

Importance of receptor geometry for initiation of signalling by the mincle family of glycan-binding receptors

A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy

Yu Liu

Department of Life Sciences
Imperial College London

February 2022

Abstract

Yu Liu

Submitted for degree of Doctor of Philosophy

February 2022

Importance of receptor geometry for initiation of signalling by the mincle family of glycan-binding receptors

Mincle family receptors bind oligosaccharide ligands through extracellular carbohydrate-recognition domains (CRDs) and initiate intracellular signalling by interacting with the common Fc receptor γ subunit (FcR γ). Mincle and dectin-2 stimulate macrophages in response to pathogen binding, while blood dendritic cell antigen 2 (BDCA-2) modulates cytokine production. The oligomeric states of these mincle family receptors and the orientations of their CRDs have been investigated to elucidate how extracellular ligand binding initiates cytosolic signalling. Chemical crosslinking and introduction and deletion of cysteine residues have been used to show that dimers of mincle in transfected mammalian cells are stabilized by disulfide bonds between cysteine residues in the neck sequence that links the CRD to the membrane. BDCA-2 forms noncovalent dimers in the absence of a neck-region cysteine residue, although a naturally occurring variant can form an interchain disulfide bond. Cysteine residues in the transmembrane portions of these receptors do not form disulfide bonds and are not required for dimer formation or association with FcR γ , but may facilitate trafficking to the cell surface. To investigate how CRDs are positioned in receptor dimers, bacterial expression systems were developed to produce extracellular domain dimers. Fusion to N-terminal dimerization domains in place of the transmembrane domain was used to improve dimerization efficiency. This approach was used to create disulfide-linked mincle extracellular domain dimers and noncovalently linked extracellular domain dimers of dectin-2 and BDCA-2. Analysis of these constructs, as well as existing crystal structures of the CRDs from these proteins and extended versions of the extracellular domain of BDCA-2, showed that there is only limited interaction of the CRDs in the dimers, but interactions can be stabilized by the presence of the neck region. The resulting orientation of sugar-binding sites in the dimers would favour crosslinking of multiple dimers by oligosaccharide ligands, causing clustering of FcR γ to initiate signalling.

Acknowledgements

I am extremely lucky to have Professor Kurt Drickamer and Dr Maureen Taylor as my supervisors who offered me the opportunity to work on this project, helped me develop good laboratory habits and problem-solving senses, surrounded me with many bright minds during my PhD years, and generously shared their beyond-spectacular perspectives on this work with me. I thank Professor Ten Feizi and Professor Erhard Hohenester who were on my progress review panel and offered me their constructive insights into my project. I thank Sabine Jégouzo for babysitting me in the lab and for introducing me to more people on this floor. Although I have not had a chance to meet Dr Hadar Feinberg, I truly appreciate her collaboration with us on solving the crystal structure of dectin-2 extracellular domain, for her effort has made this project better. I thank Dr Jong-Won Kim for kindly exchanging his knowledge on BDCA-2 project with me. I am also grateful for all the people who I encountered through the group and who lit up my days through communication on their diverse projects and interesting extracurricular life.

I thank the China Scholarship Council for funding my four-year PhD degree. I thank my postgraduate research administrator, James Ferguson, who went above and beyond his duty to make sure I was covered by the stipend extension scheme.

I thank my families in China, my UK hosts (Miranda in London and Philip in lake district), and my friends (Shiqi, Ellie, Anna, Maya, Ang, Bernadette, Danning, Rachel, Momoko, Henderson, Mario, Qinshu, Leran, and Zirui) for supporting me and cheering me up throughout the whole time.

Declaration

I hereby declare that the content of this thesis is my own work and that all additional sources of information have been cited accordingly. An image of western blot of crosslinked wild-type BDCA-2 was acquired from Dr Jong-Won Kim. Dectin-2 extracellular domain crystal structure (unpublished data) was acquired from Dr Hadar Feinberg.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Contents

Abstract	2
Acknowledgements	3
Declaration	4
Contents	5
List of figures and tables	6
List of abbreviations	9
Chapter 1 Introduction	
1.1 Glycan recognition	10
1.2 Signalling mechanisms	16
1.3 Adaptor proteins	27
1.4 Mincle family receptors	30
1.5 Plan of thesis	39
Chapter 2 Oligomeric structure of BDCA-2	
2.1 Introduction	40
2.2 Results	42
2.3 Discussion	48
Chapter 3 Identification of an interchain disulfide-forming cysteine residue in mincle	
3.1 Introduction	50
3.2 Results	51
3.3 Discussion	57
Chapter 4 Extracellular domain fragments of mincle	
4.1 Introduction	58
4.2 Results	59
4.3 Discussion	76
Chapter 5 Extracellular domain fragments of dectin-2 and BDCA-2	
5.1 Introduction	77
5.2 Results	77
5.3 Discussion	93
Chapter 6 General discussion	95
Chapter 7 Materials and methods	
7.1 Molecular biology	100
7.2 Purification of proteins from transfected HEK cells	102
7.3 Expression of protein fragments in bacteria	103
7.4 Protein analysis	105
References	109

List of Tables and Figures

Chapter 1 Introduction

- Figure 1.1. Sugar-binding site in mannose-binding protein A.
 Figure 1.2. Mechanism of G-protein coupled receptor activation.
 Figure 1.3. Mechanism of EGF receptor activation.
 Figure 1.4. Organization of CD22 (Siglec 2) and its activation by colocalization with B-cell receptor.
 Figure 1.5. Organization of CD3, the T cell receptor, and potential activation mechanism by allosteric conformational change.
 Figure 1.6. Organization of FcR γ , ζ chains, DAP12 and DAP10 on the cell membrane.
 Figure 1.7. Organization of mincle family members with sugar-binding activity.
 Figure 1.8. Sequence alignment of mincle family receptors.
 Figure 1.9. N-terminal extensions in a subset of C-type CRDs.
 Figure 1.10. Crystal structure of dectin-2 bound to high mannose oligosaccharide.
 Figure 1.11. Proposed crosslinking of dectin-2 dimers by oligosaccharide ligands.
 Table 1.1. C-type lectins involved in signalling.

Chapter 2 Oligomeric structure of BDCA-2

- Figure 2.1. Sequences of the N-terminal non-CRD regions of BDCA-2 and other mincle family receptors.
 Figure 2.2. Expression plasmids for BDCA-2 and FcR γ .
 Figure 2.3. Polypeptide sequence of human BDCA-2.
 Figure 2.4. Membrane location of wild-type BDCA-2.
 Figure 2.5. SDS-polyacrylamide gel electrophoresis analysis of BDCA-2 lacking the intramembrane cysteine residue.
 Figure 2.6. SDS-polyacrylamide gel electrophoresis of crosslinked BDCA-2.
 Figure 2.7. Alignment of N-terminal sequences of mincle and BDCA-2 Ser49Cys variant.
 Figure 2.8. SDS-polyacrylamide gel analysis of affinity purified BDCA-2 Ser49Cys variant.
 Figure 2.9. Disulfide-linked dimers of Ser49Cys form of BDCA-2 identified by SDS-polyacrylamide gel electrophoresis.
 Table 2.1. SNP frequency table for BDCA-2 Ser49Cys variant (rs143790567).

Chapter 3 Identification of an interchain disulfide-forming cysteine residue in mincle

- Figure 3.1. Polypeptide sequence and expression plasmid of wild-type mincle.
 Figure 3.2. Western blotting of affinity purified wild-type mincle and co-purified FcR γ .
 Figure 3.3. Western blotting of endoglycosidase-digested wild-type mincle to determine its location within cells.
 Figure 3.4. Mutagenesis strategy for making transmembrane cysteine and neck cysteine mutants of mincle.
 Figure 3.5. Characterization of oligomeric state and membrane location of affinity-purified neck cysteine mutant of mincle and its copurification with FcR γ .
 Figure 3.6. Characterization of oligomeric state and membrane location of affinity-purified intramembrane cysteine mutant of mincle and its copurification with FcR γ .
 Figure 3.7. BS³-crosslinking of mincle mutants.

Chapter 4 Extracellular domain fragments of mincle

- Figure 4.1. Dimeric, trimeric, and tetrameric coiled-coil structures of GCN4 dimerization domains.
 Figure 4.2. Bacterial expression system and SDS-polyacrylamide gel analysis of human mincle extracellular domain.
 Figure 4.3. Alignment of polypeptide sequences of human and bovine mincle.

Figure 4.4. Bovine mincle extracellular domain without and with GCN4 dimerization domain analyzed on SDS-polyacrylamide gels.

Figure 4.5. Gel-filtration analysis of bovine mincle extracellular domain fused to GCN4 dimerization domain.

Figure 4.6. Cloning strategy for different lengths of mincle neck following GCN4 dimerization domain.

Figure 4.7. Non-reducing SDS-polyacrylamide gels of affinity-purified GCN4/mincle extracellular domain constructs.

Figure 4.8. Gel-filtration analysis of BoMincle ECD D8.

Figure 4.9. Cloning strategy for appending different numbers of computationally optimized heptad repeats to mincle extracellular domain.

Figure 4.10. Analysis of affinity-purified Mincle ECD D9, mincle extracellular domain with 3-heptad CC-pIL-I17N.

Figure 4.11. SDS-polyacrylamide gel electrophoresis of Mincle ECD D9 subjected to resolubilization.

Figure 4.12. Analysis of affinity-purified Mincle ECD D10, mincle extracellular domain with 2-heptad CC-pIL-I17N.

Figure 4.13. Cloning strategy for insertion of protease cleavage sites.

Figure 4.14. SDS-polyacrylamide gel analysis of test protease digestion of GCN4-mincle extracellular domain fusion proteins.

Figure 4.15. SDS-polyacrylamide gel analysis of furin digests of Mincle ECD P4.

Figure 4.16. Purification and characterization of furin-cleaved Mincle ECD P4 protein.

Figure 4.17. Cloning strategy for Mincle ECD P6, mincle extracellular domain with furin cleavage site and 3-heptad CC-pIL-I17N dimerization domain.

Figure 4.18. Characterization of Mincle ECD P6.

Chapter 5 Extracellular domain fragments of dectin-2 and BDCA-2

Figure 5.1. Cloning strategy for expression of human dectin-2 extracellular domains with 3-heptad CC-pIL-I17N.

Figure 5.2. SDS-polyacrylamide gel analysis of human dectin-2 extracellular domains with 3-heptad CC-pIL-I17N.

Figure 5.3. Characterization of human dectin-2 extracellular domain with 3-heptad CC-pIL-I17N.

Figure 5.4. Introduction of Ser49Cys mutation into the extracellular domain of BDCA-2.

Figure 5.5. Cloning strategy for BDCA-2 extracellular domain with different neck lengths.

Figure 5.6. Characterization of human BDCA-2 extracellular domain constructs fused to 3-heptad CC-pIL-I17N dimerization domains.

Figure 5.7. PyMol view of BDCA-2 dimer interface and interface analysis.

Figure 5.8. PyMol view of dectin-2 dimer interface and interface analysis.

Figure 5.9. PyMol view of dectin-2 extracellular domain dimer.

Figure 5.10. Expression and characterization of human BDCA-2 CRD with half of the neck region.

Figure 5.11. SDS-polyacrylamide gel analysis of crosslinking of the extracellular domain of dectin-2.

Figure 5.12. Coverage of peptides in the uncrosslinked samples of dectin-2 and mincle extracellular domains with fused synthetic dimerization domains.

Table 5.1. Reagents tested for crosslinking the extracellular domain of dectin-2.

Chapter 6 General discussion

Figure 6.1. Summary of what dimers of mincle and BDCA-2 or dectin-2 will look like on the cell membrane.

Figure 6.2. Potential mechanism for ligand crosslinking of receptor dimers on the cell membrane.

Figure 6.3. Distorted arrangement of dectin-2 where receptors are bent over to accommodate a small oligosaccharide.

Abbreviations

BCI _P /NBT	Nitro-blue tetrazolium / 5-bromo-4-chloro-3-indolyl phosphate
BDCA-2	Blood dendritic cell antigen 2
BS ³	Bis(sulfosuccinimidyl) suberate
CRD	Carbohydrate-recognition domains
DCAR	Dendritic cell activation receptor
DCIR	Dendritic cell immunoreceptor
DC-SIGN	Dendritic cell-specific ICAM-3 grabbing nonintegrin
DMEM	Dulbecco's Modified Eagle's Medium
EDC	1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide
EDTA	Ethylenediamine tetraacetic acid
EGF	Epidermal growth factor
FcαR	Fc receptor for IgA
FcεRIα	Alpha subunit of high-affinity IgE receptor
FcγR	Gamma subunit of high-affinity IgE receptor
GlcNAc	N-Acetylglucosamine
GPVI	Glycoprotein VI
HEK	Human embryonic kidney
IFN-I	Type I interferon
IL	Interleukin
IP ₃	Inositol (1,4,5)-triphosphate
ITAM	Immunotyrosine activation motif
ITIM	Immunotyrosine inhibitory motif
ICAM-3	Intercellular adhesion molecule 3
LAT	Linker for activation of T cells
LB	Luria-Bertani
Mincle	Macrophage-inducible C-type lectin
MCL	Macrophage C-type lectin
MHC	Major histocompatibility complex
NHS	N-Hydroxysuccinimide
PCR	polymerase chain reaction
PI3K	Phosphoinositide 3-kinase
PIP2	Phosphatidylinositol (4,5)-bisphosphate
PIRα	Paired-Ig-like receptor-α
PLCγ	Phospholipase Cy
G _α Ras	Ras-like domain of G _α
SDS	Sodium dodecyl sulfate
SH	Src-homology
Sulfo-NHS	N-Hydroxysulfosuccinimide
TCA	Trichloroacetic acid
TLR	Toll-like receptor

Chapter 1 | Introduction

The receptors investigated in this thesis initiate transmembrane signalling following binding of sugar ligands. As background for the experimental work described in subsequent chapters, the four sections of this chapter introduce the structures and functions of sugar-binding receptors, general mechanisms for transmission of signalling information across membranes, the role of transmembrane adaptor proteins in signalling, and the specific family of sugar-binding receptors being investigated.

1.1 Glycan recognition

Sugars mediate a wide range of important cellular functions, typically on the cell surface. These roles depend on the presence of specific combinations of monosaccharide units. The resulting oligosaccharides, attached to proteins or lipids, are referred to as glycans. Many of the functions of glycans involve recognition by glycan-binding receptors known as lectins. In mammals, lectins can interact with both endogenous glycans and with sugars on the surfaces of micro-organisms.

Recognition of self-glycans

Critical cellular processes such as cell-cell adhesion, glycoprotein clearance, and signalling, rely on recognition of endogenous glycans. The best characterized example of cell adhesion resulting from glycan-lectin interactions arises from the selectin family. On the surface of endothelial cells, selectins recognize sialyl-Lewis^x tetrasaccharides (Neu5Ac α 2-3Gal β 1-4(Fuc α 1-3)GlcNAc) on leukocytes, which mediates the initial rolling of leukocytes along the endothelial surface. This rolling causes leukocytes to slow down, and the glycan-selectin interaction is then replaced by stronger integrin-mediated interactions that allow leukocyte extravasation into the underlying tissue (Zarbock et al., 2011).

Endocytic receptors that mediate uptake and degradation of glycoproteins contribute to selective clearance of glycoproteins from the blood. For example, the asialoglycoprotein receptor expressed by liver cells binds to exposed galactose residues on circulating glycoproteins after the terminal sialic acids are removed. This binding results in uptake and degradation of these glycoproteins and is important for the turnover of serum glycoproteins (Yang et al., 2015). A clearance function is also observed with the mannose receptor, which binds to sulphated N-galactosamine on pituitary hormones and results in rapid uptake of the hormone by liver cells (Fiete et al., 1998).

Signalling in response to self-glycans has been observed in receptors across different structural groups. Well-established models include the siglecs, which contain an extracellular domain that binds to sialylated structures that are rare in microbes but common to higher vertebrates as well as an intracellular tyrosine-based inhibitory motif that signals to dampen B

cell immune response. Such signalling is required to induce tolerance towards self-antigens (Jellusova and Nitschke, 2010). In the case of the galectins, which do not anchor to the cell membrane themselves, a different model of lectin interference with immune response has been described. For example, galectin-1 crosslinks properly glycosylated CD45 on the surface of T cells and inhibits the phosphatase function of intracellular domain of CD45, resulting in T cell death (Thiemann and Baum, 2016).

Innate immunity

Differences in the glycosylation patterns characterizing the surface of host cells and microbes has formed the basis for recognition of pathogenic glycans by host glycan-binding proteins, which leads to innate immunity. For example, serum mannose-binding protein and ficolins, which circulate in the blood, can bind to mannose or GlcNAc that often coats pathogens and initiate the complement cascade (Matsushita et al., 2000). Other lectins found on the surface of dendritic cells and macrophages, such as the mannose receptor, can mediate internalization and destruction of pathogens, while some receptors, such as dectin-1, initiate production of reactive oxygen species and upregulation of genes involved in innate immune responses (Brown, 2006).

Classification of lectins based on mechanism

Mammalian lectins fall into multiple classes defined by the structure of carbohydrate-recognition domains (CRD) that mediate the sugar-binding activity of each protein. Based on similarity in primary structures and tertiary polypeptide folds, CRDs can be grouped into evolutionarily related families. Amongst the 13 or more families of animal lectins, C-type lectins represent the largest and most diverse lectin family in animals. Although the specificity for sugars varies among the family members, a nearly invariant feature of the C-type lectins is a Ca^{2+} -binding pocket found in the primary sugar-binding site (Ng et al., 2002) (Fig. 1.1). The Ca^{2+} stabilizes the local conformation of the CRD and forms coordination bonds with key hydroxyl groups on the sugar (Weis et al., 1998). There are outliers to this rule, however. For example, the CRD of dectin-1 lacks a Ca^{2+} -binding site, but it is still classified as a C-type lectin because the sequence of its CRD shows a high degree of identity to other C-type CRDs (Ariizumi et al., 1999).

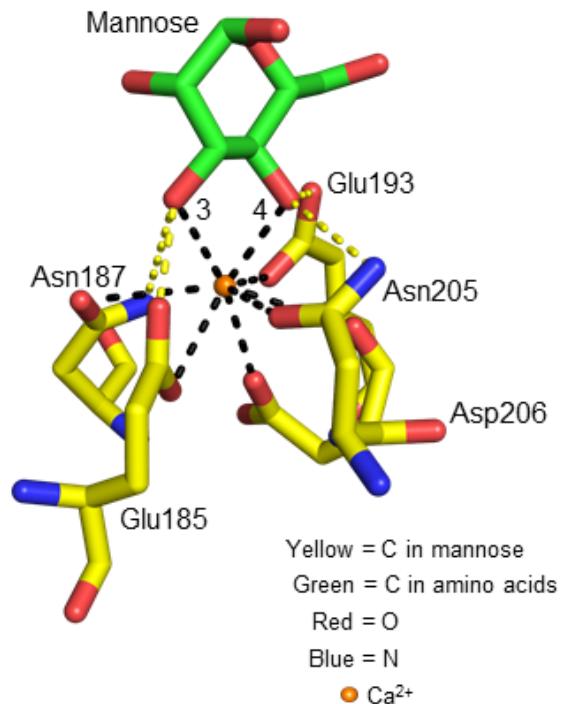


Figure 1.1. Sugar-binding site in mannose-binding protein A. Black dashed lines represent calcium coordination bonds; yellow dashed lines represent hydrogen bonds. Based on PDB file 1KX1.

C-type lectins involved in signalling

Many C-type lectins expressed on myeloid cells are involved in innate immune signalling. The cytoplasmic domains of many of these receptors contain short motifs that serve as docking points for kinases, phosphatases and adapter molecules that contain binding domains such as Src-homology (SH) domains. Some of the common motifs in C-type lectins and their roles in signalling are summarized in Table 1.1. Dectin-1, prolectin and DCIR have defined interaction partners in the cytoplasm, while the signalling pathways for the mannose receptor, DC-SIGN, CD23, and langerin are not yet clear.

Table 1.1. C-type lectins involved in signalling.

Lectin	Motif in cytoplasmic domain	Functions
Dectin-1	YXXL	Syk activation; IL-2 and IL-10 production
Prolectin	YXNX	Activation potentially through Grb2
DCIR	ITYAEV	Inhibition potentially through SHP-1, SHP-2, Syk and Src family kinases
Mannose receptor	FENTLYF	Endocytosis; involvement in signalling but lacking a signalling motif
DC-SIGN	YXXL	Resembling hemi-ITAM but no signalling function
CD23	YESI	Endocytosis
Langerin	WPREPPP	Routing of Birbeck granules

Dectin-1 is expressed on the surface of dendritic cells as a receptor for β -glucan, a glucose polymer that forms part of the yeast cell wall. Yeast recognition by dectin-1 leads to production of high levels of interleukins IL-2 and IL-10. Cells expressing a truncated version of dectin-1 lacking the cytoplasmic domain have unperturbed zymosan-binding function but are unable to generate IL-2 and IL-10 (Rogers et al., 2005). The cytoplasmic domain of dectin-1 contains a membrane-distal tyrosine in the format of YXXXL and a membrane-proximal tyrosine in the format of YXXL, but only the membrane-proximal tyrosine is phosphorylated (Kerrigan and Brown, 2010). Because of the single occurrence of phosphotyrosine and its resemblance to the duplicated YXXL/I sequences in typical immunotyrosine activation motifs (ITAM) that are targets for tyrosine kinases, the YXXL motif in dectin-1 cytoplasmic domain is described as an ITAM-like motif. The model for ITAM-like signalling pathway of dectin-1 depicted dimerization of dectin-1 caused by ligand engagement at the extracellular side of the receptors allows Syk to bridge the two monophosphorylated molecules at the cytoplasmic side of the receptor (Kerrigan and Brown, 2010). It was suggested that preference for Syk kinase instead of another cytosolic kinase, such as ZAP70, is due to the high flexibility of the linker region between the two Src homology 2 (SH2) domains in Syk. Compared with Syk, ZAP70, which has a more constrained arrangement of SH2 domains, only shows stringent binding to a dually phosphorylated ITAM but not to phosphotyrosines from two separate motifs (Hatada et al., 1995).

Prolectin is expressed in dividing B cells in germinal centres. It preferentially binds to terminal α -linked mannose or fucose residues that can be found on high-mannose or Lewis-type structures. Examination of the cytoplasmic domain of prolectin revealed at least three potential SH2-binding sequences and eight potential SH3-binding sequences. Grb2 is a ubiquitous adaptor protein that contains an SH2 domain that binds YXNX sequence that can be found in the cytoplasmic domain of prolectin. Antibody blotting following affinity purification of prolectin from transfected cells showed copurification of Grb2 with prolectin that contains phosphorylated tyrosine residues (Graham et al., 2009). Further steps in the signalling pathways leading from prolectin remain to be defined.

Dendritic cell immunoreceptor (DCIR) is expressed on the surface of dendritic cells and other antigen-presenting cells. By homology of the C-type carbohydrate recognition domain, DCIR can be grouped to a family of receptors discussed below, which includes dendritic cell activation receptor (DCAR), dectin-2, BDCA-2, macrophage C-type lectin (MCL), and mincle. However, none of these related receptors possess a consensus signalling sequence in the cytoplasmic domain except DCIR, which contains an immunotyrosine inhibitory motif (ITIM) ITYAEV. Both threonine and tyrosine in this motif can be phosphorylated, and tyrosine-phosphorylated ITIM can interact with SHP-1, SHP-2, Syk and eight of the Src family kinases (Lambert et al., 2011). DCIR signalling through the ITIM contributes to HIV-1 infection of dendritic cells, which are the primary entry sites for HIV-1 virions and is a vital step that takes the virions across the barrier formed by mucosa epithelium and closer to the T-cell zone.

DC-SIGN (dendritic cell-specific ICAM-3 grabbing nonintegrin) is mainly expressed on the surface of sub-epithelial dendritic cells and specialized macrophages in certain tissues. It functions as an adhesion receptor that stabilizes interaction between dendritic cells and T cells by binding to the glycoprotein intercellular adhesion molecule 3 (ICAM-3). DC-SIGN has also been reported to mediate the rolling of dendritic cells during their migration across the endothelial cell layer through interaction with the glycoprotein ICAM-2 (Holla and Skerra, 2011). The extracellular domain of each DC-SIGN molecule contains a C-terminal C-type CRD that allows binding to mannose- and fucose-bearing glycans, followed by an α -helical neck that promotes formation of a tetramer (Mitchell et al., 2001). The intracellular domain of DC-SIGN contains a YXX ϕ motif, in which ϕ is a large, nonpolar residue, that forms a target sequence for binding of clathrin adapter protein AP-2 (Traub, 2003). The presence of this motif is consistent with participation of DC-SIGN in a clathrin-dependent endocytic pathway (Drickamer and Taylor, 2015). In addition, it has been reported that ligation of DC-SIGN on dendritic cell surfaces leads to association with tyrosine kinases Lyn and Syk and results in cytokine production (Caparrós et al., 2006). Although the YXXL sequence found in the cytoplasmic tail of DC-SIGN resembles a hemi-ITAM as in dectin-1, tyrosine phosphorylation

has not been detected in this sequence. Thus, it has been postulated that the association of DC-SIGN with tyrosine kinases is indirect.

The mannose receptor is mainly expressed on the surface of macrophages and dendritic cells. The extracellular portion of this receptor contains an N-terminal cysteine-rich domain which recognizes the terminal sulphated *N*-acetylgalactosamine on glycoprotein hormones secreted by the pituitary. This region is followed by a fibronectin type II domain that is involved in collagen binding and eight tandemly arranged C-type lectin like domains, of which domains 4-8 are required for binding and endocytosis of ligands terminated with mannose, *N*-acetylglucosamine or fucose (Feinberg et al., 2000). The tyrosine residue and the di-aromatic YF motif in an FENTLYF sequence in the cytoplasmic domain of the mannose receptor accounts for its clathrin-mediated endocytic behaviour towards soluble ligands. Despite lack of a signalling motif, the mannose receptor appears to participate in intracellular signalling in response to pathogenic fungi. Addition of fungi to transfected cells singly expressing either mannose receptor or toll-like receptor 2 (TLR-2), a canonical pattern-recognition receptor, does not induce secretion of cytokines, but transfectants co-expressing both receptors do respond, suggesting that mannose receptor signalling is conditional on cooperation with a canonical pathogen recognition receptor via a mechanism that is yet not understood (Gazi and Martinez-Pomares, 2009).

CD23 is a cell surface receptor for the Fc domain of IgE. The extracellular portion of CD23 contains a C-terminal C-type lectin-like domain and an α -helical coiled-coil stalk region that promotes formation of receptor trimers. Although IgE is highly glycosylated and the IgE-binding domain of CD23 resides in the C-terminal C-type lectin-like region, CD23 recognizes IgE by its C ϵ 3 domain rather than its glycosylation sites (Gordon, 1994). There are two splice forms of CD23, CD23a and CD23b. The former is exclusively expressed by B cells, and the latter is expressed by IL-4-stimulated cells including B cells, macrophages, and eosinophils. The two isoforms are only different in the first 6 to 7 amino acid residues in the N-terminal cytoplasmic domains. In human CD23a, the N-terminus features a YESI motif (residue 6 to 9), which serves as a clathrin-dependent internalization signal. On the other hand, CD23b mediates IgE-dependent phagocytosis. Although the two splice forms serve different functions, a serine-to-phenylalanine mutation in the cytoplasmic tail of CD23b at a position corresponding to that of the tyrosine residue in the YESI motif of CD23a restores the ability of CD23b to mediate endocytosis (Yokota et al., 1992). CD23 is reported to negatively regulate B-cell receptor signalling. Although the exact molecular mechanisms are not known, CD23 plays a direct or indirect role in promoting B-cell contraction and actin clearance, which causes the coalescence of B cell receptor microclusters into polarized central clusters, leading to signal attenuation (Liu et al., 2016),

Langerin is an endocytic receptor expressed on the surface of Langerhans cells, which are immature dendritic cells of the epidermis and mucosal tissues. The extracellular portion of the receptor forms a trimer stabilized by an α -helical neck region in each langerin molecule. Interactions between the neck domain and the C-terminal C-type CRD of each molecule hold the CRDs in a fixed position, so that it binds selectively to mannose-containing polysaccharides of bacterial wall but not to mammalian-type high-mannose structures (Feinberg et al., 2010). Although a consensus endocytosis signal motif is not found in langerin to explain its role in antigen uptake, a proline-rich motif (WPREPPP) in the cytoplasmic tail has been assigned a role in the routing to Birbeck granules, which are antigen-processing organelles unique to Langerhans cells (van der Vlist and Geijtenbeek, 2010). Because the proline-rich domain interacts with SH3-containing adaptor proteins during vesicular trafficking, and mutation within this domain prevents langerin-mediated endocytosis of influenza A virus, it is proposed that the proline-rich domain in langerin interacts with SH3 domain-containing adaptor proteins downstream of viral internalization, which stimulates GTPase activity of Rab5 as well as endosomal trafficking and maturation (Ng et al., 2016).

1.2 Signalling mechanisms

In multicellular organisms, a cell senses specific chemical signals in its environment to make appropriate response to regulate metabolism or change its gene-expression pattern. The process through which a cell becomes aware of a chemical signal and subsequently reacts to it is called signal transduction. While chemical signals can come in various forms, from small nutrient molecules to hormones, drugs and xenobiotics, these signal molecules do not enter cells directly to exert their effects. Instead, specific cell-surface receptors are required for receiving the corresponding signals, thereby initiating the process of signalling. There are at least three different mechanisms of signal transduction: conformational changes, dimerization and clustering. In the following sections, model receptors in each category will be used as examples to illustrate these different signalling mechanisms.

Conformational changes: G-protein coupled receptors

An example of a receptor which signals through conformational change is the β -adrenergic receptor. The β -adrenergic receptor belongs to the largest family of cell surface receptors, the superfamily of G-protein coupled receptors, all of which share a seven-transmembrane topology. There are approximately 800 genes encoding proteins of this family in the human genome. Its members process enormously diverse environmental information, ranging from synaptic neurotransmitters and hormones to photons, odorants, and tastants.

When bound to its physiological ligand epinephrine (adrenaline), the β -adrenergic receptor transmits the signal by activation of a heterotrimeric G protein, which is named after its three-

subunit structure and its ability to bind guanyl nucleotides. Interaction of the G-protein with the receptor causes exchange of GDP with GTP at the nucleotide-binding site. This exchange enables G_α , the α subunit of G protein to dissociate from the $G_{\beta\gamma}$ dimer and the latter actively interacts with adenylate cyclase, which catalyses the conversion of ATP to cAMP. The increase in cytosolic concentration of cAMP causes activation of protein kinase A, which regulates a wide range of cellular events through phosphorylation of serine or threonine residues on a variety of target proteins and changes their activities (Kobilka, 2013).

The ligand-binding pocket in the β -adrenergic receptor is located at the centre of the protein and is formed by transmembrane helices 3, 6 and 7 (Fig. 1.2B). Mutagenesis studies show that amino acid residues Asp113 and Val114 in transmembrane helix 3, Phe289 and Phe290 in transmembrane helix 6, and Asn312 in transmembrane helix 7 are required for interactions with ligand (Rasmussen et al., 2007).

Crystallization studies revealed that although the binding of an agonist only induces a subtle conformational change at the ligand-binding domain of β -adrenergic receptor, the presence of agonist induces dramatic changes in the cytoplasmic domain, including an outward movement of transmembrane helices 5 and 6, by 6 Å and 11.4 Å respectively, and an inward movement of transmembrane helices 7 and 3 (Fig. 1.2C). These re-positioned transmembrane helices form a G-protein binding site at the cytoplasmic side of the receptor (Rasmussen et al., 2011a). Comparison of crystal structure of inactive GDP-bound G protein and that of activated G protein bound by β -adrenergic receptor revealed that the $\alpha 5$ -helix of the Ras-like domain of G_α (G_α Ras), rotates by approximately 130° and makes a significant 6-Å movement towards the core of β -adrenergic receptor, inserting its C-terminus into the binding site formed by transmembrane helices 5 and 6 (Rasmussen et al., 2011b). Taking together, the initiation of signalling requires an agonist-bound β -adrenergic receptor to interact with a GDP-bound G_α .

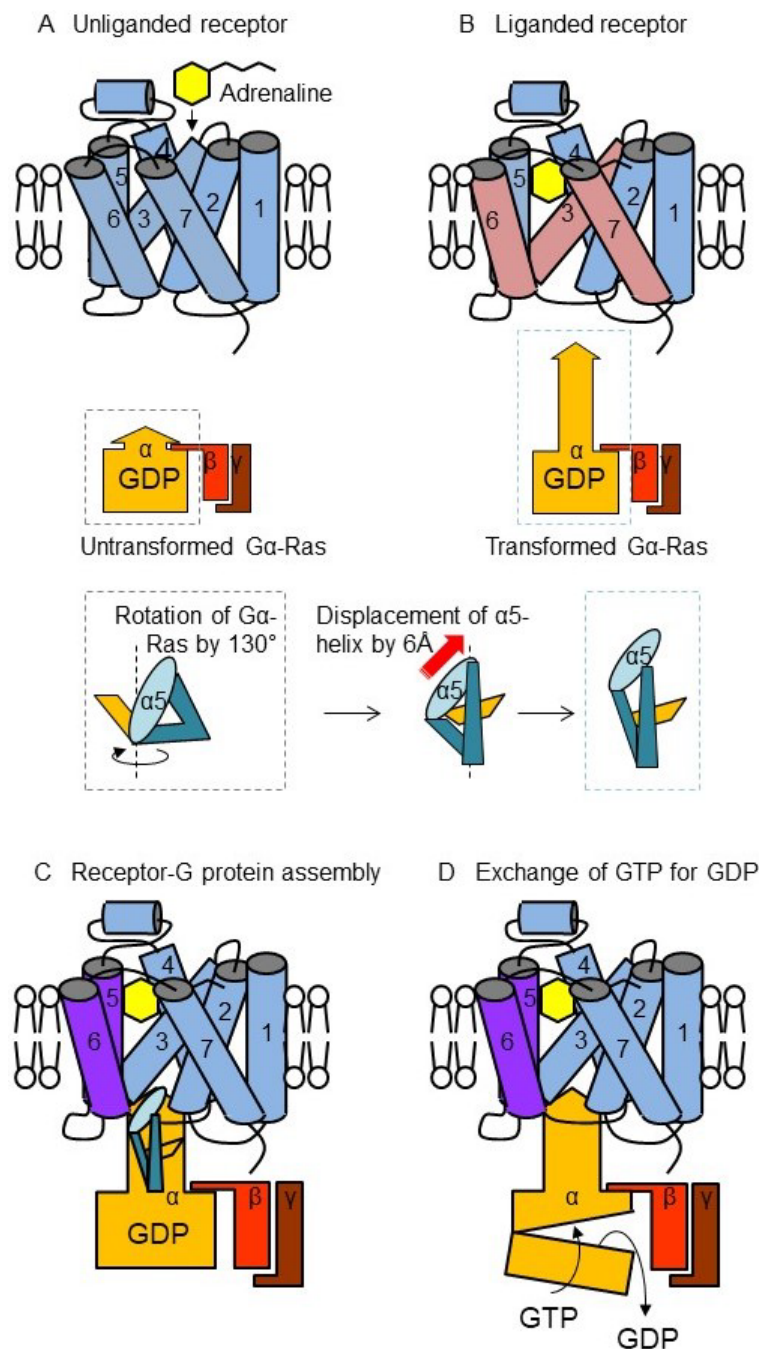


Figure 1.2. Mechanism of G-protein coupled receptor activation. (A) β-adrenergic receptor at rest. (B) Liganded β-adrenergic receptor. Adrenaline molecule interacts with the ligand-binding pocket formed by transmembrane helices 3, 6 and 7 of the β-adrenergic receptor. (C) Transformation of Gα subunit and assembly with receptor. Binding of adrenaline induces an outward displacement of transmembrane helices 5 and 6 (purple) in the β-adrenergic receptor. The α5 helix of G protein α subunit undertakes a rotation of 130° and displacement by 6 Å towards the core of β-adrenergic receptor. (D) Assembly of receptor-G protein complex releases GDP from the α subunit, and GTP binds to the nucleotide-free α subunit. This triggers the release of the βγ dimer, which regulates downstream pathway through adenylate cyclase.

It is striking that binding of a small ligand like an adrenaline molecule can lead to such a dramatic conformational change at the cytoplasmic side of β-adrenergic receptor. Comparison

of structures of the active and inactive β -adrenergic receptor suggested that perhaps agonist interactions with two serine residues were able to stabilize the ligand-binding site in a slightly different local conformation. This small local change levers rearrangements in the rest of the protein, leading to a global structural change that creates a binding site for the G protein (Rasmussen et al., 2011a).

Binding of the hormone ligand to the β -adrenergic receptor must make the subsequent conformational changes energetically favourable. At an affinity of 1.07 nM, binding of an adrenaline molecule provides -8.2 kcal/mol free energy to shift the equilibrium to the active form of the receptor. In keeping with the tight coupling of hormone binding and binding of the G protein, in the absence of a G-protein, an agonist only binds receptor with a K_i of 107 nM, which is 100-fold weaker than when the receptor is stabilized by a G protein (Rasmussen et al., 2011a).

Dimerization: tyrosine kinase growth factor receptors

An example of signalling driven by receptor dimerization is seen from activation of epidermal growth factor (EGF) receptor. EGF is secreted by a wide range of epithelial tissues. The 6-kDa polypeptide binds to receptors on the surface of epidermal and epithelial cells, resulting in stimulation of cell growth and division. The EGF receptor, designated ErbB-1, is a member of the receptor tyrosine kinase family. Fifty-eight proteins from this family are encoded in the human genome. A receptor tyrosine kinase has a cytoplasmic tyrosine kinase domain, which catalyzes transfer of a phosphate from ATP to the hydroxyl group of a tyrosine residue rather than to that of serine or threonine residues. Receptor tyrosine kinases commonly adopt an inhibitory conformation of the intracellular kinase domains, which prevents substrate from binding until ligand binds to the extracellular portion of the receptor. However, the EGF receptor is special, in that control of receptor activation is not exerted by this auto-phosphorylating inhibitory mechanism, but through inhibition of dimerization at its extracellular domains. Over-expression or mutation of EGF receptor that maintains the receptor in a persistently active state is seen in many epithelial cancers, including breast, ovarian and colorectal cancer (Burgess et al., 2003).

The EGF receptor exists as a monomer on the cell surface until the binding of ligand. When EGF is present, each receptor monomer binds to an EGF molecule, causing the receptor to change into a conformation active for dimerization. The monomers interact through the dimerization arm (Fig. 1.3B). In the absence of the EGF ligand, receptor monomers are held in a cyclic configuration, shielding the dimerization arm (Fig. 1.3A). When EGF binds, intramolecular contacts of the dimerization arm are broken, so it can insert into a binding pocket of the other monomer. This movement draws the C-terminal tail of one receptor kinase into the

active site of the kinase domain of the other subunit of the now dimeric receptor, allowing phosphorylation at the C-terminus (Fig. 1.3C).

The phosphorylated C-terminal domain of the receptor binds the SH2 domain of adaptor molecule Grb-2. The SH3 domains of the phosphorylated Grb-2 interact with the proline-rich domain of Sos, a guanine nucleotide exchange factor, and target the translocation of Sos to the plasma membrane. Sos binds to a membrane-anchored small G-protein, Ras, allowing a GDP in Ras to be exchanged for a GTP. Ras functions as a GTPase, like the G_α subunit of the heterotrimeric G protein introduced above. Bound with GTP, Ras is activated and interacts with a membrane-anchored protein kinase, Raf, which amplifies the signal transduction further by serial phosphorylation of downstream proteins, including kinases MEK and ERK (Fig. 1.3D). This pathway eventually changes the gene expression pattern to favour cell growth and division (Lemmon et al., 2014).

Crystallization studies reveal that the extracellular domain of the EGF receptor is made up of four subdomains. Subdomains I and III are involved in ligand binding, while subdomains II and IV play roles in receptor-mediated dimerization (Garrett et al., 2002). Mutagenesis studies (Dawson et al., 2005) confirm that dimerization is mediated by key residues in the dimerization arm and an adjacent loop in subdomain II of an initiating monomer. The interactions are exerted on stretches of amino acids in the subdomain II of the receiving monomer.

A key point is that contact regions for dimerization are fully shielded in a tethered, monomeric conformation. Interestingly, it was observed that mutations of contact points that are involved in dimerization in subdomain IV often contribute to an increase in ligand-binding affinity, but mutation of key amino acid residues in subdomain II have a more severe impact on ligand interaction, reducing the binding affinity of EGF by 1.5 to 5 fold. It was suggested that mutations in subdomain IV disrupt the tethered, monomeric structure, hence facilitate ligand binding. However, mutations to subdomain II impair receptor dimerization, which is thermodynamically coupled with ligand binding (Dawson et al., 2005).

With a K_D of approximately 175 nM, the binding interaction is weaker in this case than for the G-protein coupled receptors. However, energy required to drive the receptor changes comes both from the favourable binding interactions with the ligand and from the energetics of the kinase reaction.

