A structural hypothesis for the transition between bent and extended conformations of the leukocyte $\beta 2$ integrins

Shi, Minlong

2008

Shi, M. (2008). A structural hypothesis for the transition between bent and extended conformations of the leukocyte $\beta 2$ integrins. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/6564>

<https://doi.org/10.32657/10356/6564>

Nanyang Technological University

Downloaded on 26 Apr 2023 13:57:39 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**A STRUCTURAL HYPOTHESIS FOR THE
TRANSITION BETWEEN BENT AND EXTENDED
CONFORMATIONS OF THE LEUKOCYTE $\beta 2$
INTEGRIN**

**SHI MINLONG
SCHOOL OF BIOLOGICAL SCIENCES
2008**

A Structural Hypothesis For The Transition Between Bent And Extended Conformations Of Leukocyte $\beta 2$ Integrin

Shi Minlong

SCHOOL OF BIOLOGICAL SCIENCES

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

2008

A structural hypothesis for the transition between bent and extended conformations of the leukocyte $\beta 2$ integrins

Shi Minlong

A thesis presented to Nanyang Technological University
in candidacy for the degree of doctor of philosophy (PhD)

Recommended for acceptance
by the school of biological sciences

September 2007

Acknowledgements

The present thesis was carried out at the School of Biological Science, Nanyang Technological University, under the supervision of Professor S.K. Alex Law, Ph.D.

I would like to thank Professor Law for his direction, support, invaluable criticisms and the freedom he offered me during my research. Without these, I could not come to this stage. His attitude towards the science, his consideration and kindness to people impress me and four years' training under his supervision will benefit me for a long time.

Dr. Julien Lascar deserves many thanks for his nice collaborations on my projects and his training on X-ray crystallography.

I also owe many thanks to Dr. Tan Suet Mien for his support and useful helps. I also appreciate his valuable suggestion on my research and life.

My warm thanks to Dr. Li Jinming is for his generous help and suggestions for the installation of the crystallography softwares.

Thanks to all my former and present colleagues for the pleasant working atmosphere at the lab. Without them, my project could not be finished so smoothly. Many thanks go to Chan Hwee Sing, Foo Shen Yun, Cheng Ming, Manisha Cooray, Elianna Bte Mohamed Amin, Tang Renhong, Geng Xin, Li Yanfeng, Tang Xiaoyan and Florence Lim Wei Ning. They made my graduate life in Nanyang Technological University unforgettable.

In addition, I want to thank the staffs at ESRF, Grenoble, France, for providing the brilliant synchrotron beamline for the diffraction experiments.

Last but not least, I want to thank my parents sincerely for their endless love, support and understanding during my Ph.D. study.

I gratefully thank School of Biological Sciences and Nanyang Technological University for the support during my Ph.D. study.

This dissertation is dedicated to all those who care about me and for whom I care!

Contents

ACKNOWLEDGEMENTS	I
CONTENTS.....	III
LIST OF FIGURES.....	VII
ABBREVIATIONS	X
ABSTRACT.....	XII
CHAPTER ONE: INTRODUCTION.....	1
1.1 OVERVIEW	1
1.2 INTEGRIN STRUCTURE	3
1.2.1 <i>The α subunit</i>	8
1.2.1.1 The β -propeller domain.....	8
1.2.1.2 The I domain.....	10
1.2.1.3 The conformational states of I domain.....	11
1.2.1.4 The leg region of α subunit	13
1.2.2 <i>The β subunit</i>	15
1.2.2.1 The PSI domain.....	15
1.2.2.2 The hybrid domain	16
1.2.2.3 The β I domain.....	18
1.2.2.4 The I-EGF domains.....	23
1.2.2.5 The β -TD (β -tail domain) domain.....	24
1.2.3 <i>The transmembrane (TM) domains</i>	24
1.2.4 <i>The cytoplasmic domains</i>	25
1.3 THE GLOBAL CONFORMATIONAL CHANGES IN THE INTEGRIN ECTODOMAIN	29
1.4 OBJECTIVE OF MY THESIS	34
CHAPTER TWO: MATERIALS AND METHODS	35
2.1 GENERAL REAGENTS.....	35
2.1.1 <i>Enzymes</i>	35
2.1.2 <i>Commercially available kit</i>	35
2.1.3 <i>cDNA clones</i>	36
2.1.4 <i>Cells</i>	36
2.1.5 <i>Expression Vectors</i>	36
2.1.6 <i>Monoclonal antibodies</i>	37
2.2 SOLUTIONS, BUFFERS, AND MEDIA	38
2.2.1 <i>Laboratory stocks</i>	38
2.2.2 <i>Media</i>	38
2.2.3 <i>Solutions</i>	39
2.3 METHODS.....	41
2.3.1 <i>General Methods for Biochemistry</i>	41
2.3.1.1 Quantitation of DNA.....	41
2.3.1.2 Polymerase chain reaction (PCR)	41
2.3.1.3 Standard PCR protocol.....	41

Contents

2.3.1.4 Restriction endonuclease digestion	42
2.3.1.5 DNA separation by agarose gel electrophoresis	42
2.3.1.6 Purification of DNA fragments by agarose gel electrophoresis	43
2.3.1.7 DNA ligation.....	43
2.3.1.8 Preparation of <i>E.coli</i> competent cells.....	43
2.3.1.9 Transformation of plasmid DNA	44
2.3.1.10 Identification of colonies that contain recombinant plasmids of interest	44
2.3.1.11 Purification of plasmid DNA	45
2.3.1.12 Site-directed mutagenesis	45
2.3.1.13 Small scale protein induction.....	46
2.3.1.14 Large scale protein induction.....	46
2.3.1.15 Cell lysis	46
2.3.1.16 Purification of protein by Ni ²⁺ -NTA affinity column	47
2.3.1.17 Protein concentration assay	47
2.3.2 <i>General methods for cell culture</i>	48
2.3.2.1 Cell storage in liquid nitrogen	48
2.3.2.2 Cell recovery from liquid nitrogen.....	48
2.3.2.3 Culture of HEK293, HEK293T and HEK293GnTI ⁻ cells	49
2.3.3 <i>β2 fragments expression, purification and crystallization</i>	49
2.3.3.1 Transfection of HEK293 and HEK293T cells by the calcium phosphate method (for 10 cm dish).....	49
2.3.3.2 Transfection of HEK293T and HEK293 GnTI ⁻ cells by PolyFect method (for 6cm dish)	50
2.3.3.3 Creation of stable cell line	50
2.3.3.4 Recombinant protein expression and purification	51
2.3.3.5 Deglycosylation of the recombinant glycoproteins.....	52
2.3.3.6 Crystallization and post-crystallization treatments	53
2.3.3.7 Heavy atom derivatives	55
2.3.3.8 Data collection	56
2.3.3.9 Structure determination and refinement	56
2.3.4 <i>Cell based and biochemical experiments</i>	58
2.3.4.1 Harvesting transfected cells (adherent)	58
2.3.4.2 FACS analysis.....	58
2.3.4.3 Preparation of rabbit anti-mouse IgG coupled onto protein A sephrose beads	59
2.3.4.4 Labelling of free cysteines.....	59
2.3.4.5 Surface biotinylation and preparation of whole cell lysates	60
2.3.4.6 Immunoprecipitation	60
2.3.4.7 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	61
2.3.4.8 Western blot.....	61
2.3.4.9 ECL detection proteins blotted onto immobilon-P membrane	62
2.3.4.10 Coating microtitre plates with ICAMs for cell adhesion assay	62
2.3.4.11 Cell adhesion assays	63
2.4 PLASMID CONSTRUCTION DETAILS.....	63
2.4.1 <i>Integrin plasmids</i>	63

Contents

2.4.1.1 Wild type plamids	63
2.4.1.2 Hybrid domain (HD) constructs.....	64
2.4.1.3 PHE1 expression vector	64
2.4.1.4 PHE2-pIRES2-EGFP and PHE3-pIRES2-EGFP expression vector.....	66
CHAPTER THREE: THE STRUCTURE OF PHE1	68
3.1 EXPRESSION AND PURIFICATION	70
3.1.1 <i>The HD expression in E.coli</i>	70
3.1.2 <i>The PHE1 expression in E.coli BL21(DE3)</i>	74
3.1.3 <i>Transient expression and purification of HD in HEK293T cells</i>	75
3.1.4 <i>Transient expression and purification of PHE1 in HEK293T cells</i>	77
3.1.5 <i>Creation of the PHE1-HEK293 stable cell line</i>	78
3.1.6 <i>The PHE1 expression and purification</i>	78
3.2 DEGLYCOSYLATION AND MASS SPECTROMETRY	81
3.3 PHE1 CRYSTALLIZATION AND DATA COLLECTION	82
3.4 PHE1 STRUCTURE SOLUTION AND REFINEMENT	85
3.5 THE STRUCTURE OF PHE1.....	86
3.5.1 <i>Quality of the model and overall structure</i>	86
3.5.2 <i>The structure of the hybrid domain</i>	89
3.5.3 <i>The structure of the PSI domain</i>	91
3.5.4 <i>The structure of the I-EGF1 domain</i>	92
3.5.5 <i>Interactions between domains</i>	94
3.5.6 <i>Oligosaccharide chains of PHE1</i>	97
3.6 DISCUSSION.....	98
3.6.1 <i>Constructions and expressions</i>	98
3.6.2 <i>PHE1 structure</i>	99
CHAPTER FOUR: THE STRUCTURE OF PHE2 AND PHE3.....	103
4.1 EXPRESSION AND PURIFICATION	104
4.2 PHE2 AND PHE3 CRYSTALLIZATION AND DATA COLLECTION	111
4.3 PHE2 AND PHE3 STRUCTURE REFINEMENT AND DETERMINATION	114
4.4 THE STRUCTURE OF PHE2.....	114
4.4.1 <i>Quality of the model and overall structure</i>	114
4.4.2 <i>The structure of the I-EGF2 domain</i>	118
4.4.3 <i>The structural differences between PHE1 and PHE2</i>	118
4.4.4 <i>Oligosaccharide chains of PHE2</i>	122
4.4.5 <i>Residues with alternative side chain conformations</i>	123
4.5 THE STRUCTURE OF PHE3.....	124
4.5.1 <i>Quality of the model and overall structure</i>	124
4.5.2 <i>The structure of the I-EGF2 domain</i>	127
4.5.3 <i>The structure of the I-EGF3 domain</i>	129
4.5.4 <i>The conformational changes between PHE2 and PHE3 structures</i>	129
4.5.5 <i>Interaction between domains</i>	131
4.5.6 <i>Oligosaccharide chains of PHE3</i>	135
4.6 THE BENT CONFORMATION OBSERVED IN THE PHE2 CRYSTAL STRUCTURE CAN BE ADOPTED BY A COMPLETE INTEGRIN HETERODIMER.	136

Contents

4.7 THE α L β 2G33C/G486C VARIANT CAN MEDIATE ADHESION TO ICAM-1.	141
4.8 DISCUSSION.....	142
4.8.1 <i>Constructions and expressions</i>	142
4.8.2 <i>PHE2 and PHE3 structures</i>	143
4.8.3 <i>KIM127, a conformation sensitive integrin mAb</i>	146
CHAPTER FIVE: SUMMARY	149
PUBLICATIONS RELATED TO THIS STUDY	151
REFERENCES.....	152
APPENDIX 1: THE DNA SEQUENCE OF INTEGRIN β2 SUBUNIT (CD18) AND ITS AMINO ACID SEQUENCE WITH THE SIGNAL PEPTIDE HIGHLIGHTED IN RED.....	165
APPENDIX 2: THE CONSTRUCTS AND THEIR AMINO ACID SEQUENCE WITH POTENTIAL GLYCOSYLATION SITES ANALYSED BY THE VECTOR NTI AND THE VECTORS USED IN THIS THESIS FOR <i>E. COLI</i> EXPRESSION	169
APPENDIX 3: THE RECOMBINANT PROTEINS USED IN THIS THESIS WERE ANALYZED BY VECTOR NTI.....	174
APPENDIX 4: THE PRESUMED MODELS OF INTEGRIN STRUCTURES.....	175
APPENDIX 5: HUMAN INTEGRIN SUBUNITS	177

List of Figures

Figure 1. 1 The integrin family	2
Figure 1. 2 Integrin architecture.....	6
Figure 1. 3 Crystal structure of the extracellular segment of $\alpha V\beta 3$	7
Figure 1. 4 Ribbon diagrams of the αV integrin β -propeller domain.....	9
Figure 1. 5 Structural comparison of the closed and open αM I domain.....	12
Figure 1. 6 The structure of the integrin $\alpha V\beta 3$ Genu	14
Figure 1. 7 Superposition of the headpieces from the RGD liganded legless $\alpha IIb\beta 3$ (<i>green</i>) and bent, unliganded $\alpha V\beta 3$ (<i>blue</i>)	17
Figure 1. 8 Structure rearrangement of the β I domain	20
Figure 1. 9 The structure ADMIDAS metal Ca^{2+} (<i>sphere</i>) and its coordinations	21
Figure 1. 10 Model for TM and membrane-proximal regions displacements during integrin activation	26
Figure 1. 11 The NMR structure of the integrin tails.....	28
Figure 1. 12 EM images and its presumed models of integrin structures	31
Figure 1. 13 Proposed intermediates in equilibration between known conformational states.....	33
Figure 2. 1 Dehydration.....	54
Figure 2. 2 Schematic illustration of plasmid J8.1E and pcDNA3.0.	65
Figure 2. 3 Schematic illustration of pIRES2-EGFP	66
Figure 3. 1 Constructs of HD and PHE1	69
Figure 3. 2 Hybrid domain expression in <i>E.coli</i> BL21	70
Figure 3. 3 Three HD fusion protein expressions	72
Figure 3. 4 MBP-HD purification and its cleavage by Factor Xa.....	72
Figure 3. 5 Gel filtration purification of the factor Xa digested MBP-HD	73
Figure 3. 6 PHE1 and trx-PHE1 fusion protein expressions in <i>E.coli</i>	74
Figure 3. 7 HD purification.....	76
Figure 3. 8 Purification of PHE1	77
Figure 3. 9 Clones of HEK293 stable cells with the different expression level	78
Figure 3. 10 Immunoprecipitation of PHE1	79
Figure 3. 11 PHE1 purification.....	80
Figure 3. 12 The glycosylation of PHE1	81
Figure 3. 13 PHE1 crystal images.....	82
Figure 3. 14 The structure of PHE1	87
Figure 3. 15 Sequence alignment of human integrin β subunits.....	89
Figure 3. 16 The structure of the hybrid domain	90
Figure 3. 17 Stereo view of the PSI domain superposition.	92
Figure 3. 18 The structure of the I-EGF1 domain.....	93
Figure 3. 19 Hydrogen bond interactions between PHE1 domains	96
Figure 3. 20 The N-linked oligosaccharide chains of PHE1	97
Figure 3. 21 Model for the integrin $\alpha V\beta 3$ in its bent state.....	102

List of Figures

Figure 4. 1 Illustrations of recombinant protein constructs.....	104
Figure 4. 2 PHE2 and PHE3 expressed in different cell lines were detected by anti-His-tag Western blot.	105
Figure 4. 3 Immunoprecipitation of PHE2 and PHE3 with a panel of specific monoclonal antibodies.....	105
Figure 4. 4 PHE2 purification.....	108
Figure 4. 5 PHE3 purification.....	109
Figure 4. 6 Analysis of PHE2 and PHE3 on SDS PAGE before and after treatment of Endo H _f	110
Figure 4. 7 Crystals images of PHE2 (A), PHE3 (B).....	112
Figure 4. 8 The structure of the PHE2.....	115
Figure 4. 9 Independent validation of the PHE2 using the anomalous signal from the sulfur atoms	116
Figure 4. 10 Sequence alignment and topology of the I-EGF2 domain.....	117
Figure 4. 11 The polar interactions between PHE2 domains	120
Figure 4. 12 The N-linked oligosaccharide chains of PHE2 after Endo H _f digestion.....	123
Figure 4. 13 The alternative conformations of the side chains	124
Figure 4. 14 The structure of PHE3	125
Figure 4. 15 Independent validation of the PHE3 using the anomalous signal from the sulfur atoms	126
Figure 4. 16 Superposition of the I-EGF2 and I-EGF3 domains.....	128
Figure 4. 17 I-EGF3 domain sequence alignment.	128
Figure 4. 18 Movements in the quaternary structure	130
Figure 4. 19 The polar interactions between the PSI and I-EGF1 domains.....	130
Figure 4. 20 The polar interactions between I-EGF1 and I-EGF2 domains.....	133
Figure 4. 21 The polar interactions between PHE3 domains	134
Figure 4. 22 The N-linked oligosaccharide chains of PHE3 after Endo H _f digestion.....	135
Figure 4. 23 Introduction of a disulfide bond that covalently links the PSI and I-EGF2 domains in integrin $\beta 2$ subunit	137
Figure 4. 24 LFA-1 and variants expression.....	137
Figure 4. 25 The formation of the disulfide bond in LFA-1 variants was monitored using the free sulphydryl labeling reagent BMCC conjugated to biotin.....	139
Figure 4. 26 Reactivity of KIM127 with LFA-1 and variants.....	140
Figure 4. 27 Immunoprecipitation of LFA-1 in Mg ²⁺ /EGTA and at different concentrations of DTT by KIM127	141
Figure 4. 28 Adhesion of LFA-1 transfectants to ICAM-1	141
Figure 4. 29 I-EGF domains superposition and sequence alignment.....	145

List of Tables

List of Tables

Table 1. 1 $\beta 2$ Integrins on leukocytes	3
Table 2. 1 Commonly used cryo-protectants	54
Table 2. 2 The 23 most commonly used heavy-atom compounds reagents	55
Table 3. 1 PHE1 Data Collection and refinement statistics.....	83
Table 3. 2 Hydrogen bond contacts in PHE1 domains.....	95
Table 4. 1 Mass spectrometry analyses of PHE1, PHE2 and PHE3	110
Table 4. 2 Optimal conditions for PHE2 and PHE3 crystal growth	112
Table 4. 3 Data Collection and refinement statistics for PHE2 and PHE3	113
Table 4. 4 The r.m.s. deviations of domain superimpositions in integrin $\beta 2$ fragments (PHE1, PHE2 and PHE3).....	119
Table 4. 5 hydrogen bond interactions between the I-EGF1 and PSI, and I-EGF2 domains in PHE2	121
Table 4. 6 Buried surface areas between individual domains for integrin fragments PHE1, PHE2 and PHE3.....	122
Table 4. 7 hydrogen bond contacts between domains in PHE3.....	132

Abbreviations

BSA	bovine serum albumin
CCP4	collaborative computational project number 4
CD	cluster of differentiation
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECL	enhanced chemiluminescence
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	ethylene-diamine-tetracetic acid
EGTA	ethylene-glycol-bis (β -aminoethylether)-tetracetic acid
EM	electron microscope
FACS	fluorescent activated cell sorter
HEPES	N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]
ICAM	intercellular cell-adhesion molecule
Ig	immunoglobulin
IP	immunoprecipitation
IPTG	isopropyl- β -D-1-thiogalactoside
LB	Luria broth
LFA-1	leukocyte function associated antigen-1
mAb	monoclonal antibody
ME	Mg ²⁺ /EGTA solution
MW	molecular weight
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
O.D.	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulphonylfluoride
r.m.s deviation	root mean square deviation
RNA	ribonucleic acid
rpm	revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TEMED	N,N,N,N,-tetramethylethylenediamine
Tris	Tris (hydroxymethyl)-aminoethane
UV	ultraviolet
v/v	volume per volume
WT	wild-type
w/v	weight per volume

Abbreviations

One and three letter symbols for the for amino acid residues:

A	Ala	alanine
C	Cys	cysteine
D	Asp	aspartic acid
E	Glu	glutamic acid
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophan
Y	Tyr	tyrosine

Other abbreviations are defined in the text where first encountered.

Abstract

Integrins are ($\alpha\beta$) heterodimeric modular proteins which mediate cell adhesion and convey signals across the plasma membrane. Inter-domain motions have been proposed to play a key role in signal transduction by propagating structural changes through the molecule, thus controlling the activation state and adhesive properties of the integrin. It is thought that a compact bent conformation of the molecule represents its physiological low-affinity state and an extended conformation its active state.

I first tackled the PHE1 fragment of the human integrin $\beta 2$ subunit comprising the PSI/Hybrid domain/I-EGF1 fragment. It was expressed as a soluble protein and its structure was determined by x-ray diffraction at 1.8 Å resolution. The structure reveals an elongated molecule with a rigid architecture stabilized by nine disulfide bonds. The hybrid domain/PSI interface involves a buried Arg86 residue and contacts between PSI and I-EGF1 are mainly mediated by well conserved Arg39 and Trp23 residues. Preservation of key interacting residues across the various integrin β subunits sequences suggests that our structure represents a good model for the entire integrin family. Superposition with the integrin $\beta 3$ receptor in its low affinity (bent) conformation suggests that the bend in integrin $\beta 2$ subunit is present at the linkage between its I-EGF1 and I-EGF2 modules and points to this region as an important regulator of integrin mediated cell-adhesion.

Based on the PHE1 structure, we have further determined the structures of two fragments of the integrin $\beta 2$ subunit. The first fragment (PHE2), consisting of the PSI, hybrid, I-

Abstract

EGF1 and I-EGF2 domains, showed an L-shaped conformation with the bend located between the I-EGF1 and I-EGF2 domains. The second fragment (PHE3), which includes, in addition to PHE2, the I-EGF3 domain, showed an extended conformation. The major reorientation of I-EGF2 with respect to the other domains in the two structures is accompanied by a change of torsion angle of the disulfide bond between Cys461-Cys492 by 180° and the conversion of a short α -helix (residues Ser468-Cys475) into a flexible coil. Based on the PHE2 structure, we introduced a disulfide bond between the PSI and I-EGF2 domains in the integrin β 2 subunit. The resultant α L β 2 integrin (LFA-1) variant is locked in a bent state and cannot be detected with the mAb KIM127 in Mg^{2+} /EGTA, an antibody which was been shown to bind only to the extended β 2 integrins. However, it retains the binding activity to ICAM-1. These results provide a structural hypothesis for our understanding of the transition between the resting and active states of LFA-1.

Chapter One: Introduction

1.1 Overview

Integrins are cell surface receptors which play crucial roles in cell-cell, cell-extracellular matrix (ECM) and cell-pathogen interactions (Hynes, 2002). Moreover, their importance in tumor metastasis is increasingly appreciated (Gupta et al., 2006). Integrins are present in all metazoans, including sponges and cnidaria, and the number of the integrin subunits increases with the complexity of the organism (Bokel et al., 2002). In mammals, 18 α subunits and 8 β subunits associate to form at least 24 distinct integrins (Hynes, 2002) (Fig.1.1), and only specific combinations are allowed to form stable heterodimers. However the rules which regulate heterodimer formation remain a mystery. There is no detectable homology between the integrin α and β subunits. Sequence identity is about 30% among the α subunits, and 45% among the β subunits, respectively, indicating that the gene families coding for the α and β subunits evolved by gene duplication (Takada et al., 2007).

All four members of the $\beta 2$ integrins (CD11a/CD18 or $\alpha L\beta 2$; CD11b/CD18 or $\alpha M\beta 2$; CD11c/CD18 or $\alpha X\beta 2$ and CD11d/CD18 or $\alpha D\beta 2$) are expressed exclusively in the leukocytes (Gahmberg et al., 1997; Larson et al., 1990b). These integrins are critical for the transmigration of leukocytes from the circulation to injury or infection sites, activation of neutrophils and monocytes, the phagocytosis of foreign materials and neutrophil apoptosis (Larson et al., 1990a). Patients with Leukocyte Adhesion Deficiency type 1 (LAD-1) who suffer from recurrent bacterial and fungal infections to varying degrees of severity (Arnaout, 1990).

Figure 1. 1 The integrin family

The integrins expressed on leukocytes are shown with blue lines and the α subunits containing the I domain are highlighted with a red star.

Integrins on circulating leukocytes are normally in a resting state of low adhesiveness, but can be rapidly activated in response to internal (inside-out signaling) or external (outside-in signaling) (Shimaoka et al., 2002). LFA-1 is expressed on T-cells and mediates its adhesion to and migration across the endothelium once the T-cell receptor has been activated (Dustin et al., 1989; Pribila et al., 2004). A characteristic feature of the integrins is that each integrin is able to bind a repertoire of structurally diverse ligands,

Chapter One: Introduction

and many integrin ligands have the ability to bind to multiple integrin receptors. For example, the integrin $\alpha M\beta 2$ can bind to ICAM-1, iC3b, fibrinogen and factor X. ICAM-1 is also the ligand of the integrin $\alpha L\beta 2$, $\alpha M\beta 2$ and $\alpha D\beta 2$. In addition, the leukocyte integrins exist in resting and adhesive states and conformational changes of the integrin have now been generally accepted as a key feature in the conversion between the two states. The $\beta 2$ integrins on leukocytes and their major ligands are summarized in Table 1.1.

Table 1. 1 $\beta 2$ Integrins on leukocytes (from Luo et al., 2007)

Integrin	Distribution	Major ligands
$\alpha L\beta 2$, LFA-1, CD11a/CD18	Lymphocytes, NK cells, monocytes, macrophages, dendritic cells, neutrophils	ICAM-1, -2, -3, -5
$\alpha M\beta 2$, Mac-1, CR3, CD11b/CD18	Monocytes, macrophages, neutrophils, NK cells	Factor X, iC3b, heparin, fibrinogen, many others
$\alpha X\beta 2$, p150,95, CR4, CD11c/CD18	Monocytes, macrophages, NK cells, dendritic cells	iC3b, fibrinogen, heparin, many others
$\alpha D\beta 2$, CD11d/CD18	Monocytes, macrophages, eosinophils, neutrophils	ICAM-3, VCAM-1

1.2 Integrin structure

Integrin molecules are formed by two non-covalently associated α and β subunits, both are type I membrane glycoproteins. Each subunit typically contains a relatively large extracellular domain forming a globular ligand-binding “head” linked to a pair of “legs”, which anchor the integrin via two transmembrane domains and short cytoplasmic tails

Chapter One: Introduction

(Fig.1.2) (Nermut et al., 1988; Xiong et al., 2001). An exception is the integrin $\beta 4$ which has a cytoplasmic domain containing fibronectin type III domains and connected to the keratin cytoskeleton (de Pereda et al., 1999; Shimaoka et al., 2002).

The crystal structure of the ectodomain of the integrin $\alpha V\beta 3$ revealed its modular architecture (Xiong et al., 2001; Xiong et al., 2004). The interaction between the β -propeller of the α subunit and the β I domain (also referred to as β A domain, I-like domain) forms the head that contains the binding site for the Arg-Gly-Asp (RGD) ligand (Xiong et al., 2002). In addition, the α subunit leg contains three β -sandwich domains, namely, Thigh domain, Calf-1 domain and Calf-2 domain. The β subunit leg comprises a hybrid domain, a PSI (plexin/semaphorin/integrin) domain (Bork et al., 1999), and four I-EGF (integrin-epidermal growth factor) repeats followed by a β -tail domain at its C-terminal end (Fig.1.3). Remarkably, many earlier predicated structural and functional domains are confirmed in this crystal structure. For example, the four Ca^{2+} -binding regions in αV exists in the β -propeller domain as predicted by Springer (Springer, 1997). The domain formed by the N-terminal region and the mid-region in $\beta 2$ integrin, later called hybrid domain (Tan et al., 2001a), is also present in the $\alpha V\beta 3$ crystal structure. In addition, the crystal structure showed an unexpected V-shaped conformation (Xiong et al., 2001), rather than the linear or extended as depicted in previous electron microscopic images (Nermut et al., 1988), so that the ligand-binding head folded back close to the membrane proximal region. One possible reason for the severe bending was that the integrin possesses intrinsic flexibility at the Genu which permitted the bending in solution, and the crystal packing forces favored the bent conformation over others so that

Chapter One: Introduction

the crystal structure was a crystal artifact (Liddington et al., 2002). Another interpretation was that the bent conformation was the predominant form in solution especially in the presence of 5 mM Ca^{2+} (Xiong et al., 2003a). The later was supported by electron microscopic studies showing that the ecodomain of $\alpha\text{V}\beta 3$ was in bent conformation in 5 mM Ca^{2+} and became extended in the presence of Mn^{2+} or RGD (Arg-Gly-Asp) peptide (Takagi et al., 2002b). The α subunit was acutely bent at the Genu between the Thigh and Calf-1 domains. However, the bend in the β subunit was not visible in the integrin $\alpha\text{V}\beta 3$ crystal structure (Xiong et al., 2001), but the location of the bend could be deduced to be between the PSI and hybrid domains on one side, and the I-EGF3 domain on the other (Shi et al., 2005; Xiong et al., 2001).

Chapter One: Introduction

Figure 1. 2 Integrin architecture

(A) Organization of domains within the primary structure. Some α subunits contain an I domain inserted in the position denoted by the dotted lines. Cysteines and disulfide bonds are shown as lines below the stick figures. Red and blue asterisks denote Ca^{2+} and Mg^{2+} binding sites, respectively. (B) Arrangement of domains within the three-dimensional crystal structure of $\alpha\text{V}\beta 3$ (Xiong et al., 2001). Each domain is color coded as in A. (C) The structure in (B) with the I domain added (Springer et al., 2004).

Figure 1. 3 Crystal structure of the extracellular segment of $\alpha V\beta 3$

(A) Ribbon drawing of crystallized $\alpha V\beta 3$ (αV in blue and $\beta 3$ in red). (B) Model of the straightened extracellular segment of $\alpha V\beta 3$. The two tails would extend into the plasma membrane in the native integrin. Translated and rotated I-EGF3 and I-EGF4 show the approximate location of I-EGF1 and 2 (gray). The PSI tracing (gray) is approximate. Connections of the untraced domains are in dotted lines. αV was straightened by extending the structure by 135° at the Thigh-Calf-1 interface (circled) and then rotating the Calf module $\sim 120^\circ$ around its “long” axis to avoid clashes at the Thigh-Calf-1 interface. The same transformations were then applied to $\beta 3$ (residue 445 onward). Arrows point to the position of the three longest interdomain linkers 1, 2, and 3 in the structure. Amino acid domain boundaries are indicated in parenthesis. “N” and “C” indicate NH_2 - and COOH -termini, respectively (from Xiong et al., 2001). Note that the boundaries of the I-EGF domains differ from those defined by Tan, et al., and Takagi, et al., (2001), which was shown to be correct by later studies (Beglova, et al., 2002) and this thesis.

1.2.1 The α subunit

1.2.1.1 The β -propeller domain

The N-terminus of each integrin α subunit contains seven structurally homologous segments of 60 amino acids, and had been predicated to fold into a seven bladed β -propeller domain (Springer, 1997). The predicted model has been supported by mAb epitope mapping, the residues are discontinuous in the primary structure but are in close proximity in the predicted structure (Oxvig et al., 1998). Mutational analyses indicated that the β -propeller domain contributes partially to the ligand-binding site (Kamata et al., 2001). The predicted structure was confirmed by the α V β 3 crystal structure (Xiong et al., 2001) (Fig.1.4). Each of the seven blades comprises four anti-parallel β -strands, strand A is the inner strand and runs parallel to the central axis, strands B and C of the same repeat radiating outward, and strand D on the outside of the blade. Strands A, B, C, and D are connected by successive hairpin turns, and strand D of one blade is connected to strand A of the next. Three or four Ca^{2+} -binding sites locate in the hairpin turn between strand A and strand B within blades 4-7 at the bottom of the β -propeller domain which is opposite to the $\alpha\beta$ interface. The central axis is lined with amide and carbonyl oxygen groups from strand A of all blades, with a few side chains pointing to the cavity (Xiong et al., 2001).

Figure 1. 4 Ribbon diagrams of the αV integrin β -propeller domain

Views are from the bottom. Each W blade, named as W1 to W7, is shown in a different color. Cysteines in disulfides are in *red sticks*. Four Ca^{2+} ions and three glycans are in *red* and *green* spheres, respectively. (from Xiong et al., 2001).

1.2.1.2 The I domain

Among the eighteen integrin α subunits, half of them contain an ~200-amino-acid domain inserted between the blades 2 and 3 of the β -propeller domain (Arnaout et al., 1988; Michishita et al., 1993; Springer, 1997), this domain is therefore called the I (inserted) domain. It is also referred to as the A domain since it is homologous to the A domain of von Willebrand factor. The crystal and NMR structures of I domains have been reported from the integrin α L (Kallen et al., 1999; Legge et al., 2000; Qu et al., 1995; Qu et al., 1996), α M (Baldwin et al., 1998; Lee et al., 1995a; Lee et al., 1995b), α 2 (Emsley et al., 1997; Emsley et al., 2000), α X (Vorup-Jensen et al., 2003) and α 1 (Nolte et al., 1999; Rich et al., 1999) subunits. These data showed that the I domain adopts the dinucleotide-binding or Rossmann fold, with a central hydrophobic β -sheet, consisting of five parallel and one anti-parallel strands, surrounded by five to seven amphipathic α -helices, and a divalent cation-binding site, named as the MIDAS (Metal Ion-Dependent Adhesion Site) at the top of the I domain. The MIDAS physiologically binds Mg^{2+} (Fig.1.5), but it has been shown to bind other divalent cations Mn^{2+} , Cd^{2+} , Co^{2+} , Ni^{2+} , and Zn^{2+} (Baldwin et al., 1998). Binding affinity of metal ions to the I domain is regulated by integrin activation and ligands. In many integrins, Mg^{2+} or Mn^{2+} supports ligand binding, but, high concentration of Ca^{2+} has inhibitory effects (Ajroud et al., 2004). Mutation of residues involved in divalent cations binding in α L β 2 and α M β 2 integrins showed that the integrins became incapable of binding ligands which indicated the importance of the divalent cations for ligand binding (Kamata et al., 1995; McGuire et al., 1995). In addition, EDTA treated integrins lose the ability of ligand binding (Altieri, 1991).

