Structures and interactions of the leukocyte-specific integrins cytoplasmic tails

Chua, Geok Lin

2013

Chua, G. L. (2013). Structures and interactions of the leukocyte-specific integrins cytoplasmic tails. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/54936>

<https://doi.org/10.32657/10356/54936>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**STRUCTURES AND INTERACTIONS OF THE
LEUKOCYTE-SPECIFIC INTEGRINS
CYTOPLASMIC TAILS**

CHUA GEOK LIN

SCHOOL OF BIOLOGICAL SCIENCES

2013

**STRUCTURES AND INTERACTIONS OF THE
LEUKOCYTE-SPECIFIC INTEGRINS
CYTOPLASMIC TAILS**

CHUA GEOK LIN

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2013

**Structures and Interactions of the Leukocyte-Specific
Integrins Cytoplasmic Tails**

A thesis for the Degree of Doctor of Philosophy
at Nanyang Technological University

CHUA GEOK LIN

Supervised by Asst/Prof. Surajit Bhattacharyya
School of Biological Sciences
Nanyang Technological University

Acknowledgements

No amount of thanks can convey my gratitude and appreciation for my supervisor and mentor Asst. Prof. Surajit Bhattacharyya for his meticulous guidance and resolute patience during my postgraduate years and undergraduate final year project. He has opened the doors to research for me, giving me a valuable opportunity to work as a graduate student in his lab to learn valuable research and writing skills which will greatly help me in my future. I was also recommended by him to work under Asso. Prof. Tan Suet Mien as a project officer after completing my undergraduate studies.

I am also equally grateful to Asso.Prof. Tan Suet Mien for the opportunity to work in his lab and his valuable guidance during my one-year stint as a project officer. During which, I got to understand and appreciate integrin molecular biology that subsequently led to me to further my studies on integrins.

My postgraduate work cannot be so smooth sailing without the generous help from my lab colleagues, Dr. Perna Domadia, who has without reserve, taught me a lot of research techniques and selflessly shared many tips for conducting experiments; and from Dr. Rathi Saravanan for sharing her knowledge and tips which have helped me greatly, and also for the many insightful discussions we shared. I am also grateful to Monalisa Amalraj who has been source of a great help and encouragement for me during the short period of time she has been in our lab.

During my course of study, I am also fortunate to have friends who have unreservedly shared their expertise with me. I am extremely grateful to Dr. Xiao Zhou for generously sharing his knowledge of NMR with me when I ran into difficulties running my NMR experiments. I am thankful for the friendship and valuable encouragement of Manisha Cooray and Xue Zhihong, and also for their generous help during times of need, which have tide me through my postgraduate studies.

I acknowledge the help from Dr. Anirban Bhunia for providing the $\beta 2$ pdb coordinates and help in setting up the Myr- αM TOCSY and NOESY experiments. The αM , αX , $\alpha 4$, $\beta 2$ and paxillin constructs were kindly generated and provided by Dr. Tang Xiaoyan who has also purified some of the αM and $\beta 2$ used in this study. Alok Patra purified some of the αX , $\beta 2$, $\alpha 4$ and paxillin used in this study. I am extremely grateful to Dr. Ye Hong and Dr. Shin Joong for their 24/7 technical assistance and support when I had trouble with the operation of the NMR machine.

The completion of my postgraduate studies would have been impossible without the unwavering support and encouragement from my family, especially my Mum and Dad whose selfless sacrifice and understanding allowed me to selfishly and liberally pursue my interests. My husband, Jiamin, has always been there for me during my most difficult times during my undergraduate and graduate years, giving me immeasurable encouragement, support and love that have given me the strength and courage to accomplish my goals. To my family, I am eternally indebted and grateful.

Table of Contents

Acknowledgements	i
Table of Contents	iii
List of Figures	viii
List of Tables	i
Publications.....	ii
Abbreviations	iii
Abstract.....	iv
Chapter I: Introduction	1
1.1 Afflictions related to the abnormal cell surface protein, Integrin	1
1.2 Diversity of the integrin super family.....	3
1.3 Receptors and ligands of integrins.....	5
1.4 Studies on the function of integrins	5
1.5 General structure of integrins	6
1.6 The domain organization of the α -subunit	6
1.7 The domain organization of the β -subunit	7
1.8 Structural transitions of integrins	8
1.9 Factors affecting integrin structural transitions.....	9
1.10 Structural studies of the open, closed and intermediate conformations of integrin	12
1.11 Detailed structural insights into the extracellular domains of integrin	13
1.12 Recent crystal structures of the extracellular domains of the integrin molecule.....	15
1.13 Observation of global structural transitions in integrin molecules using Electron Microscopy (EM)	19
1.14 Detailed structural insights into the individual domains of integrins	23
1.15 Structural insights into the integrin β -subunit domains	23

1.16 Structural insights into the integrin receptor binding domains	25
1.17 Structural insights into the integrin transmembrane and cytoplasmic regions.....	28
1.17.1 Functional characterization of the Outer Membrane Clasp (OMC)	30
1.17.2 Functional characterization of the Inner Membrane Clasp (IMC).....	32
1.17.3 Structural characterization of the integrin OMC and IMC.....	33
1.18 Regulation of integrin affinity and avidity through inside-out and outside-in processes	51
1.19 Bi-directional convergence of outside-in and inside-out signaling on the integrin cytoplasmic tails	56
1.19.1 Sequence similarity in the cytoplasmic tails postulated sites of integrin clasps and postulated sites of ligand binding	57
1.19.2 Phosphorylation as another mechanism of cytoplasmic control of activation	59
1.19.3 Mechanism of integrin activation by talin, a well studied inside-out activator of integrins	63
1.20 Objective of study: Structural characterization of α -subunit cytoplasmic domains of β 2 integrins.....	70
1.21 Strategies to study the α -subunit cytoplasmic domains	72
1.21.1 Use of a N-terminal myristoylation	72
1.21.2 Use of membrane mimic micelles.....	74
Chapter II: Structures and interactions of the α M β 2 integrin.....	75
2.1 Overview	75
2.2 Materials and Methods.....	77
2.2.1 Synthetic and purified peptides.....	77
2.2.2 Purification and expression of ^{15}N and ^{15}N ^{13}C α M and β 2 tails.....	79
2.2.3 Circular Dichroism (CD) studies	80
2.2.4 NMR derived structure for the myristoylated α M tails.....	80

2.2.5 Rigid-body modeling of the α M β 2 complex	81
2.2.6 Paramagnetic perturbation studies on the free and myristoylated α M tails.....	82
2.2.7 Chemical shift perturbation of the α M- β 2 cytoplasmic tails.....	82
2.2.8 Assignment of backbone amides of α M and β 2	83
2.2.9 Fluorescence resonance energy transfer studies:	83
2.3 Results.....	85
2.3.1 Characteristics of the free α M, Myr- α M and Myr-P α M cytoplasmic tails.....	85
2.3.2 Properties of the free α M, Myr- α M and Myr-P α M cytoplasmic tails.....	85
2.3.3 Structural stabilization of the myristoylated peptides compared to the free α M.....	87
2.3.4 Micelle incorporation of the α M, P α M, Myr- α M and Myr-P α M	89
2.3.5 Secondary structure analyses of the α M, Myr- α M and Myr-P α M	91
2.3.6 Structural characterization of the myristoylated peptides	95
2.3.7 Backbone triple-resonance experiments for α M and β 2 cytoplasmic tails.....	101
2.3.8 Interaction analyses between the α M and β 2 tails by ^{15}N - ^1H HSQC.....	104
2.3.9 Binding interactions between β 2 and α M tails in DPC micelles	106
2.3.10 Mode of interactions of the α M and β 2 tails.....	113
2.4 Discussion	116
Chapter III: Structure and binding interface of the cytosolic tails of the α X β 2 integrin.....	123
3.1 Overview	123
3.2 Materials and Methods.....	125
3.2.1 Synthesis and purification of Myr- α X tails	125
3.2.2 Expression and purification of α X and β 2 tails	125
3.2.3 Structural derivation of the myristoylated α X tail	126

3.2.4 Paramagnetic perturbation studies on the myristoylated α X tails	127
3.2.5 Assignment of backbone amides of α X and β 2	127
3.2.6 Chemical Shift Perturbation Experiments.....	127
3.2.7 Rigid-backbone modeling of the α X β 2 complex:.....	127
3.2.8 Circular Dichroism (CD) studies	127
3.3 Results.....	129
3.3.1 NOESY spectrum of the myristoylated α X integrin in DPC micelles.....	129
3.3.2 Secondary structure analyses of the α X integrin	130
3.3.3 Structural determination of the Myr- α X integrin.....	131
3.3.4 Backbone amide resonance assignments of the α X cytoplasmic tail.....	134
3.3.5 Interaction analyses of the α X β 2 integrin	135
3.3.6 The binding interface between the α X β 2 integrin	137
3.4 Discussion	139
Chapter IV: Structures and interactions of the α 4 cytoplasmic tail and paxillin	143
4.1 Overview	143
4.2 Materials and Methods.....	147
4.2.1 Peptides and Proteins	147
4.2.2 Expression and purification of the α 4 cytoplasmic tail and paxillin.....	147
4.2.3 NMR experiments	148
4.2.4 Dissociation constant determination using Rhodamine- α 4 fluorescence titration.....	150
4.3 Results.....	152
4.3.1 Analysis of the α 4 peptides at different conditions.....	152
4.3.2 Structural characterization of the α 4 cytoplasmic tail in water	153
4.3.3 Solubility of the P234 and α 4 interaction.....	157
4.3.4 Interaction analyses between the P234 and α 4	160
4.3.5 Interaction analyses between the α 4 with P234 and LD repeats	163

4.3.6 Binding analyses of the α 4 interaction with P234 and LD repeats.....	166
4.4 Discussion	170
Chapter V: Summary and concluding remarks	174
Appendix.....	178
Appendix 1: HPLC profiles of synthetic and purified peptides of Chapter II.	178
Appendix 2: HPLC profiles of synthetic and purified peptides of Chapter III.	179
Appendix 3: HPLC profiles of synthetic and purified peptides of Chapter IV.....	180
Appendix 4: Combined changes in chemical shift ($\Delta N + \Delta H$) and change in relative intensity in the NMR titration studies using ^{15}N - ^1H HSQC experiments.	181
References.....	188

List of Figures

Chapter I: Introduction

Figure 1.1	Pairings and classification of the 24 mammalian integrin heterodimers.	4
Figure 1.2	Domain organization of the integrin heterodimer.	8
Figure 1.3	The three generic conformations of the integrin molecule and their associated states of affinity to their extracellular receptors.	9
Figure 1.4	Summary of anti-integrin antibodies.	11
Figure 1.5	Models of integrin activation	13
Figure 1.6	Crystal structures of the $\alpha V\beta 3$ integrin.	14
Figure 1.7	Recently elucidated crystal structures of the $\alpha IIb\beta 3$, $\alpha V\beta 3$ and $\alpha X\beta 2$ integrins.	18
Figure 1.8	Electron micrographs purporting the switchblade model.	21
Figure 1.9	The overall integrin conformation states of the integrins without and with the αI -domain	27
Figure 1.10	Two definitions of the Outer Membrane Clasp (OMC) and Inner Membrane Clasp (IMC) of the $\alpha IIb\beta 3$ integrin	30
Figure 1.11	Three models of transmembrane movements during integrin activation	34
Figure 1.12	Conformations of the $\alpha IIb\beta 3$ transmembrane/cytoplasmic domains and model of activation in lipid environment.	37
Figure 1.13	Conformations of the $\alpha IIb\beta 3$ interaction interface, NMR derived and modelled.	39
Figure 1.14	Conformations of the $\alpha IIb\beta 3$ transmembrane/cytoplasmic domains elucidated in acetonitrile/water.	41
Figure 1.15	Model of integrin intracellular activation.	44
Figure 1.16	NMR-derived structures of the $\alpha IIb\beta 3$ integrins cytoplasmic domains.	49
Figure 1.17	Valency and affinity modulations of integrins.	52
Figure 1.18	Regulation of integrin activation by lateral force.	53
Figure 1.19	Inside-out and Outside-in signal transduction pathways of the $\beta 2$ integrins.	55
Figure 1.20	The cytoplasmic sequences of the human α and β integrin chains.	57
Figure 1.21	Binding sites on integrin α - and β - subunits for integrin cytoplasmic domain binding proteins.	58
Figure 1.22	Conformations of the phosphorylated β -subunit cytoplasmic domains in complex with their ligands.	62

Figure 1.23	Model of talin-induced integrin activation.	66
Figure 1.24	The talin isoform 2 (talin2)/ β 1D structure.	68
Figure 1.25	Representations of the talin FERM domain.	69
Figure 1.26	Structures of the α -subunit cytoplasmic domains.	71
Figure 1.27	Sequence alignment of the leukocyte-specific α -subunit membrane proximal cytoplasmic domains.	73
 Chapter II: Structures and interaction analyses of the αMβ2 integrin		
Figure 2.1	Sequences of the leukocyte and platelet specific integrin cytoplasmic domains.	76
Figure 2.2	^1H - ^1H NOESY spectra of myristoylated and un-myristoylated α M/P α M cytoplasmic tails.	86
Figure 2.3	^1H - ^1H NOESY spectra of the free α M tails at 298 K and 278 K and myristoylated peptides in DPC micelles.	88
Figure 2.4	Paramagnetic relaxation by Mn^{2+} ions of α M/P α M in DPC micelles and in aqueous buffer.	90
Figure 2.5	Paramagnetic relaxation by Mn^{2+} ions of showing the relative changes in intensity of each amino acid residue of Myr- α M/P α M in DPC micelles.	91
Figure 2.6	CD spectra of free α M in aqueous buffer, Myr- α M in DPC micelles and Myr-P α M in DPC micelles.	92
Figure 2.7	Chemical shift deviations from random coil values of $^{13}\text{C}\alpha$ and $^1\text{H}\alpha$ of Myr- α M in DPC micelles, Myr-P α M in DPC micelles and α M in aqueous buffer at 278 K.	94
Figure 2.8	Panels of ^1H - ^1H NOESY of Myr- α M and Myr-P α M displaying the packing NOEs.	96
Figure 2.9	Panels of ^1H - ^1H NOESY of Myr- α M and Myr-P α M displaying the long range NOEs.	98
Figure 2.10	Structures of the Myr- α M and Myr-P α M.	100
Figure 2.11	Triple resonance experiments.	102
Figure 2.12	Strip plots of triple resonance experiments of α M and β 2	103
Figure 2.13	Interaction analyses of the ^{15}N α M with unlabeled β 2.	105
Figure 2.14	Interaction analyses of the ^{15}N β 2 with unlabeled α M and P α M.	106
Figure 2.15	Fluorescence absorption and emission spectrum of tryptophan and dansyl.	107

Figure 2.16	Tryptophan emission spectrum of Myr- β 2 wild type in DPC micelles titrated with Dan- α M.	108
Figure 2.17	Tryptophan emission spectrum of β 2H5W mutant titrated with Dan- α M and Dan-P α M in aqueous buffer solution.	109
Figure 2.18	Resonance energy transfer analysis of Myr- β 2H5W and Dansylated α M peptides.	111
Figure 2.19	Resonance energy transfer of Myr- β 2H5W dissolved in 50% acetonitrile and Dan- α M and Dan-P α M.	112
Figure 2.20	Ribbon representation of the modeled complex of Myr- α M to β 2 and Myr-P α M to β 2.	115
Figure 2.21	Individual structures of α IIB cytoplasmic domains.	117
Figure 2.22	Known NMR characterized complex structures of α IIB β 3 and the α L β 2 cytoplasmic tails.	118
Figure 2.23	Overlaid calcium titration TOCSY spectra of Myr- α M and Myr- α M S14P.	121
Chapter III: Structure and binding interface of the cytosolic tails of the αXβ2 integrin		
Figure 3.1	Sequence alignment of the α -subunit and β -subunit cytoplasmic sequences with the renumbering of the α X and β 2 cytoplasmic domain.	124
Figure 3.2	^1H - ^1H NOESY spectrum of Myr- α X in DPC micelles showing ^1HN connectivities to side chain protons, side chain and ring protons connectivities to other protons.	129
Figure 3.3	Paramagnetic relaxation of the Myr- α X in DPC micelles, showing the relative change in intensity of each amino acid $^1\text{HN}/^1\text{H}\alpha$ NOE after addition of 1mM MnCl_2 .	130
Figure 3.4	CD spectrum of Myr- α X in DPC micelles showing the changes in Molar ellipticity of the far UV CD range from 240 nm to 190 nm.	130
Figure 3.5	Chemical shift deviations from random coil values of the Myr- α X $^{13}\text{C}\alpha$ and $^1\text{H}\alpha$ chemical shift values.	131
Figure 3.6	Panels of the ^1H - ^1H NOESY of the Myr- α X showing packing and long range NOEs.	132
Figure 3.7	NMR derived structure of the Myr- α X in DPC micelles.	134
Figure 3.8	Strip plots of triple resonance experiments of α X.	135
Figure 3.9	Interaction analyses of the α X and β 2 cytoplasmic tails.	137
Figure 3.10	Modeled complexes of the α X and β 2 cytoplasmic tails.	138

Figure 3.11	Single structures of the Myr- α X in DPC micelles, Myr- α M in DPC micelles, α L cytoplasmic tail in solution and Myr- α IIb in DPC micelles.	140
Figure 3.12	Complex structures of the α IIb β 3 cytoplasmic tails, α L β 2 cytoplasmic tails and α M β 2 cytoplasmic tails.	142
Chapter IV: Structures and interactions of the α4 cytoplasmic tail and paxillin		
Figure 4.1	Domain organization of the paxillin molecule and its binding partners.	145
Figure 4.2	Schematic diagram of the regulation of α 4 association with paxillin via phosphorylation in a migrating cell.	146
Figure 4.3	Sequence alignment of the α - subunit and β subunit cytoplasmic domains.	152
Figure 4.4	1D-NMR spectra of the Myr- α 4 and α 4 synthetic peptides in various conditions.	153
Figure 4.5	^1H - ^1H NOESY spectrum of the α 4 cytoplasmic tails in water at 278 K.	154
Figure 4.6	Chemical shift deviations of the $^1\text{H}\alpha$ (Top panel) and $^{13}\text{C}\alpha$ (Bottom panel) from random coil.	154
Figure 4.7	Structures of the free α 4 cytoplasmic tail in water, 278K.	156
Figure 4.8	Electrostatic potential maps of α 4 cytoplasmic tail showing the front and back of the molecule.	157
Figure 4.9	Titration of ^{15}N -P234 with unlabeled α 4 cytoplasmic tail in different buffers.	158
Figure 4.10	Overlay of HNCACB and CBCA(CO)NHNH strip plots of N^4 , L^5 , S^6 , E^7 and L^8 of double-labeled $^{15}\text{N}^{13}\text{C}$ -P234	159
Figure 4.11	$^{13}\text{C}\alpha$ chemical shift deviations of P234 from random coil values.	160
Figure 4.12	^{15}N - ^1H HSQC spectra of ^{15}N -P234 titration with α 4.	161
Figure 4.13	Combined chemical shift changes ($\Delta\text{N}+\Delta\text{H}$) of each amino acid residues of P234 following titration with unlabeled α 4 at a ratio of 1:2.	162
Figure 4.14	Relative changes in intensity of each amino acid residues of P234 following titration with unlabeled α 4 at a ratio of 1:2.	163
Figure 4.15	^{15}N - α 4 cytoplasmic tail titration with P234, LD2, LD3 and LD4.	165
Figure 4.16	HPLC profile of dissolved NMR sample precipitate.	167
Figure 4.17	Fluorescence emission spectra of Rho- α 4 titration with P234, LD2, LD3 and LD4.	168
Figure 4.18	Non-linear regression analysis of the averaged normalized change in fluorescence (F_{max}) of Rho- α 4 quenching as a function of P234, LD2, LD3 and LD4 concentration.	169
Figure 4.19	Structures of known cytoskeletal remodeling proteins with LD2 (in red) and LD4 (in blue).	172

List of Tables

Chapter I: Introduction

Table 1.1	List of some integrins and their unique properties	3
Table 1.2	Structural studies of integrin cytoplasmic region	45

Chapter II: Structures and interaction analyses of the $\alpha M\beta 2$ integrin

Table 2.1	Summary of sequences for synthetic and purified α - and β - tails	78
Table 2.2	Table of long range NOEs used in structure determination	97
Table 2.3	Summary of structural statistics for the 20 final structures of αM -tail	99

Chapter III: Structure and binding interface of the cytosolic tails of the $\alpha X\beta 2$ integrin

Table 3.1	Summary of sequences for synthetic and purified α - and β - tails	126
Table 3.2	Table of long range NOEs used in structure determination	132
Table 3.3	Summary of structural statistics for the 20 final structures of Myr- αX	133

Chapter IV: Structures and interactions of the $\alpha 4$ cytoplasmic tail and paxillin

Table 4.1	List of the sequences peptides and proteins used.	147
Table 4.2	Summary of structural statistics for the 20 final structures of $\alpha 4$ in water, 278K.	154
Table 4.3	The LD repeats and residues (in red) perturbed in the ^{15}N -P234 and $\alpha 4$ titration.	165
Table 4.4	Summary of dissociation constants obtained in interaction with Rho- $\alpha 4$.	168

Publications

Chua, G.L., Tang, X., Amalraj, M., Tan, S.M and Bhattacharyya, S. **Structure and interaction analyses of the integrin α M β 2 cytoplasmic tails.** Journal of Biological Chemistry (2011), 286, 43842-43854.

Chua, G.L., Tang, X., Patra, A., Tan, S.M and Bhattacharyya, S. **Structure and binding interface of the cytosolic tails of α X β 2 integrin.** PLOS-ONE 2012;7(7):e41924. doi: 10.1371/journal.pone.0041924. Epub 2012 Jul 26.

Chua, G.L., Patra, A., Tan, S.M and Bhattacharyya, S. **NMR Structure of Integrin α 4 Cytosolic Tail and Its Interactions with Paxilin.** PLOS-ONE 2013;8(1):e55184. doi: 10.1371/journal.pone.0055184. Epub 2013 Jan 31.

Abbreviations

CD	Circular Dichroism
DPC	Dodecylphosphocoline
HSQC	Heteronuclear Single Quantum Correlation
NMR	Nuclear Magnetic Resonance Spectroscopy
NOE	Nuclear Overhauser Effect
RMSD	Root mean square deviation
NOESY	Nuclear Overhauser Spectroscopy
TOCSY	TOTAL Correlation Spectroscopy
LD	Leucine-Aspartate
LAD	Leukocyte Adhesion Deficiency
IG	Immunoglobulin
ICAM	Intercellular Adhesion Molecule
JAM	Junctional Adhesion Molecule
MIDAS	Metal Ion Dependent Adhesion Site
ADMIDAS	Adjacent Metal Ion Dependent Adhesion Site
LIMBS	Ligand-associated Metal Binding Site
SyMBS	Synergistic Metal Ion Binding Site
SDL	Specificity Determining Loop
EGF	Epidermal Growth Factor
PSI	Plexin-Semaphorin-Integrin
LIBS	Ligand-Induced Binding Sites
EM	Electron Microscopy
FILM	Fluorescent Life-time Imaging
FRET	Fluorescent Resonance Energy Transfer
OMC	Outer Membrane Clasp
IMC	Inner Membrane Clasp
Myr-	Myristoylated
KSI	Keto-Steroid Isomerase
LB	Luria-Bertani
M9	Minimal Media
IPTG	Isopropyl β -D-1-thiogalactopyranoside
DSS	2,2-Dimethyl-2-Silapentane-5-Sulfonate
RET	Resonance Energy Transfer
Dan-	Dansyl
1D	1-Dimension
2D	2-Dimension
3D	3-Dimensional

Abstract

Integrins are devoid of enzymatic activity but yet are involved in almost all aspects of mammalian physiology. Their unique bidirectional signal transduction property can enable them to continuously sense changes in their extracellular environment and respond rapidly through conformational changes that modulates their affinity for extracellular ligands. This continuous sense and response system mediated by integrin molecules is a quintessential property of the mammalian immune system, which requires leukocytes circulating in the high pressure enclosed blood vascular system to respond to minuscule changes in the release of cytokines during an immunological event. This process requires leukocytes to rapidly change from non-adhesive, symmetrical cells to highly adhesive asymmetric cells to attach to the vascular walls amid the high pressure blood flow, followed by progressive migration via adept modulation of specific integrins towards the site of injury. The rapidly changing affinity of leukocytes is modulated by the β_2 family of integrins, which consists of the $\alpha M\beta_2$, $\alpha L\beta_2$, $\alpha X\beta_2$ and $\alpha D\beta_2$ integrins; and the $\alpha 4$ integrins. Conformational changes, namely the bent with closed headpiece and clasped cytoplasmic domains; the extended with closed headpiece and clasped cytoplasmic domains; and the extended with open headpiece and separated cytoplasmic domains dictates the integrin resting and activated status respectively. Cytokine-induced activation of integrins result in the activation of the integrin inside-out signaling that changes the affinity of the integrin molecule for its extracellular ligands via these three conformational changes. The binding of the extracellular ligands results in the subsequent activation of the outside-in signaling transduction leading to the remodelling of the actin cytoskeleton which changes the ultimate morphology and adhesiveness of the leukocytes. Both these integrin inside-out and outside-in signal transduction converges on the short integrin cytoplasmic domains. It has been largely generalized for most integrins that the membrane proximal regions of the cytoplasmic domains consisting of the α -(GFFKR) and β -(KLLxxxHDRRE) motifs are involved in maintaining the cytoplasmic clasp in the resting state integrin. Mutational studies of these residues on the $\alpha I Ib\beta 3$ and $\alpha L\beta 2$ integrins have demonstrated the importance of the salt bridge between the α -subunit Arg residue and the β -subunit Asp residue; and the

interaction between the α -subunit Phe residues and β -subunit Leu and His residues. The disruption of these interactions resulting in α -/ β - cytoplasmic domain separation has been demonstrated to transmit to the other integrin domains, namely the transmembrane and extracellular domains, to cause the resting state bent integrin molecule to form the high affinity extended conformation. Conversely, the membrane distal region of the integrin cytoplasmic domains is involved in the specific recruitment of intracellular ligands that modulates this cytoplasmic clasp.

Thus far, structural efforts on the integrin cytoplasmic domains have been largely focused on the α IIB β 3 integrins which are involved in platelet aggregation that is an all-or-nothing response triggered by a cut wound. This is in contrast to the leukocyte shape-shifting ability, modulated by the β 2 and α 4 integrins, in its migration towards the site of inflammation. The α L β 2, α M β 2, α X β 2, α D β 2 and α 4 β 1 integrins are activated in an orchestrated manner in order to facilitate the distinct processes in leukocyte migration. This involves the differential recruitment of intracellular effectors which recognizes the distinct sequences in the α -subunit cytoplasmic domains found in the β 2 integrins. Moreover, it has been demonstrated that given the same stimulus, the β 2 integrins are not activated uniformly, with the α X β 2 integrin most resilient to inside-out activation via cytoplasmic tail separation. Hence, it is imminent to understand how the four different α -subunit interacts with its β 2 counterpart and how different α -cytoplasmic tails are able to recruit distinct intracellular effectors for their purpose. The understanding of how different α -subunit cytoplasmic domains interact with the β 2-subunit modulates its interaction with intracellular ligands has been generally limited by the paucity of cytoplasmic structures. The recent structural elucidation of the α L cytoplasmic tail by Bhunia et. al. (2009) characterized by a trihelical fold stabilized by electrostatic interactions, is completely distinct from all the α IIB cytoplasmic structures thus far elucidated. However, the α L β 2 cytoplasmic clasp still maintains the conserved membrane proximal interactions as the α IIB β 3 cytoplasmic clasp. Whether this can be generalized across the rest of the β 2 integrins still require further structural and interaction analyses.

The $\alpha M\beta 2$ integrin is involved in the formation of focal adhesions during leukocyte migration. It has been found to be constitutively phosphorylated at its cytoplasmic domain residue S¹¹²⁶, the phosphorylation of this residue has been demonstrated to cause differential extracellular ligand binding to $\alpha M\beta 2$. Hence, understanding the structural differences between the unphosphorylated and phosphorylated αM cytoplasmic form can provide insights into the conformational regulation brought about by the α -cytoplasmic domain. Structural analyses of the unphosphorylated and phosphorylated αM , which will be elaborated in this thesis, showed that both adopts a helical conformation at the conserved membrane proximal region followed by a loop formed by the membrane distal region. The addition of the phosphorylated group has only caused the reorganization of the membrane distal loop. Moreover, these two forms of αM are found to interact similarly with the $\beta 2$ cytoplasmic tail with only subtle differences in the electrostatic interactions at their membrane proximal regions and dissociation constants, 5.2 and 4.4 μM for interaction between $\alpha M\beta 2$ and phosphorylated $\alpha M\beta 2$. Additionally, metal ions which are found to play important roles in the modulation of the cytoplasmic interactions in $\alpha IIb\beta 3$ and $\alpha L\beta 2$ integrins, are not found to interact with any residues in unphosphorylated and phosphorylated αM cytoplasmic tails.

Similarly, in the later structural characterization of the αX cytoplasmic domain and its interaction with its $\beta 2$ counterpart, the αX cytoplasmic tail adopts a somewhat similar conformation as the αM cytoplasmic tail. A helical conformation is adopted at its membrane proximal region followed by a well-defined turn and a disordered loop at its membrane distal region. Similar $\beta 2$ -cytoplasmic tail residues are perturbed as the αM cytoplasmic tail which suggests that the conserved interactions mediated by the membrane proximal residues are important in maintaining the resting state cytoplasmic clasp. The different loop conformations adopted by the membrane distal residues of the unphosphorylated and phosphorylated αM ; and αX cytoplasmic tails could thus possibly enlist distinct intracellular effectors through the distinct structural motifs formed by the unique and unconserved membrane distal sequences of the α -subunits belonging to the $\beta 2$ family of integrins.

Although the $\alpha 4\beta 1$ integrin does not form part of the $\beta 2$ family of integrins, it plays an important role in the polarization of leukocytes to facilitate its lateral migration on the endothelial walls. Similar to the α -subunits of the $\beta 2$ integrins, it consists of the conserved α -GFFKR motif at its membrane proximal region, but unlike the α -subunits of the $\beta 2$ integrins it does not contain any proline residues and has a few hydrophobic and many charged residues at its membrane distal regions. Its direct association with the three Leucine-Aspartate (LD) motifs of paxillin from residue S⁹⁷⁸ to I⁹⁹² has been demonstrated to directly affect the migratory properties of immune cells. Hence, the structural characterization of the $\alpha 4$ -cytoplasmic tail and interaction with paxillin can provide insights into distinct α -subunit cytoplasmic structural motifs that regulate integrin function. The structure of the $\alpha 4$ cytoplasmic tail in water displayed a propensity to form an un-amphipathic helical structure with several acidic and basic residues forming positive and negative patches that could possibly be crucial for its recruitment of intracellular ligands. The paxillin region which consists of LD2, LD3 and LD4 was found to interact with the free $\alpha 4$ cytoplasmic tail. In addition, perturbation studies using these individual LD motifs were also observed to interact with the $\alpha 4$ cytoplasmic tail. In these studies, $\alpha 4$ cytoplasmic tail Q⁸⁽⁹⁷⁵⁾, S¹¹⁽⁹⁷⁸⁾ and R¹⁸⁽⁹⁸⁵⁾ was consistently observed to be perturbed. Furthermore, S¹¹⁽⁹⁷⁸⁾ and R¹⁸⁽⁹⁸⁵⁾ is found to be within the $\alpha 4$ S⁹⁷⁸ to I⁹⁹² region, which has been implicated in the $\alpha 4$ and paxillin interaction. Binding studies using rhodamine fluorescence also showed that the P234 and LD motifs interact in the micro-molar range with the $\alpha 4$ cytoplasmic tail. Hence, paxillin possibly wraps around the $\alpha 4$ cytoplasmic tail via its three LD motifs.

Chapter I: Introduction

1.1 Afflictions related to the abnormal cell surface protein, Integrin

About 30 years ago, a previously uncharacterized form of immunodeficiency was identified in five infants who had umbilical cords still attached at three weeks of birth [1]. They subsequently developed irreconcilable bacterial infections from which only one baby survived [1]. An inherited abnormality in a contractile protein was the 'suspect'. In the following year, similar symptoms of immunodeficiency were diagnosed in a five year-old boy who suffered from recurrent bacterial infections; in addition his wounds were slow to heal and non-purulent [2]. Characterization of his neutrophils unveiled a severe defect in chemotaxis and a mild impairment in phagocytosis [2]. Gel electrophoresis revealed that a cell surface molecule, gp100 (now widely known as the $\beta 2$ integrin) was missing in his neutrophils [2]. Neutrophils and other white blood cells, collectively known as leukocytes are crucial players in an immune response, they respond to local changes in the vascular system [3]. They slow down in response to chemokines released in an injured cell, and start rolling towards the site of injury, upon reaching the exact site they firmly arrest onto the endothelial walls and begin trans-endothelial migration where they encounter and phagocytose foreign antigens and bacteria. This process is crucial in preventing bacterial infections during a cut wound from becoming debilitating septicemia [3]. Cell adhesion molecules, or integrins, found on the cellular membrane of leukocytes govern leukocyte mobility and abnormalities in these molecules are known as Leukocyte Adhesion Deficiency (LAD) [4]. There are currently three types of LAD, I, II and III, the root causes of immunodeficiencies in these three types of afflictions are; abnormalities in the $\alpha L\beta 2$ integrins resulting in leukocytes that do not arrest firmly to the endothelial walls (LAD-1); defects in the fucose-specific transporter specific to the golgi apparatus leading to the production of defective integrins lacking the sialyl lewis X which binds to the selectins on the vascular endothelium, this interaction is crucial for leukocytes during the rolling phase (LAD-II); and mutations in the cytoplasmic activators of integrins resulting in LAD-1

symptoms and also together with a prolonged bleeding tendency caused by the lack of activation of $\alpha\text{IIb}\beta 3$ integrins found on platelets (LAD-III) [4].

Besides abnormality to the $\alpha\text{L}\beta 2$ and $\alpha\text{IIb}\beta 3$ integrins, mis-regulation via over-expression of $\alpha\text{Xb}2$ and $\alpha\text{M}\beta 2$ integrins on neutrophils causes redundant clustering and the eventual failure of interaction with its repertoire of ligands [5]. This also results in deficient shape changes in neutrophils during a chemotactic response and failure to traffic to sites of injury, causing sterile skin ulcers that enlarges overtime. This is a rare and painful condition known as Pyoderma Gangrenosum [5].

Additionally, mis-regulation and modifications to the $\alpha\text{M}\beta 2$ and $\alpha\text{L}\beta 2$ integrin on monocytes was also found to occur in diabetic patients. The higher oxidative stress caused by the increased levels of glucose was found to induce alterations to basement membrane proteins. Together with the abnormal increase in expression of $\alpha\text{M}\beta 2$ and $\alpha\text{L}\beta 2$ integrins found on diabetic monocytes, result in excessive interactions of the monocytes with the modified basement membrane leading to the formation of atheromatic lesions commonly found in diabetic patients [6].

Perhaps the most extensive damage that can be caused by mis-regulation or abnormalities of these cell surface molecules that enable mobility in leukocytes, is their roles in tumour initiation, progression and metastasis [7]. Integrins not only mediate interactions with the extracellular matrix proteins but also to circulating growth factors and cytokines. In addition, they are not only limitedly expressed on leukocytes but cells of the entire human body, which includes fibroblasts, endothelial, epithelial cells, and etc [7]. The collective dysfunctions of a handful of integrins in these initially normal cells contribute to tumour initiation and subsequently progression of a benign tumour by mediating angiogenesis, lymphangiogenesis, and migration [7]. $\alpha 4\beta 1$, $\alpha\text{V}\beta 3$, $\alpha\text{V}\beta 5$, $\alpha 6\beta 4$, and $\alpha\text{V}\beta 6$ integrins are normally expressed in low or undetectable levels in most adult epithelial but highly up-regulated in metastatic solid tumours. Tumour-associated blood vessels was also found to overly express $\alpha\text{V}\beta 3$ and $\alpha\text{IIb}\beta 3$ [7].

From the above non-exhaustive examples of medical implications described, integrins play important and diverse roles in the proper functioning of the human body, such as leukocyte migration [8], maintenance of the various tissue epithelium [9], control of cell division and differentiation [9]. Some examples of integrins and their functions are listed in Table 1.1. Many of these functions require continuous sensing of the cell's external environment and activation of internal cytoplasmic proteins that can affect cell shape, polarity, survival or apoptosis, and integrins integrate this bidirectional signaling despite their lack of enzymatic activity [10].

Table 1.1. *List of some integrins and their unique properties*

Integrin	Cellular localization		Ligands (extracellular)	General function
$\alpha_M\beta_2^*$	Leukocytes	Monocytes; Macrophages; Neutrophils; Dendritic cells; Lymphocytes; Natural killer cells; Myeloid leukocytes	Intercellular adhesion molecules (ICAM)-(1, 2, 4); Fibrinogen; Factor X; iC3b	Immune cell migration and recruitment in host immunity
$\alpha_L\beta_2^*$		Neutrophils; Dendritic cells; Lymphocytes; Natural killer cells	Intercellular adhesion molecules (ICAM)-(1-5); Junctional Adhesion Molecule (JAM) -1	
$\alpha_D\beta_2$		Macrophages; Eosinophils		
$\alpha_X\beta_2^*$		Monocytes; Dendritic cells	Osteopontin, iC3b	
$\alpha_4\beta_1^*$		Lymphocytes; Eosinophils; Monocytes; Natural Killer cells; Neutrophils; Non-Immune cells	VCAM; Fibronectin;	Also plays a role in development
$\alpha_{IIb}\beta_3$	Platelets		Fibrinogen	Blood coagulation
$\alpha_v\beta_3$	Bone resorbing cells; Osteoclasts;		Fibrinogen; fibronectin	Bone formation; angiogenesis

1.2 Diversity of the integrin super family

Cell-cell and cell-extracellular matrix (ECM) communications are central to integrin functions [10]. Such communications are important in multi-cellular organisms, where there are distinct tissue organizations in the different

organs, and precise regulation of cells must be strictly adhered to for normal development or functioning of the organism as a whole [11]. Hence, integrins are only found in metazoans and in more complex chordates the more integrins it possesses [11]. In addition, integrins in multi-cellular organisms are so conserved that even the sequences and genes of the *Drosophila* integrins are about as closely related as those most divergent vertebrate subunits [12].

Integrins can be found in practically all cells in the human body, and each cell type has its own specific repertoire of integrins from the 24 known mammalian integrin heterodimers [10]. Each integrin heterodimer consists of an α - and β -subunit non-covalently associated together at its extracellular domain, and they are obligate heterodimers that assemble on the endoplasmic reticulum before transportation to the cellular membrane [13]. There are currently known 18 α - and 8 β - subunits that pair specifically to form the 24 integrin heterodimers, with some subunits pairing with more than one opposing subunit [10]. $\beta 1$, $\beta 2$ and αV are examples of promiscuous subunits, capable of forming integrin heterodimers with several other α - and β - subunits [10]. The detailed pairings between α - and β - subunits can be found in Figure 1.1.

Figure 1.1 Pairings and classification of the 24 mammalian integrin heterodimers. (A) Integrin α - and β - heterodimeric pairings that are classified into the specific extracellular receptors that they recognize. Taken from, Hynes, R.O., *Integrins: bidirectional, allosteric signaling machines*. Cell, 2002. 110(6): p. 673-87. (B) The integrin α -subunit divided into two major groups with and without the Inserted (von Willebrand Factor Type A) domain. Taken from, Johnson, M.S., et al., *Integrins during evolution: evolutionary trees and model organisms*. Biochimica et biophysica acta, 2009. 1788(4): p. 779-89.

1.3 Receptors and ligands of integrins

Despite having multiple α - partners, the heterodimers formed by the $\beta 1$ subunit can recognize distinct sets of extracellular receptors and activate specific sets of signal transduction mechanisms; for example, $\alpha 3\beta 1$, $\alpha 6\beta 1$, $\alpha 7\beta 1$ and $\alpha 6\beta 4$ interacts selectively with laminin receptors [14]. Interestingly, different integrins can also share the same set of receptors to synergize signal transduction by orchestrating related sets of cytoplasmic ligands; such as the $\alpha L\beta 2$ and the $\alpha M\beta 2$ integrin can both bind to ICAM-1(Intercellular Adhesion Molecule-1) to facilitate leukocyte migration [9,14]. In addition, integrins can also bind to multiple receptors; for example, integrins consisting of αV subunits can recognize multifarious receptors such as, fibronectins, laminins, collagen, and etc [9]. Despite the largely promiscuous nature of integrins and their respective extracellular receptors, they could still be distinguished based on the receptors they mainly interact (Figure 1.1). Integrin-receptor interactions can be further specifically delineated based on the specific motifs that integrins recognize in their extracellular receptors, such as the Arg-Gly-Asp (RGD) or Leu-Asp-Val (LDV) motifs [14]. Alternatively, integrins can be classified based on their β -subunit partners, or most commonly they can be separated based on their α -subunit partners presence or absence of an Inserted I-domain (also known as the von Willebrand Factor type A domain, or αA domain) in the α -subunit. They could also be classified based on the functions they mediate.

Aside from interacting with extracellular receptors, integrins also act as a hub for intracellular ligands thereby integrating external and internal signals in a cell. Intra-cellular ligands that interact with integrins include cytoskeletal remodeling proteins, cell signaling proteins, adaptor proteins, enzymes, etc [15].

1.4 Studies on the function of integrins

Endeavours to determine the exact function of each integrin subunit have been made through the generation of knockout and chimeric animals. Knockout of the most promiscuous $\beta 1$ subunit ablates 12 integrins and the knockout mice never survive past the embryonic stage [9]; $\beta 2$ subunit

knockouts generated viable mice with LAD symptoms [16] and similarly $\beta 3$ knockouts also generated viable mice but with prolonged bleeding and placental defects [9]. Chimeric mice have been important in understanding the overlapping functions of integrins in development and hemostasis by enabling bypass of embryonic lethality in some of the integrins involved in development. For example the $\alpha 4\beta 1$ integrin, which knockout would have caused embryonic lethality and obscure its function in leukocyte migration. Additionally, integrin function has also been studied *in vivo* and *in vitro* where its sequence in relation to function has also been investigated using mutational strategies. Single-site mutagenesis at the integrin cytoplasmic region of conserved residues has been shown to modulate integrin activation [17]. Differential pairings of integrin heterodimers have also been examined to determine the effects of exchanging β -partners in the modulation of the integrin function [18,19].

1.5 General structure of integrins

Integrin allostery plays an important role in its ability to coordinate extracellular and intracellular signals. The lack of enzymatic activity of integrins is compensated for by their ability to undergo conformational changes to expose key sites for receptor and ligand binding. Hence, immaculate understanding of the integrin structure has become key to decipher integrin functions in the different processes in the human body.

In general, integrins have an extracellular, transmembrane and a short cytoplasmic domain characteristic of type I transmembrane proteins. The integrin α - and β - subunits are non-covalently bonded at the apex of its extracellular domain forming the globular headpiece, leaving the rest of the regions unconstrained, like a pair of legs, free to move apart or come together depending on the interactions of each individual domains within each α - and β -subunit.

1.6 The domain organization of the α -subunit

At the N-terminus of the α -subunit is the β -propeller which contains 7 repeats of 60 amino acids that fold into 7 radially arranged blades, and it has the

same topology as the trimeric G-protein subunit [20,21]. The α -subunits of α M, α X, α L, α D, α E, α 1, α 2, α 10 and α 11 contain an additional domain, called the I-domain, that is inserted between β -sheets 2 and 3 of the β -propeller domain and it adopts a Rossman fold with a metal ion dependent adhesion site (MIDAS) on the top of the domain that is found to bind to divalent cations such as Mg^{2+} , Ca^{2+} or Mn^{2+} [21]. The I-domain has β -strands and α -helices alternating each other, with the α -helices wrapping around the domain in counterclockwise order when viewed from the top face [22]. When present, the I-domain together with the MIDAS forms the major receptor binding site [23]. At the domain C-terminal to the β -propeller domain is the thigh domain which adopts the C2-set Immunoglobulin (Ig) fold, followed by 2 calf domains each having 2 anti parallel β -sheets, leading into the single helix transmembrane domain and finally to the short α -subunit cytoplasmic domain [20]. The domain organization of the α -subunit can be seen in Figure 1.2A.

1.7 The domain organization of the β -subunit

The N-terminus of the β -subunit makes extensive contacts with the α -subunit β -propeller domain through its β I-like domain which is inserted into the hybrid domain [20,21]. The β I-like domain consists of a central 6-stranded β -sheet surrounded by 8 helices similar to the α -subunit I domain, at its top central β strand is the MIDAS motif, an adjacent MIDAS (ADMIDAS) motif, and ligand-associated metal-binding site (LIMBS) also known as the synergistic metal ion-binding site (SyMBS) [20,24]. The β I-like domain, specifically three loops known as the specificity determining loops (SDL), together with the metal ion binding sites and some regions on the α -subunit β -propeller form the major receptor binding site in integrins that lack the α -subunit I-domain [23]. The hybrid domain that interfaces with the β I-like domain has a similar fold to the I-set Ig domain and is connected to the first of the 4 tandem cysteine-rich repeats each assuming a class I epidermal growth factor (IEGF) fold via the plexin-semaphorin-integrin (PSI) domain which has an α/β fold [20]. Following the last IEGF domain is the flexible β -tail domain leading into the single helix transmembrane domain and short cytoplasmic domain. The domain organization of the β -subunit can be seen in Figure 1.2B.

