activation. *J. Cell Biol.* 175(2):349-360.

117. Schneider D & Engelman DM (2004) Involvement of transmembrane domain interactions in signal transduction by α/β integrins. *J. Biol. Chem.* 279(11):9840-9846.

118. Lau T-L, Kim C, Ginsberg MH, & Ulmer TS (2009) The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* 28(9):1351-1361.

119. Yang J*, et al.* (2009) Structure of an integrin αIIbβ3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc. Natl. Acad. Sci. USA* 106(42):17729-17734.

120. Zhu J*, et al.* (2009) The structure of a receptor with two associating transmembrane domains on the cell surface: integrin αIIbβ3. *Mol. Cell* 34(2):234-249.

121. Vinogradova O*, et al.* (2002) A structural mechanism of integrin αIIbβ3 "inside-out" activation as regulated by its cytoplasmic face. *Cell* 110(5):587-597.

122. Bhunia A, Tang X-Y, Mohanram H, Tan S-M, & Bhattacharjya S (2009) NMR solution conformations and interactions of integrin αLβ2 cytoplasmic tails. *J. Biol. Chem.* 284(6):3873-3884.

123. Waterhouse AM, Procter JB, Martin DMA, Clamp M, & Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189-1191.

124. de Pereda JM, Lillo MP, & Sonnenberg A (2009) Structural basis of the interaction between integrin α6β4 and plectin at the hemidesmosomes. *EMBO J.* 28(8):1180-1190.

125. Anthis NJ*, et al.* (2009) The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO J.* 28(22):3623-3632.

126. Li R*, et al.* (2002) Characterization of the monomeric form of the transmembrane and cytoplasmic domains of the integrin β3 subunit by NMR spectroscopy. *Biochemistry* 41(52):15618-15624.

127. Ulmer TS, Yaspan B, Ginsberg MH, & Campbell ID (2001) NMR analysis of structure and dynamics of the cytosolic tails of integrin αIIbβ3 in aqueous solution. *Biochemistry* 40(25):7498-7508.

128. Liu S, Calderwood DA, & Ginsberg MH (2000) Integrin cytoplasmic domain-binding proteins. *J. Cell Sci.* 113(20):3563-3571.

129. Legate KR & Fässler R (2009) Mechanisms that regulate adaptor binding to β-integrin cytoplasmic tails. *J. Cell Sci.* 122(2):187-198.

130. Calderwood DA (2004) Integrin activation. *J. Cell Sci.* 117(5):657-666.

131. Lau T-L, Dua V, & Ulmer TS (2008) Structure of the integrin αIIb transmembrane segment. *J. Biol. Chem.* 283(23):16162-16168.

132. Lau T-L, Partridge AW, Ginsberg MH, & Ulmer TS (2008) Structure of the integrin β3 transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry* 47(13):4008-4016.

133. Yuan W*, et al.* (2006) CIB1 is an endogenous inhibitor of agonist-induced integrin αIIbβ3 activation. *J. Cell Biol.* 172(2):169-175.

134. Kim C*, et al.* (2012) Basic amino-acid side chains regulate transmembrane integrin signalling. *Nature* 481(7380):209-213.

135. Kim C, Ye F, Hu X, & Ginsberg MH (2012) Talin activates integrins by altering the topology of the β transmembrane domain. *J. Cell Biol.* 197(5):605-611.

136. Yau W-M, Wimley WC, Gawrisch K, & White SH (1998) The preference of tryptophan for membrane interfaces. *Biochemistry* 37(42):14713-14718.

137. Armulik A, Nilsson I, von Heijne G, & Johansson S (1999) Determination of the border between the transmembrane and cytoplasmic domains of human integrin subunits. *J. Biol. Chem.* 274(52):37030-37034.

138. Stefansson A, Armulik A, Nilsson I, von Heijne G, & Johansson S (2004) Determination of N- and C-terminal borders of the transmembrane domain of integrin subunits. *J. Biol. Chem.* 279(20):21200-21205.

139. Vinogradova O*, et al.* (2004) Membrane-mediated structural transitions at the cytoplasmic face during integrin activation. *Proc. Natl. Acad. Sci. USA* 101(12):4094-4099.

140. Ulmer TS (2010) Structural basis of transmembrane domain interactions in integrin signaling. *Cell Adhes. Migrat.* 4(2):243-248.

141. Haas TA & Plow EF (1996) The Cytoplasmic Domain of αIIbβ3. A ternary complex of the integrin alpha and beta subunits in a dvalent cation. *J. Biol. Chem.* 271(11):6017-6026.

142. Vallar L*, et al.* (1999) Divalent cations differentially regulate integrin αIb cytoplasmic tail binding to β3 and to calcium- and integrin-binding protein. *J. Biol. Chem.* 274(24):17257-17266.

143. Weljie AM, Hwang PM, & Vogel HJ (2002) Solution structures of the cytoplasmic tail complex from platelet integrin αIIb- and β3-subunits. *Proc. Natl. Acad. Sci. USA* 99(9):5878-5883.

144. Vinogradova O, Haas T, Plow EF, & Qin J (2000) A structural basis for integrin activation by the cytoplasmic tail of the αIIb-subunit. *Proc. Natl. Acad. Sci. USA* 97(4):1450-1455.

145. Metcalf DG*, et al.* (2010) NMR analysis of the αIIbβ3 cytoplasmic interaction suggests a mechanism for integrin regulation. *Proc. Natl. Acad. Sci. USA* 107(52):22481-22486.

146. Vararattanavech A, Lin X, Torres J, & Tan S-M (2009) Disruption of the integrin αLβ2 transmembrane domain interface by β2 Thr-686 mutation activates αLβ2 and promotes micro-clustering of the αL subunits. *J. Biol. Chem.* 284(5):3239-3249.

147. Chng C-P & Tan S-M (2011) Leukocyte integrin αLβ2 transmembrane association dynamics revealed by coarse-grained molecular dynamics simulations. *Proteins* 79(7):2203-2213.

148. Vararattanavech A*, et al.* (2010) A transmembrane polar interaction is involved in the functional regulation of integrin αLβ2. *J. Mol. Biol.* 398(4):569-583.

149. Gottschalk K-E (2005) A coiled-coil structure of the αIIbβ3 integrin transmembrane and cytoplasmic domains in its resting state. *Structure* 13(5):703-712.

150. Gottschalk K-E & Kessler H (2004) Evidence for hetero-association of transmembrane helices of integrins. *FEBS Lett.* 557(1-3):253-258.

151. Gottschalk KE, Adams PD, Brunger AT, & Kessler H (2002) Transmembrane signal transduction of the αIIbβ3 integrin. *Protein Sci.* 11(7):1800-1812.

152. Russ WP & Engelman DM (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* 296(3):911-919.

153. Doura AK, Kobus FJ, Dubrovsky L, Hibbard E, & Fleming KG (2004) Sequence context modulates the stability of a GxxxG-mediated transmembrane helix-helix dimer. *J. Mol. Biol.* 341(4):991-998.

154. Anbazhagan V & Schneider D (2010) The membrane environment modulates self-association of the human GpA TM domain¸ÄîImplications for membrane protein folding and transmembrane signaling. *Biochim. Biophys. Acta Biomembranes* 1798(10):1899-1907.

155. Brosig B & Langosch D (1998) The dimerization motif of the glycophorin A transmembrane segment in membranes: Importance of glycine residues. *Protein Sci.* 7(4):1052-1056.

156. Doura AK & Fleming KG (2004) Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer. *J. Mol. Biol.* 343(5):1487-1497.

157. Russ WP (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci. USA* 96(3):863-868.

158. Partridge AW, Liu S, Kim S, Bowie JU, & Ginsberg MH (2005) Transmembrane domain helix packing stabilizes integrin αIIbβ3 in the low affinity state. *J. Biol. Chem.* 280(8):7294-7300.

159. Li W*, et al.* (2005) A push-pull mechanism for regulating integrin function. *Proc. Natl. Acad. Sci. USA* 102(5):1424-1429.

160. Lin X, Tan SM, Law SKA, & Torres J (2006) Unambiguous prediction of human integrin transmembrane heterodimer interactions using only homologous sequences. *Proteins* 65(2):274-279.

161. Psachoulia E, Marshall DP, & Sansom MSP (2009) Molecular dynamics simulations of the dimerization of transmembrane α-helices. *Acc. Chem. Res.* 43(3):388-396.

162. Metcalf DG, Kulp DW, Bennett JS, & DeGrado WF (2009) Multiple approaches converge on the structure of the integrin αIIb/β3 transmembrane heterodimer. *J. Mol. Biol.* 392(4):1087-1101.

163. O'Toole TE*, et al.* (1994) Integrin cytoplasmic domains mediate inside-out signal transduction. *J. Cell Biol.* 124(6):1047-1059.

164. Hughes PE*, et al.* (1996) Breaking the integrin hinge. *J. Biol. Chem.* 271(12):6571-6574.

165. Lu CF & Springer TA (1997) The alpha subunit cytoplasmic domain regulates the assembly and adhesiveness of integrin lymphocyte function-associated antigen-1. *J. Immunol.* 159(1):268-278.

166. Hughes PE, O'Toole TE, Ylanne J, Shattil SJ, & Ginsberg MH (1995) The conserved membrane-proximal region of an integrin cytoplasmic domain specifies ligand binding affinity. *J. Biol. Chem.* 270(21):12411-12417.

167. Li R, *et al.* (2001) Oligomerization of the integrin αIIbβ3: roles of the transmembrane and cytoplasmic domains. *Proc. Natl. Acad. Sci. USA* 98(22):12462-12467.

168. Lu C, Takagi J, & Springer TA (2001) Association of the membrane proximal regions of the α and β subunit cytoplasmic domains constrains an integrin in the inactive state. *J. Biol. Chem.* 276(18):14642-14648.

169. Ma Y-Q, *et al.* (2006) Regulation of integrin αIIbβ3 activation by distinct regions of its cytoplasmic tails. *Biochemistry* 45(21):6656-6662.

170. Imai Y, *et al.* (2008) Genetic perturbation of the putative cytoplasmic membrane-proximal salt bridge aberrantly activates α4 integrins. *Blood* 112(13):5007-5015.

171. Li Y-F, Tang R-H, Puan K-J, Law SKA, & Tan S-M (2007) The cytosolic protein talin induces an intermediate affinity integrin αLβ2. *J. Biol. Chem.* 282(33):24310-24319.

172. Millon-Fremillon A, *et al.* (2008) Cell adaptive response to extracellular matrix density is controlled by ICAP-1-dependent β1-integrin affinity. *J. Cell Biol.* 180(2):427-441.

173. Luo B-H, Carman CV, Takagi J, & Springer TA (2005) Disrupting integrin transmembrane domain heterodimerization increases ligand binding affinity, not valency or clustering. *Proc. Natl. Acad. Sci. USA* 102(10):3679-3684.

174. Hato T, *et al.* (2008) Cooperative role of the membrane-proximal and -distal residues of the integrin β3 cytoplasmic domain in regulation of talin-mediated αIIbβ3 activation. *J. Biol. Chem.* 283(9):5662-5668.

175. Li R, *et al.* (2003) Activation of integrin αIIbβ3 by modulation of transmembrane helix associations. *Science* 300(5620):795-798.

176. Critchley DR & Gingras AR (2008) Talin at a glance. *J. Cell Sci.* 121(9):1345-1347.

177. Oxley CL, *et al.* (2008) An integrin phosphorylation switch: the effect of β3 integrin tail phosphorylation on Dok1 and talin binding. *J. Biol. Chem.* 283(9):5420-5426.

178. Anthis NJ, *et al.* (2009) β Integrin tyrosine phosphorylation Is a conserved mechanism for regulating talin-induced integrin activation. *J. Biol. Chem.* 284(52):36700-36710.

179. Bledzka K, *et al.* (2010) Tyrosine phosphorylation of integrin β3 regulates kindlin-2 binding and integrin activation. *J. Biol. Chem.* 285(40):30370-30374.

180. Deshmukh L, Gorbatyuk V, & Vinogradova O (2010) Integrin β3 phosphorylation dictates its complex with the Shc phosphotyrosine-binding (PTB) domain. *J. Biol. Chem.* 285(45):34875-34884.

181. Tompa P, *et al.* (2009) Close encounters of the third kind: disordered domains and the interactions of proteins. *BioEssays* 31(3):328-335.

182. Frame MC, Patel H, Serrels B, Lietha D, & Eck MJ (2010) The FERM domain: organizing the structure and function of FAK. *Nat. Rev. Mol. Cell Biol.* 11(11):802-814.

183. Campbell ID (2010) The talin FERM domain is not so FERM. *Structure* 18(10):1222-1223.

184. Fehon RG, McClatchey AI, & Bretscher A (2010) Organizing the cell cortex: the role of ERM proteins. *Nat. Rev. Mol. Cell Biol.* 11(4):276-287.

185. Ceccarelli DFJ, Song HK, Poy F, Schaller MD, & Eck MJ (2006) Crystal structure of the FERM domain of focal adhesion kinase. *J. Biol. Chem.* 281(1):252-259.

186. Elliott PR, *et al.* (2010) The structure of the talin head reveals a novel extended conformation of the FERM domain. *Structure* 18(10):1289-1299.

187. Senetar MA & McCann RO (2005) Gene duplication and functional divergence during evolution of the cytoskeletal linker protein talin. *Gene* 362(0):141-152.

188. Monkley SJ, Pritchard CA, & Critchley DR (2001) Analysis of the mammalian talin2 gene TLN2. *Biochem. Biophys. Res. Commun.* 286(5):880-885.

189. Monkley SJ, *et al.* (2011) Endothelial cell talin1 is essential for embryonic angiogenesis. *Dev. Biol.* 349(2):494-502.

190. Calderwood DA, *et al.* (2003) Integrin β cytoplasmic domain interactions with phosphotyrosine-binding domains: A structural prototype for diversity in integrin signaling. *Proc. Natl. Acad. Sci. USA* 100(5):2272-2277.

191. Calderwood DA, *et al.* (2002) The phosphotyrosine binding-like domain of talin activates integrins. *J. Biol. Chem.* 277(24):21749-21758.

192. Zhang X, *et al.* (2008) Talin depletion reveals independence of initial cell spreading from integrin activation and traction. *Nat. Cell Biol.* 10(9):1062-1068.

193. Saltel F, *et al.* (2009) New PI(4,5)P2- and membrane proximal integrin-binding motifs in the talin head control β3-integrin clustering. *J. Cell Biol.* 187(5):715-731.

194. Goult BT, *et al.* (2010) Structure of a double ubiquitin-like domain in the talin head: a role in integrin activation. *EMBO J.* 29(6):1069-1080.

195. Moore DT, *et al.* (2011) Affinity of talin-1 for the β3-integrin cytosolic domain is modulated by its phospholipid bilayer environment. *Proc. Natl. Acad. Sci. USA*.

196. Goult B, *et al.* (2008) NMR assignment of the C-terminal actin-binding domain of talin. *Biomol. NMR Assign.* 2(1):17-19.

197. Gingras AR, *et al.* (2008) The structure of the C-terminal actin-binding domain of talin. *EMBO J.* 27(2):458-469.

198. Goult BT, *et al.* (2010) The domain structure of talin: residues 1815-1973 form a five-helix bundle containing a cryptic vinculin-binding site. *FEBS Lett.* 584(11):2237-2241.

199. Patel B, *et al.* (2006) The activity of the vinculin binding sites in talin is influenced by the stability of the helical bundles that make up the talin rod. *J. Biol. Chem.* 281(11):7458-7467.

200. Critchley DR (2009) Biochemical and structural properties of the integrin-associated cytoskeletal protein talin. *Annu. Rev. Biophys.* 38(1):235-254.

201. Papagrigoriou E, *et al.* (2004) Activation of a vinculin-binding site in the talin rod involves rearrangement of a five-helix bundle. *EMBO J.* 23(15):2942-2951.

202. Hemmings L, *et al.* (1996) Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *J. Cell Sci.* 109(11):2715-2726.

203. Goult BT, *et al.* (2009) The structure of an interdomain complex that regulates talin activity. *J. Biol. Chem.* 284(22):15097-15106.

204. Goksoy E, *et al.* (2008) Structural basis for the autoinhibition of talin in regulating integrin activation. *Mol. Cell* 31(1):124-133.

205. Yan B, Calderwood DA, Yaspan B, & Ginsberg MH (2001) Calpain cleavage promotes talin binding to the β3 integrin cytoplasmic domain. *J. Biol. Chem.* 276(30):28164-28170.

206. Franco SJ, *et al.* (2004) Calpain-mediated proteolysis of talin regulates adhesion dynamics. *Nat. Cell Biol.* 6(10):977-983.

207. Franco SJ & Huttenlocher A (2005) Regulating cell migration: calpains make the cut. *J. Cell Sci.* 118(17):3829-3838.

208. de Pereda JM, *et al.* (2005) Structural basis for phosphatidylinositol phosphate kinase type Iγ binding to talin at focal adhesions. *J. Biol. Chem.* 280(9):8381-8386.

209. Chrzanowska-Wodnicka M, Smyth SS, Schoenwaelder SM, Fischer TH, & White GC (2005) Rap1b is required for normal platelet function and hemostasis in mice. *J. Clin. Invest.* 115(3):680-687.

210. Han J, *et al.* (2006) Reconstructing and deconstructing agonist-induced activation of integrin αIIbβ3. *Curr. Biol.* 16(18):1796-1806.

211. Martel V, *et al.* (2001) Conformation, localization, and integrin binding of talin depend on its interaction with phosphoinositides. *J. Biol. Chem.* 276(24):21217-21227.

212. Song X, *et al.* (2012) A novel membrane-dependent on/off switch mechanism of talin FERM domain at sites of cell adhesion. *Cell Res*.

213. Tremuth L*, et al.* (2004) A fluorescence cell biology approach to map the second integrin-binding site of talin to a 130-amino acid sequence within the rod domain. *J. Biol. Chem.* 279(21):22258-22266.

214. Xing B, Jedsadayanmata A, & Lam SCT (2001) Localization of an integrin binding site to the C terminus of talin. *J. Biol. Chem.* 276(48):44373-44378.

215. Rodius S*, et al.* (2008) The talin rod IBS2 α-helix interacts with the β3 integrin cytoplasmic tail membrane-proximal helix by establishing charge complementary salt bridges. *J. Biol. Chem.* 283(35):24212-24223.

216. Gingras AR*, et al.* (2009) Structural determinants of integrin binding to the talin rod. *J. Biol. Chem.* 284(13):8866-8876.

217. Moes M*, et al.* (2007) The integrin binding site 2 (IBS2) in the talin rod domain is essential for linking integrin β subunits to the cytoskeleton. *J. Biol. Chem.* 282(23):17280-17288.

218. Nayal A, Webb DJ, & Horwitz AF (2004) Talin: an emerging focal point of adhesion dynamics. *Curr. Opin. Cell Biol.* 16(1):94-98.

219. Puklin-Faucher E & Sheetz MP (2009) The mechanical integrin cycle. *J. Cell Sci.* 122(2):179-186.

220. Jiang G, Giannone G, Critchley DR, Fukumoto E, & Sheetz MP (2003) Two-piconewton slip bond between fibronectin and the cytoskeleton depends on talin. *Nature* 424(6946):334-337.

221. Calderwood DA, Shattil SJ, & Ginsberg MH (2000) Integrins and actin filaments: reciprocal regulation of cell adhesion and signaling. *J. Biol. Chem.* 275(30):22607-22610.

222. GarcIa-Alvarez B*, et al.* (2003) Structural determinants of integrin recognition by talin. *Mol. Cell* 11(1):49-58.

223. Wegener KL*, et al.* (2008) Structural basis for the interaction between the cytoplasmic domain of the hyaluronate receptor layilin and the talin F3 subdomain. *J. Mol. Biol.* 382(1):112-126.

224. Kong X, Wang X, Misra S, & Qin J (2006) Structural basis for the phosphorylation-regulated focal adhesion targeting of type Iγ phosphatidylinositol phosphate kinase (PIPKIγ) by talin. *J. Mol. Biol.* 359(1):47-54.

225. Haling JR, Monkley SJ, Critchley DR, & Petrich BG (2011) Talin-dependent integrin activation is required for fibrin clot retraction by platelets. *Blood* 117(5):1719-1722.

226. Coughlin SR (2005) Protease-activated receptors in hemostasis, thrombosis and vascular biology. *J. Thromb. Haemost.* 3(8):1800-1814.

227. Kinbara K, Goldfinger LE, Hansen M, Chou F-L, & Ginsberg MH (2003) Ras GTPases: integrins' friends or foes? *Nat. Rev. Mol. Cell Biol.* 4(10):767-778.

228. Bertoni A, *et al.* (2002) Relationships between Rap1b, affinity modulation of integrin αIIbβ3, and the actin cytoskeleton. *J. Biol. Chem.* 277(28):25715-25721.

229. Lee H-S, Lim CJ, Puzon-McLaughlin W, Shattil SJ, & Ginsberg MH (2009) RIAM activates integrins by linking talin to Ras GTPase membrane-targeting sequences. *J. Biol. Chem.* 284(8):5119-5127.

230. Watanabe N, *et al.* (2008) Mechanisms and consequences of agonist-induced talin recruitment to platelet integrin αIIbβ3. *J. Cell Biol.* 181(7):1211-1222.

231. Zhang XA & Hemler ME (1999) Interaction of the integrin β1 cytoplasmic domain with ICAP-1 protein. *J. Biol. Chem.* 274(1):11-19.

232. Ma Y-Q, Qin J, Wu C, & Plow EF (2008) Kindlin-2 (Mig-2): a co-activator of β3 integrins. *J. Cell Biol.* 181(3):439-446.

233. Moser M, Nieswandt B, Ussar S, Pozgajova M, & Fassler R (2008) Kindlin-3 is essential for integrin activation and platelet aggregation. *Nat. Med.* 14(3):325-330.

234. Ye F & Petrich BG (2011) Kindlin: helper, co-activator, or booster of talin in integrin activation? *Curr. Opin. Hematol.* 18(5).

235. Goult BT, *et al.* (2009) The structure of the N-terminus of kindlin-1: a domain important for αIIbβ3 integrin activation. *J. Mol. Biol.* 394(5):944-956.

236. Harburger DS, Bouaouina M, & Calderwood DA (2009) Kindlin-1 and -2 directly bind the C-terminal region of β integrin cytoplasmic tails and exert integrin-specific activation effects. *J. Biol. Chem.* 284(17):11485-11497.

237. Bledzka K, *et al.* (2012) Spatial coordination of kindlin-2 with talin head domain in interaction with integrin β cytoplasmic tails. *J. Biol. Chem.*

238. Montanez E, *et al.* (2008) Kindlin-2 controls bidirectional signaling of integrins. *Genes Dev.* 22(10):1325-1330.

239. Moser M, *et al.* (2009) Kindlin-3 is required for β2 integrin-mediated leukocyte adhesion to endothelial cells. *Nat. Med.* 15(3):300-305.

240. Ussar S, *et al.* (2008) Loss of kindlin-1 causes skin atrophy and lethal neonatal intestinal epithelial dysfunction. *PLoS Genet.* 4(12):e1000289.

241. Bialkowska K, *et al.* (2010) The integrin co-activator kindlin-3 is expressed and functional in a non-hematopoietic cell, the endothelial cell. *J. Biol. Chem.* 285(24):18640-18649.

242. Luecke H, Schobert B, Richter H-T, Cartailler J-P, & Lanyi JK (1999) Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* 291(4):899-911.

243. Gonen T, *et al.* (2005) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438(7068):633-638.

244. Stansfeld PJ & Sansom MSP (2011) Molecular simulation approaches to membrane proteins. *Structure* 19(11):1562-1572.

245. Leone V, Marinelli F, Carloni P, & Parrinello M (2010) Targeting biomolecular flexibility with metadynamics. *Curr. Opin. Struct. Biol.* 20(2):148-154.

246. Perilla JR, Beckstein O, Denning EJ, & Woolf TB (2011) Computing ensembles of transitions from stable states: Dynamic importance sampling. *J. Comp. Chem.* 32(2):196-209.

247. Arcario MJ, Ohkubo YZ, & Tajkhorshid E (2011) Capturing Spontaneous Partitioning of Peripheral Proteins Using a Biphasic Membrane-Mimetic Model. *J. Phys. Chem. B* 115(21):7029-7037.

248. Sansom MSP, Scott KA, & Bond PJ (2008) Coarse-grained simulation: a high-throughput computational approach to membrane proteins. *Biochem. Society Trans.* 036(1):27-32.

249. Scott KA*, et al.* (2008) Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* 16(4):621-630.

250. Stansfeld PJ, Hopkinson R, Ashcroft FM, & Sansom MSP (2009) PIP2-binding site in Kir channels: definition by multiscale biomolecular simulations. *Biochemistry* 48(46):10926-10933.

251. Delemotte L, Tarek M, Klein ML, Amaral C, & Treptow W (2011) Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* 108(15):6109-6114.

252. Murail S, Wallner B, Trudell James R, Bertaccini E, & Lindahl E (2011) Microsecond simulations indicate that ethanol binds between subunits and could stabilize an open-state model of a glycine receptor. *Biophys. J.* 100(7):1642-1650.

253. Dror RO*, et al.* (2009) Identification of two distinct inactive conformations of the β2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. USA* 106(12):4689-4694.

254. Vanni S, Neri M, Tavernelli I, & Rothlisberger U (2009) Observation of "ionic lock" formation in molecular dynamics simulations of wild-type β1 and β2 adrenergic receptors. *Biochemistry* 48(22):4789-4797.

255. Vanni S, Neri M, Tavernelli I, & Rothlisberger U (2011) Predicting novel binding modes of agonists to β adrenergic receptors using all-atom molecular dynamics simulations. *PLoS Comput. Biol.* 7(1):e1001053.

256. Feng Z, Hou T, & Li Y (2012) Studies on the interactions between β2 adrenergic receptor and Gs protein by molecular dynamics simulations. *J. Chem. Inf. Model.* 52(4):1005-1014.

257. Goetz A, Lanig H, Gmeiner P, & Clark T (2011) Molecular dynamics simulations of the effect of the G-protein and diffusible ligands on the β2-adrenergic receptor. *J. Mol. Biol.* 414(4):611-623.

258. Wan S & Coveney PV (2011) Molecular dynamics simulation reveals structural and thermodynamic features of kinase activation by cancer mutations within the epidermal growth factor receptor. *J. Comput. Chem.* 32(13):2843-2852.

259. Papakyriakou A, Vourloumis D, Tzortzatou-Stathopoulou F, & Karpusas M (2009) Conformational dynamics of the EGFR kinase domain reveals structural features involved in activation. *Proteins* 76(2):375-386.

260. Kästner J, Loeffler HH, Roberts SK, Martin-Fernandez ML, & Winn MD (2009) Ectodomain orientation, conformational plasticity and oligomerization of ErbB1 receptors investigated by molecular dynamics. *J. Struct. Biol.* 167(2):117-128.

261. Esteban-Martín S & Salgado J (2007) Self-assembling of peptide/membrane complexes by atomistic molecular dynamics simulations. *Biophys. J.* 92(3):903-912.

262. Babakhani A, Gorfe AA, Gullingsrud J, Kim JE, & Andrew McCammon J (2007) Peptide insertion, positioning, and stabilization in a membrane: Insight from an all-atom molecular dynamics simulation. *Biopolymers* 85(5-6):490-497.

263. Vostrikov VV, Hall BA, Greathouse DV, Koeppe RE, & Sansom MSP (2010) Changes in transmembrane helix alignment by arginine residues revealed by solid-state NMR experiments and coarse-grained MD simulations. *J. Am. Chem. Soc.* 132(16):5803-5811.

264. Popot JL & Engelman DM (2000) Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem.* 69:881-922.

265. MacKenzie KR, Prestegard JH, & Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276(5309):131-133.

266. Prakash A, Janosi L, & Doxastakis M (2010) Self-association of models of transmembrane domains of ErbB receptors in a lipid bilayer. *Biophys. J.* 99(11):3657-3665.

267. Rouse SL, Carpenter T, Stansfeld PJ, & Sansom MSP (2009) Simulations of the BM2 proton channel transmembrane domain from influenza virus B. *Biochemistry* 48(42):9949-9951.

268. Carpenter T, Bond PJ, Khalid S, & Sansom MSP (2008) Self-assembly of a simple membrane protein: coarse-grained molecular dynamics simulations of the influenza M2 channel. *Biophys. J.* 95(8):3790-3801.

269. Janosi L, Prakash A, & Doxastakis M (2010) Lipid-modulated sequence-specific association of glycophorin A in membranes. *Biophys. J.* 99(1):284-292.

270. Sengupta D & Marrink SJ (2010) Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. *Phys. Chem. Chem. Phys.* 12(40):12987-12996.

271. Chetwynd A, Wee CL, Hall BA, & Sansom MSP (2010) The energetics of transmembrane helix insertion into a lipid bilayer. *Biophys. J.* 99(8):2534-2540.

272. Jaud S, Tobias DJ, Falke JJ, & White SH (2007) Self-induced docking site of a deeply embedded peripheral membrane protein. *Biophys. J.* 92(2):517-524.

273. Kalli AC*, et al.* (2010) The structure of the talin/integrin complex at a lipid bilayer: an NMR and MD simulation study. *Structure* 18(10):1280-1288.

274. Wee CL, Balali-Mood K, Gavaghan D, & Sansom MSP (2008) The interaction of phospholipase A2 with a phospholipid bilayer: coarse-grained molecular dynamics simulations. *Biophys. J.* 95(4):1649-1657.

275. Liepiņa I, Czaplewski C, Janmey P, & Liwo A (2003) Molecular dynamics study of a gelsolin-derived peptide binding to a lipid bilayer containing phosphatidylinositol 4,5-bisphosphate. *Pept. Sci.* 71(1):49-70.

276. Psachoulia E & Sansom MSP (2008) Interactions of the pleckstrin homology domain with phosphatidylinositol phosphate and membranes: characterization via molecular dynamics simulations. *Biochemistry* 47(14):4211-4220.

277. Psachoulia E & Sansom MSP (2009) PX- and FYVE-mediated interactions with membranes: simulation studies. *Biochemistry* 48(23):5090-5095.

278. Lumb CN*, et al.* (2011) Biophysical and computational studies of membrane penetration by the GRP1 pleckstrin homology domain. *Structure* 19(9):1338-1346.

279. Blood PD & Voth GA (2006) Direct observation of Bin/amphiphysin/Rvs (BAR) domain-induced membrane curvature by means of molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* 103(41):15068-15072.

280. Yin Y, Arkhipov A, & Schulten K (2009) Simulations of membrane tubulation by lattices of amphiphysin N-BAR domains. *Structure* 17(6):882-892.

281. Hall A, Róg T, & Vattulainen I (2011) Effect of galactosylceramide on the dynamics of cholesterol-rich lipid membranes. *J. Phys. Chem. B* 115(49):14424-14434.

282. Kaiser H-J*, et al.* (2011) Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. *Proc. Natl. Acad. Sci. USA* 108(40):16628-16633.

283. Lingwood D*, et al.* (2011) Cholesterol modulates glycolipid conformation and receptor activity. *Nat. Chem. Biol.* 7(5):260-262.

284. Niemela PS, Ollila S, Hyvonen MT, Karttunen M, & Vattulainen I (2007) Assessing the nature of lipid raft Mmembranes. *PLoS Comput. Biol.* 3(2):e34.

285. Risselada HJ & Marrink SJ (2008) The molecular face of lipid rafts in model membranes. *Proc. Natl. Acad. Sci. USA* 105(45):17367-17372.

286. Kucerka N*, et al.* (2010) Cholesterol in bilayers with PUFA chains: doping with DMPC or POPC results in sterol reorientation and membrane-domain formation. *Biochemistry* 49(35):7485-7493.

287. Domanski J, Marrink SJ, & Schafer LV (2011) Transmembrane helices can induce domain formation in crowded model membranes. *Biochim. Biophys. Acta Biomembranes* (0).

288. Schafer LV*, et al.* (2011) Lipid packing drives the segregation of transmembrane helices into disordered lipid domains in model membranes. *Proc. Natl. Acad. Sci. USA*.

289. de Jong Dh, Lopez C, & Marrink SJ (2012) Molecular view on protein sorting into liquid-ordered membrane domains mediated by gangliosides and lipid anchors. *Faraday Discuss.*

290. Knotts TA, Rathore N, Schwartz DC, & de Pablo JJ (2007) A coarse grain model for DNA. *J. Chem. Phys.* 126(8).

291. de Pablo JJ (2011) Coarse-grained simulations of macromolecules: from DNA to nanocomposites. *Annu. Rev. Phys. Chem.* 62(1):555-574.

292. Born M & Oppenheimer R (1927) Zur quantentheorie der molekeln. *Annalen der Physik* 389(20):457-484.

293. Carloni P, Rothlisberger U, & Parrinello M (2002) The role and perspective of ab initio molecular dynamics in the study of biological systems. *Acc. Chem. Res.* 35(6):455-464.

294. Friesner RA & Guallar V (2004) Ab inition quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem.* 56(1):389-427.

295. Senn HM & Thiel W (2007) QM/MM studies of enzymes. *Curr. Opin. Chem. Biol.* 11(2):182-187.

296. Alves CN*, et al.* (2007) A quantum mechanics/molecular mechanics study of the protein–ligand interaction for inhibitors of HIV-1 integrase. *Chem. Eur. J.* 13(27):7715-7724.

297. Morse PM (1929) Diatomic molecules according to the wave mechanics. II. vibrational levels. *Phys. Rev. A* 34(1):57-64.

298. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98(12):10089-10092.

299. Essmann U*, et al.* (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103(19):8577-8593.

300. Allen MP & Tildesley DJ (1987) Computer simulations of liquids. *Oxford University Press, Oxford*.

301. Frenkel D & Smit B (2002) Understanding molecular simulations, second edition. *Academic Press, San Diego*.

302. Apol E*, et al.* (2011) GROMACS user manual, vesrion 4.5.4

303. David Van Der S*, et al.* (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26(16):1701-1718.

304. Hess B, Kutzner C, van der Spoel D, & Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4(3):435-447.

305. Lindahl E, Hess B, & van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7(8):306-317.

306. Jorgensen WL, Maxwell DS, & Tirado-Rives J (1996) Development and testing of the OPLS all-Atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118(45):11225-11236.

307. Jorgensen WL & Tirado-Rives J (1988) The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110(6):1657-1666.

308. Brooks BR*, et al.* (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4(2):187-217.

309. MacKerell AD, Banavali N, & Foloppe N (2000) Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56(4):257-265.

310. Perez A*, et al.* (2007) Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.* 92(11):3817-3829.

311. Berger O, Edholm O, & Jähnig F (1997) Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* 72(5):2002-2013.

312. Hess B, Bekker H, Berendsen H, & Fraaije J (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18(12):1463-1472.

313. Ryckaert J-P, Ciccotti G, & Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23(3):327-341.

314. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81(8):3684-3690.

315. Nose S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52(2):255-268.

316. Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* 31(3):1695-1697.

317. Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52(12):7182-7190.

318. Lindahl E & Sansom MSP (2008) Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* 18(4):425-431.

319. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, & de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* 111(27):7812-7824.

320. Monticelli L*, et al.* (2008) The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* 4(5):819-834.

321. Bond PJ, Holyoake J, Ivetac A, Khalid S, & Sansom MSP (2007) Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J. Struct. Biol.* 157(3):593-605.

322. Bond PJ, Wee CL, & Sansom MSP (2008) Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochemistry* 47(43):11321-11331.

323. Periole X, Cavalli M, Marrink S-J, & Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* 5(9):2531-2543.

324. Bond PJ & Sansom MSP (2006) Insertion and assembly of membrane proteins via simulation. *J. Am. Chem. Soc.* 128(8):2697-2704.

325. Atilgan AR*, et al.* (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80(1):505-515.

326. Ayton GS, Noid WG, & Voth GA (2007) Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr. Opin. Cell Biol.* 17(2):192-198.