Chapter One: Introduction

When present, the I domain is the major site of the integrin for ligand binding (Hynes 2002; (Michishita et al., 1993). The deletion of the I domain (I-less) in the α subunits does not effect the expression of $\alpha\beta$ heterodimers, but abolishes most of ligand binding capacity (Leitinger et al., 2000; Yalamanchili et al., 2000). I-less Mac-1 loses the capacities for some ligands (Factor X, and denatured proteins et al.), but still can bind iC3b at a reduced level (Yalamanchili et al., 2000), possibly mainly through the interaction blade 4 in the β -propeller domain (Li et al., 2003), suggesting that the β -propeller domain can also involve in ligand binding (Dickeson et al., 1997; Stanley et al., 1994).

1.2.1.3 The conformational states of I domain

The isolated I domains exist in at least two conformations, a “closed” structure with low affinity for the ligands, and an “open” structure with high affinity for the ligands, as first described for the α M I domain (Lee et al., 1995a; Lee et al., 1995b). The two structures have distinct coordinates of the divalent cations in the MIDAS, the different orientations of the β 6- α 7 loop and the α 7 helix (Arnaout et al., 2005; Luo et al., 2007) (Fig.1.5A). In the close conformation of α M I domain, the divalent cation Mn^{2+} is directly coordinated by the side chains of Ser142 and Ser144 in DXSXS motif (Asp140-Gly-Ser-Gly-Ser in the α M I domain) in the first loop between β 1 strand and α 1 helix (β 1- α 1 loop), Asp242 in the β 4- α 5 loop, and three water molecules completing the octahedral coordinate. One of the metal bound water molecule is fixed by forming two hydrogen bonds with the side chains of Thr209 within the α 3- α 4 loop and Asp140 in DXSXS motif (Fig.1.5B). In the open conformation, the divalent ion Mg^{2+} is directly coordinated by the side chains of

Figure 1. 5 Structural comparison of the closed and open α M I domain

(A) The superimposed structures of close (*green*) and open (*blue*) conformations. Major conformational differences are shown in *red dashed* lines. (B) Structure of the closed α M I domain MIDAS. *Yellow* sphere represents Mn^{2+} . (C) Structure of the open α M I domain MIDAS. Glu314 (*red*) from a neighboring α M I domain coordinates with the MIDAS Mg^{2+} (*yellow* sphere). Metal bound water molecules (*red* spheres) complete the octahedral coordination. (PDB: closed α M I domain, 1JLM; open α M I domain, 1IDO).

Chapter One: Introduction

Ser142 and Ser144 from the DXSXS motif, Thr209 from the $\alpha 3$ - $\alpha 4$ loop, a Glu residue from the ligand, and two water molecules which complete the octahedral coordinate (Fig.1.5C). The different divalent cation coordinates are the results of the MIDAS loops rearrangements driven by a major restructuring of the $\beta 6$ - $\alpha 7$ loop and a two-turn downward displacement of the $\alpha 7$ helix. The $\alpha 7$ helix downward displacement represents the crucial linkage to transmit the conformational changes between I domain and other integrin domains. The importance of the $\alpha 7$ helix displacement is proved by elaborate experiments using an engineered disulfide bond to shift the $\alpha 7$ helix downward by one or two turns in αL I domain. The two variants were shown to have an increase in ICAM-1 affinity, in comparison to the wild type, by ~500 and 10,000 folds, respectively, and may be considered to be in the intermediate and high affinity states (Lu et al., 2001b; Shimaoka et al., 2001; Shimaoka et al., 2003). The crucial role of the $\alpha 7$ helix displacement during I domain activation is further supported by the binding of lovastatin, an allosteric chemical inhibitor, which wedges itself in the large hydrophobic pocket between the central sheet and the $\alpha 7$ helix, and stabilizes the closed conformation by preventing the downward movement $\alpha 7$ helix (Kallen et al., 1999).

1.2.1.4 The leg region of α subunit

The remainder of the α subunit extracellular domain comprises three large β -sandwich domains of about 500 residues, designated the Thigh, Calf-1 and Calf-2 domains (Xiong et al., 2001). The known proteolytic cleavage site (Arg860 in αV subunit), which generates the light and heavy chains in certain integrins, locates in the Calf-2 domain,

Chapter One: Introduction

However, it is not visible in electron density of the $\alpha V\beta 3$ crystal structure and is presumably disordered (Xiong et al., 2001).

The α subunit Genu, located in the loop between the Thigh and Calf-1 domain and capped by a divalent cation, is well defined (Fig.1.6). Studies using activation dependent monoclonal antibodies suggest that the interface between the Genu and the Calf-1 domain is maintained during integrin activation, and the extension occurs with a rearrangement at the interface between the Genu and the Thigh domain (Xie et al., 2004).

Figure 1. 6 The structure of the integrin $\alpha V\beta 3$ Genu

The disulfide bond between Cys596 and Cys602 is *orange*. Residues involved in the hydrogen bonds (*yellow broken lines*) with Ca^{2+} (*yellow sphere*) are labeled (Xiong et al., 2001) .

1.2.2 The β subunit

The domain arrangement of the β subunit is more complex than the α subunit. The β I domain is inserted into the hybrid domain which in turn inserted into the PSI domain, which continuous with the four I-EGF domains and the β -tail domain followed by the TM (transmembrane domain) and cytoplasmic domain.

1.2.2.1 The PSI domain

Based on primary sequence alignment, the N-terminal cysteine-rich region of ~50 residues is present in many proteins, including plexins, semaphorins and integrins that mediate cell growth, migration, and differentiation. It is therefore named PSI domain (Bork et al., 1999). Though PSI domain lies N-terminal in the primary sequence, its three-dimensional location is in the middle of β subunit. The PSI domain was not built in integrin α V β 3 crystal structure (Xiong et al., 2001), because of the weak electron density. However, subsequent solved structures of the PSI domains in semaphorin 4D (SEMA4D) (Love et al., 2003) and MET (Stamos et al., 2004) facilitate the building of the β 3 integrin PSI structure (Xiong et al., 2004). Meanwhile, another structure containing integrin α IIB β 3 headpiece including the β 3 integrin PSI domain is solved (Xiao et al., 2004). The integrin PSI domain contains eight cysteines forming four disulfide bonds connected in a Cys1-Cys4, Cys2-Cys8, Cys3-Cys6, Cys5-Cys7 pattern. Unlike other PSI domains, the eighth cysteine is located after the hybrid domain, is thus separated from other seven cysteines in the primary structure.

1.2.2.2 The hybrid domain

The hybrid domain folds into an I-set Ig domain and is inserted in the PSI domain, so that it lies between the β I domain and the PSI domains. It therefore connects to each of the PSI at one side and the β I domain in the other side (Xiao et al., 2004; Xiong et al., 2004), suggesting that it may be crucial for signal intradomain transduction during integrin activation. A collection of EM studies on certain integrins indicated that hybrid domain exists in two orientations with respect to the β I domain (Iwasaki et al., 2005; Nishida et al., 2006; Takagi et al., 2001b; Takagi et al., 2002b), which was later confirmed by crystal structures of the liganded or pseudoliganded α IIb β 3 headpiece (Xiao et al., 2004) (Fig.1.7). Furthermore, in addition to these structural studies, two very different orientations of the hybrid domain with the β I domain was supported by other functional studies: The mAbs HUTS-4 and MEM148 whose epitopes located to the inner face of the hybrid domain of integrin α 5 β 1 and α L β 2, respectively, only bound the highly activated integrins (Mould et al., 2003b; Tang et al., 2005). The β 2 integrin mAb 7E4, when binding to residue Val385 in the hybrid domain inhibited the activation of the α L β 2 integrin by blocking the signal from the stalk region (Tng et al., 2004). The β 1 integrin mAb SG/19 binding to the hybrid domain was visualized by EM to prevent the hybrid domain swing-out (Luo et al., 2004b). A glycosylation site, introduced into the interface between the β I domain and the hybrid domain in α 5 β 1 or α IIb β 3 integrins, stabilized the open headpiece conformer, resulting in higher ligand binding affinity (Luo et al., 2003; Luo et al., 2004b). The downward movement of the α 7 helix in β I domain is closely related with ligand binding and hybrid domain swing-out. A α 7 helix mutation L358A in integrin α 5 β 1 not only increased the integrin ligand binding, but also induced the hybrid

Figure1. 7 Superposition of the headpieces from the RGD liganded legless α IIB β 3 (green) and bent, unliganded α V β 3 (blue)

(A) The PSI, Hybrid domains and the crucial α 7 helix are colored in *blue* and *green* in α V β 3 and α IIB β 3, respectively. The other motionless parts are in *gray*. (B) Clarification of the α 7 helix downward movement. The important ADMIDAS metal coordinating residue Met335 and the last residue of the α 7 helix in both structure are labeled, shown in sticks and colored accordingly. (PDB: bent, unliganded α V β 3, 1U8C; RGD liganded legless α IIB β 3, 1TY6).

swing-out (Mould et al., 2003b). A more detailed structure was depicted in the crystal of the liganded and pseudoliganded $\alpha\text{IIb}\beta 3$ headpiece, when compared with the $\alpha\text{V}\beta 3$ bent structure (Fig.1.7A). Together with the $\beta 6$ - $\alpha 7$ loop rearrangement, the $\alpha 7$ helix moved about one-turn downward and stopped at residue Lys350, which is found to be in the same position as residue Ser353 that represents the end of the $\alpha 7$ helix in the bent, unliganded $\alpha\text{V}\beta 3$ structure in a superposition analyses (Fig.1.7B).

1.2.2.3 The β I domain

The β I domain spans approximately residues 101-340 and inserts in the B-C loop of the hybrid domain. The β I domain was predicted to fold into a variant von Willebrand factor A-domain (Tuckwell et al., 1997). Though it has weak sequence homology to the I domain, they share a similar secondary structure, and all domains contain a putative metal-binding DXSXS motif and adopt the nucleotide-binding or Rossmann fold (Xiong et al., 2001). The β I domain involves in ligand binding in integrins without the I domain, and indirectly regulates the ligand binding in integrins containing the I domain. Compared to the I domain, the β I domain contains two extra loops, one is named SDL (specificity determining loop) (Takagi et al., 2002a; Takagi et al., 1997), as it is crucial for integrin dimer formation and ligand binding. The other forms the main interface with the β -propeller domain of the α subunit (Arnaout et al., 2005). Unlike the I domain, the β I domain has two additional metal ion binding sites on each side of MIDAS, namely, the ADMIDAS (adjacent to MIDAS) and LIMBS (ligand-induced metal binding site). The MIDAS and LIMBS are not present in the bent, unliganded integrin $\alpha\text{V}\beta 3$ crystal structure (Xiong et al., 2001), but can be clearly detected in the RGD liganded bent $\alpha\text{V}\beta 3$

Chapter One: Introduction

(Xiong et al., 2002), and the liganded or pseudoliganded $\alpha\text{IIb}\beta 3$ headpiece crystal structures (Xiao et al., 2004) (Fig.1.8A). The $\beta 3$ MIDAS is physiologically occupied by Mg^{2+} which is coordinated by the side chains of Ser121 and Ser123 from the DXSXS motif, Glu220 (equivalent to Thr209 in MIDAS of αM I domain), Asp from the ligand, and two water molecules (Fig.1.8B). Therefore, the overall geometry of the divalent cation coordination is the same as that of the I domain. The divalent cation Ca^{2+} in the LIMBS is coordinated by the side chain carboxyl oxygen of Asp158, Asn215 and Asp217 and Glu220, and the main chain carbonyl oxygen of Asp217 and Pro219 (Xiao et al., 2004; Xiong et al., 2002) (Fig.1.8C).

A divalent cation is observed in the ADMIDAS in the bent, unliganded $\alpha\text{V}\beta 3$ structure (Xiong et al., 2001), the divalent cation Ca^{2+} is coordinated by the side chain carboxyl oxygen of Asp126 and Asp127 and the main chain carbonyl oxygen of Ser123 and Met335. The hydrogen bond between the side chains of Ser337 in the $\beta 6$ - $\alpha 7$ loop and Asp127 at the starting point of the $\alpha 1$ helix may further stabilize this conformation. In addition, the side chain of Asn339 at the start of the $\alpha 7$ helix makes two hydrogen bonds with the carbonyl oxygen from the Val332 main chain and with the hydroxyl oxygen of Ser337, both from the $\beta 6$ strand of the central sheet (Fig.1.9A). Thus the position of the $\alpha 7$ helix is stabilized in the bent, unliganded $\alpha\text{V}\beta 3$ structure (Xiong et al., 2001). In the RGD liganded $\alpha\text{IIb}\beta 3$ headpiece crystal structure (Fig.1.9B), the hydrogen bond between the ADMIDAS Ca^{2+} cation and the carbonyl oxygen of Met335 is broken, and the cation moves inward $\sim 2.8\text{\AA}$ and is stabilized by forming the hydrogen bond with carboxyl

Chapter One: Introduction

Figure 1. 8 Structure rearrangement of the β I domain

(A) The superimposed structures of metals (*spheres*) and their coordinations in unliganded (*blue*), liganded bent (*cyan*) α V β 3 integrin and liganded (*green*) α IIb β 3 legless integrin are shown in sticks. All three metal ions (from left to right: LIMBS, MIDAS, and ADMIDAS and colored accordingly) are found only in liganded or pseudoliganded structures and are superimposable. The bent unliganded structure only has the ADMIDAS metal (*blue*). Water molecules are shown in crosses (*magenta*). Major conformational differences are dashed lines (*red*) with arrows. (B) Structure of MIDAS in liganded (*green*) α IIb β 3 legless integrin. *Green* sphere represents Mg^{2+} . Ligand Asp (*red*) is from the RGD peptide. Metal bound water molecules (*red* spheres) complete the octahedral coordination. (C) Structure of the LIMBS in liganded (*green*) α IIb β 3 legless integrin. *Green* sphere represents Ca^{2+} . (PDB: bent, unliganded α V β 3, 1U8C; bent, RGD liganded α V β 3, 1L5G; RGD liganded legless α IIb β 3, 1TY6).

Chapter One: Introduction

Figure1. 9 The structure ADMIDAS metal Ca^{2+} (sphere) and its coordinations

(A), unliganded (blue); (B), liganded (green) $\alpha\text{IIb}\beta 3$ legless integrin and (C), liganded bent (cyan) $\alpha\text{V}\beta 3$ integrin. The hydrogen bonds between the residues (labeled) in the $\alpha 1$ helix, $\beta 6$ - $\alpha 7$ loop and $\alpha 7$ helix are also shown by dashed lines (yellow). The water molecules are shown in small sphere (magenta). (PDB: bent, unliganded $\alpha\text{V}\beta 3$, 1U8C; bent, RGD liganded $\alpha\text{V}\beta 3$, 1L5G; RGD liganded legless $\alpha\text{IIb}\beta 3$, 1TY6).

Chapter One: Introduction

oxygen of Asp251. It pulls residues Asp126 and Asp127 at the start of the $\alpha 1$ helix inward through the Asp251-metal-Asp126/Asp127 interactions. Furthermore, the carbonyl oxygen of Met335 reforms a hydrogen bond with the amide of the side chain of Asn133 in the middle of $\alpha 1$ helix. These movements pull the $\alpha 1$ helix towards the MIDAS and push the $\alpha 7$ helix one-turn downward. Asn339 moves one-turn downward, its side chain forms a hydrogen bond with the carbonyl oxygen of Val332 from the $\beta 6$ strand. In integrin $\alpha 4\beta 7$, mutations in ADMIDAS required for rolling result in firm adhesion, but mutations in LIMBS needed for firm adhesion result in rolling, which indicate that the ADMIDAS has a negative effect, whereas the LIMBS has a positive effect in the regulation of ligand binding (Chen et al., 2003). In integrin $\alpha 5\beta 1$, mutations in ADMIDAS markedly decreased the association rate to fibronectin (Mould et al., 2003a). For most integrins, low concentration of Ca^{2+} contributes to, whereas high concentration inhibits, adhesion (Shimaoka et al., 2002).

In the RGD liganded $\alpha V\beta 3$ crystal structure, the movement of its $\beta 1$ - $\alpha 1$ loop is almost the same as that of the RGD liganded $\alpha \text{IIb}\beta 3$ legless crystal structure (Fig.1.9C). The RGD ligand was soaked into the crystal rather than co-crystallization with the integrin, it is therefore presumed that the crystal packing force prevented downward movement of the $\alpha 7$ helix and other intradomain movements. The hydrogen bond between the side chains of Asp127 and Ser337 is intact, thus driving the $\beta 6$ - $\alpha 7$ loop inward together with the motion of the ADMIDAS cation. Furthermore, the new hydrogen bond formed between the side chain hydroxyl hydrogen of Ser338 and the main chain carbonyl oxygen of the Leu128 in $\alpha 1$ helix not only stabilizes $\beta 6$ - $\alpha 7$ loop conformation, but also freezes

Chapter One: Introduction

the $\alpha 1$ helix movement. Therefore the localization of the $\beta 6$ - $\alpha 7$ loop and the motionless of the $\alpha 1$ helix prevent the hydrogen bond formation between the carbonyl oxygen of Met335 and the residue in $\alpha 1$ helix (Fig.1.9C). We should be caution in the interpretation of this structure since it may not be physiological relevant. Indeed, the EM of the RGD liganded integrin $\alpha V\beta 3$ indicated the fully extended conformation (Takagi et al., 2002) which is distinct from the RGD soaked $\alpha V\beta 3$ crystal structure.

LAD-1 (Leukocyte Adhesion Deficiency-1) is caused by defects in integrin $\beta 2$ subunit resulted in impaired integrin heterdimer formation and diminished expression or null expression on the cell surface (Shaw et al., 2001). One LAD-1 mutation, N329S, located in the β I domain in integrin $\beta 2$ subunit, converted integrin $\alpha L\beta 2$ from a resting into a high affinity conformer, and a corresponding mutation N339S in $\beta 3$ I domain also resulted in constitutive activation, which indicates that this conserved Asn may have a primary role in shaping the β I domain by stabilizing the conformation of the $\alpha 7$ helix and the $\beta 6$ - $\alpha 7$ loop in the I-like domain (Cheng et al., 2007).

1.2.2.4 The I-EGF domains

Four I-EGF domains located after the hybrid domain contain high content of cysteines, thus they have been referred to cysteine-rich repeats. The cysteine-rich repeats were initially postulated to contain three or four repeating elements with eight cysteines residues (Kishimoto et al., 1987; Law et al., 1987; Tamkun et al., 1986). Its overall structure is presumed to be similar to two tandem calcium-binding EGF domains in fibrillin (Downing et al., 1996), except that the calcium cation is replaced by a disulfide

bond in the I-EGF domains. The I-EGF domains will be discussed in full details throughout this thesis.

1.2.2.5 The β -TD (β -tail domain) domain

As the last module of the extracellular domain of β subunit, the β -TD domain locates at C-terminal of the I-EGF domains and is also cysteine-rich. It contains a four-stranded β sheet with both anti-parallel and parallel strands, an N-terminal α helix, and five disulfide bonds (Xiong et al., 2001). This domain does not affect the integrin dimer formation, however, it may restrain integrin in the bent, low affinity state (Butta et al., 2003). The β -TD domain makes only two hydrophobic contacts with the I-EGF4 domain suggesting the flexibility between them (Arnaout et al., 2002). In addition, a deadbolt model based on the small $60 \text{ \AA}^2$ contact area between the β -TD domain and the β I domain has been proposed to play a pivotal role in regulating the integrin activation (Gupta et al., 2007; Xiong et al., 2003b). This is proposed as an alternative to the switchblade model in which the extension is the key for the integrin activation (Beglova et al., 2002; Takagi et al., 2002b; Zhu et al., 2007).

1.2.3 The transmembrane (TM) domains

The TM (transmembrane domains) are not only the linkers between the ectodomain and cytoplasmic domains involved in the bidirectional signal, but also play a role in integrin activation. The disulfide bonds formation between the α IIb and β 3 subunits only occur in specific residues, which are substituted by the cysteine residues, located in the TM fragment of the resting integrin. This data suggest that TM helices interact in the low

Chapter One: Introduction

affinity state which is lost after integrin activation resulted in transmembrane domains separation (Luo et al., 2004a). Furthermore, mutagenesis studies were identified in the TM region (e.g. G708S, I714T in integrin $\beta 3$ subunit) (Partridge et al., 2005) leading to constitutively active $\alpha \text{IIb} \beta 3$ integrins, suggesting that the interactions between the integrin TM helices are involved in the regulation of activation. However, the molecular details of these interactions are still unclear (Ginsberg et al., 2005).

1.2.4 The cytoplasmic domains

Compared to the integrin ectodomain, the cytoplasmic domains of integrin α and β subunits are very short with only 20-40 residues and 45-60 residues, respectively, except for the integrin $\beta 4$ subunit (Hogervorst et al., 1990). The membrane-proximal regions of α and β cytoplasmic domains are well conserved, composed of a hydrophobic span of 4-5 residues followed by a strongly polar sequence (e.g. HD in the β subunit, KR in the α subunit). Deletion of the cytoplasmic domain of one subunit or only its membrane-proximal region results in constitutively active integrin, whereas deletions of the membrane-distal region have no effect, and may even be inhibitory to activation (Lu et al., 2001c; Lu et al., 1997; O'Toole et al., 1994; O'Toole et al., 1991). The boundaries between the membrane-proximal regions and TM domains are not clearly defined as certain residues can move into and out of the membrane bilayer depending on the affinity states (Partridge et al., 2005). Many integrin activating mutations located in these regions are thought to displace certain membrane-embedded regions out of the membrane, thereby shortening the TM domains (Armulik et al., 1999; Stefansson et al., 2004), suggesting

Chapter One: Introduction

that movement of boundary residues are involved in integrin activation (Ginsberg et al., 2005) (Fig.1.10).

Figure 1. 10 Model for TM and membrane-proximal regions displacements during integrin activation

Left panel: the low-affinity integrin is maintained through inter-subunit interactions at the TM and membrane-proximal domains. The TM helices specifically pack together and are in a long and tilted geometry. The membrane-proximal regions interact with each other, with an α subunit arginine (Arg995 in α IIb, denoted by R in the figure) forming a salt bridge (dotted line) with an aspartic acid from the β subunit (Asp723 in β 3, denoted by D in the figure). Right panel: Binding of talin to the cytoplasmic tail of the β subunit induces a series of conformational changes that give rise to a high affinity for the ligand in the extracellular domain. The talin- β tail interaction breaks the TM helical packing by shortening the TM domain and therefore creating an unfavorable crossing angle. The interaction may also directly disrupt the membrane-proximal salt bridge, resulting in a shortened TM helix and hence leading to either separation or alternative packing of the TM domains. (from Ginsberg 2005).

Chapter One: Introduction

Mutagenesis studies suggest that the salt bridge between R995 of α Ib and D723 of β 3 contributes the association between the integrin α and β membrane-proximal regions, thus stabilizing the integrin α Ib β 3 in a low affinity state (Hughes et al., 1996). The NMR structure of the integrin α Ib β 3 cytoplasmic domains confirmed the existence of this salt bridge (Vinogradova et al., 2002) (Fig.1.11A, B). The α Ib subunit cytoplasmic domain forms a short α helix followed by a turn, allowing the acidic C-terminal loop to fold back and to interact with the positively charged N-terminal region. In contrast, a long α helix is formed in β 3 cytoplasmic domain and the end C-terminal 25 residues are disordered in solution, including two NPxY motifs. The segment D740TANNPLY has the propensity to form a β strand/turn (Ulmer et al., 2001) on interaction with the PTB (phosphoryrosine-binding) domain (van der Geer et al., 1995) of the F3 domain in the talin head (Garcia-Alvarez et al., 2003; Wegener et al., 2007). The β strand consisted of β 3 residue D740-TA forms a β strand thus an anti-parallel sheet can be observed with β 3 D740-TA and the S5 from the talin F3 domain (Garcia-Alvarez et al., 2003) (Fig.1.11C).

Disruption of the contacts between the membrane-proximal region and the talin head by mutagenesis (F727A, F730A in integrin β 3 cytoplasmic domain) greatly reduced the effect of the activating mutation R995A in the integrin α Ib (Wegener et al., 2007). Thus, the membrane-proximal regions play important roles during integrin activation, and may interact with each other to restrain the integrin in a low-affinity state. The binding of talin to the integrin β cytoplasmic domain may disrupt the association between the membrane-proximal regions of the two subunits and activate integrin (Calderwood, 2004). Similarly, Talin has been shown to induce rolling α L β 2 integrin to an intermediate state capable of

binding to the ligand of intercellular adhesion molecule (ICAM)-1 but not ICAM-3 (Li et al., 2007).

Figure 1. 11 The NMR structure of the integrin tails

(A) The NMR structure of integrin $\alpha\text{IIb}\beta 3$ cytoplasmic domain, the αIIb subunit is in *cyan* and the $\beta 3$ is in *green*. The residues R995 and D723 are shown in *sticks* and their interaction is shown in dashed *yellow* line. (B) Clarification of the interaction (salt bridge) in the cytoplasmic domains. (C) The structure of talin F3 domain with the integrin $\beta 3$ chimeric tail.

FRET (fluorescence resonance energy transfer) has been used to study the interaction of the cytoplasmic tails of the $\alpha\text{L}\beta 2$ integrin. The αL subunit was fused with cyan fluorescent protein and the integrin $\beta 2$ subunit was fused with yellow fluorescent protein. FRET was shown when $\alpha\text{L}\beta 2$ integrin was in a low-affinity state suggesting the close

proximity of the α L and β 2 cytoplasmic domains. The FRET signal diminished when α L β 2 integrin is in a high-affinity state suggesting the tail separation (Kim et al., 2003).

1.3 The global conformational changes in the integrin ectodomain

The first integrin ectodomain crystal structure of α V β 3 revealed an unexpected bent conformation (Xiong et al., 2001). Although it was suggested that this bent structure was in its active state because of the truncation in the TM and cytoplasmic domains that normally restrain integrins in a low-affinity state (Dana et al., 1991; Hughes et al., 1996), it is unlikely as the crystals were obtained in the presence of 5 mM Ca^{2+} , which can stabilize integrins in their low-affinity state (Takagi et al., 2002b). Furthermore, with a bent integrin on the cell surface, the ligand binding site would face the membrane, making it topological unfavorable to interact with the extracellular ligands (Zhu et al., 2007). Subsequent experimental data have confirmed that the bent conformation exists physiologically not only in integrin α V β 3, but also in other integrins at the low-affinity state. First, EM images of recombinant integrins α V β 3 (Adair et al., 2005; Takagi et al., 2002b), α L β 2 and α X β 2 (Nishida et al., 2006) was in the bent conformation in the presence of Ca^{2+} , consistent with the crystal structure of the integrin α V β 3. Second, an engineered disulfide bond between Gly307 in the α V β -propeller domain and Arg563 in the β 2 I-EGF4 domain locked the cell surface integrin α V β 3 in a bent conformation which failed to bind ligands (Takagi et al., 2002b).

The recent and early EM studies on integrins also revealed the extended conformations (Iwasaki et al., 2005; Nermut et al., 1988; Nishida et al., 2006; Takagi et al., 2002b).

Chapter One: Introduction

Based on the gel filtration elution volume, the calculated Stoke radius of recombinant $\alpha V\beta 3$ ectodomain increased from 56Å in 1 mM Ca^{2+} to 60Å in 1 mM Mn^{2+} and expanded further to 63Å in the presence of cyclic RGD ligand (Adair et al., 2005). This is consistent with EM results that most of the integrin $\alpha V\beta 3$ molecules exist in only the bent conformation in the presence of Ca^{2+} , in both bent and extended conformations with open or closed headpiece in the presence of Mn^{2+} and in only the extended conformation with open headpiece in the presence of cyclic RGD ligand (Takagi et al., 2002b). Although the $\alpha V\beta 3$ crystal structure exists in the severe bent conformation, when the structure is computationally extended at the Genu, it resembles the shapes of integrins derived from the EM images which previously showed an elongated molecule (Xiong et al., 2003a) (Fig.1.12). All these results suggest that extension is sufficient to activate integrin (Luo et al., 2007).

However, a number of other observations is not in agreement with Genu-extension model. First, the $\alpha V\beta 3$ ectodomain crystal in the bent conformation still has the capacity to bind the cyclic RGD ligand (Xiong et al., 2002). Second, the EM image of the integrin $\alpha V\beta 3$ in 0.2 mM Mn^{2+} with a fibronectin fragment revealed the bent conformation (Adair et al., 2005). Third, the GPI-linked uPAR can stabilize the fibronectin binding to the integrin $\alpha 5\beta 1$ in its bent conformation by the interaction between the N-terminal domain of uPAR and the blade 4 of the β -propeller domain of the integrin $\alpha 5\beta 1$ which is incompatible with the extended conformation (Wei et al., 2005) (Simon et al., 2000) (Arnaout et al., 2005), as the distance between the β -propeller domain and the plasma

Figure 1.12 EM images and its presumed models of integrin structures

(A) Projection average of clasped α V β 3 particles in the presence of Ca^{2+} cation obtained by negative staining EM imaging (Takagi et al., 2002b). (B) Ribbon diagram of the bent conformation observed in the α V β 3 crystal structure (1U8C) with the I-EGF1 and I-EGF2 domains built using domain overlap superimposition with β 2 fragments, PHE1 (1YUK) and PHE3 (2P28), respectively. The α V subunit is shown in *magenta*, and the β 3 subunit in *cyan*. (C) Representative projection average of an extended integrin with closed headpiece. (D) The model of the extended integrin with closed headpiece by computationally extending from the bent crystal structure (1U8C) at the α subunit Genu between the thigh and the Calf-1 domains, superimposing with the PHE3 structure (2P28) on hybrid domain to fill in PSI and I-EGF1-3 domains, using the hybrid domain from PHE3 structure to replace the original, and finally, superimposing the fragment containing I-EGF3-4 and β -TD from the bent structure (1U8C) on the I-EGF3 domain. (E) Representative projection average of an extended integrin with an open headpiece. (F) The model of the extended integrin with open headpiece. The extended α subunit was generated as in (D), and superimposed with the open headpiece (1TY6) on the β -propeller domain, replaced with the β I and hybrid domains from 1TY6 structure and superimposed as in (D).

Chapter One: Introduction

membrane in the bent and extended integrin is around 40 Å and 200 Å, respectively. The longest dimension of the uPAR ectodomain containing D1, D2 and D3 domains is 52 Å when extended (Huai et al., 2006). A reconciliation of the disputants is that the integrins in the bent conformation somehow restrain the capacity for the small ligands, but not the large natural ligands. In contrast, the topology of the extended integrins is optimal for the easy access of their ligands at the cell surface in physiological environment (Zhu et al., 2007).

The transition of the integrins from the bent, low affinity state to the extended intermediate state with closed headpiece, and then to the high affinity state in the extended with open headpiece (hybrid domain swing-out), during the activation process, had been revealed from EM studies and two crystal structures, there may be other conformations that had not been observed yet (e.g. bent with open headpiece). We still cannot distinguish pathways involved in the inside-out and outside-in signaling. The possible quaternary structural rearrangements in the integrin extracellular domains are proposed and illustrated (Fig.1.13). Details about the intra-domain rearrangements that reflect the signals transmitted between integrin domains are still lacking.

Figure 1. 13 Proposed intermediates in equilibration between known conformational states

The upper pathways may be stimulated by ligand binding outside the cell, and the lower pathways by signals within the cell that separate the integrin alpha and beta subunit transmembrane domains. Domains in A–G are shown in solid color if known directly from crystal structures, dashed with grey if placed from crystal structures into electron microscopy image averages, and in solid grey for EGF-1 and EGF-2, which are modeled on EGF-3 and EGF-4. (from Xiao et al., 2004).

1.4 Objective of my thesis

The I-EGF domains of the β subunit provide the crucial link in signal transmission and undergo a dramatic straightening when integrins transfer from a low-affinity to an intermediate or high-affinity state, as indicated by EM studies (Nishida et al., 2006; Takagi et al., 2002b). Some mAbs (MEM48, KIM127, KIM185 etc.), whose epitopes locate in this region, either directly activate ligand binding or act as probes that bind only to the extended integrins (Humphries, 2000; Lu et al., 2001a; Takagi et al., 1997). Though structures of some parts (I-EGF3-4 in integrin $\beta 3$ subunit) of the I-EGF domains have been determined, the β subunit Genu, a flexible region that allows the conformational changes to occur, has not been defined. In order to address the molecular and structural basis of this Genu and its conformational changes during integrin activation, integrin structures containing I-EGF1 or/and I-EGF2 domains need to be obtained.

Chapter Two: Materials and Methods

2.1 General reagents

General reagents and solvents were of analytical grade, and were obtained from Sigma Aldrich Chemical Company Ltd., BDH Chemicals Ltd., Becton-Dickenson Ltd., GIBCO-BRL Ltd., Pierce Ltd. Hampton Research Ltd., unless otherwise stated. Solutions were sterilized where required, by autoclaving or by passing through a 0.22 μ m filter (Millipore).

2.1.1 Enzymes

All restriction endonucleases and other enzymes were obtained from New England Biolabs, Promega Ltd., Fermentus Ltd, Roche diagnostics unless otherwise stated and were used according to the manufacturers' instructions.