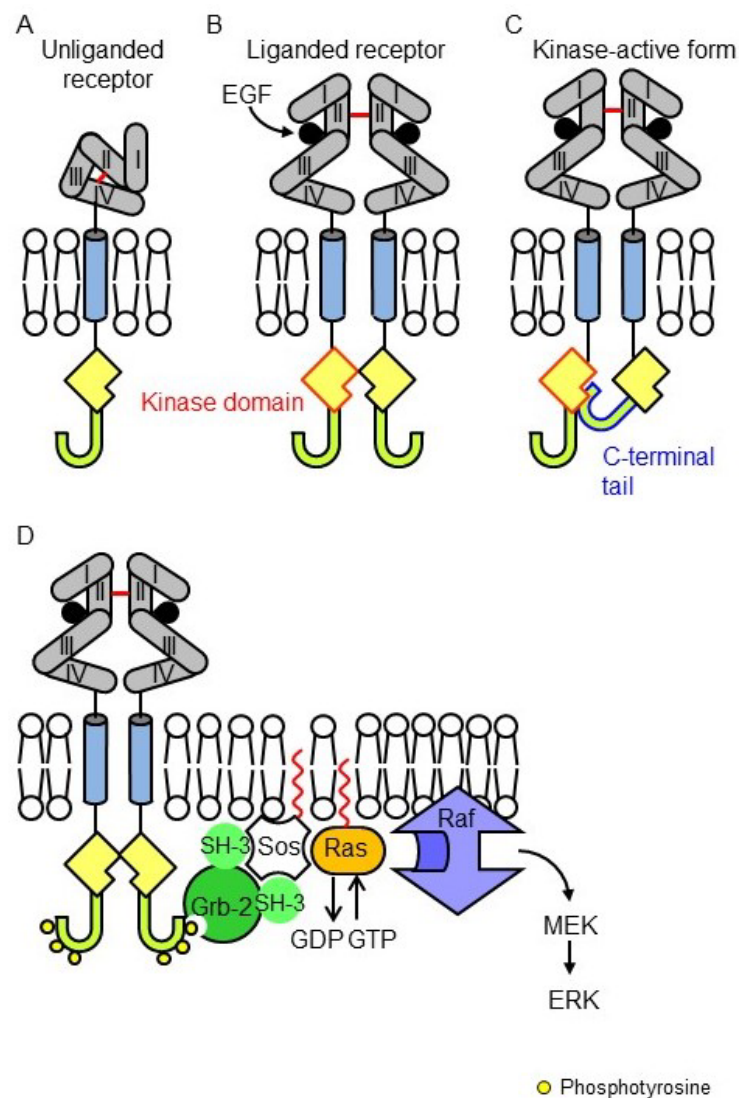


Figure 1.3. Mechanism of EGF receptor activation. (A) Unliganded receptors are monomeric. Key contact regions for dimerization in subdomains II and IV are shielded. (B) In liganded receptors, the dimerization arm in subdomain II of one monomer inserts into a binding pocket in subdomain II of the other monomer. (C) Dimerization activates the kinase function of receptors by bringing the C-terminal tail of one receptor into the active site of its partner's kinase. (D) EGF signalling. Phosphotyrosines on the receptor recruit SH2 domains for Grb-2, and leads to activation of Ras through Sos, resulting in activation of a protein kinase cascade involving Raf, MEK, and ERK.

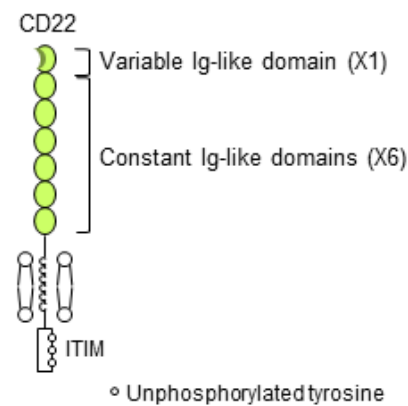
Clustering: phosphatase activation through ITIMs

An example of signalling driven by clustering is seen with the activation of CD22, also known as siglec-2. CD22 belongs to the siglec family of transmembrane sialic acid-binding immunoglobulin-like lectins. There are 14 siglecs encoded by the human genome. The extracellular domains of these receptors contain 1 to 16 immunoglobulin constant domains as spacers and are capped with an N-terminal V-set domain that binds to sialic acids. Most human siglecs bear cytoplasmic ITIMs that interact with intracellular SH2-containing tyrosine phosphatases SHP-1 or SHP-2 and lead to negative signalling (Macauley et al., 2014).

CD22 is expressed on the surface of B cells and plays an important role in dampening B-cell receptor signalling. The extracellular domain of CD22 contains an N-terminal variable immunoglobulin-like domain and six constant immunoglobulin-like domains (Fig. 1.4A). It binds to glycoconjugates bearing α 2,6-linked sialic acid either on the same cell (*cis*) or on other cells (*trans*). When a B cell is untriggered, CD22 monomers form homomultimers by binding to glycan ligands on each other in a *cis* manner. CD22 has also been suggested to interact in *cis* with CD45, a transmembrane protein with a highly glycosylated ectodomain and an intracellular tyrosine phosphatase domain, which restricts CD22 from moving to the B-cell receptor and acting as an inhibitor (Meyer et al., 2018). Upon stimulation of the B cell by an antigen, an ITAM in the B-cell receptor complex is phosphorylated by a membrane-anchored tyrosine kinase, Lyn. The phosphorylated ITAM interacts with an intracellular tyrosine kinase, Syk and activates the Syk-BLNK-BTK/PLC γ 2 pathway. PLC γ 2, phospholipase γ 2, catalyzes the hydrolysis of phosphatidylinositol (4,5)-bisphosphate (PIP₂) to inositol (1,4,5)-triphosphate (IP₃), which binds to a Ca²⁺ channel in the endoplasmic reticulum and results in an influx of Ca²⁺ into the cytoplasm.

CD22 is recruited to microdomains containing activated B-cell receptor through *trans* binding of glycan residue on the other cell. This brings the ITIM of CD22 close to Lyn and allows phosphorylation of ITIM (Fig. 1.4B). The phosphorylated ITIM interacts with the SH-2 domains of SHP-1 and Grb-2. Activated SHP-1 inhibits BTK and PLC γ 2 of the B-cell receptor signalling pathway hence suppresses further release of Ca²⁺ from the endoplasmic reticulum. Phosphorylated Grb-2 interacts with a membrane-bound ATPase, PMCA, which acts as a Ca²⁺ pump to remove the excessive Ca²⁺ from the cytoplasm.

A CD22 domain organization



B Recruitment of CD22 to microdomains containing Lyn

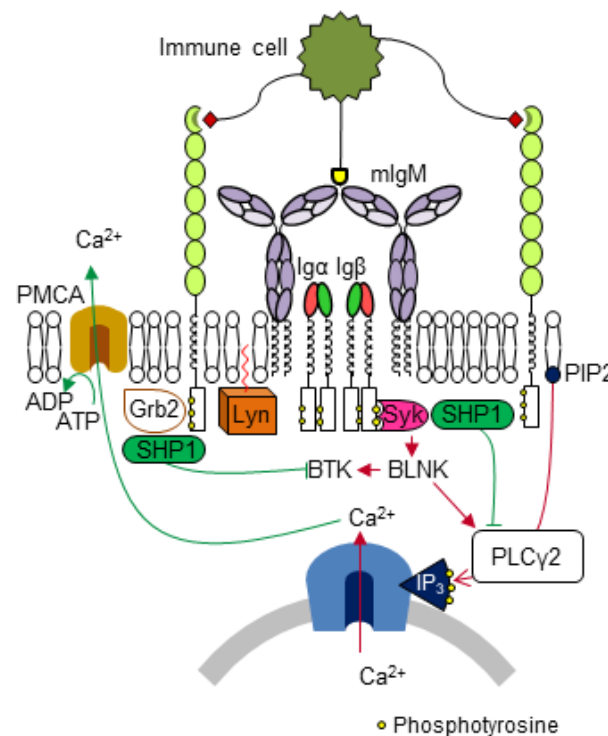


Figure 1.4. Organization of CD22 (Siglec 2) and its activation by colocalization with B-cell receptor. (A) Domain organization of CD22. (B) After B cell activation, CD22 colocalizes with B-cell receptor by *trans* binding to sialic acids (dark red diamonds) from a different cell and its ITIM becomes phosphorylated due to vicinity to membrane-anchored kinase Lyn. B-cell receptor complexes consist of membrane IgM (mIgM) and ITAM-containing Igα/Igβ heterodimer. Phosphorylated ITIMs of CD22 interact with SHP1, which suppresses BTK and PLCγ2 that were upregulated by B-cell receptor signalling. Grb2 is another effector of CD22, which activates PMCA to remove excessive Ca²⁺ from the cytoplasm. Phosphates are shown as yellow dots.

Glycan binding plays an important role in the recruitment of CD22 to the B-cell receptor. Mutations in the glycan-binding domain of CD22 that result in formation of smaller homo-oligomers of CD22, create stronger association with IgM in the B-cell receptor and stronger inhibition of B-cell receptor signalling (Gasparrini et al., 2016). Increased IgM association was

also observed when ST6Gal, the enzyme responsible for generating Sia α 2,6Gal linkages, is knocked out in CD22-expressing cells (Collins et al., 2006). It has been suggested that in both scenarios, impaired *cis*-binding liberates CD22 from a constrained micro-environment, so it gains more mobility to associate with IgM.

Activation of CD22 requires that kinase Lyn be in the vicinity. Since the distribution of Lyn on the plasma membrane is uneven, CD22 needs to move to a microenvironment that is rich in Lyn to be activated. One possible mechanism for this movement is by *trans*-interaction with glycoconjugates on an incoming cell that out-competes *cis*-interaction with ligands on the B cell. This displacement is likely to result from co-clustering with IgM at the antigen-B-cell receptor interaction interface. However, a key question yet to be answered is what high-affinity glycoconjugates on other cells might be the *trans*-interaction partners of CD22.

Allosteric conformational changes: kinase activation through ITAMs

Clustering of ITAMs, analogous to the clustering of ITIMs described for CD22, represents one mechanism by which these domains communicate information across the membrane. However, analysis of T-cell receptor activation also implicates allosteric conformational changes in triggering a kinase-dependent signalling through ITAM coupled by the receptor. The T-cell receptor is a heterodimer of α and β chains linked by a disulfide bond. Each chain has an immunoglobulin-like extracellular domain, a transmembrane domain and a short cytoplasmic tail. An antigen-binding site is formed by the variable regions in the extracellular domain of the T-cell receptor dimer, which allows it to recognize the peptide displayed by a major histocompatibility complex (MHC) molecule on the surface of an antigen-presenting cell. The membrane-spanning domains of T-cell receptor dimer interact non-covalently with three transmembrane dimeric CD3 proteins, CD3 $\epsilon\gamma$, CD3 $\epsilon\delta$, and CD3 $\zeta\zeta$, in a 1:1:1:1 stoichiometric ratio. Each of the ϵ , γ , and δ chain contains one ITAM, whereas the ζ chain bears three. The T-cell receptor-CD3 complex transmits antigen-specific information across the T-cell membrane, leading to numerous molecular changes at the cell membrane, in the cytoplasm and in the nucleus, eventually resulting in T-cell effector function (Samelson, 2002).

Upon binding of T-cell receptor to an antigen, ITAMs in CD3 are predominantly phosphorylated by a membrane-associated Src family tyrosine kinase, Lck. Two phosphorylated tyrosines in each ITAM interact with two SH2 domains of cytoplasmic tyrosine kinase, ZAP70, which is activated and phosphorylates a number of protein substrates, including a transmembrane adaptor molecule, LAT (linker for activation of T cells). LAT interacts with PLC γ 1, an isoform of phospholipase C γ that is expressed preferably in T cells rather than in B cells. PLC γ 1 catalyzes the same cleavage of phospholipid into inositol triphosphate as described for PLC γ 2, which leads to calcium influx from endoplasmic reticulum into cytoplasm. LAT also binds to the adaptor molecule Grb2, which mediates the translocation

of Sos from cytosol to plasma membrane, where it activates the small G protein Ras by facilitating GDP/GTP exchange (Fig. 1.5) (Samelson, 2002).

The mechanism by which ligand-binding at the extracellular domain of the T-cell receptor induces signal transduction across the plasma membrane is not fully understood. Two possible mechanisms of signalling have been postulated: allosteric changes in the cytoplasmic region of the receptor-adaptor complex induced by ligand binding to the receptor or conformational changes resulting from mechanical forces on T-cell receptor (Mariuzza et al., 2020).

In an allosteric conformational change model, it is proposed that T-cell receptor/peptide-MHC interaction repositions the ITAMs of CD3 from a membrane-shielded state inaccessible to cytosolic kinases to a cytoplasm-directed state that can be targeted by Lck for tyrosine phosphorylation. Structural studies reveal that CD3 ϵ is bound to the phospholipid bilayer, with its peptide backbone interacting with the hydrophilic phosphate heads and the aromatic tyrosine residues from the ITAM protruding into the lipid bilayer, which would make them inaccessible to kinases (Xu et al., 2008). In line with this finding, when T-cell receptor is at rest, CD3 ζ is associated with the negatively charged, cytoplasm-facing phospholipid layer of plasma membrane by a stretch of positively charged amino acid residues in its cytoplasmic tail. In addition, fluorescence recovery after photobleaching shows that dissociation of the ITAM-containing cytoplasmic domain of CD3 ζ from the membrane can result either from tyrosine phosphorylation or from T-cell receptor activation by anti-CD3 antibody (Zhang et al., 2011).

The alternative mechanosensing model suggests that motion of a T cell crawling over an antigen-presenting cell to bind to the peptide epitope displayed by the MHC molecule causes sliding between the two cells, generating tensile force tangential to the T-cell surface. Single-molecule assays, in which beads coated with peptide-MHC are presented to the surface of T cells show that binding of cognate peptide-MHC alone is not sufficient for activating T-cell receptor, but applying force tangentially to the binding interface of the T-cell receptor and the peptide-MHC initiates activation (Kim et al., 2009). This result would explain how a monomeric MHC molecule is sufficient to trigger the T-cell receptor-CD3 complex and drive signalling (Huang et al., 2013). The tangential force may be required to induce an allosteric long-distance change in protein dynamics (Mariuzza et al., 2020).

1.3 Adaptor proteins

In innate and adaptive immune systems, it is common that the ligand-binding functions of immune receptors and their signal-transducing elements reside in separate subunits, which are assembled into multi-subunit complexes. This arrangement allows the genes encoding the ligand-binding domains to mutate and diversify as often as necessary to meet the evolutionary challenges posed by the rapidly changing pathogen patterns while maintaining the signalling functions. In addition, it allows the same signalling subunit to be shared among receptors with different specificities (Lanier, 2009). There are seven types of dimeric adaptor molecules, namely ζ chains, FcR γ , DAP12, DAP10, CD3 $\delta\epsilon$, CD3 $\gamma\epsilon$ and CD79 $\alpha\beta$. While CD3 and CD79 heterodimers are exclusively expressed by T cells and B cells and associate with T cell receptors and B cell receptors respectively, ζ chains, FcR γ , DAP12 and DAP10 are associated with different receptors in a wide variety of cells (Berry and Call, 2017). This section will focus on the latter adaptor molecules and their association with receptors (Fig. 1.6).

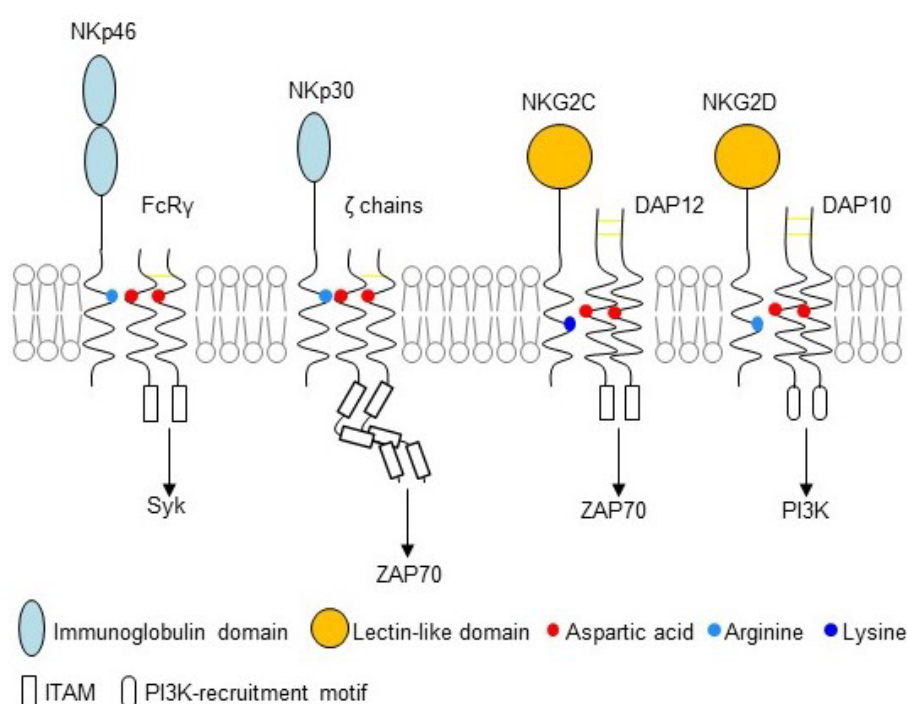


Figure 1.6. Organization of FcR γ , ζ chains, DAP12 and DAP10 on the cell membrane. Extracellular or intramembrane disulfide bonds in adaptor molecule dimers are shown as yellow lines. Oppositely charged residues in the transmembrane domains of receptors and adaptor molecules are shown interacting.

ζ chains, FcR γ , DAP12 and DAP10

Zeta chains, FcR γ , DAP12 and DAP10 are disulfide-linked type I transmembrane proteins each of which lacks a folded extracellular domain, contains an intramembrane aspartate

residue that is important for charge interactions with paired membrane receptors and possesses one or more signalling motifs in the cytoplasmic domain. Whilst the two cysteine residues that form two interchain disulfide bonds in the homodimeric DAP12 and DAP10 are located in the extracellular region, the single cysteine residues that hold the $\zeta\zeta$ and FcR γ homodimers together reside at the N-terminal side of the transmembrane domains.

In DAP10 and DAP12, an intramembrane aspartate residue is located centrally in the transmembrane domain. This residue mediates charge-pairing between DAP10 and NKG2D, and charge interactions between DAP12 and type I transmembrane receptors such as human killer cell immunoglobulin-like receptors, NKp44, TREMs, paired immunoglobulin-like receptors, signal regulatory protein, siglecs, CD300 receptors, mouse CD200 receptors, and mouse Ly49 receptors and type II transmembrane receptors including NKG2C and myeloid DAP12-associated lectin-1 (Berry and Call, 2017). Apart from NKG2D, which primarily associates with DAP10, some receptors that were originally discovered to pair with DAP12 can also interact with DAP10, such as mouse Ly49H and Ly49D, human Sirp-b1, human and mouse Siglec-15, CD300 (Call and Wucherpfennig, 2007). This interchangeability could be due to the sequence similarity in the transmembrane domain of the two adaptor molecules.

In ζ chains and FcR γ , the intramembrane aspartate residue is situated closer to the N-terminal, extracellular end of the transmembrane domain. In ζ chains, this residue directs charge interactions with transmembrane basic residues in T-cell receptor α , NKp46 and NKp30. The acidic residue in FcR γ mediates interactions with both type I transmembrane receptors such as Fc α R (human), glycoprotein VI (GPVI), PIR α (mouse) and type II transmembrane receptors such as NKp46 and mincle (Call and Wucherpfennig, 2007). Non-electrostatic interactions have been suggested to mediate association between CD16A and FcR γ or human ζ chains (Blázquez-Moreno et al., 2017). FcR γ is also reported to interact with several other receptors independent of transmembrane charge-pairing, including type I proteins such as Fc ϵ R1 α and CD64, and type II lectins such as dectin-2 and human BDCA-2 (Blázquez-Moreno et al., 2017; Geijtenbeek and Gringhuis, 2009).

As noted above, the cytoplasmic domain of each FcR γ and DAP12 molecule contains a single ITAM, whilst ζ chain has three. Each ITAM consists of two conserved YXXL/I sequences that are separated by 6 to 8 non-conserved amino acid residues. The two tyrosine residues in the ITAM can be phosphorylated by membrane-anchored Src family kinases such as Lck and Lyn. The phosphorylated ITAM subsequently recruits the SH2 domains of intracellular tyrosine kinases ZAP70 or Syk, which activate further signalling cascades (Lanier, 2009). In the cytoplasmic tail of DAP10, a YINM motif rather than an ITAM is present. When phosphorylated, this sequence binds to the SH2 domains of PI3K and adaptor protein Grb2, both of which are important for optimizing Ca²⁺ responses and inducing cytotoxicity by NKG2D.

A model for receptor-adaptor association through electrostatic interactions

Association of ζ chains, FcR γ , DAP12 and DAP10 with their corresponding receptors requires formation of a 3-helix interface in the membrane. A stoichiometric ratio of 1:2 has been established for charge-pairing of T-cell receptor α with ζ chains, Fc α R or GPVI with FcR γ , NKG2C or KIR2DS2 with DAP12, and NKG2D with DAP10 (Call et al., 2002; Feng et al., 2005; Garrity et al., 2005). The stoichiometry of these receptor-adaptor complexes was investigated by two-step non-denaturing immunoprecipitation of differentially tagged proteins and measurement of radioactively labelled proteins. In the case of association of NKG2C/CD94 heterodimer and DAP12, where CD94 is not involved in intramembrane interactions with DAP12, DAP12, biotin-tagged NKG2C/CD94 and hemagglutinin-tagged NKG2C/CD94 were expressed and assembled in an *in vitro* translation system where all methionine residues were labelled with ^{35}S (Feng et al., 2005). The protein complexes were isolated using the biotin tag under non-denaturing conditions. Following dissociation from the first matrix, a second immunoprecipitation was performed using the hemagglutinin tag. Sodium dodecyl sulfate (SDS) polyacrylamide gels showed that protein complexes from the primary immunoprecipitation contain only one of the two tags and no protein was detected after the second immunoprecipitation. This result indicates that one NKG2C receptor molecule is present in each complex. Measurement of ^{35}S -labelled methionine content revealed a NKG2C-to-DAP12 stoichiometry of 1:2. The same strategy was applied to prove a 1:1:1:1 stoichiometric ratio in the T-cell receptor $\alpha\beta$:CD3 $\delta\epsilon$:CD3 $\epsilon\gamma$: $\zeta\zeta$ complex, a 2:4 ratio of NKG2D:DAP10, and a 1:2 ratio of Fc α R:FcR γ (Call et al., 2002; Garrity et al., 2005; Feng et al., 2005).

For these dimeric adaptors, mutation of the conserved intramembrane aspartate residue in one chain to an asparagine or alanine residue disrupts assembly with their corresponding receptors, whilst mutation of the aspartate residue in both chains to asparagine or alanine almost completely eliminates receptor association. Similarly, mutagenesis of the basic residue in the receptor is sufficient to abolish protein-complex assembly (Call et al., 2002; Garrity et al., 2005; Feng et al., 2005). This charge interaction is specific, since mutation of a positively charged residue to another, such as arginine to lysine, is not well-tolerated by the corresponding adaptor (Call et al., 2010).

The 3-helix interaction interface of NKG2C-DAP12 was studied in a covalently stabilized transmembrane peptide complex system, where the transmembrane region of NKG2C was fused to the C-terminal end of a DAP12 peptide through a flexible linker sequence, and the second strand of DAP12 was captured through formation of natural disulfide bond in the extracellular stalk region (Call et al., 2010). NMR structural analysis of this trimeric complex revealed that the two DAP12 strands form an approximately symmetrical interface containing two polar residues, Asp16 and Thr20, that can interact with receptors. Although the interface extends throughout the transmembrane region, there is a lack of surface complementarity

between receptor and DAP12 dimer, which may explain the ability of DAP12 to associate with a wide range of membrane-spanning receptors. Although Thr20 does not contribute to hydrogen bonding in the interface or affect DAP12 dimerization, mutation of this threonine residue to alanine reduces NKG2C-DAP12 assembly by 70% in an *in vitro* translation system. Sequence alignments of DAP12, DAP10, CD3 δ and CD3 ϵ across species reveal a conserved pattern of DXXXT/S in the transmembrane domains. In line with the observations in DAP12, mutations of the threonine residue in CD3 δ or CD3 ϵ and mutation of serine residue in DAP10 show defects in assembly with T-cell receptor or NKG2D respectively.

Taken together, these observations suggest the conserved serine/threonine residue is functionally significant and it may contribute to charge pairing by stabilizing the aspartate in the adaptor molecule in a favourable position to interact with the basic residue in the receptor (Call et al., 2010). In contrast, ζ chains and FcR γ share a conserved DXXXF sequence rather than DXXXT/S in this region, which may explain their association with receptors by non-ionic interactions, which is not seen with DAP12, DAP10 or CD3 adaptors.

Variation from the charge-based interaction model

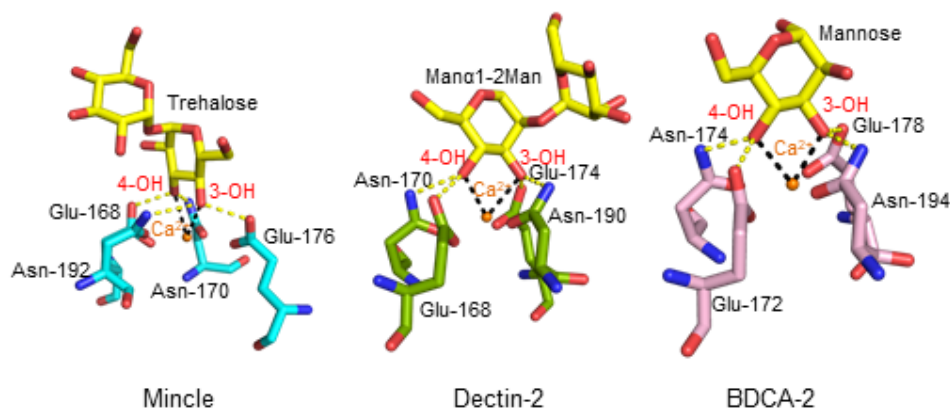
Zeta chains and FcR γ exhibit non-canonical associations with some receptors through non-ionic interactions. For example, *in vitro* translation with ^{35}S -labelling demonstrates that CD16A, which is a type I transmembrane receptor lacking an intramembrane basic residue, forms a 3-helix assembly with either ζ chains or FcR γ (Blázquez-Moreno et al., 2017). When expressed alone from transfected fibroblasts, only small amounts of receptor with immature glycosylation were observed, but when cotransfected with either ζ chains or FcR γ , abundant receptor molecules were expressed on the cell surface. This observation suggests that association with adaptor molecules prevents receptor degradation. Site-directed mutagenesis demonstrated that the critical residues are Phe194, Phe202, Asp205, Thr206, Tyr209 and Phe210. Mapping of amino acids to a helical wheel reveals that the polar or aromatic residues all co-locate to one face of the α -helix. Structural simulations were used to model the assembly of a 30-residue α -helical membrane-spanning fragment of CD16A with FcR γ or ζ chains dimers. The model reveals that the face containing the polar/aromatic residues identified by mutagenesis are favourably packed against the composite interface formed by the FcR γ or ζ chains dimers.

1.4 Mincle family receptors

The mincle receptor family consists of the structurally similar C-type lectins, mincle, dectin-2 and BDCA-2, along with the related macrophage C-type lectin (MCL). In mincle, dectin-2 and BDCA-2, the constellation of amino acids around the primary Ca^{2+} -binding site confers affinity for mannose, glucose and *N*-acetylglucosamine (Fig. 1.7A). In MCL, one of the key

Ca^{2+} ligands does not follow the pattern required for binding of either these sugars or to sugars such as galactose and *N*-acetylgalactosamine, which is consistent with the observation that sugar binding to MCL must be very weak (Balch et al., 1998; Furukawa et al., 2013).

A Primary sugar binding sites



B Domain organization

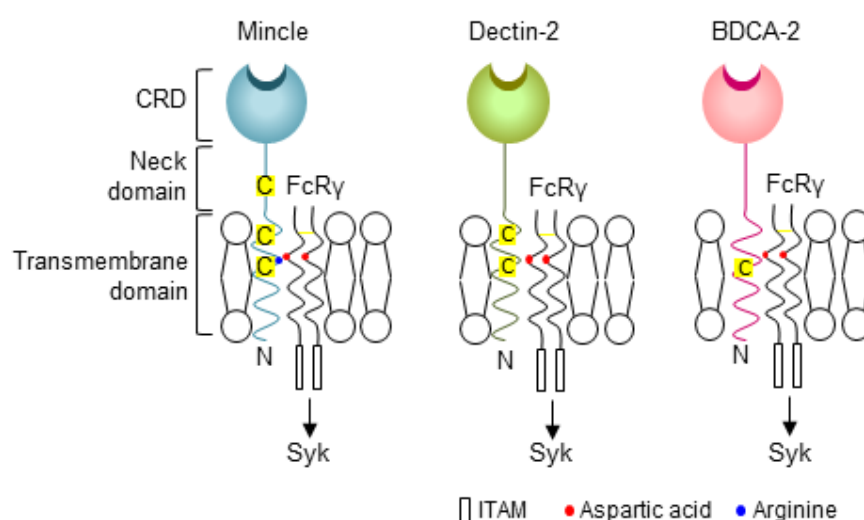


Figure 1.7. Organization of mincle family members with sugar-binding activity. (A) Primary sugar-binding sites in mincle (PDB entry: 4KZV), dectin-2 (PDB entry: 5VYB) and BDCA-2 (PDB entry: 4ZES) with bound ligands. Black dashed lines represent Ca^{2+} coordination bonds; yellow dashed lines represent hydrogen bonds. (B) Domain organization of the three receptors. Cysteine residues in the neck and transmembrane domains of mincle and dectin-2 are highlighted in yellow.

Mincle

Mincle, the macrophage-inducible C-type lectin, also known as CLEC-4E, is expressed on the surface of macrophages. Mincle is conserved across most mammalian species, including human, mouse, cow, dog, dolphin, but not in birds, reptiles or fish (Rambaruth et al., 2015).

Mincle signalling is primarily considered a defence mechanism against *Mycobacterium tuberculosis* since the receptor recognizes the cord factor, trehalose dimycolate glycolipid in the mycobacterial outer membrane, and upregulates production of nitric oxide and pro-inflammatory cytokines IL-6 and TNF α (Ishikawa et al., 2009). Binding competition assays reveal that mincle CRD binds to trehalose with an affinity of 2.2 mM (Feinberg et al., 2016).

Since mincle expression is regulated by TNF/TNF-receptor signalling in macrophages that are activated by a cord factor analogue, trehalose dibehenate, or bovine mycobacterium bacillus Calmette-Guérin, it has been suggested that mincle deficiency during acute or latent infection with *M. tuberculosis* may decrease sensing of mycobacteria and cause diminished expression of cytokines and chemokines, defective granuloma formation, and alterations in phagosomal maturation (Schick et al., 2020). Chemical synthesis reveals differential recognition of diacyl trehalose by T cell receptors and mincle. Fine tuning of the lipid composition of the synthetic analogues of trehalose dimycolate could lead to development of a vaccine that can simultaneously elicit an adaptive immune response through T cell receptor and an innate immune response through mincle (Reijneveld et al., 2021).

Although mincle was historically associated with a pro-inflammatory signalling pathway, mincle contributes in additional ways to the balance of pro- and anti-inflammatory responses. This balance can facilitate the persistence of mycobacteria inside macrophages, which leads to latent infection. Specifically, although IFN- γ stimulates macrophages to engage in antigen presentation and antimicrobial defense, trehalose dimycolate can negatively regulate IFN- γ -induced antigen presentation and antimicrobial gene expression and thus enhance mycobacterial survival inside macrophages (Huber et al., 2020).

Apart from trehalose dimycolate, mincle can also be activated by Gram-negative bacteria (Matsumoto et al., 1999) and fungal species such as *Cryptococcus deneoformans* (Sato et al., 2020) and invasive *Candida* species, although mincle plays a much more limited role in protection against fungus compared with dectin-2 (Thompson et al., 2021). Mincle also binds S-layer glycoproteins from bacteria found in the gut microbiome, including *Lactobacillus kefir* (Malamud et al., 2019) and *Lactobacillus brevis* (Prado Acosta et al., 2021). Binding of S-layer glycoproteins to mincle on macrophages causes Syk-CARD9-dependent production of both pro- and anti-inflammatory cytokines, suggesting a role of mincle in immune homeostasis.

It has also been proposed that mincle may bind endogenous ligands as well as pathogens. For example, β -glucosylceramide from injured kidney is an endogenous ligand for mincle on macrophages that surround dead or dying adipocytes (Tanaka et al., 2020). Involvement of mincle in this sterile inflammation suggests that it has a role in fibrosis in obese adipose tissue. Spliceosome-associated protein 130 released by damaged tubular cells in acute kidney injury is also an endogenous ligand of mincle, although this is apparently not a glycan-lectin interaction, because the ligand does not have any known glycosylation sites (Yamasaki, et al.,

2008). Mincle activation by spliceosome-associated protein 130 is via an miR-219c-3p-dependent mechanism, which promotes inflammation (Lv et al., 2021).

Dectin-2

Dendritic cell-associated C-type lectin-2, known as dectin-2 or CLEC-6A, is found on the surface of dendritic cells, activated macrophages and inflammatory monocytes (Ishikawa et al., 2013). The dectin-2 gene is conserved in many mammals including primates, rodents, and even-toed ungulates. Dectin-2 binds α -1,2-mannan from *Candida albicans* (Robinson et al., 2009; Saito et al., 2010), furfurman, a mannoprotein from *Malassezia* spp. (Ishikawa et al., 2013) and BI-Eng2, a fungal ligand from an attenuated vaccine strain of *Blastomyces dermatitidis*, which robustly induces IL-6 production via a dectin-2 and Card-9 dependent pathway (Wang et al., 2017). Recognition of fungal glycans elicits secretion of proinflammatory cytokines such as IL-2 and TNF, driving Th1/Th17 differentiation and the specialised T cells activate neutrophils, monocytes, macrophages and dendritic cells, constituting a defence line against fungal infection. The Man α 1-2Man disaccharide is of particular interest, because it is a common epitope in mannans on the surface of pathogenic yeasts. Binding competition assays show that dectin-2 CRD binds to Man α 1-2Man and Man α 1-4Man disaccharides with affinities of 2.6 mM and 3.1 mM respectively (Feinberg et al., 2017).

Dectin-2 is not the only C-type lectin that binds to α -1,2-linked mannose. For example, another dendritic cell surface receptor, DC-SIGN, binds to the terminal Man α 1-2Man moiety of high-mannose structures and it has an affinity of 0.56 mM for the Man α 1-2Man disaccharide, which is higher than that of dectin-2 (Feinberg et al., 2007). However, comparison of the sugar-binding sites shows that DC-SIGN has a more restricted geometry that only allows it to bind Man α 1-2Man at the non-reducing termini of oligosaccharides. In contrast, the dectin-2 sugar-binding site has extra space to accommodate additional α -linked mannose residues at the non-reducing end of the Man α 1-2Man disaccharide, which means that dectin-2 can bind to the Man α 1-2Man epitope in internal positions in *Candida albicans* mannans. Hence, although dectin-2 has a lower binding affinity for Man α 1-2Man disaccharide than DC-SIGN, extensive binding is possible through binding to repeat units in the pathogenic polysaccharides rather than the terminal sugars (Feinberg et al., 2017).

The major role of dectin-2 is to provide protection against fungal pathogens. In a study of differential involvement of C-type lectins in response to four invasive *Candida* strains, dectin-2 was shown to mediate clearance of all four species, whilst mincle was only responsible for clearance of one species (Thompson et al., 2021). The importance of the transmembrane adaptor molecule FcR γ during phagocytosis of fungal cargo by dectin-2 or mincle has been investigated, which suggests deficiency of FcR γ , dectin-2 or mincle impaired macrophage migration, engulfment and phagosome maturation (Haider et al., 2019).

A striking demonstration of the essential role of dectin-2 in protection against fungal infection is the case of an immunocompromised patient with an N170I mutation in dectin-2, who developed and died of complications of invasive aspergillosis. Structural modelling of the dectin-2 N170I mutant suggests that the change would lead to incorrect protein folding and *in vitro* experiments show minimal surface expression of functional dectin-2 (Griffiths et al., 2021).

Dectin-2 also binds to Man-LAM, a mannose-capped lipoarabinamannan on *M. tuberculosis* (Yonekawa et al., 2014) and KW3110, a strain of probiotic lactic acid bacteria that displays an α -mannose-type surface glycan (Yoshikawa et al., 2021). Similar to interaction of mincle with *L. brevis*, stimulation by KW3110 activates macrophages to secrete anti-inflammatory cytokines IL-10 and upregulates expression of dectin-2 (Yoshikawa et al., 2021).

An endogenous ligand, β -glucuronidase, also binds to dectin-2 and this binding is partially facilitated by N-linked glycans on the ligand. However, binding is not completely eliminated when the asparagine residues in all four N-glycosylation sites in β -glucuronidase are mutated to glutamine residues, so the exact mode of interaction between dectin-2 and β -glucuronidase remains unclear (Mori et al., 2017). It was proposed that recognition of self-glycans contributes to an immune surveillance function, in which abnormal cell surface glycoproteins exposing terminal mannose structures can be sensed in the same way that dectin-2 senses high-mannose glycans that are commonly expressed on yeast surfaces (Mori et al., 2017).

BDCA-2

Plasmacytoid dendritic cells are dendritic cell precursors found in peripheral blood, thymus, and secondary lymphoid tissue. These pre-dendritic cells are marked by expression of CD4, CD123, HLA-DR, CD68, ILT3, and CD45RA, but markers for myeloid dendritic cells such as CD11c and ILT-1, or for other hematopoietic lineages such as CD3, CD14, CD16, CD19, CD20, and CD56 are absent. Generation of monoclonal antibody against blood dendritic cells identified BDCA-2 and BDCA-4 as two specific markers for plasmacytoid dendritic cells (Dzionek et al., 2000).

Plasmacytoid dendritic cells express endosomal TLR-7 and TLR-9 which recognize microbial RNA or DNA respectively. Under physiological circumstances, TLR signalling induces massive production of type I interferons (IFN-I) by these cells which constitutes a major innate immune defence against viral infections. In systemic lupus erythematosus conditions, plasmacytoid dendritic cells show aberrant migration from blood and accumulate in affected organs such as skin or kidneys. Persistent stimulation of TLR-7 and TLR-9 by autoreactive immune complexes are suspected to be the primary cause of IFN-I release and systemic lupus erythematosus disease progression (Pellerin et al., 2015). BDCA-2 reduces IFN-I production regardless of the IFN-I-inducing stimulus (Cao et al., 2007). Therefore, BDCA-2 has been an appealing target for downregulating IFN-I production in systemic lupus erythematosus.

However, due to the previously unfilled knowledge gap on natural ligands of BDCA-2, the most promising BDCA-2-targeted therapy which recently entered phase I human trial, still has to rely on anti-BDCA-2 monoclonal antibodies for triggering anti-inflammatory signalling (Furie et al., 2019). Identification of several desialylated serum glycoproteins that bind to BDCA-2 at physiological concentrations provides a new perspective for targeting of BDCA-2 by exploiting its glycan-binding function, although to what extent the natural ligands can induce BDCA-2 internalization remains to be evaluated (Kim et al., 2018).

Orthologs of BDCA-2 have been identified in multiple mammalian species, including human, gorilla, chimpanzee, bats and tree shrews. Although human BDCA-2 has 50.7% amino acid identity in overall sequence with murine dectin-2, there is no orthologue of BDCA-2 in mouse (Dzionic et al., 2001). BDCA-2 has an extracellular sugar-binding domain that preferentially binds to oligosaccharides with a Gal β 1-3/4GlcNAc β 1-2Man epitope, which can be found on a subset of desialylated serum proteins including α 2-macroglobulin and immunoglobulins A1, A2, G and M (Kim et al., 2018). Although the primary sugar-binding site of BDCA-2 is occupied by a mannose residue at the reducing end of the Gal β 1-3/4GlcNAc β 1-2Man trisaccharide competition assays show that addition of an *N*-acetylglucosamine and galactose to the non-reducing end of this mannose increases affinity by nearly 1,000 fold, to 9.2 μ M. BDCA-2 CRD binds to a biantennary galactose-terminated glycan with a 20-fold higher affinity than that for a biantennary glycan with one Gal β 1-3/4GlcNAc β 1-2Man terminal structure or a Gal β 1-3/4GlcNAc β 1-2Man trisaccharide (Jégouzo et al., 2015). Taken together, these observations show that both *N*-acetylglucosamine and galactose residues in the Gal β 1-3/4GlcNAc β 1-2Man epitope contribute to binding to galactose-terminated complex N-linked glycans (Jégouzo et al., 2015).

Signalling pathways activated by mincle-family receptors

Since the mincle family receptors lack intracellular signalling motifs, association with the ITAM-bearing FcR γ adaptor molecule is required for signalling initiation. Immunoprecipitation studies show that, when activated, all three receptors associate with FcR γ and initiate either a pro-inflammatory signalling pathway through Syk/CARD9/Bcl-10/MALT1 or an anti-inflammatory signalling pathway through Syk/BLNK/PLC γ 2 (Drummond et al., 2011; Cao et al., 2007). Sequence alignment suggests that association of mincle with FcR γ is mediated by electrostatic interaction between the positively charged arginine in the transmembrane helix of mincle and the negatively charged aspartate at the similar level in the adaptor molecule (Fig. 1.7B), as seen with the Fc α R-FcR γ and GPVI-FcR γ models discussed above. However, the intramembrane basic residue is not conserved in BDCA-2 or dectin-2 and the only conserved features across all of the receptors are one of the intramembrane cysteine residues and clusters of aromatic residues at key positions (Fig 1.8). The CD16A-FcR γ model suggests that

conservation of some of these aromatic residues in the membrane regions may direct association with FcR γ . One intramembrane cysteine residue is also conserved in all of the mincle family members, suggesting that it may play a critical role.

It remains to be established how Syk is stimulated through FcR γ . A conformational change model is unlikely because the sugar-binding affinities of mincle family receptors are low compared with G-protein coupled receptors, which exhibit a binding affinity at 1 nM, providing sufficient free energy for conformational changes. The dimerization model of EGF receptor is impossible because mincle family receptors do not have intracellular tyrosine kinase domains. The force-sensing model of T-cell receptor-CD3 is also unlikely because the tangential force is generated as a consequence of relative movement between T cell and an antigen-presenting cell, but the ligand bound by the mincle family receptor is too small compared with a mammalian cell to generate a force at the same level of strength. Thus, a clustering mechanism seems most likely, but the molecular details of how this occurs remain to be established.