327. Ayton GS & Voth GA (2009) Systematic multiscale simulation of membrane protein systems. *Curr. Opin. Cell Biol.* 19(2):138-144.

328. Senn H & Thiel W (2007) QM/MM methods for biological systems atomistic approaches in modern biology. *Top. Curr. Chem.* 268:173-290.

329. Han W, Wan C-K, Jiang F, & Wu Y-D (2010) PACE force field for protein simulations. 1. full parameterization of version 1 and verification. *J. Chem. Theory Comput.* 6(11):3373-3389.

330. Han W, Wan C-K, & Wu Y-D (2010) PACE force field for protein simulations. 2. folding simulations of peptides. *J. Chem. Theory Comput.* 6(11):3390-3402.

331. Izvekov S & Voth GA (2005) A multiscale coarse-graining method for biomolecular systems. *J. Phys. Chem. B* 109(7):2469-2473.

332. Deriu MA*, et al.* (2012) Multiscale modeling of cellular actin filaments: from atomistic molecular to coarse-grained dynamics. *Proteins*:n/a-n/a.

333. Rzepiela AJ, Louhivuori M, Peter C, & Marrink SJ (2011) Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. *Phys. Chem. Chem. Phys.* 13(22):10437-10448.

334. Tai K, Stansfeld PJ, & Sansom MSP (2009) Ion-blocking sites of the Kir2.1 channel revealed by multiscale modeling. *Biochemistry* 48(36):8758-8763.

335. Wee CL, Gavaghan D, & Sansom MSP (2010) Interactions between a voltage sensor and a toxin via multiscale simulations. *Biophys. J.* 98(8):1558-1565.

336. Gorfe AA, Chang C-eA, Ivanov I, & McCammon JA (2008) Dynamics of the acetylcholinesterase tetramer. *Biophys. J.* 94(4):1144-1154.

337. Voltz K*, et al.* (2008) Coarse-grained force field for the nucleosome from self-consistent multiscaling. *J. Comput. Chem.* 29(9):1429-1439.

338. Rzepiela AJ*, et al.* (2010) Reconstruction of atomistic details from coarse-grained structures. *J. Comput. Chem.* 31(6):1333-1343.

339. Stansfeld PJ & Sansom MSP (2011) From coarse grained to atomistic: a serial multiscale approach to membrane protein simulations. *J. Chem. Theory Comput.* 7(4):1157-1166.

340. Rotkiewicz P & Skolnick J (2008) Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* 29(9):1460-1465.

341. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234(3):779-815.

342. Eswar N*, et al.* (2001) Comparative protein structure modeling using MODELLER. *Curr. Prot. Protein Sci.*, (John Wiley & Sons, Inc.).

343. van Meer G, Voelker DR, & Feigenson GW (2008) Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* 9(2):112-124.

344. Ayed A*, et al.* (2001) Latent and active p53 are identical in conformation. *Nat. Struct. Mol. Biol.* 8(9):756-760.

345. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6(3):277-293.

346. Freddolino PL, Harrison CB, Liu Y, & Schulten K (2010) Challenges in protein-folding simulations. *Nat. Phys.* 6(10):751-758.

347. Vorobyov I, Li L, & Allen TW (2008) Assessing atomistic and coarse-grained force fields for protein-lipid interactions: the formidable challenge of an ionizable side chain in a membrane. *J. Phys. Chem. B* 112(32):9588-9602.

348. Hall BA, Chetwynd AP, & Sansom MSP (2011) Exploring peptide-membrane interactions with coarse-grained MD simulations. *Biophys. J.* 100(8):1940-1948.

349. Cho W & Stahelin RV (2006) Membrane binding and subcellular targeting of C2 domains. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1761(8):838-849.

350. Rufener E, Frazier AA, Wieser CM, Hinderliter A, & Cafiso DS (2004) Membrane-bound orientation and position of the synaptotagmin C2B domain determined by site-directed spin labeling. *Biochemistry* 44(1):18-28.

351. Popot JL & Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 29(17):4031-4037.

352. Chin C-N, von Heijne G, & de Gier J-WL (2002) Membrane proteins: shaping up. *Trends Biochem. Sci.* 27(5):231-234.

353. Bowie JU (2005) Solving the membrane protein folding problem. *Nature* 438(7068):581-589.

354. White SH & Wimley WC (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* 28(1):319-365.

355. Curran AR & Engelman DM (2003) Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr. Opin. Struct. Biol.* 13(4):412-417.

356. Li E, Wimley WC, & Hristova K (2012) Transmembrane helix dimerization: beyond the search for sequence motifs. *Biochim. Biophys. Acta Biomembranes* 1818(2):183-193.

357. White SH, Ladokhin AS, Jayasinghe S, & Hristova K (2001) How membranes shape protein structure. *J. Biol. Chem.* 276(35):32395-32398.

358. Senes A, Engel DE, & DeGrado WF (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* 14(4):465-479.

359. Langosch D & Arkin IT (2009) Interaction and conformational dynamics of membrane-spanning protein helices. *Protein Sci.* 18(7):1343-1358.

360. Adamian L & Liang J (2001) Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J. Mol. Biol.* 311(4):891-907.

361. Zhou FX, Merianos HJ, Brunger AT, & Engelman DM (2001) Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. USA* 98(5):2250-2255.

362. Sal-Man N, Gerber D, & Shai Y (2005) The identification of a minimal dimerization motif QXXS that enables homo- and hetero-association of transmembrane helices in vivo. *J. Biol. Chem.* 280(29):27449-27457.

363. Sal-Man N, Gerber D, & Shai Y (2004) The composition rather than position of polar residues (QxxS) drives aspartate receptor transmembrane domain dimerization in vivo. *Biochemistry* 43(8):2309-2313.

364. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, & Engelman DM (1992) Sequence specificity in the dimerization of transmembrane α-helices. *Biochemistry* 31(51):12719-12725.

365. Smith SO*, et al.* (2001) Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers. *Biochemistry* 40(22):6553-6558.

366. Sulistijo ES & MacKenzie KR (2006) Sequence dependence of BNIP3 transmembrane domain dimerization Implicates side-chain hydrogen bonding and a tandem GxxxG motif in specific helix-helix interactions. *J. Mol. Biol.* 364(5):974-990.

367. Overton MC, Chinault SL, & Blumer KJ (2003) Oligomerization, biogenesis, and signaling is promoted by a glycophorin A-like dimerization motif in transmembrane domain 1 of a yeast G protein-coupled receptor. *J. Biol. Chem.* 278(49):49369-49377.

368. Senes A, Gerstein M, & Engelman DM (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol.* 296(3):921-936.

369. Mendrola JM, Berger MB, King MC, & Lemmon MA (2002) The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J. Biol. Chem.* 277(7):4704-4712.

370. Dews IC & MacKenzie KR (2007) Transmembrane domains of the syndecan family of growth factor coreceptors display a hierarchy of homotypic and heterotypic interactions. *Proc. Natl. Acad. Sci. USA* 104(52):20782-20787.

371. Javadpour MM, Eilers M, Groesbeek M, & Smith SO (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J.* 77(3):1609-1618.

372. Marks F, Klingmüller U, & Müller-Decker K (2009) Cellular signal processing: An introduction to the molecular mechanisms of signal transduction. *Garland Sci.*

373. Matthews EE, Zoonens M, & Engelman DM (2006) Dynamic helix interactions in transmembrane signaling. *Cell* 127(3):447-450.

374. Rath A, Melnyk RA, & Deber CM (2006) Evidence for assembly of small multidrug resistance proteins by a "Two-faced" transmembrane helix. *J. Biol. Chem.* 281(22):15546-15553.

375. Cordes FS, Bright JN, & Sansom MSP (2002) Proline-induced distortions of transmembrane helices. *J. Mol. Biol.* 323(5):951-960.

376. Senes A, Ubarretxena-Belandia I, & Engelman DM (2001) The Cα-H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA* 98(16):9056-9061.

377. Liu W, Crocker E, Siminovitch DJ, & Smith SO (2003) Role of side-chain conformational entropy in transmembrane helix dimerization of glycophorin A. *Biophys. J.* 84(2):1263-1271.

378. Xiao Zhou F, Cocco MJ, Russ WP, Brunger AT, & Engelman DM (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Mol. Biol.* 7(2):154-160.

379. Langosch D, Brosig B, Kolmar H, & Fritz H-J (1996) Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J. Mol. Biol.* 263(4):525-530.

380. Escher C, Cymer F, & Schneider D (2009) Two GxxxG-like motifs facilitate promiscuous interactions of the human ErbB transmembrane domains. *J. Mol. Biol.* 389(1):10-16.

381. Gerber D, Sal-Man N, & Shai Y (2004) Two motifs within a transmembrane domain, one for homodimerization and the other for heterodimerization. *J. Biol. Chem.* 279(20):21177-21182.

382. Sulistijo ES, Jaszewski TM, & MacKenzie KR (2003) Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent. *J. Biol. Chem.* 278(51):51950-51956.

383. Unterreitmeier S*, et al.* (2007) Phenylalanine promotes interaction of transmembrane domains via GxxxG motifs. *J. Mol. Biol.* 374(3):705-718.

384. Roy R, Peplowska K, Rohde J, Ungermann C, & Langosch D (2006) Role of the Vam3p transmembrane segment in homodimerization and SNARE complex formation. *Biochemistry* 45(24):7654-7660.

385. Berger BW*, et al.* (2009) Consensus motif for integrin transmembrane helix association. *Proc. Natl. Acad. Sci. USA*.

386. Volynsky PE*, et al.* (2010) Computer simulations and modeling-assisted ToxR screening in deciphering 3D structures of transmembrane alpha-helical dimers: ephrin receptor A1. *Phys. Biol.* 7:16014-16014.

387. Law RJ*, et al.* (2005) Membrane protein structure quality in molecular dynamics simulation. *J. Mol. Graph. Model.* 24(2):157-165.

388. Raschle T, Hiller S, Etzkorn M, & Wagner G (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr. Opin. Struct. Biol.* 20(4):471-479.

389. Cross TA, Sharma M, Yi M, & Zhou H-X (2011) Influence of solubilizing environments on membrane protein structures. *Trends Biochem. Sci.* 36(2):117-125.

390. Kaminski GA, Friesner RA, Tirado-Rives J, & Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with

accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* 105(28):6474-6487.

391. Psachoulia E, Fowler PW, Bond PJ, & Sansom MSP (2008) Helix-helix interactions in membrane proteins: coarse-grained simulations of glycophorin A helix dimerization. *Biochemistry* 47(40):10503-10512.

392. Finger C, Escher C, & Schneider D (2009) The single transmembrane domains of human receptor tyrosine kinases encode self-interactions. *Sci. Signal.* 2(89):ra56.

393. Bocharov EV*, et al.* (2010) Left-handed dimer of EphA2 transmembrane domain: helix packing diversity among receptor tyrosine kinases. *Biophys. J.* 98(5):881-889.

394. Li R*, et al.* (2004) Dimerization of the transmembrane domain of integrin αIIb subunit in cell membranes. *J. Biol. Chem.* 279(25):26666-26673.

395. Fleishman SJ & Ben-Tal N (2006) Progress in structure prediction of alpha-helical membrane proteins. *Curr. Opin. Struct. Biol.* 16(4):496-504.

396. Park H, Yoon J, & Seok C (2008) Strength of Cα-H...O=C hydrogen bonds in transmembrane proteins. *J. Phys. Chem. B* 112(3):1041-1048.

397. Williams MJ, Hughes PE, O'Toole TE, & Ginsberg MH (1994) The inner world of cell adhesion: integrin cytoplasmic domains. *Trends Cell Biol.* 4(4):109-112.

398. Daura X*, et al.* (1999) Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* 38(1-2):236-240.

399. Kim C, Lau T-L, Ulmer TS, & Ginsberg MH (2009) Interactions of platelet integrin αIIb and β3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood* 113(19):4747-4753.

400. O'Toole TE*, et al.* (1991) Modulation of the affinity of integrin αIIbβ3 (GPIIb-IIIa) by the cytoplasmic domain of αIIb. *Science* 254(5033):845-847.

401. O'Toole TE, Ylanne J, & Culley BM (1995) Regulation of integrin affinity states through an NPXY motif in the subunit cytoplasmic domain. *J. Biol. Chem.* 270(15):8553-8558.

402. Wenk MR (2010) Lipidomics: new tools and applications. *Cell* 143(6):888-895.

403. Shevchenko A & Simons K (2010) Lipidomics: coming to grips with lipid diversity. *Nat. Rev. Mol. Cell Biol.* 11(8):593-598.

404. Simons K & Ehehalt R (2002) Cholesterol, lipid rafts, and disease. *J. Clin. Invest.* 110(5):597-603.

405. Hannun YA & Obeid LM (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat. Rev. Mol. Cell Biol.* 9(2):139-150.

406. Bretscher MS & Munro S (1993) Cholesterol and the Golgi apparatus. *Science* 261(5126):1280-1281.

407. Lingwood D & Simons K (2010) Lipid rafts as a membrane-organizing principle. *Science* 327(5961):46-50.

408. Veatch SL & Keller SL (2003) Separation of liquid phases in giant vesicles of ternary mixtures of phospholipids and cholesterol. *Biophys. J.* 85(5):3074-3083.

409. Kahya N, Scherfeld D, Bacia K, Poolman B, & Schwille P (2003) Probing lipid mobility of raft-exhibiting model membranes by fluorescence correlation spectroscopy. *J. Biol. Chem.* 278(30):28109-28115.

410. Kahya N & Schwille P (2006) Fluorescence correlation studies of lipid domains in model membranes (Review). *Mol. Mem. Biol.* 23(1):29-39.

411. Erwin L (2005) How principles of domain formation in model membranes may explain ambiguities concerning lipid raft formation in cells. *Biochim. Biophys. Acta Mol. Cell Res.* 1746(3):203-220.

412. Hancock JF (2006) Lipid rafts: contentious only from simplistic standpoints. *Nat. Rev. Mol. Cell Biol.* 7(6):456-462.

413. Khalili-Araghi F*, et al.* (2009) Molecular dynamics simulations of membrane channels and transporters. *Curr. Opin. Struct. Biol.* 19(2):128-137.

414. Gaede HC & Gawrisch K (2003) Lateral Diffusion Rates of Lipid, Water, and a Hydrophobic Drug in a Multilamellar Liposome. *Biophys. J.* 85(3):1734-1740.

415. Barenholz Y & Thompson TE (1999) Sphingomyelin: biophysical aspects. *Chem. Phys. Lipids* 102(1,Äì2):29-34.

416. Lundbaek JA, Andersen OS, Werge T, & Nielsen C (2003) Cholesterol-induced protein sorting: an analysis of energetic feasibility. *Biophys. J.* 84(3):2080-2089.

417. Hofsaβ C, Lindahl E, & Edholm O (2003) Molecular dynamics simulations of phospholipid bilayers with cholesterol. *Biophys. J.* 84(4):2192-2206.

418. Falck E, Patra M, Karttunen M, Hyvonen MT, & Vattulainen I (2004) Lessons of slicing membranes: interplay of packing, free area, and lateral diffusion in phospholipid/cholesterol bilayers. *Biophys. J.* 87(2):1076-1091.

419. Samuli Ollila OH, Rog T, Karttunen M, & Vattulainen I (2007) Role of sterol type on lateral pressure profiles of lipid membranes affecting membrane protein functionality: comparison between cholesterol, desmosterol, 7-dehydrocholesterol and ketosterol. *J. Struct. Biol.* 159(2):311-323.

420. Devaux PF & Morris R (2004) Transmembrane asymmetry and lateral domains in biological membranes. *Traffic* 5(4):241-246.

421. Pomorski T & Menon A (2006) Lipid flippases and their biological functions. *Cell. Mol. Life Sci.* 63(24):2908-2921.

422. Daleke DL (2007) Phospholipid flippases. *J. Biol. Chem.* 282(2):821-825.

423. Buchan DWA, *et al.* (2010) Protein annotation and modelling servers at University College London. *Nucl. Acids Res.* 38(suppl 2):W563-W568.

424. Janik ME, Lityńska A, & Vereecken P (2010) Cell migration—The role of integrin glycosylation. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1800(6):545-555.

425. Tynan CJ, *et al.* (2011) Human epidermal growth factor receptor (EGFR) aligned on the plasma membrane adopts key features of drosophila EGFR asymmetry. *Mol. Cellular Biol.* 31(11):2241-2252.

426. Lee AG (2003) Lipid–protein interactions in biological membranes: a structural perspective. *Biochim. Biophys. Acta Biomembranes* 1612(1):1-40.

427. Marsh D (2010) Electron spin resonance in membrane research: protein–lipid interactions from challenging beginnings to state of the art. *Eur. Biophys. J.* 39(4):513-525.

428. Lee AG (2011) Biological membranes: the importance of molecular detail. *Trends Biochem. Sci.* 36(9):493-500.

429. Ramadurai S, *et al.* (2010) Influence of Hydrophobic Mismatch and Amino Acid Composition on the Lateral Diffusion of Transmembrane Peptides. *Biophys. J.* 99(5):1447-1454.

430. Powl AM, East JM, & Lee AG (2005) Heterogeneity in the binding of lipid molecules to the surface of a membrane protein: hot spots for anionic lipids on the mechanosensitive channel of large conductance MscL and effects on conformation. *Biochemistry* 44(15):5873-5883.

431. Marsh D (2008) Electron spin resonance in membrane research: protein–lipid interactions. *Methods* 46(2):83-96.

432. Hessel E, Müller P, Herrmann A, & Hofmann KP (2001) Light-induced reorganization of phospholipids in rod disc membranes. *J. Biol. Chem.* 276(4):2538-2543.

433. Joanne P, *et al.* (2009) Lipid reorganization induced by membrane-active peptides probed using differential scanning calorimetry. *Biochim. Biophys. Acta Biomembranes* 1788(9):1772-1781.

434. Windschiegl B, *et al.* (2009) Lipid reorganization induced by shiga toxin clustering on planar membranes. *PLoS ONE* 4(7):e6238.

435. Alves I, Staneva G, Tessier C, Salgado GF, & Nuss P (2011) The interaction of antipsychotic drugs with lipids and subsequent lipid reorganization investigated using biophysical methods. *Biochim. Biophys. Acta Biomembranes* 1808(8):2009-2018.

436. Varma S, Teng M, & Scott HL (2012) Nonintercalating nanosubstrates create asymmetry between bilayer leaflets. *Langmuir* 28(5):2842-2848.

437. Balasubramanian K & Schroit AJ (2003) Aminophospholipid assymmetry: a matter of life and death. *Annu. Rev. Physiol.* 65(1):701-734.

438. Negrete HO, Rivers RL, Gough AH, Colombini M, & Zeidel ML (1996) Individual leaflets of a membrane bilayer can independently regulate permeability. *J. Biol. Chem.* 271(20):11627-11630.

439. Manno S, Takakuwa Y, & Mohandas N (2002) Identification of a functional role for lipid asymmetry in biological membranes: Phosphatidylserine-skeletal protein interactions modulate membrane stability. *Proc. Natl. Acad. Sci. USA* 99(4):1943-1948.

440. Filippov A, Oradd G, & Lindblom G (2006) Sphingomyelin structure influences the lateral diffusion and raft formation in lipid bilayers. *Biophys. J.* 90(6):2086-2092.

441. Filippov A, Oradd G, & Lindblom G (2004) Lipid lateral diffusion in ordered and disordered phases in raft mixtures. *Biophys. J.* 86(2):891-896.

442. Filippov A, Oradd G, & Lindblom G (2007) Domain formation in model membranes studied by pulsed-field gradient-NMR: the role of lipid polyunsaturation. *Biophys. J.* 93(9):3182-3190.

443. East JM, Melville D, & Lee AG (1985) Exchange rates and numbers of annular lipids for the calcium and magnesium ion dependent adenosine triphosphatase. *Biochemistry* 24(11):2615-2623.

444. Esmann M & Marsh D (2006) Lipid–protein interactions with the Na,K-ATPase. *Chem. Phys. Lipids* 141(1–2):94-104.

445. Horvath LI, Knowles PF, Kovachev P, Findlay JB, & Marsh D (1997) A single-residue deletion alters the lipid selectivity of a K+ channel-associated peptide in the beta-conformation: spin label electron spin resonance studies. *Biophys. J.* 73(5):2588-2594.

446. Niemela PS*, et al.* (2010) Membrane proteins diffuse as dynamic complexes with lipids. *J. Am. Chem. Soc.* 132(22):7574-7575.

447. Hunte C & Richers S (2008) Lipids and membrane protein structures. *Curr. Opin. Struct. Biol.* 18(4):406-411.

448. Ryan TA, Myers J, Holowka D, Baird B, & Webb WW (1988) Molecular crowding on the cell surface. *Science* 239(4835):61-64.

449. Dupuy AD & Engelman DM (2008) Protein area occupancy at the center of the red blood cell membrane. *Proc. Natl. Acad. Sci. USA* 105(8):2848-2852.

450. Singh P, Tarafdar PK, Swamy MJ, & Chattopadhyay A (2012) Organization and dynamics of hippocampal membranes in a depth-dependent manner: an electron spin resonance study. *J. Phys. Chem. B* 116(9):2999-3006.

451. Pankov R, *et al.* (2005) Cholesterol distribution in plasma membranes of β1 integrin-expressing and β1 integrin-deficient fibroblasts. *Arch. Biochem. Biophys.* 442(2):160-168.

452. Hamouda AK, Chiara DC, Sauls D, Cohen JB, & Blanton MP (2005) Cholesterol interacts with transmembrane α-helices M1, M3, and M4 of the torpedo nicotinic acetylcholine receptor: photolabeling studies using [3H]azicholesterol. *Biochemistry* 45(3):976-986.

453. Shinoda T, Ogawa H, Cornelius F, & Toyoshima C (2009) Crystal structure of the sodium-potassium pump at 2.4 A resolution. *Nature* 459(7245):446-450.

454. Hanson MA, *et al.* (2008) A specific cholesterol binding site is established by the 2.8 Å structure of the human β2-Adrenergic receptor. *Structure* 16(6):897-905.

455. Sahu SK, Saxena R, & Chattopadhyay A (2012) Cholesterol depletion modulates detergent resistant fraction of human serotonin1A receptors. *Mol. Mem. Biol.*:1-9.

456. Wang X, Lupardus P, LaPorte SL, & Garcia KC (2009) Structural biology of shared cytokine receptors. *Annu. Rev. Immunol.* 27(1):29-60.

457. Balali-Mood K, Bond PJ, & Sansom MSP (2009) Interaction of monotopic membrane enzymes with a lipid bilayer: a coarse-grained MD simulation Study. *Biochemistry* 48(10):2135-2145.

Publications

The Structure of the Talin/Integrin Complex at a Lipid Bilayer: An NMR and MD Simulation Study

*Antreas C. Kalli, Kate L. Wegener, Ben T Goult[†], Nicholas J. Anthis, Iain D. Campbell, & Mark S.P. Sansom**

Department of Biochemistry,
University of Oxford,
South Parks Road,
Oxford,
OX1 3QU,
U.K.

[†]University of Leicester
Dept Biochemistry, Room 1/40
Henry Wellcome Building
Lancaster Road
Leicester, UK
LE1 9HN

*To whom communications should be addressed:
e-mail: mark.sansom@bioch.ox.ac.uk
phone: +44 1865 613306
fax: +44 1865 613238

This paper was published in Structure.

See http://www.cell.com/structure/retrieve/pii/S0969212610003047

HIGHLIGHTS

- Simulations and NMR define the interactions of the talin head domain with membranes
- Talin binds to anionic lipids in a bilayer via the F2 subdomain
- Mutations in the F2 lipid-binding site may perturb talin/integrin interactions.

SUMMARY

Integrins are cell surface receptors crucial for cell migration and adhesion. They are activated by interactions of the talin head domain with the membrane surface and the integrin β cytoplasmic tail. Here, we use Coarse-Grained MD simulations and NMR spectroscopy to elucidate the membrane-binding surfaces of the talin head (F2-F3) domain. In particular, we show that mutations in the four basic residues (K258E, K274E, R276E, and K280E) in the F2 binding surface reduce the affinity of the F2-F3 for the membrane and modify its orientation relative to the bilayer. Our results highlight the key role of anionic lipids in talin/membrane interactions. Simulation of the F2-F3 in complex with the *α/β* transmembrane (TM) dimer reveals novel information for its orientation relative to the membrane. Our studies suggest that the perturbed orientation of talin relative to the membrane in the F2 mutant would be expected to in turn perturb talin/integrin interactions.

INTRODUCTION

Integrins are heterodimeric cell surface receptors that link the exterior and interior of a cell. They are central components of focal adhesion complexes that form during cell migration, and are crucial for a variety of signal transduction events such as proliferation, survival/apoptosis, differentiation and gene expression (Hynes, 2002). They are comprised of an *α* and *β* subunit, each containing an 80-150 kDa extracellular domain, a transmembrane (TM) domain and a largely unstructured cytoplasmic domain (Wegener and Campbell, 2008). In mammals there are 18 α and 8 β subunits. These can heterodimerize to form 24 distinct integrins with specific functions. Integrins transmit signals in either direction across the cell membrane. Via the "outside-in" activation pathway the engagement of extracellular matrix ligands can cause changes to the cytoplasmic region resulting in downstream signalling events (Hynes, 2002; Schwartz and Ginsberg, 2002). In the "inside-out" activation pathway the binding of talin to the membrane proximal region of the β subunit is a key step of integrin activation (Bouaouina et al., 2008; Calderwood, 2004; Calderwood et al., 1999; Petrich et al., 2007; Tadokoro et al., 2003; Ulmer et al., 2003; Wegener et al., 2007b).

Talin consists of a rod domain (Critchley and Gingras, 2008) and a head domain. The head domain can be broken down into four subdomains: F0, F1, F2 and F3. Integrin activation is mediated via the talin head domain, with the F3 domain being sufficient for activation of the integrin αIIbβ3 (Calderwood et al., 2002). One current activation model suggests that when the talin head domain binds to the integrin β cytoplasmic tail, separation of the two TM helices of the integrin occurs, leading to conformational changes in the ectodomain and a consequent increase in the affinity of the integrin (i.e. activation) (Kim et al., 2003; Lu et al., 2001; Luo et al., 2007; Luo et al., 2004; Partridge et al., 2005). During activation, it is suggested that a positively charged surface of the talin head domain is directed towards the cytoplasmic surface of the membrane, and that interactions between the talin and negatively charged (i.e. acidic) headgroups of membrane lipids occur (Anthis et al., 2009; Wegener and Campbell, 2008). These interactions are believed to play a crucial role in the "inside-out activation" process.

Recent experimental studies have revealed structures for talin integrin tail complexes (Anthis et al., 2009; Garcia-Alvarez et al., 2003) and for the TM region of the integrins (Lau et al., 2009; Vinogradova et al., 2002; Yang et al., 2009). Furthermore, several studies have highlighted the importance of the interactions between the cytoplasmic region of the β tail of the integrin and talin head domain in integrin activation (Hughes et al., 1995; Ulmer et al., 2003; Vinogradova et al., 2002; Wegener et al., 2007b). However, the importance of the interactions between talin and the membrane in integrin activation is relatively poorly understood and a more detailed characterization is required.

Here, we investigate the interactions between a lipid bilayer and talin, and attempt to identify their role in integrin activation by combining experimental and computational techniques. In particular, nuclear magnetic resonance (NMR) and coarse-grained molecular dynamics (CG-MD) simulations are used to explore the position, orientation and interactions of wild type and mutated (at residues K258E, K274E, R276E, and K280E) talin2 F2-F3 and F2 domains (Fig. 1), with different neutral (i.e. zwitterionic) and anionic lipids. To the best of our knowledge, this is the first simulation of the complex of the talin2 F2-F3 domains and the integrin α and β TM helices. The results provide novel information concerning the orientation of talin relative to the bilayer and to the two helices. A good correlation between experimental and simulation results was obtained, and it is shown that mutations of positively charged residues in the F2 domain affect the orientation of talin with respect to the membrane providing new insights into F2-F3/membrane interactions and their role in the integrin activation mechanism.

RESULTS

NMR studies of talin/bilayer interactions

F2F3/bilayer interactions

It was recently suggested that a positively-charged patch on the surface of the talin F2 domain interacts with the membrane, and furthermore that mutations in residues K258, K274, R276, and K280 in talin2, or K256, K272, K274, and R277, in talin1 disrupt these interactions (Anthis et al., 2009). To test this hypothesis, the F2-F3-WT (or F2-WT) and the F2-F3-4E (or F2-4E) (i.e. K258E, K274E, R276E, and K280E) domains were added to liposomes and their chemical shift perturbations were monitored in ^{1}H-^{15}N-HSQC NMR spectra. Three different protein concentrations were used (75 μM, 100μM, 150 μM; see Fig. 2A), indicating the reproducibility of our results. For each concentration two spectra were obtained; a "test" spectrum with the protein alone and a spectrum of the protein in the presence of negatively charged liposomes. The induced shifts were very small at all three protein concentrations. Intensity changes were therefore used to detect binding rather than shifts. The two spectra were used to calculate a parameter R_{SN}. R_{SN} was calculated as:

$R_{SN} = (S_{PROTEIN_WITH_LIPOSOME}) / (S_{PROTEIN_WITHOUT_LIPOSOME})$

where S is the signal to noise ratio for a given spectrum. As a protein binds more tightly to a liposome the correlation time increases due to the vastly increased molecular weight of the complex, the spectral peaks broaden and R_{SN} becomes smaller (see Methods for further details). R_{SN} was thus used as an indication of interaction strength.

For F2-F3-WT in the presence of 20% DMPS/PC (1,2-dimyristoyl-*sn*-glycero-3-phospho-L-serine/L-α-phosphatidylcholine) liposomes (i.e. mixed anionic/neutral), no ^{1}H-^{15}N-HSQC spectrum was observed, i.e. $R_{SN} \approx 0$ implying that F2-F3-WT bound tightly to the liposomes. However, it was possible to get an ^{1}H-^{15}N-HSQC spectrum ($R_{SN} \approx 0.8$, Fig. 2B) from F2-F3-4E indicating that loss of the four basic residues from the surface of the F2 domain reduced binding to the liposomes.

Similar experiments were performed for F2-F3 with pure DMPS (1,2-dimyristoyl-*sn*-glycero-3-phospho-L-serine) liposomes. In this case $R_{SN} \approx 0$ for both the WT and the 4E mutant. This indicates that the strength of interaction of F2-F3 is increased as the fraction of anionic lipid is increased. *In vivo* the fraction of anionic lipid in the inner leaflet of mammalian plasma membranes is ~20 mol % (Yeung et al., 2006), similar to our 20% DMPS/PC condition.

F2/bilayer interactions

We also explored the interaction of the WT and F2-4E domains with liposomes. For the F2-WT domain in the presence of DMPS liposomes no spectrum was observed at any protein concentration (i.e. $R_{SN} \approx 0$), implying a strong F2-WT interaction with the liposomes. In contrast, in the ^{1}H-^{15}N-HSQC spectra of the F2-4E with and without DMPS liposomes, the shift induced in the peaks was very small (Supplementary information 1,2), suggesting that the F2-4E mutations in the F2 domain inhibited binding to the negatively charged lipids. Interestingly the R_{SN} value for F2-4E with DMPS liposomes (0.68) is lower than the R_{SN} for F2-4E (0.92) and for F2-F3-4E with 20% DMPS/PC (0.84). Given the qualitative nature of the results, this suggests that the interaction of the F2 domain with DMPS liposomes can be classified as "strong". In contrast, R_{SN} for the interaction of the F2 with 20% DMPS/PC liposomes is larger for both F2-WT and F2-4E than for the other systems. Therefore the interaction of F2 with 20% DMPS/PC can be classified as "weak". By similar criteria, the interaction of the F2-F3 with 20% DMPS/PC liposome can be classified as "intermediate" and with DMPS as "very strong". It is important to note that there is therefore some residual interaction of the protein with the bilayer for the F2-F3-4E mutant. This may reflect either 'non-productive' binding of F2 (see below) and/or interaction of F3 with the bilayer.

MD simulations

Having shown that the binding of the talin F2-F3 domain to bilayers depends on the interaction between anionic lipids and basic sidechains in F2, we used CG-MD simulations (summarized in Table 1) to develop a molecular model of these interactions, and to explore the possible role of interactions with the β-integrin TM domain.

F2-F3/bilayer and F2/bilayer interactions

In the first set of simulations we compared the binding of WT and F2-F3-4E with a preformed POPC/POPG(palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine/palmitoyl-oleoyl-phosphatidyl glycerol at a ration of 3:2; i.e. mixed zwitterionic/anionic) lipid bilayer. Five simulations were run with the F2-F3 domain initially displaced 100 Å away from a bilayer (Fig. 3A) for both the WT and the mutant. In all five simulations the WT domain initially diffused in the aqueous environment before binding to the bilayer at the lipid/water interface; it then remained bound for the remainder of the 1 μs simulation (Fig. 3AB). In contrast, for the 4E mutant only three simulations out of five resulted in binding of F2-F3 to the bilayer (Fig. 3C) consistent with a weakened talin/membrane interaction.

For the F2-F3-WT simulations, the pattern of interacting residues in the preferred binding surface (this appeared in 3 of 5 simulations) is consistent with the published F2-F3-β1D complex structure ((Anthis et al., 2009), and see below). Thus, lysine and arginine residues of the F2 subdomain of F2-F3 make large numbers of contacts with the lipids (Fig. 4A). In contrast, in the three F2-F3-4E mutant simulations which yielded a complex, a different pattern of interactions between the protein and the bilayer was observed (Fig. 4B). Therefore, the 4E mutations of F2 not only reduce the affinity of F2-F3 for the membrane but they also perturb the orientation of protein relative to the membrane when bound. This in turn might be anticipated to perturb talin/integrin interactions.

We also compared the interaction of the isolated F2 WT domain and 4E mutant with a preformed POPC/POPG bilayer. The F2 WT bound to the bilayer in all five simulations performed. The preferred binding surface (appearing in 3 of 5 simulations; Fig. 4C) was consistent with the F2-F3-WT simulations (above) and the structure of the F2-F3-β1D complex(Anthis et al., 2009). In contrast, simulations of the F2-4E mutant with a lipid bilayer did not yield the preferred binding surface in any of the four out of five simulations in which the F2-4E finally bound the bilayer (Fig. 4D). As for F2-F3, this suggests that the mutations inhibit the F2 domain from interacting productively with the bilayer.

We have also performed control simulations (for both F2-F3 and for F2) in which we made the quadruple K258E, K274E, R276E, and K280E mutation but set the charge on the four glutamate sidechains to zero (this may be viewed as an approximation to setting these four sidechains to a protonated state). For both F2 and for F2F3 this yielded only 'non-productive' complexes, i.e. the protein bound to the bilayer, but with a different orientation from the preferred orientation in the wildtype. Thus the K/R motif in F2 is needed for the 'productive' binding mode of the protein to be adopted.

Together, these four sets of simulations suggest a strong electrostatic interaction between the basic surface of the F2 subdomain and anionic lipid headgroups. The pattern of interacting lysine and arginine residues is conserved between the F2 and F2-F3 simulations and residues L262, D263, E266, Q275, G277 and R281 form the preferred surface made the largest number of contacts in both simulations.

The Talin F2F3/integrin β1D TM Complex in a Bilayer

We wished to extend our analysis of talin/bilayer interactions to include the possible reciprocal effects of the interaction of talin F2-F3 with the TM domain and intracellular juxtamembrane region of integrin β. This was possible because of the known X-ray structure

(albeit in the absence of a lipid bilayer) of F2-F3 bound to the β1D TM helix and tail (PDB 3G9W). We therefore took the F2-F3-β1D structure (converted to a CG structure) and positioned it in a preformed POPC/POPG bilayer with the β1D TM helix oriented such that the positively charged patch of F2-F3 interacted with the membrane surface; this resulted in the TM helix spanning the bilayer with a tilt angle of ~15°. This F2-F3-β1D/bilayer complex was then simulated for 0.5 μs (Fig. 3D). During the simulation, the β1D tilt angle increased by ~20° to a final value of ~35°. This increase in β1D tilt angle might be crucial in disrupting the interaction between the two integrin tails and lead to tail separation and integrin activation.