Figure 1.2 Domain organization of the integrin heterodimer. **(A)** Domain organization of the I-domain containing integrin. **(B)** Domain organization of integrins with no I-domain. (Taken and adapted from Luo, B.H. et. al. (2007) *Structural basis of integrin regulation and signaling*. *Annu Rev Immunol.* 25: p. 619-647.)

1.8 Structural transitions of integrins

Interactions between the domains in the α - and β -subunit dictate integrin conformations and hence its affinity to its extracellular receptors. Three known integrin conformations, the 'bent integrin with closed headpiece and clasped cytoplasmic tails' (Figure 1.3A); the 'extended integrin with closed headpiece and clasped cytoplasmic tails' (Figure 1.3B); and the 'extended with open headpiece and separated cytoplasmic tails' (Figure 1.3C) are possibly conserved across the 24 known integrins. Global integrin conformational changes have been found to be correlated to its' low and high affinity binding to their respective extracellular receptors (Figure 1.3). These three aforementioned conformations are not entirely definitive as conformations with varying degrees of extension have been observed to exist, but it still remains unclear which integrin conformation binds to the ligand [25]. Electron microscopic studies on α V β 3, α Ib β 3, α L β 2 and α X β 2 suggest that these three conformations exist in dynamic equilibrium to each other, and different integrins vary markedly in their susceptibility to activation due its intrinsic differences amongst each integrin, for example, difference in α -subunit results in the greater propensity for the α L β 2 to extend than the α X β 2 integrin [26,27]. The presence of external factors can also favour the formation of certain

conformations. Hence, receptor binding to the integrin molecule has been suggested to depend on the rate of integrin conformational equilibration, difference in conformation population and the intrinsic binding affinities of these conformational states [28].

Figure 1.3 The three generic conformations of the integrin molecule and their associated states of affinity to their extracellular receptors. Taken from, Zhu, J., et al., *Closed headpiece of integrin $\alpha IIb\beta 3$ and its complex with an $\alpha IIb\beta 3$ -specific antagonist that does not induce opening. Blood, 2010. 116(23): p. 5050-9.*

1.9 Factors affecting integrin structural transitions

The structural transitions from the bent low affinity state to the extended high affinity state are known as, integrin activation or priming. Both extracellular and intracellular factors are found to be involved. Extracellular factors such as metal ions, presence of high concentrations of the integrin receptor, use of antibodies that bind to specific regions of the integrin extracellular domain can prime integrin activation to the extended conformation. Indirect priming from intracellular factors which perturb the integrin intracellular regions involves cytokine activation pathways [17]. Integrin priming extracellularly and intracellularly are termed outside-in and inside-out signaling, respectively and both involve the recruitment of distinctive sets of proteins and global structural changes.

The presence of metal ions such as Mn^{2+} , Mg^{2+} , or Ca^{2+} which bind to the metal-ions binding sites, namely the MIDAS, ADMIDAS and LIMBS of the integrin receptor binding site can prime or inhibit integrins to its extended

conformations. Both Mn^{2+} and Mg^{2+} are found to prime integrins to its high affinity extended state, whereas Ca^{2+} at high concentrations is found to stabilize the bent low affinity conformation [22]. Integrins can also be primed by the presence of high concentrations of its extracellular ligands or small fragments containing receptor specific binding motifs, which could also possibly trigger integrin clustering and activation. Antibodies generated to bind to specific regions of the integrin extracellular domain, known as the ligand-induced binding sites (LIBS), especially at regions maintaining the bent conformation, can destabilize contacts in the bent conformation and promote the formation of the high affinity extended form (Figure 1.4) [29]. Similarly there are also inhibitory antibodies that favor the bent conformation (Figure 1.4). Some of these antibodies have also been useful in detecting specific sites in the integrin molecule when it is in the bent or extended conformation. Besides activation via extracellular factors, integrins can also be primed to its high affinity extended form through cytokine activation where the binding of cytokine activated proteins to the cytoplasmic region of the integrin molecule causes disruption of the α - and β -subunit cytoplasmic region interaction which can propagate to the integrin transmembrane and extracellular regions leading to extension.

Figure 1.4 Summary of anti-integrin antibodies, displaying Non-functional antibodies (**Topmost panel**) that recognizes the individual subunits, Inhibitory antibodies (**Leftmost panel**) favouring the bent integrin conformation and Activatory antibodies (**Rightmost panel**) promoting the activated extended conformation. *From Byron, A., et al., Anti-integrin monoclonal antibodies. Journal of cell science, 2009. 122(Pt 22): p. 4009-11.*

1.10 Structural studies of the open, closed and intermediate conformations of integrin

X-ray crystallography, electron microscopy (EM) and Nuclear Magnetic resonance (NMR) spectroscopy have provided detailed insights into the structure and conformational flexibility of integrins. In combination with the molecular biology studies of integrins in whole cell systems, the correlation between the integrin molecule conformation and affinity state has been determined. In addition, computational modeling studies have also enabled structural gaps in the integrin crystal structures, EM and NMR studies to be combined for the generation of hypothetical models of integrin activation and deactivation. Currently two models of integrin activation, namely the switchblade and deadbolt model have been proposed (Figure 1.5). Briefly, the switchblade model proposes that upon activation, either through the integrin headpiece or at the cytoplasmic regions, the headpiece of the integrin molecule initially near to the plasma membrane flicks away in a 'jack-knife'-like manner. However, the headpiece can remain in a closed conformation or transit to the open conformation with high affinity when the integrin headpiece comes into contact with its receptor. Upon receptor binding, the hybrid domain then swings outwards with respect to the β -I like domain. This model proposes that the hybrid domain is in close contact with the PSI domain [22]. In contrast, the alternate deadbolt model suggests that the hairpin loop (Or CD loop) in the β -tail domain is in close contact with the β I-like and hybrid domains in the bent and closed headpiece state (Figure 1.5). This contact is proposed to function as a deadbolt which prevents the β I-like domain from displacing the hybrid domain. Structural changes in the transmembrane and cytoplasmic regions, due to binding of intracellular ligands, are transmitted to the β -tail region which will destabilize the deadbolt and allow the displacement of the hybrid domain, exposing the ligand binding site. Current studies that support the deadbolt model also suggested the possibility of integrins in the bent conformation binding to its receptor [30]. The validation of these models can only be achieved by affirming whether integrin conformational changes were due to ligand binding or is

a prerequisite to ligand binding. Additionally, whether can these proposed models be applied to all 24 integrin heterodimers requires further investigation.

Figure 1.5 Models of integrin activation, the topmost panel showing the switchblade transitions, while the bottommost panel showing the deadbolt transitions. Taken and adapted from Hynes, R.O., *Integrins: bidirectional, allosteric signaling machines*. *Cell*, 2002. 110(6): p. 673-87; and Arnaout, M.A., et al., *Structure and mechanics of integrin-based cell adhesion*. *Current opinion in cell biology*, 2007. 19(5): p. 495-507.

1.11 Detailed structural insights into the extracellular domains of integrin

The first detailed insight into the structure of the integrin molecule comes from the α V β 3 crystal structure from Xiong et. al (2001), which showed the integrin headpiece fold $\sim 135^\circ$ back to the lower leg regions forming a V-shaped structure pivoting at a 'genu' formed between the thigh, calf and disordered I-EGF domains of the α V β 3 [20]. It is also observed that the disordered CD loop of the β -tail domain is in close contact with the hybrid domain which the authors explained

to be the key forces maintaining the bent conformation (deadbolt model). And in the following year, the crystal structure of the $\alpha V\beta 3$ integrin was elucidated by the same group in complex with its receptor the cyclic pentapeptide Arg-Gly-Asp (RGD) (Figure 1.6A) [31]. However, the presence of the RGD peptide and even the use of Mn^{2+} in the RGD $\alpha V\beta 3$ crystal, which is previously found to be a strong activator of integrins [32], did not induce great changes in the orientation of αV relative to $\beta 3$ as corresponding to the change to high affinity state in the presence of its receptor, except slight rearrangements in the βI -like and the αV propeller domain (Figure 1.6B) [31]. In addition, not all domains are observed clearly in these 2 crystal structures, namely the I-EGF and PSI domains were disordered and the two metal ions, Ca^{2+} and Mg^{2+} , which were supposed to be found in the MIDAS and ADMIDAS of the βI -like domain were not observed.

Figure 1.6 Crystal structures of the $\alpha V\beta 3$ integrin. **(A)** Crystal structure of the integrin $\alpha V\beta 3$ extracellular domain displaying the ribbon diagram of the unliganded $\alpha V\beta 3$. The $\alpha V\beta 3$ integrin is bent at the genu (arrow), I-EGF 1 and 2 are not visible in the structure, only their approximate locations are indicated at E1 and E2. **(B)** Ribbon diagram of the crystal structure in complex with its high affinity ligand RGD. The RGD peptide is displayed as a ball-and-stick model with carbons in green, amides in blue and oxygens in red. The specificity determining loop (SDL) of the binding site of the $\alpha V\beta 3$ is indicated. Taken from Arnaout, M.A., et al., *Structure and mechanics of integrin-based cell adhesion. Current opinion in cell biology*, 2007. 19(5): p. 495-507.

Furthermore, the proximity of the CD loop to the hybrid domain observed in this structure has been contentious. In one study, the deletion of this loop in

$\alpha V\beta 3$ and $\alpha IIb\beta 3$ integrins did not cause constitutive integrin activation and eventual receptor binding [33]. And in another using the I-domain containing $\alpha M\beta 2$, mutation and deletion of this CD loop is found to be activating [34]. Whether or not the presence of an I-domain could affect the lower legs domain architecture requires further structural examinations.

1.12 Recent crystal structures of the extracellular domains of the integrin molecule

Just recently a crystal structure of the extracellular domain of the $\alpha IIb\beta 3$ integrin is elucidated by Zhu et. al. 2008, and previously unobserved Ca^{2+} and Mg^{2+} ions were captured in this crystal structure concluding that metal binding does not necessarily require prior changes in integrin binding affinity (Figure 1.7A) [35]. The overall arrangement of domains in this $\alpha IIb\beta 3$ crystal structure is similar to the $\alpha V\beta 3$ crystals [20], except for differences in angles between the domains [35]. The 2008 $\alpha IIb\beta 3$ crystal further revealed the details of the ‘genu’ which was found to occur between I-EGF 1 and I-EGF 2, concurring with the crystal structure of $\beta 2$ leg fragments [36], and is flanked by the PSI domain and a knob like projection from the thigh domain, the rest of the I-EGF domains 2 – 4 were observed extending in a straight orientation with 90° left handed twist between successive domains [35]. The I-EGF domains 2 – 4 are observed to be buried between domains of the upper $\beta 3$ leg and the αIIb leg. Some parts of the βI -like and β propeller domains also help to bury the lower $\beta 3$ leg I-EGF domains [35]. In this structure, the previously undefined part of I-EGF 4, termed the the β ankle was observed to be disordered. The authors suggested that the gimbal-like connection of the I-EGF domains permits rotation and the observed disordered loops at the β -knee, namely the I-EGF 1 and 2 domains implied flexibility at the PSI/I-EGF 1 and I-EGF 2/I-EGF 3 interfaces allowing for genu-linear extension [37].

The major contentions of this $\alpha IIb\beta 3$ structure were its lack of the last few leg residues, the use of disulphide bonds at the C-termini to obtain the crystal structure [37]. These issues were raised by Xiong et. al. (2009), following the

elucidation their $\alpha V\beta 3$ structure which consists the last few αV and $\beta 3$ subunit leg residues and four transmembrane residues [37]. The 2009 $\alpha V\beta 3$ crystal structure with residues extending into the transmembrane region was suggested to be a closer replica to the physiological bent inactive conformation, and this structure compared to the previous $\alpha V\beta 3$ [20,31] and $\alpha I\text{Ib}\beta 3$ [35] crystal structures without the last few residues of the ectodomain and transmembrane fragments is more compact and less active in solution (Figure 1.7B) [37]. This $\alpha V\beta 3$ structure further defined the extracellular domains adjacent to the cellular domains, including the I-EGF 1 and 2 domains, unlike the $\alpha I\text{Ib}\beta 3$ crystal structure [35] which showed these domains to be disordered. This study postulated that the I-EGF 2 and thigh domain interface is the stabilizing factor of the inactive state, and also suggest that genu-linearity may not be necessary for receptor binding.

Until now, only integrins without I-domains have been successfully crystallized and solved. The structural details of I-domain containing $\alpha X\beta 2$ extracellular domains with its C-terminal $\alpha\beta$ legs disulphide bonded, has only been recently unraveled. The structure displays marked differences from the $\alpha V\beta 3$ crystal structure, namely in the β -tail domains, the backbone of the 'genu', the closer packing of thigh and calf domains, the orientation of the calf-1 and calf-2 domains (Figure 1.7C) [35,38]. Most importantly, it revealed how the αX I-domain is accommodated in the context of the whole $\alpha X\beta 2$ extracellular domain. The αX I-domain in the structure is observed to be linked to the β -propeller via flexible linkers and extends away from the rest of the domain with small contact with the SDL of the β I-like domain [38].

To date all integrin crystal structures elucidated are incomplete lacking the full transmembrane and cytoplasmic domains and all of them are found crystallized in the compact bent conformation even in conditions, i.e. addition of Mn^{2+} and RGD, supposed to activate integrins to their high affinity elongated state. Additionally, the different integrin extracellular domains used and the use of disulphide constraints on the $\alpha\beta$ legs [35,38] have complicated efforts of

comparisons and conclusions. Crucially, the lack of complete domains failed to take into effects of transmembrane and cytoplasmic conformations, and the nature of data collection in X-ray crystallography which uses non-physiological buffers for protein crystallization, the nature of the protein in terms of solubility and flexibility, and crystal packing forces may stabilize the energetically low bent conformation and occlude the true nature of integrin conformational flexibility.

Figure 1.7 Recently elucidated crystal structures of the $\alpha\text{IIb}\beta 3$, $\alpha\text{V}\beta 3$ and $\alpha\text{X}\beta 2$ integrins. (A) Crystal structure of the $\alpha\text{IIb}\beta 3$ elucidated by Zhu, J., et al., *Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces. Molecular cell*, 2008. 32(6): p. 849-61. (B) Crystal structure of the $\alpha\text{V}\beta 3$ elucidated by Xiong, J.P., et al., *Crystal structure of the complete integrin alphaVbeta3 ectodomain plus an alpha/beta transmembrane fragment. The Journal of cell biology*, 2009. 186(4): p. 589-600. (C) Crystal structure of the I-domain integrin $\alpha\text{X}\beta 2$ elucidated by Xie, C., et al., *Structure of an integrin with an alphaI domain, complement receptor type 4. The EMBO journal*, 2010. 29(3): p. 666-79.

1.13 Observation of global structural transitions in integrin molecules using Electron Microscopy (EM)

In conjunction with the structural elucidation of the integrins extracellular domains, EM has also been used to observe integrin structural agility in different activating and deactivating conditions. Unlike the use of X-ray crystallography, EM is not hindered by packing contacts and each integrin can be viewed as an isolated particle. The EM studies can be separated into two arguments each purporting either the switchblade or deadbolt mode of integrin activation.

The studies purporting the switchblade model concluded that leg separation leads to integrin extension from the bent conformation which is necessary for receptor binding, and upon binding the integrin molecule adopts the extended and open headpiece due to the swing out of the hybrid domain. In these studies, using negatively stained EM, the $\alpha 5 \beta 1$ [39,40], $\alpha V \beta 3$ [27] and $\alpha IIb \beta 3$ [27] integrin extracellular domains are attached to coiled-coil helices, mimicking the physiological transmembrane and cytoplasmic regions (Figure 1.8) [27,39,40]. The clasped integrin can be unclasped through the use of a cleavage site engineered to occur between the C-terminal of the β subunit extracellular domain and the coiled coil. The release of constraints at the β -subunit, through cleavage of the coiled-coil, is hypothesized to mimic leg separation and results in the high affinity integrin conformation observed as extended particles in on the electron micrographs. The release of this constrain at the C-terminal of β -subunit resulted in extended conformations with higher binding affinities to their receptors, and when bound to their receptors an extended conformation with greater separation due to the swing out of the hybrid domain is observed (Figure 1.8A) [40]. Interestingly, the clasped and unclasped integrins behave almost similarly in the presence of activating (1mM Mn^{2+} or 1mM Mg^{2+}/Ca^{2+}) and deactivating (5 mM Ca^{2+}) metal ions, except that the unclasped integrins have a higher percentage of extended high affinity forms during activating conditions in the electron micrographs. Bent and extended forms are observed to exist together in different percentages depending on conditions and a continuum of bent and extended forms are observed. These

observations suggest the existence of equilibrium amongst integrin conformations, and the presence of external factors can tilt or promote the formation of certain conformations. In another similar study, using I-domain containing integrins, $\alpha\text{L}\beta 2$ and $\alpha\text{X}\beta 2$, further suggested that this equilibrium may be different for each integrin molecule (Figure 1.8E, F and G) [26,41]. This study has cleavage sites connected to both the α and β C-termini and the release of the clasp showed that the β -leg adopts more conformations compared to the almost invariant α -leg in most integrin particles (Figure 1.8E and F) [26]. Apart from showing that leg separation in I-domain integrins results in extension and activation, the EM study on $\alpha\text{L}\beta 2$ and $\alpha\text{X}\beta 2$ also demonstrated that by using activating monoclonal antibodies that target the integrin receptor binding I-domain or β -leg regions can also cause extension despite being clasped at the $\alpha\beta$ C-termini (Figure 1.8G) [26,41]; hence, implying the reversibility of conformational changes whether mediated extracellularly or intracellularly. Additionally, it is also revealed that the many intermediate forms, extended with a closed headpiece, have low affinity for their receptors thus emphasizing the open headpiece requirement in receptor binding [41]. The hybrid domain swing out contributed by the open headpiece and structural flexibility at the lower β -leg as opposed to the α -leg stability was also mapped in a 3D structure using electron tomography in combination with previously obtained structure [42]. These studies conclude that allostery from the leg and to the hybrid domain, and structural changes in the receptor binding domain (β -like/ β propeller domain or I-domain) can be transmitted to the flexible β -leg.

Figure 1.8 Electron micrographs purporting the switchblade model. Taken from Springer, T.A. and M.L. Dustin, *Integrin inside-out signaling and the immunological synapse. Current opinion in cell biology*, 2012. 24(1): p. 107-15.

In contrast, EM studies supporting the deadbolt model hypothesized modest structural changes of the bent conformation under activating conditions, and receptor binding hardly caused the reported drastic angle change in the hybrid domains. Cryo-EM was used to observe the full length $\alpha IIb\beta 3$ obtained from human blood, and a less compact, bent conformation with different arrangement leg domains from the 2001 $\alpha V\beta 3$ structure was observed [43]. Following transmission-EM and single particle image analysis on the extracellular domain of the $\alpha V\beta 3$ in the presence and absence of its receptor, Fibronectin, under metal-ions

activating and deactivating conditions failed to find extended conformations [44]. In another study using cryoelectron tomography of full length $\alpha\text{IIb}\beta 3$ incorporated into small liposomes found no significant height change between the observed integrin particles, as postulated by the switchblade model, in spite the presence of Mn^{2+} ions [45]. In support of modest changes during integrin activation, the Fluorescent Resonance Energy Transfer (FRET) distance between the dyes conjugated at the $\alpha 4\beta 1$ headpiece and the surface of the plasma membrane was only 50 Å compared to the measured distance of 200 Å from the headpiece to the membrane proximal extracellular domain of the extended integrin [46]. In a more recent study using an improved FRET technique known as Fluorescent life-time imaging (FILM), which minimizes the problem of dye flip-flop between the outer and inner leaflets of the membrane, also did not find any significant height change in the un-activated and activated full length $\alpha\text{IIb}\beta 3$ integrin on live cells [37]. A modeling study combining experimentally determined crystal [20,31], NMR [47], transmembrane and cytoplasmic [48] structures with known hydrodynamic dynamic data [26,40,41] of bound and unbound integrins to reconstruct the $\alpha\text{V}\beta 3$ extracellular domain and full length $\alpha\text{IIb}\beta 3$ integrin found the that integrins are extended in the resting state and may collapse fully or partially to the bent conformation depending in conditions. This study postulates that receptors can bind to integrins regardless, bent or extended, with only modest changes and that the β -legs have to be rigid in order for changes to be transmitted to the hybrid domain for receptor binding [49]. However, these bent EM observations have been further contradicted in the recent EM revelations of the intact full length $\alpha\text{IIb}\beta 3$ integrin embedded on lipid nanodiscs and detergent [50].

The greatest contentions against conclusions of the switchblade model is its lack of transmembrane and cytoplasmic domains, although the transmembrane mimic coiled-coils were used, differences in electron micrographs of full length integrins strongly suggests that the effects of these domains cannot be discounted. Similarly, arguments presented by the deadbolt supporters are also controversial due to different EM techniques used, i.e manually picked [26,27] and randomly

picked structures [43,44]. It is also hard to compare between studies using different type of integrin heterodimers, which have intrinsically disparate functions in the cell and hence different conformation requirements. Additionally, EM characterization also uses non-physiological buffer components and cannot distinguish different orientations in the same conformation; hence conformations observed have to be evaluated carefully.

1.14 Detailed structural insights into the individual domains of integrins

The earliest crystal structures [20,31] determined has certain domains in the leg regions of the β -subunit, such as the hybrid, PSI, and I-EGF 2 and 3 domains, which are securing 'genu', are disordered in the crystal structure thus obscuring the detailed interface that is formed between these domains. Although, some of these lapses in structural details were circumvented in the recent crystal structures determined [35,37,38], there are still disputes as to which domains are important in stabilizing the bent structure. Structural lucidity in these domains is thus crucial to understand the forces in the integrin molecule maintaining the bent, low affinity conformation. Additionally, detailed structural characterization of the receptor binding domain at the interface of the α -subunit β -propeller and the β subunit I-like domain could provide mechanistic insights into the conformational changes that occur during receptor binding. And the lack of the transmembrane and cytoplasmic regions in the integrin crystal structures, regrettably due to technical difficulties that has not yet been overcome in crystallizing the insoluble transmembrane region, requires them to be studied as separate modules. Detailed structural insights into these regions of the integrin molecule are required for deciphering the mode of integrin activation.

1.15 Structural insights into the integrin β -subunit domains

NMR solution structures of the I-EGF 2 and 3 domains of the β 2-subunit further enhanced the α V β 3 crystal structure at the β -leg domains [47]. These two modules are observed to be stabilized by cysteine bonds and are related to each

other by a two-fold axis and are positioned in an extended arrangement with respect to each other [47]. Superposition of the NMR structure on the $\alpha V\beta 3$ crystal structure shows that these 2 domains are buried in the hybrid and PSI domains, and have a similar orientation as the I-EGF 3 and 4 domains in the bent $\alpha V\beta 3$ crystal structure [20]. Release of contacts from these buried domains is proposed to trigger a switchblade extension of the integrin, purporting the switchblade mode of integrin activation [47]. The PSI domain has also been crystallized to reveal its structural features, the disulphide bonds stabilizing the domain and is suggested to interface with the I-EGF-2 domain [51]. Further studies on other extracellular domains using X-ray crystallography on the hybrid/PSI/I-EGF 1-3 domains excluding the β I-like domain of the $\beta 2$ integrin have also shown that the truncated molecule are maintained by disulphide bridges and can become elongated in shape in the addition of the I-EGF 2 and 3 domains[52]. This truncated $\beta 2$ structure correlates well to the $\alpha V\beta 3$ structure and also showed that the ‘genu’ occurs at the I-EGF 1 and 2 domains [20], with the PSI domain observed to interface extensively with the hybrid and I-EGF 1 domains. Interestingly, when this truncated $\beta 2$ structure was expanded to consist of the I-EGF 2 and 3 domains, the additional domains resulted in an extended conformation instead of the bent conformation [36]. The addition of I-EGF 3 domains seems to disrupt disulphide bonds and domain architecture, causing a change in the angle between the I-EGF 1 and 2 domains which leads to an extended structure. The authors postulate that the release of contacts between the hybrid, PSI and I-EGF 1 domains due to the twist between the I-EGF 1 and 2 domains could be part of the mechanism for the conversion of the bent, low affinity state to the extended high affinity state [36]. These studies show that the β -legs are highly flexible, as evident from the disorder observed in the crystal structure of the $\alpha V\beta 3$ integrin [20] , and on some EM studies, they are observed as a blur compared to the fixed and stable α -subunit [42]. However, there are also EM studies described in Section 1.13 supporting a rigid β -leg and flexible α -leg.

1.16 Structural insights into the integrin receptor binding domains

Besides structural distortions in the $\alpha V\beta 3$ structures, conformational changes in the integrin receptor binding site located in the β -I like domain and the α -subunit β -propeller in the presence of their receptors has not been clear in the $\alpha V\beta 3$ crystals [20,31]. The lack of structural changes in the $\alpha V\beta 3$ RGD crystal complex could not provide insights into the possible drastic conformational changes that can occur during receptor binding. Hence, in order to elucidate the structural changes that will occur during receptor binding, crystal fragments containing the receptor binding headpiece, which includes the αIIb subunit β -propeller domain and the $\beta 3$ subunit I-like, hybrid and PSI domain, were used. These crystal fragments were soaked either in the absence or presence of RGD mimics in order to determine the distinct structures representative of the bent, closed headpiece, low affinity form versus the open, open headpiece, high affinity form (Figure 1.9 A,B and C) [53]. In the open, high affinity form, the $\alpha 7$ helix of the $\beta 3$ I-like domain is observed to move downwards causing the hybrid and PSI domains to swing away from the αIIb subunit, see Figure 1.9C [53]. This downward movement of the $\alpha 7$ helix is caused by the disruption of the coordination of the Ca^{2+} ion with a methionine residue at the $\beta 3$ I-like ADMIDAS which results in the remodeling of adjacent loops connected to the $\alpha 7$ helix [22]. The lack of significant structural changes in the $\alpha V\beta 3$ crystal structure [31], and the dramatic observation of the $\beta 3$ hybrid domain swing out has been suggested to be due to the missing lower β legs present in the $\alpha V\beta 3$ crystal structure, which acts as constraints in the bent structure.

Until recently [38], there has been no crystal structure available for integrins containing the α I-domain which serves as the major receptor binding domain for these integrins. Studies on only the I-domain previously have provided detailed insights into the conformational changes during receptor binding. Most crystal studies on I-domain containing integrins have been focused on only the I-domain of the α -subunit, in the presence and absence of their receptors, and of conformational mutants favouring the different affinity states [23]. Thus far, the

conformations of the α M I-domain in the presence of Mg^{2+} [54] and Mn^{2+} [55], the α 2 I-domain in complex with a triple helical collagen peptide [56], and the α L I-domain with mutationally introduced disulfide bonds bound to ICAM 1 and 3 have been crystallized [57]. The α I-domain differs from its structurally similar counterpart in the β I-like domain by 2 additional segments which only occurs in the β I-like domain (Figure 1.9D, E and F) [22]. One of these segments forms the interface with the β propeller and the other is the specificity determining loops (SDL). Studies in the presence and absence activating metal ions showed changes in the coordination of residues at the MIDAS site in the α I-domain leading to the open and closed α I-domain conformations representative of the high and low affinity states[54,55]. The change in the coordination of the MIDAS residues following metal ion binding is coupled with backbone and side chains rearrangements of the loops in tandem with the perturbed residues. However, unlike the I-domain less integrins, transmission of conformational changes that has occurred after extracellular receptor binding in the I-domains is transmitted to the β -I like domain first before the downwards displacement of the α 7 helix (Figure 1.9F) [58], which will cause the hybrid domain to swing out similar to high affinity conformation observed in the I-domain less integrins (Figure 1.9C) [53]. The dramatic changes observed in a singular domain are postulated to translate to the flexible integrin leg domains that will eventually lead to the high affinity, extended integrin form.

Figure 1.9 The overall integrin conformation states of the integrins without (A,B,C) and with (D,E,F) the α I-domain. The bent and extended conformations with closed headpiece (A,B/D,E) with low affinity. Extension, followed by hybrid domain swing out at its interface with the β I domain (C) is connected through the β I α 7-helix to rearrangements at the β I interface with the β -propeller domain that greatly (~ 1000 -fold) increase affinity for ligand in the extended-open conformation. Similar interdomain rearrangements in α I integrins (F) result in activating a binding site for an internal ligand, Glu310 in α L, which pulls down the α I α 7-helix to activate a similar increase in affinity (~ 1000 – $10,000$ -fold) of the α L I domain for its extrinsic ligand ICAM-1. Taken from Springer, T.A. and M.L. Dustin, *Integrin inside-out signaling and the immunological synapse. Current opinion in cell biology*, 2012. 24(1): p. 107-15.

1.17 Structural insights into the integrin transmembrane and cytoplasmic regions

The transmembrane and cytoplasmic regions are missing from the constructs used for crystallization and for some of the EM studies; however, it has become increasingly important to include these regions in the evaluation of integrin structural changes in the light of several conflicting studies that are in support of receptor binding to both the distinct bent and extended conformations. The different conformations observed in crystal structures of the $\alpha\text{IIb}\beta 3$ ectodomain with C-terminal disulphide coiled-coils [35] vs. the $\alpha\text{V}\beta 3$ ectodomain with parts of the transmembrane residues [37]; and the EM images of full length $\alpha\text{IIb}\beta 3$ integrin [43] vs. c-terminal disulphide locked $\alpha\text{IIb}\beta 3$ integrin ectodomains [39] demonstrated the importance of the transmembrane and cytoplasmic domains in determining integrin conformations. Additionally, studies of the conformational changes of full length integrins on cells detected using monoclonal antibodies that support genu-linearity during activation and with findings using FILM that the headpiece of the full length $\alpha\text{V}\beta 3$ do not move much away from the cellular membrane [37] further underscored the importance of the transmembrane and cytoplasmic domains in the evaluation of integrin activation pathways. It is generally agreed upon that the separation of the transmembrane and cytoplasmic regions leads to integrin activation and increased binding to their extracellular receptors, whether or not it leads to global conformational changes is still very much debatable.

Before structural data were available, the integrin transmembrane region was predicted based on the presence of conserved sequences in other homologous transmembrane proteins, and the first charged residue after the predicted transmembrane region was taken as the beginning of the cytoplasmic domain. However, structural insights into these regions is starting to indicate a vague transmembrane-cytoplasmic border at the lipid/aqueous bilayer in order to facilitate intermediary conformations that occur during integrin activation processes. The integrin transmembrane domain spans the cellular membrane, and the interaction between the α - and β - subunit transmembrane regions is postulated to dictate the

activation status of the integrin molecule. Sequence analysis of the integrin transmembrane regions reveals motifs (GxxxG) homologous to the Glycophorin A (GpA) protein that is known to form a membrane spanning helical coiled-coil interface between two GpA proteins [59]. These motifs are the GxxxG motif in the α -subunit and the SxxxA motif found in the β -subunit, which suggests hetero-dimerization tendencies between the α - and β - subunit to maintain the resting state integrin; and homo-dimerization tendencies with other integrin transmembrane domains important in integrin clustering processes [59]. Recent gains in structural data of the full length transmembrane and cytoplasmic regions has separated the transmembrane interaction regions into the outer membrane clasp (OMC) and inner membrane clasp (IMC), see Figure 1.10, which are suggested to hold key transmembrane residues that interact and keep the transmembrane clasped in the inactivated state [60]. Additionally, the interactions between the integrin relatively short α - and β -cytoplasmic regions (except β 4 cytoplasmic domain), which forms part of the IMC, also modulates the separation of the transmembrane regions and hence conformational changes in the ectodomains leading to activation. The importance of the cytoplasmic regions has been elegantly demonstrated using FRET of two different dyes attached to the C-terminus the α - and β - cytoplasmic tails; where FRET is increased in the inactivated state with clasped transmembrane and cytoplasmic regions and is subsequently reduced when they separate [61].

Figure 1.10 Two definitions of the Outer Membrane Clasp (OMC) and Inner Membrane Clasp (IMC) of the $\alpha_{IIb}\beta_3$ integrin. (A) From Ulmer, T.S., *Structural basis of transmembrane domain interactions in integrin signaling. Cell adhesion & migration*, 2010. 4(2): p. 243-8. (B) Yang, J., et al., *Structure of an integrin $\alpha_{IIb}\beta_3$ transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. Proceedings of the National Academy of Sciences of the United States of America*, 2009. 106(42): p. 17729-34.

1.17.1 Functional characterization of the Outer Membrane Clasp (OMC)

Due to the technical difficulties of studying the insoluble integrin transmembrane regions with most structural methods, early characterization of the α -/ β -subunit transmembrane interfaces involve using mutagenesis methods to map out the transmembrane regions and interactions. The N- and C- terminus region of the integrin α - and β -subunit transmembrane region is mapped out by using a glycosylation strategy which makes use of the glycosylation enzyme, oligosaccharyl transferase to transfer a glycan group to a Asn acceptor site. Glycosylation efficiency is reduced when the acceptor site is moved closer to the transmembrane segment [62]. The β -transmembrane region is mapped to 22

residues whereas the α -transmembrane region is mapped to about 25 to 29 residues depending on sequence of the different integrin α -subunits, the longer spanning α -transmembrane region is suggested to tilt in order to accommodate the extra residues, whereas the β -transmembrane region is positioned vertically into the membrane [62]. However, the recent structural characterizations of the α IIb β 3 integrin transmembrane and cytoplasmic domains have shown this to be otherwise [63,64]

The heterodimeric interaction interface between the α - and β -transmembrane is characterized by a series of cysteine scanning and mutagenesis experiments conducted on intact α IIb β 3 integrins on the surface of mammalian cells. Transmembrane residues in the α - and β -subunit are randomly mutated and activating mutations that disrupt α - and β -transmembrane interactions are postulated to be at the $\alpha\beta$ transmembrane interface [65]. Mutations of these transmembrane interfacial residues to Leu also disrupted transmembrane interactions and are activating [66]. Cysteine-scanning experiments also showed cysteine mutations of these residues results in the formation of disulphide bonds that locks the integrin in the clasped, inactivated state, where even activating mutations at the cytoplasmic region could not rescue these disulphide bonded mutants [66,67]. Additionally, homo- and hetero- oligomerization motifs in the transmembrane regions have also been characterized and verified using a series of bacterial reporter systems, such as the GALLEX [65] and TOXCAT [68,69] assay, which fuses the short transmembrane regions to bacterial sequences that reports the oligomerization status either through the formation a paired or unpaired product that represses or promotes the formation of a bacterial gene product.

The residues involved in the α IIb β 3 heterodimeric transmembrane packing are found to be the α IIb W968, G972, G976 and L980; and the β 3 V696, L697, V700, M701, I704, L705, G708, and L709. However, only mutations of the α IIb G972, G976 and β 3 G708 are consistently found to be the most activating and structurally disruptive [70]. Only the α IIb residues G972 and G976 correspond to

the α - GxxxG motif of the transmembrane OMC but not the β 3 G708 which does not correspond to the conserved β - SxxxA motif. Heterodimerization studies have shown that the Ser and Ala of the β 3-S⁶⁹⁹VMGA⁷⁰³ (or SxxxA) motif to be pointing away from the $\alpha\beta$ transmembrane interface and do not participate in the $\alpha\beta$ heterodimerization interface [71]. Additionally, the motifs involved in heterodimerization are likely to be different for different integrins. In the α IIB β 3 transmembrane interface, α IIB G972 and G976 corresponding to the GxxxG motif at the OMC, are involved in the heterodimerization with its β 3 partner. However, in β 2 integrins, instead of the α -transmembrane GxxxG motif at the OMC, a SxxxG motif is present. The α -subunit Ser is found to be important in maintaining heterodimeric interactions between the $\alpha\beta$ transmembrane OMC with the formation of a hydrogen bond with the β 2-transmembrane Thr at the OMC region [72].

Integrin transmembrane regions not only heterodimerize to maintain the clasped inactivated state, they also mediate homooligomerization during the clustering of integrins, an important process in the formation of focal adhesion sites in the migrating cell. The tendency of transmembrane regions to homooligomerize has been shown in peptides containing α IIB and β 3 transmembrane and cytoplasmic regions in Dodecylphosphocholine (DPC) micelles [73]. It is found that the α -transmembrane GxxxG motif at the OMC not only mediates heterodimerization but also homodimerization, however, the packing of the α -subunit with different residues in the β -transmembrane distinguishes between these 2 processes. Asn scanning analysis of α M β 2 transmembrane regions reveals that homodimerization in α M and β 2 transmembrane segments are driven by GxxxG motifs in both α M and β 2 [74]. Homooligomerization motifs are likely to involve similar motifs in different integrins as the GxxxG motif has been found in several transmembrane proteins that homooligomerizes such as the GpA transmembrane dimer [59].

1.17.2 Functional characterization of the Inner Membrane Clasp (IMC)

The IMC which forms the lower transmembrane region that constrains the integrin in its resting state, consists of the membrane proximal regions of the

cytoplasmic domain. Experiments that replace the cytoplasmic domain completely with acidic and basic coiled-coil peptides also showed that the close association at the α - and β - cytoplasmic region with opposing charged acidic and basic peptides locks the integrin in the bent conformation which do not react to activation reporter antibodies [75]. Conversely, separation at the cytoplasmic regions leads to integrin activation [75]. However only interactions between the membrane proximal (MP) residues of the cytoplasmic domain, α -(GFFKR) and β -(KLLxxxHDRRE) have been shown to be important in maintaining the resting state of the integrin, disruption of the interaction between these two MP sequences, via truncation after the α -G residue [76], deletion of cytoplasmic α -(FFKR) and/or β -(LLxxxHD) [77], or mutation of the conserved cytoplasmic α -(----R) and β -(-----D---) that disrupts salt bridge formation between them results in constitutively activated integrins. Alanine scanning experiments also showed that the membrane proximal α -(-FF-R) and α -R. β -D salt bridge are important in maintaining the integrin resting state [78]. Besides these integrin membrane proximal residues, it is becoming more important to consider the first conserved Lys/Arg on the postulated β -cytoplasmic region, where mutation of this basic charged amino acid to the non-charged Ala results in constitutive activated integrins on mammalian cells [79]. This Lys/Arg is proposed to ‘snorkel’ via their positive charge near the negatively charged phospholipid headgroups and bring the β -subunit close enough to the α -(-FF--) motifs, so as to affect the crossing angles of the transmembrane regions which maintains the integrin in the clasped inactivated state [79].

1.17.3 Structural characterization of the integrin OMC and IMC

The functional experiments described in Section 1.17.1 and 1.17.2 provide a general idea of how the integrin α - and β - subunit transmembrane and cytoplasmic domains interact and separate to form the clasped heterodimer in the inactivated and activated integrin. However, how the α - and β - subunit are topologically arranged in the plasma membrane can only be visualized in structural studies. Detailed structural insights in addition to the above mentioned functional experiments could provide a more accurate picture of the integrin transmembrane and cytoplasmic control of integrin activation. Structural elucidation using X-ray,

NMR or EM, of the transmembrane regions, which forms the integrin OMC, has been difficult thus far. Most structural visualizations of the transmembrane regions have been based on modelling calculations and focused on the $\alpha\text{IIb}\beta 3$ integrin. Unlike the transmembrane region, the integrin cytoplasmic domains which forms the IMC, due to their short sequences have been relatively well characterized by NMR studies.

The transmembrane structures of the resting state integrin have been modeled for the $\alpha\text{IIb}\beta 3$ transmembrane regions based on cryo-EM data [43], modeled with or without experimental restraints described in Section 1.17.1 and 1.17.2[48,80,81], or based purely on evolutionarily conserved data of heterodimeric transmembrane proteins [82]. Some of the modeled structures have also used NMR derived cytoplasmic domain structures as part of the modeling calculation but still did not manage to fit in the conserved salt bridge interaction between α -Arg and β -Asp [81]. Moreover, these modeled structures are distinct from each other with different α - and β - transmembrane crossing angles, hence several transmembrane activation models of how the α -/ β - transmembrane domains interact to effect separation and global extracellular changes have been proposed. The three proposed models suggests global extracellular changes can be effected by the piston, scissor or separation movements of the α - and β - transmembrane domains (Figure 1.11) [59]. In the piston model, the transmembrane domains move vertically during integrin activation; and in the scissor model, the transmembrane domains interact with a small crossing angle in the resting state and increases following activation; and finally in the separation model, the transmembrane simply separates upon activation [59].

Figure 1.11 Three models of transmembrane movements during integrin activation, α -subunit transmembrane helix in cyan, and β -subunit transmembrane helix in red. The **inactive** transmembrane conformation can undergo **piston**, **scissor** and **separation** like movements during separation. Taken from Wegener, K.L. and I.D. Campbell, *Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions* (review). *Molecular membrane biology*, 2008. 25(5): p. 376-87.

However, which of these models reflect the actual physiological states during integrin activation are still controversial due to the lack of consideration of the extracellular and cytoplasmic domains until the recent structural elucidation, using NMR, of the α IIB β 3 integrin transmembrane and cytoplasmic structure in bicelles and the subsequent modeling calculations that ensued. These studies support the model of separation. The individual α Ib- and β 3- subunit NMR structures containing the integrin transmembrane and the membrane proximal cytoplasmic residues were obtained in bicelles which represented the activated separated integrin conformation. The α Ib structure which consists of Ala⁹⁵⁸ to Pro⁹⁹⁸ adopts a straight linear 24-residue α -helical structure from residue Ile⁹⁶⁶ to Lys⁹⁸⁹ that spans the bilayer followed by an atypical backbone reversal that packs the two conserved membrane proximal Phe⁹⁹³ and Phe⁹⁹² against the transmembrane residue Trp⁹⁸⁸ (Figure 1.12) [63]. The β 3 structure consisting of residues P⁶⁸⁵-F⁷²⁷ of which 30 residues, I⁶⁹³-H⁷⁷², forms a linear α -helix embedded in the bicelles [64]. The 30 β 3 transmembrane residues linearly spanning the lipid bilayer meant that it has to be tilted in order for it to be accommodated into bilayer; such a tilting would bring the first K⁷¹⁶ to lie at lipid/aqueous interface which facilitates its snorkelling (Figure 1.12). The importance of the snorkelling K⁷¹⁶ in maintaining the resting state integrin has been demonstrated in mutagenesis studies, where mutation to Ala resulted in constitutively activated integrins [79]. Following the structural elucidation of the dissociated integrin transmembrane and cytoplasmic domains, the α Ib β 3 complex is then calculated by the enrichment of close inter-proton distances provided by the selective methyl labelling of the transmembrane residues and use of some experimental restraints, such as the R⁹⁹⁵ and D⁷²³ salt bridge [83]. The α Ib- and β 3- transmembrane is observed to cross within their N-terminal halves at 25° forming the OMC (Figure 1.12A) [83]. This complex structure concurred with the earlier mutagenesis experiments described in section 1.17.1 and 1.17.2 that characterized the transmembrane heterodimeric interface, showing the packing interactions of the α Ib G⁹⁷², G⁹⁷³ with β 3 G⁷⁰⁸ at the OMC (Figure 1.12B). However, at the IMC, instead of the transmembrane

residues forming this clasp, the α IIb conserved F⁹⁹² and F⁹⁹³ turns and immerse into the bilayer and come into close contact with the β 3 transmembrane W⁷¹⁵, which also brings the conserved R⁹⁹⁵ and D⁷²³ close enough to form a salt bridge in the low dielectric environment of the lipid head groups (Figure 1.12C) [83]. A model of activation based on the structure of this complex β has been proposed (Figure 1.12D). The N-terminal of the transmembrane which consist of the OMC connected to the extracellular lower leg assembly and the IMC maintains the integrin in resting conformation. The resting state integrin can be activated via the direct disruption of the OMC or IMC. Disruption of the OMC occurs via receptor binding to the extracellular domain brings about conformational changes that separates the associated extracellular leg domains, whereas the IMC is removed by the binding of intracellular ligands that disrupts the salt bridge and conserved phenylalanine packing (Figure 1.12D). Both processes lead to separation of the transmembrane and cytoplasmic regions.

Figure 1.12 Conformations of the $\alpha\text{IIb}\beta 3$ transmembrane/cytoplasmic domains and model of activation in lipid environment. (A) Overall conformation of the $\alpha\text{IIb}\beta 3$ transmembrane and cytoplasmic regions. (B) Residues in contact with each other at the OMC. (C) Residues in contact with each other at the IMC. (D) Model of integrin activation from the bent integrin (middle panel) conformation to the activated extended conformation activated intracellularly by talin (left panel) and extracellularly (right panel). Taken from Ulmer, T.S., *Structural basis of transmembrane domain interactions in integrin signaling. Cell adhesion & migration*, 2010. 4(2): p. 243-8.

