Sampo Kukkurainen

Molecular Dynamics of Talin-Integrin Interaction
SAMPO KUKKURAINEN

MOLECULAR DYNAMICS OF TALIN-INTEGRIN INTERACTION
ACADEMIC DISSERTATION
University of Tampere, Faculty of Medicine and Life Sciences
Tampere Graduate Program in Biomedicine and Biosciences (TGPBB)
Doctoral Programme in Medicine and Life Sciences

Supervised by
Associate Professor Vesa P. Hytönen
University of Tampere
Finland

Reviewed by
Adjunct Professor Aki Manninen
University of Oulu
Finland

Dr Jarkko Valjakka
University of Tampere
Finland

Professor Mohammad R. K. Mofrad
University of California, Berkeley
CA, USA

The originality of this thesis has been checked using the Turnitin Originality Check service in accordance with the quality management system of the University of Tampere.

Copyright © 2018 Sampo Kukkurainen

Acta Electronica Universitatis Tamperensis 1944
ISBN 978-952-03-0889-6 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Printed at BoD – Books on Demand, Norderstedt, Germany
ABSTRACT

Cells sense their environment and adhere to substrates with the help of cell surface receptors, integrins. The activity of integrins is controlled by talin, a large intracellular adaptor protein binding to the cytoplasmic tail of integrin. Talin both activates integrin and connects it to the force-bearing machinery of the cell. The talin-integrin complex mediates mechanical signals between the extracellular matrix and the intracellular actin cytoskeleton, and the tension dictates the coupling of further adaptor and signaling proteins to talin. These mechanical and biochemical signals affect the shape and motion of the cell, and regulate gene expression. Despite numerous studies on talin-integrin interactions, it is not yet fully clear how talin activates integrin, and the complete structure of the integrin-binding talin head domain remains to be solved.

In this thesis I focus on the interactions of integrin and the talin head domain using molecular dynamics simulations. In the first study we analyzed how mechanical load affects the talin-integrin binding interface in a set of constant force pulling simulations. We found that the mechanical stability of the talin-integrin interface requires a set of anti-parallel $\beta$-sheet-like hydrogen bonds, but not the binding of the $\beta$-integrin NPxY motif.

In the second and third studies, we studied the talin-\(\alpha \text{IIb}\beta 3\) integrin complex in the presence of a lipid bilayer. These simulations were carried out using two different talin models: an extended talin head conformation based on a published crystallographic structure, and our proposed compact talin conformation, whose design is based on homologous 3-D structures. Simulations of the extended talin conformation showed that the integrin dimer could become perturbed, and even activated, by acidic phospholipids in the cell membrane. With our own, compact domain model, our MD simulations predicted an interaction between a flexible loop and integrin. Experiments carried out \textit{in vitro} and \textit{in cellulo} supported the proposed compact conformation of the talin head and provided evidence for the predicted interaction between the talin loop and integrin.


# TABLE OF CONTENTS

Abbreviations ............................................................................................................................... 9
One-letter codes of amino acids .................................................................................................................. 13

List of original publications............................................................................................................... 14

1 Introduction .................................................................................................................... 15

2 Review of the literature.................................................................................................. 16
   2.1 Cell-matrix adhesion ........................................................................................................ 16
      2.1.1 Architecture of cell-matrix adhesions ................................................................................. 16
      2.1.2 Integrin clustering ............................................................................................................... 17
      2.1.3 Phosphoinositides in cell adhesion ....................................................................................... 18
   2.2 Integrin ........................................................................................................................ 18
      2.2.1 Integrin ectodomains ......................................................................................................... 19
      2.2.2 Integrin transmembrane domains ....................................................................................... 20
      2.2.3 β-integrin cytoplasmic domains .......................................................................................... 21
      2.2.4 Integrin activation ............................................................................................................... 22
   2.3 Talin ............................................................................................................................. 23
      2.3.1 The talin head FERM domain ............................................................................................ 25
      2.3.2 Talin-integrin interface ........................................................................................................ 28
      2.3.3 Talin regulation ................................................................................................................... 30
   2.4 Characterization methods for protein structures .......................................................... 32
      2.4.1 Small-angle X-ray scattering .............................................................................................. 34
      2.4.2 Molecular dynamics ............................................................................................................ 36

3 Aims of the study .............................................................................................................. 38

4 Materials and methods .................................................................................................... 39
   4.1 Modeling (I, II, III) ............................................................................................................. 39
   4.2 Molecular dynamics (I, II, III) ............................................................................................. 41
      4.2.1 Steered molecular dynamics (SMD) with NAMD (I) .......................................................... 41
      4.2.2 Molecular dynamics with GROMACS (II, III) .................................................................. 42
   4.3 Small-angle X-ray scattering (III) ..................................................................................... 43
   4.4 Biosensor analysis for talin-integrin binding (I, III) .......................................................... 45
   4.5 Cell assays (III) ................................................................................................................. 46
      4.5.1 Integrin activation (III) ....................................................................................................... 47
      4.5.2 Integrin clustering (III) ........................................................................................................ 47
      4.5.3 Cysteine-cysteine crosslinking (III) .................................................................................... 47
   4.6 Structural characterization (III) ........................................................................................... 48
      4.6.1 Electron microscopy (III) ................................................................................................... 48
      4.6.2 Mass spectrometry (III) ..................................................................................................... 48
Results ............................................................................................................................. 49
5.1 Talin-integrin binding ................................................................................................. 49
  5.1.1 Steered molecular dynamics of β-integrins and talin F2F3.............................. 49
  5.1.2 αIIbβ3 integrin, talin and a lipid bilayer .......................................................... 54
5.2 Lipid interactions ...................................................................................................... 55
  5.2.1 Talin-lipid binding ............................................................................................ 56
  5.2.2 Integrin-lipid binding ....................................................................................... 58
5.3 Talin FERM .............................................................................................................. 59
5.4 F1 loop function ...................................................................................................... 62

Discussion ...................................................................................................................... 64
6.1 Why study talin-integrin interaction? ..................................................................... 64
6.2 Talin-integrin binding ............................................................................................. 65
6.3 Lipid interactions ..................................................................................................... 66
6.4 Talin FERM ............................................................................................................. 67
6.5 How does talin activate integrin? .......................................................................... 67

Conclusions .................................................................................................................... 69

Acknowledgments ......................................................................................................... 70

References ..................................................................................................................... 71

Original publications ..................................................................................................... 83
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5GC1</td>
<td>a rat monoclonal antibody, recognizes CD31 Ig domains</td>
</tr>
<tr>
<td>9EG7</td>
<td>a rat monoclonal antibody, recognizes ligand-bound β1 integrin</td>
</tr>
<tr>
<td>ABS1, ABS2,</td>
<td>actin binding sites in talin</td>
</tr>
<tr>
<td>ABS3</td>
<td></td>
</tr>
<tr>
<td>ADMIDAS</td>
<td>adjacent metal ion-dependent adhesion site, a site in integrin ectodomains</td>
</tr>
<tr>
<td>AKT</td>
<td>protein kinase B</td>
</tr>
<tr>
<td>BDPLT16</td>
<td>bleeding disorder platelet-type 16, autosomal dominant Glanzmann's thrombasthenia</td>
</tr>
<tr>
<td>Cdk5</td>
<td>cyclin-dependent kinase 5</td>
</tr>
<tr>
<td>CFP</td>
<td>cyan fluorescent protein</td>
</tr>
<tr>
<td>cryo-EM</td>
<td>cryogenic electron microscopy</td>
</tr>
<tr>
<td>Da</td>
<td>dalton, ( \sim 1.66 \times 10^{-24} ) g</td>
</tr>
<tr>
<td>Dab1</td>
<td>disabled homolog 1</td>
</tr>
<tr>
<td>Dab2</td>
<td>disabled homolog 2</td>
</tr>
<tr>
<td>DLC1</td>
<td>deleted in liver cancer 1</td>
</tr>
<tr>
<td>Dmax</td>
<td>maximum dimension</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Dok1</td>
<td>docking protein 1</td>
</tr>
<tr>
<td>DOPC</td>
<td>1,2-dioleoyl-sn-glycero-3-phosphocholine</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>Edc/NHS</td>
<td>1-ethyl-3-(3-dimethylaminopropyl)carbodiimide/N-hydroxysuccinimide</td>
</tr>
<tr>
<td>EOM</td>
<td>Ensemble Optimization Method, a flexible modeling tool for SAXS data in the ATSAS package</td>
</tr>
<tr>
<td>EPB41L3</td>
<td>band 4.1-like protein 3, alternative names: differentially expressed in adenocarcinoma of the lung protein 1 (DAL-1), 4.1B</td>
</tr>
<tr>
<td>EPS8</td>
<td>epidermal growth factor receptor kinase substrate 8</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>ERM</td>
<td>ezrin, radixin, moesin</td>
</tr>
<tr>
<td>F0, F1, F2, F3</td>
<td>domains of the talin head</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FAK</td>
<td>focal adhesion kinase</td>
</tr>
<tr>
<td>FERM</td>
<td>protein 4.1, ezrin, radixin, moesin-like domain, composed of subdomains F1, F2 and F3</td>
</tr>
<tr>
<td>Gα13</td>
<td>guanine nucleotide-binding protein subunit α13</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione S-transferase</td>
</tr>
<tr>
<td>HDX</td>
<td>hydrogen/deuterium exchange</td>
</tr>
<tr>
<td>HUVEC</td>
<td>human umbilical vein endothelial cell</td>
</tr>
<tr>
<td>I(s)</td>
<td>scattering intensity</td>
</tr>
<tr>
<td>IBS2</td>
<td>integrin-binding site 2 in talin</td>
</tr>
<tr>
<td>ILK</td>
<td>integrin-linked kinase, a pseudokinase</td>
</tr>
<tr>
<td>IMC</td>
<td>inner membrane clasp, a sequence motif at the border of integrin transmembrane and cytosolic domains</td>
</tr>
<tr>
<td>JAK1</td>
<td>tyrosine-protein kinase JAK1, Janus kinase 1</td>
</tr>
<tr>
<td>JAK2</td>
<td>tyrosine-protein kinase JAK2, Janus kinase 2</td>
</tr>
<tr>
<td>KANK1</td>
<td>KN motif and ankyrin repeat domain-containing protein 1</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>Kd</td>
<td>dissociation constant</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton, a thousand daltons (see also Da)</td>
</tr>
<tr>
<td>KRIT1</td>
<td>Krev interaction trapped protein 1, alternative name: Cerebral cavernous malformations 1 protein (CCM1)</td>
</tr>
<tr>
<td>MD</td>
<td>molecular dynamics</td>
</tr>
<tr>
<td>MIDAS</td>
<td>metal-ion dependent adhesion site, a site in integrin ectodomains</td>
</tr>
<tr>
<td>MP</td>
<td>membrane-proximal</td>
</tr>
<tr>
<td>NAMD</td>
<td>Nanoscale Molecular Dynamics, a molecular dynamics program</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>nickel-nitrilotriacetic acid</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance spectroscopy</td>
</tr>
<tr>
<td>NPxY</td>
<td>conserved PTB-domain-binding motif: asparagine-proline-variable-tyrosine</td>
</tr>
</tbody>
</table>
Numb | protein numb homolog
--- | ---
NxxY | conserved PTB-domain-binding motif: asparagine-variable-variable-tyrosine
OMC | outer membrane clasp, a sequence motif in the integrin transmembrane domains
p(r) | pair distance distribution function
PDB | Protein Data Bank
PH domain | pleckstrin homology domain
PI(3)P | phosphatidylinositol 3-phosphate
PI(4)P | phosphatidylinositol 4-phosphate
PI(4,5)P2 | phosphatidylinositol (4,5)-bisphosphate
PI3K | phosphatidylinositol-4,5-bisphosphate 3-kinase
PI5K1γ | phosphatidylinositol 4-phosphate 5-kinase type 1γ
PIP2 | phosphatidylinositol (4,5)-bisphosphate
PIP3 | phosphatidylinositol (3,4,5)-trisphosphate
PIPK1γ | phosphatidylinositol 4-phosphate 5-kinase type 1γ
PKC | protein kinase C
PSI domain | domain found in plexins, semaphorins and integrins
PTB domain | phosphotyrosine-binding-like domain
PX domain | phox (phagocytic oxidase) homology domain
R1-R13 | α-helical bundles in talin rod
RALS | right-angle light scattering
Rap1 | Ras-related protein 1
Rg | radius of gyration
RGD | arginine-glycine-aspartate, a sequence motif present in some integrin ligands
RhoA | Ras homolog gene family member A, a small GTPase
RIAM | Rap1-GTP-interacting adapter molecule, alternative name: Amyloid β A4 precursor protein-binding family B member 1-interacting protein (APBB1IP)
s | $4\pi \sin(\theta)/\lambda$, momentum transfer, often also denoted q
SAXS | small-angle X-ray scattering
SDS-PAGE | sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC       size-exclusion chromatography
SKI-7     kistrin-CD31 fusion protein, recognizes active β3 integrin
SMD       steered molecular dynamics
SNX17     sorting nexin-17
Src       proto-oncogene tyrosine-protein kinase Src
SyMBS     synergistic metal ion-binding site, a site in integrin ectodomains
talA      talin-A, a talin homolog in Dictostelium discoideum
talB      talin-B, a talin homolog in Dictostelium discoideum
TEM       transmission electron microscopy
TIAM1     T-lymphoma invasion and metastasis-inducing protein 1
TIRF      total internal fluorescence microscopy
TM        transmembrane
TYK2      non-receptor tyrosine-protein kinase TYK2
UV        ultraviolet
VASP      vasodilator-simulated phosphoprotein
VE-β3     β3 integrin with high-affinity talin binding mutation D740TAN743/VE
VMD       Visual Molecular Dynamics, a molecular visualization program
Vp        Porod volume
WT        wild-type
XFEL      X-ray free electron laser
Å         ångström, 10^{-10} m
### One-letter codes of amino acids

<table>
<thead>
<tr>
<th>A</th>
<th>alanine</th>
<th>M</th>
<th>methionine</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>cysteine</td>
<td>N</td>
<td>asparagine</td>
</tr>
<tr>
<td>D</td>
<td>aspartic acid, aspartate</td>
<td>P</td>
<td>proline</td>
</tr>
<tr>
<td>E</td>
<td>glutamic acid, glutamate</td>
<td>Q</td>
<td>glutamine</td>
</tr>
<tr>
<td>F</td>
<td>phenylalanine</td>
<td>R</td>
<td>arginine</td>
</tr>
<tr>
<td>G</td>
<td>glycine</td>
<td>S</td>
<td>serine</td>
</tr>
<tr>
<td>H</td>
<td>histidine</td>
<td>T</td>
<td>threonine</td>
</tr>
<tr>
<td>I</td>
<td>isoleucine</td>
<td>V</td>
<td>valine</td>
</tr>
<tr>
<td>K</td>
<td>lysine</td>
<td>W</td>
<td>tryptophan</td>
</tr>
<tr>
<td>L</td>
<td>leucine</td>
<td>Y</td>
<td>tyrosine</td>
</tr>
</tbody>
</table>
LIST OF ORIGINAL PUBLICATIONS


1 Equal contribution
Cells use a variety of adhesion structures to sense and attach to their environment. In this study I have examined a protein-protein interaction taking place in the cell’s attachment to the extracellular matrix. Integrins are cell membrane receptors present in all our cells, and they respond to mechanical cues from both the inside and outside of the cell. Integrins can get activated from the inside of the cell by the talin protein. Talin is a large, elongated protein that not only activates integrin, but also functions as a hub connecting integrin to the force-bearing actin machinery of the cell and feeding into different signaling pathways in the cell.

Integrins are a large family of heterodimeric receptors, while there are only two talins in human. This thesis focuses on the interaction of the talin-1 head domain with the cytoplasmic regions of integrin αIIbβ3 and a variety of β-integrins. The studies were conducted using mainly computational structural models and molecular dynamics simulations. Molecular dynamics simulations of individual talin-integrin complexes, in the presence or absence of a cell membrane model, provide a simplified yet powerful tool to analyze and predict the atomic details of a system, that would be very hard to dissect in cellulo or in vitro. As an example, we were able to analyze the atomic details of talin-integrin complexes under tension. While the molecular dynamics models presented in this thesis lack the complexity of an adhesion site in a living cell, ranging from local chemistry to integrin co-activators like kindlin, the simulations proved powerful in predicting the details of interactions within the talin head domain and between integrin and talin.
2 REVIEW OF THE LITERATURE

2.1 Cell-matrix adhesion

The cells of multicellular organisms require constant contact to other cells or to the extracellular matrix (ECM). The ECM is a complex network of secreted fibrous proteins, for example fibronectin or collagen. Cells use the ECM as a substrate to attach to, and they sense and respond to the stiffness of the ECM e.g. by adhering or migrating. Most cell-ECM contacts take place in adhesion structures involving integrins as the cell-matrix adhesion molecules. Two important proteins in the control of integrin activity are the cytoplasmic adaptors talin and kindlin.

2.1.1 Architecture of cell-matrix adhesions

Nascent adhesions are small, short-lived assemblies of active integrins in complex with talin and kindlin, formed at the leading edge of a migrating cell (Bachir et al., 2014). The integrins in nascent adhesions become transiently coupled to the actin cytoskeleton via talin, and some of the nascent adhesion sites mature further into focal adhesions at sites along bundled actin filaments. Focal adhesions are larger adhesion sites in the lamella rather than at the edge of lamellipodia. The structure of a mature focal adhesion has been elucidated in a study using interferometric photoactivated localization microscopy (Figure 1) (Kanchanawong et al., 2010): A focal adhesion site has a membrane-proximal signaling layer, which harbors tyrosine kinases such as focal adhesion kinase (FAK) and their substrates such as paxillin (Bellis, Miller and Turner, 1995). Further into the cell, it has a force transduction layer containing talin and vinculin, but also the ILK-pinch-parvin complex (Radovanac et al., 2013; Jansen, Atherton and Ballestrem, 2017), and on the membrane-distal side an actin-regulatory layer containing zyxin, α-actinin and vasodilator-simulated phosphoprotein (VASP) (Krause et al., 2003; Kanchanawong et al., 2010). Focal adhesions are connected to actomyosin stress fibers, and both tension and myosin II-generated contraction of the actin filaments are required for the growth of the focal adhesion (Burridge and Guilluy, 2016). Focal adhesions containing β1 integrin can translocate further towards the cell center and mature into fibrillar adhesions. Fibrillar adhesions primarily bind to fibronectin and contain actins and integrins interlinked with tensin (Pankov et al., 2000). Unlike focal adhesions, fibrillar adhesions are not connected to stress fibers and do not mediate high tension (Zamir et al., 2000).
2.1.2 Integrin clustering

Activation of integrins changes their affinity to extracellular ligands. Integrin activity appears to modulate also the affinity of integrin to cytoplasmic adaptors such as talin (Wang et al., 2018). Active integrins can become clustered into adhesion sites, and this clustering reinforces the linkage between the extracellular matrix and the cytoskeleton (Cluzel et al., 2005; Shattil, Kim and Ginsberg, 2010). How integrins become clustered is not fully clear. Integrin clustering requires the spacing between integrin ligands to be maximally 60 nm (Roca-Cusachs, Iskratsch and Sheetz, 2012). The initial clustering is reported to require integrin activation but little tension: Nascent adhesions with ~100 nm clusters of ~50 β3 integrins form at integrin activation, independent of substrate rigidity, but dependent of talin (Changede et al., 2015). Talin contains two β-integrin binding sites and forms dimers in the cell, hence a talin dimer might recruit up to four integrins (Klapholz and Brown, 2017). The linkage to actin filaments mediated by talin appears to stabilize the nascent adhesion sites (Comrie, Babich and Burkhardt, 2015). Moreover, the integrin coactivator kindlin has been indicated to promote integrin clustering (Ye et al., 2013). In the absence of traction, β3 integrin clusters disassemble and are recycled via clathrin-mediated endocytosis (Yu et al., 2015).
2.1.3 Phosphoinositides in cell adhesion

The eukaryotic cell membrane, or plasma membrane, is a phospholipid bilayer that envelops the cell and controls the traffic of particles between the cell and its environment. The composition of this membrane affects the mechanical properties of the cell. Lipids not only enclose the cell, but function also as signaling molecules. Phosphoinositides form only a small fraction of the total of lipids in the membranes of cells, but they have signaling roles in various processes at the cell membrane and in different organelles (reviewed in (Balla, 2013)). Phosphorylation of the inositol ring of phosphatidylinositol at sites 3, 4 and 5 generates lipids with different negative charges, binding specificities, and roles in cell signaling. For example, phosphatidylinositol 3-phosphate (PI(3)P) is found in endosomes and binds to the phox homology (PX) domains of e.g. sorting nexins, whereas PI(4)P is present mainly in Golgi membranes and binds e.g. to pleckstrin homology (PH) domains (Di Paolo and De Camilli, 2006). The major phosphoinositide in cell membranes is phosphatidylinositol 4,5-bisphosphate (PI(4,5)P2, henceforth referred to as PIP2). PIP2 serves as a substrate for phospholipase C, that cleaves it into the second messengers inositol trisphosphate and diacylglycerol, which in turn control the release of Ca\(^{2+}\) from the endoplasmic reticulum and the activity of protein kinase C, respectively. PIP2 is primarily produced from PI(4)P by PI5K1\(\gamma\) (PI5K1\(\gamma\)). Talin binds and activates PI5K1\(\gamma\) and can thereby increase the local concentration of PIP2 in the inner leaflet of the cell membrane at nascent adhesions (Di Paolo et al., 2002).

PIP2 promotes focal adhesion formation by activation and clustering of focal adhesion kinase (Goni et al., 2014), adhesion of the talin head domain to the membrane (Saltel et al., 2010; Moore et al., 2012) and opening of its inactivated form (Ye, McLean and Sligar, 2016). PIP2 binding to vinculin has been found to promote the protein’s oligomerization and activation (Palmer et al., 2009; Chinthalapudi et al., 2014). Kindlin binds different phosphoinositides via its pleckstrin homology (PH) domain and prefers PI(3,4,5)P3 (PIP3) lipids over PIP2, suggesting that its membrane localization is regulated by phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) activity (Liu et al., 2011; Qu et al., 2011).

2.2 Integrin

Humans have 24 different α-β integrin heterodimers, composed of a combination of 18 α and 8 β integrins. Integrins recognize their specific substrates in the extracellular matrix, adhere to the substrate and convey mechanical signals across the cell membrane. α and β integrins both have large, modular ectodomains, an α-helical single-pass transmembrane domain and a cytoplasmic domain with interaction sites for intracellular adaptor proteins. The majority of integrins connect fibers of the extracellular matrix, e.g. fibronectin, to actin in so-called focal
adhesions. In addition to focal adhesions, integrin α6β4 is found in hemidesmosomes, where it links laminins of the basal lamina to intermediate filaments within the cell. Moreover, the mucosal T lymphocyte integrin αEβ7 functions in cell-cell contacts by binding E-cadherins expressed on epithelial cells (Robinson et al., 2001).