Human Mincle	MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQL----	59
Human MCL	MCLMGLEKPQSKLEGGMHPQLIPSVIAVVFILLGLGVCFIASCLVTHHNSRCRGTGVHK---	60
Human Dectin-2	MMQEQ--QPQSTEKRGW-LSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSEL---	57
Human BDCA-2	MVPEE--EPQDREKGLWWFQLKVMMAVVSILLLSVCFTVSSVVPNFMYSKTVKRLSKLREY	61
CRD →		
Human Mincle	PE [*] NFTELSCYNYGSGS-----VKNCCLNWEYFQSSCYFFSTDTISWALS [*] LK [*] CSAMGAHL	115
Human MCL	LEHHAKLKCIKEKSELKSAEGSTWNCCPIDWRAFQSNCFPLTD [*] NKTWAESE [*] R [*] CSGMGAHL	122
Human Dectin-2	HSYHSSLTCFSEGTKV-----PAWGCPASWKSFGSSCYFISSEK [*] VWSKSEQNCVEMGAHL	114
Human BDCA-2	QQYHPSLTCVMEGKDI-----EDWSCCPTPWTSFQSSCYFISTGMQSWTKSQK [*] NCSVMGADL	118
Human Mincle	VVINSQEEQEFLSYKKPKMREFFIGLSDQVVEGQWQWDGTPLTKSLSFWDVGE [*] PNNIAT	175
Human MCL	MTISTEAEQNFIIQFLDRRLSYFLGLRDENAKGQWRWVDQTPFNPRRVFWHKNE [*] PDNSQ-	181
Human Dectin-2	VVFNTAEQNFIVQQL [*] NESFSYFLGLSDPQGN [*] NNWQWIDKTPYEKNVRFWHLGE [*] PNHS--	172
Human BDCA-2	VVINTREEQDFIIQNLK [*] NSSYFLGLSDPGRRHWQWVDQTPYNE [*] NVTFWHSGE [*] PNNL--	176
Human Mincle	LED [*] CATMRDSSNPRQNW [*] NDVTCFLNYFRICEMVGINPLNKGKSL	219
Human MCL	GENCVLVYNQD-KWAWNDVPCNFEASRICKIPGTTLN-----	218
Human Dectin-2	AEQCASIVFWKPTGWGW [*] NDVICETRRNSICEMNKIYL-----	209
Human BDCA-2	DERCAIINFRSSEEWG [*] NDI [*] HCHVPQKSICKMKKIYI-----	213

Transmembrane domain

N-linked glycosylation sites

Basic residue that interacts with FcR γ

Aromatic residues that may interact with FcR γ

Primary sugar-binding sites

Variant residue in MCL binding site

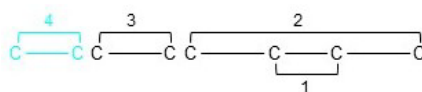
Figure 1.8. Sequence alignment of mincle family receptors. Asterisk and dot indicate perfectly conserved and well-conserved position in the transmembrane domain and primary sugar-binding site of mincle family receptors.

Structural features of mincle-family members

In spite of the fact that the mincle family members bind distinct oligosaccharide ligands, the CRDs share many common features. The surface of a C-type CRD formed from the N- and C-termini of the polypeptide is commonly involved in interactions between CRDs in receptor oligomers. This surface is on the opposite side of the CRD from the sugar-binding site. In the trimers of mannose-binding protein and langerin, interactions between the N-terminal portion of the CRDs in one monomer and the neck of another monomer are important for both oligomer formation and spacing of the CRDs. These interactions hold the CRDs in fixed orientations, which enhance their binding to dense sugar patterns on relatively flat surfaces (Weis et al., 1992; Feinberg et al., 2010).

While a minimal CRD is present in all C-type lectins, the mincle family receptors are amongst the few that have an N-terminal extension to the minimal CRD containing an extra disulfide bond (Fig. 1.9A/B) (Feinberg et al., 2016; Feinberg et al., 2017). Structural analysis shows that this extra disulfide bond locks a stretch of amino acids in a β -hairpin structure, which extends these CRD (Fig. 1.9C). However, none of the crystal structures of C-type lectin oligomers involve CRDs with extended N-termini such as those seen in the mincle family members.

A



B

Human Mincle	LSCYNYGS-----GSVKNCPLNWEYFQSSC
Mouse Mincle	LSCYSEAS-----GSVKNCPLNWKHYQSSC
Human MCL	LKCIKEKSELKSAEGSTWNC CPIDWRAFQSN
Mouse MCL	VT CIREE-PQPGATGGTWT C PVS WRAFQSN
Human Dectin-2	LTCFSEGV---STKVPWAGCCPASWKSFGSSC
Mouse Dectin-2	LTCFSEGT-----MEKMWGCCPNHWKSFGSSC
Human BDCA-2	LTCVMEGK-----DIEDWSC CPTPWSFQSSC
Human DCIR	LEC VKNM---PVEETAWS C PKNWKSFSNC
Mouse DCIR1	LNCTKSVS---PMEDKVWSC C PKDWRLFGSH
Mouse DCIR2	LNCIKNGS---LMEDKVWSC C PKDWKPFVSH
Mouse DCIR3	LECTKAS---LLEDKVWSC C PKDWKPFGSY
Mouse DCIR4	LHCVKNHS---SVEDKVWSC C PKNWKPFDSH
Mouse DCAR1	LICFSEGT---TVSEKVWSC C PKDWKPFGSY
Human ASGPR-H1	LSCQMAALQGN---GSERT C PVNWVEHERS
Mouse ASGPR-H1	LSCQMAAFRGN---GSERIC C PINWVEYEGS

C

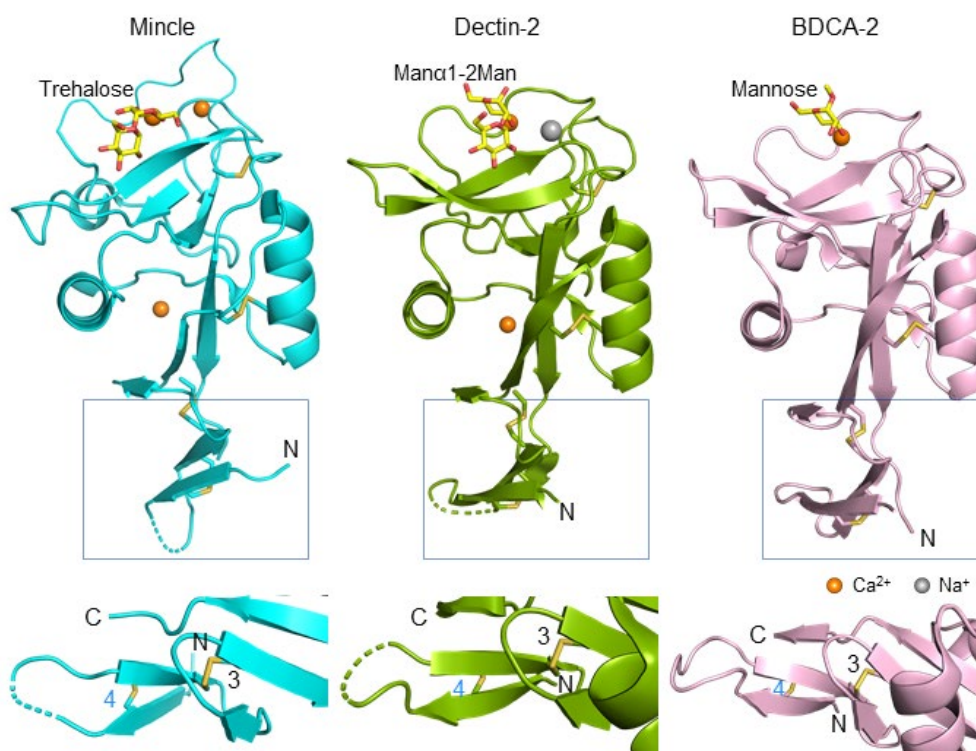


Figure 1.9. N-terminal extensions in a subset of C-type CRDs. (A) Positions of conserved disulfide bonds in minimal C-type CRDs, numbered 1 – 3, and extra disulfide bond found in extended CRDs is shown in cyan. (B) Sequences of extended CRDs, with cysteine residues that form conserved disulfide bond 3 highlighted in yellow and cysteine residues in the extra disulfide bond highlighted in cyan. (C) Structure of the extended CRDs of mincle (cyan), dectin-2 (green) and BDCA-2 (pink) based on PDB files 5KTI, 5VYB, 4ZES. Enlarged views of the boxed regions, containing the hairpin loop formed by the N-terminal extensions are shown at the bottom. The extra disulfide bond is labelled for each receptor.

In contrast to the CRDs, the neck region between the transmembrane domain and the CRD is poorly conserved in the mincle family members (Fig. 1.8). A cysteine residue is present in the neck regions of mincle and MCL, but not in dectin-2 and BDCA-2. N-linked glycosylation sites are also present in some cases but not in others. In addition, the neck sequences lack the characteristic heptad repeat of hydrophobic residues associated with coiled-coil formation in other C-type lectins. The absence of a helical neck and the presence of an extended CRD in mincle family receptors might allow flexibility in both the neck region and the orientation of sugar-binding sites, which would enable the receptors to engage avidly with glycans with a broader range of spacings (Feinberg et al., 2016). However, there is no experimental evidence regarding the conformation of the neck or its flexibility.

Dectin-2 CRDs crosslinked by two branches of an oligosaccharide

Crystals of dectin-2 CRD in complex with $\text{Man}_9\text{GlcNAc}_2$ oligosaccharide shows strong electron density for the portion containing two trisaccharide branches joined by a mannose residue, $\text{Man}\alpha 1\text{-}2\text{Man}\alpha 1\text{-}3(\text{Man}\alpha 1\text{-}2\text{Man}\alpha 1\text{-}3\text{Man}\alpha 1\text{-}6)\text{Man}$. The electron density is clear for the two $\alpha 1\text{-}2$ -linked mannose residues at the non-reducing ends. The remaining electron density corresponds to the $\text{Man}\alpha 1\text{-}3(\text{Man}\alpha 1\text{-}6)\text{Man}$ moiety, which can be fitted into both orientations about the crystallographic 2-fold axis, meaning the $\text{Man}_9\text{GlcNAc}_2$ oligosaccharide can be bound in both orientations. Consequently, the electron density-fitting model shows that two dectin-2 CRD molecules can be crosslinked by two branches of the $\text{Man}_9\text{GlcNAc}_2$ oligosaccharide with different covalent structures: $\text{Man}\alpha 1\text{-}2\text{Man}\alpha 1\text{-}3\text{Man}$ in the D1 branch and $\text{Man}\alpha 1\text{-}2\text{Man}\alpha 1\text{-}3\text{Man}\alpha 1\text{-}6$ in the D2 branch (Fig. 1.10). In each branch, the sub-terminal mannose occupies the primary sugar binding site and makes hydrophobic interactions with Trp182 in the secondary sugar-binding site of each dectin-2 CRD (Feinberg et al., 2017). The structural analysis of dectin-2 CRDs crosslinked by a single $\text{Man}_9\text{GlcNAc}_2$ oligosaccharide suggests a model in which signalling is initiated when receptors are clustered by crosslinking of a single ligand (Fig. 1.11).

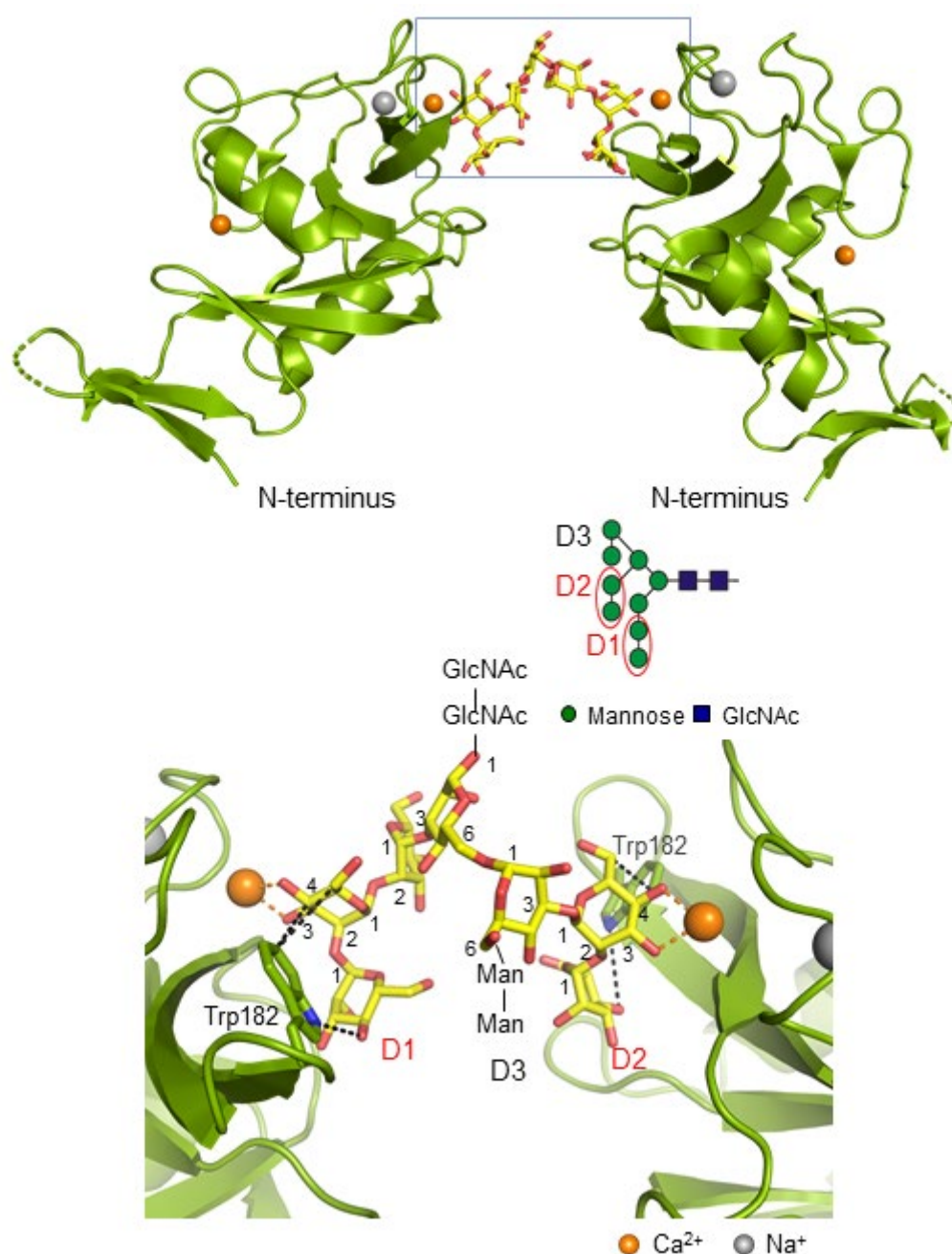


Figure 1.10. Crystal structure of dectin-2 bound to high mannose oligosaccharide. Two dectin-2 CRDs can be crosslinked by the D1 branch and the D2 branch of a single Man₉GlcNAc₂ oligosaccharide ligand (PDB entry: 5VYB). The boxed region is shown in an enlarged view at the bottom. Carbon backbones of the sugar molecules are shown in yellow.

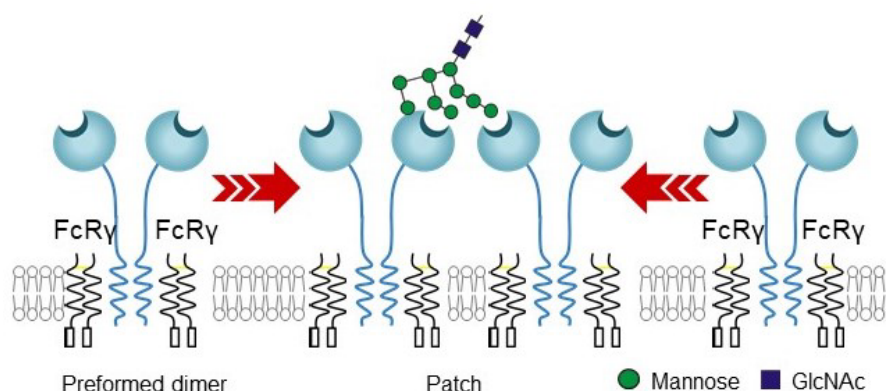


Figure 1.11. Proposed crosslinking of dectin-2 dimers by oligosaccharide ligands. Two preformed dectin-2 dimers are brought close together by binding to two branches of a single oligosaccharide ligand.

1.5 Plan of thesis

The material discussed in this chapter suggests that ligand-induced clustering of mincle family receptors, and hence the ITAM motifs in their associated FcRγ polypeptides, is likely to initiate signalling. Aspects of this model are investigated in the following chapters by analysis of the existence and arrangement of receptor oligomers. In Chapters 2 and 3, expression in eukaryotic cells and analysis of purified receptor fragments are used to demonstrate the existence of preformed receptor dimers and to define the role of disulfide bonds in stabilization of these dimers. In Chapters 4 and 5, bacterial expression systems are developed to mimic the natural dimers, providing a basis for interpreting existing structural information and for undertaking additional structural analysis. Chapter 6 describes the resulting model of rigid receptor dimers that can be brought together into larger clusters by multivalent ligands.

Chapter 2 | Oligomeric structure of BDCA-2

2.1 Introduction

Blood dendritic cell antigen 2 (BDCA-2) is uniquely found on the surface of plasmacytoid dendritic cells, which account for up to 0.5% of the human peripheral blood cells and produce a high level of type I interferon (IFN-I) in response to viral infection (Dzionek et al., 2000; Reizis et al., 2011; Swiecki and Colonna, 2015). It binds in a Ca^{2+} -dependent manner to glycans containing the $\text{Gal}\beta 1\text{-}3/4\text{GlcNAc}\beta 1\text{-}2\text{Man}$ epitope, which are commonly found on a subset of serum proteins, such as α_2 -macroglobulin and immunoglobulins A1, A2, G and M (Jégouzo et al., 2015; Kim et al., 2018). It has a well-documented role in suppressing inflammation in autoimmune diseases, such as systemic lupus erythematosus. A monoclonal antibody against BDCA-2 has been developed to enhance anti-inflammatory BDCA-2 signalling and thus to reduce surface expression of the pro-inflammatory receptor CD32 (Pellerin et al., 2015).

Comparison of BDCA-2 and other mincle family members reveals that the portions of dectin-2 and BDCA-2 that precede the CRD are closely related to each other, while mincle and MCL are similar to each other (Fig. 2.1). However, the neck regions of these receptors are not highly conserved and do not have obvious structural motifs. For example, N-linked glycosylation sites and cysteine residues are present in mincle and human MCL, but not in human dectin-2 or BDCA-2.

Cow Mincle	MNSSTSPASQTERRCFSFQ	LFSWAIAGTSILLLSACFITRCVV	TSIFQLCDEKKFQPPGDSME
Human Mincle	MNSSKSSETQCTERGCFSSQ	MFLWTVAGIPILFLSACFITRCVV	TERIFQTCDEKKFQLPENFTE
Human MCL	MGLEKPSKLEGGMHPQL	LIPSVIAVVFILLLGVCFIASCLV	THHNSRCRGRTGVHKLHHAK
Human Dectin-2	MMQEQQPQSTKRGWLSRL	WSVAGISIALLSACFIVSCVV	TYHFTYGETGKRLSELHSYHSS
Human BDCA-2	MVPEEPPQDREKGLWWFQLK	VWSMAVVSILLLSVCF	TVSSVVP

35

Figure 2.1. Sequences of the N-terminal regions that precede the CRDs of BDCA-2 and other mincle family receptors. Transmembrane domains are highlighted in green. Cysteine residues are highlighted in yellow. Sequences are aligned based on the positions of the intramembrane cysteine residues. Positively charged residues that are presumed to be within the membrane are shaded in red. N-linked glycosylation sites in the neck regions are shown in magenta.

Prior to this project, a human embryonic kidney cell line co-transfected with wild-type human BDCA-2 and FcR γ had been generated for studying the organization of BDCA-2 (Fig. 2.2). Co-expression was achieved using vectors containing cytomegalovirus promoters and an internal ribosome re-entry site from the encephalomyocarditis virus (EMCV) to allow co-selection with two different antibiotic resistance genes. HEK 293 cells were chosen for transfection because they enable correct formation of disulfide bonds of the polypeptide in the oxidative environment of endoplasmic reticulum. The mammalian cell system also allows

appropriate processing of the N-linked glycosylation to occur. The state of processing of the glycans provides information on the intracellular trafficking of the receptor to the plasma membrane. Methods of affinity purification of BDCA-2 had been established and sensitive rabbit polyclonal antiserum for detecting BDCA-2 had been generated. On western blots, a band corresponding to a covalently linked dimer of BDCA-2 was shown to be an artefact resulting from oxidation of an intramembrane cysteine residue during purification. The dimer could be eliminated by pretreating the cell lysates with alkylating reagents such as *N*-ethylmaleimide before isolating the protein. However, chemical crosslinking experiments showed that alkylated BDCA-2 can still be crosslinked into a dimer (Kim, 2018). Thus, it was known that wild-type BDCA-2 naturally forms a dimer that is not disulfide-linked. To confirm that the covalently linked dimer of BDCA-2 generated during purification resulted from reaction of the intramembrane cysteine residue (Cys35), a transfected HEK cell line co-expressing BDCA-2 Cys35Ala, lacking this intramembrane cysteine residue, along with FcR γ , was generated (Kim, 2018).

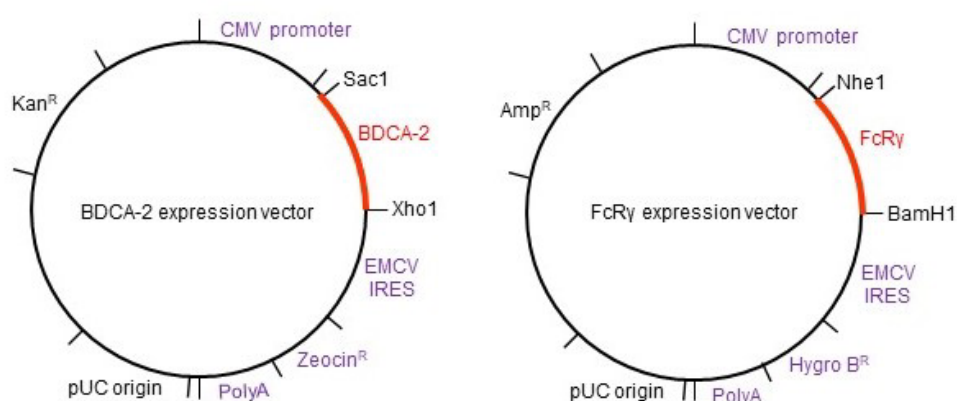


Figure 2.2. Expression plasmids for BDCA-2 and FcR γ . Expression of BDCA-2 and FcR γ is driven by the cytomegalovirus (CMV) promoter. Genes encoding resistance to zeocin or hygromycin B following an internal ribosome re-entry site (IRES).

This chapter describes further biochemical analysis of the role of cysteine residues in the oligomeric structure of BDCA-2. The results clarify the different states of processing of N-linked glycans attached to the polypeptide and confirm that the intramembrane cysteine is not required for dimer assembly, association with FcR γ or trafficking out of the endoplasmic reticulum. In addition, a naturally occurring variant of BDCA-2 with a cysteine residue in the extracellular neck region has been shown to form covalent dimers.

2.2 Results

Glycan processing in wild-type BDCA-2

Following extraction of BDCA-2 from transfected HEK 293 cells and affinity purification on an immobilized glycopeptide-agarose column, SDS-polyacrylamide gel electrophoresis reveals two bands (Kim, 2018). Following PNGase digestion, the two bands collapse into a single, lower-molecular weight band, indicating that these bands represent different glycosylation states of BDCA-2. The extracellular domain of BDCA-2 contains three sites for N-linked glycosylation, all of which are found in the CRD (Fig. 2.3).

The basis for the difference in glycosylation patterns was examined by Endo H_f digestion in parallel with PNGase F digestion. High-mannose glycans can be removed from glycoproteins by Endo H_f digestion, while glycoproteins with complex glycans are resistant to Endo H, but the glycans can be removed by PNGase (Fig. 2.4A). Digestion of purified BDCA-2 confirms that both forms can be fully deglycosylated by PNGase, but only the lower band of BDCA-2 can be digested by Endo H (Fig. 2.4B). Thus, the lower band represents a form bearing high mannose oligosaccharides, while the glycans on the upper band have been processed to complex glycans. Since high-mannose glycosylation occurs in the endoplasmic reticulum and further processing occurs after the membrane-targeted glycoprotein is transported into the Golgi apparatus, the difference in glycosylation patterns indicates part of the BDCA-2 extracted during this experiment had been transported to the cell surface, but part of the protein potentially retained in the endoplasmic reticulum.

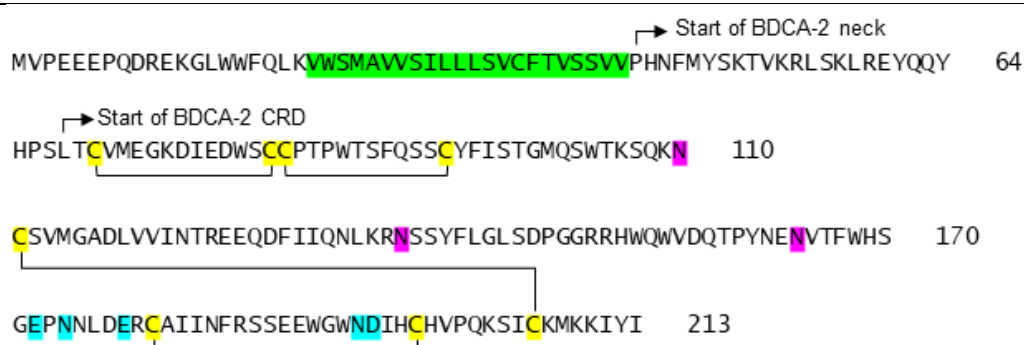


Figure 2.3. Polypeptide sequence of human BDCA-2. Transmembrane domains are highlighted in green. N-linked glycosylation sites are highlighted in magenta. Intramolecular disulfide bonds of the CRD are indicated by black lines. Residues that coordinate Ca²⁺ at the primary binding site are highlighted in cyan.

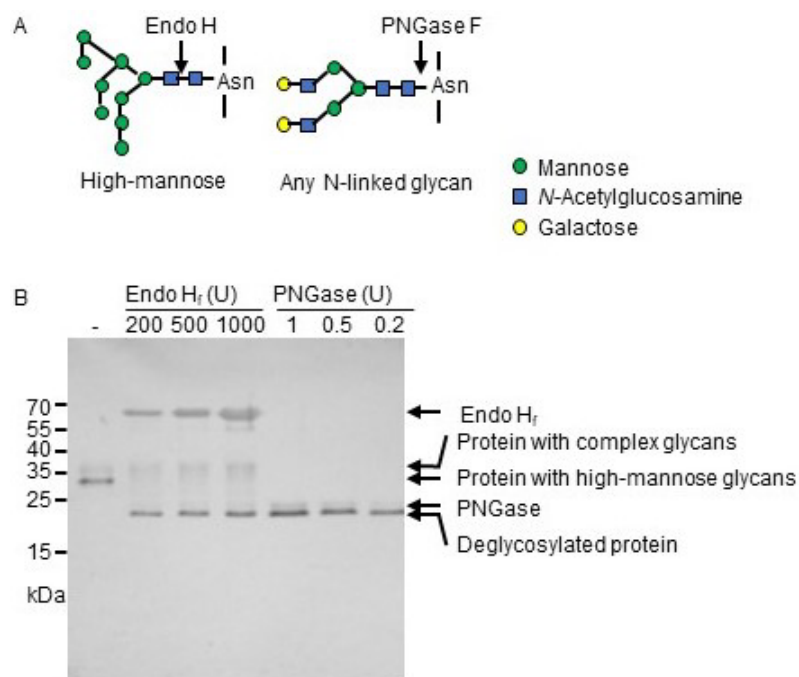


Figure 2.4. Membrane location of wild-type BDCA-2. (A) Deglycosylation reaction catalyzed by Endo H_r and PNGase F. (B) SDS-polyacrylamide gel electrophoresis of BDCA-2 digested with endoglycosidases. One 75 cm² flask of confluent cells were used. Purified BDCA-2 was dissolved in 0.5% SDS and 40 mM dithiothreitol. NP-40 and PNGase F or Endo H_r were added to the denatured protein. After 15 min incubation at 37 °C, samples were separated on an SDS-polyacrylamide gel and blotted onto nitrocellulose and visualized with antibody to BDCA-2. Experiment was performed twice.

Characterization of BDCA-2 lacking an intramembrane cysteine residue

Affinity purification of both wild-type BDCA-2 and the Cys35Ala mutant lacking the intramembrane cysteine residue, demonstrated co-elution of FcRγ in the major fractions (Fig. 2.5). A 50-kDa band was consistently detected on the anti-FcRγ blot, but not on the anti-BDCA-2 blot. This result suggests that the high-molecular weight band is not a complex of BDCA-2 and FcRγ, but so far it has not been possible to confirm the identity of this protein. FcRγ forms disulfide-linked dimers through a cysteine residue in the short extracellular domain. However, since *N*-ethylmaleimide and 2-mercaptoethanol were used during sample preparation, it is unlikely that this 50-kDa band was a disulfide-bonded complex between FcRγ and a nearby protein due to oxidation of the free cysteine.

Comparing the results for wild-type BDCA-2 and the Cys35Ala mutant lacking the intramembrane cysteine shows that this mutation does not affect association with FcRγ, because co-purified FcRγ can still be detected (Fig. 2.5A/B). The presence of two bands, representing glycosylation with either high mannose or complex N-linked glycans, also suggests that movement of the mutant out of the endoplasmic reticulum is not impaired.

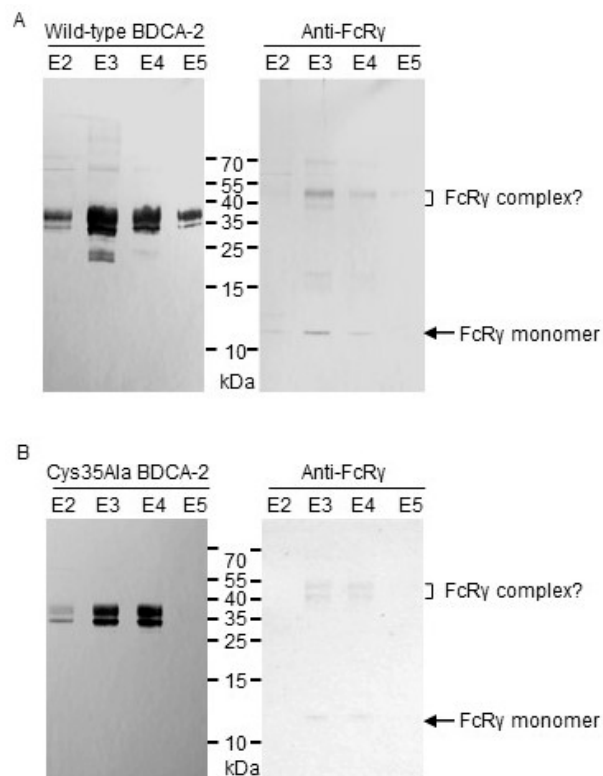


Figure 2.5. SDS-polyacrylamide gel electrophoresis of BDCA-2 lacking the intramembrane cysteine residue. BDCA-2 affinity purified from lysates of two 75 cm² flasks of confluent cells expressing (A) wild type or (B) mutant (Cys35Ala) BDCA-2 was analyzed by western blotting of SDS-polyacrylamide gels with anti-BDCA-2 and anti-FcRγ antibodies. Each experiment was performed twice.

Chemical crosslinking with *bis*(sulfosuccinimidyl) suberate (BS³) showed that removal of the intramembrane cysteine residue does not affect dimerization of BDCA-2 (Fig. 2.6C). The crosslinking pattern for the Cys35Ala mutant is identical to that seen before with the wild-type protein (Fig. 2.6B; Kim, 2018), demonstrating that dimer formation in the cells is independent of the presence of the intramembrane cysteine residue. For both proteins, the two glycosylated forms of BDCA-2 were crosslinked to generate two forms of dimeric BDCA-2. The fact that both the intracellular form of the protein, with high mannose oligosaccharides, and the form with processed complex glycans can both be crosslinked suggests that dimerization occurs in the endoplasmic reticulum.

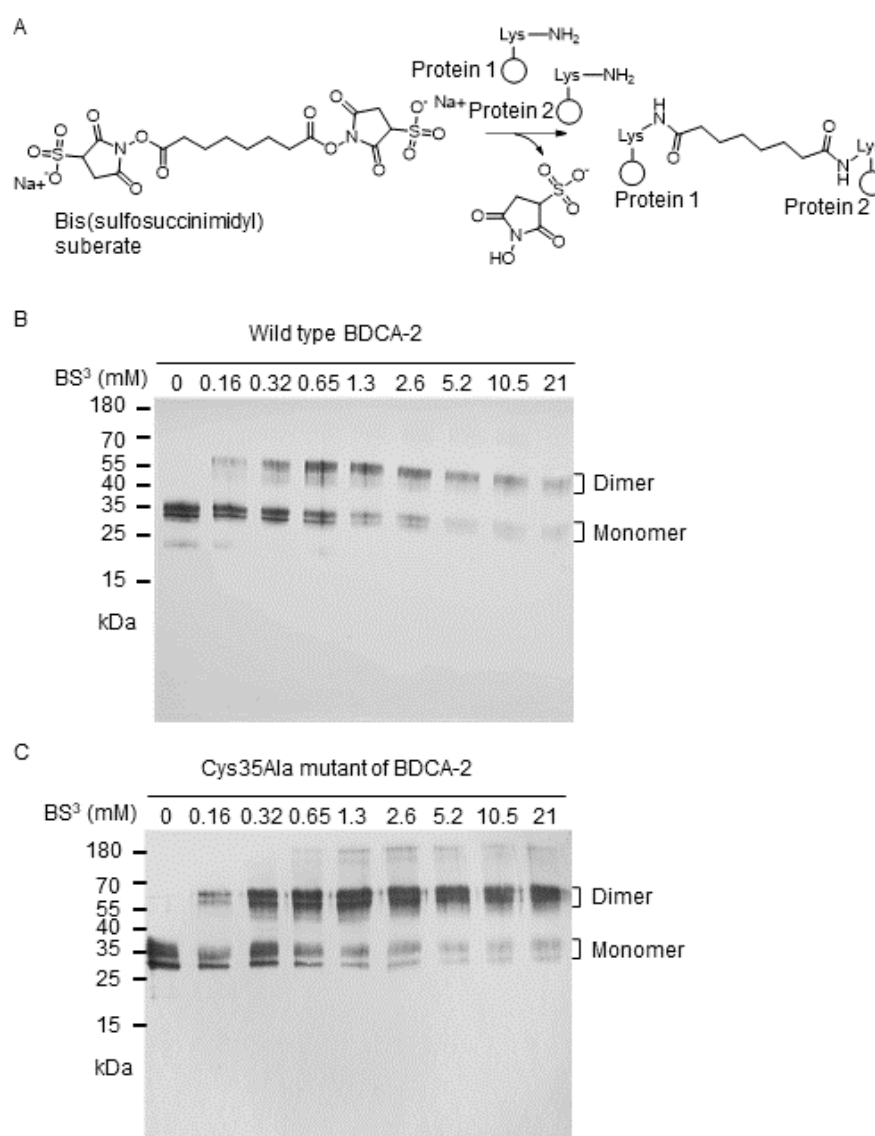


Figure 2.6. SDS-polyacrylamide gel electrophoresis of crosslinked BDCA-2. (A) The mechanism of BS³ crosslinking. (B) Wild-type BDCA-2 (data generated by Jong-Won Kim) and (C) the Cys35Ala mutant were treated with BS³. Each construct was purified from one 75 cm² flask of confluent cells. After incubation at room temperature for 30 min, samples were TCA-precipitated and separated by SDS-polyacrylamide gel electrophoresis in the presence of 2-mercaptoethanol and blotted onto nitrocellulose. The blots were developed with anti-BDCA-2 antiserum. Crosslinking experiment was performed twice for wild-type BDCA-2 and once for mutant BDCA-2.

Characterization of naturally occurring disulfide-bonded variant of BDCA-2

Data mining performed by Yu H Kwok revealed single nucleotide polymorphisms (SNP) (rs143790567) in the human BDCA-2 gene that lead to nonsynonymous amino acid changes Ser49Gly or Ser49Cys. The Ser49Cys variant is of particular interest, since a cysteine residue at this position would be comparable to the cysteine residue in the neck region of mincle, which contributes to interchain disulfide bond in mincle as described in Chapter 3 (Fig. 2.7). The Genome Aggregation Database recorded that the Ser49Cys allele occurs at a frequency of 0.000003984 (1 out of 251004 genomes). The individual who is heterozygous for this allele comes from a non-Finnish European ethnic background (Table 2.1).

These observations suggested that the Ser49Cys version of BDCA-2 might form an interchain disulfide bond. The ability for forming such a disulfide bond would provide evidence that, in the BDCA-2 dimer, the two copies of the neck region around residue 49 are close to each other. In order to study whether the Ser49Cys variant dimerizes in the same way mincle does, a full-length construct of BDCA-2 Ser49Cys variant was created for transfection into HEK cells. Affinity purification of the BDCA-2 Ser49Cys variant shows unperturbed FcR γ copurification and the characteristic double bands of BDCA-2, indicating that some receptors transit the Golgi *en route* to the plasma membrane whilst others were trapped in the endoplasmic reticulum (Fig. 2.8).

Comparison of reducing and non-reducing SDS-polyacrylamide gels show that the BDCA-2 Ser49Cys variant forms a disulfide-linked dimer that is dissociated in the presence of 2-mercaptoethanol (Fig. 2.9). Since the position of the cysteine in BDCA-2 Ser49Cys variant is analogous to that in the neck region of wild-type mincle, this result provides direct evidence that the BDCA-2 variant is likely to be displayed in a similar orientation as mincle on the membrane.

49

```

BDCA-2 wild type  MVPEEEPQDREKGLWWFQLK VWSMAVVSILLLSVCFTVSSVVP-HNFMYSKTVKRLSKLREYQQY
BDCA-2 variant   MVPEEEPQDREKGLWWFQLK VWSMAVVSILLLSVCFTVSSVVP-HNFMYSKTVKRLSKLREYQQY
Human Mincle     MNSSKSSETQCTERGCFSSQ MFLWTVAGIPIILFLSACFITRCVVTERIFQTCDEKKFQLPENFTE
  
```

Figure 2.7. Alignment of N-terminal sequences of mincle and BDCA-2 Ser49Cys variant. Transmembrane domains are highlighted in green and the amino acids on either side of position 49 are aligned as indicated by grey shading. Cysteine residues are highlighted in yellow.

Table 2.1. SNP frequency table for BDCA-2 Ser49Cys variant (rs143790567). Data is obtained from gnomAD v2.1.1.

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
European (non-Finnish)	1	113628	0	0.000008801
European (Finnish)	0	21648	0	0.000
African/African American	0	16254	0	0.000
Latino/Admixed American	0	34376	0	0.000
Ashkenazi Jewish	0	10072	0	0.000
East Asian	0	18350	0	0.000
South Asian	0	30570	0	0.000
Other	0	6106	0	0.000
XX	1	115314	0	0.000008672
XY	0	135690	0	0.000
Total	1	251004	0	0.000003984

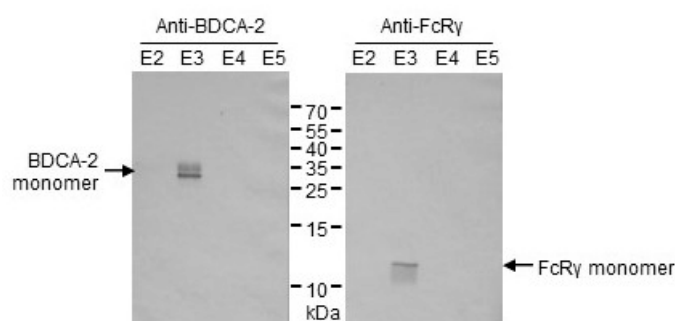


Figure 2.8. SDS-polyacrylamide gel analysis of affinity purified BDCA-2 Ser49Cys variant. BDCA-2 purified from one 75 cm² flask of confluent cells by affinity chromatography was examined by western blotting. Experiment was performed twice.

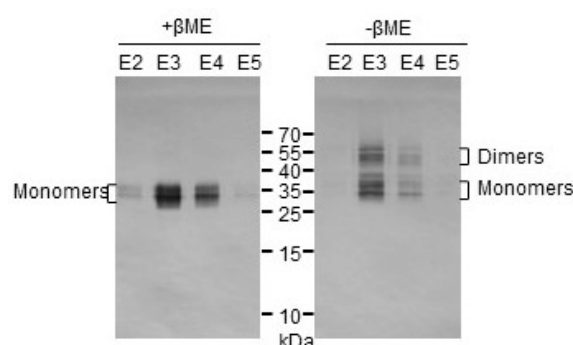


Figure 2.9. Disulfide-linked dimers of Ser49Cys form of BDCA-2 identified by SDS-polyacrylamide gel electrophoresis. BDCA-2 was purified from four 75 cm² flasks of confluent cells by affinity chromatography, TCA-precipitated, separated on SDS-polyacrylamide gels. Blots were developed with anti-BDCA-2 antiserum. Experiment was performed twice.

2.3 Discussion

Previous characterization of wild type BDCA-2 expressed in HEK cells, which employed N-ethylmaleimide to block free sulfhydryl groups, indicated that the intramembrane sulfhydryl group does not play an essential role in stabilizing receptor dimers or the complex with FcRγ. The fact that the sulfhydryl groups can form disulfide bonds under oxidizing conditions in the absence of the alkylating reagent also largely ruled out the possibility that this group is conjugated to another molecule, such as a fatty acid. However, the earlier experiments did not rule out the possibility that this group might be important in assembly of the complex. The experiments on wild type BDCA-2 and the artificially generated mutant lacking an intramembrane cysteine residue reported here demonstrate that this cysteine residue is not needed for dimer formation, association with FcRγ or movement out of the endoplasmic reticulum. The role of the conserved cysteine residue at the same position in the intramembrane portion of all of the mincle-family receptors therefore remains unexplained.

In contrast, analysis of the naturally occurring variant with a cysteine residue in the neck region provides important insights into the arrangement of BDCA-2 polypeptides. The ability of the variant form to make disulfide bonds provides direct evidence that the neck regions of two BDCA-2 polypeptides must interact closely in the dimer. This arrangement is likely to be similar to the arrangement of mincle, as described in Chapter 3. The observation that the neck is likely to form part of the dimer interaction provides a basis for further analysis of the arrangement of the extracellular domain described further in Chapter 5.