Interestingly, a similar α IIB β 3 transmembrane-cytoplasmic region complex conformation for the resting state is obtained in a separate modelling study which made use of experimental disulphide data as distance restraints to model α IIB β 3 transmembrane-cytoplasmic region. This modelling study is different from previous modelling studies as it is completely derived from experimental data without biased inclusion of the conserved R⁹⁹⁵ and D⁷²³ salt bridge, and the modelling program takes into consideration the effects of hydrophobic lipid bilayer in modulating the interactions between the integrin α - and β - transmembrane and membrane proximal residues [67]. This complex structure has a similar transmembrane inter-helical packing and angle of α - and β - transmembrane crossing; in addition, the unusual α IIB F^{992/993} reverse packing to the α IIB-transmembrane is also observed (Figure 1.13). The main differences between this Rosetta/disulphide modelled complex (Figure 1.13B) [67] and the NMR/experimental restraints derived structure (Figure 1.13A) [83] lies in the β 3 transmembrane helix and the α IIB F^{992/993} packing with the β 3 transmembrane. The β 3-transmembrane region of the Rosetta/disulphide modelled complex embedded in the lipid bilayer is shorter, 23 residues compared to the 29 residues in the NMR structure, suggesting that β 3 is no longer tilted in the bilayer; and the α IIB F^{992/993} packs closer to the snorkelling K⁷¹⁶, which also brings the canonical R⁹⁹⁵ and D⁷²³ salt bridge residues closer to each other (Figure 1.13) [83]. Extrapolating the transmembrane and cytoplasmic regions to include the crystal structure of the recently elucidated α IIB β 3 ectodomain [35] showed that the transmembrane is linked flexibly to the lower legs of the ectodomain. The flexible linkage implies that a large scale movement, such as the separation of the transmembrane and cytoplasmic regions are required for global changes in the extracellular domains [67]. The differences between these two complex structures have been postulated to be transitory states that occur during the integrin activation process which involves the transmembrane and cytoplasmic regions. The NMR structure of the individual transmembrane and cytoplasmic domains in bicelles represent the monomeric, final state of integrin activation, whereas the Rosetta/disulphide complex structure represents the resting state integrin where the shorter lipid

embedded $\beta 3$ -transmembrane brings the snorkelling $\beta 3$ -K⁷¹⁶ closer to the α IIB-F^{992/993}. The NMR complex structure with longer lipid embedded $\beta 3$ -transmembrane is suggested to be the intermediary state of separation with the snorkelling $\beta 3$ -K⁷¹⁶ positioned further away [70]. There are also several other modelling studies that have accurately predicted the transmembrane and cytoplasmic inter-helical interface, although the slight differences that exist in these structures could be thought as intermediary states that occur during integrin activation [82,84].

Figure 1.13 Conformations of the α IIB $\beta 3$ interaction interface, NMR derived and modelled. (A) The NMR derived α IIB $\beta 3$ transmembrane/cytoplasmic structure zoomed into the IMC interface by Lau, T.L., et al., *The structure of the integrin α IIB $\beta 3$ transmembrane complex explains integrin transmembrane signalling. The EMBO journal*, 2009. 28(9): p. 1351-61. (B) The disulphide/Rosetta modelled α IIB $\beta 3$ complex at the IMC interface by Zhu, J., et al., *The structure of a receptor with two associating transmembrane domains on the cell surface: integrin α IIB $\beta 3$. Molecular cell*, 2009. 34(2): p. 234-49.

The NMR-derived α IIB $\beta 3$ transmembrane/cytoplasmic structure lacked the rest of the cytoplasmic residues and exhibited a completely different cytoplasmic conformation compared to the previous NMR-derived structures of the integrin cytoplasmic regions, see Table 1.2. The truncation at the cytoplasmic regions that resulted in shorter sequences could possibly contribute to additional conformational flexibility and energetic favourability for excursions of the α IIB-

F^{992/993} motif into the hydrophobic core of the bilayer. A NMR structure of the transmembrane and full length cytoplasmic region that is published not long after the Lau, et. al (2009) structure, did not observe this unusual α Ib-F^{992/993} reverse turn and presented a different integrin activation mechanism with contrasting integrin associated and dissociated states (Figure 1.14) [85]. The structure of the α Ib β 3 transmembrane-cytoplasmic domain complex has been obtained in a 1:1 mixture of organic solvent CD₃CN and water. The α Ib region E⁹⁶⁰-E¹⁰⁰⁸ and the β 3 region K⁶⁸⁹-T⁷⁶² were used, compared to the Lau et. al. (2009) and the Rosetta/disulfide complex structure, more residues assume a helical conformation that extends well into the cytoplasmic regions (Figure 1.14A) [85]. The helical portion of the α Ib extends from I⁹⁶⁶-N⁹⁶⁶ well into the distal cytoplasmic regions, and for β 3, it extends from I⁶⁹³-A⁷³⁷. The structure of the cytoplasmic regions concurred with several NMR structures of the cytoplasmic domains. Instead of the left-handed turn at the α Ib-GFF motif that packs back to the α Ib transmembrane residues, a straight right handed helix which extends from the transmembrane is observed. Despite this obvious disparity, the backbone for the α Ib β 3 transmembrane region and an almost similar crossing angle of 30° between the α - and β -transmembrane helices occurring close to the N-terminal half of the transmembrane is almost similar to the NMR structure in bicelles by Lau et. al.(2009) (Figure 1.14B). Additionally, the widely observed OMC packing in experimental and structural data involving the α Ib-G^{972/973} and β 3-G⁷⁰⁸ is too observed. The N-terminal α - and β -regions that connects to the extracellular domains are spatially close to each other and adopt short loop conformations that can be superposed to the C-terminal legs of the α Ib β 3 crystal structure by Zhu et. al. (2008) [35,85].

Figure 1.14 Conformations of the $\alpha_{IIb}\beta_3$ transmembrane/cytoplasmic domains elucidated in acetonitrile/water. **(A)** The full $\alpha_{IIb}\beta_3$ transmembrane/cytoplasmic complex with interfacial residues at the OMC and IMC displayed derived. **(B)** Overall conformations of the full $\alpha_{IIb}\beta_3$ transmembrane/cytoplasmic complex. **(C)** Superposition of the $\alpha_{IIb}\beta_3$ heterocomplex structure calculated by Lau et al. (2008) in bicelles (green), and Yang et al. (2009) in perdeuterated acetonitrile (red) and in aqueous solution (blue). Taken from Yang et al. (2009). *Structure of an integrin $\alpha_{IIb}\beta_3$ transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. Proc Natl Acad Sci U S A., 106(42): p. 17729-17734.*

The similarities, albeit the GFF reverse turn, suggest a physiologically competent $\alpha_{IIb}\beta_3$ transmembrane and cytoplasmic domain structure despite being

obtained in a non-physiological hydrophobic environment. The possibility of the GFF reverse turn in the calculated structure [85] is further tested using the modelling software, CS-Rosetta, also produced a population of structures containing the reverse turn. The authors conclude this finding with the lack of observed inter-NOEs (Nuclear Overhauser Effects) between the α IIB-F^{992/993} and β 3 transmembrane in their NMR spectra that the NMR structure of the transmembrane-cytoplasmic domain in bicelles [63] represents the intermediate conformation leading to separation instead of being directly involved in the resting integrin IMC. The F^{992/993} has been shown in some mutational studies to facilitate the binding of intracellular negative regulators such as Calcium-Integrin binding (CIB) protein to the α IIB cytoplasmic domain, hence its reversal which places them and the salt bridge into the hydrophobic milieu of the bilayer would seem to impede the interaction [85]. Hence, more in-depth study into the possibility of the GF^{992/993} reversal would be required. As this motif (GFFKR) is conserved across all the 18 integrin α -subunits, it would be interesting to see if the GFF reversal into the hydrophobic milieu of the bilayer is possible for the rest of the α -subunits. A more complex mechanism of integrin activation involving several states seems more feasible instead of the direct disruption of the IMC proposed by Lau et. al. (2009) and Wang et. al. (2010). This model of integrin activation suggests the requirement of intracellular proteins to drive the stabilization of the resting or promote the activated state (Figure 15).

However, a more recent study radically challenged these previously determined integrin intracellular borders, placing them further down into the cytoplasmic membrane proximal regions of the α IIB and β 3 for both the active and inactive integrin states [86]. Unlike prior methods that uses *in vitro* means, this study have subjected the full length α IIB β 3 into a cellular system and determined its intracellular borders using a Substituted Cysteine Scanning Accessibility Method (SCSAM) [86]. They have sequentially mutated each and every residue which constitute the OMC and IMC with the polar unchanged amino acid, cysteine, and assessed its aqueous exposure to Biotin Maleimide (BM) labeling [86]. Their study

has localized the α Ib membrane proximal region, $^{989}\text{KVGFFKR}^{995}$ together with downstream residues up to P998 into the hydrophobic core of the lipid bilayer in the activated integrin state. And in the inactivated integrin state, lipid partitioning of the α Ib cytoplasmic tail also involve the membrane proximal region up to K^{994} , one residue short of the R^{995} that is directly involved in the salt bridge formation with $\beta 3 \text{ D}^{723}$. Conversely, this dramatic intracellular border changes for the active and inactive integrin were not observed in the $\beta 3$ cytoplasmic tail. Nonetheless, this study still observed a greater hydrophobic core partitioning of the $\beta 3$ membrane proximal residues up to F^{727} which includes D^{723} that is directly involved in the canonical salt bridge [86], . During the resting state, the α Ib membrane proximal $^{991}\text{GFFKR}^{995}$ interacts with $\beta 3 \text{ K}^{716}$ within the hydrophobic core forming a clasp that keeps receptor inactive. Upon intracellular activation, the breakage of the intracellular clasp together with the extension of the α Ib extracellular domains pulls the transmembrane upwards causing four more consecutive residues lower down the membrane proximal region to repartition into the lipid bilayer bringing the negatively charged stretch of α Ib $^{1001}\text{EEDEEGE}^{1007}$ stretch close to the lipid bilayer charged head groups [86].

The different solution conditions (micelles [87], acetonitrile and bicelles [83,85]) used in structure calculations yielded two distinct α Ib conformations in the resting state. The *in vivo* study by Kurtz et. al. (2012) seems to favour the α Ib GFF reverse turn structure in bicelles but whether or not spontaneous (without cysteine mutation) lipid transversions of the membrane proximal residues which includes several highly charged residues such as K^{989} , K^{994} and R^{995} in α Ib or K^{716} , D^{723} , R^{724} , K^{725} , and E^{726} in $\beta 3$ can occur without substitution to a polar, uncharged cysteine residue require further investigation. Further, the integrin mode of activation based solely on α Ib $\beta 3$ needs to be correlated to the rest of the 24 integrins.

Figure 1.15 Model of integrin intracellular activation, β -subunit in blue and α -subunit in red. (A) Integrin maintained at resting state by an inhibitor (shaded in red oval) at the membrane proximal clasp. (B) Activated cellular signal dissociates the inhibitor leading to the intermediate state of the transmembrane/cytoplasmic heterodimer. (C) The cytoplasm-exposed hydrophobic residues (in cyan) can now insert or anchor into the membrane. (D) Integrin regulators such as talin (in green) further perturb the clasp, leading to the dissociation of the clasp and subsequent unwinding of the transmembrane coiled-coil. Taken from Yang, J., et al., *Structure of an integrin α IIb β 3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(42): p. 17729-34.

Indeed, the helical propensity of the cytoplasmic GFFKR stretch has been demonstrated in the NMR structural characterization of the cytoplasmic domains (Table 1.2).

Table 1.2 Structural studies of the integrin cytoplasmic region		
Year/Author	Region	Method
(2009) Yang et al. <i>PNAS</i> .	$\alpha_{IIb}\beta_3$ full transmembrane and cytoplasmic region* α_{IIb} (residues E ⁹⁶⁰ -E ¹⁰⁰⁸) β_3 (residues K ⁶⁸⁹ -T ⁷⁶²)	Solution state NMR heterocomplex structure determination of the clasped cytoplasmic conformation in deuterated acetonitrile
(2009) Zhu et al. <i>Molecular cell</i> .	$\alpha_{IIb}\beta_3$ intact receptor*	Rosetta modelling based on crystal structure obtained from Zhu et al.(2008) and experimental disulfide scanning restraints
(2008) Lau et al. <i>EMBO</i> .	$\alpha_{IIb}\beta_3$ transmembrane and cytoplasmic region* α_{IIb} (residues A ⁶⁵⁸ -P ⁹⁹⁸) β_3 (residues P ⁶⁸⁵ -F ⁷²⁷)	Solution state NMR heterocomplex structure determination of the clasped transmembrane and cytoplasmic conformation in bicelles, in addition, simulated annealing of transmembrane α/β interface from mutational data.
(2008) Lau et al. <i>JBC</i> .	α_{IIb} transmembrane and cytoplasmic region α_{IIb} (residues A ⁶⁵⁸ -P ⁹⁹⁸)	Solution state NMR in bicelles
(2008) Lau et al. <i>Biochemistry</i> .	β_3 transmembrane and cytoplasmic region β_3 (residues P ⁶⁸⁵ -F ⁷²⁷)	Solution state NMR in bicelles
(2008) Bhunia et al. <i>JBC</i> .	α_L and β_2 cytoplasmic region α_L (residues K ¹⁰⁸⁸ -D ¹¹⁴⁵) β_2 (residues K ⁷⁰² -S ⁷⁴⁷)	Solution state NMR
(2004) Vinogradova et al. <i>PNAS</i> .	α_{IIb} and β_3 cytoplasmic region α_{IIb} (residues K ⁹⁸⁹ -E ¹⁰⁰⁸) β_3 (residues K ⁷¹⁶ -D ⁷⁶²)	Solution state NMR in Dodecylphosphocholine micelles
(2002) Vinogradova et al. <i>Cell</i> .	$\alpha_{IIb}\beta_3$ cytoplasmic region* α_{IIb} (residues M ⁹⁸⁷ -R ¹⁰⁰⁸) β_3 (residues K ⁷¹⁶ -D ⁷⁶²)	Solution state NMR in the presence of ¹⁵ N/ ¹³ C labelled partner fused maltose-binding protein (MBP)
(2002) Weljie et al. <i>PNAS</i> .	$\alpha_{IIb}\beta_3$ cytoplasmic region α_{IIb} (residues M ⁹⁸⁷ -R ⁹⁹⁷) β_3 (residues K ⁷¹⁶ -D ⁷⁴⁰)	Solution state NMR
(2000) Vinogradova et al. <i>PNAS</i> .	α_{IIb} cytoplasmic region Myristoylated- α_{IIb} (residues K ⁹⁸⁹ -E ¹⁰⁰⁸)	Solution state NMR in Dodecylphosphocholine micelles
(2000) Hwang et al. <i>JMR</i> .	α_{IIb} cytoplasmic region (part transmembrane) α_{IIb} (residues V ⁹⁸⁴ -E ¹⁰⁰⁸)	Solution state NMR in trifluoroacetic acid (TFA) solvent
The transmembrane and cytoplasmic interface of the integrin molecule is based on the numbering by provided by the individual journals. *Heterocomplex structure is denoted without 'and'		

The structural characterizations of the short cytoplasmic domain particularly of the $\alpha\text{IIb}\beta 3$ integrin showed that without their transmembrane counterparts, they mostly adopt random coil structures in solution. Several strategies have been used to mimic the integrin transmembrane domain in order to elucidate the cytoplasmic domains structures prior to the structural elucidation of the transmembrane-cytoplasmic NMR structures [63,64,85]. The NMR structure of the αIIb V⁹⁸⁴-E1⁰⁰⁸ which consists of part of the transmembrane region and the full length cytoplasmic region was first determined in 45% Trifluorethanol (TFE), a chemical known to stabilize alpha-helical conformations. And indeed the transmembrane residues and part of the cytoplasmic domains consisting of the conserved GFFKR motif assumed a helical structure[88]. Subsequent structural characterization which have used full transmembrane and cytoplasmic sequences in DPC micelles detected homooligomerization between individual αIIb and $\beta 3$ subunits which occluded structural characterization efforts[73]. Coiled-coil constructs linking αIIb and $\beta 3$ cytoplasmic domains fused at N-terminus of the coiled-coil with a disulphide bond were used to circumvent homooligomerization tendencies, yet detected random coil structures with little/no interaction between the cytoplasmic tails not even the conserved salt bridge [89].

Strangely, NMR-derived structures of truncated peptides of consisting αIIb M⁹⁸⁷-R⁹⁹⁷ and $\beta 3$ (K⁷¹⁶-D⁷⁴⁰) mixed together in high concentrations and low salt environment detected interactions between αIIb and $\beta 3$ truncated peptides [90]. However, two distinct structural conformations of the cytoplasmic $\alpha\text{IIb}\beta 3$ complex are detected in the peptide mixture which the authors proposed to be the different intermediate states of interaction. These conformations displayed interactions between αIIb (V⁹⁹⁰, F⁹⁹⁴) and $\beta 3$ (L⁷¹⁷-I⁷²¹), which consists of the cytoplasmic membrane proximal residues but in the absence of the conserved salt bridge (Figure 1.16A). N-terminal myristoylation strategy has also been used to determine αIIb cytoplasmic domain structures in DPC micelles [91]. The myristate group mimics the transmembrane domain and stabilized the alpha-helical nature of membrane proximal residues. In addition to the structural characterization,

functional effects of membrane distal residues in the myristoylated- α Ib cytoplasmic tail are determined by mutating P⁹⁹⁸/P⁹⁹⁹ to Ala. The myristoylated wild-type and myristoylated α Ib mutant displayed distinct structural conformations at the C-terminal half of the cytoplasmic region. Both assumed a helical conformation at the membrane proximal regions of the cytoplasmic tails, however, the wild-type myristoylated peptide assumed a more compact structure with its membrane distal regions folding back to the helical stretch (Figure 1.16B). The functional consequence of the mutation is examined by adding these myristoylated peptides to platelets; and the effects of these peptides on the fibrinogen binding efficiency of the intact α Ib β 3 is observed [91]. The hydrophobic acyl chains of the myristate group enables penetration of the peptides through the cellular membrane and can thus effect changes in the intact α Ib β 3 cytoplasmic region. Only the myristoylated wild-type α Ib peptide inhibited fibrinogen binding by intact α Ib β 3 on platelets [91]. The mutation of the membrane distal residues resulted in a different conformation that could affect the cytoplasmic interactions of the intact α Ib β 3, this activating conformation of the P^{998/999}A mutant is proposed to be the ‘open’ conformation for α Ib cytoplasmic tail [91].

Interactions between the α - and β - cytoplasmic tails are inherently weak and favoured by certain conformations. The conflicting finding of null interaction [73,89] and interacting cytoplasmic regions [90] could be due to unsuitable N-terminal modifications that result in conformations that do not favour heterodimeric α/β cytoplasmic interactions [92]. Hence, DPC micelles that were used to characterize the full length transmembrane and cytoplasmic regions that did not detect interactions between the cytoplasmic regions, was probably due to competition between the DPC micelles which were also found to interact with the membrane proximal residues involved in cytoplasmic interactions [92]. Hence, a further study which determined the interaction interface of the α Ib β 3 cytoplasmic region have determined the structure of free α Ib- K⁹⁸⁹-E¹⁰⁰⁸ and β 3-K⁷¹⁶-T⁷⁶² cytoplasmic domains in the presence of their binding partners β 3 or α Ib

cytoplasmic regions attached to the Maltose binding protein (MBP) via a long linker which does not affect the structural integrity of the cytoplasmic tails or interaction interface[92]. The complex structure obtained is proposed to represent the resting conformation of the cytoplasmic tails as the structure of the free tails was calculated in the presence of their interacting partners (Figure 1.16C). Additionally, titration between the free cytoplasmic domains and MBP-linked cytoplasmic domains were used to characterize the interaction interface of the α IIB β 3 cytoplasmic domains. The α IIB cytoplasmic domain structure concurs with the compact structure obtained using myristoylated α IIB in DPC micelles [91], assuming a helical conformation at the membrane proximal cytoplasmic region and C-terminal loop back. Similarly, the β 3 is also found to be helical at the membrane proximal region from K⁷¹⁶-K⁷³⁸ with the remaining forming a disordered loop with a turn at the membrane distal NPLY region [92]. The interface of interaction consists of a hydrophobic interface formed between α IIB-V⁹⁹⁰. β 3-L⁷¹⁸/I⁷¹⁹ and α IIB-F⁹⁹². β 3-I⁷²¹/H⁷²² followed by an electrostatic interface formed between the conserved salt bridge residues α IIB-R⁹⁹⁵. β 3-D⁷²³/E⁷²⁶ (Figure 1.16C) [92]. This interface was easily abolished by mutation of the salt bridge or addition of a known integrin intracellular activator, talin [92].

Figure 1.16 NMR-derived structures of the α IIb β 3 integrins cytoplasmic domains. (A) Two distinct conformations of the α IIb β 3 cytoplasmic tails (1KUP and 1KUZ). (B) The resting and active conformations of the Myristoylated- α IIb cytoplasmic tails (1DPK and 1DPQ). (C) The conformation of the α IIb β 3 complex (1M8O). (D) Conformation of the α IIb cytoplasmic tail in DPC micelles (1S4W). Structures are downloaded from the Protein Data Bank (PDB) and prepared using PYMOL.

Following the characterization of the integrin resting state conformation, the structures of the dissociated tails in the integrin activated in membrane-mimic conditions were determined. Unmodified cytoplasmic domains of the $\alpha\text{IIb}\beta 3$, $\alpha\text{V}\beta 3$ and $\alpha 5\beta 1$ integrin were characterized in DPC micelles, only $\alpha\text{IIb}\beta 3$ is structurally characterized [87]. The αIIb and $\beta 3$ structures in DPC micelles are found to be helical at the membrane proximal region, and the $\alpha\text{IIb-K}^{989}\text{-F}^{993}$ and most regions of the $\beta 3\text{-K}^{716}\text{-I}^{721}$ make extensive contacts with the hydrophobic methylene groups of DPC indicating this regions are inserted in the membrane (Figure 1.16D) [87]. Membrane proximal residues of the $\alpha\text{V}\beta 3$ and $\alpha 5\beta 1$ are also found to interact with the hydrophobic acyl chains in the DPC micelles [87]. The αIIb and $\beta 3$ cytoplasmic domain structures determined separated in DPC micelles are perceived to be the activated state of the integrin during separation, where the insertion of the membrane proximal residues abrogates the $\alpha\beta$ cytoplasmic interface, leading to separation [87].

Most structural characterizations of the integrin cytoplasmic regions have been based on the platelet integrin, $\alpha\text{IIb}\beta 3$. Apart from the conserved GFFKR motif at the membrane proximal regions and some similar residues across the 18 integrin α -subunits, they generally have variegated sequences. The conformational effects of the cytoplasmic membrane distal residues of the α -subunit have been observed in the Lau, et.al. (2008) [63] and Vinogradova et. al. (2000) structure [87]. As most integrin α -subunits share the same β -partners, structural variation of the α -subunit could possibly be a determining factor in modulating the membrane proximal residues. A recent structural elucidation of the αL cytoplasmic domain revealed a tri-helical fold stabilized with electrostatic interactions [93]. The membrane proximal region containing the conserved $\alpha\text{L-G}^{1090}\text{FFKR}^{1094}$ motif assumes an alpha-helical conformation with the R^{1094} facing out of the compact αL structure. Investigations of the $\alpha\text{L}\beta 2$ cytoplasmic interaction interface concurred with the interaction $\alpha\text{IIb}\beta 3$ cytoplasmic interface [92], involving only membrane proximal residues which includes the conserved salt bridge between $\alpha\text{L-R}^{1094}\beta 2\text{-D}^{709}/\text{E}^{712}$ and hydrophobic packing of the $\alpha\text{L-F}^{1092}\beta 2\text{-I}^{705}$ [93].

All structural characterization of the cytoplasmic domains concurred with the helical structural integrity of α -subunit membrane proximal residues in the formation of the $\alpha\beta$ cytoplasmic clasp, or at least in the case of the $\alpha\text{IIb}\beta 3$ and $\alpha\text{L}\beta 2$ integrins. The membrane proximal residues form right-handed helical conformations in the resting state [92,93] instead of a left-handed loop reported in the bicelles $\alpha\text{IIb}\beta 3$ transmembrane-cytoplasmic domain structure [83]. Transitions of the membrane-proximal cytoplasmic residues into the hydrophobic milieu of the bilayer are also possible, and perhaps not captured in the single representation of averaged conformations derived from NMR studies. The energetic possibility of the conserved GFF excursions into the hydrophobic acyl chains of the bilayer needs to be further investigated in other integrins, especially in the finding of the tri-helical fold αL cytoplasmic domain. Whether the reverse-turn into the bilayer occurs spontaneously or needs to be energetically sponsored by intracellular proteins still needs to be further ascertained, which will provide clearer mechanistic insights into the cytoplasmic control of integrin activation. The importance of understanding integrin activation mechanism is highlighted by the numerous and systemic implications to human health brought about by aberrant integrin activation or inactivity.

1.18 Regulation of integrin affinity and avidity through inside-out and outside-in processes

Integrins are undoubtedly regulated by its changes in conformation which leads to its activation as shown by structural evidences but the global perspective of integrins on a single cell requires another dimension of understanding. Integrins are dynamically regulated in the cell, especially in leukocytes that undergo drastic cell shape changes to travel to their target sites. Different types of integrins are present on a single cell, most often quiescent until activated and only specific subsets of integrins are engaged for specific cellular behaviours. Aside from the monomeric changes in affinity coupled to changes in integrin conformation; overall integrin affinity to receptor, termed avidity or valency, can be further enhanced through clustering processes that are receptor-binding dependent or

independent (Figure 1.17) [94]. Biased cellular re-localization of activated integrins through clustering processes is crucial in creating cellular polarity requisite in cellular differentiation, adhesion and migration. Integrins clustering can be receptor-binding dependent, whereby they relocate following the removal of cytoskeletal restraints that constrain them to certain parts of the cellular membrane, and by diffusing and binding to their extracellular receptors; it can also be due to the reshaping of membrane during cell spreading that can enhance complementarity between integrin and receptor-bearing surfaces. Contrastingly, in receptor-binding independent reorganization, integrins are relocated via intracellular processes that do not require them to be in contact with their extracellular receptors, such as vesicular trafficking of integrins to the cellular membrane, integrin recruitment through lipid rafts, or homotypic associations between transmembrane domains of neighbouring integrins upon activation (Figure 1.17) [94].

Figure 1.17 Valency and affinity modulations of integrins. By clustering of the integrins the avidity becomes high enough for functional adhesion (left). An intermediate affinity may be achieved by straightening out of the integrin, but high affinity needs opening of the binding site (right). It should be pointed out that it is not known for sure that the integrins need to straighten out in order to exhibit high affinity. Taken from *Gahmberg, C.G., et al., Regulation of integrin activity and signalling. Biochimica et biophysica acta, 2009. 1790(6): p. 431-44.*

Regulation of integrin affinity and avidity is especially important to circulating blood cells in the enclosed and highly pressured vasculature, where leukocytes and platelets can rapidly change their quiescent, non-interacting status to shape-morphing cells capable of adhesion and migration. A plethora of integrins are involved in this process: In leukocytes, the $\alpha 4\beta 1$ and $\beta 2$ integrins are involved

in their migratory processes, slow rolling, adhesion strengthening and trans-endothelial migration; and in platelets, the $\alpha\text{IIb}\beta 3$ integrin is involved in the rapid change of platelet adhesiveness during blood clot formation. Changes in overall cellular adhesiveness due to changes in integrin affinity and avidity are precisely regulated by inside-out and outside-in signaling processes [17,24]. Priming or inside-out signaling are events that precede or occur independently of receptor binding serve to enhance receptor binding efficiency of the integrin molecule. Receptor-induced activation or outside-in signaling propagates intracellular signals that result from integrin conformational changes and affinity, which will lead to receptor binding. Inside-out signaling pathways are known to eventually lead to integrin clustering which will transit into the outside-in signaling processes. The activation of outside-in signaling pathways leads to the formation of highly organized focal adhesomes that contribute to adhesion strengthening required during leukocyte migratory processes, such as cell adherence, cell spreading, and other aspects of cell migration [17,24]. The engagement of intracellular proteins that remodel the actin cytoskeleton may also exerts a lateral force on the β -subunit cytoplasmic tail that can be allosterically transmitted to the extracellular domain thereby changing the integrin activation status (Figure 1.18) [95,96].

Figure 1.18 Regulation of integrin activation by lateral force. (A) In the bent conformation, integrins have low affinity for ligand. (B) At sites where actin filaments are formed lateral translocation on the cell surface and buffeting cause integrin extension. Both open and closed headpiece conformations are putatively present. (C) Binding to an immobilized extracellular receptor greatly increases the lateral force and markedly favors the high-affinity, open headpiece conformation. (D) Disassembly of the actin cytoskeleton removes the lateral force. Tensile force between the ligand and the integrin cytoplasmic domains favors the closed headpiece conformation and ligand dissociation. (E) Extracellular receptor dissociates, further favoring the closed headpiece conformation. (F) In the absence of ligand and tensile force, the bent conformation is favored, completing the cycle, and the integrin returns to the same state as shown in (A). Taken from Zhu, J., et al., *Closed headpiece of integrin $\alpha\text{IIb}\beta 3$ and its complex with an $\alpha\text{IIb}\beta 3$ -specific antagonist that does not induce opening*. *Blood*, 2010. 116(23): p. 5050-9.

Yet, despite them being two related processes, inside-out and outside-in signaling processes are two distinct processes that are activated and regulated by two distinct sets of proteins [24]. Figure 1.19 depicts both the extracellular and intracellular proteins known to be involved in the inside-out and outside-in signaling pathways. The inside-out signaling pathway in circulating leukocytes is activated by the release of chemokines or cytokines into the blood vasculature or exposure of immobilized chemokines during an immunological event. Interestingly, amongst the leukocytes, inside-out activation of integrins is cell-type specific and is activated through different chemokines and thus receptors found on the leukocytes, i.e. chemokine signaling can occur through the G protein coupled receptors (GPCRs) expressed on all leukocytes [97]; and on lymphocytes, chemokines activation occurs through the B-cell receptor/T-cell receptor (BCR/TCR) leading to $\alpha\text{L}\beta 2$ and $\alpha 4$ -dependent adhesion [17]; whereas in T-cells, TCR-CD3 engagement activates the $\alpha\text{L}\beta 2$ integrin [17]. Despite the varied means of integrin activation, all lead to the formation of activation complexes that directly modulate the interaction of the integrin cytoplasmic domains.

The outside-in signaling that follows the inside-out activation of integrin molecules, involve binding of the activated integrins to their extracellular receptors which leads to their clustering. The clustering of integrin molecules results in the recruitment of intracellular proteins to its cytoplasmic domains, which effect the remodeling of the actin cytoskeleton leading to morphological changes in the cell required for migratory processes.

Figure 1.19 Inside-out and outside-in signal transduction pathways of the β_2 integrins. Immobilized chemokine or TCR engagement triggers membrane proximal signalling events that could lead to the activation of PLC γ . In turn activated PLC γ catalyses the formation of DAG and IP $_3$ (inositol 3,4,5-trisphosphate) from PIP $_2$ (phosphatidylinositol 4,5-bisphosphate). IP $_3$ induces an increase in intracellular Ca $^{2+}$ concentration. Ca $^{2+}$ and DAG activate CalDAG GEF1 that catalyses the exchange of GDP for GTP, activating the GTPase Rap1. RIAM, the effector of Rap1, may regulate the ligand-binding affinity of the β_2 integrins via talin. The other effectors of Rap1, namely RAPL and Mst, regulate the translocation of the integrins. Dok1 and filamin A are negative regulators of integrin activation. Kindlin3 is a co-activator, together with talin, of β_2 integrin activation. The mechanism by which kindlin3 serves as a co-activator may involve its association with migfilin, and migfilin disrupts the interaction between filamin A and the β_2 integrin tail. The β_2 integrins modulate gene transcription. When integrin $\alpha_L\beta_2$ is engaged by ligand, JAB-1 dissociates from the β_2 tail, translocates into the nucleus and regulates AP-1 transcriptional responses. Clustering of the β_2 integrins induces SFK activation, leading to ERK1/2 signalling or PKC signalling that can regulate gene transcription. In neutrophils, it has been shown that clustering of the β_2 integrins induces ERK1/2 signalling and PKB activation, which delay the onset of spontaneous apoptosis. β_2 integrin-induced SFK activation also regulates actin reorganization via the Syk/Vav1/3 pathway. Other proteins that are reported to directly associate with the integrin β_2 tail are RACK1, α -actinin, cytohesin 1, radixin, 14-3-3 ζ and Syk. Cytohesin 1 and 14-3-3 ζ can regulate actin reorganization via the RhoGTPases Rac, Cdc42 and Rho. The cytoplasmic tails of the β_2 integrins are phosphorylated at specific Ser and Thr residues by PKCs. For many of these phosphorylation sites, their functions remain to be determined. The best characterized to date is Thr758 of the β_2 tail. Phosphorylation of β_2 Thr758 leads to high-affinity binding with 14-3-3 ζ but not filamin A. This phospho-switch may be one of the mechanisms that regulate the activation status of the β_2 integrins. Red circle indicates that the protein interacts directly with the integrin β_2 tail. Red circle beside RAPL indicates direct interaction with the integrin α_L tail. It should be noted that cytoplasmic proteins contributing to integrin inside-out signalling could also be involved in outside-in signalling. Taken from Tan, S.M., *The leukocyte beta2 (CD18) integrins: the structure, functional regulation and signalling properties*. Bioscience reports, 2012. 32(3): p. 241-69.

1.19 Bi-directional convergence of outside-in and inside-out signaling on the integrin cytoplasmic tails

The inside-out and outside-in signaling are two discrete processes and yet only the integrin cytoplasmic domains are the ultimate receiver and transmitter in the integrin bidirectional signaling. The docking of intracellular proteins regulates the integrin global conformational changes and hence affinity to its extracellular ligands. Consequently, binding to their extracellular ligands precisely regulate its anchorage to specific cytoskeletal proteins. Inside-out and outside-in integrin activation might also possibly involve distinct structural changes in the integrin molecule, as disulphide constraints at the integrins transmembrane regions that prohibit cytoplasmic tail separation inhibits cytokine-activated inside-out integrin activation but not outside-in integrin activation by Mn^{2+} ions [98]. Hence, precise regulation of the integrin cytoplasmic region to distinguish between the two inter-related signalling processes is obligatory. The molecular basis for inside-out signalling has been relatively well examined compared to the outside-in signalling pathways due to the difficulty in generating and measuring outside-in integrin signals in the absence of inside-out signals required for integrin activation [99]. Generally, measurement of inside-out signals involves activating intact integrins with cytokines and analyzing the extracellular receptor binding efficiency; conversely, outside-in signalling is examined by the cellular spreading in absence of extracellular stimulus by cytokines [99].

The integrins cytoplasmic domains lack enzymatic activity and rely on its sequence motifs to recruit intracellular proteins that specifically modulate integrin structure and valency changes. It has been demonstrated that distinct sequences in the integrin cytoplasmic domains recruit inside-out and outside-in effectors. Mutation of specific residues in the $\beta 3$ cytoplasmic region can favour either inside-out or outside-in effectors, where mutation of Y⁷⁴⁷ abolished inside-out signalling but not outside-in signal transduction [99]. Besides sequence recognition, phosphorylation of Ser or Thr present in the cytoplasmic domains provides another dimension of complexity in integrin intracellular ligand recognition [100].

Additionally, these intracellular effectors may also recognize overlapping sequences in the integrin cytoplasmic domains which are scrupulously spatially and temporally regulated to prevent spurious ligand recognition [100].

1.19.1 Sequence similarity in the cytoplasmic tails postulated sites of integrin clasps and postulated sites of ligand binding

Sequence alignment of the integrin cytoplasmic domains show that they are highly conserved in the membrane proximal region (Figure 1.20). The membrane proximal regions consisting of the α -(KxGFFKR) and β -(β -(KLLxxxHDDRRE)) chains play an important role in maintaining the resting state integrin, although structural effects in the membrane distal regions have been shown to affect the interaction at the membrane proximal regions [101].

Figure 1.20 The cytoplasmic sequences of the human α and β integrin chains (β 4 is not shown). The potential phosphorylation sites are marked in red. The established phosphorylation sites are numbered. The conserved membrane proximal sequences in the α chain are marked in brown. The functionally important NPXY(F) sequences in the β chain are marked in green and the threonine containing important phosphorylation sites in magenta. Taken from *Gahmberg, C.G., et al., Regulation of integrin activity and signalling. Biochimica et biophysica acta, 2009. 1790(6): p. 431-44.*

These cytoplasmic regions contain binding sites for intracellular signaling proteins of the integrin inside-out and outside-in signalling transduction cascade (Figure 1.21). Interestingly, the different integrins share the same ligand that can be recruited to highly conserved sequence determinants in the cytoplasmic domains, such as talin which is a well known activator of β_1 , β_2 and β_3 integrins and bind to the NPxY motif conserved in these three integrins [102]. However, certain differences in binding efficiency do exist due to subtle differences in the conserved and encompassing sequences [101]. There are also proteins that covets the same sequences and compete with each other for the same binding site, such as filamin and talin where filamin is a negative regulator of talin activation [103].

Figure 1.21 Binding sites on integrin α and β subunits for integrin cytoplasmic domain binding proteins. The reported interacting sites on integrin tails for integrin-tail-binding proteins are indicated by boxes and arrows. Conserved amino acid sequences among integrin α and β cytoplasmic domains are illustrated: uppercase letters correspond to near-invariant residues; lowercase letters correspond to residues conserved in at least three subunits; O represents conserved hydroxylated residues; π and ϕ represent conserved apolar and polar residues, respectively; and dashes represent unconserved residues and/or gaps. Taken from Liu, S., et al., *Integrin cytoplasmic domain-binding proteins. Journal of cell science*, 2000. 113 (Pt 20): p. 3563-71.

he intracellular ligands that bind to the cytoplasmic domains can be separated into proteins that have binding sites in the α - or β - chains. There are, however, examples that bind to both the membrane proximal and distal regions, such as talin, which is recruited via the membrane distal NPxY motif and disrupts the conserved salt bridge at the membrane proximal regions [104]. Intracellular ligands that bind to the α -chain includes RapL, which is a inside-out signalling effector that recognizes two Lys K^{1097/1099} that occur after the α L-GFFKR motif[105]; CIB, which recognizes the α IIb-membrane proximal residues [106]; paxillin which recognizes α 4-E⁹⁸³-Y⁹⁹¹ whereby mutation to Ala abrogates interaction[18]; and several examples described in Figure 1.21. Proteins that interact with the β - cytoplasmic domains mostly consists of a FERM (4.1, Erzin, Radixin, Moesin) domain which interacts with the β -NPxY motif. These proteins includes, talins, Kindlins, Filamins, Radixin, Myosin X, ICAP-1A, Dok-1 and Numb [102]. Other β -cytoplasmic binding proteins that recognize other regions include Cytohesin I which recognizes the β 2-W⁷²³-D⁷³¹ motif, α -actinin that recognizes the β 2 membrane proximal regions and several others described in Figure 1.21.

1.19.2 Phosphorylation as another mechanism of cytoplasmic control of activation

Phosphorylation of the α - and β - cytoplasmic chains also play important roles in the regulated recruitment of proteins in integrin inside-out and outside-in signalling. Ser, Thr, and Tyr present in the α - and β - cytoplasmic domains are possible sites for phosphorylation to occur [107]. Figure 1.20 obtained from the review by Gahmberg et. al. (2009), displays the alignment of α - and β - chain cytoplasmic sequences and the possible sites of phosphorylation is shown in red colour. Phosphorylation could possibly occur differently for each integrin and the exact role it plays in the cell is also not exactly known or thoroughly investigated. Nonetheless, studies on the phosphorylation of some integrin cytoplasmic domains, mostly on the β 2 and β 3 integrins have shown that it recruits a set of intracellular effectors distinct from the sequence motifs. Phosphorylation sites in the α -

cytoplasmic domain seems to occur sporadically at the membrane distal regions and do not seem to form any part of the conserved phosphorylation motifs, whereas phosphorylation sites in the β -cytoplasmic domains seems to consistently occur in conserved motifs such as the two NxxY and triple Thr motifs. It has been noted that some of the α -cytoplasmic chains are found to be constitutively phosphorylated such as the α L-S¹¹⁴⁰ and α M-S¹¹²⁶, and plays an important role in inducing extracellular domain changes in the α L β 2 and α M β 2 in their binding to its extracellular receptors, such as ICAM-1. However, constitutive phosphorylation in the β -chain has not been discovered, and most likely involves intracellular effectors that recognize the conserved phosphorylation motifs [100]. External inducers which are found to propagate β -cytoplasmic chain phosphorylation includes phorbol esters, T-cell receptor antibodies, chemokines, etc [108]. It is also possible for different residues to be activated due to different extracellular stimulus providing an additional regulated control at the cytoplasmic domain, i.e. β 2-S⁷⁵⁶ is found to be strongly phosphorylated in phorbol ester treated T-cells, whereas, β 2-Y⁷³⁵ is phosphorylated in Interleukin-2 treated Natural killer cells [109]. Phosphorylation effects in different integrins have also been found to be different. Phosphorylation of the Tyr in the β 3-N⁷⁴⁴PLY and N⁷⁵⁶ITY motif is requisite in inducing integrin signalling in platelet aggregation process but not so in the corresponding β 1-N⁷⁸⁰PIY and N⁷⁹²PKY motifs, where the β 1-Tyr are involved in other interactions than phosphorylation[110].

Following activation, activated Protein Kinase C (PKCs) in the cell will be recruited proximal to the integrin cytoplasmic tails to effect phosphorylation on specific residues. Protein kinases, specifically PKC δ and PKC ϵ , have been found to phosphorylate S⁷⁴⁵ and T⁷⁵⁸ on β 2 cytoplasmic fragments [109]. The permeabilization of these phosphorylated peptides into T-cells are also found to activate Rac-1 (Ras-related C3 botulinum toxin substrate 1) and Cdc42 (Cell division control protein 42), both activated forms are well-known to initiate actin reorganization in cell adhesion [108]. The phosphorylation of these residues has also been found to recruit 14-3-3 proteins that are required for actin rearrangements

in cell spreading and adhesion (outside-in) processes. A recent investigation into the phosphorylated and non-phosphorylated $\beta 2$ -T⁷⁵⁸ have revealed a switching mechanism between talin, filamin and 14-3-3 proteins in the T-cell adhesion process [111]. Talin is found not to be regulated by the phosphorylation of this $\beta 2$ -T⁷⁵⁸ residue which forms the first Thr of the TTT motif of the β -cytoplasmic domain. However, phosphorylation of this residue increases the 14-3-3 binding affinity to the $\beta 2$ cytoplasmic domain which can outcompete and displace the bound talins. Filamin which binds only to the unphosphorylated $\beta 2$ cytoplasmic domains acts as a inhibitor of phosphorylation. Hence the phosphorylation of $\beta 2$ -T⁷⁵⁸, which displaces filamin and promotes the binding of the 14-3-3 which outcompetes talin, mediates the downstream actin reorganization process in outside-in signalling [111]. Subsequent structural characterization, using X-ray crystallography, have elucidated the different conformations of the cytoplasmic $\beta 2$ peptide with and without T⁷⁵⁸ phosphorylation in filamin and 14-3-3 domains respectively (Figure 1.22A) [111]. A similar mode of phosphorylation switch in the $\beta 3$ integrins have also been elucidated but involving different intracellular effectors and phosphorylation motif. Instead of the first Thr of the TTT motif, the Tyr of the $\beta 3$ N⁷⁴⁴PxY motif is involved in the switch between the integrin cytoplasmic activator talin and negative regulator Dok-1 (Docking Protein 1) [112]. Talin is recruited to the un-phosphorylated $\beta 3$ N⁷⁴⁴PxY and the subsequent phosphorylation of this Y⁷⁴⁷ results in the increased binding efficiency of Dok-1 and displacement of talin [112]. Structural conformations of phosphorylated $\beta 3$ -Y⁷⁴⁷ interaction with the Dok-1 proteins have also been elucidated using a combination of X-ray crystallography and subsequent modelling of $\beta 3$ peptide-Dok-1 complex using NMR titration studies (Figure 1.22B) [112]. A recent structural characterization, using NMR, of phosphorylated $\beta 3$ chains gave further structural and mechanistic insights into the diverse recruitment of intracellular effectors mediated through phosphorylation at multiple sites [113]. Shc (Src homology 2 domain) plays a key role in MAPK (Mitogen-activated Protein Kinases) signalling, and it is found to interact directly with the two Tyr phosphorylated $\beta 3$ -N⁷⁴⁴PLY and N⁷⁵⁶ITY motifs (Figure 1.22C) [113].

Figure 1.22 Conformations of the phosphorylated β -subunit cytoplasmic domains in complex with their ligands, phosphorylation are circled in orange. (A) Conformation of the phosphorylated $\beta 2$ membrane distal peptide phosphorylated at T⁷⁵⁸ in complex with 14-3-3 protein. Taken from Takala, H., et al., *Beta2 integrin phosphorylation on Thr758 acts as a molecular switch to regulate 14-3-3 and filamin binding*. *Blood*, 2008. 112(5): p. 1853-62. (B) Conformation of the phosphorylated $\beta 3$ membrane distal peptides phosphorylated at Y⁷⁴⁷ in complex with Dok-1. Taken from Oxley, C.L., et al., *An integrin phosphorylation switch: the effect of beta3 integrin tail phosphorylation on Dok1 and talin binding*. *The Journal of biological chemistry*, 2008. 283(9): p. 5420-6. (C) Conformation of the $\beta 3$ membrane distal residues phosphorylated at two sites Y⁷⁴⁷ and Y⁷⁵⁹ in complex with the Shc protein. Taken from Deshmukh, L., V. Gorbatyuk, and O. Vinogradova, *Integrin [114]3 phosphorylation dictates its complex with the Shc phosphotyrosine-binding (PTB) domain*. *The Journal of biological chemistry*, 2010. 285(45): p. 34875-84.