2.1.2 Commercially available kit

Plamid Maxi Kit	QIAGEN Ltd.
QIAprep Spin Miniprep Kit	QIAGEN Ltd.
QIAquick PCR purification Kit	QIAGEN Ltd.
Crystallization Basic kit for proteins	Sigma Aldrich Chemical Company
Crystallization Extension kit	Sigma Aldrich Chemical Company
Heavy Atom Kit	Sigma Aldrich Chemical Company
Nextal PEGs Suite	QIAGEN Ltd.
Nextal PEGs II Suite	QIAGEN Ltd

Chapter Two: Materials and Methods

2.1.3 cDNA clones

CD11a (integrin α L)	From Prof. Law SK (School of Biological Sciences, NTU)
CD18 (integrin β 2)	From Prof. Law SK (School of Biological Sciences, NTU)
pMJ1 (integrin β 2 without I-like domain)	From Dr. Tan SM (School of Biological Science, NTU)

2.1.4 Cells

<i>E.coli</i> BL21 (DE3)	Novagen
Origami (DE3)	Novagen
HEK293T cells (human embryonic kidney cell with SV40 large T antigen)	Purchased from ATCC
HEK293 cells (human embryonic kidney cell)	Purchased from ATCC
HEK293 GnT ^I (deficient in N-acetylglucosaminyltransferase-I)	From Prof. Gobind Khorana (Departments of Biology and Chemistry, MIT)

2.1.5 Expression Vectors

pcDNA3.0	Invitrogen
pIRES2-EGFP	BD Biosciences
pET24a	Novagen

Chapter Two: Materials and Methods

pET32a	Novagen
pMAL-c2X	New England Biolabs
pGEX-5X1	Amersham

2.1.6 Monoclonal antibodies

MHM24	anti- α L	From Prof. McMichael (John Radcliffe Hospital, Oxford, UK)
KB43	anti- α X	From K. Pulford (LRF Diagnostic Unit, Oxford, UK)
MHM23	anti- β 2 heterodimer dependent	From Prof. McMichael (John Radcliffe Hospital, Oxford, UK)
MEM48	anti- β 2 activating mAb	From Prof. Horejsi (Prague, Czech Republic)
H52	anti- β 2	From Prof. LAW SK (School of Biological Sciences, NTU)
7E4	anti- β 2	From Beckman Instruments
MEM148	anti- β 2	From Prof. Horejsi (Prague, Czech Republic)
KIM127	anti- β 2	From Dr. Robinson (UCB, CellTech, UK)
KIM185	anti- β 2 activating mAb	From Dr. Robinson (UCB, CellTech, UK)
KIM202	anti- β 2	From Dr. Robinson (UCB, CellTech, UK)
KIM89	anti- β 2	From Dr. Robinson (UCB, CellTech, UK)

Chapter Two: Materials and Methods

2.2 Solutions, buffers, and media

2.2.1 Laboratory stocks

10 mg/ml BSA	Stored at -20°C
0.1 M DTT	Stored at -20°C
0.5 M EDTA, pH 8.0	Stored at RT
100 mM EGTA, pH 7.4	Stored at RT
0.1% ethidium bromide	Stored at 4°C
1 M HEPES, pH 7.4	Stored at 4°C
1 M MgCl ₂	Stored at RT
3 M NaAc, pH 4.8	Stored at RT
5 M NaCl	Stored at RT
0.1 M Sodium Bicarbonate Buffer	Stored at RT
2x protein solubilisation buffer	0.16 M Tris pH 8.0, 8 M Urea, 1.6% (w/v) SDS, bromophenol blue 0.08% (w/v)
10% (w/v) SDS	Stored at RT
10x TBE	0.9 M Tris pH 8.0, 0.9 M Boric acid 25 mM EDTA
1M Tris-HCl, pH 8.0	Stored at RT
1M Tris-HCl, pH 7.4	Stored at RT

2.2.2 Media

All media were sterilized by autoclaving unless otherwise stated.

Chapter Two: Materials and Methods

40mg/ml ampicillin	40 mg/ml in ddH ₂ O, filtered (0.22µm) and stored at -20°C
25mg/ml Kanamycin	25 mg/ml in ddH ₂ O, filtered (0.22µm) and stored at -20°C
LA broth	LB medium containing 60 µg/ml ampicillin
LA plate	LB agar plate containing 60 µg/ml ampicillin
293 and 293T cell culture media	DMEM (JRH) containing 10% (v/v) heat-inactivated FBS, 100 IU/ml penicillin, 100 µg/ml streptomycin
LB medium	1% (w/v) Bacto-tryptone (BD), 0.5% (w/v) yeast extract (BD), 1% (w/v) NaCl
LB agar	LB medium plus 1.5% (w/v) Bacto-agar (BD)

2.2.3 Solutions

blocking buffer (western)	PBS containing 1% (w/v) non-fat milk, 0.1% (v/v) Tween20
blotting buffer (western)	12.5 mM Tris pH 8.0, 96 mM glycine, 10% (v/v) ethanol
cell freezing media	10% DMSO in heat-inactivated FBS
immunoprecipitation buffer	50 mM Tris pH 7.5, 150 mM NaCl, 1% (v/v) NP-40, 0.5 mM PMSF, 2.5 mM iodoacetamide

Chapter Two: Materials and Methods

100 mM PMSF	prepared in ethanol, stored at -20 °C
10x SDS-PAGE running buffer	0.25 M Tris, 1.9 M glycine, 1% (w/v) SDS
4x resolution gel buffer	0.25 Tris, 0.4% (w/v) SDS, pH 8.8
4x stacking gel buffer	0.5M Tris, 0.4% (w/v) SDS, pH 6.8
sodium bicarbonate buffer	1.36 g sodium carbonate, 7.35 g sodium bicarbonate, pH 9.2
SOC medium	2% Bactotryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 10 mM MgSO ₄ , 20 mM Glucose, add H ₂ O to 1000 ml and autoclave.

Chapter Two: Materials and Methods

2.3 Methods

2.3.1 General Methods for Biochemistry

2.3.1.1 Quantitation of DNA

The concentration of DNA was determined by its absorbance at wavelength 260nm based on the calculation; 50 mg/ml double-stranded DNA gives an OD₂₆₀ of 1. The OD₂₈₀ was also read and the ratio of OD₂₆₀/OD₂₈₀ was calculated to estimate the purity of the DNA solutions. An OD₂₆₀/OD₂₈₀ ratio between 1.6-2.0 was considered satisfactory.

2.3.1.2 Polymerase chain reaction (PCR)

DNA fragment between two oligonucleotide primers were amplified by PCR through repeated cycles of three incubation steps where the DNA template is denatured, allowed to anneal with the primers and a new strand synthesized using a thermo-stable DNA polymerase. A typical cycle included denaturation at 94°C for 1 min, annealing for 1 min at a temperature ~5°C below the lower T_m of the two primers, and finally an extension at 72°C for an appropriate time (normally 1 min per kb). T_m refers to the melting temperature for each oligonucleotide calculated according to its composition allowing 2°C for each A or T and 4°C for each C or G.

2.3.1.3 Standard PCR protocol

PCR was routinely performed in a 50 µl reaction volume containing less than 1 µg template DNA, 1 µM of each oligonucleotide primer, 200 µM of each dNTP, and 1 U DNA polymerase. *Taq* polymerase (Fermentas) was used for PCR colony screening. Pfu DNA polymerase (Fermentas), which possesses a 3'-5' proofreading activity resulting in

Chapter Two: Materials and Methods

a twelve fold increase in fidelity of DNA synthesis over *Taq* DNA polymerase, was used for high fidelity DNA synthesis. Each polymerase has its own reaction buffer, normally supplied by the manufacturer. The reaction mixtures were subjected to a varying number of cycles of amplification using the DNA Thermal Cycler (MJ research).

2.3.1.4 Restriction endonuclease digestion

Restriction endonuclease digestions were usually carried out in a 40-100 μ l reaction volume, with 2-4 U of enzyme used for up to 500 ng DNA, for 1-18 h at an appropriate temperature. Commercially available 10x buffers were used according to the manufacturers' instructions. Digests with more than one enzyme were carried out simultaneously in a suitable buffer. When the buffer requirements were incompatible, restriction digestions were performed sequentially with individual buffer for each enzyme.

2.3.1.5 DNA separation by agarose gel electrophoresis

A 50 ml 0.8% (w/v) agarose gel was routinely used for analysis of 0.1-8 kb DNA fragments. Agarose was melted in 1x TBE, cooled to 55°C, before adding ethidium bromide to a final concentration of 1 mg/ml followed by casting in a mini-gel apparatus. The gel was left to set at RT for at least 30 min. Electrophoresis was carried out in horizontal apparatus with the gel submerged in 1x TBE. DNA samples were loaded with one-fifth volume glycerol loading dye. A standard DNA ladder was also run to allow estimation of the sizes of the sample DNA fragments. DNA fragments were visualized by fluorescence over a UV light (302 nm, UV transilluminator TM-20, UVP), under which

Chapter Two: Materials and Methods

DNA/EB complexes fluoresce and the image was recorded with a Mitsubishi video copy processor.

2.3.1.6 Purification of DNA fragments by agarose gel electrophoresis

When a particular fragment of DNA in a mixture of fragments was required e.g. in ligations or as a PCR template, it was routinely separated from other fragments by agarose gel electrophoresis. Gel slice containing the DNA fragments to be purified was cut from the gel using a razor blade, carefully avoiding their exposure to the UV light. DNA was extracted using a QIAquick Gel Extraction Kit.

2.3.1.7 DNA ligation

DNA vectors with complementary ends to be used for ligation were prepared by restriction enzyme digestion and where necessary treated with alkaline phosphatase and purified using agarose gel electrophoresis. Vector DNA (~ 10 ng) and insert DNA (20-40 ng) were ligated by incubation at RT for 2-3 h, using 1U of T4 DNA ligase with ligase buffer provided together with the enzyme in a 20 µl reaction volume. A reaction with uncut vector, a reaction without insert DNA fragment and a reaction without both insert DNA fragment and T4 DNA ligase were included in the experiments as controls.

2.3.1.8 Preparation of *E.coli* competent cells

The *E.coli* strain DH5α was used to prepare competent cells in advance which were stored at -80°C. A fresh plate of cells was prepared by streaking out cells from frozen stocks and growing overnight at 37°C. On the second day, an individual colony was

Chapter Two: Materials and Methods

picked and grown in 10 ml LB broth culture overnight. On the third day, 5 ml of overnight culture was transferred into each of two flasks containing 500 ml LB broth.

Incubate at 37°C with aeration until the culture reaches OD₅₅₀ of 0.5. This should take approximately 2 h. The cells were transferred to centrifuge bottles and spin at 4°C for 8 min at 8000 rpm. Pellets were gently resuspended in 250 ml ice cold 0.1 M CaCl₂ and combined into a single bottle. Cells were resuspended again in 250 ml ice cold 0.1 M CaCl₂ and centrifuged at 5000 rpm for 10 min at 4°C. Finally the pellet was resuspended in 43 ml ice cold 0.1 M CaCl₂ in ddH₂O with 7 ml sterile glycerol. Competent cells were distributed into convenient aliquots (0.2 ml) in cold microfuge tubes. Cells were store at -80°C. A portion of the cells was saved to assay for viability, purity and competence.

2.3.1.9 Transformation of plasmid DNA

For each transformation, only 50-100 µl of competent cells are necessary. Competent cells were defrozed on ice. DNA was added to cells (the volume of DNA should not exceed 30% of the cell volume). The mixture was incubated on ice for 30 min followed by a heat shock in a 42°C water bath for 1 min. 0.5 ml LB broth was added to the tube. The tube was incubated at 37°C with constant shaking for 30 min to 1h. 50-300 µl of the mixture was streaked out onto plates containing the appropriate antibiotics.

2.3.1.10 Identification of colonies that contain recombinant plasmids of interest

After transformation of competent cells, colonies of interest were identified using both PCR screening and restriction digestion. Each colony to be tested was used to inoculate 15 µl LB. 2 µl of this inoculated LB broth was added to a 20 µl PCR reaction containing

Chapter Two: Materials and Methods

0.25 U *Taq* polymerase. A negative control and, where possible, a positive control were included in the experiment. Colonies containing the recombinant plasmid of interest were identified by the size of their PCR products using agarose gel electrophoresis. Positives from PCR screening were confirmed using restriction digestion.

2.3.1.11 Purification of plasmid DNA

For small scale purification of plasmid DNA, 5ml LB with appropriate antibiotics was inoculated with a single colony that contains recombinant plasmid from an agar plate and incubated at 37°C with constant shaking overnight. A QIAprep Spin Miniprep Kit (QIAGEN) was used to extract DNA according to the manufacturers' instructions.

For large scale purification of plasmid DNA, a Plasmid Maxi Kit (QIAGEN) was used according to the manufacturers' instructions. This kit routinely yielded between 200-750 µg DNA from 100 ml of overnight bacterial cultures.

2.3.1.12 Site-directed mutagenesis

Point mutations were made using QuikChange Site-Directed Mutagenesis Kit (Stratagene), and following the manufacturer's protocols. Overlay primers with the desired mutations were used in the long PCR cycle. The restriction enzyme Dpn I was used to digest the original template. The Dpn I treated PCR product was transformed in competent *E.coli* and plated onto LB agar plates with appropriate antibiotics. All constructs were verified by sequencing (Research Biolabs).

Chapter Two: Materials and Methods

2.3.1.13 Small scale protein induction

The plasmid containing different target gene was transformed into *E.coli* BL21 (DE 3) or origami and cultured on the LB plate with antibiotic. A single colony was inoculated into LB/antibiotic and cultured at 37 °C, 250 rpm, when the OD600 get to 0.6-1.0, IPTG was added to 0.1 mM final concentration, then the cells were cultured for another 3 hours. The cells were mixed with SDS-loading dye and heat at 100 °C for 10 min and loaded into SDS-PAGE for analysis.

2.3.1.14 Large scale protein induction

A single colony was inoculated into 25 ml LB with antibiotic in 125 ml flask and cultured at 37 °C, 250 rpm overnight, 5 ml of the overnight culture was transferred into 500 ml LB medium with antibiotic in a 2 L flask and cultured at 37 °C, 250 rpm. IPTG was added to 0.1 mM final concentration when the OD600 get to 0.6-0.8, and the induction was performed at 20°C, 220 rpm for 3 h. The cells were harvested by centrifuging at 5000 g for 10 min for protein purification. The cell pellet can be stored at -20°C for the purification later.

2.3.1.15 Cell lysis

Cell pellets were broken by sonication. Cells were first completely resuspended in resuspension buffer (1g pellet adding with 10 ml resuspension buffer) and 1 mg/ml lysozyme was added to ease cell breakage. The cells solution was incubated on ice for 30 min, cell were sonicated for 15 min at 20% power. In order to confirm if the cells were broken completely, protein concentration was measured in the supernatant, if there is no

Chapter Two: Materials and Methods

increase in protein concentration when the sonication time is extended, the sonication should be finished.

2.3.1.16 Purification of protein by Ni^{2+} -NTA affinity column

The sonicated cell lysate was cleared by centrifuging at 25,000g for 30 min; the supernatant was used for further purification. The Ni^{2+} -NTA resin was loaded into the disposable column from BioRad, 2 ml resin was used for the purification of 1 liter of LB cultured cell lysate. About 10 ml of resuspension buffer was used to equilibrate the resin to remove the resin storage buffer. The supernatant obtained from the lysate was loaded to the column and then followed by 30-40 ml washing with washing buffer to remove the unspecific binding proteins. To elute the His tag fusion protein, 5 ml of elution buffer was used. The elution fractions were loaded into SDS-PAGE for analysis. During this purification step, DTT and EDTA are avoided to be used in any of the buffers.

2.3.1.17 Protein concentration assay

The dye from Bio-Rad was used for the assay. The protein concentration was checked by the following way: Prepare dye reagent by diluting 1 volume of Dye Reagent Concentrate with 4 volumes of distilled, deionized water; Filter through Whatman #1 filter (or equivalent) to remove particulates; This diluted reagent may be used for approximately 2 weeks when kept at room temperature; Prepare three to five dilutions of a protein standard, which is representative of the protein solution to be tested; The linear range of the assay for BSA is 0.2 to 0.9 mg/ml, whereas with IgG the linear range is 0.2 to 1.5 mg/ml; Pipet 100 μl of each standard and sample solution into a clean, dry test tube.

Chapter Two: Materials and Methods

Protein solutions are normally assayed in duplicate or triplicate; Add 5.0 ml of diluted dye reagent to each tube and vortex; Incubate at room temperature for at least 5 minutes; Absorbance will increase over time; Samples should incubate at room temperature for no more than 1 hour; Measure absorbance at 595 nm; Plot the standard curve based upon the known protein concentration; Get the relationship between the protein concentration and the OD₅₉₅; Using the OD₅₉₅ of protein sample to calculate the protein concentration. Also the purified protein concentration can be detected by checking OD₂₈₀. Diluting sample and measuring the OD₂₈₀, and using the analysis data from Vector NTI, the concentration of the protein sample can be obtained.

2.3.2 General methods for cell culture

All cell lines were maintained in 5% CO₂ at 37°C in a humidified tissue culture incubator using Nunc tissue culture flasks or dishes.

2.3.2.1 Cell storage in liquid nitrogen

Cells were sedimented at 400 g for 5 min, resuspended in cell freezing media at a concentration of 5x10⁶ cells/ml and dispensed into Cryo Vials (Greiner). Cells were frozen in a NALGEN™ Cryo 1°C Freezing Container (Nalgen) for 24 h at -80°C to achieve a 1°C/min cooling rate. Thereby the vials were transferred into liquid nitrogen for long-term storage.

2.3.2.2 Cell recovery from liquid nitrogen

Cells were removed from the liquid nitrogen storage and quickly brought to 37°C in a

Chapter Two: Materials and Methods

water bath. Cells were washed in 10 ml warmed media to remove DMSO. Each cell pellet was resuspended in complete media and cultured in 75 cm² flask.

2.3.2.3 Culture of HEK293, HEK293T and HEK293GnTI cells

HEK293, HEK293T and HEK293GnTI cells were maintained in DMEM with 10% (v/v) heat-inactivated FBS, 100 IU/ml penicillin and 100 µg/ml streptomycin in a 5% CO₂ and humidified tissue culture incubator at 37°C. Cells were passaged when they were subconfluent. Cells were washed once in PBS, incubated in 0.25% (w/v) trypsin (GIBCO) for 5 min at 37°C, followed by tapping of each flask to dislodge the adherent cells. Trypsin was inactivated by adding full media and cells were directly seeded into fresh culture media in new flasks at desired cell density.

2.3.3 β2 fragments expression, purification and crystallization

2.3.3.1 Transfection of HEK293 and HEK293T cells by the calcium phosphate method (for 10 cm dish)

Cells were split into 10 cm dishes the night before to give 50-70% confluence at the day of transfection. On the next day, 10 ml of fresh media containing 25µM chloroquine was added to replace the old media one hour prior to the transfection. 10µg DNA was added to ddH₂O (1095 µl total) in a 15 ml sterile tube, followed by adding in 155 µl 2M CaCl₂. 1250 µl of 2xHBS (8.0g NaCl, 0.37g KCl, 201mg Na₂HPO₄•7H₂O, 1.0g glucose, 5.0g HEPES/500ml pH 7.05 and filter sterile, store at 4°C) was added to that tube drop by drop with gentle mixing. Finally, this mixture was added directly to the cells drop by drop. This step should be done within 1 min after adding 2xHBS. The cells were then

Chapter Two: Materials and Methods

incubated for 7-11 h. After incubation, 10 ml of fresh media was added to replace the old media containing chloroquine. Cells, or culture supernatant, can be harvested 48-72 h after transfection.

2.3.3.2 Transfection of HEK293T and HEK293 GnTI⁻ cells by PolyFect method (for 6cm dish)

One day before transfection, cells were plated in 3ml of cell growth media to achieve a 50-70% confluency at the time of transfection. On the next day, 4 µg DNA were diluted by 150 µl of serum and antibiotics free media. 35 µl of PolyFect Transfection Reagent was added to the diluted DNA solution, and mixed gently and incubated for 5-10 minutes at room temperature to allow complex formation. During this period, aspirated the media from the dish gently and added 3ml fresh cell growth media. After adding 1ml of cell growth media, the complex solution was immediately added to cells gently and evenly. Cells were incubated at 37°C and 5% CO₂ for 48-72h to allow for gene expression. Cells or cultured media were harvested for further uses.

2.3.3.3 Creation of stable cell line

a. The pcDNA3 coding for PHE1-His was transfected to HEK293 cell (as described in section 2.3.3.1). After 10 hours of transfection, the cells were washed once with PBS and were incubated in 10 ml fresh DMEM medium. After 36 hours, the medium was changed to selective medium (fresh DMEM medium with 1mg/ml G418 antibiotic). The selective medium was changed every 3 days for two weeks. The cells were split to approximately 100 cells per 10cm dish in selective medium using dilution. Once G418 resistant foci

Chapter Two: Materials and Methods

were identified, they were picked and expanded in 24-well plates. Supernatants were collected from the plates and protein expression was analysed by Western blot using anti-His tag antibody.

b. The pIRES2-EGFP coding for PHE2-His and PHE3-His was transfected to HEK293GnTI⁻ cell (as described in section 2.3.3.2). After 12 hours of transfection, the medium was changed to selective medium (fresh DMEM medium with 1mg/ml G418 antibiotic). After 2 weeks of antibiotic selection, the top 0.3% of fluorescent cells was separated by fluorescence activated cell sorting (FACS) Aria flow cytometry-based cell sorter (BD Biosciences). These cells were allowed to grow in selective media for several days and re-sorted into a 96-well plate with one cell per well. The cultured supernatants were collected and protein expression was analysed by Western blot using anti-His tag antibody. The cells secreting the highest recombinant proteins were maintained in culture media.

2.3.3.4 Recombinant protein expression and purification

Cells with the highest expression were grown in 150 mm tissue culture dishes in DMEM supplemented with Penicillin/Streptomycin and 10% FBS. The spent media were harvested every 10 days, centrifuged 10,000g for 20 min to remove the cell debris, and exchanged into buffer A (50 mM Tris-HCl, 150 mM NaCl, pH 7.5). The recombinant proteins were purified using Ni-NTA agarose beads (Qiagen), and were concentrated to 1-2 ml, and subjected to size exclusion chromatography on a Superdex75 16/180 column (Amersham) in buffer A. If the recombinant proteins were not pure enough for

Chapter Two: Materials and Methods

crystallization, an additional purification step of anion-exchange chromatography was carried out after Ni-NTA purification as following: The proteins were dialyzed against 30 mM Tris HCl at pH 7.5 overnight at 4°C and loaded onto a MonoQ 5/5 column (Amersham), mounted on an Akta FPLC system (Amersham). The flow-through containing the protein, as showed by SDS-PAGE, was pooled and concentrated for further purification using size exclusion chromatography.

2.3.3.5 Deglycosylation of the recombinant glycoproteins

a. Using peptide N-glycosidase F

A mixture (2.5µl) containing 5% (w/v) SDS and 10% (v/v) 2-mercaptoethanol was added to 18µl (18µg) pure protein solution (PHE1). The mixture was boiled at 100°C for 10 min. 2.5µl of the 10% (v/v) Nonidet P40 in 0.5M sodium phosphate (pH7.5) and 2ul (10 units) of peptide N-glycosidase F (New England Biolabs) were added. The mixture was incubated for 1hour at 37°C.

b. Using Endo H_f

Purified PHE2 and PHE3 were concentrated to 1 mg/ml in buffer A and treated with Endo H_f (New England Biolabs, 4,000 units/ml) overnight at room temperature, and further purification by size exclusion chromatography. Molecular masses were determined by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry.

Chapter Two: Materials and Methods

2.3.3.6 Crystallization and post-crystallization treatments

After purification, the recombinant proteins were concentrated to 15-20mg/ml. Crystallization was performed using the hanging drop vapor diffusion method at 18°C with sparse crystal screen kit1 and 2 (Sigma), Nextal PEGs I, II suites (Qiagen). One microliter of the protein was mixed with an equal volume of the reservoir buffer and equilibrated against 1ml reservoir buffer. Initial crystal conditions were obtained and optimized to give the best quality crystals for data collection, together with seeding. Seeding: The initial crystal was removed to a stabilizing solution which had a 10-20% higher concentration of precipitant than the mother liquor to prevent the crystal from dissolving, and was crushed evenly with a needle or a glass fiber with a ball at the tip. The resulting suspension was added to a tube containing 100 µl of the harvesting buffer and was centrifuged for a five minutes at around 100 g. The supernatant containing seeds were used for seeding at a series of dilutions with mother liquor from 10^{-3} - 10^{-5} . All the diffraction experiments were carried out at cryogenic temperatures (100K) to reduce the thermal motion and radiation-induced damage. The crystals were soaked in the cryo solution for a short time (seconds or minutes), then plunged directly into liquid nitrogen (rather than flash cooling in the nitrogen gas). Several cryo-protectants were optimized for best results (Table 2.1) (Rodgers, 2001).

One of the post-crystallization treatments is the dehydration of protein crystals to improve their diffraction quality and the resolution limit. The normal protocol for crystal dehydration can be performed by transferring the crystals into a dehydrating solution, that is usually the original mother liquor or with a higher concentration of precipitant or

Chapter Two: Materials and Methods

supplemented with cryoprotective agents such as PEG400, PEG600, MPD or glycerol. A more gentle dehydrating method consists of the serial transfer of the cover slip holding the crystal droplet over reservoirs containing increasing concentrations of dehydrating solution with incubations of 8–12 h over each condition (Fig.2.1.) (Heras et al., 2005).

Table 2. 1 Commonly used cryo-protectants

Cryo-protectant	Suggested concentration
DMSO	2-20%
Erythritol	>50% (w/v)
Ethylene glycol	15-45%
Glycerol	15-45%
Inositol	20-50%
MPD	25-45%
PEG 200-600	30-50%
Raffinose	>50% (w/v)
Sucrose	>50% (w/v)
Ammonium sulphate	50% saturation
Iso-propanol	>70%
Xylitol	20-50%

Figure 2. 1 Dehydration

Transfer cover slip to reservoirs containing serial increases of dehydrating solution

Chapter Two: Materials and Methods

Table 2. 2 The 23 most commonly used heavy-atom compounds reagents

The first column gives the number of times the reagent has appeared in the heavy-atom data bank (Carvin, 2001).

Used times	Compound
287	Potassium tetrachlorophlatinate(II)
111	Potassium dicyanoaurate(I)
103	Uranyl acetate
101	Mercury(II) acetate
98	Mercury(II) chloride
85	Ethylmercurythiosalicylate (EMTS)
82	Potassium tetraiodomercurate(II)
81	Para-Chloromercuriobenzenesulfonate (PCMBS)
75	Trimethyllead(IV) acetate
73	Potassium pentafluorooxyuranate(VI)
73	Phosphatotris(ethylmercury)
61	Potassium tetranitritoplatinum(II)
60	Uranyl nitrate
58	Potassium tetracyanoplatinate(II)
57	Dichlorodiammineplatinum(II)
51	Potassium hexachloroplatinate(IV)
51	Methylmercury chloride
44	Potassium tetrachloroaurate(III)
42	Para-Chloromercurybenzoate (PCMB)
39	Lead(II) acetate

2.3.3.7 Heavy atom derivatives

Several commonly used heavy atom compounds such as K_2PtCl_4 , $HgCl_2$, $PbAc$, etc. had been tried (Table 2.2). Finally, K_2PtCl_4 proved to be successful. The Pt derivative was prepared by soaking the crystals into the mother liquor with 1mM K_2PtCl_4 for 12 hours.

2.3.3.8 Data collection

The crystals were soaked into the cryoprotection solution, harvested with cryo-loops and the latter were mounted on a goniometer head and cooled to 100K in a nitrogen gas stream (Oxford Cryosystems). All X-ray reflection data were collected in the oscillation method. The diffraction data from a single cryocooled crystal were recorded on an ADSC charge-coupled device (CCD) detector (ADSC Corporation, Poway, CA) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) or on a Raxis IV Image Plate detector using CuK radiation from a Micromax-007 rotating anode operating at 20 mA and 40 kV. A complete dataset consists of a consecutive series of these images covering a total rotation of 30° to 360°, depending on the symmetry (space group) of the crystal. A dataset was also collected at a wavelength of 1.907 Å and anomalous difference Fourier synthesis was performed to confirm the location of the disulfide bonds. Data were processed using programs MOSFLM (Leslie, 2006) and SCALA(1994).

2.3.3.9 Structure determination and refinement

The PHE1 structure was determined by the single isomorphous replacement method using anomalous scattering (SIRAS) from a K₂PtCl₄ derivative. Two heavy atom binding sites were located using the program SOLVE (Terwilliger, 2003; Terwilliger et al., 1999), and an initial map was calculated after solvent flattening. Phases calculated from the partial model combined with the experimental SIRAS phases were used to calculate a new map that was modified using the program DM (CCP4, 1997). The whole procedure was iterated, and the envelope was modified. Once about 70% of the residues had been

Chapter Two: Materials and Methods

traced, the program ARP/WARP (Perrakis et al., 1997) was used for final phase improvement and model building. The model was refined by slow cool energy minimization and B-factor refinement protocols of the program CNS (Brunger et al., 1998) with the Engh and Huber force field constants using the maximum likelihood amplitude target. Manual inspection and correction of the model were made using the program O (Jones et al., 1991). The free *R*-factor was calculated from 5% of the measured unique data, randomly chosen, that were not included in the refinement. The coordinates and structure factors have been deposited in the Protein Data Bank with code 1YUK.

The PHE2 and PHE3 structures were determined by molecular replacement using the program PHASER (Storoni et al., 2004) with PHE1 (PDB, 1YUK) (Shi et al., 2005) and PHE2 (PDB, 2P26) lacking the I-EGF2 domain, as search models, respectively. Molecular replacement for the I-EGF3 domain was performed using its NMR structure 1L3Y (Beglova et al., 2002) as a search probe. Both PHE2 and PHE3 Models were improved by alternating cycles of model building with program Coot (Emsley et al., 2004) and refinement with REFMAC (CCP4, 1997). TLS refinement was introduced in the last refinement step (Winn et al., 2003). The free *R*-factor was calculated from 5% of the measured unique data, randomly chosen, that were not included in the refinement. The PHE2 and PHE3 coordinates and structure factors have been deposited in the Protein Data Bank with code 2P26 and 2P28, respectively. The stereochemistry of the obtained structures was checked with PROCHECK (Laskowski et al., 1993). Solvent Accessible Surfaces Area (SASA) was calculated using program AREAIMOL (CCP4, 1997) with a

Chapter Two: Materials and Methods

1.7Å radius sphere as the probe. Conformational differences were analyzed using the DynDom server (<http://www.cmp.uea.ac.uk/dyndom/main.jsp>). Sequence alignments were performed using CLUSTALW (Thompson et al., 1994). Superposition of structures and rms deviation calculations were carried out using the program Lsqkab (CCP4, 1997). Structural figures were created using the program Pymol (written by Dr. Delano, available at pymol.sourceforge.net), except other stated.

2.3.4 Cell based and biochemical experiments

2.3.4.1 Harvesting transfected cells (adherent)

48-72 h after transfection, the cells in each plate were washed in PBS and detached in 5 ml of 0.5 mM EDTA in PBS by incubation at RT for 10 min with gentle rocking. Cells were collected and mixed with 10 ml full media and transferred to centrifuge tubes. Cells were spun down (400g, 5 min, 4 °C) for subsequent analysis.

2.3.4.2 FACS analysis

Cells were incubated with 20 µg/ml primary mAb in DMEM for 1 h at 4°C. The cells were then washed twice and incubated with FITC-conjugated sheep anti-mouse F(ab')₂ secondary antibody (1:400 dilution; Sigma) for 45 min at 4°C. Stained cells were washed once in DMEM and fixed in 1% (v/v) formaldehyde in PBS. Cells were analyzed on a FACSCalibur flow cytometer (Becton Dickinson). Data were analyzed using CellQuest pro software (Becton Dickinson). Expression index was calculated by (% cells gated positive) x (geomean fluorescence intensity).

Chapter Two: Materials and Methods

2.3.4.3 Preparation of rabbit anti-mouse IgG coupled onto protein A sephrose beads

Protein A sephrose (PAS) (Sigma) (1g) was swelled by rotating in an excess of PBS overnight at 4°C. The swelled PAS beads, which occupied a bed volume of 4 ml, was sedimented by centrifugation (3000 g, 5 min, 4°C), washed twice, and finally resuspended to a 25 % (v/v) suspension in PBS. Rabbit anti-mouse IgG (RaM, Sigma) (4 mg) was added and the mixture was rotated for 1 h at 4°C. The PAS-RaM bead were washed once in PBS, once in IP buffer, and resuspended to a 25 % (v/v) suspension in IP buffer. The PAS-RaM suspension was stored at 4°C.

2.3.4.4 Labelling of free cysteines

Free cysteine labeling was performed as described (Luo et al., 2004c). Briefly, transfected cells were washed once Tris-buffered saline (TBS) (20 mM Tris-HCl, pH 7.4, 200 mM NaCl) containing 1mM Ca^{2+} , and labeled with 400 μM of biotin-BMCC (1-biotinamido-4 (4' [maleimidoethylcyclohexane]-carboxamido) butane) (Pierce) for 30 min at room temperature. Biotin-BMCC solution must be prepared immediately before use at approximately 8.5 mM in DMSO by dissolving 2.5 mg Biotin-BMCC in 0.55 ml DMSO. Cells were washed three times with TBS with 1mM Ca^{2+} , and lysed by adding 1 ml immunoprecipitation buffer with the addition of 1 mM PMSF and incubated on ice for 20 min. Cell debris were removed by centrifugation at 12000 g at 4°C for 10 min, and the cell lysates were transferred to fresh microfuge tubes. The lysates were stored at -20°C.