Chapter 3 | Identification of an interchain disulfide-forming cysteine residue in mincle

3.1 Introduction

Mincle, the macrophage-inducible C-type lectin was first reported as a transcriptional target of NF-IL6, a nuclear factor that can activate macrophages in a NO-independent pathway in defence against *Listeria* infections and tumour cytotoxicity (Matsumoto et al., 1999; Tanaka et al., 1995). Strong conservation of mincle in mammals, but not beyond this in birds or reptiles, suggests that mincle may have evolved in mammals as a consequence of selective pressures to overcome exposure to different sets of viruses, bacteria and fungi, as seen with the divergence of other mammalian-specific lectins in the process of evolution (Rambaruth et al., 2015).

Attempts to study mincle organization on the cell surface include an investigation by Lobato-Pascual and colleagues (2013), which showed disulfide bonded heterodimers of mincle and MCL can be formed in HEK cells cotransfected with both receptors. However, due to the low expression level of mincle when it was transfected alone or with FcRy in HEK transfectants, the oligomeric state of mincle itself has not previously been investigated. Thanks to the established affinity purification method for bacterially expressed mincle CRD (Feinberg et al., 2013), we were able to purify full-length mincle from HEK cells coexpressing mincle and FcRy and could perform analytical studies on the purified proteins.

This chapter describes the identification of mincle as a naturally occurring disulfide-linked homodimer. Since two of the potential inter-chain disulfide-forming cysteine residues reside in the transmembrane domain and one is in the stalk region (Fig. 2.1), it was important to define which cysteine residues are responsible for mincle dimerization. Although it has been reported that transmembrane cysteine residues are capable of forming inter-chain disulfide bonds, such as $\zeta\zeta$ dimer in the T cell receptor complex (Rutledge et al., 1992), the role of the neck cysteine residue in mincle dimerization was investigated initially because it is exposed to a more oxidizing environment. The significance of the neck cysteine residue in forming mincle homodimers was investigated by examining purified mutant proteins on reducing or non-reducing SDS-polyacrylamide gels, followed by immunoblotting with anti-mincle antibody. Understanding how mincle oligomerizes on the cell surface provides a model to be compared with related receptors, such as MCL, BDCA-2 and dectin-2. It also helps to build a picture of how the ligand-receptor interaction triggers immune signalling.

3.2 Results

Characterization of wild-type mincle

A plasmid encoding full-length mincle was digested with Not1 and Xba1, and the digestion product was inserted into a pSF-CMV-EMCV-Zeo vector (Oxford Genetics). Expression of mincle is driven by the cytomegalovirus promoter and selection is produced by the zeocin resistance gene that follows an internal ribosome re-entry site from the encephalomyocarditis virus (Fig. 3.1).

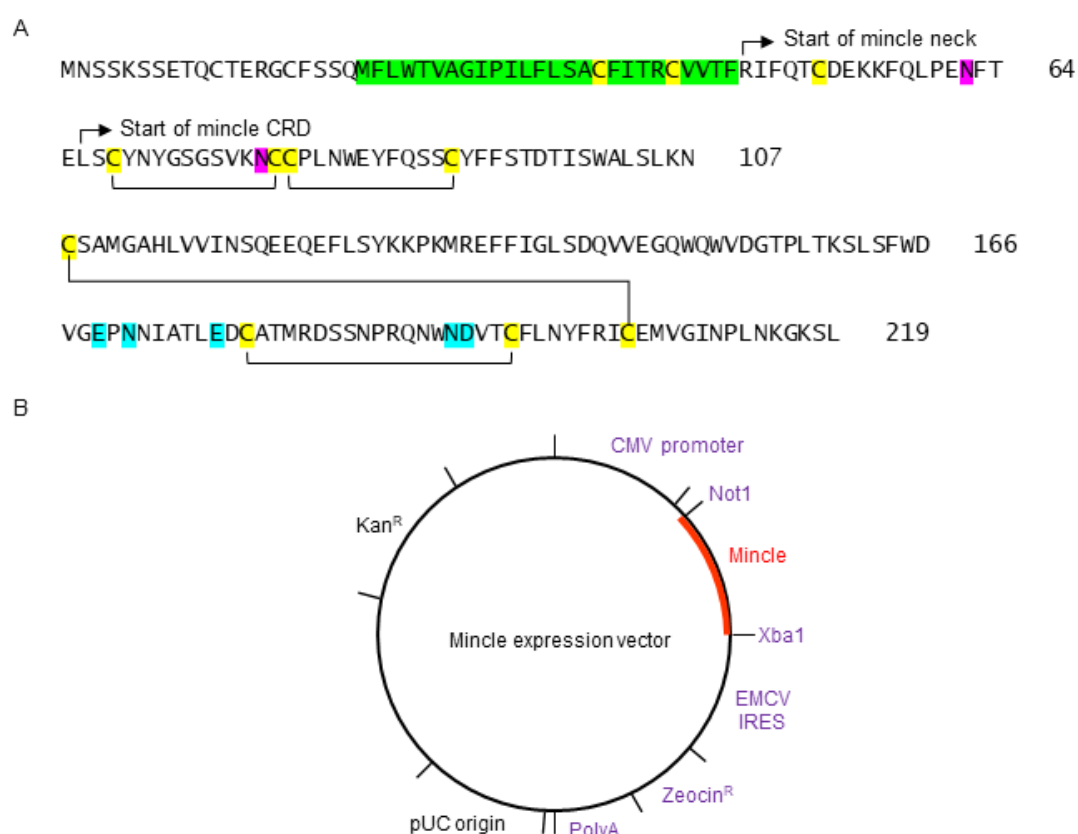


Figure 3.1. Polypeptide sequence and expression plasmid for wild-type mincle. (A) In the mincle sequence, the transmembrane domain is highlighted in green, cysteine residues are highlighted in yellow, glycosylation sites are highlighted in magenta and amino acids that form the primary sugar-binding site are highlighted in cyan. Locations of intramolecular disulfide bonds are indicated by lines. (B) Scheme of expression plasmid. Full-length cDNA for wild-type mincle was inserted between the Not1 and Xba1 sites.

Following cotransfection of mincle and FcRy expression plasmids into HEK 293 cells, cells expressing mincle and FcRy were selected for survival of single colonies that express dual resistance to zeocin and hygromycin. Semi-quantitative analysis of expression levels of mincle and FcRy was achieved by western blotting and cell lines expressing higher levels of both proteins were expanded and used for affinity purification on trehalose-Sepharose. Affinity purification of mincle showed coelution of FcRy in the major fractions (Fig. 3.2), which is

consistent with observations by Yamasaki and colleagues (2008). It was suggested that this association is dependent on the opposite charges on the receptor and the adaptor molecule, since a mutation of the positively charged arginine residue in mincle to a neutral isoleucine disabled the coupling with FcR γ (Yamasaki et al., 2008).

The intracellular localization of the mincle was investigated using glycosidases to monitor the state of processing of the two N-linked glycans attached to mincle. SDS-polyacrylamide gel analysis revealed that the glycans on purified wild-type mincle were removed upon treatment with PNGase, resulting in a shift of 10 kDa, consistent with loss of 2 N-linked glycans, and leaving a polypeptide of $M_r = 25.1$ kDa, as expected for unglycosylated mincle. In contrast, there was very little Endo H $_r$ -susceptible glycoprotein, which would correspond to material in the early secretory pathway, with high mannose oligosaccharides. These results suggest that almost all of the mincle molecules have undergone complete glycosylation and have been transported to the plasma membrane (Fig. 3.3).

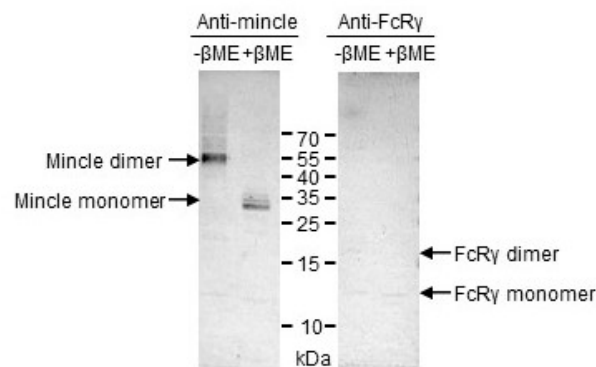


Figure 3.2. Western blotting of affinity purified wild-type mincle and co-purified FcR γ . Mincle was purified on a 2-ml trehalose-Sepharose column from lysates of one 75 cm² flask of confluent cells. Elution fractions 5 to 7 were pooled, precipitated with TCA, and separated on SDS-polyacrylamide gels in the absence or presence of 2-mercaptoethanol. Blots were developed with anti-mincle and anti-FcR γ antibodies. Experiment was performed three times.

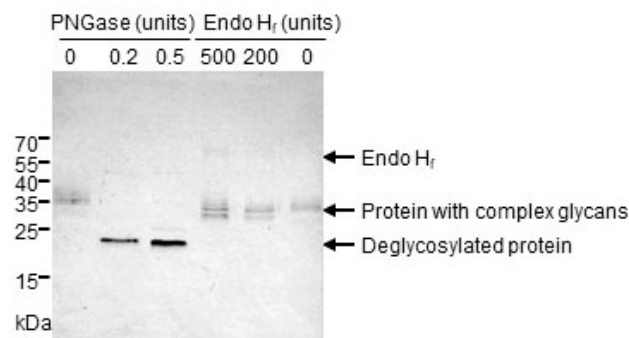
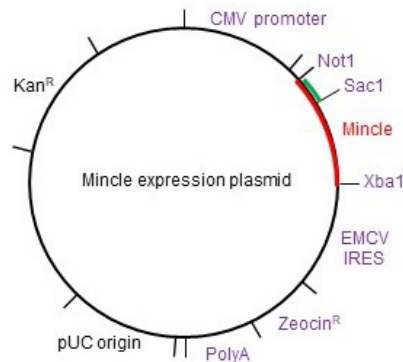


Figure 3.3. Western blotting of endoglycosidase-digested wild type mincle to determine its location within cells. Wild-type mincle was purified from one 75 cm² flask of confluent cells, TCA-precipitated, and dissolved in 0.5% SDS and 40 mM dithiothreitol. The denatured protein was digested with PNGase F or Endo H_f for 15 min at 37 °C and analyzed by western blotting with anti-mincle antibody. Experiment was performed once.

Characterization of mincle lacking cysteine in the extracellular neck region

In order to determine which cysteine residue is responsible for mincle dimerization, mutants with either only the single neck cysteine mutated to alanine (Cys52Ala) or with both of the transmembrane cysteine residues replaced by alanine (Cys37Ala and Cys42Ala) were generated as shown in Fig. 3.4. DNA segments from the NotI site to the SacI site that code for mutation of cysteine residues in either the transmembrane domain or the neck domain (Fig. 3.4) were synthesized and inserted into NotI and SacI-digested expression plasmid. The resulting plasmids were co-transfected with FcR γ into HEK 293 cells.

A



B Wild-type mincle

MetAsnSerSerLysSerSerGluThr
 gcggccgc cagtgtgatggatatctgcagaattcgcccttaactcgagatgaattcatctaaatcatctgaaaca
 Not1
 GlnCysThrGluArgGlyCysPheSerSerGlnMetPheLeuTrpThrValAlaGlyIleProIleLeuPheLeu
 caatgcacagagagaggatgcttctcttcccaaagtgtcttatggactgttgctgggatcccatcctatttctc
 SerAlaCysPheIleThrArgCysValValThrPheArgIlePheGlnThrCysAspGlyLysLysPheGlnLeu
 agtgccgtttcatcaccagatgtgtgtgacatttcgcacatctttcaaacctgtgatgagaaaaagtttcagcta
 ProGluAsnPheThrGluLeu
 cctgagaatttcacagagctc
 Sac1

C Transmembrane cysteine mutant of mincle

MetAsnSerSerLysSerSerGluThr
 gcggccgc cagtgtgatggatatctgcagaattcgcccttaactcgagatgaattcatctaaatcatctgaaaca
 Not1
 GlnCysThrGluArgGlyCysPheSerSerGlnMetPheLeuTrpThrValAlaGlyIleProIleLeuPheLeu
 caatgcacagagagaggatgcttctcttcccaaagtgtcttatggactgttgctgggatcccatcctatttctc
 SerAlaAlaPheIleThrArgAlaValValThrPheArgIlePheGlnThrCysAspGlyLysLysPheGlnLeu
 agtgccgctttcatcaccagagcgtgtgtgacatttcgcacatctttcaaacctgtgatgagaaaaagtttcagcta
 ProGluAsnPheThrGluLeu
 cctgagaatttcacagagctc
 Sac1

D Neck cysteine mutant of mincle

MetAsnSerSerLysSerSerGluThr
 gcggccgc cagtgtgatggatatctgcagaattcgcccttaactcgagatgaattcatctaaatcatctgaaaca
 Not1
 GlnCysThrGluArgGlyCysPheSerSerGlnMetPheLeuTrpThrValAlaGlyIleProIleLeuPheLeu
 caatgcacagagagaggatgcttctcttcccaaagtgtcttatggactgttgctgggatcccatcctatttctc
 SerAlaCysPheIleThrArgCysValValThrPheArgIlePheGlnThrAlaAspGlyLysLysPheGlnLeu
 agtgccgtttcatcaccagatgtgtgtgacatttcgcacatctttcaaaccgcgatgagaaaaagtttcagcta
 ProGluAsnPheThrGluLeu
 cctgagaatttcacagagctc
 Sac1

Figure 3.4. Mutagenesis strategy for making transmembrane cysteine and neck cysteine mutants of mincle. (A) Map of mincle expression plasmid showing region modified in cysteine mutants. (B) Sequence of wild-type mincle cDNA encoding cytoplasmic, transmembrane and neck domains. (C) and (D) Sequences of synthetic replacement cDNAs, flanked by Not1 and Sac1 sites, that encode mutated forms of transmembrane and neck domains. Transmembrane domain of the protein sequence is highlighted in green. Mutations in the DNA sequences are coloured in red.

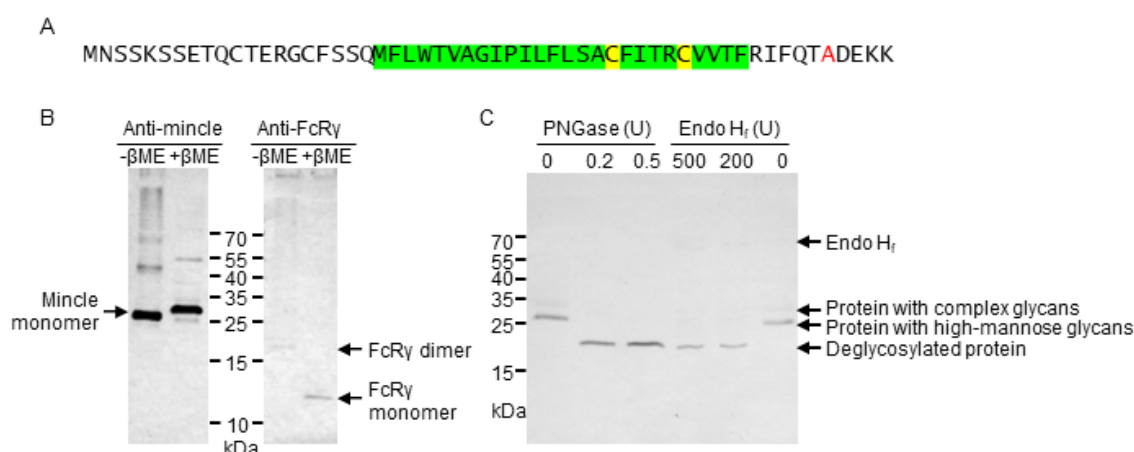


Figure 3.5. Characterization of oligomeric status and membrane location of affinity-purified neck cysteine mutant of mincle and its copurification with FcR γ . (A) Sequence of mincle with mutated neck cysteine. (B) Western blot analysis of mutant mincle purified by affinity chromatography on trehalose-Sepharose and TCA-precipitated. One 75 cm² flask of confluent cells were used. Blots were developed with anti-mincle and anti-FcR γ antibodies. (C) Western blot analysis of mutant mincle digested with endoglycosidases after purification from one 75 cm² flask of confluent cells. After 15 min incubation at 37 °C, samples were separated on an SDS-polyacrylamide gel in the presence of 2-mercaptoethanol. Mincle was detected with anti-mincle antibody. Each experiment was performed twice.

Characterization of mincle lacking the cysteine residue in the neck revealed that elimination of this cysteine residue almost completely abolished dimer formation (left panel, Fig. 3.5B), demonstrating that this disulfide is essential for dimerization. Association with FcR γ has not been affected in this mutant (right panel, Fig. 3.5B). However, endoglycosidase digestion revealed that all of the mutant mincle is susceptible to Endo H_f digestion, indicating that this mutant bears high-mannose glycans (Fig. 3.5C). Thus, in the absence of formation of disulfide-crosslinked dimers, the protein is retained in the endoplasmic reticulum instead of being expressed on the cell surface.

Chemical crosslinking was used to determine if the mutant mincle forms dimers that are not stabilized by interchain disulfide bonds. After crosslinking, SDS-polyacrylamide gel electrophoresis did not show trapping of dimer by the crosslinking reagent (Fig. 3.7A). The results suggest in the absence of the neck cysteine residue, mincle does not form a stable non-disulfide linked dimer. This contrasts with the results in the previous chapter showing that BDCA-2 dimer does not depend on formation of disulfide bonds.

Characterization of mincle lacking intramembrane cysteine residues

The results in the previous section show that the two intramembrane cysteine residues do not support dimer formation in the absence of the neck region cysteine. In order to see if they play a role in dimer formation, the effect of mutating the intramembrane cysteine residues was explored. Affinity purification showed that dimers are still observed in the absence of both of

the intramembrane cysteine residues (left panel, Fig. 3.6B), although at a reduced level compared to wild type. As with the neck region mutant, copurification of FcR γ was still observed (right panel, Fig. 3.6B). Glycosylation analysis of mincle lacking the intramembrane cysteine residues revealed three protein bands in the untreated sample (Fig. 3.6C). The top band and a small proportion of the middle band were Endo H $_r$ -resistant while the majority of the middle band and all of the bottom band were shifted down upon Endo H $_r$ treatment. The segregation of the middle band could be explained if it is composed of two species of protein molecules. The difference in molecular weights between the Endo H $_r$ -resistant species in the middle band and the intact protein could be due to proteolytic cleavage, while the molecular-weight difference between the Endo H $_r$ -prone species in the middle band and the intact protein could be explained by deglycosylation of high-mannose glycans. The existence of the bottom band could be explained by a combination of proteolysis and deglycosylation of high-mannose glycans. Chemical crosslinking with BS³ also confirmed the presence of dimer (Fig. 3.7B).

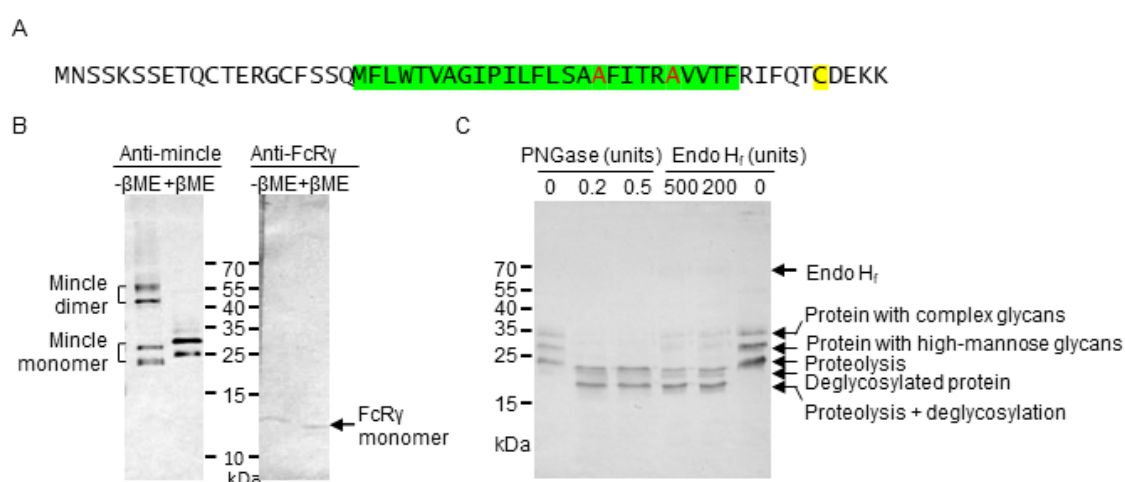


Figure 3.6. Characterization of oligomeric status and membrane location of affinity-purified intramembrane cysteine mutant of mincle and its copurification with FcR γ . (A) Sequence of mincle intramembrane cysteine mutant. (B) Western blot analysis of intramembrane cysteine mutant of mincle purified from one 75 cm² flask of confluent cells by affinity chromatography on trehalose-Sepharose and separated on SDS-polyacrylamide gels in the absence or presence of 2-mercaptoethanol. (C) Western blot analysis of intramembrane cysteine mutant of mincle digested with endoglycosidases after purification. One 75 cm² flask of confluent cells were used. Samples were separated on an SDS-polyacrylamide gel in the presence of 2-mercaptoethanol. Each experiment was performed twice.

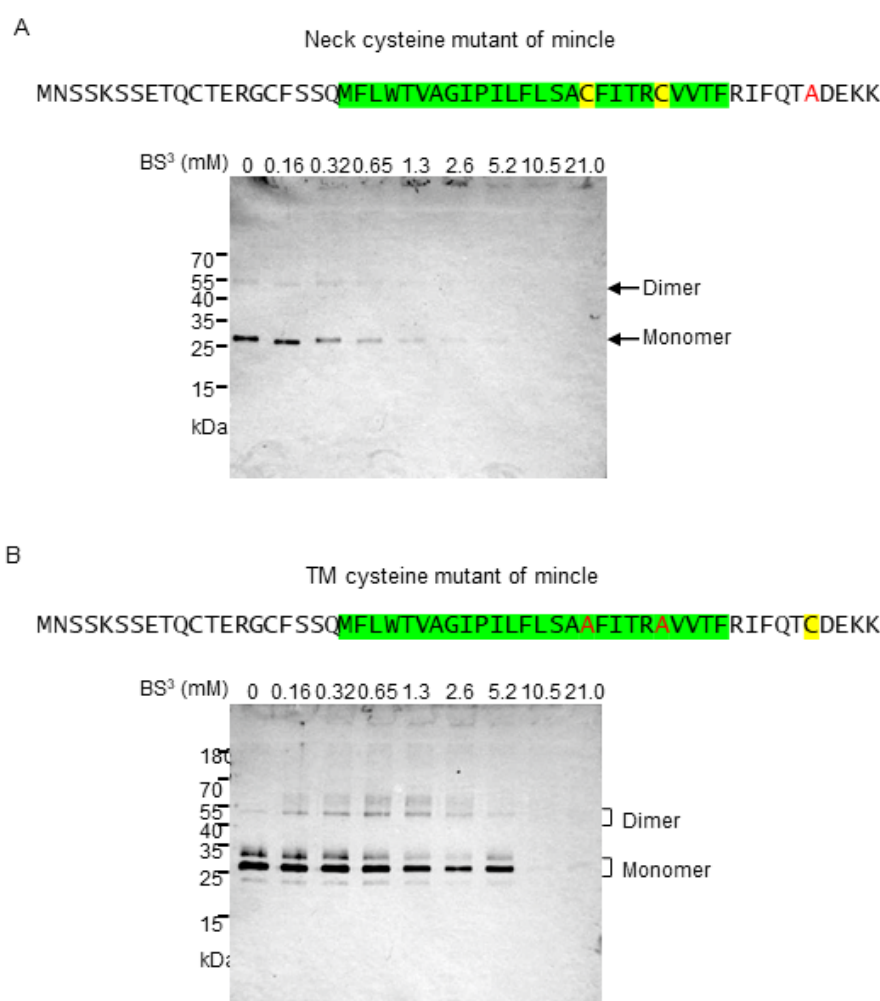


Figure 3.7. BS³-crosslinking of mincle mutants. Western blot analysis of mincle lacking the cysteine residue in the neck (A) or cysteine residues in the transmembrane domain (B). One 75 cm² flask of confluent cells were used for each mutant. Proteins were purified from by affinity chromatography, transferred to 150 mM NaCl, 25 mM HEPES pH 7.8, 2.5 mM EDTA, and treated with crosslinking reagent. After 30 min incubation at room temperature, samples were precipitated by trichloroacetic acid and separated on SDS-polyacrylamide gels in the presence of 2-mercaptoethanol. Blots were developed with anti-mincle antibody. Each experiment was performed once.

3.3 Discussion

In summary, the work presented in this chapter shows for the first time that wild-type mincle forms a disulfide-linked homodimer. Further investigation of the role of cysteine residues in the transmembrane domain and the neck region of mincle by creating HEK cell lines transfected with mincle mutants confirmed that the neck cysteine residue is indispensable for forming the interchain disulfide bond. Although the transmembrane cysteine residues are not necessary for mincle dimerization, emergence of heterogeneous species in the transmembrane cysteine mutant of mincle suggests these cysteines may help with the correct processing of mincle and transport to the cell surface.

Chapter 4 | Extracellular domain fragments of mincle

4.1 Introduction

The results described in the previous two chapters show that interactions in the neck regions of the mincle family receptors form part of the dimer interface in these receptors and that this interface includes an interchain disulfide bond in the cases of mincle and the variant form of BDCA-2. Bacterial systems for expression of the dimeric extracellular domains could provide an important tool to facilitate further analysis of how the dimers are organized and how the CRDs are presented at the cell surface. This chapter describes several different approaches that were used to develop and optimize a system for expressing receptor dimers. Mincle was chosen as the target for the initial experiments based on the evidence presented in Chapter 3, showing that dimer formation is accompanied by formation of an interchain disulfide bond in the neck of mincle. Formation of disulfide-bonded covalent dimers of mincle polypeptides could thus be used as an initial indication of successful expression of the target oligomer.

As discussed in more detail below, much of this work was based on appending dimerization domains to mincle. Coiled coils of α -helices are widely used to engineer multi-stranded structures (Doñate et al., 2000). The α -helices adopt characteristic seven-residue repeats, *abcdefg*, where the first and fourth positions are hydrophobic residues. In this way, the hydrophobic side chains from one α -helix can be packed into holes formed by four other residues from a neighbouring helix (Woolfson et al., 2012). The oligomeric state of coiled coils is significantly impacted by their hydrophobic cores, such that a coiled-coil structure will adopt a dimeric topology when positions *d* are leucine, trimeric topology when both positions are isoleucine, and tetrameric topology when positions *a* and *d* are leucine and isoleucine (Fig. 4.1) (Harbury et al., 1993).

More recent studies have demonstrated that autonomous assembly of coiled-coil structures from two or more inter-winding α -helices is initiated at a specific region of the peptide sequence, termed a trigger sequence, towards the C terminus of the coiled-coil structure. This theory was developed by Fletcher and colleagues (2012), who addressed the importance of placing an asparagine residue at position *a* in the first heptad of the trigger sequence instead of a default isoleucine to allow formation of stable dimeric coiled coils. Others have reported that when an E-*abcdK* motif is changed to R-*abcdE* in the trigger sequence of a GCN4 coiled coil, its topology is changed from a dimer to a trimer (Ciani et al., 2010). Using these concepts, amino acid sequences for forming 2-helix, 3-helix, and 4-helix coiled-coil structures have been extensively characterized (Fletcher et al. 2012).

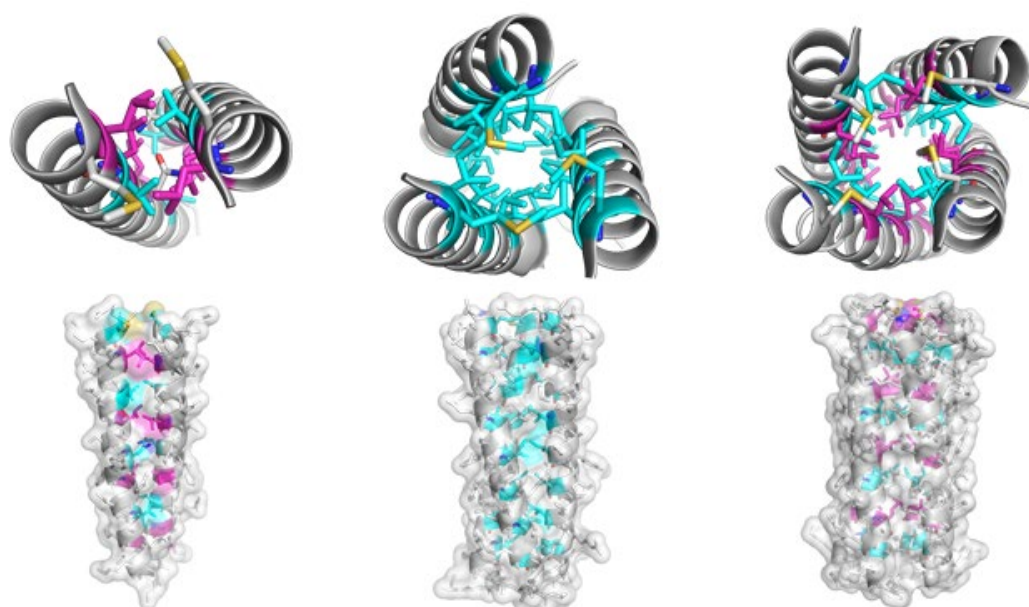


Figure 4.1. Dimeric, trimeric, and tetrameric coiled-coil structures of GCN4 dimerization domains. In the dimeric structure (PDB entry: 2ZTA), leucine residues at position *d* are coloured magenta. In the trimeric structure, isoleucine residues at both positions *a* and *d* are coloured cyan (PDB entry: 1GCM). In the tetrameric structure, isoleucine residues at position *a* are coloured cyan and leucine residues at position *d* are coloured magenta (PDB entry: 1GCL).

4.2 Results

Direct expression of mincle extracellular domain

Earlier studies on human mincle CRD employed an expression system in which the protein was fused to a bacterial signal sequence and secreted into the periplasm, where it folds into a native conformation with intrachain disulfide bonds (Jégouzo et al., 2014). Previous attempts to adapt this system for expression of the extracellular domain were unsuccessful. Therefore, an alternative expression system was tested, in which the protein is expressed without a signal sequence and accumulates in the cytoplasm as inclusion bodies that can be solubilized in guanidine and renatured. During the renaturation process, disulfide bonds are formed by air oxidation. Initial attempts to express a disulfide-linked dimer of the mincle extracellular domain were undertaken with an expression construct containing the portion of the human mincle cDNA coding for residue Thr45 to Leu219 in a T7 expression system (Fig. 4.2A). Purification of correctly folded protein was undertaken on a trehalose-Sepharose column.

Although the protein was successfully purified on the affinity column, only a small amount of dimer formation was observed on a non-reducing SDS-polyacrylamide gel and the majority of protein remained as a monomer (Fig. 4.2B). Crosslinking of the purified protein with Cu^{2+} -o-phenanthroline was tested as a way to induce disulfide bond formation. This reagent catalyzes air oxidation of sulfhydryl groups (Steck, 1972). Because a low concentration of human mincle may have contributed to inefficient interchain disulfide bond formation, the purified protein was

concentrated ten-fold, and then subjected to treatment with Cu^{2+} -*o*-phenanthroline. However, neither a high concentration nor crosslinking with Cu^{2+} -*o*-phenanthroline enhanced disulfide bond formation (Fig. 4.2C).

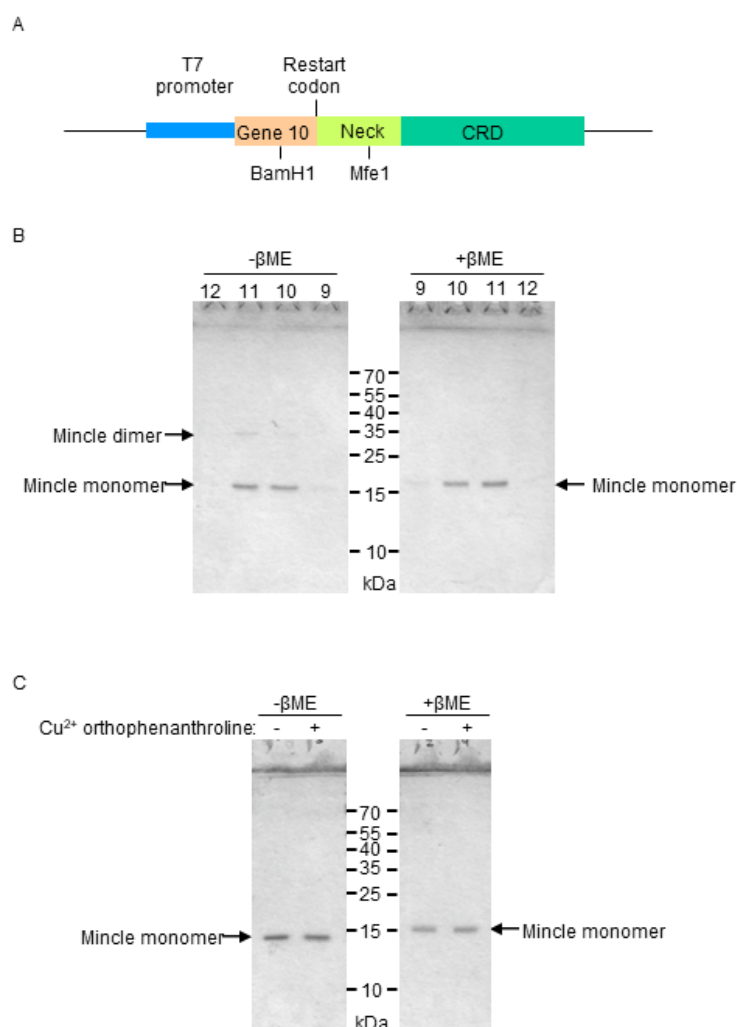


Figure 4.2. Bacterial expression system and SDS-polyacrylamide gel analysis of human mincle extracellular domain. (A) Expression of human mincle extracellular domain under the control of a T7 promoter. (B) SDS-polyacrylamide gel electrophoresis of human mincle extracellular domain renatured by dialysis and affinity-purified over a 10-ml trehalose-Sepharose column. Protein was eluted from the column with EDTA, in 1-ml fractions. Aliquots (15 μ l) of elution fractions 9 to 12 were separated on SDS-polyacrylamide gels in the absence or presence of 2-mercaptoethanol. (C) SDS-polyacrylamide gel electrophoresis of Cu^{2+} -*o*-phenanthroline-treated human mincle extracellular domain. Sample was analyzed in the absence or presence of 2-mercaptoethanol. Gels were stained with Coomassie blue.

Preparation of mincle extracellular domain dimers with a GCN-4 dimerization domain

In the case of mincle, a dimer-forming coiled-coil structure is desired. The coiled-coil dimerization domain from yeast transcription factor GCN-4 is 33 amino acids in length, consisting of four heptads. Experiments with this dimerization domain were undertaken with the extracellular portion of cow mincle instead of human mincle, because previous studies showed that cow mincle CRD is expressed much more efficiently in the T7 expression system. Mincle sequences in cow and human share a high degree of identity (Fig. 4.3) and show very similar Ca^{2+} - and pH-dependent binding properties with trehalose ligand (Jégouzo et al., 2014). The structure of cow mincle CRD has been successfully used to model the human protein, suggesting the orientation of CRDs in a cow mincle extracellular domain dimer will show how CRDs are projected in the human orthologue (Feinberg et al., 2016).

As with human mincle, non-reducing SDS-polyacrylamide gels of affinity-purified cow mincle extracellular domain without a GCN4 dimerization domain shows very little disulfide dimer is formed (Fig. 4.4A). In contrast, analysis of cow mincle extracellular domain fused to a GCN4 dimerization domain shows the majority of mincle formed disulfide-linked dimers that can be dissociated in the presence of 2-mercaptoethanol (Fig. 4.4B). Gel-filtration analysis revealed that despite some aggregation, the major fraction has an apparent molecular weight of 47 kDa (Fig. 4.5), which matches to the expected molecular weight of dimeric mincle extracellular domain plus the GCN4 sequence (2 x 23.5 kDa).

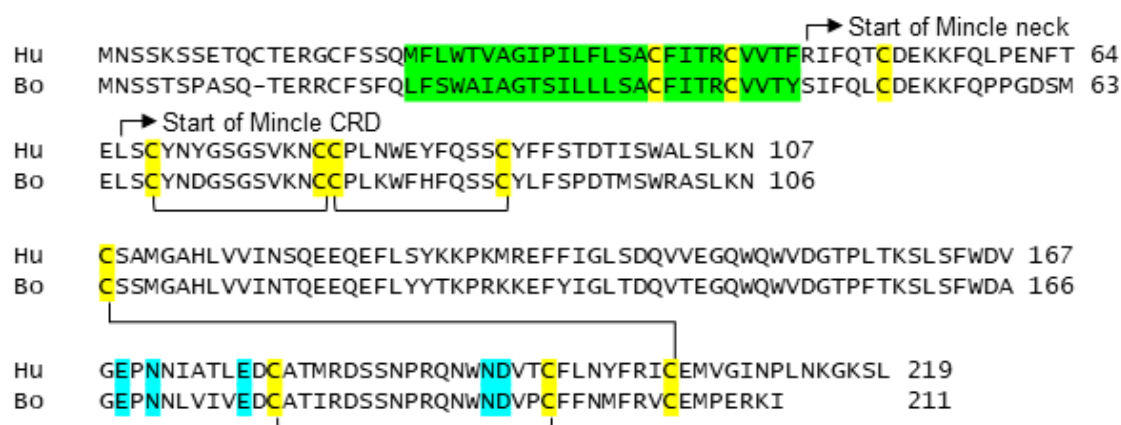


Figure 4.3. Alignment of polypeptide sequences of human and bovine mincle. Transmembrane domains are highlighted in green. Intramolecular disulfide bonds of the CRD are indicated by black lines. Residues that coordinate Ca^{2+} at the primary sugar binding site are highlighted in cyan.

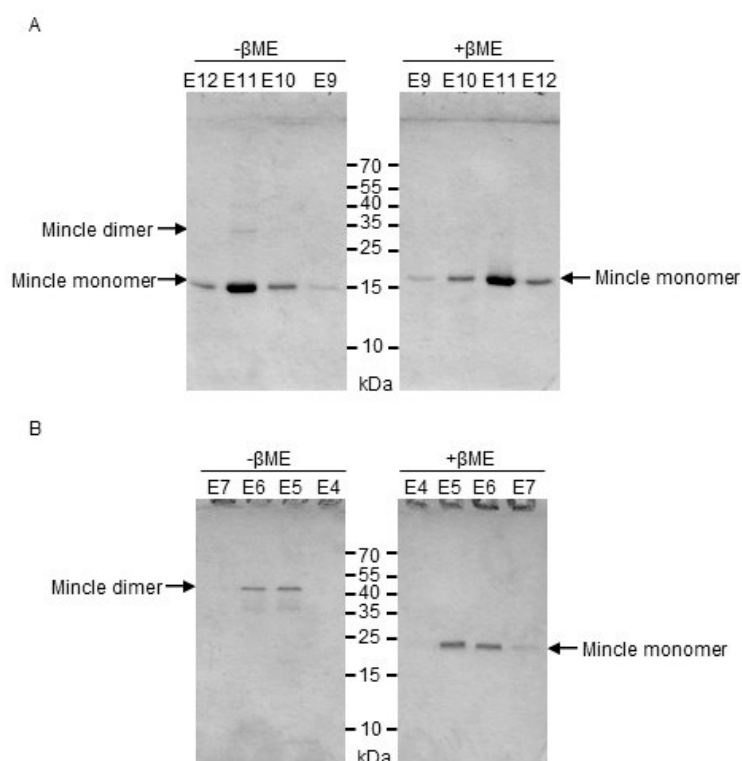


Figure 4.4. Bovine mincle extracellular domain without and with GCN4 dimerization domain analyzed on SDS-polyacrylamide gels. (A) A 2-L preparation of mincle extracellular domain was renatured by direct dialysis and purified over a 10-ml trehalose-Sepharose column. Aliquots (15 μ l) of 1-ml elution fractions (E9-E12) were separated on an SDS-polyacrylamide gel in the absence or presence of 2-mercaptoethanol. (B) A 2-L preparation of mincle extracellular domain with GCN4 dimerization domain was renatured by batch dilution at a 1:5 ratio and purified over a 2-ml trehalose-Sepharose column. Aliquots (15 μ l) of 0.5-ml elution fractions (E4-E7) were separated on an SDS-polyacrylamide gel in the absence or presence of 2-mercaptoethanol. Gels were stained with Coomassie blue.

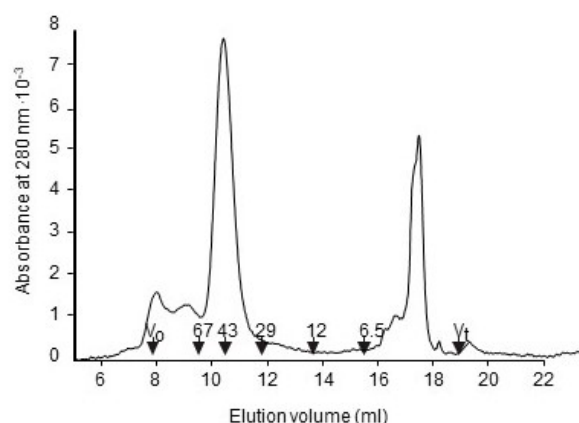


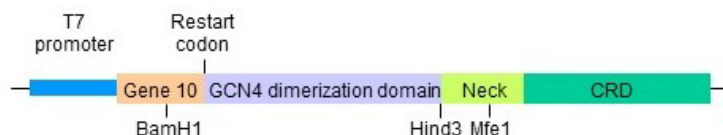
Figure 4.5. Gel-filtration analysis of bovine mincle extracellular domain fused to GCN4 dimerization domain. Elution fraction 5 (0.5 ml) was used to load the Superdex S75 column. Elution positions of standards on Superdex S75 column are indicated in kDa at the bottom.

Attempts to concentrate the initial GCN4-stabilized dimer by dialysis against water and lyophilization were unsuccessful, since the lyophilized protein could not be re-dissolved in

aqueous buffer (data not shown). The undesired solubility profile compared to the isolated CRD could arise from either the presence of portions of the neck from mincle or from the hydrophobicity of the 4-heptad dimerization domain. To test if reducing the neck of mincle improves solubility, the expression vector was modified to include a Hind3 restriction site at the end of the coding sequence for the GCN4 dimerization domain, which facilitates generation of expression constructs with different lengths of mincle neck. Synthetic oligonucleotide pairs were then generated to replace the portion of the expression vector between the Hind3 site and a naturally occurring Mfe1 site in the portion of the cDNA encoding the neck region of mincle (Fig. 4.6A).