Phosphorylation switch for the α -cytoplasmic domain also plays important roles in leukocyte adhesion and migratory processes. The α -cytoplasmic chain E⁹⁸³-Y⁹⁹¹ region facilitates binding to the intracellular protein, paxillin [115]. Interestingly, phosphorylation of α 4-S⁹⁸⁸ which occurs outside of the delineated paxillin interaction region was found to abrogate interaction [116]. Reversible phosphorylation of the α 4-S⁹⁸⁸ plays an important role in the leukocyte migration [117]. At the leading edge of the migrating leukocyte α 4-S⁹⁸⁸ is phosphorylated and paxillin is unbound, whereas unphosphorylated α 4-S⁹⁸⁸ at the tailing edge interacts with paxillin [117]. Disruption of this reversibility in α 4-S⁹⁸⁸ phosphorylation by mutation to a unphosphorylatable mutant α 4-S⁹⁸⁸A has been found to interfere mononuclear leukocyte recruitment to sites of inflammation [118].

1.19.3 Mechanism of integrin activation by talin, a well studied inside-out activator of integrins

Talins (talin-1 and talin-2) are well characterized intracellular integrin activators involved in the inside-out signalling transduction [102]. Talin is ~270 kDa consisting of an N-terminal ~47kDa head and ~220 kDa flexible rod domain. The talin head domain consists of a FERM domain composed of 3 subdomains (F1, F2 and F3) and an F0 domain, and the talin rod domain is composed of a series of helical bundles that contain multiple binding sites for the F actin-binding protein, vinculin and a second integrin-binding site [102]. The C-terminus contains a THATCH (talin/HIP1R/Sla2p actin tethering C-terminal homology) domain that mediates dimerization and provides a linkage between talin and F-actin [102]. The intact talin is inefficient in activating the integrin molecule as the rod domain acts as a restraint to prevent the exposure of the talin head domain. Only the removal of the head-rod restraint during a signalling event initiated by an external stimulus can it expose the head domain for binding to the cytoplasmic β -subunit [102]. The talin FERM/head domain plays an important role in the activation of the integrin molecule. A conserved phosphotyrosine binding (PTB) binding site in the F3 domain is found to interact specifically with the cytoplasmic β -subunit membrane distal motif, NPxY, followed by the binding to the membrane proximal motif [102].

The associated integrin cytoplasmic tails have been shown to be crucial in the maintenance of the integrin resting state. talin, especially its head domain, has been shown in several studies to disrupt the associated integrin cytoplasmic tails to promote activating conformational changes in the entire integrin molecule [102]. FRET between C-terminal α L- and β 2- cytoplasmic tails each attached to cyan and yellow fluorescent proteins is reduced during the overexpression of the talin head domain in intact cells [61]. The interaction of talin and the β 3-cytoplasmic tails can also be disrupted with the overexpression of β 3 cytoplasmic tails [119]. The addition of the talin head domain to complex mixtures containing α Ib β 3, α V β 3, or α 5 β 1 cytoplasmic peptides has been observed to abolish intermolecular NOEs between the α - and β -cytoplasmic tails [87]. A recent study on the intact α Ib β 3 integrin embedded on lipid membranes has further unequivocally demonstrated the activating properties of the talin head domain on the α Ib β 3 integrin global conformational changes. Intact α Ib β 3 integrins are individually synthesized on lipid membrane nanodiscs to mimic physiological conditions, where bent conformations are detected by cryo-EM; and following the addition of the talin head domain, extended conformations are detected [120]. Additionally, the talin head domain is also found has a higher binding affinity for the β 2 cytoplasmic tail compared to the interaction between α L- and β 2 cytoplasmic domains, thus reinforcing its role in integrin activation through the disruption of cytoplasmic interactions [93]. The talin rod domain that conceals the activating talin head domain also plays an important role in the downstream processes following integrin activation and has been found to be required in the formation of focal adhesions [102].

Interfaces between the integrin cytoplasmic tail regions involved in its interaction with talin, characterized by X-ray crystallography and NMR studies, have been shown to comprise of membrane proximal residues on both α - and β -chains and the β -chain first NPxY motif at the membrane distal region [121]. The corresponding binding site for the integrin cytoplasmic tails in the talin, is at the F3 domain the N-terminal talin head domain, although this might be different for different integrins due to inherent differences in sequence among the 24 integrins, i.e., the F3 in the head domain has been found to be sufficient

to activate $\beta 3$ integrin but not in the $\beta 1$ integrins which require the other F0/F1/F2 domains [122]. The F3 contains a Phospho-Tyrosine Binding (PTB) domain that interacts with the $\beta 3$ -NPxY motif. Structural elucidation, using X-ray crystallography, of chimeras containing the $\beta 3$ cytoplasmic membrane distal regions and the talin-F2/F3 (PTB) has demonstrated that the $\beta 3$ -NPxY and talin-F3 interaction is mediated by acidic and hydrophobic interactions [121], and the addition of a negatively charged moiety through Tyr phosphorylation of the $\beta 3$ -NPxY motif can abrogate this interaction [112].

The talin head domain not only interacts with the $\beta 3$ -NPxY motif, it is also found to disrupt membrane proximal residues of β - cytoplasmic tails [87]. Structural characterization using NMR and mutagenesis studies have pinpointed the flexible loop of the talin-F3 domain to be involved in its interaction with the $\beta 3$ membrane proximal cytoplasmic region [104]. This unique loop, not present in other PTB proteins such as the Dok-1, Shc, Numb, etc, forms a hydrophobic pocket that accepts the $\beta 3$ membrane proximal Phe, F⁷²⁷ and F⁷³⁰ [104]. Mutagenesis on the α IIB $\beta 3$ membrane proximal (α IIB-R⁹⁹⁵ and $\beta 3$ -F^{727/730}) and distal residues ($\beta 3$ -Y⁷⁴⁷) have showed that the membrane proximal $\beta 3$ -F^{727/730} residues play an important role in integrin activation. The interaction between talin-F3 and $\beta 3$ -F^{727/730} membrane proximal residues disrupts the putative salt bridge α IIB-R⁹⁹⁵. $\beta 3$ -D⁷²³ and seems to place the talin-F3 close to the plasma membrane [104]. This interaction of the talin-F3 with the charged head groups of the plasma membrane serves to shorten the $\beta 3$ transmembrane helix and bring it further away from the α IIB transmembrane helix [104]. In short, the β -NPxY membrane distal motif serves to form the first strong linkage with the activated talin-F3 domain, which puts it in close proximity with the membrane proximal residues which promotes separation of the integrin α - and β - transmembrane and cytoplasmic regions (Figure 1.23) [104].

Figure 1.23 Model of talin-induced integrin activation. (A) The talin F3 domain (surface representation; colored by charge), freed from its autoinhibitory interactions in the full-length protein, becomes available for binding to the integrin. (B) F3 engages the membrane distal part of the β 3-integrin tail (in red), which becomes ordered, but the α - β integrin interactions that hold the integrin in the low-affinity conformation remain intact. (C) In a subsequent step, F3 engages the membrane proximal portion of the β 3 tail while maintaining its membrane distal interactions. Consequences of this additional interaction are (1) destabilization of the putative integrin salt bridge; (2) stabilization of the helical structure of the membrane proximal region; and (3) electrostatic interactions between F3 and the acidic lipid head groups. The net result is a change in the position of the transmembrane helix, which is continuous with the membrane proximal- β -tail helix. This position change causes a packing mismatch with the α IIb-transmembrane helix, separation or reorientation of the integrin tails, and activation. Mutants of F3 that have compromised interactions with the membrane proximal region and other PTB domains that lack a membrane proximal-binding site stall at point B, consistent with their dominant-negative behavior. Taken from Wegener, K.L., et al., *Structural basis of integrin activation by talin*. *Cell*, 2007. 128(1): p. 171-82.

Talin-integrin interactions are inherently weak hence the requirement to generate talin-cytoplasmic tail chimeras to artificially increase the local concentration in structure determinations [104,121]. The use of chimeras, though have provided insights into the talin F3/ β 3-Tail interactions, have precluded the use of full length cytoplasmic sequences and restricted talin interaction to either the cytoplasmic membrane proximal or distal regions. Moreover, meaningful interactions brought about by the other domains, F0/F1/F2, cannot be thoroughly investigated. The importance in the F0/F1/F2 domains in integrin activation of the integrins, other than the β 3 integrins, has been reported in several studies [122]. Interactions between the F0/F1/F2 and the plasma membrane have been suggested to play an important factor in integrin activation. Additionally, sequence variation amongst the β -

cytoplasmic tails could mean that talins interact differently with different integrins. Indeed, not all other integrin β -cytoplasmic tails have the same binding affinity as the $\beta 3$ with talin [123], interaction between the talin isoform 2 and $\beta 1$ slice variant $\beta 1D$ was found to have the highest binding affinity, thus allowing for the structural characterization of the full length $\beta 1D$ cytoplasmic tail with talin F2/F3 using X-ray crystallography. The structure precisely revealed the β -cytoplasmic and talin F3 residues involved in their interaction and is consistent with previous structural characterizations using chimeras and a different β - subunit (Figure 1.24) [124]. The talin F3-K³²⁷ is found to form a salt bridge with the $\beta 1$ -D⁷⁵⁹ which corresponds to the canonical cytoplasmic salt bridge position (Figure 1.24A). The talin2 F2 domain is connected to the F3 domain in tandem via flexible linker which brings this domain close to the plasma membrane. A positively charged patch, known as the membrane orientation patch (MOP), consisting of several Lys in the F2 have been identified and is found to interact with the cell membrane (Figure 1.24A) [124]. Talin F2 interaction with the cell membrane is an important factor in the $\beta 3$ integrin activation, and changes in the composition of the phospholipid bilayer have been observed to affect the F2 binding affinity. The positioning of the F2 MOP due its interaction with the plasma membrane tilts the β - cytoplasmic domain leading to the separation of the transmembrane helices (Figure 1.24B). This mode of integrin activation is in support of the separation model for integrin transmembrane and cytoplasmic domains and adds on to the mechanistic insights into the integrin activation models postulated by Lau et. al. (2009) [83] and Yang et. al. (2009) [85].

Figure 1.24 The talin isoform 2 (talin2)/ β 1D structure. (A) One heterodimer from the crystal structure of talin2 F2–F3 bound to the β 1D integrin tail. Labelling is for talin2/ β 1D with talin1/ β 3 numbering in parentheses. Highlighted residues interact with the membrane or form a key integrin/talin salt bridge. (B) The talin2/ β 1D structure was merged with the β 3 transmembrane segment from Lau et. al. (2008) and aligned to the calculated membrane tilt angle of 25° . The electrostatic potential is mapped on talin, illustrating the juxtaposition of several positively charged residues (MOP) next to the membrane surface. Taken from Anthi, N.J., et al., *Structural diversity in integrin/talin interactions*. *Structure*, 2010. 18(12): p. 1654-66.

As mentioned, not only the talin-F2/F3 is important in integrin activation, the talin F1 domain is found to consist of a large flexible loop which interacts with the charged head groups of the plasma membrane [125]. This flexible loop is proposed to act as ‘fly-cast’ that latches onto the plasma membrane, and upon encountering the membrane, the loop will fold into a helical conformation with the side of helix containing basic residues interaction with the negatively-charged phospholipids [125]. The folding into a helical structure is suggested to draw the F1 domain closer to the membrane to facilitate subsequent interactions of the F2 and F3 domain with the membrane [125]. The F0 domain which is also required for efficient β 1- and β 3- integrin activation is fully exposed in the activated talin head is hypothesized to interact with other intracellular proteins involved in the integrin activation cascade [122].

Many other intracellular proteins that interact with the integrin β -cytoplasmic tails contains the FERM (F1/F2/F3) domain, but unlike talins, they

do not have the same ability to activate integrins. It was previously thought that all FERM domain-containing proteins are compact structures with their domains arranged in a cloverleaf manner (Figure 1.25), however, a recent crystal structure of the full talin head domain consisting of the F0/F1/F2/F3 revealed a linear arrangement for the domains (Figure 1.25) [126]. Despite the unique conformation from its contemporary FERM containing-proteins, the individual proteins folds in the talin F1/F2/F3 domains remains structurally conserved. The differences lies the sequences of the linker regions between each domains, the F3 domain and the presence of two unique flexible loops in the F1 and F3 domain. This linear arrangement facilitates optimal engagement of the F2 and F3 domains following the ‘flycast’ latch of the F1 flexible loop onto the plasma membrane. Subsequent interaction of the F3 domain with the β -cytoplasmic membrane distal NPxY motif allows the flexible loop in F3 to engage into the β -cytoplasmic membrane proximal region [126]. Hence, conferring talin the unique ability to activate integrins but not the other FERM containing proteins, such as Dok-1, Filamin, etc; and whether or not the talin FERM domain interactions with the rest of the 8 β -integrins and their isoforms still require further investigations [123].

Figure 1.25 Representations of the talin FERM domain. (Leftmost panel) Conventional representation of the talin FERM (F1/F2/F3), in orange, arranged in a clover-leaf manner. Taken from Moser, M., et al., *The tail of integrins, talin, and kindlins. Science*, 2009. 324(5929): p. 895-9. (Rightmost panel) Recent structural elucidation of the the talin FERM domain. Taken from Elliott, P.R., et al., *The Structure of the talin head reveals a novel extended conformation of the FERM domain. Structure*, 2010. 18(10): p. 1289-99.

1.20 Objective of study: Structural characterization of α -subunit cytoplasmic domains of $\beta 2$ integrins

Integrin cytoplasmic tails lack enzymatic activity, yet sequence conservation and, importantly, variation have enabled them to act as signalling hubs recruiting diverse intracellular effectors in their inside-out and outside-in signalling cascade. Undeniably, extracellular conformational changes are also important in the investigation of integrin activation as a whole. Although there are still controversies surrounding the changes occurring in the extracellular and transmembrane regions during integrin activation, we believe that the enrichment of the integrin cytoplasmic structures can provide clues into integrin activation mechanisms. Despite being the smallest domain but yet being able to affect the greatest conformational changes by mere modulation of one amino acid residue in the conserved cytoplasmic salt bridge; or engage a different set of intracellular effectors via phosphorylation in its sequence motifs; and sequence variation resulting in diverse differences in binding affinities to similar sets of intracellular proteins underscores its importance in integrin activation. Its small size also allows various avenues of methods to be used and ease of study to address the effects of sequence variation and structural conservation across the 24 diverse sets of integrin heterodimers. As each integrin mediates its own unique function in a tissue specific manner, structural differences are bound to exist, and whether proposed activation mechanisms based on the $\alpha \text{IIb}\beta 3$ integrin could apply across the 24 integrins remains an open question.

Currently, the 18 α -subunit cytoplasmic domains are structurally scarce with the exception of the $\alpha \text{IIb}\beta 3$ integrin which saw the αIIb -cytoplasmic domains being characterized in various constructs and conditions described in section 1.17c. Mechanisms of integrin activation have been largely based on the $\alpha \text{IIb}\beta 3$ integrin, which have included two different conformations of the αIIb -cytoplasmic domain elucidated in two different construct and conditions (Figure 1.26A or B). It is still uncertain whether these two conformations are intermediates of the integrin activation mechanism and how this can be translated in the compact tri-helical fold of the αL -cytoplasmic domain requires

further investigations (Figure 1.26C). The three helical bundles in α L stabilized by electrostatic forces and hydrophobic packing most possibly suggest that spontaneous helix-reverse turn transition is energetically impossible in α L. This process most possibly requires external factors, such as intracellular proteins, to stabilize the transition. These intracellular proteins could have 2 faces, one of which interacts with the membrane proximal region of the cytoplasmic tail at a higher affinity than the forces that hold them together causing them to be disordered, and the other face brings them closer to the hydrophobic core of the bilayer. In addition, the extension of the integrin extracellular domain could also assist in the partitioning of the GFFKR motif into the hydrophobic core. However, these remain as speculations without evidence from further studies.

Figure 1.26 Structures of the α -subunit cytoplasmic domains. (A) Structure of the α IIb β 3 transmembrane/cytoplasmic complex elucidated in bicelles (2K9J). (B) Structure of the α IIb β 3 transmembrane/cytoplasmic complex elucidated in 50% acetonitrile/H₂O (2KNC). (C) Structure of the α L cytoplasmic domain elucidated in solution (1K80). Structures were downloaded from the Protein Data Bank and prepared with PYMOL.

β 2 integrins consisting of the α L β 2 (LFA-1, CD11aCD18), α M β 2 (Mac-1, CR3, CD11bCD18), α X β 2 (p150, 95, CR4, CD11cCD18) and α D β 2 (CD11dCD18) integrins presents a valuable opportunity to study the effect of having a different α -subunit in the context of the same β 2-subunit. Apart from the common β 2 subunit they share, they also mediate functions exclusively on leukocytes. Leukocytes have to undergo shape-shifting changes involving rapid changes in cellular adhesiveness mediated by integrins. As the β 2 integrins each have a different extracellular binding affinity for the same ligand and can selectively recruits certain intracellular ligands to effect their function, they mediate diverse functions on the leukocytes [127]. The sequential activation of different leukocyte integrins enables the white blood cells to assume different

extracellular receptor binding affinities to transform from non-adherent circulating white blood cells, to shape-shifting adherent cells capable of targeted migration against a high pressure circulatory system [3]. Moreover, it is highly conserved with the well studied $\alpha\text{IIb}\beta 3$ integrin. Through the understanding of how different α -subunits perturb their β -subunit partner, we can understand how the β -subunit can be re-modeled to transmit structural changes to the transmembrane and extracellular regions, and the 'grasp' of the α -subunit on the $\beta 2$ -subunit that contributes to the differing tightness of the α - and β - cytoplasmic region 'handshake' in different integrins. Hence, the numerous structural intermediates proposed to exist in the integrin cytoplasmic unclasping process can then be ascertained.

Hence, we intend to elucidate the structures of the cytoplasmic regions of the α -subunits of the $\beta 2$ integrins in particular; and determine the residues involved in the α - and β -subunit cytoplasmic 'handshake'.

1.21 Strategies to study the α -subunit cytoplasmic domains

1.21.1 Use of a N-terminal myristoylation

The short α -cytoplasmic is intrinsically flexible in solution and adopts a random coil structure. Several structural studies of the $\alpha\text{IIb}\beta 3$ cytoplasmic tails have demonstrated that the α -cytoplasmic membrane proximal regions assume a helical conformation, in the presence of its transmembrane domain [85]; when associated with a N-terminal myristoylation [87,91]; or in the presence of its interacting β -partner [92]. Interestingly, in the framework of a long cytoplasmic sequence, such as αL cytoplasmic domain which contains 58 residues, the membrane proximal region also adopts a helical conformation [93]. The α -cytoplasmic membrane distal regions have been observed in most structural studies to adopt a random coil that packs back to the membrane proximal regions in the associated cytoplasmic conformation[91]. An analysis on the membrane proximal sequences of the α -subunits of the $\beta 2$ integrins (Figure 1.27), shows that it contains several helix promoting amino acids, Ala (A), Glu (E), Asp (D), Leu (L), Val (V), Met (M), Arg (R), Lys (K), Gln (Q) and Asn (N)

[128]. Pro (P) and Gly (G), which disrupt helices, are found around the start of the membrane distal regions.

```

αL 1087 K V G F F K R N L K E K M E A G R G V P N G I P A E D S E Q L A
αM 1114 K L G F F K R Q Y K D M M S E G G P P G A E P Q
αX 1110 K V G F F K R Q Y K E M M E E A N G Q I A P E N G T Q T P S P P
αD 1108 K L G F F K R H Y K E M L E D K P E D T A T F S G D D F S C V A
αIIb 989 K V G F F K R N R P P L E E D D E E G E
α4 1001 K A G F F K R Q Y K S I L Q E E N R R D S W S Y I N S K S N D D

```

Figure 1.27 Sequence alignment of the leukocyte-specific α -subunit membrane proximal cytoplasmic domains, in blue are amino acids with helix promoting residues and in red are helix disruptors.

N-terminal addition of a 14-carbon acyl chain, myristate, has been used in the structural determination of the α IIb cytoplasmic domain [91]. The myristate moiety is highly hydrophobic and readily targets into the hydrophobic hydrocarbon core of the phospholipid membrane bilayer, thus it acts as an anchor to the cytoplasmic peptides covalently attached to it. N-terminal myristoylation is also a natural occurring cellular modification that is catalyzed by N-myristoyltransferase in eukaryotes and viruses [129]. This attached moiety participates in sub cellular localization by facilitating protein-membrane and protein-protein interactions [129]. Additionally, N-terminal acylation has been reported to stabilize secondary structures in intrinsically dynamic helices and β -sheets to promote proper folding [130]. It is possible that the helical integrin transmembrane domains stabilize the membrane proximal regions to form helical structures. Hence, the short dynamic α -cytoplasmic domains, in the absence of its transmembrane, with propensity to form helical structures exist in a dynamic state, transiting between random coil and helical structures. This property is observed in the diverse integrin cytoplasmic structures obtained in solution [89,90]. The N-terminal acylation can then circumvent the absence of the integrin transmembrane and facilitate structural characterization.

The physiological competence of using a N-terminal acylation in the cellular context has been demonstrated in several studies. N-terminal acylated α IIb integrin cytoplasmic domains containing membrane proximal and distal regions have been found to activate full length α IIb β 3 binding to fibrinogen in platelets. The palmitoylated α IIb membrane proximal sequence KVGFFKR

have been demonstrated to activate $\alpha\text{IIb}\beta 3$ resulting in platelet aggregation [131]. Similarly, addition of the myristoylated αIIb cytoplasmic domain having membrane distal mutations, $\alpha\text{IIb-P}^{998/999}\text{A}$ [91], or palmitoylated αIIb membrane distal region, $\text{R}^{997}\text{PPLEED}$ [132] to platelets is found to activate the fibrinogen binding of its intact $\alpha\text{IIb}\beta 3$. Cell permeable peptides have also been recently considered as a possible therapeutic agent in integrin related malaises [133].

1.21.2 Use of membrane mimic micelles

Dodecylphosphocholine (DPC) consists of a hydrophobic acyl chain of 12 hydrocarbons covalently bonded to a positively charged choline group and negatively charged phosphate group. Approximately 55 DPC molecules are required to form a single micelle in a solution containing DPC at/above its critical micellar concentration of ~ 1.2 mM. DPC micelles provides a reasonable approximation to the mammalian membrane environment and its rotational tumbling times (τ_c) in solution is rapid enough to enable the yield of high resolution NMR spectra [134]. Additionally, a ‘NMR-silent’ perdeuterated option is available to enable for the characterization of the tethered peptide without complications of signals from the protons of the DPC acyl chain [134].

Hence, a N-terminal myristoylation of the α -cytoplasmic domain acting as a transmembrane surrogate will be used to tether the integrin α -chain cytoplasmic domain to the DPC micelles to allow for structural characterization in a near physiological environment.

Chapter II: Structures and interactions of the $\alpha M\beta 2$ integrin

2.1 Overview

The $\alpha M\beta 2$ (Mac-1, CD11bCD18) I-domain containing integrin is found to be primarily expressed on myeloid, natural killer, and $\gamma\delta$ T-cells. It mediates various functions in the immune system, such as phagocytosis; leukocyte adhesion and trans-endothelial migration; immune tolerance; and survival of neutrophils [24]. Its ability to mediate these functionally distinct processes lie in its extracellular ligand binding sites other than its I-domain which allows the $\alpha M\beta 2$ extracellular domain to bind to multifarious ligands that do not share conserved motifs. Extracellular ligands, such as ICAM-1, fibrinogen, Junctional adhesion molecule (JAM)-3, denatured Bovine serum albumin (BSA), neutrophil inhibitory factor from hook worm, microbial lipopolysaccharides (LPS) and zymosan, have been found to interact with the extracellular domains of the $\alpha M\beta 2$ integrins [24]. The involvement of the $\alpha M\beta 2$ integrin in many important processes in the immune system also meant diverse disease complications in the event of a malfunctioning integrin. An association between the abnormality in the chromosome arm that encodes $\alpha M\beta 2$ and an increased risk of the auto-immune disease Systemic lupus erythematosus (SLE) was discovered [135], un-regulated infiltration of phagocytes, due to aberrant $\alpha M\beta 2$, into the central nervous system leads to the development of multiple sclerosis [136], and the over-expression of $\alpha M\beta 2$ on monocytes, associated with diabetic patients, was found promote the formation of atheromatic lesions [6].

The function of the $\alpha M\beta 2$ integrin is mediated by the bidirectional inside-out and outside-in signaling transduction that converges on the recruitment of distinct sets of intracellular ligands onto its cytoplasmic domains. Distinct motifs and modifications in the αM cytoplasmic tails is found to regulate $\alpha M\beta 2$ extracellular binding in its outside-in signaling and recruitment of specific sets of intracellular ligands. Switching of the αM cytoplasmic domain with αL or αX abrogates its ability to recruit the Hemopoietic cell kinase (HCK) [127] and truncated αM cytoplasmic domain or replacement with the αL tail results in impaired protection of transfected K563 cells from

apoptosis [137]. In addition, phosphorylation of the α M cytoplasmic domain has also been found to regulate extracellular ligand binding [109]. Unlike the regulated phosphorylation of β -subunit cytoplasmic tail, the α -subunits are generally found to be constitutively phosphorylated [107]. Conformational rearrangements of the α -subunit extracellular binding site have been found to be correlated to phosphorylation of the α -subunit cytoplasmic tails. The α L β 2 integrin is found to be constitutively phosphorylated at α L-S¹¹⁴⁰, and mutation of the S¹¹⁴⁰ to Alanine affects its ability to induce conformational changes to enable binding to its extracellular ligand ICAM-1 [107]. Similarly, the α M β 2 integrin is found to be constitutively phosphorylated on S¹¹²⁶, mutation of this residue to Alanine resulted in impairment in cellular extravasation [109]. Peculiarly, the inability to phosphorylate α M-S¹¹²⁶ resulted in selective recognition of its extracellular ligands. This α M-S¹¹²⁶A mutation only disrupted its ability to recognize ICAM-1 and ICAM-2 but not iC3b and denatured BSA [109]. These findings underscore the importance of the α M cytoplasmic tail in the regulation of extracellular ligand binding in outside-in signaling. In order to understand the conformational changes associated with α M-S¹¹²⁶ phosphorylation, the structure of two forms of α M cytoplasmic domains (K¹¹¹⁴ – Q¹¹³⁷), and its interaction with β 2 (K⁷⁰² – S⁷⁴⁷) will be determined using NMR and RET analysis. For ease of reference, the cytoplasmic tails of the α M β 2 integrin, consisting of the α M (K¹¹¹⁴ to Q¹¹³⁷) and β 2 (K⁷⁰² to S⁷⁴⁷), will be renumbered from one for ease of reference (Figure 2.1).

(A)

```

αM 1114(1) KLGFFKRQYKDMMS1137(24)EGGPPGAEPQ
αL 1087 KVGFFKRN1144LKEKMEAGRGVPNGIPAEDSEQLASGQEAGDPGCLKPLHEKDSESGGGKD1145
αX 1110 KVGFFKRQYKEMMEEANGQIAPENGQTTPSP1144SEK
αD 1108 KLGFFKRHYKEMLEDKPEDTATFSGDDFSCVAPNVPL1145S
αIIb 989 KVGFFKRNRP1008LEEDDEEGE
α4 1001 KAGFFKRQYKSILQEENRRDSWSYINSK1032SNDD

```

(B)

```

β2 702(1) KAL747(46)IHLSDLREYRRFEKEKLKSQWNND•NPLFKSATTVMNPKFAES
β3 716 KLLITIH762DRKEFAKFEEERARAKWD789TANNPLYKEATSTFTNITYRGT
β1 742 KLLMI789IHDRREFAKFEKEKMNAKWDGTGENPIYKSAVTTVVNPKYEGK

```

Figure 2.1 Sequences of the leukocyte and platelet specific integrin cytoplasmic domains. (A) Cytoplasmic sequences of the integrin α -subunits. (B) Cytoplasmic sequences of the integrin β -subunits.

2.2 Materials and Methods

2.2.1 Synthetic and purified peptides

The crude peptides of the Myristoylated (Myr), Dansylated (Dan) and free integrin cytoplasmic tails in Table 2.1 were purchased from GL Biochem (Shanghai, China) and further purified using Waters reverse phase High Performance Liquid Chromatography system (HPLC) (Massachusetts, USA) connected to a C₁₈ column (300 Å pore size, 5µM particle size). A linear gradient of acetonitrile/water gradient with a flow rate of 2 ml/min was used to elute the peptides, and the major peak fractions were collected and lyophilized into powder form.

αM and β2 tail regions, were respectively cloned into the pET-31b(+) vector (Novagen EMD, San Diego). Singly labeled with ¹⁵N- or doubly labeled ¹⁵N¹³C-β2 was purified based on the purification scheme for Ketosteroid isomerase (KSI) recombinant proteins that was previously used by Bhunia et al. (2009) [93]. However, a modified purification scheme described in the next section was used for the αM tail. Instead of using cyanogen bromide to cleave after methionine to release the peptide from the KSI tag, an Asp-Pro sequence was introduced between the KSI and the αM tail to make use of the Asp-Pro bond vulnerability in 90% formic acid [138]. Table 2.1 shows the list of protein sequences with modifications and additional amino acids that were used in this study.

Synthetic peptides and purified proteins that have been further purified with HPLC are shown in Appendix 1.

<i>Table 2.1 Summary of sequences for synthetic and purified α- and β- tails</i>		
<i>Peptide</i>	<i>Method</i>	<i>Sequence</i>
Myr- α M	Synthetic	Myr- ¹ KLGF ^{<u>FF</u>} KRQYKDMMS <u>EG</u> GPPGAEPQ ²⁴
Myr-P α M	Synthetic	Myr- ¹ KLGF ^{<u>FF</u>} KRQYKDMMS-PO ₄ <u>EG</u> GPPGAEPQ ²⁴
α M	Synthetic	¹ KLGF ^{<u>FF</u>} KRQYKDMMS <u>EG</u> GPPGAEPQ ²⁴
P α M	Synthetic	¹ KLGF ^{<u>FF</u>} KRQYKDMMS-PO ₄ <u>EG</u> GPPGAEPQ ²⁴
β 2	Synthetic	¹ KALIHLSDLREYRRFEKEKLKSQWNNDNPLFKSATTTVI ⁺ NPKFAES ⁴⁶
¹⁵ N α M/ ¹⁵ N- ¹³ C α M	Expressed	<u>P</u> ¹ KLGF ^{<u>FF</u>} KRQYKDMMS <u>EG</u> GPPGAEPQ ²⁴ <u>D</u>
¹⁵ N β 2/ ¹⁵ N- ¹³ C β 2	Expressed	<u>L</u> ¹ KALIHLSDLREYRRFEKEKLKSQWNNDNPLFKSATTTVI ⁺ NPKFAES ⁴⁶ <u>GQGH</u> <u>HHHHH</u>
β 2H5W	Synthetic	¹ KALI <u>W</u> LSDLREYRRFEKEKLKSQ <u>A</u> NNDNPLFKSATTTVI ⁺ NPKFAES ⁴⁶
Myr β 2H5W	Synthetic	Myr- ¹ KALI <u>W</u> LSDLREYRRFEKEKLKSQ <u>A</u> NNDNPLFKSATTTVI ⁺ NPKFAES ⁴⁶
Dan- α M	Synthetic	Dansyl- ¹ KLGF ^{<u>FF</u>} KRQYKDMMS <u>EG</u> GPPGAEPQ ²⁴
Dan-P α M	Synthetic	Dansyl- ¹ KLGF ^{<u>FF</u>} KRQYKDMMS- <u>PO</u> ₄ <u>EG</u> GPPGAEPQ ²⁴

**Sequences underlined are amino acids that have been changed from the original sequence*

**Sequences underlined and bolded are extra amino acids added in the generation of the pET-31b- tail clones*

⁺ Methionine in the endogenous β ₂ sequence has been replaced by isoleucine to facilitate purification by preventing intrasequence cyanogen bromide cleavage after methionine.

2.2.2 Purification and expression of ^{15}N and ^{15}N ^{13}C αM and $\beta 2$ tails

The αM and $\beta 2$ cytoplasmic regions were cloned into the pET-31b(+) vector (Novagen EMD, San Diego) described previously in Bhunia et. al. (2009) [93]. Unlike the $\beta 2$ construct which has a Met cleavage site inserted prior to the integrin cytoplasmic sequence in the KSI recombinant protein, Asp and Pro were introduced for the αM construct as the cleavage site to recover the cytoplasmic peptide from KSI [138].

Both the recombinant pET-31b(+) αM and $\beta 2$ cytoplasmic region with ampicillin resistance were respectively transformed into BL21(DE3) cells. The transformed cells were cultured overnight in Luria-Bertani (LB) broth. Depending on whether unlabeled, singly labeled with ^{15}N - or doubly labeled with $^{15}\text{N}^{13}\text{C}$ - was needed, the overgrown culture was seeded in 1:100 volume ratio in LB for unlabeled proteins or isotope-enriched minimal media (M9), containing ^{15}N ammonium chloride without/with ^{13}C -glucose for labeled proteins. The seeded culture was then placed in a 37 °C shaking incubator at 150 rpm until an optical density at 600 nm ($\text{O.D}_{600\text{nm}}$) of 0.7 was reached. Isopropyl β -D-1-thiogalactopyranoside (IPTG) at 1mM/litre of media was then used to initiate plasmid protein expression. Following which, the induced culture was then left to grow for the next 18 hrs at 25 °C with a shaking speed of 150 rpm. The cells were then collected, via centrifugation at 5000 rpm for 20 mins, and resuspended in 20 mM Tris-Cl, 0.5 M NaCl buffer pH 8.0. The resuspended cells were then lysed via sonication on ice to release the cellular contents that contain the recombinant fusion proteins. As the KSI recombinant proteins is targeted to the inclusion bodies, the cell pellets were collected via centrifugation at 14 000 rpm for 30 mins and re-solubilized using 20 mM Tris-Cl, 0.5 M NaCl urea buffer pH 8.0 to release the KSI recombinant proteins. Unsolubilized cellular debris were then eliminated by centrifuging at 14 000 rpm for 30 mins. The supernatant containing the solubilized KSI recombinant proteins was then extracted using Nickel-NitriloTriAcetic (NTA) acid (QIAGEN) beads by making use of the six histidine tags attached to the N-terminus of the KSI- αX protein or the C-terminus of the KSI- $\beta 2$ protein. The KSI fusion proteins were then eluted with 20 mM Tris-Cl, 0.5 M NaCl and 8 M urea buffer pH 8.0 now containing 0.5 M imidazole. The eluates were then

dialyzed against large volumes water at 4 °C for 2 days to remove the urea, causing the formation of KSI-protein precipitates, which were collected by centrifuging at 5000 rpm for 30 mins.

Both the KSI- α M and KSI- β 2 precipitates were collected and weighed for 2 different methods of cleavage due to the different amino acids introduced prior to cytoplasmic tails. KSI- α M was cleaved with 90 % formic acid at a ratio of 1 mg KSI- α X to 1ml of 90% formic acid, purged with N₂ gas to cleave the Asp-Pro connector between KSI and α M for 22 hrs in the dark. Whereas the KSI- β 2 with a Met connector between KSI and β 2, was first dissolved in 70 % formic acid and 37.5 mg of cyanogen bromide was used to cleave every 1 mg of KSI- β 2 [93]. The reaction was also purged in N₂ gas and left in the dark for the next 22 hrs. For both reactions, formic acid and/or cyanogen bromide were neutralized with sodium hydroxide using a rotary evaporator, leaving a film of precipitate. The precipitate was dissolved in 10 mM sodium phosphate buffer pH 6.5 and further purified using HPLC. The identities of the cleaved peptides were verified with mass spectrometry.

2.2.3 Circular Dichroism (CD) studies

A 0.1 mm pathlength cuvette (Hellma, Müllheim, Germany) was used and circular dichroic measurements were taken by Chirascan CD spectrometer (Applied Photophysics, Leatherhead, UK). CD spectra of 0.2 mM of Myr- α M and Myr- α MS14P in 10 mM sodium phosphate buffer pH 6.5 containing 50 mM DPC; and 0.3 mM of synthetic α M in 10 mM sodium phosphate buffer, pH 6.5 were acquired from 190 -240 nm with slit width of 0.1 nm at 0.1 time steps. Three spectra for the same sample were acquired, averaged and subtracted from the corresponding background buffer contributions using the software Pro-Data Viewer (Applied Photophysics, Leatherhead, UK). The final CD spectra are converted to molar ellipticity and smoothed.

2.2.4 NMR derived structure for the myristoylated α M tails

All NMR spectra were recorded on a Bruker DRX 600-MHz instrument equipped with an actively shielded cryo-probe. 10% Deuterium oxide and 2 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added to all NMR samples. All spectra were directly/indirectly referenced to DSS ¹H.

0.7 mM of HPLC purified peptide powder of; Myr- α M, Myr-P α M and α M were dissolved in 10 mM sodium phosphate buffer pH 5.6 separately. Except for α M, 200 mM deuterated dodecylphosphocholine (DPC) (CIL, Massachusetts, USA) was added to the solution. 2-dimensional proton Total Correlation Spectroscopy (TOCSY) and Nuclear Overhauser Spectroscopy (NOESY) experiments were carried out at 308 K for the myristoylated peptides and 298 K for the free α M. TOCSY mixing times were fixed at 50 ms and NOESY mixing times were maintained at 200 ms for the three peptides. ^{13}C - ^1H Heteronuclear Single Quantum Coherence (HSQC) experiment was also carried out for the three peptides at 100% D_2O to obtain $^{13}\text{C}\alpha$ chemical shifts. Raw NMR data was processed using TOPSPIN 2.1 and analyzed with SPARKY [139].

Backbone dihedral angles were then predicted based on the $\text{H}\alpha$, amide proton and $\text{C}\alpha$ chemical shifts of the assigned NOESY and ^{13}C - ^1H HSQC spectra using the web server PREDITOR [140].

The sequence specific resonance assignment strategy was used to assign the NOESY spectrum [141]. NOE restraints were then derived from NOE intensities of the assigned NOESY peaks. NOE intensities were qualitatively categorized into strong, medium and weak signals and translated to upper bound distance limits of 2.5 Å, 3.5 Å and 5.0 Å respectively. These distances were used together with the predicted dihedral angles to carry out several rounds of structure calculations using CYANA (Combined assignment and dynamics for NMR applications) 2.1 [142]. Distance restraints were also adjusted for every calculation cycle based on NOE violations. Of the 100 structures, 20 lowest energy structures were selected for evaluation and analyses. PROCHECK-NMR [143] was employed to evaluate the stereochemical quality of the structure ensembles and figures were prepared using PYMOL [144], MOLMOL [145], and Discovery Studio Visualizer 2.0.

2.2.5 Rigid-body modeling of the α M β 2 complex

The calculated myristoylated α M structures were placed close to β 2 based on the chemical shift perturbation data and sent to the RosettaDock Server [146] for the modeling of closed conformation of the α M/P α M complex

with $\beta 2$. However, the RosettaDock Server does not accept modified amino acid group and hence does not take into account the charges of the phosphate group added to the calculated Myr- α MS14P structure. Hence the model of Myr-P α M. $\beta 2$ complex was generated by a local docking refinement module, with recommended side chain rotamer libraries, in Rosetta 3.2.1.

2.2.6 Paramagnetic perturbation studies on the free and myristoylated α M tails

0.7 mM α M, P α M, Myr- α M and Myr-P α M were separately dissolved in 10 mM sodium phosphate buffer, pH 5.6, with 200 mM DPC. NOESY experiments with 200 ms mixing times were carried out at 308K on both these peptides to obtain the free spectra and the spectra following the additions of Manganese Chloride (MnCl_2) respectively up to a final concentration of 1 mM [147]. The change in $\text{H}\alpha$ proton signal intensities, I , following the additions of the paramagnetic probes were then normalized to the signal intensities of the un-titrated spectra, I_0 , and plotted for each amino acid.

2.2.7 Chemical shift perturbation of the α M- $\beta 2$ cytoplasmic tails

The interaction interface between $\beta 2$ with α M/P α M was determined by ^{15}N -filtered NMR experiments, where ^1H - ^{15}N HSQC experiments were carried out at 298 K to detect the backbone amides of either ^{15}N - α M or ^{15}N - $\beta 2$ which was later titrated with their respective unlabeled partners of HPLC purified synthetic $\beta 2$ or α M tail up to 8 times concentration of the labeled tail. ^{15}N -labeled α M or $\beta 2$ purified cytoplasmic tails were dissolved in 10 mM sodium phosphate buffer, pH 6.5 to obtain a concentration of 0.1 mM for α M and 0.2 mM for $\beta 2$.

The chemical shift changes, after addition of the unlabeled titrant, of the backbone amide nitrogen and proton of each residue were calculated using the following equation [93]:

$$\Delta^1\text{H} + \Delta^{15}\text{N}$$

where $\Delta^1\text{H}$ is the absolute value of the change in amide proton chemical shift from the un-titrated ^1H - ^{15}N HSQC spectrum, and $\Delta^{15}\text{N}$ the absolute value of the

change in backbone amide nitrogen chemical shift from the untitrated ^1H - ^{15}N HSQC spectrum.

2.2.8 Assignment of backbone amides of αM and $\beta 2$

Two complementary pairs of triple resonance experiments, the HNCA with HN(CO)CA and the HNCACB with CBCA(CO)NH were carried out on 0.5 mM of purified ^{15}N ^{13}C αM and $\beta 2$ dissolved in 10 mM sodium phosphate buffer, pH 6.5. The experiments were carried out at 298 K for αM and 278 K for $\beta 2$. Backbone amide resonances of the ^{15}N HSQC were identified by the sequential assignment of intra- and inter- $^{13}\text{C}\alpha$ and $\text{C}\beta$ resonances of the triple resonance experiments [148,149]. The corresponding 298 K $\beta 2$ backbone amide resonances were obtained by performing ^{15}N HSQC at different temperatures, 283 K, 288 K, 293 K and 298 K.

2.2.9 Fluorescence resonance energy transfer studies:

All fluorescence studies were conducted using a 0.5 cm path length luminescence spectroscopy microcell (Perkin-Elmer, California, USA) on a Varian Cary Eclipse spectrophotometer. The sample was excited at 295 nm with 5 nm slit width, and the emission of the $\beta 2$ tryptophan and the dansyl group of the Dan- αM /P αM was collected from 310 nm to 580 nm at 10 nm slit width.

A variation of synthetic $\beta 2$ with His 5 changed to tryptophan and the distal Trp 24 to alanine ($\beta 2\text{H5W}$), and its myristoylated version (Myr $\beta 2\text{H5W}$) were bought from GL Biochem (Shanghai, China). The α/β cytoplasmic tail interaction is monitored by making use of the resonance energy transfer from the $\beta 2$ mutant Trp 5 to the dansyl group of the synthetic N-terminus dan- αM /P αM .

Myr $\beta 2\text{H5W}$ was dissolved in 10 mM sodium phosphate buffer, pH6.5, with 1.2 mM DPC and $\beta 2\text{H5W}$ was dissolved in 10 mM sodium phosphate buffer, pH 6.5. All experiments were carried out at 1 μM of $\beta 2$ peptides, except the experiment between $\beta 2\text{H5W}$ with Dan- αM , where 5 μM of $\beta 2\text{H5W}$ was used. The fluorescence energy transfer from the $\beta 2$ peptides' tryptophan maximal emission at ~ 355 nm to the dansylated- αM /P αM was quantified by

the decrease in the intensity of the β_2 tryptophan emission following the additions of dansylated- α M/- α MS14P up to five times the concentration of β_2 . The change in fluorescence intensity normalized to the intensity of the final titration point for each experiment, Normalized $F_{\max}$, was plotted as a function of concentration of the dansylated peptides [150]. A non-linear regression function in Origin7.0 was used to determine the dissociation constants, K_d , for the interaction between the myristoylated β_2 mutant with the dansylated- α M/ α MS14P:

$$\text{Normalized } \Delta F_{\max} = \frac{F_{\max} \cdot [\alpha M]}{K_d + [\alpha M]}$$

2.3 Results

2.3.1 Characteristics of the free α M, Myr- α M and Myr-P α M cytoplasmic tails

The effects of the N-terminal C₁₄ myristate group on the α M cytoplasmic tails, which acts as a probable the integrin transmembrane domain surrogate [91] were examined in a series of experiments.

2.3.2 Properties of the free α M, Myr- α M and Myr-P α M cytoplasmic tails

Early experiments have shown that the interaction between the α - and β -cytoplasmic domains are very weak and are hence conducted in buffers with low/null ionic strength [92,151]. Moreover, many studies have shown that electrostatic interactions predominate integrin membrane proximal cytoplasmic interactions [78,85,92]. The myristoylated and non-myristoylated peptides were dissolved in low ionic strength buffer and DPC micelles and the myristoylated peptides demonstrated a narrow chemical shift dispersion of $\approx 8.0 - 8.5$ ppm, broadened and overlapped NMR resonances and lack of NOEs in the ^1H - ^1H NOESY spectra (Figure 2.2 A and B), which possibly suggest that they have aggregated via their hydrophobic acyl chains in buffer. However in the presence of the deuterated DPC micelles, the un-myristoylated (α M and P α M) and myristoylated peptides (Myr- α M and Myr-P α M) were not aggregated in solution and instead demonstrated a wide chemical shift dispersion of $\approx 6.5 - 9.0$ ppm, sharp line-widths with large number of NOE connectivities characteristic of a well folded structure (Figure 2.2 C and D). Although the free α M/P α M was soluble in the DPC micelles, it did not have as wide a chemical shift dispersion as its myristoylated counterpart and its spectrum is characterized with overlapping resonances from $\approx 7.7 - 8.6$ ppm (Figure 2.2 C and D).

Figure 2.2 ^1H - ^1H NOESY spectra of myristoylated and un-myristoylated α M/P α M cytoplasmic tails. (A) Myr- α M dissolved in sodium phosphate buffer. (B) Myr- α M dissolved in sodium phosphate buffer. (C) (In black) Myr- α M dissolved in DPC micelles and (In red) α M dissolved in DPC micelles. (D) (In black) Myr-P α M dissolved in DPC micelles and (In red) P α M dissolved in DPC micelles.