The F2-F3/lipid interactions in the pre-inserted F2-F3-β1D simulations were compared (Fig. 5). For this analysis the number of contacts (as defined by a distance cut-off of 8 Å (Hall and Sansom, 2009) between the protein residues and the lipids were counted over the last 50 ns of each simulation, and then normalised. Residues L262, D263, E266, Q275, G277 and R281 appeared to form a significant number of contacts with the lipids, in addition to the strong interactions between the lipids and the surface lysines and arginine residues in the talin F2 domain. Interestingly, the same residues (i.e. binding surface) formed a significant number of contacts with the bilayer in the simulations with F2-F3-WT and F2-WT (above). Significantly, if one analyses the lipids in contact with protein e.g. the F2-F3 simulations, the ratio of PC:PG contacts is 1:2, despite an overall bilayer composition of 3:2. Thus the protein binding region of the bilayer is enriched in anionic lipids about threefold relative to the bilayer as a whole.

Self-assembly simulations with the Talin F2-F3/Integrin αIIb-β1D TM Complex

We extended our study to a self-assembly simulation (Fig. 6) of lipids with a model of the complex formed by the F2F3-WT domain bound to the αIIb-β1D TM integrin helix heterodimer. This model was generated by adding the β1D TM domain and the αIIb TM and cytoplasmic domains to the F2-F3-β1D X-ray structure (PDB 3G9W) (Anthis et al., 2009), based on the NMR structure (PDB 2K9J) (Lau et al., 2009) of the α/β TM heterodimer. The CGMD self-assembly protocol allows us to explore the orientation of a complex membrane protein in a lipid bilayer without assuming an initial location for the protein (Scott et al., 2008). Note that the elastic network model used in this simulation does not permit dissociation of the α and β subunits. In the self assembly simulation we observed that the protein complex reached a final position and orientation in the lipid bilayer identical to that observed in simulations of the interaction of F2, F2-F3, and F2-F3-β1D discussed above. The

β1D tilt angle was ~27°.

To test the effect of the 4E mutations on the F2 binding surface, the self assembly simulations were repeated for F2-F3-4E-αIIb-β1D. Ten separate CG-MD self-assembly simulations were performed. However, in none of them did F2F3-4E reach a final binding surface similar to the "correct" binding surface observed in the WT simulations of F2, F2-F3, and of F2-F3-αIIb-β1D. Thus the self-assembly simulations on the complete complex support the picture obtained from NMR and simulation studies of the talin subdomains that electrostatic interactions between the surface of F2 and anionic lipids drive the talin/integrin interaction with a cell membrane.

DISCUSSION

A combination of experiment and simulation has revealed the nature of the interactions of the talin F2-F3 domain with a lipid bilayer membrane. We have defined residues within the F2 domain which play a key role in interactions with negatively charged lipid headgroups of the bilayer. Thus, four mutations (K258E, K274E, R276E, and K280E) in the binding surface of F2, which switch the net charge from +4 to -4, alter the nature of the interactions between talin and the bilayer. In the presence of these mutations the F2-F3 domain binds the membrane with a modified orientation which may be anticipated to perturb the interaction of the two TM helices of the integrin heterodimer.

Using NMR to characterize the interactions between talin and variously charged liposomes, it has been demonstrated that negatively charged headgroups in the bilayer promote the binding of talin to the membrane. Analysis of the simulations confirms this result since the number of contacts between either the F2-F3 domain or the F2 domain and the anionic headgroups of POPG is consistently larger than the number of contacts with the zwitterionic headgroups of POPC. In particular, the interactions of the F2-F3 domains were 3-fold higher with POPG/POPC than with POPC alone. In addition, the presence of F3 in combination with F2 seems to enhance the binding of talin to the membrane. Thus, in the presence of the F3 domain the protein is capable of binding to 20% anionic liposomes in the NMR studies, unlike the F2 subdomain alone. Therefore, anionic lipids in the bilayer and the presence of the F3 domain alongside the F2 domain are two key factors that enhance the interaction of talin with a membrane.

This is of likely mechanistic significance. It has been suggested (Anthis et al., 2009) that the formation of a salt bridge between the β integrin TM helix/tail with talin may be a crucial step of integrin activation. In particular, a salt bridge is thought to be formed between

K327 of F3 in talin2 (or K324 in talin1) and D759 of β1D (or D723 in β3). This interaction requires formation of a stable talin F2-F3 complex with the bilayer, correctly oriented to interact with the β-integrin TM domain/tail. Our simulations of the binding of the mutant talin F2-F3-4E reveal a different orientation relative to the bilayer which might thus be expected to inhibit the formation of this talin/β tail salt bridge, and thus to inhibit integrin activation. It is known that disruption of the αIIb R995-β3 D723 salt bridge of the integrin heterodimer is not sufficient alone for integrin activation (Tadokoro et al., 2003; Wegener et al., 2007b); rather talin binding to the β integrin tail is required to fully activate the integrin by replacing the αIIb R995-β3 D723 salt bridge with the talin/β1D bridge (Anthis et al., 2009). However, we should note that in our F2-F3-β1D coarse-grained simulations K327 of F3 and D759 of β1D were held in contact throughout the whole simulation by the elastic network model, and thus a multi-scale approach (see below) may be needed to test the predicted destabilization of the interaction by the perturbed bilayer/talin interaction.

The surface of F2-F3 which binds to the membrane consists of a stretch of basic and acidic residues. Residues K258, K274, R276, and K280 appear to make a substantial number of contacts with anionic lipids in the simulations explaining why mutations in these residues disrupt interaction with the membrane. This binding surface is consistent with NMR (Wegener et al., 2007a) and recent TIRF microscopy studies (Saltel et al., 2009). It is significant that the 'correct' orientation of F2-F3 at the bilayer was seen in our simulation *without* the β1D TM helix. This supports our suggestion that the F2/bilayer interaction positions F3 so as to facilitate interactions with β1D. A perturbed orientation of F2-F3 relative to the membrane (due to disruption of the talin/membrane interaction) would thus be expected to result in a less productive interaction of the F3 with β1D.

This study raises an interesting issue of relevance to ongoing efforts in integrative structural biology. It might be contended that visualisation alone could reveal the likely bilayer interaction surface of F2-F3. However, such an approach simply enables one to formulate a hypothesis which may be subsequently testing using simulation. Indeed, a more rigorous computational approach to evaluation of structural hypotheses is essential if one is to use the outcomes of integrative structural biology studies (Alber et al., 2008) as platforms for further investigations into higher levels of biological organization and function. Thus, we have used simulations to test an initial structural hypothesis alongside experiments designed to validate the simulations.

From a more methodological perspective, this study demonstrates that CG-MD simulations may be used to predict the position and the orientation of a protein domain bound

to a lipid bilayer. This builds upon previous CG-MD studies of integral membrane proteins (Scott et al., 2008) and of membrane bound enzymes (Balali-Mood et al., 2008; Wee et al., 2008). In particular, the self assembly CG-MD simulations for F2-F3-αIIb-β1D reported here extend this general approach to more complex membrane/protein assemblies containing both transmembrane elements and domains which interact with the bilayer surface.

Of course, there are limitations to the accuracy of CG-MD simulations (Allen, 2007). In order to improve on this we will extend the current studies to embrace multi-scale simulations (Ayton and Voth, 2009), by converting the CG models back to atomistic resolution, allowing their refinement by conventional MD simulations. As has been seen for e.g. the S4 helix of Kv channels (Wee et al., 2010) this allows for a more quantitative approach to protein/bilayer interactions.

ACKNOWLEDGMENTS

This research was funded by grants from the Wellcome Trust and the BBSRC. KLW was funded by the NIH Cell Migration Consortium. We also thank the David Critchley and Mark Ginsberg groups for discussions and some of the talin constructs used.

FIGURES

Figure 1

A Atomistic structure (shown in cartoon and VDW formats) and **B** coarse-grained (CG) model of the F2-F3 domain. In the atomistic model the F3 domain is shown in yellow and the F2 in cyan. In the VDW model (transparent) the colours for atoms are: cyan=carbon; red=oxygen; blue=nitrogen; yellow=sulphur. In the coarse-grained model the colours for CG particles are: cyan=apolar; red=polar; blue=positively charged; bronze=negatively charged; yellow=neutral. **C** Sequence of the talin2 F2F3 domain with the mutated residues indicated by arrows. The F2 domain is underlined in cyan and the F3 domain in yellow.

Figure 2

A Protein/liposome titrations. The ratio R_{SN} (see Methods for details) is shown for F2F3-4E with 20% DMPS/PC liposomes (purple), for F2-4E with DMPS liposomes (orange), for F2-WT with DMPS/PC liposomes (blue) and for F2-4E with DMPS/PC liposomes (red). Values close to 1 imply weak interaction with the liposomes, while values near zero imply strong interactions. **B** The value of R_{SN} is compared for the different proteins at a concentration of 100 μM. The same colour code as in the previous figure is used. The ratio R_{SN} is mean value ± standard deviation from the signal to noise ratio with and without the liposomes for every peak in each system. See also Figures S1 and S2.

Figure 3

A The F2-F3 simulation. Snapshots of the simulation systems are shown at 0 ns, 450 ns and 1 μs. The F2 domain is shown in cyan, and the F3 domain in yellow. The bilayer lipid phosphates are in grey. The lipid tails and waters are omitted for clarity. In **B** and **C** the progress of the F2-F3 and F2-F3-4E simulations respectively (five of each, indicated by the differently coloured lines) are shown as the distance of the centre of mass of the protein from the centre of mass of the bilayer as a function of time. **D** shows snapshots from the F2-F3-β1D simulations, with the β1D subunit in red and other colours as in **A**.

Figure 4

Contacts between the F2-F3 domain (**A,B**) and the F2 domain (**C,D**) and lipids. In **A** the contacts seen in the preferred orientation of the F2-F3 domain are shown as the backbone particles of the residues which made the highest number of contacts with the lipids, namely:

L262 (green), E266 (cyan), Q275 (orange) and G277 (purple). The yellow spheres correspond to the key basic residues (K258, K274, R276, and K280) in F2. In **B** the F2-F3-4E simulation is shown with the mutated residues (258, 274, 276, and 280) again shown in yellow. The phosphate groups are shown as grey spheres. Corresponding diagrams for the F2 and F2-4E simulations are shown in **C** and **D** respectively. For F2 the residues showing the highest number of contacts are: K256 (cyan), L262 (green), Q275 (orange) and G277 (purple). {In each case a distance cutoff of 8 Å between the outermost side-chain particle and any lipid particles was used to define a contact.

Figure 5

A Normalized average number of contacts as a function of residue number for the F2-F3 domain and lipids in the F2F3β1D simulation. Red lines correspond to the contacts between the lipid headgroups and the F2-F3 domain, and blue lines to the contacts between the lipid tails and the protein. **B** Snapshot, corresponding to the last frame of the simulation, of the F2-F3-β1D simulation system. The normalized average number of lipid/protein contacts was mapped onto the structure and is shown on a colour scale from blue (low number of contacts) through white to red (high number of contacts).

Figure 6

Bilayer self assembly simulation for F2-F3-αIIb-β1D. **A** shows the progress of the simulation. In the initial setup randomly placed lipids (POPC and POPG at a ratio of 3:2, shown in bonds formats with the hydrophobic tails in cyan, glycerol in green, phosphate in brown and choline in blue) surround the protein complex (F2 in cyan, F3 in yellow, αIIb in blue, β1D in red). Waters and counter ions are omitted for clarity. **B** shows the final frame (200 ns) of a F2-F3-αIIb-β1D- SA simulation.

REFERENCES

Alber, F., Forster, F., Korkin, D., Topf, M., and Sali, A. (2008). Integrating diverse data for structure determination of macromolecular assemblies. Ann. Rev. Biochem. *77*, 443-477.

Allen, T.W. (2007). Modeling charged protein side chains in lipid membranes. J. Gen. Physiol. *130*, 237-240.

Anthis, N.J., Wegener, K.L., Ye, F., Kim, C., Goult, B.T., Lowe, E.D., Vakonakis, I., Bate, N., Critchley, D.R., Ginsberg, M.H., and Campbell, I.D. (2009). The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. EMBO J. *28*, 3623-3632.

Ayton, G.S., and Voth, G.A. (2009). Systematic multiscale simulation of membranes protein systems. Curr. Opin. Struct. Biol. *19*, 138-144.

Balali-Mood, K., Bond, P.J., and Sansom, M.S.P. (2008). The interaction of monotopic membrane enzymes with a lipid bilayer: a coarse-grained MD simulation study. Biochem. *48*, 2135-2145.

Bouaouina, M., Lad, Y., and Calderwood, D.A. (2008). The N-terminal domains of talin cooperate with the phosphotyrosine binding-like domain to activate Œ≤1 and Œ≤3 Integrins. J. Biol. Chem. *283*, 6118-6125.

Calderwood, D.A. (2004). Talin controls integrin activation. Biochem. Soc. Trans. *32*, 434-437.

Calderwood, D.A., Yan, B., de Pereda, J.M., Alvarez, B.a.G.Ã., Fujioka, Y., Liddington, R.C., and Ginsberg, M.H. (2002). The phosphotyrosine binding-like domain of talin activates integrins. J. Biol. Chem. *277*, 21749-21758.

Calderwood, D.A., Zent, R., Grant, R., Rees, D.J.G., Hynes, R.O., and Ginsberg, M.H. (1999). The talin head domain binds to integrin Œ≤ subunit cytoplasmic tails and regulates integrin activation. J. Biol. Chem. *274*, 28071-28074.

Critchley, D.R., and Gingras, A.R. (2008). Talin at a glance. J. Cell Sci. *121*, 1345-1347.

Garcia-Alvarez, B., de Pereda, J.M., Calderwood, D.A., Ulmer, T.S., Critchley, D., Campbell, I.D., Ginsberg, M.H., and Liddington, R.C. (2003). Structural determinants of integrin recognition by talin. Mol. Cell *11*, 49-58.

Hall, B.A., and Sansom, M.S.P. (2009). Coarse-grained MD simulations and protein-protein interactions: the cohesin-dockerin system. J. Chem. Theor. Comput. *5*, 2465–2471.

Hughes, P.E., O'Toole, T.E., Yl√§nne, J., Shattil, S.J., and Ginsberg, M.H. (1995). The conserved membrane-proximal region of an integrin cytoplasmic domain specifies ligand binding affinity. J. Biol. Chem. *270*, 12411-12417.

Hynes, R.O. (2002). Integrins: bidirectional, allosteric signaling machines. Cell *110*, 673-687.

Kim, M., Carman, C.V., and Springer, T.A. (2003). Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. Science *301*, 1720-1725.

Lau, T.L., Kim, C., Ginsberg, M.H., and Ulmer, T.S. (2009). The structure of the integrin alphaIIbbeta3 transmembrane complex explains integrin transmembrane signalling. EMBO J. *28*, 1351-1361.

Lu, C., Takagi, J., and Springer, T.A. (2001). Association of the membrane proximal regions of the Œ± and Œ≤ subunit cytoplasmic domains constrains an integrin in the inactive state. J. Biol. Chem. *276*, 14642-14648.

Luo, B.-H., Carman, C.V., and Springer, T.A. (2007). Structural basis of integrin regulation and signaling. Ann. Rev. Immunol. *25*, 619-647.

Luo, B.-H., Springer, T.A., and Takagi, J. (2004). A specific interface between integrin transmembrane helices and affinity for ligand. PLoS Biol. *2*, e153.

Partridge, A.W., Liu, S., Kim, S., Bowie, J.U., and Ginsberg, M.H. (2005). Transmembrane domain helix packing stabilizes integrin alphaIIbbeta3 in the low affinity state. J. Biol. Chem. *280*, 7294-7300.

Petrich, B.G., Marchese, P., Ruggeri, Z.M., Spiess, S., Weichert, R.A.M., Ye, F., Tiedt, R., Skoda, R.C., Monkley, S.J., Critchley, D.R., and Ginsberg, M.H. (2007). Talin is required for integrin-mediated platelet function in hemostasis and thrombosis. J. Exp. Med. *204*, 3103-3111.

Saltel, F., Mortier, E., Hytonen, V.P., Jacquier, M.-C., Zimmermann, P., Vogel, V., Liu, W., and Wehrle-Haller, B. (2009). New PI(4,5)P2- and membrane proximal integrin-binding motifs in the talin head control β3-integrin clustering. J. Cell Biol. *187*, 715-731.

Schwartz, M.A., and Ginsberg, M.H. (2002). Networks and crosstalk: integrin signalling spreads. Nat. Cell Biol. *4*, E65-E68.

Scott, K.A., Bond, P.J., Ivetac, A., Chetwynd, A.P., Khalid, S., and Sansom, M.S.P. (2008). Coarse-grained MD simulations of membrane protein-bilayer self-assembly Structure *16*, 621-630.

Tadokoro, S., Shattil, S.J., Eto, K., Tai, V., Liddington, R.C., de Pereda, J.M., Ginsberg, M.H., and Calderwood, D.A. (2003). Talin binding to integrin β tails: a final common step in integrin activation. Science *302*, 103-106.

Ulmer, T.S., Calderwood, D.A., Ginsberg, M.H., and Campbell, I.D. (2003). Domain-specific interactions of talin with the membrane-proximal region of the integrin Œ≤3 subunit,Ä†. Biochem. *42*, 8307-8312.

Vinogradova, O., Velyvis, A., Velyviene, A., Hu, B., Haas, T.A., Plow, E.F., and Qin, J. (2002). A structural mechanism of integrin αIIbβ3 "inside-out" activation as regulated by Its cytoplasmic face. Cell *110*, 587-597.

Wee, C.L., Balali-Mood, K., Gavaghan, D., and Sansom, M.S.P. (2008). The interaction of phospholipase A2 with a phospholipid bilayer: coarse-grained molecular dynamics simulations. Biophys. J. *95*, 1649-1657.

Wee, C.L., Gavaghan, D., and Sansom, M.S.P. (2010). Interactions between a Kv channel voltage sensor and toxin via simulations. Biophys. J. *(in press)*.

Wegener, K.L., and Campbell, I.D. (2008). Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (Review). Mol. Memb. Biol. *25*, 376-387.

Wegener, K.L., Partridge, A.W., Han, J., Pickford, A.R., Liddington, R.C., Ginsberg, M.H., and Campbell, I.D. (2007a). Structural basis of integrin activation by talin. Cell *128*, 171-182.

Wegener, K.L., Partridge, A.W., Han, J., Pickford, A.R., Liddington, R.C., Ginsberg, M.H., and Campbell, I.D. (2007b). Structural basis of integrin activation by talin. Cell *128*, 171-182.

Yang, J., Ma, Y.Q., Page, R.C., Misra, S., Plow, E.F., and Qin, J. (2009). Structure of an integrin αiibβ3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. Proc. Natl. Acad. Sci. USA *106*, 17729-17734.

Yeung, T., Terebiznik, M., Yu, L., Silvius, J., Abidi, W.M., Philips, M., Levine, T., Kapus, A., and Grinstein, S. (2006). Receptor activation alters inner surface potential during phagocytosis. Science *313*, 347-351.

SUPPLEMENTARY INFORMATION

Figure S1, related to Figure 2:

Overlaid ^{1}H-^{15}N-HSQC NMR spectra of the F2-4E domain in the absence (blue) and F2-4E in the presence (red) of DMPS liposomes at a concentration of 100 μM. The figure shows that the shifts of the peaks induced by the addition of DMPS liposomes were very small.

Figure S2, related to Figure 2:

^{1}H-^{15}N-HSQC NMR spectra of the F2-4E **A** in the absence of and **B** in the presence of DMPS liposomes (100 μM) showing the reduction of the intensities upon the addition of liposomes.

TABLES

Table S1: Summary of CG simulations

Simulation	System	No. of particles	Duration (ns)	Final RMSD† (Å)
F2	F2-WT + 10289 W + POPC/POPG bilayer*	13955	5x500	0.20±0.02
F2-4E	F2-4E + 10243 W + POPC/POPG bilayer*	13921	5x500	0.18±0.03
F2-F3	F2-F3-WT + 13145 W + POPC/POPG bilayer*	17020	5x1000	0.28±0.05
F2-F3-4E	F2-F3-4E + 13095 W + POPC/POPG bilayer*	16982	5x1000	0.29±0.07
F2-F3-β1D	F2-F3-WT-β1D + 12047 W + POPC/POPG bilayer*	16066	500	0.28
F2-F3-αIIb-β1D-SA	F2-F3-WT-αIIb-β1D + 5858W + POPC + POPG§	10785	200	0.14
F2-F3-4E-αIIb-β1D-SA	F2-F3-4E-αIIb-β1D + 5774W + POPC + POPG§	10734	10x200	0.13±0.01

*These simulations used a pre-formed lipid bilayer. §These simulations used self-assembly to form the bilayer in the presence of the protein, starting with randomly positioned and oriented lipid molecules. The final root mean square deviation (RMSD) was calculated over the whole trajectory for each simulation using the backbone particles, and measured relative to the starting structure for that simulation. After calculating the RMSD for all repeats of simulation of a given system the mean value of the RMSD was calculated.

METHODS

Experimental procedures:

Protein expression and purification for NMR:

U-^{15}N-labeled F2, F2-4E and F2-F3 domains were expressed as GST-fusion proteins in M9 minimal media using an ^{15}N source. Cells were harvested by centrifugation and resuspended in phosphate-buffered saline. Lysozyme (1 mg/mL), 1 M $MgSO_4$ (10mL/mL) and 1 mg/mL DNAse1 solution (10 µL/mL) were added. After three freeze/thaw cycles, 10% Triton was added and the sample was centrifuged at 40000 G for 10 min at 4°C. A glutathione-separose 4B (Roche Applied Science) column was set up for purification and the protein was eluted in fractions of 15 mM glutathione in 50 mM Tris (~ pH 5.8). Glutathione S-transferase fusion protease $3C^{pro}$ was added to cleave the purified fusion protein. The protein was further purified by gel filtration chromatography into NMR buffer (50 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, pH 7.0). The F2-F3-4E domain was expressed as a His-tag fusion protein and purified as mentioned above, but with Talon resin (Takara Bio) in 50 mM sodium phosphate, 300 mM NaCl, 0.035% β-mercaptoethanol, pH 7.0 buffer. TEV protease was used to cleave the polyhistidine tags and gel filtration chromatography into pH 7.0 NMR buffer was used for further purification.

Liposomes:

To construct the DMPS/PC 1:4 mol %, where DMPS is L-α-phosphatidylcholine/1,2-dimyristoyl-*sn*-glycero-3-phospho-L-serinenad and PC is L-α-phosphatidylcholine (Egg, Chicken) with predominant species 16:0/18:1, DMPS and PC lipids (in chloroform) were mixed at the required ratio. The chloroform was removed using a stream of argon gas and the sample was attached to a vacuum pump overnight. 1 mL of 50mM phosphate, 100 mM NaCl and 1 mM DTT was added to the liposomes and the sample was agitated. At 40 ^{o}C the sample was rotated for 1 hour. The LMVs (large multilamellar vesicles) were disrupted by five freeze-thaw cycles and an Avanti mini-extruder kit was utilized to extrude the sample in order to size the liposomes, following the protocol specified on the Avanti website (http://avantilipids.com/) which states that unilamellar liposomal suspensions with a low polydispersity can only be prepared with membranes having a pore size of ≤200 nm. Accordingly, the suspension was extruded through a 100 nm polycarbonate membrane. The sample was maintained at a temperature of 40 ^{o}C (i.e. above the phase transition temperature of the lipids). The sample was extruded back and forth 20 times. Vesicle suspensions were stored for a maximum of 3-4 days at 4 °C in order to avoid any significant change in vesicle

size distribution. The DMPS liposomes were constructed by using the same procedure as described above.

NMR spectroscopy:

NMR experiments were performed at 40 °C on 600 MHz NMR spectrometers with Oxford Instruments superconducting magnets and GE Omega consoles. All experiments with both wild and mutated F2-F3 and F2 constructs with and without liposomes were performed using samples of 75, 100 and 150 μM of U-^{15}N-labeled protein in 5% D_2O / 95% H_2O. All samples were buffered with 50mM phosphate and 100 mM NaCl. In addition, 5mM DTT was added to prevent oxidation. The samples with liposomes contained 10 mM lipids. All samples were set to pH 6.5. For each sample a two dimensional heteronuclear single quantum coherence (HSQC) spectrum was acquired. The resonance assignments of the F2 domain were made using three dimensional (3D) gradient-enhanced f-NOESY-HSQC (τ_m= 100 ms) and {^{1}H-^{15}N}-TOCSY-HSQC (τ_m= 70 ms) spectra. The NMR assignments of the F2-F3 domains and F2 domain that were used in this study have been submitted to the BiomagresBank: F2 (196-309), BMRB ID=16930 F2F3 (196-405), BMRB ID=169n32. NMR experiments were processed using NMRpipe (Delaglio et al., 1995) and visualized using SPARKY (www.cgl.ucsf.edu/home/sparky). The ratio R_{SN} which characterizes the strength of the interactions, was calculated using the equation

$$R_{SN} = \frac{S_{protein\ with\ liposome}}{S_{protein\ without\ liposome}},$$

where S is the signal to noise ration for each spectrum. The signal to noise was calculated using NMRPipe and Sparky (www.cgl.ucsf.edu/home/sparky). Individual peak R_{SN} values were calculated using the signal to noise ratio of the peak concerned, with and without liposomes. The overall R_{SN} value for a particular condition, was calculated as the mean of the individual peak values ± the standard deviation. For the R_{SN} calculations, 106 and 201 peaks for F2 and F2-F3, were used respectively.

CG-MD simulations:

The crystal structure of F2-F3 bound to the β1D TM and tail domain (PDB: 3G9W) and the structure of the αIIbβ3 complex (PDB: 2K9J) were converted to a CG representation (Bond et al., 2007; Bond and Sansom, 2006; Bond et al., 2008) providing initial structures for the simulations. The CG simulations used a local variant (Bond et al., 2008) of the MARTINI (Monticelli et al., 2008) coarse-grained forcefield in which there is an approximate 4:1

mapping of heavy atoms to CG particles. The secondary and tertiary structure of the protein domains was modelled as an elastic network by imposing a harmonic force constraint (force constant = 10 kJ/mol/Å^2) (Atilgan et al., 2001) between all backbone particles that were within a cutoff distance of 7 Å.

In the CG-MD simulations with the preformed POPC/POPG bilayer, the F2-F3 or F2 domain was displaced 100 or 80 Å away from a preformed bilayer containing POPC and POPG at a ratio of 3:2 (154 POPC and 102 POPG molecules). Each system was solvated with CG water particles and with CG sodium particles to neutralize the system. The system was energy minimized, and after 5 ns equilibration (with restraints applied to Ca atoms, force constant = 10 kJ/mol/Å^2); production simulations were run for 0.5 ns and 1 μs respectively for F2 and F2-F3. Five simulations were performed for each system with all simulations starting from the same configuration but with different initial velocities. The wild type and the mutated protein simulations of each system started from the same initial configuration. In the self-assembly simulations, POPC and POPG lipids at a ratio of 3:2 (154 POPC and 102 POPG) were randomly positioned in a cubic box along with the protein, CG water particles, and counterions. The simulation was run for 200 ns after an initial 400 steps of energy minimization.

All CG-MD simulations were performed using GROMACS 3.3.3 (www.gromacs.org) (Lindahl et al., 2001; van der Spoel et al., 2005). A Berendsen thermostat (Berendsen et al., 1984) was used for temperature coupling with a coupling constant of 1.0 ps. The reference temperature was 310 K. Electrostatic/Coulombic interactions utilized a relative dielectric constant of 20. Lenard-Jones and Electrostatic/Coulombic interactions were shifted to zero between 9 Å and 12 Å and 0 and 12 Å, respectively. A Berendsen barostat with a coupling constant of 1.0 ps, a compressibility value of 5.0×10^{-6} bar^{-1} and a reference pressure of 1 bar was used. The integration time step was 40 fs and coordinates were saved and analyzed every 400 ps. Visualization was via VMD (Humphrey et al., 1996).

REFERENCES

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys. J. *80*, 505-515.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684-3690.

Bond, P.J., Holyoake, J., Ivetac, A., Khalid, S., and Sansom, M.S.P. (2007). Coarse-grained molecular dynamics simulations of membrane proteins and peptides. J. Struct. Biol. *157*, 593–605.

Bond, P.J., and Sansom, M.S.P. (2006). Insertion and assembly of membrane proteins via simulation. J. Amer. Chem. Soc. *128*, 2697-2704.

Bond, P.J., Wee, C.L., and Sansom, M.S.P. (2008). Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. Biochem. *47*, 11321-11331.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277-293.

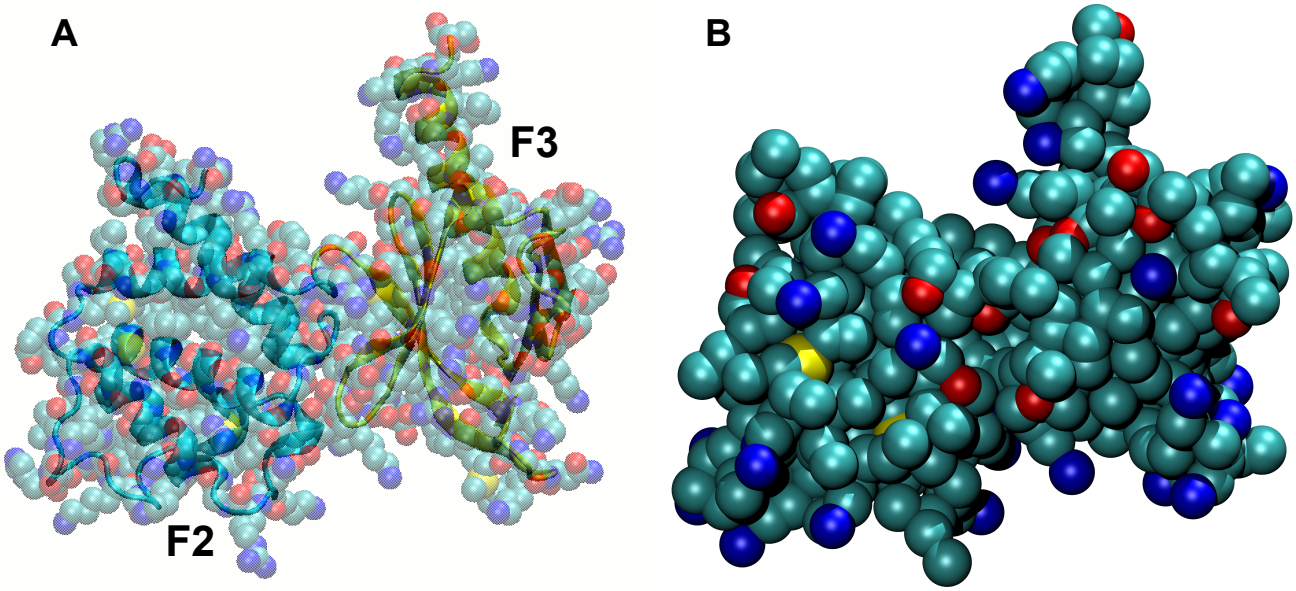
Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD - Visual Molecular Dynamics. J. Molec. Graph. *14*, 33-38.

Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Molec. Model. *7*, 306-317.

Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.J. (2008). The MARTINI coarse grained force field: extension to proteins. J. Chem. Theor. Comp. *4*, 819-834.

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J. Comput. Chem. *26*, 1701-1718.