This thesis focuses heavily on the platelet integrins αIIbβ3 and αVβ3 because of their role as model integrins in studies on integrin structure and activation. Major findings in the mechanisms of integrin activation have been gained by electron microscopy studies of αIIbβ3 integrin (Ye et al., 2010; Eng et al., 2011; Choi et al., 2013; Dai et al., 2015). These studies show that αIIbβ3 integrins undergo conformational changes as a response to inside-out or outside-in signaling events: resting, inactive integrin ectodomains reside in a bent, closed conformation, but are extended and open in their active form. αIIbβ3 is expressed exclusively in platelets, with around 80,000 molecules per cell (Wagner et al., 1996). Inactive αIIbβ3 integrins on rolling platelets bind fibrinogen, a blood glycoprotein that promotes vascular repair, both by being cleaved into fibrin to form a blood clot and by interaction with αIIbβ3. αIIbβ3-bound fibrinogen then links platelets together, and shear stress on the platelets finally outside-in activates αIIbβ3 integrins, as reviewed in (Sorrentino et al., 2015).

Recent results suggest that the picture of the conformational regulation of integrins, derived from αIIbβ3 integrin, may be an oversimplification. While the precise control of β3 integrin activation is critical to avoid bleeding or thrombosis (Estevez, Shen and Du, 2014), studies suggest β1 integrins to be constitutively extended (Zhang et al., 2008; Lu et al., 2016; Su et al., 2016; Miyazaki, Iwasaki and Takagi, 2018). Electron microscopy of the ectodomains of αVβ3, α2β1, α3β1, α5β1, α6β1 and α6β4 integrins showed that αVβ3 has a distinct conformational space: αVβ3 resided in a bent and closed conformation and was strongly activated by Mn2+, whereas the other integrins resided mostly in intermediate states between the bent-closed and extended-open conformations (Miyazaki, Iwasaki and Takagi, 2018). The integrin family members are also functionally different: for example, αVβ3 integrins are located in peripheral focal adhesions and required for mechanotransduction, whereas α5β1 integrins are required for cell spreading and traction force generation (Roca-Cusachs et al., 2009; Lin et al., 2013; Schaufler et al., 2016). In a similar fashion, a knock-out study showed that epithelial cells can bind collagen via α2β1 integrins but αV integrins are critical for talin recruitment to adhesion sites and adhesion maturation (Teräväinen et al., 2013).

2.2.1 Integrin ectodomains

α-integrins have an ectodomain composed of a β-propeller, thigh, calf-1 and calf-2 domains, a transmembrane helix and a cytoplasmic tail. The integrins α1, α2, α10, α11, αL, αM, αX, αD, and αE contain also an α1 domain insert in the β-propeller domain, as reviewed in (Barczyk, Carracedo and Gullberg, 2009). The β-integrin headpiece has a ligand-binding βI domain (a von Willebrand Factor A domain) inserted into a hybrid domain, which in turn is inserted into a PSI domain (a domain found in plexins, semaphorins and integrins). The leg of β
integrin has four cysteine-rich epidermal growth factor-like (EGF) domains and a $\beta$-tail domain. The ligand-binding site is formed at the $\beta$-propeller-$\beta$I-domain interface, except in $\alpha$I domain-containing integrins, which bind their ligands via the $\alpha$I domain.

Integrins have a Mg$^{2+}$ binding metal ion-dependent adhesion site (MIDAS) composed of the I-domains of $\alpha$ and $\beta$ integrins (Lee et al., 1995). The Mg$^{2+}$ ion is coordinated by an acidic side chain from the ligand, e.g. RGD (arginine-glycine-aspartate), along with acidic residues from both integrin subunits (Xiong et al., 2009). In integrins lacking the $\alpha$I domain, MIDAS is located within the $\beta$I domain. MIDAS is flanked by an adjacent metal ion-dependent adhesion site (ADMIDAS), that, when bound to a Ca$^{2+}$ ion, inhibits integrin activation, and a synergistic metal ion-binding site (SyMBS), that also binds a divalent cation (Barczyk, Carracedo and Gullberg, 2009). The MIDAS ion is coordinated by an acidic side chain of the ligand (Xiong et al., 2001, 2002; Zhang and Chen, 2012), and in $\alpha$5$\beta$1 it prefers a Mg$^{2+}$ ion over Ca$^{2+}$ (Zhu et al., 2010). ADMIDAS and SyMBS preferentially bind Ca$^{2+}$ (Zhu et al., 2010).

2.2.2 Integrin transmembrane domains

The transmembrane domains of an integrin dimer in the inactive conformation are aligned in parallel. The hydrophobic transmembrane segment of $\beta$-integrin is longer than in $\alpha$-integrin, and studies suggest that $\alpha$-integrin passes perpendicular through the membrane, while $\beta$-integrin does at a 25°-30° angle (Lau et al., 2009; Kalli et al., 2011). The $\alpha$ and $\beta$ transmembrane domains cross each other in a right-handed orientation, and their interactions are stabilized by glycine motifs in both subunits. In the case of $\alpha$IIb$\beta$3 integrin, the amino acids $\alpha$IIb G972 and G976 and $\beta$3 G708 form the core of the outer membrane clasp (Lau et al., 2009). An inner membrane clasp is formed near the cytoplasmic border of the membrane by hydrophobic interactions involving the $\alpha$IIb residues F992 and F993, $\beta$3 residues W715 and I719, while $\alpha$IIb R995 and $\beta$3 D723 (Lau et al., 2009) build an interdomain salt bridge. Mutations disrupting the helix packing at the OMC or the contacts at the IMC have shown these motifs to be critical for keeping integrin inactive (Hughes et al., 1996; Luo et al., 2005; Lau et al., 2009; Berger et al., 2010). Moreover, the mutation D723H in $\beta$3 and the mutations R995W and R995Q in $\alpha$IIb have been associated with the bleeding disorder platelet-type 16 (BDPLT16, autosomal dominant Glanzmann thrombasthenia), where the $\alpha$IIb$\beta$3 integrin is rendered constitutively active or partially active (Peyruchaud et al., 1998; Ghevaert et al., 2008; Kunishima et al., 2011). The D723H mutation impairs proplatelet formation and promotes the formation microtubule-driven protrusions, possibly by downregulating the small GTPase RhoA (Ras homolog gene family member A) (Ghevaert et al., 2008; Schaffner-Reckinger et al., 2009).

The transmembrane helix tilt and the activation of $\beta$3 integrin have been reported as being controlled by the K716 residue in the $\beta$3 integrin IMC. While the K716 is buried within the hydrophobic core of the cell membrane, its positively charged side chain amino group is thought to “snorkel” amongst the polar head groups of membrane phospholipids, thus
orienting the integrin transmembrane helices to keep the integrin inactive (Lau et al., 2008; Kim et al., 2011, 2012).

### 2.2.3 β-integrin cytoplasmic domains

The cytoplasmic domains of β integrins bind various cytoplasmic adaptor proteins. All β integrins, with the exception of β4 and β8, have a short cytoplasmic tail containing an NPxY and an NxxY motif (Figure 2). The membrane-proximal NPxY motif binds to the focal adhesion adaptor protein talin, and the membrane-distal NxxY motif binds to kindlin, another focal adhesion adaptor (Calderwood et al., 1999; Ma et al., 2008; Montanez et al., 2008; Moser et al., 2008). The β1A tail interacts with the phosphotyrosine-binding-like (PTB) domain of talin, epidermal growth factor receptor kinase substrate 8 (EPS8) and disabled homolog 1 (Dab1); the β2 tail with docking protein 1 (Dok1), and talin; the β3 tail with protein numb homolog (Numb), Dab1, disabled homolog 2 (Dab2), EPS8, tensin, Dok1, and talin; the β5 tail with Numb, Dab1, Dab2, EPS8, tensin, Dok1, and talin; and the β7 tail with tensin, Dok1, and talin (Calderwood et al., 2003). Moreover, alternative splicing gives rise to at least 5 different β1 integrins and 3 different β3 integrins, all with different cytoplasmic domains.

Integrin activity is controlled by phosphorylation of the NPxY tyrosine. Src tyrosine kinases, present in the cell and in Rous sarcoma virus, bind to the C-terminus of β3 integrin (Xiao et al., 2013) and phosphorylate at least β1 (Hirst et al., 1986) and β3 integrins (Tapley et al., 1989). Phosphorylation of integrin β increases Dok1 binding and decreases talin binding to integrins (Anthis, Haling, et al., 2009).

![Figure 2](image.png)

**Figure 2.** The cytoplasmic domains of β-integrins, except β4 and β8, have a conserved sequence with potentially talin-binding motifs. Cytoplasmic tails of talin-binding β integrins aligned with Muscle 3.8 (Edgar, 2004). The NPxY/NxxY sites are marked with a red bar above the alignment. Integrin β3 residues contributing to the inner membrane clasp (Lau et al., 2009) are marked with blue color. The talin-binding sequences characterized in the crystallographic structure of the β3-talin-1 complex (García-Alvarez et al., 2003) and the β1D-talin-2 complex (Anthis, Wegener, et al., 2009) are highlighted in green.
2.2.4 Integrin activation

Cryogenic electron microscopy of the platelet integrin αIIbβ3 embedded in a lipid nanodisc has shown it to exist in four main states: bent, two intermediate states, and a fully upright state (Xu et al., 2016). All states are found independently of talin, but the presence of talin and extracellular ligands shifts the balance of the four states towards the activated, upright state (Xu et al., 2016). In the active conformation, the α and β transmembrane domains are separated at a distance of 8 nm (Xu et al., 2016). Integrin activation has two major steps: leg extension and binding site opening, resulting in an extended-open conformation that binds ligands with a thousand-fold affinity compared to the bent conformation (Zhu, Zhu and Springer, 2013; Li and Springer, 2017) (for a simplified view of integrin activation, please see Figure 3).

![Figure 3.](image.png)

Integrins have an active, extended conformation and an inactive, bent conformation. Several intermediate conformations have been described, but they are omitted here for clarity. Left: inactive integrin. Right: active, extended integrin bound to an extracellular ligand. Blue: α-integrin. Green: β-integrin. Cyan: extracellular matrix protein.

Talin activates integrin in so-called inside-out activation, and was long thought to be the only activator of integrin (Calderwood et al., 1999). However, more recent findings have shown integrin activation to be controlled also by kindlin (Ma et al., 2008; Moser et al., 2008) and force generated in the actin network (Nordenfelt, Elliott and Springer, 2016; Li and Springer, 2017). Intermolecular interactions by integrin homomeric or heteromeric oligomerization have also been suggested to drive integrin activation during integrin clustering (Li et al., 2003; Ye, Kim and Kim, 2014). As visualized by electron microscopy, the talin head alone cannot completely activate αIIbβ3 integrin in a purified system with lipid nanodiscs: the majority of the integrin molecules are in an inactive, bent conformation, and the talin head only increases the number of extended integrin molecules to 22%. Additionally, it may cause smaller changes in the conformation of inactive, bent integrins (Ye et al., 2010). Importantly, these conformational changes do not require kindlin (Ye et al., 2010).
Outside-in activation of integrin takes place when the ectodomains of integrin bind to an extracellular ligand and tensile force from the bound ligand extends the ectodomains (Chen et al., 2017). Alternatively, outside-in signaling can take place via the G protein-mediated pathway, where the trimeric G protein Gz13 binds to the ExE sequence in the membrane-proximal helix of β-integrin (Shen et al., 2013) and to talin F3 (Srinivasan et al., 2015; Schiemer et al., 2016). Gz13 binding to the integrin β tail activates Src (Gong et al., 2010), which phosphorylates β3 NPxY and NxxY tyrosines. Gz13 binding also negatively regulates the activation of RhoA (Gong et al., 2010), that promotes actin polymerization and actomyosin contraction as reviewed in (Shen, Delaney and Du, 2012) and is critical to directed cell migration (Goulimari et al., 2005). Integrins can also be artificially activated outside-in with manganese ions: Mn$^{2+}$ replaces the Ca$^{2+}$ ion of the β-integrin ADMIDAS site and increases the fibronectin-binding affinity of integrin αIIbβ3 (Kirchhofer et al., 1990). However, crystallographic (Xiong et al., 2002) and electron microscopy studies of αIIbβ3 and αVβ3 integrins (Adair et al., 2005; Ye et al., 2008; Dai et al., 2015) show that the addition of manganese does not cause the integrin ectodomains to swing out; instead manganese-treated integrins remain in a bent conformation.

2.3 Talin

Humans have two talins, talin-1 and talin-2, both composed of a head and a rod (Figure 4). Talin was first purified from chicken gizzard smooth muscle and initially named 215K due to its apparent molecular weight in electrophoresis: the molecular weight of chicken talin 1-2541 is 272 kDa, and the rod alone is 222 kDa. It later got the name talin, from Latin “talus” (ankle), due to its location between actin filaments and the cell’s “feet”, the adhesion sites (Burridge and Connell, 1983). Talin-1 is ubiquitously expressed and critical to integrin activation (Nieswandt et al., 2007; Petrich et al., 2007), and the name “talin” is often used synonymously to talin-1. Human talin-1 and talin-2 are 2541 and 2542 residues in length, respectively, and 76% identical. On gene level, talin-1 and talin-2 differ more: the talin-1 gene is only approximately 30 kb in length, whereas the ancestral talin-2 gene contains multiple introns and exceeds 200 kb in length (Monkley, Pritchard and Critchley, 2001). No splicing variants of talin-1 are known, whereas talin-2 has 4 different forms (Debrand et al., 2009).
Talin is thought to have emerged in amoebozoans, while an integrin adhesome may have developed first in opisthokonts and apusozoans (Senetar and McCann, 2005; Sebé-Pedrós et al., 2010). The amoeba Dictyostelium discoideum has two talin genes, talA and talB, with overlapping functions in substrate adhesion, motility and cytokinesis, and also possesses β-integrin-like proteins (sib) (Cornillon et al., 2006), but no known α-integrins. Talins are found also in some Fungi such as the chytrid Spizellomyces punctatus (Sebé-Pedrós et al., 2010). The only talin gene of the fruit fly Drosophila melanogaster (rhea) is critical for integrin clustering and integrin-cytoskeleton coupling (Brown et al., 2002), and in the nematode Caenorhabditis elegans the loss of talin results in paralysis and defects in gonad formation (Cram, Clark and Schwarzbauer, 2003). The talin gene has duplicated early in chordates, resulting in two talins present in mammals and three copies in fish as a result of further duplication (Senetar and McCann, 2005). In zebrafish, talin-1 has been shown to be critical for cardiac Z-disc stabilization (Wu et al., 2015). In mouse, talin-1 is critical for embryonic development (Monkley et al., 2000), whereas talin-2 knockout mice have only a mild dystrophic phenotype (Debrand et al., 2012; Calderwood, Campbell and Critchley, 2013).

The talin rod contains 62 α-helices divided into 13 bundles of 4-5 helices each and a C-terminal dimerization helix. While the head domain functions mainly as the integrin activator,
the rod connects in the integrin-bound talin to the force-generating actin network. The talin rod has two actin-binding sites, ABS2 and ABS3. ABS2 is composed of rod bundles R4-R8 and is required for focal adhesion maturation (Hemmings et al., 1996; Atherton et al., 2015). ABS3 is also known as the I/LWEQ domain and is formed by the R13 bundle and the dimerization helix. Talin dimerization via the dimerization helix is essential to ABS3 actin binding (Gingras et al., 2008). Structural studies of talin have elucidated the structures of all talin domains (Gingras et al., 2005, 2006, 2008, 2009, 2010; Elliott et al., 2010; Goul, Bouaouina, et al., 2010; Goul, Gingras, et al., 2010).

2.3.1 The talin head FERM domain

The talin head (residues 1-433) contains a FERM domain (residues 85-405) (Figure 5). FERM (band 4.1R, ezrin, radixin, moesin) domains are characterized by a ubiquitin-like F1 domain, an acetyl-CoA binding-like α-helical F2 domain and a phosphotyrosine-binding (PTB) F3 domain. FERM domains are found in a variety of membrane-cytoskeletal interacting proteins, and currently approximately 50 members of the family are known (Frame et al., 2010). FERM domains emerged in unikonts in the form of talin and myosin, and more FERM domain proteins appeared later in metazoans (Ali and Khan, 2014). FERM-containing proteins can be divided into 3 main categories: 1) classical ERMs (ezrin, radixin, moesin, merlin) and kinases, 2) kindlins and talins, and 3) myosins and KRIT1 (Krev interaction trapped protein 1) (Frame et al., 2010). The ERM proteins – ezrin, radixin and moesin – bind actin and connect it to the cell membrane, thus controlling e.g. cell shape, polarization, migration and division (Fehon, McClatchey and Bretscher, 2010; Lallemand and Arpin, 2010). Moesin (membrane-organizing extension spike protein) is found in actin protrusions, e.g. filopodia, and controls actin assembly at cell membrane edges. The C-terminal domain of ERM proteins binds to their FERM domain, blocking the binding to actin and membrane proteins (Pearson et al., 2000). The autoinhibition is released by PIP2 binding and consequential phosphorylation (Fievet et al., 2004; Pelaseyed et al., 2017). FERM proteins share a cloverleaf-like three-lobe structure with conserved contacts in the F1-F2, F2-F3 and F1-F3 interfaces. In addition to the FERM core, inserts and extra domains exist in many FERM proteins.
In comparison to typical FERM domains, the talin head has an additional ubiquitin-like F0 domain, to which the F1 is connected by a very short linker (Gould, Bouaouina, et al., 2010) (Figure 6). In addition, talin contains a large loop insert within the F1 domain. In FERM evolution, talin belongs to the same branch with kindlins (kindlin-1, kindlin-2, kindlin-3). Both talins and kindlins have an F0 domain, but kindlin’s F1 loop insert is larger, and a pleckstrin homology (PH) domain has been inserted into the kindlin F2 domain. To date, there is no complete 3-D structure of either talin or kindlin available. However, a F1 loop-truncated form of talin-1 has been shown by crystallography to form an atypical FERM domain with the individual domains F0, F1, F2 and F3 arranged in an extended outlay (Elliott et al., 2010; Chinthalapudi, Rangarajan and Izard, 2018), whereas an F1-loop- and PH-domain-deleted kindlin-2 has been crystallized in a FERM shape (Li et al., 2017). Thus the talin head conformation appears to differ from that of its relatives. While the crystallization of the loop-truncated talin head produced a rather extended conformation, small-angle X-ray scattering data as well as molecular dynamics simulations have suggested the linker between domains F1 and F2 to function as a hinge that allows the structure to bend into a V shape (Elliott et al., 2010; Kalli, Campbell and Sansom, 2013).
Figure 6. Comparison of the talin head FERM and moesin. FERM proteins, such as moesin, possess the domains F1, F2 and F3 (PDB ID: 1EF1 (Pearson et al., 2000)). For talin, the NMR structures of F0 (PDB ID: 2KC1 (Goult, Bouaouina, et al., 2010)), F1 (PDB ID: 2KC2 (Goult, Bouaouina, et al., 2010)) and F3 (PDB ID: 2H7D (Wegener et al., 2007)) and the crystal structure of F2 (PDB ID: 3IVF (Elliott et al., 2010)) are shown. Both talins and kindlins contain an additional N-terminal F0 domain. The sequence alignment of talin (top row) and moesin (bottom row) shows the talin secondary structure as determined with the DSSP algorithm (Kabsch and Sander, 1983) in the Protein Data Bank (Berman et al., 2000). Blue arrows: $\beta$-strands. Red bars: $\alpha$-helices.

The talin F3 domain is a phosphotyrosine-binding (PTB)-like as well as a PH-like domain. PH domains bind phosphoinositides mainly by residues in the $\beta$1-$\beta$4 strands (Martin, 1998) (Figure 7). Similar to PH-domain L1,2 loops, the talin L1,2 loop is rich in basic residues. Like typical PTB domains, talin F3 binds NPxY motif-containing peptides. It specifically binds the membrane-proximal NPxY motif in the cytoplasmic tails of $\beta$1, $\beta$3, $\beta$5, $\beta$6, and $\beta$7 integrins; however, it prefers the nonphosphorylated form of the NPxY motif (Oxley et al., 2008; Anthis, Haling, et al., 2009). In addition to the conserved PTB-like binding site for integrin, talin F3 has also a membrane-proximal binding site that is required for integrin activation (Patil et al., 1999; Wegener et al., 2007). Studies suggest that residues in the membrane proximal region may directly interfere with the integrin heterodimer (Wegener et al., 2007; Anthis, Wegener, et al., 2009; Saltel et al., 2009); however, these residues may also regulate integrin activity through autoinhibition by talin head-rod binding (Banno et al., 2012; Zhang et al., 2016). In addition to the PTB domain, the talin F2 domain may bind $\beta$3 integrin (Calderwood et al., 2002), and a secondary integrin binding site (IBS2) has been identified in the rod domain bundles R11 and R12 (Tremuth et al., 2004). The IBS2 site is not involved in integrin activation, but it may function in linkage of integrins to the cytoskeleton (Moes et al., 2007). Its access to integrins is controlled by the actomyosin machinery orienting the talin rod domains (Klapholz et al., 2015).
2.3.2 Talin-integrin interface

Earlier studies have found talin to bind the cytoplasmic tail of canonical β1 (β1A), the splicing variant β1D, β2, β3, β5, and β7. Moreover, studies conflict on talin binding to αIIb integrin (Calderwood et al., 2003; García-Alvarez et al., 2003; Goult, Bououina, et al., 2010).

The talin-β3 integrin complex has been crystallized as a fusion protein (García-Alvarez et al., 2003), apparently to overcome problems in the production of stable talin-integrin complexes. In this structure, the 11 β3 residues forming its membrane-distal talin-binding motif (WDTANNPLYKE) and containing the 1st (membrane-proximal) NPxY have been fused to the N-terminus of a talin F2F3 double-domain. The asymmetric unit of this crystal structure contains two antiparallelly aligned integrin-talin fusions, where the integrin peptide of fusion protein 1 is bound to the talin F3 domain of fusion protein 2, and vice versa. This structure (PDB ID: 1MK7, Figure 8C) is discussed in detail in Article I of this thesis.
Figure 8. The three currently available 3-D structures of talin-β-integrin complexes show slightly different binding conformations for the bound integrin peptide. The structures in panels B and D show both the membrane-proximal and membrane-distal binding sites, whereas the integrin peptide in panel C contains only the membrane-distal stretch. A) Superimposition of 3 talin-integrin binding structures presented in B (white integrin), C (dark green integrin) and D (light green integrin). B) Crystal structure of talin-2 F2F3 domains in complex with integrin β1D cytoplasmic tail, solved at a 2.165 Å resolution (PDB ID: 3G9W (Anthis, Wegener, et al., 2009)). C) Crystal structure of talin-1 F2F3 in complex with integrin β3, produced as an integrin-talin fusion protein and determined at 2.2 Å resolution (PDB ID: 1MK7 (García-Alvarez et al., 2003)). D) NMR structure of talin F3 in complex with an integrin β3 (green)-PIPKIγ (cyan) hybrid peptide (PDB ID: 2H7E (Wegener et al., 2007)). The view is slightly rotated around the X-axis between the superimposed image (A) and the images of individual talin-integrin complexes (B, C, D).