2.3.3 Structural stabilization of the myristoylated peptides compared to the free α M

Free α M cytoplasmic tails dissolved in buffer displayed distinct characteristics in the ^1H - ^1H NOESY spectra recorded at room temperature 298 K and low temperature 278 K (Figure 2.3 A and B). The α M ^1H - ^1H NOESY spectrum at 298 K had a narrow chemical shift dispersion $\approx 8.0 - 9.5$ ppm with mostly intra-residues NOE connectivities characteristic of a random coil structure (Figure 2.3A). The maximum possible number of NOEs observed depends on the molecular correlation time which is inverse to the rate of molecular tumbling. At higher temperatures, molecules tumble rapidly and structure is less stable by decreasing the temperature to 278 K, more NOEs could be observed together with a slight increase in the chemical shift dispersion, arising from structural stabilization (Figure 2.3B). Sequential and medium range NOEs could be identified in the 278 K ^1H - ^1H NOESY spectrum.

Further structural stabilization of the myristoylated peptides in DPC micelles could be observed in the ^1H - ^1H NOESY (Figure 2.3 C and D), with even wider dispersion of resonances of $\approx 6.5 - 9.0$ ppm and a large number of NOEs between the backbone/side chain and side chain/side chain protons (Figure 2.3 C and D). Sequential NOEs from residue (i-i+1), (i-i+2), (i-i+3) and (i-i+4), medium and long range NOEs could be identified in the NOESY spectra which are characteristic of a folded helical structure [141].

Figure 2.3 ^1H - ^1H NOESY spectra of the free αM tails at 298 K and 278 K and myristoylated peptides in DPC micelles. ^1H - ^1H NOESY spectrum of Free αM cytoplasmic tails in (A) 298 K and (B) 278 K. ^1H - ^1H NOESY spectrum myristoylated αM (C) and P αM (D) in DPC micelles.

2.3.4 Micelle incorporation of the α M, PaM, Myr- α M and Myr-PaM

The topology of the free and myristoylated peptides in DPC micelles were next examined using the paramagnetic agent, MnCl_2 [152]. Paramagnetic agents have an unpaired electron spin, whereupon interaction with the ^1H nucleus will enhance the rate of relaxation and broadening of the ^1H resonance resulting in a decrease in signal intensity [152,153]. MnCl_2 , which is soluble in a hydrophilic environment, perturbs residues that are solvent exposed. MnCl_2 additions to free α M and PaM dissolved in DPC micelles resulted in the reduction in intensities along the cytoplasmic tail, with residues Q⁸ to G¹⁷ having the greatest reduction of signal intensity by 80% (Figure 2.4 A). This indicated that most of the amino acid residues of the free α M and PaM associate in DPC micelles and suggest that they adopt a horizontal topology on the surface of the micelles. Additionally, the MnCl_2 perturbation of the free α M cytoplasmic tails in the lipid-free buffer show that the N-terminal residues L² to F⁵ are protected from exposure to the paramagnetic Mn^{2+} ions (Figure 2.4 B). The random perturbation of the free α M tails in DPC and restricted diminution in intensities at the N-terminal residues connotes the entire stretch of the α M tail could possibly be interacting with the DPC micelles by virtue of the hydrophobic residues it contains. In addition, the hydrophobic stretch of α M at its N-terminus consisting of the GFFKR motif are protected from the Mn^{2+} ions in lipid-free buffers indicating an intrinsic tendency of these motif to associate together.

Figure 2.4 Paramagnetic relaxation by Mn^{2+} ions of α M/P α M in DPC micelles and in aqueous buffer. Relative changes in intensity of each residues of (A) α M in DPC micelles, (B) P α M in DPC micelles, (C) α M in buffer and (D) P α M in buffer.

In contrast, $MnCl_2$ induced almost 90% reduction of signal intensities in both the myristoylated peptides from residues R⁷ to Q²⁴, with limited decrease of intensities by 40% from L² to F⁵ (Figure 2.5). Hence, reiterating the lipophilic nature of the N-terminus L² to F⁵ stretch, however, this time as a consequence of the potential insertion of the α M/P α M tails into the DPC micelles by myristoylation. The prominent reduction in signal intensities experienced by the majority of residues in the myristoylated peptides concludes that they are exposed in the aqueous milieu. Therefore, the N-terminus myristate group could act as an anchor into the lipid core of the DPC micelles analogous to the α M transmembrane domain.

Figure 2.5 Paramagnetic relaxation by Mn^{2+} ions of showing the relative changes in intensity of each amino acid residue of Myr- α M (■)/P α M(□) in DPC micelles.

2.3.5 Secondary structure analyses of the α M, Myr- α M and Myr-P α M

Far UV CD spectroscopy was used to gauge the secondary conformations of the myristoylated peptides in DPC micelles. Distinct absorption of circularly polarized light by different secondary structure elements, such as the α -helix and β -sheet, displays characteristic absorption CD spectra that could be used to approximate the secondary structure of the protein or peptide [154]. The structural effects of the myristate moiety on the α M tails dissolved in DPC micelles could be observed in the CD spectra acquired for the free and myristoylated α M tails. The free α M tail adopts a random coil conformation in solution as evident from the negative molar ellipticity observed at ≈ 200 nm of the CD spectrum recorded at room temperature (Figure 2.6 A). However, in the presence of the myristate moiety that anchors the α M cytoplasmic tail into the DCP micelles, it adopts an α -helical conformation as indicated by their characteristic absorption at 225 nm and 208 nm (Figure 2.6 B and C). The α -helical content of Myr- α M and Myr-P α M was calculated based on the equation by Morris et. al (1973) and was estimated to be 86% and 23% respectively. The NMR spectrum (Figure 2.3) for both Myr- α M and Myr-P α M detected no presence of β -sheet structures, but detected a coil back which contributes to the random coil or β -sheet content in the calculation for percentage α -helical and β -sheet structures. The introduction of a phosphate moiety in the α M sequence

possibly resulted in a change of Myr-PaM assuming helical populations and hence the lower percentage helical content.

Hence, asserting the suitability of analysis of the myristoylated peptides as a physiological mimic of the α M cytoplasmic tails.

Figure 2.6 CD spectra of (A) free α M in aqueous buffer, (B) Myr- α M in DPC micelles and (C) Myr-PaM in DPC micelles.

Deviations from random coil chemical shifts of the individual amino acid backbone $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ chemical shifts obtained from the assigned ^1H - ^1H NOESY and ^{13}C - ^1H HSQC can be used to detect local secondary structure elements [155]. Formation of secondary structures such as the helix and β -sheets causes up- or down-field changes in chemical shift due to their distinct hydrogen-bonding patterns. α -helical structures causes down-field shift for $^{13}\text{C}\alpha$ resonances and a upfield shift for $^1\text{H}\alpha$ resonances whereas β -sheets displays a reverse trend in chemical shift deviations [155]. The $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ chemical shift deviations for the myristoylated peptides demonstrated helical type deviations of $^1\text{H}\alpha$ and for $^{13}\text{C}\alpha$ resonances for the N-terminal residues F⁴ to S¹⁴ irrespective of phosphorylation, and the lack of deviation for the C-terminal residues E¹⁵-Q²⁴ suggest that they form part of a random coil conformation (Figure 2.7 A and B). Chemical shift deviations for the free αM cytoplasmic tails at 278 K (Figure 2.7 C) were less pronounced than the myristoylated peptides but nonetheless suggest a weakly forming helical structure at its N-terminus as characterized by the negative $^1\text{H}\alpha$ and positive $^{13}\text{C}\alpha$ deviations, which indicates the intrinsic helix-promoting behaviour of the αM membrane proximal sequence. All random coil $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ chemical shifts for neighbouring residues (i-1, i-2, i+1 and i+2) of aromatic residues (F and Y), which were displayed to have ≈ 0.1 ppm deviation from standard chemical shift values were corrected for [156] before chemical shift deviation analyses. In addition, the unique ^1H and $^{13}\text{C}\alpha$ chemical shift of the phosphorylated serine was also used [157].

Figure 2.7 Chemical shift deviations from random coil values of $^{13}\text{C}\alpha$ (Top panel) and $^1\text{H}\alpha$ (Bottom panel) of (A) Myr- αM in DPC micelles, (B) Myr-P αM in DPC micelles and (C) αM in aqueous buffer at 278 K.

2.3.6 Structural characterization of the myristoylated peptides

Structure determination in solution-state NMR begins with the identification of individual amino acid spin systems from the signals present in the TOCSY spectrum. The identification of the location of each amino acid in the 2D- ^1H - ^1H TOCSY spectrum can aid in the identification of the sequence-specific position of each amino acid in the 2D- ^1H - ^1H NOESY spectrum. Unlike the signals observed in the TOCSY spectrum, signals in the NOESY spectrum, known specifically as Nuclear Overhauser Effects (NOESs), are observed due to $< 5 \text{ \AA}$ spatial proximity of two protons. Hence, the unambiguous identification of the signals, via the 'sequence-specific walk' in the NOESY spectrum followed by a qualitative scheme of distance assignment in correlation to observed signal intensity can provide the 3-dimensional (3D) structure of the peptide with the help of a computational program which places these inter-proton distances on the peptide [141].

The myristoylated peptides, Myr- αM and Myr-P αM in DPC micelles had a large number of NOEs and spectral dispersion suitable for structure determination (Figure 2.3 C and D). Secondary structure analyses have demonstrated that the myristoylated cytoplasmic tails are mostly helical at its N-terminal portion. Sequence-specific resonance assignments of their ^1H - ^1H NOESY spectra revealed sequential and medium NOEs from residues (i-i+1, i-i+2, i-i+3 and i-i+4) from L² to S¹⁴ together with many other NOEs arising from backbone/backbone backbone/side chain and side chain/side chain couplings. Most peculiarly, the ring protons from F⁴, F⁵ and Y⁹ are distinctly dispersed from 6.5 to 7.5 ppm allowing for their unambiguous assignment and identification of their medium and long range interactions with the rest of the amino acids. NOEs were observed between the aliphatic side chains of L² with the ring protons of F⁴ and F⁵, ring proton NOEs between F⁴/F⁵ and F⁵/Y⁹ (Figure 2.8A), and NOEs between the aliphatic side chains of M¹²/M¹³ and ring protons of Y⁹ (Figure 2.8 B). These observed aliphatic/aromatic and aromatic/aromatic NOEs are observed in both Myr- αM and Myr-P αM (Figure 2.8).

Figure 2.8 Panels of ^1H - ^1H NOESY of Myr- α M and Myr-P α M displaying the packing NOEs. (A) Packing NOEs between the ring protons of F⁵ and Y⁹ in the ^1H - ^1H NOESY panel of Myr- α M and Myr-P α M. (B) Packing NOEs from L² aliphatic side chains with the F⁵ aromatic ring protons; and packing of M^{12/13} aliphatic side chains with Y⁹ aromatic ring protons in the ^1H - ^1H NOESY panel of Myr- α M and Myr-P α M.

Two different sets of long range NOEs were detected and used for structural determination with fewer unambiguous long range NOEs detected for the phosphorylated form (Table 2.2).

Table 2.2 Table of long range NOEs used in structure determination

Table of long range NOEs used to determine the structure of Myr- α M		
1	² Leu C δ H1	⁷ Arg NH
2	² Leu C δ H2	⁷ Arg NH
3	⁹ Tyr C ϵ H	²⁴ Gln C γ H
4	⁹ Tyr C δ H	²⁴ Gln C β H
5	⁹ Tyr C δ H	²⁴ Gln NH
6	¹² Met C γ H	²⁴ Gln NH
7	¹² Met C β H	²⁴ Gln NH
8	¹³ Met C γ H	²⁴ Gln NH
9	¹³ Met C β H	²⁴ Gln NH
10	¹⁴ Ser NH	²⁴ Gln C β H
11	¹⁴ Ser NH	²⁴ Gln C γ H
12	¹⁵ Glu C β H	²⁴ Gln NH
13	¹⁸ Pro C β H	¹² Met NH
14	¹⁸ Pro C β H	¹³ Met NH
Table of long range NOEs used to determine the structure of Myr-P α M		
1	¹² Met C γ H	²⁴ Gln NH
2	¹² Met C β H	²⁴ Gln NH
3	¹³ Met C γ H	²⁴ Gln NH
4	¹³ Met C β H	²⁴ Gln NH
5	¹⁴ Ser C β H	²⁴ Gln NH
6	¹⁴ Ser NH	²⁴ Gln C α H
7	²³ Pro C β H	¹² Met NH
8	²³ Pro C β H	¹³ Met NH

Long range NOEs between Y⁹/Q²⁴, M¹²/Q²⁴, M¹³/Q²⁴, M¹²/P¹⁸, S¹⁴/Q²⁴ and M¹²/E²² observed for Myr- α M suggest that the C-terminal residues loops back to the N-terminal half (Figure 2.9 A). The fewer long range NOEs between the C-terminal loop and the N-terminal helix for the Myr-P α M indicates a slightly less rigid C-terminal loop packing possibly due to the introduction of a large negatively charged phosphate moiety (Figure 2.9 B).

Figure 2.9 Panels of ^1H - ^1H NOESY of Myr- α M and Myr-P α M displaying the long range NOEs. **(A)** Long range NOEs of the Myr- α M. **(B)** Long range NOEs of the Myr-P α M.

Distance restraints obtained via qualitative observation of the NOE intensity together with dihedral Φ and Ψ angles predicted based on chemical shift values of $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ were input into the automated structure calculation software, CYANA [142], followed by several rounds of refinements to generate the 3-dimensional coordinates of the Myr- α M and Myr-P α M. The summary of structural statistics of the 20 lowest energy conformers is listed in a Table 2.3.

Table 2.3 Summary of structural statistics for the 20 final structures of αM-tail		
Structure	Myr-αM	Myr-PaM
Distance restraints		
Total	204	171
Intraresidue ($i - j = 0$)	35	36
Sequential ($ i - j = 1$)	70	58
Medium-range ($2 \leq i - j \leq 4$)	85	69
Long-range ($ i - j \geq 5$)	14	8
Angular restraints (Φ)	20	20
Angular restraints (Ψ)	23	23
Distance restraints violations		
Number of violations	2	5
Average violation ($\text{\AA}$)	≤ 0.22	≤ 0.25
Maximum violation ($\text{\AA}$)	≤ 0.45	≤ 0.31
Average target function values	3.15(2.91-3.45)	2.82(2.68-2.96)
Deviation from mean structure		
All backbone (Ca, C', and N)/(Residue 1-14) ($\text{\AA}$)	0.14 (0.06)	0.39 (0.08)
All heavy atoms/(Residue 1-14) ($\text{\AA}$)	1.05 (0.83)	1.02 (0.94)
Ramachandran plot for the mean structure		
% residues in the most favourable	87	87
% residues in the additionally allowed region	13	13
% residues in the disallowed region	0	0

For both Myr- α M and Myr-PaM structure ensembles, their backbones were highly superimposed (Figure 2.10 A and B). Both myristoylated α M tails are helical from F⁴ to S¹⁴ concurring with earlier secondary structure analyses based on chemical shifts. The F⁴ - S¹⁴ helix display an amphipathic conformation with the hydrophobic residues F⁴, F⁵, Y⁹, M¹² and M¹³ facing one face of the helix and the hydrophilic residues R⁷, K¹⁰, D¹¹, S¹⁴/pS¹⁴ facing the other (Figure 2.10).

The structural effect on the α M cytoplasmic tail due to phosphorylation lies in the conformation of the C-terminal loop from E¹⁵ to Q²⁴. The C-terminal loop of the Myr- α M packs compactly with its N-terminal helix as defined by the van der Waals' interactions between the helix M¹²/M¹³ and C-terminal loop residues P¹⁸/P²³/Q²⁴ (Figure 2.10 D). The C-terminal loop of Myr-PaM does not loop back to its N-terminal helix as it is poorly defined due to the fewer loop-helix packing NOEs detected (Figure 2.10 F). The negatively charged phosphate moiety in the α M cytoplasmic tail could possibly disrupt the packing of the α M C-terminal loop due to cationic interaction with K¹⁰.

Figure 2.10 Structures of the Myr- α M and Myr-P α M. Superposition of the lowest 20 conformers of the (A) Myr- α M and (B) Myr-P α M. Blue ribbon representation of the Myr- α M with polar side chains represented in sticks and denoted (C) and in spheres (D) displaying the hydrophobic residues and C-terminal loop packing. Red ribbon representation of the Myr-P α M with polar side chains represented in sticks and denoted (E) and in spheres (F) displaying the hydrophobic residues and C-terminal loop packing.

2.3.7 Backbone triple-resonance experiments for α M and β 2 cytoplasmic tails

Each individual signal in the ^{15}N - ^1H HSQC spectrum represents a single amino acid, and can be identified with the specific resonant frequencies in corresponding backbone triple resonance experiments on $^{15}\text{N}/^{13}\text{C}/^1\text{H}$ uniformly labeled proteins/peptides. The direct and large scalar coupling between ^{15}N - ^1H , ^{13}C - ^1H and ^{15}N - ^{13}C nuclei are recorded in the backbone triple-resonance experiments with specific connectivities between the preceding and next amino acid (Figure 2.11A). These experiments take advantage of J-couplings, or spin-spin couplings between these nuclei that are one to three bonds apart. Triple resonance experiments usually begin with the net magnetization of the proton spins, followed by a series of pulse sequences to effect energy transitions for the ^{15}N , ^{13}C and ^1H nuclei. Known J-coupling values between various nuclei pair are exploited to direct net magnetization from one type of nucleus to another (Figure 2.11A). Backbone triple-resonance experiments are typically carried out in pairs of experiment with one experiment giving rise to correlations to both the preceding (Residue i-1) and the intra-residue amino acid (Residue i), while its partner giving rise only to inter-residue (i-1) correlation (Figure 2.11B). This direct sequential connectivity between the preceding and next amino acid can be reliably and robustly identified in the pair of backbone triple resonance experiments. The types of doublet combination of triple-resonance experiment allowing for sequential assignment are shown in Figure 2.11B.

(A)

(B)

Figure 2.11 Triple resonance experiments. (A) Scalar couplings between the ^{15}N , ^{13}C and ^1H nuclei. (B) Triple resonance experiments with the left panel showing experiments observing connectivities between the preceding and next amino acids and their complementary experiments observing signals from the preceding amino acid. Taken and adapted from Rule, G.S. and T.K. Hitchens, *Fundamentals of protein NMR spectroscopy. Focus on structural biology* 2006, Dordrecht: Springer. xxvi, 530 p.

The backbone amide resonance assignments for the αM cytoplasmic tail were identified using the triple resonance combination, HNCA and HN(CO)CA experiments (Figure 2.12), while the $\beta 2$ cytoplasmic tails resonance assignments were identified from the HNCACB and CBCA(CO)NH experiments (Figure 2.12).

Figure 2.12 Strip plots of triple resonance experiments of α M and β 2. $^{15}\text{N}-^1\text{H}$ Strip plots of α M residues S¹⁴ to G¹⁷ from the HNCA (red peaks) and HN(CO)CA spectra (blue peaks) on the left panel, and $^{15}\text{N}-^1\text{H}$ strip plots of β 2 residues S²² to N²⁶ from the HNCACB (red and green peaks) and CBCA(CO)NH spectra (blue peaks) on the right panel. Horizontal lines correlates residues $\text{C}\alpha/\text{C}\beta$ to the preceding amino acid.

2.3.8 Interaction analyses between the α M and β 2 tails by ^{15}N - ^1H HSQC

The interaction between the unmodified cytoplasmic α M and β 2 tails are examined in aqueous buffer solution. The backbone amide resonances of the ^{15}N - ^1H HSQC spectrum of the α M and β 2 are assigned using two pairs of complementary triple resonance experiments, the HNCA/HN(CO)CA and HNCACB/CBCA(CO)NH. The chemical shift changes of the ^{15}N labeled α M or β 2 were monitored at various concentrations of its unlabeled counterpart in ratios of 1:1, 1:2 and 1:8. Only at 1:8 ratio considerable combined chemical shift changes of the ^1HN and ^{15}N resonances could be observed for both ^{15}N α M or β 2 titrations, indicating the intrinsically weak interactions between the α M β 2 cytoplasmic tails (Figure 2.13 and 2.14). The α M residues L², R⁷, K¹⁰, D¹¹, M¹³ and S¹⁴ showed a much higher change in combined chemical shift (≥ 10 Hz) than the rest of the residues when it was titrated in over excess of β 2 (Figure 2.13).

Figure 2.13 Interaction analyses of the ^{15}N αM with unlabeled $\beta 2$. **(A)** ^{15}N - ^1H HSQC of ^{15}N αM in black contour and titrated spectrum with up to 8 times unlabeled $\beta 2$ in grey contour. **(B)** Combined chemical shift changes ($\Delta\text{N} + \Delta\text{H}$) of each amino acid of the ^{15}N - αM titrated with unlabeled $\beta 2$ in 1:8 molar ratio. **(C)** Ribbon representation of the Myr- αM structure, in spheres are residues perturbed ≥ 10 Hz.

Conversely, the $\beta 2$ residues H5, L6, E11 and Y12 showed the greatest chemical shift changes in the titration of both unlabeled αM or PaM peptides

(Figure 2.14). The residues perturbed as a result of the presence of its binding partner are located at the α M and β 2 membrane proximal cytoplasmic regions.

Figure 2.14 Interaction analyses of the ^{15}N β 2 with unlabeled α M and PaM. (A) ^{15}N - ^1H HSQC of ^{15}N β 2 in black contour and titrated spectrum with up to 8 times unlabeled α M in grey contour. (B) Combined chemical shift changes ($\Delta\text{N} + \Delta\text{H}$) of each amino acid of the ^{15}N - β 2 titrated with unlabeled α M in 1:8 molar ratio. (C) ^{15}N - ^1H HSQC of ^{15}N β 2 in black contour and titrated spectrum with up to 8 times unlabeled PaM in grey contour. (D) Combined chemical shift changes ($\Delta\text{N} + \Delta\text{H}$) of each amino acid of the ^{15}N - β 2 titrated with unlabeled PaM in 1:8 molar ratio.

2.3.9 Binding interactions between β 2 and α M tails in DPC micelles

Fluorescence spectroscopy is a highly sensitive technique that allows the use of micro-molar concentrations of reagents in solution. The fluorescence phenomenon, **Resonance Energy Transfer (RET)** between a donor and an acceptor molecule can be used to approximate the binding kinetics between integrins cytoplasmic domains [150]. RET is a fluorescence phenomenon whereby energy transfer occurs without the emission or re-absorption of a photon but is a result of interactions between donor-acceptor dipoles [150]. Unlike the usual fluorescence phenomena that depends on interactions of the fluorophore with other molecules in the surrounding solvent shell, intervening solvent and molecule has little effect on RET efficiency which depends primarily on the donor and acceptor distances [150]. RET occurs between donor molecules in the excited and acceptor molecules in the ground state, and

typically have overlapping donor emission and acceptor absorption spectrum (Figure 2.15).

Figure 2.15 Fluorescence absorption and emission spectrum of tryptophan (Trp, green line) and dansyl (red dotted line). The overlapping emission of Trp and absorption of Dansyl is shaded in orange. Taken from Lakowicz, J.R., *Principles of fluorescence spectroscopy* 1983, New York: Plenum Press. xiv, 496 p.

Energy transfer occurs only between donors and acceptors that are within their characteristic Förster distance and thus can be used as a proximity indicator. The efficiency of energy transfer between the donor and acceptor is dependent on the extent of donor emission and acceptor absorption spectral overlap; quantum yield of the donor; orientation of the donor and acceptor transition dipoles; and most importantly the distance between the donor and acceptor molecules [150]. This can be quantified in the equation:

$$k_T(r) = \frac{1}{\tau_D} \left(\frac{R_0}{r} \right)^6$$

Where, τ_D is the decay time of the donor in the absence of acceptor, R_0 is the Förster distance, and r is the donor-acceptor distance [150].

The amino acid Trp has a distinct excitation wavelength (~295 nm) and emission profile (~350 nm) that overlaps the excitation wavelength of the Dansyl moiety (Figure 2.15) and they have a Förster distance of 21-24 Å [150]. Hence, energy transfer between these two chemical groups can be used to approximate the binding kinetics by measuring the decreasing fluorescence intensity of a fixed amount of donor molecule (Trp) during a titration with increasing acceptor (Dansyl) concentrations.

The binding kinetics of the α M/P α M and β 2 in the presence of DPC micelles were examined by resonance energy transfer between tryptophan and a dansyl moiety. However, the only tryptophan in the β 2 is located distally from the site interaction and could possibly further undermine the weak cytoplasmic interactions as observed in the lack of tryptophan fluorescence quenching. Myristoylated- β 2 wild type peptide does not undergo resonance energy transfer with the dansylated α M/P α M peptides as observed from its increasing intensity upon addition of Dan- α M (Figure 2.16). Hence, the distal W²⁴ is mutated to an Alanine and the β 2-H⁵ which is among the highly perturbed in its interaction with α M is mutated to Tryptophan which will act as the fluorescence donor to the dansyl moiety which will be added to the N-terminus of the α M and P α M.

Figure 2.16 Tryptophan emission spectrum of Myr- β 2 wild type in DPC micelles titrated with Dan- α M.

The mutated β 2 peptide, β 2H5W, predictably, interacts weakly with the dansylated Dan- α M and Dan-P α M peptides in aqueous buffer solution as observed from the lack of change in tryptophan fluorescence intensity (Figure 2.17A) and increasing tryptophan fluorescence intensity (Figure 2.17B).

Figure 2.17 Tryptophan emission spectrum of β 2H5W mutant titrated with (A) Dan- α M and (B) Dan-P α M in aqueous buffer solution.

Hence, a N-terminal myristoylated- β 2H5W incorporated into the DPC micelles was used to observe its interaction with the Dan- α M and Dan-P α M peptides. Resonant energy transfer due to close proximity between the tryptophan-Dansyl Förster distance ($\leq 24 \text{ \AA}$) resulted in a concentration dependent quenching in the H^5W fluorescence intensity between the Myr- β 2H5W and dansyl moiety of both Dan- α M and Dan-P α M (Figure 2.18 A and B) [150]. The incorporation of the Myr- β 2H5W into the DPC micelles possibly resulted in a stabilization of secondary structure elements required for cytoplasmic interaction and hence decreased the energy of entropy that has to be overcome in the randomly structured peptide in solution to interact with its binding partner. Corresponding to the requirement of the membrane to increase binding affinity between the cytoplasmic tails, interaction between acetonitrile solubilized Myr- β 2H5W and Dansylated peptides showed negligible fluorescence energy transfer (Figure 2.19).

Figure 2.18 Resonance energy transfer analysis of Myr-β2H5W and Dansylated αM peptides. Tryptophan emission spectrum of Myr-β2H5W mutant titrated with (A) Dan-αM and (B) Dan-PαM in DPC micelles. Non-linear least squares analysis of the maximum fluorescence intensity as a function of (C) Dan-αM concentration and (D) Dan-PαM concentration.

Figure 2.19 Resonance energy transfer of Myr- β 2H5W dissolved in 50% acetonitrile and (Top panel) Dan- α M and (Bottom panel) Dan-P α M.

The changes in tryptophan fluorescence intensity in the DPC micelle-inserted Myr- β 2H5W due to fluorescence energy transfer to the dansyl moiety of the Dan- α M and Dan-P α M demonstrated that the binding interaction between the cytoplasmic tails is possibly increased due to structural stabilizations of the N-terminal regions attributed by lipid anchoring of the Myr- β 2H5W. The concentration dependent quenching of the Myr- β 2H5W tryptophan due to the dansylated peptides were used to determine the equilibrium dissociation constants (K_d) (Figure 2.18 C and D). K_d values obtained from non-linear regression analysis of the changes in tryptophan fluorescence intensity as a function of dansylated peptides concentration were $5.2 \pm 0.6 \mu\text{M}$ and $4.4 \pm 0.7 \mu\text{M}$ for Dan- α M and Dan-P α M respectively.

2.3.10 Mode of interactions of the α M and β 2 tails

The interaction between the integrin α - and β - cytoplasmic domains is weak in nature and is found experimentally to be stabilized by a few interactions, such as the conserved α -R and β -D salt bridge and the presence of the two Phe present in the conserved α -GFFKR motif [77,78]. The weak interaction precludes the structural determination of the interaction interface using NMR spectroscopy, although NMR titration experiments can arbitrarily map the corresponding residues involved in the interaction. Structural insights into such weak and transient interactions can be obtained with modeling in combination with experiment data [158]. RosettaDock, a convenient and freely available online resource, allows users to submit their protein docking tasks to their server for calculations [146]. The RosettaDock has been benchmarked in the **Critical Assessment of PRediction of Interactions (CAPRI)** to have achieved full-atom models with Root Mean Square Deviations (RMSDs) better than 1.6 Å, and producing several structures that were the most atomically accurate models submitted by any group in the CAPRI challenge [159]. RosettaDock requires a starting position of the two proteins with an interface that must be a reasonable guess based on experimental data as its algorithm will search for a set of conformations near the given starting conformation. It uses a Monte Carlo-based side chain rotamer sampling (packing), and gradient-based rigid-body minimization to find a local energy minimum. The two structures to be docked are perturbed via a ± 3 Å horizontal and a ± 8 Å vertical translation along the axis of the interaction interface; and a 8° tilting which will complete a 360° spin around the axis between the centers of the two proteins, followed by the optimization of side chain configurations [146]. The RosettaDock server will then generate 1000 independent docked structures from this modeling algorithm and a plot of the total energy of each individual structures showing the profile of the energetic binding funnel [146]. The top 10 scoring models of the 1000 structures will be presented with detailed scoring information [146].

The structures of the myristoylated α M/PaM and the β 2 obtained from Bhunia et. al. (2009) [93] were first iteratively docked based on the ^{15}N - ^1H chemical shift changes. These models were then energetically refined using the

RosettaDock rigid-body with side chain optimization docking protocol. The N-terminal membrane proximal regions of the α M and β 2 assumes a helical conformation as reported in other structural studies of the α Ib β 3 [85,91,92] and α L β 2 [93] integrins. The helical membrane proximal regions of the α M/P α M and β 2 are orientated in parallel to each other with their C-terminal membrane distal regions looping away from the interacting helices (Figure 2.20). The interaction between α M/P α M and β 2 cytoplasmic tails is characterized by hydrophobic packing between the aliphatic side chains of the β 3-L³I⁴ and the aromatic side chains of α M/P α M-F⁴F⁵, together with charge-charge interactions between α M/P α M and β 2 side chains. For both α M/P α M interactions with β 2, the canonical salt bridge interaction between α M-R⁷ and β 2-D⁸ is observed. The E¹¹ residue in β 2, which is found to interact with its α L partner by Bhunia, et. al. (2009) [93], is too found to be in close proximity to the α M/P α M K¹⁰. The addition of a phosphate moiety to the α M cytoplasmic tail, seems to result in a rearrangement of potential ionic interactions between with the β 2 cytoplasmic tails (Figure 2.20). In the un-phosphorylated α M, residues D¹¹ and E¹⁵ are found to form multiple ionic and hydrogen bonds with the β 2-R¹⁰ and R¹⁴ residues. The addition of a phosphate moiety resulted fewer bifurcated ionic interactions with the phosphate group found to be in close ionic proximity with the β 2-R¹⁴ (Figure 2.20 B).

Figure 2.20 Ribbon representation of the modeled complex of (A) Myr- α M (in blue) to β 2 (in black) and (B) Myr-P α M (in red) to β 2 (in black) with small panels on the right displaying residues in proximity for ionic interactions and H-bonding (green lines). In spheres are hydrophobic residues and the side chains of the membrane proximal residues are represented in sticks.

2.4 Discussion

The cytoplasmic domains of the integrin heterodimers play an important role in integrin activation status. They do not have catalytic properties but serve as a docking hub to recruit specific proteins despite their relatively short sequences compared to the rest of the integrin molecule. The regulation of integrin activation via the β -subunit cytoplasmic tails through recruitment of multifarious intracellular ligands, such as talin, migfilin, filamin, etc, through specific motifs and modifications (phosphorylation) of the β -subunit. Comparatively, much less is known about the regulation of the α -subunit cytoplasmic tails. The β -subunit cytoplasmic tails are generally more conserved than the α -subunit cytoplasmic tails with many recognizable motifs, such as its membrane proximal motif and membrane distal triple TTT and NPXY motifs. In the α -subunit cytoplasmic regions, only the N-terminal membrane-proximal residues are highly conserved, whereas their C-termini are characterized by differing sequence length and variegated sequences that are unrelated to each other (Figure 2.1). Different α -subunits play different roles in the cellular context where exchanging α -cytoplasmic tails of the $\alpha 2\beta 1$ integrin with that of the $\alpha 4$ and $\alpha 5$ cytoplasmic tails resulted in altered extracellular ligand binding properties of cells that they were transfected in [160]. The different α -subunits cytoplasmic domains of the $\beta 2$ integrins are also shown to recruit specific intracellular effectors, where $\alpha M\beta 2$ but not $\alpha L\beta 2$ or $\alpha X\beta 2$ recruits Src kinase Hck [127]. Further highlighting the differential recruitment of proteins by different α -cytoplasmic domains is demonstrated in the calcium and integrin-binding protein (CIB) interaction with $\alpha IIB\beta 3$ but not to its closely related $\alpha V\beta 3$ integrin [106]. The understanding of integrin regulation by the α -subunit is relatively handicapped by the lack of α -subunit cytoplasmic tail structures. Most functional and structural studies, however, concur on the role of the cytoplasmic membrane proximal regions to maintain the resting state integrin, leaving the membrane distal regions for the selective recruitment of intracellular ligands.

Currently, the majority of integrin cytoplasmic domains structural efforts are focused on the α IIB β 3 integrin. Thus far, the α IIB cytoplasmic tail has been individually characterized with or without N-terminal myristoylation in the presence of DPC micelles (Figure 2.21 A and B) [91] and with its transmembrane counterpart in bicelles (Figure 2.21 C) [63] to represent the activated state when the cytoplasmic tails are separated.

Figure 2.21 Individual structures of α IIB cytoplasmic domains. (A) Myristoylated α IIB in DPC micelles (Vinogradova, et al.(2000)). (B) Myristoylated α IIB PP/AA inactive mutant in DPC micelles (Vinogradova, et al.(2000)). (C) Transmembrane/truncated cytoplasmic domain in bicelles (Lau, et al.(2008)).

The resting state α IIB β 3 cytoplasmic tail conformation has also been characterized in aqueous buffer solutions [90]; together with its transmembrane domains in micelles [73]; N-terminal attached coiled-coils that has been linked by disulfide bonds [89]; in the presence of its β 3 cytoplasmic binding partner N-terminal attached to a Maltose-binding protein (Figure 2.22 A) [92]; characterized in partiality with its transmembrane counterpart in bicelles (Figure 2.22 C) [63]; and in a hydrophobic 50% acetonitrile environment (Figure 2.22 D) [85]. In addition, the α L- cytoplasmic domain of the β 2 integrins has been structurally characterized and demonstrated a compact tri-helical fold stabilized by electrostatic interactions between the three helices (Figure 2.22 B).

Figure 2.22 Known NMR characterized complex structures of α IIb β 3 and the α L β 2 cytoplasmic tails. **(A)** Structure of the α IIb β 3 complex elucidated with the aid of a N-terminal MBP attachment using transferred NOESY. **(B)** Structure of the α L β 2 cytoplasmic tail complex elucidated in aqueous buffer solution. **(C)** Structure of the transmembrane/truncated cytoplasmic tail α IIb β 3 complex in bicelles. **(D)** Structure of the transmembrane/cytoplasmic domains of the α IIb β 3 complex in 50% acetonitrile.

In most of these structural characterizations, the N-terminal (α IIb-V⁹⁹⁰ to R⁹⁹⁵) membrane proximal regions adopts an α -helical conformation followed by a C-terminal membrane distal loop, with the exception of the α IIb transmembrane-truncated cytoplasmic domains elucidated by Lau, et. al. (2008) [63] which demonstrated a disordered α IIb cytoplasmic domain with an unusual reverse turn exhibited by the conserved G⁹⁹¹FF motif (Figure 2.22 C). In our study, the myristoylated α M cytoplasmic tails elucidated in DPC micelles adopts a longer helical conformation (L²⁽¹¹¹⁴⁾ to E⁽¹⁵⁾¹¹²⁷) at its N-terminal followed by a loop that packs to the N-terminal helix stabilized by non-polar

interactions. The structure of the $\alpha\text{M}\beta 2$ complex structure draws several similarities to the $\alpha\text{L}\beta 2$ complex elucidated in Bhunia, et al. (2009) [93], which also demonstrated N-terminal hydrophobic packing and subsequent ionic and electrostatic interactions involving similar $\beta 2$ amino acids such as $\alpha\text{M}\text{-R}^{7(1120)}\text{-}\beta 2\text{-D}^{8(709)}$ and $\alpha\text{M}\text{-K}^{10(1123)}\text{-}\beta 2\text{-E}^{11(712)}$ (Figure 2.20) vs. $\alpha\text{L}\text{-R}^{1094}\text{-}\beta 2\text{-D}^{709}$ and $\alpha\text{L}\text{-R}^{1094}\text{-}\beta 2\text{-D}^{709}$ (Figure 2.22 B). The addition of a phosphate moiety to the αM membrane distal $\text{S}^{14(1126)}$ only resulted in lesser packing of the C-terminal loop. Hence, different membrane distal cytoplasmic sequences of the αIb , αL , and αM resulted in different structural conformations that might act as distinct recognition sites for different intracellular ligands.

It is also possible that the transmembrane helix and interactions with its binding partner could stabilize cytoplasmic helical conformations as demonstrated in the complex $\alpha\text{Ib}\beta 3$ transmembrane/cytoplasmic structure [85] elucidated in acetonitrile and the characterization of the $\alpha\text{Ib.MBP-}\beta 3$ complex [92]. Additionally, myristoylated αIb cytoplasmic structures mostly have a helical membrane proximal region. The GFF reverse turn observed could be due to a shorter cytoplasmic sequence used in the construct that might have an energetic bias towards excursions into the lipid core of the bicelles. A disordered αIb full length cytoplasmic tail was also observed in a NMR study which characterized part of its transmembrane domain linked via a disulphide bond to its $\beta 3$ part transmembrane/cytoplasmic counterpart in phospholipid bicelles. However, despite the disordered αIb cytoplasmic domain observed, α - and β - cytoplasmic interactions were still maintained by the membrane proximal residues $\alpha\text{Ib V}^{990}\text{-N}^{996}$ and $\beta 3 \text{L}^{713}\text{-E}^{726}$, albeit only $\beta 3\text{-D}^{723}$ was observed to be perturbed in the complex. Hence, suggesting that the αIb sequence is highly malleable to structural changes, such that the resting to activated state integrin is most possibly stabilized by intrinsic (sequence) and extrinsic (modification and ligands) factors. The left-handed GFF turn in the αIb could thus be an intermediate in the integrin activation pathway, where binding of intracellular ligands to sequences results in the unfolding of the membrane proximal helix allowing it to transverse into the lipid core and disrupt the membrane proximal electrostatic interactions that hold the cytoplasmic clasp together. The helical stabilization effects of the free αM

cytoplasmic tails following the addition of a myristate group and use of DPC micelles, and the observation that the Myr- α M/P α M residues L²⁽¹¹¹⁵⁾ to F⁵⁽¹¹¹⁸⁾ has partial hydrophobic protection from the paramagnetic Mn²⁺ ions, suggests the possibility of trespassing the lipid core if provided with energetic favourability for unfolding.

Despite the use of different constructs to elucidate the molecular structures, interactions between the α and β cytoplasmic tails as observed by structural studies are mostly predominated by electrostatic interactions between the canonical salt bridge α -R and β -D and other polar residues (Figure 2.20 and 2.22), and are characterized by weak binding kinetics which has been detected in low salt conditions [151]. Molecular models of the α M/P α M and β 2 complexes are shown to be stabilized by N-terminal hydrophobic packing that could possibly be stabilized in the presence of DPC micelles and electrostatic interactions at the subsequent membrane proximal regions. Only subtle differences exist in the way phosphorylated and unphosphorylated α M interacts with the β 2 cytoplasmic tail. However, these complex models can only provide limited insight into the physiological cytoplasmic interaction as they are not modeled in the presence of its membrane mimic but could nonetheless serve to reveal some insights into the cytoplasmic interactions between the α M/P α M and β 2.

Interactions between α -subunit cytoplasmic tails are different for different integrins or related integrins. Surface Plasmon Resonance (SPR) analysis of the interaction between GST- β 3 cytoplasmic tail and free α IIB cytoplasmic tail reveals an interaction of K_d value of 7.7 μ M and 50 μ M in the presence or absence of calcium ions [161]. Isothermal calorimetric analyses of the binding α L and β 2 cytoplasmic tails in aqueous solution yielded a dissociation constant of 2.6 μ M [93]. The over excess titration of free cytoplasmic α M/P α M or β 2 tails in our study demonstrates the intrinsically weak interactions in aqueous solution which could be stabilized by lipid anchoring into DPC micelles, yielding values of 5.2 μ M and 4.4 μ M respectively for α M/Myr- β 2H5W and P α M/Myr- β 2H5W interactions. Hence, integrin cytoplasmic interactions in the cell could be stronger than anticipated

from studies involving only cytoplasmic domains, due to the presence of its transmembrane and extracellular domains; and phospholipid bilayer. In addition, they are also modulated by metal ions, such as calcium, and phosphorylation. In $\alpha\text{IIb}\beta 3$, calcium ions are found to increase the strength of the α - and β -cytoplasmic interactions, these calcium ions most possibly interact with the αIIb membrane distal loop which contains several acidic residues; and in $\alpha\text{L}\beta 2$, calcium ions are found to interact with the αL -cytoplasmic domain, and its structure is found to be characterized by many negative patches that could function as ion binding sites [93]. However, the membrane distal cytoplasmic portion of the αM -subunit does not contain as much acidic residues as αIIb cytoplasmic tail (Figure 2.1), and it was not found to interact with calcium ions as observed in NMR titration studies where addition of CaCl_2 does not cause chemical shift and resonance intensity changes to both Myr- αM and Myr-P αM dissolved in DPC micelles (Figure 2.23).

Figure 2.23 Overlaid calcium titration TOCSY spectra of (A) Myr- αM and (B) Myr- αMS14P . Resonances in red is the un-titrated spectra, following addition of 0.5 mM CaCl_2 (in blue) and finally 1 mM CaCl_2 (in green)

The lack of difference in the dissociation constants obtained from the $\beta 2$ resonance energy transfer to unphosphorylated and phosphorylated αM suggest that the differences in extracellular ligand binding in the $\alpha\text{M}\beta 2$ and

$\alpha M^{pS14(1126)}\beta 2$ could be due to recruitment of different sets of intracellular effectors that might possibly be involved in lateral force activation resulting in distinct exposure of extracellular ligand binding sites. Hence, it would be interesting to uncover intracellular effectors that could bind specifically to the phosphorylated αM which could further provide directions in future structural characterizations of the αM cytoplasmic tail in order to understand its differential extracellular ligand binding modes.

Chapter III: Structure and binding interface of the cytosolic tails of the $\alpha X\beta 2$ integrin

3.1 Overview

The $\alpha X\beta 2$ integrin consisting of two fragments, one with a molecular weight of 150 kDa and the other 95 kDa non-covalently associated, is also known as the p150/95 particle and complement receptor type 4 (CR4) due to its role in the innate immune system. The $\alpha X\beta 2$ integrin is found to be expressed on macrophages, dendritic cells, with lower amounts on other myeloid cells, and activated B-lymphocytes, natural killer cells and cyto-toxic T-cells [162]. The activated $\alpha X\beta 2$ integrin binds to wide variety of extracellular ligands, such as osteopontin, plasminogen, heparin, and many other ligands with exposed acidic residues [163]. In addition, it shares many extracellular ligands with the $\alpha M\beta 2$ integrin, such as iC3b, ICAMs, fibrinogen, etc [163]. The $\alpha X\beta 2$ integrin plays an important role in host defense and function in chemokine degranulation, chemotactic migration and phagocytosis of complement opsonized particles. Its main role in immunity has been demarcated in the αX knockout mice which do not exhibit gross abnormalities but have lesser propensity to develop adipose tissue inflammation induced by high fat diet [24]. In addition, the over-expression of the $\alpha X\beta 2$ integrins have been found to cause painful sterile ulcers in patients with defective neutrophil migration [5].

The importance of the outside-in and inside-out signal convergence on the non-catalytic cytoplasmic domains of integrins necessitates further understanding of the integrin cytoplasmic domains. Despite the comparatively high sequence conservation of the $\beta 2$ integrins, inside-out activation does not affect them uniformly, with the $\alpha X\beta 2$ integrin found to be less susceptible to intracellular activation compared to the rest of the $\beta 2$ integrins [164]. The cytoplasmic regions of the α -subunits of the $\beta 2$ integrins are only highly conserved at their N-terminal membrane proximal and show great sequence variety at their membrane distal regions. Structural characterization of the α -subunit cytoplasmic domains could possibly provide insights into the differential intracellular activation threshold for the $\beta 2$ integrins. Thus far, the

α L [93] and α M cytoplasmic regions and their interaction with the β 2 cytoplasmic tail has been structurally characterized. Hence, we will continue to elucidate the structure and binding interface of the α X β 2 integrin.