Figure 1:



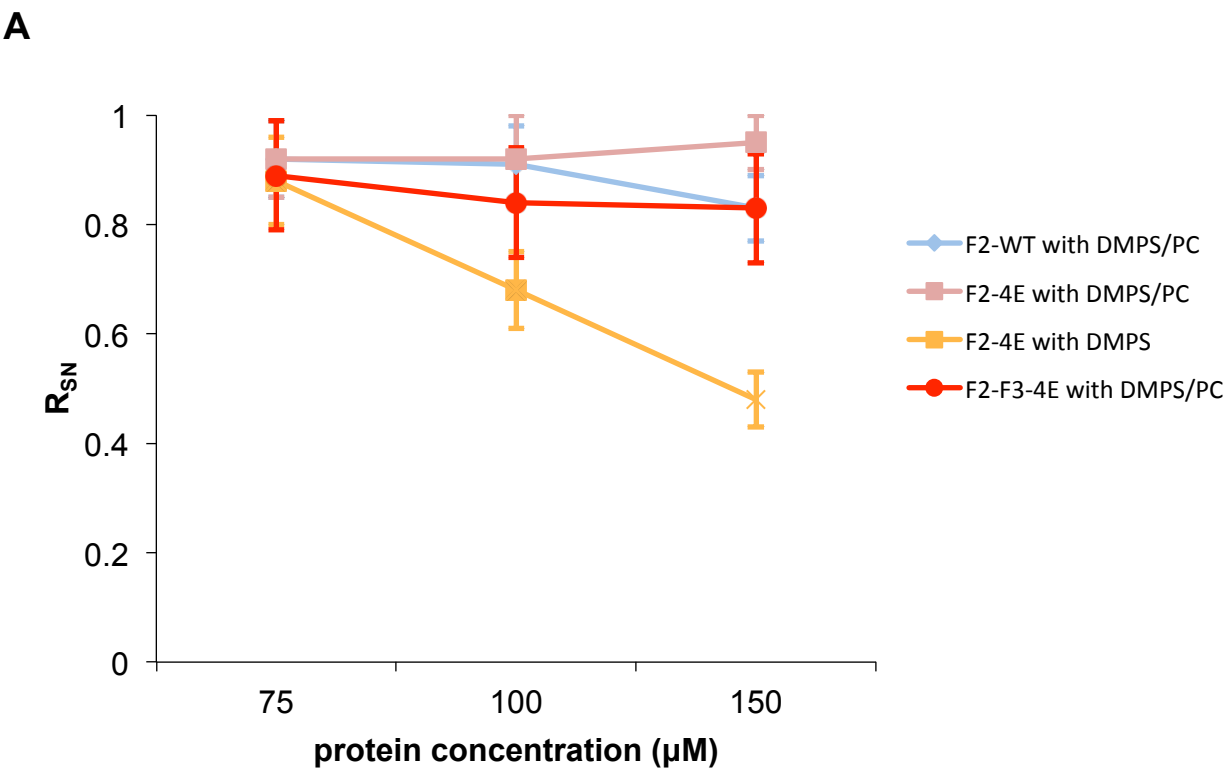
C

```
Tal2 198 KFF YSDQNVDSRD PVQLNLLYVQ ARDDILNGSH PVSFEKACEF GGFQAQIQFG PHVEHKH 257

Tal2 258 KPG FLDLKEFLPK EYIKQRGAEK RIFQEHKNCG EMSEIEAKVK YVKLARSLRT YGVSFFL 317
           E                 E  E    E
Tal2 318 VKE KMKGKNKLVP RLLGITKDSV MRVDEKTKEV LQEWPLTTVK RWAASPKSFT LDFGEYQ 374

Tal2 378 ESY YSVQTTEGEQ ISQLIAGYID IILKKKQS 408
```

Figure 2:


A
1
0.8
0.6
0.4
0.2
0
R_{SN}
75
100
150
protein concentration (μM)
F2-WT with DMPS/PC
F2-4E with DMPS/PC
F2-4E with DMPS
F2-F3-4E with DMPS/PC

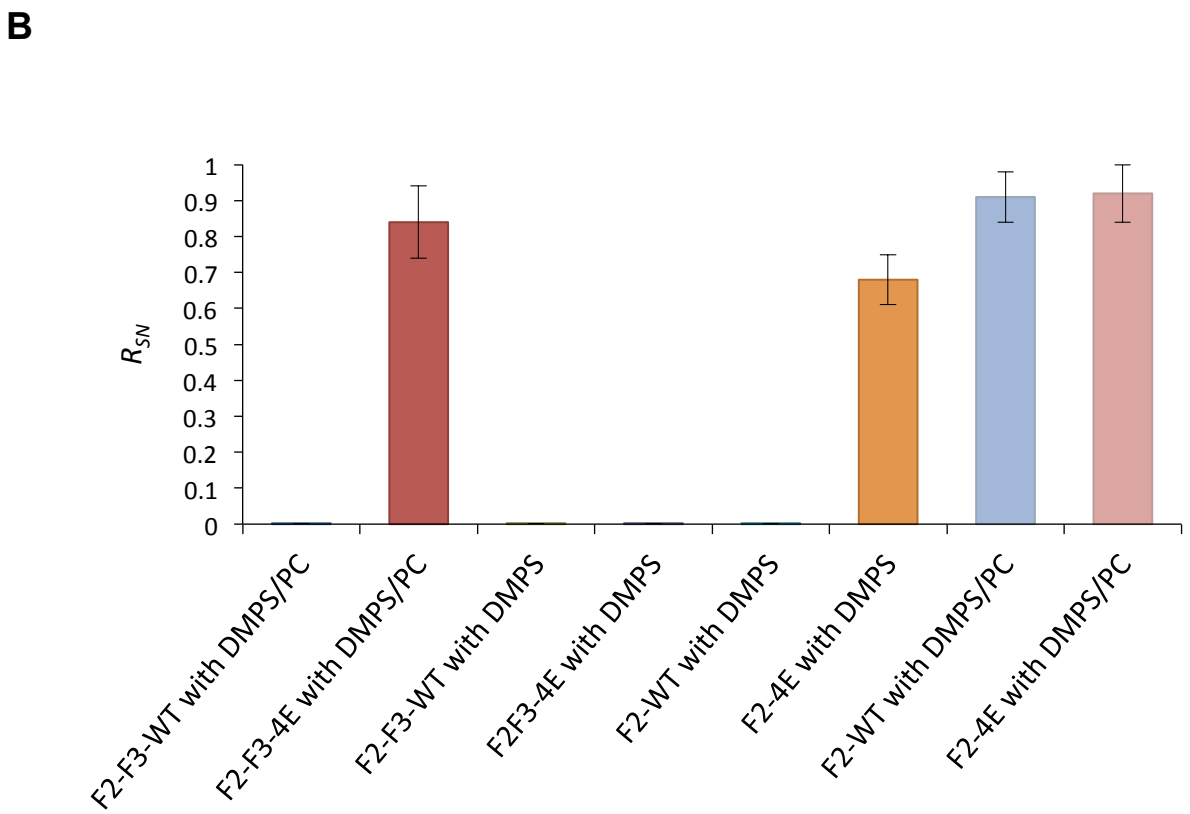

B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
R_{SN}
F2-F3-WT with DMPS/PC
F2-F3-4E with DMPS/PC
F2-F3-WT with DMPS
F2F3-4E with DMPS
F2-WT with DMPS
F2-4E with DMPS
F2-WT with DMPS/PC
F2-4E with DMPS/PC

Figure 3:

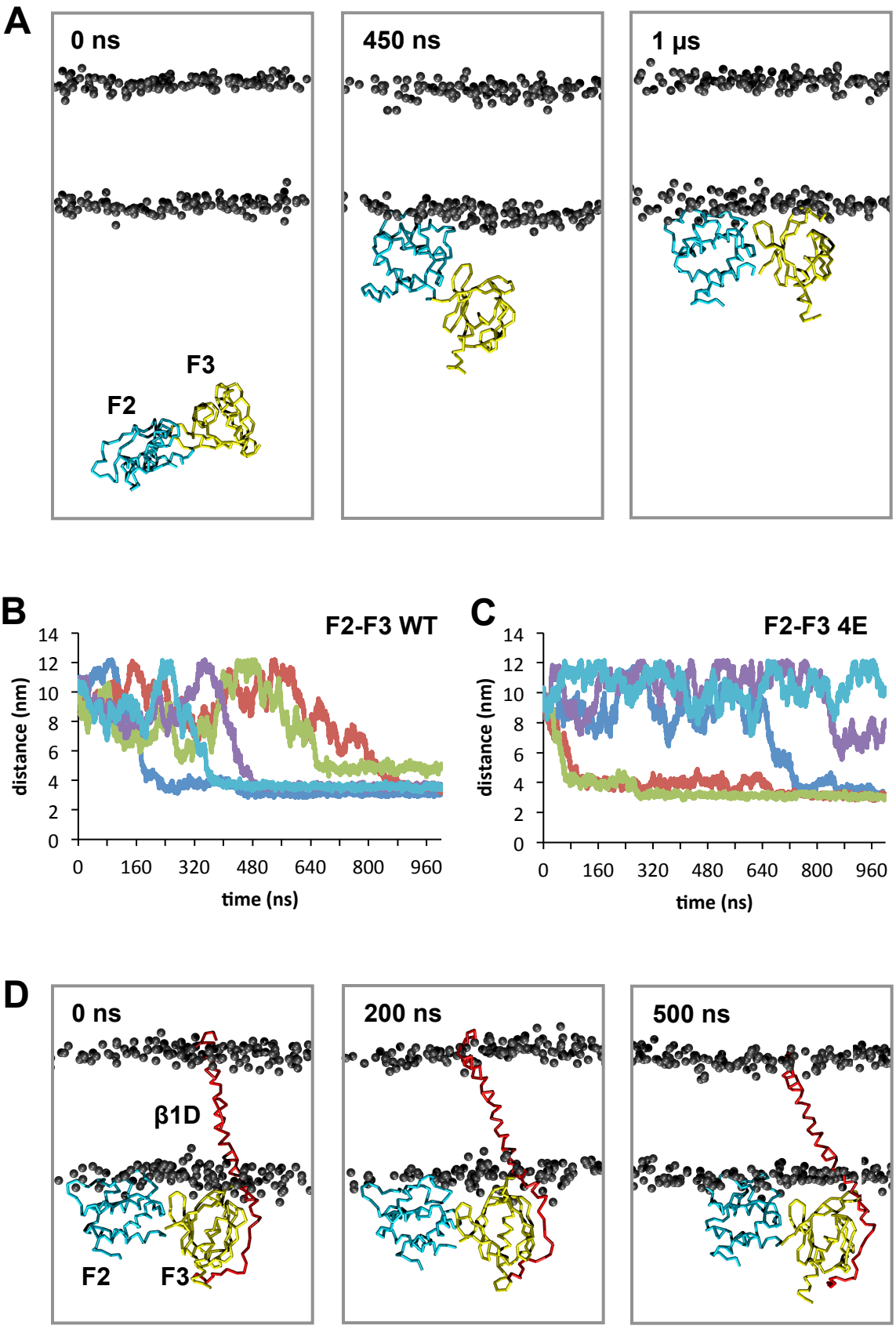

A
0 ns
450 ns
1 µs
F3
F2
B
F2-F3 WT
distance (nm)
time (ns)
C
F2-F3 4E
distance (nm)
time (ns)
D
0 ns
200 ns
500 ns
β1D
F2
F3

Figure 4:

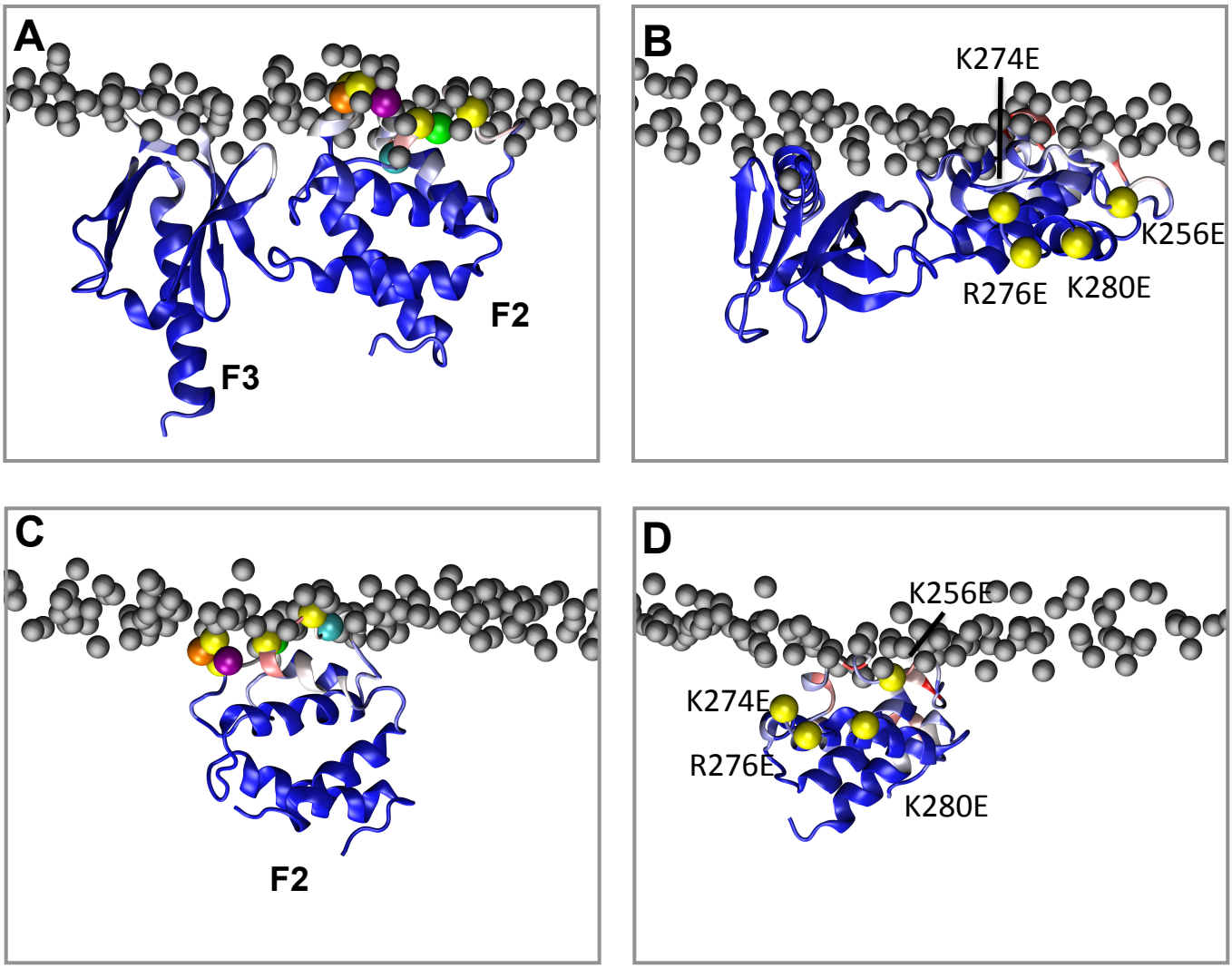

A
F2
F3
B
K274E
K256E
R276E
K280E
C
F2
D
K256E
K274E
R276E
K280E

Figure 5:

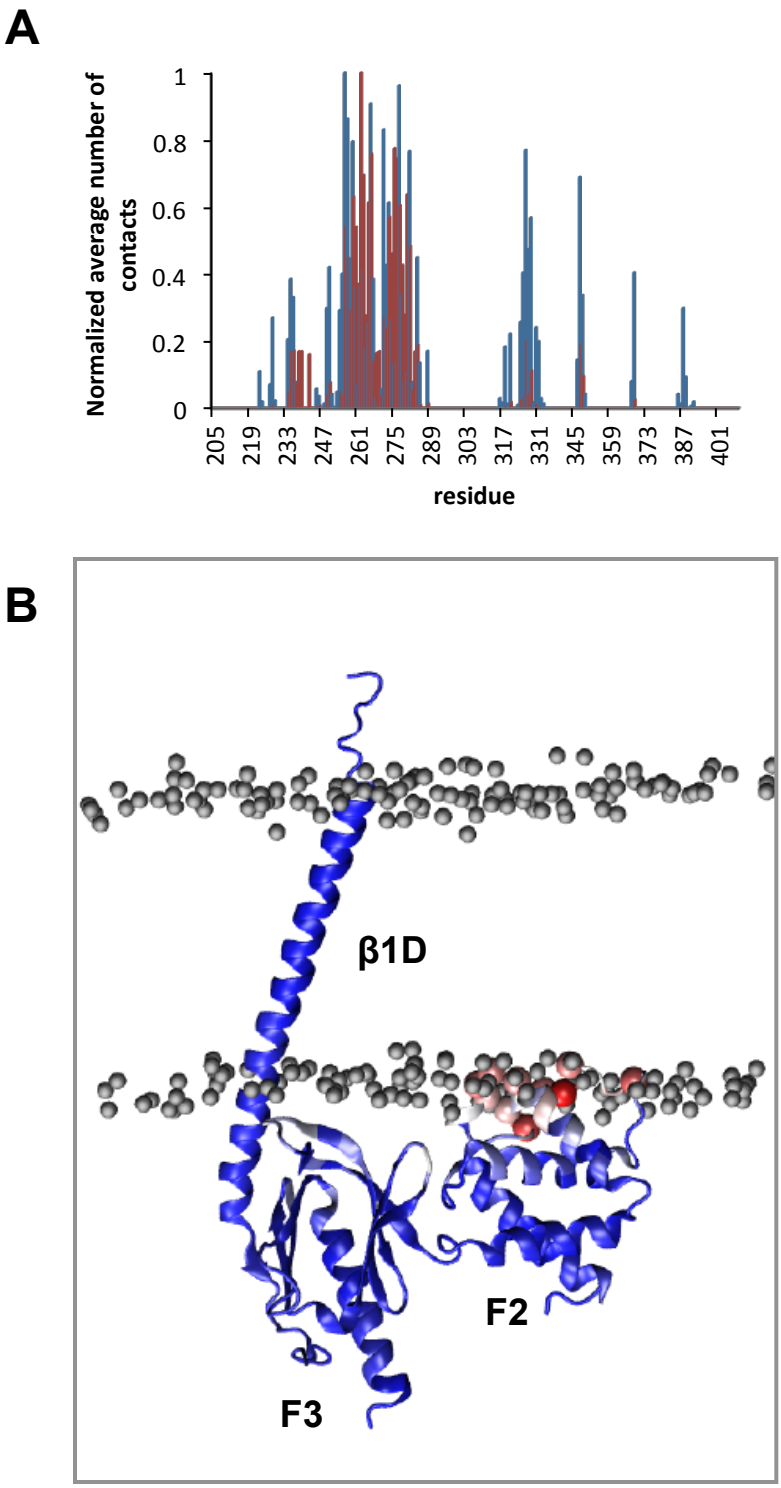
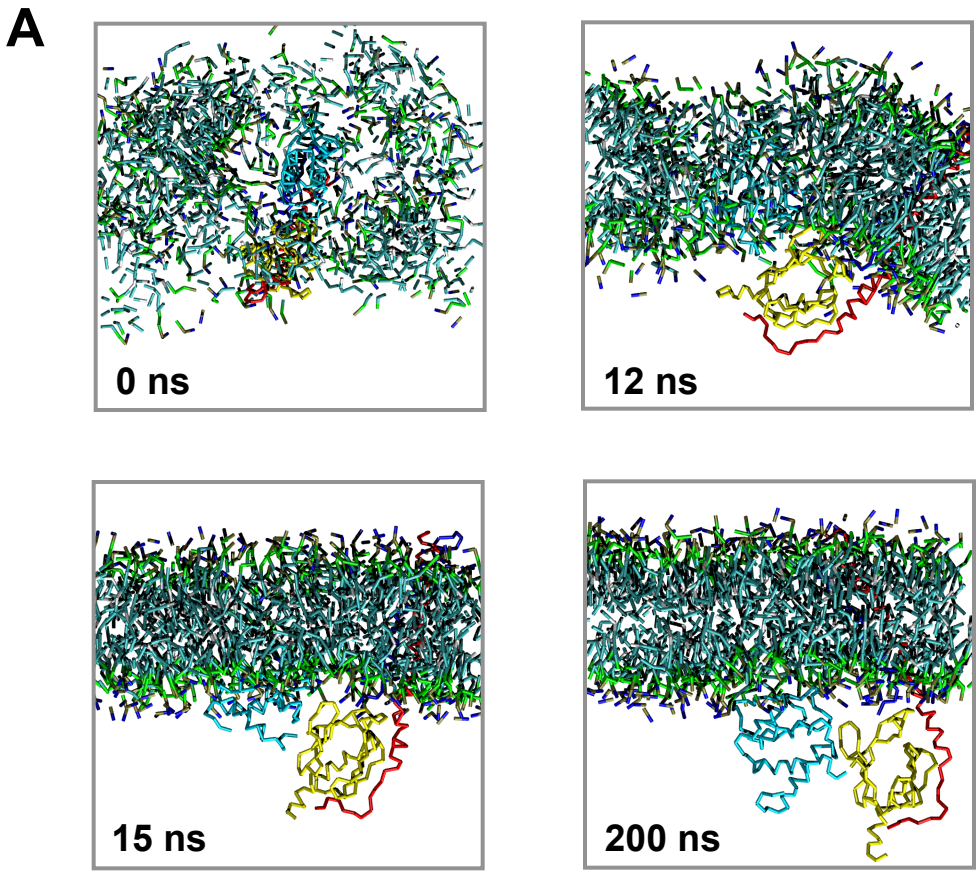
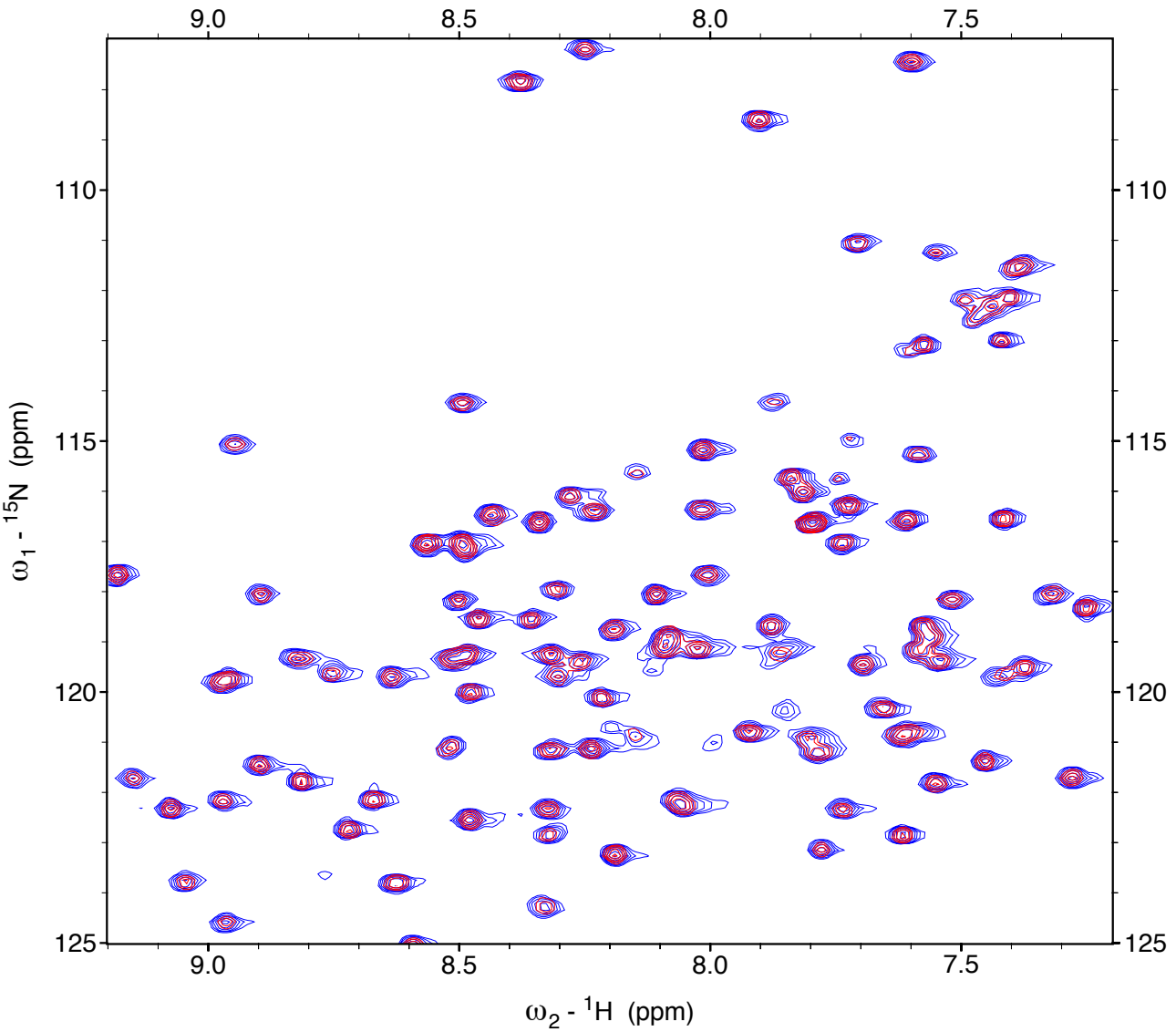

A
Normalized average number of contacts
1
0.8
0.6
0.4
0.2
0
205
219
233
247
261
275
289
303
317
331
345
359
373
387
401
residue
B
β1D
F2
F3

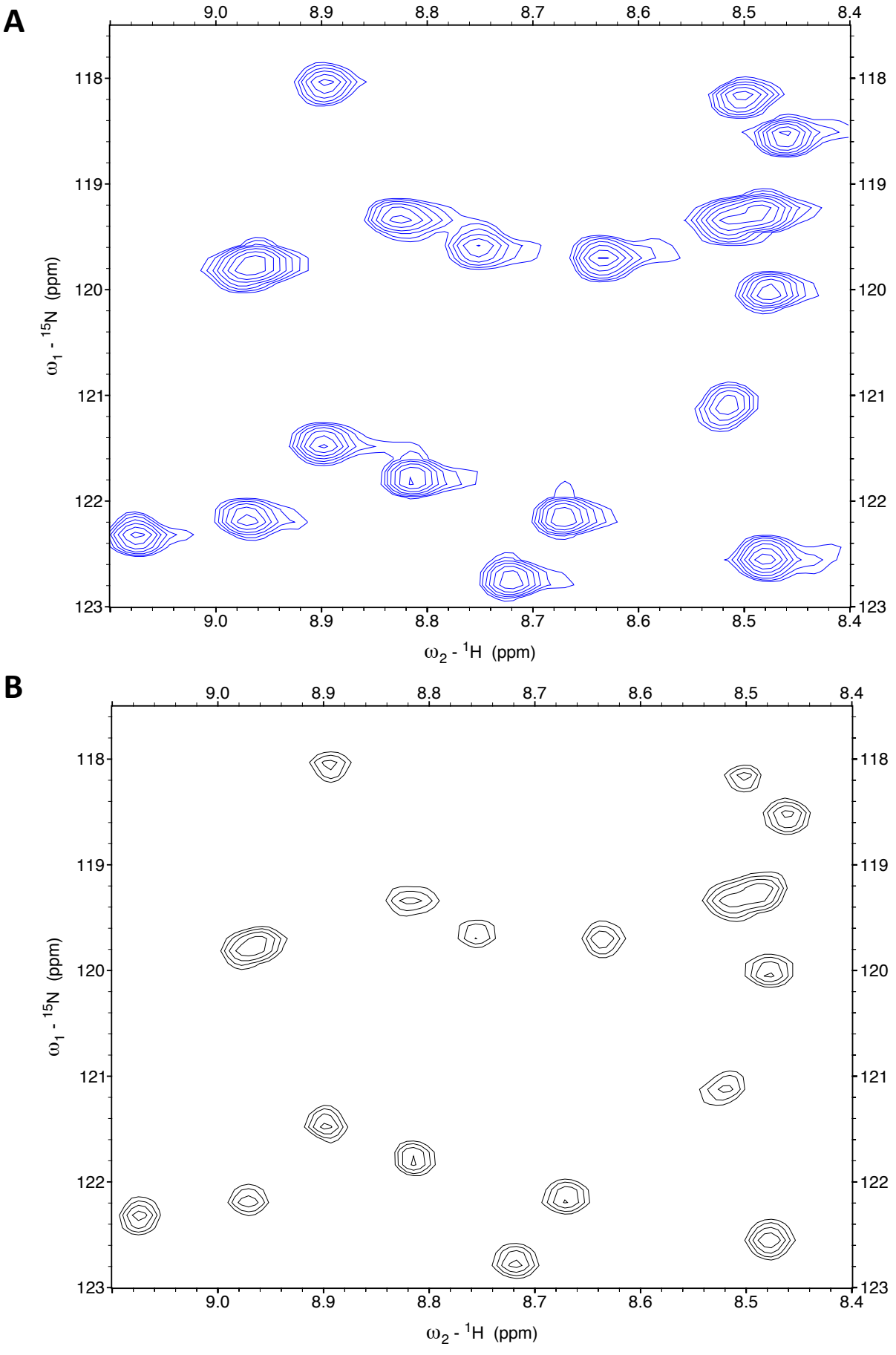
Figure 6:



Supplementary figure S1:



Supplementary figure S2:



A HELIX HETERODIMER IN A LIPID BILAYER: STRUCTURE PREDICTION OF THE STRUCTURE OF AN INTEGRIN TRANSMEMBRANE DOMAIN VIA MULTISCALE SIMULATIONS

*Antreas C. Kalli, Benjamin A. Hall, Iain D. Campbell & Mark S.P. Sansom**

Department of Biochemistry,
University of Oxford,
South Parks Road,
Oxford,
OX1 3QU,
U.K.

*To whom communications should be addressed:
e-mail: mark.sansom@bioch.ox.ac.uk
phone: +44 1865 613306

This paper was published in structure.

See http://www.sciencedirect.com/science/article/pii/S0969212611002565

HIGHLIGHTS

- Multiscale simulation have been used to model transmembrane helix heterodimers
- This method is applied to the integrin αIIb/β3 heterodimer
- The model exhibits right-handed packing of the helices, in agreement with NMR structures
- Flexibility of the αIIb/β3 interface suggests of a role for alternative packing modes in transbilayer signalling.

SUMMARY

Dimerization of transmembrane (TM) α-helices of membrane receptors plays a key role in signalling. We show that molecular dynamics simulations yield models of integrin TM helix heterodimers which agree well with available NMR structures. We use a multiscale simulation approach, combining coarse-grained and subsequent atomistic simulation, to model the dimerization of wild type (WT) and mutated sequences of the αIIb and β3 integrin TM helices. The WT helices formed a stable, right-handed dimer with the same helix-helix interface as in the published NMR structure (PDB:2K9J). In contrast, the presence of disruptive mutations perturbed the interface between the helices, altering the conformational stability of the dimer. The αIIb/β3 interface was more flexible than that of e.g. glycophorin A, This is suggestive of a role for alternative packing modes of the TM helices in transbilayer signalling.

INTRODUCTION

Lateral association of transmembrane (TM) α-helices within a lipid bilayer has been studied from a number of perspectives (Bowie, 2005; Langosch and Arkin, 2009; Senes et al., 2004). In particular, TM helix dimerization is important in mechanisms of signalling across membranes by membrane bound receptors, which have ectodomains that bind extracellular ligands and are linked to intracellular signalling domains via single TM helices (Marks et al., 2009). Mutations in TM helices of receptors can perturb signalling and hence result in disease phenotypes (Li and Hristova, 2006). TM helix dimerization also provides a simple model of the lateral association step in membrane protein folding (Bowie, 2005), helping to cast light on the role of simple sequence motifs, such as those involving glycine residues, in providing stable interactions between TM helices (Curran and Engelman, 2003; Javadpour et al., 1999; Russ and Engelman, 2000; Senes et al., 2004).

Molecular dynamics simulations and related computational approaches provide useful tools to model the structure and dynamics of membrane proteins (Khalili-Araghi et al., 2009; Lindahl and Sansom, 2008) and can provide insights into TM helix packing (Psachoulia et al., 2008; Psachoulia et al., 2010). For example, a range of simulation techniques including atomistic and coarse-grained MD simulations (Braun et al., 2004; Cuthbertson et al., 2006; Petrache et al., 2000; Psachoulia et al., 2008; Sengupta and Marrink, 2010)], Monte Carlo methods (Vereshaga et al., 2005), potential of mean force calculations (Janosi et al., 2010), and knowledge-based methods (Chugunov et al., 2007) have been used to study the dimerization of TM helices such as glycophorin A (GpA). GpA contains a GxxxG motif in

the TM region which is shown to be important for the packing and the stability of GpA homodimer (Anbazhagan and Schneider, 2010; Brosig and Langosch, 1998; Doura and Fleming, 2004; Doura et al., 2004; Russ and Engelman, 1999; Russ and Engelman, 2000).

Integrins provide an important and well studied example of cell surface receptors whose TM helices are involved in TM signalling (Ginsberg et al., 2005; Harburger and Calderwood, 2009; Hynes, 2002). In mammals, there are 18 different integrin α subunits which may heterodimerize with 8 different β subunits to form 24 different integrins. Integrins exist in equilibrium between two states, an inactive low affinity and an active high affinity states (Wegener and Campbell, 2008). Transmembrane helix-helix association plays a crucial role in maintaining the inactive state (Lau et al., 2009; Luo et al., 2004; Schneider and Engelman, 2004; Yang et al., 2009). In the presence of activating proteins such as talin, integrins are believed to activate either by dissociation or by changes in packing of the two TM domains (see e.g. (Anthis et al., 2009; Wegener et al., 2007).

Two recent NMR structures of the αIIbβ3 TM dimer explore the role of the packing of the two integrin TM helices in maintaining the integrin inactive state (Lau et al., 2009; Yang et al., 2009). One structure (pdb id 2K9J; (Lau et al., 2009)) was obtained in a phospholipid bicelle, the other (pdb id 2KNC; (Zhu et al., 2009)) in a non-aqueous solvent system. Although showing some differences in the interactions beyond the TM region in the immediate cytoplasmic region of the integrin subunits, the packing of the TM helices is similar in the two structures. In particular the so-called outer membrane clasp (OMC) region shows close packing of the two helices mediated by a Gx3G motif in the αIIb helix. Mutational changes (e.g. αIIbG972L, αIIbG976L/I and β3G708L/I) in the dimer interface affect dimerization, probably because of induced steric clashes (Berger et al., 2010; Li et al., 2005; Luo et al., 2005).

The integrin TM heterodimer is a more complex system than GpA, and thus provides a valuable test of computational approaches to modelling TM helix dimers of membrane receptors. We have used a multiscale MD approach (combining CG-MD and AT-MD simulations) to explore the αIIb/β3 integrin TM helix heterodimer. We show that simulation methods predict a structure similar to that observed in NMR studies and that mutations suggested to disrupt the dimer, perturb the packing of the TM helices.

RESULTS

Dimerization of TM helices by CG-MD

Our starting point was to use CG-MD simulations to self-assemble integrin TM helix dimers to explore the nature of the predicted helix packing, and in particular to define the role of the Gx_3G motif of the αIIb helix, within the outer membrane clasp – see above. To this end we started with the sequences of the TM domains of the αIIb and β3 subunits (Fig. 1A), as defined in a number of structural studies (Lau et al., 2009). Note that this excludes the non-helical inner membrane clasp from the C-terminus of the αIIb TM helix associated with a GFFKR motif (Lau et al., 2009).

The CG-MD simulations were performed using a high throughput methodology (Hall et al., 2011) which enabled automatic running of multiple self assembly simulations, this permitting statistical analysis over an ensemble of ca. 100 structures. In each simulation (Fig. 1B) the two helices were initially positioned at a distance of ~60 Å away from one another. This separation between the helices is therefore significantly larger than the cutoff distance for electrostatic and van der Waals interactions, and so no inter-helix interactions occur at the beginning of the simulation. This ensures that the initial position does not favour dimer formation. During the simulation the helices diffuse randomly in the bilayer relatively to one another until an encounter leads to formation of a stable dimer occurs, usually within a few hundred nanoseconds (Supporting Information Fig. S5A). Visual inspection of the helix dimers reveals them to be largely right-handed (RH), as expected for packing involving a Gx_3G motif (Russ and Engelman, 2000), although there was a degree of variation in helix packing, both within individual simulations as a function of time and between simulations within the ensemble. This dynamic aspect of TM helix packing can also be seen if one analyses the distribution of helix crossing angles for the whole CG-WT ensemble (Fig. 1C); this shows a bimodal distribution, as suggested by preliminary studies (Psachoulia et al., 2010). The major interaction mode (~70% of the structures) corresponds to right-handed packing of the helices, with a mean crossing angle of ~-30°. We note that this agrees well with the RH packing of the αIIbβ3 TM dimer seen in NMR structures determined in lipid bicelles and in non-aqueous (CH_3CN/H_2O) solution (both -27°) (Lau et al., 2009; Yang et al., 2009). The minor mode (30%) corresponding to left-handed (LH) packing of the TM helices, was not observed by NMR.

Visualization of selected RH and LH structures from the CG-WT ensemble (Fig. 1D) reveals a significant difference in the helix/helix interface; in the RH structure the two glycines of the Gx_3G motif of αIIb (G972 and G976) are directed *towards* the β3 helix whilst

in the LH dimer they are directed *away* from the other helix. To further investigate the packing of the two helices, spatial distributions of the αIIb helix relative to the β3 helix were calculated for the CG-WT simulations. The trajectories for all of the CG-WT simulations (see Table 1) were concatenated and the backbone particles of the β3 helix were fitted to a reference structure in order to calculate the probability density of the backbone particles of the αIIb helix at a given point in the bilayer plane around the β3 helix (Fig. 2A). This revealed three maxima, one major (RH1) and two minor (RH2 and LH). These were shown to correspond to three packing modes of the helices, namely the major RH helix dimer (RH1) as already discussed, a minor RH dimer (RH2), and the LH dimer. To test if the initial position of the helices affects the dimerization of the two helices, similar simulations with the two helices in the same initial positions but with the β3 helix rotated by a 15^{o} increment in each subsequent simulation were performed (~100 simulations). The results were identical (see Supporting Information Fig. S5B) with the above simulation (i.e CG-WT) and therefore for the remaining study we continued using the protocol as described in the methods section.

In addition to these studies of helical transmembrane regions involving the OMC (Lau et al., 2009), we explored the role of residues in the IMC (see Supporting Information Fig. S1). Extending the C-terminus of the αIIb subunit by 3 residues to include aromatic residues of the IMC did not greatly perturb the TM helix packing, and a slightly modified RH conformation was observed to be the major form. Extension of the α-helix by a further two residues to include two basic residues resulted in perturbed helix packing and strong interactions with lipid phosphate groups. The loosening of the native packing in the presence of the IMC was also suggested by recent CG simulations using the MARTINI forcefield (Chng and Tan). Thus the remainder of this study focussed on the TM region as defined in Fig. 1.

Model assessment and refinement by atomistic MD simulations

In order to refine and evaluate further the model(s) of the αIIb/β3 TM helix dimer from the CG-MD self-assembly simulations, the CG structures were converted to full atomistic models, using CG2AT a fragment based procedure which has been previously tested for a number of membrane protein systems (Stansfeld and Sansom, 2011). This then allowed us to perform atomistic MD simulations of the models of the TM helix dimer. Such simulations have been used previously to aid in evaluation of models of membrane proteins (Psachoulia et al., 2010; Volynsky et al., 2010). We performed 3 simulations of duration 30 ns for each structure generated by CG2AT (Fig. 3; Table 1).