The second talin-integrin complex with solved 3-D structure (Wegener et al., 2007) (Figure 8D) makes use of the fact that PIPKIγ binds to talin with higher affinity than integrins (de Pereda et al., 2005). An integrin-PIPKIγ hybrid peptide was complexed with the talin F3 domain, and the structure was determined with nuclear magnetic resonance (NMR) spectroscopy. The hybrid peptide contains the integrin β3 inner membrane clasp residues, the membrane-proximal helix and part of the membrane-distal talin recognition site (K716LLITIHDRKFEAKFEERARAKW739) followed by part of the mouse PIPKIγ talin recognition site (V643pYSPLHYSA652, where pY is a phosphotyrosine). Combining the higher-affinity PIPKIγ talin-binding motif with the membrane-proximal helix of integrin β3, Wegener et al. showed how the membrane-proximal and membrane-distal binding motifs cooperate to bind talin and identified key residues that contribute to the binding of the membrane-proximal helix (Wegener et al., 2007). For example, two phenylalanine side chains...
in the MP-helix become buried in hydrophobic pockets formed by the L1,2 residues K318, N323 and L325 (F727) and between β-strands 1 and 7, formed by residues K318, L325, S379 and Q381 (F730).

Talin-2 and β1D integrin form a complex with high enough affinity to be co-crystallized as done by Anthis et al. (Anthis, Wegener, et al., 2009) (Figure 8B). This structure contains an integrin peptide with inner membrane clasp residues, the membrane-proximal helix and the membrane-distal talin recognition site, which was not produced as a fusion or hybrid sequence.

### 2.3.3 Talin regulation

Talin activity is controlled at many levels from conformational changes to interaction with adaptors and signaling proteins. In talin autoinhibition, the R9 domain blocks membrane-interaction sites in the talin head F3 domain (Goksoy et al., 2008), possibly by formation of a double-donut shape, where the whole two rod domains of a dimeric talin become wrapped around the heads (Goult, Xu, et al., 2013). Release of the R9 domain enables talin binding to the barbed end of an actin fiber via its ABS1 site in talin head F2 and F3, and this in turn is suggested to block actin fiber elongation (Ciobanasu et al., 2018). Accumulation of PIP2 to the adhesion site would then release the actin and allow binding to the cell membrane (Ciobanasu et al., 2018). RIAM (Rap1 interacting adhesion molecule) promotes integrin activation, but is dispensable in the process (Lee et al., 2009; Stritt et al., 2015), and its role is to release the head-rod interaction: it binds both to the head (Yang et al., 2014) and to rod R8 and R3 (Chang et al., 2014). RIAM’s binding site in the talin head lies in the F3 domain (Yang et al., 2014). RIAM has a Ras association domain binding to the small GTPase Ras-related protein 1 (Rap1), which takes RIAM to the membrane (Lee et al., 2009). Rap1 also binds the talin F0 domain in a GTP-dependent manner to recruit talin to the membrane (Zhu et al., 2017). Gα13 is another talin activator that releases the R9-F3 contact by binding to the N-terminal part of the F3 domain (Schiemer et al., 2016).

Talin is also regulated through proteolytic modifications. Proteolytic cleavage by calpain II between glutamine residues 433 and 434 in the neck domain promotes focal adhesion disassembly (Fox et al., 1985; Franco et al., 2004). Proteolysis at a secondary calpain II cleavage site between the R13 bundle and the C-terminal dimerization helix has been proposed to promote FAK-dependent cell cycle progression and focal adhesion turnover (Bate et al., 2012). Moreover, calpain II cleavage between residues P1902 and A1903 in the rod bundle R10 and subsequent arginylation at A1903 of the C-terminal 70-kDa fragment have been suggested to take place in cadherin-mediated cell adhesion (Zhang, Saha and Kashina, 2012). Cathepsin H cleaves the talin head F0 domain and decreases αVβ3 activation, possibly promoting focal adhesion disassembly and cell migration (Jevnikar et al., 2013). Finally, posttranslational modifications regulate the subcellular location, conformation, and function of talin (Loh and Su, 2016). Phosphorylation of S425 by Cdk5 promotes focal adhesion
disassembly and increases β1 activation (Huang et al., 2009; Jin et al., 2015). S446 can be phosphorylate, possibly by protein kinase A (Ratnikov et al., 2005), and controls the calpain cleavage of the talin head (Y. Li et al., 2016). Threonine phosphorylations in the F1 domain loop appear to have a contrary effect: a combination of T144A and T150A reportedly increases the calpain-mediated cleavage of talin and focal adhesion disassembly (Y. Li et al., 2016).

Figure 9. Talin has several potential binding partners in integrin-mediated adhesion. Known talin-binding proteins are shown at their approximate binding sites in talin (yellow). All interactions may not be present at the same time. Actin (blue filaments) bind to ABS2 (R4-R8) and ABS3 (R13 and the dimerization helix). The talin-actin linkage is strengthened as a response to mechanical force (red arrows) by the increased binding of vinculin (turquoise), which has multiple binding sites along the talin rod. Several binding sites have been indicated for RIAM, marked with arrows. Calpain II proteolytically cleaves both the head-rod linker and the linker connecting R13 and the dimerization helix; the approximate cleavage sites are marked with scissors. Talin binding to integrin (green and blue transmembrane receptor) is strengthened by PIP2 lipids (red lipids in the membrane) binding to positively charged patches in the talin head. An additional β3 integrin (transparent green) marks the secondary integrin binding site (IBS2) on the talin rod, although integrin binding to IBS2 would require the talin rod to be aligned along the membrane plane and would not be possible in the configuration presented in the figure.
While the primary function of talin appears to be to link integrin to actin, several other proteins have binding sites in talin as well, and even a secondary binding site for integrins $\beta_3$ and $\beta_1$ (IBS2) has been located in the rod bundles R11 and R12 (Tremuth et al., 2004; Rodius et al., 2008; Gingras et al., 2009). Vinculin is among the more studied talin binders, and functions to enforce the linkage between talin and actin. Exposure to mechanical force uncovers buried vinculin binding sites in the talin rod (Hytonen and Vogel, 2008; del Rio et al., 2009). Talin rod binds also KANK1 (KN motif and ankyrin repeat domain-containing protein 1) in R7 (Bouchet et al., 2016), DLC1 (Deleted in liver cancer 1) in R8 (Li et al., 2011) and synemin in R7 and R8 (Sun et al., 2008), RIAM (Rap1-GTP-interacting adapter molecule) in R2-R3, R8 and R11 (Goult, Zacharchenko, et al., 2013), and moesin in R11-R13 (Beaty et al., 2014). The talin head binds at least $\beta$-integrins, PIP5K1$\gamma$ (de Pereda et al., 2005), RIAM (Yang et al., 2014), layilin (Borowsky and Hynes, 1998; Wegener et al., 2008), actin (Lee et al., 2004; Ciobanasu et al., 2018), G$\alpha_13$ (guanine nucleotide-binding protein subunit $\alpha_13$) (Schiemer et al., 2016), TIAM1 (T-lymphoma invasion and metastasis-inducing protein 1) (Wang et al., 2012), Rap1b (Ras-related protein Rap1b) (Plak et al., 2016; Zhu et al., 2017), paxillin (Gao et al., 2017) and focal adhesion kinase (Lawson et al., 2012). Some of talin’s binding partners are presented in Figure 9.

Talin interacts with lipids, and PIP2 lipids have been indicated open a binding interface between the talin rod and head and to make the rod accessible for proteases (Martel et al., 2001; Goksoy et al., 2008; Song et al., 2012). Specific membrane interaction patches have been identified in talin head domains F3 and F2 (Anthis, Wegener, et al., 2009; Saltel et al., 2009; Arcario and Tajkhorshid, 2014). Moreover, the talin F1 loop contributes to lipid binding (Goult, Bouaouina, et al., 2010). The F1 loop residues 144-152 and 154-167 have $\alpha$-helical propensity, and the helix has been suggested to function as a fly-cast mechanism to drag the talin head to the membrane (Goult, Bouaouina, et al., 2010).

### 2.4 Characterization methods for protein structures

X-ray crystallography has gained a role as the standard method of resolving protein structure, and its popularity continues. With X-rays discovered already in 1895 by Wilhelm Röntgen and the first biological macromolecule, jack bean urease, crystallized in 1926 by James Sumner, research on using X-rays to solve the structures of biological macromolecules began in the 1940’s (Shi, 2014). This led to the discovery of the DNA double helix by fiber diffraction (Watson and Crick, 1953) and the first protein crystal structure of myoglobin at 6 Å resolution (Kendrew et al., 1958). In 1971 the Protein Data Bank (Berman et al., 2000) was founded to store crystal structures of macromolecules, and still to date the majority of its over 130,000 structures have been solved by X-ray crystallography (Table 1).

In X-ray crystallography, the sample is first crystallized, and the large, symmetric crystal (up to millimeters in diameter) is exposed to radiation from a strong X-ray source. The incident beam of X-rays scatters from the crystalline sample. The outgoing X-rays have the
same wavelength and energy as the incoming X-rays, but a different direction. The molecular structure is calculated using the angles and intensities of the diffracted X-rays. Recent advances in protein crystallography include the use of X-ray free electron lasers (XFEL). This method utilizes a very bright electron source and very short exposure times, minimizing radiation damage. Instead of a single crystal, the sample is provided as 25-500 nm nanocrystals that are streamed to the laser beam (Bogan, 2013).

Nuclear magnetic resonance (NMR) spectroscopy is a group of methods based on the magnetic properties of atoms exposed to a strong magnetic field. The protein sample is typically labeled with at least one isotope with non-zero spin, e.g. $^{15}$N or $^{13}$C. When exposed to a magnetic field, the atoms will change their spin and upon removal from the magnetic field return to their low-energy state. Atoms residing close to each other will resonate, resulting in a shift in their detected spectra. Atoms can correlate either via chemical bonds (e.g. the $^{15}$N and $^1$H atoms of an amide group) by polarization of the bond electrons, or through the space between hydrogen nuclei (Wider, 2000). NMR has been utilized in protein structural research from the 1950’s (Saunders, Wishnia and Kirkwood, 1957), although the first protein structure was solved by NMR decades later (Braun et al., 1981).

Transmission electron microscopy using negative staining or electron cryomicroscopy (cryo-EM) has a lower resolution than X-ray crystallography. A major benefit of electron microscopy is that it does not require crystallization of the sample. Transmission electron microscopy images are captured with a charge-coupled device camera or a direct electron detector. In cryo-EM, the sample is rapidly frozen from a liquid state, yielding a near-native conformation. Negative staining increases the contrast by the introduction of heavy metals, and thereby electrons, to the sample, which is then dried before imaging. Uranyl acetate for example is widely used in negative staining, although the solution is acidic and can flatten the structure (Mast and Demeestere, 2009; De Carlo and Harris, 2011). The effect of ammonium molybdate to the sample can be milder than uranyl acetate, but it also provides less contrast (De Carlo and Harris, 2011). Some protein structural methods are compared in Table 1.
### Table 1. Comparison of protein structural characterization methods. The list is not exhaustive.

<table>
<thead>
<tr>
<th>Method</th>
<th>Protein sample requirements</th>
<th>Resolution limit</th>
<th>Sample preparation</th>
<th>Challenges or sources of error</th>
<th>Structures in public databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray crystallography</td>
<td>structured protein, must form crystals</td>
<td>atomistic, no hydrogens</td>
<td>crystallization</td>
<td>crystallization artefacts, misinterpretation of electron density</td>
<td>125846 a,d</td>
</tr>
<tr>
<td>NMR spectroscopy</td>
<td>small particle size (&lt; 50 kDa)</td>
<td>atomistic</td>
<td>isotope labeling in protein production</td>
<td>high protein concentration required</td>
<td>12215 a,d</td>
</tr>
<tr>
<td>Electron microscopy (EM)</td>
<td>more suitable for large particles</td>
<td>negative staining: ~1 nm b cryo-EM: 2 Å c</td>
<td>sample in buffer</td>
<td>negative staining may flatten the structure</td>
<td>2139 a,d</td>
</tr>
<tr>
<td>Small-angle X-ray scattering</td>
<td>monodisperse sample</td>
<td>~1 nm</td>
<td>sample in buffer</td>
<td>oligomerization, interparticle interaction, radiation damage</td>
<td>1075 models c</td>
</tr>
<tr>
<td>Molecular dynamics simulations</td>
<td>structure or model available</td>
<td>atomistic</td>
<td>building a model system for simulation</td>
<td>errors in model or force field</td>
<td>no universally adopted database</td>
</tr>
</tbody>
</table>

---

a) As of 24 May 2018.
b) (Rames, Yu and Ren, 2014)
c) (Vonck and Mills, 2017)
d) Protein Data Bank: [www.rcsb.org](http://www.rcsb.org) (Berman et al., 2000)
e) Small Angle Scattering Biological Data Bank: [www.sasbdb.org](http://www.sasbdb.org) (Valentini et al., 2015)

### 2.4.1 Small-angle X-ray scattering

Structures of proteins too large for NMR and too flexible to form symmetric crystals can be determined by low-resolution methods like electron microscopy or small-angle scattering (SAS). For small-angle scattering, one can utilize neutrons (SANS) or X-rays (SAXS). While X-ray crystallography provides atom-resolution data, and NMR can even tell the positions of hydrogens, SAXS only gives an approximate shape of the molecule. SAXS does not require the formation of crystals and no labeling is involved, which makes it a relatively inexpensive and straightforward method. SAXS is suitable for particles in the size range of one kilodalton to hundreds of megadaltons (Mertens and Svergun, 2010).

In SAXS the unlabeled protein sample flows through a capillary. A collimated, monochromatic X-ray beam passes through the capillary and scatters from the protein in the sample. The scattered intensity is recorded as 2-D images on a detector (Figure 10). As the individual particles in the sample can be in any orientation, they scatter the incident X-rays in all directions. The intensity of the scattering is hence radially averaged to yield a scattering profile $I(s)$ per $s$, where $\lambda$ is the wavelength of the X-ray beam, $s$ is the momentum transfer...
\[ s = \frac{4\pi \sin \theta}{\lambda}, \] often expressed in Å\(^{-1}\) or nm\(^{-1}\), and \( \theta \) is the scattering angle between the incident X-ray and the point of observation. Scattering from several, perhaps tens of 1-s exposures, are averaged to yield a scattering profile. To inspect the scattering of the protein in the sample, the scattering profile of buffer alone is subtracted. SAXS can be carried out using an X-ray generator with \( \lambda \) in the range of 0.1-0.2 nm or a synchrotron with \( \lambda \) in the range of 0.03-0.4 nm. Small-angle neutron scattering (SANS) with thermal neutrons can have \( \lambda \) in the range of 0.1-1 nm.

**Figure 10.** Small-angle X-ray scattering. An incoming X-ray beam passes through a glass capillary containing the sample solution. Scattering of the X-rays from the sample is recorded by a detector, radially averaged, background-subtracted, and utilized for further analyses.

Guinier analysis is carried out as one of the first analyses to check the data quality at very small angles, but also to provide an estimate of forward scattering intensity \( (I(0)) \) and the compactness of the particle \( (R_g, \text{radius of gyration}) \). Zero angle scattering cannot be directly measured, both because it is shadowed by the beamstop at the instrument and because it would be impossible to distinguish scattering at 0° from radiation that passes directly through the sample without scattering. However, \( I(0) \) can be extrapolated from the slope of the Guinier region (reviewed in Svergun et al. 2013 p. 82). In addition to compactness \( (R_g) \), the primary data can describe the length \( (D_{\text{max}}, \text{maximum dimension}) \), hydrated volume \( (V_P, \text{Porod volume}) \), an approximate molecular weight \( (\text{from } V_P \text{ or by comparison of } I(0) \text{ to a sample of known molecular weight (Jeffries et al., 2016)}) \) and how ordered the protein is \( (\text{Kratky plot, } I(s)s^2 \text{ per } s) \). The approximate shape, size and possible multimericity can be determined from the pair distribution function, calculated by indirect Fourier transformation from the \( I(s) \) profile.

Monodisperse protein samples can be modeled \emph{ab initio} using dummy residues or beads. The modeling is started with a sphere of beads that exceeds the size of the protein \( (R = D_{\text{max}}/2) \), its theoretical scattering is calculated and compared to the scattering profile of the sample. New models are iteratively generated and their theoretical scattering profiles compared to the measured data to yield a shape that fits the data. With low resolution SAXS data (maximum \( s \) in the range of 2-3 nm\(^{-1}\)) only approximate shapes can be modeled.
Previous structural information of the protein can be utilized in several ways in SAXS data analysis: The measured scattering profile can be compared to the theoretical scattering profile of a crystal or NMR structure or for example to a homology model. The domain organization of a multidomain protein or the setup of a multisubunit protein can be studied from SAXS scattering by fitting the individual domain or subunit structures to the scattering profile, either by rigid-body analysis or by hybrid methods combining both rigid body fitting and \textit{ab initio} modeling of flexible regions (Kikhney and Svergun, 2015).

### 2.4.2 Molecular dynamics

Force fields are functions that can be used to explain and predict the behavior of protein structures. The properties of individual amino acid residues and dipeptides have been extensively studied both experimentally and by quantum mechanics, and this information has been transferred to a function describing the energy associated with stretching and rotation of covalent bonds, electrostatic and van der Waals behaviors of atoms. Standard force fields for proteins are not reactive and do not allow the polarization of charges. In practice, this means that covalent bonds do not form or break and electrons do not transfer, instead charge distributions are described with constant partial charges. Force fields differ in the number of terms contributing to the total energy, how water is modeled and e.g. what types of molecules are included in the core distribution of the force field. As force fields are generally used for molecular dynamics simulations, their design is a balance between realistic description of the system and computational efficiency.

Molecular dynamics simulations allow the study of small systems, typically a protein molecule in a solution containing water and neutralizing ions. Water is typically described with an all-atom water model, e.g. TIP3P (Jorgensen \textit{et al.}, 1983), that is compatible with the force field chosen for the protein and other molecules in the system. The topology and the properties of atoms and angles are given as input to a molecular dynamics software such as NAMD (Phillips \textit{et al.}, 2005) or GROMACS (Hess, Kutzner, Spoel, & Lindahl, 2008), both of which have been utilized in this thesis. The molecular dynamics software moves each particle (typically atoms) of the system and then calculates their energy in repeated cycles for a determined number of steps. The length of each cycle, called time step, is typically 1-2 femtoseconds, meaning that the production of even a single second of simulation is beyond the reach of current computational power. State-of-the-art all-atom simulations can reach up to one millisecond time scale in simulations of small proteins such as ubiquitin (Lindorff-Larsen \textit{et al.}, 2016). The primary result of a molecular dynamics simulation is a trajectory: a collection of 3-D structures, “frames”, collected at a defined frequency from the simulation.

Researcher have utilized various molecular dynamics simulation set-ups to study talin-integrin interaction in solution (Neumann and Gottschalk, 2016; Yuan \textit{et al.}, 2017) and the interactions of talin and integrin with membrane lipids (Kalli \textit{et al.}, 2011, 2016; Kalli, Campbell and Sansom, 2011; Mehrbod and Mofrad, 2013; Mehrbod, Trisno and Mofrad, 2013; Arcario
and Tajkhorshid, 2014; Provasi et al., 2014). Moreover, molecular dynamics studies have shed light on how integrin ectodomains become extended upon fibronectin binding with (Wang et al., 2017) and without manganese (Puklin-Faucher et al., 2006).
3 AIMS OF THE STUDY

The objective of these studies was to study the atomic details of the talin-integrin interaction. The specific aims of each article were:

I. to analyze the specificity and relative affinity of NPxY motif-containing $\beta$-integrins in complex with talin-1.

II. to study how PI(4,5)P2 lipids in the cell membrane interact with the talin-integrin complex, and whether they have a specific role in integrin activation.

III. to find out whether the complete talin head domain can exist in a canonical FERM fold, and what the role of the F1 loop insert is within the talin head.
4 MATERIALS AND METHODS

The materials and methods used in the studies are described in detail in Articles I-III and summarized briefly in this chapter. Sections 4.1-4.3 present the experiments I have carried out mostly or completely myself (molecular modeling, molecular dynamics and small-angle X-ray scattering). Section 4.4 briefly introduces biosensor experiments that had an important role in Articles I and III and were carried out mostly by colleagues. Section 4.5 introduces the cell biology done in collaboration with Prof Bernhard Wehrle-Haller in Article III. Section 4.6 briefly introduces the biophysical methods used for protein structural characterization in Article III.

4.1 Modeling (I, II, III)

The talin-integrin complexes in Article I were constructed using the crystal structure (PDB: 1MK7) of chicken talin head F2 and F3 domains in complex with a 14-residue peptide of the integrin β3 cytoplasmic domain (R736AKWDTANNPLYKE749), complexed by expressing the two proteins as a chimera (García-Alvarez et al., 2003). Integrin β1, β5, β6, and β7 peptides were modeled by exchanging the side chains of the β3 peptide in Bodil (Lehtonen et al., 2004).

The mouse talin F2-F3 (residues 209-400) and F0-F3 (residues 2-398, del30) domains in Article II were modeled using the crystal structure of the mouse talin head with the F1 loop omitted (residues 139-168; PDB: 3IVF, (Elliott et al., 2010) using Modeller (Šali and Blundell, 1993) and DeepView (Guex and Peitsch, 1997).

