Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL

Atsushi Furukawa[1, 2, †], Jun Kamishikiryo[3, †], Daiki Mori[4, †], Kenji Toyonaga[4], Yuki Okabe[1, ‡], Aya Toji[1], Ryo Kanda[1], Yasunobu Miyake[4], Toyoyuki Ose[1], Sho Yamasaki[4, *] and Katsumi Maenaka[1,2, *]

[1]Laboratory of Biomolecular Science, Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan, [2]CREST, Japan Science and Technology Agency, Saitama, Japan, [3]Faculty of Pharmaceutical Sciences, Fukuyama University, Fukuyama, Japan, [4]Division of Molecular Immunology, Research Center for Infectious Diseases, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan

*Correspondence:

Katsumi Maenaka (maenaka@pharm.hokudai.ac.jp)

Laboratory of Biomolecular Science, Faculty of Pharmaceutical Sciences, Hokkaido University, Kita-12, Nishi-6, Kita-ku, Sapporo 060-0812, Japan

Sho Yamasaki (yamasaki@bioreg.kyushu-u.ac.jp)

Division of Molecular Immunology, Research Center for Infectious Diseases, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan

[†] These authors equally contributed to this work.

[‡] Present address: Department of Functional Biological Chemistry, Division of Science, Fukuoka University, Fukuoka, Japan

Running title: Structures and functions of C-type lectin receptors, Mincle and MCL.

Key words: C-type lectin receptor, cell surface receptors, X-ray crystallography, mycobacteria, innate immunity, glycolipid, adjuvant

Abbreviations: PRRs, Pattern recognition receptors, CLRs, C-type lectin receptors, TLRs, toll-like receptors, Mincle, Macrophage inducible C-type lectin, MCL, Macrophage C-type lectin, LPS, lipopolysaccharide, TDM, trehalose-6,6'-dimycolate, CARD9, caspase recruitment domain family member 9.

Abstract

Mincle (Macrophage inducible C-type lectin, CLEC4E) and MCL (Macrophage C-type lectin, CLEC4D) are receptors for cord factor, TDM (Trehalose-6,6'-dimycolate), a unique glycolipid of *Mycobacterium tuberculosis* cell surface components, and activate immune cells to confer adjuvant activity. Although the receptor-TDM interactions require both sugar and lipid moieties of TDM, the mechanisms of glycolipid recognition by Mincle and MCL remained unclear. We here report the crystal structures of Mincle, MCL and Mincle-citric-acid complex. The structures revealed that these receptors are capable of interacting with sugar in a Ca^{2+}-dependent manner, as observed in other C-type lectins. However, Mincle and MCL uniquely possess shallow hydrophobic regions found adjacent to their putative sugar-binding sites, which reasonably locate for recognition to fatty acid moieties of glycolipids. Functional studies using mutant receptors as well as glycolipids ligands support this deduced binding mode. These results give insight on the molecular mechanism of glycolipid recognition through C-type lectin receptors, which may provide clues to rational design for effective adjuvants.

Significance Statement

Here we report the crystal structures of human C-type lectin receptors, Mincle (Macrophage inducible C-type lectin, CLEC4E) and MCL (Macrophage C-type lectin, CLEC4D), both of which are receptor for mycobacterial glycolipid adjuvant, cord factor (also called trehalose-6,6'-dimycolate; TDM). Our structural and functional studies clearly revealed the simultaneous recognition of sugar and lipid moieties by these C-type lectin receptors on myeloid cells, distinct from other C-type lectin receptors. Since better adjuvants are desired for enhancing vaccination effects on the medical treatments for infectious diseases, cancer, and etc, the structures provide a framework for rational design of more effective adjuvants than TDM.

/body **Introduction**

Pattern recognition receptors (PRRs) play important roles in innate immunity. PRRs recognize nucleotides, sugars, lipopolysaccharides (LPS), other pathogen components and self-ligands, and consequently trigger intracellular signaling cascades that initiate innate and adaptive immune responses (1). Among them, toll-like receptors (TLRs) are well-characterized receptors, in terms of their ligand specificities, ligand recognition mechanisms and signaling pathways (2-4). The C-type lectin receptors (CLRs) are also a large family of PRRs (5-7). The term 'C-type lectin' was introduced to distinguish a group of Ca^{2+}-dependent lectins from other lectins. In the CLRs, two amino acids harboring long carbonyl side chains separated by a proline in a *cis* conformation coordinate a Ca^{2+} ion, which forms hydrogen bonds with monosaccharides and determines the binding specificity. The CLRs have broad recognition abilities towards not only saccharides but also proteins (5, 7-9). For instance, human NKR-P1interacts with Lectin-like transcript-1, and some members of the CD94/NKG2 family interact with HLA-E.

Macrophage inducible C-type lectin (Mincle, also called CLEC4E) is a type II transmembrane C-type lectin receptor that is expressed in macrophages, dendritic cells and monocytes upon stimulation (10). We have reported that Mincle is an FcRγ-coupled activating receptor that recognizes pathogenic fungus and mycobacteria (11-13). Detailed investigations of the ligands of Mincle revealed that Mincle binds glycolipids, such as trehalose-6,6'-dimycolate (TDM) from *M. tuberculosis*, and novel glyceroglycolipids from Malassezia fungus. The Malassezia and *M. tuberculosis* ligands are recognized through the carbohydrate recognition domain (CRD) in the extracellular

region of Mincle (11, 12). The binding of TDM to Mincle leads to the phosphorylation of the ITAM in the FcRγ chain, which provides a binding site for the Syk tyrosine kinase. Syk activates the caspase recruitment domain family member 9 (CARD9)-mediated NF-kB signaling pathway, to promote the expression of TNF and IL-6. A recent report revealed that Mincle plays a nonredundant role in T cell immune responses to infection by microbes and in the adjuvanticity of mycobacterial cord factor and its synthetic analog, trehalose-dibehenate (TDB) (14, 15).

MCL (also called Clec4D) is another C-type lectin receptor expressed in myeloid cells (16, 17). Recently, we found that MCL is also an FcRγ-coupled activating receptor that binds to TDM (15). MCL is distinct from Mincle, in the following manners: 1) The expression of Mincle is inducible, whereas MCL is constitutively expressed in myeloid cells. 2) MCL shows weaker binding affinity to TDM than that of Mincle. 3) EPN motif, a typical glucose/mannose binding motif, is conserved in Mincle but not in MCL.

We now report the crystal structures of Mincle and MCL, as well as Mincle complexed with citric acid. They have similar overall structures to other typical CLRs, but exhibit characteristic conformations in the vicinity of the Ca^{2+} binding motif. A patch of hydrophobic amino acids located adjacent to the carbohydrate binding site may likely contribute to the recognition of the fatty acid chain of TDM. The mutational analysis essentially supports this TDM binding model, and may also explain the different affinities of MCL and Mincle.