An initial assessment of the conformational stability of the three WT models (RH1, RH2 and LH; see above) was made by comparing their Cα root mean squared deviations from the initial structures in the corresponding simulations (see Fig. 3A). From this it can be seen that the Cα RMSD for the RH1 dimer rapidly reaches a plateau at a value of less than 3 Å, as is generally the case for 'stable' membrane protein models (Law et al., 2005). We note that equivalent simulations starting from the NMR structure of the αIIb/β3 TM helix dimer (see more detailed discussion below) yielded a comparable Cα RMSD after 30 ns of ~2.5 Å. An additional simulation, using the OPLS forcefield, was performed. The yielded a a Cα RMSD of less that 3 Å, characteristic of a stable dimer (Supporting Information Fig. S7). In contrast, the Cα RMSDs for the RH2 and the LH dimer simulations rise steadily over the course of the simulations, reaching ~5 Å for RH2 and ~6 Å for LH after 30 ns. This behaviour is characteristic of an 'unstable' membrane protein model. Similarly, the AT-RH1 simulations yielded a relatively narrow helix crossing angle distribution (see below) characteristic of a stable RH dimer. In contrast, the crossing angle distribution was considerably wider for both the RH2 and LH dimer simulations than for RH1, ranging from -60° to 0° for the RH2 dimer, and from +10° to +60° for the LH dimer, suggesting 'looser' helix packing (Supporting Information Fig. S6). Previous atomistic MD simulations of the αIIbβ3 NMR structure (i.e. containing the GFFKR motif) also reveal a conformationally stable dimer (Kalli et al., 2011).

Effects of mutations presumed to be at the helix/helix interface

Various studies have suggested that mutations in glycine residues thought to form the αIIb/β3 TM helix/helix interface (e.g. αIIbG972L, αIIbG976L/I and β3G708L/I mutations) disrupt helix dimerization (Berger et al., 2010; Li et al., 2005; Luo et al., 2005). We therefore explored the structure and conformational stability of these mutants via multiscale (i.e. CG then AT) simulation.

At the level of CG simulations the relative locations of the two helices in the bilayer plane (Fig. 2C) indicate changes in the mode of helix packing in the mutant simulations, that disfavour the RH1 mode which is predominant in WT. For each of the mutant simulations, a CG structure from the favoured position in the spatial distribution analysis for each system was converted to an atomistic structure and further (atomistic) MD simulations were initiated (see Table 1 for details). For all of the mutant dimers the Cα RMSD rose quickly to significantly higher values (5 to 7 Å after 30 ns) than for the AT-RH1 simulations (3 Å). The crossing angle distributions for all the mutant dimer simulations were also wider compared to the WT, ranging from -90° to 0° (Supporting Information Fig. S2). Thus, the atomistic

simulations confirm the suggestion from CG of weaker helix packing interactions and conformationally less stable dimers in the mutants.

Comparison to the NMR structure

It is valuable to compare the most stable structure predicted by the simulations (i.e. AT-RH1) to the structures from NMR studies of the αIIb/β3 TM helix dimer, especially the structure solved in a membrane-like environment (2K9J (Lau et al., 2009)). The environments in the simulation (a DPPC bilayer in the fluid phase; Fig. 4A) and in the NMR studies (a phospholipid bicelle, containing DHPC alongside POPC or DMPC) are similar but non-identical (Cross et al., 2011; Raschle et al., 2010). In addition to comparing the AT-RH1 simulation with the experimental NMR structure, we also compared it with a simulation (AT-NMR; Table 1) initiated from the NMR structure (of *just* the residues in the TM helix dimer, i.e. the same residues as in AT-RH1) in a DPPC bilayer.

Comparing one of the final structures from the AT-RH1 simulations with the NMR structure (2K9J) yields a Cα RMSD of 2.2 Å between the simulated dimer and the same (TM helix) region of the NMR structure (see Supporting Information Fig. S3). Furthermore, both structures have a crossing angle of -30±3° and a helix/helix interface formed by the same residues (Fig. 4B). These crossing angles are consistent with those seen in the NMR (2K9J) ensemble, for which an average value 25±3° is observed.

Both AT-RH1 and AT-NMR simulations yielded a low Cα RMSD from start to end of the simulations (2.5 to 3 Å; Fig. 3A), indicative of a conformationally stable dimer. The crossing angle distributions for the two simulations are very similar (Fig. 4C), as are the spatial distribution maps (data not shown). Pairwise comparison of structures from the AT-RH1 and AT-NMR trajectories yielded Cα RMSDs of 1-3 Å for the latter halves of each trajectory indicating that both simulations have converged to the same structure within the range of dynamic fluctuations observed in the simulations.

DISCUSSION

Biological implications

The most significant outcome of this study is that the helix/helix interface between the αIIb and β3 integrin TM helices observed in multiscale simulations is the same as that identified in NMR studies (Lau et al., 2009). This demonstrates the predictive capabilities of multiscale MD simulations for identification of helix packing interactions in receptor TM helix dimers even for a heterodimer. It is of particular interest that the correct relative depths

of insertion of the two helices were obtained as well as the correct register for the interfacial contacts for the heterodimer.

The current study has focused on packing of the *transmembrane* regions of αIIb and β3, and so it mainly refers to interactions involving the 'outer membrane clasp' which includes the Gx3G motif of αIIb. The Gx_3G motif has been extensively studied in the context of helix/helix interactions in membrane proteins (Popot and Engelman, 2000; Russ and Engelman, 2000), especially in the model membrane protein glycophorin A (GpA) (MacKenzie et al., 1997). It should be noted that tight RH helix packing in GpA occurs via a close homo-dimeric interaction between two identical Gx3G motifs. In the integrin dimer the interaction is more complex due to the heteromeric nature of the dimer, with only the αIIb helix donating a Gx3G motif. The 'exposed' surface on the αIIb helix corresponding to this motif is packed between the bulky hydrophobic sidechains of Met701 and Ile704 of β3 (Fig. 5) with G708 of β3 packing against the hydrophobic sidechain of Leu980 in αIIb. These interactions are largely conserved in other integrin α and β TM helix sequences (Lau et al., 2009) and the results are in good agreement with structure/function analyses of integrin TM helix interactions via mutagenesis studies (Luo et al., 2005) and TOXCAT (Li et al., 2005) and related assays (Berger et al., 2010) of helix/helix interactions in bacterial membrane systems. Besides the αIIb/β3 heterodimer, recent studies (Chng and Tan; Vararattanavech et al., 2010) for the TM region of the αLβ2 integrin suggest similar packing to that seen in our studies for the helices in the outer membrane clasp (OMC) region (because of the presence of a GxxxG-like motif, i.e SxxxG). This, in combination with the fact that sequence alignment of the different integrin α subunits reveals the presence of a small-xxx-small motif in many integrin α subunits, suggest that this is a general property of integrins (Vararattanavech et al.). In addition, experimental (Lau et al., 2009; Li et al., 2005; Yang et al., 2009) and computational data (Kalli et al., 2011) suggest that disruption of OMC interactions is a key step in switching integrins from an inactive to an active state. Therefore, the conclusions of this study can be generalised for a variety of integrins and help to explain, in combination with the existing experimental and computational data, the role of the OMC in integrin activation.

The αIIb/β3 interface exhibits a somewhat greater degree of flexibility than the GpA/GpA interface, both in the CG and AT-MD simulations. This is not surprising as GpA forms an exceptionally stable helix/helix dimer, compared to e.g. receptor tyrosine kinase TM helix dimers (Finger et al., 2009). In contrast, transmembrane signalling for integrins is likely to involve changes in helix packing, possibly via a 'scissoring' movement (Kalli et al., 2011).

This would be more compatible with a flexible helix/helix interface and/or the existence of alternative packing modes for the TM helices.

We note that alternative modes of TM helix interactions have been observed in NMR structures and discussed in relation to TM signalling by EphA2 receptor tyrosine kinases (Bocharov et al., 2010). This is an aspect of TM helix interaction which merits further pursuit as many widely-used assays for such interactions (e.g. the TOXCAT assay (Dews and MacKenzie, 2007; Finger et al., 2009; Li et al., 2004; Russ and Engelman, 1999)) may average over the range of possible multiple modes of interaction.

Simulation methods and receptor structural biology

Our results demonstrate that a multi-scale MD approach can be used to model and characterize the packing and the conformational stability of TM helix dimers. The use of a high-throughput method ensures better sampling in the CG-MD simulations. The subsequent use of atomistic simulations leads to a refined structure with high degree of similarity with the NMR structures, as discussed above. Thus we combine the strengths of the two methods with CG simulations used for assessing assembly and crossing angles etc. of TM helices whilst the more accurate AT simulation are used to assess the relative stabilities of the resultant dimers. This approach is a valuable addition to the ongoing efforts to use computational approaches to model membrane proteins structures and in particular the conformation, dynamics and energetics of helix/helix interactions in the TM domains of receptors.

The integrin heterodimer is a good test case for modelling TM helix dimerization, as it contains a more complex interaction motif than that in GpA or other similar homodimers (e.g. syndecans (Dews and MacKenzie, 2007; Psachoulia et al., 2010). There have been a number of approaches adopted to modelling the integrin TM helix heterodimer (see e.g. (Gottschalk, 2005; Gottschalk and Kessler, 2004) for an early approach) and indeed more recent studies combining modelling and data-based filtering (Metcalf et al., 2009) have shown good agreement with NMR structures.

In the current study we have shown that a purely simulation based approach including explicit treatment of the bilayer environment can yield an 'NMR accuracy' structure for the integrin TM helix dimer. There have also been recent studies using simulations to select between alternative models (LH and RH) of the EphA2 helix homodimer. In combination, such studies indicate that simulation-based approaches can complement experimental data, providing structural insights into membrane receptor dimers (Jura et al., 2009), and thereby

enabling us to establish the link between receptor structure and mechanisms of cellular function.

METHODS

Coarse-grained MD simulations (CG-MD)

CG simulations used a local modification (Bond et al., 2008) of the MARTINI (Monticelli et al.) forcefield, CG models were constructed for the α-helical regions of αIIb and β3 integrins. These were inserted, parallel and at a distance of ~60 Å to one another, in a phospholipid bilayer containing ~128 DPPC molecules. After the insertion the N and C termini of the two helices were located in the lipid head group region. The N and C termini of the peptide were not acetylated or amidated. A high throughput framework was used to run the CG-MD simulations which allowed ~100 simulations of 500ns each to be performed (Hall et al., 2011). All CG-MD simulations started with the same initial position but with different initial velocities. Analysis for the convergence of the CG simulations suggests that our results converge after using 50% of the simulations for each system (Supporting Information Fig. S4). CG-MD simulations were performed using GROMACS (www.gromacs.org) (Lindahl et al., 2001; van der Spoel et al., 2005). The electrostatic/Coulombic interactions were shifted to zero between 0 and 12 Å and the Lenard-Jones interactions between 9 and 12 Å. A Berendsen thermostat (Berendsen et al., 1984) and barostat (reference temperature 323 K and reference pressure 1 bar) was used for temperature (coupling constant 1 ps) and pressure (coupling constant 10 ps) coupling. The integration step was 20 fs.

From CG to atomistic representation

The conversion from CG to atomistic representation was made as described previously (Stansfeld and Sansom). Briefly, the lipids were constructed by alignment of the coarse-grained particles with energy-minimized atomistic fragments. The CG protein was converted to atomistic using PULCHRA http://cssb.biology.gatech.edu/skolnick/files/PULCHRA/) and MODELLER (http://www.salilab.org/modeller/) and then was energy-minimized using the conjugant gradient method.

Atomistic MD simulations

After the conversion to atomistic representation three simulations for each system starting from the same initial position but with different initial configurations (i.e different initial velocities) were performed. The simulations were performed using the GROMOS96

43a1 forcefield (Scott et al., 1999). The LINCS algorithm (Hess et al., 1997) was used to constrain bond lengths and long-range electrostatic were modelled up to 10 Å using the Ewald Particle Mesh (PME). The same cut-off distance was used for the van der Waals interactions. The Parrinello-Rahman barostat and the Berendsen thermostat were used for pressure and temperature coupling. The reference temperature was 323 K. Before every production simulation every system was energy minimized using the conjugant gradient method and subsequently equilibrated with the protein Cα atoms restrained for 0.5 ns (force constant = 1000 kJ/mol/A^2). Production simulations for 30 ns were performed. The analysis was performed using GROMACS, VMD (Humphrey et al., 1996) and locally written codes.

ACKNOWLEDGEMENTS

This work was funded by the Wellcome Trust (via grants to MSPS and AK) and via the BBSRC via the Oxford Centre for Integrative Systems Biology. Our thanks to Jason Schnell for his interest in this work and his insightful comments on the manuscript.

TABLES

Table 1: Summary of simulations

Simulation	Old name	System	Duration (ns)
CG-WT	αβ1CG	αIIb/β3 WT	~100x500
CG-αG972L	αβ2CG	αIIb (G972L)/β3	~100x500
CG-αG976L	αβ3CG	αIIb (G976L)/β3	~100x500
CG-αG972L,G976L	αβ4CG	αIIb (G972L-G976L)/β3	~100x500
CG-βG708I	αβ5CG	αIIb /β3 (G708I)	~100x500
AT-RH1	αβ1AT	αIIb/β3 WT, RH1 dimer	3x50
AT-RH2	αβ3AT	αIIb/β3 WT, RH2 dimer	3x30
AT-LH	αβ2AT	αIIb/β3 WT, LH dimer	3x30
AT-αG972L	αβ4AT	αIIb (G972L)/β3	3x30
AT-αG976L	αβ5AT	αIIb (G976L)/β3	3x30
AT-αG972L,G976L	αβ6AT	αIIb (G972L,G976L)/β3	3x30
AT-βG708I	αβ7AT	αIIb /β3 (G708I)	3x30
AT-NMR	αβNMR	αIIb /β3 (PDB:2K9J)	3x50

*All simulations used a DPPC lipid bilayer. System sizes ranged from ~8500 particles (CG-MD simulations) to ~41,500 atoms (atomistic simulations)

FIGURES

Figure 1:

Coarse-grained simulations to self assemble an integrin α/β TM helix dimer.

A. Sequence of the αIIb and β3 helices used in the simulations. In each sequence key residues at the helix/helix interface are in green and those that have been mutated are underlined.

B. An illustration of the progress of self assembly of a TM helix dimer by CG-MD simulations. The initial system (0 ns) consists of an αIIb (blue) and a β3 (red) helix in a PC bilayer (shown as the phosphate particles in grey), separated by an interhelix distance of ~60 Å. Subsequent snapshots illustrate the helix dimer at 300 and 500 ns.

Helix crossing angle distribution (**C)** and representative structures (**D**) from the αβ1-CG simulation (see Table 1 for details). A positive value of the crossing angle corresponds to left-handed (LH) helix packing and a negative value to right-handed (RH) packing. The αIIb helix is shown in blue, the β3 helix in red, and the αIIbG972 and αIIbG976 residues in green. The orange spheres represent the phosphate headgroups. Note that in The RH helix dimer the αIIb G972 and G976 residues are directed towards the helix interface whilst in the LH dimer they are directed away from the interface (i.e. towards the viewer).

Figure 2:

Packing of the helices in the dimers assembled by CG-MD simulations.

A. Spatial distributions of the αIIb helix relative to the β3 helix for the CG-WT simulations. The trajectories for all of the WT CG-MD simulations (see Table 1) were concatenated and the backbone particles of the β3 helix were fitted to a reference structure. The diagram shows the probability density of finding the backbone particles of the αIIb helix at a given point in the bilayer plane around the β3 helix, colour coded such that blue represents low probability through to red then green for a high probability. Three maxima are seen in this distribution, one major (RH1) and two minor (RH2 and LH). These correspond to the three representative structures shown in **B** for the major RH helix dimer (RH1), a minor RH dimer (RH2), and a LH dimer.

C. Spatial distribution analysis for CG simulations in which key glycine residues thought to be at the helix/helix interface are mutated to leucine or isoleucine (see Table 1 for details). In each case the distribution is altered as the RH1 dimer is disfavoured.

Figure 3:

Atomistic simulations of αIIb/β3 helix dimers. In each graph the conformational stability of the α/β helix dimer is analysed in terms of the Cα root mean square deviation (RMSD) from the initial structure as a function of time. In **A** the RMSDs are shown for the wild type helix dimer (the CG-RH1 simulations (pink, pale green, grey) vs. the NMR structure (red, green, black)), and in **B** the RMSDs are shown as average across the 3 repeat simulations for simulations of the mutants and the RH2 and LH models from the wild type CG simulations (see Table 1 for details).

Figure 4:

Comparison of the AT-RH1 structure as predicted by simulation in a DPPC lipid bilayer with that of the NMR structure (2K9J; obtained in lipid bicelles (Lau et al., 2009)).

A. Snapshot of the AT-RH1 simulation system after CG2AT conversion of the dimer from the corresponding CG-WT simulation. The helices are shown in a DPPC bilayer; water molecules have been omitted for clarity. **B.** The NMR (2K9J) and simulated (final frame from one of the AT-RH1 simulations) structures of the αIIb/β3 TM helix dimer compared. The αIIb helix is shown in blue, the β3 helix in red, and the αIIbG972, αIIbG976, and β3G708 residues in yellow. **C.** Comparison of the crossing angle distributions from the AT-RH1 (blue) and AT-NMR (red) simulations. Note that the crossing angle of the initial NMR structure (2K9J) was -27°.

Figure 5:

Comparison of the helix/helix interface in the NMR (2K9J) structure and a structure from the end of one of the AT-RH1 simulations. The G972,G976 surface (cyan) of the αIIb helix (blue) can bee seen to pack against the M701 (yellow) and I704 (green) sidechains of the β3 helix (red). The G708 surface (cyan) of β3 is also seen to pack against αIIb. These key interactions are seen in both the NMR and the simulation structure.

REFERENCES

Anbazhagan, V., and Schneider, D. (2010). The membrane environment modulates self-association of the human GpA TM domain-Implications for membrane protein folding and transmembrane signaling. Biochim. Biophys. Acta - Biomembranes *1798*, 1899-1907.

Anthis, N.J., Wegener, K.L., Ye, F., Kim, C., Goult, B.T., Lowe, E.D., Vakonakis, I., Bate, N., Critchley, D.R., Ginsberg, M.H., and Campbell, I.D. (2009). The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. EMBO J. *28*, 3623-3632.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684-3690.

Berger, B.W., Kulp, D.W., Span, L.M., DeGrado, J.L., Billings, P.C., Senes, A., Bennett, J.S., and DeGrado, W.F. (2010). Consensus motif for integrin transmembrane helix association. Proc. Natl. Acad. Sci. USA *107*, 703-708.

Bocharov, E.V., Mayzel, M.L., Volynsky, P.E., Mineev, K.S., Tkach, E.N., Ermolyuk, Y.S., Schulga, A.A., Efremov, R.G., and Arseniev, A.S. (2010). Left-handed dimer of EphA2 transmembrane domain: helix packing diversity among receptor tyrosine kinases. Biophys. J. *98*, 881-889.

Bond, P.J., Wee, C.L., and Sansom, M.S.P. (2008). Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. Biochem. *47*, 11321-11331.

Bowie, J.U. (2005). Solving the membrane protein folding problem. Nature *438*, 581-589.

Braun, R., Engelman, D.M., and Schulten, K. (2004). Molecular dynamics simulations of micelle formation around dimeric Glycophorin A transmembrane helices. Biophys. J. *87*, 754-763.

Brosig, B., and Langosch, D. (1998). The dimerization motif of the glycophorin A transmembrane segment in membranes: importance of glycine residues. Prot. Sci. *7*, 1052-1056.

Chng, C.-P., and Tan, S.-M. Leukocyte integrin αLβ2 transmembrane association dynamics revealed by coarse-grained molecular dynamics simulations. Proteins: Structure, Function, and Bioinformatics, n/a-n/a.

Chugunov, A.O., Novoseletsky, V.N., Nolde, D.E., Arseniev, A.S., and Efremov, R.G. (2007). Method to assess packing quality of transmembrane α-helices in proteins. 1. Parametrization using structural data. J. Chem. Inform. Model. *47*.

Cross, T.A., Sharma, M., Yi, M., and Zhou, H.X. (2011). Influence of solubilizing environments on membrane protein structures. Trends Biochem. Sci. *36*, 117-125.

Curran, A.R., and Engelman, D. (2003). Sequence motifs, polar interactions and conformational changes in membrane proteins. Curr. Opin. Struct. Biol. *13*, 412-417.

Cuthbertson, J.M., Bond, P.J., and Sansom, M.S.P. (2006). Transmembrane helix-helix interactions: comparative simulations of the glycophorin A dimer. Biochem. *45*, 14298-14310.

Dews, I.C., and MacKenzie, K.R. (2007). Transmembrane domains of the syndecan family of growth factor coreceptors display a hierarchy of homotypic and heterotypic interactions. Proc. Natl. Acad. Sci. USA *104*, 20782-20787.

Doura, A.K., and Fleming, K.G. (2004). Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer. J. Mol. Biol. *343*, 1487-1497.

Doura, A.K., Kobus, F.J., Dubrovsky, L., Hibbard, E., and Fleming, K.G. (2004). Sequence context modulates the stability of a GxxxG-mediated transmembrane helix-helix dimer. J. Mol. Biol. *341*, 991-998.

Finger, C., Escher, C., and Schneider, D. (2009). The single transmembrane domain of human tyrosine kinases encode self interactions. Science Signal. *2*, ra56, 51-58.

Ginsberg, M.H., Partridge, A., and Shattil, S.J. (2005). Integrin regulation. Curr. Opin. Cell Biol. *17*, 509-516.

Gottschalk, K.E. (2005). A coiled-coil structure of the alpha llb beta 3 integrin transmembrane and cytoplasmic domains in its resting state. Structure *13*, 703-712.

Gottschalk, K.E., and Kessler, H. (2004). Evidence for hetero-association of transmembrane helices of integrins. FEBS Lett. *557*, 253-258.

Hall, B.A., Chetwynd, A.P., and Sansom, M.S.P. (2011). Exploring peptide-membrane interactions with coarse-grained MD simulations. Biophys. J. *100*, 1940-1948.

Harburger, D.S., and Calderwood, D.A. (2009). Integrin signalling at a glance. J. Cell Sci. *122*, 159-163.

Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. (1997). LINCS: A linear constraint solver for molecular simulations. J. Comp. Chem. *18*, 1463-1472.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD - Visual Molecular Dynamics. J. Molec. Graph. *14*, 33-38.

Hynes, R.O. (2002). Integrins: bidirectional, allosteric signaling machines. Cell *110*, 673-687.

Janosi, L., Prakash, A., and Doxastakis, M. (2010). Lipid-modulated sequence-specific association of glycophorin A in membranes. Biophys. J. *99*, 284-292.

Javadpour, M.M., Eilers, M., Groesbeek, M., and Smith, S.O. (1999). Helix packing in polytopic membrane proteins: Role of glycine in transmembrane helix association. Biophys. J. *77*, 1609-1618.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X.W., and Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293-1307.

Kalli, A.C., Campbell, I.D., and Sansom, M.S.P. (2011). Multiscale Simulations Suggest a Mechanism for Integrin Inside-out Activation. Proceedings of the National Academy of Sciences *In press*.

Khalili-Araghi, F., Gumbart, J., Wen, P.C., Sotomayor, M., Tajkhorshid, E., and Schulten, K. (2009). Molecular dynamics simulations of membrane channels and transporters. Curr. Opin. Struct. Biol. *19*, 128-137.

Langosch, D., and Arkin, I.T. (2009). Interaction and conformational dynamics of membrane-spanning protein helices. Prot. Sci. *18*, 1343-1358.

Lau, T.-L., Kim, C., Ginsberg, M.H., and Ulmer, T.S. (2009). The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling. EMBO J. *28*, 1351-1361.

Law, R.J., Capener, C., Baaden, M., Bond, P.J., Campbell, J., Patargias, G., Arinaminpathy, Y., and Sansom, M.S.P. (2005). Membrane protein structure quality in molecular dynamics simulation. J. Mol. Graph. Mod. *24*, 157-165

Li, E., and Hristova, K. (2006). Role of receptor tyrosine kinase transmembrane domains in cell signaling and human pathologies. Biochem. *45*, 6241-6251.

Li, R.H., Gorelik, R., Nanda, V., Law, P.B., Lear, J.D., DeGrado, W.F., and Bennett, J.S. (2004). Dimerization of the transmembrane domain of integrin alpha(IIb) subunit in cell membranes. Journal of Biological Chemistry *279*, 26666-26673.

Li, W., Metcalf, D.G., Gorelik, R., Li, R., Mitra, N., Nanda, V., Law, P.B., Lear, J.D., DeGrado, W.F., and Bennett, J.S. (2005). A push-pull mechanism for regulating integrin function. Proc. Natl. Acad. Sci. USA *102*, 1424-1429.

Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Molec. Model. *7*, 306-317.

Lindahl, E., and Sansom, M.S.P. (2008). Membrane proteins: molecular dynamics simulations. Curr. Opin. Struct. Biol. *18*, 425-431.

Luo, B.-H., Carman, C.V., Takagi, J., and Springer, T.A. (2005). Disrupting integrin transmembrane domain heterodimerization increases ligand binding affinity, not valency or clustering. Proc. Natl. Acad. Sci. USA *102*, 3679-3684.

Luo, B.-H., Springer, T.A., and Takagi, J. (2004). A specific interface between integrin transmembrane helices and affinity for ligand. PLoS Biol. *2*, e153.

MacKenzie, K.R., Prestegard, J.H., and Engelman, D.M. (1997). A transmembrane helix dimer: structure and implications. Science *276*, 131-133.

Marks, F., Klingmüller, U., and Müller-Decker, K. (2009). Cellular Signal Processing (Garland Science).

Metcalf, D.G., Kulp, D.W., Bennett, J.S., and DeGrado, W.F. (2009). Multiple approaches converge on the structure of the integrin αIIb/β3 transmembrane heterodimer. J. Mol. Biol. *392*, 1087-1101.

Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.J. (2008). The MARTINI coarse grained force field: extension to proteins. J. Chem. Theor. Comp. *4*, 819-834.

Petrache, H.I., Grossfield, A., MacKenzie, K.R., Engelman, D.M., and Woolf, T.B. (2000). Modulation of glycophorin A transmembrane helix interactions by lipid bilayers: Molecular dynamics calculations. J. Mol. Biol. *302*, 727-746.

Popot, J.L., and Engelman, D.M. (2000). Helical membrane protein folding, stability, and evolution. Ann. Rev. Biochem. *69*, 881-922.

Psachoulia, E., Bond, P.J., Fowler, P.W., and Sansom, M.S.P. (2008). Helix-helix interactions in membrane proteins: coarse grained simulations of glycophorin helix dimerization. Biochem. *47*, 10503-105012.

Psachoulia, E., Marshall, D., and Sansom, M.S.P. (2010). Molecular dynamics simulations of the dimerization of transmembrane α-helices. Acc. Chem. Res. *43*, 388–396.

Raschle, T., Hiller, S., Etzkorn, M., and Wagner, G. (2010). Nonmicellar systems for solution NMR spectroscopy of membrane proteins. Curr. Opin. Struct. Biol. *20*, 471-479.

Russ, W.P., and Engelman, D.M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. Proc. Natl. Acad. Sci. USA *96*, 863-868.

Russ, W.P., and Engelman, D.M. (2000). The GxxxG motif: a framework for transmembrane helix-helix association. J. Mol. Biol. *296*, 911-919.

Schneider, D., and Engelman, D.M. (2004). Involvement of transmembrane domain interactions in signal transduction by alpha/beta integrins. J. Biol. Chem. *279*, 9840-9846.

Scott, W.R.P., Hunenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, T., Kruger, P., and van Gunsteren, W.F. (1999). The GROMOS biomolecular simulation program package. J. Phys. Chem. A *103*, 3596-3607.

Senes, A., Engel, D.E., and DeGrado, W.F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. Curr. Opin. Struct. Biol. *14*, 465-479.

Sengupta, D., and Marrink, S.J. (2010). Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. Phys. Chem. Chem. Phys. *12*, 12987-12996.

Stansfeld, P.J., and Sansom, M.S.P. (2011). From coarse-grained to atomistic: a serial multi-scale approach to membrane protein simulations. J. Chem. Theor. Comp. *(in press)*.

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J. Comput. Chem. *26*, 1701-1718.

Vararattanavech, A., Chng, C.-P., Parthasarathy, K., Tang, X.-Y., Torres, J., and Tan, S.-M. (2010). A transmembrane polar interaction is involved in the functional regulation of integrin αLβ2. J. Mol. Biol. *398*, 569-583.

Vereshaga, Y.A., Volynsky, P.E., Nolde, D.E., Arseniev, A.S., and Efremov, R.G. (2005). Helix interactions in membranes: Lessons from unrestrained Monte Carlo simulations. J. Chem. Theory Comput. *1*, 1252-1264.

Volynsky, P.E., Mineeva, E.A., Goncharuk, M.V., Ermolyuk, Y.S., Arseniev, A.S., and Efremov, R.G. (2010). Computer simulations and modeling-assisted ToxR screening in deciphering 3D structures of transmembrane alpha-helical dimers: ephrin receptor A1. Physical Biol. *7*, 15.

Wegener, K.L., and Campbell, I.D. (2008). Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (Review). Mol. Memb. Biol. *25*, 376-387.

Wegener, K.L., Partridge, A.W., Han, J., Pickford, A.R., Liddington, R.C., Ginsberg, M.H., and Campbell, I.D. (2007). Structural basis of integrin activation by talin. Cell *128*, 171-182.

Yang, J., Ma, Y.-Q., Page, R.C., Misra, S., Plow, E.F., and Qin, J. (2009). Structure of an integrin αIIbβ3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. Proc. Natl. Acad. Sci. USA *106*, 17729-17734.

Zhu, J., Luo, B.H., Barth, P., Schonbrun, J., Baker, D., and Springer, T.A. (2009). The structure of a receptor with two associating transmembrane domains on the cell surface: integrin alpha(IIb)beta(3). Molec. Cell *34*, 234-249.

SUPPLEMENTARY INFORMATION

Figure S1, related to Figure 2 :

A. Sequences used in the simulations. The sequence of the TM helix of the αIIb subunit (residues 964-990) is shown as well as the two extended αIIb sequences (964-993 and 964-995; see main text for details).

B. Spatial distributions of the αIIb helix relative to the β3 helix for the CG-WT simulations using the αIIb TM helix (964-990) and using the extended (i.e. 964-993 and 964-995) αIIb helices. Other details are as for Fig. 2 of the main text.

Figure S2, related to Figure 3:

Comparison of the crossing angle distributions from the AT-RH1 wild type (grey filled histogram) and for the AT simulations of various mutants (coloured lines as labelled). The crossing angle was calculated using the last 15 ns of all simulations.

Figure S3, related to Figure 4:

Comparison of the helix/helix interface of structures from the end of one of the AT-NMR (green) and one of the AT-RH1 (cyan) simulations. The Cα RMSD between the two structures is 2.2 Å. The G972 & G976 residues of the αIIb helix and the M701 & I704 residues of the β3 helix are indicated.

Figure S4, related to Figure 2:

Convergence analysis. The crossing angle distribution (A) and the spatial distribution of the αIIb helix (B) form the CG-WT simulations is shown by using 25%, 50%, 75% and 100% of the simulations. The systems used for the 25%, 50% and 75% subsets were chosen randomly. This analysis suggests that the simulation converges after 50%.

Figure S5, related to Figure 2:

A.Example of the motion of the αIIb helix relative to the β3 helix in the plane of the lipid bilayer. The position of the center of mass of the αIIb helix from the CG-WT simulation is shown. The β3 helix is fixed in the xy plane. The coordinates (x,y) of the mobile helix are drawn in red to show the dimerization of the two helices. The starting position of the αIIb helix is shown in green and the final in blue.

B. Spatial distributions of the αIIb helix relative to the β3 helix for the WT sequence with the β3 helix successively rotated by 15^{o} at the beginning of each simulation. Again ~100 simulations were performed. The spatial distribution is identical with that from the WT-CG simulation.

Figure S6, related to Figure 5:

Helix/helix interface from a structure from the end of one of A the AT-RH2 and **B** the AT-LH simulations. The G972,G976 surface (cyan) of the αIIb helix (blue) can be seen to pack away from the M701 (yellow) and I704 (green) sidechains of the β3 helix (red). The G708 surface is shown in cyan. These key interactions are seen in both the NMR and the AT-RH1 simulation structure.

Figure S7, related to Figure 3:

Conformational stability of the α/β helix dimer from an AT-RH1 simulation using the OPLS forcefield as analysed in terms of the Cα root mean square deviation (RMSD) from the initial structure as a function of time. The RMSD was calculated using the αIIb 966-988 and β 694-720 residues.