The complete mouse talin head (residues 1-405) in Article III was modeled using the separate domains of the F1-loop-deleted talin head crystal structure (PDB: 3IVF). The domains were arranged into a FERM fold using the merlin crystal structure (PDB: 1H4R (Kang et al., 2002)) as a template. The modeling was carried out in Modeller. A 23-ns molecular dynamics simulation was carried out with the FERM model using GROMACS 4.5.5 (Hess et al., 2008) and the setup described in section 4.2.2. The resulting conformation was used as the initial model to build a talin-integrin complex.
<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Protein 1 (res. visible)</th>
<th>Protein 2 (res. visible)</th>
<th>Method</th>
<th>Notes, resolution</th>
<th>Used as a template for</th>
</tr>
</thead>
<tbody>
<tr>
<td>1MK7</td>
<td>talin-1 (209-400)</td>
<td>integrin β3 (739-749)</td>
<td>X-ray</td>
<td>β3/talin chimera; 2.2 Å</td>
<td>β1A, β1D, β3, β5, β6, and β7 (I); β3 membrane-distal site (II,III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3G9W</td>
<td>talin-2 (198-408)</td>
<td>integrin β1D (752-789)</td>
<td>X-ray</td>
<td>2.165 Å</td>
<td>β1A and β1D (I); β3 cytoplasmic domain (II,III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3IVF</td>
<td>talin-1 (2-133, 183-398)</td>
<td>(none)</td>
<td>X-ray</td>
<td>ΔF1-loop; 1.94 Å</td>
<td>talin F0-F3(ΔF1-loop) (II); individual domains as templates for FERM model (III); individual domains fitted to SAXS data with EOM (III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1H4R</td>
<td>merlin (20-313)</td>
<td>(none)</td>
<td>X-ray</td>
<td>1.8 Å</td>
<td>FERM domain organization (III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2H7E</td>
<td>talin-1 C336S (309-405)</td>
<td>integrin β3 (716-739)/ PIP5K1γ (643-652) chimera</td>
<td>NMR</td>
<td>average structure</td>
<td>F3 in FERM model (III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2KC2</td>
<td>talin-1 (86-202)</td>
<td>(none)</td>
<td>NMR</td>
<td></td>
<td>F1 in FERM model (III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2KMA</td>
<td>talin-1 (1-138, 169-202)</td>
<td>(none)</td>
<td>NMR</td>
<td>ΔF1-loop</td>
<td>F0F1 in FERM model (III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2K9J</td>
<td>integrin αIIb (958-998)</td>
<td>integrin β3, C687S (685-727)</td>
<td>NMR</td>
<td>embedded in bicelles</td>
<td>αIIbβ3 TM domains and IMC (II,III)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2KNC</td>
<td>integrin αIIb (960-1008)</td>
<td>integrin β3 (689-762)</td>
<td>NMR</td>
<td>organic solvent</td>
<td>αIIbβ3 cytoplasmic domains (II,III)</td>
</tr>
</tbody>
</table>

* (García-Alvarez et al., 2003)
* (Anthis, Wegener, et al., 2009)
* (Elliott et al., 2010)
* (Kang et al., 2002)
* (Wegener et al., 2007)
* (Goult, Bouaquina, et al., 2010)
* (Lau et al., 2009)
* (Yang et al., 2009)
The integrin αIIbβ3 model in Articles II and III was created using the NMR structures of integrin αIIbβ3 transmembrane and cytosolic domains in a 50/50 mixture of organic solvent and water (PDB: 2KNC (Yang et al., 2009)) or embedded in small bicelles (PDB: 2K9J, (Lau et al., 2009)), the crystal structures of the talin-2-β1D-integrin interface (PDB: 3G9W, (Anthis, Wegener, et al., 2009)) and the talin-1-β3-integrin interface (PDB: 1MK7, (García-Alvarez et al., 2003)). Homology modeling was carried out using Modeller. The free C-terminal tail of β3-integrin (residues 750-762) was modeled as an α-helix.

In the control simulations of Article III point mutations in the talin head domain were introduced using PyMOL (Schrodinger). The model for examining the role of the F1 loop deletion (del30) was prepared in Modeller. I used the final conformation of the 100-ns wild-type talin head simulation as the template for the talin head without the F1 loop. The NMR structure of the loop-deleted F1 domain (PDB: 2KMA (Goult, Bouaouina, et al., 2010)) served as template for the shortened F1 loop. Please see Table 2 for details of the templates.

### 4.2 Molecular dynamics (I, II, III)

Molecular dynamics simulations were carried out using systems of proteins in solution (Article I) or proteins and a lipid bilayer in solution (Articles II and III) using all-atom protein force fields and explicit water. Molecular dynamics simulation data were analyzed using the GROMACS tools (Hess et al., 2008; Pronk et al., 2013) and VMD (Humphrey, Dalke and Schulten, 1996).

#### 4.2.1 Steered molecular dynamics (SMD) with NAMD (I)

The different talin-β-integrin complexes under mechanical stress were studied in a set of steered molecular dynamics simulations. The system for simulation, consisting of a β-integrin peptide and talin F2-F3 domains in 150 mM NaCl, was prepared using the Psfgen tool in VMD. The CHARMM22 force field (MacKerell et al., 1998) was used for the protein and ions, and the TIP3 model (Jorgensen et al., 1983) was used for the water molecules. Two energy minimizations were run using the conjugate gradient method in NAMD (Phillips et al., 2005): the first 4000 steps had all protein atomic positions fixed, the second 4000 steps were without positional constraints. The temperature of the system was gradually raised to 310 K in a 31-ps molecular dynamics simulation. A short (1 ns) equilibration simulation was performed before starting the steered molecular dynamics simulations. The molecular dynamics simulations were conducted at 310 K temperature and 1 atm pressure with Berendsen pressure and temperature coupling (Berendsen et al., 1984) in NAMD.

The N-terminal Cα atom of the integrin peptide acted as a pulling point to simulate traction from ligand-bound integrin extracellular domains. In the cell talin is connected to the force-generating actin machinery via its C-terminal rod bundles; however, to avoid the unraveling
of the talin F3 C-terminal α-helix or the breakage of the integrin binding site, the force
directed onto talin was divided to three Cα atoms instead of the C-terminal residue of its F3
domain. The three selected Cα atoms belonged to N355, located in a 3_10 helix in the NPxY
tyrosine binding site, G371 in the L6,7 loop and the C-terminal L400 residue. Talin and
integrin were pulled to opposite directions along the integrin N-terminal Cα–talin L400 Cα
axis. Constant force at the magnitude of 200 pN, 300 pN, 400 pN and 500 pN was used in
two simulations each and 600 pN in one simulation.

4.2.2 Molecular dynamics with GROMACS (II, III)

In Article II we carried out six 500 ns simulations of the talin-αIIbβ3-integrin complex in the
context of a lipid bilayer consisting of either 1,2-dioleoyl-sn-glycero-3-phosphocholine
(DOPC) or a 9:1 (n:n) mixture of DOPC and phosphatidylinositol-4,5-bisphosphate (PIP2).
Two different talin constructs were used, mouse talin-1 domains F2F3 (2+2 simulations) and
the F1-loop-truncated talin-1 head in the crystallographic structure PDB: 3IVF (Elliott et al.,
2010) (1+1 simulations). Integrin without talin in a DOPC:PIP2 (9:1) lipid bilayer was
simulated in two 750 ns control simulations. All simulations were carried out in GROMACS
4.5.5 (Hess et al., 2008; Pronk et al., 2013).

The simulations were carried out in 150 mM KCl, with counter ions added to neutralize
the charge of the system. We used the OPLS all-atom force field (Jorgensen and Tirado-Rives,
1988) for the proteins and ions, and a compatible, lipid-refined version of the OPLS all-atom
force field for the lipids (Maciejewski et al., 2014; Kulig, Pasenkiewicz-Gierula and Róg, 2016).
For water, we used the TIP3P water model (Jorgensen et al., 1983). We applied periodic
boundary conditions with the minimum image convention, utilized a timestep of 2 fs and
constrained the length of covalent bonds involving hydrogens with the LINCS algorithm
(Hess et al., 1997). Lennard-Jones interactions were cut off at 1 nm. The particle mesh Ewald
method was used for electrostatic interactions, using a real space cutoff of 1 nm, B-spline
interpolation and a direct sum tolerance of 10^-6. Pressure was kept constant at 1 bar with the
temperature was kept at 310 K using the v-rescale method (Bussi, Donadio and Parrinello,
2007) and coupling the temperatures of the solvent (water and ions) and solute (protein and
lipids) separately.

In Article III we studied the conformation of talin head and its integrin binding in the
context of the 100% DOPC and DOPC:PIP2 (9:1) bilayers in 150 mM KCl using the same
molecular dynamics simulation settings as in Article II. Three talin-integrin complexes (1 x
DOPC, 2 x DOPC:PIP2) were started from the same initial conformation. For the repeat of
the DOPC simulation, the initial conformation of the talin-integrin complex was adopted
from a DOPC:PIP2 simulation at 20 ns, where the talin F2 domain had formed a contact
with the lipid bilayer. Control simulations of the talin head alone in 150 mM KCl were run
using the same settings as with the lipid membrane simulations, but with isotropic pressure
coupling. Control simulations were run with the wild-type talin head (t1-405), the F1-F2 interface mutant t1-405(R194E) and the F1-F2 interface charge inversion double mutant t1-405(R194E,D222K) for 100 ns. The t1-405(R194E) simulation was run using GROMACS 4.6.6. Three 200 ns simulations of the wild-type talin head and three 200 ns simulations of a F1-loop-deleted (del30) talin head were run with Gromacs 2018.1, using the same settings as previously.

4.3 Small-angle X-ray scattering (III)

The talin head constructs t1-405, t1-405(del30), t1-405(del37/GAG) and t1-405(R194E) were studied both by batch SAXS and by SAXS with in-line size exclusion chromatography (SEC) (Figure 11). The SAXS data was collected at the PETRA-III storage ring of the Deutsches Elektronen-Synchrotron (DESY) at the P12 beamline operated by the European Molecular Biology Laboratory (EMBL), Hamburg outpost (Blanchet et al., 2015).
Figure 11. SAXS was combined with in-line size exclusion chromatography to separate any oligomeric fractions from the monomeric sample. The monomeric protein was detected from right-angle light scattering. SAXS analyses were carried out with background-subtracted SAXS data of the monomeric protein.

For batch SAXS, a series of 4 dilutions was prepared in 50 mM sodium phosphate, 150 mM NaCl, pH 7.2: e.g. for t1-405 the concentrations were 0.78 mg ml\(^{-1}\), 1.51 mg ml\(^{-1}\), 2.86 mg ml\(^{-1}\) and 5.37 mg ml\(^{-1}\). SAXS data of buffer alone was collected before and after each protein sample, and in the data analysis stage the average of the buffer data was subtracted from the scattering data of the protein sample. The Autorg program of the ATSAS package (Petoukhov et al., 2007) assigned the protein in one to two dilutions of each construct to be possibly aggregated. Moreover, the model-free parameters showed an increase in particle size with increasing concentration, suggesting that oligomerization took place in the sample.

To overcome the problem of oligomerization, we utilized in-line size exclusion chromatography that separated the protein in the sample by molecular size. Refractive index, right-angle light scattering (RALS) and ultraviolet (UV) data were collected downstream of
the size exclusion column (Superdex 200 Increase 5/150, GE Healthcare). The flow rate was 0.25 ml min⁻¹. SEC-SAXS data were collected in 0.995 s exposures, 2400 frames per sample at +10 °C, and scattering intensities from the image data were radially averaged. To determine the monomeric fraction of the data, the buffer background was first subtracted from the total data. The forward scattering intensity I(0) and the radius of gyration (Rg) were then calculated for each frame of the dataset and used for determining the location of protein peaks in the data. I(0) and RALS data were fitted using the Signal package in GNU Octave (Eaton et al., 2015) to confirm the molecular weight of the sample at the presumed monomeric peak using the UV data. In search for a good peak/buffer combination for background subtraction, sets of pre-peak and post-peak scattering data representing the buffer (31 frames from each side of the monomeric peak) were subtracted from the monomeric peak (in slices of 20 frames) using the Dataver, Datop and Autorg tools in a semi-automated manner with Perl scripts. Promising peak/background combinations were selected based on parameters such as the radius of gyration, the quality of the data as reported by Autorg, the length of the Guinier range, forward scattering intensity and the shapes of the log I(s)/s and p(r) curves.

The Ensemble Optimization Method (Bernadó et al., 2007; Tria et al., 2015) (EOM), implemented in the ATSAS package (Franke et al., 2017), was used in the analysis of the conformational flexibility of the proteins. EOM is meant for the analysis of flexible or even intrinsically disordered proteins, and it generates an ensemble of models describing the conformational space of the protein in solution. Initially, EOM generates a large pool of conformations (typically 10,000) to cover the whole conformational space of the protein. The conformations can be generated by \textit{ab initio} modeling or as a hybrid model composed of rigid body structures and \textit{ab initio}-modeled regions. Theoretical scattering profiles are calculated for all the models in the pool. A genetic algorithm is utilized to select an ensemble from the pool to represent the dynamics of the protein in the sample (Bernadó et al., 2007; Tria et al., 2015).

The X-ray structure of the talin head residues 1-400(del30) (PDB ID: 3IVF (Elliott et al., 2010)) was separated into individual domains, and the F1 domain was further split at the F1 loop into fragments comprised of residues P87-L133 and W173-K196. The theoretical scattering profiles of these fragments were calculated with Crysol (Svergun, Barberato and Koch, 1995). The positions of the F1 domain fragments were fixed in the EOM modeling run, whereas the positions and orientations of the F0, F2 and F3 domains were not constrained. The F1 loop, inter-domain linkers and N- and C-terminal residues were modeled \textit{ab initio} using the random coil method.

### 4.4 Biosensor analysis for talin-integrin binding (I, III)

Binding of β-integrin to talin was measured using the Octet biosensor with Ni-NTA (nickel-nitrilotriacetic acid) sensors. H₆-tagged talin was immobilized on the sensors via Edc/NHS (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide/N-hydroxysuccinimide) coupling. In Article I we measured the binding of the peptides spanning the cytoplasmic tails up to the
end of the (putative) talin-binding sequence of integrins β3 and β6 to talin head. A randomized β6 sequence (β6-scramble) was used as a control (Table 3). These integrin peptides alone gave no detectable binding signal in Octet, and were therefore complexed with streptavidin via an N-terminal biotin. The binding of the streptavidin-bound biotinylated peptides to the talin-functionalized surface was measured at four different concentrations.

In Article III, we utilized the Octet biosensor to analyze the integrin binding of different talin head constructs. The cytoplasmic domains of wild-type β3 integrin and a high-affinity β3 integrin (β3-VE) (Pinon et al., 2014) were used, both as C-terminal fusions to glutathione S-transferase (GST).

<table>
<thead>
<tr>
<th>Name</th>
<th>Article</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>β3</td>
<td>I</td>
<td>HDRKEFAKFEERARAKWDTANNPLYKE</td>
</tr>
<tr>
<td>β6</td>
<td>I</td>
<td>HDRKEVAKFEAESKAKWTGTPLYRG</td>
</tr>
<tr>
<td>β6-scramble</td>
<td>I</td>
<td>KDWGTEHRQALINSVFYFKAGKERKTPARE</td>
</tr>
<tr>
<td>GST-β3</td>
<td>III</td>
<td>GSKLLITIHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNITYRGT</td>
</tr>
<tr>
<td>GST-β3-VE</td>
<td>III</td>
<td>GSKLLITIHDRKEFAKFEERARAKWVENPLYKEATSTFTNITYRGT</td>
</tr>
</tbody>
</table>

Talin-integrin binding was measured at five different GST-integrin concentrations. To allow the comparison of data between experiments, the binding signals were normalized within the GST-β3 and GST-β3-VE groups. A GST control without integrin peptide was subtracted as background from all datasets, and the data were normalized to the wild-type talin head (t1-405) at a GST-integrin concentration of 1250 nM. A linear regression analysis assuming one site-specific binding was carried out in GraphPad Prism 5.02 (GraphPad Software, La Jolla, CA).

### 4.5 Cell assays (III)

The experiments described in this section were carried out by our collaborators in the laboratory led by Prof Wehrle-Haller at the University of Geneva using αVβ3 integrin. The β3 integrin construct used was C-terminally GFP-tagged mouse β3 integrin, and the murine αV integrin was untagged. Assays with full-length talin were carried out using a chimeric talin construct consisting of the human talin-1 head and the murine talin-1 rod. All talin constructs were expressed with an N-terminal CFP-tag. Kindlin-1, utilized in clustering and cysteine-cysteine crosslinking assays, was N-terminally labeled with TagRFP.
4.5.1 Integrin activation (III)

The activation of integrin was measured by expressing integrin (β1 or β3) and talin (wild-type or mutated) in murine melanoma cells. The cell population was divided in two. β1 integrins expressed on the cell surface were detected by a hamster anti-β1 antibody, and the other cell population was incubated with a rat monoclonal antibody binding to the active, extended β1 integrin (9EG7 (Lenter et al., 1993)). In a similar manner, total β3 integrin was detected by a hamster anti-β3 antibody. To detect active β3 integrins, cells were incubated with the SKI-7 fusion protein composed of a snake venom disintegrin and a CD31 Ig domain (Ballestrem et al., 2001), followed by incubation with a rat anti-CD31 monoclonal antibody (5GC1). β1 and β3 integrins expressed on the cells were detected by flow cytometry using phycoerythrin-labeled anti-hamster antibodies. To detect active β3 integrins, cells were incubated with the SKI-7 fusion protein composed of a snake venom disintegrin and a CD31 Ig domain (Ballestrem et al., 2001), followed by incubation with a rat anti-CD31 monoclonal antibody (5GC1). β1 and β3 integrins expressed on the cells were detected by flow cytometry using phycoerythrin-labeled anti-hamster antibodies. The rat antibodies 9EG7 targeting active β1 and 5GC1 targeting the active β3-bound SKI-7 were detected with phycoerythrin-labeled anti-rat F(ab')2 fragments in fluorescence-activated cell sorting (FACS). The integrin activation index was calculated as the ratio of the activation-specific antibodies to the total integrin-binding antibodies as described by Saltel et al. (Saltel et al., 2009). The only kindlin in the experiment was endogenous, and thus the results measure the ability of different talin forms to extend integrins at low kindlin levels.

4.5.2 Integrin clustering (III)

In order to test the integrin activation and clustering efficiency of talins in the presence of kindlin integrin clustering assays were carried out. Murine melanoma cells were transfected with CFP-talin, integrin-GFP and TagRFP-kindlin. They were then plated on fetal calf serum-coated glass and fixed with paraformaldehyde either directly or after stimulation with manganese. Total internal fluorescence microscopy (TIRF) was utilized to quantitate integrin clusters at adhesion sites located near the cell membrane. Clustering was measured using a clustering index described by Saltel et al. (2009). Intensity histograms of background-subtracted cell images were normalized by cell area and averaged. The clustering index was then determined from the number of pixels brighter than an arbitrarily selected threshold level, resulting in an index that describes the area occupied by integrin clusters as a percentage of the total cell area.

4.5.3 Cysteine-cysteine crosslinking (III)

Potential novel interactions between talin and the αIIbβ3 integrin dimer were tested in monkey kidney fibroblasts (COS-7) expressing αV, TagRFP-kindlin-1, wild-type or cysteine-mutated β3-GFP-integrin and wild-type or cysteine-mutated talin head-CFP. Palmitoylation of the free cysteine residues was blocked with 20 µg ml⁻¹ 2-bromopalmitate in culture medium.
for 17 hours. Oxidizing conditions were created by incubating the cells in a medium containing 200 µM CuSO₄, 100 µM phenanthroline and 0.02 % saponin to allow the Cu-phenanthroline to penetrate the cell membrane. After 10 minutes the oxidation solution was removed and the cells were lysed. The lysate was analyzed in SDS-PAGE under non-reducing conditions and detected with an anti-GFP antibody.

4.6 Structural characterization (III)

Mass spectrometry was utilized to determine the exact composition of the talin head constructs as described in Paper III. The fold of the talin head was studied using hydrogen/deuterium exchange mass spectrometry in collaboration with Prof. Janne Jänis (University of Eastern Finland) and electron microscopy in collaboration with Prof. Holland Cheng (University of California, Davis).

4.6.1 Electron microscopy (III)

The shape of the talin head construct t1-405 was studied by transmission electron microscopy (TEM). The sample was immobilized on carbon grids, negatively stained with ammonium molybdate and trehalose and air-dried at room temperature. Fifteen images taken with a CCD camera attached to the TEM were processed for single particle analysis, and the background noise was reduced by low-pass filtering to 10 Å and high-pass filtering to 300 Å. A total of 1300 particles were selected semi-automatically, classified using reference-free classification and then averaged. The class averages were used to generate a low-resolution initial model, which was then refined over 10 cycles, resulting in a map with a final resolution of 15.2 Å by 0.143 FSC criteria. This experiment was carried out by our collaborators in Prof. Holland Cheng’s laboratory at the University of California at Davis.

4.6.2 Mass spectrometry (III)

The compactness of the talin head fold was studied using hydrogen/deuterium exchange (HDX) mass spectrometry. The protein samples were diluted 1:4 (V/V) in D₂O, and hence the maximum theoretical deuterium intake was 80%. The samples were infused to the electrospray ionization source at a constant flow, and HDX mass spectra were collected at 125-s intervals for up to an hour. The maximum number of exchangeable hydrogens was calculated as the sum of the backbone amide groups and certain side chains. The relative deuterium intake at each time point was calculated as the mass difference between the deuterated and undeuterated protein divided by 0.8 × the theoretical maximum number of exchangeable hydrogens.
5 RESULTS

5.1 Talin-integrin binding

A complex of talin and integrin was present in all three studies of this thesis. According to the current view, integrin $\beta$ binds to talin with both its membrane-proximal helix and the membrane-distal site described in the García-Alvarez et al. structure (García-Alvarez et al., 2003). In Article I, we studied the $\beta$-integrin membrane-distal binding site, using a 14-residue stretch and a model based on the García-Alvarez et al. $\beta$3-integrin-talin-1 complex. This first article analyses the contacts in detail, and while we did not attempt to determine the affinity of the binding, we detected differences in the stability of the complexes, and the contribution of the modeled interactions to the stability. Article II focused primarily on the lipid binding of talin and integrin and Article III on the conformation of talin head and the interactions of its F1 loop. However, I will revisit the data of Articles II and III in this chapter to provide a commentary on the talin F3-$\beta$3 integrin binding details omitted from those articles.

5.1.1 Steered molecular dynamics of $\beta$-integrins and talin F2F3

A set of nine SMD simulations were carried out for each talin-integrin complex, two 200 pN, 300 pN, 400 pN and 500 pN simulations each and one 600 pN simulation. In the García-Alvarez structure, integrin is bound to talin as a short antiparallel $\beta$-strand comprising the integrin $\beta$3 residues D$^{740}$TA$^{742}$. $\beta$-integrin’s first NPxY motif, N$^{744}$PLY$^{747}$, inserts its tyrosine into a groove formed by the K357 side chain and the 3$^{10}$-helix residues L353, T354 and N355, and F370 in the L6,7 loop. The W739 side chain sits in a pocket formed by the talin residues R358, Y377, T367 and A360.

Results from the simulations varied markedly. Common to all simulations, except those of $\beta$5, were the contact between W739 and R358 and the two D740-W359 backbone bonds defining the minimum of the antiparallel $\beta$-strand. In the case of $\beta$5 integrin, the W739 residue is a tyrosine and did not stably stack with the R358. The conserved W739 and NPxY motifs are in most $\beta$ integrins separated by four residues (three residues in $\beta$2), whereas in other talin-binding peptides the linker between the motifs is only two residues long, allowing a more compact binding. Our analysis followed four residue pairs along the talin-integrin interface: W739-R358, D740-W359, A742-K357 and Y747-N355 (Figure 13). As the data from SMD simulations were pooled together, the analysis showed that the $\beta$3 complex did not favor a bound state of the A742-K357 hydrogen bond.
**Figure 12.** β integrin peptides studied in this thesis. Residues buried within the cell membrane are shown in gray. The membrane-proximal and membrane-distal talin-binding sites are marked above the alignment. The membrane-distal peptide studied in Article I is marked by a bar below the alignment. The β3:D723 residue contributing to the αβ salt bridge in the integrin heterodimer is marked in blue. The NPxY and NxxY motifs are highlighted with yellow and gray shading, respectively. β3, β3-VE, β5, β6, and β7 residue numbers are counted without the signal peptide.