Results

Preparation, crystallization and structural determination of MCL

The extracellular domain of human MCL (residues 61-215, Fig. 1) was expressed in *E. coli* as inclusion bodies, and was refolded *in vitro* by a dilution method. Ca^{2+} ions were required in the refolding procedure, and the crude, refolded MCL was purified by sequential gel filtration chromatography steps (Fig. S1A and C). The purified MCL was crystallized by the hanging drop method with 0.1 M Bis-Tris propane, pH 6.5, 0.2 M potassium thiocyanate and 20% (w/w) PEG 3350 (PACT, 64). Crystals of the MCL protein (the space group was I centered orthorhombic (*I*222), and the unit cell parameters were a = 85.19 Å, b = 96.06 Å, c = 104.53 Å) were obtained, and the data set was collected to the resolution limit of 2.2 Å at the BL32XU beam line at SPring-8 (Table S1). The crystal structure of MCL has two α-helices (α1 and α2) and eleven β -strands (β1 to β11) (Fig. 2A), which is typical structural organization of CLRs and partly similar to the solution structure of MCL registered in the Protein Databank (PDB: 2LS8) (Fig. S1E). Two MCL molecules exist in the asymmetric unit. The gel filtration analysis showed the mixture of the peaks (Fig. S1A), suggesting that MCL may have some conformational variation.

Preparation, crystallization and structural determination of Mincle

Using a similar refolding method to that for MCL, we also prepared the extracellular domain of Mincle (residues 74-219) (Fig. 1). The expression, refolding and purification were successful. However, the crystallization was not successful, because the refolded Mincle was not sufficiently soluble at high concentrations. To improve the protein solubility, we performed site-directed mutagenesis, and changed the hydrophobic amino

134 acids presumably located on the surface of the Mincle protein to hydrophilic amino
135 acids, as found in the corresponding residues of MCL. Among them, the mutant with
136 the substitution of isoleucine to lysine at residue 99 (I99K mutant) formed good crystals
137 by the hanging drop method with two conditions. One is 1 M lithium chloride, 0.1 M
138 citric acid (pH 4), and 20% (w/v) PEG6000, and the other is 0.2M NH_4SO_4, 0.1M
139 Bis-Tris (pH 5.5), 25% (w/v) PEG3350. These diffraction data were collected to the
140 resolution limit of 1.3 Å and 1.35 Å at the BL5A and BL17A beam lines at KEK
141 (Tsukuba, Japan), respectively. Both crystals have the same space group, primitive
142 trigonal ($P3_1$), and the unit similar cell parameters (Table S1). Mincle exhibits typical
143 CLR fold, as shown in Fig. 2B. The asymmetric unit contained one molecule of Mincle
144 and no physiologically important packing was detected. This is consistent with the gel
145 filtration analysis showing that Mincle behaves as a monomer, although the eluted time
146 is later than the expected one likely due to affinity of Mincle to the glucose-based
147 dextran resin of Supderdex column (Fig. S1B and D).
148
149 **Structural comparison between Mincle, MCL and other C-type lectins**
150 MCL and Mincle superimposed well on each other (root mean square deviation (r. m. s.
151 d.) 1.5 Å for 124 Cα atoms) (Fig. 2A and B, and Fig. S2B). However, Mincle has two
152 calcium ions, while MCL has only one. A DALI analysis indicated that MCL and
153 Mincle share high homology with mouse collectin (2OX9) (r. m. s. d. of 1.3 Å for 120
154 Cα atoms, 28% identity (MCL), and 1.07 Å for 118 Cα atoms, 35% identity (Mincle))
155 (18) and DC-SIGNR (1K9J) (r. m. s. d. of 1.7 Å for 114 Cα atoms, 37% identity (MCL)
156 and 1.3 Å for 121 Cα atoms, 45% identity (Mincle)), which has been extensively
157 studied as an entry receptor of HIV (19, 20). Since collectin recognizes fucose-based

158 oligosaccharides, rather than glucose- or mannose-based ones, we chose to compare the
159 structural features of MCL, Mincle, and DC-SIGNR (Fig. 2C and D). The entire
160 structures and the positions of the amino acid residues in the putative CRD are similar.
161 Specifically, the positions of the Ca^{2+} ions (site 1) are the same among the three proteins.
162 The glutamic-acid-proline-asparagine (EPN) motif (residues 169–171 in Mincle) is
163 often observed in C-type lectins, and contributes to carbohydrate recognition via a Ca^{2+}
164 ion-mediated binding network (Fig. 3A and B). In contrast, the EPD motif of MCL
165 (residues 173–175) is an unusual sequence among the C-type lectins (6) (Fig. 1).
166 However, the Ca^{2+} ion and other amino acids involved in carbohydrate recognition are
167 located in this region, as in other C-type lectins (Fig. 2 and 3A-C). These results
168 indicated that Mincle and MCL recognize carbohydrates through these motifs in slightly
169 different, but similar manners.
170 The regions surrounding the Ca^{2+}-bound sites in MCL and Mincle are distinctly
171 different from that in DC-SIGNR. In DC-SIGNR, two additional bound Ca^{2+} ions are
172 observed close to the site (red, **2** and **3**) (Fig. 2C), and they stabilize the typical protein
173 conformation of the C-type lectins (6). The Ca (**2** and **3**) ions push the loop (residues
174 312-317) close to the Ca (**1**) ion (Fig. 2D, red dotted circle). In contrast, the
175 corresponding loops in MCL and Mincle are located far from the Ca (**1**) ion. The
176 asparagine/aspartate residues just after the EPD/EPN sequences are conserved (Fig. 1).
177 The directions of the asparagines in MCL (residue 176) and Mincle (residue 172) are
178 different from those in other CLRs, such as DC-SINGR (Fig. 3A-C). The asparagine in
179 DC-SIGNR is used to bind the Ca (**2**) ion, and therefore the side chain faced to the
180 opposite direction of the Ca (**1**) ion. In contrast, neither Mincle nor MCL coordinates Ca
181 (**2** and **3**) ions, and their asparagine side chains extend in different directions, as

compared to other C-type lectins.

Calcium binding and ligand recognition

In the crystals of Mincle grown in 1 M lithium chloride, 0.1 M citric acid (pH 4), and 20% (w/v) PEG6000, a strong electron density in addition to that of the Mincle protein was observed close to the Ca(**1**) ion and matched a citric acid molecule (Fig. 3D). The superimposition of the amino acids in the Ca^{2+} ion binding regions of the ligand complex structures of Mincle and DC-SIGNR revealed the well conserved locations of the oxygen atoms of the ligands, citrate and mannose (equatorial 3- and 4-OH groups), respectively (Fig. 3E). Because the chemical property of the sugar moiety is different from the citric acid, we cannot simply compare the recognition modes but these data may support the idea that Mincle can utilize this Ca ion to bind nucleophiles for the sugar moieties of TDM and Malassezia ligands, in essentially the same manner as generally observed in CLRs including DC-SIGNR.