A total of 8 different versions of the expression construct were created. Versions D1 to D3 contain one of the transmembrane cysteine residues, D4 and D5 cover part of the transmembrane domain, and versions D6-D8 encode shorter versions of the neck (Fig. 4.6B-D). Each of the constructs was expressed, renatured and purified on trehalose-Sepharose. Non-reducing SDS-polyacrylamide gels showed that D1, D2, D3, D4, D5 and D6 formed multiple oligomeric states, whilst the majority of D7 and D8 formed dimers (Fig. 4.7). The ratio of dimer to monomer is significantly lower in D1, D2, D3, than in D4 and D5. Comparing D6, D7 and D8, the yield and the homogeneity of dimer improves with shorter necks. Consistent with the appearance of disulfide-linked dimers on non-reducing SDS-polyacrylamide gels, gel-filtration analysis of mincle extracellular domain D8 revealed the peak of elution corresponding to a molecular weight of 51 kDa (Fig. 4.8), which is close to the expected molecular weight of dimeric mincle extracellular domain D8, 45 kDa (2×22.7 kDa). Based on the combined results from SDS-polyacrylamide gels and gel-filtration analysis, the neck length of the mincle ECD D8 construct was chosen for protein expression and purification in subsequent experiments.

A Organization of GCN4-mincle expression vector



B GCN4 dimerization domain

MetLysGlnLeuGluAspLysValGluGluLeuLeuSerLys
atgaagcagctggaagataaagttgaagaactgctgagcaaa

AsnTyrHisLeuGluAsnGluValAlaArgLeuLysLysLeu
aactaccacctggaaaacgaagttgcgcgtctgaaaagagctt
Hind3

C Bovine mincle transmembrane domain and start of neck

LeuLeuLeuSerAlaCysPheIleThrArgCysValValThrTyrSerIlePheGlnLeu
ctgctcctgagtgccctgttttatcaccagatgtgttgacatatagcatctttcaattg
Mfe1

D Inserted synthetic oligonucleotides

BoMincleD1	LysLeuThrArgCysValValThrTyrSerIlePheGlnLeu
D1 forward	agcttaccagatgtgttgacatatagcatctttc
D1 reverse	atggctctacacaacactgtatatcgtagaaaggttaa
BoMincleD2	LysLeu ArgCysValValThrTyrSerIlePheGlnLeu
D2 forward	agcttagatgtgttgacatatagcatctttc
D2 reverse	tctacacaacactgtatatcgtagaaaggttaa
BoMincleD3	LysLeu CysValValThrTyrSerIlePheGlnLeu
D3 forward	agctttgtgttgacatatagcatctttc
D3 reverse	acacaacactgtatatcgtagaaaggttaa
BoMincleD4	LysLeu ValValThrTyrSerIlePheGlnLeu
D4 forward	agcttggttgacatatagcatctttc
D4 reverse	caacactgtatatcgtagaaaggttaa
BoMincleD5	LysLeu ValThrTyrSerIlePheGlnLeu
D5 forward	agcttgacatatagcatctttc
D5 reverse	cactgtatatcgtagaaaggttaa
BoMincleD6	LysLeu ThrTyrSerIlePheGlnLeu
D6 forward	agcttatagcatctttc
D6 reverse	tgtatatcgtagaaaggttaa
BoMincleD7	LysLeu TyrSerIlePheGlnLeu
D7 forward	agctttatagcatctttc
D7 reverse	atatcgtagaaaggttaa
BoMincleD8	LysLeu SerIlePheGlnLeu
D8 forward	agcttagcatctttc
D8 reverse	tcgtagaaaggttaa

Figure 4.6. Cloning strategy for different lengths of mincle neck following GCN4 dimerization domain. (A) Organization of the expression vector for bovine mincle with an N-terminal GCN4 dimerization domain. Key restriction sites are indicated. (B) GCN4 dimerization domain, with heptad repeat hydrophobic residues highlighted in grey. The second line corresponds to the trigger sequence. (C) Sequence of bovine mincle at the end of the transmembrane domain, highlighted in green, and the beginning of the neck domain. Cysteine residues are shaded yellow. (D) Forward and reverse oligonucleotides inserted between the Hind3 and Mfe1 sites are shown, along with the encoded amino acid sequences.

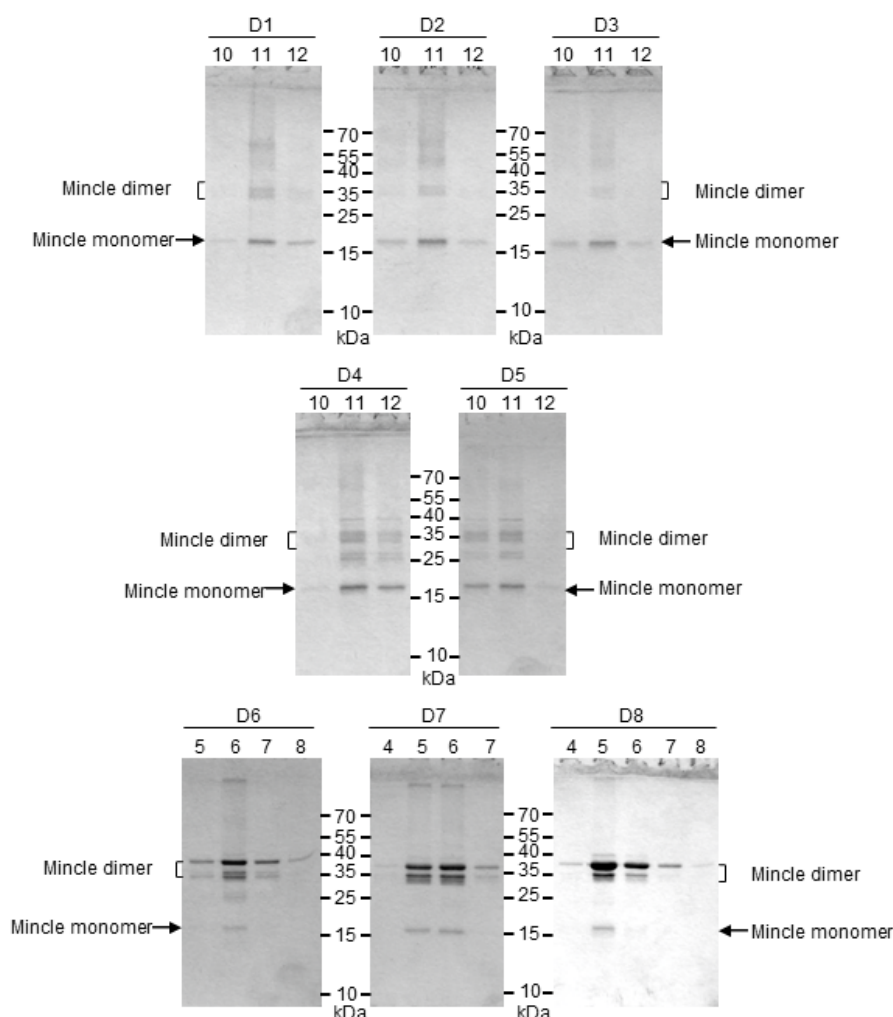


Figure 4.7. Non-reducing SDS-polyacrylamide gels of affinity-purified GCN4-mincle extracellular domain constructs. Each construct was prepared from 2 l of bacterial culture. Renaturation of constructs D1, D2, D3, D4 and D5 followed a direct dialysis protocol and purification was performed over a 10-ml trehalose-Sepharose column, from which fifteen 1-ml fractions were collected. The majority of protein was eluted in fractions 10 to 12. Renaturation of constructs D6, D7 and D8 followed a batch dilution protocol and purification was performed over a 2-ml trehalose-Sepharose column, where ten 0.5-ml fractions were collected. The majority of protein was eluted in fractions 4 to 8. Samples were separated on SDS-polyacrylamide gels in the absence of 2-mercaptoethanol. Gels were stained with Coomassie blue.

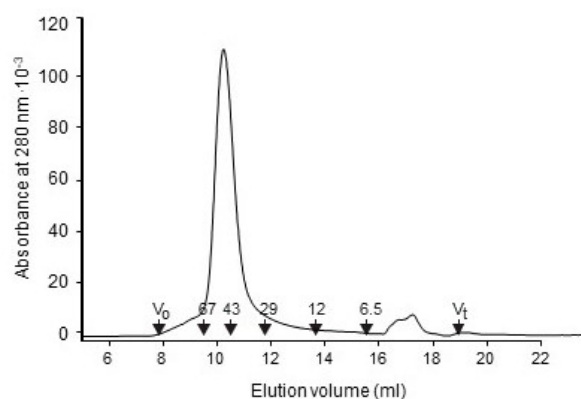


Figure 4.8. Gel-filtration analysis of BoMincle ECD D8. Elution fraction 5 (0.5 ml) from a 2-L preparation was loaded on Superdex S75 column. Elution positions of standards on Superdex S75 column are indicated in kDa at the bottom.

Replacing GCN-4 dimerization domain with computationally-optimized dimerization domains

Although the modified GCN4 fusion proteins with shorter necks showed improved properties following initial purification, after dialysis and lyophilization the protein still exhibited poor solubility at high concentration (data not shown). This result in the previous section suggested that the solubility issues of the mincle dimers might result from the behaviour of the GCN4 domain itself. Therefore, computationally optimized dimerization domains were explored as an alternative way to promote dimerization of mincle extracellular domain.

Based on the observation that the neck length in Mincle ECD D8 encouraged formation of more homogenous dimers than other neck lengths, this region was fused to computationally-optimized dimerization domains based on the sequence designated CC-pIL-I17N (Fletcher et al. 2012). The cloning strategy involved synthesis of oligonucleotide pairs that encode the 3-heptad version of CC-pIL-I17N. These oligonucleotides are flanked with a BamH1 site and a Hind3 site, which allows them to replace the GCN4 coding sequence in the GCN4-mincle extracellular domain construct. A new feature of these synthetic nucleotide sequences is the inclusion of an Msc1 site immediately after the start codon for the dimerization domain sequence. This Msc1 site facilitated subsequent replacement of the coding sequence for the 3-heptad dimerization domain with synthetic oligonucleotide pairs that encode different numbers of heptad-repeats (Fig. 4.9).

Oligonucleotides for 3 heptad (Mincle ECD D9)

```

AspProIleLeuGluAspAsp***MetAlaAsnGlu IleAlaAlaLeuLysGlnGlu
gatccgatcttgaggatgattaaatggccaacgag atcgctgcactgaaacaggag
gctagaacctctactaatttaccgggtgctc tagcgacgtgactttgtcctc
BamH1 Msc1

IleAlaAlaLeuLysLysGlu AsnAlaAlaLeuLysGln Gly LysLeu
atcgctgcactgaagaaagag aacgctgcactgaaacag ggc a
tagcgacgtgacttccttctc ttgcgacgtgactttgtc ccg ttcga
Hind3

```

Oligonucleotides for 2 heptad (Mincle ECD D10)

```

MetAlaAsnGlu IleAlaAlaLeuLysGlnGlu IleAlaAlaLeuLysGln Gly LysLeu
ccaacgag atcgctgcactgaaacaggag atcgctgcactgaagcag ggc a
gggtgctc tagcgacgtgactttgtcctc tagcgacgtgacttcgtc ccg ttcga

```

Figure 4.9. Cloning strategy for appending different numbers of computationally optimized heptad repeats to mincle extracellular domain. Restriction sites on the oligonucleotide pairs encoding the 3-heptad version of CC-pIL-I17N were designed for insertion between the BamH1 and Hind3 sites of BoMincle ECD D8. The oligonucleotide pairs that encode 2-heptad version of CC-pIL-I17N were inserted between the Msc1 and Hind3 sites of the 3-heptad vector.

Affinity-purified bovine mincle extracellular domain fused to the 3-heptad CC-pIL-I17N (Mincle ECD D9) ran as a single band on non-reducing SDS-polyacrylamide gels, consistent with formation of a stable disulfide-linked dimer (Fig. 4.10A). Gel-filtration analysis confirmed that the main species eluted has a molecular weight of 43 kDa (Fig. 4.10B), which is close to the expected molecular weight of dimeric Mincle ECD D9 construct, 44 kDa (2 x 22.2 kDa). Following dialysis against water and lyophilization, Mincle ECD D9 displayed significantly better solubility in water than GCN4-mincle extracellular domain constructs. Since protein is the least soluble at a pH near its isoelectric point, aliquots of 1M NaOH is added to the re-dissolved Mincle ECD D9 (isoelectric point: 5.49) until solution turned clear and the dissolved protein showed a pH between 7 and 8. The protein was allowed sufficient time to go into solution (overnight), after which the sample was centrifuged and the supernatant was separated from the pellet. SDS-polyacrylamide gel showed the ratio of dissolved protein to undissolved protein was 1:1 (Fig. 4.11). Aggregates can be seen from the wells where resuspended pellet was loaded, but a band corresponding to dimeric mincle in the resuspended pellet lanes suggested more protein entered solution after resuspension of the pellet in 10 mM Tris-Cl pH 7.8.

To determine if the full 3-heptad dimerization domain was required for forming mincle extracellular domain dimer, a truncated version with 2 heptad repeats (Mincle ECD D10) was also generated. Since only the two heptad repeats from the N-terminus of CC-pIL-I17N were preserved in Mincle ECD D10 construct and the 2-heptad trigger sequence, which was considered important for initiating coiled coil formation, was missed out (Fig. 4.9), it was expected that the dimerization function might be perturbed. Non-reducing SDS-polyacrylamide gel electrophoresis shows that, although some dimer does form, about half of the protein remains as monomer (Fig. 4.12A). A second peak corresponding to monomer is also observed on the gel-filtration column (Fig. 4.12B). The larger species has an apparent molecular weight of 49 kDa on the gel filtration profile, close to the expected molecular weight, 43 kDa (2 x 21.5 kDa), suggesting mincle extracellular domain can still be brought together to form a dimer with only two heptad repeats.

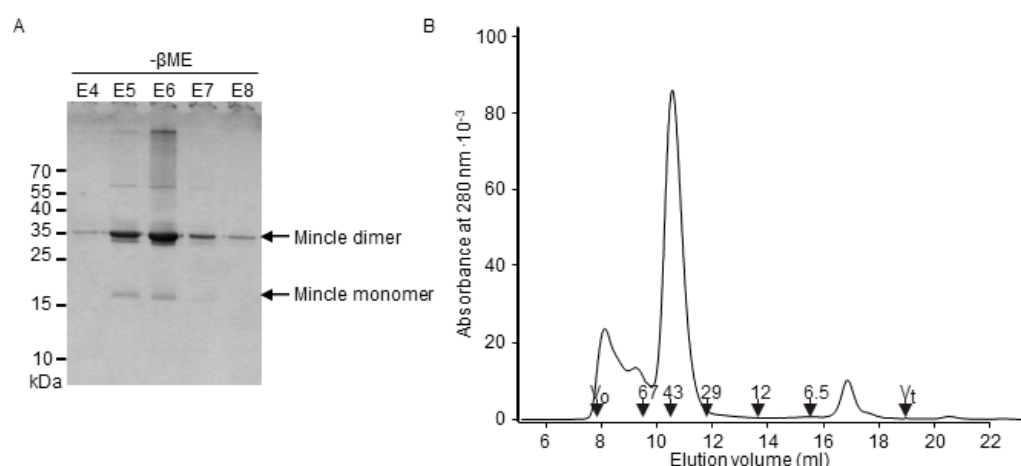


Figure 4.10. Analysis of affinity-purified Mincle ECD D9, mincle extracellular domain with 3-heptad CC-pIL-I17N. (A) SDS-polyacrylamide gel electrophoresis of Mincle ECD D9 renatured by gradual dilution in the presence of Triton X-100 and purified on a 2-ml trehalose-Sepharose column. Aliquots (15 μ l) of 0.5-ml elution fractions (E4-E8) were analyzed on an SDS-polyacrylamide gel in the absence of 2-mercaptoethanol. Gel was stained with Coomassie blue. (B) Superdex S75 gel filtration analysis of Mincle ECD D9 purified from 4 L of bacterial culture, lyophilized and redissolved in 0.5 ml of gel-filtration buffer. Elution positions of standards are indicated in kDa at the bottom.

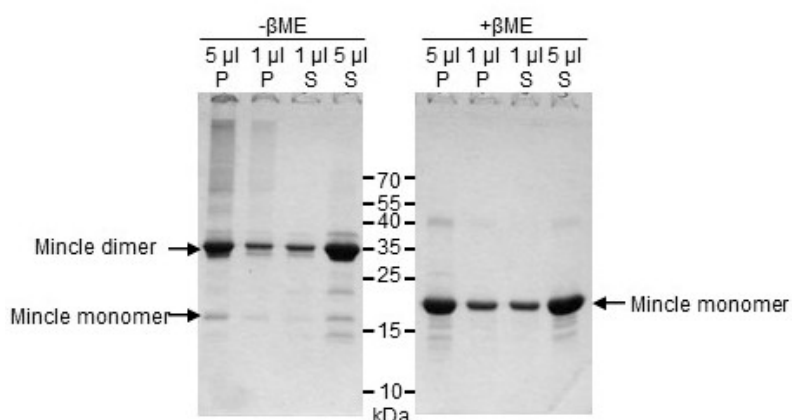


Figure 4.11. SDS-polyacrylamide gel electrophoresis of Mincle ECD D9 subjected to resolubilization. Lyophilized protein equivalent to a 1-L preparation of mincle was dissolved in 25 μ l of water. pH was measured to be between 7 and 8. Tris-Cl pH 7.8 was added to a final concentration of 10 mM. Supernatant (S) was separated from resuspended pellet (P) by centrifugation at 14,000 rpm for 5 min. The pellet was resuspended in 25 μ l of 10 mM Tris-Cl pH 7.8. Aliquots of supernatant and pellet were resolved by separation on an SDS-polyacrylamide gel, which was stained with Coomassie blue.

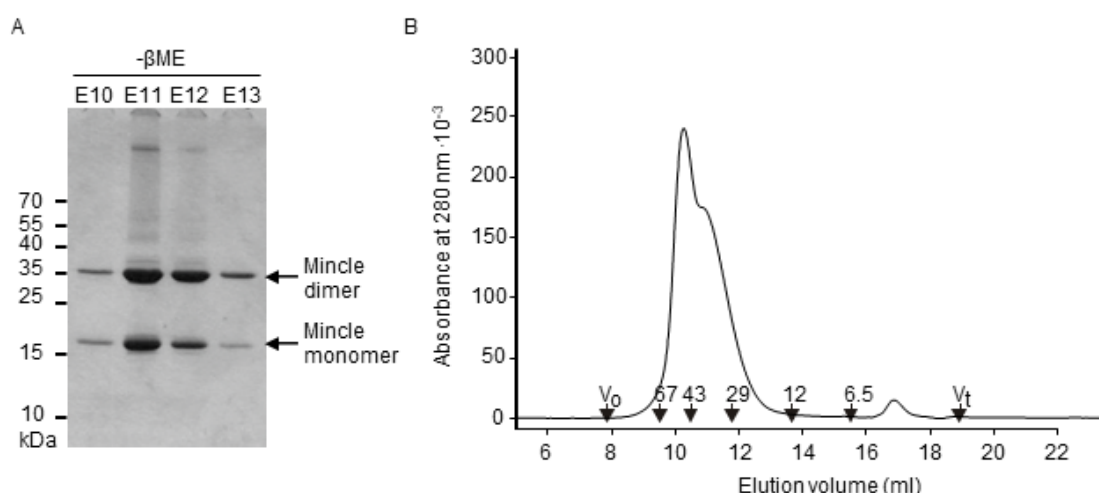


Figure 4.12. Analysis of affinity-purified Mincle ECD D10, mincle extracellular domain with 2-heptad CC-pIL-I17N. (A) SDS-polyacrylamide gel analysis of mincle prepared from 2-L culture, renatured by gradual dilution in the presence of Triton X-100 and purified over a 10-ml trehalose-Sepharose column. Aliquots (15 μ l) of 1-ml elution fractions (E10-E13) were separated in the absence of 2-mercaptoethanol. Gel was stained with Coomassie blue. (B) Superdex S75 gel filtration analysis of mincle purified from 6.5 L of bacterial culture. Sample was dialyzed against water, lyophilized and redissolved in 0.5 ml of gel-filtration buffer. Elution positions of standards are indicated in kDa at the bottom.

Introduction of a protease cleavage site for removal of dimerization domain

An ideal model of mincle extracellular domain for structural analysis would involve minimal modification by the dimerization sequence. Formation of the disulfide-linked dimer is directed by the dimerization domain, but once formed the disulfide bond should stabilize the dimer even in the absence of the fused domain. Insertion of a protease cleavage site between the C-terminus of the dimerization domain and the N-terminus of the mincle neck provides a way to remove the non-mincle sequence after dimer formation. In initial experiments, the Hind3 to Mfe1 restriction fragment of the Mincle ECD D8 plasmid was replaced with double-stranded oligonucleotides coding for different protease cleavage sites (Fig. 4.13). Protein constructs Mincle ECD P1, P2, P3, P4, and P5 were produced and tested for digestion by the relevant proteases (Fig. 4.14).

Enterokinase cleaves specifically after the lysine residue in the sequence DDDDK, although digestion at other basic residues has been reported in some cases (Liew et al., 2005). A non-reducing SDS-polyacrylamide gel of an enterokinase digest of Mincle ECD P1, containing the DDDDK sequence, showed that treatment with 0.2 ng of enzyme resulted in partial cleavage of dimeric Mincle ECD P1, but with an increase in the amount of enzyme that was used, more product emerged in the monomeric form rather than staying as a dimer (Fig. 4.14A). This result suggests that the protein was prone to non-specific cleavage by enterokinase and the reaction conditions need to be carefully monitored so that it can reach completion without converting too much substrate into a secondary product. Thus,

enterokinase and its cleavage site were not an ideal choice for removing the dimerization domain.

Factor Xa cleaves preferentially after the arginine residue of the sequences IEGR or IDGR, although occasional cleavage at other basic residues has been reported (Eaton et al. 1986). The Mincle ECD P2 protein, which contains an IDGR sequence, was successfully produced. However, the protease did not show any activity towards this construct (data not shown). Thus, this cleavage site and protease were also not considered a viable combination for producing a cleaved dimer.

Thrombin cleaves specifically between the arginine and glycine residues in the sequence LVPR//GS. Non-reducing SDS-polyacrylamide gel electrophoresis of thrombin test digestion of mincle extracellular domain containing this sequence, Mincle ECD P3, showed that 4 µg of enzyme digested part of the mincle dimer, but with an increase in the amount of enzyme used, both the substrate and the cleaved dimer decreased, suggesting they were degraded by the excess of thrombin (Fig. 4.14B). Thus, the combination of thrombin and its cleavage site was not pursued further.

Furin cleaves specifically after the second arginine residue of an RXXR sequence. Often, a lysine or arginine residue is favoured at the second X position. Test treatment with furin showed efficient cleavage, with the digestion product remaining as a disulfide-linked dimer (Fig. 4.14C). Since furin did not favour the RARR (P5) sequence over the RAAR (P4) sequence, Mincle ECD P4 was selected for optimization of protease cleavage conditions. Firstly, the effects of temperature and detergent on furin digestion were tested. Non-reducing SDS-polyacrylamide gels showed that an increase in temperature from 30 °C to 37 °C did not have a significant impact on protease digestion (Fig. 4.15A). However, Triton X-100 at 0.5% in the digestion buffer resulted in significantly less undigested substrate remaining in the pellet from the reaction mixture. This result suggests that 0.5% Triton helped to maintain the substrate in solution and allowed the enzyme to cleave before the substrate aggregated.

The effects of prolonged digestion time and amount of enzyme on the completion of furin cleavage were also examined. Non-reducing SDS-polyacrylamide gel electrophoresis showed that after 18 h, the amount of product formed no longer responded to an increase in the incubation time. However, the amount of enzyme used during the first 18 h made a significant impact on the yield of cleaved dimer (Fig. 4.15B). Since the cleavage effect exerted by 3.2 U of furin during the first 18 h was not less than that exerted by 0.8 U of furin over 66 h, an intermediate 2 U of furin would be sufficient to drive the 18-h digestion to completion. To prevent protein degradation and denaturation, 30 °C was favoured over 37 °C for the large-scale furin cleavage. Thus, in the optimized protocol, the reaction was carried out so that each 20 µl of substrate (at the same concentration as the test experiments) was digested with 2 U of furin, in the presence of 0.5% Triton X-100, at 30 °C.

After protease cleavage, the intact mincle extracellular domain dimer could be recovered on a trehalose-Sepharose affinity column and eluted in fractions 5 to 7 (Fig. 4.16A). Gel filtration confirmed that the uncleaved Mincle ECD P4 has an apparent molecular weight of 46 kDa, matching exactly to the expected molecular weight, 46 kDa (2×23.2 kDa), and the furin-cleaved protein has an apparent molecular weight of 38 kDa, which matched exactly to the expected molecular weight, 38 kDa (2×19.3 kDa) (Fig. 4.16B).

The optimized protocol successfully removed GCN4 dimerization domain from intact disulfide-linked mincle extracellular domain dimer. However, after dialysis against water and lyophilization, gel-filtration analysis showed that barely any lyophilized protein was dissolved in the running buffer (Fig. 4.18C). Thus, the digestion was performed on mincle extracellular domain dimer with the less hydrophobic CC-pIL-I17N dimerization domain. Fortunately, the aforementioned cloning strategy made it possible to replace the GCN4 dimerization domain in Mincle extracellular domain P4 to the 3-heptad version of CC-pIL-I17N, leading to the generation of Mincle ECD P6 construct (Fig. 4.17).

Non-reducing SDS-polyacrylamide gel showed the majority of affinity-purified Mincle P6 formed disulfide-linked dimer (Fig. 4.18A) and after furin cleavage, a disulfide-linked dimer could be repurified by affinity chromatography (Fig. 4.18B). Affinity-purified, cleaved dimer was dialyzed against water and lyophilized. Solubility of the lyophilized protein was checked by dissolving the protein in a buffer containing 10 mM Tris-Cl pH 7.8 and 100 mM NaCl, compatible with the gel filtration running buffer, which also contained 2.5 mM EDTA. The sample was spun at 14,000 RPM for 5 min, and the supernatant was analyzed on a Superdex S75 column. Taking into consideration the different amounts of starting material, the gel-filtration result for lyophilized furin-cleaved Mincle ECD P6 shows that more protein went into solution than was seen for furin-cleaved Mincle ECD P4 (Fig. 4.18C). The cleaved dimer of Mincle ECD P6 has an apparent molecular weight of 40 kDa, close to the expected molecular weight of 38 kDa (2×19.3 kDa).

A Restriction sites in pBoMincle ECD D8 with GCN4 dimerization sequence

BoMincleD8	LysLeu	SerIlePheGlnLeu
aacgaagttg	cgctctgaaa	agcatctttc
	aaagcttt	caattg
	Hind3	Mfe1

B Inserted synthetic oligonucleotides

Enterokinase		
BoMincleP1	LysLeu	SerIlePheGlnLeu
P1 forward	agctt	agcatctttc
P1 reverse	actactgctactgttc	tcgtagaaa
		gttaa
Factor Xa		
BoMincleP2	LysLeu	SerIlePheGlnLeu
P2 forward	agctt	agcatctttc
P2 reverse	ataacttccggca	tcgtagaaa
		gttaa
Thrombin		
BoMincleP3	LysLeu	GlySerIlePheGlnLeu
P3 forward	agctt	ggtagcatctttc
P3 reverse	acacggagca	ccatc
		gtagaaa
		gttaa
Furin 1		
BoMincleP4	LysLeu	SerIlePheGlnLeu
P4 forward	agctt	agcatctttc
P4 reverse	agcagtcgatct	tcgtagaaa
		gttaa
Furin 2		
BoMincleP5	LysLeu	SerIlePheGlnLeu
P5 forward	agctt	agcatctttc
P5 reverse	agcagttctgca	tcgtagaaa
		gttaa

Figure 4.13. Cloning strategy for insertion of protease cleavage sites. (A) Sequence of pBoMincle ECD D8 at the junction between the GCN4 dimerization domain and the mincle neck. (B) Oligonucleotides P1 to P5 encode different protease cleavage sites between the end of the dimerization domain and the start of the mincle neck. Recognition sites for the proteases are shown in red.

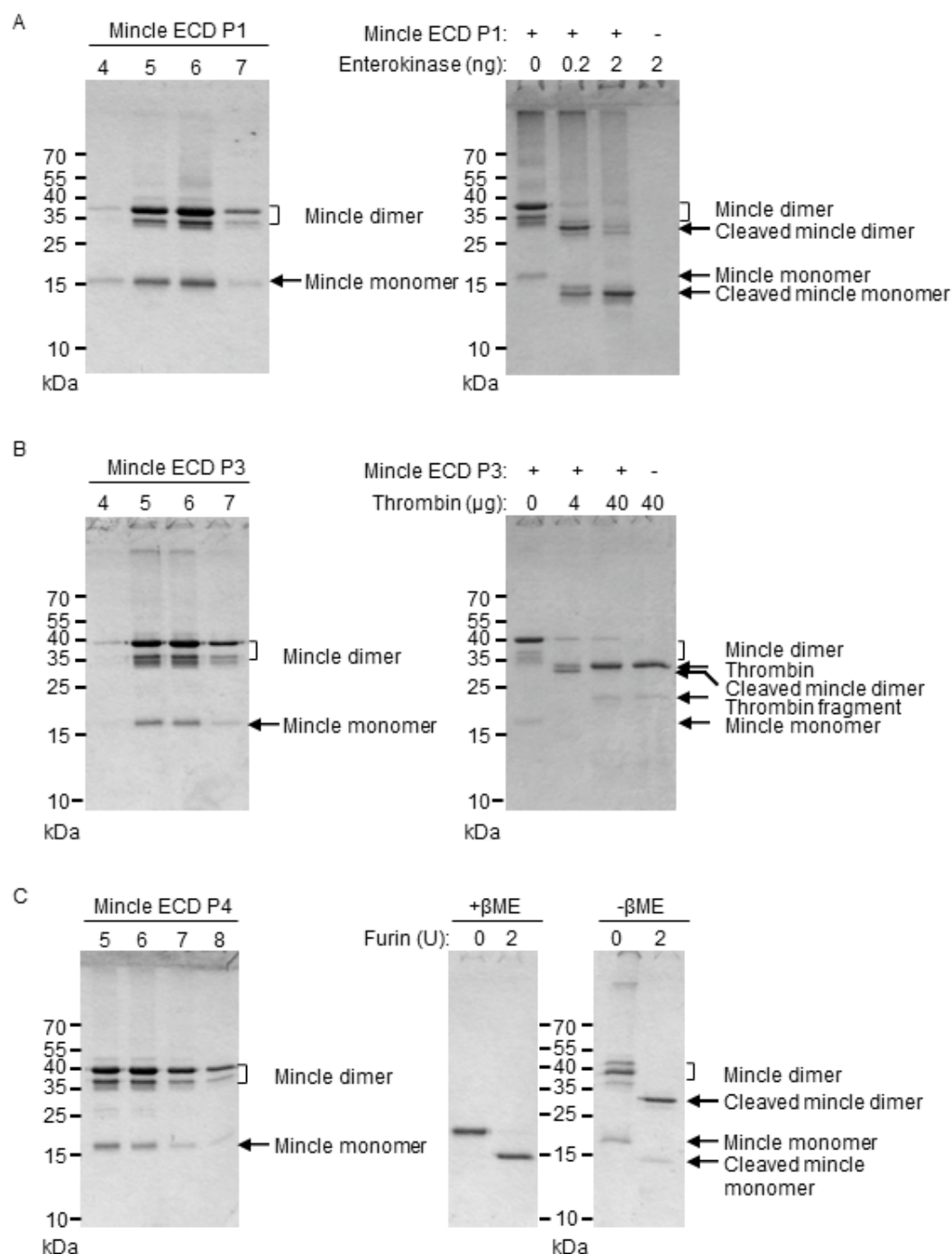


Figure 4.14. SDS-polyacrylamide gel analysis of test protease digestion of GCN4-mincle extracellular domain fusion proteins. Each construct was prepared from 2 L of bacterial culture, renatured by batch dilution, and purified on a 2-ml trehalose-Sepharose column, from which ten 0.5-ml fractions were collected. In each panel, the gel on the left shows elution fractions from purification of the modified expression construct on trehalose-Sepharose, and gels on the right show digestion with proteases. Samples (15 μ l) were mixed with equal volumes of 2X sample buffer in the absence of 2% 2-mercaptoethanol and heated at 100 $^{\circ}$ C for 5 min. Gels were stained with Coomassie blue. (A) Mincle ECD P1 digested with enterokinase at 37 $^{\circ}$ C for 3 h. (B) Mincle ECD P3 digested with thrombin at 37 $^{\circ}$ C for 18 h. (C) Mincle ECD P4 substrate was incubated with furin at 30 $^{\circ}$ C for 18 h.

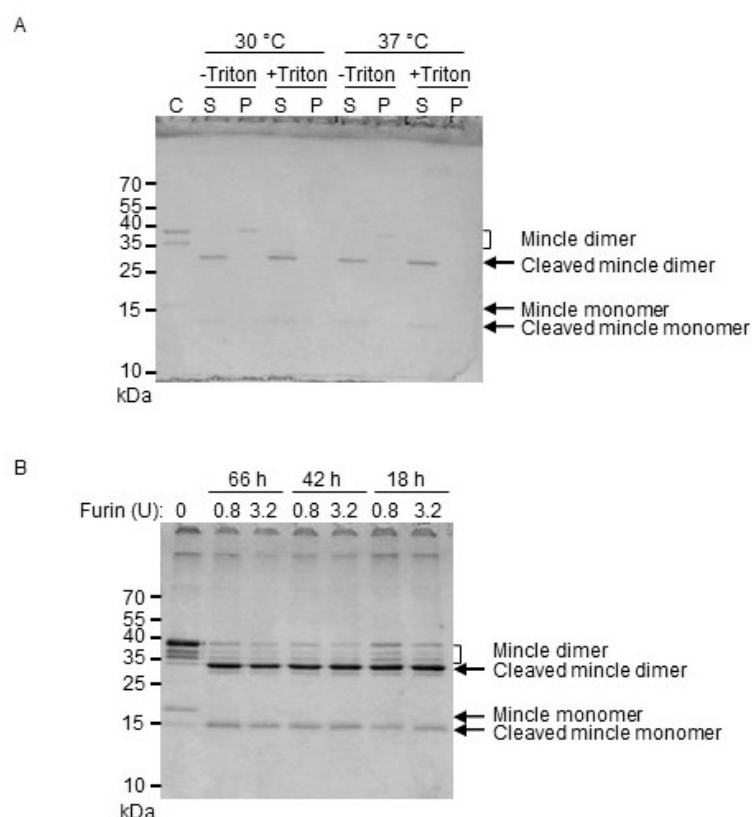


Figure 4.15. SDS-polyacrylamide gel analysis of furin digests of Mincle ECD P4. (A) Digestion in the presence or absence of Triton X-100. Mincle ECD P4 substrate was digested for 18 h. Supernatant (S) was separated from pellet (P) after centrifugation. (B) Digestion with different concentrations of furin. Mincle ECD P4 substrate was incubated at 37 °C. Gels were stained with Coomassie blue.

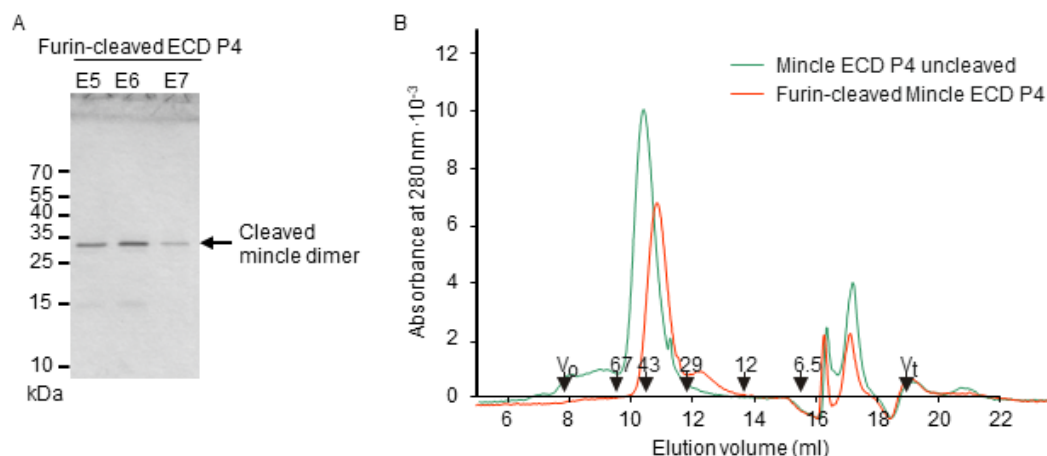


Figure 4.16. Purification and characterization of furin-cleaved Mincle ECD P4 protein. (A) SDS-polyacrylamide gel analysis of furin-cleaved mincle re-purified over a 2-ml trehalose-Sepharose column. Aliquots (15 μ l) of 0.5-ml elution fractions (E5-E7) were separated in the absence of 2-mercaptoethanol. Gel was stained with Coomassie blue. (B) Superdex S75 gel filtration analysis of affinity-purified, uncleaved mincle (0.5 ml of E8 from a 2-L preparation of Mincle ECD P4) and affinity-purified, furin-cleaved mincle (0.45 ml of pooled E5 and E6 from 4.16A). Elution positions of standards are indicated in kDa at the bottom.

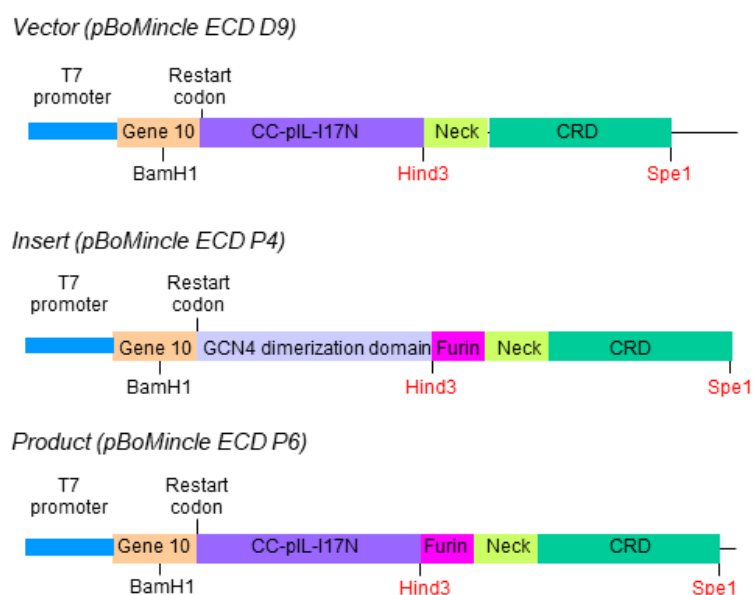


Figure 4.17. Cloning strategy for Mincle ECD P6, mincle extracellular domain with furin cleavage site and 3-heptad CC-pIL-117N dimerization domain. Hind3 and Spe1 sites were used to move the DNA fragment encoding the furin cleavage site and extracellular domain of mincle from pBoMincle ECD P4 plasmid into pBoMincle ECD D9.

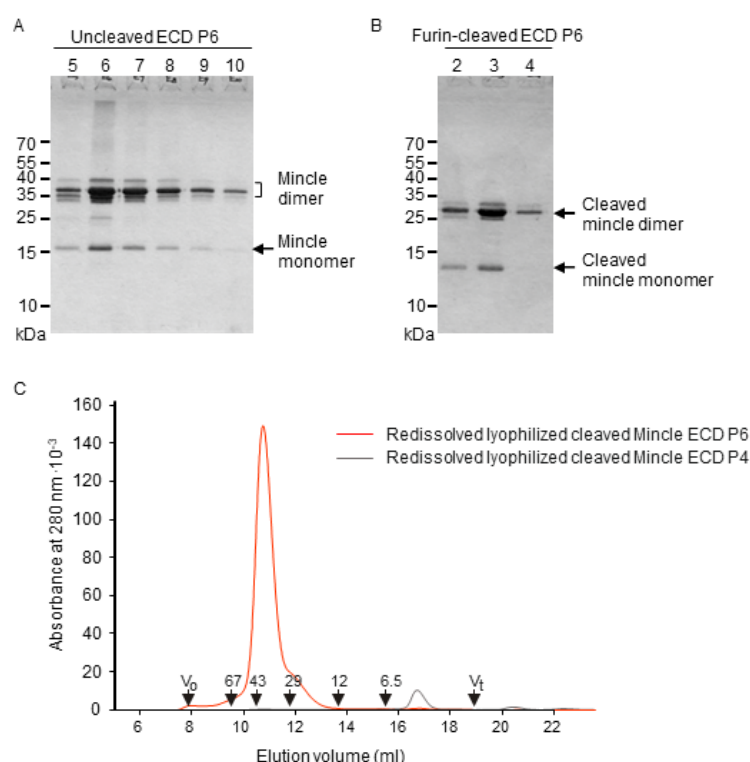


Figure 4.18. Characterization of Mincle ECD P6. (A) SDS-polyacrylamide gel analysis of mincle prepared from 6-L culture, renatured by gradual dilution in the presence of Triton X-100 and purified over a 2-ml trehalose-Sepharose column. Elution fractions 5 to 10 were pooled (3 ml) for preparative furin digestion. (B) SDS-polyacrylamide gel analysis of furin-cleaved Mincle ECD P6 re-purified over a 1-ml trehalose-Sepharose column. Aliquots (15 μ l) of elution fractions 2 to 4 were separated in the absence of 2-mercaptoethanol. Gel was stained with Coomassie blue. (C) Superdex S75 gel filtration analysis of re-dissolved samples of furin-cleaved Mincle ECD P6 and Mincle ECD P4. Elution positions of standards are indicated in kDa at the bottom.