Chapter Two: Materials and Methods

2.3.4.5 Surface biotinylation and preparation of whole cell lysates

Harvested cells were pelleted by centrifugation (400 g, 5 min, 4 °C), washed twice in ice cold PBS and resuspended in ice cold PBS at $\sim 2.5 \times 10^6$ cells/ml in microfuge tubes. 200 μ l cells were mixed with 200 μ l 1 mg/ml sulphy-NHS-biotin (Pierce) that was freshly prepared in PBS, and incubated on ice for 30 min. Reaction was quenched by adding 5 ml 10 mM Tris-HCl pH 8.0 in PBS and washing the cells twice in the same solution. Cells were resuspended in 200 μ l PBS and transferred to a 1.5 ml eppendorf tube. The labeled cells were lysed by adding 1 ml immunoprecipitation buffer with the addition of 1 mM PMSF and incubated on ice for 20 min. Cell debris were removed by centrifugation at 12000 g at 4°C for 10 min, and the cell lysates were transferred to fresh microfuge tubes. The lysates were stored at -20°C.

2.3.4.6 Immunoprecipitation

0.5 ml of biotinylated cell lysates or 0.25 ml cultured media were precleared by adding 3 μ g of irrelevant mAb as appropriate, and incubated at 4°C for 45 min with rotation. 40 μ l PAS-RaM suspension was added and the mixture was rotated at 4°C for 30 min. The PAS-RaM was sedimented by centrifugation (2000 g, 5 min, 4 °C) and the supernatant was transferred to a fresh tube. 3 μ g of appropriate mAb was added to the precleared cell lysate and incubated at 4°C for 45 min with rotation. Subsequently, 70 μ l PAS-RaM suspension was added and the mixture was incubated at 4°C for 1 h with rotation. The PAS-RaM was sedimented by centrifugation (2000 g, 4°C, 3 min) and the supernatant was discarded. The PAS-RaM was washed in 500 μ l IP-buffer for three times. Thereafter, 100 μ l protein solubilization buffer containing 40 mM DTT was added to the PAS-RaM

Chapter Two: Materials and Methods

which was then vortexed, heated at 100°C for 5 min. The immunoprecipitated proteins were resolved by directly running on SDS-PAGE gels, or were stored at -20°C until used.

2.3.4.7 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Laemmli SDS-PAGE was performed with slight modifications. Protein sample (12 µl for minigel system) was mixed with an equal volume of sample loading buffer (2x) under non-reducing conditions or in the presence of 40 mM DTT under reducing conditions. Electrophoresis was carried out at 200V in a Mini Electrophoresis Set (Biorad) in SDS-PAGE electrophoresis buffer.

2.3.4.8 Western blot

Proteins separated by SDS-PAGE were transferred onto a PVDF membrane (Immobilon-P, Millipore). The PVDF membrane was prepared according to the manufacturer's instructions: the membrane was equilibrated by washing in ethanol for 10 sec, then water for 10 min until miscible, and then in blot buffer (12 mM Tris-HCl, 95 mM Glycine, 10 % (v/v) ethanol) for 10 min. One western blot sponges, the PVDF membrane, the SDS-PAGE gel and finally the second sponge, was assembled and placed in a semi-dry western blot apparatus (Biorad), bubbles between each layer were removed carefully. Western blot was performed at 25 V for 30 min. After protein transfer, the PVDF membrane was transferred to blocking buffer and was rotated at RT for an hour or overnight at 4°C.

Chapter Two: Materials and Methods

2.3.4.9 ECL detection proteins blotted onto immobilon-P membrane

Proteins had been transferred onto PVDF membranes by western blot. For surface labeled studies, proteins were labeled with a streptavidin-HRP (Horse radish peroxidase) conjugate (Amersham) and were then detected by enhanced chemiluminescence (ECL). For non-surface labeled analysis, proteins were detected by diluted primary antibody whose optimal concentration should be determined by titration. Followed by washing and horseradish peroxidase (HRP) conjugated secondary antibody, and were detected by enhanced chemiluminescence (ECL). Membrane was removed from blocking buffer and were washed three times in PBS-T (PBS, 0.1% (v/v) Tween-20 (Sigma)) at RT for 10 min. Membrane with surface labeled proteins was then incubated with the streptavidin-HRP conjugate (1:1000 dilution in PBS-T) at RT for 60 min with gentle agitation, and was washed again for three times in PBS-T at RT for 10 min. Membrane with non-surface labeled proteins was incubated with primary antibody at RT for 60 min with gentle agitation, and was washed again for three times in PBS-T at RT for 10 min. Followed by the horseradish peroxidase (HRP) conjugated secondary antibody for 45 min at RT with gentle agitation, and washed for three times in PBS-T at RT for 10 min. Thereafter, all the membranes were developed using an ECL Plus Detection Kit (Amersham) according to the manufacturer's recommendations. Membrane was then exposed to Kodak X-Omat film and developed using a Kodak X-OMAT ME-1 processor.

2.3.4.10 Coating microtitre plates with ICAMs for cell adhesion assay

Goat anti-human IgG (Fc specific) (Sigma) was diluted to 5 µg/ml in sodium bicarbonate buffer, pH 9.2, and 50 µl was added to each well of a microtitre plate (Polysorb, Nunc

Chapter Two: Materials and Methods

Immuno-Plate). The microtitre plates were left at 4 °C overnight. On the next day, the solution was discarded and the wells were washed twice with 150 µl PBS per well. A solution of 0.5 % (w/v) BSA in PBS was added to each well (150 µl per well) and the plates were incubated at 37°C for 30 min. The plates were washed with PBS once (150 µl per well) before the addition of 50 µl per well of ICAM-1/Fc or ICAM-3/Fc (1 µl /ml in PBS containing 0.1 % BSA). After 2-3 h incubation at RT, the plates were washed twice in RPMI/HEPES/FBS (RPMI1640 supplemented with 10 mM HEPES, pH 7.4 and 5% (v/v) heat inactivated FBS) before use.

2.3.4.11 Cell adhesion assays

Plates were prepared as described in 2.3.3.19. Cells were then incubated with BCECF dye (2',7'-bis-(2-carboxyethyl)-5(and -6) carboxy fluorescein, acetoxymethyl ester) (Molecular Probes), at the concentration of 1 µg/ml for 20 min at 37°C. Labeled cells were transferred to wells (2x10⁴ cells/well) alone or with stimulation/inhibition reagents and incubated for 30 min at 37°C. Nonadherent cells were removed by washing. Cell fluorescence, which indicated the number of cells adhering to ligand-coated wells was measured using FL600 fluorescence plate reader (Bio-Tek).

2.4 Plasmid construction details

2.4.1 Integrin plasmids

2.4.1.1 Wild type plamids

The $\beta 2$ cDNA in J8.1E (Douglass et al., 1998) was used as a template for construction of the $\beta 2$ in mammalian expression vector pcDNA3.0 (Invitrogen) (Fig.2.2). KpnI and SpeI

Chapter Two: Materials and Methods

were used to digest $\beta 2$ cDNA from J8.1E and the fragment was ligated into pcDNA3.0 vector which was digested with KpnI and XbaI (SpeI and XbaI digested DNA fragments have compatible ends). αL cDNA plasmids were in pcDNA 3.0 vector with the restriction cutting sites KpnI and XbaI (The whole mRNA of the $\beta 2$ subunit is available in appendix section).

2.4.1.2 Hybrid domain (HD) constructs

The vector pcDNA3.0 (Invitrogen) (Fig.2.2) and PET24a (Appendix section) were used for HD expression. PCR products of HD with hexa-histidine at the C-terminal were directly linked to the signal sequence using overlapping PCR technique, and digested with HindIII/XbaI and were ligated into pcDNA3.0 which was digested with the same enzymes (more detailed information about the constructs is included in the appendix section).

2.4.1.3 PHE1 expression vector

The vector pcDNA3.0 (Invitrogen) and PET24a were used for PHE1 expression. PCR products of PHE1 with hexa-histidine at the C-terminal were digested with HindIII/XbaI and were ligated into pcDNA3.0 which was digested with the same enzymes. Specific primers are listed below, and the restriction enzyme sites are underlined (more detailed information about the constructs is included in the appendix section).

F 5' CCC AAG CTT GGG CAG ACT GGT AGC 3'

R 5' GC TCT AGA TCA GTG GTG GTG GTG GTG GTG CTC ACA GTT TTT
CC3'

Chapter Two: Materials and Methods

Figure 2. 2 Schematic illustration of plasmid J8.1E and pcDNA3.0.

The $\beta 2$ cDNA in J8.1E (Douglass et al., 1998) was used as a template for the construct of the $\beta 2$ in mammalian expression vector pcDNA3.0. KpnI/SpeI was used to digest $\beta 2$ cDNA from J8.1E and the fragment was ligated into pcDNA3.0 vector which was digested with KpnI and XbaI (SpeI and XbaI digested DNA fragments have compatible ends). αL plamid was in pcDNA3.0 vector with the restriction cutting sites KpnI and XbaI.

Chapter Two: Materials and Methods

Restriction Map and Multiple Cloning Site (MCS) of pIRES2-EGFP Vector. Unique restriction sites are in bold. Note that the *Eco47 III* site has not been confirmed in the final construct.

Figure 2. 3 Schematic illustration of pIRES2-EGFP

The restriction enzymes *NheI* and *XhoI* were used for both PHE2 and PHE3 constructs.

2.4.1.4 PHE2-pIRES2-EGFP and PHE3-pIRES2-EGFP expression vector

The vector pIRES2-EGFP (BD bioscience) (Fig.2.3) was used for PHE2 and PHE3 expression. The restriction enzymes *NheI* and *XhoI* were used for both PHE2 and PHE3. PCR products of PHE2 or PHE3 with hexa-histidine at the C-terminal were digested with *NheI/XhoI* and were ligated into pIRES2-EGFP vector that was digested with the same enzymes. Specific primers for are listed below and the restriction enzyme sites are underlined.

Chapter Two: Materials and Methods

PHE2:

F 5' AC CAT GCT AGC GGG CAG ACT GGT AGC 3'

R 5' CCG CTC GAG TCA GTG GTG GTG GTG GTG GTG CTC GCA GTA CTG
CCC G 3'

PHE3:

F 5' AC CAT GCT AGC GGG CAG ACT GGT AGC 3'

R 5' CCG CTC GAG TCA GTG GTG GTG GTG GTG GTG CTG GCA CGC TGA GC
3'

Chapter Three: The structure of PHE1

The integrin $\beta 2$ subunit can be expressed in the absence of any integrin α subunits on COS-7 cells but can only be detected by monoclonal antibodies whose epitopes map to regions outside the β I (I-like) domain (Huang et al., 1997a; Tan et al., 2001a). This suggests that the proper folding of the β I domain requires the interaction with the β -propeller domain of the associated α subunit of the integrin headpiece (Fig.3.1A). To avoid the complexity introduced, a $\beta 2$ subunit without the β I domain ($\beta 2\Delta I$) has been constructed. When transfected to COS-7 cells, as expected, the $\beta 2\Delta I$ subunit failed to associate with the αL subunit (Tan et al., 2001a), the surface expression of $\beta 2\Delta I$ was detected using a panel of monoclonal antibodies outside of the β I domain, which indicates that the $\beta 2\Delta I$ subunit can be expressed with the proper folding at least in mammalian cells. Therefore, two constructs were chosen for the initial study, the hybrid domain (HD), and PHE1, which contains the PSI (residues Gln1-Asp58 and Glu424-Arg426), hybrid (residues Pro59-Ala100 and Lys340-Cys423), and I-EGF1 (residues Cys427-Glu460) domains. Hexa-histidine tag was added at the C-terminal of both constructs to facilitate downstream recombinant proteins purification (Fig.3.1B). In an initial attempt, I tried to express those two constructs in *E.coli* as this host has many advantages over other expression systems.

Chapter Three: The structure of PHE1

Figure 3. 1 Constructs of HD and PHE1

(A) Schematic representation of the domain organization within an integrin $\alpha\beta$ heterodimer. The α -leg contains three β -sandwich domains and the β -leg comprises an N-terminal plexin, semaphorin and integrin (PSI) domain, a unique immunoglobulin-like hybrid domain, four integrin-epidermal growth factor (I-EGF) and a β -tail domain (β -TD). (B) Illustrations of protein constructs. Both hybrid domain and PHE1 are devoid of the β I domain. The sequence numbering scheme is indicated. The same color codes for the domains are conserved throughout: yellow, hybrid domain; green, PSI domain; red, I-EGF1 domain; purple, I-EGF2; pink, I-EGF3; The β -TD, β -Tail Domain; TM, Transmembrane helix; Cyt, Cytoplasmic domain are not colored. (His)₆, hexa-histidine tag.

3.1 Expression and purification

3.1.1 The HD expression in *E.coli*

The PCR amplified hybrid domain coding region was cloned into the pET24a vector to express the HD (Hybrid domain with hexa-histidine tag at the C-terminal) (Analyses for the constructs using vector NTI are shown in Appendix 2-3). The recombinant plasmids were transformed into *E.coli* BL21 (DE3). After the induction with 0.1 mM IPTG at 20 °C, the expression of HD was analyzed from the whole cell lysate, the supernatant and the pellet. The HD was found exclusively in the pellet, but not in the soluble fraction. Anti-his-tag Western blot analysis confirmed the expression of a protein with a molecular weight (MW) of around 16 kDa which only presented in the pellet (Fig.3.2).

Figure 3. 2 Hybrid domain expression in *E.coli* BL21

(A) HD protein expression induced with 0.1 mM IPTG at 20 °C in *E.coli* BL21. The induction was checked by 12% SDS-PAGE. P: pellet; S: supernatant; W: whole lysate; M: Protein markers. (B) Anti-His-tag Western blot analysis for the same induction.

Chapter Three: The structure of PHE1

In order to circumvent insolubility problem, HD was expressed as fusion proteins with thioredoxin reductase (Trx)-HD, Maltose Binding Protein (MBP)-HD and Glutathione S-Transferase (GST)-HD (vector information is supplied in appendix 4). In all cases, a factor Xa cleavage site was incorporated between the carrier proteins and HD. The fusion proteins were expressed in BL21 (ED3). GST-HD (42kDa) was found in the whole cell lysate and pellet but not in the supernatant. However, significant amount of soluble complexes were found both in the MBP-HD (58.5kDa) and Trx-HD (28kDa) (Fig.3.3). Therefore, the fusion proteins of MBP-HD and Trx-HD were expressed for further analysis. After Ni^{2+} -nitrolotriacetic acid (Ni-NTA) affinity chromatography purification, MBP-HD was cleaved with factor Xa for 18h at RT, and most of the MBP-HD was converted to MBP and HD (16kDa) as shown by SDS-PAGE (Fig.3.4).

The factor Xa digested MBP-HD protein mixture was subjected to gel filtration on Sephacryl S-100HR (Fig.3.5A), and the results were analyzed on SDS-PAGE (Fig.3.5B). Peak1 contained a mixture of MBP-HD at 58.5kDa and HD at 16kDa. Thus it was concluded that upon cleavage from MBP fusion protein, HD formed an aggregate and was eluted in the void volume. Only MBP (43kDa) was found in Peak2 peak as monomer. Similar results were also obtained from the Trx-HD experiments (data not shown).

Chapter Three: The structure of PHE1

Figure 3. 3 Three HD fusion protein expressions

The expressions of three fusion proteins (MBP-HD, GST-HD and Trx-HD) induced with 0.1 mM IPTG at 20 °C in *E.coli*. The inductions were checked by SDS-PAGE. P: pellet; S: supernatant; W: whole cell lysate; and M: protein markers.

Figure 3. 4 MBP-HD purification and its cleavage by Factor Xa

(A) Ni-NTA purification. W: whole cell lysate; P: pellet; S: supernatant; FT: Flow through that did not bind to the Ni-NTA resin; W1: Wash fraction of the resin using 15 mM imidazole; E1-3: Elution fractions of MBP-HD from the Ni-NTA resin using 500 mM imidazole; M: protein markers. (B) MBP-HD was digested by factor Xa at RT. The digested mixtures were analysed at different intervals on 12% SDS-PAGE.

Chapter Three: The structure of PHE1

Figure3. 5 Gel filtration purification of the factor Xa digested MBP-HD

(A) Gel filtration analysis for digested MBP-HD. The digested mixture was loaded on a Sephacryl S-100HR column mounted on FPLC and pre-equilibrated with 20mM Tris, pH8.0, 150mM NaCl, 10% glycerol. (B) After gel filtration, peaks were analyzed on 12% SDS-PAGE. S: the sample before loading; P1 P2 were the eluted peaks, M: protein markers.

Chapter Three: The structure of PHE1

3.1.2 The PHE1 expression in *E.coli* BL21(DE3)

The PHE1 was also cloned in to the pET24a vector, and the recombinant plasmids were transformed into *E.coli* BL21 (DE3). After the induction with 0.1 mM IPTG at 20°C, the expression of PHE1 (27kDa) was analyzed from the whole cell lysate, the supernatant and the pellet. The majority of PHE1 was found in the pellet but not in the soluble fraction (Fig.3.6A).

Figure 3. 6 PHE1 and trx-PHE1 fusion protein expressions in *E.coli*.

(A) The expression of PHE1 in *E.coli* BL21 (DE3). (B) The expression of Trx- PHE1 in *E.coli* BL21 (DE3) or in origami host. (-): whole cell lysate without IPTG induction; P: pellet; S: supernatant; W: whole cell lysate; M: Protein markers.

Chapter Three: The structure of PHE1

In order to increase the PHE1 solubility, the fusion protein Trx-PHE1 was constructed in the pET32a vector and expressed in *E.coli* BL21 (DE3) at 20 °C or in the origami strain at 18 °C. The origami strain, which bears mutations in both the thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes, is considered to be able to enhance the formation of the disulfide bond in the cytoplasm. Thus it is suitable for the expression of proteins with disulfide bonds, such as PHE1. In either case, no apparent soluble Trx-PHE1 fragment was found in the supernatant (Fig.3.6B). Thus, HD and PHE1 can be expressed in *E.coli*, but they do not appear to fold properly. We therefore switch to the mammalian expression system which is the original system for integrin expression.

3.1.3 Transient expression and purification of HD in HEK293T cells

The HD was constructed in pcDNA3 vector with the native signal peptide of the integrin β 2 subunit at its N-terminal. After transfection to HEK293T cells, the cultured medium was collected after seven days and cleared by centrifugation. HD was purified by a Ni-NTA column (Fig.3.7A) followed by a gel filtration column (Sephacryl S-100HR, 16/60) on FPLC, and HD eluted at around 57 ml after sample injection, which indicated that HD existed as a monomeric form (Fig.3.7B). The fractions (A5-A7) containing the HD were pooled and concentrated before running on SDS-PAGE. Due to the glycosylation modification, the HD band was shown at around 19kDa with 4kDa higher than the theoretical calculated molecular weight. It should be mentioned that high molecular weight contaminated proteins were still present (Fig.3.7C), however, what made up the contaminants was not further investigated.

Chapter Three: The structure of PHE1

Figure 3. 7 HD purification

(A) Ni-NTA purification. FT: Flow through that did not bind to the beads. E1-3: Elution fractions of HD from the Ni-NTA beads using 500 mM imidazole. (B) The gel filtration A280 spectrum of HD after Ni-NTA column purification. (C) After gel filtration, the fractions were analysed on 12% SDS-PAGE. S: the sample for gel filtration, A1-3: the fractions collected from gel filtration column, A5/7: the concentrated fraction from A5 to A7, M: Protein markers.

Chapter Three: The structure of PHE1

3.1.4 Transient expression and purification of PHE1 in HEK293T cells

The PHE1 containing its original signal peptide was also constructed into the pcDNA3 vector. After transfection to HEK293T cells, the cultured medium was collected after ten days and cleared by centrifugation. The cultured medium was purified by Ni-NTA column followed by a gel filtration column (Superdex-75 10/30) on FPLC. The PHE1

Figure 3. 8 Purification of PHE1

(A) The gel filtration analysis of PHE1 expressed in 293T after Ni-NTA purification. (B) The fractions (A1 to A4) were checked by 12% SDS-PAGE. S: the sample for gel filtration, A1-4: gel filtration fractions; M: Protein markers.

Chapter Three: The structure of PHE1

peak was found in the fractions A3 and A4 at ~12.0 ml after sample injection (Fig.3.8A) which suggested that PHE1 is monomer in the solution. Moreover, the PHE1 in fraction A4 is pure as shown by SDS-PAGE results (Fig.3.8B).

3.1.5 Creation of the PHE1-HEK293 stable cell line

Based on transient expression experimental data, PHE1 was shown to be more promising, and was therefore chosen for further studies. In order to obtain enough amount of protein for crystal structure study, HEK293 transfectants were screened for high expressors. Stable clones were created by dilution, expression levels of PHE1 were determined using anti-His-tag Western blot of the supernatant (Fig.3.9). The clone with the highest secretion level (clone 4) was selected to produce the PHE1 recombinant protein.

Figure 3. 9 Clones of HEK293 stable cells with the different expression level

Clones of HEK293 transfectants with the different expression level were selected by anti-His-tag Western blot. Lanes 1-12 represented different clones, + was the transient transfection to HEK293T and – represented HEK293 cell cultured media without transfection.

3.1.6 The PHE1 expression and purification

The recombinant proteins (PHE1) recovered from the cultured media were shown to express conformational epitopes of the mAbs H52, 7E4, KIM202, MEM148 (Tan et al.,

Chapter Three: The structure of PHE1

2001a; Tang et al., 2005; Tng et al., 2004) and KIM89 (Foo, 2006) indicating that PHE1 was folded into the native conformations (Fig.3.10).

Figure 3. 10 Immunoprecipitation of PHE1

(A) Schematic representation of the domain organization in integrin β subunit. It should be noted that the epitopes of 7E4, KIM89, KIM202, H52 and MEM148, have been mapped to the hybrid domain of the β_2 integrin subunit, that of MEM48 has been mapped to I-EGF2 and I-EGF3 (Lu et al., 2001a). (B) Immunoprecipitation of PHE1 with a panel of specific monoclonal antibodies. MHM24 binding to the integrin α L subunit, and MEM48 were used as negative controls.

The cultured media was harvested once a week, and cell debris was removed by centrifugation (12,000g 15 min). PHE1 were purified from the cultured supernatant by Ni-NTA affinity chromatography. Fractions containing PHE1 were loaded to a gel-filtration column (Sephacryl S-100HR 16/60, Amersham Biosciences) on FPLC. The high purity of PHE1 was suitable for crystal screen or other assays (Fig.3.11). Generally, 2 mg of pure PHE1 can be obtained from 1 liter cultured media. It is concentrated to 100 μ l with a concentration of 20 mg/ml.

Chapter Three: The structure of PHE1

Figure 3. 11 PHE1 purification

(A) Ni-NTA purification. FT: Flow through that did not bind to the Ni-NTA resin. W1-2: Wash fractions using 15 mM imidazole. E1-4: Elution fractions of PHE1 from the Ni-NTA resin using 500 mM imidazole. (B) The gel filtration elution profile of PHE1 on Sephacryl S-100HR column. (C) The gel-filtration fractions of PHE1. S: the sample for gel filtration; A1-8: the gel filtration collected fractions (D) PHE1: 1 μ l of pure and concentrated PHE1 (total 100 μ l) before crystal screening. M: Protein markers.

Chapter Three: The structure of PHE1

3.2 Deglycosylation and Mass spectrometry

Mass spectrometry analysis of PHE1 expressed in HEK293 cells revealed a single species with a molecular mass of 29 KDa, a value 4 KDa higher than its calculated molecular mass using Vector NTI software (Fig.3.12A). After treatment with Peptide N-Glycosidase F (PNGase F), the molecular mass was reduced to the theoretical protein molecular weight of 25 KDa resolved on 10% SDS-PAGE, suggesting that the additional mass was due to glycosylation modification (Fig.3.12B). The carbohydrate on PHE1 expressed in this way is endoglycosidase H_f (Endo H_f) resistant, as is expected for fully glycosylated proteins (Song et al., 2005).

Figure 3. 12 The glycosylation of PHE1

(A) The single peak in mass spectrometry analysis indicated that the protein molecular weight is 29 KDa. It was 4 KDa bigger than the theoretical molecular weight. (B) Proteins digested by PNGase F were analyzed by SDS-PAGE (10%). B and A were the PHE1 proteins before and after PNGase F digestion; M: Protein markers.

Chapter Three: The structure of PHE1

3.3 PHE1 crystallization and data collection

With the pure concentrated PHE1 protein (20 mg/ml) in 20 mM Tris buffer pH 8.0, the initial screening for crystal formation was performed by vapor diffusion using the hanging drop method at 18°C. Many overlapping thin 2-D plate crystals appeared after 2 days in the buffer of 0.1 M sodium cacodylate, pH 6.5, containing 0.2 M calcium acetate, 18% PEG8000 (corresponding to conditions #46 of Sigma Screen 1) (Fig.3.13A). Conditions were optimized including protein concentrations, precipitant concentrations and pH, and the way of seeding (Stura et al., 1991). After seeding in the same buffer, the crystals grew to dimensions of $\sim 0.05 \times 0.3 \times 0.3 \text{ mm}^3$ compared to the original (Fig. 3.13B).

Figure 3. 13 PHE1 crystal images

A, before optimization; B, after optimization.

Chapter Three: The structure of PHE1

Table 3. 1 PHE1 Data Collection and refinement statistics

PHE1	Native	K ₂ Pt Cl ₄
Soaking concentration/time	NA*	1mM; 12 hrs
Wavelength (Å)	0.9760	1.5418
Cell parameters (Å), P2 ₁	<i>a</i> =58.6 <i>b</i> =31.8 <i>c</i> =74.9 β=91.2°	<i>a</i> =58.6 <i>b</i> =31.4 <i>c</i> =75.5 β=91.9°
Resolution range (Å)	30-1.80	19.6-2.80
No. of observed reflections	182,914	26,551
No. of unique reflections ^a	25,996 (3,671)	7,034 (702)
Completeness ^a	99.5% (97.0%)	99.8% (98.4%)
Multiplicity ^a	7.0 (6.1)	3.8 (3.8)
Rmerge (%) [¶]	0.068 (0.478)	0.088 (0.289)
R _{iso} ^b	NA	0.300
I/σ(I)	19.2 (3.2)	7.8 (3.0)
Solvent content (%)	55%	55%
No of sites R _{cullis} centrics, acentrics) ^c	NA	2, 0.70, 0.75
Phasing power ^d	NA	1.33, 1.47
Figure of merit ^e 20-2.8 Å	NA	0.38
Rmsd bond lengths(Å)	0.007	NA
Rmsd bond angles(°)	1.269	NA
Rwork [¶] (%)	20.6	NA
Rfree [#] (%)	23.2	NA
Ramachandran plot (%)	90.6/9.4/0.0/0.0	NA

Chapter Three: The structure of PHE1

* NA: Not applicable.

^aThe numbers in parentheses refers to the last (highest) resolution shell.

[¶] $R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_{h,i} I_{hi}$, where I_{hi} is the i th observation of the reflection h , while $\langle I_h \rangle$ is its mean intensity; ^b $R_{iso} = \sum_h ||F^{PH} - F^P| / \sum_h |F^P|$

^c $R_{cullis} = \sum_h ||F^{PH} - F^P| - |F_{H(calc)}|| / \sum_h |F^{PH} - F^P|$ where $F_{H(calc)}$ is the calculated heavy atom structure factor for acentric and centric reflections respectively

^dPhasing power is the r.m.s heavy atom structure factor divided by the r. m. s. lack of closure.

^eMean value of Figure of merit before density modification and phase combination

^{¶¶} $R_{work} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$.

[#] Rfree was calculated with 5% of reflections excluded from the whole refinement procedure.

Chapter Three: The structure of PHE1

Before data collection, crystals were transferred to a cryoprotecting solution containing in the mother liquor supplemented with 25% (v/v) glycerol. Crystals diffracted to 1.8 Å on the ID-14-4 beamline at ESRF, Grenoble, France. Data from a potassium tetrachloroplatinate (K_2PtCl_4) derivative were collected on the Raxis IV Image Plate detector using CuK radiation from a Rigaku Micromax-007 rotating anode operating at 20 mA and 40 kV. Data were processed with the programs MOSFLM (Leslie, 2006) and SCALA (CCP4, 1997). Crystal parameters and data collection statistics are summarized in Table 3.1. The PHE1 crystals contained one monomer per asymmetric unit, with solvent contents of 53.9% and V_m of 2.7, respectively.

3.4 PHE1 structure solution and refinement

Attempts to solve the structure of PHE1 by molecular replacement using the hybrid domain of $\beta 3$ integrin with all the residues mutated to Ala residues as the model were unsuccessful. The PHE1 structure was determined by the single isomorphous replacement method using single isomorphous replacement combined with anomalous scattering (SIRAS) from one K_2PtCl_4 derivative. Two Pt sites were identified using program SOLVE v2.05 (Terwilliger et al., 1999) and SIRAS phases were calculated in the two possible hands, with program MLPHARE from the CCP4 suite. A clearer solvent/protein contrast was visible in the electron density map with the negative hand (to 2.8 Å). After solvent flattening using the program density modification (DM), a partial model was built and calculated phases were recombined with experimental SIRAS phases using program SIGMAA. The initial R_{free} and R were 53.68% and 51.01%, respectively, for the reflections in the 8-2.5 Å resolution range. Approximately ten cycles of model

Chapter Three: The structure of PHE1

building and minimization with Crystallography & NMR System (CNS) and phase recombination yielded a partial model for PHE1 with $R_{\text{free}}=37.8\%$ and $R=42.02\%$, respectively. Then the availability of a higher resolution native data diffracted to $1.8 \text{ \AA}$ allowed the automatic rebuilding of the model with ARP-WARP (Perrakis et al., 1997). The geometries of all the final models were checked with PROCHECK (Laskowski et al., 1993) and all parameters were within acceptable ranges. A summary of refinement statistics is shown in Table 3.1. The coordinates and structure factors have been deposited in the Protein Data Bank with the accessible code of 1YUK.

3.5 The structure of PHE1

3.5.1 Quality of the model and overall structure

Overall, the chain traces of PHE1 are unambiguous with clear electron density including the carbohydrates observed on both of the two potential N-glycosylation sites, corresponding to Asn28 in the PSI domain and Asn94 in the hybrid domain, and two histidine residues in the C-terminal hexa-histidine tag. Two segments are poorly defined: the four residues (Al100a-L340ys-Leu-Ser) that link the split fragments of the hybrid domain could not be traced in the electron density maps. SDS-PAGE analysis of the dissolved protein crystals ruled out proteolytic cleavage at this position, thus these residues are presumably mobile. Likewise, five residues in the loop (Asp68-His-Glu-Gly-Gly72) are not visible in the electron density map and were therefore not included in the final model. (Fig.3.14).

Chapter Three: The structure of PHE1

Figure 3. 14 The structure of PHE1

PHE1 structure with its domains labeled. The nine disulfide bonds in *orange*, the two N-linked carbohydrates and their corresponding Asn residues are represented as sticks. Their oxygen and nitrogen atoms are shown in *red* and *blue*, respectively. Disordered region is indicated in dashed line. The hexa-histidine tag is at the C-terminal in *brown*.

Chapter Three: The structure of PHE1

Chapter Three: The structure of PHE1

Figure 3. 15 Sequence alignment of human integrin β subunits

Secondary structure elements are marked *above* the alignment in *green* for the PSI domain, *yellow* for the hybrid domain, *red* for the I-EGF1 domain. The structural elements for the integrin $\beta 3$ PSI and hybrid domains are marked in *blue*. The conserved residues contributing to the contacts between the PSI and hybrid domains are highlighted in *light blue*. All conserved cysteines are highlighted in *yellow* and nine disulfide pairs are marked with *arrows*. The two *N*-glycosylation sites are indicated by *red inverted triangles*. Amino acid sequence GenBankTM accession numbers, labeled as *ITB* (integrin β), are as follows: 1, 124963; 2, 124966; 3, 124968; 4, 13638154; 5, 124970; 6, 13432176; 7, 124973; 8, 4504779 (obtained from GenBankTM).

The PSI domain is located centrally in the structure and makes extensive interactions with both the hybrid and I-EGF1 domains. The whole structure is further stabilized by nine disulfide bonds. A superposition of the PSI/hybrid tandem with the equivalent domains in the $\alpha\text{IIb}\beta 3$ integrin structure (PDB, 1TYE) (Xiao et al., 2004) gives an r.m.s deviation of 2.2 Å for 169 equivalent C α atoms. This illustrates that the relative orientation of the hybrid and PSI domains in the $\beta 2$ and $\beta 3$ integrins is well preserved.

3.5.2 The structure of the hybrid domain

The N- and C-terminal sequences flanking the β I domain of the integrin $\beta 2$ cluster in space (Fig.3.16B) form the hybrid domain which adopts a β -sandwich fold. In the PHE1 construct, the β I domain is absent and the 4 residues connecting the β B and β C strands of the hybrid domain are not resolved in the electron density map. Interestingly, in spite of the absence of the whole β I domain, the hybrid domain superimposes well with the corresponding domain of the $\beta 3$ integrin with an r.m.s deviation of 1.6 Å for 115 equivalent C α atoms, this suggests that the β I domain does not influence the conformation of the hybrid domain. Actually, the invisible four-residue linker (Ala100-Lys340LeuSer) in PHE1 structure forms the loop between strands B and C,

Figure 3. 16 The structure of the hybrid domain

(A) Topological graph of the hybrid domain. The β I domain is inserted in the loop in dashed line between B and C strands (B) A ribbon diagram of the hybrid domain structure. The domain is constructed from two discontinuous segments of the primary sequence (residues 59-100; 340-423), which assemble into an I-set Ig-like domain. Compared with I-set Ig domains, the hybrid domain has an extra (X) N-terminal strand and two short strands within the DE loop. Two disulfide bonds (Cys364-Cys378 and Cys389-Cys423) stabilized the domain fold are shown in *orange*. The location of the β I domain between B and C strands is shown as dotted line. The epitope of mAb 7E4, Val385 (Tng et al., 2004) and the epitope of mAb MEM148, Pro374 (Tang et al., 2005), resolved in the structure are represented as sticks. (C) Superposition of the hybrid domain of integrin β 2 in (*yellow*) and β 3 (*blue*) (PDB, 1TY6).