The calcium binding site in the human and mouse Mincles includes the EPN motif, a well conserved in the mannose-recognizing C-type lectins, as described above. We examined whether the EPN motif in Mincle is involved in direct TDM recognition using soluble Mincle protein (Mincle-Ig). Mincle-Ig ($Mincle^{WT}$), but not control Ig, selectively bound to plate-coated TDM, as previously reported (11, 12). This recognition was shown to require the EPN motif, as the binding was eliminated by introducing a mutation of EPN into QPD, a putative galactose-recognition sequence ($Mincle^{QPD}$) (21). Substitution of EPN motif into MCL-type EPD ($Mincle^{EPD}$) also impaired the binding capacity, although their reactivities to anti-hIgG were comparable (Fig. 3F and Fig. S3A). This data suggested that EPN in human Mincle is indispensable

for TDM recognition, as previously shown in mouse Mincle (11, 12). In contrast, the direct binding of MCL to TDM was much weaker than that of Mincle (Fig. 3G), consistent with the previous report that MCL recognizes TDM with less affinity than Mincle (15). Mutation of EPD sequence of into QPD (MCL^{QPD}) did not have large impact on the TDM binding in higher concentrations. Unexpectedly, however, the EPD to EPN mutation in MCL, which was expected to coordinate the Ca ion location well and facilitate carbohydrate binding, did not improve the affinity for TDM (Fig. 3G). These results suggested that the TDM binding site of MCL might be distinct from that used by Mincle. Furthermore, the side chain of Arg183 in Mincle is in a suitable position to interact with the hydroxyl groups of TDM, based on the crystal structure of Mincle (22) (Fig. 4A). This arginine residue of Mincle is well conserved from fishes to mammals. In contrast, the valine (Val186) at the corresponding position of human MCL is conserved among placentalia, however, its side chain cannot reach the putative carbohydrate recognition site (Fig. 4B). To verify the role of Arg183 in TDM recognition, we introduced the R183V mutation in Mincle and tested its function in an NFAT-GFP reporter assay. This mutation reduced the NFAT-GFP activity in the reporter cell assay, suggesting that Arg183 of Mincle is crucially involved in the ligand recognition (Fig. 4C).

Taken together, these results strongly suggested that the binding mode of the two OH groups of citrate acid to Ca^{2+} reflects the equatorial 3- and 4-OH groups of mannose and glucose of Mincle/MCL ligands, in a similar, but slightly different manner than the CLRs (6).

Putative lipid recognition sites

To determine whether Mincle and MCL utilize unique amino acids for their interactions with the lipid regions of glycolipids, we verified the characteristics of the surfaces surrounding the putative sites for the Ca^{2+}-mediated sugar binding. A series of hydrophobic regions were specifically found in Mincle and MCL, but not in other C-type lectins in the vicinity of the putative sugar-binding sites (dotted circles in Fig. 4A and B, yellow surfaces in Fig. 4D-F, and yellow-shaded amino acid residues in the box enclosed with a blue line in Fig. 1). The regions are composed of Val195, Thr196, Phe198, Leu199, Tyr 201 and Phe 202 in Mincle, and Val197, Pro198 and Phe201 in MCL. The Mincle has larger hydrophobic areas than MCL, while DC-SGNR has only much smaller one than both Mincle and MCL (Fig. 4D-F). If the trehalose part of TDM is placed on the sugar binding site of Mincle, as in the binding mode of mannose to DC-SIGNR, then the mycolic acid attached to the 6-*O* of the glucose of TDM (Fig. 4A and B, red arrow) is oriented toward the hydrophobic regions of Mincle and MCL, as described above. To investigate whether the hydrophobic region of Mincle contributes to the recognition of TDM, the Ala substitutions of both Phe198 and Leu199 in this region was introduced in reporter cells expressing Mincle (Fig. 4C). The cells expressing the Mincle$^{F198A/L199A}$ mutant exhibited reduced NFAT activity in response to TDM. Moreover, we replaced the hydrophobic region of Mincle (residues 195-202) with the corresponding region of another CLR, Dectin-2 (residues 192-199), which lacks the hydrophobic residues (23, 24). The reporter cells expressing this Mincle-Dectin-2 chimeric molecule (Mincle$^{MD\ chimera}$) still retained the activity against anti-Mincle mAb, 13D10-H11 (Fig. S3C), which recognizes the conformational epitope on Mincle (Fig. S3D), indicating that the mutation as well as other mutations of this study did not make remarkable effect on the overall protein folding and stability on the

cell surface. However, the TDM recognition of Mincle$^{MD\ chimera}$ was severely compromised (Fig. 4C). As described above, the Oδ atom of the corresponding residue Asn172 just after the EPN motif in Mincle does not face toward the Ca ion, which is an unusual type of Ca coordination among the C-type lectins (Fig. 3A and B). Instead, the Nδ atom of Asn172 forms a hydrogen bond with the Oδ atom of Thr196 of the hydrophobic patch (Fig. 4A). The reporter cells expressing the mutant Mincle (Mincle^{N172Q}), which has only one additional methylene group, showed reduced NFAT-GFP activity (Fig. 4C). This result may suggest that the N172Q mutation indirectly influences the hydrophobicity of the putative lipid binding patch via the side chain of Thr196.

In order to further examine the effect of a set of acyl chains, we performed surface plasmon resonance binding assays using a set of trehalose-based glycolipids, which have a single acyl chain with different carbon lengths (C8, C10, C12). These glycolipids have a single and short tail and thus are expected to be water-soluble while retaining the ligand activity. The single acyl chains with trehalose (C10 and C12) bound to Mincle (Figs. 4G and S5). The affinity of C8 to Mincle is much lower than those of C10 and C12 (Fig. 4G). The crystal structure clearly indicated that the 10-carbon acyl chain with trehalose is reasonably accommodated within the hydrophobic portion in Mincle (Fig. S6).

Discussion

We have determined the crystal structures of the ectodomains of Mincle and MCL, which confirmed that the overall structures of Mincle and MCL are similar to those of other CLRs. Furthermore, we have also solved the crystal structure of Mincle

278 complexed with citric acid, which revealed that the binding mode to citric acid
279 essentially resembles that of glucose/mannose recognition by typical CLRs. We further
280 performed the competition binding of glycolipids with citric acid as Fig. S7, clearly
281 showing that the citric acid inhibits the glycolipid binding to Mincle, while the acetic
282 acid does not. Notably, other mannose-binding c-type lectin, codakine, bound the
283 similar positions of oxygens of glycerol and glycan in Ca^{2+} ion-mediated manner (25).
284 The citric acid is likely accommodated at this position to block the ligands and hydroxyl
285 groups are likely utilized following the coordination of Ca^{2+} ions generally observed in
286 CLRs.
287 Glycolipids play pivotal roles in innate immunity, as exemplified by the functions of
288 CD1-mediated natural killer T cells (NKT) (26, 27). The CD1 family molecules display
289 a variety of glycolipids toward semi-invariant NKT cell receptors to activate NKT cells.
290 The structural analyses of CD1 family proteins have revealed that the lipid parts of the
291 glycolipids are deeply accommodated inside the hydrophobic cores of the proteins (28,
292 29), and thus only the sugar moieties are exposed, for recognition by NKT cell receptors
293 (Fig. S6). On the other hand, our present study showed that the putative TDM binding
294 sites of Mincle and MCL include hydrophobic loops uniquely found in Mincle and
295 MCL, which distinguish them from the other C-type lectins (Fig. 4A, B and D-F). These
296 loops form shallow hydrophobic patches extending from the corresponding position of
297 the 6-OH of glucose on the structure of the mannose complex of DC-SIGNR, which is
298 attached to mycolic acid, in the case of TDM (Fig. S6). The mutational study suggested
299 that these CLRs directly recognize the acyl groups of the glycolipid TDM using this
300 shallow hydrophobic region, which is close to the Ca^{2+} binding site (Fig. 4D and E).
301 Notably, the SPR binding study using a set of glycolipids clearly showed that the single