4.3 Discussion

Structural analysis of how the mincle dimer is constructed will require a significant quantity of protein soluble at high concentrations. The primary outcome of these experiments has been development of a protocol which allows production of a mincle dimer suitable for such analysis. An initial batch of protein has been produced and provided to a collaborating crystallography group, but so far no crystals have been obtained.

Although a crystal structure is the long-range target, the characteristics of the different expression constructs generated here provide some insights into the nature of the interactions that lead to dimer formation. Two interesting points can be noted. First, the fact that the naturally occurring disulfide bond can be formed when the N-termini of the polypeptides have been brought into close proximity suggests that they are interacting in a native-like way in the extracellular domain construct. In a membrane, association of the neck regions would probably be enhanced as a result of the orientation of the polypeptides in the membrane, which would place the neck cysteine residues at a common distance from the membrane. In this sense, the fused dimerization domains are acting as a replacement of the transmembrane domain. A second point is that the experiments employing the extracellular domain without a fused dimerization sequence suggest that interactions between the CRDs, which make up the major portion of the extracellular domain, are not sufficient to stabilize the dimer. Taken together, these results are consistent with a picture of dimers in which the neck domains interact, but the CRDs may point away from each other, and might make little or no contact. These themes will be explored further in the context of other mincle family members in the next chapter.

Chapter 5 | Extracellular domain fragments of dectin-2 and BDCA-2

5.1 Introduction

Previous expression of a portion of the extracellular domain of dectin-2 containing part of the neck domain resulted in monomeric protein (Ambati et al., 2019). Successful creation of mincle extracellular domain dimers described in the preceding chapter provided a template for generating dimers of additional mincle family members. For both dectin-2 and BDCA-2, optimized, synthetic dimerization domains were employed to stabilize the dimers and versions were created in which the dimerization domain could be removed by furin digestion. Characterization of the dimeric forms of these proteins generated in this way, along with analysis of existing crystal structures of the CRDs from these two proteins, provides a basis for modelling the organization of the dimers as well as investigating their structures experimentally.

5.2 Results

Dectin-2 extracellular domain fused to synthetic dimerization domain

To generate expression plasmids for human dectin-2 extracellular domain, a synthetic cDNA was designed to encode a furin cleavage site followed by the coding sequence of dectin-2 from residue His45 to Leu209. Restriction sites were incorporated to allow fusion with the dimerization domains in the mincle extracellular domain expression plasmids (Fig. 5.1). An expression plasmid for dectin-2 extracellular domain with the 3-heptad CC-pIL-I17N dimerization domain and furin cleavage site (HuDectin-2 ECD D11) was generated first. Replacing a short region of the cDNA with annealed, synthetic oligonucleotides resulted in a further plasmid from which the furin cleavage site was removed (HuDectin-2 ECD D12).

Non-reducing SDS-polyacrylamide gel electrophoresis showed that both affinity-purified dectin-2 proteins ran as single bands with the molecular weight of a monomer, which was expected, as there was no cysteine residue in the neck of dectin-2 that would support formation of interchain disulfide bond (Fig. 5.2). After dialysis against water and lyophilization, affinity-purified Dectin-2 ECD D11 displayed good solubility in water (Fig. 5.3A), consistent with the good solubility previously observed with Mincle ECD D9, which also contains a 3-heptad CC-pIL-I17N dimerization domain. Gel-filtration analysis showed that the main species of Dectin-2 ECD D12 and D11 have apparent molecular weights of 43 kDa and 46 kDa respectively (Fig. 5.3B), which match closely to the expected molecular weight for their dimers, 44 kDa (D12: 2 x 22 kDa) and 45 kDa (D11: 2 x 22.5 kDa). Thus, the appended CC-pIL-I17N sequence allowed dectin-2 to form stable soluble non-covalently linked dimers.

Although the HuDectin-2 ECD D11 protein could be cleaved to remove the dimerization domain, evidence from the extracellular domain expressed without a fused dimerization

domain indicated that the neck alone would not stabilize dimers in solution. Therefore, no attempt was made to cleave the product with furin.

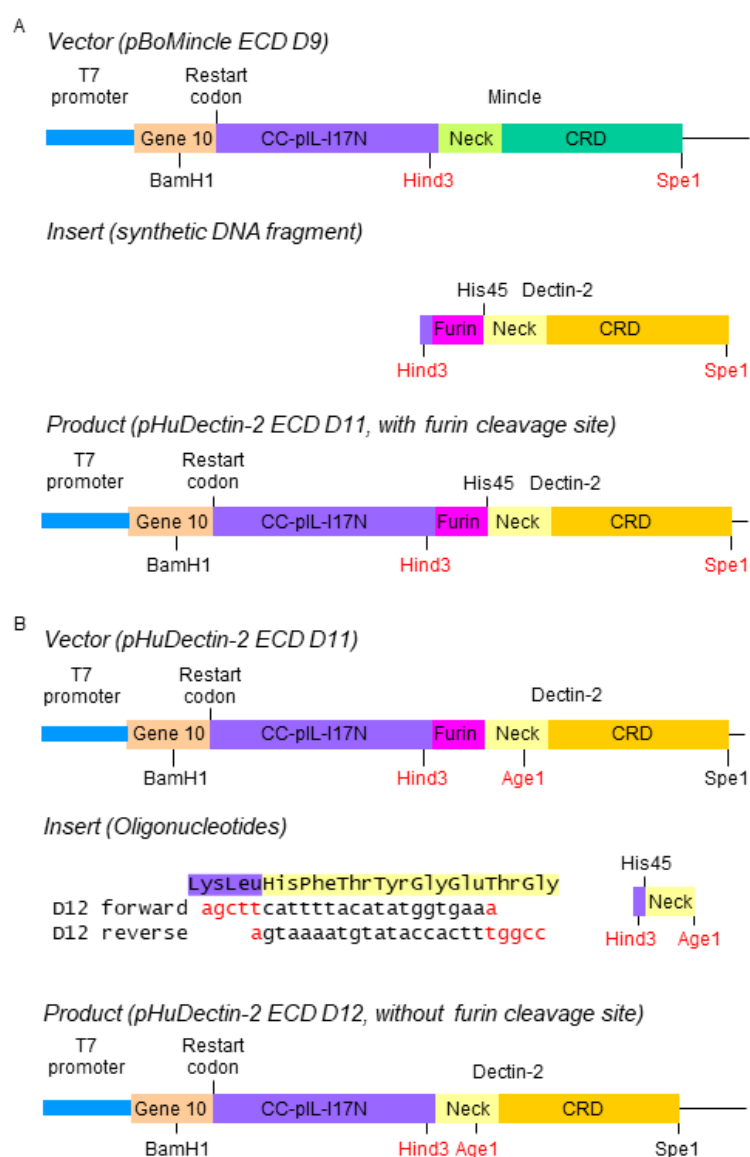


Figure 5.1. Cloning strategy for expression of human dectin-2 extracellular domains with 3-heptad CC-pIL-I17N. (A) HuDectin-2 ECD D11 was achieved by moving a synthetic DNA fragment encoding a furin cleavage site and the extracellular domain of dectin-2 into the expression plasmid for BoMincle ECD D9 after restriction digestion with Hind3 and Spe1. (B) HuDectin-2 ECD D12 was obtained by using annealed oligonucleotides to replace the sequence encoding the linker region to a version without the furin cleavage site.

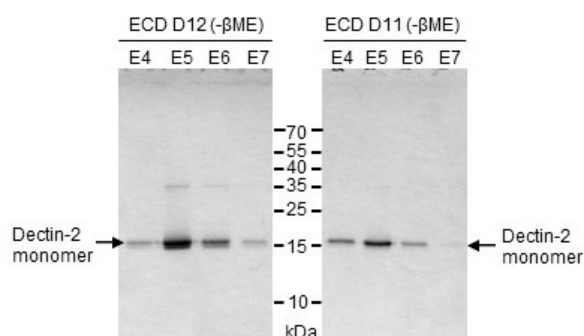


Figure 5.2. SDS-polyacrylamide gel analysis of human dectin-2 extracellular domains with 3-heptad CC-pIL-I17N. Dectin-2 ECD D12 was prepared from 6 L of bacterial culture. Dectin-2 ECD D11 was prepared from 2 L of bacterial culture. Both proteins were renatured by gradual dilution in the presence of 1% Triton X-100 and purified over 2-ml mannose-Sepharose columns. Aliquots (15 μ l) of 0.5-ml elution fractions (E4-E7) were separated on gels run in the absence of 2-mercaptoethanol and stained with Coomassie blue.

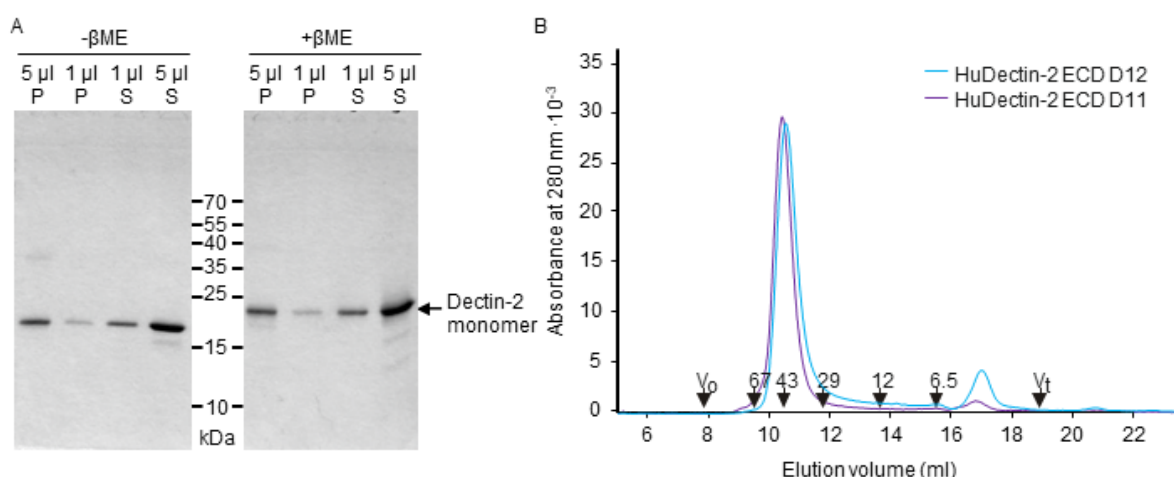


Figure 5.3. Characterization of human dectin-2 extracellular domain with 3-heptad CC-pIL-I17N. (A) SDS-polyacrylamide gel electrophoresis of Dectin-2 ECD D11 subjected to resolubilization. Lyophilized protein equivalent to a 2-L preparation of Dectin-2 ECD D11 was dissolved in 25 μ l of water. pH was measured to be between 7 and 8. Tris-Cl pH 7.8 was added to a final concentration of 10 mM. Supernatant (S) was separated from resuspended pellet (P) after centrifugation at 14,000 rpm for 5 min. The pellet was resuspended in 25 μ l of 10 mM Tris-Cl pH 7.8. Aliquots of supernatant and pellet were resolved by separation on an SDS-polyacrylamide gel, which was stained with Coomassie blue. (B) Superdex S75 gel filtration analysis of resolubilized Dectin-2 ECD D11 (from a 3-L preparation) and Dectin-2 ECD D12 (from a 4-L preparation). Sample was dialyzed against water, lyophilized and redissolved in 0.5 ml of gel-filtration buffer. Elution positions of standards are indicated in kDa at the bottom.

BDCA-2 extracellular domain dimers fused to synthetic dimerization domain

As with dectin-2, the predominant form of BDCA-2 lacks a cysteine in the neck region, so it would not be expected to form interchain disulfide bonds. However, as reported in Chapter 2, a BDCA-2 variant with a neck cysteine did form a disulfide-linked dimer when expressed in HEK cells. Since the location of cysteine in the neck of BDCA-2 variant is similar to that in mincle neck and both receptors form homodimers that are covalently linked by interchain disulfide bonds, the dimerization-followed-by cleavage approach employed for mincle was potentially viable for BDCA-2 as well.

A strategy was developed to generate four versions of the BDCA-2 extracellular domain, with either serine or cysteine at position 49, in the neck, and starting at two alternative points at the junction between the transmembrane domain and the neck. Introduction of the Ser49Cys mutation was achieved by PCR amplification of the portion of a full-length cDNA encoding the extracellular domain. In addition to making this substitution, the 5' primer was designed to facilitate fusion with the dimerization domain and to introduce a BstB1 restriction site (Fig. 5.4). The cloned cDNA fragment was moved into a vector containing the 3-heptad CC-pIL-I17N, giving rise to BDCA-2 ECD D13. The sequence at the 5' end could then be modified using pairs of oligonucleotides to generate the alternative constructs (Fig. 5.5).

Both wild-type extracellular domain constructs, which differ by one amino acid residue at the start of the neck domain, run as single bands on SDS-polyacrylamide gels in the presence or absence of 2-mercaptoethanol (Fig. 5.6A/B), which was expected, because there was no neck cysteine for forming an interchain disulfide bond. Gel-filtration analysis of BDCA-2 ECD S1 and BDCA-2 ECD S2 showed the main species eluted had an apparent molecular weight of 45 kDa, which matches the expected molecular weight of dimeric BDCA-2 ECD S1 and S2, 46 kDa (2 x 22.8 kDa). These results show that with the appended 3-heptad dimerization domain wild-type BDCA-2 extracellular domain stabilizes non-covalent dimers.

For the corresponding cysteine mutants, non-reducing SDS-polyacrylamide gels showed that the majority of affinity-purified BDCA-2 ECD C1 and BDCA-2 ECD C2 exists as disulfide-linked dimers that can be dissociated in the presence of 2-mercaptoethanol (Fig. 5.6C/D). Gel-filtration analysis of these proteins showed that the main peaks have apparent molecular weights of 45 kDa and 44 kDa respectively, which matches closely the expected molecular weight of dimeric BDCA-2 ECD C1, 46 kDa (2 x 23 kDa), and dimeric BDCA-2 ECD C2, 45 kDa (2 x 22.8 kDa). For both the proteins with serine at position 49 and the proteins with cysteine at this position, a slightly shorter neck resulted in a higher yield of purified protein in its native dimeric state. However, the overall yield was low.

```

                                Furin|                                Cys
                                LysLeuArgAlaAlaArgHisAsnPheMetTyrSerLysThrValLys
                                cacaattttatgtatagcaaaactgtcaag
bdca2mut9  aaaagcttcgtgcagctagacacaattttatgtattgtaaaactgtcaag
                                Hind3

ArgLeuSerLysLeuArgGluTyrGlnGlnTyrHisProSerLeuThrCysValMetGlu
aggctgtccaagttacgagagtatcaacagtatcatccaagcctgacctgctcatggaa
aggcttctgaagttacgagagtatcaacagtat
                                BstBI

GlyLysAspIleGluAspTrpSerCysCysProThrProTrpThrSerPheGlnSerSer
ggaaaggacatagaagattggagctgctgcccaacccttggaacttcatttcagtcagtagt

CysTyrPheIleSerThrGlyMetGlnSerTrpThrLysSerGlnLysAsnCysSerVal
tgctactttatttctactgggatgcaatcttggaactaagagtcaaaagaactgttctgtg

MetGlyAlaGluLeuValValIleAsnThrArgGluGluGlnAspPheIleIleGlnAsn
atgggggctgatctgggtggtgatcaacaccaggaagaacaggatttcattcagagaat

LeuLysArgAsnSerSerTyrPheLeuGlyLeuSerAspProGlyGlyArgArgHisTrp
ctgaaaagaaattcttcttattttctggggctgtcagatccagggggctcggcgacattgg

GlnTrpValAspGlnThrProTyrAsnGluAsnValThrPheTrpHisSerGlyGluPro
caatgggttgaccagacaccatacaatgaaaatgtcacattctggcactcaggtgaaccc

AsnAsnLeuAspGluArgCysAlaIleIleAsnPheArgSerSerGluGluTrpGlyTrp
aataaccttgatgagcgttggtgcgataataaatttcggttcttcagaagaatggggctgg

AsnAspIleHisCysHisValProGlnLysSerIleCysLysMetLysLysIleLysIle***
aatgacattcactgtcatgtacctcagaagtcaatttgcaagatgaagaagatctacatataa
                                cagttaaacgttctacttcttctagatgtatattaa
                                bdca2mut10

aagggcgaattcactagt  (PCR-TOPO vector)
                                EcoR1

```

Figure 5.4. Introduction of Ser49Cys mutation into the extracellular domain of BDCA-2. 5' primer bdca2mut9 and 3' primer bdca2mut10, highlighted in cyan, were used with a full-length BDCA-2 clone as template. The PCR product resulting from amplification was cloned into pCR-TOPO vector. An EcoR1 site in this vector adjacent to the cloning site is shown in green.

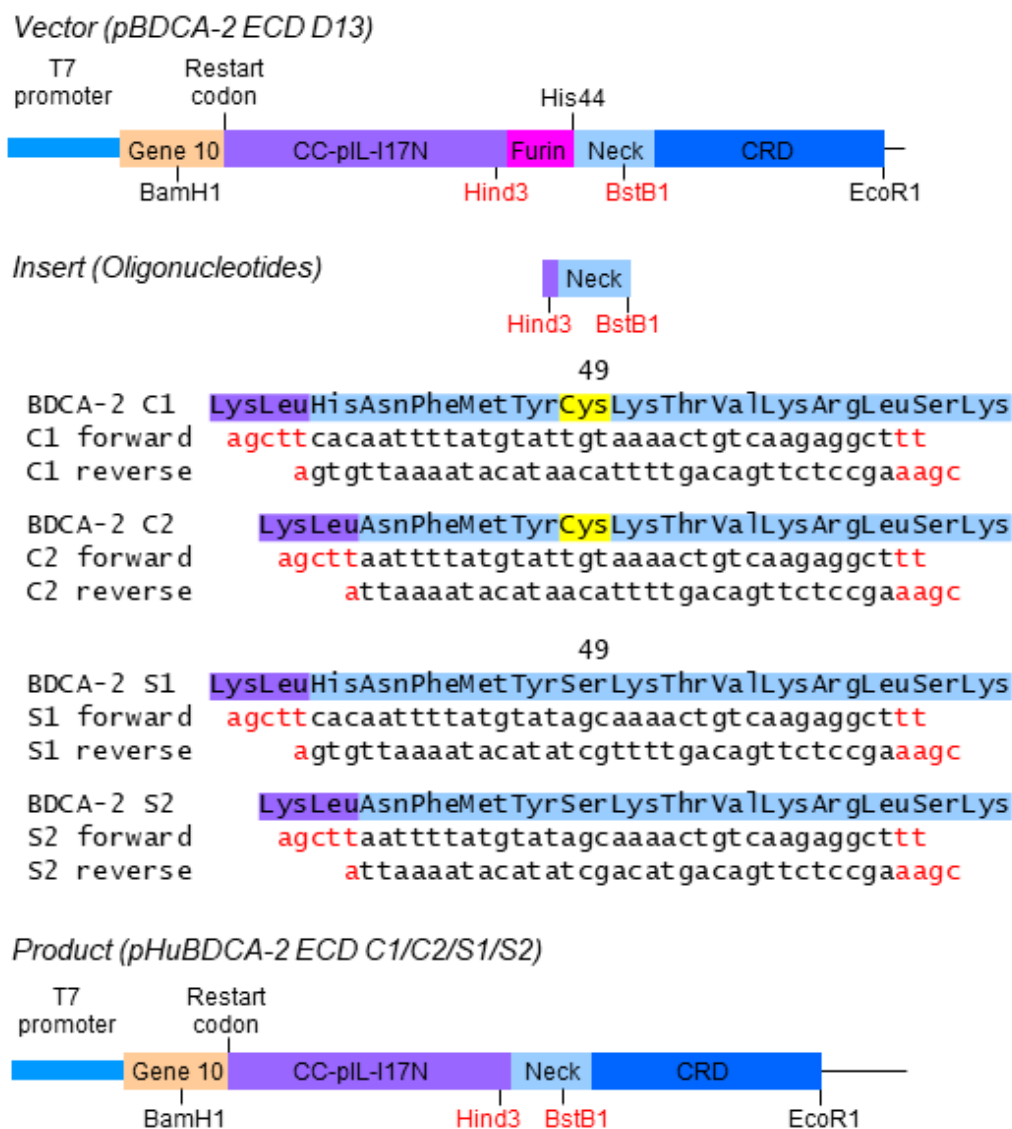


Figure 5.5. Cloning strategy for BDCA-2 extracellular domain with different neck lengths. Expression plasmid BDCA-2 ECD D13 was created with PCR amplified cDNA and Mincle ECD D9 plasmid digested with Hind3 and EcoR1 was used as a vector. Oligonucleotide pairs C1 and C2 were used to remove the furin cleavage site and fuse the dimerization at two alternative points in the neck of BDCA-2. Oligonucleotide pairs S1 and S2 were used to create the wild-type equivalents with a serine residue at position 49.

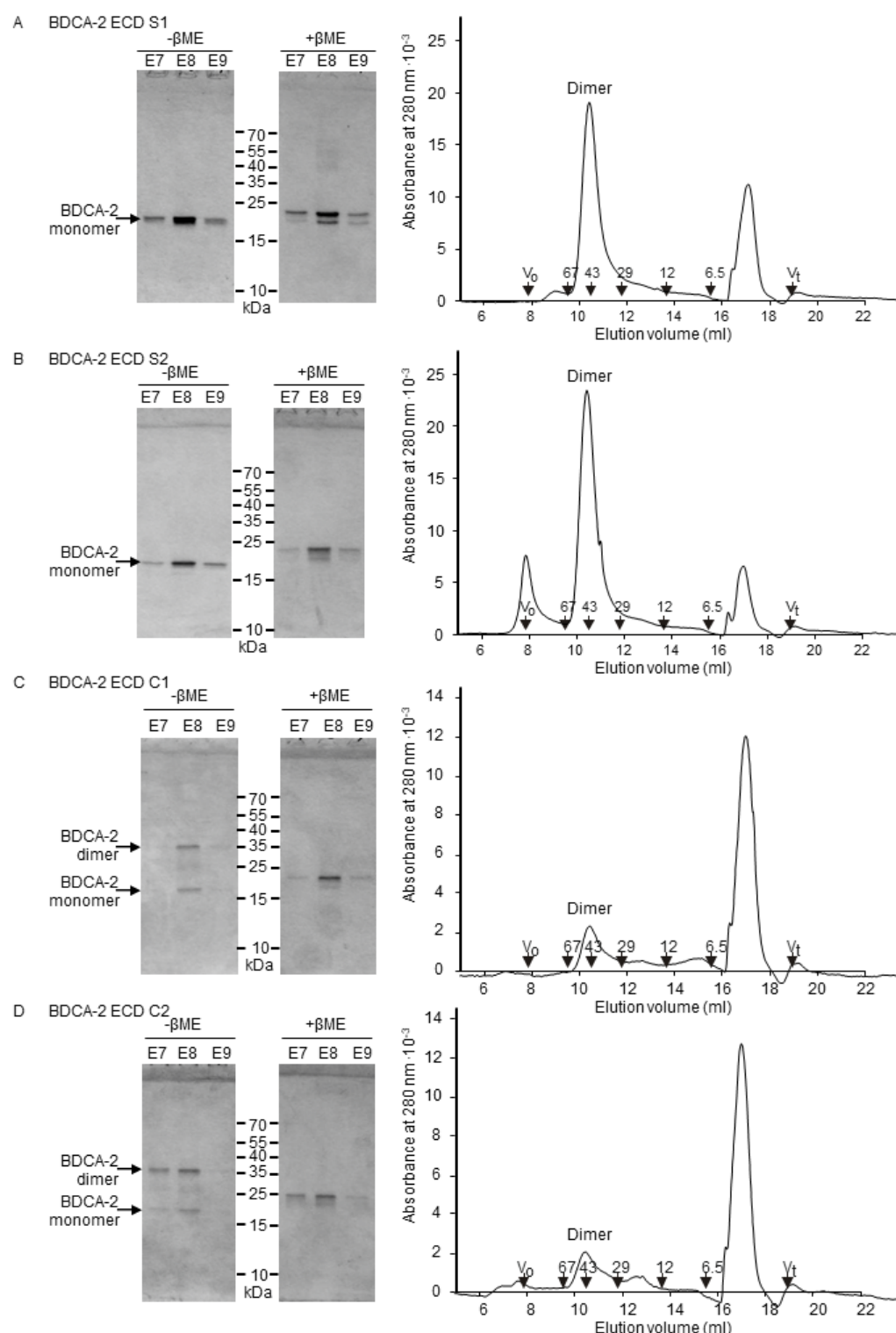


Figure 5.6. Characterization of human BDCA-2 extracellular domain constructs fused to 3-heptad CC-pIL-17N dimerization domains. Each construct was prepared from 2 L of bacterial culture, renatured by batch dilution, and purified over a 5-ml desialylated egg-glycopeptide column, from which ten 1-ml fractions were collected. Aliquots (15 μ l) of elution fractions (E7-E9) were separated on SDS-polyacrylamide gels that were stained with Coomassie blue. Half of fraction 8 (0.5 ml) of each affinity-purified construct was subject to Superdex S75 gel filtration analysis. Elution positions of standards are indicated in kDa at the bottom of each elution profile.

Potential dimers interactions observed in crystal structures

The results described in this chapter for BDCA-2 and in the previous chapter for mincle suggest that the dimer interfaces in these receptors involve only limited portions of the CRD and are dependent on the neck region. Based on these considerations, existing crystal structures of the CRDs of mincle, BDCA-2 and dectin-2 were examined to see if there are interactions present that might represent native interactions in a dimer. In each case, PyMol was used to generate all of the symmetry mates from the unit cell in the crystal. For two of the proteins, dectin-2 and BDCA-2, potential interactions of the N-terminal ends of the CRDs were observed.

PDB entry 4ZES contains two copies of the CRD from BDCA-2 in the asymmetric unit, chain A and chain B. In PyMol, symmetry mates of the asymmetric unit were generated based on the $P22_12_1$ space group. A possible BDCA-2 dimer interface can be observed between chain A of the original copy (4ZES) and chain B of a symmetry mate (Fig. 5.7A). The coordinates of these two copies of the CRD were saved as a separate PDB file, which was submitted for interface analysis by PDBePISA. Interface surface areas of 351 and 378 Å² were calculated for the two subunits. These numbers are not identical, because the two copies of BDCA-2 that form the dimer being visualized are not related by a crystallographic symmetry axis.

PDB entry 5VYB contains only a single dectin-2 CRD in the asymmetric unit. It was crystallized in the $C222_1$ space group. In this case, a potential dectin-2 dimer interface very similar to the one observed in BDCA-2 was identified between the original copy and one of the symmetry mates (Fig. 5.8A). In this case, the PDBePISA software indicated that the N-terminus of each subunit forms an interface with a surface area of 366 Å². The arrangement of this dimer is strikingly similar to that seen for BDCA-2, in spite of the different crystal packing and the fact that the subunits of the dectin-2 dimer are related by a crystallographic symmetry axis while those in the BDCA-2 dimer are not.

Concentrated solutions of several of the dimeric proteins containing the extracellular domain of mincle and dectin-2 were produced for screening of potential crystallization conditions. Crystals of Dectin-2 ECD D12 in the I_2 space group, with bound α -methyl mannoside, diffracted to 2.36 Å and contained two dectin-2 polypeptides in the asymmetric unit (Hadar Feinberg, personal communication). The geometry of the CRDs is almost identical to that observed for the isolated CRD (Fig. 5.9A). Two additional amino acids (SerSer) at the N-terminus of each subunit are visible in the new crystal structure, contributing to a more extensive surface area of 453 Å² (Fig. 5.9B). Overlay of the extracellular domain dimer and the CRD dimer also shows very small amount of flexibility in the relative positions of the CRDs. Since the coiled-coil dimerization domain is likely formed, the fact it is absent from the crystal

indicates that its position is not ordered in the crystal, which in turn suggests that the neck region of dectin-2 is flexible.

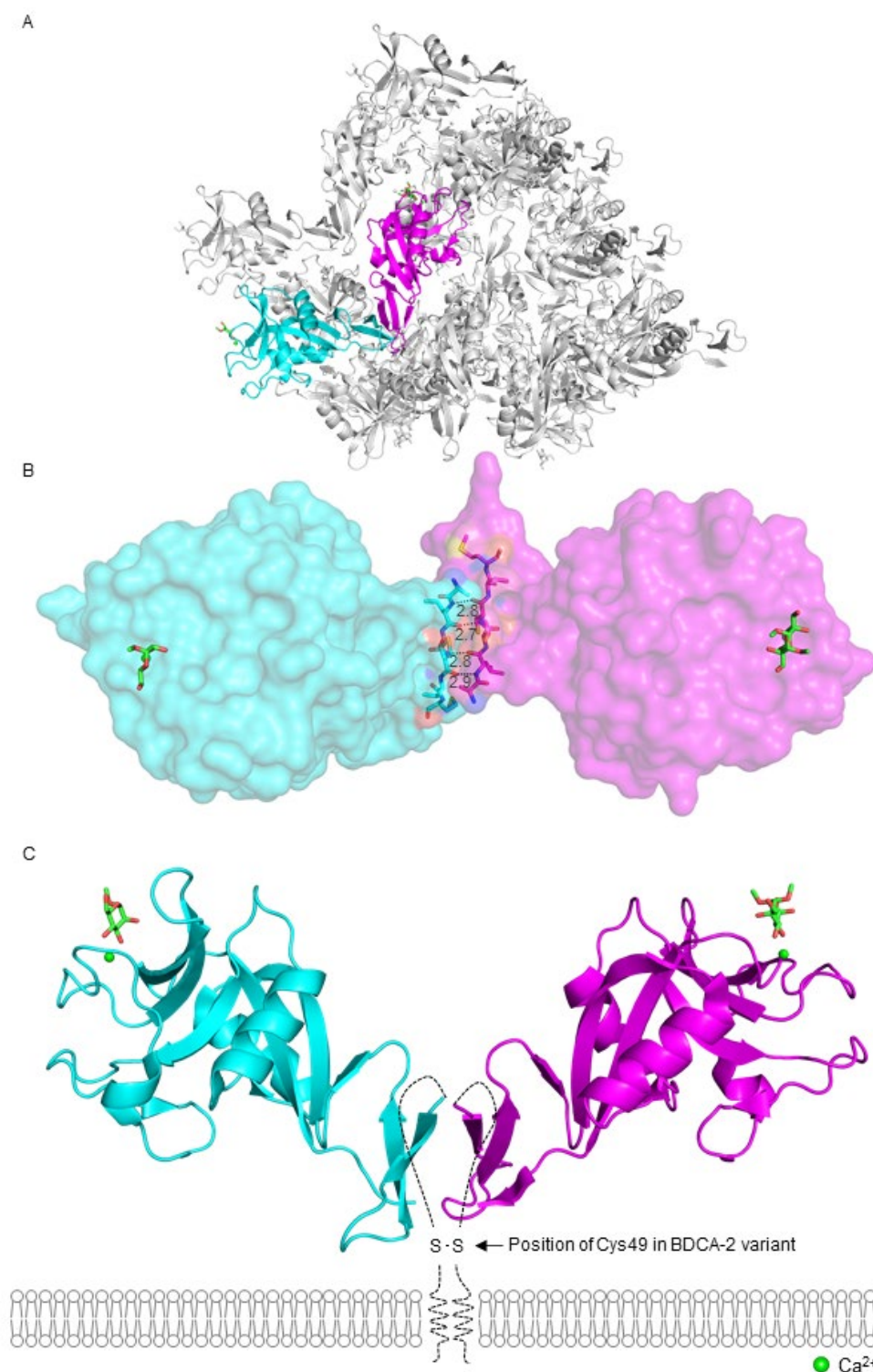


Figure 5.7. PyMol view of BDCA-2 dimer interface and interface analysis. (A) after generation of symmetry mates from PDB 4ZES (grey), a possible interface was identified between chain A of the original copy (magenta) and chain B of symmetry mate 0300-100 (cyan). (B) Hydrogen bonds between the N-termini of the two subunits are shown. In this view, the two sugar binding sites, occupied by a mannose residue, are facing out of the page. (C) The proposed dimer is shown in a sideways orientation. In order for the sugar-binding sites to face away from the membrane, the neck polypeptide would need to loop around the interface between the N-terminal portions of the CRDs.

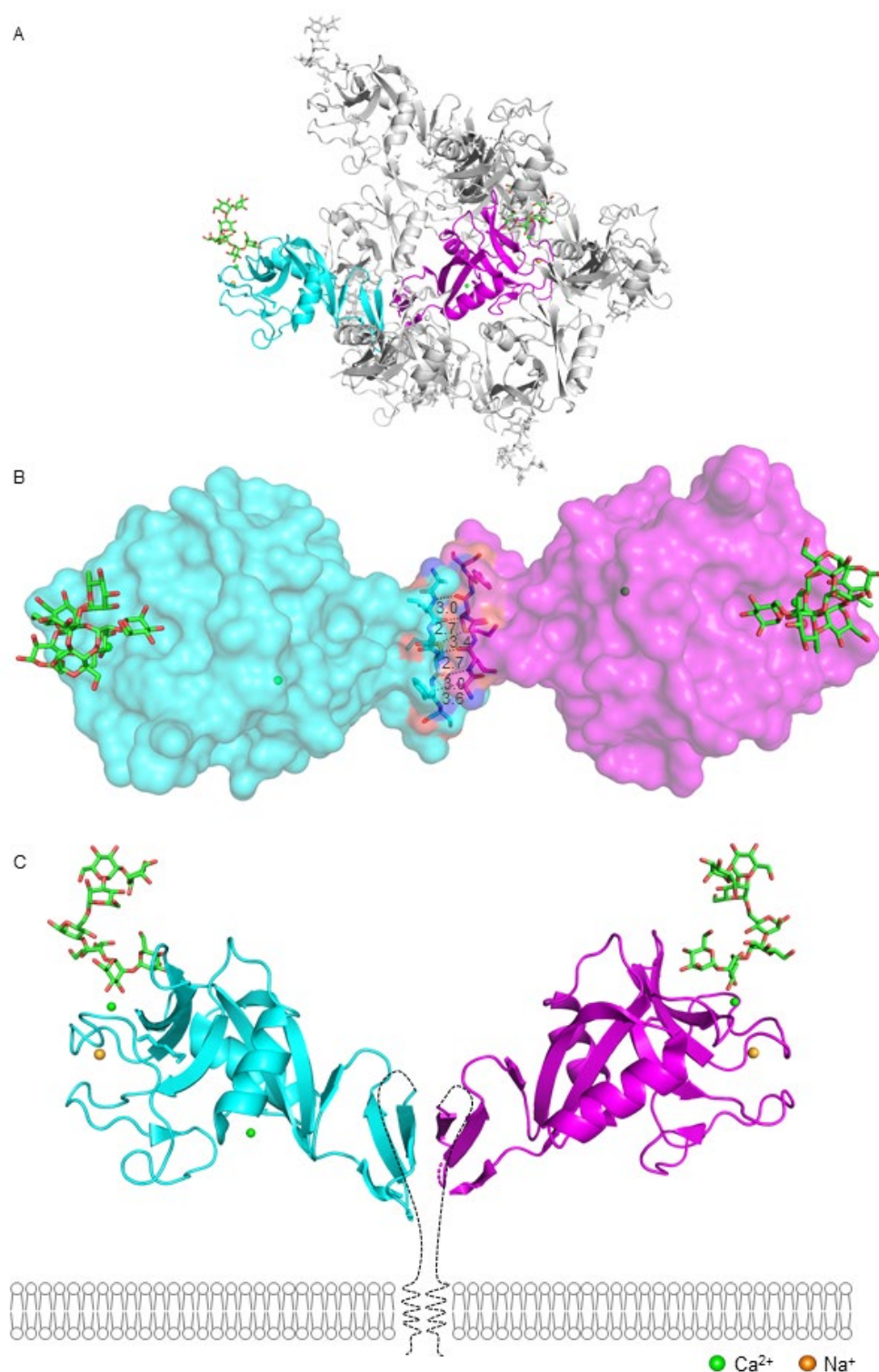


Figure 5.8. PyMol view of dectin-2 dimer interface and interface analysis. (A) After generation of symmetry mates from PDB 5VYB (grey), a candidate interface was identified between the original copy (magenta) and symmetry mate 03010000 (cyan). (B) Hydrogen bonds formed between the N-termini of subunits are shown. Bound oligosaccharide ligands are shown facing out of the page. (C) Side view of the proposed dimer showing its potential orientation relative to the membrane.

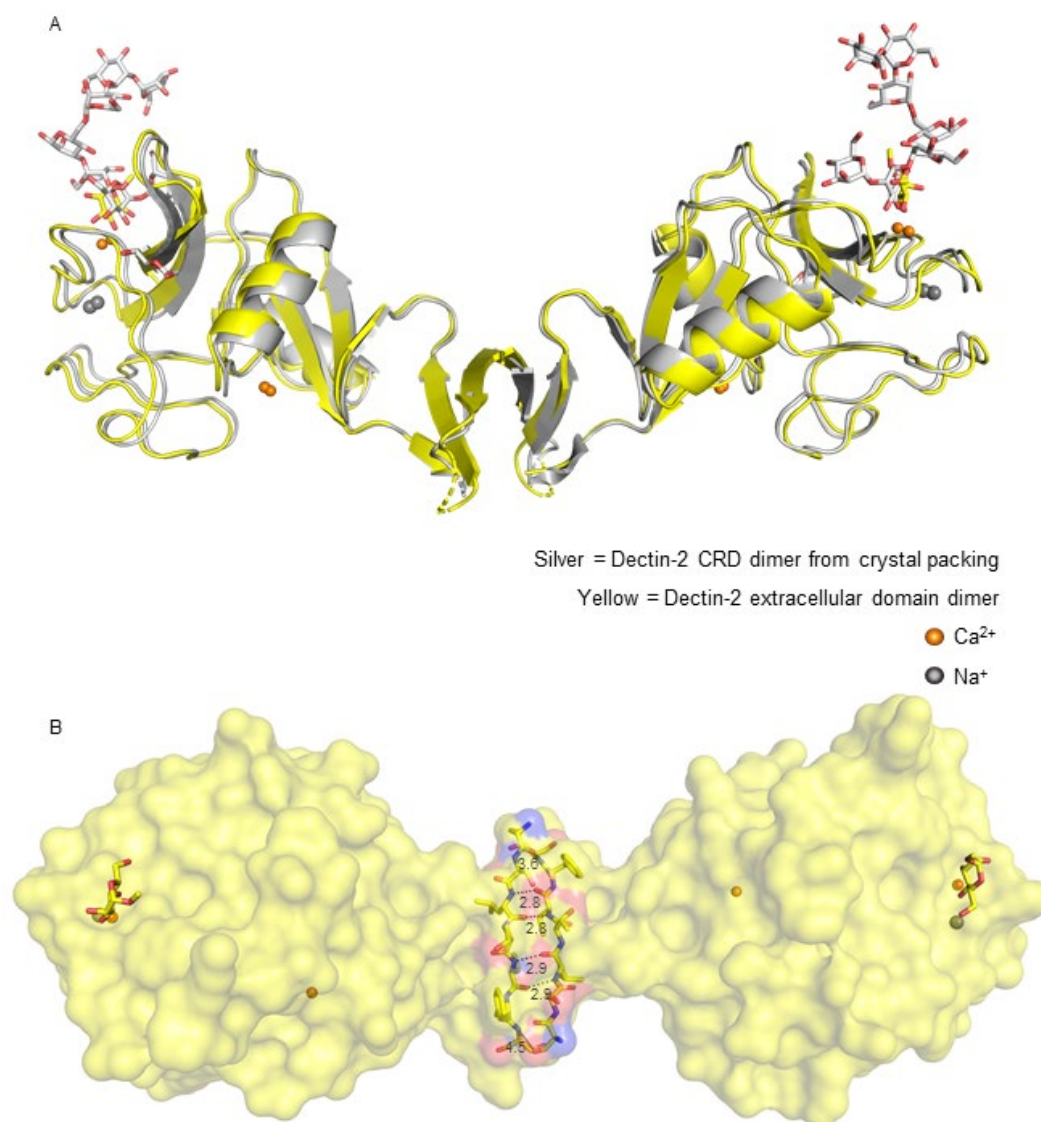


Figure 5.9. Pymol view of dectin-2 extracellular domain dimer. (A) Overlay of the interface observed in dectin-2 extracellular domain dimer (yellow) and one that was predicted by crystal packing of PDB entry 5VYB (silver). N-termini of dectin-2 extracellular domain subunits form longer β sheets due to extension by two serine residues. (B) Hydrogen bonds formed between the N-termini of subunits are shown. Bound α -methyl mannoside ligands are shown facing out of the page.

The CRD of human mincle has been crystallized (Furukawa et al., 2013). However, the expressed fragment used in these studies started at residue 74, which excludes all of the neck and the first part of the CRD, so the regions expected to be involved in dimer formation are missing. For cow mincle, several crystal structures are also short forms, lacking the N-terminal part of the CRD. However, the full CRD is included in two crystal structures. These structures (PDB 5KTH and 5KTI), with a single CRD in the asymmetric unit in space group $P3_121$, have been examined in the same way as for dectin-2 and BDCA-2. None of the symmetry mates

show any interactions that involve the N-terminal portion of the CRD as seen for dectin-2 and BDCA-2.

In a study of 32 homodimers, which have interface surface areas ranging from 368 to 4746 Å², the interface surface area has a linear correlation with the molecular weight of the subunit with a correlation coefficient $r = 0.69$ (Jones and Thornton, 1996). According to this correlation, monomeric BDCA-2 CRD and dectin-2 CRD, which both have an approximate molecular weight of 17,000 Da, would require an interface area of around 1250 Å² to form stable homodimers. Thus, the interface areas predicted from the crystal packing of just the CRDs probably fall short of the required area for holding the dimer together.

The crystal structures suggest that extending the portion of the neck domain included in the expressed construct might increase the stability of the dimer. One existing BDCA-2 construct from other work, which encodes residues Ser49 to Ile213 and thus contains roughly half of the neck domain, has been expressed and purified (Fig. 5.10A). The protein appears as a single band on SDS-polyacrylamide gels in the presence of 2-mercaptoethanol (Fig. 5.10B). Gel-filtration analysis of this extended version of BDCA-2 CRD showed the main species elutes as a somewhat broad peak at an apparent molecular weight of 38.2 kDa, which approaches the expected molecular weight of a dimer (2 x 19.5 kDa) (Fig. 5.10D). Compared to the monomeric CRD and the covalent dimer discussed above (Fig. 5.10C), the peak is broader, and it shifts to a lower molecular weight with dilution. These features indicate that the complex is unstable and is probably an equilibrium between monomeric and dimeric species. Because there appears to be room in the CRD crystals to accommodate some additional polypeptide, this fragment containing half of the neck domain is being screened under the crystallization conditions used for the CRDs.