The α X integrin subunit from K¹¹¹⁰ to K¹¹⁴⁵ and β 2 integrin subunit from K⁷⁰² to S⁷⁴⁷ will be renumbered from one to facilitate reference to α X β 2 integrin cytoplasmic regions (Figure 3.1) [24].

```

 $\alpha$ X 1110(1) K V G F F K R Q Y K E M M E E A N G Q I A P E N G T Q T P S P P S E K
 $\alpha$ M 1114 K L G F F K R Q Y K D M M S E G C P P C A E P Q
 $\alpha$ L 1087 K V G F F K R N L K E K M E A G R G V P N G I P A E D S E Q L A S G Q E A G D P G C L K P L H E K D S E S G G I
 $\alpha$ D 1108 K L G F F K R H Y K E M L E D K P E D T A T F S G D D F S C V A P N V P L S
 $\alpha$ IIb 989 K V G F F K R N R P P L E E D D E E G E
 $\alpha$ 4 1001 K A G F F K R Q Y K S I L Q E E N R R D S W S Y I N S K S N D D

 $\beta$ 2 702(1) K A L I H L S D L R E Y R R F E K E K L K S Q W N N D • N P L F K S A T T T V M N P K F A
 $\beta$ 3 716 K L L I T I H D R K E F A K F E E E R A R A K W D T A N N P L Y K E A T S T F T N I T Y R
 $\beta$ 1 742 K L L M I I H D R R E F A K F E K E K M N A K W D T G E N P I Y K S A V T T V V N P K Y E

```

Figure 3.1 Sequence alignment of the α -subunit and β -subunit cytoplasmic sequences with the renumbering of the α X and β 2 cytoplasmic domain.

3.2 Materials and Methods

3.2.1 Synthesis and purification of Myr- α X tails

The myristoylated form of the α X cytoplasmic tails, abbreviated Myr- α X, were purchased in the crude form from GL Biochem (Shanghai, China). The crude peptide was clarified by a linear acetonitrile/water gradient, at a flow rate of 2ml/min, passing through a C18 column (300 Å pore size, 5 µM particle size) connected to a Waters reverse phase High Performance Liquid chromatography (HPLC) system (Massachusetts, USA). The major peak fraction was collected and lyophilized. The peptide powder was later reconstituted in the required buffers.

3.2.2 Expression and purification of α X and β 2 tails

The α X(K¹¹¹⁰ - K¹¹⁴⁴) and β 2 (K⁷⁰² - S⁷⁴⁷) cytoplasmic regions were cloned into the pET-31b(+) vector (Novagen EMD, San Diego) described previously in Bhunia et. al. (2009) [93]. Unlike the β 2 construct which has a Met cleavage site inserted prior to the integrin cytoplasmic sequence in the KSI recombinant protein, Asp and Pro were introduced for the α X construct as the cleavage site to recover the cytoplasmic peptide from KSI.

α X and β 2 was purified using the same method to purify α M and β 2 as described in chapter 2.2.2.

The exact sequences of the synthetic and purified α X and β 2 tails used in this study are summarized in Table 3.1.

Table 3.1 Summary of sequences for synthetic and purified α - and β - tails

Peptide	Method	Sequence
Myr- α X	Synthetic	Myr- ¹ KVGFFKRQYKEMMEEANGQIAPENGQTTPSPSEK ³⁵
α X	Synthetic	¹ KVGFFKRQYKEMMEEANGQIAPENGQTTPSPSEK ³⁵
β 2	Synthetic	¹ KALIHLSDLREYRRFEKEKLKSQWNNDNPLFKSATTTVI* ¹ NPKFAES ⁴⁶
¹⁵ N α X/ ¹⁵ N-	Expressed	<u>P</u> ¹ KVGFFKRQYKEMMEEANGQIAPENGQTTPSPSEK ³⁵ <u>D</u>
¹⁵ N β 2/ ¹⁵ N-	Expressed	<u>L</u> ¹ KALIHLSDLREYRRFEKEKLKSQWNNDNPLFKSATTTVI* ¹ NPKFAES ⁴⁶ <u>GOGHHHHHHH</u>

*Sequences underlined are amino acids that have been changed from the original sequence

*Sequences underlined and bolded are extra amino acids added in the generation of the pET-31b- tail clones

+ Methionine in the endogenous β ₂ sequence has been replaced by isoleucine to facilitate purification by preventing intrasequence cyanogen bromide cleavage after methionine.

3.2.3 Structural derivation of the myristoylated α X tail

All NMR experiments were carried out on a Bruker DRX 600-Mhz instrument equipped with an actively shielded cryo-probe. 10% deuterium oxide and 2 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added to all NMR samples. All NMR spectra are directly/indirectly referenced to the ¹H of DSS.

2-dimensional proton Total Correlation Spectroscopy (TOCSY) and Nuclear Overhauser Spectroscopy (NOESY) experiments were carried out at 308K for 0.7 mM of purified Myr- α X dissolved in 10 mM sodium phosphate buffer pH 5.6, containing 200 mM deuterated dodecylphosphocholine (DPC) (CIL, Massachusetts, USA) added. TOCSY mixing time was fixed at 50 ms and NOESY mixing time was kept at 200 ms. In addition, ¹³C-¹H Heteronuclear Single Quantum Coherence (HSQC) experiment was also carried out at 100% D₂O to obtain ¹³C α chemical shifts for the myristoylated peptide. Raw NMR data was processed using TOPSPIN 2.1 and analyzed with SPARKY[139].

Structure calculation procedures, data processing and analysis are the same as described in chapter 2.2.4.

3.2.4 Paramagnetic perturbation studies on the myristoylated α X tails

2-dimensional proton TOCSY with 50 ms mixing time was carried out on 0.7 mM Myr- α X dissolved in 10 mM sodium phosphate buffer containing 200 mM deuterated DPC, in the absence and presence of 1 mM MnCl_2 . The intensity of the intra-residue C α H/NH cross peaks after the addition of MnCl_2 was normalized to the peak intensity of the unperturbed sample.

3.2.5 Assignment of backbone amides of α X and β 2

Standard 3-dimensional HNCACB and CBCA(CO)NH experiments were performed on both 0.5 mM of $^{15}\text{N}/^{13}\text{C}$ α X and β 2 dissolved in 10 mM sodium phosphate buffer pH 6.5 at 298K, to obtain the backbone amide resonances for the 2-dimensional ^{15}N - ^1H HSQC of ^{15}N - α X and ^{15}N - β 2.

3.2.6 Chemical Shift Perturbation Experiments

2-dimensional ^{15}N - ^1H HSQC was used to determine the interactions between the α X and β 2 cytoplasmic tails. ^{15}N - α X or ^{15}N - β 2 dissolved in 10 mM sodium phosphate buffer pH 6.5 was titrated with unlabeled β 2 and α X respectively up to two times the concentration of the labeled sample. The change in amide nitrogen and proton chemical shift of each amide peak from the un-titrated spectra was calculated using the following equation: $\Delta^1\text{H} + \Delta^{15}\text{N}$, where $\Delta^1\text{H}$ is the absolute value of the change in amide proton chemical shift after addition of the titrant and $\Delta^{15}\text{N}$ is the absolute value of the change in backbone amide nitrogen chemical shift.

3.2.7 Rigid-backbone modeling of the α X β 2 complex:

The calculated Myr- α X structure and the β 2 structure from Bhunia, et. al. (2009) [93] was placed in close contact to each other based on the α X and β 2 residues perturbed in the chemical shift perturbation experiment to form the approximated starting structure for input into the RosettaDock Server [146].

3.2.8 Circular Dichroism (CD) studies

0.2 mM Myr- α X dissolved in 10 mM sodium phosphate buffer pH 5.6 containing 60 mM of DPC was loaded onto a 0.1 mm path length cuvette (Hellma, Müllheim, Germany) for circular dichroic measurements using the

Chirascan CD spectrometer (Applied Photophysics, Leatherhead, UK). The range of signal acquisition was set at 190 to 240 nm with slit width of 0.1 nm and the spectrum was acquired at 0.1 s time steps. Three spectra were acquired for the myristoylated peptide and was then averaged and subtracted from the corresponding background buffer contributions. The final CD spectrum was converted to molar ellipticity and smoothed.

3.3 Results

3.3.1 NOESY spectrum of the myristoylated α X integrin in DPC micelles

The myristoylation strategy implemented in the α Mb2 to mimic the integrin transmembrane domain is also used in the structural characterization of the α X cytoplasmic tail. In the presence of DPC micelles, the ^1H - ^1H 2D NOESY of the myristoylated α X (Myr- α X) is as well dispersed as the myristoylated α M peptides, consisting of sharp resonances that spread from 6.5 to 9.0 ppm (Figure 3.2) Numerous NOEs between backbone/side chain, side chain/side chain protons were observed with sequential and medium range connectives (i+1, i+2, i+3, i+4) which can be identified and assigned, thus suggesting a helical conformation for Myr- α X in DPC micelles. The ring protons of the Myr- α X F⁴, F⁵ and Y⁹ are also well dispersed allowing for unambiguous identification.

Figure 3.2 ^1H - ^1H NOESY spectrum of Myr- α X in DPC micelles showing ^1HN connectivities to side chain protons, side chain and ring protons connectivities to other protons.

The topology of the Myr- α X in DPC micelles is similarly investigated with the paramagnetic Mn^{2+} ions that rapidly relaxes NMR FID of the solvent exposed backbone HN/HCA protons. V² to F⁵ demonstrated limited solvent exposure of approximately $\leq 50\%$ compared to the rest of the α X residues

(Figure 3.3), indicating that Myr- α X is anchored to the lipid core of the DPC micelles with V² to F⁵ closer to the DPC micelles and the rest of the residues exposed into the aqueous buffer environment.

Figure 3.3 Paramagnetic relaxation of the Myr- α X in DPC micelles, showing the relative change in intensity of each amino acid $^1\text{HN}/^1\text{H}\alpha$ NOE after addition of 1mM MnCl_2 .

3.3.2 Secondary structure analyses of the α X integrin

Secondary structure elements of the Myr- α X in DPC micelles are investigated in the absorption far UV CD spectroscopy. Positive maxima at 190 nm and two negative minima at 208 nm and 220 nm observed at the far UV range of the CD spectrum indicates that the Myr- α X adopts a stabilized helical conformation in the presence of DPC micelles (Figure 3.4).

Figure 3.4 CD spectrum of Myr- α X in DPC micelles showing the changes in Molar Ellipticity of the far UV CD range from 240 nm to 190 nm.

Further analyses of the individual amino acids $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ chemical shift deviations from random coil values demonstrated that the Myr- αX F⁴ to A¹⁶ showed positive $^{13}\text{C}\alpha$ deviations and negative $^1\text{H}\alpha$ deviations from denoting a helical segment at the N-terminal region (Figure 3.5) [155]. In contrast, Myr- αX N¹⁷ to K³⁵ which forms the C-terminal membrane distal regions did not deviate much from random coil values suggesting that it adopts a loop conformation. Characteristic $^1\text{H}\alpha/^1\text{HN}$ (i+1, i+2, i+3, i+4) and $^1\text{HN}/^1\text{HN}$ (i+1, i+2) NOEs for helical conformations were also observed from residues F⁴ to A¹⁶, in addition to a number of NOEs involving backbone/side chain and side chain/side chain protons. Unambiguous identification of F⁴, F⁵ and Y⁹ ring protons also suggest that they interact with each other and packs with the aliphatic side chains of V² (Figure 3.6 A). A few long range NOEs involving the Myr- αX C-terminal amino acids N¹⁷/T²⁸, A¹⁶/N²⁴ and N¹⁷/E²³ define the arrangement of the C-terminus loop (Figure 3.6 C, D and E). In general, the N-terminus membrane proximal region of the Myr- αX F⁴ – N¹⁷ forms a helical conformation followed by a loop that is packed with well defined contacts with the helical portion.

Figure 3.5 Chemical shift deviations from random coil values of the Myr- αX $^{13}\text{C}\alpha$ (Top panel) and $^1\text{H}\alpha$ (Bottom panel) chemical shift values.

3.3.3 Structural determination of the Myr- αX integrin

The structure of the Myr- αX in DPC micelles is elucidated based on distance restraints derived from observed NOE intensities and the predicted dihedral (Ψ , Φ) angles from $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ chemical shifts. The packing NOEs of F⁴, F⁵ and Y⁹ (Figure 3.6A and B) and long range NOEs (Table 3.2 and

Figure 3.6C, D and D) observed in the ^1H - ^1H NOESY spectrum were gradually input into the structures and refined to obtain the twenty lowest energy conformers.

Table 3.2 Table of long range NOEs used in structure determination		
1	Asn ¹⁷ NδH	Thr ²⁸ NH
2	Ala ¹⁶ CβH	Asn ²⁴ NδH
3	Glu ²³ CβH	Asn ¹⁷ NδH

Figure 3.6 Panels of the ^1H - ^1H NOESY of the Myr- αX showing packing and long range NOEs. **(A)** Packing NOEs between the aliphatic side chains of V² and F⁵. **(B)** Packing NOEs between F⁵ and Y⁹ ring protons. Long range NOEs between the **(C)** E²³C β H and N¹⁷N δ H; **(D)** M¹²C δ H and I²⁰NH; and **(E)** A¹⁶C β H and N²⁴N δ H.

The backbone superposition of the twenty lowest energy conformers is shown in Figure 3.6A, and the summary of structure statistics is listed in Table 3.3.

Table 3.3 Summary of structural statistics for the 20 final structures of Myr- α X

Structure	Myr- α X
Distance restraints	
Total	263
Intraresidue ($i - j = 0$)	74
Sequential ($ i - j = 1$)	90
Medium-range ($2 \leq i - j \leq 4$)	96
Long-range ($ i - j \geq 5$)	3
Angular restraints (Φ)	29
Angular restraints (Ψ)	28
Distance restraints violations	
Number of violations	1
Average violation ($\text{\AA}$)	0.25
Maximum violation ($\text{\AA}$)	≤ 0.31
Average target function values	3.86 (3.37-4.56)
Deviation from mean structure	
All backbone (Ca, C', and N) / (Residue 1-17) ($\text{\AA}$)	1.14 / (0.11)
All heavy atoms / (Residue 1-17) ($\text{\AA}$)	1.89 / (1.14)
Ramachandran plot for the mean structure	
% residues in the most favourable	82
% residues in the additionally allowed region	18
% residues in the disallowed region	0

The structure of the Myr- α X is manifested by a N-terminal amphipathic helical portion from residue F⁴ to N¹⁷, with one face of the helix comprising of polar/ionic residues R⁷, Q⁸, K¹⁰, E¹¹, E¹⁴ and E¹⁵ (Figure 3.7C) and the other comprising hydrophobic residues V², F⁵, Y⁹ and M¹³ (Figure 3.7B). The Myr- α X N-terminal helix is then followed by a well defined turn, consisting of Q¹⁹ to T²⁶ as indicated by the close superposition of the structural ensemble for these residues (Figure 3.7A). This turn is defined by long-range packing NOEs used in structural calculations, and is found in the final structure to bring P²⁹ and T²⁸ side chains close to the hydrophobic aromatic and aliphatic side chains of Y⁹ and M¹³ (Figure 3.7B). The C-terminal residues S³⁰ to K³⁵ mostly defined by sequential NOEs and are hence found to be the most disordered motif compared with the rest of the α X cytoplasmic tail.

Figure 3.7 NMR derived structure of the Myr- α X in DPC micelles. (A) Superposition of the top twenty lowest energy conformers. (B) Single structure of Myr- α X in ribbon representation, displaying residues residing at hydrophobic face of the N-terminal helix and the packing of C-terminal turn in spheres. (C) Single structure of the Myr- α X in ribbon and stick representation, with polar residues labeled residing at the hydrophilic face of the N-terminal residues.

3.3.4 Backbone amide resonance assignments of the α X cytoplasmic tail

The backbone amides of the double-labeled $^{15}\text{N}^{13}\text{C}$ α X are assigned with the two complementary triple resonance experiments HNCACB and CBCA(CO)NH. A stretch of α X cytoplasmic residues E²³ to Q²⁷ amide resonances identified with the sequence specific resonance assignments based on the $^{13}\text{C}\alpha/\text{C}\beta$ values of current and preceding residues are shown in Figure 3.8. β 2 amide resonances were previously assigned in section 2.3.7.

Figure 3.8 Strip plots of triple resonance experiments of αX . ^{15}N - ^1H strip plots of HNCACB (in red and green) and CBCA(CO)NH (in blue) showing the $^{13}\text{C}\alpha/\text{C}\beta$ values of αX cytoplasmic residues E²³ to Q²⁷. Horizontal lines correlates the $^{13}\text{C}\alpha/\text{C}\beta$ values of the preceding residues.

3.3.5 Interaction analyses of the $\alpha X\beta 2$ integrin

The cytoplasmic interaction interface of the $\alpha X\beta 2$ integrin is determined by the specific changes in combined ^1H and ^{15}N chemical shifts following the additions of their respective binding partners. ^{15}N αX or ^{15}N $\beta 2$ was titrated with unlabeled $\beta 2$ and αX respectively up to a molar ratio of 1:2 (Figure 3.9A and B), and individual ^{15}N - ^1H amide resonances of each backbone amino acid were recorded before and after titration and plotted as a function of amino acid to observe residue specific perturbation (Figure 3.9 C and D). The ^{15}N - ^1H

HSQC amide resonances of the α X and β 2 cytoplasmic tails was assigned with the aid of two complementary triple resonance experiments HNCACB and CBCA(CO)NH. The α X cytoplasmic tail residues significantly perturbed by β 2 were Y⁹, K¹⁰, E¹¹, M¹², M¹³, E¹⁴, E¹⁵ and N¹⁷ which comprise of the hydrophilic face of the N-terminal amphipathic helix (Figure 3.9 C). Interestingly, T²⁸ from the well defined loop in the Myr- α X structure is also perturbed by the addition of β 2 which suggest a possibility of the α X C-terminal loop to remodel in the presence of its binding partner. The reverse titration to determine the β 2 cytoplasmic residues involved in the interaction with α X revealed that β 2 cytoplasmic residues H⁵, L⁶, S⁷, D⁸, L⁹, E¹¹, Y¹² and R¹⁴ showed greater chemical shift changes compared to the rest of the β 2 cytoplasmic regions (Figure 3.9 D). β 2 H⁵ and R¹⁴ HSQC cross-peak intensity is also observed to be severely diminished (Figure 3.9 B) compared to the rest of the highly perturbed β 2 cytoplasmic residues, possibly as a result of conformational exchange between the free and bound states. The cytoplasmic residues of the α X and β 2 that are perturbed mostly comprise of the N-terminal membrane proximal regions.

Figure 3.9 Interaction analyses of the αX and $\beta 2$ cytoplasmic tails. (A) ^{15}N - ^1H HSQC of the αX cytoplasmic tail, contours in black is the spectrum of the free αX and contours in grey is the αX spectrum titrated with unlabeled $\beta 2$ in 1:2 molar ratio. (B) ^{15}N - ^1H HSQC of the $\beta 2$ cytoplasmic tail, contours in black is the spectrum of the free $\beta 2$ and contours in grey is the $\beta 2$ spectrum titrated with unlabeled αX in 1:2 molar ratio. (C) Combined chemical shift changes of ^{15}N and ^1H ($\Delta^{15}\text{N} + \Delta^1\text{H}$) for each amino acid of the αX cytoplasmic tail. (D) Combined chemical shift changes of ^{15}N and ^1H ($\Delta^{15}\text{N} + \Delta^1\text{H}$) for each amino acid of the $\beta 2$ cytoplasmic tail.

3.3.6 The binding interface between the $\alpha\text{X}\beta 2$ integrin

Based on the chemical shift changes obtained in the ^{15}N - ^1H HSQC titration experiments. The Myr- αX in DPC micelles structure elucidated in this study is placed close to the $\beta 2$ coordinates obtained from Bhunia, et al. (2009) [93] based on the residues of αX and $\beta 2$ showing significant chemical shift changes in the ^{15}N - ^1H HSQC. This starting point structure is then input into the RosettaDock server for rigid-body energetic optimization [146]. The docked cytoplasmic $\alpha\text{X}\beta 2$ structure assumes a parallel orientation at their N-terminal helical regions comprising of residues αX -K¹ to A¹⁶ and $\beta 2$ -K¹ to F¹⁵, followed by their C-terminal loops that do not participate in the $\alpha\text{X}\beta 2$ binding interface (Figure 3.10). The binding interface between the αX and $\beta 2$ cytoplasmic region at the N-terminal membrane proximal region is found to be sustained by hydrophobic packing interactions between the αX -F⁴ aromatic ring and the $\beta 2$

aliphatic alkyl chains of L³ and I⁴, this is followed by electrostatic interactions between α X residues (R⁷, K¹⁰, E¹¹) and β 2 residues (S⁸, E¹¹, R¹⁰) respectively. Multiple hydrogen bonds and/or salt bridges are also observed to be formed between α X residues (E¹⁴, E¹⁵) and β 2 residues (R¹⁰, R¹⁴) (Figure 3.10). In addition, α X-Q⁸ is found to be in close proximity with β 2-S⁷ (Figure 3.10).

Figure 3.10 Modeled complexes of the α X (in blue ribbon) and β 2 (in black ribbon) cytoplasmic tails. Hydrophobic residues are represented in spheres and the rest of the membrane proximal side chains are represented in sticks. The interactions between the membrane proximal domain are shown in the small panels with green dotted lines denoting distances close enough for ionic interactions.

3.4 Discussion

The α X cytoplasmic tail is 11 amino acids longer than the α M cytoplasmic tail and they share similar sequences at their N-terminal regions comprising of α X K¹⁽¹¹¹⁰⁾ – E¹⁵⁽¹¹²⁴⁾ and α M K¹⁽¹¹¹⁴⁾ – E¹⁵⁽¹¹²⁸⁾ (Figure 3.11 A and B). The characterization of Myr- α X in DPC micelles exhibited several similarities to the study of the α M peptides in Chapter II. Their ¹H-¹H NOESY showed similar width of dispersion of ¹HN/¹Ha resonances at the fingerprint region, and a well dispersed ring proton resonances that can be unambiguously identified. The N-terminal amphipathic helical membrane proximal regions are manifested in both Myr- α X and Myr- α M, both have their membrane proximal phenylalanines, tyrosine and methionines packing together at the hydrophobic face of the amphipathic helix and polar residues, namely α M-R⁷, K¹⁰, D¹¹, S¹⁴ and α X-R⁷, K¹⁰, E¹¹, E¹⁴ at the hydrophilic face of the helix (Figure 3.10 A and B). However, the Myr- α X assumes a longer helix than Myr- α M by two residues. The C-terminal loop of the Myr- α M is formed by non-polar packing between Y⁹⁽¹¹²²⁾, M¹²⁽¹¹²⁵⁾, M¹³⁽¹¹²⁶⁾ and Q²⁴⁽¹¹³⁷⁾, similarly the well defined turn of the Myr- α X consisting of Q¹⁹ to T²⁶ is also sustained by non-polar packing of Y⁹, M¹³, N¹⁷, and P²⁹ (Figure 3.10 A and B). The difference of the Myr- α X and Myr- α M C-terminal loop lies in the Myr- α X S³⁰-K³⁵ residues which forms a disordered loop that does not pack closely to the N-terminal helix. Structural similarities and disparities of the Myr- α X and Myr- α M structures can be drawn to existing structures of the α L and α IIb cytoplasmic structures (Figure 3.11 C and D). The N-terminal α L cytoplasmic regions comprising of K¹⁽¹⁰⁸⁷⁾ to K¹²⁽¹⁰⁹⁸⁾ assumes a limited helical conformation from K⁶⁽¹⁰⁹³⁾ to K¹²⁽¹⁰⁹⁸⁾, presumably due to the lack of transmembrane stabilization elements, followed by two helices that packs the α L cytoplasmic regions compactly (Figure 3.11 C). The shorter myristoylated α IIb cytoplasmic tail structure too assumes a limited helical N-terminal membrane proximal region from K⁹⁸⁹-N⁹⁹⁷ albeit with a shorter helix compared to the Myr- α X and Myr- α M structure, followed by a loop that packs to the helical portion via electrostatic interaction between K⁶⁽⁹⁹⁴⁾, D¹⁶⁽¹⁰⁰⁴⁾ and R⁹⁽⁹⁹⁷⁾, D¹⁵⁽¹⁰⁰³⁾ (Figure 3.11 D). The distinct cytoplasmic membrane distal loop conformations of the α X, α M, α L and α IIb probably suggest that they are perturbed by different intracellular factors which possibly

make use of non-polar and polar residues to disrupt loop packing for binding and hence regulation of integrin activation. It has also been shown that the mutation of membrane distal residues of the α Ib R⁹⁹⁷PP motifs to alanines rendered α Ib β 3 to be constitutively active [165], and this has been structurally and functionally characterized by Vinogradova et al, (2000) [91].

Figure 3.11 Single structures of the (A) Myr- α X in DPC micelles, (B) Myr- α M in DPC micelles, (C) α L cytoplasmic tail in solution and (D) Myr- α IIb in DPC micelles.

The interaction interface between the α X and β 2 cytoplasmic domains hold similarities to determined complexes of the α Ib β 3, α Lb2 and α Mb2, interactions between parallel membrane proximal helices sustained by electrostatic interactions with the canonical salt bridge between α -R and β -D (Figure 3.12). Longer α -subunit cytoplasmic sequences assume a longer helical

conformation that forms more electrostatic contacts with its binding partner. The sequence and structural conservation of residues involved in maintaining the resting state integrin found in the $\alpha X\beta 2$, $\alpha M\beta 2$, $\alpha L\beta 2$ and $\alpha IIb\beta 3$ in purport of functional mutation studies of the conserved membrane proximal residues [77,78] probably suggest that the resting state cytoplasmic clasp can be remodeled by membrane distal sequences or through the recruitment of intracellular effectors. Hence, the mechanistic understanding of the regulation of integrin activation at its cytoplasmic region can be furthered by the identification activating or deactivating cytoplasmic membrane distal mutations or specific intracellular effectors.

Figure 3.12 Complex structures of the (A) α IIb β 3 cytoplasmic tails, (B) α L β 2 cytoplasmic tails and (C) α M β 2 cytoplasmic tails.

Chapter IV: Structures and interactions of the $\alpha 4$ cytoplasmic tail and paxillin

4.1 Overview

The $\alpha 4\beta 1$ integrin, also known as the Very Late Antigen-4 (VLA4) or CD49D/CD29 is predominantly expressed in lymphoid cells, neural crest cells and in developing skeletal muscle. Its presence in differentiated lymphoid cells and nascent cells denote its dual roles in immune responses and development. The $\alpha 4\beta 1$ integrin dual functionality has been demonstrated in engineered mouse mutants, where the complete knockout of the $\alpha 4$ -subunit resulted in embryonic lethality [166] and in a mouse mutant with a single amino acid mutation at its cytoplasmic domain Y⁹⁹¹A created viable and fertile mice but with impaired mononuclear cell recruitment during induced inflammation [118]. The $\alpha 4\beta 1$ integrin regulates cell migration, cytoskeletal organization and gene expression differently from other α -subunits and are hence indispensable for embryogenesis, hematopoiesis and immune responses. The $\alpha 4\beta 1$ integrin is generally involved in cellular migratory processes and is found to be important in neural crest cell migration and leukocyte migration. Its cytoplasmic domains has been found to be implicated in its migratory properties [18], where the replacement of the $\alpha \text{IIb}\beta 3$ cytoplasmic domains with the $\alpha 4\beta 1$ cytoplasmic domains in immune cells resulted in increased cell migration and $\alpha \text{IIb}\beta 3$ dependent phosphorylation of Focal adhesion kinase (FAK), and at the same time reduced cell spreading and focal adhesion [18]. As such, aberrant regulation of $\alpha 4\beta 1$ integrins usually results in abnormal migratory behaviours of leukocytes found in many autoimmune diseases such as multiple sclerosis, asthma, rheumatoid arthritis, Crohn's disease, etc [167]. It is also implicated in cancer metastasis where it is also found to be expressed in its high affinity conformation on highly metastatic human melanoma cells and facilitate melanoma cell migration [168].

The $\alpha 4\beta 1$ integrin plays an integral role together with the $\beta 2$ integrins in leukocyte migratory processes during the trigger of an immunological event. Its extracellular ligands includes fibronectin, Vascular-cell adhesion molecule-1

(VCAM-1) and Mucosal Addressin Cell adhesion molecule-1 (MAdCAM-1) found on the extracellular matrix and other cells. Similar to the $\beta 2$ integrins, the affinity for extracellular ligands of the $\alpha 4\beta 1$ integrin is regulated by conformation changes induced by inside-out signaling mechanisms [169]. Unique sequences at its cytoplasmic domains recruit intracellular effectors which remodel the actin cytoskeleton for migratory processes. Indeed, the cytoplasmic domain of the $\alpha 4$ integrin subunit has been found to promote cell migration and antagonize cell spreading and focal adhesion formation when joined to the αIIb - subunit [170]. The migratory function has been attributed to the direct binding of the $\alpha 4$ cytoplasmic tail with its intracellular ligand, paxillin. This interaction is only unique to the $\alpha 4$ cytoplasmic tail and not to αIIb , $\alpha 3$, $\alpha 5$ and $\alpha 6$ cytoplasmic domains [18]. In addition, only a short sequence S⁹⁷⁸ to I⁹⁹² in the $\alpha 4$ cytoplasmic is found to directly mediate its interaction with paxillin [115]. Paxillin is a 68 kDa adaptor protein that connects the $\alpha 4\beta 1$ integrin to proteins that modulates the actin cytoskeleton. The ability of paxillin to bind to diverse intracellular targets mediating structural and signaling functions lies in its multi-domain module that function as unique recognition sequences for intracellular proteins. It comprises of five leucine-rich LD motifs with the consensus sequence (LDXLLLLXXL) and several proline and glycine rich regions, which has characteristic of SH3-binding motifs at its N-terminal halve, followed by four double zinc finger Lin-11, Isl-1, Mec-3 (LIM) domains at its C-terminus (Figure 4.1) [171]. Serine, threonine and tyrosine phosphorylation sites are found to exist throughout the paxillin molecule (Figure 4.1).

Figure 4.1 Domain organization of the paxillin molecule and its binding partners. LD repeats 1 to 5 are found near the N-terminus of the molecule followed by the four LIM domains. Possible sites of Ser (PS) and Tyr (PY) phosphorylation are denoted in circles. Taken from Brown MC, Turner CE: *Paxillin: Adapting to change. Physiol Rev* (2004) 84(4):1315-1339.

Paxillin is found to directly interact with intracellular ligands such as, FAK, pp60, c-Src, PTP-PEST, Crk, p95 PKL, PIX and PAK, which have been found to be associated in the regulation of cell migration, cytoskeletal organization and gene expression [171]. Hence, its interaction with the $\alpha 4$ cytoplasmic domain regulates the migratory properties of the cells it is present in. Paxillin interacts with $\alpha 4$ via a 100 amino acid fragment consisting of LD2, LD3 and LD4 repeats [172]. This $\alpha 4$ -paxillin interaction is found to be regulated by phosphorylation at the $\alpha 4$ cytoplasmic domain at S⁹⁸⁸ which is found in its short sequence (S⁹⁷⁸ to I⁹⁹²) of interaction with paxillin. A phosphorylation-mimic mutant S⁹⁸⁸D has been found to blocked paxillin binding that inhibited cell migration and promoted cell spreading. Additionally, mutation of S⁹⁸⁸ to the non-phosphorylatable amino acid alanine also inhibited cell migration and further reduced cell spreading [116]. The S⁹⁸⁸A non-phosphorylatable mutant possibly resulted in constitutive interaction with paxillin and thus the reversible interaction of $\alpha 4$ -paxillin regulated by phosphorylation is responsible for cellular migratory behaviour [116]. The spatial location of phosphorylated and non-phosphorylated $\alpha 4$ -S⁹⁸⁸ has been found to be accumulated along the leading edge and lateral edges of migrating cells respectively, where paxillin associates with the non-phosphorylated $\alpha 4$

cytoplasmic tails [173]. The specific spatial localization of the $\alpha 4$ -paxillin complex, regulated by phosphorylation, in the leading and lateral edges of the migrating cell control cellular migration and polarization by spatial discrimination in the formation of actin scaffolds necessary for migration (Figure 4.2) [3].

Figure 4.2 Schematic diagram of the regulation of $\alpha 4$ association with paxillin via phosphorylation in a migrating cell. Taken from *Rose DM, Alon R, Ginsberg MH: Integrin modulation and signaling in leukocyte adhesion and migration. Immunol Rev (2007) 218(126-134).*

Thus far, the structure of the $\alpha 4$ cytoplasmic tail and its interaction with paxillin has not been well characterized. Hence, we would be defining the structure of the $\alpha 4$ cytoplasmic domain and mapping its interaction with the 100 amino acid region in paxillin consisting of the three LD repeats. The $\alpha 4$ cytoplasmic domain unlike the α -subunits of the $\beta 2$ integrins does not contain any prolines and would therefore assume a different structural conformation that would enable it to recruit distinct sets of intracellular effectors.

4.2 Materials and Methods

4.2.1 Peptides and Proteins

The sequence of $\alpha 4$ cytoplasmic tail from K⁹⁶⁸ to D⁹⁹⁹ will be renumbered from one for ease of reference (Table 4.1). Additionally, the paxillin region G¹³⁹ to F²⁷⁷ will be abbreviated P234 and renumbered from one.

The peptides of the synthetic $\alpha 4$, Myr- $\alpha 4$, N-terminal rhodamine labeled $\alpha 4$ (Rho- $\alpha 4$) and the paxillin leucine-rich repeats (LD) 2, 3, 4 listed on table 4.1 were purchased from GL Biochem (Shanghai, China).

Table 4.1 List of the sequences peptides and proteins used	
Synthetic $\alpha 4$	⁹⁶⁸⁽¹⁾ KAGFFKRQYKSILQEENRRDSWSYINSKSND ⁹⁹⁹
Synthetic Myr- $\alpha 4$	Myristoylated- ⁹⁶⁸⁽¹⁾ KAGFFKRQYKSILQEENRRDSWSYINSKSND ⁹⁹⁹
Synthetic Rho- $\alpha 4$	Rhodamine- ⁹⁶⁸⁽¹⁾ KAGFFKRQYKSILQEENRRDSWSYINSKSND ⁹⁹⁹
Purified ¹⁵ N/ ¹⁵ N ¹³ C $\alpha 4$	M* ⁹⁶⁸⁽¹⁾ KAGFFKRQYKSILQEENRRDSWSYINSKSND ⁹⁹⁹
Purified paxillin (P234) unlabeled/ ¹⁵ N/ ¹⁵ N ¹³ C	¹³⁹⁽¹⁾ MGSNLSELDRLLELNAVQHNPFGPADEANSSPPLPGALS PLYGVPETNSPLGGKAGPLTKEKPKRNGGRGLEDVRPSVESLL <u>DELESSVPSPVPAITVNQEMSSPQRTSTQQQTRISASSATREL</u> <u>DELMASLSDFKF</u> ²⁷⁷
LD2	¹⁴¹⁽¹⁾ NLSELDRLLELNAVQHN ¹⁵⁸
LD3	²¹³⁽¹⁾ VRPSVESLLDELESSVPSPV ²³²
LD4	²⁶²⁽¹⁾ ATRELDLMASLSDFKF ²⁷⁷ MAQ

Underlined sequences are the Leucine-rich repeats.

* Additional sequences that is removed at the final stage of purification

4.2.2 Expression and purification of the $\alpha 4$ cytoplasmic tail and paxillin

The $\alpha 4$ cytoplasmic regions K⁹⁶⁸ to D⁹⁹⁹ is cloned into the pET-31b(+) vector (Novagen EMD, San Diego) described previously in Bhunia et. al. (2009). It has a Met cleavage site inserted prior to the $\alpha 4$ cytoplasmic sequence.

The recombinant plasmid is transformed into BL21(DE3) cells. $\alpha 4$ was purified using the same method to purify $\beta 2$ as described in chapter 2.2.2.

The paxillin protein region G¹³⁹ to F²⁷⁷ was cloned into the pET24a(+) vector, which will include a C-terminal 6x-His tag that will aid in protein purification. The recombinant vector is first transformed into BL21(DE3) cells. Depending on whether unlabeled, singly labeled with ¹⁵N- or doubly labeled with ¹⁵N¹³C- was needed, the overgrown culture was seeded in 1:100 volume ratio in LB for unlabeled proteins or isotope-enriched minimal media (M9), containing ¹⁵N ammonium chloride without/with ¹³C-glucose for labeled proteins. The seeded culture was then placed in a 37 °C shaking incubator at 150 rpm until an optical density (O.D) of 0.5 was reached. Isopropyl β -D-1-thiogalactopyranoside (IPTG) at 0.5 mM/litre of media was then used to initiate plasmid protein expression. The IPTG induced culture is then left to grow overnight at 18 °C for 18 hours. The induced cells were then centrifuged at 5000 rpm, 4°C, for 20 mins. The pelleted cells were then resuspended in 20 mM Tris-Cl, 0.5 M NaCl buffer pH 8.0 and lysed via sonication on ice to release the recombinant proteins. The cell suspension was later clarified from cell debris by centrifuging at 14 000 rpm, 4°C, for 30 minutes. The supernatant was then incubated with Nickel-NitriloTriAcetic (NTA) acid (QIAGEN) beads at 4°C on a roller. The proteins bound to the Nickel-NTA beads were then washed with 20 mM Tris-Cl, 0.5 M NaCl and 20 mM imidazole pH 8.0 twice and eluted with 20 mM Tris-Cl, 0.5 M NaCl and 300 mM imidazole pH 8.0. The eluates were dialysed in 20 mM Tris-Cl, 0.15 M NaCl, pH 7.0 for 1 hour at room temperature. The dialysate is then passed through the HPLC machine connected to a C18 column at a flow rate of 2 ml/min. A gradient of 10%/90% Acetonitrile-water was used to elute the protein.

The HPLC profiles of the purified $\alpha 4$ cytoplasmic tail and P234 are shown in Appendix 3.

4.2.3 NMR experiments

All NMR experiments are spectra were recorded on a Bruker DRX 600-MHz instrument equipped with an actively shielded cryo-probe unless stated

otherwise. 10% Deuterium oxide and 2 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added to all NMR samples. All spectra were directly/indirectly referenced to DSS 1H.

1D-NMR analysis of $\alpha 4$ peptides: 1D-NMR spectra of 0.5 mM Myr- $\alpha 4$ dissolved in 200 mM, 10 mM, 2 mM, and 1.2 mM of deuterated DPC (CIL, Massachusetts, USA) were respectively recorded at 308 K. In addition, the 1D-NMR spectra of 0.5 mM synthetic $\alpha 4$ cytoplasmic tails dissolved in 200 mM of deuterated DPC was also recorded in 308 K. The synthetic $\alpha 4$ dissolved in water pH 5.6 at 278 K was also recorded on a Bruker DRX 700-MHz instrument.

NMR derived structure for the $\alpha 4$ cytoplasmic tail at 5°C: 2D-TOCSY and 2D NOESY were recorded for 0.5 mM of $\alpha 4$ was dissolved in water, pH 5.6 at 278 K. TOCSY mixing times were fixed at 50 ms and NOESY mixing times were maintained at 200 ms. The ^{13}C - ^1H Heteronuclear Single Quantum Coherence (HSQC) experiment was also carried out for the three peptides at 100% D₂O to obtain $^{13}\text{C}\alpha$ chemical shifts. Structure calculation, NMR data processing and analysis conducted on $\alpha 4$ is the same as described in chapter 2.2.4.

Backbone resonance assignments for $\alpha 4$ and P234: Two complementary sets of the triple backbone resonance experiments, the HNCACB which records the inter- and intra amide backbones of the current and preceding amino acids' $\text{C}\alpha$ and $\text{C}\beta$ resonances; and the CBCA(CO)NH which records the preceding amino acids inter-amide backbones $\text{C}\alpha$ and $\text{C}\beta$ resonances were used to achieve the backbone amide resonances of $^{15}\text{N}^{13}\text{C}$ - $\alpha 4$ and $^{15}\text{N}^{13}\text{C}$ -paxillin dissolved in 10 mM sodium phosphate buffer, pH 5.6 and 10 mM sodium phosphate buffer, with 50 mM L-Arg and 50 mM L-Glu pH 5.6 respectively. The 10 mM sodium phosphate buffer, with 50 mM L-Arg and 50 mM L-Glu pH 5.6 was used to dissolve the paxillin to achieve a higher concentration without significantly affecting the chemical shifts of the amide resonances [174].

NMR titration studies of ^{15}N -paxillin with unlabeled $\alpha 4$ cytoplasmic tails:

200 μM of ^{15}N -paxillin was dissolved in various buffers of 10 mM sodium phosphate buffer, pH 5.6; 10 mM MOPS buffer, pH 6.4; 10 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer with 50 mM L-Arg, pH 6.4; and 10 mM MOPS buffer with 50 mM L-Arg and 50 mM L-Glu, pH 6.4. ^{15}N - ^1H HSQC spectrum of the free ^{15}N -paxillin and unlabeled $\alpha 4$ spectrum were recorded for each titration occurring at the different buffers.

NMR titration studies of ^{15}N - $\alpha 4$ with unlabeled P234 and LD repeats:

The residues perturbed in the ^{15}N - $\alpha 4$ were determined in a series of titration with P234, LD2, LD3 and LD4 conducted in 10 mM sodium phosphate buffer, pH 6.0 at 298 K. The changes in chemical shift and signal intensity of ^1HN and ^{15}N of the backbone amides were recorded for each set of titration, where the intensity of each amino acid is normalized to the un-titrated ^{15}N - $\alpha 4$ spectrum, and the combined chemical shift changes of ^1HN and ^{15}N ($\Delta\text{H}+\Delta\text{N}$) from the un-titrated ^{15}N - $\alpha 4$ spectrum were recorded for each titration point. For every titration, 0.1 mM ^{15}N - $\alpha 4$ was used and titrated up to 0.75 mM for P234; and up to 1 mM for LD2, 3, and 4.

4.2.4 Dissociation constant determination using Rhodamine- $\alpha 4$ fluorescence titration

1 μM of Rho- $\alpha 4$ was dissolved in 10 mM sodium phosphate buffer, pH 6.0 and titrated with P234, LD2, LD3 and LD4 respectively up to a concentration of 8 μM . The titration experiment was conducted using a 1 cm path length luminescence spectroscopy cuvette (Perkin-Elmer, California, USA) on a Varian Cary Eclipse spectrophotometer. The sample was excited at 544 nm with 5 nm slit width, and the emission of the rhodamine labeled $\alpha 4$ (Rho- $\alpha 4$) was collected from 550 nm to 650 nm at 10 nm slit width. The quenching of the N-terminal rhodamine was recorded and normalized to the intensity of the final titration point for each experiment. Two sets of intensity readings for each titrant was taken and normalized to the initial fluorescence intensity value. The normalized F_{max} for both sets of were first averaged before being plotted as a function of concentration of the dansylated peptides. A non-

linear regression function in Origin7.0 was used to determine the dissociation constants, K_d .

$$\text{Normalized } \Delta F_{\text{max}} = \frac{F_{\text{max}} \cdot [\alpha M]}{K_d + [\alpha M]}$$

4.3 Results

4.3.1 Analysis of the $\alpha 4$ peptides at different conditions

Apart from the conserved membrane proximal KxGFFKR motif, the $\alpha 4$ cytoplasmic tail sequence is entirely unique from other α -subunits of the α IIb β 3 and $\beta 2$ integrins. It is a short 32 amino acid sequence which does not contain any prolines (Figure 4.3), consequently it is expected to adopt a different structural conformation compared to the rest of the α -subunit cytoplasmic tails but could possibly maintain similar electrostatic interactions with its $\beta 1$ counterpart gauging based on the conserved sequences in the β -subunits (Figure 4.3).

```

 $\alpha 4$  968(1) K A G F F K R Q Y K S I L Q E E N R R D S W S Y I N S K S N D D
 $\alpha L$  1087 K V G F F K R N L K E K M E A G R G V P N G I P A E D S E Q L A S G Q E A G D P G C L K P L H E K D S E S G G
 $\alpha M$  1114 K L G F F K R Q Y K D M M S E G G P P G A E P Q
 $\alpha X$  1110 K V G F F K R Q Y K E M M E E A N G Q I A P E N G T Q T P S P P S E K
 $\alpha D$  1108 K L G F F K R H Y K E M L E D K P E D T A T F S G D D F S C V A P N V P L S
 $\alpha I I b$  989 K V G F F K R N R P P L E E D D E E G E

 $\beta 1$  742 K L L M I I H D R R E F A K F E K E K M N A K W D T G E N P I Y K S A V T T V V N P K Y E G K
 $\beta 2$  702 K A L I H L S D L R E Y R R F E K E K L K S Q W N N D . N P L F K S A T T T V M N P K F A E S
 $\beta 3$  716 K L L I T I H D R K E F A K F E E E R A R A K W D T A N N P L Y K E A T S T F T N I T Y R G T

```

Figure 4.3 Sequence alignment of the α - subunit and β subunit cytoplasmic domains.