302 acyl chain is sufficient for Mincle binding. In addition, importantly, at least a C10
303 length of the lipid moiety is required (Fig. 4G). These observations might suggest that
304 Mincle recognizes only the sugar-proximal part of the acyl chain of glycolipids. The
305 hydrophobic patch branches out from the potential sugar binding site (downward and to
306 right in Fig. 4D and Fig. S6). These might confer the sites accommodating the branched,
307 acyl chains in mycolic acids, such as TDM and trehalose mono-mycolate (TMM) (11).
308 The recently discovered ligands of Mincle (44-1 and 44-2), which also have branched
309 acyl chains, may interact similarly with TMM and TDM (13). Therefore, the recognition
310 of glycolipids by Mincle and presumably MCL seems to be significantly distinct from
311 those of lipid recognition proteins, such as CD1 and Toll-like receptor 4/MD2 complex,
312 which have deep hydrophobic grooves to accommodate the acyl moieties of glycolipids
313 (Fig. S6). Thus, a minimum acyl chain length is required for glycolipid recognition by
314 CLRs. The unique modes of CLR-glycolipid recognition would be advantageous for
315 host defense responses, because they may allow receptors to recognize these bipolar
316 ligands even within a microbial cell wall or in the micellar form in aqueous solution.
317 Future study for the co-crystallization with glycolipids harboring short branched acyl
318 chains, which might have increased binding affinity, would elucidate the lipid binding
319 modes.
320 The production of NO and IL6 by Bone Marrow derived Macrophages (BMMφ),
321 which express Mincle and MCL, was reportedly changed by stimulation with several
322 lengths of acyl chains, revealing the importance of the acyl chain length (30). The
323 fungal glycolipids, recently identified as Mincle ligands, have more complicated and
324 branched lipid moieties. The structural and functional data presented here showed that
325 Mincle and MCL probably require an acyl chain longer than 10 carbons for glycolipid

326 recognition, thus clearly providing important clues for the design of better adjuvants
327 than TDM.
328 The present study indicated that Mincle has a higher affinity for TDM than MCL,
329 which is consistent with our *in vitro* binding study (Fig. 3F and G). The crystal
330 structures of MCL and Mincle clearly revealed that MCL has a smaller hydrophobic
331 patch next to the putative Ca^{2+}-mediated sugar binding site, as compared to that in
332 Mincle. The different sizes of these hydrophobic sites might explain the affinity
333 differences of two CLRs observed in the binding data.
334 The typical CLRs that simply recognize sugars, such as DC-SIGNR and CEL-IV,
335 exhibit remarkably low affinities (Kd ~mM) (31, 32). They require multiple valencies of
336 sugar ligands to mediate signaling. However, the SPR analysis revealed that Mincle
337 seemed to show higher affinity, suggesting that it can detect small numbers of
338 glycolipids on fungal surfaces. On the other hand, MCL showed much lower affinity
339 than Mincle, but essentially the same ligand specificity. It is plausible that
340 MCL-mediated signaling requires multiple valencies of glycolipid ligands. Therefore,
341 Mincle and MCL may play distinct roles in physiological events.
342 CLRs often form homodimers or heterodimers on the cell surface. As described above,
343 the multivalent ligands on the bacterial surface likely induce the multimerization of the
344 CLRs (either monomeric or dimeric structures), which may mediate efficient signaling.
345 The recent report by Lobato-Pascual et al. demonstrated that Mincle and MCL form the
346 disulfide-linked heterodimer associated with the FcεRIγ chain (33). The heterodimeric
347 complex formation between Mincle and MCL through the N-terminal β strand and/or
348 stalk regions, as previously reported for maltose-binding protein (34), for efficient
349 recognition/signaling would be an intriguing issue to be addressed.

Materials and Methods

Plasmid construction

The *E. coli* expression plasmids encoding the partial extracellular domain of human Mincle (residues 74-219), pET22-Mincle, and the extracellular domain (residues 61–215) of human MCL, pET22-MCL, were constructed (the detail in Supplemental information) (35).

In order to improve the solubility and crystallization of Mincle, we synthesized, purified and crystallized several mutated Mincle proteins. Among them, the I99K mutant was produced with a high yield and generated good crystals.

Preparation of recombinant proteins

The pET22-Mincle, pET22-Mincle I99K and pET22-MCL plasmids were transformed into *Escherichia coli* strain BL21(DE3) plysS, and the protein was obtained as inclusion bodies. The protein was solubilized in a buffer containing 6 M guanidine-HCl, 50 mM MES, pH 6.5, 100 mM NaCl, and 10 mM EDTA for 12 hours at 4°C. 1h after the addition of dithiothreitol (10mM), the solubilized proteins were slowly diluted into 1 liter of buffer, containing 0.1 M Tris-HCl, pH 8.5, 1 M l-arginine, 2 mM EDTA, 6.3 mM cystamine, 3.7 mM cysteamine, and 0.1 mM phenylmethylsulfonyl fluoride. The refolding mixture was purified by gel filtration chromatography. The buffer was finally exchanged to 20 mM Tris-HCl, pH 8.0, with 5 mM $CaCl_2$ for crystallization.

Crystallization and Structure Determination

Crystals of purified Mincle I99K and MCL were grown at 20°C (reservoir solutions: 1 M lithium chloride, 0.1 M citric acid (pH 4), 20% (w/v) PEG6000 and 0.1 M Bis-Tris propane, pH 6.5, 0.2 M potassium thiocyanate, 20% (w/w) PEG 3350, respectively) by the hanging drop vapor-diffusion method. Crystals were equilibrated in a cryo-protectant consisting of reservoir solution supplemented with 16% (v/v) glycerol. X-ray data were collected on beamlines, BL32XU in SPring-8 and BL5C in KEK. The data were processed with HKL2000 (36) or XDS (37). The structure was solved by molecular replacement with PHASER (38), using CD69 as the search model (PDB: 1FM5). Several rounds of model building in COOT (39) and refinement in PHENIX (40) were performed. The final refinement statistics are provided in Table S1. The coordinates for the refined Mincle, Mincle-citrate complex, and MCL structures have been deposited in the Protein Data Bank (accession codes 3WH3, 3WH2 and 3WHD, respectively).

Binding assay using Ig-fusion proteins

The MCL-Ig and Mincle-Ig fusion proteins were prepared as described previously. Briefly, the C terminus of the extracellular domain of human MCL (residues 42-215), human Mincle (residues 46-219), or their mutants was fused to the N terminus of the hIgG1 Fc region. The Ig-fusion proteins were incubated with 0.2 μg/well of plate-coated TDM or plate-coated anti-human IgG, and the bound proteins were detected by using HRP-labeled anti-human IgG.

Reporter Assay

Reporter cells were prepared as described previously (11, 35). Briefly, 2B4-NFAT-GFP reporter cells were transfected with FcRγ, together with Mincle and mutants. The reporter cells were stimulated with various concentrations of plate-coated TDM or anti-human Mincle antibody (13D10H11). The activation of NFAT-GFP was monitored by flow cytometry.