**Figure 13.** The binding of β-integrin membrane-distal binding site residues to talin was analyzed by measuring the distances of residues pairs along the integrin-talin interface in Article I, using the pooled data from all SMD simulations. A) A hydrophobic contact between the β3:W739 and talin:R358 side chains was present in all data sets. β5 has a tyrosine instead of the β3:W739 tryptophan, and this tyrosine was more flexible, giving a slightly shifted peak in the distance histogram. B) The β3:D740-talin:W359 double backbone hydrogen bonds appeared to be critical for the mechanical stability of the complex. C) A β3:A742-talin:K357 backbone hydrogen bond provided stability in some of the simulations. D) NPxY binding was stable only in some complexes and simulations. Note: The Y747-N355 distance was calculated for Y747 sidechain and N355:C, not N355:O as marked in panel D. Reproduced from Kukkurainen et al. 2014 with permission from the Royal Society of Chemistry.
A single simulation of the β6-talin complex dominated the whole β6 dataset. While the two backbone hydrogen bonds between β3:D740 and talin:W359 had appeared critical to the mechanical stability of other integrin-talin complexes, the β6 200 pN simulation showed a stable complex involving only one backbone hydrogen bond between β6:Q734 (β3:D740) and talin:W359, and the NPLY remained unbound (Figure 14). To the best of our knowledge β6 integrin has not been reported to bind talin. As the lifetime of the β6-talin complex observed in one simulation by far exceeded the lifetimes of all other β-talin complexes, we proceeded to confirm the interaction of β6 and talin experimentally. Using a set of β3, β6 and randomized β6 peptides, we measured integrin-talin binding with an Octet biosensor. Indeed, we could confirm β6 to bind talin, whereas the randomized β6 peptide and β3 showed negligible binding.

For β1A and β1D integrins, we ran sets of SMD simulations also with an initial structure adapted from the β1D-talin-2 structure published by Anthis et al. (Anthis, Wegener, et al., 2009) (Figure 15). This initial conformation was markedly different from the García-Alvarez structure and showed no β-sheet-like backbone hydrogen bonding. The angle of NPxY (NPIY) tyrosine insertion into the tyrosine-binding groove on F3 was different from the García-Alvarez structure. In the García-Alvarez initial conformation β1D formed a more stable complex with talin than β1A, with three and two backbone hydrogen bonds, respectively. The Anthis et al. conformation, however, had only one β-sheet-like backbone hydrogen bond (D776:N-W359:O). This hydrogen bond was unstable and broke quickly in all 300-600 pN simulations, while it survived up to 5 ns (β1A) and 7 ns (β1D) at 200 pN. In the Anthis et al. conformation, the β1A and β1D complexes were very similar in stability (please see Figure 13).
Figure 14. Integrin β6 binds talin. In Article I, β6 was computationally predicted and biophysically confirmed to be a talin-binding integrin. The β6-integrin-talin complex was very stable under constant force pulling at 200 pN. A) Hydrogen bonding in the 200 pN simulation. Backbone hydrogen bonds are marked with an asterisk. B) and C) Snapshots from the 200 pN simulation. D) Octet biosensor measurements showed talin to bind β6 integrin. The scrambled β6 sequence (β6-SCR) and β3 integrins showed negligible binding. Adapted from Kukkurainen et al. 2014 with permission from the Royal Society of Chemistry.
Figure 15. Comparison of talin interactions of β1A and β1D in the García-Alvarez configuration and β1D in the Anthis configuration, showing hydrogen bonding and snapshots from representative simulations in Article I. Dissociation of β1A and β1D integrin from talin-1 in pulling simulations at a 300 pN constant force. The occupancies of hydrogen bonds in the course of representative simulations. The single-residue difference (G778 vs. Q778) between the membrane-distal talin-binding sequences in β1A (panels A-C) and β1D (panels D-F) was reflected in the higher mechanical stability of the β1D-talin complex in the García-Alvarez et al. configuration, but not in the Anthis et al. configuration, as visualized for β1D (panels G-I). The Q778 residue in β1D is highlighted in yellow. Asterisks mark hydrogen bonds between peptide backbone groups. Adapted from Kukkurainen et al. 2014 with permission from the Royal Society of Chemistry.
5.1.2 αIIbβ3 integrin, talin and a lipid bilayer

The details of β3 integrin binding to the talin F3 domain were not analyzed in Articles II and III. However, to validate the results of Article I I carried out a similar analysis of the membrane-distal residue pairs as in Figure 13 using the datasets collected for Articles II and III. The results of the analysis are shown in Figure 16.

The long simulations of integrin in complex with FERM-folded talin showed instability in the binding interface. The initial configuration of the simulations was modeled based on the PDB structures 1MK7 (García-Alvarez et al., 2003), 2H7E (Wegener et al., 2007) and 3G9W (Anthis, Wegener, et al., 2009) and accommodated both MP-helix binding and a β-sheet-like set of backbone hydrogen bonds (D740:N-W359:O, W359:N-D740:O, A742:N-K357:O). As discussed in section 5.1.1 on the membrane-distal peptide binding, the Anthis structure contains a bound MP-helix and its NPxY motif (NPIY) is inserted in a different angle compared to the García-Alvarez structure, with the residue W739 (β1D: W775) in contact with R358, but with no β-sheet-like backbone hydrogen bonding taking place. The NPxY tyrosine did not form stable contacts to its binding pocket in any of the simulation set-ups, in agreement with the results from Article I. The backbone hydrogen bonds between β3 D740 and talin W359 were very stable, with the exception of one of the FERM domain simulations with a DOPC:PIP2 membrane.

Simulations of the F2F3 domains and of the linear arrangement of F0-F3 domains showed a stable talin-integrin binding interface in the presence of a lipid bilayer. While the initial integrin binding configuration was the same in the FERM and non-FERM simulations and the membrane-proximal helix retained its contacts to the F3 domain in all simulations, the β-sheet-like binding of the membrane-distal binding stretch became unstable in the FERM domain simulations.

Our MD simulations of αIIbβ3 with different talin models and a lipid bilayer, ten in total, were not conclusive in regard to the integrin αIIb cytoplasmic tail conformation and interactions. We originally modeled the αIIb tail as an extended coil having no contact with the talin F2 and F3 domains, and it formed a turn in all simulations. In one FERM domain simulation with a DOPC:PIP2 bilayer the αIIb tail crossed the F2 domain and formed interactions with both the F2 domain and the lysine-rich L3,4 loop of F3.
Figure 16. Inter-residue distances along the integrin-talin interface (unpublished analysis of data pooled from the MD simulations of Articles II and III). To review the results of Article I, inter-residue distances of selected residue pairs along the talin-integrin binding interface were analyzed and plotted in histograms (bin width=0.2 Å, occupancy in percentage) using Gnumeric and Graphpad Prism. The X-axis shows the aom-atom distances along the membrane-proximal and membrane-distal binding sites of the talin-β3-integrin complex, and the Y-axis shows the occupancy at each distance. For each contact, data from Article II were pooled together into extended talin (F2F3 and F0-F3 del30), and data from Article III were pooled into FERM talin. Simulations with DOPC as the sole lipid are shown in gray, and simulations with 10% PIP2 in the membrane are shown in red. The distance was calculated for one frame per 100 ps, totaling 20,000 frames for the FERM domain talin with 10% PIP2 and 15,000 frames for other conditions. Note the different Y-axis scale for D740(β3)-W359(tal) and the different X-axis scale for Y747(β3)-N355(tal). Equivalent distances in talin-integrin NMR and crystal structures are marked with arrowhead, and their corresponding PDB IDs are shown in the legend.

5.2 Lipid interactions

A lipid bilayer was included in the talin-integrin complex simulations in Articles II and III. While the focus of Article II was on lipid interactions, I will report here lipid interactions from both Article II and Article III datasets. We used two simplified lipid bilayer models: 1) 100% 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 2) 9:1 DOPC:phosphatidylinositol 4,5-bisphosphate (PIP2). DOPC is a charge-neutral lipid, whereas PIP2 has two negatively charged phosphate groups thought to bind positively charged patches in the talin head F2 and F3 domains and in the F1 domain loop. Article II specifically discusses the lipid interactions of the F2F3 domains with integrin, as no F1 loop was included in the data. Article III, on the other hand, briefly discusses the lipid contacts of the F1 loop.
5.2.1 Talin-lipid binding

The FERM domain simulations were started from a conformation, where the center of mass of the F2 domain was at approximately 3.5 nm distance from the P8 (phosphate atom) layer of the membrane. In both PIP2 simulations the F2 domain reached the membrane quickly (at 11 ns and 78 ns, the distance between protein and lipid ≤ 3 Å). The phosphate groups of the PIP2 lipid protrude approximately 4 Å further from the membrane plane than the ammonium group of DOPC, and hence the distance between protein and lipids in the initial configuration was effectively longer in the simulations without PIP2. In the first run of the non-PIP2 containing simulation, the F2 domain failed to contact the lipid bilayer, instead its F0 domain reached the membrane and remained in contact with it for the rest of the 500 ns simulation. A repeat of the non-PIP2-containing simulation was started from a configuration adopted from the first PIP2-containing simulation, and this time the F2-membrane contact formed at 11 ns of simulation and no F0-membrane contact was observed.

In the simulations of the F2F3 domains a single PIP2 molecule occupied a PH-domain-like binding site formed by the F3 residues K320, K328 and R339, in a similar fashion as observed for dibutyryl-PIP2 bound to the phosphoinositide binding site in the PH domain of the Arf-GAP ASAP1 (PDB ID: 5C39 (Jian et al., 2015)). Using the β2 strand (L325:Cα and T333:Cα) of the F3 as a measure, F3 subdomains with PIP2 bound in a PH-like mode were at a ~60° angle from the membrane plane. PH-like PIP2 binding was not observed in FERM domain simulations. Instead, on average two PIP2 lipids bound to lysines in the L1,2 loop (Figure 18). In the extended conformation of loop-deleted talin, PIP2 lipids formed stable contacts with K322, K343 and K364, one PIP2 lipid per residue.

The previously described (Anthis, Wegener, et al., 2009; Saltel et al., 2009; Song et al., 2012) basic patch on talin F2, composed of residues K256, K272, K274 and R277, was found to attract several PIP2 lipids (Figure 17A). In addition to the previously identified PIP2 binding residues, we found K254 and K263 to contribute to PIP2 binding by the extended (F2F3 or F0-F3(del30)) talin forms; however, previous reports suggest that this site does not affect talin F2F3 interaction with the membrane (Song et al., 2012).

The F1 loop was included only in the simulations of the FERM folded talin model. It was found to bind PIP2 lipids especially with its R146 and K149 residues (occupancies 1.23 and 0.93, respectively). In addition, lower occupancy (0.42-0.49) binding was observed for K147, K156 and K157.

The F0 domain contacted the lipid bilayer in both simulations of the extended conformation. Moreover, in one of the non-PIP2-containing FERM domain simulations the F2 domain failed to contact the lipid bilayer. Instead, in this particular simulation the F0 domain bound to the membrane, resulting in the twisting of the whole talin-integrin complex.
Figure 17. All the domains of the F2F3 and extended F0-F3 talins were in contact with lipids in the simulations of Article II. The percentage of time that each talin residue spent within 3.5 Å of A) a PIP2 lipid or B) a DOPC lipid is color-coded on the talin surface. White: no contact, red: contact with PIP2, green: contact with DOPC. The location of the lipid-interacting residues in the talin sequence and the color scales are shown in the graphs on the right. Note: the residue R277 has been incorrectly marked as K277 in panel A. Reprinted with permission from Orłowski et al. 2015. Copyright 2015 American Chemical Society.
5.2.2 Integrin-lipid binding

In Article II, we found that PIP2 interfered with the D723-R995 salt bridge of the integrin αIIbβ3 inner membrane clasp (IMC), resulting in opening of the clasp. This observation was made in two out of the three simulations with PIP2 lipids; however, in the third PIP2 simulation the clasp opened without interference from PIP2. Control simulations of integrin in a PIP2-containing lipid bilayer without talin showed no similar effect. The IMC remained intact in two of the simulations with a DOPC membrane, and in one it was broken by the negatively charged αIIb integrin that folded back to interact with R995. In the simulation of t2-398(del30) PIP2 lipids bound to the L1,2 loop of the F3 domain interfered with the IMC salt bridge, whereas in the simulation of F2F3 PIP2 was bound to talin in a PH-like fashion and the interfering PIP2 molecule approached the integrin dimer from the opposite side (Figure 19). These results indicate that the presence of the F2F3 domains alone or an extended F0-F3 construct can effectively destabilize the αIIbβ3 IMC, in line with the results of Mehrbod and coworkers (Mehrbod, Trisno and Mofrad, 2013). PIP2 may destabilize the IMC further, and the lysine-rich L1,2 loop in the F3 domain may help the process by attracting PIP2 lipids to the vicinity of the IMC.

Figure 18. PIP2 binds to the F3 domain both in a PH-like and in a PTB-like fashion. The lipid binds to talin F3 in MD simulations (light brown) and to the PH and PTB domains of PDK1 and Dab1 (white). The F2F3 simulation snapshot is from t=249 ns and the FERM simulation snapshot from t=686 ns.
Figure 19. The non-FERM talin simulations predict the perturbation of the integrin inner membrane clasp by PIP2 lipids. A) Schematic figure of the simulation. B) Snapshot from a simulation with F2F3 talin. C) Snapshot from a simulation with talin F0-F3(del30). Reprinted with permission from Orłowski et al. 2015. Copyright 2015 American Chemical Society.

5.3 Talin FERM

We explored the possibility of a talin head FERM fold using a homology model of the mouse talin-1 head residues 1-405. As templates we used the 3-D structures of talin-1 domains (Elliott et al., 2010; Goult, Bouaouina, et al., 2010) and adapted a FERM fold from the crystal structure of merlin (Kang et al., 2002). To date, there are no reports of the talin head adopting a FERM domain conformation, and hence our major concern was the reliability of the model: Is the FERM model sufficiently stable to represent a real, existing domain conformation?

We ran a 0.5 µs and a 1.0 µs simulation of the talin FERM model in complex with an αIIbβ3 integrin dimer embedded in a DOPC membrane and two 1.0 µs simulations of a similar system with a DOPC:PIP2(9:1) membrane. Problems with the structure arose in two of the simulations: In the 0.5 µs long simulation without PIP2 the F0 domain, but not F2, formed stable contacts to the membrane. This F0-membrane interaction is not supported by
the literature, and as it appeared to both distort the protein complex and prevent the formation of the known F2-membrane interactions (Saltel et al., 2009), we stopped the simulation at 0.5 µs. In one of the simulations with DOPC:PIP2 membrane, the talin F1-F2 interface rapidly opened after 415 ns of simulation, but the interface slowly closed in the course of the subsequent ~400 ns (Figure 20).

**Figure 20.** The initial talin FERM model was compact, but lost some of its compactness in the course of the 0.5-1 µs simulations. The radius of gyration was calculated for the complete talin head (residues 1-405) using the g_gyrate tool in GROMACS 4.5.5.

Small-angle X-ray scattering experiments were carried out with the talin head forms t1-405, t1-405(del30), t1-405(del37/GAG) and the F1-F2 interface mutant t1-405(R194E) (Figure 21). In-line size-exclusion chromatography with RALS and UV detection showed the majority of the protein in solution to be monomeric. The SAXS data about the conformation of the talin head was not conclusive. Ensemble modeling with EOM suggested the majority of the t1-405 proteins in solution to have a FERM-like shape (Figure 21B). However, all talin forms except t1-405(del37/GAG) were somewhat flexible. This suggests that, whether or not the talin head has a FERM fold, this FERM fold is not fully stable in solution.

The talin head crystal structure (PDB ID: 3IVF (Elliott et al., 2010)) has been produced with a loop-deleted talin construct comprising the residues 1-400(del30) and yielded an atypical, extended conformation of the FERM domain. Moreover, earlier findings suggest that ~30 residues in the F1 loop are unstructured in the context of the talin head F1F2F3 (residues 86-410) (Barsukov et al., 2003). As our FERM model and experimental results were in contradiction with both of these previously reported findings, we proceeded to study talins t1-405, t1-400(del30) and t1-405(del37/GAG) in hydrogen/deuterium exchange (HDX) mass spectrometry. A slow exchange of hydrogens would indicate the structure to be compact in a way that protects buried residues from contact with the solvent. Indeed, the results of the HDX experiments showed that during the measurement intact talin head 1-405 exchanged approximately 20 percent fewer hydrogens than the construct used in the crystal structure (t1-400(del30)). The other loop-deleted talin, t1-405(del37/GAG), was more accessible to solvent than wild-type t1-405, but less than t1-400(del30). SAXS results supported this finding: We found t1-405(del37/GAG) to be the most compact of the talin forms studied.
Figure 21. Experiments were conducted using a set of talin head constructs. A) Biophysical characterization was carried out for talin forms with S405 (KKKKS) as the last residue, with the exception of t1-435 (not shown) and t1-400(del30). Cell assays were carried out for the same talin forms, but with Q435 as the C-terminal residue, indicated with a tail of hyphens. All constructs had an N-terminal tag (solid line). B) SAXS analysis of the wild-type talin head t1-405 suggests the protein to be somewhat flexible, and more than half of the molecules of the sample to exist in a FERM-like conformation. Four conformations were selected in EOM to represent fractions of the conformational ensemble: a) 56%, b) 18%, c) 18%, and d) 9%. Adapted from Article III Supplementary Information.
5.4 F1 loop function

The F1 loop has been previously described as a membrane anchor that, by formation of an α-helix when in contact with negatively charged membrane lipids, attaches the talin head to the lipid bilayer (Goult, Bouaouina, et al., 2010).

When modeling the F1 loop in a FERM-folded talin head, only one of the ten conformations of the NMR structure of talin F1 fit the FERM domain fold without steric clashes and was therefore selected. The complete talin head 1-405 including an F1 loop was simulated for 23 ns in solution, before the αIIBβ3 integrin dimer was added to the model. Already in this initial conformation the talin D154 and αIIB R995 residues were at a sufficiently short distance for charged interaction (2.6 Å between charged heavy atoms). In all simulations of the FERM talin integrin and lipid membrane an R995-D154 salt bridge stabilized and remained throughout the simulation. The neighboring E155 residue of the F1 loop occasionally formed contacts with the αIIB R997 residue.

The human talin F1 loop (residues M134-E170) contains 11 positively charged residues (9×K, 2×R) and 11 negatively charged residues (7×E, 4×D). The positively charged residues have been found to anchor talin F1 to the membrane and are suggested to stabilize an α-helical conformation upon contact with negatively charged phospholipids (Goult, Bouaouina, et al., 2010). The L152 side chain dug deep into the hydrophobic core of the membrane in all four simulations. The D154 residue formed a stable salt bridge with the R995 residue in all simulations, suggesting a mechanism of integrin activation by disruption of the integrin inner membrane clasp. Indeed, we found charge neutralization of the D154 and its neighboring residue E155 to markedly impair integrin clustering (Figure 22; Article III Figure 3). Moreover, by cysteine crosslinking we showed that the F1 loop directly interacts with integrin, further supporting our model of F1-loop-mediated integrin activation (Article III Figure 4).

Integrin binding assays with the Octet biosensor showed that the deletion of the F1 loop from a talin head 1-405 does not negatively affect β3 or VE-β3 integrin binding. Instead, the affinity appeared to be enhanced by the loop deletion (Table 4). Moreover, the F2F3 domains alone (t206-405) bound wild-type β3 integrin with a much higher affinity than any other talin form tested, yet the differences between talin forms were less pronounced when probed with the high-affinity VE mutant of the β3 integrin (Pinon et al., 2014).
Figure 22. Schematic figure of the predicted location of the F1 loop in the talin-integrin complex and representative snapshots from a molecular dynamics simulation. MD simulations predicted the F1 loop to insert the L152 side chain into the lipid bilayer and to form a salt bridge between talin D154 and βIIb R995. An additional salt bridge between talin E155 and βIIb R997 was occasionally observed. Adapted from Article III.

Table 4. Dissociation constants (± standard error, number of experiments > 3) of talin-β3-integrin complexes as measured for the talin head and a GST fusion of the complete β3 cytoplasmic domain. The data were collected by immobilizing talin head on the sensor surface of the Octet biosensor and measuring the binding of 0, 20, 80, 320, 1250, and 5000 nM GST-integrin fusion. The dissociation constant was calculated using the one-site specific binding model in GraphPad Prism.

<table>
<thead>
<tr>
<th>talin</th>
<th>WT-β3 Kd</th>
<th>VE-β3 Kd</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1-405</td>
<td>18.7 ±1.9 µM</td>
<td>0.36 ±0.04 µM</td>
</tr>
<tr>
<td>t1-405(del30)</td>
<td>14.0 ±4.7 µM</td>
<td>0.24 ±0.03 µM</td>
</tr>
<tr>
<td>t1-405(del37/GAG)</td>
<td>8.7 ±1.3 µM</td>
<td>0.33 ±0.05 µM</td>
</tr>
<tr>
<td>t206-405</td>
<td>0.10 ±0.07 µM</td>
<td>0.25 ±0.05 µM</td>
</tr>
</tbody>
</table>
6 DISCUSSION

6.1 Why study talin-integrin interaction?

Integrins have been successfully targeted by therapeutics for thrombosis prevention after percutaneous coronary intervention, ulcerative colitis, Crohn’s disease and multiple sclerosis, and the ClinicalTrials.gov registry currently lists 47 studies of integrin-targeting therapeutics for a variety of conditions from cancers to schizophrenia. So far, almost all integrin-targeting drugs in use or in late-stage trials target the extracellular ligand-binding site (Ley et al., 2016), and understanding the intracellular mechanisms of integrin activation may help to develop novel therapeutics to target other steps of the integrin activation process (Shen et al., 2013), for example talin-integrin binding.

Only a few associations of talins and disease have been reported. The S339L point mutation in the F3 domain of talin-2 head causes fifth finger camptodactyly, a type of rare dominantly inherited syndrome manifesting in contraction of the proximal interphalangeal joint of one or both fifth fingers with associated heart problems such as bradycardia (Deng et al., 2016). Structural analysis suggests that the S339L mutation blocks Ƣ1 integrin binding, affecting the conformation of the lysine-rich loop containing K327 (K324 in talin-1) (Yuan et al., 2017). Moreover, although talin is a non-secreted protein, elevated talin-1 levels in serum can be found in multiple sclerosis (Muto et al., 2017), colon cancer (Bostanci et al., 2014) and in urine in Kawasaki disease (Kentsis et al., 2012). Talin-2 is also present in the cerebrospinal fluid of epilepsy patients, but not in healthy individuals (Xiao et al., 2009).