Figure 1:

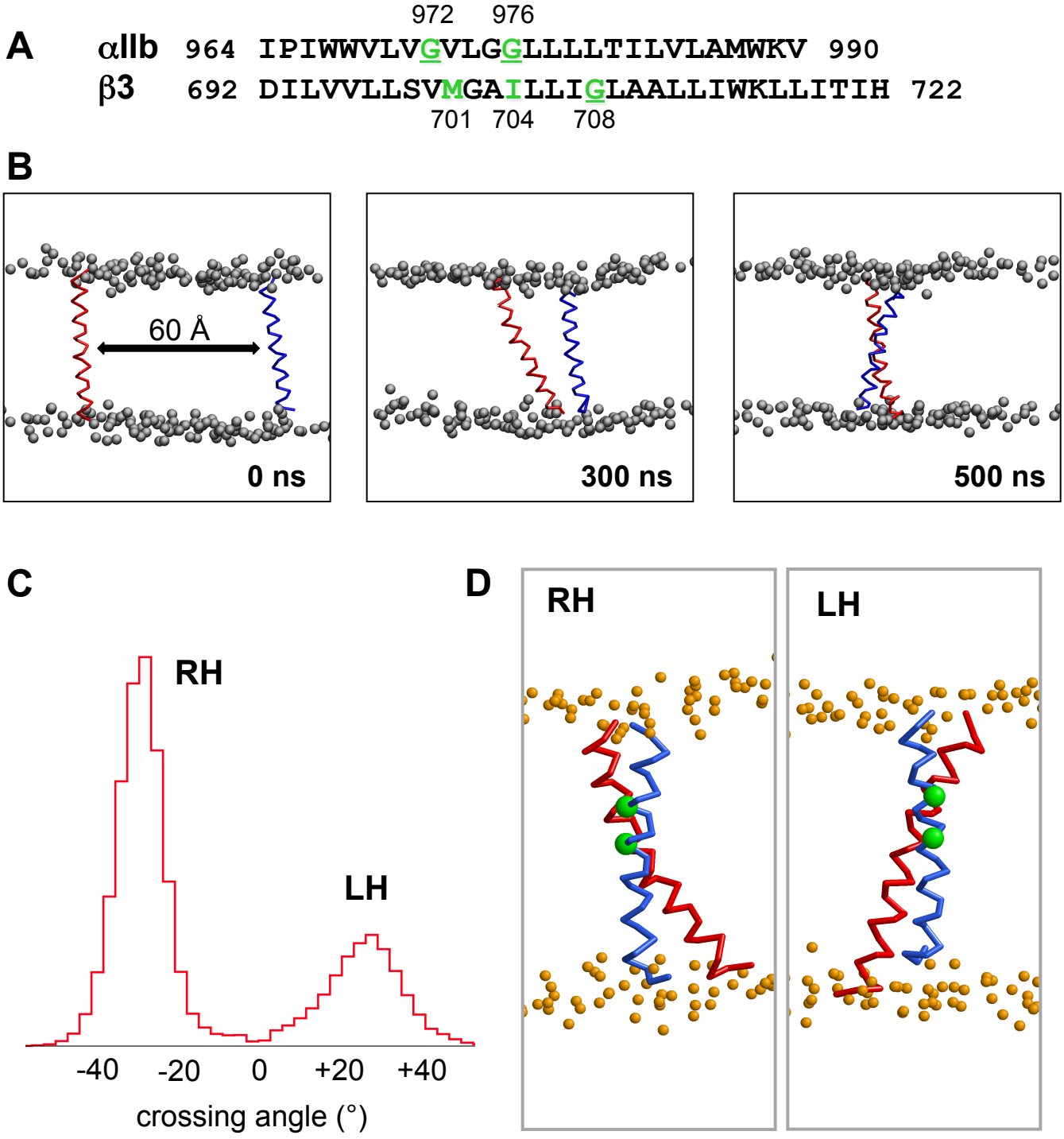
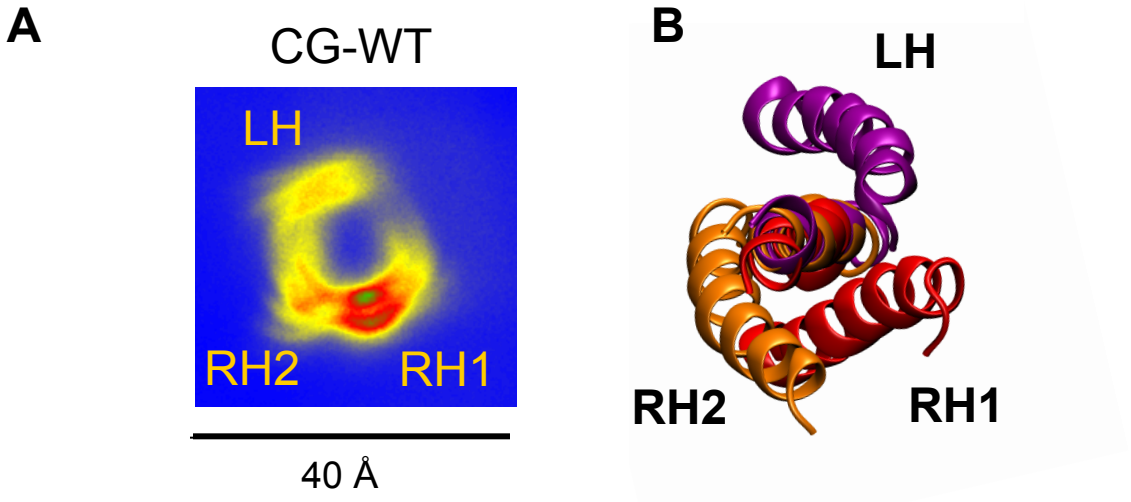

A
αIIb 964 IPIWWVLVGVLGGLLLLTILVLAMWKV 990
β3 692 DILVVLLSVMGAILLIGLAALLIWKLLITIH 722
972 976
701 704 708
B
60 Å
0 ns
300 ns
500 ns
C
RH
LH
-40 -20 0 +20 +40
crossing angle (°)
D
RH
LH

Figure 2:


A
CG-WT
LH
RH2
RH1
40 Å
B
LH
RH2
RH1

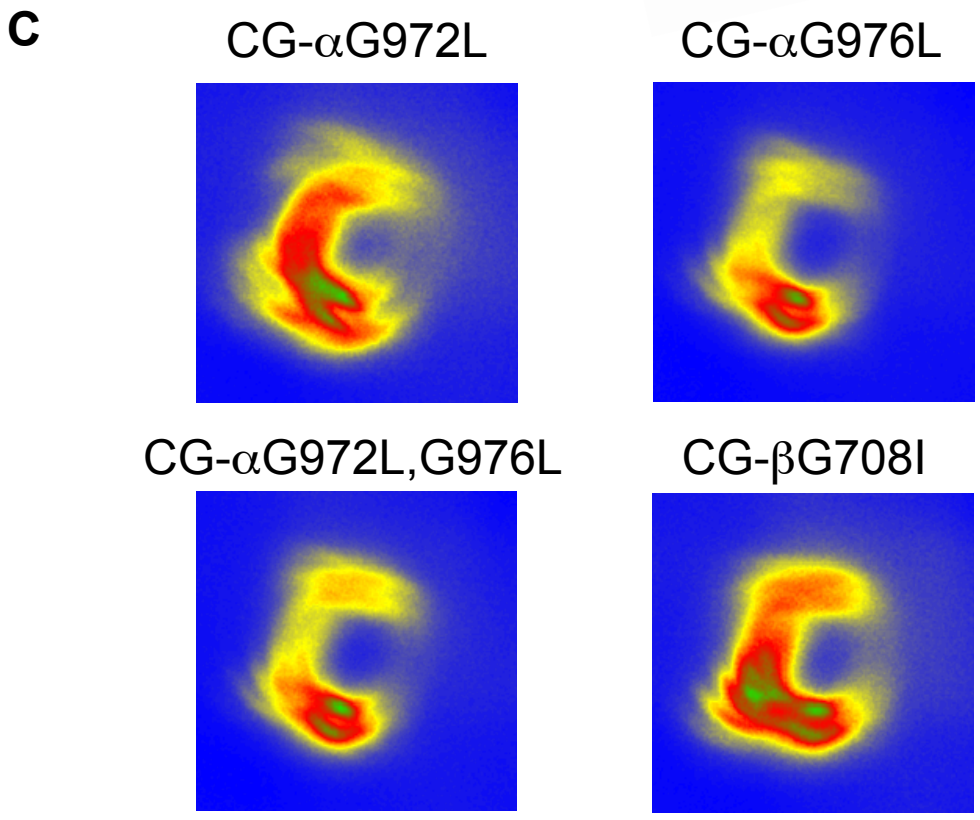

C
CG-αG972L
CG-αG976L
CG-αG972L,G976L
CG-βG708I

Figure 3:

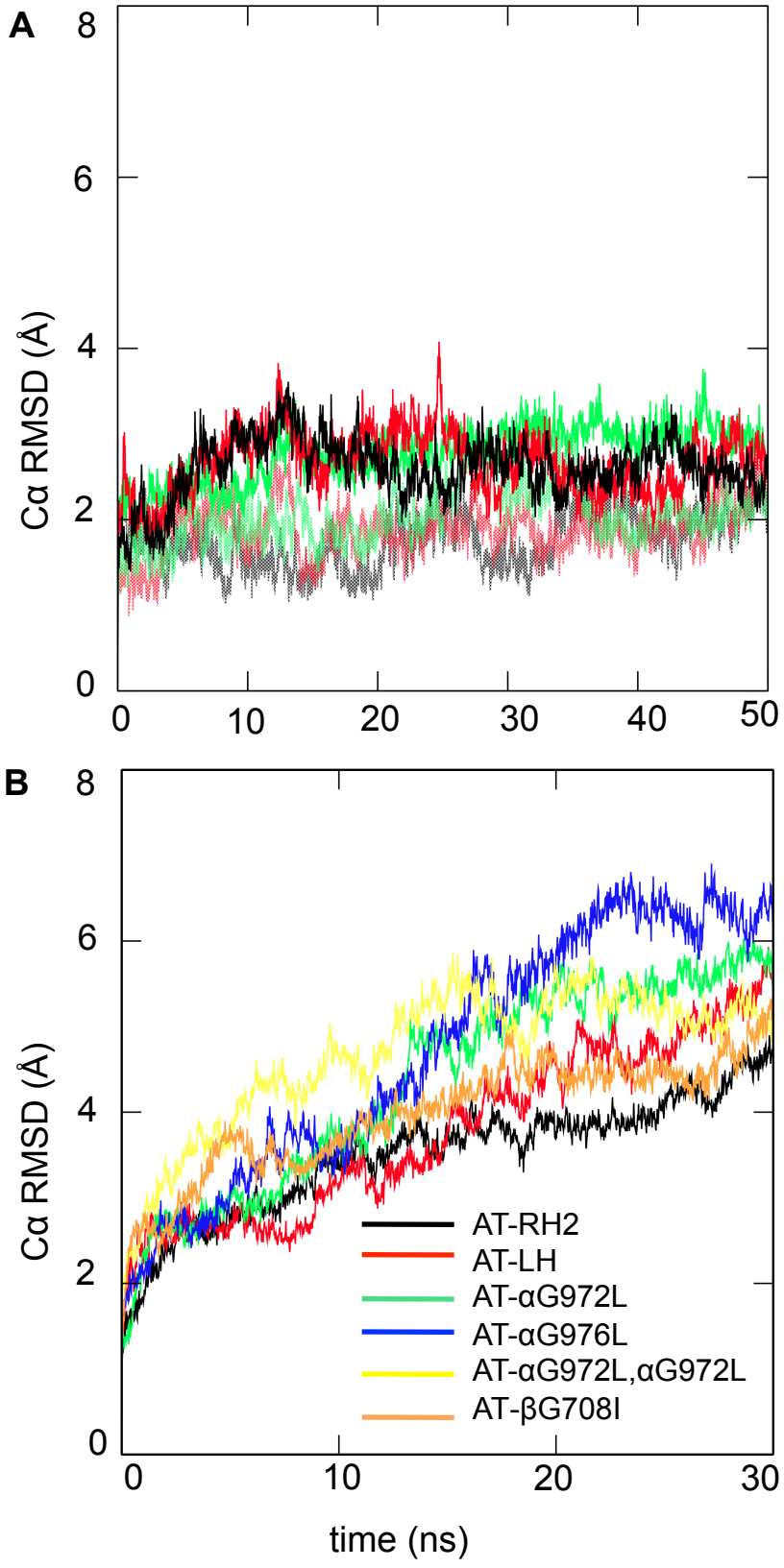

A
B
Cα RMSD (Å)
time (ns)
AT-RH2
AT-LH
AT-αG972L
AT-αG976L
AT-αG972L,αG972L
AT-βG708I

Figure 4:

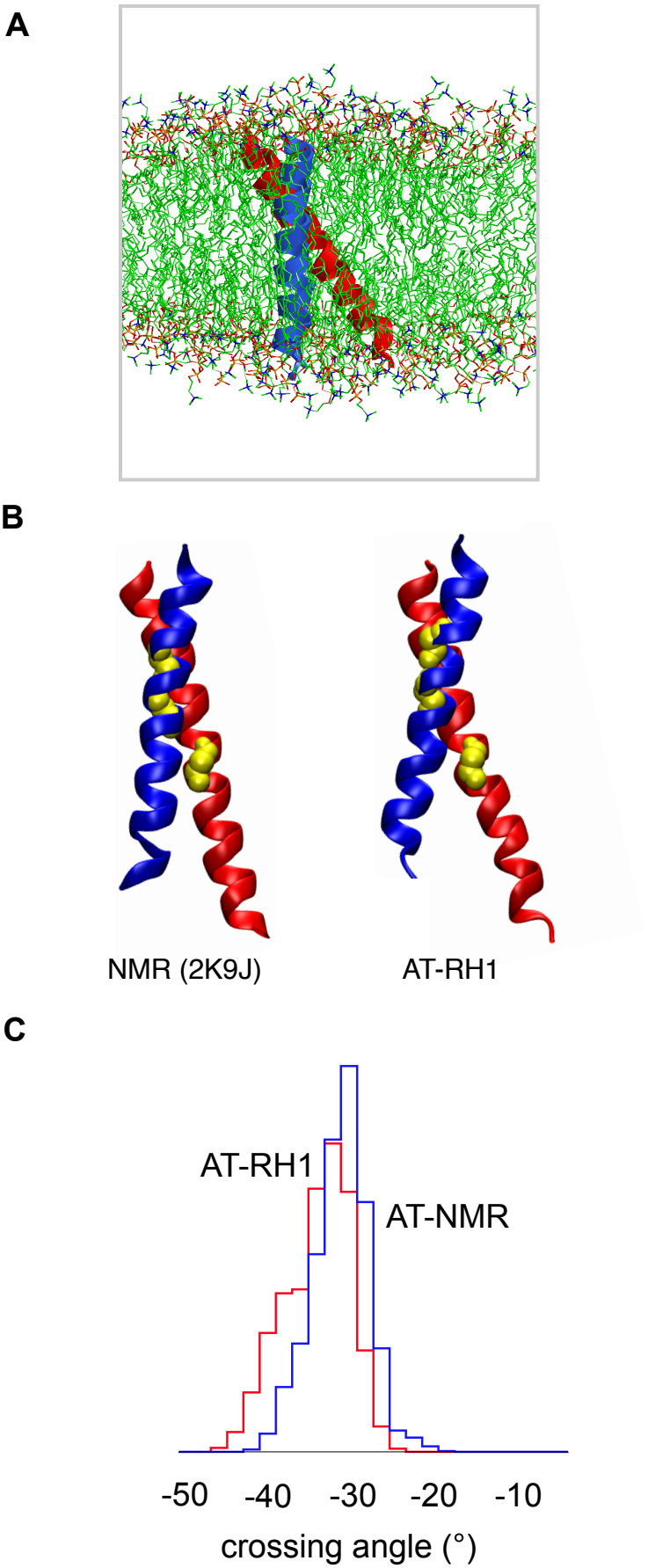


Figure 5

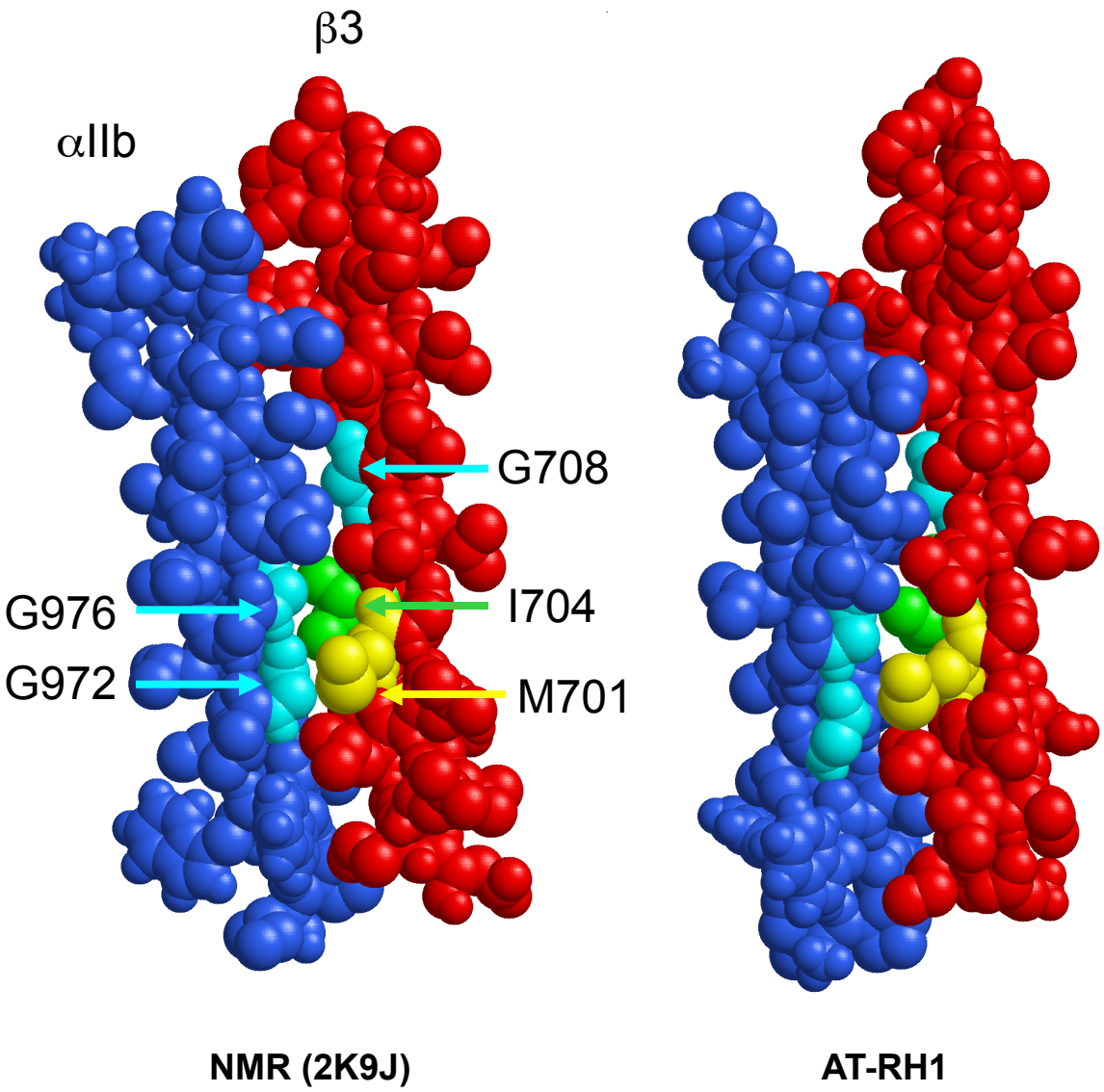


Figure S1:

A

αIIb	964	IPIWWVLVGVLGGLLLLTILVLAMWKV	990
	964	IPIWWVLVGVLGGLLLLTILVLAMWKVGFF	993
	964	IPIWWVLVGVLGGLLLLTILVLAMWKVGFFKR	995
β3	692	DILVVLLSVMGAILLIGLAALLIWKLLITIH	722

B

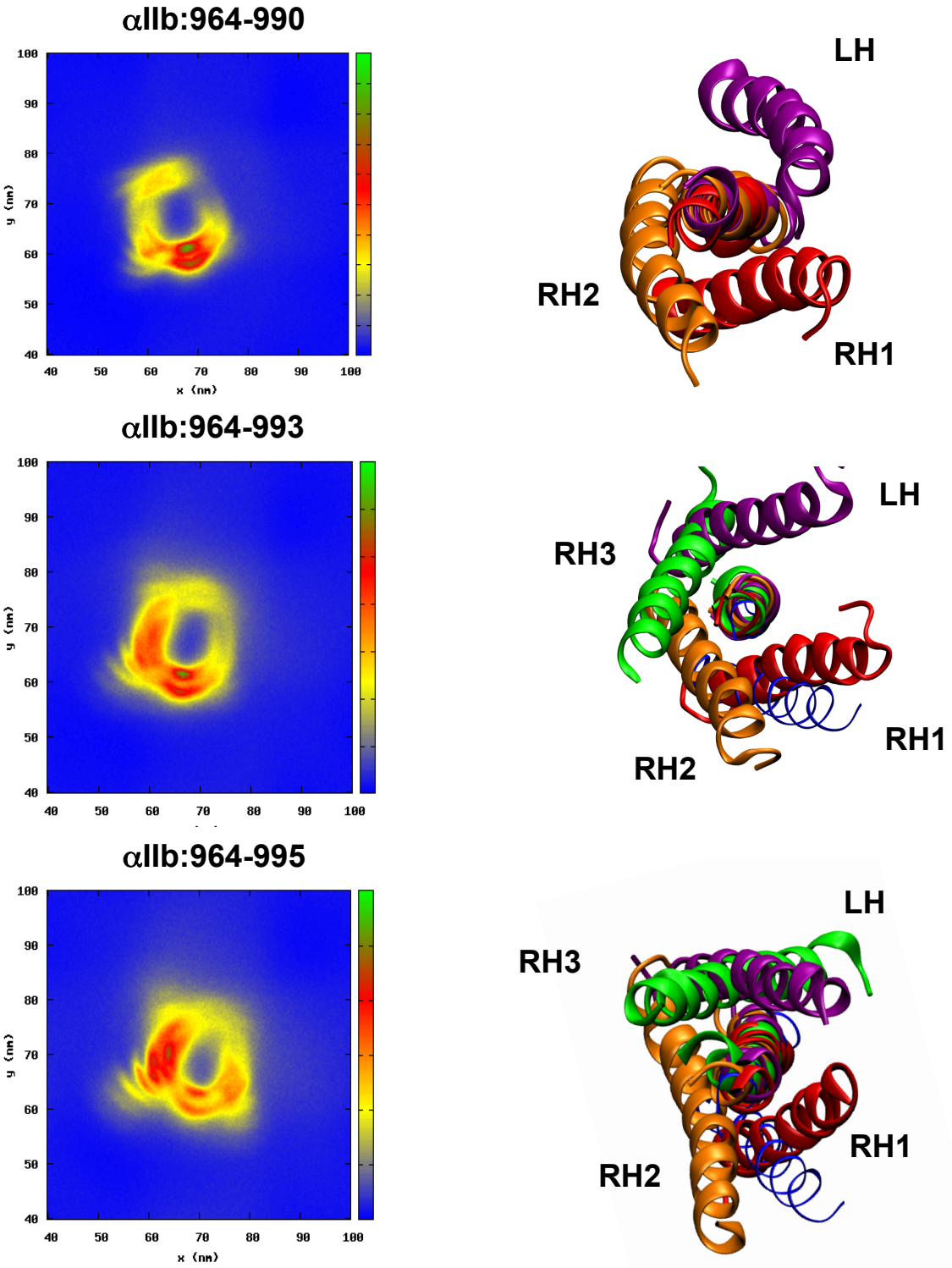


Figure S2:

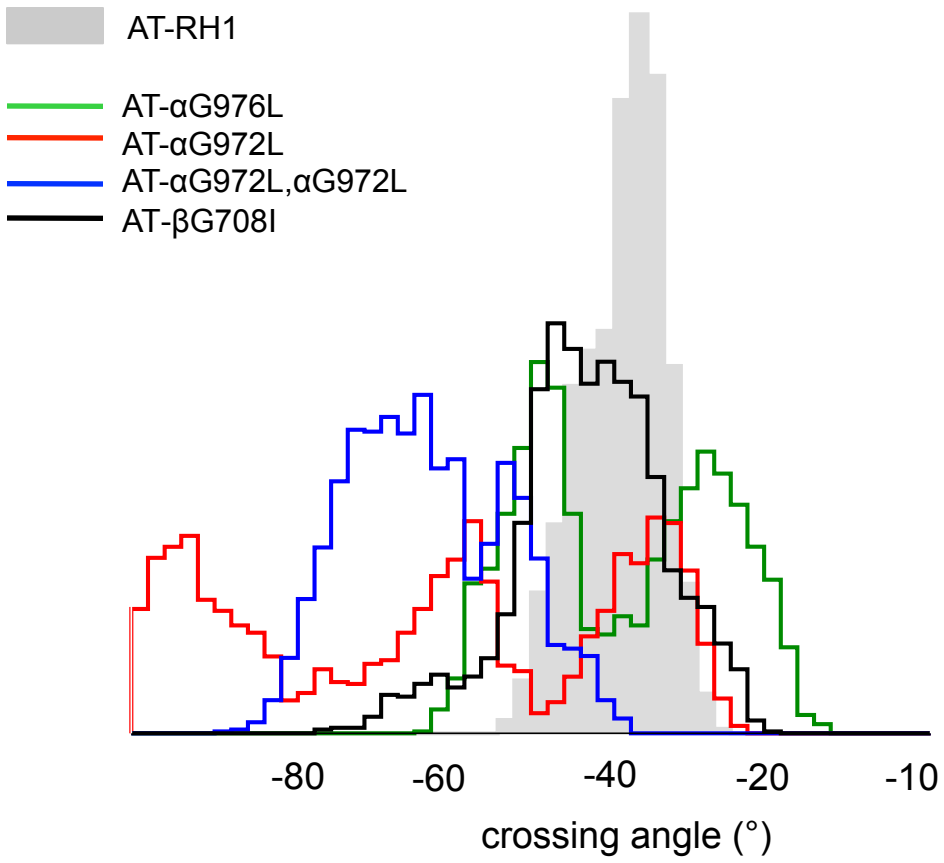

AT-RH1
AT-αG976L
AT-αG972L
AT-αG972L,αG972L
AT-βG708I
-80
-60
-40
-20
-10
crossing angle (°)

Figure S3:

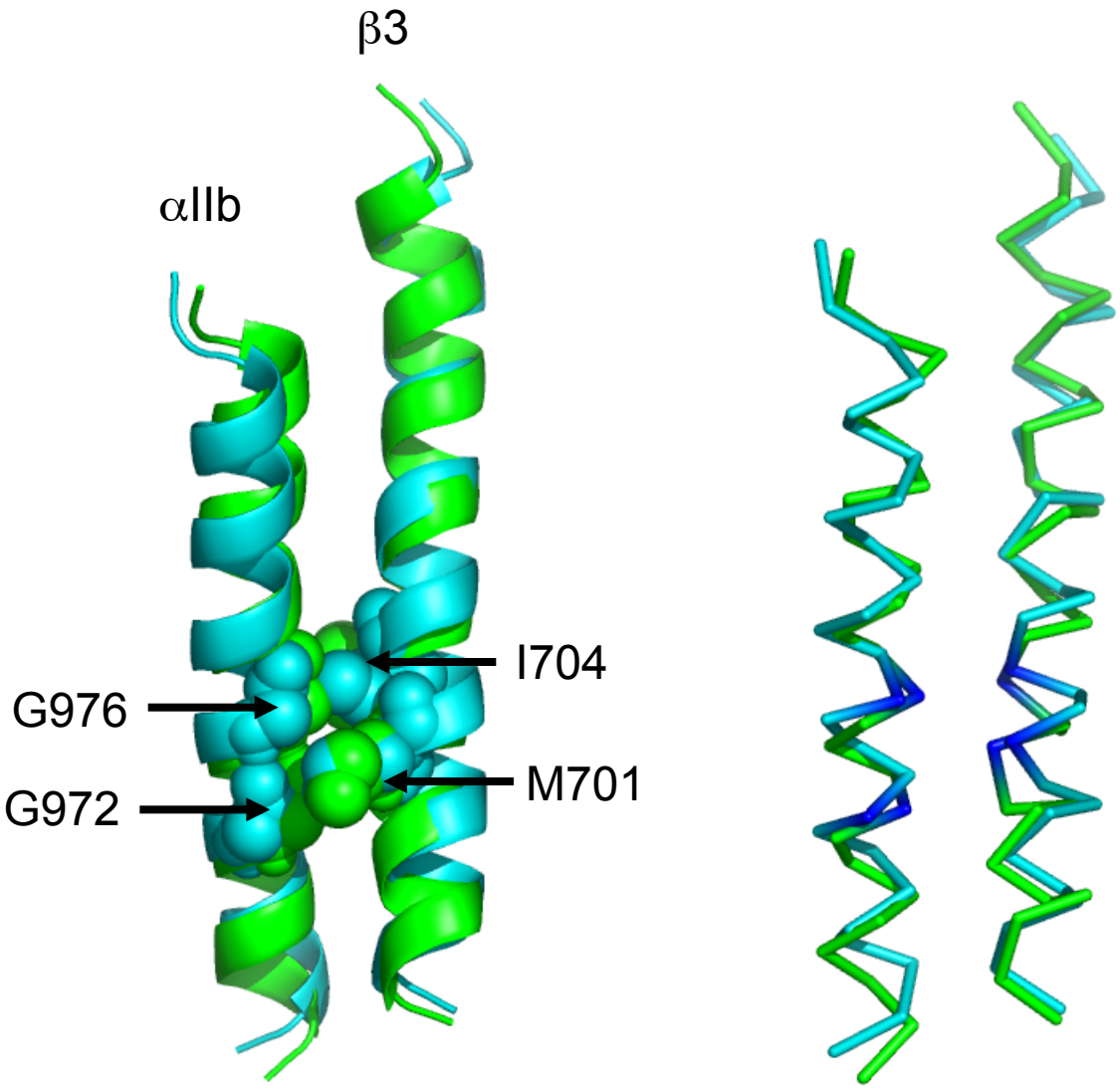
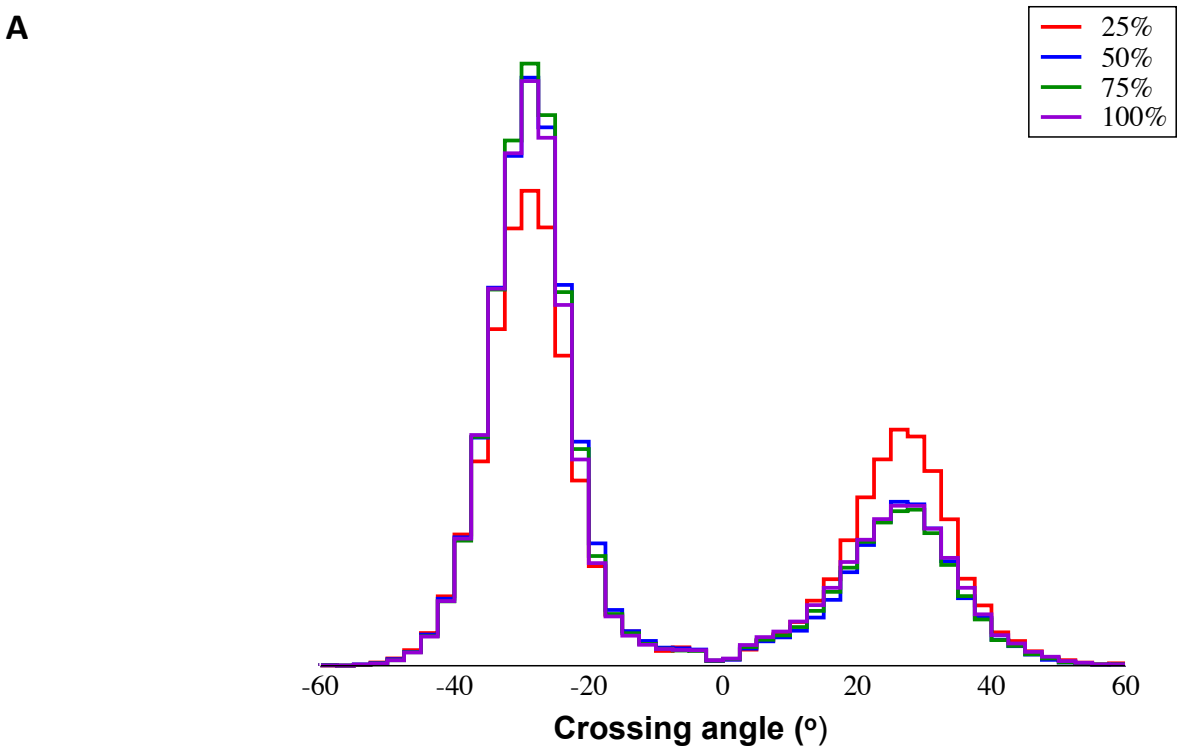


Figure S4:


A
25%
50%
75%
100%
-60
-40
-20
0
20
40
60
Crossing angle (°)

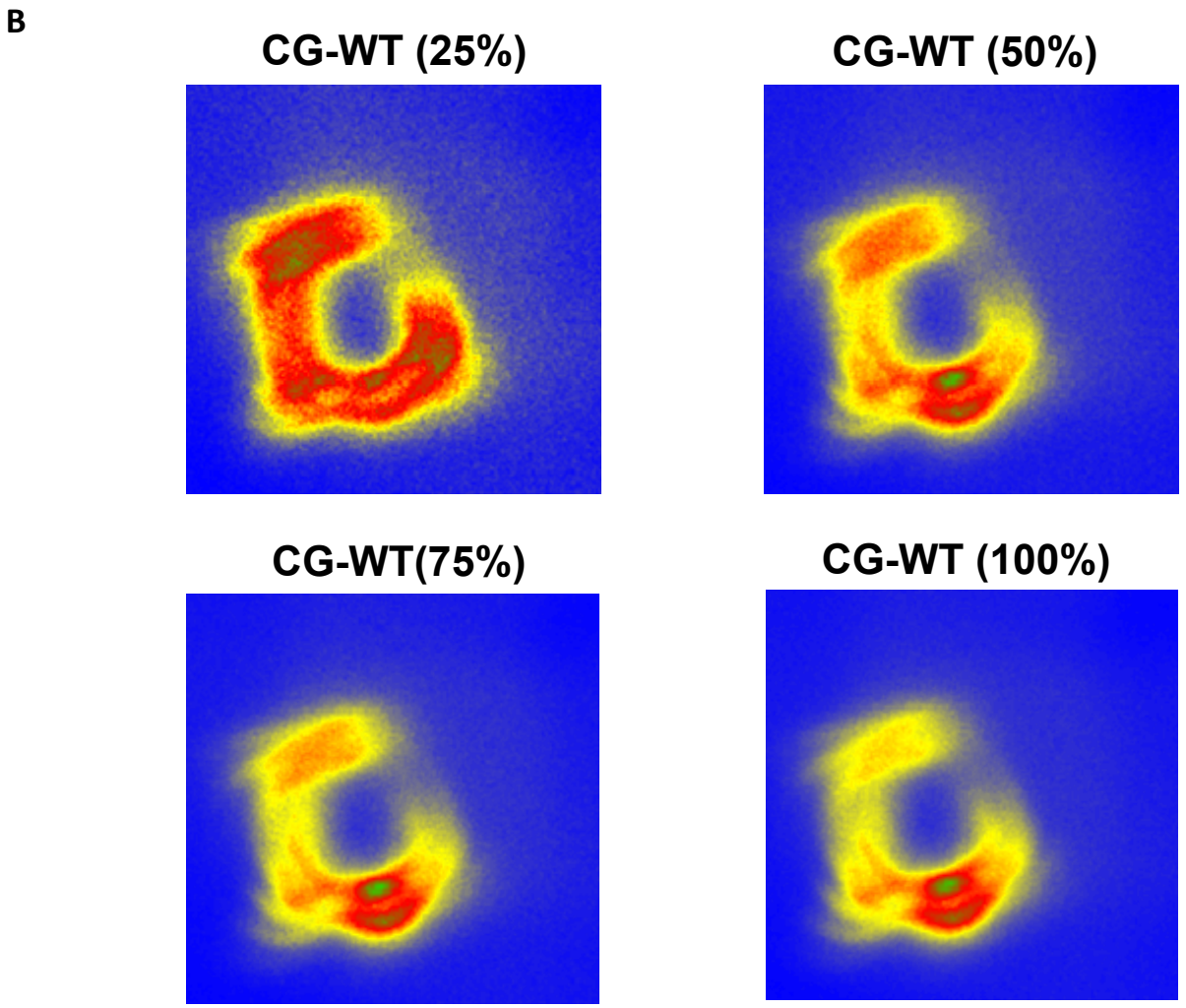
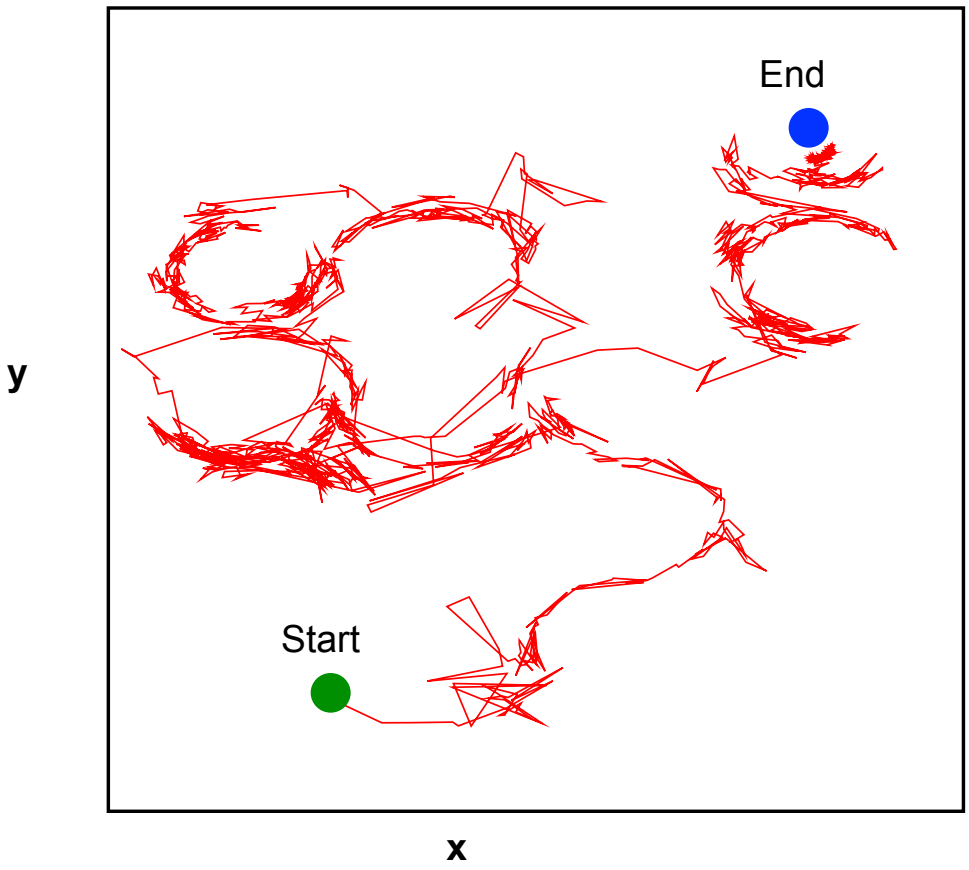