SPR analysis

The SPR analysis was performed similarly as described previously for other cell surface receptors (41). Briefly, Mincle and MCL were each dissolved in 10 mM sodium acetate (pH 4), containing 5 mM $CaCl_2$ with or without 5% dimethyl sulfoxide (DMSO). SPR experiments were performed with a BIAcore T3000 (GE Healthcare). All of the proteins were covalently immobilized on the CM5 sensor chip by amine-coupling (GE Healthcare). β2 microglobuline was used as a negative control protein. All glycolipids (C12, C10 and C8), and trehalose as a negative control, were injected in 10 mM HEPES, pH 7.4, containing 150 mM NaCl and 5 mM $CaCl_2$. The data were analyzed using the BIAevaluation software, version 4.1 (GE Healthcare).

Note Added in Proof. While this paper was under revision, Feinberg *et al.* (42) reported the crystal structure of bovine Mincle complexed with torehalose, whose binding mode is similar to that of human Mincle for glycolipids we proposed here.

Acknowledgements

We thank Dr. S. Kita, M. Nagata and M. Shiokawa for their kind help with the data

collection. We also thank the beamline staff at the Photon Factory (Tsukuba, Japan) and Spring 8 (Harima, Japan) for technical help during the X-ray data collection. This work was supported by grants to K. M. from the Ministry of Education, Culture, Sports, Science and Technology and the Ministry of Health, Labor and Welfare of Japan.

References

1. Geijtenbeek TB & Gringhuis SI (2009) Signalling through C-type lectin receptors: shaping immune responses. *Nat Rev Immunol* 9(7):465-479.
2. Barton GM & Kagan JC (2009) A cell biological view of Toll-like receptor function: regulation through compartmentalization. *Nat Rev Immunol* 9(8):535-542.
3. Choe J, Kelker MS, & Wilson IA (2005) Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* 309(5734):581-585.
4. Akira S & Takeda K (2004) Toll-like receptor signalling. *Nat Rev Immunol* 4(7):499-511.
5. Kuroki K, Furukawa A, & Maenaka K (2012) Molecular recognition of paired receptors in the immune system. *Front Microbiol* 3:429.
6. Zelensky AN & Gready JE (2005) The C-type lectin-like domain superfamily. *FEBS J* 272(24):6179-6217.
7. Sancho D & Reis e Sousa C (2012) Signaling by myeloid C-type lectin receptors in immunity and homeostasis. *Annu Rev Immunol* 30:491-529.
8. Sancho D & Reis e Sousa C (2013) Sensing of cell death by myeloid C-type lectin receptors. *Curr Opin Immunol* 25(1):46-52.
9. Miyake Y, Ishikawa E, Ishikawa T, & Yamasaki S (2010) Self and nonself recognition through C-type lectin receptor, Mincle. *Self Nonself* 1(4):310-313.
10. Matsumoto M*, et al.* (1999) A novel LPS-inducible C-type lectin is a transcriptional target of NF-IL6 in macrophages. *J Immunol* 163(9):5039-5048.
11. Ishikawa E*, et al.* (2009) Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J Exp Med* 206(13):2879-2888.
12. Yamasaki S*, et al.* (2009) C-type lectin Mincle is an activating receptor for pathogenic fungus, Malassezia. *Proc Natl Acad Sci U S A* 106(6):1897-1902.

13. Ishikawa T, *et al.* (2013) Identification of Distinct Ligands for the C-type Lectin Receptors Mincle and Dectin-2 in the Pathogenic Fungus Malassezia. *Cell Host Microbe* 13(4):477-488.
14. Schoenen H, *et al.* (2010) Cutting edge: Mincle is essential for recognition and adjuvanticity of the mycobacterial cord factor and its synthetic analog trehalose-dibehenate. *J Immunol* 184(6):2756-2760.
15. Miyake Y, *et al.* (2013) C-type Lectin MCL Is an FcRγ-Coupled Receptor that Mediates the Adjuvanticity of Mycobacterial Cord Factor. *Immunity*.
16. Arce I, Martínez-Muñoz L, Roda-Navarro P, & Fernández-Ruiz E (2004) The human C-type lectin CLECSF8 is a novel monocyte/macrophage endocytic receptor. *Eur J Immunol* 34(1):210-220.
17. Graham LM, *et al.* (2012) The C-type lectin receptor CLECSF8 (CLEC4D) is expressed by myeloid cells and triggers cellular activation through Syk kinase. *J Biol Chem* 287(31):25964-25974.
18. Feinberg H, Taylor ME, & Weis WI (2007) Scavenger receptor C-type lectin binds to the leukocyte cell surface glycan Lewis(x) by a novel mechanism. *J Biol Chem* 282(23):17250-17258.
19. Feinberg H, Mitchell DA, Drickamer K, & Weis WI (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* 294(5549):2163-2166.
20. Geijtenbeek TB, *et al.* (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100(5):587-597.
21. Drickamer K (1992) Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* 360(6400):183-186.
22. Fujihashi M, Peapus DH, Kamiya N, Nagata Y, & Miki K (2003) Crystal structure of fucose-specific lectin from Aleuria aurantia binding ligands at three of its five sugar recognition sites. *Biochemistry* 42(38):11093-11099.
23. Robinson MJ, *et al.* (2009) Dectin-2 is a Syk-coupled pattern recognition receptor crucial for Th17 responses to fungal infection. *J Exp Med* 206(9):2037-2051.
24. Saijo S, *et al.* (2010) Dectin-2 recognition of alpha-mannans and induction of Th17 cell differentiation is essential for host defense against Candida albicans. *Immunity* 32(5):681-691.
25. Gourdine JP, *et al.* (2008) High affinity interaction between a bivalve C-type lectin and a biantennary complex-type N-glycan revealed by crystallography and microcalorimetry. *J Biol Chem* 283(44):30112-30120.

26. Kawano T, *et al.* (1997) CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides. *Science* 278(5343):1626-1629.
27. Moody DB & Besra GS (2001) Glycolipid targets of CD1-mediated T-cell responses. *Immunology* 104(3):243-251.
28. Gadola SD, *et al.* (2002) Structure of human CD1b with bound ligands at 2.3 A, a maze for alkyl chains. *Nat Immunol* 3(8):721-726.
29. Koch M, *et al.* (2005) The crystal structure of human CD1d with and without alpha-galactosylceramide. *Nat Immunol* 6(8):819-826.
30. Khan AA, *et al.* (2011) Long-chain lipids are required for the innate immune recognition of trehalose diesters by macrophages. *Chembiochem* 12(17):2572-2576.
31. Probert F, Whittaker SB, Crispin M, Mitchell DA, & Dixon AM (2013) Solution NMR analyses of the C-type carbohydrate recognition domain of DC-SIGNR reveal different binding modes for HIV-derived oligosaccharides and smaller glycan fragments. *J Biol Chem*.
32. Hatakeyama T, *et al.* (2011) Galactose recognition by a tetrameric C-type lectin, CEL-IV, containing the EPN carbohydrate recognition motif. *J Biol Chem* 286(12):10305-10315.
33. Lobato-Pascual A, Saether PC, Fossum S, Dissen E, & Daws MR (2013) Mincle, the receptor for mycobacterial cord factor, forms a functional receptor complex with MCL and FcεRIγ. *Eur J Immunol*.
34. Weis WI & Drickamer K (1994) Trimeric structure of a C-type mannose-binding protein. *Structure* 2(12):1227-1240.
35. Yamasaki S, *et al.* (2008) Mincle is an ITAM-coupled activating receptor that senses damaged cells. *Nat Immunol* 9(10):1179-1188.
36. Z. O & W. M (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **V**olume 2**76:** Macromolecular Crystallography, part A:307-326.
37. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.
38. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
39. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
40. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.