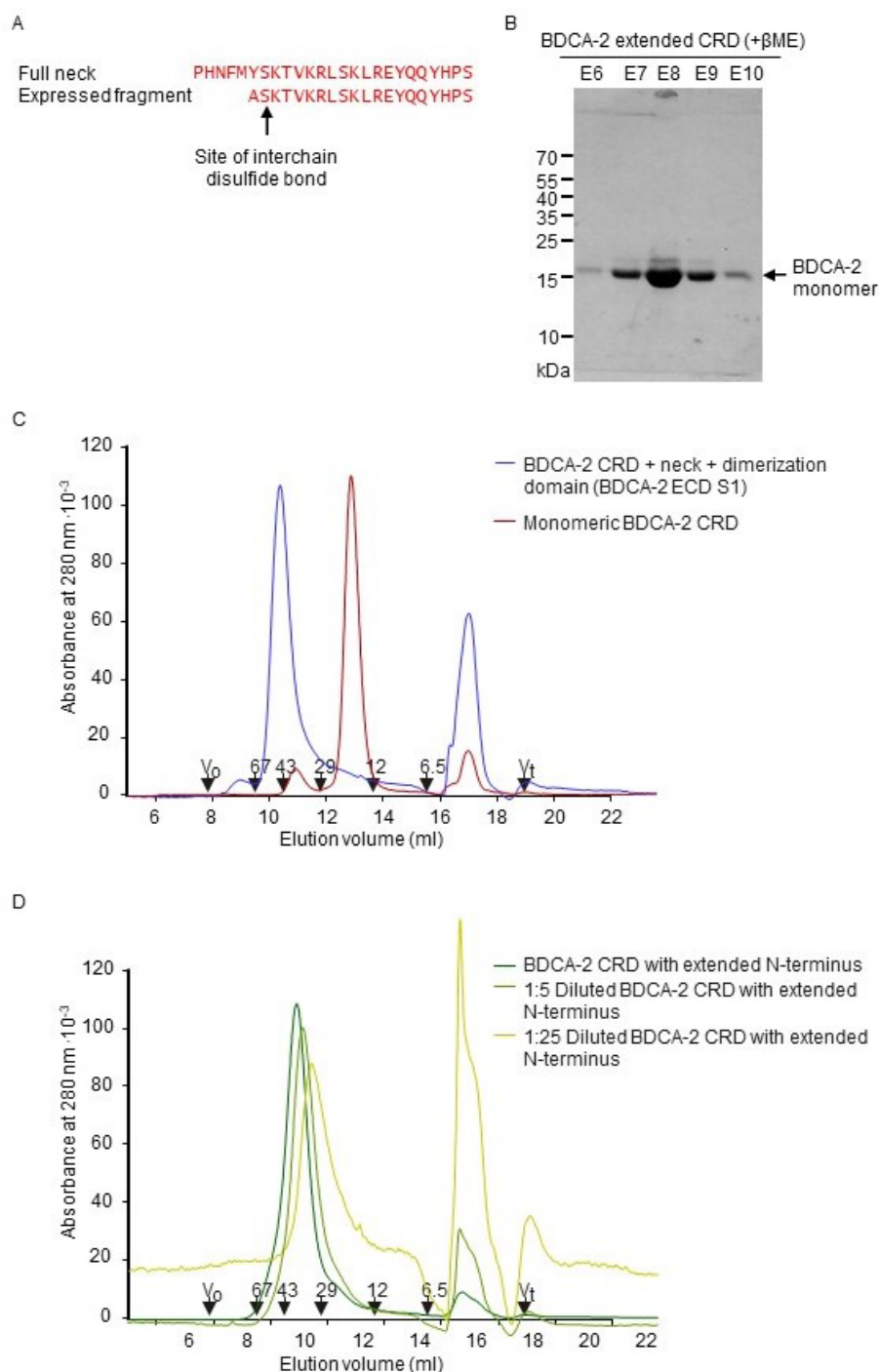


Figure 5.10. Expression and characterization of human BDCA-2 CRD with half of the neck region. (A) Sequence of the portion of BDCA-2 expressed. The position of cysteine in the naturally occurring variant of BDCA-2 is shown by the arrow. (B) SDS-polyacrylamide gel analysis of purified protein eluted from the glycopeptide affinity column. Aliquots (15 μ l) of 1-ml elution fractions (E6-E10) were separated on SDS-polyacrylamide gels. Gel was run in the presence of 2-mercaptoethanol and stained with Coomassie blue. (C) Superdex S75 gel filtration analysis of BDCA-2 with the full neck and dimerization domain (0.5 ml of fraction 8) and the CRD alone. (D) Superdex S75 gel filtration analysis of affinity-purified BDCA-2 extended CRD from fraction 8 at various dilutions. Absorbance values for the dilute samples have been multiplied by the dilution factor to make the comparison. Elution positions of standards are indicated in kDa at the bottom.

Other experimental approaches to analysis of the dimer interface

Crosslinking mass spectrometry has also been investigated as an alternative way for identifying regions of dimer interactions. Chemical crosslinkers with different spacer arm lengths were tested to find the shortest crosslinking reagent that could covalently crosslink the dectin-2 extracellular domain dimer (Table 5.1). Most of these crosslinkers are only reactive towards primary amines, except 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, which reacts with a carboxyl group at one end and a primary amine at the other end, resulting in an amide bond between the two crosslinked residues. Hence, it is called zero-length crosslinking. SDS-polyacrylamide electrophoresis in the presence of 2-mercaptoethanol showed that even the zero-length crosslinking reagent can crosslink dectin-2 extracellular domain dimer (Fig. 5.11).

The use of mass spectrometry to identify peptides from dectin-2 and mincle extracellular domains with synthetic dimerization domains was tested with uncrosslinked samples. To prepare peptides from uncrosslinked proteins for mass spectrometry analysis, proteins were linearized by reduction and carboxymethylation. Chymotrypsin rather than trypsin was chosen for proteolytic digestion for two reasons. First, lysine and arginine residues appear at high frequencies throughout sequences of dectin-2 and mincle, so tryptic digestion would create many small peptide fragments. Peptides of similar masses but different sequences would be result in a crowding of the low mass region of the mass spectrum. This problem can be circumvented by using chymotrypsin, because the aromatic residues are less common. Second, since crosslinking would modify lysine residues, the modified regions would no longer be digested by trypsin, confounding the analysis.

Uncrosslinked peptides digested by chymotrypsin were detected using matrix-assisted laser desorption ionization mass spectrometry. Analysis of mass spectrometry profiles and sequence data from mass spectrometry/mass spectrometry experiments revealed a good coverage of peptides in all three regions of the dimerization domain, neck domain and CRD of both dectin-2 and mincle (Fig. 5.12). Although the results suggest that this approach might lead to identification of portions of the dimers that are close to each other, the findings described above suggested that the interaction region was likely to be relatively short, so that finding crosslinked peptide could become quite difficult. The strategy was not investigated further.

Table 5.1. Reagents tested for crosslinking the extracellular domain of dectin-2.

Chemical crosslinking reagent	Spacer arm (Å)	Reactive with
<i>bis</i> (sulfosuccinimidyl) suberate (BS ³)	11.4	Primary amines
disuccinimidyl tartrate	6.4	Primary amines
1,5-difluoro-2,4-dinitrobenzene	3	Primary amines
1-ethyl-3-(3-dimethylaminopropyl) carbodiimide	0	Carboxyl + amine

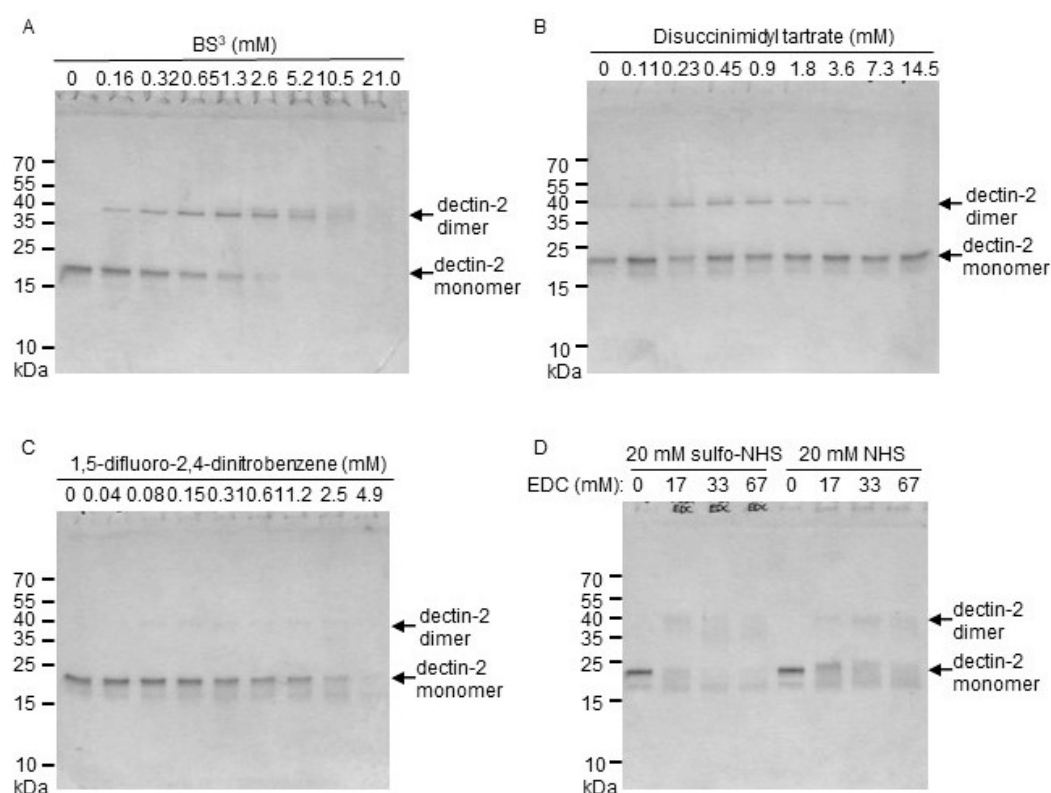


Figure 5.11. SDS-polyacrylamide gel analysis of crosslinking of the extracellular domain of dectin-2. Crosslinking was performed with bifunctional reagents (A) BS³, (B) disuccinimidyl tartrate or (C) 1,5-difluoro-2,4-dinitrobenzene or (D) by condensation reaction with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) in the presence of *N*-hydroxysuccinimide (NHS) or *N*-hydroxysulfosuccinimide (sulfo-NHS). All reactions were stopped by adding 2X sample buffer containing 2% 2-mercaptoethanol and were immediately separated on 17.5% SDS-polyacrylamide gels that were stained with Coomassie blue.

Mincle ECD D10

10 20 30 40 50 60
 ANEIAALKQEIAALKQKGLSIFQLCDEKKFQPPGDSMELSCYNDGSGSVKNCCPLKWFHF
 KQEIAAL QLCDEKKFQPPGDSMELSCY FHF
 70 80 90 100 110 120
 QSSCYLFSPDTMSWRASLKNCSMGAHLVVINTQEEQEFLLYYTKPRKKEFYIGLTDQVTEGQW
 QSSCY TKPRKKEF YIGLTDQVTEGQW
 130 140 150 160 170 180
 QWVDGTPFTKLSLFWDAEEPNNLVTVEDCATIRDSSNPRQNWNDVPCFFNMFRVCEMPERKI
 SFWDAEEPNNLVTVEDCATIRDSSNPRQNW
 DAEEPNNLVTVEDCATIRDSSNPRQNWNDVPCFF RVCCEMPERKI

Dectin-2 ECD D12

10 20 30 40 50 60
 ANEIAALKQEIAALKKENAALKQKGLHFTYGETGKRLSELHSYHSSLTCFSEGTKVPAWGCCPA
 TYGETGKRLSELHSY
 70 80 90 100 110 120
 SWKSGSSCYFISSEKVVWSKSEQNCVEMGAHLVVFNTAEQNFIVQQLNESFSYFLGLSDPQG
 FISSEKVV NTEAEQNFIVQQL
 130 140 150 160 170 180 190
 NNNWQWIDKTPYEKNVRFWHLGEPNHSAEQCASIVFWKPTGWGWNDVICETRRNSICEMNKIYL
 HLGEPNHSAEQCASIVFW
 Dimerization domain Neck CRD

Figure 5.12. Coverage of peptides in the uncrosslinked samples of dectin-2 and mincle extracellular domains with fused synthetic dimerization domains. Polypeptide sequences of dectin-2 or mincle are segregated into regions of dimerization domain, neck domain and CRD. Peptides identified from molecular ion and fragmentation data are highlighted in magenta.

5.3 Discussion

The results for crystallization of the dectin-2 extracellular domain show that the geometry of the CRDs in the dimer matches closely to the arrangement deduced from analysis of the crystal packing for the isolated CRD. The close similarity of these structures in different crystals forms suggests that the arrangement of CRDs in the extracellular domain dimer is likely to be rigid. The extended interface in the new crystal expands the 4-stranded antiparallel β sheet across the dimer interface as observed for the BDCA-2 and dectin-2 CRDs. However, although the interface is more extensive, at 453 Å², it still falls short of the surface area predicted to be required for holding a dimer together. The two additional serine residues visible in the dimer crystals project to the side, indicating that the neck could bend around the interface to link to the transmembrane domain.

The observation that BDCA-2 CRD with an extended N-terminus can form a non-covalent dimer suggests that interactions critical to dimer formation may be found further toward the N-terminus of the neck region. The N-terminus of the extended BDCA-2 construct

ends exactly at Ser49, which is the site of the neck cysteine residue in the naturally occurring BDCA-2 variant. The fact that this fragment forms a partially stable dimer suggests that the stretch of 19 amino acids from this point to the beginning of the CRD is sufficient to allow interactions within the neck region and may be enough to facilitate formation of the interchain disulfide bond in the BDCA-2 variant. Ser49 is 6 amino acids away from the membrane. Taken together, these results provide an image in which the CRDs of a rigid receptor dimer, with both of the glycan-binding sites facing upward, are spaced from the cell surface by a stretch of 25 amino acid residues that form a flexible stalk.

Chapter 6 | General discussion

The oligomeric states of the three mincle family receptors mincle, dectin-2 and BDCA-2 and the orientations of their CRDs have been investigated.

Preformed dimers on cell surfaces

Mammalian cell expression systems have been employed to show that these receptors are likely to exist as preformed dimers on cell surfaces. Receptors linked to DAP-10 and DAP-12 are often drawn as monomers that can come together to make a dimer at the event of binding to a ligand. For example, when a NKG2D homodimer binds to its monomeric ligand, subunits A and B interact with the $\alpha 2$ and $\alpha 1$ helices respectively (Zuo et al., 2018). However, this is not likely the case for mincle family receptors as it is clear from the mammalian cell experiments that both BDCA-2 and mincle are expressed as preformed dimers on the cell surface. Thus, it is hypothesized that initiation of transmembrane signalling results from bringing together dimers into larger clusters.

The mammalian cell expression system has also been employed to elucidate the role of cysteine residues in the transmembrane and neck regions in dimer formation. Mincle has been shown to be expressed as a disulfide-bonded dimer linked through a single neck cysteine residue in each polypeptide. Although the two intramembrane cysteine residues in mincle are not involved in dimerization, they may be required for efficient processing of the polypeptide on its way to the cell surface. The dimerization mechanism for BDCA-2 is different from that for mincle, since BDCA-2 is a non-covalently linked dimer on the cell surface (Kim, 2018). However, a single-nucleotide polymorphism found in one individual gives rise to a BDCA-2 variant in which a serine residue in the neck region is replaced by a cysteine residue, corresponding to the interchain disulfide bond-forming cysteine residue in the neck region of mincle. Characterization of this BDCA-2 variant confirms formation of a disulfide-linked dimer as seen with mincle.

These results demonstrate that, although they are superficially different, the dimers formed by mincle and BDCA-2 are likely to involve similar interactions. Based in its similarity to BDCA-2, the same is likely true for dectin-2. In all three cases, key interface residues must lie in the neck region. Previous information that formation of a three-helix interface is required for assembly of a receptor molecule with an FcR γ dimer (Feng et al., 2005) would suggest a stoichiometric ratio of 2 receptor polypeptide to 4 FcR γ polypeptides for mincle family receptors.

Orientation of CRDs in receptor dimers

Although the neck region is important for dimer formation, the presence of the full neck and carbohydrate-recognition domains is not sufficient for stabilizing a dimer of mincle and

forming the interchain disulfide bond. Appending an N-terminal dimerization domain greatly improves dimerization efficiency. This attachment may act as a substitute for the transmembrane domain by providing critical interactions beyond the neck region and aligning the neck cysteine residues from the two mincle subunits. The ability of dimerization domains to stabilize dimers of dectin-2 and BDCA-2 as well as mincle led to identification of interactions between the CRDs seen in crystal structures of the CRDs from dectin-2 and BDCA-2. The importance of the neck domain was confirmed by demonstration that part of the BDCA-2 neck is sufficient to form partially dissociating dimers.

In spite of the similarities, the dimer interface in mincle may be somewhat different to the interfaces in dectin-2 and BDCA-2 due to the disulfide bond and an N-linked glycosylation site in the neck region of mincle. The CRDs of dectin-2 and BDCA-2 contain one (Asn131) and three (Asn110, Asn137, and Asn164) N-linked glycosylation sites respectively (Fig. 6.2B). Since none of these sites is close to the dimer interface, N-linked glycosylation is not likely to interfere with dimer formation in dectin-2 and BDCA-2 on the membrane. In the case of mincle, the N-linked glycosylation site in the CRD, at Asn107, corresponds to Asn110 of BDCA-2 and would thus not interfere with dimer formation. The other N-linked glycosylation site of mincle, Asn62, resides in the neck region, 3 amino acids before the start of the CRD. Since this site would be near where the polypeptide comes out of the dimer interface and starts to loop back around to the membrane, the glycosylation seems unlikely to interfere with dimer formation. However, N-linked glycans in IgA and NKp30 have been reported to promote oligomerization or binding to NKp30 ligands (Göritzer et al., 2020; Skořepa et al., 2020). In a similar way, the N-linked glycosylation site in the neck region of mincle could serve a biological function such as interfering with receptor clustering or dictating the orientation of clustering. (Fig. 6.1).

The geometry at which the binding sites are arranged in the BDCA-2 and dectin-2 dimers puts major constraints on how multivalent ligands would interact with these receptors. Two factors argue against the suggestion that signal initiation results from binding of multiple sugar residues in oligosaccharide ligands to the two binding sites in a dimer. First, the similar geometry in BDCA-2 and dectin-2 suggests that the orientations are likely rigid and thus the relative orientation of the CRDs is not likely to be changed by the ligands. Second, binding to the two subunits would probably be unfavourable, because the distance across the dimer is approximately 70 Å and would require a precise geometric fit to the binding sites (Fig. 6.2). However, it is possible to bring the binding sites in subunits of two different dimers within 13 Å of each other and these interactions could form in multiple different orientations. The resulting selection against binding of multivalent ligands to two subunits in a single dimer and preferential binding to two separate dimers is essential to initiate signalling by clustering, since binding of a ligand to two sites on the same receptor would not do anything.

The crystal structure of dectin-2 with a bivalent oligosaccharide provides some insight into potential ways that CRDs could be brought together by oligosaccharide ligands. However, the geometry of the dimers observed is distorted from the likely arrangement on the cell surface, since the two dimers must be bent around to accommodate the relatively small oligosaccharide (Fig. 6.3). While this distortion probably reflects the small size of the oligosaccharide ligand, in which three sugar residues bridge between the sugars that interact with Ca^{2+} in the two CRDs, it is also possible that binding in this orientation could cause bending of the membrane, which might be an alternative mechanism of initiating signalling.

Knowledge of receptor dimer geometry will facilitate design of bivalent or multivalent stimulatory ligands with appropriate spacers. Such ligands would need to be tested on cells, since interaction of the bridging ligand would only need to stabilize dimer formation in the constrained, two-dimensional environment of the membrane. In addition, the dectin-2 results suggest that multivalent binding with inappropriate geometry might be observed when the CRD is in solution and not on the cell surface. Synthetic ligands that cause clustering on cells could have utility in modulation of the immune response. For example, mincle stimulation has been targeted for adjuvant action in vaccines and BDCA-2 stimulation is being targeted for suppression of interferon α/β secretion in systemic lupus erythematosus and other autoimmune diseases (Decout et al., 2017; Furie et al., 2019).

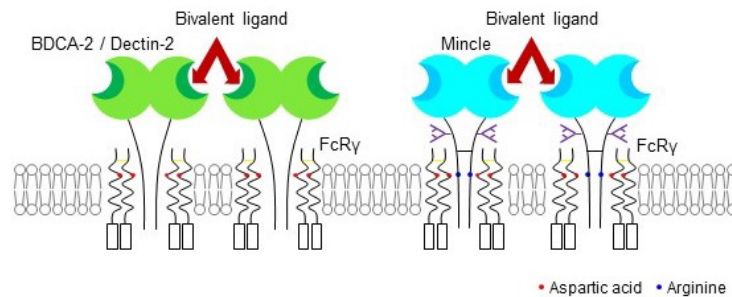


Figure 6.1. Summary of what dimers of mincle and BDCA-2 or dectin-2 may look like on the cell membrane. N-linked glycosylation in the neck region of mincle is shown in purple. Bivalent/multivalent ligands that bind to carbohydrate-recognition domains of adjacent dimers are shown as red arrows. Disulfide bond in the neck region of mincle is shown as a black line.

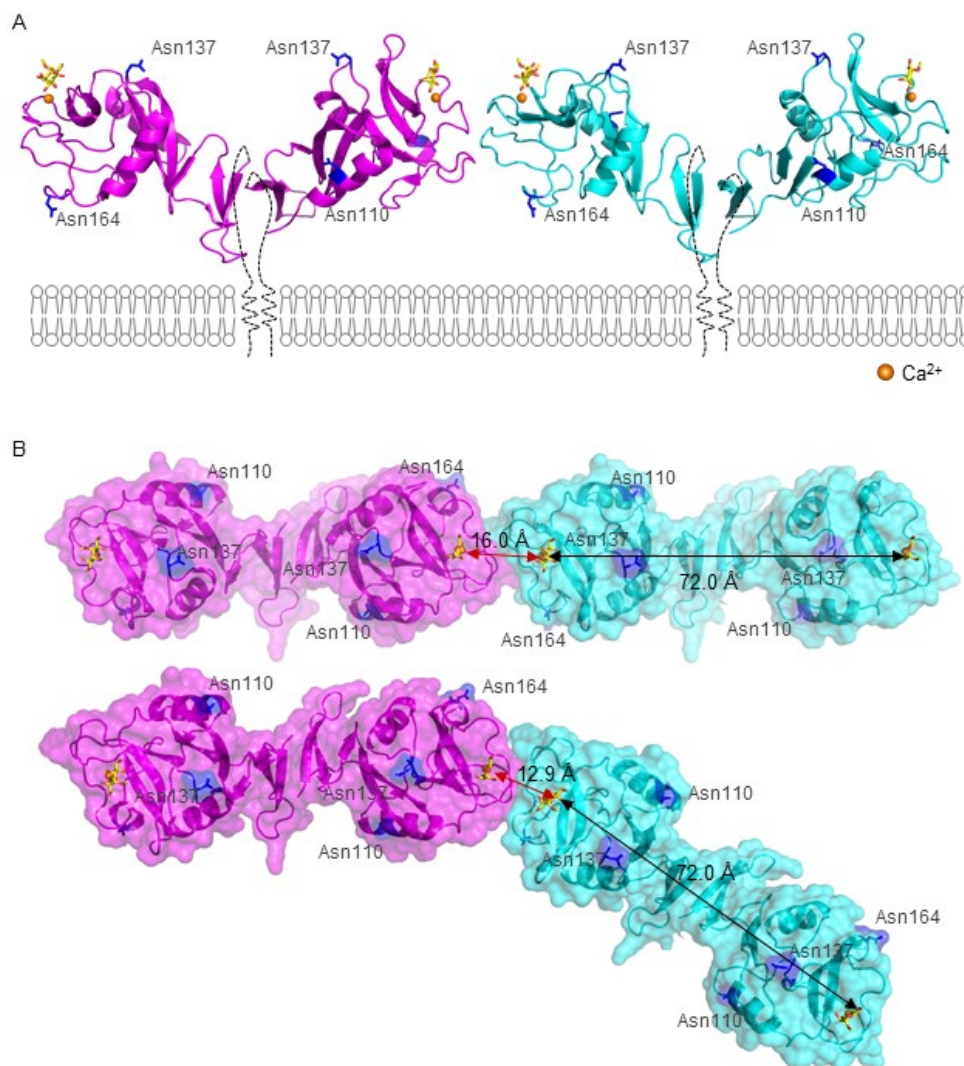


Figure 6.2. Potential mechanism for ligand crosslinking of receptor dimers on the cell membrane. (A) Side view of the proposed dimer showing its potential orientation relative to the membrane. Stretches of amino acids that loop around to the membrane are shown as black dashed lines. (B) Top view looking down on the membrane, comparing the spacing of binding sites within a dimer (red arrow) and between adjacent dimers (black arrow) in two orientations. Asparagine residues bearing N-linked glycosylation sites are shown as sticks in blue in both panels. Bound oligosaccharides are shown facing out of the page. Based on PDB entry 4ZES.

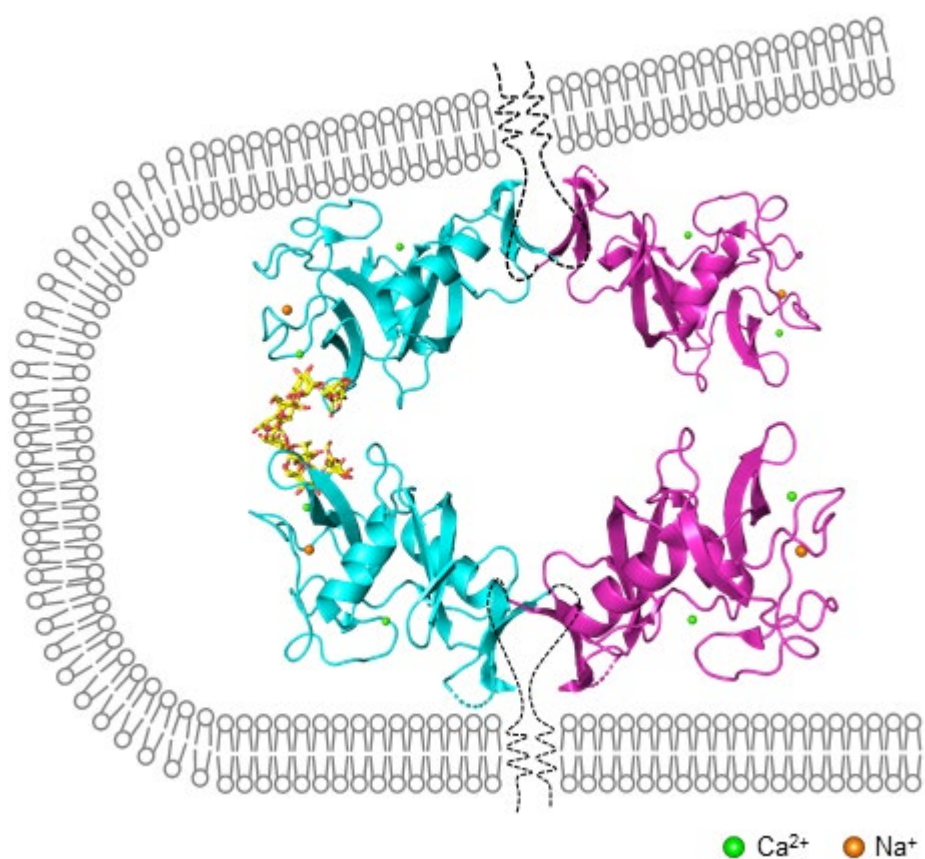


Figure 6.3. Distorted arrangement of dectin-2 where receptors are bent over to accommodate a small oligosaccharide. Based on PDB entry 5VYB. Ligand-induced dimer is shown in cyan and the other subunits in each preformed dimer are shown in magenta. Carbon backbones of the sugar molecules are yellow.

Chapter 7 | Materials and Methods

7.1 Molecular Biology

Modification of expression plasmids for expression in HEK cells

Mutagenesis of the BDCA-2 neck serine residue to cysteine was achieved by a two-step polymerase chain reaction (PCR). Forward and reverse primers were dissolved in deionized water to 100 μ M and diluted in deionized water to make 10 μ M working stocks. The PCR mix contained a final concentration of 0.22 μ M deoxynucleotide triphosphates, and 1X Advantage-HF 2 polymerase. For the first PCR step, a pCR2.1-TOPO plasmid containing the wild-type BDCA-2 cDNA was used as a template (Jégouzo et al., 2015). Forward primer bdca2mut5 and reverse primer bdca2mut7 were added at concentrations of 5 μ M to create the front part of the construct with the neck mutation. Forward primer bdca2mut6 and reverse primer bdca2mut8 were added at concentrations of 5 μ M to produce the back part of the construct. The reactions were heated at 95°C for 2 min, followed by 40 cycles of 30-s denaturation at 95°C followed by 2-min renaturation and elongation at 65°C. The PCR products were eluted from a 2% agarose gel, lyophilized and resuspended in 100 μ l of 10 mM Tris-Cl, 1 mM EDTA, pH 8.0. For the second PCR step, forward primer bdca2mut5 and reverse primer bdca2mut8 were added at concentrations of 5 μ M to the PCR mix along with the PCR products encoding the front part and back parts of the construct from the first PCR step. Amplified DNA was isolated directly from agarose gels and cloned into vector pCR-TOPO (Invitrogen) following the manufacturer's protocol.

Fragments encoding receptors were generated by Not1 and Sac1 digestion of vectors containing the PCR products or synthetic cDNAs and isolated on 1.5% agarose gels. Ligation was performed by combining insert with wild-type mincle/pSF-CMV-EMCV-Zeo expression plasmid digested with Not1 and Sac1 and treated with calf-intestinal alkaline phosphatase in a ligation reaction containing 1X T4 DNA ligase buffer and 40 units/ μ l T4 DNA ligase (New England Biolabs). After incubation for 2.5 h at 18°C, followed by 1 h at room temperature, the products were used to transform *Escherichia coli* strain DH5 α (Invitrogen). Single colonies were picked and grown in Luria-Bertani (LB) medium. Plasmid DNA isolated from individual minipreps was sequenced to check for integrity (Eurofin Genomics).

Transfection of HEK 293 cells

Single colonies of *E. coli* containing the expression constructs were picked and grown in 100 ml of LB medium containing 50 μ g/ml kanamycin. Cultures were grown at 37°C with shaking for 12-16 hours. Bacterial growth was monitored so that the OD₆₀₀ did not exceed 4. Purification of the expression plasmids employed the NucleoBond Xtra Midi plasmid protocol (Macherey-Nagel). Purified expression plasmids were sterilized by ethanol precipitation.

Transfection of HEK 293 cells was carried out with the Lipofectamine 3000 reagent (Invitrogen). To make up DNA solution for transfection, the concentration of sterilized DNA was determined by measurement of A_{260} . BDCA-2 and FcR γ plasmids (1.25 μ g each) were added to 125 μ l of OptiMEM medium with 5 μ l of p3000 reagent. The plasmids were mixed with 3.75 μ l of lipofectamine 3000 reagent in 125 μ l of OptiMEM. A twice-concentrated DNA solution was prepared by doubling the amounts of each plasmid and mixing with doubled portions of lipofectamine 3000 reagent in the same volumes of OptiMEM. DNA solutions at two concentrations were incubated at room temperature for 15 min, added to HEK 293 cells at 90% confluence and incubated at 37°C, 5% CO₂ for 5 hours, after which the medium was changed to Dulbecco's Modified Eagle's Medium (DMEM, Life Technologies) supplemented with 10% fetal calf serum and 10 U/ml penicillin-streptomycin (Gibco). On the next day, each plate of transfected cells was passaged onto a 90 mm plate and grown in fresh DMEM with 10% fetal calf serum and 10 U/ml penicillin-streptomycin. After 48 hours, cells were moved to DMEM containing 10% fetal calf serum, 10 U/ml penicillin-streptomycin, 300 μ g/ml zeocin (Invitrogen) and 200 μ g/ml hygromycin (Invitrogen). Following selection for 10 - 14 days, single colonies of zeocin- and hygromycin-resistant cells were picked and transferred to a 24-well plate before being passaged into 25 cm² or 75 cm² flasks.

Creation of mincle, dectin-2 and BDCA-2 extracellular domain expression constructs

Pairs of forward and reverse oligonucleotides encoding cow mincle neck domain, and cow mincle neck domain with an N-terminal GCN4 dimerization domain and a factor Xa cleavage site were synthesized (Thermo Fisher). For kinase treatment and annealing of oligonucleotides, 100 μ M stocks of oligonucleotides were prepared in deionized water. Stock solutions were diluted to 25 μ M in 20 μ l of 1X T4 polynucleotide kinase buffer containing 1 mM ATP and 10 units of T4 polynucleotide kinase (New England Biolabs). Following incubation for 1 h at 37°C, pairs of oligonucleotide were pooled and purified by phenol-chloroform extraction, followed by ethanol precipitation and lyophilization. Lyophilized oligonucleotides were dissolved in 30 μ l of 50 mM Tris-Cl pH 7.4 containing 10 mM MgCl₂, denatured by heating for 2 min at 100°C and cooled to room temperature to allow annealing of forward and reverse oligonucleotides. Ligation was performed by mixing diluted phosphorylated oligonucleotides with expression vector pT5T (Eisenberg et al., 1990) digested with appropriate restriction enzymes and treated with calf-intestinal alkaline phosphatase. Ligation reaction containing 1X T4 DNA ligase buffer and 40 units/ μ l T4 DNA ligase (New England Biolabs) were incubated for 2.5 h at 18°C, followed by 1 h at room temperature, and were used to transform *E. coli* strain DH5 α . All of the final plasmids were checked by DNA sequencing.

7.2 Purification of proteins from transfected HEK cells

Affinity resins

Preparation of trehalose-Sepharose and mannose-Sepharose resins followed the divinyl sulfone coupling methods (Fornstedt and Porath, 1975). Desialylated egg-glycopeptide agarose column was prepared as described by Jégouzo et al. (2015).

BDCA-2

Medium was removed from two 75-cm² flasks of HEK BDCA-2 / FcRy transfectants. Cells were rinsed twice with 10 ml of phosphate-buffered saline and harvested. Cells were lysed by dounce homogenization (7 ml; Wheaton) in 5 ml of cold lysis buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM CaCl₂, 10 mM *N*-ethylmaleimide, 1% Triton X-100 (Roche), and 50 µl of protease inhibitor (Merck Millipore). The lysate was centrifuged at 16,000 g, 4 °C for 10 min. The supernatant was loaded onto a 1-ml desialylated egg-glycopeptide column, which was rinsed with 5 ml of wash buffer (0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM CaCl₂, and 0.1% Triton X-100) and eluted in five 0.5-ml fractions with elution buffer (0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM EDTA and 0.1% Triton X-100). Eluted fractions 2 to 5 were precipitated by addition of ½ volume of 30% trichloroacetic acid (TCA) and incubated for 10 min at 4 °C. TCA-precipitated fractions were centrifuged at 16,000 g for 5 min at room temperature, washed twice with 0.5 ml of ethanol:ether (1:1) and lyophilized for 5 min. Pellets were dissolved in 1X gel sample buffer with or without 1% 2-mercaptoethanol for SDS-polyacrylamide gel electrophoresis.

Mincle

Medium was removed from two 75-cm² flasks of HEK mincle / FcRy transfectants. Cells were rinsed twice with 10 ml of phosphate-buffered saline and harvested. Cells were lysed by homogenization in a 1-ml syringe in 1 ml of cold lysis buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM CaCl₂, 1% Triton X-100 (Roche), 10 µl of protease inhibitor (Merck Millipore), and 0.5 µl of Benzonase nuclease (Sigma). Lysate was centrifuged at 16,000 g, 4 °C for 10 min. Supernatant was loaded onto a 2-ml trehalose-Sepharose column. Fractions 1 to 3 were collected by rinsing the column three times with 1-ml wash buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM CaCl₂, and 0.1% Triton X-100. A further 0.5-ml wash was collected with the first 0.5-ml elution (fraction 4) in elution buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, 2.5 mM EDTA and 0.1% Triton X-100, followed by collection of three 1-ml elution fractions (fractions 5 to 7). Eluted fractions 5 to 7 were precipitated by addition of TCA and prepared for SDS-polyacrylamide gel electrophoresis as above.

7.3 Expression of protein fragments in bacteria

Expression and purification of extracellular domain fragments

Proteins were expressed from pT5T plasmids in *E. coli* strain BL21 (DE3). Cells were grown in LB medium containing 50 µg/m ampicillin with shaking to an A_{550} of 0.7 at 37 °C, isopropyl β-D-thiogalactoside was added to 50 µM and incubation was continued for a further 2.5 h at 37 °C. Cultures were harvested by centrifugation at 4,000 rpm for 15 min at 4 °C in a Beckman JLA-8.1000 rotor. Cell pellets were washed by resuspension in 10 mM Tris-Cl pH 7.8 and sonicated with four bursts of 60 sec following by 30-min cooling intervals. Inclusion bodies were collected at 10,000 rpm for 15 min at 4 °C in a Thermo Scientific Fiberlite F21-8x50y rotor, resuspended in 6 M guanidine-HCl, 100 mM Tris-Cl, pH 7.0, by homogenization: 100 ml for a 6-L preparation or 30 ml for a 2-L preparation. 2-Mercaptoethanol was added to 0.01% followed by incubation for 30 min at 4 °C.

After 30-min centrifugation at 35,000 rpm in a Beckman type 70.1 Ti rotor, the supernatant was subject to direct dialysis or dilution dialysis. For direct dialysis, the supernatant was dialyzed against 2 l of a dialysis buffer containing 0.5 M NaCl, 25 mM Tris-Cl 7.8, 25 mM CaCl₂, with constant stirring at 4 °C, and buffer was changed three times, followed by removal of precipitates by 30-min centrifugation at 18,000 rpm in a Thermo Scientific Fiberlite F21-8x50y rotor. Alternative renaturation schemes were either batch dilution or gradual dilution. For batch dilution, the supernatant was dripped slowly into 4 volumes of dialysis buffer with stirring, to achieve a 1 in 5 dilution, followed by dialysis against three changes of the same buffer. For gradual dilution, dialysis buffer was slowly added to the supernatant with stirring to achieve a 5-fold dilution, followed by dialysis with three changes. For renaturation of mincle extracellular domain without dimerization domain or with GCN4 dimerization domain, direct dialysis and batch dilution schemes were attempted where specified. For renaturation of mincle extracellular domain with synthetic dimerization domain, Triton X-100 was added to 1% to the guanidine HCl-solubilized protein before gradual dilution. Gradual dilution was employed for dectin-2 renaturation. Batch dilution was used for BDCA-2 renaturation.

After centrifugation and filtering through glass wool, the supernatant of mincle was passed through either a 10-ml trehalose-Sepharose column or a 2-ml trehalose-Sepharose column. Dectin-2 was passed through a 2-ml mannose-Sepharose column. The 10-ml sugar-Sepharose columns were rinsed with 2 + 8 ml of 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 25 mM CaCl₂, and eluted in fifteen 1-ml fractions of 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 2.5 mM EDTA. The 2-ml sugar-Sepharose columns were rinsed with 2 + 8 ml of 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 25 mM CaCl₂, and eluted in ten 0.5-ml fractions of 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 2.5 mM EDTA. For BDCA-2, the centrifuged supernatant was passed through a 5-ml desialylated egg-glycopeptide agarose column, rinsed with 2 + 8 ml of 0.15 M

NaCl, 25 mM Tris-Cl pH 7.8, and 25 mM CaCl₂, and eluted in ten 1-ml fractions of 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 2.5 mM EDTA.

For structural analysis, fractions of affinity-purified receptor extracellular domain with appended synthetic dimerization domain were pooled, dialyzed against water, lyophilized, and re-dissolved in water. The pH of the protein solution was adjusted to between 7 and 8 using 1 M NaOH. The protein concentration was determined by absorbance at 280 nm. Neutralized protein solution was re-lyophilized and 5 mg of BDCA-2 extended CRD, 5.2 mg of Dectin-2 ECD D12, 1.1 mg of Mincle ECD D9 and 1.2 mg of Mincle ECD D10 were supplied to Hadar Feinberg (Stanford University) for a crystallization screen and structure determination.

Cu²⁺-o-phenanthroline catalysis of inter-chain disulfide bond formation

Purified human mincle extracellular domain (estimated 10 µg) was dialyzed against water, lyophilized and re-dissolved in 1/10 its original volume (100 µl) of a buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 2.5 mM CaCl₂. Ten-times concentrated protein was incubated without or with a mix of 1 mM 1,10-phenanthroline and 0.5 mM cupric acetate at room temperature for 30 min before addition of 2X sample buffer without or with 2% 2-mercaptoethanol and separation on an SDS-polyacrylamide gel.

Protease cleavage

Stock concentrations of proteases were: 2 µg/ml enterokinase (New England Biolabs); 1 mg/ml Factor Xa (New England Biolabs); and 40 mg/ml thrombin, prepared freshly by dissolving 2 mg of lyophilized thrombin (Sigma-Aldrich) in 50 µl of water. Enzymes diluted at a 1:10 ratio were prepared by mixing the stock enzyme solution with digestion buffer containing 0.15 M NaCl, 25 mM Tris-Cl pH 7.8, and 2.5 mM CaCl₂. Substrate was prepared by adding CaCl₂ to 5 mM to the 6th eluted fractions of affinity-purified Mincle ECD P1, P2, and P3. Aliquots (20 µl, estimated 0.6 µg) of substrate were incubated without enzyme or with 1 µl of 1:10 diluted enzymes or 1 µl of enzyme stocks. For enterokinase digestion and factor Xa digestion, samples were incubated at 37 °C for 3 h. For thrombin digestion, samples were incubated at 37 °C for 18 h. Negative controls minus substrate were prepared by mixing 10 µl of digestion buffer with 10 µl of 1:10 diluted enzymes. Reactions were stopped by mixing sample with equal volume of 2X sample buffer without or with 2% 2-mercaptoethanol and heating at 100 °C for 5 min. Portions (20 µl) of each sample were resolved on SDS-polyacrylamide gels.