Talin has been found to be upregulated in several types of cancer such as metastatic oral squamous cell carcinoma (Lai 2010) metastatic prostate cancer (Sakamoto et al., 2010), metastatic breast cancer (Beaty et al., 2014), nasopharyngeal carcinoma (Xu et al., 2015) and bevacizumab-treated glioblastoma multiforme (Kang et al., 2015). Malignant tumors modify their extracellular matrix, providing more stiffness of the substrate, upregulating focal adhesion proteins such as talin, and thus allowing active signaling through the integrin-talin mediated complexes. Talin promotes the invasion and anoikis resistance of human prostate cells (Sakamoto et al., 2010), making it a potential target for drug development against prostate cancer metastasis. In prostate cancer talin-1 has been found to promote invasion and anoikis resistance by facilitating the phosphorylation of adhesion proteins such as Src and FAK and downstream activation of the PI3K-AKT pathway (Sakamoto et al., 2010). The PI3K-AKT pathway promotes proliferation and protects the cell from apoptosis and is hence a widely studied pathway in cancers and a target of pharmaceutical development (Osaki, Oshimura and Ito, 2004; Porta, Paglino and Mosca, 2014; Nitulescu et al., 2016). Also the mechanisms of talin-integrin-mediated adhesion are therefore a potential target for cancer therapy.
FERM domains have not been widely targeted in the search of drug molecules, and there are currently no FERM domain-targeting drugs on the market. Inhibitors of focal adhesion kinase 2 (FAK2)-regulated glioma migration have been developed and tested in vitro by Meurice and coworkers (2010). The compounds bind to the FAK2 F3 domain, a site equivalent to the membrane-distal β-integrin binding site in talin (Meurice et al., 2010). Furthermore, the vanilloid compounds divanillin and diapocynin have antimetastatic effects on hepatocellular carcinoma, and they have been suggested to control the autophosphorylation of focal adhesion kinase 1 (FAK) by interaction with the Y397 binding site in the FAK FERM domain. However, the actual binding site was predicted by molecular docking and not confirmed in vitro (Jantaree et al., 2017). In a recent study of talin, a talin head-binding N-acylurea derivative was found to inhibit angiogenesis in human umbilical vein endothelial cells and to decrease the phosphorylation of the talin downstream kinases FAK, AKT (protein kinase B) and ERK (extracellular signal-regulated kinase) (Lim et al., 2017).

6.2 Talin-integrin binding

In Article I, the talin-β-integrin binding study focuses on the membrane-distal integrin-talin interface and omits the membrane-proximal helix that also interacts with the talin F3 domain (Wegener et al., 2007; Anthis, Wegener, et al., 2009; Saltel et al., 2009). The talin-integrin complex becomes subjected to mechanical stress from the integrin ectodomains and from actin filaments bound to talin rod. There are multiple possible models of how talins, integrins and actins become interconnected (Klapholz and Brown, 2017), and the configuration of actin, talin and integrin should affect the direction of the actual force vector. The simulations, in which the talin model was allowed to freely align perpendicular to the force vector, were likely an oversimplification of the complex system in vivo, where the lipid membrane would sterically bar the rotation of the talin head.

In a similar study, Neumann and Gottschalk analyzed the effect of the pulling direction on the rotation angle of the talin F3 domain in a complex of β1D integrin and talin-2 (Neumann and Gottschalk, 2016). As initial conformation they used the Anthis et al. (Anthis, Wegener, et al., 2009) conformation, where no β-sheet-like hydrogen bonds exist between the integrin and talin (see also 5.1.1). Pulling the integrin-F3 complex parallel or perpendicular to the membrane plane, they observed different binding configurations: pulling with a parallel force vector preserved a binding mode similar to the initial configuration, but with a perpendicular pulling direction β1D and talin F3 formed a β-sheet-like, mechanostable complex very similar in conformation to our results reported in Article I.

The majority of talin-integrin studies focus on β3 and β1 integrins, for example αIIbβ3 or αVβ3 integrin. The integrins β1A, β1D, β2, β3, β5, and β7 have been identified as talin binders, whereas β6 binding to talin had not been previously recognized. Our original hypothesis of β6 being a talin binder was based on the sequence similarity in the potential membrane-distal binding sites, where the spacing between the conserved tryptophan and tyrosine residues is
identical to β3 (β3: W739 and Y747). While Article I does not discuss the membrane-proximal binding site, it is noteworthy that β6 also shares high similarity with the MP sites of all known talin binders (Figure 2, Figure 12). Furthermore, β6 binds kindlin-1 (Bandyopadhyay et al., 2012), which together with talin activates integrins (Harburger, Bouaouina and Calderwood, 2009). Taken together, our results suggest that talin binds and activates β6 integrin and that the β6 integrin membrane-distal talin-binding site may adopt a mechanically stable binding conformation distinct from those reported for β3 and β1D (García-Alvarez et al., 2003; Anthis, Wegener, et al., 2009). However, as we later discovered, we were not the first to report the talin binding of the β6 integrin cytoplasmic tail: β6 integrin-talin binding has been shown at least in pull-down experiments in a study (Tseng et al., 2014) that was published already while Article I was in the review process.

6.3 Lipid interactions

Earlier studies have suggested the interactions of talin with PIP2 to be not specific, but rather charge-driven (Moore et al., 2012). More recent results, however, show that the lipid binding by F2F3 is more specific and favors PIP2, whereas the interactions of the F0F1 domains are driven by charge rather than the combination of charge and geometry (Ye, McLean and Sligar, 2016). The Moore and Ye studies agree well on the affinities of F2F3 and F0F1 to a 10% PIP2 bilayer: F2F3 is bound with higher affinity than F0F1. Moreover, the F0F1 and F2F3 domain pairs have little cooperativity in PIP2 binding, as F2F3 alone binds PIP2 with higher affinity than the complete talin head (Moore et al., 2012; Ye, McLean and Sligar, 2016). Our F2F3 domain simulations suggested that one PIP2 lipid can occupy a PH-domain-like phosphoinositide binding pocket formed by the F3 residues R328, R339 and K320. The NMR study of Goksoy and coworkers also shows the F3 domain to bind PIP2 and the L6,7 loop and the β7 strand (residues ~370-380) to be perturbed more by PIP2 binding than the lysine-rich L1,2 loop (Goksoy et al., 2008). The potential PH-domain-like PIP2 binding might explain the reported PIP2-specificity of the F2F3 double domain but not F0F1 (Ye, McLean and Sligar, 2016). However, in our MD simulations, PIP2 bound also several other residues outside the PH-domain-like site, as the F3 domain contains several positively charged residues that can readily interact with the membrane, especially in the lysine-rich L1,2 loop.

Phosphoinositides thus appear to bind talin to the cell membrane as reported several times earlier (Moore et al., 2012; Song et al., 2012; Ye, McLean and Sligar, 2016). MD simulation studies have suggested that the local lipid environment of integrins might regulate integrin activation or clustering (Mehrabd and Mofrad, 2013; Kalli et al., 2016). We found indications that talin, along with PIP2 molecules, could perturb the transmembrane and cytosolic domains of αIIbβ3 integrin. An increase in the local concentration of PIP2 by talin-activated PI5K1 might thus destabilize the inactive conformation of integrin. However, we did not observe direct interference of the integrin IMC by PIP2 molecules with FERM-folded talin.
Our results suggest that FERM-folded talin lands on the membrane plane using all three subdomains of the FERM fold: L1,2 of the F3 subdomain, K277 and surrounding residues in F2 and residues 146-149 and 156-157 in the F1 loop (Article III Supplementary Figure S2).

6.4 Talin FERM

The consensus of the talin research is that the talin head is an atypical FERM domain as shown in the crystal structure of a loop-deleted talin-1 head (residues 1-400(del30); (Elliott et al., 2010)). The crystallization of the complete talin head is not trivial, as shown by the futile attempts of other groups (García-Alvarez et al., 2003; Elliott et al., 2010) and our collaborators (unpublished), and for structure determination by NMR the talin head is too large. This leaves low-resolution structural analyses to determine the shape of the talin head. Both electron microscopy and SAXS analyses showed the talin head to resemble a FERM domain more than the crystal structure of the loop-deleted talin head. Moreover, experiments in cells show that the talin F1 and F2 domains have an interface interaction that cannot be explained by an extended talin fold. In addition, we show that the F1 loop binds β3 integrin in a way that would not be possible without a compact conformation. However, as we still do not have a high-resolution 3-D structure of the talin head explaining these findings, we cannot claim with certainty that the talin head primarily exists as a FERM fold.

6.5 How does talin activate integrin?

One of the recurring questions in all three studies in this thesis and in the vast literature of integrin-mediated adhesion is “How does integrin become activated?” The titles of two articles in this thesis boldly state the importance of PIP2 (Article II) and the talin head F1 loop (Article III) in integrin activation. The current model, based on β3 integrin, is that the integrin ectodomains become extended during activation. Less clear is, however, how the binding of talin to the β-integrin cytoplasmic tail can bring about such a massive change in integrin conformation and how this extended conformation becomes stabilized. Competing, or perhaps complementary, hypotheses involve the lateral separation of the integrin transmembrane domains and the tilting of β-integrin (Luo, Springer and Takagi, 2004; Kim et al., 2012). The potential mechanisms of integrin activation by talin have been explored in several MD simulation studies using coarse-grained or all-atom set-ups and different lipid bilayer compositions; however, none of the earlier studies have included PIP2 in the simulations. Kalli et al. simulated a talin-2 F2F3 (Kalli, Campbell and Sansom, 2011) and a complete, extended talin-1 head (Kalli, Campbell and Sansom, 2013) and proposed a mechanism in which integrin becomes activated by a scissor-like movement of the transmembrane helices. Mehrbod and coworkers report talin F3 binding to the β3 cytoplasmic tail to loosen the αHb-
β3 contacts at the IMC (Mehrbod, Trisno and Mofrad, 2013). Based on a multi-microsecond simulation of the talin F2F3-αIIbβ3 complex Provasi and coworkers proposed that integrin activation involves transmembrane rotation rather than tilting or separation (Provasi et al., 2014). Arcario and Tajkhorshid’s study of a talin F2F3 at a phosphoserine-containing lipid bilayer suggested talin to become anchored to the membrane by phenylalanine residues in the F2 domain (Arcario and Tajkhorshid, 2014). Both Provasi et al. and Arcario and Tajkhorshid suggested that the talin F3 rotates relative to F2. Simulating a complex of talin-2 F3 and integrin β1D without α integrin and lipid bilayer, Neumann and Gottschalk suggested that force applied on the talin-integrin complex would be sufficient to break the integrin IMC salt bridge by rotating the F3 (Neumann and Gottschalk, 2016).

Our simulations suggest the talin F3 domain to loosen the IMC, in line with Mehrbod and coworkers findings, but also that PIP2 lipids may interfere with the IMC salt bridge. Furthermore, our simulations with a FERM-folded talin head model show that the F1 loop interacts with the IMC and that this interaction may involve a salt bridge between the loop and α-integrin. This may suggest a mechanism where the F1 loop interferes with the IMC contacts, destabilizing the inactive conformation of the integrin. However, as the results from Article II suggest, the F3 alone or with PIP2 lipids might be sufficient to perturb the IMC contacts, and hence the role of the F1 loop might be the stabilization of the activated conformation of integrin.

The strength of modeling and molecular dynamics simulations in understanding integrin activation is the ability to generate model systems that would be very difficult to study in vitro or in cellulo. Moreover, molecular dynamics can provide atomic resolution models of biological events, and the conditions (e.g. tension) of the simulation can be freely tuned. Molecular dynamics simulations alone may come with artefacts from modeling or force fields, but can provide novel hypotheses, which in turn can be confirmed in laboratory. As an example, the unfolding of talin rod bundles and subsequent unmasking of vinculin binding sites was first predicted by molecular dynamics simulations (Lee, Kamm and Mofrad, 2007; Hytönen and Vogel, 2008) and later confirmed by single cell biophysical experiments and cell biology (del Rio et al., 2009). Likewise, predictions from molecular dynamics simulations have been successfully combined with atomic force microscopy to analyze the mechanical stability of talin rod bundles (Haining et al., 2016).

While the studies in this thesis provide answers to our research questions, we have not fully solved the mystery of integrin activation, and not even the question of talin head fold. Does the talin head have a stable solution FERM fold and, if not, what controls its organization? How does the crystal structure conformation of talin head fit in the picture? Molecular modeling may hold the key in solving the details.
The central novel findings of the studies presented in this thesis are the following:

1) **Mechanostability of the membrane-distal talin-integrin interface largely depends on the $\beta$-sheet-like backbone hydrogen bonding and not on the NPxY tyrosine.** The NPxY tyrosine dissociated both in the presence and absence of external force.

2) **Talin F3 and PIP2 lipids together may interfere with the integrin inner membrane clasp.** Talin F3 appears to loosen the interactions of the inner membrane clasp, and to recruit PIP2 lipids to the site. The PIP2 lipids in turn may perturb the inner membrane clasp contacts.

3) **The talin head F1 loop directly interacts with $\beta3$ integrin.** Interactions of the F1 loop with the integrin inner membrane clasp were predicted from MD simulations and confirmed with $\beta3$ integrin and talin head in cells.

4) **The talin head has structural flexibility and appears to be intermediate between a globular and an elongated particle in solution.** This shape may indicate the talin head to have a FERM fold in solution.
ACKNOWLEDGMENTS

My greatest appreciation goes to my supervisor Dr Vesa Hytönen, who is not only an enthusiastic researcher with an incredible ability to focus on everything at the same time, but also a truly pleasant person. I would also like to thank my supervisor Dr Jarkko Valjakka for delightful, quirky discussions on protein structure, and Prof Emer Markku Kulomaa for welcoming me in his research group and creating a warm working atmosphere.

This PhD project was funded by the Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB) and by the Academy of Finland via Vesa Hytönen’s project funding. I would also like to acknowledge CSC – IT Center for Science, Finland, for computational resources. The work presented in Articles II and III was granted access to the HPC resources of EPCC made available within the Distributed European Computing Initiative by the PRACE-2IP, receiving funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. RI-283493.

I would like to thank my thesis committee members Dr Tomi Airenne and Prof Janne Jänis for providing valuable insight and reality checks in the PhD project. I would also like to acknowledge Janne Jänis for the great collaboration in mass spectrometry. I have enjoyed the collaboration and correspondence with Professor Bernhard Wehrle-Haller, without whom Article III would not have been possible. I would also like to thank Prof Viola Vogel, Marie-Claude Jacquier, Prof Ilpo Vattulainen, Dr Tomasz Róg, Dr Adam Orłowski, Dr Holland Cheng, Mo Baikoghli, Dr Anne Tuukkanen, Dr Marko Pesu, Dr Hannu Turpeinen, Dr Markku Varjosalo and Xiaonan Liu for collaborations in projects in this thesis and beyond.

I would like to thank all the current and former members of the Protein Dynamics and Molecular Biotechnology groups. I wish especially to thank Latifeh Azizi for an amazing collaboration in the talin head project, Rolle Rahikainen for being my peer in PhD thesis writing and for the inspiring discussions on talins, adhesions and kinases; Dr Tiina Riihimäki, Dr Barbara Taskinen and Dr Soili Lehtonen for enjoyable collaborations in avidin-related projects; and Dr Sanna Auer, Dr Anssi Nurminen, Dr Jenni Leppiniemi, Dr Jenita Rauta, Magdaléna von Essen, Dr Vasy Mykuliak, Dr Juha Määttä and Dr Paula Turkki for keeping up the fun, friendly atmosphere.

I would like to thank my family and especially my father for all the support. Lastly, I wish to thank Kaisa, Anna, and all my friends for all the good moments, and for being there.
REFERENCES


ORIGINAL PUBLICATIONS
Introduction

Cells sense and respond to the rigidity of their environment,1,2 and the interactions with their surroundings control the molecular composition of adhesion complexes, cellular shape and movement.3,4 Focal adhesions are complex structures containing cell membrane receptor integrins and more than 100 different proteins with signaling and scaffolding functions5–6 (Fig. 1A). They mediate cell attachment to the extracellular matrix (ECM), and transmit mechanical forces between the cytoskeleton and ECM.7,8

Integrins exist in conformations that differ in ligand binding ability. They form weak interactions with ligands in the ECM, which cleaves talin and thereby facilitates the disassembly of the adhesion.9 The integrin-bound talin functions as a scaffold that harbor other focal adhesion proteins, and links the integrin-matrix connection to the force bearing structures of the cell by direct binding to actin filaments.10 This scaffold is also the target of adhesion turnover regulation by calpain, which cleaves talin and thereby facilitates the disassembly of the adhesion.11,12

The initial interaction between the integrin-talin complex and actin is weak and breaks at 2 pN at a 60 nm s−1 loading rate.26,27 The ECM–actin contact grows in size and gains strength through the binding of other adhesion proteins, such as vinculin,19–21 and by the lateral association or clustering of integrins, particularly β1.22–25 Early integrin–ECM interactions and clustering may take place independently of the force generating machinery of the cell,11,23,28 but further steps of adhesion formation and maturation are regulated by tension between the ECM and the cytoskeleton.36,37 Within the cell, tension is generated as actin assembly at the adhesion sites pushes the membrane outwards,38–40 and myosin-generated traction forces pull actin filaments inwards from the adhesion sites.6,40 Moreover, substrate stiffness guides the adhesion processes.40 Under tension between the ECM and cytoskeleton, the talin rod domain stretches and shortens cyclically.41 In this process, the talin–integrin interface is exposed to mechanical load in the range of up to 30 pN at physiological loading rates,42 which is sufficient to expose buried vinculin binding sites in the talin rod domain and to thereby reinforce talin–actin binding.43–45

Humans have 24 different integrins, each composed of one of 18 α chains and one of 8 β chains.19,46 In addition,
Fig. 1  (A) Schematic representation of the talin–integrin complex at the cell membrane. Focal adhesions connect the extracellular matrix (ECM) to the actin cytoskeleton (purple) through an integrin heterodimer. Talin binds to the cytoplasmic tail of β-integrin in the early steps of focal adhesion formation. Several other proteins contribute to the signaling and functional diversity at adhesion sites. (a) Talin–integrin complex at a lipid bilayer. Talin subdomains F2 and F3 are shown in gray, and integrin β3 cytoplasmic tail in green. Figure rendered using the model of αⅡbβ3/β1D chimeric integrin and talin-2 published by Kalli and coworkers.59 (b) A closer view of talin-F3 subdomain in complex with β3-integrin in the 1MK7 structure. (B) β-Integrin cytoplasmic domains aligned using MUSCLE 3.8.103 The sequence stretch used in the simulations is highlighted with green in the sequence of the structural template of the models, integrin β3. The conserved tryptophan residue (tyrosine in β5) is labeled as Trp. The talin-binding membrane-proximal (MP) NPxY motif and the membrane-distal (MD) NxxY motif are highlighted in yellow. Residue numbers are provided for the C-terminal residues. (C) Dynamics of the of the integrin tail docked to the F3 domain of talin. Main chain presentation of integrins (green) shown every 5 ps in the last 200 ps before starting the SMD simulations.
alternative splicing has been found for example in β1 integrin. Moreover, the two talins – talin-1 and talin-2 – may serve separate functions, as talin-2 knockout develops a myopathy, whereas a talin-1 knockout is embryonic lethal. Talin-1 is known to bind β1-integrin variants β1A–β1D, but also integrins β2, β3, β5, β6, and β7. The talin-binding site in β-integrin cytoplasmic tails is composed of a membrane-proximal χ-helix with conserved charged and aromatic residues, and a membrane-distal site with a conserved tryptophan (W739 in β3) and an NPxY motif (Fig. 1A and B). It has been proposed that the membrane-distal site provides the first contact to talin, followed by membrane-proximal χ-helix binding and subsequent reorganization of the contacts between the z and β integrin transmembrane helices.

Integrins β1A, β1D, and β3 have different affinities for talin, and show differences in binding conformation. Moreover, a recent report suggests that β1 and β3 integrins may have separate functions in adhesion, with β1 forming strong adhesions through clustering, and β3 functioning in mechanotransduction and binding only transiently to talin. The cytoplasmic domain of β-integrins also serves as a binding site for a number of other cytoplasmic proteins, among which at least α-actinin, filamin, and tensin also directly connect integrins to the actin cytoskeleton. While talin predominantly links β3 integrins to actin filaments within the peripheral cell adhesion sites, α-actinin can replace talin through its overlapping binding site on integrin tails. α-Actinin thereby contributes to the adhesion maturation, dynamic force generation and force release. Filamin also competes with talin for β-integrin binding, and increased filamin binding blocks integrin activation by talin. Consequently, the sequence differences of β-integrins (Fig. 1B) and the mechanical stability of the integrin–talin linkage may contribute to how adhesion site formation and breakdown is regulated by mechanical signals.

The binding conformation of the membrane-distal site has been shown for the β3-talin1 complex, and of the membrane-distal and membrane-proximal sites for the β1D-talin2 complex. Yet, such structural knowledge of other β-integrins is not available. Moreover, what happens in integrin–talin complexes as they are exposed to mechanical load is poorly understood, leaving a gap in the understanding of the force-regulated formation and maturation of adhesion sites. In an attempt to bridge this gap, we utilized all-atom molecular dynamics simulations to study the interactions in the membrane-distal talin binding sites – the presumed first interactions that occur between talin and integrin – of different β-integrins under tension between the extracellular matrix and the cytoskeleton.

β6 is among the less studied integrins, and has not been previously reported to bind talin-1. Integrin β6 is exclusively epithelial, generally downregulated in healthy cells and upregulated in tissue injury and tumors (reviewed by Bandyopadhyay and Raghavan). β6 has been shown to interact with kindlin-1, a protein that assists talin in the activation of integrin, and studies show that β6 and talin-1 localize to focal adhesions. Moreover, latent TGFβ1 activation by αvβ6 has been shown to require a connection to both the ECM and the cytoskeleton, suggesting a force-driven activation mechanism and a potential role for talin in the process. We were therefore intrigued to find that it formed a relatively stable complex with talin-1 in simulation, thereby motivating us to study the binding of both β6 and our reference molecule β3 to talin in vitro.

Results and discussion

Initial conformations

The cytoplasmic domains of integrins β1A, β1D, β3, β5, β6, and β7 contain the key residues (W739(β3) and NPxY(β6) required for interaction with talin-1 (Fig. 1B). As the structure of the β3-integrin cytoplasmic tail in complex with talin-1 is known (PDB ID: 1MK7), we utilized this structural information to model the talin-1 complexes of the six different β-integrins. Steered molecular dynamics (SMD) simulations using these complexes revealed how differences in β-integrin cytoplasmic tail sequences affect the stability of their complexes with talin. We first ran three 1 ns simulations without external force for each complex before starting the SMD pulling simulation. This was done in order to let the β3-talin-1 structure based models settle, and to obtain baseline trajectories. The baseline trajectories obtained from these simulations were analyzed by measuring the area of the talin–integrin interface, the number of hydrogen bonds between the integrin peptide and the talin head, and the root mean square fluctuation (RMSF) for Cα atoms (Table 1), and of the three structures obtained the one with the largest interface area and most hydrogen bonds was selected as the starting structure for the SMD simulations with external force. The baseline measurements showed that the initial conformations were relatively stable for β3 and β5 peptide complexes with talin. Visual inspection of the complexes showed that β7 integrin peptide took a distinct, loosely bound conformation (Fig. 1C). The crystal structure of the β1D-talin-2 complex (PDB ID: 3GW9) shows a different talin binding conformation for the integrin peptide. This structure has been solved for talin-2 instead of talin-1, and it is unclear whether the same binding conformation would exist in talin-1. To address this question, we prepared another set of simulations for the β1A and β1D complexes, from here on called β1A-3GW and β1D-3GW.