523 41. Tabata S, *et al.* (2008) Biophysical characterization of O-glycosylated CD99
524 recognition by paired Ig-like type 2 receptors. *J Biol Chem* 283(14):8893-8901.
525 42. Feinberg H, *et al.* (2013) Mechanism for recognition of an unusual
526 mycobacterial glycolipid by the macrophage receptor mincle. *J Biol Chem*.
527 43. Gouet P, Robert X, & Courcelle E (2003) ESPript/ENDscript: Extracting and
528 rendering sequence and 3D information from atomic structures of proteins.
529 *Nucleic Acids Res* 31(13):3320-3323.
530 44. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics
531 of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad*
532 *Sci U S A* 98(18):10037-10041.
533

534 **Figure legends**

535 **Fig. 1.** Structure-based sequence alignment of Mincle, MCL, DC-SIGNR and Dectin-2.
536 The sequence alignment of the ecto-domains of Mincle, MCL, DC-SIGNR and Dectin-2
537 (h and m indicate that human and mouse, respectively) is shown, as depicted with
538 ESPript (43). Identical residues are highlighted in red, and similar residues are framed
539 in blue. The secondary structure elements (α: α-helix, β: β-strand, T: turn) of Mincle
540 and MCL are shown above the sequences. The box enclosed by the thick black line
541 indicates the EPN motifs, which are usually involved in carbohydrate recognition by
542 C-type lectins. The box enclosed by the blue line indicates the hydrophobic amino acids
543 loops and yellow-shaded amino acids residues are hydrophobic residues within Mincle
544 and MCL. The asterisks below the sequences indicate the residues involved in calcium
545 binding in Mincle and MCL. The red filled circles below the sequences indicate the
546 residues changed to other amino acid residues in the mutational studies. The numbers
547 under the cysteine residues indicate disulfide bond formation with the cysteine residue
548 with the same number.

549

550 **Fig. 2.** Structures of MCL and Mincle, and structural comparison with DC-SIGNR.

(**A** and **B**) Cartoon models of overall structures of MCL (**A**) and Mincle (**B**). The secondary structure elements are shown. Gradient rainbow color from blue to red indicates N- to C-terminal. The yellow and cyan spheres are Ca^{2+} ions in MCL and Mincle, respectively. (**C** and **D**) Overall structures (**C**) and putative ligand binding sites (**D**), close-up view of black box of **C**, of MCL (yellow), Mincle (cyan) and DC-SIGNR (pink) are shown. Yellow, blue and pink spheres are Ca^{2+} ions in Mincle, MCL and DC-SIGNR, respectively. Red dotted circle indicates the large structural difference of loops among these CLRs (see Text).

Fig. 3. Structural comparison of the putative ligand binding sites in MCL, Mincle and DC-SIGNR, and *in vitro* binding assays of Mincle and MCL mutants.

(**A**-**C**) Close-up views of the putative ligand binding sites of MCL (yellow) (**A**), Mincle (cyan) (**B**), and DC-SIGNR (pink) (**C**) are shown. The amino acid residues involved in and close to Ca^{2+} ion binding are shown as stick models. Interactions with Ca^{2+} ions are shown in black dotted lines. (**D**) Composite OMIT map (2Fo-Fc) for citric acid in Mincle. The electron-density map is contoured at 1.0σ, and the resolution is 1.3 Å. The citric acid is shown with the O atoms colored red and the C atoms in green. Putative amino acids involved in Ca^{2+} binding are depicted by sticks. (**E**) The superimposed structures of Mincle (cyan, the same as (**D**)) and DC-SIGNR (pink) are shown. The stick model indicates the mannose (the O atoms colored red and the C atoms in pink) in the DC-SIGNR complex. (**F** and **G**) Mincle-Ig, mutated Mincle-Igs or hIgG (**F**) and MCL-Ig, mutated MCL-Igs or hIgG (**G**) were incubated with plate-coated TDM. Bound proteins were detected by anti–hIgG-HRP.

Fig. 4. Unique amino acid residues in MCL and Mincle, reporter assays of mutant Mincles, and SPR analysis.

(**A**-**B**) The superimposed structures of Mincle (cyan) and DC-SIGNR (**A**), and MCL (yellow) and DC-SIGNR (pink) (**B**) are shown. Arrowheads indicate the oxygen atom connected with mycolic acid in TDM (mannose binding to DC-SIGNR is shown in the figure). Dotted circles indicate the hydrophobic loops found in Mincle and MCL. A sequence comparison between human Mincle and Dectin-2 is shown. (**C**) Analyses of Mincle and its mutants were performed. Reporter cells expressing human Mincle or its mutants were stimulated with TDM for 18 h. (**D**-**F**) Electrostatic potentials of Mincle (**D**), MCL (**E**) and DC-SGINR (**F**) are shown. Electrostatic surface potentials were calculated using the program APBS (44) and represented by PyMOL, with the color of the surface potentials in the scale ranging from negatively charged (-4.0 kbT/ec, red) to positively charged amino acids (4.0 knT/ec blue). Black spheres are Ca^{2+} ions. The yellow surface indicated the hydrophobic site. (**G**) SPR analysis of Mincle and several lengths of acyl chains with trehalose were performed. The C12, C10 and C8 glycolipids used in this experiment are shown.