The stock concentration of furin (New England Biolabs) was 2,000 U/ml. Substrate was prepared by adding CaCl₂ to 5 mM to the 5th eluted fraction of affinity-purified Mincle ECD P4. For the optimized protocol, 5 U of furin was added per 100 µl of substrate and reaction was incubated at 30 °C for 18 h. The minus enzyme control contained 20 µl of substrate and was kept 4 °C. For SDS-polyacrylamide gel electrophoresis, 20 µl aliquots of control and digested

samples were mixed with an equal volume of 2X sample buffer in the presence or absence of 2% 2-mercaptoethanol, heated at 100 °C for 5 min, and 20 µl of each sample was analyzed on an SDS-polyacrylamide gel.

To test effects of temperature and detergent on furin digestion, 5 U of furin was added per 100 µl of substrate. Samples were incubated in the presence or absence of 0.5% Triton X-100, and reactions were incubated at 30 °C or at 37 °C for 18 h. After incubation, each sample was centrifuged at 16,000 g for 5 min, and supernatant was separated from pellet. The minus enzyme control contained 20 µl of substrate and was kept 4 °C. For SDS-polyacrylamide gel electrophoresis, 20 µl of control and supernatant of digested samples (containing 2 U of furin) were mixed with equal volume of 2X sample buffer, and pellets were resuspended in 20 µl of 1X sample buffer. After heating at 100 °C for 5 min, 20 µl of each sample was analyzed on an SDS-polyacrylamide gel.

To test effects of prolonged digestion time and amount of enzyme on completion of reaction, 0.8 U or 3.2 U of furin were added per 20 µl of substrate. Triton X-100 was added to all samples to 0.5%. The minus enzyme control contained 20 µl of substrate and was kept 4 °C until SDS-polyacrylamide gel electrophoresis. After incubating at 37 °C for 18 h, 42 h, or 66 h, reactions were stopped by mixing with an equal volume of 2X sample buffer and heating at 100 °C for 5 min. Aliquots (20 µl) of each sample were resolved on SDS-polyacrylamide gels.

7.4 Protein analysis

PNGase F / Endo H_f treatment

Following affinity purification of BDCA-2, fractions 2 to 5 were pooled and TCA precipitated. Pellets were resuspended in 0.5% SDS containing 40 mM dithiothreitol, and heated at 100 °C for 10 min. For PNGase treatment, the sample was incubated with 0.2, 0.5 or 1 unit of PNGase F (Roche) in 50 mM sodium phosphate pH 7.5 and 1% NP-40 at 37 °C for 15 min. For Endo H_f treatment, the sample was incubated with 200, 500 or 1,000 units of Endo H_f (New England Biolabs) in 50 mM sodium acetate pH 6 at 37 °C for 15 min. Sample incubation in the absence of either enzyme were carried out as a control. Reactions were stopped by addition of equal volumes of 2X gel sample buffer containing 2% 2-mercaptoethanol and heating at 100 °C for 5 min.

Following affinity purification of mincle, fractions 5 to 7 were pooled, TCA precipitated, treated with PNGase F or Endo H_f, and prepared for SDS-polyacrylamide gel electrophoresis as described above.

Chemical crosslinking

Crosslinking reagent *bis*(sulfosuccinimidyl) suberate (BS³; Pierce) was dissolved in deionized water to 12 mg/ml and serially diluted in water. Cell lysate of one 75 cm² flask of

confluent HEK transfectants co-expressing BDCA-2 or mincle was fractionated by affinity chromatography on a 1-ml desialylated egg-glycopeptide column or a 1-ml trehalose-Sepharose column respectively. After application of the sample, the column was rinsed with 5 ml of wash buffer (0.15 M NaCl, 25 mM HEPES pH 7.8, 2.5 mM CaCl₂, and 0.1% Triton X-100) and eluted in five 0.5-ml fractions with elution buffer (0.15 M NaCl, 25 mM HEPES pH 7.8, 2.5 mM EDTA and 0.1% Triton X-100). Dilutions of BS³ were added to affinity-purified BDCA-2 or mincle at a 1:10 ratio and incubated at room temperature for 30 min. TCA precipitation was carried out as described above and pellets were dissolved in 1X gel sample buffer containing 1% 2-mercaptoethanol.

Dectin-2 extracellular domain for chemical crosslinking with bifunctional reagents was re-purified by affinity chromatography on a 1-ml mannose-Sepharose column that was washed with 3 ml of 150 mM NaCl, 25 mM HEPES pH 7.8, 2.5 mM CaCl₂ and eluted with 0.5-ml aliquots of 150 mM NaCl, 25 mM HEPES pH 7.8, 2.5 mM EDTA. BS³ and disuccinimidyl tartrate (Pierce) were dissolved in water to 12 mg/ml and 50 mg/ml respectively and serially diluted in water. 1,5-Difluoro-2,4-dinitrobenzene (Pierce) was dissolved in ethanol to 10 mg/ml and serially diluted in ethanol. Diluted reagents were added to 15 µl of dectin-2 (estimated 3.5 µg) at a 1:10 ratio and incubated at room temperature for 30 min. Dectin-2 extracellular domain (1.5 ml, estimated 0.35 mg) for crosslinking with 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide (Fluorochem), was dialyzed against 100 mM NaCl, 50 mM MES pH 6.5 with three buffer changes. *N*-Hydroxysuccinimide (Sigma-Aldrich) and *N*-hydroxysulfosuccinimide (Sigma-Aldrich) were dissolved in deionized water to 23 mg/ml and 43 mg/ml respectively and added to aliquots (15 µl) of dectin-2 at a 1:10 ratio. 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide was dissolved in deionized water to 192 mg/ml, added to the reactions at ratios of 1:15, 1:30 or 1:60, and incubated at room temperature for 1 h. Reactions were stopped by adding 2X sample buffer containing 1% 2-mercaptoethanol and were immediately analyzed on a 17.5% SDS-polyacrylamide gel.

Computational analysis

Molecular structures were visualized, analyzed and rendered for figures in PyMol. Buried surface area was analyzed using the PDBePISA software on the EMBL-EBI website (https://www.ebi.ac.uk/msd-srv/prot_int/pistart.html).

Gels and blotting

SDS-polyacrylamide gel electrophoresis was performed on gels containing 17.5% acrylamide according to Laemmli (1970), followed by Coomassie blue staining or western blotting onto nitrocellulose (Burnette, 1981). Blots were blocked in 5% bovine serum albumin in Tris-buffered saline, pH 7.4 for 30 min at room temperature, followed by overnight incubation

with primary antibody or antiserum. Blots were washed in four changes of Tris-buffered saline for 5 min, 10 min, 10 min, and 5 min with 0.1% Triton X-100 in the middle two washes. After incubation with an alkaline phosphatase-conjugated secondary antibody for 60 min at room temperature, blots were washed as above and developed in nitro-blue tetrazolium / 5-bromo-4-chloro-3-indolyl phosphate (BCIP/NBT; Merck).

For detecting BDCA-2, rabbit polyclonal anti-BDCA-2 antibody that was raised against the BDCA-2 CRD expressed in *E. coli* was used at 1:5000 dilution as a primary antibody and protein A conjugated to alkaline phosphatase (Millipore) was used at 1:5000 dilution as a secondary reagent. Mouse monoclonal anti-CLECSF9 antibody (Abnova) was used at 1:10,000 dilution as a primary antibody for detecting mincle and goat anti-mouse IgG and IgM antibody conjugated to alkaline phosphatase (Jackson ImmunoResearch Laboratories) was used at 1:5000 dilution as a secondary antibody. For detecting FcR γ , rabbit polyclonal anti-FcR γ antibody (Millipore) was used at 1:2000 dilution as a primary antibody and protein A conjugated to alkaline phosphatase was used at 1:5000 dilution as a secondary reagent.

Gel-filtration analysis

Gel-filtration analysis was performed on a 1 x 30 cm Superdex-S75 column (GE Healthcare Life Sciences). For analysis of resolubilized samples, affinity-purified protein was dialyzed against water, lyophilized, redissolved in gel filtration elution buffer containing 10 mM Tris-Cl, pH 7.8, 100 mM NaCl and 2.5 mM EDTA, and centrifuged at 16,000 g for 5 min. The supernatant was loaded onto the Superdex-S75 column, which was eluted with gel filtration elution buffer at a flow rate of 0.5 ml/min. Absorbance was monitored at 280 nm. Elution positions of the following standard proteins (with M_r as indicated) were determined: Aprotinin (6.5 kDa), cytochrome c (12.4 kDa), bovine erythrocyte carbonic anhydrase (29 kDa), ovalbumin (43 kDa), and bovine serum albumin (67 kDa).

Mass spectrometry

To prepare peptides from uncrosslinked human dectin-2 extracellular domain with a 3-heptad dimerization domain or uncrosslinked bovine mincle extracellular domain with a 2- or 3-heptad dimerization domain for mass spectrometry, 100 μ l of affinity-purified dectin-2 (66 μ g) or each affinity-purified mincle construct (10 μ g) was incubated in 400 μ l of 6 M guanidine-HCl, pH 7, to which dithiothreitol was added to 5 mM. Reactions were incubated for 20 min at 56 °C. To carboxymethylate free thiol groups, sodium iodoacetate was added to a final concentration of 15 mM and samples were incubated for 20 min at room temperature in the dark. Carboxymethylated samples were dialyzed against deionized water with three changes and lyophilized.

For in-solution chymotrypsin digestion, a stock solution of sequencing-grade chymotrypsin (Promega) was prepared by dissolving to 0.1 µg/µl in 0.5% NH₄HCO₃. A digestion buffer was prepared by diluting ProteaseMAX Surfactant (Promega) to 0.01% in 0.5% NH₄HCO₃. Reactions were set up by adding 2.5 µl of stock chymotrypsin solution to lyophilized mincle in 50 µl of digestion buffer and adding 6 µl of stock chymotrypsin solution to lyophilized dectin-2 in 50 µl of digestion buffer. Samples were incubated overnight at 25 °C. To inactivate chymotrypsin, a 10% solution of trifluoroacetic acid was prepared in deionized water and this solution was added to make the solution 0.5% trifluoroacetic acid followed by mixing and lyophilization. To retrieve and enrich peptides, a 10 µl Omix C18 tip (Agilent Technologies) that had been wetted with 50% acetonitrile and washed with 0.1% trifluoroacetic acid was used to aspirate and dispense each sample 10 times, followed by 10 washes with 0.1% trifluoroacetic acid, and the hydrophobic peptides were eluted with 10 µl of 50% acetonitrile in 0.05% trifluoroacetic acid.

Matrix was prepared by dissolving 10 mg of α-cyano-4-hydroxycinnamic acid in 1 ml of 50% acetonitrile, 0.05% trifluoroacetic acid in water. Peptide solution (0.5 µl) was spotted on to a sample plate on which 0.5 µl of matrix has been spotted on and allowed to dry. Mass spectrometry data were collected on an Applied Biosystems 4800 time-of-flight mass spectrometer. Peak picking for peptide mass fingerprint data was performed using Data Explorer software.

Solubility test

Affinity-purified receptor extracellular domain dimers were pooled, dialyzed against water and lyophilized. Lyophilized protein was re-dissolved in 25 µl of water and the pH was adjusted to between 7 and 8 using pH paper. Tris-Cl buffer, pH 7.8 was added to a final concentration of 10 mM and the sample was kept at 4 °C for 16 h. The sample was centrifuged at 16,000 g for 5 min and the pellet was resuspended in 25 µl of 10 mM Tris-Cl, pH 7.8.

References

Ambati, S., Ellis, E. C., Lin, J., Lin, X., Lewis, Z. A. and Meagher, R. B. (2019) Dectin-2-targeted antifungal liposomes exhibit enhanced efficacy. *mSphere* **4**, e00715-19.

Ariizumi, K., Shen, G. L., Shikano, S., Xu, S., Ritter 3rd, R., Kumamoto, T., Edelbaum, D., Morita, A., Bergstresser, P. R. and Takashima, A. (2000) Identification of a novel, dendritic cell-associated molecule, dectin-1, by subtractive cDNA cloning. *J. Biol. Chem.* **275**, 20157-20167.

Balch, S. G., McKnight, A. J., Seldin, M. F. and Gordon, S. (1998) Cloning of a novel C-type lectin expressed by murine macrophages. *J. Biol. Chem.* **273**, 18656-18664.

Berry, R. and Call, M. E. (2017) Modular activating receptors in innate and adaptive immunity. *Biochemistry* **56**, 1383-1402.

Blázquez-Moreno, A., Park, S., Im, W., Call, M. J., Call, M. E. and Reyburna, H. T. (2016) Transmembrane features governing Fc receptor CD16A assembly with CD16A signaling adaptor molecules. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E5645-E5654.

Brown, G. D. (2006) Dectin-1: a signalling non-TLR pattern-recognition receptor. *Nat. Rev. Immunol.* **6**, 33-43.

Burgess, A. W., Cho, H. S., Eigenbrot, C., Ferguson, K. M., Garrett, T. P., Leahy, D. J., Lemmon, M. A., Sliwkowski, M. X., Ward, C. W. and Yokoyama, S. (2003) An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol. Cell.* **12**, 541-552.

Burnette, W. N. (1981) "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195-203.

Call, M. E., Pyrdol, J., Wiedmann, M. and Wucherpfennig, K. W. (2002) The organizing principle in the formation of the T cell receptor-CD3 complex. *Cell* **111**, 967-979.

Call, M. E. and Wucherpfennig, K. W. (2007) Common themes in the assembly and architecture of activating immune receptors. *Nat. Rev. Immunol.* **7**, 841-850.

- Call, M. E., Wucherpfennig, K. W. and Chou, J. J. (2010) The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nat. Immunol.* **11**, 1023-1029.
- Cao, W., Zhang, L., Rosen, D. B., Bover, L., Watanabe, G., Bao, M., Lanier, L. L. and Liu, Y. (2007) BDCA2/FcεRIγ complex signals through a novel BCR-Like pathway in human plasmacytoid dendritic cells. *PLoS Biol.* **5**, e248.
- Caparrós, E., Munoz, P., Sierra-Filardi, E., Serrano-Gómez, D., Puig-Kröger, A., Rodríguez-Fernández, J. L., Mellado, M., Sancho, J., Zubiaur, M. and Corbí A. L. (2006) DC-SIGN ligation on dendritic cells results in ERK and PI3K activation and modulates cytokine production. *Blood* **107**, 3950-3958.
- Ciani, B., Bjelić, S., Honnappa, S., Jawhari, H., Jaussi, R., Payapilly, A., Jowitta T., Steinmetz, M. O. and Kammerer, R. A. (2010) Molecular basis of coiled-coil oligomerization-state specificity. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19850-19855.
- Collins, B.E., Smith, B.A., Bengtson, P. and Paulson, J.C. (2006) Ablation of CD22 in ligand-deficient mice restores B cell receptor signaling. *Nat. Immunol.* **7**, 199-206.
- Dawson, J.P., Berger, M.B., Lin, C.C., Schlessinger, J., Lemmon, M.A. and Ferguson, K.M. (2005) Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. *Mol. Cell. Biol.* **25**, 7734-7742.
- Decout, A., Silva-Gomes, S., Drocourt, D., Barbe, S., André, I., Cueto, F. J., Lioux, T., Sancho, D., Pérouzel, E., Vercellone, A., Prandi, J., Gilleron, M., Tiraby, G. and Nigou, J. (2017) Rational design of adjuvants targeting the C-type lectin Mincle. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2675-2680.
- Doñate, F., Kelly, C.R., Ruf, W. and Edgington, T. S. (2000) Dimerization of tissue factor supports solution-phase autoactivation of factor VII without influencing proteolytic activation of factor X. *Biochemistry* **39**, 11467-11476.
- Drickamer K. (1992) Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* **360**, 183–186.

Drickamer, K. and Taylor, M. E. (2015) Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* **34**, 26-34.

Drummond, R. A., Saijo, S., Iwakura, Y. and Brown, G. D. (2011) The role of Syk/CARD9 coupled C-type lectins in antifungal immunity. *Eur. J. Immunol.* **41**, 276-281.

Dzionic A., Fuchs A., Schmidt P., Cremer S., Zysk M., Miltenyi S., Buck D. W. and Schmitz J. (2000) BDCA-2, BDCA-3, and BDCA-4: Three markers for distinct subsets of dendritic cells in human peripheral blood. *J. Immunol.* **165**, 6037–6046.

Dzionic, A., Sohma, Y., Nagafune, J., Cella, M., Colonna, M., Facchetti, F., Günther, G., Johnston, I., Lanzavecchia, A., Nagasaka, T., Okada, T., Vermi, W., Winkels, G., Yamamoto, T., Zysk, M., Yamaguchi, Y. and Schmitz, J. (2001) BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction. *J. Exp. Med.* **194**, 1823-1834.

Eaton, E., Rodriguez, H. and Vehar, G. A. (1986) Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and activated protein C with activation and inactivation of factor VIII coagulant activity. *Biochemistry* **25**, 505–512.

Eisenberg S. P., Evans R. J., Arend W. P., Verderber E., Brewer M. T., Hannum C. H., Thompson R. C. (1990) Primary structure and functional expression from complementary DNA of a human interleukin-1 receptor antagonist. *Nature* **343**, 341-346.

Feinberg, H., Park-Snyder, S., Kolatkar, A. R., Heise, C. T., Taylor, M. E. and Weis, W. I. (2000) Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor. *J. Biol. Chem.* **275**, 21539-21548.

Feinberg, H., Castelli, R., Drickamer, K., Seeberger, P. H. and Weis, W. I. (2007) Multiple modes of binding enhance the affinity of DC-SIGN for high mannose N-linked glycans found on viral glycoproteins. *J. Biol. Chem.* **282**, 4202-4209.

Feinberg, H., Powlesland, A. S., Taylor, M. E. and Weis, W. I. (2010) Trimeric structure of langerin. *J. Biol. Chem.* **285**, 13285-13293.

Feinberg, H., Jégouzo, S. A. F., Rowntree, T. J., Guan, Y., Brash, M. A., Taylor, M. E., Weis, W. I. and Drickamer, K. (2013) Mechanism for recognition of an unusual mycobacterial glycolipid by the macrophage receptor mincle. *J. Biol. Chem.* **288**, 28457-28465.

Feinberg, H., Rambaruth, N. D. S., Jégouzo, S. A. F., Jacobsen, K. M., Djurhuus, R., Poulsen, T. B., Weis, W. I., Taylor, M. E. and Drickamer, K. (2016) Binding sites for acylated trehalose analogs of glycolipid ligands on an extended carbohydrate recognition domain of the macrophage receptor Mincle. *J. Biol. Chem.* **291**, 21222-21233.

Feinberg, H., Jégouzo, S. A. F., Rex, M. J., Drickamer, K., Weis, W. I. and Taylor, M. E. (2017) Mechanism of pathogen recognition by human dectin-2. *J. Biol. Chem.* **292**, 13402-13414.

Feng, J., Garrity, D., Call, M. E., Moffett, H. and Wucherpfennig, K. W. (2005) Convergence on a distinctive assembly mechanism by unrelated families of activating immune receptors. *Immunity* **22**, 427-438.

Fiete, D. J., Beranek, M. C. and Baenziger, J. U. (1998) A cysteine-rich domain of the "mannose" receptor mediates GalNAc-4-SO₄ binding. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2089-2093.

Fletcher, J. M., Boyle, A. L., Bruning, M., Bartlett, G. J., Vincent, T. L., Zaccai, N. R., Armstrong, C. T., Bromley, E. H. C., Booth, P. J., Brady, R. L., Thomson, A. R. and Woolfson, D. N. (2012) A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth. Biol.* **1**, 240-250.

Fornstedt, N. and Porath, J. (1975) Characterization studies on a new lectin found in seeds of *Vicia ervilia*. *FEBS Lett.* **57**, 187-191.

Furie, R., Werth, V. P., Merola, J. F., Stevenson, L., Reynolds, T. L., Naik, H., Wang, W., Christmann, R., Gardet, A., Pellerin, A., Hamann, S., Auluck, P., Barbey, C., Gulati, P., Rabah, D., Franchimont, N. (2019) Monoclonal antibody targeting BDCA2 ameliorates skin lesions in systemic lupus erythematosus. *J. Clin. Invest.* **129**, 1359-1371.

Furukawa, A., Kamishikiryo, J., Mori, D., Toyonaga, K., Okabe, Y., Toji, A., Kanda, R., Miyake, Y., Ose, T., Yamasaki, S. and Maenaka, K. (2013) Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17438-17443.

Garrett, T.P., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H.J., Walker, F., Frenkel, M.J., Hoyne, P.A., Jorissen, R.N., Nice, E.C., Burgess, A.W. and Ward, C.W. (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110**, 763-773.

Garrity, D., Call, M. E., Feng, J. and Wucherpfennig, K. W. (2005) The activating NKG2D receptor assembles in the membrane with two signaling dimers into a hexameric structure. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7641-7646.

Gasparrini, F., Feest, C., Bruckbauer, A., Mattila, P.K., Müller, J., Nitschke, L., Bray, D., and Batista, F.D. (2016) Nanoscale organization and dynamics of the siglec CD22 cooperate with the cytoskeleton in restraining BCR signalling. *EMBO J.* **35**, 258-280.

Gazi, U. and Martinez-Pomares, L. (2009) Influence of the mannose receptor in host immune responses. *Immunobiology* **214**, 554-61.

Gordon, J. (1994) B-cell signalling via the C-type lectins CD23 and CD72. *Immunol. Today* **15**, 411-417.

Geijtenbeek, T. B. H. and Gringhuis, S. I. (2009) Signalling through C-type lectin receptors: shaping immune responses. *Nat. Rev. Immunol.* **9**, 465-479.

Göritzer, K., Goet, I., Duric, S., Maresch, D., Altmann, F., Obinger, C. and Strasser, R. (2020) Efficient N-glycosylation of the heavy chain tailpiece promotes the formation of plant-produced dimeric IgA. *Front. Chem.* **8**, 346.

Graham, S. A., Jégouzo, S. A. F., Yan, S., Powlesland, A. P., Brady, J. P., Taylor, M. E. and Drickamer, K. (2009) Prolectin, a glycan-binding receptor on dividing B cells in germinal centers. *J. Biol. Chem.* **284**, 18537-18544.

Griffiths, J. S., White, P. L., Czubala, M. A., Simonazzi, E., Bruno, M., Thompson, A., Rizkallah, P. J., Gurney, M., da Fonseca, D. M., Naglik, J. R., Ingram, W., Wilson, K., van de Veerdonk, F. L., Barnes, R., Taylor, P. R., and Orr, S. J. (2021) A human Dectin-2 deficiency associated with invasive aspergillosis. *J Infect Dis.* **224**, 1219-1224.

Harbury, P. B., Zhang, T., Kim, P. S. and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401-1407.

Haider, M., Dambuza, I. M., Asamaphan, P., Stappers, M., Reid, D., Yamasaki, S., Brown, G. D., Gow, N. A. R., and Erwig, L. P. (2019) The pattern recognition receptors dectin-2, mincle, and FcR γ impact the dynamics of phagocytosis of *Candida*, *Saccharomyces*, *Malassezia*, and *Mucor* species. *PLoS One*. **14**, e0220867.

Hatada, M. H., Lu, X., Laird, E. R., Green, J., Morgenstern, J. P., Lou, M., Marr, C. S., Phillips, T. B., Ram, M. K., Theriault, K., Zoller M. J. and Karas J. L. (1995) Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* **377**, 32-38.

Holla, A. and Skerra, A. (2011) Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. *Protein Eng. Des. Sel.* **24**, 659-669.

Huang, J., Brameshuber, M., Zeng, X., Xie, J., Li, Q., Chien, Y., Valitutti, S. and Davis, M. (2013) A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4⁺ T cells. *Immunity* **39**, 846-857.

Huber, A., Killy, B., Grummel, N., Bodendorfer, B., Paul, S., Wiesmann, V., Naschberger, E., Zimmer, J., Wirtz, S., Schleicher, U., Vera, J., Ekici, A. B., Dalpke, A., and Lang, R. (2020) Mycobacterial cord factor reprograms the macrophage response to IFN- γ towards enhanced inflammation yet impaired antigen presentation and expression of GBP1. *J Immunol.* **205**, 1580-1592.

Ishikawa, E., Ishikawa, T., Morita, Y. S., Toyonaga, K., Yamada, H., Takeuchi, O., Kinoshita, T., Akira, S., Yoshikai, Y. and Yamasaki, S. (2009) Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J. Exp. Med.* **206**, 2879-2888.

Ishikawa, T., Itoh, F., Yoshida, S., Saijo, S., Matsuzawa, T., Gono, T., Saito, T., Okawa, Y., Shibata, N., Miyamoto, T. and Yamasaki, S. (2013) Identification of distinct ligands for the C-type lectin receptors Mincle and Dectin-2 in the pathogenic fungus *Malassezia*. *Cell Host Microbe* **13**, 477-488.

Jégouzo, S. A. F., Harding, E. C., Acton, O., Rex, M. J., Fadden, A. J., Taylor, M. E., and Drickamer, K. (2014) Defining the conformation of human mincle that interacts with mycobacterial trehalose dimycolate. *Glycobiology* **24**, 1291-1300.

Jégouzo, S. A. F., Feinberg, H., Dungarwalla, T., Drickamer, K., Weis, W. I. and Taylor, M. E. (2015) A novel mechanism for binding of galactose-terminated glycans by the C-type

carbohydrate recognition domain in blood dendritic cell antigen 2. *J. Biol. Chem.* **290**, 16759-16771.

Jellusova, J. and Nitschke, L. (2010) Regulation of B cell functions by the sialic acid-binding receptors Siglec-G and CD22. *Front. Immunol.* **2**, 96.

Jones, J. and Thornton, J. M. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13-20.

Kerrigan, A. M. and Brown, G. D. (2010) Syk-coupled C-type lectin receptors that mediate cellular activation via single tyrosine based activation motifs. *Immunol. Rev.* **234**, 335-352.

Kijimoto-Ochiai, S. (2002) CD23 (the low-affinity IgE receptor) as a C-type lectin: a multidomain and multifunctional molecule. *Cell. Mol. Life Sci.* **59**, 648-664.

Kim, J. W. (2018) Signalling initiation by blood dendritic cell antigen 2, a novel immunoglobulin receptor on plasmacytoid dendritic cells. PhD thesis. Imperial College London.

Kim, J. W., Budzak, J., Liu, Y., Jégouzo, S. A., Drickamer K. and Taylor, M. E. (2018) Identification of serum glycoprotein ligands for the immunomodulatory receptor blood dendritic cell antigen 2. *Glycobiology* **28**, 592-600.

Kim, S. T., Takeuchi, K., Sun, Z. J., Touma, M., Castro, C. E., Fahmy, A., Lang, M. J., Wagner, G. and Reinherz, E. L. (2009) The $\alpha\beta$ T-cell receptor is an anisotropic mechanosensor. *J. Biol. Chem.* **284**, 31028-31037.

Kobilka, B. (2013) The structural basis of G-protein-coupled receptor signaling (Nobel Lecture). *Angew. Chem. Int. Ed.* **52**, 6380-6388.

Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lambert, A. A., Barabé, F., Gilbert, C. and Tremblay, A. J. (2011) DCIR-mediated enhancement of HIV-1 infection requires the ITIM-associated signal transduction pathway. *Blood* **117**, 6589-6599.

- Lanier, L. L. (2009) DAP10- and DAP12-associated receptors in innate immunity. *Immunol. Rev.* **227**, 150-160.
- Lemmon, M. A., Schlessinger, J. and Ferguson, K. M. (2014) The EGFR family: not so prototypical receptor tyrosine kinases. *Cold Spring Harb. Perspect. Biol.* **6**, a020768.
- Liew, O. W., Chong, J. P. C., Yandle, T. G. and Brennan, S. O. (2005) Preparation of recombinant thioredoxin fused N-terminal proCNP: analysis of enterokinase cleavage products reveals new enterokinase cleavage sites. *Protein Expr. Purif.* **41**, 332-340.
- Liu, C., Richard, K., Wiggins, M., Zhu, X., Conrad, D. H. and Song, W. (2016) CD23 can negatively regulate B-cell receptor signaling. *Sci. Rep.* **6**, 25629.
- Lobato-Pascual, A., Saether, P. C., Fossum, S., Dissen, E. and Daws, M. R. (2013) Mincle, the receptor for mycobacterial cord factor, forms a functional receptor complex with MCL and FcεRI-γ. *Eur. J. Immunol.* **43**, 3167-3174.
- Lv, L., Wang, C., Li, Z., Cao, J., Zhong, X., Feng, Y., Chen, J., Tang, T., Ni, H., Wu, Q., Wang, B., Lan, H., and Liu, B. (2021) SAP130 released by damaged tubule drives necroinflammation via miRNA-219c/Mincle signaling in acute kidney injury. *Cell Death Dis.* **12**, 866.
- Macauley, M.S., Crocker, P.R. and Paulson, J.C. (2014) Siglec-mediated regulation of immune cell function in disease. *Nat. Rev. Immunol.* **14**, 653-666.
- Malamud, M., Carasi, P., Assandri, M. H., Freire, T., Lepenies, B., Serradell, M. de LA. (2019) S-layer glycoprotein from *Lactobacillus kefir* exerts its immunostimulatory activity through glycan recognition by Mincle. *Front. Immunol.* **10**, 1422.
- Mariuzza, R.A., Agnihotri, P. and Orban, J. (2019) The structural basis of T cell receptor (TCR) activation: an enduring enigma. *J. Biol. Chem.* **295**, 914-925.
- Matsumoto, M., Tanaka, T., Kaisho, T., Sanjo, H., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Akira, S. (1999) A novel LPS-inducible C-type lectin is a transcriptional target of NF-IL6 in macrophages. *J. Immunol.* **163**, 5039-5048.
- Matsushita, M., Endo, Y. and Fujita, T. (2000) Cutting edge: complement-activating complex of ficolin and mannose-binding lectin-associated serine protease. *J. Immunol.* **164**, 2281-2284.

Meyer, S.J., Linder, A.T., Brandl, C. and Nitschke, L. (2018) B cell Siglecs—news on signaling and its interplay with ligand binding. *Front. Immunol.* **9**, 2820.

Mitchell, D. A., Fadden, A. J. and Drickamer, K. (2001) A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR: subunit organization and binding to multivalent ligands. *J. Biol. Chem.* **276**, 28939-28945.

Mori, D., Shibata, K., and Yamasaki, S. (2017) C-Type lectin receptor Dectin-2 binds to an endogenous protein β -glucuronidase on dendritic cells. *PLoS One*. **12**, e0169562.

Ng, K. K-S., Kolatkar, A. R., Park-Snyder, S., Feinberg, H., Clark, D. A., Drickamer, K. and Weis, W. I. (2002) Orientation of bound ligands in mannose-binding proteins. Implications for multivalent ligand recognition. *J. Biol. Chem.* **277**, 16988-16095.

Ng, W. C., Londrigan, S. L., Nasr, N., Cunningham, A. L., Turville, S., Brooks, A. G. and Reading, P. C. (2015) The C-type lectin langerin functions as a receptor for attachment and infectious entry of influenza A virus. *J. Virol.* **90**, 206-221.

Prado Acosta, M., Goyette-Desjardins, G., Scheffel, J., Dudeck, A., Ruland, J., and Lepenies, B. (2021) S-Layer from *Lactobacillus brevis* modulates antigen-presenting cell functions via the Mincle-Syk-Card9 axis. *Front Immunol.* **12**, 602067.

Pellerin, A., Otero, K., Czerkowicz, J. M., Kerns, H. M., Shapiro, R. I., Ranger, A. M., Otipoby, K. L., Taylor, F. R., Cameron, T. O., Viney, J. L. and Rabah, D. (2015) Anti-BDCA2 monoclonal antibody inhibits plasmacytoid dendritic cell activation through Fc-dependent and Fc-independent mechanisms. *EMBO Mol. Med.* **7**, 464-476.

Rambaruth, N. D., Jégouzo, S. A. F., Marlor, H., Taylor, M. E. and Drickamer, K. (2015) Mouse Mincle: characterization as a model for human mincle and evolutionary implications. *Molecules* **20**, 6670–6682.

Rasmussen, S., Choi, H., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., Schertler, G.F.X., Weis, W.I., and Kobilka, B.K. (2007) Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383-387.

Rasmussen, S. G., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., and Kobilka, B. K. (2011a) Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**, 175-180.

Rasmussen, S. G., DeVree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skinotitis, G., Weis, W. I., Sunahara, R. K., and Kobilka, B. K. (2011b) Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-555.

Reijneveld, J. F., Holzheimer, M., Young, D. C., Lopez, K., Suliman, S., Jimenez, J., Calderon, R., Lecca, L., Murray, M. B., Ishikawa, E., Yamasaki, S., Minnaard, A. J., Moody D. B., and Van Rhijn, I. (2021) Synthetic mycobacterial diacyl trehaloses reveal differential recognition by human T cell receptors and the C-type lectin Mincle. *Sci Rep.* **11**, 2010.

Reizis, B., Bunin, A., Ghosh, H. S., Lewis, K. L. and Sisirak, V. (2011) Plasmacytoid dendritic cells: recent progress and open questions. *Annu. Rev. Immunol.* **29**, 163-183.

Robinson, M. J., Osorio, F., Rosas, M., Freitas, R. P., Schweighoffer, E., Gross, O., Verbeek, J. S., Ruland, J., Tybulewicz, V., Brown, G. D., Moita, L. F., Taylor, P. R. and Reis e Sousa, C. (2009) Dectin-2 is a Syk-coupled pattern recognition receptor crucial for Th17 responses to fungal infection. *J. Exp. Med.* **206**, 2037-2051.

Rogers, N. C., Slack, E. C., Edwards, A. D., Nolte, M. A., Schulz, O., Schweighoffer, E., Williams, D. L., Gordon, S., Tybulewicz, V. L., Brown, G. D. and Reis e Sousa, C. (2005) Syk-dependent cytokine induction by dectin-1 reveals a novel pattern recognition pathway for C type lectins. *Immunity* **22**, 507-517.

Rutledge T., Cosson P., Manolios N., Bonifacino J. S. and Klausner R. D. (1992) Transmembrane helical interactions: zeta chain dimerization and functional association with the T cell antigen receptor. *EMBO J.* **11**, 3245-3254.

Saijo, S., Ikeda, S., Yamabe, K., Kakuta, S., Ishigame, H., Akitsu, A., Fujikado, N., Kusaka, T., Kubo, S., Chung, S., Komatsu, R., Miura, N., Adachi, Y., Ohno, N., Shibuya, K., Yamamoto, N., Kawakami, K., Yamasaki, S., Saito, T., Akira, S. and Iwakura, Y. (2010) Dectin-2

recognition of alpha-mannans and induction of Th17 cell differentiation is essential for host defense against *Candida albicans*. *Immunity* **32**, 681-691.

Samelson, L.E. (2002) Signal transduction mediated by the T cell antigen receptor: the role of adapter proteins. *Annu. Rev. Immunol.* **20**, 371-394.

Sato, Y., Sato, K., Yamamoto, H., Kasamatsu, J., Miyasaka, T., Tanno, D., Miyahara, A., Kagesawa, T., Oniyama, A., Kawamura, K., Yokoyama, R., Kitai, Y., Umeki, A., Ishizuka, S., Takano, K., Shiroma, R., Nakahata, N., Kawakami, K., Kanno, E., Tanno, H., Yamasaki, S., Hara, H., Ishii, K., and Kawakami, K. (2020) Limited role of Mincle in the host defense against infection with *Cryptococcus deneoformans*. *Infect Immun.* **88**, e00400-20.

Schick, J., Schäfer, J., Alexander, C., Dichtl, S., Murray, P. J., Christensen, D., Sorg, U., Pfeffer, K., Schleicher, U., and Lang, R. (2020) Cutting edge: TNF is essential for mycobacteria-induced MINCLE expression, macrophage activation, and Th17 adjuvanticity. *J Immunol.* **205**, 323-328.

Skořepa, O., Pazicky, S., Kalousková, B., Bláha, J., Abreu, C., Ječmen, T., Rosůlek, M., Fish, A., Sedivy, A., Harlos, K., Dohnálek, J., Skálová, T. and Vaněk, O. (2020) Natural killer cell activation receptor NKp30 oligomerization depends on its N-glycosylation. *Cancers* **12**, 1998.

Steck, T. L. (1972) Cross-linking the major proteins of the isolated erythrocyte membrane. *J. Mol. Biol.* **66**, 295-305.

Swiecki, M. and Colonna, M. (2015) The multifaceted biology of plasmacytoid dendritic cells. *Nat. Rev. Immunol.* **15**, 471-485.

Tanaka, M., Ikeda, K., Suganami, T., Komiya, C., Ochi, K., Shirakawa, I., Hamaguchi, M., Nishimura, S., Manabe, I., Matsuda, T., Kimura, K., Inoue, H., Inagaki, Y., Aoe, S., Yamasaki, S., and Ogawa, Y. (2014) Macrophage-inducible C-type lectin underlies obesity-induced adipose tissue fibrosis. *Nat. Commun.* **5**, 4982.

Tanaka, M., Saka-Tanaka, M., Ochi, K., Fujieda, K., Sugiura, Y., Miyamoto, T., Kohda, H., Ito, A., Miyazawa, T., Matsumoto, A., Aoe, S., Miyamoto, Y., Tsuboi, N., Maruyama, S., Suematsu, M., Yamasaki, S., Ogawa, Y., Suganami, T. (2020) C-type lectin Mincle mediates cell death-triggered inflammation in acute kidney injury. *J. Exp. Med.* **217**, e20192230.

- Tanaka, T., Akira, S., Yoshida, K., Umemoto, M., Yoneda, Y., Shirafuji, N., Fujiwara, H., Suematsu, S., Yoshida, N. and Kishimoto, T. (1995) Targeted disruption of the NF-IL6 gene discloses its essential role in bacteria killing and tumor cytotoxicity by macrophages. *Cell* **80**, 353-361.
- Thiemann, S. and Baum, L. G. (2016) Galectins and immune responses: just how do they do those things they do? *Annu. Rev. Immunol.* **34**, 243-264.
- Thompson, A., da Fonseca, D. M., Walker, L., Griffiths, J. S., Taylor, P. R., Gow, N. A. R., and Orr, S. J. (2021) Dependence on Mincle and Dectin-2 varies with multiple *Candida* species during systemic infection. *Front Microbiol.* **12**, 633229.
- Traub, L. M. (2003). Sorting it out: AP-2 and alternate clathrin adaptors in endocytic cargo selection. *J. Cell. Biol.* **163**, 203-208.
- van der Vlist, M. and Geijtenbeek, T. B. H. (2010) Langerin functions as an antiviral receptor on Langerhans cells. *Immunol. Cell. Biol.* **88**, 410-415.
- Wang, H., Lee, T., Fites, S. J., Merkhofer, R., Zarnowski, R., Brandhorst, T., Galles, K., Bruce Klein, B., and Wüthrich, M. (2017) Ligation of Dectin-2 with a novel microbial ligand promotes adjuvant activity for vaccination. *PLoS Pathog.* **13**, e1006568.
- Weis, W. I., Drickamer, K. and Hendrickson, W. A. (1992) Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127-134.
- Weis, W. I., Taylor, M. E. and Drickamer, K. (1998) The C-type lectin superfamily in the immune system. *Immunol. Rev.* **163**, 19-34.
- Woolfson, D. N., Bartlett, G. J., Bruning, M. and Thomson, A. R. (2012) New currency for old rope: from coiled-coil assemblies to α -helical barrels. *Curr. Opin. Struct. Biol.* **22**, 1–10.
- Xu, C., Gagnon, E., Call, M. E., Schnell, J. R., Schwieters, C. D., Carman, C. V., Chou, J. J. and Wucherpfennig, K. W. (2008) Regulation of T cell receptor activation by dynamic membrane binding of the CD3 ϵ cytoplasmic tyrosine-based motif. *Cell* **135**, 702-713.
- Yamasaki, S., Ishikawa, E., Sakuma, M., Hara, H., Ogata, K. and Saito, T. (2008) Mincle is an ITAM-coupled activating receptor that senses damaged cells. *Nat. Immunol.* **9**, 1179-1188.

Yang, W. H., Aziz, P. V., Heithoff, D. M., Mahan, M. J., Smith, J. W. and Marth, J. D. (2015) An intrinsic mechanism of secreted protein aging and turnover. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 13657-13662.

Yokota, A., Yukawa, K., Yamamoto, A., Sugiyama, K., Suemura, M., Tashiro, Y., Kishimoto, T. and Kikutani, H. (1992) Two forms of the low-affinity Fc receptor for IgE differentially mediate endocytosis and phagocytosis: identification of the critical cytoplasmic domains. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5030-5034.

Yonekawa, A., Saijo, S., Hoshino, Y., Miyake, Y., Ishikawa, E., Suzukawa, M., Inoue, H., Tanaka, M., Yoneyama, M., Oh-Hora, M., Akashi, K., and Yamasaki, S. (2014) Dectin-2 is a direct receptor for mannose-capped lipoarabinomannan of mycobacteria. *Immunity*. **41**, 402-413.

Yoshikawa, M., Yamada, S., Sugamata, M., Kanauchi, O., and Morita, Y. (2021) Dectin-2 mediates phagocytosis of *Lactobacillus paracasei* KW3110 and IL-10 production by macrophages. *Sci Rep.* **11**, 17737.

Zarbock, A., Ley, K., McEver, R. P. and Hidalgo, A. (2011) Leukocyte ligands for endothelial selectins: specialized glycoconjugates that mediate rolling and signaling under flow. *Blood* **118**, 6743-6751.

Zhang, H., Cordoba, S., Dushek, O. and van der Merwe, P. A. (2011) Basic residues in the T-cell receptor ζ cytoplasmic domain mediate membrane association and modulate signaling. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19323-19328.

Zuo, J., Mohammed, F. and Moss, P. (2018) The biological influence and clinical relevance of polymorphism within the NKG2D ligands. *Front. Immunol.* **9**, 1820.