B
CG-WT (25%)
CG-WT (50%)
CG-WT(75%)
CG-WT (100%)

Figure S5:

A



B

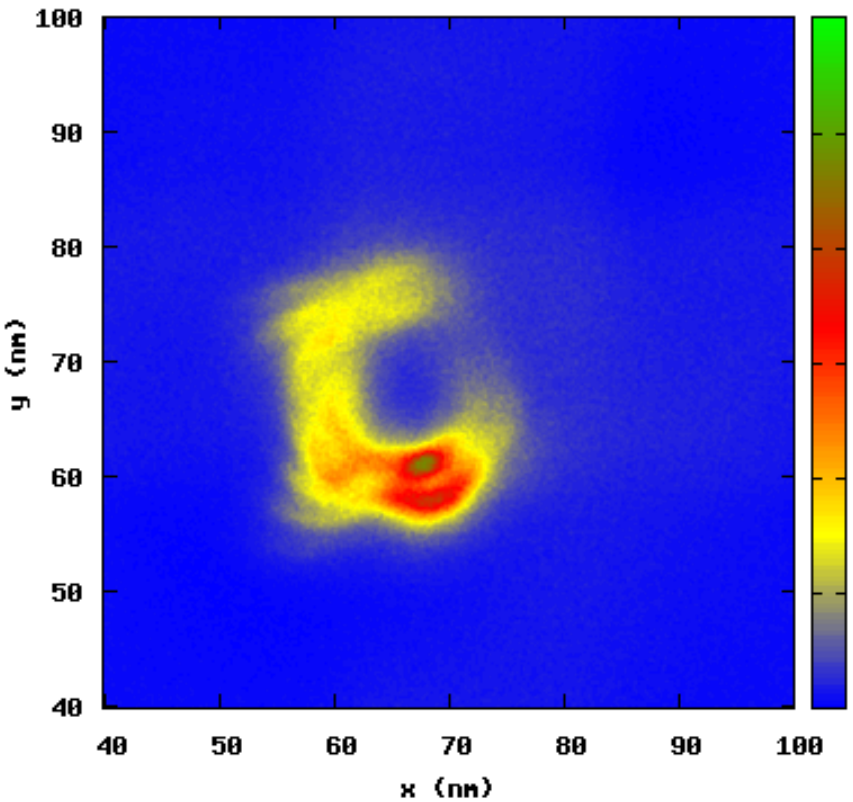


Figure S6:

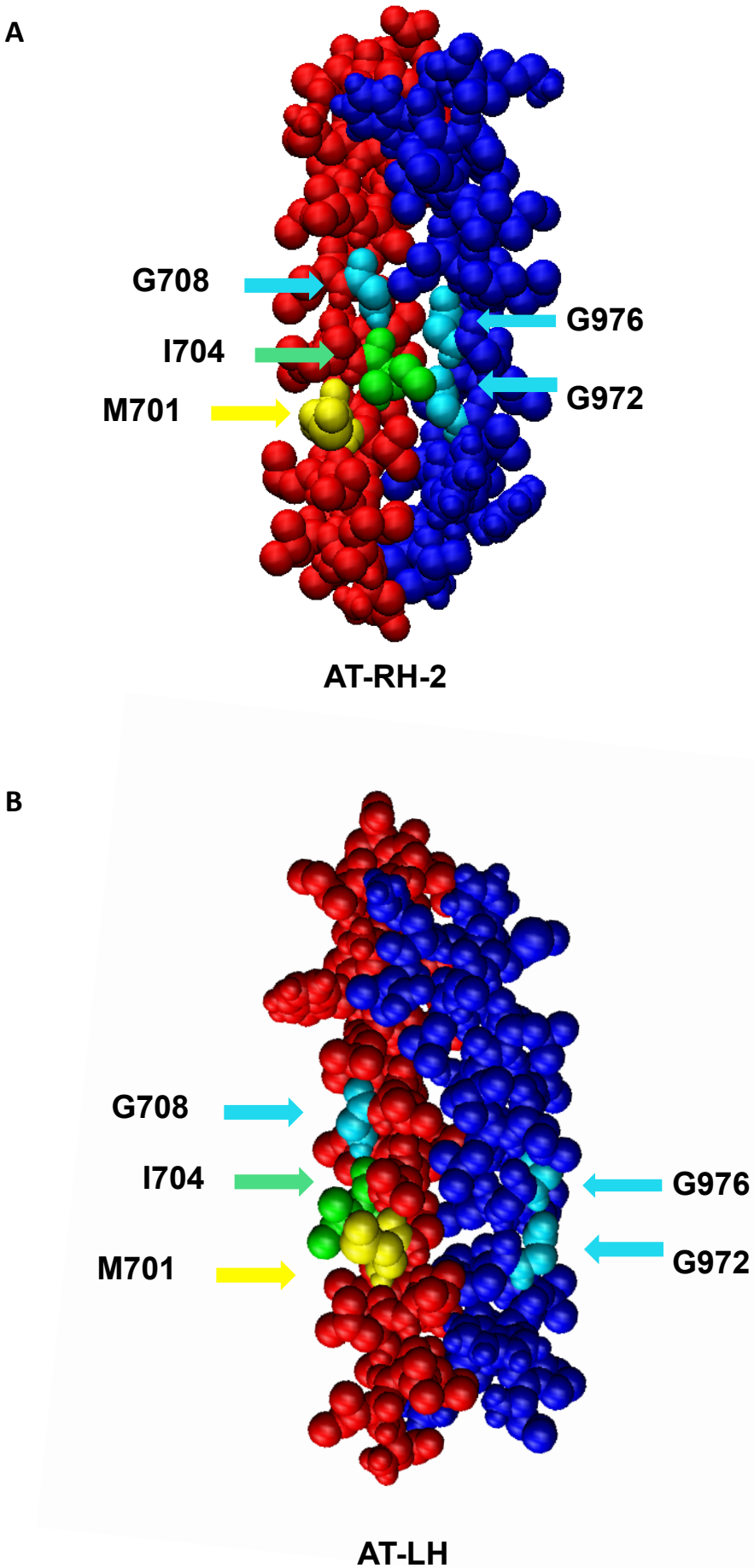
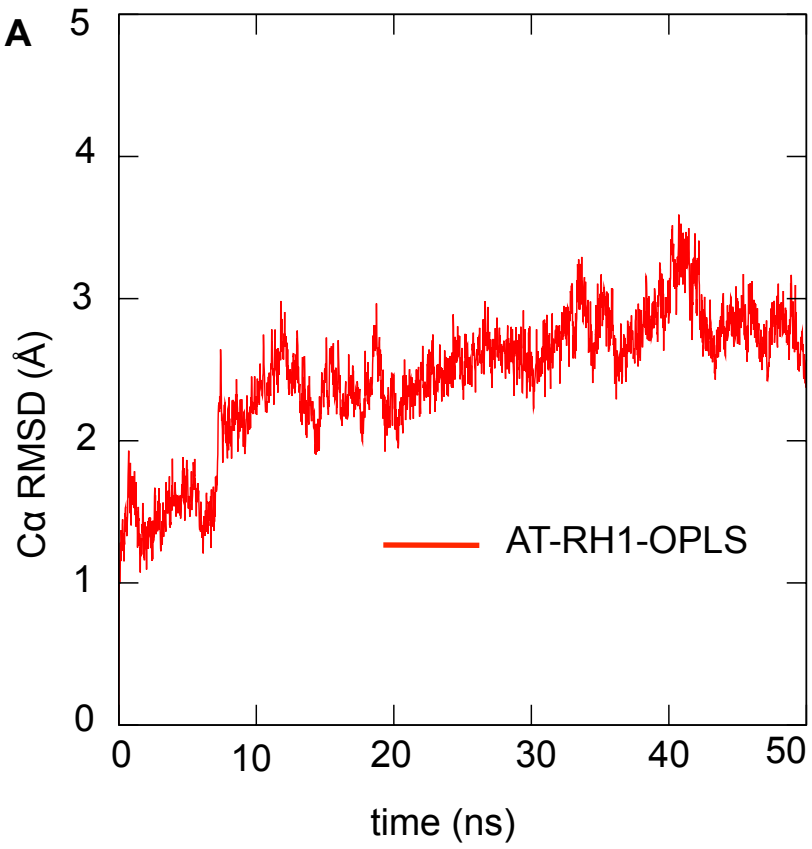

A
G708
I704
M701
G976
G972
AT-RH-2
B
G708
I704
M701
G976
G972
AT-LH

Figure S7:



Multiscale Simulations Suggest a Mechanism for Integrin Inside-out Activation

*Antreas C. Kalli, Iain D. Campbell & Mark S.P. Sansom**

Department of Biochemistry,
University of Oxford,
South Parks Road,
Oxford,
OX1 3QU,
U.K.

*To whom communications should be addressed:
e-mail: mark.sansom@bioch.ox.ac.uk
phone: +44 1865 613306

This paper was published in PNAS. See http://www.pnas.org/content/108/29/11890.abstract

ABSTRACT

Integrins are large cell-surface adhesion receptors that can be activated to a high affinity state by the formation of an intracellular complex between the integrin β-subunit tail, the membrane and talin. The F2 and F3 sub-domains of the talin head play a key role in formation of this complex. Here, activation of the integrin αIIb/β3 dimer by the talin head domain was probed using multiscale molecular dynamics (MD) simulations. A number of novel insights emerge from these studies including: i) the importance of the integrin αIIb subunit F992 and F993 residues in stabilizing the 'off' state of the αIIb/β3 dimer; ii) a crucial role for negatively charged groups in the F2-F3/membrane interaction; iii) binding of the talin F2-F3 domain to negatively charged lipid headgroups in the membrane induces a reorientation of the β TM domain; iv) an increase in the tilt angle of the β TM domain relative to the bilayer normal helps to destabilize the α/β TM interaction and promote a scissor-like movement of the integrin TM helices. These results, combined with various published experimental observations, suggest a model for the mechanism of inside-out activation of integrins by talin.

INTRODUCTION

Integrins are heterodimeric (αβ) cell surface receptors; each subunit contains a large extracellular (ecto) domain, an α-helical transmembrane (TM) domain, and a cytoplasmic domain (1). Integrins are crucial for signal transduction events involved in cell adhesion,

migration and differentiation (2-5). In mammals, combinations of the 18 α and 8 β subunits may form at least 24 different integrins. Transmembrane signaling can occur via outside-in (6) and inside-out pathways. In inside-out activation, the cytoplasmic tail of the integrin β subunit forms a complex with a cytoplasmic protein, talin, at the membrane surface (7-14).

Talin consists of a flexible rod and a globular head that has four sub-domains (F0 to F3) (15, 16). The F3 sub-domain alone is sufficient for αIIbβ3 integrin activation (17) but interactions between a positively charged patch in the F2 surface (14, 18) and negatively charged groups in the membrane are also crucial for integrin activation and clustering (19). Structural and functional studies suggest that the F0-F1 sub-domain also contributes to the activating membrane complex (16, 20) partly via a flexible loop in the F1 domain. PIP_2 is also believed to play a role (19).

The inactive state of integrins is maintained by inter-subunit associations in the transmembrane (TM) (21-26) and cytoplasmic regions (27-29) of the two integrin subunits. Integrin activation is believed to involve dissociation or destabilization of the two integrin TM helices (29-33), followed by rearrangement of the ectodomain to an extended conformation (34-37). Recent NMR and modeling studies (24) revealed similar structures for the αIIb and β3 TM helices in a membrane-like environment. Two main interaction surfaces between the two TM helices of the αIIb/β3 complex were identified in the NMR structure. These were called an *outer membrane clasp* (OMC), in which close packing of the helices is facilitated by a Gx_3G motif in αIIb, and an *inner membrane clasp* (IMC), which involves interactions of F992 and F993 in αIIb with the β3 tail together with a salt-bridge between αIIbR995 and β3D723 (24).

Different models have been proposed for structural changes in the membrane region during activation, including 'piston' and 'scissors' movements of the TM helices (38). However, more recent studies (7, 23, 32, 39-41) have been interpreted in terms of a model in which the binding of talin to the integrin β cytoplasmic tail and the TM domain disrupts the OMC and IMC interactions between the two TM helices, leading to helix separation and subsequent switching of the integrins to a high affinity state (14). Resolution of the underlying mechanism of activation thus requires information about the *conformational dynamics* of integrin/talin/membrane interactions that is not readily available from experimental methods.

Molecular dynamics (MD) simulations (42) can play a key role in understanding the dynamic interactions of membrane proteins with their environment (43). They have been used to investigate the interactions of transmembrane helices with one another in a bilayer

environment (44) as well as the interactions of peripheral proteins with membranes (45-48). MD studies of the F2-F3 domain of talin (18) in a bilayer demonstrated a key role for electrostatic interactions between the protein and anionic lipid headgroups.

Here, we use a multi-scale MD simulation approach (49, 50) to explore the conformational dynamics of the interaction of the integrin TM helix dimer with talin F2-F3 and a lipid bilayer. The aim was to understand the inside-out activation mechanism. On the basis of these studies we suggest how talin can weaken the α/β TM association and promote an integrin conformation that corresponds to a high affinity state.

RESULTS AND DISCUSSION

Conformation and Dynamics of the Integrin αIIb/β3 Transmembrane Helix Dimer

As a first step towards understanding how talin interactions may modulate the behavior of the integrin αIIb/β3 transmembrane (TM) helix dimer, we simulated the conformational dynamics of the TM helix dimer. We started with the coordinates of the NMR structure (PDB 2K9J) which is probably a reasonable approximation of the low-affinity integrin 'off' state. CG-MD simulations were used to optimize the position of the TM helix dimer relative to a POPC/POPG bilayer, whilst restraining the αIIb TM helix to be perpendicular to the bilayer. This orientation of αIIb is suggested by the NMR structures of the αIIb helix (51) and the αIIb/β3 dimer in bicelles (24). The resultant CG system (i.e. helix dimer plus bilayer) was converted to atomistic resolution (50), to form the starting point for 3x100 ns simulations.

In these **αβ1-AT** (Table 1) simulations, the TM dimer remained stable (Fig. 1A). Both IMC interactions (i.e. the αIIbR995-β3D723R salt-bridge and the αIIb(F992,F993)-β3 aromatic cluster interactions) and the OMC interaction, involving the αIIb Gx_3G motif, were maintained throughout all three simulations. The tilt angles for the αIIb and the β3 helices during the αβ1-AT simulations were ~5° and ~30°, respectively, and the helix dimer maintained its right-handed packing with a helix crossing angle of ~-35° (Fig. 1C).

Effect of Mutations on the Stability of the Integrin αIIb/β3 Transmembrane Helix Dimer

The two components of the IMC (i.e. the salt bridge and the aromatic cluster) were further explored by simulations of mutants. Thus, in simulation **αβ2-AT** (Table 1), the αIIbR995D mutation disrupted the salt bridge but did not appear to greatly alter the conformational stability of the dimer, as the hydrophobic interactions of the aromatic cluster of the IMC remained intact, and helix packing in the OMC was not perturbed. Thus, the αIIbR995D mutant dimer *retained* its right-handed helix packing, with a crossing angle of ~-35° (Fig.

1C), i.e. similar to for the WT helix dimer. This is in good agreement with experimental results which suggest that disruption of the αIIb R995 and β3 D723 salt bridge weakens, but does not eliminate, the α/β association (24, 52).

The contribution of the aromatic cluster of the IMC to the conformational stability of the αIIb/β3 dimer was also explored by mutation of αIIb (F992A, F993A) in simulation **αβ3-AT**; Table 1. This double mutation significantly perturbed the packing of the helices, reducing their crossing angle to ~-10° (Fig. 1C). The β3 tilt angle relative to the bilayer normal remained at ~30°, i.e. similar to the WT, but the tilt angle of the αIIb helix increased to ~20° (cf. ~5° in the WT). This increase in the tilt angle reflects the change in the packing of the two helices. This change is such that the β3 residues M701 and L705 in the OMC and residues L712 and K716 in the IMC, which interact with the αIIb helix in the WT, no longer form these contacts, resulting in looser packing of the two helices. This is in agreement with mutational data that suggest that substitution of either F992A or F993A activates integrins (30).

In simulation **αβ4-AT**, both components of the IMC were removed by mutation (as in **αβ3-AT**) of the aromatic residues and deletion of the salt-bridging arginine of the α subunit.. In all three **αβ4-AT** simulations the distance between the residue 992, 992 region of αIIb and the β3 helix increased. This altered the packing of the two helices, resulting in a wider distribution of crossing angles. However, disruption of the packing of the OMC clasp was observed only in one out of the three **αβ4-AT** simulations. This suggests that release of the IMC alone may not be sufficient to disrupt all the αIIb/β3 TM interactions.

The OMC is formed by a glycophorin-like interaction (53), involving a Gx_3G motif formed by residues G972 to G976 of the αIIb helix. Mutation of either G972 to serine (**αβ5-AT**; Table 1) or of G976 to leucine (**αβ6-AT**) has been reported to induce integrin activation (24, 54) but neither mutation showed significant perturbation of the dimer in our 100 ns atomistic simulations. However, multi-scale self-assembly simulations of αIIb/β3 TM helix interactions, using an approach previously tested for glycophorin (44, 55), suggest that mutations of the glycine residues in the OMC do indeed perturb helix packing.

Interactions with Talin Perturb the Integrin Transmembrane Helices

We next explored how interactions with the F2-F3 domains of the talin head modulate the TM helices in a lipid bilayer. If signaling from inside to outside is indeed mediated by a change in packing interactions of the TM helices then one would anticipate that F2-F3 bilayer interactions (18) could perturb TM/TM and/or TM/bilayer interactions in the simulations. A

short (30 ns) atomistic simulation of F2-F3 bound to the PC/PG bilayer, revealed a small relative movement (~6°) of the F2 sub-domain relative to F3. This movement optimizes contacts with the lipids and is made possible by a flexible linker between the two sub-domains. More detailed analysis of F2-F3/lipid contacts confirmed earlier results (18), showing a preference for the F2-F3 domain to interact with negative lipid headgroups.

We next investigated how the close association of the F2-F3 domain with an anionic bilayer might modulate the integrin TM domains, by performing simulations of F3 and of F2-F3 in complex with the β TM helix. The results were compared with the β TM helix alone (**β-AT, β-F3-AT**, and **β-F2F3-AT** in Table 1). The β helix was modeled on the structures of the αIIb/β3 TM helix dimer (pdb:2K9J (24)) and of talin F2-F3 bound to the cytoplasmic tail of the integrin β1D (pdb:3G9W (14)). In the β-F2-F3 simulations, there was a significant increase (~15°) in the tilt angle of the β TM helix relative to that observed in either the **β-AT** or **αβ-AT** simulations (Fig. 2). Further analysis of this change revealed no significant movement of the F3 sub-domain relative to the β helix. In contrast, the F2 sub-domain rotates ~10° relative to F3, possibly to optimize the interactions with the bilayer. Examination of H-bonds and close contacts reveals that the F2-F3/bilayer interface in the **β-F2F3-AT** simulation is similar to that identified in simulations of F2-F3 and bilayer alone. A comparable movement of the β TM helix was seen in the **β-F3-AT** simulations. Again an increase in β TM helix tilt angle of ~15° was seen over the course of the simulation.

The β TM helix also adopted a similar tilted orientation (~37°) in simulations (**β-AT**) of just the β TM in a bilayer. Taken together, these three simulations indicate that the β TM helix prefers to adopt a higher tilt angle relative to the bilayer than in the αβ TM dimer. The drive for the β TM helix to a higher tilt angle may arise because the hydrophobic helix length (~40 Å) is better accommodated; the tilted orientation also places the W715 sidechain in the lipid headgroup region and allows the K716 sidechain to interact with lipid phosphate groups. This tendency of the β3 TM helix to adopt a tilted orientation in the membrane is consistent with other studies of integrin TM domains (19, 56). The higher tilt angle of the β TM also allows the talin F2-F3 sub-domain to optimize its interactions with the anionic bilayer surface.

Changes in Integrin TM Helix Packing Induced by Talin F2-F3

We next sought to explore the effect of the talin F2-F3 subdomain on TM helix helix packing and/or the β TM tilt. To do this, the NMR structure of the αIIb/β3 TM helix dimer

(pdb:2K9J (24)); and the X-ray structure of F2-F3 in complex with the cytoplasmic tail of integrin β1D (pdb:3G9W (14)) were combined. The backbone coordinates of the helical region of β1D from the F2-F3-β1D structure and the β3 helix from α/β TM dimer structure were aligned. After the alignment, the β3 residues from T720 to F727 were removed thus keeping the equivalent β1D residues in the positions found in pdb:3G9W. This resulted in a chimeric structure (β3 residues 688-719/β1D residues 753-787) for the β subunit that combined the TM domain of β3 and the C-terminal tail of β1D. We note that the TM domains of β3 and β1D are very similar in sequence (90% similarity; 52% identity) and that the crucial salt-bridge in the IMC, between β1D D759 and αIIb R995, is preserved.

To explore the stability of the IMC αIIb R995/β1D D759 salt-bridge in this model of αβ-F2-F3, four atomistic simulations, each of 20 ns duration, were performed for the model in a POPC/POPG bilayer. Rapid rearrangement of the salt-bridge was observed, due to the interactions of the αIIbR995D sidechains with a positively charged loop (residues M322-L328) in the talin F3 subdomain. An additional control was performed, using a second model with a slightly different β chain chimera (β3 residues 688-727/β1D residues 764-787). In this model, the homologous salt-bridge is formed between the sidechains of β3 D723 and αIIb R995. A similar rapid reorganization of the salt-bridge was again observed in the presence of F2-F3, resulting in interactions of the αIIb R995 with the charged loop (residues M322-L328) on the F3 sub-domain.

In order to explore the possibility of longer timescale (>0.1 μs) changes in the conformation of the αβ/F2-F3 complex, a multi-scale approach was taken, combining CG simulations to explore longer timescale events, with subsequent atomistic MD simulations to refine the resultant models. A CG representation of the αβ/F2-F3 model structure was embedded in a POPC/POPG bilayer using our self-assembly protocol (57). ENM restraints between the αIIb subunit and the β/F2-F3 complex were removed. Five simulations (each of 5 μs) were performed. These simulations suggested that talin F2-F3 destabilizes the αIIb/β dimer; the number of contacts between the αIIb and β helices was greatly reduced at the dimer interface with both the IMC and OMC clasps being disrupted. In particular, the close packing of the two helices via the αIIb Gx_3G motif and the interactions of the αIIbF992, F993 aromatic cluster and the αIIbR995 with the β tail were disrupted. The αIIb helix maintained its orientation perpendicular to the bilayer (tilt angle ~10°) whilst the β helix tilt increased (tilt angle ~28°).

Analysis of the movements of the β and F2-F3 domains relative to the α TM helix suggests how F2-F3 may perturb the αIIb/β interactions. Each of the 3 CG simulations

interface in this contact region involves αIIb L978, L979, I982, A986, V990 and the β I707, G708, A711, T715, M718 residues (see Supporting Information Fig 2).

We note that during these movements the interactions of the F2-F3 and of the β C-terminal tail with the bilayer were preserved. In particular, the F2-F3 domains interacted preferentially with the negatively charged lipids in the membrane. This suggests that protein/lipid interactions must be considered when analyzing possible activation mechanisms of integrins.

Suggested Mechanism

Structural studies (7, 14, 16, 24, 25, 58) have greatly increased our understanding of integrin activation from inside the cell. However, static structures fall short of revealing dynamic mechanisms of activation. Using multiscale MD simulations we have revealed further details of the mechanism whereby the talin head domain, the lipid bilayer membrane and the α/β TM helix dimer together interact to enable inside-out activation of integrins. In particular, we propose a mechanism whereby talin F2-F3 interactions lead to a conformational change in the αβ TM helix dimer which may act as a trigger for inside-out activation of the integrin ecto-domain. In this model (see Fig. 4A) the first stage is when F2-F3 binds to anionic lipids via electrostatic interactions (18). This is determined by the F2 sub-domain which forms basic sidechain/anionic headgroup interactions. Talin F3 binds to the β integrin C-terminal tail in three areas: the membrane distal NPxY motif, the β1D W775 (β3 W739) residue and the β tail membrane proximal region. Previous experimental studies have shown that a tyrosine to alanine mutation in the distal NPxY motif or the β3 R358A mutation inhibits integrin activation (7, 8, 59-61). Thus, F2-F3 associates with the α/β TM helix dimer via an interaction between F3 and the cytoplasmic tail of β. A recent NMR study (62) suggested that, in the resting integrin off state, the β TM region is bent at residue R724, allowing the C terminal cytoplasmic β tail to interact with the membrane via two helices. Promotion of a helical structure in the R724 region would extend the β helix and allow binding of F3. This, together with F2-F3 interactions with the membrane bilayer, then leads to a "scissoring" motion of the α/β TM helix dimer about the IMC such that the OMC is disrupted, and the N-termini (extracellular) separation of the TM helices is increased (Fig. 3A).

The scissoring motion that was revealed in the αβF2-F3o-AT simulations reflects the tendency of the β TM helix to be more tilted in a lipid bilayer than it is in the 'off' state. This tilting motion may be seen quite clearly by monitoring the distances between the α and β

TMs. As shown in Fig. 4AB, distances in the IMC region (d3, d4) are maintained whereas the distances (d1, d2) at the N-termini of the helices increase by ca. 6 Å from start to end of the simulation.

We have used the coordinates of the αIIb/β3 structure determined by NMR in bicelles (pdb: 2K9J) here. Other structures have been reported, determined either in organic solvents (pdb: 2KNC) (25, 27) or from computational approaches (26, 63). In 2KNC, which was obtained using a 1:1 mixture of acetonitrile and water, the IMC part is different from both 2K9J and the computational models, suggesting that the lipid environment is important for the formation of the IMC interactions. It is of interest, however, that 2KNC is rather like αβF2-F3p suggesting that it might relate to an intermediate state between 'off' and 'on' Supporting Information Fig. 3.

The scissor model for inside-out activation of integrins suggests how the integrin ectodomain might be stimulated to go from a bent inactive state to an extended active state (34-36). Talin weakens αIIb/β3 interactions, thus promoting a scissor movement of the TM part of the two helices with a perturbed IMC as the centre of the scissor movement. This is in agreement with the recent suggestion that disruption of the α/β dimer involves competition between an α/β tail salt bridge and a talin/β tail salt bridge as well as a 20^{o} reorientation of the β integrin tail in a plane perpendicular to the membrane There is ample experimental data to suggest that OMC and IMC interactions are required for integrin activation (23-25, 30-32, 54, 64-67). Our results have clarified the role of the IMC and OMC clasps and explain much of the experimental data. Deletion or mutation of residues in the GFFKR motif in the α tail lead to integrin activation (30, 41, 66, 67). Similarly, deletion of the β tail cytoplasmic region leads to activation (39, 65). Further, leucine or isoleucine mutations in the GxxxG motif in the OMC clasp cause steric clashes and thus disruption of the packing of the α/β dimer (24, 31, 54). Our results are also in good agreement with mutational experimental results that suggest that disruption of the F3/ β tail interface or on the talin F2-F3 interface with the membrane reduce the ability of talin to activate integrins (7-14). Fluorescent resonance energy transfer (FRET) studies (29) using αL and β2 integrin subunits tagged with fluorescent proteins suggested significant reduction in the FRET signal in the presence of talin and this was interpreted in terms of tail separation. A scissor movement around the IMC might be expected to separate the scissor handles as well as the blades. Fig 4B (d4) does not indicate an increase in separation but the αIIbR995-β3D723R is broken and the largely cytoplasmic α and β tails are flexible so increased tail separation is not inconsistent with our model.

Recent structures of integrin ectodomains shows that the membrane-proximal C-terminal regions of the two ectodomain subunits are in close proximity (26, 36, 68). It is therefore plausible that the scissor movement which causes dissociation of the OMC clasp promotes a rearrangement of the integrin ectodomain.

In summary, our results suggest a mechanism for integrin inside-out activation that explains the role of the talin head domain, the membrane and the α/β TM interactions. We have shown that the α/β TM helix dimer is conformationally stable in a lipid bilayer on a 100 ns timescale, that the β TM helix has a propensity to tilt relative to the lipid bilayer and that this tilt motion is resisted by the presence of the α TM helix. Binding of talin destabilizes interactions between the TM helices and permits a scissor-like motion in part driven by the propensity of the β helix to tilt. This model shows how protein/membrane interactions can disrupt the TM helix interactions of complex receptors and potentially lead to structural changes in ectodomains. It also demonstrates the strength of multiscale simulations for developing models of the conformational dynamics of a complex membrane receptor.

MATERIALS AND METHODS

Coarse-grained MD simulations

Coarse-grained (CG) MD simulations were performed using a local variant (69) of the Martini force field (70) with an elastic network applied to backbone particles using a cut-off distance of 7 Å, and the protein inserted in a POPC/POPG bilayer. The system was solvated, energy minimized for 200 steps, and equilibrated for 5 ns with the protein backbone particle restrained. After that, 5 production simulations were carried out for 5 μs each.

All CG-MD simulations were performed using GROMACS 3.3.3 (www.gromacs.org) (71, 72). A Berendsen thermostat (73) (coupling constant of 1.0 ps; reference temperature 310 K) and barostat (coupling constant of 1.0 ps, compressibility value of 5.0×10^{-6} bar^{-1}, reference pressure 1 bar) were used. The integration time step was 40 fs. Lennard-Jones and Coulombic interactions were shifted to zero between 9 Å and 12 Å, and 0 and 12 Å respectively. CG to atomistic conversion used a fragment-based approach (50).

Atomistic MD simulations

Atomistic MD simulations were performed using the GROMOS96 43a1 force field (74). The Parinello-Rahman barostat (75) and Berendsen thermostat (73). The LINCS algorithm was used to constrain bond lengths (76) and particle mess Ewald particle mesh was used to model electrostatics up to 10 Å. A 10 Å cut-off distance was also used for van der Waals

interactions. Systems were equilibrated for 2 to 5 ns with the protein Cα atoms restrained, followed by unrestrained MD simulations as in Table 1. Analyses used Gromacs, VMD (77) and locally written codes.

ACKNOWLEDGMENTS

This research was funded by the Wellcome Trust. IDC acknowledges support from the NIH Cell Migration consortium.

FIGURE LEGENDS

Figure 1

The αIIbβ3 TM helix dimer.

A Structure of the αIIb (blue) β3 (red) TM helix dimer (PDB 2K9J: 24). Interacting residues in the OMC are in magenta, in the IMC clasp in orange, and the αIIb R995 to β3 D756 salt-bridge is in green.

B Snapshot (100 ns) from the αβ1-AT simulation (i.e. WT; see Table 1 for details) showing the two TM helices, the interacting residues of the OMC (magenta) and the phosphorus atoms of the POPC/POPG lipid bilayer (grey).

C Helix crossing angle distributions for the **αβ1-AT** (blue), **αβ2-AT** (i.e. αIIbR995D; red), and **αβ3-AT** (i.e. αIIb F992A,F993A; green) simulations. A negative crossing angle corresponds to right-handed helix packing, a positive angle to left-handed packing.

Figure 2

Tilting of the β3 TM helix in β3/talin simulations.

A Initial (blue) and final (t = 70 ns; red) orientation of the F2-F3-β complex in one of the β-F2-F3 simulations.

B The β3 TM helix tilt angle relative to the bilayer normal for **β-F2F3-AT** (black) and **β-F3-AT** (red) simulations.

Figure 3

Simulations of the αIIbβ3 TM helix dimer in complex with talin F2-F3.

A Movements of the α and β TM domains in the **αβ-F2F3-AT** simulations. Three structures are shown, the 'closed' structure at the start of the CG-MD simulations (**αβ-F2-F3c-AT**), a representative structure from the 'open' cluster after the CG-MD simulations (**αβ-F2-F3o-AT**), and the final 'open' structure from 100 ns AT-MD simulation. In these diagrams only the Cα traces of the TM helices (α in blue; β in red) are shown, along with the phosphorus atoms of the nearby lipid headgroups (grey). The OMC residues and the aromatic residues of the IMC are shown in green.

B Initial (blue) and final (t = 100 ns; red) configuration of the αβ-F2-F3 complex in the **αβ-F2-F3o-AT** simulation. The αIIb TM helix from the initial and the final structures of the simulation are superimposed.

C Snapshot corresponding to the final (t = 100 ns) configuration of the **αβ-F2-F3o-AT** simulation showing that the interactions in both the IMC and the OMC are disrupted. The F2 sub-domain of talin is omitted for clarity.

Figure 4

A Proposed mechanism of integrin inside-out activation. The positively charged patch on the F2 and F3 sub-domains of talin (**1**) interacts with the negatively charged headgroups on the membrane surface (**2**), orienting talin for binding to the β cytoplasmic tail (**3**). Upon talin binding (**4**), the αIIbR995-β3D723 salt-bridge is disrupted because of the interactions of the αIIbR995 with the talin F3 acidic loop. Strengthening of interactions between the positively charged surface patch of F2 and the membrane surface rotates the talin F2 sub-domain ~34° in a plane perpendicular to the bilayer normal, i.e. parallel to the lipid-water interface. This in turn rotates F3 (~26°) and the β tail (~30°) in the opposite direction and in the same plane. This, in combination with a ~15° increase in the tilt angle of the β helix relative to the bilayer normal, disrupts the interactions in both the IMC and the OMC of the αβ dimer (**5**). This is suggested to promote a switch of the integrin ectodomain to an active extended state.

B and C illustrate the proposed 'scissors' motion of the integrin α and β TM helices that occurs on activation by talin. **B** indicates four inter-helix distances (d1 to d4) which characterize the scissors motion and **C** shows the time evolution of d1 to d4 over the course of one of the three **αβ-F2-F3o-AT** simulations.

TABLES

Table 1: Summary of simulations and nomenclature

Simulation	Proteins	Coordinates	Duration (ns)
αβ1-AT	αIIb/β3 TM	2K9J	3x100
αβ2-AT	αIIb(R995D)/β3 TM		3x100
αβ3-AT	αIIb(F992A,F993A)/β3 TM		3x100
αβ4-AT	αIIb(F992A,F993A,Δ995-998)/β3 TM		3x100
αβ5-AT	αIIb(G972S)/β3 TM		1x100
αβ6-AT	αIIb(G976L)/β3 TM		1x100
F2F3-AT	talin F2F3	3G9W	1x30
β-AT	β* TM	model based on	3x70
β-F3-AT	β* TM/talin F3	2K9J β3 + 3G9W	3x70
β-F2F3-AT	β* TM/talin F2F3		3x70
αβ-F2F3-CG	αIIb/β* TM/talin F2F3	model based on	5x5000
αβ-F2F3c-AT	αIIb/β* TM/talin F2F3 closed	2K9J + 3G9W	3x100
αβ-F2F3p-AT	αIIb/β* TM/talin F2F3 partial		3x100
αβ-F2F3o-AT	αIIb/β* TM/talin F2F3 open		3x100

TM = transmembrane helix. All simulations were of the protein complex in a POPC/POPG lipid bilayer. System sizes ranged from ~40,000 (αβ simulations) to ~100,000 (αβ-F2F3 simulations). β* indicates a chimeric β chain (β3/β1D) - see main text for details).