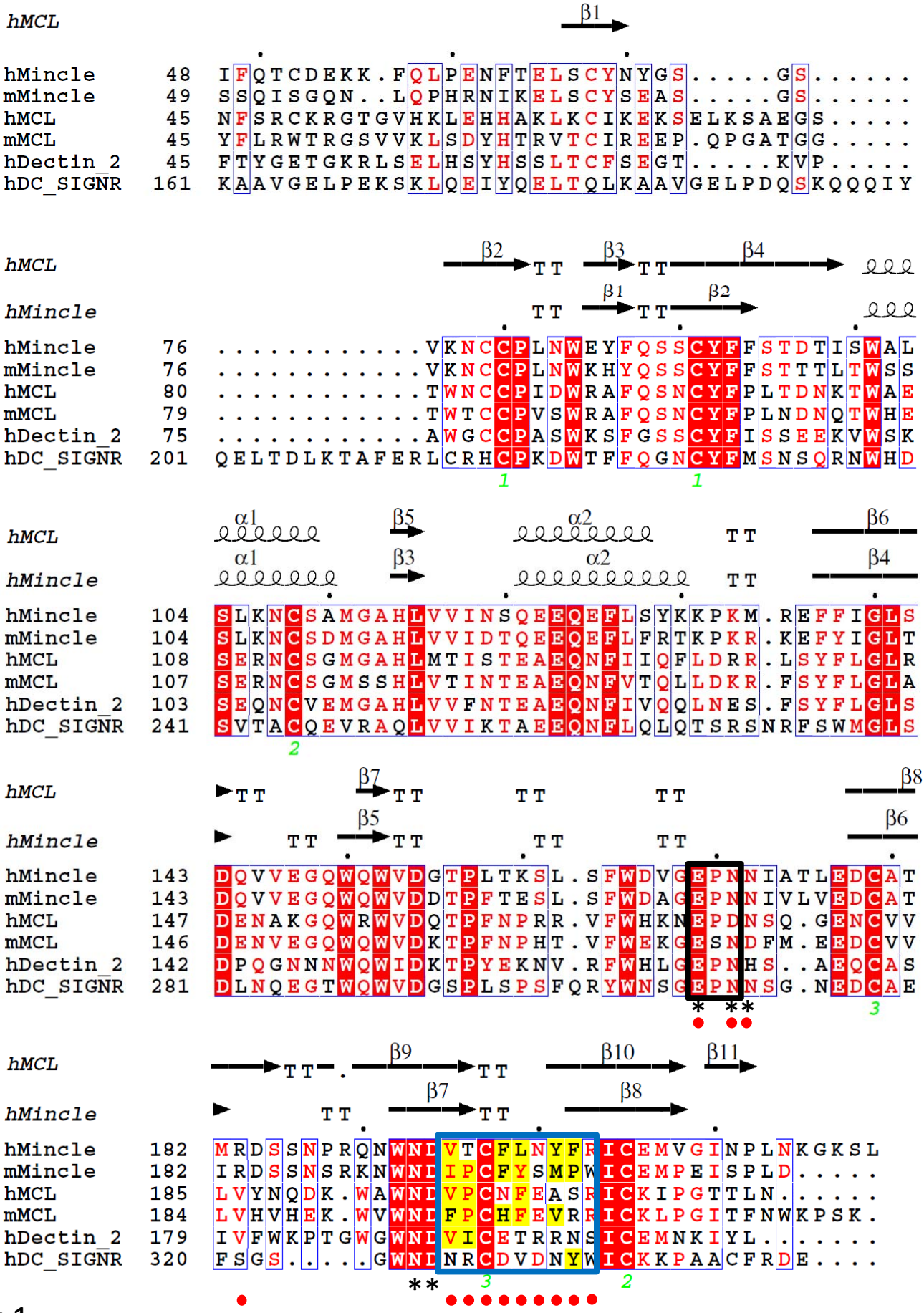

hMCL
β1
hMincle 48 IFQTCDEKK.FQLPENFTELSCYNYGS.....GS.....
mMincle 49 SSQISGQN..LQPHRNIKELSCYSEAS.....GS.....
hMCL 45 NFSRCKRGTGVHKLEHHAKLKCIKEKSELKSAEGS.....
mMCL 45 YFLRWTRGSVVKLSDYHTRVTCIREEP.QPGATGG.....
hDectin_2 45 FTYGETGKRLSELHSYHSSLTCFSEGT.....KVP.....
hDC_SIGNR 161 KAAVGELPEKSKLQEIYQELTQLKAAVGELPDQSKQQQIY
hMCL
β2
β3
β4
hMincle
β1
β2
hMincle 76 .....VKNCCPLNWEYFQSSCYFFSTDTISWAL
mMincle 76 .....VKNCCPLNWKHYQSSCYFFSTTTLTWSS
hMCL 80 .....TWNCCPIDWRAFQSNCYFPLTDNKTWAE
mMCL 79 .....TWTCCPVSWRAFQSNCYFPLNDNQTWHE
hDectin_2 75 .....AWGCCPASWKSFGSSCYFISSEEKVWSK
hDC_SIGNR 201 QELTDLKTAFERLCRHCPKDWTFFQGNCYFMSNSQRNWHD
1 1
hMCL
α1
β5
α2
β6
hMincle
α1
β3
α2
β4
hMincle 104 SLKNCSAMGAHLVVINSQEEQEFLSYKKPKM.REFFIGLS
mMincle 104 SLKNCSDMGAHLVVIDTQEEQEFLFRTKPKR.KEFYIGLT
hMCL 108 SERNCSGMGAHLMTISTEAEQNFIIQFLDRR.LSYFLGLR
mMCL 107 SERNCSGMSSHLVTINTEAEQNFVTQLLDKR.FSYFLGLA
hDectin_2 103 SEQNCVEMGAHLVVFNTEAEQNFIVQQLNES.FSYFLGLS
hDC_SIGNR 241 SVTACQEVRAQLVVIKTAEEQNFLQLQTSRSNRFSWMGLS
2
hMCL
β7
β8
hMincle
β5
β6
hMincle 143 DQVVEGQWQWVDGTPLTKSL.SFWDVGEPNNIATLEDCAT
mMincle 143 DQVVEGQWQWVDDTPFTESL.SFWDAGEPNNIVLVEDCAT
hMCL 147 DENAKGQWRWVDQTPFNPRR.VFWHKNEPDNSQ.GENCVV
mMCL 146 DENVEGQWQWVDKTPFNPHT.VFWEKGESNDFM.EEDCVV
hDectin_2 142 DPQGNNNWQWIDKTPYEKNV.RFWHLGEPNHS..AEQCAS
hDC_SIGNR 281 DLNQEGTWQWVDGSPLSPSFQRYWNSGEPNNSG.NEDCAE
3
hMCL
β9
β10
β11
hMincle
β7
β8
hMincle 182 MRDSSNPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL
mMincle 182 IRDSSNSRKNWNDIPCFYSMPWICEMPEISPLD.....
hMCL 185 LVYNQDK.WAWNDVPCNFEASRICKIPGTTLN.....
mMCL 184 LVHVHEK.WVWNDFPCHFEVRRICKLPGITFNWKPSK.
hDectin_2 179 IVFWKPTGWGWNDVICETRRNSICEMNKIYL.....
hDC_SIGNR 320 FSGS.....GWNDNRCDVDNYWICKKPAACFRDE....
3 2