<table>
<thead>
<tr>
<th>Integrin</th>
<th>H-bonds ± SD (Å)</th>
<th>RMSF ± SD (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β1A</td>
<td>571 ± 38</td>
<td>4.1 ± 1.4</td>
</tr>
<tr>
<td>β1D</td>
<td>564 ± 53</td>
<td>3.7 ± 1.4</td>
</tr>
<tr>
<td>β3</td>
<td>601 ± 33</td>
<td>4.2 ± 1.3</td>
</tr>
<tr>
<td>β5</td>
<td>642 ± 65</td>
<td>4.6 ± 2.2</td>
</tr>
<tr>
<td>β6</td>
<td>579 ± 69</td>
<td>2.8 ± 1.0</td>
</tr>
<tr>
<td>β7</td>
<td>514 ± 66</td>
<td>2.0 ± 1.1</td>
</tr>
<tr>
<td>β1A-3GW</td>
<td>749 ± 46</td>
<td>4.1 ± 1.4</td>
</tr>
<tr>
<td>β1D-3GW</td>
<td>790 ± 58</td>
<td>4.5 ± 1.7</td>
</tr>
</tbody>
</table>
respectively, using the integrin conformation from the 3G9W structure but complexed with talin-1 from 1MK7. The baseline simulations without external force showed milder changes in the talin-1–integrin interface of β1A-3G9W and β1D-3G9W, compared to the interface of β1A or β1D prepared with the 1MK7 talin-1 structure. Moreover, the integrin–talin binding interface area was larger in β1A-3G9W and β1D-3G9W (Table 1).

**Contact lifetimes in constant force pulling simulations**

The complexes were subjected to constant force pulling in a set of molecular dynamics simulations using forces ranging from 200 pN to 600 pN. Several individual contacts in the talin–integrin interface contribute to the force resistance of the complex. The individual contacts were dynamically switching between bound and unbound states, and therefore their lifetimes do not have clearly defined endpoints. This is why we evaluated the force dependency of these contacts by measuring distances between selected atoms in the talin–integrin interface throughout the constant force pulling simulations. The distances were plotted as histograms showing the time in picoseconds spent at each distance (Fig. 2A–D). We then identified on-states for the contacts showing sharp peaks in the distance histograms, and estimated the lifetimes of these contacts by counting the observations of the contact within a defined cut-off distance (Fig. S1–S4, ESI†). Four contacts along the talin–integrin interface were used for this measurement. The packing of the side chain of the conserved integrin tryptophan residue with R358(tal) was assessed by measuring the distance between the centers of mass of the tryptophan and R358(tal) side chains (Fig. 2A). Next to the tryptophan residue, the D740(b3) and W359(tal) residues form 1–2 backbone hydrogen bonds with each other (Fig. 3), and the distance between their Cα atoms was used as an indirect measure of this contact (Fig. 2B). The A742(b3) residue forms an additional backbone hydrogen bond to K357(tal) in some of the simulations (Fig. 3B), and this extended backbone hydrogen bonding was measured using the A742(b3)–K357(tal) Cα–Cα distance (Fig. 2C). Finally, the binding of the integrin NPxY tyrosine side chain to its binding groove was measured using the distance of the Y747(b3) side chain center of mass from the backbone carbonyl oxygen of N355(tal) (Fig. 2D).

The histogram of the D740(b3)–W359(tal) distance showed a sharp peak around 6 Å (Fig. 2B), which indicates the presence of a bound state. The Y747(b3)–N355(tal) forms a less pronounced peak (Fig. 2D), which reflects the less tight binding and the variety of conformations that the NPxY motif took in the simulations. For each simulation, the total time spent within a cutoff distance of 5 Å (Fig. S1, ESI†) or 7 Å (Fig. S2–S4, ESI†).

Using the D740(b3)–W359(tal) distance as a measure (Fig. 2B), our results grouped β3, β5 and β6 together as forming the most stable complexes with talin, whereas β7 dissociated most rapidly.

**Fig. 2** Talin–integrin distances along the binding interface in the simulations. Solid lines: simulations based on the β3–talin-1 structure (1MK7). Dashed lines: simulations based on the β1D–talin-2 structure (3G9W). Distances between two residues in four contacts (A–D) are shown as histograms using a bin width of 0.2 Å. The atoms or centers of mass used in distance measurement are shown in the 1MK7 structure, with a purple sphere representing the measuring point in integrin (green) and a blue sphere representing the measuring point in talin (white). The definition of the bound state shown in Fig. S1–S4 (ESI†) is highlighted with gray. (A) Histogram of distances between the side chain centers of mass of the conserved tryptophan W739(b3) (Y743 in β5) and R358(tal). (B) Histogram of Cα–Cα distances between D740(b3) and W359(tal). (C) Histogram of Cα–Cα distances between A742(b3) and K357(tal). (D) Histogram of distances between NPxY tyrosine side chain center of mass and N355(tal) backbone oxygen atom.
The talin complexes of β1A, β1A-3G9W, and β1D-3G9W all had short lifetimes, and β1D formed clearly more stable complexes than the other β1 peptides in the simulations.

The membrane-proximal site may function as a second binder after the first interaction with the membrane-distal site is formed, and its importance for talin binding has been estimated to be greater for β3 than for β1A integrin. Although we excluded the membrane-proximal binding site from this study, β3 appeared to bind talin more tightly than β1A. According to previous reports, the affinity of β3 to talin-1 is slightly above that of β1A, with β1D having the highest affinity. Taken together, the membrane-distal site of β3 appeared to be more force-resistant in our simulations than one might predict from equilibrium affinities.

In the case of integrins β3, β5, and β6, individual 200 pN simulations and one 300 pN simulation showed particularly long (>30 ns and approximately 17 ns, respectively) bound states. To identify the interactions that might account for the stability of the complex, we focused on the hydrogen bonding patterns of these simulations. Although the number of very stable (lifetime more than 50% of simulation length) hydrogen bonds in these simulations varied between 2 and 4, only one backbone hydrogen bond was common to all these simulations, namely the bond between D740(β3) amide group and W359(talin) carbonyl oxygen (Table 2).

**Hydrophobic contacts**

To evaluate the contribution of hydrophobic effect in the talin–integrin interaction under mechanical stress, we analyzed the buried area of hydrophobic residues within the interface (Fig. S5, ESI†). The buried hydrophobic area remained relatively constant in most simulations, and started to decrease as the hydrogen bonds along the talin–integrin interface were ruptured. The buried hydrophobic area of β1–talin complexes was clearly larger in the 3G9W-based simulations (~500 Å²) than in the 1MK7-based simulations (~300 Å²), but the 3G9W-based simulations failed to show greater force resistance. Furthermore, two simulations with a relatively stable bound state, i.e. the β5–talin and β6–talin complexes (Fig. S5, ESI† gray), showed a slight increase in buried hydrophobic area.

---

**Fig. 3** β3–Integrin–talin complex in a representative 300 pN SMD simulation. (A) Dynamics of hydrogen bond formation and breakage between integrin β3 and talin. The hydrogen bonds were determined every 5 ps using cutoff angle and distance constraints of 30.1° and 3.51 Å, respectively. Hydrogen bonds in on-state are indicated as black bars. Left: donor–acceptor atom pair, integrin in green and talin in black. Hydrogen bonds formed by main chain atoms are marked with *. An on-state percentage for each hydrogen bond is shown on the right. Snapshots (B–C) from the simulation illustrate the hydrogen bonding interactions at 1 ns and 2 ns. Green: main chain presentation of integrin. White: talin shown as cartoon model.
Support the importance of the W739(**a** W739A mutation has been reported to inhibit 3.3 Å (1MK7, residue is replaced by a tyrosine (Fig. 1). Earlier studies also to form between the A742(**a** to talin R358(tal) (Fig. 2A), and together with the D740(**b** (3G9W (3G9W correspond to bound states with one or two backbone hydrogen bonds formed between D740(**a** and W359(tal). For example, the corresponding residues in the β5–talin complex, A746(β5) and K357(tal), formed a hydrogen bond that appeared to contribute to the stability of the complex in two long simulations (Table 2). Intriguingly, this A742(β3) position is also the only spot that differs between the β1A and β1D membrane-distal binding site sequences (Fig. 1B). Furthermore, in the 1MK7-based simulations of β1 integrins, the backbone hydrogen bonding between β1D and talin was clearly more force resistant than between β1A and talin (Fig. 2B, C, 4A and D).

**Role of glutamine residue Q778 in the muscle-specific β1D**

Integrin β1D reportedly\(^{5,55,60,85}\) binds more tightly to the talin head compared to β1A. In our 1MK7 structure based simulations, β1A dissociated on average four times as rapidly from talin as β1D (Fig. S2, ESI†). In the 3G9W-based conformation, however, both β1 integrin variants showed a similarly weak force resistance as β1A in the 1MK7 conformation (Fig. S2, ESI†). What might then explain the observed higher force resistance of β1D in the 1MK7 conformation?

The crystal structure of the β1D–talin-2 complex\(^{55}\) shows that integrin β1D binds to talin-2 in a conformation different from that of β3 and talin-1, with Cz RMSD 5.3–5.4 Å for this 14-residue segment of β1D integrin. The differences in integrin backbone conformation bring the integrin into loose contact with talin in terms of backbone hydrogen bonding; the aspartic acid in integrin (β3: D740, β1D: D776) and W359 in talin are bound by two hydrogen bonds in 1MK7, and one in 3G9W.

The Q778(β1D) side chain in 3G9W is oriented towards the solution and does not form interactions with talin-2, whereas its counterpart A742(β3) in the 1MK7 structure is buried. In addition, the region between W775(β1D) and the NPxY motif in the β1D in the 3G9W structure is loosely bound, whereas β3 in 1MK7 is in closer contact with talin in this region. This difference was also seen with the simulations of β1D: the β1D-3G9W simulations showed no hydrogen bonding between Q778(β1D) and talin-1, whereas in the 1MK7-based simulations, the Q778(β1D) side chain formed a network of hydrogen bonds to T354(tal), I356(tal), and L353(tal) (Fig. 4F–D). Intriguingly, this side in the 1MK7 and 3G9W structures is occupied by the NPxY asparagine (N744 in β3, N760 in β1D), and the NPxY asparagine was also bound in our simulations of the 1MK7-based β1D, β6, and β7. With the combination of backbone and side chain hydrogen bonds via Q778, the 1MK7-based β1D–talin complex was relatively force resistant, but the binding of the NPxY was weak in this conformation (Fig. 2D and S4, ESI†). This may suggest that in β1D, the Q778 side chain can occupy the binding site of the NPxY asparagine N780(β1D) (Fig. 4E), and thus drive the NPxY motif to dissociate from its binding groove.

**The mechanical stability of NPxY motif binding varies between integrins and their conformations**

The NPxY motif showed weak binding in our simulations: in all 1MK7-based systems the motif dissociated before hydrogen bonds between the D740(β3) and W359(tal) were broken. In the

**Table 2** Lifetimes (ns) of selected stable talin–integrin hydrogen bonds from the simulations: β3–Integrin numbering is used for the residues of all integrins. Side chains participating in hydrogen bonding are marked with ‘sc’.

<table>
<thead>
<tr>
<th>H-bond</th>
<th>β3 (200 pN)</th>
<th>β3 (200 pN)</th>
<th>β5 (200 pN)</th>
<th>β5 (300 pN)</th>
<th>β6 (200 pN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D740–W359</td>
<td>30.0</td>
<td>34.1</td>
<td>28.3</td>
<td>14.6</td>
<td>52.3</td>
</tr>
<tr>
<td>W359–D740</td>
<td>30.6</td>
<td>28.3</td>
<td>27.2</td>
<td>14.6</td>
<td>—</td>
</tr>
<tr>
<td>A742–K357</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>9.5</td>
<td>—</td>
</tr>
<tr>
<td>A361–K738</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>38.6</td>
<td>—</td>
</tr>
<tr>
<td>W359–T741</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>39.8</td>
<td>—</td>
</tr>
<tr>
<td>T743–K357</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>42.5</td>
<td>—</td>
</tr>
<tr>
<td>N744(sc)–T354</td>
<td>—</td>
<td>—</td>
<td>13.4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>N744(sc)–I356</td>
<td>—</td>
<td>—</td>
<td>11.5</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

During the simulation, yet the area remained smaller than in the 3G9W-based simulations of β1A and β1D. The entropy-driven hydrophobic effect may thus be of lesser importance for the mechanical stability, in agreement with earlier studies.\(^{79–81}\) However, reliable estimation of entropic contribution under applied mechanical force would require extensive validation and is beyond the scope of the current study.

**Backbone hydrogen bonding and conserved tryptophan are important for mechanical stability**

All published talin–integrin complex structures\(^{15,55,61}\) share a close contact between D740(β3) and W359(tal) backbones, and their Cz–Cz distance varies between 5.1 Å (2H7E\(^{15}\)) and 6.1 Å (3G9W\(^{55}\)). Our analysis of the Cz–Cz distance for D740(β3) and W359(tal) showed two pronounced peaks (Fig. 2B) that correspond to bound states with one or two backbone hydrogen bonds between the two residues. The first peak, around 5.6 Å, corresponds to the 2-bond conformation seen in the 1MK7 structure (5.4–5.5 Å), and the second, around 6 Å, to a 3G9W-like 1-bond conformation (5.9–6.1 Å). The backbone hydrogen bonds formed between D740(β3) and talin appeared to be stabilized by the interactions of its neighboring residue, W739(β3). The W739(β3) forms a conserved\(^{15,55,61}\) interaction with talin R358(tal) (Fig. 2A), and together with the D740(β3)–W359(tal) hydrogen bonds it formed a cluster of contacts that appeared to determine the lifetime of the talin–integrin complex. This was also true for β5, in which the tryptophan residue is replaced by a tyrosine (Fig. 1). Earlier studies also support the importance of the W739(β3)–R358(tal) interaction: a R358A mutation in talin-1 reportedly reduces the binding of integrin β3,\(^{61,82}\) β1α, and β1D,\(^{80}\) and the corresponding mutation in Drosophila abolishes the talin–integrin interaction.\(^{83}\) Similarly, a W739A mutation has been reported to inhibit β3-integrin to talin-1.\(^{84}\)

The crystal structure of the β3–talin-1 complex\(^{61}\) shows that the A742(β3) and K357(tal) residues are sufficiently close to be able to form a third backbone hydrogen bond between talin and β3–integrin. Further analysis of crystal structures shows that the Cz–Cz distance of this residue pair varies between 5.3 Å (1MK7, β3–integrin–talin-1) and 9.9 Å (3G9W: chains B,C; β1D–intigrin–talin-2). The 1MK7 structure allows a close contact to form between the A742(β3) and K357(tal) residues, as was observed in many of the longer simulations. This contact was less stable than the D740(β3)–W359(tal) contact, and dissociated before D740(β3)–W359(tal). For example, the corresponding residues in the β5–talin complex, A746(β5) and K357(tal), formed a hydrogen bond that appeared to contribute to the stability of the complex in two long simulations (Table 2). Intriguingly, this A742(β3) position is also the only spot that differs between the β1A and β1D membrane-distal binding site sequences (Fig. 1B). Furthermore, in the 1MK7-based simulations of β1 integrins, the backbone hydrogen bonding between β1D and talin was clearly more force resistant than between β1A and talin (Fig. 2B, C, 4A and D).
b1A-3G9W and b1D-3G9W simulations, on the other hand, binding of the NPxY motif was more stable than in the 1MK7-based b1 simulations (Fig. 2D and S4, ESI†). This implies that in the 3G9W conformation, the NPxY motif dominates the binding, but does not provide as high force resistance as the D776(b1)–W359(tal) backbone hydrogen bonding in the 1MK7 conformation.

This first NPxY motif in β-integrins is reportedly important for talin binding: mutation of Y783(b1) in b1A or b1D to alanine reduces talin binding,60 and the same mutation to Y759(b7) abolishes it completely.56 However, mutation of this tyrosine to alanine has a milder effect on talin binding affinity in b3 than in b1A and b1D,60 and structural analysis shows diversity in how the tyrosine is inserted in its binding site.55,60,61 Our results thus suggest a most intriguing possibility: while the NPxY motif is important for the formation of the talin–integrin complex, its contribution to the force resistance of the complex may be considerably smaller. Moreover, it might be possible for the NPxY motif to interact with other integrin adaptors, such as filamin64,86 while integrin still remains complexed with talin.

**β6-Integrin binds talin**

While the majority of the talin–β6-integrin complexes analyzed under mechanical stress showed similarly poor stability as talin–β1A-integrin, one of the 200 pN simulations of the talin–β6-integrin complex produced a tightly bound conformation (Fig. 5). This complex was characterized by a stable (over 60% of the time in on-state) backbone hydrogen bonding pattern of Q734(b6)–W359(tal), A361(tal)–K732(b6), W359(tal)–T735(b6), and T737(b6)–K357(tal) (Table 2 and Fig. 5). The NPxY tyrosine Y741(b6) did not form stable interactions to its binding groove, and a dynamic salt bridge connected the R742(b6) to D372(tal).

**Fig. 4** Hydrogen bonding of b1 variants to talin in representative 300 pN constant force pulling simulations. Simulations of complexes based on the 1MK7 are shown in A–C (b1A) and D–F (b1D). A simulation of b1D based on the 3G9W structure is shown in G–I. Hydrogen bonds were determined once every 5 ps (b1A) or 10 ps (b1D, b1D-3G9W) and plotted in A, D, and G. Yellow color highlights residue Q778(b1D) in the b1D panels. Left: donor–acceptor pair, integrin in green and talin in black. Hydrogen bonds formed by main chain atoms are marked with *. An on-state percentage for each hydrogen bond is shown on the right. Snapshots (B–C, E–F, H–I) from the simulations illustrate the hydrogen bonding interactions at the marked time points. Green: main chain presentation of integrin. White: talin shown as cartoon model. Integrin residues contributing to hydrogen bonding with talin are labeled.
To experimentally confirm that β6-integrin can indeed bind to talin, we immobilized his-tagged talin head on a Ni-NTA surface and studied the binding of streptavidin-conjugated biotinylated integrin peptides. The biosensor data were analyzed by subtracting the signal obtained for peptide-free streptavidin analyzed simultaneously with peptide-conjugated streptavidins. Our biosensor experiment revealed clearly more β6 binding to the talin surface as compared to the β6-SCR peptide and β3 (Fig. 6). The binding on-rate was found to be fast in terms of kinetics, and virtually all of the bound streptavidin–peptide conjugate was released from the sensor within few seconds after moving the sensor to bare buffer. These findings suggest that β6-integrin can bind talin in vitro with higher affinity than β3. β6-Integrin has to date not been reported to be a talin binding protein. However, the β6-integrin cytoplasmic tail contains the conserved tryptophan and the NPxY motif recognized as important for talin binding (Fig. 1B), and recent studies show that integrin β6 colocalizes with talin-1 at adhesion sites.74,76 Integrin αvβ6 reportedly activates latent TGFβ1 by binding to the latency-associated protein (LAP),73 and the connection to TGFβ1
signaling has been shown for tissue fibrosis,\textsuperscript{87,88} acute lung injury,\textsuperscript{89} pulmonary emphysema,\textsuperscript{90} and for example colon and cervical squamous cell carcinomas.\textsuperscript{91,92} Moreover, depletion of functional \( \alpha v \beta 6 \) reportedly causes lung emphysema, skin infection, and periodontal disease in mice,\textsuperscript{72,90,93} possibly by interfering with latent TGF\( \beta 1 \) activation.\textsuperscript{88} Therefore, understanding also the interactions of the \( \alpha v \beta 6 \) cytoplasmic tail may help to develop therapies for a variety of diseases.

**Experimental**

**Homology modeling and molecular dynamics**

Talin–integrin \( \beta 3 \) chimera (PDB ID: 1MK7,\textsuperscript{81} chains B and C) was used as a template in homology modeling. Homology modeling was carried out using Homodige in the Bodil Modeling Environment.\textsuperscript{94} All models were inspected visually and side-chain rotamers of integrin residues differing in sequence from \( \beta 3 \) were manually selected from the rotamer library\textsuperscript{95} implemented in Bodil. Side-chain rotamers of K357(tal) and E386(tal), and I396(tal) were adjusted in some of the complexes to allow contact with integrin and to abolish steric clashes.

The obtained models were hydrogenated using Psfgen tool in VMD\textsuperscript{96} and moved to a box filled with explicit TIP3 water molecules.\textsuperscript{97} Water molecules resolved in 1MK7 were included in the system. Physiological ionic strength and neutral total charge was obtained by adding 150 mM Na\textsuperscript{+} and Cl\textsuperscript{-} ions to the system. The system was then subjected to two 4000-step conjugate gradient energy minimizations with NAMD\textsuperscript{98} using the CHARMM22 force field,\textsuperscript{99} first with all protein atoms fixed, and then with all atoms released to move. Temperature of the system was gradually increased to 310 K under 1 atm pressure in 31 ps using the Berendsen barostat.\textsuperscript{100} The system was then subjected to equilibrium simulations under 310 K constant temperature and 1 atm pressure.

**Steered molecular dynamics**

To study the force-resistance of the complex, constant force was applied to the complex. Force was applied to the N-terminal C\textsubscript{z} atom of the integrin fragment and to the C\textsubscript{z} atoms of residues L400(tal), G371(tal) and N355(tal) in talin F3 domain.

The charged group of integrin C-terminus was found to make contacts to integrins \( \beta 3, \beta 5, \) and \( \beta 6 \) in the SMD simulations. We thus ran a new set of simulations with termini for each of these three complexes to avoid the problem of false ionic contacts at the termini, and the terminus-free data were used for the analyses.

**Molecular dynamics data analysis**

**Hydrogen bonding analysis.** The trajectory from NAMD simulation was subjected to hydrogen bonding analysis in VMD. The analysis was based on the script by Anishkin,\textsuperscript{101} and cutoff values 3.51 Å for donor–acceptor distance and 30.1° for deviation from 180° for the donor-hydrogen-acceptor angle were applied. For on-state percentage calculations of hydrogen bonds, the end of the simulation was defined as the last frame containing at least one hydrogen bond between talin and integrin.

**Solvent–buried area of the complex.** The area buried from solvent in each talin–integrin complex was calculated with a 1.4 Å scanning probe using the ‘sasa’ function in VMD. The last 200 frames of the simulations without constant force pulling were used for the analysis. Average and standard deviation were calculated from three simulations for each \( \beta \)-integrin–talin complex.

**RMSF calculation.** The last 200 frames of the simulations without constant force pulling were used superimposed using the C\textsubscript{z} atoms of talin residues 311–395 as a reference. RMSF was calculated for integrin C\textsubscript{z} atoms in the 200 frames. Average and standard deviation were calculated from three simulations for each \( \beta \)-integrin–talin complex.