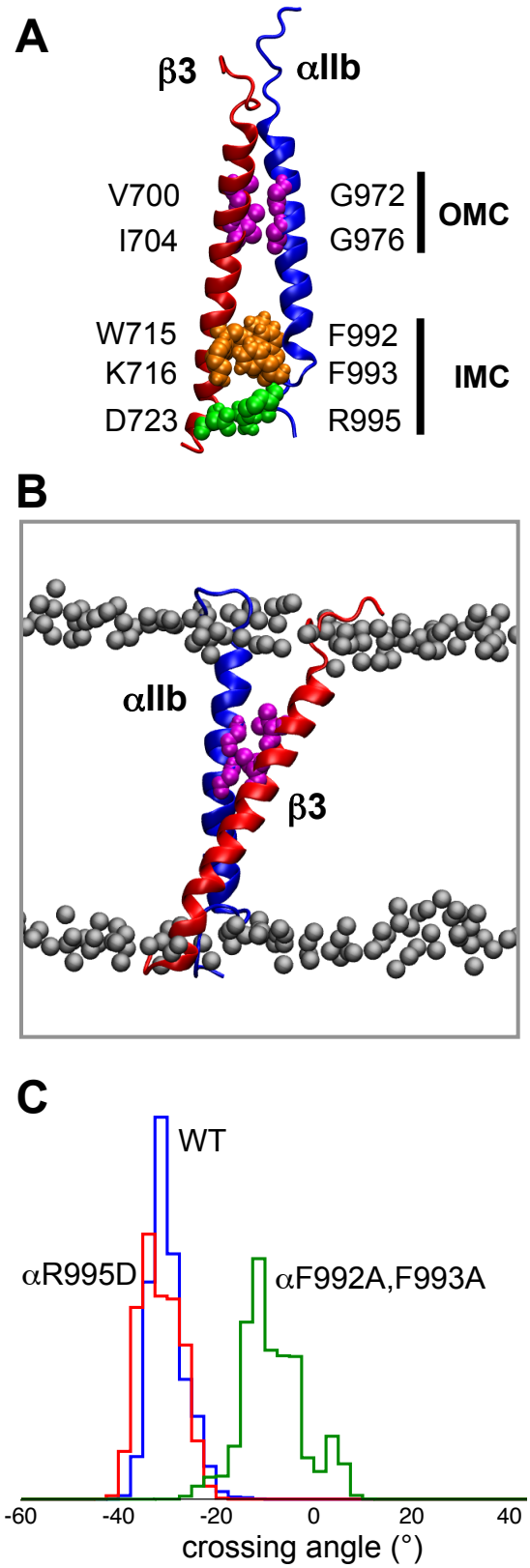
REFERENCES

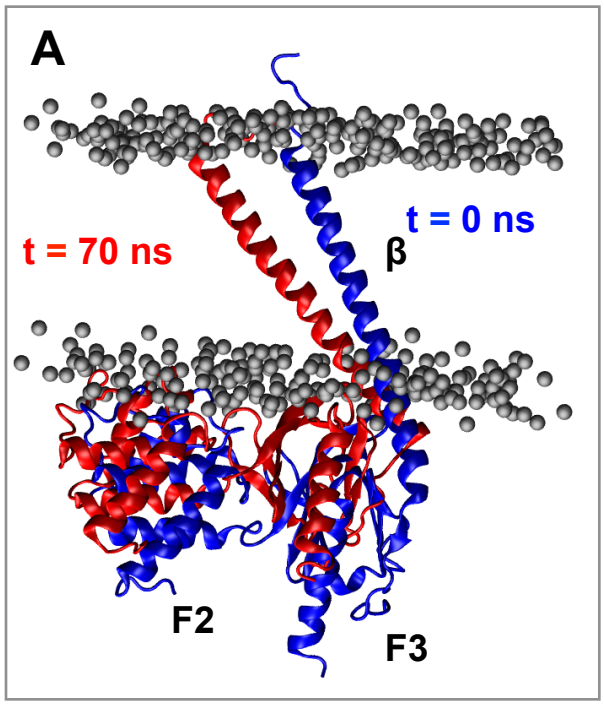
1. Wegener KL & Campbell ID (2008) Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (Review). *Mol. Mem. Biol.* 25(5):376 - 387.
2. Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* 110(6):673-687.
3. Ginsberg MH, Partridge A, & Shattil SJ (2005) Integrin regulation. *Curr. Opin. Cell Biol.* 17(5):509-516.
4. Harburger DS & Calderwood DA (2009) Integrin signalling at a glance. *J. Cell Sci.* 122(2):159-163.
5. Bennett JS, Berger BW, & Billings PC (2009) The structure and function of platelet integrins. *J. Thromb. Haemost.* 7(s1):200-205.
6. Gong H*, et al.* (2010) G protein subunit Gα13 binds to integrin αIIbβ3 and mediates integrin "outside-in" signaling. *Science* 327(5963):340-343.
7. Wegener KL*, et al.* (2007) Structural basis of integrin activation by talin. *Cell* 128(1):171-182.
8. Tadokoro S*, et al.* (2003) Talin binding to integrin β tails: a final common step in integrin activation. *Science* 302(5642):103-106.
9. Petrich BG*, et al.* (2007) Talin is required for integrin-mediated platelet function in hemostasis and thrombosis. *J. Exp. Med.*:jem.20071800.
10. Calderwood DA*, et al.* (1999) The talin head domain binds to integrin β subunit cytoplasmic tails and regulates integrin activation. *J. Biol. Chem.* 274(40):28071-28074.
11. Ulmer TS, Calderwood DA, Ginsberg MH, & Campbell ID (2003) Domain-specific interactions of talin with the membrane-proximal region of the integrin β3 subunit. *Biochemistry* 42(27):8307-8312.
12. Bouaouina M, Lad Y, & Calderwood DA (2008) The N-terminal domains of talin cooperate with the phosphotyrosine binding-like domain to activate β1 and β3 integrins. *J. Biol. Chem.* 283(10):6118-6125.
13. Calderwood DA (2004) Talin controls integrin activation. *Biochem. Soc. Trans.* 32(Pt3):434-437.
14. Anthis NJ*, et al.* (2009) The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO J.* 28(22):3623-3632.
15. Critchley DR & Gingras AR (2008) Talin at a glance. *J. Cell Sci.* 121(9):1345-1347.
16. Elliott PR*, et al.* (2010) The structure of the talin head reveals a novel extended conformation of the FERM domain. *Structure* 18(10):1289-1299.
17. Calderwood DA*, et al.* (2002) The phosphotyrosine binding-like domain of talin activates integrins. *J. Biol. Chem.* 277(24):21749-21758.
18. Kalli AC*, et al.* (2010) The structure of the talin/integrin complex at a lipid bilayer: an NMR and MD simulation study. *Structure* 18(10):1280-1288.
19. Saltel F*, et al.* (2009) New PI(4,5)P2- and membrane proximal integrin-binding motifs in the talin head control β3-integrin clustering. *J. Cell Biol.* 187(5):715-731.
20. Goult BT*, et al.* (2010) Structure of a double ubiquitin-like domain in the talin head: a role in integrin activation. *EMBO J.* 29(6):1069-1080.
21. Adair BD & Yeager M (2002) Three-dimensional model of the human platelet integrin αIIbβ3 based on electron cryomicroscopy and x-ray crystallography. *Proc. Natl. Acad. Sci. USA* 99(22):14059-14064.
22. Schneider D & Engelman DM (2004) Involvement of transmembrane domain interactions in signal transduction by α/β integrins. *J. Biol. Chem.* 279(11):9840-9846.

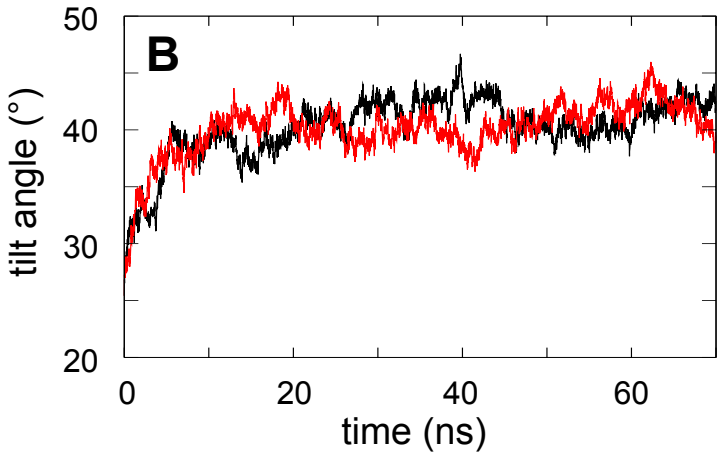
23. Luo B-H, Springer TA, & Takagi J (2004) A specific interface between integrin transmembrane helices and affinity for ligand. *PLoS Biol.* 2(6):e153.
24. Lau T-L, Kim C, Ginsberg MH, & Ulmer TS (2009) The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* 28(9):1351-1361.
25. Yang J*, et al.* (2009) Structure of an integrin αIIbβ3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc. Natl. Acad. Sci. USA* 106(42):17729-17734.
26. Zhu J*, et al.* (2009) The structure of a receptor with two associating transmembrane domains on the cell surface: integrin αIIbβ3. *Mol. Cell* 34(2):234-249.
27. Vinogradova O*, et al.* (2002) A structural mechanism of integrin αIIbβ3 "inside-out" activation as regulated by its cytoplasmic face. *Cell* 110(5):587-597.
28. Bhunia A, Tang X-Y, Mohanram H, Tan S-M, & Bhattacharjya S (2009) NMR solution conformations and interactions of integrin αLβ2 cytoplasmic tails. *J. Biol. Chem.* 284(6):3873-3884.
29. Kim M, Carman CV, & Springer TA (2003) Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science* 301(5640):1720-1725.
30. Hughes PE*, et al.* (1996) Breaking the integrin hinge. *J. Biol. Chem.* 271(12):6571-6574.
31. Li W*, et al.* (2005) A push-pull mechanism for regulating integrin function. *Proc. Natl. Acad. Sci. USA* 102(5):1424-1429.
32. Partridge AW, Liu S, Kim S, Bowie JU, & Ginsberg MH (2005) Transmembrane domain helix packing stabilizes integrin αIIbβ3 in the low affinity state. *J. Biol. Chem.* 280(8):7294-7300.
33. Zhu J*, et al.* (2007) Requirement of α and β subunit transmembrane helix separation for integrin outside-in signaling. *Blood* 110(7):2475-2483.
34. Adair BD*, et al.* (2005) Three-dimensional EM structure of the ectodomain of integrin αVβ3 in a complex with fibronectin. *J. Cell Biol.* 168(7):1109-1118.
35. Rocco M, Rosano C, Weisel JW, Horita DA, & Hantgan RR (2008) Integrin conformational regulation: uncoupling extension/tail separation from changes in the head region by a multiresolution approach. *Structure* 16(6):954-964.
36. Zhu J*, et al.* (2008) Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces. *Mol. Cell* 32(6):849-861.
37. Nogales A*, et al.* (2009) Three-dimensional model of human platelet integrin αIIbβ3 in solution obtained by small angle neutron scattering. *J. Biol. Chem.* 285(2):1023-1031.
38. Williams MJ, Hughes PE, O'Toole TE, & Ginsberg MH (1994) The inner world of cell adhesion: integrin cytoplasmic domains. *Trends Cell Biol.* 4(4):109-112.
39. Lu C, Takagi J, & Springer TA (2001) Association of the membrane proximal regions of the α and β subunit cytoplasmic domains constrains an integrin in the inactive state. *J. Biol. Chem.* 276(18):14642-14648.
40. Kim S, Chamberlain AK, & Bowie JU (2003) A simple method for modeling transmembrane helix oligomers. *J. Mol. Biol.* 329(4):831-840.
41. Luo B-H, Carman CV, & Springer TA (2007) Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.* 25(1):619-647.
42. Adcock SA & McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* 106(5):1589-1615.
43. Ash WL, Zlomislic MR, Oloo EO, & Tieleman DP (2004) Computer simulations of membrane proteins. *Biochim. Biophys. Acta Biomembranes* 1666(1-2):158-189.

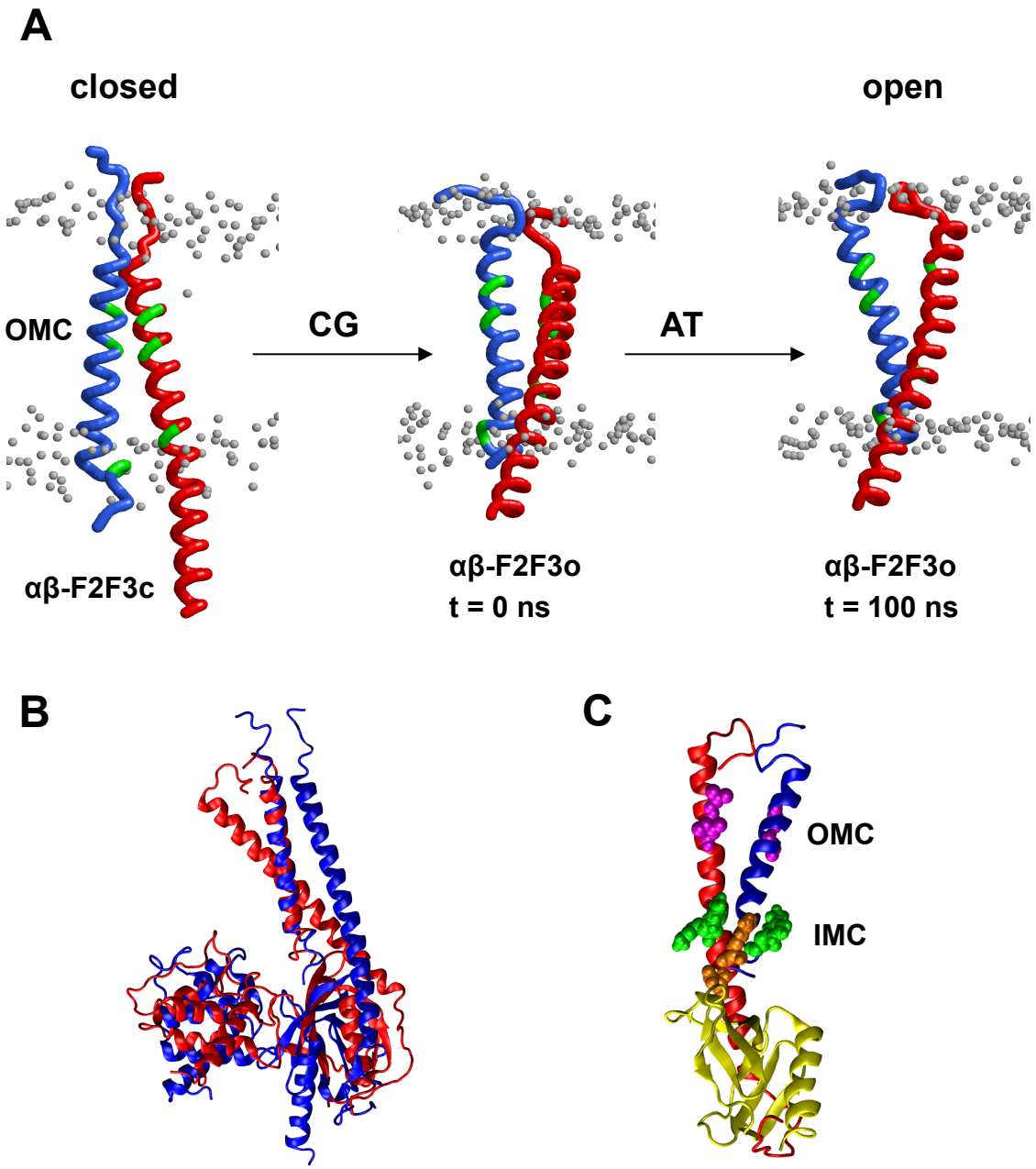
44. Psachoulia EMI, Marshall DP, Sansom MSP, & Spectus CON (2010) Dimerization of Transmembrane α-Helices. *Accounts of Chemical Research* 43(3).
45. Liepiņa I, Czaplewski C, Janmey P, & Liwo A (2003) Molecular dynamics study of a gelsolin-derived peptide binding to a lipid bilayer containing phosphatidylinositol 4,5-bisphosphate. *Pept. Sci.* 71(1):49-70.
46. Jaud S, Tobias DJ, Falke JJ, & White SH (2007) Self-induced docking site of a deeply embedded peripheral membrane protein. *Biophys. J.* 92(2):517-524.
47. Psachoulia E & Sansom MSP (2008) Interactions of the pleckstrin homology domain with phosphatidylinositol phosphate and membranes: characterization via molecular dynamics simulations. *Biochemistry* 47(14):4211-4220.
48. Psachoulia E & Sansom MSP (2009) PX- and FYVE-mediated interactions with membranes: simulation studies. *Biochemistry* 48(23):5090-5095.
49. Ayton GS & Voth GA (2009) Systematic multiscale simulation of membrane protein systems. *Curr. Opin. Cell Biol.* 19(2):138-144.
50. Stansfeld PJ & Sansom MSP (2011) From coarse grained to atomistic: a serial multiscale approach to membrane protein simulations. *J. Chem. Theory Comput.* 7(4):1157-1166.
51. Lau T-L, Dua V, & Ulmer TS (2008) Structure of the integrin αIIb transmembrane segment. *J. Biol. Chem.* 283(23):16162-16168.
52. Kim C, Lau T-L, Ulmer TS, & Ginsberg MH (2009) Interactions of platelet integrin αIIb and β3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood* 113(19):4747-4753.
53. MacKenzie KR, Prestegard JH, & Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276(5309):131-133.
54. Luo B-H, Carman CV, Takagi J, & Springer TA (2005) Disrupting integrin transmembrane domain heterodimerization increases ligand binding affinity, not valency or clustering. *Proc. Natl. Acad. Sci. USA* 102(10):3679-3684.
55. Psachoulia E, Fowler PW, Bond PJ, & Sansom MSP (2008) Helix-helix interactions in membrane proteins: coarse-grained simulations of glycophorin A helix dimerization. *Biochemistry* 47(40):10503-10512.
56. Lau T-L, Partridge AW, Ginsberg MH, & Ulmer TS (2008) Structure of the integrin β3 transmembrane segment in phospholipid bicelles and detergent micelles. *Biochemistry* 47(13):4008-4016.
57. Scott KA*, et al.* (2008) Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* 16(4):621-630.
58. Goult BT*, et al.* (2009) The structure of an interdomain complex that regulates talin activity. *J. Biol. Chem.* 284(22):15097-15106.
59. O'Toole TE, Ylanne J, & Culley BM (1995) Regulation of integrin affinity states through an NPXY motif in the subunit cytoplasmic domain. *J. Biol. Chem.* 270(15):8553-8558.
60. GarcIa-Alvarez B*, et al.* (2003) Structural determinants of integrin recognition by talin. *Mol. Cell* 11(1):49-58.
61. Ulmer TS, Yaspan B, Ginsberg MH, & Campbell ID (2001) NMR analysis of structure and dynamics of the cytosolic tails of integrin αIIbβ3 in aqueous solution. *Biochemistry* 40(25):7498-7508.
62. Metcalf DG*, et al.* (2010) NMR analysis of the αIIbβ3 cytoplasmic interaction suggests a mechanism for integrin regulation. *Proc. Natl. Acad. Sci. USA* 107(52):22481-22486.

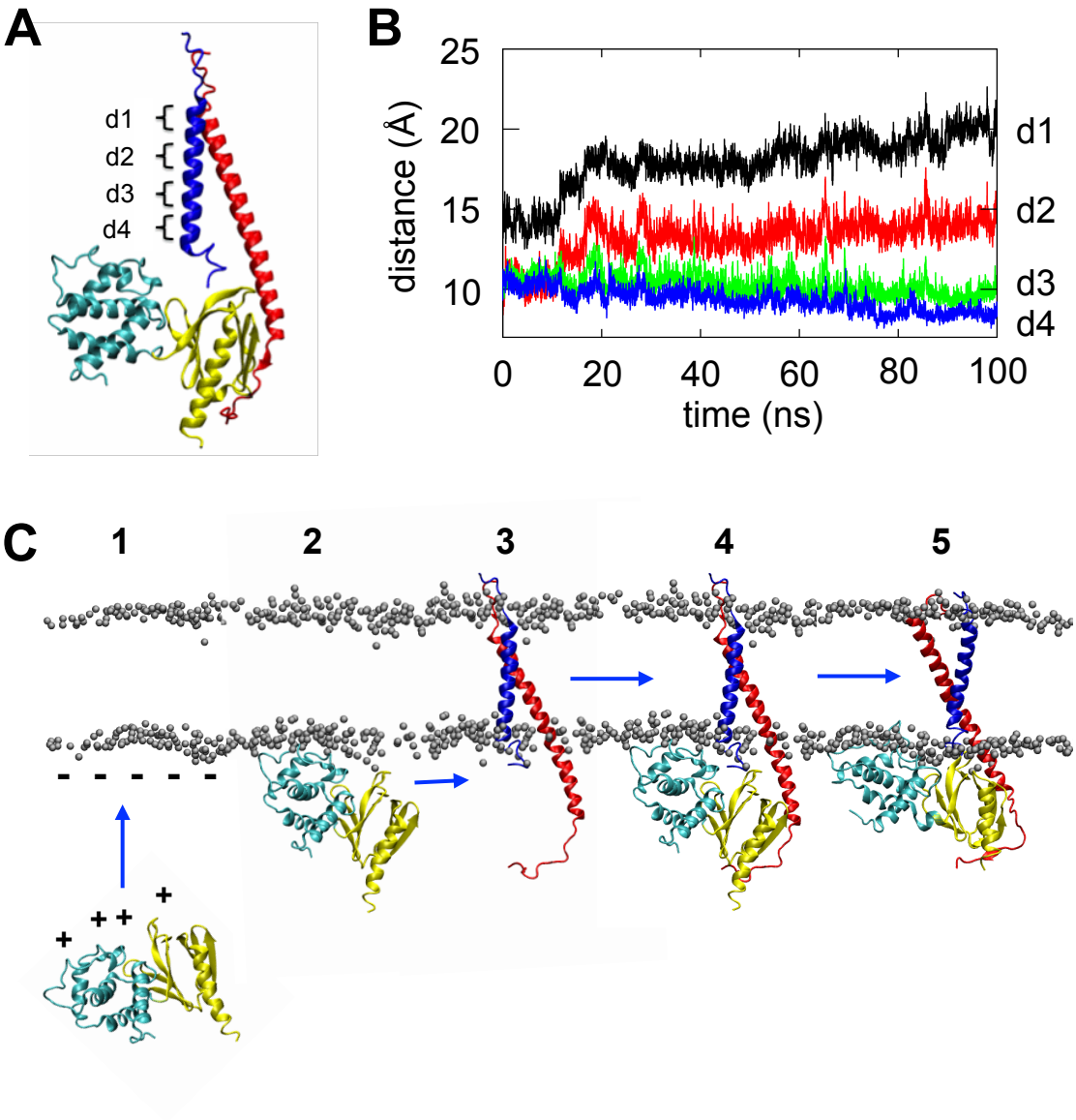
63. Metcalf DG, Kulp DW, Bennett JS, & DeGrado WF (2009) Multiple approaches converge on the structure of the integrin αIIb/β3 transmembrane heterodimer. *J. Mol. Biol.* 392(4):1087-1101.
64. Berger BW*, et al.* (2009) Consensus motif for integrin transmembrane helix association. *Proc. Natl. Acad. Sci. USA*.
65. Hughes PE, O'Toole TE, Ylanne J, Shattil SJ, & Ginsberg MH (1995) The conserved membrane-proximal region of an integrin cytoplasmic domain specifies ligand binding affinity. *J. Biol. Chem.* 270(21):12411-12417.
66. O'Toole TE*, et al.* (1994) Integrin cytoplasmic domains mediate inside-out signal transduction. *J. Cell Biol.* 124(6):1047-1059.
67. O'Toole TE*, et al.* (1991) Modulation of the affinity of integrin αIIbβ3 (GPIIb-IIIa) by the cytoplasmic domain of αIIb. *Science* 254(5033):845-847.
68. Xiong J-P*, et al.* (2009) Crystal structure of the complete integrin αVβ3 ectodomain plus an α/β transmembrane fragment. *J. Cell Biol.* 186(4):589-600.
69. Bond PJ, Wee CL, & Sansom MSP (2008) Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochemistry* 47(43):11321-11331.
70. Monticelli L*, et al.* (2008) The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* 4(5):819-834.
71. Lindahl E, Hess B, & van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7(8):306-317.
72. David Van Der S*, et al.* (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26(16):1701-1718.
73. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81(8):3684-3690.
74. Scott WRP*, et al.* (1999) The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* 103(19):3596-3607.
75. Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52(12):7182-7190.
76. Hess B, Bekker H, Berendsen H, & Fraaije J (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18(12):1463-1472.
77. Humphrey W, Dalke A, & Schulten K (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* 14(1):33-38.

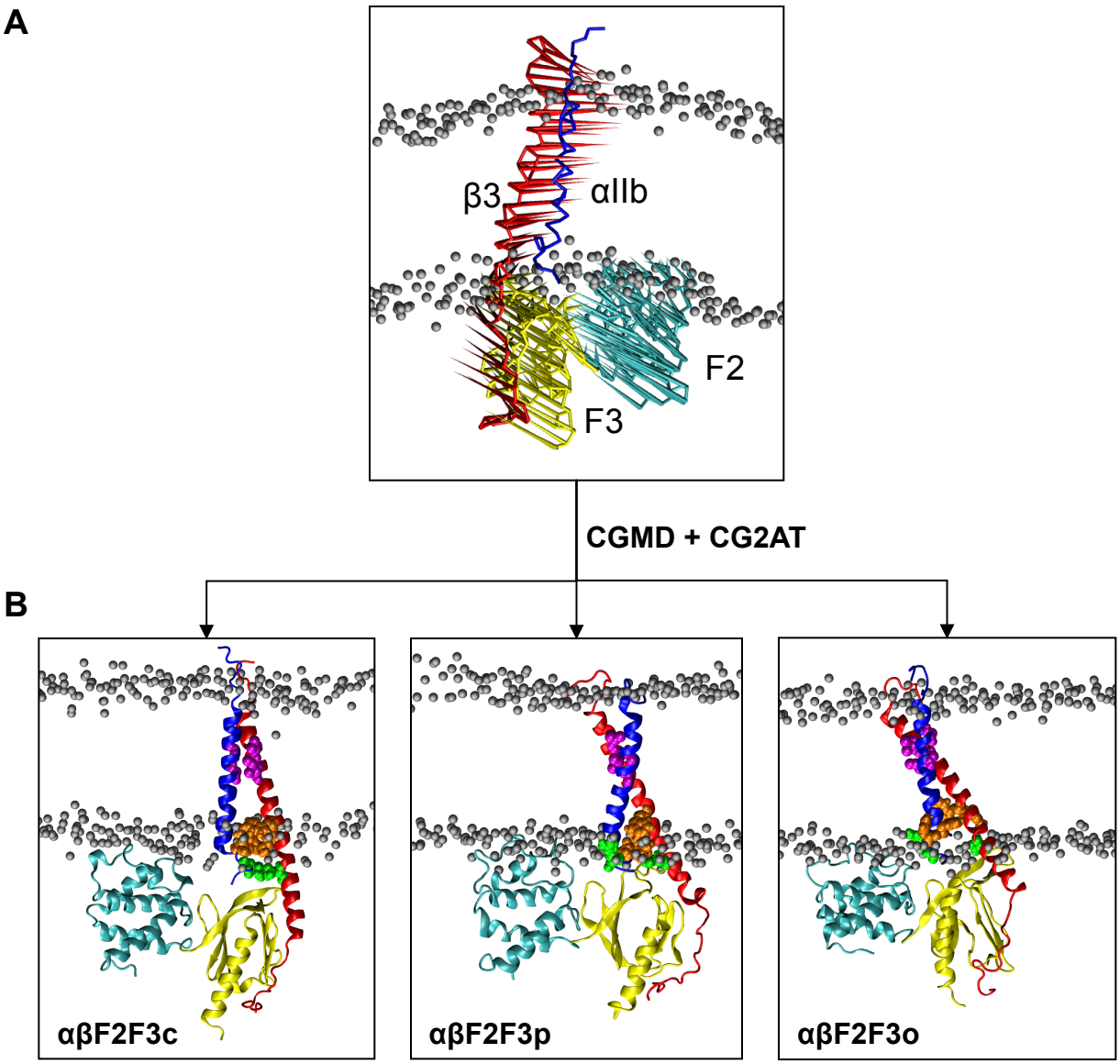

A
β3
αIIb
V700
I704
G972
G976
OMC
W715
K716
D723
F992
F993
R995
IMC
B
αIIb
β3
C
WT
αR995D
αF992A,F993A
-60
-40
-20
0
20
40
crossing angle (°)

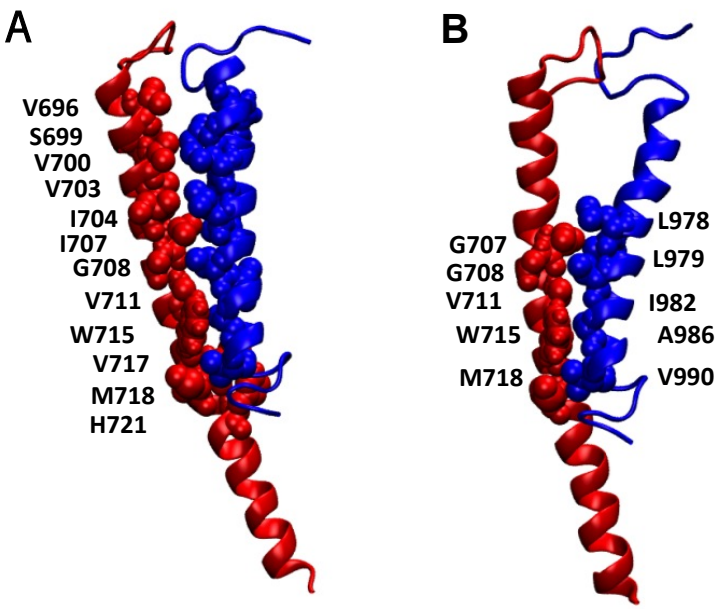

A
t = 70 ns
t = 0 ns
β
F2
F3

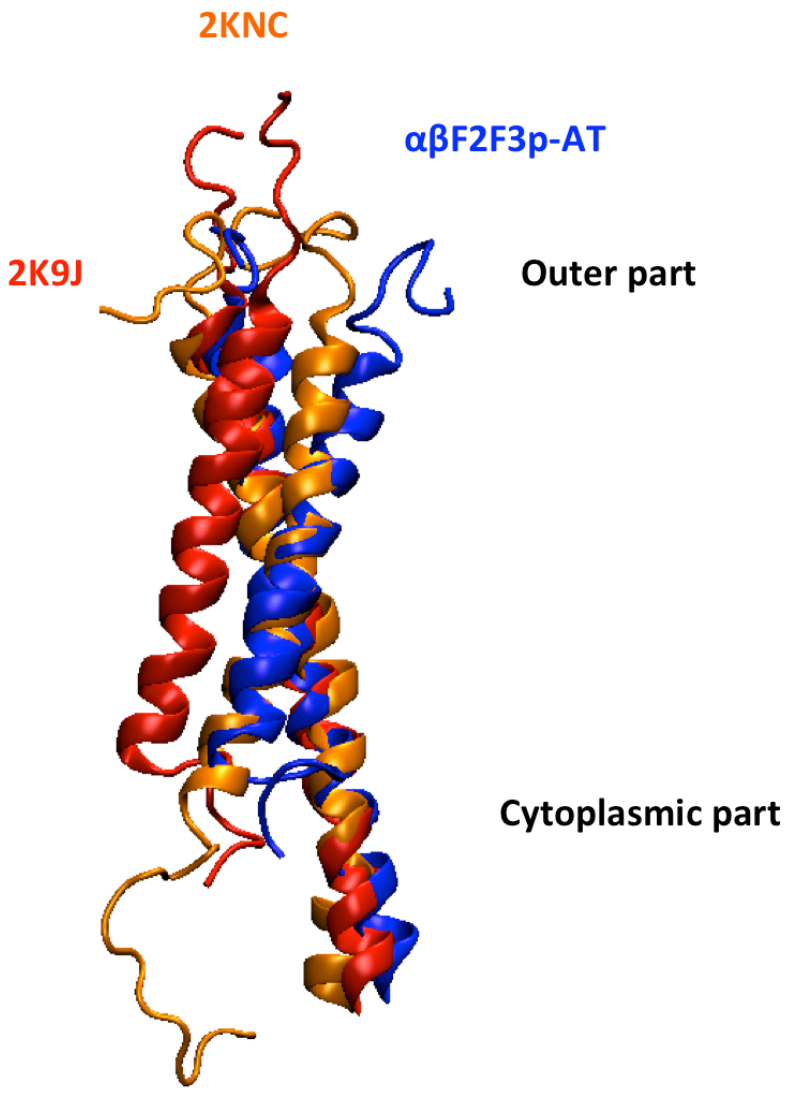

B
tilt angle (°)
50
40
30
20
0
20
40
60
time (ns)

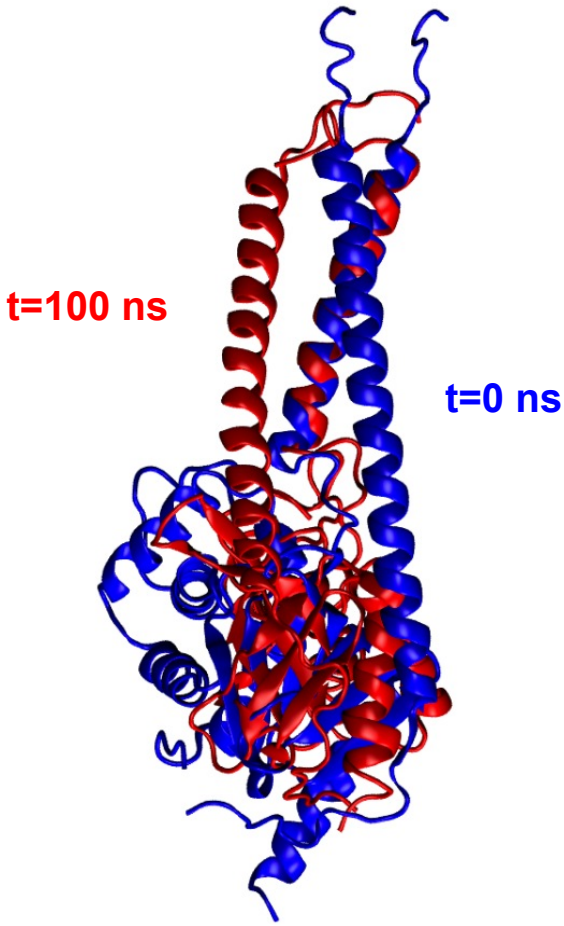

A
closed
open
OMC
CG
AT
αβ-F2F3c
αβ-F2F3o
t = 0 ns
αβ-F2F3o
t = 100 ns
B
C
OMC
IMC


A
d1
d2
d3
d4
B
distance (Å)
25
20
15
10
0
20
40
60
80
100
time (ns)
d1
d2
d3
d4
C
1
2
3
4
5
- - - - -
+ + +
+ +


A
β3
αIIb
F2
F3
CGMD + CG2AT
B
αβF2F3c
αβF2F3p
αβF2F3o


A
V696
S699
V700
V703
I704
I707
G708
V711
W715
V717
M718
H721
B
G707
G708
V711
W715
M718
L978
L979
I982
A986
V990


2KNC
αβF2F3p-AT
2K9J
Outer part
Cytoplasmic part


t=100 ns
t=0 ns