Chapter Three: The structure of PHE1

which is verified by the subsequent solved structure PHE2 (PDB, 2P26). Two disulfide bonds in the hybrid domain are conserved and superposable to that of $\beta 3$ in the $\alpha V\beta 3$ ectodomain structure. This suggests that the β I domain does not influence the conformation of the hybrid domain and that other domains could be inserted at the same position, opening the possibility of creating chimeric receptor molecules. The loop connecting the βX and βA strands is significantly shorter in integrin $\beta 2$ compared to the corresponding loops in $\beta 1$, $\beta 3$, $\beta 5$, $\beta 6$ and $\beta 7$ integrins (Fig.3.15; Fig.3.16B) and is presumably flexible. In the $\beta 3$ integrin structure, the same loop is longer and is stabilized through contacts with residues from the β I domain (Xiao et al., 2004; Xiong et al., 2001).

3.5.3 The structure of the PSI domain

The PSI domains of integrin $\beta 2$ (Protein Data Bank accessible code: 1YUK) and $\beta 3$ (PDB, 1TY3) have similar structures (Shi et al., 2005; Xiao et al., 2004) with a residual r.m.s deviation between the two domains of 1.8 Å for 54 equivalent C α atoms after superposition. All four disulfide bonds are conserved and superposable including the one formed by the second cysteine residue (Cys11), and the eighth cysteine residue (Cys425), separated by the hybrid domain and β I domain in the linear sequence. The two cysteines corresponded to Cys13 and Cys435 in the $\beta 3$ integrin (Fig.3.8; Fig.3.10). This long-range disulfide bridge also superimposes well with the highly conserved intra-domain disulfide that links the second cysteine residue to the last cysteine residue in the PSI domain of the MET (PDB, 1SHY) (Stamos et al., 2004) and semaphorins 4D (PDB, 1OLZ) (Love et al., 2003). These two cysteines align perfectly in all PSI domains, which suggests that Cys425 in $\beta 2$ integrin (Cys435 in $\beta 3$ integrin) is the eighth cysteine of the PSI domain,

Chapter Three: The structure of PHE1

and an integral part of the integrin PSI domain based on the structural conservation. Therefore, the domain arrangement of β subunit is complex: The β I domain is inserted in the hybrid domain, which is in turn inserted in the PSI domain (Fig. 3.1). Differences however exist in the PSI domain between integrin $\beta 2$ and $\beta 3$ subunits: The N-terminal $\alpha 1$ helix in $\beta 3$ integrin is missing in $\beta 2$ integrin, an additional 3/10 helix between $\beta 1$ and $\beta 2$ strands, and a $\alpha 3a$ helix between $\alpha 3$ helix and $\beta 3$ strand are present in $\beta 2$ PSI domain (Fig. 3.17).

Figure 3. 17 Stereo view of the PSI domain superposition.

Green, integrin $\beta 2$; *blue*, integrin $\beta 3$. The four conserved disulfide bridges resolved in the structure are represented as sticks in *orange*. The amino and carboxyl termini of each domain are indicated. The N' and C' refer to termini for residues 424-426.

3.5.4 The structure of the I-EGF1 domain

The I-EGF1 domain of the integrin β subunits was not resolved in previous structural studies and our data provide the first view at atomic resolution of this domain. Unlike the integrin EGF domains 2, 3 and 4, as well as those of the ten β -integrin EGF-like repeat domains (TIED) proteins (Berg et al., 1999; Takagi et al., 2001a) which have eight

Chapter Three: The structure of PHE1

cysteine residues, the I-EGF1 structure of $\beta 2$ subunit has only a total of six cysteines engaged in three disulfide bonds (Fig.3.15; Fig.3.18).

Figure 3. 18 The structure of the I-EGF1 domain

(A) Topological graph of the integrin $\beta 2$ I-EGF1. The three disulfide bonds are indicated. It should be noted that the cysteines are numbered according to the typical integrin-EGF domain with eight cysteines; in this case, C2 and C4 are missing. The Arg432 and Leu443 residues which replace C2 and C4, respectively, according to the published sequence alignment (Beglova et al., 2002; Takagi et al., 2001a) are in a yellow background. (B) Superposition of the I-EGF1 domain (*red*, PDB, 1YUK,) with the I-EGF3 from integrin $\beta 3$ (*blue*, PDB, 1U8C) and the EGF domain of P-selectin (*turquoise*, PDB, 1FSB). The disulfide bonds C3-C6 and C7-C8 are well conserved constraining the fold of the C-terminal part of the EGF module. The disulfide bond (C2-C4) is absent in the I-EGF1 domain. Note that the first disulfide bond (C1-C4) is missing in the P-selectin. Large conformational variations are observed between the N-terminal ends of the three structures.

Chapter Three: The structure of PHE1

In the NMR structure of the integrin $\beta 2$ I-EGF3 domain (Beglova et al., 2002), eight cysteines were shown to arrange in the C1-C5, C2-C4, C3-C6 and C7-C8 pattern. Amino-acid sequences alignment shows that I-EGF1 belongs to this group of EGF domains but with the C2-C4 disulfide pair missing, and not to the laminin EGF domains which have the C1-C3, C2-C4, C5-C6 and C7-C8 disulfide arrangement (Takagi et al., 2001a). A superposition with the integrin $\beta 2$ I-EGF3 NMR structure reveals that the absence of the second and fourth cysteine residues are accompanied by a large movement of the polypeptide chain (including the basic patch formed by Arg-428, Arg-432 and Arg-434) that projects away from the central β -sheet formed by the $\beta 1$ and $\beta 2$ strands (Fig.3.18). This is presumably due to the release of the structural constraint brought by the C2-C4 disulfide bond in the I-EGF3 domain, which is absent in the I-EGF1 domain. Interestingly, the closest structural homologue of I-EGF1 is the EGF domain of P-selectin (PDB, 1FSB) (Freedman et al., 1996) whose polypeptide chain follows a similar path with 33 equivalent residues superimpose with an r.m.s deviation of 1.7 Å (Fig.3.18).

3.5.5 Interactions between domains

The interactions between the PSI and the hybrid domain, and PSI and the I-EGF1 domain are stabilized through the formation of extensive interfaces of 820 Å² and 1100 Å², respectively. The interface between the PSI and hybrid domains in the integrin $\alpha \text{IIb} \beta 3$ is similar with a comparable buried surface area of 860 Å² and the involvement of a highly invariant Arg residue in multiple contacts with main chain atoms of the PSI domain in both cases. The side chain of Arg86 (Arg93 in integrin $\beta 3$) is deeply buried in the PSI/hybrid interface forming six hydrogen bonds with main chain atoms from Ile15,

Chapter Three: The structure of PHE1

Gly18, Cys21, Pro59 and one with the side chain atom of Asp58, for comparison, the corresponding contacts involved with Arg93 in $\beta 3$ integrin are also shown (Fig.3.19A). Interactions are listed in Table 3.2. Together with the four disulfide bonds in the PSI domain and one (Cys398-Cys423) in the hybrid domain, the hybrid and PSI domains are held in a rigid orientation with respect to each other in both the $\beta 2$ and $\beta 3$ integrins. Given the strict conservation of the Arg86 residue, the relative orientation of PSI and hybrid domains is likely to be conserved in all integrin β subunits (Fig.3.15). The PSI domain and I-EGF1 domains are also found to interact extensively (Fig.3.19B), and their polar interactions are given in Table 3.2. Hydrogen bonds are formed between the guanido group of Arg39 and the main chain carbonyl oxygen of Gly446 and between the side chain of Trp23 in the PSI domain and the main chain carbonyl group of Cys445. Interestingly, Trp23, Arg39, Cys445 and Gly446 four residues are highly invariant among the integrin β subunit (Fig.3.15), suggesting that these interactions may be stable in the interface between the PSI domain and the I-EGF1 domain.

Table 3. 2 Hydrogen bond contacts in PHE1 domains

PSI domain	Distance (Å)	Hybrid domain	PSI domain	Distance (Å)	I-EGF1 domain
Pro59 O	2.46	Arg86 NE	Ser9 O	2.90	Arg449 NH2
Gly18 O	3.13	Arg86 NE1	Trp23 NE1	2.90	Cys445 O
Ile15 O	3.43	Arg86 NE1	Glu25 OE1	2.92	Arg428 N
Cys21 O	3.09	Arg86 NE1	Asn28 OD1	3.09	Asn458 ND2
Asp58 O	2.47	Arg86 NE2	Thr30 O	3.10	Lys457 NZ
Pro59 O	3.12	Arg86 NE2	Gly31O	3.06	Lys457 NZ
Cys21 O	3.14	Arg86 NE2	Arg39 NH2	3.21	Gly446 O

Chapter Three: The structure of PHE1

Figure 3. 19 Hydrogen bond interactions between PHE1 domains

Hydrogen bond interactions are represented by broken lines, and residues taking part in the interactions are labeled. Contacted residues are shown as sticks. (A) Interactions at the interface between the hybrid domain (yellow) and the PSI domain (green) in $\beta 2$ subunit (left panel) and in $\beta 3$ subunit (right panel), respectively. The invariant Arg86 residue (Arg93 in $\beta 3$ subunit) is located centrally making extensive polar contacts indicated by broken lines. (B) Interactions between the PSI domain (green) and I-EGF1 domain (red) in $\beta 2$ subunit.

Chapter Three: The structure of PHE1

3.5.6 Oligosaccharide chains of PHE1

Two N-glycosylation sites on Asn28 and Asn94 residues in Asn-X-Ser/Thr motifs were predicted using Vector NTI software. Electron density for three sugar residues (N-GlcNAc, GlcNAc, Fuc) bound to Asn28 and two sugar residues (N-GlcNAc, GlcNAc) bound to Asn94 are clearly observed at 1.0 σ contour level (Fig.3.20), although the indication of additional branched sugar residues can be visible at 0.5 σ contour level. The sugar chain in GP-II protein was shown to participate in the crystal lattice arrangement (Choi et al., 1999) while no such involvement of the carbohydrate in intermolecular bond formation was observed in PHE1.

Figure 3. 20 The N-linked oligosaccharide chains of PHE1

The electron density maps (contoured at 1.0 σ , *gray*) surrounding the oligosaccharide chain linked to Asn28 (A) and Asn94 (B) are shown, respectively. The structures of the oligosaccharide chains are depicted as sticks.

Chapter Three: The structure of PHE1

3.6 Discussion

3.6.1 Constructions and expressions

In $\alpha V\beta 3$ ectodomain crystal structure at the resolution of 3.1 Å, the α and β subunits combine together to form a globular ligand-binding headpiece connected to two rod-like legs that are unexpectedly bent at the 'Genus'. The Genu of α subunit, located at the junction between the Thigh and Calf1 domains, together with a Ca^{2+} divalent cation, is well defined. However, the Genu of β subunit is not defined because of the poor electron density in the PSI, and the I-EGF1-2 domains suggesting that these domains are flexible and severely bent in this region (Xiong et al., 2001). In order to investigate the unsolved region of the β subunit, we studied the structure of the PHE1 fragment. However, the domain arrangement in integrin β subunit is not linear, therefore, the hybrid domain was also included into the construction and the β I domain is removed because correct folding of the β I domain needs its association with the β -propeller domain in the α subunit and *vice versa* (Huang et al., 1997a; Huang et al., 1997b). Although one report claimed that the β I domain can be expressed in *E. coli* with a native structure, it was not supported by the evidence such as the binding of the mAbs whose conformational epitopes located in the β I domain (Baneres et al., 2000). Indeed, the experimental results were not reproduced by another group (Takagi et al., 2002a). In our hands, the expression level of the secreted recombinant protein (PHE1 plus the β I domain) is significantly lower than that of PHE1 in HEK293 cells and we failed to obtain sufficient amount for crystal structure study (unpublished results). This low expression of PHE1 plus the β I domain may be caused by the incorrect folding of the β I domain. Previous experimental data also indicated that all the extracellular domains except for the β I domain in integrin β subunit

Chapter Three: The structure of PHE1

can be expressed with retention of immunological activities (Tan et al., 2001a). To circumvent the complexity introduced by the β I domain, a β 2 subunit with the β I domain deleted (β 2 Δ I) was constructed.

The hybrid domain and PHE1 were chosen for the study and were initially expressed in *E. coli* expression system, a convenient and economical means of producing recombinant proteins. Both of them were insoluble, which indicated that these two constructs cannot be folded correctly. Chaperones and the post-translational modifications may be required, and proper disulfide bonds may not be formed in *E. coli* because of its reduced extracellular environment. Although the origami host strain, which would greatly enhance the disulfide bond formation, together with the maltose-binding protein (Riggs et al., 1990) or thioredoxin (LaVallie et al., 1993) fusion tag, has been used to express the hybrid domain or PHE1 under various conditions, the proteins were only obtained in aggregated forms. While considerable efforts were taken to avoid inclusion body formation or to obtain monomer from the aggregates, sufficient amount of the pure target proteins in their soluble forms suitable for crystallography was not obtained.

With the β 2 original signal sequence, the recombinant proteins (PHE1 and hybrid domain) were expressed in a mammalian system as the secreted proteins. Significant amount of pure PHE1 was obtained after purification for crystallography studies.

3.6.2 PHE1 structure

The PHE1 structure suggests that the PSI and hybrid domains represent a rather rigid unit that may act as a lever to orientate the β I domain with respect to the lower leg. This

Chapter Three: The structure of PHE1

investigation is consistent with those obtained from the headpiece of integrin $\alpha\text{IIb}\beta 3$ (Xiao et al., 2004). We should emphasize that the structure of the PHE1 is only a fragment of the integrin $\beta 2$ subunit, the presence of other domains in the $\beta 2$ subunit, and the α subunit, may alter the interdomain interaction (see Chapter Four).

Although the disulfide linkages are conserved in the integrin structures determined so far (Beglova et al., 2002; Xiao et al., 2004; Xiong et al., 2001) several pairs of cysteine residues are found in close vicinity (*e.g.* Cys11-Cys425 is spatially close to Cys427-Cys445, and Cys3-Cys21 is close to Cys14-Cys40; Fig 3.14). Given small adjustments in the three-dimensional structure, sulfhydryl groups are at the right distance to form alternate disulfide bonds. Thus, we cannot completely rule out the possibility of an alternative disulfide arrangement that would introduce subtle alterations in the structure, possibly acting as relays to transmit signals.

Previous long-range disulfide bond Cys3-Cys425 in $\beta 2$ integrin, and Cys5-Cys435 in $\beta 3$ integrin had been proposed for the integrin β subunit (Calvete et al., 1991). Based on the PSI domain in other structures (Love et al., 2003; Stamos et al., 2004), the PHE1 structure, together with those of the $\beta 3$ integrin (Xiao et al., 2004; Xiong et al., 2004), firmly established that the long disulfide bond is between Cys11-Cys425 in $\beta 2$ integrin, and corresponding to Cys13-Cys435 in $\beta 3$ integrin. In the linear sequence, the Cys425 of $\beta 2$ is located C-terminal of the hybrid domain, such that the hybrid domain is inserted into the last loop of the PSI domain. This unexpected architecture through which the PSI and hybrid domains are connected, and the cysteine pairings in such close proximity, may

Chapter Three: The structure of PHE1

explained the incorrect predications of cysteine pairing in integrin β subunits in previous publication (Calvete et al., 1991).

Having defined the relative orientations of the hybrid, PSI, and I-EGF1 domains, and given that the PSI and I-EGF1 domains are located side by side, what can we infer about the probable location of the I-EGF2 domain? We performed a superposition of the PSI/hybrid domains of the PHE1 structure with the corresponding domains resolved in the bent conformation of the integrin $\alpha V\beta 3$ structure (Xiong et al., 2004) (Fig.3.21). Based on the resulting locations of the C-terminal end of the I-EGF1 domain and the N-terminal end of the I-EGF3, a severe bent is likely to exist between the I-EGF1 domain and I-EGF2 domain (Fig. 3.21).

Thus, we have identified one critical region for the structural change associated to the regulation for the integrin-mediated cell adhesion. This hypothesis can be further investigated by analyzing the structures of the integrin fragments including the I-EGF2 domain.

Figure 3. 21 Model for the integrin α V β 3 in its bent state

The PSI and hybrid domains from the PHE1 structure (shown as *ribbons* in *green* and *yellow*, respectively) were superposed on the corresponding domains present in the crystal structure of the extracellular portion of α V β 3 (Protein Data Bank code 1JV2). The I-EGF2 domain (labeled in *italics*) was modeled according to the I-EGF3 structure (Protein Data Bank code 1L3Y) using Quanta (Accelrys Inc., San Diego, CA) and displayed with program O (Jones et al., 1991). The C-terminal ends of the α and β subunits are labeled as α C and β C, respectively. The putative hinge region located between the I-EGF1 and I-EGF2 domains is indicated by an *arrow*. β -TD, β -tail domain; HD, hybrid domain.

Chapter Four: The structures of PHE2 and PHE3

Chapter Four: The structure of PHE2 and PHE3

Using X-ray crystallography, we have determined the crystal structure of I-EGF1 as part of a fragment of the $\beta 2$ subunit, consisting of the PSI, hybrid, and I-EGF1 domains which was named as PHE1. Superposition of the PHE1 structure onto the bent $\alpha V\beta 3$ ectodomain crystallographic structure (Xiong et al., 2004) suggested that the bend in the β subunit may lie between the I-EGF1 and I-EGF2 domains, a conclusion also supported by electron microscopic studies of the $\alpha L\beta 2$ and $\alpha X\beta 2$ integrins (Nishida et al., 2006). However a clear picture of the modules making up the “Genu” (Fig.3.1) in the β subunit has not emerged yet and the atomic basis for its flexibility is still largely unresolved. In order to further understand the molecular mechanism of integrin activation, it is crucial to obtain the crystal structure of the I-EGF2 domain, and to determine how it is connected the neighboring integrin modules, both in the bent and extended states. Accordingly, we expressed the PHE2 and PHE3 fragments, both from the $\beta 2$ integrin subunit, the construction of PHE2 and PHE3 were similar to that reported for PHE1 (Shi et al., 2005), except that the I-EGF2 domain (residues Cys461-Glu513) is also included in PHE2, and both I-EGF2 and I-EGF3 domains (residues Cys461-Gln552) in PHE3 (Fig.4.1).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 1 Illustrations of recombinant protein constructs

Both PHE2 and PHE3 are devoid of the β I domain. The sequence numbering scheme is indicated. The same color codes for the domains are conserved throughout: *yellow*, hybrid domain; *green*, PSI domain; *red*, I-EGF1 domain; *purple*, I-EGF2; *pink*, I-EGF3; The β -TD, β -Tail Domain; TM, Transmembrane helix; Cyt, Cytoplasmic domain are not colored. (His)₆, hexa-histidine tag.

4.1 Expression and purification

Initially, PHE2 was expressed and purified as did PHE1. Briefly, after constructed in pcDNA3 vector, PHE2 expression was verified through a transient expression in HEK293T cells (Fig.4.2). Then, stable clones were established in HEK293 cells and the PHE2 protein was purified. However, only tiny crystals were obtained which were not suitable for x-ray diffraction. As there is an additional glycosylation site in the I-EGF2 domain, an Asn to Gln substitution was performed and PHE2_{N479Q} was also expressed from the HEK293 cells and purified, but no crystal was obtained. In order to minimize the effects of the carbohydrate on the crystallization, HEK293-GnTI cells were chosen to

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 2 PHE2 and PHE3 expressed in different cell lines were detected by anti-His-tag Western blot.

Figure 4. 3 Immunoprecipitation of PHE2 and PHE3 with a panel of specific monoclonal antibodies

It should be noted that the epitopes of KIM89, MEM148, KIM202, 7E4, and H52 have been mapped to integrin β 2 hybrid domain, that of MEM48 has been mapped to I-EGF2 and I-EGF3 domains (Lu et al., 2001a), that of KIM127 has been mapped to I-EGF2 (Lu et al., 2001a) and MHM24 binding to the integrin α L subunit (Tan et al., 2001a) was used as a negative control.

Chapter Four: The structures of PHE2 and PHE3

express PHE2 and PHE3, which cells lack N-acetylglucosaminyltransferase I (GnTI) activity, and consequently unable to synthesize the complex N-glycans (Reeves et al., 2002). Thus, the N-linked carbohydrates are of the Man5 (Man₅GlcNAc₂) type (Crispin et al., 2006), and the molecular weight of the N-linked carbohydrates is lower than that when expressed in HEK293 cells. Indeed, PHE2 and PHE3, when expressed from HEK293-GnTI⁻ cells, have lower MW than those expressed in HEK293 and HEK293T cells (Fig.4.2). PHE2 and PHE3 were inserted into the vector pIRES2-EGFP with a C-terminal hexa-histidine tag to facilitate downstream recombinant proteins purification. After transfection, the HEK293-GnTI⁻ cells were sorted by FACS Aria flow cytometry-based cell sorter (Mancia et al., 2004; Patel et al., 2004) according to the GFP expression. About 20 GFP⁺ cells were spread into a 10 cm dish, single clones were selected and clones with the highest expression level for PHE2 and PHE3, detected by the anti-his-tag Western bolt, were chosen to produce the recombinant proteins.

The recombinant proteins (PHE2 and PHE3) recovered from the cultured media were shown to express conformational epitopes of the mAbs H52, 7E4, KIM202, MEM148 (Tan et al., 2001a; Tang et al., 2005; Tng et al., 2004) and KIM89 (Foo, 2006) indicating that they are in their native conformations (Fig.4.3). In addition, the PHE3 protein is also recognized by the mAbs MEM48 and KIM127. This agrees with the previous findings that the expression of the epitopes of these two antibodies requires both I-EGF2 and I-EGF3 domains (Lu et al., 2001a; Takagi et al., 2001a; Tan et al., 2001b).

Chapter Four: The structures of PHE2 and PHE3

The cultured media was harvested after ten days, and the cell debris was cleared by centrifugation. PHE2 and PHE3 were purified from the culture supernatant by Ni-NTA affinity chromatography (Fig.4.4A; Fig.4.5A) followed by a gel-filtration column (Superdex 75, Amersham Biosciences). In addition, PHE3 was further purified by a Mono Q anion exchange column before the gel filtration to remove the contaminated (Fig. 4.5B,C). The expressed PHE2 and PHE3 proteins secreted from HEK293-GnTI cells were sensitive to Endo H_f (Fig. 4.6), as observed in other proteins e.g. sRPTP (Crispin et al., 2006) and sCD48 (Chang et al., 2007) in the same expression system. Purified PHE2 and PHE3 were concentrated to approximately 1 mg/ml and treated with Endo H_f (New England Biolabs) (4,000 units/ml) overnight at room temperature. After the deglycosylation, they were further purified by gel filtration and both were shown to elute later than the uncleaved proteins (Fig.4.4B; Fig.4.5D), in agreement with having the carbohydrates removed. Both proteins were of high purity (Fig.4.4D; Fig.4.5F), and concentrated to approximately 15 mg/ml in 20 mM Tris-HCl, pH 7.2 for preparation of crystals and other studies.

The molecular masses of PHE2 and PHE3 were determined by matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry (Table 4.1). The Majorities of Endo H_f digested PHE2 and PHE3 are 154 Da and 133 Da smaller than the calculated MW after Endo H_f digestion, which suggested that the Endo H_f digestion is complete. It should be noticed that the MW peaks of the dissolved crystals are broad, and the Masses of PHE2 and PHE3 crystals are 10-257 Da and 511-902 Da bigger than the calculated after Endo H_f digestion, respectively.

Chapter Four: The structures of PHE2 and PHE3

Figure 4.4 PHE2 purification

(A) Ni-NTA purification. E1-4: Elution fractions of PHE2 from the Ni-NTA resin using 500 mM imidazole. (B) The gel filtration elution profiles of PHE2 before (*red* line) and after (*blue* line) Endo H_f digestion on a Superdex75 16/180 column mounted on FPLC. PHE2 eluted at volume of 127 ml and 133 ml before and after Endo H_f digestion, respectively. (C) The gel-filtration even fractions (A2-A10) of PHE2 after Endo H_f digestion. (D) PHE2: 1 μ l of purified PHE2 before crystal screening. M: Protein markers.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 5 PHE3 purification

(A) Ni-NTA purification. E1-5: Elution fractions of PHE3 from the Ni-NTA resin using 500 mM imidazole. (B) The Mono Q elution fractions analyzed on 10% SDS-PAGE. Only PHE3 did not bind to the column and eluted at the flow through (A1) fraction. (C) The elution profile of the Mono Q purification. (D) The gel filtration A₂₈₀ spectrum of PHE3 before (*red* line) and after (*blue* line) Endo H_f digestion on a Superdex75 column. PHE3 eluted at volume of 57 ml and 60 ml before and after Endo H_f digestion, respectively. (E) The gel-filtration fractions of PHE3 after after Endo H_f digestion. (F). PHE3: 1 μ l of purified PHE3 before crystal screening. M: Protein markers.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 6 Analysis of PHE2 and PHE3 on SDS PAGE before and after treatment of Endo H_f

M: Molecular masses of standard protein markers are also displayed. *B*, before Endo H_f treatment; *A*, after Endo H_f treatment.

Table 4. 1 Mass spectrometry analyses of PHE1, PHE2 and PHE3

Protein	Calculated MW (Da)	MS before PNG-F	Calculated MW after Endo H _f (Da)	MS after Endo H _f	Dissolved crystal MS
PHE1	25347	29676	25753	NA*	29217-29902
PHE2	31017	NA	31626	31472 35602/Major	31636-31993
PHE3	35126	NA	35735	36367/Minor	36246-36637

* NA: Not applicable.

Chapter Four: The structures of PHE2 and PHE3

4.2 PHE2 and PHE3 crystallization and data collection

Crystallization was performed by vapor diffusion using the hanging drop method at 18°C. For PHE2, small plate crystals were observed after a week in 0.1 M sodium acetate, pH 4.6, 25% PEG2000MME (corresponding to conditions of #6 in Qiagen PEGs Suite) (Fig.4.7A1). After optimization, single plate crystals grew up to the size of $\sim 0.05 \times 0.3 \times 0.4 \text{ mm}^3$ after a week in 0.2 M magnesium acetate, pH 4.6, 0.1 M sodium acetate and 18% PEG 3350 (Fig.4.7A2). For PHE3, the crystals appeared as plates in 0.2 M ammonium sulfate, 15% PEG4000 (corresponding to conditions of #55 in Qiagen PEGII Suite), but were in aggregated forms (Fig.4.7B1). After optimization, single hexagon plate crystals with dimensions of $\sim 0.1 \times 0.4 \times 0.5 \text{ mm}^3$ (Fig.4.7B2) were observed in a solution containing 0.2 M ammonium sulfate, 15% PEG 4000 and 5% isopropanol. The conditions that gave the best crystal shape and quality were summarized in Table 4.2. The crystals were equilibrated to higher PEG concentration by 7% in every 12 hour. Three rounds of increment were carried out so that the final PEG concentrations were 39% and 36% for the PHE2 and PHE3, respectively (Heras et al., 2005). The PHE2 and PHE3 crystals were cooled to 100 K in a nitrogen gas stream (Oxford Cryosystems). The PHE2 crystals diffracted to 1.75 Å on the Raxis IV Rigaku Micromax 007 with a rotating anode, and the PHE3 crystals diffracted to 2.2 Å on the ID-23-1 beamline at ESRF, Grenoble, France. Data were processed using programs MOSFLM (Leslie, 2006) and SCALA (CCP4, 1997). Crystal parameters and data collection statistics are summarized in Table 4.3. Both crystals contain one monomer per asymmetric unit, with solvent contents of 34% and 48% and V_m of 1.86 and 2.35, respectively.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 7 Crystals images of PHE2 (A), PHE3 (B)

1, before optimization; 2, after optimization.

Table 4. 2 Optimal conditions for PHE2 and PHE3 crystal growth

Proteins	Concentration	Buffer	Salt	Precipitate	Additive
PHE2	15 mg/ml	0.1 M sodium acetate pH4.6	0.2 M magnesium acetate	18% PEG3350	NA
PHE3	15 mg/ml	NA*	0.2 M ammonium sulfate	15% PEG4000	5% 2-propanol

*NA: Not applicable.

Chapter Four: The structures of PHE2 and PHE3

Table 4. 3 Data Collection and refinement statistics for PHE2 and PHE3

Integrin fragment	PHE2	PHE3 ^a	
Space group	P2 ₁	P6 ₅ 22	P6 ₅ 22
Cell parameters (a, b, c) (Å)	58.5 30.8 65.2	52.3 52.3 423.9	52.2 52.2 424.1
(abγ) (°)	90. 94.3 90.	90. 90. 120.	90. 90. 120.
Wavelength (Å)	1.5418	0.976	1.907
X-ray source	Rigaku Micromax-007 rotating anode	ESRF ID23-1	ESRF ID23-1
Resolution (Å)	1.75	2.2	3.0
Number of reflections (total/unique)	144663/21822	177586/18866	207994/7895
Completeness (%)	91.3 (55.5) ^c	99.5 (99.3)	100.0 (100.0)
Multiplicity	6.6 (4.5)	9.4 (8.7)	26.3
I/σ(I)	23.4 (2.0)	22.2 (2.5)	31.0 (6.4)
Rmerge (%) [¶]	6.5 (52.0)	8.0 (60.7)	12.1
Rmsd bond lengths(Å)	0.007	0.009	NA ^b
Rmsd bond angles(°)	1.20	1.22	NA
Rwork ^{¶¶} (%)	20.4 (40.4)	26.1 (30.2)	NA
Rfree [#] (%)	25.4 (59.2)	30.8 (40.5)	NA
Ramachandran plot	86.4/13.6	85.2/14.8	NA
Residues in most favored and additional allowed regions (%)			
Average Temperature Factors (Å ²)			
Main chain	18.0	22.3	NA
Side-chains	19.5	23.4	NA
Water (total number)	33.7 (268)	76.1 (200)	NA

^a The same crystal of PHE3 was used for both data sets which were collected successively at high (left) and low (right) energy. The second PHE3 data set was collected to maximize the anomalous signal from the sulfur atoms. An anomalous difference Fourier map was calculated to confirm the trace of the polypeptide chain (see Fig.4.15). The PHE3 structure was refined using the high energy data set only.

^b N/A: not applicable

^c The numbers in parentheses refers to the last (highest) resolution shell.

[¶] $R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_{h,i} I_{hi}$, where I_{hi} is the i th observation of the reflection h , while $\langle I_h \rangle$ is its mean intensity.

^{¶¶} $Rwork = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$.

[#] Rfree was calculated with 5% of reflections excluded from the whole refinement procedure.

Chapter Four: The structures of PHE2 and PHE3

4.3 PHE2 and PHE3 structure refinement and determination

The PHE2 and PHE3 structures were determined by molecular replacement using the program PHASER (Storoni et al., 2004) and separate modules as search models. The geometries of all the final models were checked with PROCHECK (Laskowski et al., 1993) and all parameters were within acceptable ranges. A summary of the refinement statistics for crystals is shown in Table 4.3. The coordinates and structure factors of PHE2 and PHE3 have been deposited in the Protein Data Bank with accessible code 2P26 and 2P28, respectively.

4.4 The structure of PHE2

4.4.1 Quality of the model and overall structure

The PHE2 structure has approximate overall dimensions of 85 Å x 30 Å x 30 Å. The main chain, including a complete C-terminal hexa-histidine tag, is clearly defined in the PHE2 electron density at the resolution of 1.75Å. The I-EGF2 domain forms an acute angle with the I-EGF1 domain which gives PHE2 an approximate L-shaped conformation (Fig.4.8). As independent confirmations of the correctness of the polypeptide chain traces, we used the anomalous signal from sulfur atoms to confirm the positions of the thirteen disulfide-bridges present in the PHE2 structure (Fig.4.9). In addition to the two N-acetylglucosamines observed on residues Asn28 and Asn94 as in PHE1, the electron density for an N-acetylglucosamine residue on Asn479 in the I-EGF2 domain is also visible.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 8 The structure of the PHE2

The I-EGF2 is roughly perpendicular to the PSI/hybrid tandem, shown as a vertical arm here. Thirteen disulfide bonds, the three N-linked acetylglucosamine residues and their corresponding Asn residues are represented as sticks. Especially, the Cys461-Cys492 disulfide bridge in I-EGF2 is labeled, and the hexa-histidine tag is in *brown*. The first three histidines (His514-His516), Lys506, Ile507 and Tyr508, form two anti-parallel β -strands.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 9 Independent validation of the PHE2 using the anomalous signal from the sulfur atoms

Anomalous Fourier electron density maps, with coefficients D_{ano} (anomalous difference) and phases calculated from the refined models shifted by 90° are displayed at 3σ levels. The anomalous maps were cut using masks covering the entire PHE2 molecule. The same color code for the various integrin $\beta 2$ domains (PSI, hybrid, I-EGF1 and 2) is used. The 13 disulfide bonds that are present in the PHE2 structure coincide with peaks in the anomalous difference Fourier maps. In the anomalous map, individual sulfur atoms are well resolved. The peaks with heights ranging from 3.5 to 4.9σ are labeled with the corresponding disulfide bridge.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 10 Sequence alignment and topology of the I-EGF2 domain

(A) Secondary structure elements are marked *above* the alignment *purple* for the I-EGF2. Eight conserved cysteines are highlighted in *yellow* and the disulfide pairs are marked with *arrows*. The *N*-glycosylation sites are indicated by a *red inverted triangle*. Amino acid sequence GenBankTM accession numbers, labeled as *ITB* (integrin β), are as follows: 1, 124963; 2, 124966; 3, 124968; 4, 13638154; 5, 124970; 6, 13432176; 7, 124973; 8, 4504779. (B) A cartoon diagram of the structure of the I-EGF2 domain in PHE2. The four conserved disulfide bridges, the N-linked N-acetylglucosamine residue and its corresponding residue Asn479, the epitope of KIM127 (Gly504, Leu506, and Tyr508) resolved in the structure are represented as sticks. (C) Topological graph of the integrin β 2 I-EGF2 domain. The four evolutionarily-conserved disulfide bonds are indicated. The N-glycosylation site on Asn479 is indicated by a *gray triangle*.