**Contact stability.** Contact stability along the integrin–talin interface was analyzed using the C\textsubscript{z}–C\textsubscript{z} distance of \( \beta 3\)-integrin D740(\( \beta 3 \)) and A742(\( \beta 3 \)) from talin W359(tal) and K357(tal), respectively. In addition, we calculated the distance between the centers of mass of W739(\( \beta 3 \)) and R358(tal) side chains, and the distance of Y747(\( \beta 3 \)) side chain center of mass from N355(tal) carbonyl oxygen. For other integrins, the residues at corresponding positions were used. The distances were collected into histograms using a bin width of 0.2 Å. Bound state lifetimes were estimated by counting the observations of C\textsubscript{z} distances within a 7 Å cutoff.

**Production of recombinant talin head**

The production of his-tagged talin head protein is described in ref. 102. In brief, his-tagged human talin-head was generated by inserting residues 1–406 of human talin1 into pTrcHisC vector (Life Technologies, Carlsbad, California, United States) at the BamHI site and confirmed by DNA sequencing. The protein was expressed in *E. coli* BL21-Star cells at 37 °C. Lysate was prepared using homogenization (Emuliflex C, Avestin Inc, Ottawa, Canada) in 20 mM NaPO\textsubscript{4}, 1 M NaCl, 20 mM imidazole pH 7.4. After clarification by centrifugation, the lysate was applied on HisTrap FF crude 5 ml affinity column using ÄKTA Purifier (GE Healthcare), washed, and eluted with linear imidazole gradient 0–700 mM. Eluted fractions were further purified by cation exchange chromatography using Hitrap SP FF 1 ml column (GE Healthcare). For this purpose, talin-containing fractions were pooled together and diluted (1:10) in 20 mM Tris-HCl 20 mM NaCl pH 7.5 (running buffer). Elution was performed by preparing linear NaCl gradient by mixing running buffer and 20 mM Tris-HCl, 1 M NaCl, pH 7.5. Talin head was eluted at around 550 mM NaCl. Fractions containing talin protein were further concentrated through 30 K filter and analyzed on SDS-PAGE gel. The protein was estimated to be over 95% pure.

**Interaction analysis with Octet biosensor**

N-terminally biotinylated and C-terminally amidated \( \beta 3 \) (HDRKEFAKFFEEARKWDTANNPLYKE), \( \beta 6 \) (HDRKEVAK-FEAERSKAKWQTGTPNLVYRG) and scrambled \( \beta 6 \) (KDWTEHR-QALNSVYFKAGKERKTPARE) peptides were purchased from Caslo (Lyngby, Denmark). Biosensor analyses were performed...
by using Fortebio Octet RED384 instrument (Pall, Menlo Park, CA) using Ni-NTA sensors. Temperature of 25 °C and stirring speed of 1000 rpm were used throughout the experiment. Sensors were chemically activated by immersing them in 0.1 M EDC 0.05 M NHS in H₂O for 100 seconds. His-tagged talin head domain (50 μg ml⁻¹) was used to biofunctionalize the biosensors in 50 mM NaPO₃ 150 mM NaCl pH 7.2, resulting in a binding response of ~8 nm after 300 s incubation. The remaining activated groups were then quenched by 1 M ethanol-amine (pH 8.5) for 100 seconds. Because binding of peptides directly on talin-functionalized surface provided negligible responses, the peptides were first conjugated to streptavidin (Anaspec, Fremont, CA, USA) using molar ratio of 1 peptide per remaining activated groups were then quenched by 1 M ethanol-amine (pH 8.5) for 100 seconds. His-tagged talin head domain (50 μg ml⁻¹) was used to biofunctionalize the biosensors in 50 mM NaPO₃ 150 mM NaCl pH 7.2, resulting in a binding response of ~8 nm after 300 s incubation. The remaining activated groups were then quenched by 1 M ethanol-amine (pH 8.5) for 100 seconds. Because binding of peptides directly on talin-functionalized surface provided negligible responses, the peptides were first conjugated to streptavidin (Anaspec, Fremont, CA, USA) using molar ratio of 1 peptide per streptavidin tetramer. These conjugates were then applied on talin-functionalized surface in various concentrations and each concentration was assayed for 200 s.

Visualization
Molecular structure details were rendered using PyMOL (Schrödinger, LLC), histograms were prepared with Gnumeric (https://projects.gnome.org/gnumeric), and images were further processed with GIMP [http://www.gimp.org/] and Inkscape (http://www.inkscape.org/). Residues were numbered according to expressed protein sequence without signal peptide (UniProt: P05556-1, P05556-5, P05106, P18084, P18564, P626010, β7). Amino acid residue positions in the different β-integrins are referred to using β numbering for clarity.

Conclusions
These are the first data from a simulation to predict force resistance in the binding of β-integrin membrane-distal talin-binding stretch to talin, that is, the assumedly first contacts in the process of integrin activation. The results show a diversity of binding conformations for the β-integrin, and underlines the importance of a close talin contact for the peptide between the conserved tryptophan and NPxY motifs in β-integrin. Our findings also suggest the NPxY motif might have a minor role in the force resistance of the talin-integrin complex. Moreover, our data suggest that integrin β6 may bind talin-1 with moderate affinity.

Acknowledgements
We thank the Academy of Finland and Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB) for funding, Dr Bernhard Wehrle-Haller for comments, and the Swiss Center for Scientific Computing (CSCS), ETH Zurich, and CSC - IT Center for Science Ltd for providing the computational resources for this study. We thank Dr Stuart Prince for proofreading the manuscript text.

Notes and references
PIP2 and Talin Join Forces to Activate Integrin

Adam Orłowski,†‡ Sampo Kukkurainen,‡,†,‡,§ Annika Pöyry,† Sami Rissanen,† Ilpo Vattulainen,†,§ Vesa P. Hytönen,‡,¶ and Tomasz Różycki,†

1Department of Physics, Tampere University of Technology, P.O. Box 692, FI-33101 Tampere, Finland
2BioMediTech, University of Tampere, FI-33520 Tampere, Finland
3Fimlab Laboratories Ltd., FI-33520 Tampere, Finland
4Department of Physics and Chemistry, MEMPHYS—Center for Biomembrane Physics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

Supporting Information

ABSTRACT: Integrins are major players in cell adhesion and migration, and malfunctions in controlling their activity are associated with various diseases. Nevertheless, the details of integrin activation are not completely understood, and the role of lipid in the process is largely unknown. Herein, we show using atomistic molecular dynamics simulations that the interplay of phosphatidylinositol 4,5-bisphosphate (PIP2) and talin may directly alter the conformation of integrin αβ3. Our results provide a new perspective on the role of PIP2 in integrin activation and indicate that the charged PIP2 lipid headgroup can perturb a clasp at the cytoplasmic face of the integrin heterodimer.

1. INTRODUCTION

Multicellular organisms provide well-defined and tightly controlled mechanisms for cell–cell and cell–extracellular matrix interactions. The group of receptors called integrins plays a central role in these mechanisms. Integrins and integrin-mediated processes are essential for normal cell functions such as signaling, cell migration, adhesion to the local extracellular environment, and leukocyte function. Moreover, integrins play a role in cancer progression and metastasis, and certain tumor types have been found to exhibit higher levels of specific integrins. This makes the integrin-associated signal complex an important spot for cancer therapy development.

Integrins are heterodimeric transmembrane receptors that in mammals are composed of 18 different α and 8 β subunits. These proteins respond to both extracellular (outside-in) and intracellular (inside-out) stimuli, connect the extracellular matrix to the cytoskeleton, and pass signals across the plasma membrane in both directions. This requires a properly controlled integrin activation mechanism that involves conformational changes within the integrin heterodimer. In the case of the inside-out signaling, the changes eventually lead to an extended integrin conformation with a high affinity for extracellular ligands. Proper control of integrin activation and thus cellular communication with the external environment is crucial for many physiologically relevant processes. Perturbation of this equilibrium can lead to constitutive activation of the integrin and result in bleeding disorders.

Talin, a 2541-residue-long cytoplasmic protein, is one of the triggers that binds to integrin and activates it (inside-out) in the first stages of cell attachment. Importantly, the conformational changes induced by talin in the transmembrane part of the integrin αβ complex are crucial for the inside-out activation process. More specifically, the N-terminal FERM domain of talin binds to the NPxY motif of the integrin β tail, induces a reorganization of the integrin heterodimer, and contributes to integrin activation. Attachment of talin to a membrane is enhanced by a lipid known as phosphatidylinositol 4,5-bisphosphate (PIP2), which can enforce talin to undergo conformational changes that expose the integrin β-tail binding site in its head domain.

The role of PIP2 in integrin activation is intriguing given that PIP2 is a major phosphoinositide of the inner plasma membrane. Moreover, talin itself binds and activates phosphatidylinositol phosphate kinase type I, and may thereby regulate the local concentration of PIP2 in the membrane. PIP2 is able to regulate many important cellular processes including vesicular trafficking, platelet activation, and organization of the cytoskeleton, and it may control several steps in focal adhesion. PIP2 can also affect protein conformation, target cytoplasmic proteins to the membrane, and stabilize protein oligomers.

In this article, we focus on clarifying the importance of lipid membrane composition in the conformational modulation of integrins. In particular, we focus on the role of PIP2 in the...
Table 1. Composition of the Simulated Systems

<table>
<thead>
<tr>
<th>system</th>
<th>integrin</th>
<th>talin (F0–F3)</th>
<th>talin (F2–F3)</th>
<th>DOPC</th>
<th>PIP2</th>
<th>water</th>
<th>K⁺</th>
<th>Cl⁻</th>
<th>simulation time (µs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIP₂F₀F₁−F₂</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>844</td>
<td>96</td>
<td>113310</td>
<td>703</td>
<td>319</td>
<td>0.5</td>
</tr>
<tr>
<td>PIP₂F₁F₀−F₂</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>440</td>
<td>50</td>
<td>60559</td>
<td>365</td>
<td>169</td>
<td>0.5</td>
</tr>
<tr>
<td>PIP₂F₁F₂−F₀</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>440</td>
<td>50</td>
<td>60559</td>
<td>365</td>
<td>169</td>
<td>0.5</td>
</tr>
<tr>
<td>NO-PIP₂F₀F₁</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>950</td>
<td>0</td>
<td>114470</td>
<td>320</td>
<td>320</td>
<td>0.5</td>
</tr>
<tr>
<td>NO-PIP₂F₁F₀</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>439</td>
<td>0</td>
<td>56631</td>
<td>158</td>
<td>162</td>
<td>0.5</td>
</tr>
<tr>
<td>NO-PIP₂F₂F₀</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>439</td>
<td>0</td>
<td>56631</td>
<td>158</td>
<td>162</td>
<td>0.5</td>
</tr>
<tr>
<td>PIP₂CONTROL−A</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>420</td>
<td>50</td>
<td>56182</td>
<td>358</td>
<td>155</td>
<td>0.75</td>
</tr>
<tr>
<td>PIP₂CONTROL−B</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>420</td>
<td>50</td>
<td>56182</td>
<td>358</td>
<td>155</td>
<td>0.75</td>
</tr>
</tbody>
</table>

*F₀–F₃ and F₂–F₃ correspond to the talin domains included in the systems. The abbreviation PIP refers to systems with PIP₂ (and DOPC) lipids, while NO-PIP refers to systems without PIP₂ lipids. The subscripts A and B stand for two independent runs of the same system. The control systems with PIP₂ but without talin are represented by PIP₂CONTROL.*
Further away from the cytoplasm, the interaction of the $\beta 3$ residue Gly708 with the $\alpha$IIb G972xxxG976 motif establishes the outer membrane clasp (OMC). Importantly, mutations of residues in either of these interaction sites (IMC and OMC) result in perturbed interactions of the transmembrane (TM) helices that can lead to a constantly active integrin. In this respect, our particular interest is the salt bridge Arg995-Asp723 of the IMC.

By using atomistic molecular dynamics simulations, we investigated the interactions between PIP2 and the talin-integrin complex. The number of hydrogen bonds between PIP2 and each of the protein domains in the complex (integrin $\alpha$IIb, $\beta 3$, and talin) are shown in Table 2. A more detailed overview of the interactions between the integrin complex residues and PIP2 lipids is shown in Figure 1 and Table S1 and Table S2, where the number of contacts between the atoms of integrin $\alpha$IIb and $\beta 3$ and atoms of PIP2 during the last 300 ns of the simulations are reported. Two residues were defined as being in contact when two heavy atoms of different molecules were located at a distance of 0.35 nm or less. Some residues in both $\alpha$IIb and $\beta 3$ integrin show a very high affinity toward PIP2. This applies to residues that are already outside the transmembrane helix in the water phase, such as Lys994, Arg995, and Arg997 in integrin $\alpha$IIb, Arg724, Lys729, and Arg736 in integrin $\beta 3$. Interestingly, both from visual inspection (see Figure 2) and from the analyses of the number of hydrogen bonds between the residue Arg995 and PIP2 in time (Table 2), one can conclude that the PIP2 headgroup can find its way to the positively charged residue Arg995 and establish a bond with it through negatively charged phosphate groups. This interaction led to the breakage of the Arg995-Asp723 salt bridge (Figure 2B–C, Figure 3), and it took place spontaneously after less than 300 ns of the simulation time in two out of three simulations, where PIP2 was present. It has been shown that this salt bridge is crucial for the stability of the integrin $\alpha$IIb$\beta 3$ dimer, and once it is broken, integrin activation can take place. The local concentration of PIP2 might thus translate to integrin activation.

3.2. Opening of the Arg995-Asp723 Salt Bridge—Step toward Integrin Activation? Having observed the breakage of the Arg995-Asp723 salt bridge, we were tempted to ask whether in the simulations, where PIP2 breaks the Arg995-Asp723 salt bridge, one can see any signs of the initial steps of the integrin activation process. In order to characterize the integrin activation mechanism in more detail, we analyzed a number of parameters suggested to be associated with the activation process. Specifically, opening of the integrin dimer at the IMC and OMC and tilting of the transmembrane domain have been proposed to take place in the integrin activation process, although recent simulation studies are not conclusive on the mechanism of activation. We therefore analyzed the integrin conformation by measuring the tilt angle of the $\alpha$-helix in reference to the bilayer normal, representing the helix as a vector between the center of mass of the residues at the bottom and the corresponding center of mass of the residues at the top of the helix (Figure S4, Table S4).

The distance between the residues is expected to increase at the IMC and OMC interfaces when the interactions are disrupted. In our analyses, we assume that the distance between the Ca atoms of the residues Val700 and Gly972 measures the stability of OMC, and similarly between Lys716 and Phe993 of IMC. Time evolutions of these distances are shown in Figure S3, and the average values over the last 300 ns of the simulations are depicted in Table S3. The OMC distances remained in the same range whether or not Arg995 and Asp723 were in contact, with the exception of the system PIP CONTROL−A. In PIP CONTROL−A, a DOPC lipid tail entered between Val700 and Gly972 during the first 30 ns of the simulation, possibly suggesting that the OMC contact is not very tight. Separation of the IMC was observed for PIP F0−F3− but its distance remained shorter than in the system NO-PIP F3−A. Therefore, we observe no clear correlation between the existence of the Arg995-Asp723 salt bridge (Figure 3) and the separation of the integrin transmembrane helices.

Another potential indicator of integrin activation is the tilt angle of the integrin $\beta 3$ helix, which in the nonactive state has been shown to be $\sim 25^\circ$ and is stabilized by the residue Lys716, which keeps the helix in a certain orientation by snorkeling with NH$_3^+$ toward the phosphate headgroup region. Time evolution (Figure S4) and average over the last 300 ns of the simulations (Table S4) show that, in three of the cases, the tilt angle is approximately $25^\circ$. In the systems NO-PIP F3−A and NO-PIP F3−B, it is higher (33.37° and 29.43°, respectively); while in the systems PIP F3−A, PIP CONTROL−A, and PIP CONTROL−B, it is lower (18.12°, 18.00°, and 22.22°, respectively). However, these differences in tilt angle are small, and it is difficult to associate the changes in tilt angle with the existence of the Arg995-Asp723 salt bridge.

Experimental and computational data have been reported to be in favor of the view that talin-induced breakage of the Arg995-Asp723 salt bridge is involved in integrin activation. Meanwhile, a recent simulation study questioned the paradigm as to the separation of integrin TM helices upon activation. The authors speculated on the insignificance of the Arg995-Asp723 interaction in talin-mediated integrin activation. Our simulations indicate that this interaction can be disrupted in the presence of ionic lipids such as PIP2, which were not present in any of the previous MD simulation studies of integrin activation by talin. The models we simulated in this work contain only a small part of the integrin, and there are many factors causing stress to the $\alpha$IIb$\beta 3$ interface, including conformational changes of the integrin extracellular domain and its interactions with extracellular proteins. Overall, although we can claim the connection between PIP2 and the initialization of inside-out integrin activation process, the data is not sufficient to clarify
Figure 1. Interactions between integrin and PIP2 lipids. Lipid contact occupancies are mapped on integrin structures in PyMOL (Schrödinger) and presented in separate histograms for the α and β integrins. Occupancy was defined as the percentage of time (during the last 300 ns of the simulation time) when the integrin residue formed contacts with PIP2. Two residues were defined as being in contact when two heavy atoms of different molecules were located at a distance of 0.35 nm or less. The results for PIP_{F2F3} and PIP_{CONTROL} are averages over the two independent simulations (A, B).
the impact of the PIP2-induced breakage of the Arg995-Asp723 salt bridge on integrin activation.

3.3. Role of Talin in PIP2-Mediated Integrin Activation—Joint Action of Talin and PIP2? Interactions of talin with lipids have been shown to be an important factor in integrin activation, but some of the related mechanisms are still not clear. To shed light on this matter, in our simulations we included different talin configurations, containing domains F2−F3 (residues 209−400) and F0−F3 (residues 2−398 Δ139−168), to characterize the interactions that occur between talin and a lipid membrane. Similarly to previous analyses of the contacts between integrin tails and PIP2, we carried out contact and hydrogen bond analyses for talin and PIP2/DOPC (Figures 4 and S5; Tables S5, S6).

Our data show that there are PIP2 binding spots present in the F0 subdomain: the residues Lys15-Met17, Arg35, and Asp55-Lys58 can bind strongly to PIP2. Subdomains F2 and F3 seem to be very well characterized regarding their affinity toward negatively charged lipids, and indeed, our results show that there are regions within this area that interact very strongly with PIP2, including the residues Glu252-Lys284, Lys318-Arg328, Arg339, Lys343, Ile348, and Lys364. Similar regions tend to bind to DOPC (Figure 4), which is in line with recent studies on the talin−integrin in different lipid bilayer compositions.

Figure 2. Schematic picture of the simulated system and central findings. (A) Schematic picture of the talin head domain in complex with the integrin αIIb/β3 transmembrane and cytoplasmic domains. Integrin αIIb (blue) and β3 (red) chains are embedded in a lipid bilayer (gray). Talin head subdomains F3 (green), F2 (yellow), F1 (orange), and F0 (olive green) bind the cytoplasmic tail of integrin β3. PIP2 is shown in the van der Waals (vdW) sphere representation. (B) Snapshot of the Arg995-PIP2 contact from the simulation system PIP2F3-A. The PIP2 lipid interfering with the Asp723-Arg995 interaction is shown in vdw spheres, and other lipids in gray transparent lines. (C) Snapshot of the Arg995-Asp723 salt bridge that is broken by the PIP2−Arg995 interaction from the simulation system PIP2F0−F3. Snapshots were made using the VMD package.

Figure 3. Presence of the Arg995-Asp723 salt bridge. The results show the distance between the Arg995 and Asp723 charged groups. (A) Systems with a membrane containing PIP2 and DOPC lipids. (B) Systems with a membrane containing only DOPC. (C) Systems with a membrane containing PIP2 and DOPC but without talin included in the protein complex. The results shown here were smoothened using a running average over 50 points from a single frame.
and PIP2 lipid headgroup phosphates (see Figure 2C). In the second case where we observed the breakage of the salt bridge mediated by PIP2 (system PIP\textsubscript{F2F3−A}), talin was not in contact with the PIP2 lipid (for details see Figure 2B). This raises a question whether talin only indirectly facilitates the opening of the IMC by bringing PIP2 lipids into contact with integrin, or...
whether it has an active role in the process. In the simulations without talin, Arg995 in integrin $\alpha$IIb did not form contacts to PIP2 lipids, yet its neighboring residue Lys994 was constantly in contact with a PIP2 lipid (Figure 1 and Table S1). This suggests that talin is not required for PIP2 lipids to localize near the integrin dimer and the Arg995-Asp723 salt bridge. However, the binding of talin to $\beta$3 integrin may expose the Arg995-Asp723 salt bridge to an attack by PIP2.

4. CONCLUSIONS

Schematic representation of our main results is shown in Figure 5. We propose that PIP2 can have an impact on integrin activation through interaction with talin and Arg995, in a manner which leads to a breakage of the Arg995-Asp723 salt bridge. Our simulations did not show other clear signs of structural changes in the interaction between the $\alpha$IIb and $\beta$3 chains after the breakdown of the ionic bond.

The integrin activation process involves multiple cytoplasmic proteins, and it is possible that the system simulated here is too simplified for the evaluation of further dissociation of the integrin tails and integrin activation. Nevertheless, the fact that experimentally it has been well established that the lack of this salt bridge leads to integrin activation renders our considerations reasonable. Moreover, we recognized other regions of integrin involved in the interactions with PIP2 that might be of high importance in the conformational changes of integrin complexes.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b06457. Structures of the lipids considered and additional data (PDF)

Figure 5. Schematic representation of (A) the suggested PIP2-talin-integrin interaction, and (B) the suggested activation of the integrin complex mediated by PIP2 and talin. DOPC is depicted with a black and PIP2 with a yellow headgroup. Arrows in the panels A and B indicate the direction of PIP2 movement and possible helix dissociation caused by the breaking of the Arg995-Asp723 salt bridge.

AUTHOR INFORMATION

Corresponding Author
E-mail: tomasz.rog@gmail.com. Tel.: +358 40 198 1010. Fax: +358 3 3115 3015.

Author Contributions
The two leading authors contributed equally to this work.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Academy of Finland (Center of Excellence program, project no. 272130 (IV); grants 136288 and 273192 (VPH)), the European Research Council (Advanced Grant CROWDED-PRO-LIPIDS (IV)), the Competitive Research Funding of the Tampere University Hospital (VH), and the Sigrid Juselius Foundation (IV). CSC–IT Center for Science (Espoo, Finland) is acknowledged for computational resources. Additionally, the Tampere Graduate Program in Biomedicine and Biotechnology (S.K.) and the Graduate School program of Tampere University of Technology (S.R.) are thanked for financial support. This work was granted access to the HPC resources of EPCC made available within the Distributed European Computing Initiative by the PRACE-2IP, receiving funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grand agreement no. RI-283493.

REFERENCES


(23) Song, X.; Yang, J.; Hribar, J.; Ye, S.; Perera, H. D.; Goksoy, E.; Dwivedi, P.; Plow, E. F.; Zhang, R.; Qin, J. A Novel Membrane-
Like Domain in the Talin Head: a Role in Integrin Activation. EMBO J. 2010, 29, 1069–1080.