The $\alpha 4$ cytoplasmic tail exhibited different characteristics from the αM and αX cytoplasmic tails. Unlike the myristoylated αM and αX cytoplasmic peptides, the Myr- $\alpha 4$ peptides in DPC micelles have a shorter dispersion of resonances from 6.6 to 8.5 ppm (Figure 4.4 C to F). The resonances were also broadened, in addition to the lack of observation or broadened tryptophan (W^{22}) ϵ proton at around 10 ppm suggest that the Myr- $\alpha 4$ might have aggregated in the presence of DPC micelles, with the tryptophan (W^{22}) buried into the micelles. The decrease in DPC micelle concentration did not alleviate the NMR-signal line broadening or help widen the range of dispersion, but instead decreased its solubility in DPC micelles. This could possibly due to the increased aggregating behaviour due to the decreased number of DPC micelles for each Myr- $\alpha 4$. In the case of the Myr- $\alpha 4$, membrane tethering into the DPC micelles did not occur in a similar fashion as the myristoylated αM or αX rendering the Myr- $\alpha 4$ unsuitable for further structural characterization. Instead, the unmodified $\alpha 4$ dissolved in water enabled NMR signals with narrow line-widths to be obtained with a wider dispersion of low field resonances scattered

from 6.5 to 8.9 ppm (Figure 4.4 C). In addition, dissolving the unmodified $\alpha 4$ in DPC micelles also yielded sharp NMR signals but only slightly improved the signal dispersion (Figure 4.4 A). The different 1D proton NMR profile of the free $\alpha 4$ in water and DPC micelles suggest that different structures are formed. Additionally, the $\alpha 4$ - $W^{22}He$ could possibly be interacting with the DPC micelles judging from the downfield shift of the 1He compared to the $\alpha 4$ in water.

Figure 4.4 1D-NMR spectra of the Myr- $\alpha 4$ and $\alpha 4$ synthetic peptides in various conditions; (A) $\alpha 4$ in 200 mM DPC, 308 K, (B) $\alpha 4$ in water at 278K, (C) Myr- $\alpha 4$ in 1.2 mM DPC, 308 K, (D) Myr- $\alpha 4$ in 2 mM DPC, 308 K, (E) Myr- $\alpha 4$ in 10 mM DPC, 308 K and (F) Myr- $\alpha 4$ in 200 mM DPC, 308 K.

4.3.2 Structural characterization of the $\alpha 4$ cytoplasmic tail in water

The relatively wide chemical shift dispersion of the free $\alpha 4$ cytoplasmic tail in water at 278K enabled further structural characterization. No long range NOEs were observed in the NOESY spectrum, however, mostly sequential and medium range (i+1, i+2, i+3 and i+4) NOEs were observed. A resonance dispersion for amide protons of 6.8 to 9.0 ppm suggested that the free $\alpha 4$ in 278 K may assume folded structures (Figure 4.5). The side chains of the F4, F5, Y9 and Y24 were also clearly demarcated from 6.8 to 7.7 ppm (Figure 4.5),

however, ambiguity among these resonances precluded packing NOEs to be unambiguously assigned and were not used for structural calculations.

Figure 4.5 ^1H - ^1H NOESY spectrum of the $\alpha 4$ cytoplasmic tails in water at 278 K.

The chemical shift deviation of the $^1\text{H}\alpha$ and $^{13}\text{C}\alpha$ from random coil values demonstrated negative and positive deviations respectively for residues Q^{14} to S^{27} , connoting a stabilized helical stretch for this region (Figure 4.6) [149].

Figure 4.6 Chemical shift deviations of the $^1\text{H}\alpha$ (Top panel) and $^{13}\text{C}\alpha$ (Bottom panel) from random coil.

The dihedral angle values of Φ (-120° to -30°) and Ψ (-120° to 120°) values were used for the $\alpha 4$ residues together with the unambiguously assigned NOEs in the NOESY spectrum for structural characterization. The structural

statistics of the twenty lowest energy structures obtained are shown in Table 4.2.

Table 4.2 Summary of structural statistics for the 20 final structures of $\alpha 4$ in water, 278K

Structure	$\alpha 4$
Distance restraints	
Total	300
Intraresidue ($i - j = 0$)	107
Sequential ($ i - j = 1$)	95
Medium-range ($2 \leq i - j \leq 4$)	98
Long-range ($ i - j \geq 5$)	0
Angular restraints (Φ)	31
Angular restraints (Ψ)	31
Distance restraints violations	
Number of violations	1
Average violation ($\text{\AA}$)	0.24
Maximum violation ($\text{\AA}$)	≤ 0.24
Average target function values	1.18 (0.92 – 1.44)
Deviation from mean structure	
All backbone (C α , C', and N)/($\text{\AA}$)	1.15
All heavy atoms/($\text{\AA}$)	1.86
Ramachandran plot for the mean structure	
% residues in the most favourable	93
% residues in the additionally allowed region	7
% residues in the disallowed region	0

Superposition of the twenty lowest energy structures, showed that the structure is quite flexible at its N-terminal, possibly due to the absence of a transmembrane stabilization factor (Figure 4.7 A). The structure of the free $\alpha 4$ cytoplasmic tail assumes a random coil conformation at the membrane proximal region from K¹ to Y⁹ followed by a helical stretch from K¹⁰ to D³². Unlike the αM and αX cytoplasmic regions, the $\alpha 4$ does not adopt an amphipathic helical conformation followed by a loop. The $\alpha 4$ cytoplasmic sequence does not consist of prolines, in addition, it only consist of a few hydrophobic residues, I¹², L¹³, Y²⁴ and I²⁵ amidst the aplenty acidic and basic amino acid residues (Figure 4.7 B), which probably attributed to its non-amphipathic helical nature.

Figure 4.7 Structures of the free $\alpha 4$ cytoplasmic tail in water, 278K. **(A)** Superposition of the twenty lowest energy structures. **(B)** Ribbon representation of a single structure with side chains represented as sticks and hydrophobic residues in spheres.

The electrostatic potential map of the $\alpha 4$ cytoplasmic tails showed that the N-terminus is predominated by basic residues from $K^1 - Q^{14}$, $R^{18/19}$ followed by a small negative patch consisting of the acidic residues $E^{15/16}$ and D^{20} (Figure 4.8). The difference in cytoplasmic conformation from the αM and αX cytoplasmic tail probably dictates binding to a different subcellular set of ligands.

Figure 4.8 Electrostatic potential maps of $\alpha 4$ cytoplasmic tail showing the front and back of the molecule.

4.3.3 Solubility of the P234 and $\alpha 4$ interaction

The LD repeats in paxillin play an important role in the interaction with the $\alpha 4$ cytoplasmic tail. Hence, the paxillin construct, P234, consisting of the LD 2,3 and 4 was used to determine its interaction with the $\alpha 4$ cytoplasmic tail. ^{15}N - ^1H HSQC titration of ^{15}N -P234 with $\alpha 4$ resulted in complete loss in backbone amide signal intensity of many amino acids and those that remain after the titration demonstrated reduced peak intensity and minor chemical shift changes. These are indicative of that the P234 and $\alpha 4$ interaction occurs at the intermediate ($K = 100 \text{ s}^{-1}$) exchange time scale [149]. Different buffers were used to determine this interaction and it was observed that different titration ratios of P234 to $\alpha 4$ is achieved in the different buffer compositions before sample precipitation and severe loss of signal intensities (Figure 4.9). In 10 mM MOPS buffer, 200 μM of ^{15}N -P234 was titrated up to 300 μM unlabeled $\alpha 4$ (Figure 4.9 A). Further changing buffer conditions to include 50 mM of L-Arg (Figure 4.9 B) or 50 mM L-Arg/L-Glu (Figure 4.9 C) to increase solubility of the complex, yet was found to reduce the titration ratio of the P234- $\alpha 4$ interaction to 1:1. Although L-Arg [174] and L-Glu has been reported to

drastically increase protein solubility without affecting the chemical shift of the NMR signal [175], in the P234- $\alpha 4$ interaction it seems to shift the chemical equilibrium such that a lower $\alpha 4$ concentration could cause precipitation. Hence, a simplified buffer consisting of only 10 mM sodium phosphate buffer was used and a titration ratio of 1:2 can be achieved, without much loss in NMR signal intensity in order to facilitate the characterization of the P234 interaction interface with $\alpha 4$.

Figure 4.9 Titration of ^{15}N -P234 with unlabeled $\alpha 4$ cytoplasmic tail in different buffers. Black peaks represent the untitrated spectrum while the red peaks represent the titrated spectrum. (A) 10 mM MOPS buffer pH 6.4, (B) 10 mM MOPS buffer pH 6.4 with 50 mM L-Arg, (C) 10 mM MOPS buffer pH 6.4 with 50 mM L-Arg and 50 mM L-Glu, and (D) 10 mM sodium phosphate buffer, pH 6.0.

Backbone amide assignments of the double-labeled $^{15}\text{N}^{13}\text{C}$ -P234 and $\alpha 4$ cytoplasmic tail were accomplished with the triple resonance experiments HNCACB and CBCA(CO)NH. The resonance assignments based on the $^{13}\text{C}\alpha/\text{C}\beta$ values of the current and preceding amino acids of a segment of P234 consisting of residues N⁴ to L⁸ is shown in Figure 4.10.

Figure 4.10 Overlay of HNCACB (red and green peaks) and CBCA(CO)NH (blue peaks) NH strip plots of N⁴, L⁵, S⁶, E⁷ and L⁸ of double-labeled ¹⁵N¹³C-P234. The triple resonance spectra for P234 are analyzed using CcpNMR analysis.

Subsequent analysis of the ¹³C α deviations from random coil values showed that the P234 mostly assumes a random coil structure, with residues that comprise the LD2 and LD4 showing positive deviations indicating that these regions have a tendency to form helical structures (Figure 4.11).

Figure 4.11 $^{13}\text{C}\alpha$ chemical shift deviations of P234 from random coil values. LD repeats are underscored with grey bars.

4.3.4 Interaction analyses between the P234 and $\alpha 4$

The interaction interface of the ^{15}N -P234 with $\alpha 4$ was analyzed based on the titration data obtained from ^{15}N -P234 and $\alpha 4$ interaction in 10 mM sodium phosphate buffer (Figure 4.12). Only at the final titration ratio, significant chemical shift and intensity changes in the backbone amides were observed (See Appendix 4). Additionally, some previously absent peaks were observed. The greatly reduced peak intensities with subtle chemical shift changes for a portion of amide peaks point out that these residues are undergoing exchange between the bound and unbound form of the P234 and $\alpha 4$ interaction at an intermediate exchange time scale (Figure 4.12). The appearance of new peaks, which were unidentifiable, suggest that some of the residues are undergoing slow exchange and also predicting a relatively tight binding interaction, in the micro-molar range, between P234 and the $\alpha 4$ cytoplasmic tail.

Analyses of the combined chemical shift changes in amide ^1H and ^{15}N ($\Delta\text{N}+\Delta\text{H}$) showed that most of the residues with ≥ 10 Hz changes falls in the LD2, 3 and 4 regions (Figure 4.13), with some other residues E²⁹, G³⁸, G⁵⁵, L⁶⁰, K⁶², R⁶⁷, N⁶⁸, V¹⁰⁰, N¹⁰¹ and L¹⁴¹ also showing significant changes. The residues perturbed in the LD2, LD3 and LD4 regions mostly involve hydrophobic residues, such as leucine, valine and alanine; and uncharged polar residues such as serine and glutamine. In the LD3 region, negatively charged amino acids E^{74/81/86} also showed significant chemical shift perturbations in addition to the hydrophobic and polar amino acids. The residues perturbed fall in the LD2, 3, 4 and other regions possibly suggest direct interaction with $\alpha 4$ and structural re-arrangement.

Figure 4.13 Combined chemical shift changes ($\Delta\text{N}+\Delta\text{H}$) of each amino acid residues of P234 following titration with unlabeled $\alpha 4$ at a ratio of 1:2. The LD repeats are underscored with grey bars and amino acids showing ≥ 10 Hz changes are denoted in black and grey letters. Amino acids denoted in grey letters form part of the LD repeats.

The addition of $\alpha 4$ to the ^{15}N -P234 at the final titration ratio of 1:2 (P234: $\alpha 4$) greatly diminished the NMR signal intensity to ≤ 0.25 as observed from the relative change in intensity for each backbone amide (Figure 4.14) with a few peaks N¹⁶, A²⁷, S¹¹³, R¹¹⁹ and D¹³⁸ disappearing completely. The

greatly diminished NMR signal intensity suggests that the interaction occurs in an intermediate exchange rate.

Figure 4.14 Relative changes in intensity of each amino acid residues of P234 following titration with unlabeled $\alpha 4$ at a ratio of 1:2. The LD repeats are underscored with grey bars and amino acids showing complete loss of intensity denoted.

4.3.5 Interaction analyses between the $\alpha 4$ with P234 and LD repeats

Interaction analyses of the ^{15}N - $\alpha 4$ cytoplasmic domain with P234 and the LD repeats also showed similar characteristics as the ^{15}N -P234 titration with $\alpha 4$, with the final titration ratio showing the greatest changes in terms of combined chemical shift and relative intensity changes. The reverse titration of ^{15}N - $\alpha 4$ cytoplasmic tail with the unlabeled P234 to determine the corresponding interaction interface revealed that residues in the $\alpha 4$ region from Q⁸ – R¹⁹ and S²³ – N³⁰ demonstrated combined chemical shift changes of greater than 10 Hz (Figure 4.15 A). Additionally, the relative intensity changes showed reduced relative intensity ≤ 0.5 for the residues Q⁸ – R¹⁹, however, residues W²² – N²⁶ which comprise of the S²³ – N³⁰ segment displayed protection from signal diminution with relative intensities ≥ 0.5 (Figure 4.15 A). Interestingly, G³ did not exhibit any change in combined chemical shift or relative intensity. In order

to demarcate the $\alpha 4$ residues that interact with the specific LD segments in paxillin. ^{15}N - $\alpha 4$ cytoplasmic tails were titrated against LD2, 3 and 4 respectively. Titration with LD2, LD3 and LD4 demonstrated lower combined chemical shift changes compared to P234, and intensity changes were distinct from the full length P234 (Figure 4.15 B, C and D). The interaction with each LD repeats were also distinct from each other, with some residues such as Q⁸, S¹¹ and R¹⁸ were consistently perturbed in all titration with P234, LD2, 3 and 4.

The $\alpha 4$ cytoplasmic tail titration with LD2 showed that residues, Q⁸, S¹¹, L¹³, N¹⁷, R¹⁸, W²², Y²⁴, N²⁶, K²⁸ and N³⁰ were more perturbed compared to the rest of the residues (Figure 4.15 B). Most of the $\alpha 4$ cytoplasmic residues with the exception of A² and Y⁹ showed a decrease in NMR signal intensity of $\approx 50\%$. The combined chemical shift changes of A² and Y⁹ did not show much changes and are most probably not involved in the interaction with LD2 (Figure 4.15 B). Titration with the LD3 repeat also showed similar residues, Q⁸, S¹¹, L¹³, R¹⁸, N²⁶ and N³⁰ perturbed, in addition to other residues Y⁹, N¹⁷, S²¹, Y²⁴, S^{27/29} and D³¹ (Figure 4.15 C). The combined chemical shift changes were generally higher when LD3 was added compared to LD2 (Figure 4.15 C). Relative intensity diminution across the $\alpha 4$ cytoplasmic residues was generally maintained at $\approx 50\%$ with A², Q¹⁴ and R¹⁹ relatively unchanged (Figure 4.15 C). Last but not least, titration of the $\alpha 4$ cytoplasmic tail with LD4 showed lower combined chemical shift changes compared to LD2 and LD3, with fewer amino acids, Q⁸, N¹⁷, R^{18/19}, S²⁷ and D^{31/32} showing relatively significant chemical shift changes (Figure 4.15 D). Further, titration with LD4 showed a greater diminution in relative intensity of $\approx 70\%$ in most residues except G³, F⁴, Y²⁴ and I²⁵ (Figure 4.15 D). In general, the relative intensity is decreased across most $\alpha 4$ cytoplasmic residues with a few exceptions, which were probably not involved in its interaction with the LD repeats. Titration with LD4 showed a much greater reduction in NMR signal intensity and lower chemical shift changes compared to LD2 and LD3, which possibly suggest that the LD4 repeat engages a different set of $\alpha 4$ residues with distinct interaction kinetics. In summary, LD2 and LD3 interaction with $\alpha 4$ demonstrated a greater chemical shift changes across the $\alpha 4$ cytoplasmic residues are most possibly implicated in the P234 and $\alpha 4$ interaction.

Figure 4.15 ^{15}N - $\alpha 4$ cytoplasmic tail titration with (A) P234, (B) LD2, (C) LD3 and (D) LD4. Right panel are the ^{15}N - ^1H HSQC spectra and the left panel showing the combined chemical shift changes and relative change in intensity of the $\alpha 4$ titration with P234, LD2, LD3 or LD4 in the ratio of 1:2.

Looking retrospectively into the ^{15}N -P234 and $\alpha 4$ titration data, the residues in the LD repeats that are perturbed consists of hydrophobic and polar amino acids (Table 4.3). The residues, Q⁸, S¹¹ and R¹⁸, which were consistently perturbed in P234, LD2, LD3 and LD4, are possibly involved in bringing the three LD repeats around the $\alpha 4$ cytoplasmic tail through van der waals' interaction and hydrogen bonding.

Table 4.3 The LD repeats and residues (in red) perturbed in the ^{15}N -P234 and $\alpha 4$ titration.

LD2	¹⁴¹⁽¹⁾ NLSELDRLLELNAVQH ¹⁵⁸
LD3	²¹³⁽¹⁾ VRPSVESLLDELESSVSPV ²³²
LD4	²⁶²⁽¹⁾ ATRELDELMA ²⁷⁷ SLSDFKF ²⁷⁷ MAQ

4.3.6 Binding analyses of the $\alpha 4$ interaction with P234 and LD repeats

Interaction analyses using NMR titration studies between the $\alpha 4$ with the P234 and LD repeats was observed to greatly diminish NMR signal intensity, and this was caused by precipitation. HPLC analysis of the precipitated NMR sample in the ^{15}N -P234 and $\alpha 4$ titration experiment reconstituted in 10% acetonitrile showed that the precipitate contained both P234 and $\alpha 4$ (Figure 4.16). The high concentration of proteins and peptides prerequisite of NMR experiments of the $\alpha 4$ interaction with P234 and the LD repeats likely caused the precipitation to occur.

Figure 4.16 HPLC profile of dissolved NMR sample precipitate. In red, is the absorbance at 220 nm and in green is the absorbance at 278 nm.

Fluorescence spectroscopy is an excellent tool for interaction analyses as it is a highly sensitive technique that allows micro-molar concentrations of reagents to be used, hence allowing for the determination of dissociation constants in complex or aggregate formation [150]. The interactions between $\alpha 4$ cytoplasmic tail with P234 and the LD repeats were determined using the quenching of the rhodamine conjugated to the N-terminus of $\alpha 4$ cytoplasmic domain [176]. Rhodamine undergoes self quenching at high concentrations or when several rhodamine labeled peptides are in close proximity when they aggregate [176]. Aggregation of the Rho- $\alpha 4$ peptides attributed to the increasing concentrations of P234, LD2, LD3 and LD4 causes quenching of the N-terminus attached Rhodamine (Figure 4.17).

Figure 4.17 Fluorescence emission spectra of Rho- α 4 titration with (A) P234, (B) LD2, (C) LD3 and (D) LD4.

The extent of quenching was normalized to the initial fluorescence intensity of the Rho- α 4 and plotted against the increasing titrant concentration (Figure 4.17). Non-linear regression analysis of the averaged normalized intensity as a function of titrant concentration was used to determine the dissociation constants for the Rho- α 4 interaction with P234, LD2, LD3 and LD4 respectively (Figure 4.18 and Table 4.4).

Figure 4.18 Non-linear regression analysis of the averaged normalized change in fluorescence (F_{max}) of Rho- $\alpha 4$ quenching as a function of (A) P234, (B) LD2, (C) LD3 and (D) LD4 concentration.

Table 4.4. Summary of dissociation constants obtained in interaction with Rho- $\alpha 4$

Region	Kd/(μM)
P234	0.59 ± 0.10
LD2	1.11 ± 0.05
LD3	2.22 ± 0.15
LD4	1.20 ± 0.07

4.4 Discussion

The $\alpha 4$ cytoplasmic tail exhibited a distinct conformation compared to the myristoylated αM and αX , αL and αIIb cytoplasmic tails thus far elucidated. The αM , αX and αIIb cytoplasmic domains adopted an amphipathic membrane proximal helix followed by a C-terminal loop, whereas the αL cytoplasmic tail adopted a tri-helical fold. Instead, the $\alpha 4$ cytoplasmic tail assumed a disordered structure at the membrane proximal region followed by a helix (Figure 4.7), which could possibly be due to the lack of N-terminal helical stabilization via its transmembrane domain and the distinct absence of helix disrupting amino acids, prolines and glycines. The $\alpha 4$ cytoplasmic tail too exhibited different characteristics from the αM and αX cytoplasmic tails apart from the conserved KxGFFKR motif, it does not contain many hydrophobic residues throughout the sequence (Figure 4.3), and hence the membrane distal helix formed from the $\alpha 4$ -K¹⁰ to D³² does not have an amphipathic nature. Rather, distinct positive and negative patches at specific locations predominates the $\alpha 4$ cytoplasmic tail speckled with a small hydrophobic patches contributed by I¹², L¹³, Y²⁴ and I²⁵ (Figure 4.7). Additionally, the $\alpha 4$ cytoplasmic tail when myristoylated did not contribute to its stabilization in DPC micelles instead resulted in aggregation. The amino acids of the $\alpha 4$ cytoplasmic tail were mostly helix stabilizing and the free $\alpha 4$ cytoplasmic tail structure determined at 278 K was more structurally stabilized than the free αM cytoplasmic tail at the same temperature, hence enabling the structure of the cytoplasmic tail to be determined. The distinct structure adopted by the $\alpha 4$ cytoplasmic tail reiterates the notion of differential recruitment of different intracellular ligands to affect unique signaling mechanisms in the various integrins.

The unique recruitment of paxillin to the $\alpha 4$ cytoplasmic tail is a crucial process in leukocyte migration. The binding site of paxillin in the $\alpha 4$ cytoplasmic tail has been found to be localized to the S¹¹⁽⁹⁷⁸⁾ to I²⁵⁽⁹⁹²⁾ region with E¹⁶⁽⁹⁸³⁾, S²¹⁽⁹⁸⁸⁾, Y²⁴⁽⁹⁹¹⁾ playing important roles in cell migration [115]. Correspondingly the binding site of the $\alpha 4$ cytoplasmic tail in paxillin has been found to be localized in its N-terminal region from P¹⁰⁰ to K²⁷⁷, which consists of the LD2, LD3 and LD4 repeats [172]. The interaction interface was

attempted to be determined using NMR titration studies with a paxillin construct containing the three LD repeats, P234, and the individual LD repeats. The $\alpha 4$ cytoplasmic tail residues Q⁸, S¹¹ and R¹⁸ were consistently perturbed in all titrations with the truncated forms of paxillin. Additionally, E¹⁶⁽⁹⁸³⁾ and Y²⁴⁽⁹⁹¹⁾ were found to be substantially perturbed in the ¹⁵N- $\alpha 4$ titration with P234 and LD3 which possibly implicates the LD3 repeat in the $\alpha 4$ -P234 interaction. It is highly likely that the 15 Kda P234 wraps around the $\alpha 4$ cytoplasmic tails via its LD2, LD3 and LD4 motif as all individual repeats were found to contact certain $\alpha 4$ cytoplasmic residues. Upon binding, P234 imaginably changes its conformation which probably resulted in other residues that are not part of the LD repeats to be significantly perturbed in the ¹⁵N-P234 titration with $\alpha 4$ (Figure 4.13).

The paxillin LD repeats, LD2 and LD4, have been structurally defined in crystal and NMR structures in complexes with cytoskeletal remodeling proteins such as the C-terminal calponin homology domain of the alpha-parvin [177], focal adhesion targeting (FAT) domain of the focal adhesion kinase (FAK) [178] and the cerebral cavernous malformation protein (CCM3) (Figure 4.19) [179]. The LD2 and LD4 repeats in these crystal structures mostly assumed an amphipathic helix, with leucines residing at the hydrophobic face and the remaining charged residues on the other. The interaction interface of the LD2 and LD4 with the alpha-parvin and FAT domain relies on the hydrophobic leucine face of the amphipathic helix. P234 and $\alpha 4$ NMR titration analysis reveals that hydrophobic and polar residues in the LD repeats were involved in its interaction with the $\alpha 4$ cytoplasmic tail, possibly through hydrophobic interactions and hydrogen bonding (Table 4.3), and in LD3, the negatively charged glutamine acid is implicated in its interaction with the LD repeats. Additionally, these residues of the LD repeats do not entirely lie on the hydrophobic face that was found to interact alpha-parvin and the FAT domain possibly allowing for the consecutive and/or orchestrated interaction with the $\alpha 4$ cytoplasmic domain and recruitment of other cytoskeletal proteins. It would be interesting to determine if the binding of alpha-parvin and the FAT domain to paxillin occurs consecutively with the $\alpha 4$ cytoplasmic tail.

Figure 4.19 Structures of known cytoskeletal remodeling proteins with LD2 (in red) and LD4 (in blue). Residues that were found to be perturbed in the titration with the ^{15}N - $\alpha 4$ cytoplasmic tails are highlighted in spheres and denoted. (A) LD2 and (B) LD4 in complex with the FAT domain of FAK (PDB:2L6F). (C) LD2 in complex with CCM3 (PDB:3RQF) and (D) LD4 in complex with CCM3 (PDB:3RQG). (E) LD4 in complex with the FAT domain of PyK (PDB:3GM1). (F) LD4 in complex with the CH domain of alpha-parvin.

The binding of the Rho- $\alpha 4$ cytoplasmic tail with P234 was found to be in the micro-molar range $\approx 0.59 \mu\text{M}$. The dissociation constant obtained were for interaction of the free tails without considerations of the phospholipid bilayer. The influence of the phospholipid bilayer probably plays a minor role in the interaction between P234 and $\alpha 4$ cytoplasmic tail, as the residues

involved were located at the membrane distal region which are presumably exposed to the aqueous milieu of the cytoplasmic. The slightly weaker binding of the individual fragments of LD2, LD3 and LD4 compared to the longer P234 indicates that LD motifs play an important role in the $\alpha 4$ interaction.

Chapter V: Summary and concluding remarks

The leukocyte specific integrins, namely the $\beta 2$ integrins and $\alpha 4$ integrins are involved in the generation of immune responses, specifically the migration of leukocytes to specific inflammatory sites [3]. Chemokines and cytokines induced activation of integrins resulting in the subsequent activation of the unique integrin bidirectional (inside-out and outside-in) signaling transduction [3]. The dominant processes of leukocyte movement towards the target sites first involve transient rolling interactions with the vascular endothelium via its C-type lectins with specific fucosylated sialoglycoproteins on the extracellular matrix; closer to the site of inflammation, chemoattractants which are released by the injured site activates the leukocyte integrins, $\alpha L\beta 2$, $\alpha M\beta 2$ and $\alpha 4\beta 1$ integrins to have high affinity for its endothelial ligands resulting in the formation of shear resistant adhesions; the activation of these integrins also result in its cytoplasmic domain association with intracellular effectors such as talin, which remodels the actin cytoskeleton, and paxillin which binds specifically to the $\alpha 4$ -subunit cytoplasmic domain resulting in the polarization of the leukocyte to form a leading lamellopodia and trailing uropod to facilitate the lateral migration of leukocytes to the junction of two or more endothelial cells; finally, trans-endothelial migration mediated by specific integrins with its junctional ligands involving distinct signaling pathways from lateral migration [3].

The involvement of the $\beta 2$ and $\alpha 4$ integrins in leukocyte migratory processes has implicated them in many inflammatory/autoimmune disorders where the lack of $\beta 2$ and $\alpha 4$ integrin activities resulted in leukocyte adhesion deficiencies characterized by deficient leukocyte migratory behaviours contributing to the lack of an immune response [4]. Conversely, the over activity of the $\beta 2$ and $\alpha 4$ integrins resulted in autoimmune diseases characterized by the hyper-mobility and hyper-activation of leukocyte migration behaviours leading to spurious localization of leukocytes in normal tissues contributing to autoimmune diseases such as Systemic lupus erythematosus (SLE), multiple sclerosis, rheumatoid arthritis, asthma, etc [180]. It is becoming apparent that the modulation of leukocyte trafficking is a very

effective strategy for treating a host of inflammatory diseases. Several strategies to therapeutically target integrins includes gene replacement therapy for the defective $\beta 2$ integrins in LAD and; to design in inhibitors to block extracellular ligand binding sites, block downstream integrin signaling or modulation of integrin expression for the over-active integrin molecule. Thus far, leukocyte adhesion deficiency can be cured by bone marrow or stem cell replacement for the production of normal leukocytes to replace the defective leukocytes; and could be managed by using prophylactic antibiotics or administration of fucose to prevent bacterial infections and improve phagocytic functions respectively [181]. In contrast to the lack of immune activity in LAD, autoimmune disorders caused by the over-active migration and spurious accumulation of immune cells are managed by the administration of integrin inhibitors, such as Natalizumab, which is a monoclonal antibody, that binds to the $\alpha 4$ integrin subunit that is used to treat multiple sclerosis and Crohn's disease; and Efalizumab, also a monoclonal antibody that targets the $\alpha L\beta 2$ integrin is used to treat psoriasis [182].

However, the integrin therapeutics used to control integrin activation, especially integrin antagonists is their potential to also activate the integrin receptor through allosteric changes in the β -subunit [183]. Additionally due to the immune suppressive properties of the integrin antagonists, extensive usage can lead to the development of progressive multifocal leukoencephalopathy commonly found in immune deficient patients [182]. Hence, the identification of novel targets to achieve selective integrin inhibition, particularly in targets that distinguish between the integrin inside-out and outside-in signaling to control integrin activation and actin cytoskeleton reorganization could possibly aid the development of selective integrin therapeutics with fewer side effects. The inside-out and outside-in signal transduction converges on the integrin cytoplasmic domains has been demonstrated to be regulated diversely through unique sequences in the cytoplasmic domain. The inactivated, resting state integrin is maintained by the membrane proximal clasps, consisting of the conserved α -KxGFFKR and β -HDRKE motifs, where recruitment of activated intracellular effectors through unique sequences at the cytoplasmic membrane distal regions can modulate the separation of the cytoplasmic clasps. Membrane

distal sequences of the α -cytoplasmic tails show more sequence variation than their β -subunit counterpart, and it has been demonstrated that different α -subunits are capable of recruiting their own set of intracellular effectors, such as $\alpha 4$ with paxillin and αIIb with CIB. In support of making use of the integrin α -subunit cytoplasmic tails as a therapeutic target, a small non-cytotoxic inhibitor is found to directly modulate the interaction of $\alpha 4$ with paxillin resulting in reduced T-cell migration *in vivo* [184]. Moreover, cell-permeable peptides with N-terminal myristoylation containing wild type αIIb -cytoplasmic tail sequence or a mutant sequence has been demonstrated to modulate $\alpha \text{IIb}\beta 3$ integrin activity on platelets [91], or N-terminal palmitoylated membrane proximal KVGFFKR sequence or mutated of the phenylalanines with small hydrophilic amino acids can cause a switch in biological activity resulting in inhibition of platelet aggregation, adhesion, ADP secretion or thromboxane synthesis [133].

Hence, understanding the structure attributed by the unique sequences of the α -cytoplasmic domains can provide some insights into the regulation of the concerted recruitment of intracellular effectors mediating the unique bidirectional signal transduction. Importantly, intracellular borders at the C-terminus of the $\alpha \text{IIb}\beta 3$ integrin have to be first clarified in the light of conflicting findings in the current literature [79,83,85,86]. Hydrogen-Deuterium exchange experiments can be conducted on synthetic αIIb and $\beta 3$ transmembrane and cytoplasmic domains reconstituted in phospholipid bilayers. Mass spectrometry can be used to identify the residues which amide protons have been exposed to the aqueous solution containing D_2O [185]. Ideally, amide protons that are exposed to the solution will exchange with D_2O and change in molecular weight, hence membrane proximal residues that are submerged in the hydrophobic core of the bilayer or exposed can be distinguished.

The transmembrane and membrane proximal residues of the αIIb and $\beta 3$ located in the hydrophobic core can be further localized using phospholipid analogs with their acyl chains attached to a paramagnetic doxyl moiety [186]. αIIb and $\beta 3$ transmembrane and membrane proximal segments can be first selectively labeled with ^{15}N or ^{13}C enriched amino acids. Magic-angle spinning

NMR can then be used to determine their paramagnetic relaxation rate. The difference in paramagnetic relaxation rates due to the location of the enriched amino acid within the hydrophobic acyl chains induced by the attached doxyl moiety can enable the topology and location of the isotope enriched amino acids to be determined [186].

Only after delineation of the α Ib β 3 intracellular borders can integrin inside-out signaling mechanism be made known, and the diverse methods used to determine the structures can then be generalized and used across the 24 known mammalian integrins.

Appendix

Appendix 1: HPLC profiles of synthetic and purified peptides of Chapter II.

Figure A. HPLC profiles of synthetic peptides. Absorbance at 220 nm (Intensity/a.u) is plotted as a function of time/min.

Figure B. HPLC profiles of purified peptides. Absorbance at 220 nm (Intensity/a.u) is plotted as a function of time/min.

Appendix 2: HPLC profiles of synthetic and purified peptides of Chapter III.

Figure A. HPLC profiles of purified peptides. Absorbance at 220 nm (Intensity/a.u) is plotted as a function of time/min.

Appendix 3: HPLC profiles of synthetic and purified peptides of Chapter IV.

Figure A. HPLC profiles of purified peptides. Absorbance at 220 nm (Intensity/a.u) is plotted as a function of time/min.

Appendix 4: Combined changes in chemical shift ($\Delta N + \Delta H$) and change in relative intensity in the NMR titration studies using ^{15}N - ^1H HSQC experiments.

<i>Table A. 200 μM ^{15}N-P234 titration with unlabeled $\alpha 4$.</i>						
	+ 100 μM $\alpha 4$		+ 200 μM $\alpha 4$		+ 400 μM $\alpha 4$	
	$\Delta N + \Delta H$	Relative intensity	$\Delta N + \Delta H$	Relative intensity	$\Delta N + \Delta H$	Relative intensity
N4N-H	0.014	0.664679	0.057	1.182057	12.91279	0.540277
L5N-H	0.006	0.769568	0.001	1.185341	5.85875	0.742151
S6N-H	0.009	0.829201	0.005	1.209474	5.274573	0.458932
E7N-H	0.027	0.869081	0.017	1.238565	2.834171	0.368467
L8N-H	0.03	0.85691	0.004	1.220776	1.755532	0.410602
D9N-H	0.059	0.862497	0.014	1.175847	6.194709	0.421999
R10N-H	0.017	1.169353	0.009	1.183957	6.141875	0.358148
L11N-H	0.095	0.934703	0.033	1.0993	13.56575	0.453106
L12N-H	0.066	0.846971	0.022	1.233045	8.36794	0.457585
L13N-H	0.07	0.846971	0.024	1.233045	7.350112	0.457585
E14N-H	0.14	0.696999	0.04	1.191086	3.648639	0.27432
L15N-H	0.098	0.934703	0.017	1.08561	13.99142	0.453106
N16N-H	0.046	0.7506	0.014	1.22512	2.355662	0.085935
A17N-H	0.081	0.889252	0.044	1.182851	8.262271	0.324129
V18N-H	0.092	0.875593	0.016	1.146601	1.51229	0.181269
Q19N-H	0.007	0.695347	0.027	1.073781	16.13275	0.409486
G24N-H	0.005	0.744995	0.01	1.131749	6.329283	0.372866
F25N-H	0.006	0.8049	0.052	1.050734	1.581077	0.333232
A27N-H	0.003	0.74561	0.018	1.14479	0.60811	0.01
D28N-H	0.029	0.577892	0.033	1.064983	9.036857	0.263685
E29N-H	0.065	0.868697	0.014	1.224634	14.71318	0.412149
A30N-H	0.007	0.754636	0.004	1.079689	0.904183	0.252486
N31N-H	0.013	0.664932	0.012	1.068039	9.985898	0.45589
S32N-H	0.027	0.696854	0.006	1.027055	1.929988	0.126088
S33N-H	0.062	0.944142	0.039	1.214663	4.248769	0.225317
L36N-H	0.023	0.873313	0.006	1.234997	9.150502	0.339269
G38N-H	0.01	0.755663	0.011	1.165687	23.05418	0.510562
A39N-H	0.018	0.749238	0.007	1.145711	5.660365	0.277718
L40N-H	0.038	0.587586	0.008	1.163314	5.449029	0.261926
S41N-H	0.019	0.742667	0.001	1.095156	7.236467	0.255832
L43N-H	0.139	0.696999	0.03	1.191086	3.648639	0.27432
Y44N-H	0.018	0.785913	0.026	1.100912	0.304053	0.282377
G45N-H	0.016	0.740981	0.002	1.115548	7.987123	0.492635
V46N-H	0.005	0.82629	0.004	1.155462	7.638212	0.321227
E48N-H	0.024	0.826358	0.017	1.10225	5.412147	0.299854

T49N-H	0.015	0.714174	0.017	1.099831	5.266597	0.292481
N50N-H	0.09	0.777669	0.043	0.913022	5.168905	0.193716
S51N-H	0.024	0.696151	0.013	1.068651	2.128373	0.397122
L53N-H	0.007	0.644007	0.013	1.100409	9.126572	0.320036
G54N-H	0.005	0.734022	0.002	1.164924	4.277674	0.331335
G55N-H	0.004	0.677492	0.01	1.201898	20.65967	0.253247
K56N-H	0.038	0.587966	0.015	1.148151	3.709449	0.267081
A57N-H	0.045	0.684548	0.017	1.203596	1.147426	0.312841
G58N-H	0.004	0.75799	0.006	1.164465	7.135774	0.339934
L60N-H	0.003	0.759503	0.018	1.169533	11.45333	0.523771
T61N-H	0.023	0.772408	0.01	1.095067	0.843373	0.210692
K62N-H	0.032	0.8733	0.002	1.234978	10.18428	0.339264
E63N-H	0.011	0.719672	0.018	1.115477	1.086615	0.356236
K64N-H	0.026	0.69367	0.024	1.073889	2.894982	0.228436
R67N-H	0.01	0.695347	0.004	1.015336	12.80412	0.449735
N68N-H	0.009	0.64808	0.033	1.088872	17.02098	0.588629
G69N-H	0.008	0.692573	0	1.107378	6.215638	0.397353
G70N-H	0.001	0.703776	0.011	1.280037	1.20026	0.31574
R71N-H	0.065	0.675802	0.024	1.203013	2.781337	0.388181
G72N-H	0.002	0.699575	0.01	1.102404	4.59768	0.434142
L73N-H	0.033	0.630202	0.006	1.02826	1.147426	0.160332
E74N-H	0.007	0.787394	0.005	1.191182	10.3059	0.305344
V76N-H	0.02	0.997939	0.018	1.400665	7.68307	0.32599
R77N-H	0.034	0.812929	0.013	1.188586	1.694722	0.127044
S79N-H	0.021	0.829622	0.007	1.211417	5.25862	0.324522
E81N-H	0.15	0.939429	0.032	1.213684	32.26352	0.39963
S82N-H	0.019	0.99762	0.009	1.175845	2.120396	0.378414
L83N-H	0.047	0.973636	0.016	1.396868	14.17583	0.443982
E86N-H	0.076	0.971122	0.03	1.223741	13.18493	0.449123
E88N-H	0.078	0.681902	0.025	1.200187	2.781337	0.387269
S89N-H	0.03	0.663716	0.039	0.989751	9.316981	0.415837
S90N-H	0.027	0.761397	0.026	1.221292	8.497537	0.523273
V91N-H	0.057	0.848256	0.056	1.152029	14.17386	0.205306
S93N-H	0.027	0.752276	0.022	1.055606	18.25615	0.283536
A97N-H	0.01	0.751463	0.005	1.095232	0.425675	0.391859
I98N-H	0.096	0.699014	0.043	1.153862	4.756182	0.290037
T99N-H	0.025	0.89024	0.023	1.14336	2.234041	0.241382
V100N-H	0.033	0.886608	0	1.179545	10.75551	0.7716
N101N-H	0.025	0.697238	0.01	1.143969	16.62721	0.26433
Q102N-H	0.03	0.632966	0.002	1.14441	5.898633	0.330462
G103N-H	0.005	0.742347	0.008	1.126984	4.148076	0.372866
E104N-H	0.038	0.820145	0.008	1.169842	2.286875	0.333663
M105N-H	0.016	0.686485	0.018	1.136652	0.721751	0.300801
S106N-H	0.018	0.646792	0.001	1.052604	0.060811	0.172208

S107N-H	0.025	0.573211	0.007	1.137807	3.101343	0.250657
Q109N-H	0.005	0.784722	0.014	1.16181	3.434301	0.278325
R110N-H	0.011	0.718305	0.013	1.17361	2.522141	0.166926
V111N-H	0.018	0.67389	0.005	0.983341	1.633911	0.183685
T112N-H	0.01	0.720577	0.001	1.126277	7.897407	0.249863
S113N-H	0.009	0.72279	0.013	1.112846	3.199035	0.061207
T114N-H	0.065	0.861091	0.059	1.135545	4.848899	0.39459
Q115N-H	0.013	0.692955	0	1.100329	2.781337	0.346143
Q116N-H	0.003	0.731016	0.004	1.163876	11.07251	0.41448
Q117N-H	0.08	0.641273	0.063	1.077301	14.58658	0.310968
T118N-H	0.057	0.867722	0.013	1.201301	1.329858	0.272728
R119N-H	0.008	0.753515	0.006	1.156928	0	0.010104
I120N-H	0.098	0.847878	0.029	1.169182	11.04361	0.322933
S121N-H	0.013	0.741759	0.023	1.208878	7.188608	0.297767
A122N-H	0.03	0.726723	0.013	1.105947	2.242017	0.030249
S123N-H	0.006	0.671974	0.001	1.193765	0.608106	0.132327
S124N-H	0.023	0.98896	0.025	1.574775	13.1321	0.315524
A125N-H	0.032	0.680487	0.006	1.247808	7.849549	0.364204
T126N-H	0.027	0.750737	0.024	1.079396	6.754957	0.284383
R127N-H	0.021	0.797037	0.032	1.112898	1.147426	0.119537
E128N-H	0.093	0.926795	0.04	1.333703	3.154177	0.231082
L129N-H	0.091	0.934703	0.024	1.08561	13.68737	0.453106
A134N-H	0.061	0.848256	0.088	1.164986	14.59953	0.205306
S135N-H	0.03	0.868495	0.021	1.243804	12.56087	0.326258
L136N-H	0.052	0.89509	0.028	1.284499	0.243243	0.316806
S137N-H	0.055	0.72245	0.005	1.14689	7.788738	0.278548
D138N-H	0.072	0.603722	0.005	1.067967	1.633911	-0.00724
F139N-H	0.066	0.596686	0.038	1.160017	0.486485	0.337081
K140N-H	0.145	0.560003	0.062	1.174952	8.132674	0.356041
L141N-H	0.162	0.808369	0.003	1.254329	16.42385	0.396074
E142N-H	0.052	0.675802	0.016	1.203013	2.234041	0.388181

<i>Table B. 200 μM ^{15}N-α4 titration with unlabeled P234</i>				
	+400 μ M P234		+800 μ M P234	
	Δ N+ Δ H	Relative intensity	Δ N+ Δ H	Relative intensity
A2	4.393814	0.765492	5.006999	0.218172
G3	2.38557	0.704635	5.002718	0.251833
F4	0	1	0	1
F5	1.990798	0.64034	6.238067	0.190744
K6	2.330738	0.617798	7.420384	0.099606
R7	3.461121	0.682313	6.690561	0.198241
Q8	2.944728	0.633105	20.24963	0.335662
Y9	2.934461	0.739334	3.809432	0.328172
K10	3.648039	0.685001	6.664641	0.199022
S11	0	1	19.31603	0.275327
I12	2.813337	0.68944	4.29921	0.340708
L13	1.294667	0.648468	10.3376	0.470947
Q14	1.278816	0.591792	10.08579	0.42476
E15	2.054199	0.665471	14.76202	0.291718
E16	2.877236	0.654897	11.66407	0.337207
N17	3.64086	0.630923	4.806331	0.25033
R18	2.289868	0.706856	11.98457	0.374727
R19	0	1	9.278804	0.466841
S21	1.600615	0.749867	3.958568	0.487223
W22	0.304053	0.798085	4.621309	0.580218
W22HE	0	1	2.948519	0.752894
S23	8.120203	0.723873	12.96531	0.431797
Y24	0	1	14.18053	0.587316
I25	0	1	15.64028	0.559553
N26	10.50479	1.143586	12.8755	0.640699
S27	3.887095	0.789172	8.417488	0.422523
K28	3.641957	0.685001	6.120039	0.199022
S29	1.133968	0.678717	11.68291	0.160128
N30	0.787048	0.805373	11.63166	0.389435
D31	3.480359	0.742901	6.440242	0.621494
D32	0	1	0	1

Table C.200μM¹⁵N-α4 titration with unlabeled LD2				
	+400 μM LD2		+800 μM LD2	
	ΔN+ΔH	Relative intensity	ΔN+ΔH	Relative intensity
A2	2.476186	0.688249	0	1
G3	2.218988	0.66495	2.418866	0.474396
F4	5.719083	0.701013	3.00185	0.477389
F5	1.497735	0.649053	2.942137	0.487682
K6	0	1	3.50199	0.439442
R7	0	1	5.162121	0.416419
Q8	0	1	11.18286	0.400446
Y9	4.461503	0.685028	0	1
K10	0	1	5.752482	0.416419
S11	12.26041	0.675848	9.651637	0.466777
I12	4.324924	0.688392	4.383445	0.48317
L13	7.908164	0.694637	6.408439	0.408761
Q14	6.456385	0.632947	4.438972	0.401817
E15	7.879054	0.648764	4.889768	0.422176
E16	7.104066	0.689136	5.663556	0.438542
N17	8.837666	0.625216	11.95615	0.446832
R18	6.113255	0.679073	10.26184	0.570968
R19	7.266264	0.679492	5.916971	0.49479
S21	5.809004	0.735986	3.239707	0.562851
W22	5.339664	0.727066	7.236854	0.582306
W22HE	2.707566	0.652605	2.711855	0.488233
S23	5.356518	0.645247	4.902531	0.462328
Y24	3.938034	0.66076	7.637312	0.493925
I25	4.358923	0.59233	5.816886	0.458779
N26	7.427657	0.552464	8.149511	0.438917
S27	10.70475	0.74706	5.205486	0.546879
K28	0	1	6.51391	0.421318
S29	6.624862	0.571617	4.378161	0.435736
N30	7.911552	0.688043	9.006341	0.447638
D31	5.82904	0.675027	6.138779	0.46356
D32	5.715695	0.634174	4.751396	0.411297