Chapter Four: The structures of PHE2 and PHE3

4.4.2 The structure of the I-EGF2 domain

The I-EGF2 domain of the integrin β subunits was not resolved in previous structural studies and our data provide the first view at atomic resolution of this domain. There are eight conserved cysteines arranged in the C1-C5, C2-C4, C3-C6, C7-C8 pattern which also appeared in the I-EGF3 domain of known structures (Xiong et al., 2001; Beglova et al., 2002; also see PHE3) and the I-EGF4 domain of the integrin β 3 subunit (Xiong et al., 2001). The region between C1 and C2 in the I-EGF2 domain consists of 9-13 residues (e.g. 13 residues spanning Gln462-Ser474 in the β 2 integrin) which is much longer when compared to four residues in I-EGF3 domains and six residues in I-EGF4 domains of integrin β subunits. The residues from Ser468 to Cys475 fold into a two-turn α -helix (Fig.4.10B). The epitope of KIM127 (Lu et al., 2001a; Robinson et al., 1992) is formed by the residues Gly504, Leu506 and Tyr508 located in the loop between C7 and C8 (see later). In addition, the C-terminal hexa-histidine tag was fully traced and the first three histidines (His514-His516), together with Lys506, Ile507 and Tyr508, form two anti-parallel β -strands (β 3 and β 4) (Fig.4.10B).

4.4.3 The structural differences between PHE1 and PHE2

Each of the three domains found in both PHE2 and PHE1 (Shi et al., 2005), i.e. the PSI, hybrid and I-EGF1 domains, retains similar structures with r.m.s. deviations in atomic positions less than 0.82 Å (Table 4.4). Their relative orientation is essentially preserved between these two fragments, so that the hydrogen bond interaction between the PSI domain and hybrid domain is the same between PHE1 and PHE2 structures (Fig.3.19). In addition to the hydrogen bond interactions between the PSI and I-EGF1 domains in

Chapter Four: The structures of PHE2 and PHE3

PHE1, PHE2 shows more polar contacts: the side chain hydroxyl of Ser9 forms two hydrogen bonds with the side chain of Arg449 in PHE2, while only one in PHE1 (Fig.4.11A). Because of the presences of the I-EGF2 domain, the loop (Thr30 to Arg39) in the PSI domain moves towards the I-EGF1 domain by approximately 2.0 Å; thereby, several of its residues make contacts with Lys457 of in the I-EGF1 domain and Asp489 in the I-EGF2 domain (Fig.4.11B, Table 4.5).

Table 4. 4 The r.m.s. deviations of domain superimpositions in integrin β 2 fragments (PHE1, PHE2 and PHE3)

Pair wise comparison between integrin fragments: r.m.s. deviation (Å) *			
Individual Domains	PHE1/PHE2	PHE1/PHE3	PHE2/PHE3
PSI	0.51	0.86	0.92
Hybrid	0.82	0.79	0.90
I-EGF1	0.19	1.57	1.65
I-EGF2	NA	NA	2.37
PSI Hybrid	0.86	1.13	1.07
PSI Hybrid I-EGF1	0.84	2.68	2.51
PSI Hybrid I-EGF1, 2	NA	NA	5.14

Note also the r.m.s. deviation of I-EGF3 domain between PHE3 and the integrin β 3 (Xiong et al., 2001) is 2.04 Å. NA: not applicable. *Estimated overall coordinate errors for the PHE2 and PHE3 molecules are 0.16 Å and 0.26 Å respectively.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 11 The polar interactions between PHE2 domains

Hydrogen bond interactions are represented with broken lines, and residues taking part in the polar interactions are labeled. Contacted residues are shown as sticks. (A) Interactions at the interfaces between the PSI domain (green) and I-EGF1 domain (red) in PHE2 and the additional hydrogen bond, compared to that of PHE1, is indicated by a red dash arrow. (B) Interactions at the interfaces between the PSI, I-EGF1, and I-EGF2 (purple) domains are shown in yellow. (C) For comparison, the corresponding interactions between the PSI and I-EGF1 domains in PHE1 are represented in gray.

Chapter Four: The structures of PHE2 and PHE3

Table 4. 5 hydrogen bond interactions between the I-EGF1 and PSI, and I-EGF2 domains in PHE2

PSI	Distance (Å)	I-EGF1	Distance (Å)	I-EGF2
Ser 9 O	3.41	Arg449 NH1		
Ser 9 O	2.71	Arg449 NH2		
Gln25 OE1	2.78	Arg428 N		
Asn28 OD1	2.8	Asn458 ND2		
Trp23 NE1	2.81	Cys445 O		
Arg39 NH2	2.96	Gly446 O		
Gly31 O	2.82	Lys457 NZ	3.14	Asp489 OD1
Asp34 O	2.69	Lys457 NZ	3.16	Asp489 OD2
Thr30 O	2.81	Lys457 NZ		
		Gly453 O	3.22	Thr463 OG1
		Gly453 O	2.96	Thr463 N
		Tyr454 OH	2.9	Gln462 OE1
		Ile455 N	2.96	Cys461 O
		Ile455 O	3.16	Cys461 N

The side chain of Asp489 points from the $\beta 1$ strand of the I-EGF2 domain into the crevice between the PSI and I-EGF1 domains, making hydrogen bonds with the amino group of Gly33 and the side chain of Lys457, and the latter, in turn, forms three polar contacts with the main chain carbonyl atoms (Gly31, Asp34, and Thr30) in the loop in the PSI domain. A water molecule is trapped to mediate an additional interaction with the side chain carboxyl atoms of Asp489 that may contribute to adjust the localization of

Chapter Four: The structures of PHE2 and PHE3

the side chain. However, there are only two hydrogen bonds between the side chain of Lys457 and the main chain carbonyl group of Thr30 and Gly31 in PHE1 (Fig.4.11C). The buried surfaces between the domains are modest, the interface between the I-EGF2 domain and the PSI domain and I-EGF1 domain is 405 Å² and 740 Å², respectively. The interfaces between the PSI and hybrid domains, and I-EGF1 domains are 840 Å² and 1220 Å², respectively, which are similar to those of PHE1 (Table 4.6).

Table 4. 6 Buried surface areas between individual domains for integrin fragments PHE1, PHE2 and PHE3

Interacting Domains	PHE1	PHE2	PHE3
PSI and Hybrid	815	840	875
PSI and I-EGF1	1100	1220	1230
PSI and I-EGF2	NA	405	0
I-EGF1 and I-EGF2	NA	740	535
I-EGF2 and I-EGF3	NA	NA	615

Note also the buried surface areas between I-EGF3 and I-EGF4 in integrin β3 subunit (Xiong et al., 2001) is 445 Å². NA: not applicable.

4.4.4 Oligosaccharide chains of PHE2

Three N-glycosylation sites on Asn28 Asn94 and Asn479 residues in Asn-X-Ser/Thr motifs were predicted. Electron density for an N-acetylglucosamine residue is visible at each predicated glycosylation sites at 1.0 σ contour level (Fig.4.12).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 12 The N-linked oligosaccharide chains of PHE2 after Endo H_f digestion

The electron density maps contoured at 1.0 σ surrounding an N-acetylglucosamine linked to Asn28 (A), Asn94 (B) and Asn479 (C) are showed in *gray*, respectively. The N-acetylglucosamine residue structures are depicted as *gray* sticks.

4.4.5 Residues with alternative side chain conformations

The high quality electron density map at the resolution of 1.75 Å permitted the identification of four residues (Glu2, Ser37, Arg43 and Gln400) with alternative side chain conformations (Fig.4.13A-D) on the surface of the PHE2 structure and which were modeled as single side chain conformations in PHE1 structure at the resolution of 1.8 Å. It should be mentioned that the electron density for residues Asp68-Gly72 in the hybrid domain is relatively weak. However, a continuous density connecting the main chain is still clear at 0.6 σ contour levels, and alternative conformations of the side chain of His69 is clear (Fig.4.13E). For simplicity, all alternative conformations are modeled with the occupancy of 0.5. The electron density maps also suggested minor alternative conformations for other residues in PHE2 structure, the density was not strong enough to warrant the inclusion of two conformations, therefore these residues were modeled as single conformations.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 13 The alternative conformations of the side chains

The alternative conformations of Glu2, Ser37, Arg43 and Gln400 along with the difference electron density map contoured at 1.0 σ levels in *blue*, while the map for His69 is displayed at 0.5 σ levels in *blue*. Those pictures are generated using software of the coot (Emsley et al., 2004).

4.5 The structure of PHE3

4.5.1 Quality of the model and overall structure

In contrast to PHE2, the PHE3 fragment adopts an extended conformation with dimensions of approximately 128 Å $\times$ 30 Å $\times$ 30 Å (Fig. 4.14). Overall, the main chain trace is clearly defined and four histidine residues from the hexa-histidine tag are visible at the C-terminal end.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 14 The structure of PHE3

The structure of the PHE3 features an extended conformation with its domains and secondary structure elements labeled. The Cys461-Cys492 disulfide bridge in I-EGF2 is labeled, as well as the segment of the I-EGF2 domain that undergoes an order (α -helix in PHE2) to disorder (coil in PHE3) transition. Seventeen disulfide bonds, the one N-linked acetylglucosamine residues and its corresponding Asn resolved in the structure are represented as sticks. The hexa-histidine tag is in *brown*.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 15 Independent validation of the PHE3 using the anomalous signal from the sulfur atoms

Anomalous Fourier electron density maps, with coefficients Dano (anomalous difference) and phases calculated from the refined models shifted by 90° are displayed at 3σ levels. The anomalous maps were cut using masks covering the entire PHE3 molecule. The same color code for the various integrin $\beta 2$ domains (PSI, hybrid, I-EGF1, 2 and 3) is used. Seventeen disulfide bridges that are present in the PHE3 structure coincide with peaks in the anomalous difference Fourier maps. In the anomalous map, individual sulfur atoms can be observed in the PHE3 structure but less distinct than in the PHE2 structure (Fig 4.9), especially for two disulfide bonds (C437-C448 and C451-C459) in the I-EGF1 domain. The peaks with heights ranging from 2.8 to 4.1σ are labeled with the corresponding disulfide bridge.

Chapter Four: The structures of PHE2 and PHE3

An anomalous Fourier map was computed from a low energy data set collected using the same PHE3 crystal that confirmed the location of the disulfide bonds present (Fig. 4.15; Table 4.3). The electron density for an N-acetylglucosamine residue is observed only on residue Asn94 in the hybrid domain. Some segments are poorly defined: residues 69-72 in the hybrid domain, residues 432-433 in the I-EGF1 domain, residues 467 in the I-EGF2 domain. The side chains of residues 466, 468, 469 lacking the electron density, are therefore not included in the refined PHE3 model.

4.5.2 The structure of the I-EGF2 domain

Superposition of individual domains from PHE3 with the corresponding segments in the PHE1 and PHE2 showed higher r.m.s. deviations compared to those observed between PHE1 and PHE2 (Table 4.4), which indicates that all the domain structure in itself may be intact without distinct changes. The most interesting difference between the two structures lies in the I-EGF2 domain. When the I-EGF2 domain from PHE3 was superposable to that of PHE2, four disulfide bridges and two center β strands are well fitted and its r.m.s. deviation is 2.37 (Fig.4.16A). The disulfide bond between Cys461 and Cys492 (C1 to C5 within the I-EGF2 domain) adopts two different conformations in PHE2 and PHE3. The dihedral angle of the disulfide bond (ξ_3 torsion angle) switches from the right-handed ($\xi_3 \sim 100^\circ$) in PHE2 to the left-handed ($\xi_3 \sim -80^\circ$) stable conformer (Fig.4.16A). Concomitantly, the short α -helix (residues Ser468-Cys475) in PHE2 becomes disordered in the elongated integrin structure (PHE3).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 16 Superposition of the I-EGF2 and I-EGF3 domains

(A) Superposition of the I-EGF2 domain in the bent (*cyan*) and extended (*dark blue*) conformations, highlighting the variation in the torsion angle occurring within the C1-C5 disulfide bond that links Cys461 and Cys492. The segment corresponding to the α -helix becomes very mobile in the extended conformation. The N-linked glycan at position Asn479 is also shown as sticks. (B) Superposition of the I-EGF3 domain in integrin $\beta 2$ (*pink*) and $\beta 3$ (*gray*) (Xiong et al., 2001). The four conserved disulfide bridges, the epitope of MEM48 (Leu534, Phe536, and His543) resolved in the structure are represented as sticks. His-tag is shown in *brown*.

Figure 4. 17 I-EGF3 domain sequence alignment.

Secondary structure elements are marked *above* the alignment in *pink* for the I-EGF3 of integrin $\beta 2$ subunit. The structural elements for the integrin $\beta 3$ I-EGF3 domain are marked in *blue*. All conserved cysteines are highlighted in *yellow* and the disulfide pairs are marked with *arrows*.

Chapter Four: The structures of PHE2 and PHE3

4.5.3 The structure of the I-EGF3 domain

The I-EGF3 domain also contains the central β sheet and four disulfide bonds is in agreement with its NMR structure (Beglova et al., 2002). The r.m.s. deviation of the I-EGF3 domain between $\beta 2$ and $\beta 3$ subunit is 2.04 Å (Table 4.4). The epitope of an integrin activated mAb MEM48 (Lu et al., 2001a) involving the residues Leu534, Phe536 and His543 locates in the central β sheet of I-EGF3 domain (Fig.4.16B; Fig.4.17), and all their side chains extend outward.

4.5.4 The conformational changes between PHE2 and PHE3 structures

The PSI/hybrid tandem remains invariant in all structures reported for integrin $\beta 2$ subunit including PHE1, PHE2 and PHE3. However, the I-EGF1-2 domains in PHE2 and PHE3 show remarkable differences in their relative orientation to the PSI/hybrid tandem. With reference to the invariant PSI/hybrid tandem, two maneuvers are required to bring the I-EGF1 and I-EGF2 domains from the PHE2 conformation to the PHE3 conformation: First, a rotation of 40° of the I-EGF1 domain about the axis passing through residues Lys427-Ser431 and a 0.2 Å shift (a value which falls within the experimental errors), followed by a rotation of 69° of the I-EGF2 domain about the axis passing through residues Ser468-Glu472 with a translation of 0.8 Å (Fig.4.18).

Based on the available crystal structures: PHE1 (PDB, 1YUK), PHE2 (PDB, 2P26), PHE3 (PDB, 2P28), integrin $\alpha II\beta 3$ open headpiece (PDB, 1TY6) and bent inactive integrin $\alpha V\beta 3$ ectodomain (PDB, 1U8C), the possible distinct conformers of the whole integrin extracellular structures were modeled (see Appendix 4).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 18 Movements in the quaternary structure

Movements in the quaternary structure, which bring the I-EGF1 and I-EGF2 domains in coincidence between the bent PHE2 form (A) and the extended PHE3 (C) structure. Panel (B) represents a hypothetical intermediate structure, which has not been observed experimentally. The same color code for individual domains is used. The values of rotation angles and the position of their axis are indicated.

Figure 4. 19 The polar interactions between the PSI and I-EGF1 domains

Hydrogen bond interactions in PHE3 (*right*) are represented with broken lines, and residues taking part in the interactions are labeled and showed in sticks. For comparison, the hydrogen bond interaction between PSI and I-EGF1 domains in PHE2 (*left*) are also shown in the same orientation.

Chapter Four: The structures of PHE2 and PHE3

4.5.5 Interaction between domains

The PSI/hybrid tandem remains constant among $\beta 2$ fragments (PHE1, PHE2 and PHE3) with r.m.s. deviations between equivalent C α positions of 0.86-1.13 Å (Table 4.4). The hydrogen bonds between the PSI domain and hybrid domain in PHE3 are the exactly the same as those of PHE1 or PHE2 (Fig.3.19), though this fragment adopts two distinct conformations. Because of the conformational changes between PHE2 and PHE3, which are only related to the reorientations of I-EGF1 and I-EGF2 domains, the contact area between the I-EGF1 and PSI domains in PHE3 remains similar to that of PHE1 or PHE2. However, the contact area between the I-EGF1 domain and the I-EGF2 domain is reduced from 740 Å² in PHE2 to 535 Å² in PHE3, and the contact area of 405 Å² between the PSI domain and the I-EGF2 domain in PHE2 no longer exists in PHE3 (Table 4.6). Accordingly, the hydrogen bond interactions between Lys457 of the I-EGF1 domain, Asp489 of the I-EGF2 domain, and the main chain amide groups projecting from the loop of the PSI domain in PHE2 are disrupted in PHE3. Besides these changes, some polar contacts between the PSI domain and the I-EGF1 domain are abolished (Table 4.7) when compared to that of PHE1 or PHE2, and only two hydrogen bonds between Trp23 and Cys445, and between Arg39 and Gly446 exist (Fig.4.19). These four residues are highly invariant among integrin β subunit sequences (Fig.3.15), suggesting that these interactions may be well conserved.

The relative position of the I-EGF1 and I-EGF2 domains is also different in PHE2 and PHE3. Only three hydrogen bonds are detected in this region in PHE3 (Fig.4.20B): two are formed by the main chains of Ile455 and Cys461, the third is formed between the

Chapter Four: The structures of PHE2 and PHE3

main chain carbonyl group of Gly453 and the main chain amino group of Thr463. Two additional hydrogen bonds are observed between the side chain hydroxyl of Tyr454 and the side chain carbonyl oxygen of Gln462 and between main chain carbonyl group of Gly453 and side chain hydroxyl of Thr463 in PHE2 (Fig.4.20A). A list of these interactions in PHE3 is given in Table 4.7.

Table 4. 7 hydrogen bond contacts between domains in PHE3

PSI	Distance (Å)	I-EGF1	Distance (Å)	I-EGF2	Distance (Å)	I-EGF3
Trp23 NE1	2.93	Cys445 O		Gly504 O	2.68	Thr516 OG1
Arg39 NH2	2.94	Gly446 O		Leu506 N	3.00	Thr516 N
Gly31 O	3.44	Lys457 NZ		Leu506 O	3.45	Thr516 N
Asp34 O	2.75	Lys457 NZ		Tyr508 N	3.14	Cys514 O
Thr30 OG1	3.18	Lys457 NZ		Tyr508 O	3.06	Cys514 N
				Tyr508 OH	3.04	Cys514 N
		Gly453 O	3.24	Thr463 N		
		Ile455 N	2.97	Cys461 O		
		Ile455 O	3.22	Cys461 N		

The interface between the I-EGF2 and I-EGF3 covers 615 Å² in PHE3 (Table 4.6) and includes at its centre hydrogen-bonding residues Gly504, Leu506 and Tyr508. Interestingly, KIM127, a mAb which only recognizes the extended β 2 integrin (Nishida et al., 2006), has been mapped to these three residues, and all residues interact intimately with the I-EGF3 module to form an intricate hydrogen bond network (Fig.4.21).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 20 The polar interactions between I-EGF1 and I-EGF2 domains

The hydrogen bond interactions at the interface between the I-EGF1 (*red*) and I-EGF2 (*purple*) domains in PHE2 (A) and in PHE3 (B). Hydrogen bonds are represented by broken lines, and residues taking part in the interactions are labeled and showed in sticks.

Chapter Four: The structures of PHE2 and PHE3

The side chain of Thr516 points to the loop between C7 and C8 of I-EGF2 and make two hydrogen bonds with the main chain carbonyl group of Gly504 and the main chain amino group of Leu506, its main chain amino group is in the hydrogen-bond distance of the main chain carbonyl group of Leu506. The main chains of Tyr508 and Cys514 form two parallel hydrogen bonds with each other, and the side chain hydroxyl of Tyr508 deeply penetrates the I-EGF3 domain, making a hydrogen bond with the main chain carbonyl group of Cys537. This hydrogen network may stabilize the interface between I-EGF2 and I-EGF3 domains.

Figure 4. 21 The polar interactions between PHE3 domains

The hydrogen bond interactions at the interface between the I-EGF2 (*purple*) and I-EGF3 (*pink*) domains in PHE3. Hydrogen bonds are represented by broken lines, and residues taking part in the interactions are showed in sticks and labeled.

Chapter Four: The structures of PHE2 and PHE3

4.5.6 Oligosaccharide chains of PHE3

Three N-glycosylation sites on Asn28, Asn94 and Asn479 residues within Asn-X-Ser/Thr motifs were also predicted in PHE3 using Vector NTI software. Electron density for only one N-acetylglucosamine residue is visible at the potential glycosylation site on Asn94 in the hybrid domain at 1.0 σ contour levels (Fig. 4.22). Although Asn28 and Asn479 are also predicted to be glycosylation sites, the electron density for the carbohydrates are not well defined to identify the carbohydrate residues. This is possibly due to the flexible nature of the oligosaccharides in PHE3 structure.

Figure 4. 22 The N-linked oligosaccharide chains of PHE3 after Endo H_f digestion

The electron density maps contoured at 1.0 σ contour levels surrounding an N-acetylglucosamine linked to Asn94 (B) and the possible disordered oligosaccharide chain linked to Asn28 (A) and Asn479 (C) are showed, respectively. The N-acetylglucosamine residue structure is depicted as *gray sticks*.

Chapter Four: The structures of PHE2 and PHE3

4.6 The bent conformation observed in the PHE2 crystal structure can be adopted by a complete integrin heterodimer.

The PHE2 and PHE3 molecules are only fragments of a complete integrin $\beta 2$ subunit. To rule out possible artifacts introduced by crystal packing forces and to assess whether the bent conformation displayed by the PHE2 molecule can also be adopted by a complete integrin receptor, an engineering disulfide bond introduced by site-directed mutagenesis would constrain the integrin in the bent conformation observed in PHE2 structure. In the PHE2 crystal structure, the α -carbon atoms of residues Gly33 from the PSI domain and Gly486 and Leu487 from the I-EGF2 domain are separated by distances of 4.21 and 5.05 Å respectively, which are compatible with the distance spanned by the α -carbon atoms within disulfide bonds (Fig.4.23). Two double substitutions were made in the full-length $\beta 2$ integrin subunit, with cysteines replacing Gly33 and Gly486 in one ($\beta 2G33C/G486C$) and Gly33 and Leu487 in the other ($\beta 2G33C/L487C$). In addition, one single mutant with Gly33 substituted by cysteine residues ($\beta 2G33C$) was also constructed as a control.

The αL and $\beta 2$ cDNA plasmids were transfected into 293T cells. All three $\beta 2$ variants ($\beta 2G33C$, $\beta 2G33C/G486C$ and $\beta 2G33C/L487C$) supported LFA-1 (Leukocyte Function Associated Antigen-1, $\alpha L\beta 2$ integrin) surface expression, as detected by flow cytometry using mAb MHM23, which is specific for the heterodimer (Fig.4.24).

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 23 Introduction of a disulfide bond that covalently links the PSI and I-EGF2 domains in integrin $\beta 2$ subunit

The locations of residues Gly33, Gly486 and Leu487 in the PHE2 structure are shown. Gly33 and Gly486, Leu487 (marked with a *green* and two *purple* spheres, respectively), when substituted by the cysteines, only G33C and G486C form a disulfide bond. The distances between the α -carbon atoms of residues Gly33 and Gly486, Gly33 and Leu487 are also shown ($\text{\AA}$).

Figure 4. 24 LFA-1 and variants expression

The cell-surface expression of LFA-1 and variants on 293T transfectants were detected by MHM23 (a $\beta 2$ integrin heterodimer-specific mAb) or by an irrelevant mAb KB43 (integrin αX -specific mAb), followed by flow cytometry. Shaded histogram represents the labeling of the LFA-1 and variants transfectants, and unshaded histogram is the background.

Chapter Four: The structures of PHE2 and PHE3

The transfectants were surface labeled with biotin-BMCC, a reagent containing maleimide that binds to free sulfhydryl groups. LFA-1 was immunoprecipitated from solubilized membranes of the transfectants with MHM23, and was analyzed for biotin labeling by Western blot (Fig.4.25, upper panel). Minimal signal was detected from the bands of the wild-type $\alpha\text{L}\beta 2$. In contrast, the $\beta 2\text{G33C}$ and $\beta 2\text{G33C/L487C}$ bands were labeled, indicating the presence of free sulfhydryl in these variant $\beta 2$ subunits. The intensity of the $\beta 2\text{G33C/G486C}$ band was significantly lower than those of the other two mutants, suggesting that a disulfide bond was formed between the two introduced cysteine residues in the majority of the molecules. The same samples were also analyzed for the presence of the αL subunit by Western blot using the anti- αL mAb clone 27 (Fig.4.25, lower panel). The results showed that the αL subunits are present in equal amounts in the four transfectants, suggesting that the different levels of free cysteines in the four samples were not due to differences in extent of immunoprecipitation.

Expression of the KIM127 epitope in LFA-1 is associated with an extended conformation (Lu et al., 2001a; Nishida et al., 2006). KIM127 can bind to the extended integrin induced by $\text{Mg}^{2+}/\text{EGTA}$, as stated by wild-type $\beta 2$ integrin (Dransfield et al., 1992). To test the reactivity of KIM127 with the LFA-1 variants including an additional control of $\beta 2\text{G486C}$, transfectants were surface-labeled with biotin (Tan et al., 2001a) and the labeled cells were incubated at 37°C with KIM127 in presence of $\text{Mg}^{2+}/\text{EGTA}$. Cells were subsequently lysed and the biotin-labeled LFA-1 was precipitated with protein A-Sepharose. KIM127 was able to immunoprecipitate the wild-type LFA-1, the

Chapter Four: The structures of PHE2 and PHE3

α L β 2G33C, α L β 2G486C and α L β 2G33C/L487C LFA-1 variants, but not the α L β 2G33C/G486C LFA-1 variant (Fig.4.26).

Figure 4. 25 The formation of the disulfide bond in LFA-1 variants was monitored using the free sulfhydryl labeling reagent BMCC conjugated to biotin

LFA-1 was immunoprecipitated with mAb MHM23, resolved on a non-reducing 7.5% SDS-PAGE gel, followed by detection of biotin-BMCC labeled LFA-1 (upper panel). The reducing agent DTT-treated wild-type LFA-1 was included as a positive control. The amount of LFA-1 and variants immunoprecipitated with MHM23 was comparable, as indicated on the Western blot with the α L-specific antibody clone 27 (lower panel).

This was not due to inefficient labeling of LFA-1 variant α L β 2G33C/G486C because the control mAb MHM23 precipitated LFA-1 and its three variants. Thus, the lack of reactivity of KIM127 with LFA-1 variant α L β 2G33C/G486C suggests that the engineered disulfide bridge, as a result of β 2G33C/G486C mutations, restrains LFA-1 in a conformation that is incapable of full extension in the presence of the activating reagent Mg^{2+} /EGTA. It should be noted that wild-type LFA-1 under physiological divalent cation concentration in culture medium is not precipitated by KIM127 as previously reported

Chapter Four: The structures of PHE2 and PHE3

(Cheng et al., 2007). Although there was a weak band of the integrin $\beta 2$ subunit, there was no band for the integrin αL subunit. Therefore, the weak band represented free $\beta 2$ subunit which can be expressed by itself. Furthermore, the MEM48 (integrin activate antibody) can only precipitate the wild-type (WT) but not the $\alpha L\beta 2G33C/G486C$ variant. In contrast, the disulfide bond does not affect the $\alpha L\beta 2G33C/G486C$ variant precipitation by integrin activate antibody KIM185 whose epitope locates in the I-EGF4 domain of $\beta 2$ integrin (Lu et al., 2001). An attempt was made to reduce the introduced disulfide bond in the $\alpha L\beta 2G33C/G486C$ variant with varying concentrations of DTT. Even at 30 mM DTT, the $\alpha L\beta 2G33C/G486C$ cannot be precipitated with KIM127 in the presence of $Mg^{2+}/EGTA$, although under these conditions, the precipitation of wild-type $\alpha L\beta 2$ is easily detected (Fig.4.27). At 100 mM DTT, the epitope is abolished in the wild-type, presumably by the reduction of other disulfide bonds.

Figure 4. 26 Reactivity of KIM127 with LFA-1 and variants

Biotinylated surface-expressed LFA-1 and variants on 293T transfectants were immunoprecipitated with KIM127, MHM23 (positive control mAb), KIM185 and MEM48 with or without ME. The subunits of LFA-1 were resolved on a 7.5% SDS-PAGE gel followed by detection.

Chapter Four: The structures of PHE2 and PHE3

Figure 4. 27 Immunoprecipitation of LFA-1 in Mg^{2+} /EGTA and at different concentrations of DTT by KIM127

Wild-type immunoprecipitation is not affected up to 30 mM DTT. No precipitation was observed for the variant. The introduced disulfide is not reduced up to 30 mM DTT.

4.7 The $\alpha L\beta 2G33C/G486C$ variant can mediate adhesion to ICAM-1.

Unlike the wild-type LFA-1, the $\alpha L\beta 2G33C/G486C$ variant can mediate the adhesion of 293T transfectants to ICAM-1 coated surface in the absence of Mg^{2+} /EGTA (Fig.4.28).

Figure 4. 28 Adhesion of LFA-1 transfectants to ICAM-1

293T transfectants expressing wild-type LFA-1 or the $\alpha L\beta 2G33C/G486C$ variants were assayed for ICAM-1 coated surface in the presence or absence of Mg^{2+} /EGTA. Forward bars represent results obtained in the presence of the LFA-1 specific blocking mAb MHM24.

Chapter Four: The structures of PHE2 and PHE3

4.8 Discussion

4.8.1 Constructions and expressions

PHE2, which has an additional glycosylation site in the I-EGF2 domain, was initially expressed in HEK293 cells as we did for PHE1. However, only small needle form crystals were obtained. Numerous ways were tried to increase the sizes of the PHE2 crystals, but they were still too small to be mounted for diffraction. It is known that the glycan moieties often perturb crystallization because of its chemical and conformational heterogeneity (Davis et al., 1993) and it is possible that the extra glycosylation may prevent the PHE2 forming well diffracted crystals. Therefore, several attempts were made to deglycosylate the recombinant proteins. Asn to Gln substitution is a useful way to eliminate the N-glycosylation and to produce low glycosylated recombinant proteins suitable for crystallization, and some successful cases have been reported. For example, recombinant human butyrylcholinesterase with Asn to Gln substitution was successfully expressed in CHO cells, and tetragonal crystals were obtained and diffracted to 2.0 Å (Nachon et al., 2002). Therefore, we conducted a PHE2 construct with N479Q substitution in the I-EGF2 domain (PHE2_{N479Q}). A stable HEK293 cell line was also created, PHE2_{N479Q} was expressed at the level similar to that of native PHE2, but crystals were not obtained.

In order to solve the glycosylation modification problem, The GnTI⁻ HEK293S cells was used to express the recombinant proteins PHE2 and PHE3. Folding and initial glycosylation of the proteins expressed in GnTI⁻ HEK293S cells would proceed normally, but processing of the N-glycans is restricted so that they may be removed with

Chapter Four: The structures of PHE2 and PHE3

endoglycosidase (endo) H (Butters et al., 1999; Davis et al., 1995; Davis et al., 1993). Endo H cleaves between the GlcNAc residues in the di-N-acetylchitobiose core of oligomannose and hybrid-type N-glycans, leaving single GlcNAc residue at each glycosylation site. Furthermore, in contrast to proteins expressed in HEK293 cells, GnTII deficient HEK293S-derived glycoproteins contain only traces of core fucose (Crispin et al., 2006), thus allowing endo H cleavage to proceed more efficiently. Crystals grew from endo H_f treated PHE2, PHE3 expressed in GnTII deficient HEK293S cells diffracted to 1.8Å, 2.2Å, respectively.

4.8.2 PHE2 and PHE3 structures

The determined structures of the PHE2 and PHE3 fragments of the integrin β 2 subunit show two distinct conformations. An acute bend is observed in the PHE2 structure between the I-EGF1 domain and the I-EGF2 domain, forming an L-shaped conformation. While PHE3 assumes an extended conformation, faithfully reflecting the extended conformer of the integrin. This acute linkage supports our previous conjecture that the Genu of integrin β subunit is between the I-EGF1 and I-EGF2 domains (Shi et al., 2005).