<i>Table D.200μM¹⁵N-α4 titration with unlabeled LD3</i>				
	+400 μ M LD3		+800 μ M LD3	
	Δ N+ Δ H	Relative intensity	Δ N+ Δ H	Relative intensity
A2	0	2	3.248584	0.681149
G3	0.544105	0.786502	6.04138	0.354958
F4	0.748468	0.810215	1.700503	0.361065
F5	2.478879	0.828236	5.44155	0.358304
K6	2.108732	0.831634	5.131115	0.399183
R7	4.066732	0.864102	6.716678	0.312728
Q8	4.630675	0.75794	15.89259	0.275378
Y9	3.109612	0.724723	8.032178	0.329019
K10	4.107704	0.864102	6.716678	0.312728
S11	8.415285	0.829213	12.5159	0.267509
I12	3.632386	0.777746	7.171463	0.312883
L13	2.644663	0.74032	10.22664	0.306527
Q14	0.579794	1.211141	6.920441	0.685067
E15	2.304921	1.070277	5.666549	0.271403
E16	4.756182	0.816681	10.70585	0.339269
N17	4.801245	0.763501	11.53606	0.246124
R18	8.307721	0.796825	12.5835	0.313575
R19	2.206629	0.820034	2.206629	1.640067
S21	2.349779	0.866052	7.505819	0.310661
W22	3.203521	1.375878	2.57079	0.544007
W22HE	1.729913	1.373072	5.717685	0.400527
S23	1.579481	0.738354	5.301488	0.31546
Y24	1.50222	0.699829	7.793721	0.252975
I25	4.562789	0.818232	6.092121	0.291349
N26	1.821326	0.831417	8.891701	0.38925
S27	2.991086	0.792391	15.57827	0.387887
K28	4.234609	0.864102	7.100583	0.312728
S29	2.888103	0.775457	8.707974	0.358668
N30	4.108201	0.774314	11.13541	0.270312
D31	4.195129	0.716448	11.95157	0.28934
D32	2.143022	0.723981	8.108041	0.269408

<i>Table D. 200 μM ^{15}N-α4 titration with unlabeled LD4.</i>				
	+400 μ M LD4		+800 μ M LD4	
	Δ N+ Δ H	Relative intensity	Δ N+ Δ H	Relative intensity
A2	2.84105	0.592759	5.708122	0.432108
G3	1.681762	0.612319	0	1
F4	0	1	0	1
F5	3.193254	0.342003	6.197995	0.192136
K6	3.421941	0.319823	5.731048	0.245186
R7	2.169842	0.544329	4.5582	0.353579
Q8	4.146986	0.38313	10.43161	0.150606
Y9	0	1	3.192156	0.275325
K10	2.217693	0.547386	4.608444	0.350569
S11	0	1	2.85979	0.284061
I12	2.046128	0.499188	3.533596	0.269405
L13	3.058073	0.457362	5.286333	0.226559
Q14	2.376598	0.491105	3.102235	0.225331
E15	0	1	1.717648	0.240717
E16	0	1	3.372195	0.271457
N17	3.961261	0.438214	7.840475	0.212543
R18	3.314575	0.500563	5.762946	0.252071
R19	0	1	7.162886	0.275704
S21	2.693209	0.5098	2.779639	0.312484
W22	1.509099	0.514099	2.42734	0.278813
W22HE	1.249706	0.431006	3.75839	0.216244
S23	0	1	4.112687	0.201135
Y24	0	1	0	1
I25	1.927697	0.497299	0	1
N26	1.919223	0.494857	2.549656	0.19389
S27	5.615508	0.534173	5.05376	0.246635
K28	2.276111	0.539326	3.80046	0.345407
S29	1.582372	0.505635	4.255047	0.264458
N30	3.879514	0.510628	3.202423	0.338515
D31	1.335141	0.485451	5.518408	0.232688
D32	5.534061	0.46448	16.10653	0.251048

References

1. Hayward AR, Harvey BA, Leonard J, Greenwood MC, Wood CB, Soothill JF: **Delayed separation of the umbilical cord, widespread infections, and defective neutrophil mobility.** *Lancet* (1979) **1**(8126):1099-1101.
2. Crowley CA, Curnutte JT, Rosin RE, Andre-Schwartz J, Gallin JI, Klemmner M, Snyderman R, Southwick FS, Stossel TP, Babior BM: **An inherited abnormality of neutrophil adhesion. Its genetic transmission and its association with a missing protein.** *N Engl J Med* (1980) **302**(21):1163-1168.
3. Rose DM, Alon R, Ginsberg MH: **Integrin modulation and signaling in leukocyte adhesion and migration.** *Immunol Rev* (2007) **218**(126-134).
4. Hanna S, Etzioni A: **Leukocyte adhesion deficiencies.** *Ann N Y Acad Sci* (2012) **1250**(50-55).
5. Adachi Y, Kindzelskii AL, Cookingham G, Shaya S, Moore EC, Todd RF, 3rd, Petty HR: **Aberrant neutrophil trafficking and metabolic oscillations in severe pyoderma gangrenosum.** *J Invest Dermatol* (1998) **111**(2):259-268.
6. Kostidou E, Koliakos G, Kaloyianni M: **Increased monocyte alpha1, alpha2 and beta2 integrin subunits in diabetes mellitus.** *Clin Biochem* (2009) **42**(7-8):634-640.
7. Desgrosellier JS, Cheresch DA: **Integrins in cancer: Biological implications and therapeutic opportunities.** *Nat Rev Cancer* (2010) **10**(1):9-22.
8. Shimizu Y, Rose DM, Ginsberg MH: **Integrins in the immune system.** *Adv Immunol* (1999) **72**(325-380).
9. De Arcangelis A, Georges-Labouesse E: **Integrin and ecm functions: Roles in vertebrate development.** *Trends Genet* (2000) **16**(9):389-395.
10. Hynes RO: **Integrins: Bidirectional, allosteric signaling machines.** *Cell* (2002) **110**(6):673-687.
11. Johnson MS, Lu N, Denessiouk K, Heino J, Gullberg D: **Integrins during evolution: Evolutionary trees and model organisms.** *Biochim Biophys Acta* (2009) **1788**(4):779-789.
12. Meighan CM, Schwarzbauer JE: **Temporal and spatial regulation of integrins during development.** *Curr Opin Cell Biol* (2008) **20**(5):520-524.

13. Springer TA, Thompson WS, Miller LJ, Schmalstieg FC, Anderson DC: **Inherited deficiency of the mac-1, lfa-1, p150,95 glycoprotein family and its molecular basis.** *J Exp Med* (1984) **160**(6):1901-1918.
14. Humphries JD, Byron A, Humphries MJ: **Integrin ligands at a glance.** *J Cell Sci* (2006) **119**(Pt 19):3901-3903.
15. Liu S, Calderwood DA, Ginsberg MH: **Integrin cytoplasmic domain-binding proteins.** *J Cell Sci* (2000) **113** (Pt 20):3563-3571.
16. Scharffetter-Kochanek K, Lu H, Norman K, van Nood N, Munoz F, Grabbe S, McArthur M, Lorenzo I, Kaplan S, Ley K, Smith CW *et al*: **Spontaneous skin ulceration and defective t cell function in cd18 null mice.** *J Exp Med* (1998) **188**(1):119-131.
17. Abram CL, Lowell CA: **The ins and outs of leukocyte integrin signaling.** *Annu Rev Immunol* (2009) **27**:339-362.
18. Liu S, Thomas SM, Woodside DG, Rose DM, Kiosses WB, Pfaff M, Ginsberg MH: **Binding of paxillin to alpha4 integrins modifies integrin-dependent biological responses.** *Nature* (1999) **402**(6762):676-681.
19. Hyland RH, Douglass WA, Tan SM, Law SK: **Chimeras of the integrin beta subunit mid-region reveal regions required for heterodimer formation and for activation.** *Cell Commun Adhes* (2001) **8**(2):61-69.
20. Xiong JP, Stehle T, Diefenbach B, Zhang R, Dunker R, Scott DL, Joachimiak A, Goodman SL, Arnaout MA: **Crystal structure of the extracellular segment of integrin alpha vbeta3.** *Science* (2001) **294**(5541):339-345.
21. Springer TA, Wang JH: **The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion.** *Adv Protein Chem* (2004) **68**:29-63.
22. Luo BH, Carman CV, Springer TA: **Structural basis of integrin regulation and signaling.** *Annu Rev Immunol* (2007) **25**:619-647.
23. Luo BH, Springer TA: **Integrin structures and conformational signaling.** *Curr Opin Cell Biol* (2006) **18**(5):579-586.
24. Tan SM: **The leucocyte beta2 (cd18) integrins: The structure, functional regulation and signalling properties.** *Biosci Rep* (2012) **32**(3):241-269.
25. Askari JA, Buckley PA, Mould AP, Humphries MJ: **Linking integrin conformation to function.** *J Cell Sci* (2009) **122**(2):165-170.

26. Nishida N, Xie, C., Shimaoka, M., Cheng, Y., Walz, T., Springer, T.A.: **Activation of leukocyte beta2 integrins by conversion from bent to extended conformations.** *Immunity* (2006) **25**(4):583.
27. Takagi J, Petre, B.M., Walz, T., Springer, T.A.: **Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling.** *Cell* (2002) **110**(5):599-611.
28. Luo BH, Carman, C.V., Springer, T.A.: **Structural basis of integrin regulation and signaling.** *Annu Rev Immunol* (2007) **25**:619-647.
29. Byron A, Humphries JD, Askari JA, Craig SE, Mould AP, Humphries MJ: **Anti-integrin monoclonal antibodies.** *J Cell Sci* (2009) **122**(Pt 22):4009-4011.
30. Arnaout MA, Goodman SL, Xiong JP: **Structure and mechanics of integrin-based cell adhesion.** *Curr Opin Cell Biol* (2007) **19**(5):495-507.
31. Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaout MA: **Crystal structure of the extracellular segment of integrin alpha vbeta3 in complex with an arg-gly-asp ligand.** *Science* (2002) **296**(5565):151-155.
32. Altieri DC: **Occupancy of cd11b/cd18 (mac-1) divalent ion binding site(s) induces leukocyte adhesion.** *J Immunol* (1991) **147**(6):1891-1898.
33. Zhu J, Boylan B, Luo BH, Newman PJ, Springer TA: **Tests of the extension and deadbolt models of integrin activation.** *J Biol Chem* (2007) **282**(16):11914-11920.
34. Gupta V, Gylling A, Alonso JL, Sugimori T, Ianakiev P, Xiong JP, Arnaout MA: **The beta-tail domain (betatd) regulates physiologic ligand binding to integrin cd11b/cd18.** *Blood* (2007) **109**(8):3513-3520.
35. Zhu J, Luo BH, Xiao T, Zhang C, Nishida N, Springer TA: **Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces.** *Mol Cell* (2008) **32**(6):849-861.
36. Shi M, Foo SY, Tan SM, Mitchell EP, Law SK, Lescar J: **A structural hypothesis for the transition between bent and extended conformations of the leukocyte beta2 integrins.** *J Biol Chem* (2007) **282**(41):30198-30206.
37. Xiong JP, Mahalingham B, Alonso JL, Borrelli LA, Rui X, Anand S, Hyman BT, Rysiok T, Muller-Pompalla D, Goodman SL, Arnaout MA: **Crystal structure of the complete integrin alphavbeta3 ectodomain**

- plus an α/β transmembrane fragment. *J Cell Biol* (2009) 186(4):589-600.
38. Xie C, Zhu J, Chen X, Mi L, Nishida N, Springer TA: **Structure of an integrin with an α domain, complement receptor type 4.** *Embo J* (2010) 29(3):666-679.
39. Takagi J, Erickson HP, Springer TA: **C-terminal opening mimics 'inside-out' activation of integrin $\alpha 5 \beta 1$.** *Nat Struct Biol* (2001) 8(5):412-416.
40. Takagi J, Strokovich K, Springer TA, Walz T: **Structure of integrin $\alpha 5 \beta 1$ in complex with fibronectin.** *Embo J* (2003) 22(18):4607-4615.
41. Chen X, Xie C, Nishida N, Li Z, Walz T, Springer TA: **Requirement of open headpiece conformation for activation of leukocyte integrin $\alpha x \beta 2$.** *Proc Natl Acad Sci U S A* (2010) 107(33):14727-14732.
42. Iwasaki K, Mitsuoka K, Fujiyoshi Y, Fujisawa Y, Kikuchi M, Sekiguchi K, Yamada T: **Electron tomography reveals diverse conformations of integrin $\alpha i i \beta 3$ in the active state.** *J Struct Biol* (2005) 150(3):259-267.
43. Adair BD, Yeager M: **Three-dimensional model of the human platelet integrin $\alpha i i \beta 3$ based on electron cryomicroscopy and x-ray crystallography.** *Proc Natl Acad Sci U S A* (2002) 99(22):14059-14064.
44. Adair BD, Xiong JP, Maddock C, Goodman SL, Arnaout MA, Yeager M: **Three-dimensional em structure of the ectodomain of integrin $\{\alpha\}\{\beta\}3$ in a complex with fibronectin.** *J Cell Biol* (2005) 168(7):1109-1118.
45. Ye F, Liu J, Winkler H, Taylor KA: **Integrin $\alpha i i \beta 3$ in a membrane environment remains the same height after mn^{2+} activation when observed by cryoelectron tomography.** *J Mol Biol* (2008) 378(5):976-986.
46. Chigaev A, Buranda T, Dwyer DC, Prossnitz ER, Sklar LA: **FRET detection of cellular $\alpha 4$ -integrin conformational activation.** *Biophys J* (2003) 85(6):3951-3962.
47. Beglova N, Blacklow SC, Takagi J, Springer TA: **Cysteine-rich module structure reveals a fulcrum for integrin rearrangement upon activation.** *Nat Struct Biol* (2002) 9(4):282-287.
48. Gottschalk KE: **A coiled-coil structure of the $\alpha i i \beta 3$ integrin transmembrane and cytoplasmic domains in its resting state.** *Structure* (2005) 13(5):703-712.

49. Rocco M, Rosano C, Weisel JW, Horita DA, Hantgan RR: **Integrin conformational regulation: Uncoupling extension/tail separation from changes in the head region by a multiresolution approach.** *Structure* (2008) **16**(6):954-964.
50. Springer TA, Dustin ML: **Integrin inside-out signaling and the immunological synapse.** *Curr Opin Cell Biol* (2012) **24**(1):107-115.
51. Xiong JP, Stehle T, Goodman SL, Arnaout MA: **A novel adaptation of the integrin psi domain revealed from its crystal structure.** *J Biol Chem* (2004) **279**(39):40252-40254.
52. Shi M, Sundramurthy K, Liu B, Tan SM, Law SK, Lescar J: **The crystal structure of the plexin-semaphorin-integrin domain/hybrid domain/i-egf1 segment from the human integrin beta2 subunit at 1.8-a resolution.** *J Biol Chem* (2005) **280**(34):30586-30593.
53. Xiao T, Takagi J, Collier BS, Wang JH, Springer TA: **Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics.** *Nature* (2004) **432**(7013):59-67.
54. Lee JO, Rieu P, Arnaout MA, Liddington R: **Crystal structure of the a domain from the alpha subunit of integrin cr3 (cd11b/cd18).** *Cell* (1995) **80**(4):631-638.
55. Lee JO, Bankston LA, Arnaout MA, Liddington RC: **Two conformations of the integrin a-domain (i-domain): A pathway for activation?** *Structure* (1995) **3**(12):1333-1340.
56. Emsley J, Knight CG, Farndale RW, Barnes MJ, Liddington RC: **Structural basis of collagen recognition by integrin alpha2beta1.** *Cell* (2000) **101**(1):47-56.
57. Song G, Yang Y, Liu JH, Casasnovas JM, Shimaoka M, Springer TA, Wang JH: **An atomic resolution view of icam recognition in a complex between the binding domains of icam-3 and integrin alphabeta2.** *Proc Natl Acad Sci U S A* (2005) **102**(9):3366-3371.
58. Yang W, Shimaoka M, Salas A, Takagi J, Springer TA: **Intersubunit signal transmission in integrins by a receptor-like interaction with a pull spring.** *Proc Natl Acad Sci U S A* (2004) **101**(9):2906-2911.
59. Wegener KL, Campbell ID: **Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (review).** *Mol Membr Biol* (2008) **25**(5):376-387.
60. Ulmer TS: **Structural basis of transmembrane domain interactions in integrin signaling.** *Cell Adh Migr* (2010) **4**(2):243-248.
61. Kim M, Carman CV, Springer TA: **Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins.** *Science* (2003) **301**(5640):1720-1725.

62. Stefansson A, Armulik A, Nilsson I, von Heijne G, Johansson S: **Determination of n- and c-terminal borders of the transmembrane domain of integrin subunits.** *J Biol Chem* (2004) **279**(20):21200-21205.
63. Lau TL, Dua V, Ulmer TS: **Structure of the integrin α IIb transmembrane segment.** *J Biol Chem* (2008) **283**(23):16162-16168.
64. Lau TL, Partridge AW, Ginsberg MH, Ulmer TS: **Structure of the integrin β 3 transmembrane segment in phospholipid bicelles and detergent micelles.** *Biochemistry* (2008) **47**(13):4008-4016.
65. Partridge AW, Liu S, Kim S, Bowie JU, Ginsberg MH: **Transmembrane domain helix packing stabilizes integrin α IIb β 3 in the low affinity state.** *J Biol Chem* (2005) **280**(8):7294-7300.
66. Luo BH, Carman CV, Takagi J, Springer TA: **Disrupting integrin transmembrane domain heterodimerization increases ligand binding affinity, not valency or clustering.** *Proc Natl Acad Sci U S A* (2005) **102**(10):3679-3684.
67. Zhu J, Luo BH, Barth P, Schonbrun J, Baker D, Springer TA: **The structure of a receptor with two associating transmembrane domains on the cell surface: Integrin α IIb β 3.** *Mol Cell* (2009) **34**(2):234-249.
68. Zhu H, Metcalf DG, Streu CN, Billings PC, Degrado WF, Bennett JS: **Specificity for homooligomer versus heterooligomer formation in integrin transmembrane helices.** *J Mol Biol* (2010) **401**(5):882-891.
69. Berger BW, Kulp DW, Span LM, DeGrado JL, Billings PC, Senes A, Bennett JS, DeGrado WF: **Consensus motif for integrin transmembrane helix association.** *Proc Natl Acad Sci U S A* (2010) **107**(2):703-708.
70. Wang W, Luo BH: **Structural basis of integrin transmembrane activation.** *J Cell Biochem* (2010) **109**(3):447-452.
71. Schneider D, Engelman DM: **Gallex, a measurement of heterologous association of transmembrane helices in a biological membrane.** *J Biol Chem* (2003) **278**(5):3105-3111.
72. Vararattanavech A, Lin X, Torres J, Tan SM: **Disruption of the integrin α IIb transmembrane domain interface by β 2 thr-686 mutation activates α IIb and promotes micro-clustering of the α IIb subunits.** *J Biol Chem* (2009) **284**(5):3239-3249.
73. Li R, Babu CR, Lear JD, Wand AJ, Bennett JS, DeGrado WF: **Oligomerization of the integrin α IIb β 3: Roles of the**

- transmembrane and cytoplasmic domains. *Proc Natl Acad Sci U S A* (2001) **98**(22):12462-12467.
74. Parthasarathy K, Lin X, Tan SM, Law SK, Torres J: **Transmembrane helices that form two opposite homodimeric interactions: An asparagine scan study of alpha and beta2 integrins.** *Protein Sci* (2008) **17**(5):930-938.
75. Lu C, Takagi J, Springer TA: **Association of the membrane proximal regions of the alpha and beta subunit cytoplasmic domains constrains an integrin in the inactive state.** *J Biol Chem* (2001) **276**(18):14642-14648.
76. O'Toole TE, Katagiri Y, Faull RJ, Peter K, Tamura R, Quaranta V, Loftus JC, Shattil SJ, Ginsberg MH: **Integrin cytoplasmic domains mediate inside-out signal transduction.** *J Cell Biol* (1994) **124**(6):1047-1059.
77. Hughes PE, O'Toole TE, Ylanne J, Shattil SJ, Ginsberg MH: **The conserved membrane-proximal region of an integrin cytoplasmic domain specifies ligand binding affinity.** *J Biol Chem* (1995) **270**(21):12411-12417.
78. Hughes PE, Diaz-Gonzalez F, Leong L, Wu C, McDonald JA, Shattil SJ, Ginsberg MH: **Breaking the integrin hinge. A defined structural constraint regulates integrin signaling.** *J Biol Chem* (1996) **271**(12):6571-6574.
79. Kim C, Schmidt T, Cho EG, Ye F, Ulmer TS, Ginsberg MH: **Basic amino-acid side chains regulate transmembrane integrin signalling.** *Nature* (2012) **481**(7380):209-213.
80. Gottschalk KE, Adams PD, Brunger AT, Kessler H: **Transmembrane signal transduction of the alpha(iiib)beta(3) integrin.** *Protein Sci* (2002) **11**(7):1800-1812.
81. Gottschalk KE, Kessler H: **Evidence for hetero-association of transmembrane helices of integrins.** *FEBS Lett* (2004) **557**(1-3):253-258.
82. Lin X, Tan SM, Law SK, Torres J: **Unambiguous prediction of human integrin transmembrane heterodimer interactions using only homologous sequences.** *Proteins* (2006) **65**(2):274-279.
83. Lau TL, Kim C, Ginsberg MH, Ulmer TS: **The structure of the integrin alpha(iiib)beta3 transmembrane complex explains integrin transmembrane signalling.** *Embo J* (2009) **28**(9):1351-1361.
84. Metcalf DG, Kulp DW, Bennett JS, DeGrado WF: **Multiple approaches converge on the structure of the integrin alpha(iiib)/beta3 transmembrane heterodimer.** *J Mol Biol* (2009) **392**(4):1087-1101.

85. Yang J, Ma YQ, Page RC, Misra S, Plow EF, Qin J: **Structure of an integrin α IIb β 3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation.** *Proc Natl Acad Sci U S A* (2009) **106**(42):17729-17734.
86. Kurtz L, Kao L, Newman D, Kurtz I, Zhu Q: **Integrin α IIb β 3 inside-out activation: An in situ conformational analysis reveals a new mechanism.** *J Biol Chem* (2012) **287**(27):23255-23265.
87. Vinogradova O, Vaynberg J, Kong X, Haas TA, Plow EF, Qin J: **Membrane-mediated structural transitions at the cytoplasmic face during integrin activation.** *Proc Natl Acad Sci U S A* (2004) **101**(12):4094-4099.
88. Hwang PM, Vogel HJ: **Structures of the platelet calcium- and integrin-binding protein and the α IIb-integrin cytoplasmic domain suggest a mechanism for calcium-regulated recognition; homology modelling and nmr studies.** *J Mol Recognit* (2000) **13**(2):83-92.
89. Ulmer TS, Yaspan B, Ginsberg MH, Campbell ID: **Nmr analysis of structure and dynamics of the cytosolic tails of integrin α IIb β 3 in aqueous solution.** *Biochemistry* (2001) **40**(25):7498-7508.
90. Weljie AM, Hwang PM, Vogel HJ: **Solution structures of the cytoplasmic tail complex from platelet integrin α IIb- and β 3-subunits.** *Proc Natl Acad Sci U S A* (2002) **99**(9):5878-5883.
91. Vinogradova O, Haas T, Plow EF, Qin J: **A structural basis for integrin activation by the cytoplasmic tail of the α IIb-subunit.** *Proc Natl Acad Sci U S A* (2000) **97**(4):1450-1455.
92. Vinogradova O, Velyvis A, Velyviene A, Hu B, Haas T, Plow E, Qin J: **A structural mechanism of integrin α (IIb) β (3) "inside-out" activation as regulated by its cytoplasmic face.** *Cell* (2002) **110**(5):587-597.
93. Bhunia A, Tang XY, Mohanram H, Tan SM, Bhattacharjya S: **Nmr solution conformations and interactions of integrin α IIb β 2 cytoplasmic tails.** *J Biol Chem* (2009) **284**(6):3873-3884.
94. Carman CV, Springer TA: **Integrin avidity regulation: Are changes in affinity and conformation underemphasized?** *Curr Opin Cell Biol* (2003) **15**(5):547-556.
95. Astrof NS, Salas A, Shimaoka M, Chen J, Springer TA: **Importance of force linkage in mechanochemistry of adhesion receptors.** *Biochemistry* (2006) **45**(50):15020-15028.
96. Campbell ID, Humphries MJ: **Integrin structure, activation, and interactions.** *Cold Spring Harb Perspect Biol* (2011) **3**(3).

97. Hogg N, Patzak I, Willenbrock F: **The insider's guide to leukocyte integrin signalling and function.** *Nat Rev Immunol* (2011) **11**(6):416-426.
98. Luo BH, Springer TA, Takagi J: **A specific interface between integrin transmembrane helices and affinity for ligand.** *PLoS Biol* (2004) **2**(6):e153.
99. Zou Z, Chen H, Schmaier AA, Hynes RO, Kahn ML: **Structure-function analysis reveals discrete beta3 integrin inside-out and outside-in signaling pathways in platelets.** *Blood* (2007) **109**(8):3284-3290.
100. Gahmberg CG, Fagerholm SC, Nurmi SM, Chavakis T, Marchesan S, Gronholm M: **Regulation of integrin activity and signalling.** *Biochim Biophys Acta* (2009) **1790**(6):431-444.
101. Hato T, Yamanouchi J, Tamura T, Yakushijin Y, Sakai I, Yasukawa M: **Cooperative role of the membrane-proximal and -distal residues of the integrin beta3 cytoplasmic domain in regulation of talin-mediated alpha iib beta3 activation.** *J Biol Chem* (2008) **283**(9):5662-5668.
102. Moser M, Legate KR, Zent R, Fassler R: **The tail of integrins, talin, and kindlins.** *Science* (2009) **324**(5929):895-899.
103. Kiema T, Lad Y, Jiang P, Oxley CL, Baldassarre M, Wegener KL, Campbell ID, Ylanne J, Calderwood DA: **The molecular basis of filamin binding to integrins and competition with talin.** *Mol Cell* (2006) **21**(3):337-347.
104. Wegener KL, Partridge AW, Han J, Pickford AR, Liddington RC, Ginsberg MH, Campbell ID: **Structural basis of integrin activation by talin.** *Cell* (2007) **128**(1):171-182.
105. Katagiri K, Maeda A, Shimonaka M, Kinashi T: **Rapl, a rap1-binding molecule that mediates rap1-induced adhesion through spatial regulation of lfa-1.** *Nat Immunol* (2003) **4**(8):741-748.
106. Barry WT, Boudignon-Proudnon C, Shock DD, McFadden A, Weiss JM, Sondek J, Parise LV: **Molecular basis of cib binding to the integrin alpha iib cytoplasmic domain.** *J Biol Chem* (2002) **277**(32):28877-28883.
107. Fagerholm SC, Hilden TJ, Gahmberg CG: **P marks the spot: Site-specific integrin phosphorylation regulates molecular interactions.** *Trends Biochem Sci* (2004) **29**(9):504-512.
108. Nurmi SM, Autero M, Raunio AK, Gahmberg CG, Fagerholm SC: **Phosphorylation of the lfa-1 integrin beta2-chain on thr-758 leads to adhesion, rac-1/cdc42 activation, and stimulation of cd69 expression in human t cells.** *J Biol Chem* (2007) **282**(2):968-975.

109. Fagerholm S, Morrice N, Gahmberg CG, Cohen P: **Phosphorylation of the cytoplasmic domain of the integrin cd18 chain by protein kinase c isoforms in leukocytes.** *J Biol Chem* (2002) **277**(3):1728-1738.
110. Czuchra A, Meyer H, Legate KR, Brakebusch C, Fassler R: **Genetic analysis of beta1 integrin "activation motifs" in mice.** *J Cell Biol* (2006) **174**(6):889-899.
111. Takala H, Nurminen E, Nurmi SM, Aatonen M, Strandin T, Takatalo M, Kiema T, Gahmberg CG, Ylänne J, Fagerholm SC: **Beta2 integrin phosphorylation on thr758 acts as a molecular switch to regulate 14-3-3 and filamin binding.** *Blood* (2008) **112**(5):1853-1862.
112. Oxley CL, Anthis NJ, Lowe ED, Vakonakis I, Campbell ID, Wegener KL: **An integrin phosphorylation switch: The effect of beta3 integrin tail phosphorylation on dok1 and talin binding.** *J Biol Chem* (2008) **283**(9):5420-5426.
113. Deshmukh L, Gorbatyuk V, Vinogradova O: **Integrin {beta}3 phosphorylation dictates its complex with the shc phosphotyrosine-binding (ptb) domain.** *J Biol Chem* (2010) **285**(45):34875-34884.
114. Tabeta K, Georgel P, Janssen E, Du X, Hoebe K, Crozat K, Mudd S, Shamel L, Sovath S, Goode J, Alexopoulou L *et al*: **Toll-like receptors 9 and 3 as essential components of innate immune defense against mouse cytomegalovirus infection.** *Proc Natl Acad Sci U S A* (2004) **101**(10):3516-3521.
115. Liu S, Ginsberg MH: **Paxillin binding to a conserved sequence motif in the alpha 4 integrin cytoplasmic domain.** *J Biol Chem* (2000) **275**(30):22736-22742.
116. Han J, Rose DM, Woodside DG, Goldfinger LE, Ginsberg MH: **Integrin alpha 4 beta 1-dependent t cell migration requires both phosphorylation and dephosphorylation of the alpha 4 cytoplasmic domain to regulate the reversible binding of paxillin.** *J Biol Chem* (2003) **278**(37):34845-34853.
117. Rose DM: **The role of the alpha4 integrin-paxillin interaction in regulating leukocyte trafficking.** *Exp Mol Med* (2006) **38**(3):191-195.
118. Feral CC, Rose DM, Han J, Fox N, Silverman GJ, Kaushansky K, Ginsberg MH: **Blocking the alpha 4 integrin-paxillin interaction selectively impairs mononuclear leukocyte recruitment to an inflammatory site.** *J Clin Invest* (2006) **116**(3):715-723.
119. Calderwood DA, Tai V, Di Paolo G, De Camilli P, Ginsberg MH: **Competition for talin results in trans-dominant inhibition of integrin activation.** *J Biol Chem* (2004) **279**(28):28889-28895.

120. Ye F, Hu G, Taylor D, Ratnikov B, Bobkov AA, McLean MA, Sligar SG, Taylor KA, Ginsberg MH: **Recreation of the terminal events in physiological integrin activation.** *J Cell Biol* (2010) **188**(1):157-173.
121. Garcia-Alvarez B, de Pereda JM, Calderwood DA, Ulmer TS, Critchley D, Campbell ID, Ginsberg MH, Liddington RC: **Structural determinants of integrin recognition by talin.** *Mol Cell* (2003) **11**(1):49-58.
122. Bouaouina M, Lad Y, Calderwood DA: **The n-terminal domains of talin cooperate with the phosphotyrosine binding-like domain to activate beta1 and beta3 integrins.** *J Biol Chem* (2008) **283**(10):6118-6125.
123. Anthis NJ, Wegener KL, Critchley DR, Campbell ID: **Structural diversity in integrin/talin interactions.** *Structure* (2010) **18**(12):1654-1666.
124. Anthis NJ, Wegener KL, Ye F, Kim C, Goult BT, Lowe ED, Vakonakis I, Bate N, Critchley DR, Ginsberg MH, Campbell ID: **The structure of an integrin/talin complex reveals the basis of inside-out signal transduction.** *Embo J* (2009) **28**(22):3623-3632.
125. Goult BT, Bouaouina M, Elliott PR, Bate N, Patel B, Gingras AR, Grossmann JG, Roberts GC, Calderwood DA, Critchley DR, Barsukov IL: **Structure of a double ubiquitin-like domain in the talin head: A role in integrin activation.** *Embo J* (2010) **29**(6):1069-1080.
126. Elliott PR, Goult BT, Kopp PM, Bate N, Grossmann JG, Roberts GC, Critchley DR, Barsukov IL: **The structure of the talin head reveals a novel extended conformation of the ferm domain.** *Structure* (2010) **18**(10):1289-1299.
127. Tang RH, Law SK, Tan SM: **Selective recruitment of src family kinase hck by leukocyte integrin alphabeta2 but not alphalbeta2 or alphaxbeta2.** *FEBS Lett* (2006) **580**(18):4435-4442.
128. Pace CN, Scholtz JM: **A helix propensity scale based on experimental studies of peptides and proteins.** *Biophys J* (1998) **75**(1):422-427.
129. Martin DD, Beauchamp E, Berthiaume LG: **Post-translational myristoylation: Fat matters in cellular life and death.** *Biochimie* (2011) **93**(1):18-31.
130. Poschner BC, Langosch D: **Stabilization of conformationally dynamic helices by covalently attached acyl chains.** *Protein Sci* (2009) **18**(8):1801-1805.
131. Stephens G, O'Luanaigh N, Reilly D, Harriott P, Walker B, Fitzgerald D, Moran N: **A sequence within the cytoplasmic tail of gpiib independently activates platelet aggregation and thromboxane synthesis.** *J Biol Chem* (1998) **273**(32):20317-20322.

132. Ginsberg MH, Yaspan B, Forsyth J, Ulmer TS, Campbell ID, Slepak M: **A membrane-distal segment of the integrin alpha iib cytoplasmic domain regulates integrin activation.** *J Biol Chem* (2001) **276**(25):22514-22521.
133. Bernard E, Parthasarathi L, Cho MK, Aylward K, Raab M, Daxecker H, O'Dushlaine CT, Shields DC, Devocelle M, Keyes T, Cosgrave L *et al*: **Ligand switching in cell-permeable peptides: Manipulation of the alpha-integrin signature motif.** *ACS Chem Biol* (2009) **4**(6):457-471.
134. Kallick DA, Tessmer MR, Watts CR, Li CY: **The use of dodecylphosphocholine micelles in solution nmr.** *J Magn Reson B* (1995) **109**(1):60-65.
135. Nath SK, Han S, Kim-Howard X, Kelly JA, Viswanathan P, Gilkeson GS, Chen W, Zhu C, McEver RP, Kimberly RP, Alarcon-Riquelme ME *et al*: **A nonsynonymous functional variant in integrin-alpha(m) (encoded by itgam) is associated with systemic lupus erythematosus.** *Nat Genet* (2008) **40**(2):152-154.
136. Bullard DC, Hu X, Schoeb TR, Axtell RC, Raman C, Barnum SR: **Critical requirement of cd11b (mac-1) on t cells and accessory cells for development of experimental autoimmune encephalomyelitis.** *J Immunol* (2005) **175**(10):6327-6333.
137. Pluskota E, Soloviev DA, Szpak D, Weber C, Plow EF: **Neutrophil apoptosis: Selective regulation by different ligands of integrin alphabeta2.** *J Immunol* (2008) **181**(5):3609-3619.
138. Lamed R, Kenig R, Morag E, Yaron S, Shoham Y, Bayer EA: **Nonproteolytic cleavage of aspartyl proline bonds in the cellulosomal scaffoldin subunit from clostridium thermocellum.** *Appl Biochem Biotechnol* (2001) **90**(1):67-73.
139. Goddard TD, Kneller DG: **Sparky. Graphical nmr assignment and integration program.** In: University of California, San Francisco.
140. Berjanskii MV, Neal S, Wishart DS: **Preditor: A web server for predicting protein torsion angle restraints.** *Nucleic Acids Res* (2006) **34**(Web Server issue):W63-69.
141. Wüthrich K: *Nmr of proteins and nucleic acids.* Wiley, New York (1986).
142. Güntert P: **Automated nmr protein structure calculation with cyana.** *Methods Mol Biol* (2004) **278**(353-378).
143. Laskowski RA, M.M.W., Moss DS, Thornton JM: **Procheck-a program to check the stereochemical quality of protein structures.** *J App Cryst* (1993) **26**(283-291).

144. Delano WL, and Lam, J. W.: **Pymol v0.99**. In: DeLano Scientific LLC.
145. Koradi R: **Molmol 2k.2**. In: E. Z. Institut fuer Molekularbiologie und Biophysik, Faellanden, Switzerland ((1999-2003)).
146. Lyskov S, Gray JJ: **The rosettadock server for local protein-protein docking**. *Nucleic Acids Res* (2008) **36**(Web Server issue):W233-238.
147. Hilty C, Wider G, Fernandez C, Wuthrich K: **Membrane protein-lipid interactions in mixed micelles studied by nmr spectroscopy with the use of paramagnetic reagents**. *Chembiochem* (2004) **5**(4):467-473.
148. Ikura M, Kay LE, Bax A: **A novel approach for sequential assignment of 1h, 13c, and 15n spectra of proteins: Heteronuclear triple-resonance three-dimensional nmr spectroscopy. Application to calmodulin**. *Biochemistry* (1990) **29**(19):4659-4667.
149. Rule GS, Hitchens TK: *Fundamentals of protein nmr spectroscopy*. Springer, Dordrecht (2006).
150. Lakowicz JR: *Principles of fluorescence spectroscopy*. Springer, New York (2006).
151. Haas TA, Plow EF: **The cytoplasmic domain of alphaIIb beta3. A ternary complex of the integrin alpha and beta subunits and a divalent cation**. *J Biol Chem* (1996) **271**(11):6017-6026.
152. Clore GM, Tang C, Iwahara J: **Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement**. *Curr Opin Struct Biol* (2007) **17**(5):603-616.
153. Cavanagh J: *Protein nmr spectroscopy : Principles and practice*. Academic Press, Amsterdam ; Boston (2007).
154. Bulheller BM, Rodger A, Hirst JD: **Circular and linear dichroism of proteins**. *Phys Chem Chem Phys* (2007) **9**(17):2020-2035.
155. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD: **1h, 13c and 15n random coil nmr chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects**. *J Biomol NMR* (1995) **5**(1):67-81.
156. Schwarzing S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ: **Sequence-dependent correction of random coil nmr chemical shifts**. *J Am Chem Soc* (2001) **123**(13):2970-2978.
157. Bienkiewicz EA, Lumb KJ: **Random-coil chemical shifts of phosphorylated amino acids**. *J Biomol NMR* (1999) **15**(3):203-206.
158. van Dijk AD, Boelens R, Bonvin AM: **Data-driven docking for the study of biomolecular complexes**. *Febs J* (2005) **272**(2):293-312.

159. Kaufmann KW, Lemmon GH, Deluca SL, Sheehan JH, Meiler J: **Practically useful: What the rosetta protein modeling suite can do for you.** *Biochemistry* (2010) **49**(14):2987-2998.
160. Chan BM, Kassner PD, Schiro JA, Byers HR, Kupper TS, Hemler ME: **Distinct cellular functions mediated by different α integrin subunit cytoplasmic domains.** *Cell* (1992) **68**(6):1051-1060.
161. Vallar L, Melchior C, Plancon S, Drobecq H, Lippens G, Regnault V, Kieffer N: **Divalent cations differentially regulate integrin α IIb cytoplasmic tail binding to β 3 and to calcium- and integrin-binding protein.** *J Biol Chem* (1999) **274**(24):17257-17266.
162. Kishimoto TK, Hollander N, Roberts TM, Anderson DC, Springer TA: **Heterogeneous mutations in the β subunit common to the α 1, α 5, and α 10 glycoproteins cause leukocyte adhesion deficiency.** *Cell* (1987) **50**(2):193-202.
163. Vorup-Jensen T, Carman CV, Shimaoka M, Schuck P, Svitel J, Springer TA: **Exposure of acidic residues as a danger signal for recognition of fibrinogen and other macromolecules by integrin α 5 β 2.** *Proc Natl Acad Sci U S A* (2005) **102**(5):1614-1619.
164. Vorup-Jensen T, Ostermeier C, Shimaoka M, Hommel U, Springer TA: **Structure and allosteric regulation of the α 5 β 2 integrin α domain.** *Proc Natl Acad Sci U S A* (2003) **100**(4):1873-1878.
165. Yamanouchi J, Hato T, Tamura T, Fujita S: **Suppression of integrin activation by the membrane-distal sequence of the integrin α IIb cytoplasmic tail.** *Biochem J* (2004) **379**(Pt 2):317-323.
166. Arroyo AG, Yang JT, Rayburn H, Hynes RO: **Differential requirements for α 4 integrins during fetal and adult hematopoiesis.** *Cell* (1996) **85**(7):997-1008.
167. Jackson DY: **α 4 integrin antagonists.** *Curr Pharm Des* (2002) **8**(14):1229-1253.
168. Mould AP, Askari JA, Craig SE, Garratt AN, Clements J, Humphries MJ: **Integrin α 4 β 1-mediated melanoma cell adhesion and migration on vascular cell adhesion molecule-1 (vcam-1) and the alternatively spliced α 4 region of fibronectin.** *J Biol Chem* (1994) **269**(44):27224-27230.
169. Chigaev A, Zwartz G, Graves SW, Dwyer DC, Tsuji H, Foutz TD, Edwards BS, Prossnitz ER, Larson RS, Sklar LA: **α 4 β 1 integrin affinity changes govern cell adhesion.** *J Biol Chem* (2003) **278**(40):38174-38182.
170. Rose DM, Han J, Ginsberg MH: **α 4 integrins and the immune response.** *Immunol Rev* (2002) **186**:118-124.

171. Brown MC, Turner CE: **Paxillin: Adapting to change.** *Physiol Rev* (2004) **84**(4):1315-1339.
172. Liu S, Kiosses WB, Rose DM, Slepak M, Salgia R, Griffin JD, Turner CE, Schwartz MA, Ginsberg MH: **A fragment of paxillin binds the alpha 4 integrin cytoplasmic domain (tail) and selectively inhibits alpha 4-mediated cell migration.** *J Biol Chem* (2002) **277**(23):20887-20894.
173. Goldfinger LE, Han J, Kiosses WB, Howe AK, Ginsberg MH: **Spatial restriction of alpha4 integrin phosphorylation regulates lamellipodial stability and alpha4beta1-dependent cell migration.** *J Cell Biol* (2003) **162**(4):731-741.
174. Baynes BM, Wang DI, Trout BL: **Role of arginine in the stabilization of proteins against aggregation.** *Biochemistry* (2005) **44**(12):4919-4925.
175. Golovanov AP, Hautbergue GM, Wilson SA, Lian LY: **A simple method for improving protein solubility and long-term stability.** *J Am Chem Soc* (2004) **126**(29):8933-8939.
176. Penzkofer A, Lu Y: **Fluorescence quenching of rhodamine 6g in methanol in high concentration.** *Chemical Physics* (1986) **103**(399-405).
177. Lorenz S, Vakonakis I, Lowe ED, Campbell ID, Noble ME, Hoellerer MK: **Structural analysis of the interactions between paxillin ld motifs and alpha-parvin.** *Structure* (2008) **16**(10):1521-1531.
178. Hoellerer MK, Noble ME, Labesse G, Campbell ID, Werner JM, Arold ST: **Molecular recognition of paxillin ld motifs by the focal adhesion targeting domain.** *Structure* (2003) **11**(10):1207-1217.
179. Li X, Ji W, Zhang R, Folta-Stogniew E, Min W, Boggon TJ: **Molecular recognition of leucine-aspartate repeat (ld) motifs by the focal adhesion targeting homology domain of cerebral cavernous malformation 3 (ccm3).** *J Biol Chem* (2011) **286**(29):26138-26147.
180. Luster AD, Alon R, von Andrian UH: **Immune cell migration in inflammation: Present and future therapeutic targets.** *Nat Immunol* (2005) **6**(12):1182-1190.
181. Elhasid R, Rowe JM: **Hematopoietic stem cell transplantation in neutrophil disorders: Severe congenital neutropenia, leukocyte adhesion deficiency and chronic granulomatous disease.** *Clin Rev Allergy Immunol* (2010) **38**(1):61-67.
182. Cox D, Brennan M, Moran N: **Integrins as therapeutic targets: Lessons and opportunities.** *Nat Rev Drug Discov* (2010) **9**(10):804-820.

183. Mousa SA: **Anti-integrin as novel drug-discovery targets: Potential therapeutic and diagnostic implications.** *Curr Opin Chem Biol* (2002) **6**(4):534-541.
184. Kummer C, Petrich BG, Rose DM, Ginsberg MH: **A small molecule that inhibits the interaction of paxillin and alpha 4 integrin inhibits accumulation of mononuclear leukocytes at a site of inflammation.** *J Biol Chem* (2010) **285**(13):9462-9469.
185. Wales TE, Engen JR: **Hydrogen exchange mass spectrometry for the analysis of protein dynamics.** *Mass Spectrom Rev* (2006) **25**(1):158-170.
186. Vogel A, Scheidt HA, Huster D: **The distribution of lipid attached spin probes in bilayers: Application to membrane protein topology.** *Biophys J* (2003) **85**(3):1691-1701.