Three cysteine-rich repeats have been postulated from the first integrin β subunit sequence and each contained eight cysteines (Law et al., 1987; Tamkun et al., 1986). Subsequently, the fourth I-EGF domain containing seven cysteines, located directly at the N-terminal of the other three cysteine-rich repeats, was identified and the other three repeats were supposed to contain eight cysteines (Kishimoto et al., 1987). However, the definition of domain boundaries is incorrect due to their adjacent regions are also

Chapter Four: The structures of PHE2 and PHE3

cysteine rich and the sequence conservation among the four cysteine-rich repeats are low, except for the cysteines and a few glycines (Fig.4.29B). Thus, various possible combinations have been suggested (Arnaout, 2002; Law et al., 1987; Takagi et al., 2001a; Tan et al., 2001b). The $\alpha V\beta 3$ crystal structure (Xiong et al., 2001) greatly facilitated the correct boundary determination of other domains, except for the I-EGF domains. It proposed that there are six cysteines in each I-EGF domains connected in the C1-C3, C2-C4, C5-C6 pattern with an additional disulfide bond between consecutive I-EGF domains (Arnaout et al., 2002). Based on epitope expression and series of truncated $\beta 2$ variants, the boundaries of the I-EGF domains had been experimentally defined (Takagi et al., 2001a; Tan et al., 2001b). Takagi et al (2001a) also showed the I-EGF domains found in the integrins are distinct from those in laminin, and suggested that the eight cysteines in each integrin I-EGF domain were connected in the C1-C5, C2-C4, C3-C6, C7-C8 form except the I-EGF1 domain that only contained six cysteines forming three disulfide bonds in C1-C5, C3-C6, C7-C8 pattern (Takagi et al., 2001a). This model was further supported by the expression of the $\beta 2$ I-EGF3 domain and its NMR structure (Beglova et al., 2002). Our work on PHE1, PHE2 and PHE3 structures substantiated this model and provided definite structures for the I-EGF1 and I-EGF2 domains for the first time. A structure superposition (Fig.4.29A) has been carried out among available structures of I-EGF domains, $\beta 2$ I-EGF3 domain (Beglova et al., 2002), $\beta 3$ I-EGF3/I-EGF4 domains (Xiong et al., 2004) and $\beta 2$ I-EGF1/I-EGF2/I-EGF3 domains in this thesis study.

Chapter Four: The structures of PHE2 and PHE3

A

B

		1		2		3		4	5	6	7		8
	27	CRDQS		R		DRSLCHG		KGFLECGICRC		DTGY		IGKNCE	
$\beta 2$ I-EGF2	461	CQTQGRSSQELEGS		CRKDNNSSIICSG		LGDCVCGQCLCHTSDVPGKLIYGQYCE							
$\beta 2$ I-EGF3	514	CDTIN		CERY		NGQVCGGPGRGLCF		CGKCR		HPGF		EGSACQ	
$\beta 2$ I-EGF4	553	CERTT		EG		CLNP		RRVECSG		RGRRCNVCEC		HSGY	
												QLPLCQ	

Figure 4. 29 I-EGF domains superposition and sequence alignment

(A) Superposition of $\beta 2$ I-EGF domains 1, 2, 3 (this work) and integrin $\beta 3$ I-EGF3-4 domain (Xiong et al., 2001) highlighting the structural conservation of the core composed by the central 2-stranded β -sheet. I-EGF2 contains a long stretch of residues between cysteines C1 and C5 which contains an α -helix (shown in *gray*) in the bent (PHE2) form but becomes disordered in the elongated (PHE3) conformation. (B) Structure based sequence alignment of integrin $\beta 2$ I-EGF domains.

Chapter Four: The structures of PHE2 and PHE3

The interface of PSI/hybrid domain remains invariant in the bent and extended integrin $\beta 2$ fragments structures, as visible in integrin $\beta 3$ fragments structures (Xiao et al., 2004). Arg86 (Arg93 in integrin $\beta 3$ subunit) from the hybrid domain inserts its side chain deep into the PSI domain, making extensive polar contacts (Fig.3.19A). Moreover, the stability of this interface is enhanced by the disulfide bonds in the proximity. Thus, a movement of the hybrid domain is directly linked to the PSI domain movement and *vice versa*. Thus these two domains may be considered as a rigid unit.

In the PHE2 and PHE3 structures, the I-EGF2 domain assumes two different conformations involving alternative dihedral torsion angles of a disulfide bond between Cys461 and Cys492 and the presence or absence of a short α -helix (Fig.4.16A). This difference may represent the conformational switch within the leg of the integrin $\beta 2$ subunit associated with a large reorientation of the I-EGF1 and I-EGF2 domains. An energy barrier of *ca* 7 kcal/mol has to be overcome in order to switch the disulfide bond from one favorable conformer to the other (Creighton 1993). However, this kinetic energy barrier would not contribute to affinity regulation when the system is at thermodynamic equilibrium. Whether an exogenous disulfide isomerase is involved in the transition remains to be clarified (Jordan et al., 2005). It is also possible that the integrin may have intrinsic disulfide isomerase activity (O'Neill et al., 2000).

4.8.3 KIM127, a conformation sensitive integrin mAb

The extended PHE3 structure may represent the extended Genu in the β subunit, in agreement with EM observations on $\alpha L\beta 2$ and $\alpha X\beta 2$ (Nishida et al., 2006). It also

Chapter Four: The structures of PHE2 and PHE3

expresses the epitopes of a panel of conformation-sensitive mAbs, in particular, with KIM127, which does not detect the bent LFA-1 conformers (Lu et al., 2001a; Nishida et al., 2006). By contrast, it is not immediately apparent from the lack of KIM127 reactivity with PHE2 because it is bent, since the expression of the KIM127 epitope requires I-EGF3, which is absent in PHE2 (Takagi et al., 2001a; Tan et al., 2001b). To verify that the L-shaped conformation in the PHE2 structure also occurs in an intact integrin, we introduced two cysteine residues, one to replace Gly33 in the PSI domain and the other to replace Gly486 in the I-EGF2 domain of the $\beta 2$ subunit. The distance between the C α atoms of these two residues is 16 Å in the extended (PHE3) form but only 4.2 Å in the bent (PHE2) structure (Fig.4.23). Thus, we expect that the disulfide bond would only be formed if these two cysteine residues are brought in close proximity, as seen in the compact PHE2 structure. Formation of this engineered disulfide bond in this LFA-1 variant was confirmed. In addition, the epitope of KIM127 was not expressed in the presence of Mg²⁺/EGTA, which is also expected if the LFA-1 variant is locked by the engineered disulfide into a bent conformation. Thus, these data provide direct evidence for the presence of a bent conformation adopted by LFA-1 integrins at the cell surface that resembles the PHE2 structure. On the other hand, in the absence of the engineered disulfide bond such as in the LFA-1 variants with α L β 2G33C, α L β 2G486C and α L β 2G33C/L487C substitutions, the extended conformations can be detected by the KIM127 monoclonal antibody in the presence of Mg²⁺/EGTA. These data also give additional support that mAb KIM127 is conformation sensitive and only binds to extended integrin, not bent one, as postulated previously (Nishida et al., 2006).

Chapter Four: The structures of PHE2 and PHE3

The MEM48 epitope, which also requires both I-EGF2 and I-EGF3 for expression in wild-type LFA-1 and is not sensitive to Mg^{2+} /EGTA, is not expressed in the $\alpha L\beta 2G33C/G486C$ variant. This is not unexpected since the introduced disulfide bond would lock I-EGF2 into a conformation that does not allow the I-EGF2 and I-EGF3 to bind MEM48 by an induced-fit mechanism. On the other hand, the expression of the epitope of KIM185, which maps to the I-EGF4/ β -TD domains (Lu et al., 2001a), is not affected. Our attempt to reduce the introduced disulfide bond on the $\alpha L\beta 2G33C/G486C$ variant was not successful. The expression of the KIM127 epitope was not affected in 30 mM DTT in the wild-type LFA-1. Under these conditions, re-expression of the KIM127 epitope was not detected on the $\alpha L\beta 2G33C/G486C$ variant, suggesting that the introduced disulfide between the PSI and I-EGF2 domains are not accessible for reduction. At 100 mM DTT, expression of the epitope was abolished in the wild-type LFA-1, presumably by the reduction of other disulfide bonds in the $\beta 2$ subunit. The $\alpha L\beta 2G33C/G486C$ variant can mediate adhesion to ICAM-1, suggesting that the ligand binding site is intact. Furthermore, it also suggests that leg extension is not an absolute requirement for ligand binding, at least under this particular assay system.

Chapter Five: Summary

In this thesis, three fragments (PHE1, PHE2 and PHE3) of integrin $\beta 2$ subunit have been expressed and crystallized for structural and functional studies. Different expression systems have been tested and the mammalian expression system is successfully applied to express all three recombinant proteins. Crystals displaying good diffraction quality were obtained and the structures are solved at high resolutions.

PSI domain is located in the middle of the crystal structure of PHE1, which contained PSI, hybrid domain and I-EGF1 domain, and makes the extensive contact areas with the hybrid domain and the I-EGF1 domain, respectively, suggested that the whole structure might be considered as a rigid unit to transmit the inside-out and outside-in signals. Superposition of the PHE1 structure onto the $\alpha V\beta 3$ crystal structure indicated that the bend in integrin $\beta 2$ subunit must locate between the I-EGF1 and I-EGF2 domains, which was subsequently confirmed by the PHE2 crystal structure. The I-EGF structures solved in PHE1, PHE2 and PHE3, together with I-EGF4 structure from the $\alpha V\beta 3$ crystal structure, substantiate not only the boundaries of the I-EGF domains, but also the connecting pattern of the disulfide bonds within I-EGF domains.

The following solved crystal structures of the $\beta 2$ Genu in both bent (PHE2) and extended (PHE3) conformations provide a structural basis, represented by conformational changes located in the $\beta 2$ Genu, for the hypothesis of how integrins switch form a bent to a extended form. In addition, we showed that LFA-1 can be locked into a bent

Chapter Five: Summary

conformation by a disulfide bond strategically introduced between the PSI and I-EGF2 domains, the “bent” LFA-1 has the capacity for the ligand binding, which may provide a further twist in the understanding towards the allostery mechanisms responsible for the integrin activation.

However, PHE2 and PHE3 are only fragments of the integrin $\beta 2$ subunit, how this conformational change is coupled with others described, including leg separation (Kim et al., 2003), removal of a deadbolt (Xiong et al., 2003b), swing-out of the hybrid domain (Xiao et al., 2004), and conformational changes in the I-domains of the α and β subunits (Alonso et al., 2002; Arnaout et al., 2005; Luo et al., 2007; Staunton et al., 2006; Yang et al., 2004) will require further examination. In order to obtain the overall atomic-level conformational changes which accompany integrin activation, the crystal structures, at least the whole $\beta 2$ integrin ectodomain in both bent and extended conformations are necessary. For the universal mechanism of integrin activation, these must also be done in other integrins.

Publications related to this study

1. **Shi, M.**, Foo, S.Y., Tan, S.M., Mitchell, E.P., Law, S.K., Lescar, J. (2007). A Structural hypothesis for the Transition between Bent and Extended Conformations for the Leukocyte $\beta 2$ Integrins. *J. Biol. Chem.* 282, 30198-30206.
2. **Shi, M.**, Sundramurthy, K., Liu, B., Tan, S.M., Law, S.K., Lescar, J. (2005). The crystal structure of the PSI/Hybrid domain/I-EGF1 segment from human integrin beta2 subunit at 1.8 Å resolution. *J. Biol. Chem.* 280, 30586-30593.
3. Cheng, M., Foo, S.Y., **Shi, M.**, Tang, R.H., Kong, L.S., Law, S.K., Tan, S.M. (2007). Mutation of a conserved asparagines in the I-like domain promotes constitutively active integrin $\alpha L\beta 2$ and $\alpha I Ib\beta 3$. *J. Biol. Chem.* 282, 18225-18232.

References

(1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.

Adair, B. D., Xiong, J. P., Maddock, C., Goodman, S. L., Arnaout, M. A., Yeager, M. (2005). Three-dimensional EM structure of the ectodomain of integrin $\{\alpha\}V\{\beta\}_3$ in a complex with fibronectin. *J Cell Biol* 168, 1109-1118.

Ajrout, K., Sugimori, T., Goldmann, W. H., Fathallah, D. M., Xiong, J. P., Arnaout, M. A. (2004). Binding Affinity of Metal Ions to the CD11b A-domain Is Regulated by Integrin Activation and Ligands. *J Biol Chem* 279, 25483-25488.

Alonso, J. L., Essafi, M., Xiong, J. P., Stehle, T., Arnaout, M. A. (2002). Does the integrin αA domain act as a ligand for its βA domain? *Curr Biol* 12, R340-342.

Altieri, D. C. (1991). Occupancy of CD11b/CD18 (Mac-1) divalent ion binding site(s) induces leukocyte adhesion. *J Immunol* 147, 1891-1898.

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

Arnaout, M. A. (1990). Leukocyte adhesion molecules deficiency: its structural basis, pathophysiology and implications for modulating the inflammatory response. *Immunol Rev* 114, 145-180.

Arnaout, M. A. (2002). Integrin structure: new twists and turns in dynamic cell adhesion. *Immunol Rev* 186, 125-140.

Arnaout, M. A., Goodman, S. L., Xiong, J. P. (2002). Coming to grips with integrin binding to ligands. *Curr Opin Cell Biol* 14, 641-651.

Arnaout, M. A., Mahalingam, B., Xiong, J. P. (2005). Integrin structure, allostery, and bidirectional signaling. *Annu Rev Cell Dev Biol* 21, 381-410.

Arnaout, M. A., Remold-O'Donnell, E., Pierce, M. W., Harris, P., Tenen, D. G. (1988). Molecular cloning of the α subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the α subunits of integrins. *Proc Natl Acad Sci U S A* 85, 2776-2780.

Baldwin, E. T., Sarver, R. W., Bryant, G. L., Jr., Curry, K. A., Fairbanks, M. B., Finzel, B. C., Garlick, R. L., Heinrichson, R. L., Horton, N. C., Kelley, L. L., *et al.* (1998). Cation binding to the integrin CD11b I domain and activation model assessment. *Structure* 6, 923-935.

References

- Baneres, J. L., Roquet, F., Martin, A., Parello, J. (2000). A minimized human integrin $\alpha(5)\beta(1)$ that retains ligand recognition. *J Biol Chem* 275, 5888-5903.
- Beglova, N., Blacklow, S. C., Takagi, J., Springer, T. A. (2002). Cysteine-rich module structure reveals a fulcrum for integrin rearrangement upon activation. *Nat Struct Biol* 9, 282-287.
- Berg, R. W., Leung, E., Gough, S., Morris, C., Yao, W. P., Wang, S. X., Ni, J., Krissansen, G. W. (1999). Cloning and characterization of a novel beta integrin-related cDNA coding for the protein TIED ("ten beta integrin EGF-like repeat domains") that maps to chromosome band 13q33: A divergent stand-alone integrin stalk structure. *Genomics* 56, 169-178.
- Bokel, C., Brown, N. H. (2002). Integrins in development: moving on, responding to, and sticking to the extracellular matrix. *Dev Cell* 3, 311-321.
- Bork, P., Doerks, T., Springer, T. A., Snel, B. (1999). Domains in plexins: links to integrins and transcription factors. *Trends Biochem Sci* 24, 261-263.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-921.
- Butta, N., Arias-Salgado, E. G., Gonzalez-Manchon, C., Ferrer, M., Larrucea, S., Ayuso, M. S., Parrilla, R. (2003). Disruption of the beta3 663-687 disulfide bridge confers constitutive activity to beta3 integrins. *Blood* 102, 2491-2497.
- Butters, T. D., Sparks, L. M., Harlos, K., Ikemizu, S., Stuart, D. I., Jones, E. Y., Davis, S. J. (1999). Effects of N-butyldeoxynojirimycin and the Lec3.2.8.1 mutant phenotype on N-glycan processing in Chinese hamster ovary cells: application to glycoprotein crystallization. *Protein Sci* 8, 1696-1701.
- Calderwood, D. A. (2004). Talin controls integrin activation. *Biochem Soc Trans* 32, 434-437.
- Calvete, J. J., Henschen, A., Gonzalez-Rodriguez, J. (1991). Assignment of disulphide bonds in human platelet GPIIIa. A disulphide pattern for the beta-subunits of the integrin family. *Biochem J* 274 (Pt 1), 63-71.
- Carvin, D., Islam S.A., Sternberg M.J.E and Blundell T.L. (2001). The preparation of heavy-atom derivatives of protein crystals for use in multiple isomorphous replacement and anomalous scattering. *International Tables for Crystallography (Volume F: Crystallography of Biological macromolecules Editors: MG Rossmann and E Arnold, Cordrecht: Kluwer Academic Publishers, The Netherlands. 247-255. .*

References

- Chang, V. T., Crispin, M., Aricescu, A. R., Harvey, D. J., Nettleship, J. E., Fennelly, J. A., Yu, C., Boles, K. S., Evans, E. J., Stuart, D. I., *et al.* (2007). Glycoprotein structural genomics: solving the glycosylation problem. *Structure* *15*, 267-273.
- Chen, J., Salas, A., Springer, T. A. (2003). Bistable regulation of integrin adhesiveness by a bipolar metal ion cluster. *Nat Struct Biol* *10*, 995-1001.
- Cheng, M., Foo, S. Y., Shi, M. L., Tang, R. H., Kong, L. S., Law, S. K., Tan, S. M. (2007). Mutation of a conserved asparagine in the I-like domain promotes constitutively active integrins alphaLbeta2 and alphaIIb beta3. *J Biol Chem* *282*, 18225-18232.
- Crispin, M., Harvey, D. J., Chang, V. T., Yu, C., Aricescu, A. R., Jones, E. Y., Davis, S. J., Dwek, R. A., Rudd, P. M. (2006). Inhibition of hybrid- and complex-type glycosylation reveals the presence of the GlcNAc transferase I-independent fucosylation pathway. *Glycobiology* *16*, 748-756.
- Dana, N., Fathallah, D. M., Arnaout, M. A. (1991). Expression of a soluble and functional form of the human beta 2 integrin CD11b/CD18. *Proc Natl Acad Sci U S A* *88*, 3106-3110.
- Davis, S. J., Davies, E. A., Barclay, A. N., Daenke, S., Bodian, D. L., Jones, E. Y., Stuart, D. I., Butters, T. D., Dwek, R. A., van der Merwe, P. A. (1995). Ligand binding by the immunoglobulin superfamily recognition molecule CD2 is glycosylation-independent. *J Biol Chem* *270*, 369-375.
- Davis, S. J., Puklavec, M. J., Ashford, D. A., Harlos, K., Jones, E. Y., Stuart, D. I., Williams, A. F. (1993). Expression of soluble recombinant glycoproteins with predefined glycosylation: application to the crystallization of the T-cell glycoprotein CD2. *Protein Eng* *6*, 229-232.
- de Pereda, J. M., Wiche, G., Liddington, R. C. (1999). Crystal structure of a tandem pair of fibronectin type III domains from the cytoplasmic tail of integrin alpha6beta4. *Embo J* *18*, 4087-4095.
- Dickeson, S. K., Walsh, J. J., Santoro, S. A. (1997). Contributions of the I and EF hand domains to the divalent cation-dependent collagen binding activity of the alpha2beta1 integrin. *J Biol Chem* *272*, 7661-7668.
- Douglass, W. A., Hyland, R. H., Buckley, C. D., Al-Shamkhani, A., Shaw, J. M., Scarth, S. L., Simmons, D. L., Law, S. K. (1998). The role of the cysteine-rich region of the beta2 integrin subunit in the leukocyte function-associated antigen-1 (LFA-1, alphaLbeta2, CD11a/CD18) heterodimer formation and ligand binding. *FEBS Lett* *440*, 414-418.

References

- Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D., Handford, P. A. (1996). Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. *Cell* 85, 597-605.
- Dransfield, I., Cabanas, C., Craig, A., Hogg, N. (1992). Divalent cation regulation of the function of the leukocyte integrin LFA-1. *J Cell Biol* 116, 219-226.
- Dustin, M. L., Springer, T. A. (1989). T-cell receptor cross-linking transiently stimulates adhesiveness through LFA-1. *Nature* 341, 619-624.
- Emsley, J., King, S. L., Bergelson, J. M., Liddington, R. C. (1997). Crystal structure of the I domain from integrin $\alpha 2\beta 1$. *J Biol Chem* 272, 28512-28517.
- Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., Liddington, R. C. (2000). Structural basis of collagen recognition by integrin $\alpha 2\beta 1$. *Cell* 101, 47-56.
- Emsley, P., Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- Freedman, S. J., Sanford, D. G., Bachovchin, W. W., Furie, B. C., Baleja, J. D., Furie, B. (1996). Structure and function of the epidermal growth factor domain of P-selectin. *Biochemistry* 35, 13733-13744.
- Gahmberg, C. G., Tolvanen, M., Kotovuori, P. (1997). Leukocyte adhesion--structure and function of human leukocyte $\beta 2$ -integrins and their cellular ligands. *Eur J Biochem* 245, 215-232.
- Garcia-Alvarez, B., de Pereda, J. M., Calderwood, D. A., Ulmer, T. S., Critchley, D., Campbell, I. D., Ginsberg, M. H., Liddington, R. C. (2003). Structural determinants of integrin recognition by talin. *Mol Cell* 11, 49-58.
- Ginsberg, M. H., Partridge, A., Shattil, S. J. (2005). Integrin regulation. *Curr Opin Cell Biol* 17, 509-516.
- Gupta, G. P., Massague, J. (2006). Cancer metastasis: building a framework. *Cell* 127, 679-695.
- Gupta, V., Gylling, A., Alonso, J. L., Sugimori, T., Ianakiev, P., Xiong, J. P., Arnaout, M. A. (2007). The beta-tail domain (betaTD) regulates physiologic ligand binding to integrin CD11b/CD18. *Blood* 109, 3513-3520.
- Heras, B., Martin, J. L. (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr* 61, 1173-1180.

References

- Hogervorst, F., Kuikman, I., von dem Borne, A. E., Sonnenberg, A. (1990). Cloning and sequence analysis of beta-4 cDNA: an integrin subunit that contains a unique 118 kd cytoplasmic domain. *Embo J* 9, 765-770.
- Huai, Q., Mazar, A. P., Kuo, A., Parry, G. C., Shaw, D. E., Callahan, J., Li, Y., Yuan, C., Bian, C., Chen, L., *et al.* (2006). Structure of human urokinase plasminogen activator in complex with its receptor. *Science* 311, 656-659.
- Huang, C., Lu, C., Springer, T. A. (1997a). Folding of the conserved domain but not of flanking regions in the integrin beta2 subunit requires association with the alpha subunit. *Proc Natl Acad Sci U S A* 94, 3156-3161.
- Huang, C., Springer, T. A. (1997b). Folding of the beta-propeller domain of the integrin alphaL subunit is independent of the I domain and dependent on the beta2 subunit. *Proc Natl Acad Sci U S A* 94, 3162-3167.
- Hughes, P. E., Diaz-Gonzalez, F., Leong, L., Wu, C., McDonald, J. A., Shattil, S. J., Ginsberg, M. H. (1996). Breaking the integrin hinge. A defined structural constraint regulates integrin signaling. *J Biol Chem* 271, 6571-6574.
- Humphries, M. J. (2000). Integrin structure. *Biochem Soc Trans* 28, 311-339.
- Hynes, R. O. (2002). Integrins: bidirectional, allosteric signaling machines. *Cell* 110, 673-687.
- Iwasaki, K., Mitsuoka, K., Fujiyoshi, Y., Fujisawa, Y., Kikuchi, M., Sekiguchi, K., Yamada, T. (2005). Electron tomography reveals diverse conformations of integrin alphaIIb beta3 in the active state. *J Struct Biol* 150, 259-267.
- Jones, T. A., Zou, J. Y., Cowan, S. W., Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47 (Pt 2), 110-119.
- Jordan, P. A., Stevens, J. M., Hubbard, G. P., Barrett, N. E., Sage, T., Authi, K. S., Gibbins, J. M. (2005). A role for the thiol isomerase protein ERP5 in platelet function. *Blood* 105, 1500-1507.
- Kallen, J., Welzenbach, K., Ramage, P., Geyl, D., Kriwacki, R., Legge, G., Cottens, S., Weitz-Schmidt, G., Hommel, U. (1999). Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain. *J Mol Biol* 292, 1-9.
- Kamata, T., Tieu, K. K., Irie, A., Springer, T. A., Takada, Y. (2001). Amino acid residues in the alpha IIb subunit that are critical for ligand binding to integrin alpha IIb beta 3 are clustered in the beta-propeller model. *J Biol Chem* 276, 44275-44283.

References

- Kamata, T., Wright, R., Takada, Y. (1995). Critical threonine and aspartic acid residues within the I domains of beta 2 integrins for interactions with intercellular adhesion molecule 1 (ICAM-1) and C3bi. *J Biol Chem* 270, 12531-12535.
- Kim, M., Carman, C. V., Springer, T. A. (2003). Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science* 301, 1720-1725.
- Kishimoto, T. K., O'Connor, K., Lee, A., Roberts, T. M., Springer, T. A. (1987). Cloning of the beta subunit of the leukocyte adhesion proteins: homology to an extracellular matrix receptor defines a novel supergene family. *Cell* 48, 681-690.
- Larson, R. S., Hibbs, M. L., Springer, T. A. (1990a). The leukocyte integrin LFA-1 reconstituted by cDNA transfection in a nonhematopoietic cell line is functionally active and not transiently regulated. *Cell Regul* 1, 359-367.
- Larson, R. S., Springer, T. A. (1990b). Structure and function of leukocyte integrins. *Immunol Rev* 114, 181-217.
- Laskowski, R. A., Moss, D. S., Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231, 1049-1067.
- LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., McCoy, J. M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Biotechnology (N Y)* 11, 187-193.
- Law, S. K., Gagnon, J., Hildreth, J. E., Wells, C. E., Willis, A. C., Wong, A. J. (1987). The primary structure of the beta-subunit of the cell surface adhesion glycoproteins LFA-1, CR3 and p150,95 and its relationship to the fibronectin receptor. *Embo J* 6, 915-919.
- Lee, J. O., Bankston, L. A., Arnaout, M. A., Liddington, R. C. (1995a). Two conformations of the integrin A-domain (I-domain): a pathway for activation? *Structure* 3, 1333-1340.
- Lee, J. O., Rieu, P., Arnaout, M. A., Liddington, R. (1995b). Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* 80, 631-638.
- Legge, G. B., Kriwacki, R. W., Chung, J., Hommel, U., Ramage, P., Case, D. A., Dyson, H. J., Wright, P. E. (2000). NMR solution structure of the inserted domain of human leukocyte function associated antigen-1. *J Mol Biol* 295, 1251-1264.
- Leitinger, B., Hogg, N. (2000). Effects of I domain deletion on the function of the beta2 integrin lymphocyte function-associated antigen-1. *Mol Biol Cell* 11, 677-690.
- Leslie, A. G. (2006). The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62, 48-57.

References

- Li, Y., Zhang, L. (2003). The fourth blade within the beta-propeller is involved specifically in C3bi recognition by integrin alpha M beta 2. *J Biol Chem* 278, 34395-34402.
- Li, Y. F., Tang, R. H., Puan, K. J., Law, S. K., Tan, S. M. (2007). The cytosolic protein talin induces an intermediate affinity integrin alpha Lbeta 2. *J Biol Chem*.
- Liddington, R. C., Ginsberg, M. H. (2002). Integrin activation takes shape. *J Cell Biol* 158, 833-839.
- Love, C. A., Harlos, K., Mavaddat, N., Davis, S. J., Stuart, D. I., Jones, E. Y., Esnouf, R. M. (2003). The ligand-binding face of the semaphorins revealed by the high-resolution crystal structure of SEMA4D. *Nat Struct Biol* 10, 843-848.
- Lu, C., Ferzly, M., Takagi, J., Springer, T. A. (2001a). Epitope mapping of antibodies to the C-terminal region of the integrin beta 2 subunit reveals regions that become exposed upon receptor activation. *J Immunol* 166, 5629-5637.
- Lu, C., Shimaoka, M., Ferzly, M., Oxvig, C., Takagi, J., Springer, T. A. (2001b). An isolated, surface-expressed I domain of the integrin alphaLbeta2 is sufficient for strong adhesive function when locked in the open conformation with a disulfide bond. *Proc Natl Acad Sci U S A* 98, 2387-2392.
- Lu, C., Takagi, J., Springer, T. A. (2001c). Association of the membrane proximal regions of the alpha and beta subunit cytoplasmic domains constrains an integrin in the inactive state. *J Biol Chem* 276, 14642-14648.
- Lu, C. F., Springer, T. A. (1997). The alpha subunit cytoplasmic domain regulates the assembly and adhesiveness of integrin lymphocyte function-associated antigen-1. *J Immunol* 159, 268-278.
- Luo, B. H., Carman, C. V., Springer, T. A. (2007). Structural basis of integrin regulation and signaling. *Annu Rev Immunol* 25, 619-647.
- Luo, B. H., Springer, T. A., Takagi, J. (2003). Stabilizing the open conformation of the integrin headpiece with a glycan wedge increases affinity for ligand. *Proc Natl Acad Sci U S A* 100, 2403-2408.
- Luo, B. H., Springer, T. A., Takagi, J. (2004a). A specific interface between integrin transmembrane helices and affinity for ligand. *PLoS Biol* 2, e153.
- Luo, B. H., Strokovich, K., Walz, T., Springer, T. A., Takagi, J. (2004b). Allosteric beta1 integrin antibodies that stabilize the low affinity state by preventing the swing-out of the hybrid domain. *J Biol Chem* 279, 27466-27471.

References

- Luo, B. H., Takagi, J., Springer, T. A. (2004c). Locking the beta3 integrin I-like domain into high and low affinity conformations with disulfides. *J Biol Chem* 279, 10215-10221.
- Mancia, F., Patel, S. D., Rajala, M. W., Scherer, P. E., Nemes, A., Schieren, I., Hendrickson, W. A., Shapiro, L. (2004). Optimization of protein production in mammalian cells with a coexpressed fluorescent marker. *Structure* 12, 1355-1360.
- McGuire, S. L., Bajt, M. L. (1995). Distinct ligand binding sites in the I domain of integrin alpha M beta 2 that differentially affect a divalent cation-dependent conformation. *J Biol Chem* 270, 25866-25871.
- Michishita, M., Videm, V., Arnaout, M. A. (1993). A novel divalent cation-binding site in the A domain of the beta 2 integrin CR3 (CD11b/CD18) is essential for ligand binding. *Cell* 72, 857-867.
- Mould, A. P., Barton, S. J., Askari, J. A., Craig, S. E., Humphries, M. J. (2003a). Role of ADMIDAS cation-binding site in ligand recognition by integrin alpha 5 beta 1. *J Biol Chem* 278, 51622-51629.
- Mould, A. P., Barton, S. J., Askari, J. A., McEwan, P. A., Buckley, P. A., Craig, S. E., Humphries, M. J. (2003b). Conformational changes in the integrin beta A domain provide a mechanism for signal transduction via hybrid domain movement. *J Biol Chem* 278, 17028-17035.
- Nachon, F., Nicolet, Y., Viguie, N., Masson, P., Fontecilla-Camps, J. C., Lockridge, O. (2002). Engineering of a monomeric and low-glycosylated form of human butyrylcholinesterase: expression, purification, characterization and crystallization. *Eur J Biochem* 269, 630-637.
- Nermut, M. V., Green, N. M., Eason, P., Yamada, S. S., Yamada, K. M. (1988). Electron microscopy and structural model of human fibronectin receptor. *Embo J* 7, 4093-4099.
- Nishida, N., Xie, C., Shimaoka, M., Cheng, Y., Walz, T., Springer, T. A. (2006). Activation of leukocyte beta2 integrins by conversion from bent to extended conformations. *Immunity* 25, 583-594.
- Nolte, M., Pepinsky, R. B., Venyaminov, S., Kotliansky, V., Gotwals, P. J., Karpusas, M. (1999). Crystal structure of the alpha1beta1 integrin I-domain: insights into integrin I-domain function. *FEBS Lett* 452, 379-385.
- O'Neill, S., Robinson, A., Deering, A., Ryan, M., Fitzgerald, D. J., Moran, N. (2000). The platelet integrin alpha IIbeta 3 has an endogenous thiol isomerase activity. *J Biol Chem* 275, 36984-36990.

References

- O'Toole, T. E., Katagiri, Y., Faull, R. J., Peter, K., Tamura, R., Quaranta, V., Loftus, J. C., Shattil, S. J., Ginsberg, M. H. (1994). Integrin cytoplasmic domains mediate inside-out signal transduction. *J Cell Biol* 124, 1047-1059.
- O'Toole, T. E., Mandelman, D., Forsyth, J., Shattil, S. J., Plow, E. F., Ginsberg, M. H. (1991). Modulation of the affinity of integrin alpha IIb beta 3 (GPIIb-IIIa) by the cytoplasmic domain of alpha IIb. *Science* 254, 845-847.
- Oxvig, C., Springer, T. A. (1998). Experimental support for a beta-propeller domain in integrin alpha-subunits and a calcium binding site on its lower surface. *Proc Natl Acad Sci U S A* 95, 4870-4875.
- Partridge, A. W., Liu, S., Kim, S., Bowie, J. U., Ginsberg, M. H. (2005). Transmembrane domain helix packing stabilizes integrin alphaIIbbeta3 in the low affinity state. *J Biol Chem* 280, 7294-7300.
- Patel, S. D., Rajala, M. W., Rossetti, L., Scherer, P. E., Shapiro, L. (2004). Disulfide-dependent multimeric assembly of resistin family hormones. *Science* 304, 1154-1158.
- Perrakis, A., Sixma, T. K., Wilson, K. S., Lamzin, V. S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr* 53, 448-455.
- Pribila, J. T., Quale, A. C., Mueller, K. L., Shimizu, Y. (2004). Integrins and T cell-mediated immunity. *Annu Rev Immunol* 22, 157-180.
- Qu, A., Leahy, D. J. (1995). Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin. *Proc Natl Acad Sci U S A* 92, 10277-10281.
- Qu, A., Leahy, D. J. (1996). The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. *Structure* 4, 931-942.
- Reeves, P. J., Callewaert, N., Contreras, R., Khorana, H. G. (2002). Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc Natl Acad Sci U S A* 99, 13419-13424.
- Rich, R. L., Deivanayagam, C. C., Owens, R. T., Carson, M., Hook, A., Moore, D., Symersky, J., Yang, V. W., Narayana, S. V., Hook, M. (1999). Trench-shaped binding sites promote multiple classes of interactions between collagen and the adherence receptors, alpha(1)beta(1) integrin and Staphylococcus aureus cna MSCRAMM. *J Biol Chem* 274, 24906-24913.

References

- Robinson, M. K., Andrew, D., Rosen, H., Brown, D., Orllepp, S., Stephens, P., Butcher, E. C. (1992). Antibody against the Leu-CAM beta-chain (CD18) promotes both LFA-1- and CR3-dependent adhesion events. *J Immunol* 148, 1080-1085.
- Shaw, J. M., Al-Shamkhani, A., Boxer, L. A., Buckley, C. D., Dodds, A. W., Klein, N., Nolan, S. M., Roberts, I., Roos, D., Scarth, S. L., *et al.* (2001). Characterization of four CD18 mutants in leucocyte adhesion deficient (LAD) patients with differential capacities to support expression and function of the CD11/CD18 integrins LFA-1, Mac-1 and p150,95. *Clin Exp Immunol* 126, 311-318.
- Shi, M., Sundramurthy, K., Liu, B., Tan, S. M., Law, S. K., Lescar, J. (2005). 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* 280, 30586-30593.
- Shimaoka, M., Lu, C., Palframan, R. T., von Andrian, U. H., McCormack, A., Takagi, J., Springer, T. A. (2001). Reversibly locking a protein fold in an active conformation with a disulfide bond: integrin alphaL I domains with high affinity and antagonist activity in vivo. *Proc Natl Acad Sci U S A* 98, 6009-6014.
- Shimaoka, M., Takagi, J., Springer, T. A. (2002). Conformational regulation of integrin structure and function. *Annu Rev Biophys Biomol Struct* 31, 485-516.
- Shimaoka, M., Xiao, T., Liu, J. H., Yang, Y., Dong, Y., Jun, C. D., McCormack, A., Zhang, R., Joachimiak, A., Takagi, J., *et al.* (2003). Structures of the alpha L I domain and its complex with ICAM-1 reveal a shape-shifting pathway for integrin regulation. *Cell* 112, 99-111.
- Simon, D. I., Wei, Y., Zhang, L., Rao, N. K., Xu, H., Chen, Z., Liu, Q., Rosenberg, S., Chapman, H. A. (2000). Identification of a urokinase receptor-integrin interaction site. Promiscuous regulator of integrin function. *J Biol Chem* 275, 10228-10234.
- Song, G., Yang, Y., Liu, J. H., Casasnovas, J. M., Shimaoka, M., Springer, T. A., Wang, J. H. (2005). An atomic resolution view of ICAM recognition in a complex between the binding domains of ICAM-3 and integrin alphaLbeta2. *Proc Natl Acad Sci U S A* 102, 3366-3371.
- Springer, T. A. (1997). Folding of the N-terminal, ligand-binding region of integrin alpha-subunits into a beta-propeller domain. *Proc Natl Acad Sci U S A* 94, 65-72.
- Springer, T. A., Wang, J. H. (2004). The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion. *Adv Protein Chem* 68, 29-63.
- Stamos, J., Lazarus, R. A., Yao, X., Kirchhofer, D., Wiesmann, C. (2004). Crystal structure of the HGF beta-chain in complex with the Sema domain of the Met receptor. *Embo J* 23, 2325-2335.

References

- Stanley, P., Bates, P. A., Harvey, J., Bennett, R. I., Hogg, N. (1994). Integrin LFA-1 alpha subunit contains an ICAM-1 binding site in domains V and VI. *Embo J* 13, 1790-1798.
- Staunton, D. E., Lupher, M. L., Liddington, R., Gallatin, W. M. (2006). Targeting integrin structure and function in disease. *Adv Immunol* 91, 111-157.
- Stefansson, A., Armulik, A., Nilsson, I., von Heijne, G., Johansson, S. (2004). Determination of N- and C-terminal borders of the transmembrane domain of integrin subunits. *J Biol Chem* 279, 21200-21205.
- Storoni, L. C., McCoy, A. J., Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60, 432-438.
- Takada, Y., Ye, X., Simon, S. (2007). The integrins. *Genome Biol* 8, 215.
- Takagi, J., Beglova, N., Yalamanchili, P., Blacklow, S. C., Springer, T. A. (2001a). Definition of EGF-like, closely interacting modules that bear activation epitopes in integrin beta subunits. *Proc Natl Acad Sci U S A* 98, 11175-11180.
- Takagi, J., DeBottis, D. P., Erickson, H. P., Springer, T. A. (2002a). The role of the specificity-determining loop of the integrin beta subunit I-like domain in autonomous expression, association with the alpha subunit, and ligand binding. *Biochemistry* 41, 4339-4347.
- Takagi, J., Erickson, H. P., Springer, T. A. (2001b). C-terminal opening mimics 'inside-out' activation of integrin alpha5beta1. *Nat Struct Biol* 8, 412-416.
- Takagi, J., Kamata, T., Meredith, J., Puzon-McLaughlin, W., Takada, Y. (1997). Changing ligand specificities of alphavbeta1 and alphavbeta3 integrins by swapping a short diverse sequence of the beta subunit. *J Biol Chem* 272, 19794-19800.
- Takagi, J., Petre, B. M., Walz, T., Springer, T. A. (2002b). Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* 110, 599-511.
- Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horwitz, A. F., Hynes, R. O. (1986). Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell* 46, 271-282.
- Tan, S. M., Robinson, M. K., Drbal, K., van Kooyk, Y., Shaw, J. M., Law, S. K. (2001a). The N-terminal region and the mid-region complex of the integrin beta 2 subunit. *J Biol Chem* 276, 36370-36376.

References

- Tan, S. M., Walters, S. E., Mathew, E. C., Robinson, M. K., Drbal, K., Shaw, J. M., Law, S. K. (2001b). Defining the repeating elements in the cysteine-rich region (CRR) of the CD18 integrin beta 2 subunit. *FEBS Lett* 505, 27-30.
- Tang, R. H., Tng, E., Law, S. K., Tan, S. M. (2005). Epitope mapping of monoclonal antibody to integrin alphaL beta2 hybrid domain suggests different requirements of affinity states for intercellular adhesion molecules (ICAM)-1 and ICAM-3 binding. *J Biol Chem* 280, 29208-29216.
- Terwilliger, T. C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* 374, 22-37.
- Terwilliger, T. C., Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* 55, 849-861.
- Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-4680.
- Tng, E., Tan, S. M., Ranganathan, S., Cheng, M., Law, S. K. (2004). The integrin alpha L beta 2 hybrid domain serves as a link for the propagation of activation signal from its stalk regions to the I-like domain. *J Biol Chem* 279, 54334-54339.
- Tuckwell, D. S., Humphries, M. J. (1997). A structure prediction for the ligand-binding region of the integrin beta subunit: evidence for the presence of a von Willebrand factor A domain. *FEBS Lett* 400, 297-303.
- Ulmer, T. S., Yaspan, B., Ginsberg, M. H., Campbell, I. D. (2001). NMR analysis of structure and dynamics of the cytosolic tails of integrin alpha IIb beta 3 in aqueous solution. *Biochemistry* 40, 7498-7508.
- van der Geer, P., Wiley, S., Lai, V. K., Olivier, J. P., Gish, G. D., Stephens, R., Kaplan, D., Shoelson, S., Pawson, T. (1995). A conserved amino-terminal Shc domain binds to phosphotyrosine motifs in activated receptors and phosphopeptides. *Curr Biol* 5, 404-412.
- Vinogradova, O., Velyvis, A., Velyviene, A., Hu, B., Haas, T., Plow, E., Qin, J. (2002). A structural mechanism of integrin alpha(IIb)beta(3) "inside-out" activation as regulated by its cytoplasmic face. *Cell* 110, 587-597.
- Vorup-Jensen, T., Ostermeier, C., Shimaoka, M., Hommel, U., Springer, T. A. (2003). Structure and allosteric regulation of the alpha X beta 2 integrin I domain. *Proc Natl Acad Sci U S A* 100, 1873-1878.

References

- Wegener, K. L., Partridge, A. W., Han, J., Pickford, A. R., Liddington, R. C., Ginsberg, M. H., Campbell, I. D. (2007). Structural basis of integrin activation by talin. *Cell* *128*, 171-182.
- Wei, Y., Czekay, R. P., Robillard, L., Kugler, M. C., Zhang, F., Kim, K. K., Xiong, J. P., Humphries, M. J., Chapman, H. A. (2005). Regulation of alpha5beta1 integrin conformation and function by urokinase receptor binding. *J Cell Biol* *168*, 501-511.
- Winn, M. D., Murshudov, G. N., Papiz, M. Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* *374*, 300-321.
- Xiao, T., Takagi, J., Collier, B. S., Wang, J. H., Springer, T. A. (2004). Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* *432*, 59-67.
- Xie, C., Shimaoka, M., Xiao, T., Schwab, P., Klickstein, L. B., Springer, T. A. (2004). The integrin alpha-subunit leg extends at a Ca²⁺-dependent epitope in the thigh/genu interface upon activation. *Proc Natl Acad Sci U S A* *101*, 15422-15427.
- Xiong, J. P., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D. L., Joachimiak, A., Goodman, S. L., Arnaout, M. A. (2001). Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* *294*, 339-345.
- Xiong, J. P., Stehle, T., Goodman, S. L., Arnaout, M. A. (2003a). Integrins, cations and ligands: making the connection. *J Thromb Haemost* *1*, 1642-1654.
- Xiong, J. P., Stehle, T., Goodman, S. L., Arnaout, M. A. (2003b). New insights into the structural basis of integrin activation. *Blood* *102*, 1155-1159.
- Xiong, J. P., Stehle, T., Goodman, S. L., Arnaout, M. A. (2004). A novel adaptation of the integrin PSI domain revealed from its crystal structure. *J Biol Chem* *279*, 40252-40254.
- Xiong, J. P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S. L., Arnaout, M. A. (2002). Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* *296*, 151-155.
- Yalamanchili, P., Lu, C., Oxvig, C., Springer, T. A. (2000). Folding and function of I domain-deleted Mac-1 and lymphocyte function-associated antigen-1. *J Biol Chem* *275*, 21877-21882.
- Yang, W., Shimaoka, M., Salas, A., Takagi, J., Springer, T. A. (2004). Intersubunit signal transmission in integrins by a receptor-like interaction with a pull spring. *Proc Natl Acad Sci U S A* *101*, 2906-2911.
- Zhu, J., Boylan, B., Luo, B. H., Newman, P. J., Springer, T. A. (2007). Tests of the extension and deadbolt models of integrin activation. *J Biol Chem* *282*, 11914-11920

Appendix 1: The DNA sequence of integrin $\beta 2$ subunit (cd18) and its amino acid sequence with the signal peptide highlighted in *red*

ATGCATAAGCTCTAGACGTGACCCAGGGCAGACTGGTAGCAAAGCCCCACGCCAGCCAGGAGCACCGCCGAGGACTCCAGCACACCG
 -90 XbaI -70 -60 -50 -40 -30 -20 -10
 TACGTATTCTGAGATCTGCACTGGGTCCCCTCTGACCATCGTTTCGGGGGTGCGGGTGGTCTCTGTCGGGCTCTGAGGTCGTGTGGC

-20 -10 1
M L G L R P P L L A L V G L L S L G C V L S Q E C1 T K F
 AG GGACATGCTGGGCCTGCGCCCCCACTGCTCGCCCTGGTGGGGCTGCTCTCCCTCGGGTGC GTCCTCTCTCAGGAGTGCACGAAGTTC
 1|2 1 10 20 30 40 50 2|3 70 80
 TC CCTGTACGACCCGGACGCGGGGGGTGACGAGCGGGACCACCCGACGAGAGGGAGCCACG CAGGAGAGAGTCTCACGTGCTTCAAG

10 20 30
 K V S S C2 R E C3 I E S G P G C4 T W C5 Q K L N F T G P G D P D
 AAGTCAGCAGCTGCCGGGAATGCATCGAGTCGGGGCCCGGCTGCACCTGGTGCAGAAAGCTG AACTTCACAGGGCCGGGGATCCTGAC
 90 100 110 120 130 140 3|4 160 BamHI
 TTCCAGTCGTCGACGGCCCTTACGTAGCTCAGCCCCGGGCGACGTGGACCACGGTCTTTCGAC TTGAAGTGTCCCGCCCCCTAGGACTG

40 50 60
 S I R C6 D T R P Q L L M R G C7 A A D D I M D P T S L A E T Q
 TCCATTCTGCTGCGACACCCGGCCACAGCTGCTCATGAGGGGCTGTGCGGCTGACGACATCATGGACCCACAAGCCTCGCTGAAACCCAG
 180 190 200 210 220 230 240 250 260
 AGGTAAGCGACGCTGTGGGCCGGTGTGACGAGTACTCCCCGACACGCCGACTGCTGTAGTACCTGGGGTGTTCGGAGCGACTTTGGGTC

70 80 90
 E D H N G G Q K Q L S P Q K V T L Y L R P G Q A A A F N V T
 GAAGACCACAATGGGGGCCAGAAGCAGCTGTCCCCACAAAAAGTGACGCTTTACCTGCGACCAG GCCAGGCAGCAGCGTTCAACGTGACC
 270 280 290 300 310 320 4|5 340 350
 CTTCTGGTGTACCCCCGGTCTTCGTCGACAGGGGTGTTTTTCACTGCGAAATGGACGCTGGTTC CGGTCCGTCGTCGCAAGTTGCACTGG

100 110 120
 F R R A K G Y P I D L Y Y L M D L S Y S M L D D L R N V K K
 TTCCGGCGGGCCAAAGGGCTACCCCATCGACCTGTACTATCTGATGGACCTCTCCTACTCCATGCTTGATGACCTCAGGAATGTCAAGAAG
 360 StyI 380 390 400 410 420 Bsu36I 440
 AAGGCCGCCCGGTTCCCGATGGGGTAGCTGGACATGATAGACTACCTGGAGAGGATGAGGTACGAACTACTGGAGTCTTACAGTTCTTC

130 140 150
 L G G D L L R A L N E I T E S G R I G F G S F V D K T V L P
 CTAGGTGGCGACCTGCTCCGGGCCCTCAACGAGATCACCGAGTCCGGCCGATTG GCTTCGGGTCTTCGTGGACAAGACCGTGCTGCCG
 450 460 470 480 490 5|6 510 520 530
 GATCCACCGCTGGACGAGGCCCGGGAGTTGCTCTAGTGGCTCAGGCCGGCGTAAC CGAAGCCCAGGAAGCACCTGTTCTGGCACGACGGC

160 170 180
 F V N T H P D K L R N P C8 P N K E K E C9 Q P P F A F R H V L
 TTCGTGAACACGACCCCTGATAAGCTGCGAAACCCATGCCCCAACAAGGAGAAAGAGTGCCAGCCCCCGTTTGCTTCAGGCACGTGCTG
 540 550 560 570 580 590 600 610 620
 AAGCACTTGTGCGTGGGACTATTTCGACGCTTTGGGTACGGGGTGTTCCTCTTCTCACGGTCGGGGGCAAACGGAAGTCCGTGCACGAC

Appendix

190 200 210
 K L T N N S N Q F Q T E V G K Q L I S G N L D A P E G G L D
 AAGCTGACCAACAACCTCCAACCAAGTTTCAGACCGAGGTCGGGAAGCAGCTGATTTCGGAAACCTGGATGCACCCGAGGGTGGGCTGGAC
 630 640 650 660 670 BspEI 690 700 710
 TTCGACTGGTTGTTGAGGTTGGTCAAAGTCTGGCTCCAGCCCTTCGTCGACTAAAGGCCTTTGGACCTACGTGGGCTCCCACCCGACCTG

220 230 240
 A M M Q V A A C10 P E E I G W R N V T R L L V F A T D D G F H
 GCCATGATGCAGGTCGCCGCTGCCCG GAGGAAATCGGCTGGCGCAACGTCACGCGGCTGCTGGTGTTCGCACTGATGACGGCTTCCAT
 720 730 6|7 750 760 770 780 790 800
 CGGTACTACGTCCAGCGGCGGACGGGC CTCCTTTAGCCGACCGGCTTGCACTGCGCCGACGACCACAAACGGTGACTACTGCCGAAGGTA

250 260 270
 F A G D G K L G A I L T P N D G R C11 H L E D N L Y K R S N E
 TTCGCGGGCGACGAAAGCTGGGCGCCATCCTGACCCCAACGCGCCGCTGTACCTGGAGGACAACCTTGTACAAGAGGAGCAACGAA
 810 820 830 840 850 860 870 880 890 EcoRI
 AAGCGCCCGCTGCCTTTTCGACCCGCGGTAGGACTGGGGGTGTGCGCGCGACAGTGGACCTCCTGTTGAACATGTTCTCCTCGTTGCTT

280 290 300
 F D Y P S V G Q L A H K L A E N N I Q P I F A V T S R M V K
 TTC GACTACCCATCGGTGGGCCAGCTGGCGCACAAGCTGGCTGAAAACAACATCCAGCCCATCTTCGCGGTGACCAGTAGGATGGTGAAG
 7|8 910 920 930 940 950 960 970 980
 AAG CTGATGGGTAGCCACCCGGTCGACCGCGTGTTCGACCGACTTTTGTGTGAGGTGCGGTAGAAGCGCCACTGGTCATCCTACCACTTC

310 320 330
 T Y E K L T E I I P K S A V G E L S E D S S N V V H L I K N
 ACC TACGAGAACTCACCGAGATCATCCCCAAGTCAGCCGTGGGGGAGCTGTCTGAGGACTCCAGCAATGTGGTCCATCTCATTAAGAAT
 8|9 1000 1010 1020 1030 1040 1050 1060 1070
 TGG ATGCTCTTTGAGTGGCTCTAGTAGGGGTTTCAGTCGGCACCCCTCGACAGACTCCTGAGGTCGTTACACCAGGTAGAGTAATCTTA

340 350 360
 A Y N K L S S R V F L D H N A L P D T L K V T Y D S F C12 S N
 GCTTACAAT AAATCTCCTCCAGGGTCTTCCTGGATCACAACGCCCTCCCCGACACCCTGAAAGTCACCTACGACTCCTTCTGCAGCAAT
 BsmI 9|10 1100 1110 1120 1130 1140 1150 1160
 CGAATGTTA TTTGAGAGGAGTCCCAGAAGGACCTAGTGTTCGGGAGGGGCTGTGGGACTTTCAGTGGATGCTGAGGAAGACGTCGTTA

370 380 390
 G V T H R N Q P R G D C13 D G V Q I N V P I T F Q V K V T A T
 GGAGTGACGCACAGGAACCCAGCCAGAGGTGACTGTGATGGCGTGCAGATCAATGTCCCG ATCACCTTCCAGGTGAAGGTCACGGCCACA
 1170 1180 1190 1200 1210 1220 10|11 1240 1250
 CCTCACTGCGTGTCTTGGTTCGGGTCTCCACTGACACTACCGCACGTCTAGTTACAGGC TAGTGGAAGGTCCACTTCCAGTGCCGGTGT

400 410 420
 E C14 I Q E Q S F V I R A L G F T D I V T V Q V L P Q C15 E C16 R
 GAGTGCATCCAGGAGCAGTCGTTTGTTCATCCGGGCGCTGGGCTTCACGGACATAGTGACCGTGCAGGTCCTTCCCCAGTGTGAGTGCCGG
 1260 1270 1280 1290 1300 1310 1320 1330 1340
 CTCACGTAGGTCTCGTCAGCAAACAGTAGGCCCGCAGCCGAAGTGCCTGTATCACTGGCACGTCCAGGAAGGGGTCACACTCACGGCC

430 440 450
 C17 R D Q S R D R S L C18 H G K G F L E C19 G I C20 R C21 D T G Y I G
 TGCCGGGACCAGAGCAGAGACCGCAGCCTCTGCCATGGCAAGGCTTCTTGGAGTGGCGCATCTGCAG GTGTGACACTGGCTACATTGGG
 1350 1360 1370 NcoI 1390 1400 11|12 1420 1430
 ACGGCCCTGGTCTCGTCTCTGGCGTCGGAGACGGTACCGTTCCCGAAGAACCCTACGCCGTAGACGTC CACACTGTGACCGATGTAACCC

Appendix

460 470 480
 K N C22 E C23 Q T Q G R S S Q E L E G S C24 R K D N N S I I C25 S G
 AAAAAGTGTGAGTGCCAGACACAGGGCCGGAGCAGCCAGGAGCTGGAAGGAAGCTGCCGGAAGGACAACAAGTCCATCATCTGCTCAGGG
 1440 1450 1460 1470 1480 1490 1500 1510 1520
 TTTTGTGACACTCACGGTCTGTGTCCCGCCTCGTCGGTCTCGACCTTCCTTCGACGGCCTTCCTGTTGTTGAGGTAGTAGACGAGTCCC

490 500 510
 L G D C26 V C27 G Q C28 L C29 H T S D V P G K L I Y G Q Y C30 E C31 D T
 CTGGGGGACTGTGTCTGCGGGCAGTGCCCTGTGCCACACCAGCGACGTCCCCGGCAAGCTGATATACGGGCAGTACTGCGAGTGTGACACC
 1530 1540 1550 1560 AatII 1580 1590 1600 1610
 GACCCCTGACACAGACGCCCCGTACGGACACGGTGTGGTTCGCTGCAGGGGCCGTTCTGACTATATGCCCGTCATGACGCTCAGCTCAGTGTGG

520 530 540
 I N C32 E R Y N G Q V C33 G G P G R G L C34 F C35 G K C36 R C37 H P G F
 ATCAAGTGTGAGCGCTACAACGGCCAGGTCTGCGGCGGCCCGG GGAGGGGGCTCTGCTTCTGCGGGAAGTGCCGCTGCCACCCGGGCTTT
 1620 1630 1640 1650 12|13 1670 1680 1690 1700
 TAGTTGACACTCGCGATGTTGCCGGTCCAGACGCCGCCGGGCC CCTCCCCGAGACGAAGACGCCCTTCACGGCGACGGTGGGCCCCGAAA

550 560 570
 E G S A C38 Q C39 E R T T E G C40 L N P R R V E C41 S G R G R C42 R C43
 GAGGGCTCAGCGTGCCAGTGCGAGAGGACCACTGAGGGCTGCCTGAACCCGCGGCGTGTGAGTGTAGTGGTCGTGGCCGGTGCCGCTGC
 1710 1720 1730 1740 1750 SacII 1770 1780 1790
 CTCCCGAGTCGCACGGTCACGCTCTCCTGGTGACTCCCGACGGACTTGGGCGCCGCACAACCTCACATCACCAGACCCGGCCACGGCGACG

580 590 600
 N V C44 E C45 H S G Y Q L P L C46 Q E C47 P G C48 P S P C49 G K Y I S C50
 AACGTATGCGAGTGCCATTACAGGTACCAGCTGCCTCTGTGCCAGGAGTGCCCCGGCTGCCCCCTACCCCTGTGGCAAGTACAT CTCCTGC
 1800 1810 1820 1830 1840 1850 1860 1870 13|14
 TTGCATACGCTCACGGTAAGTCCGATGGTTCGACGGAGACACGGTCTCACGGGGCCGACGGGGAGTGGGACACCGTTTCATGTA GAGGACG

610 620 630
 A E C51 L K F E K G P F G K N C52 S A A C53 P G L Q L S N N P V K
 GCCGAGTGCCTGAAGTTTCGAAAAGGGCCCCCTTTGGGAAGAACTGCAGCGCGGCGTGTCCGGGCCCTGCAGCTGTGCAACAACCCCGTGAAG
 1890 BstBI 1910 1920 1930 1940 1950 1960 1970
 CGGCTCACGGACTTCAAGCTTTTCCCGGGGAAACCTTCTTGACGTCGCGCCGCACAGGCCCGGACGTCGACAGCTTGTGTTGGGGCACTTC

640 650 660
 G R T C54 K E R D S E G C55 W V A Y T L E Q Q D G M D R Y L I Y
 GGCAGGACCTGCAAGGAGAGGGACTCAGAGGGCTGTGCGGTGGCTACACGCTGGAGCAGCAGGACGGGATGGACCGCTACCTCATCTAT
 1980 1990 2000 2010 2020 2030 2040 2050 2060
 CCGTCCTGGACGTTCTCTCCCTGAGTCTCCCGACGACCCACCGGATGTGCGACCTCGTCGTCTGCCCCTACCTGGCGATGGAGTAGATA

670 680 690
 V D E S R E C56 V A G P N I A A I V G G T V A G I V L I G I L
 GTGGATGAGAGCCGAG AGTGTGTGGCAGGCCCAACATCGCCGCCATCGTCGGGGGCACCGTGGCAGGCATCGTGCTGATCGGCATTCTC
 2070 14|15 2090 2100 2110 2120 2130 2140 2150
 CACCTACTCTCGGCTC TCACACACCGTCCGGGGTTGTAGCGGCGGTAGCAGCCCCGTGGCACCCTCGTAGCAGACTAGCCGTAAGAG

700 710 720
 L L V I W 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
 CTGCTGGTCATCTGGAAGGCTCTGATCCACCTGAGCGACCTCCGGGAGTACAGGCGCTTTGAGAAGGAGAAGCTCAAGTCCCAGTGGAAAC
 2160 2170 2180 2190 2200 2210 2220 2230 2240
 GACGACCAGTAGACCTTCCGAGACTAGGTGGACTCGCTGGAGGCCCTCATGTCCGCGAAACTCTTCTCTTCGAGTTACAGGTTACCTTG

```

      730                               740
N   D N P L F K S A T T T V M N P K F A E S *
AAT GATAATCCCCTTTTCAAGAGCGCCACCACGACGGTCATGAACCCCAAGTTTGCTGAGAGTTAGGAGCACTTGGTGAAGACAAGGCCG
15|16      2260      2270      2280      2290      2300      2310      2320      2330
TTA CTATTAGGGGAAAAGTTCTCGCGGTGGTGCTGCCAGTACTTGGGGTTCAAACGACTCTCAATCCTCGTGAACCACTTCTGTTCCGGC

TCAGGACCCACCATGTCTGCCCCATCACGCGGCCGAGACATGGCTTGCCACAGCTCTTGAGGATGTCACCAATTAACCAGAAATCCAGTT
2340      2350      2360      2370      2380      2390      2400      2410      2420
AGTCCTGGGTGGTACAGACGGGGTAGTGCGCCGGCTCTGTACCGAACGGTGTCGAGAACTCCTACAGTGGTTAATTGGTCTTTAGGTCAA

ATTTTCCGCCCTCAAAATGACAGCCATGCCTACTAGTGAGCTC
2430      2440      2450      SpeI  SstI
TAAAAGGCGGGAGTTTACTGTTCGGTACGGATGATCACTCGAG

```

Codes

```

XXXXXX:    Unique Restriction Enzyme Sites

X|(X+1):    Exon Boundaries

CX:         Conserved Cysteines from C1 to C56

X:          Single Residue Changes in LAD Patients

XXXXXX:    Leader Sequence/Transmembrane Segment

```

Figure A.1 The map of the pET24a vector for expression his-tag protein expression.

Figure A.2 The map of the pET32b vector for expression trx fusion protein expression.

Figure A.3 The map of the pGEX_5X1 vectors for GST fusion protein expression.

Figure A.4 The map of the pMAL-c2X vector for expression MBP fusion protein expression.

Appendix 3: The recombinant proteins used in this thesis were analyzed by Vector NTI

A: hybrid domain

Analysis	Entire Protein
Length	132 aa
Molecular Weight	14755.90 m.w.
1 microgram =	67.770 pMoles
Molar Extinction coefficient	3040
1 A[280] corr. to	4.85 mg/ml
A[280] of 1 mg/ml	0.21 AU
Isoelectric Point	7.93
Charge at pH 7	1.52

B: PHE1

Analysis	Entire Protein
Length	227 aa
Molecular Weight	25347.47 m.w.
1 microgram =	39.452 pMoles
Molar Extinction coefficient	11690
1 A[280] corr. to	2.17 mg/ml
A[280] of 1 mg/ml	0.46 AU
Isoelectric Point	7.18
Charge at pH 7	0.49

C: PHE2

Analysis	Entire Protein
Length	280 aa
Molecular Weight	31017.45 m.w.
1 microgram =	32.240 pMoles
Molar Extinction coefficient	15210
1 A[280] corr. to	2.04 mg/ml
A[280] of 1 mg/ml	0.49 AU
Isoelectric Point	6.61
Charge at pH 7	-1.50

D: PHE3

Analysis	Entire Protein
Length	319 aa
Molecular Weight	35126.90 m.w.
1 microgram =	28.468 pMoles
Molar Extinction coefficient	17450
1 A[280] corr. to	2.01 mg/ml
A[280] of 1 mg/ml	0.50 AU
Isoelectric Point	6.87
Charge at pH 7	-0.49

Appendix 4: The presumed models of integrin structures

(A) Ribbon diagram of the bent conformation observed in the $\alpha V\beta 3$ crystal structure (PDB, 1U8C) with the I-EGF1 and I-EGF2 domains built using domain overlap superimposition with $\beta 2$ fragments, PHE1 (PDB, 1YUK) and PHE2 (PDB, 2P26), respectively.

(B) The model of the half-extended integrin with closed headpiece by superimposing with the PHE2 structure (2P26) on hybrid domain to fill in PSI and I-EGF1-2 domains, using the hybrid domain from PHE2 structure to replace the original, and superimposing the PHE3 fragment on the I-EGF2 domain, followed by superimposing the fragment containing I-EGF3-4 and β -TD from the bent structure (1U8C) on the I-EGF3 domain, finally, unbending manually at the α subunit Genu between the thigh and the Calf-1 domains to parallel the Calf-1-2 to the I-EGF4 and β -TD.

(C) The model of the fully extended integrin with closed headpiece by superimposing with the PHE3 structure (2P28) on hybrid domain to fill in PSI and I-EGF1-3 domains, the following is generated as in (B).

(D) The model of the extended integrin with open headpiece. The extended α subunit was generated as in (C), and superimposed with the open headpiece (1TY6) on the β -propeller domain, replaced with the β I and hybrid domains from 1TY6 structure and superimposed as in (C). The same color codes for the domains are conserved throughout: blue, β I domain; yellow, hybrid domain; green, PSI domain; red, I-EGF1 domain; magenta, I-EGF2; pink, I-EGF3; cyan, I-EGF4 and β -TD domains; gray, α subunit.

Appendix 5: Human Integrin subunits

Table: Human integrin subunits

Gene symbol	Protein name	Synonyms	Gene accession number	Location (chromosome)	Protein accession number
<i>ITGA1</i>	$\alpha 1$	CD49a	NM_181501	5q11.2	P56199
<i>ITGA2</i>	$\alpha 2$	CD49b, $\alpha 2$ subunit of very late antigen 2 (VLA-2)	NM_002203	5q23-q31	P17301
<i>ITGA2B</i>	αIIb	GTA, CD41, GP2B, HPA3, CD41b, GPIIb	NM_000419	17q21.32	P08514
<i>ITGA3</i>	$\alpha 3$	CD49c, $\alpha 3$ subunit of VLA-3	NM_002204, NM_005501	17q21.33	P26006
<i>ITGA4</i>	$\alpha 4$	CD49d, $\alpha 4$ subunit of VLA-4	NM_000885	2q31.3	AAB25486
<i>ITGA5</i>	$\alpha 5$	CD49e, fibronectin receptor alpha	NM_002205	12q11-q13	P08648
<i>ITGA6</i>	$\alpha 6$	CD49f, ITGA6B	NM_000210	2q31.1	P23229
<i>ITGA7</i>	$\alpha 7$		NM_002206	12q13	Q86W93
<i>ITGA8</i>	$\alpha 8$		NM_003638	10p13	P53708
<i>ITGA9</i>	$\alpha 9$		NM_002207	3p21.3	Q13797
<i>ITGA10</i>	$\alpha 10$		NM_003637	1q21	O75578
<i>ITGA11</i>	$\alpha 11$		NM_001004439, NM_012211	15q23	Q9UKX5
<i>ITGAD</i>	αD		NM_005353	16p11.2	Q13349
<i>ITGAE</i>	αE	CD103, human mucosal lymphocyte antigen 1 α	NM_002208	17p13	P38570
<i>ITGAL</i>	αL	CD11a (p180), lymphocyte function-associated antigen 1 (LFA-1) α subunit	NM_002209	16p11.2	P20701
<i>ITGAM</i>	αM	Mac-1, CD11b, complement receptor 3 (CR3) subunit	J03925, NM_000632,	16p11.2	P11215
<i>ITGAV</i>	αV	CD51, MSK8, vitronectin receptor α (VNR α)	NM_002210	2q31-q32	P06756
<i>ITGAX</i>	αX	CD11c, CR4 subunit	NM_000887	16p11.2	P20702
<i>ITGB1</i>	$\beta 1$	Fibronectin receptor β , CD29, MDF2, MSK12	BC020057	10p11.2	P05556
<i>ITGB2</i>	$\beta 2$	Leukocyte cell adhesion molecule, CD18, CR3 subunit, CR4 subunit	NM_000211	21q22.3	P05107
<i>ITGB3</i>	B3	CD61; GP3A; GPIIIa, platelet glycoprotein IIIa	NM_000212	17q21.32	P05106
<i>ITGB4</i>	$\beta 4$	CD104	NM_000213	17q25	P16144
<i>ITGB5</i>	$\beta 5$		NM_002213	3q21.2	P18084
<i>ITGB6</i>	$\beta 6$		NM_000888	2q24.2	P18564
<i>ITGB7</i>	$\beta 7$		NM_000889	12q13.13	P26010
<i>ITGB8</i>	$\beta 8$		NM_002214	7p15.3	P26012