Fig.1

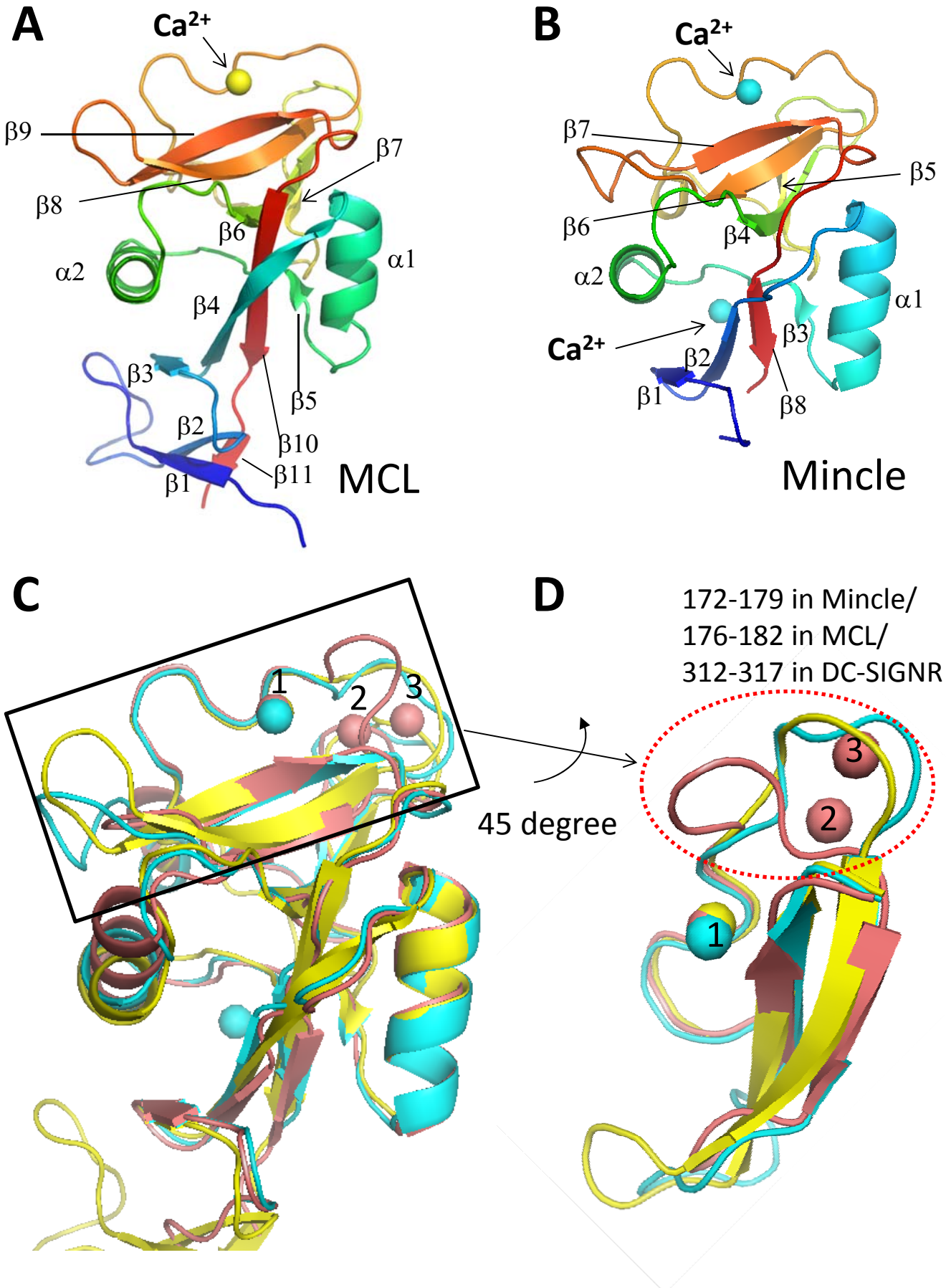

A
Ca2+
β9
β7
β8
β6
α2
α1
β4
β3
β5
β2
β10
β1
β11
MCL
B
Ca2+
β7
β5
β6
β4
α2
α1
Ca2+
β3
β2
β1
β8
Mincle
C
1
2
3
45 degree
D
172-179 in Mincle/
176-182 in MCL/
312-317 in DC-SIGNR
3
2
1

Fig.2

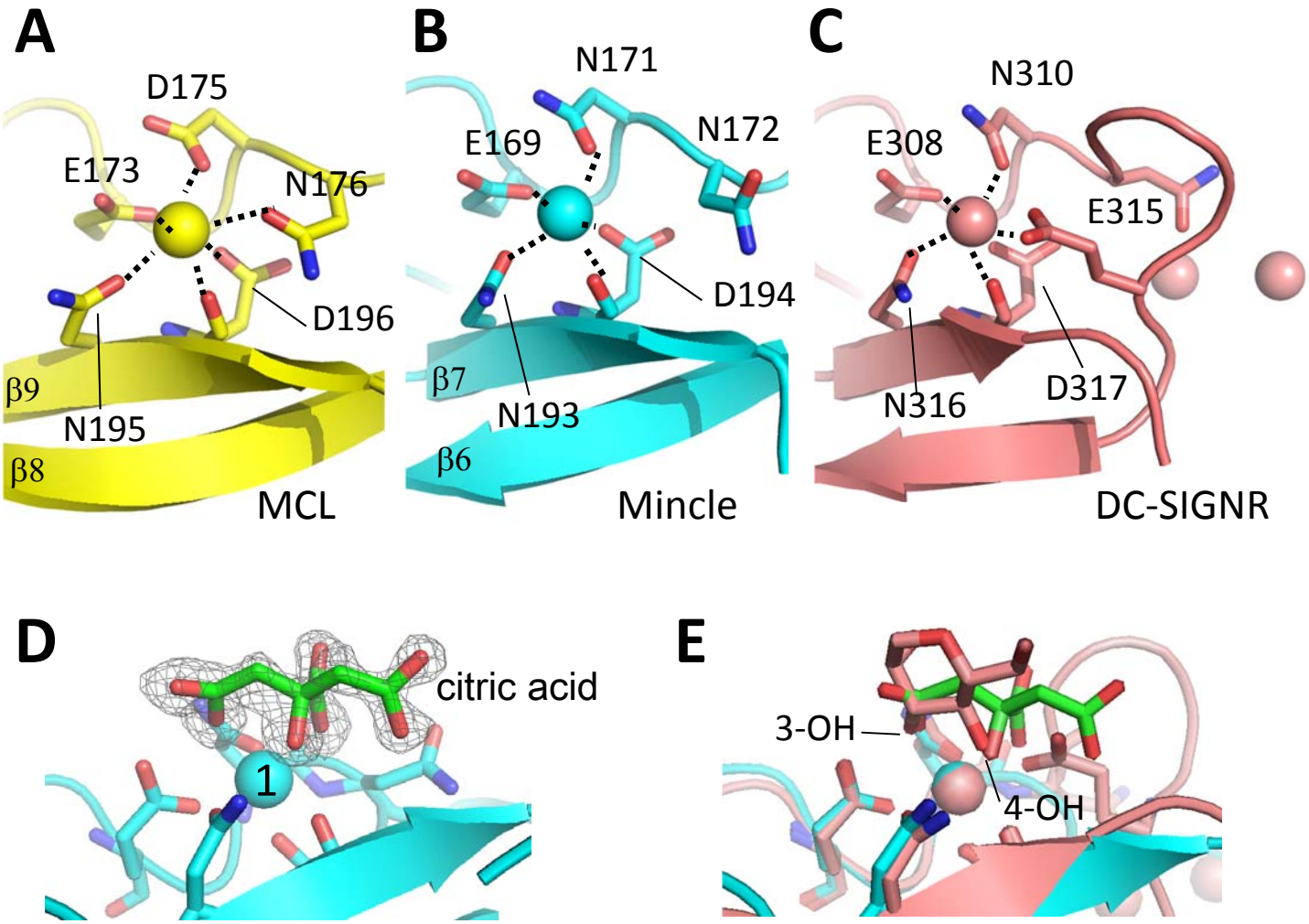
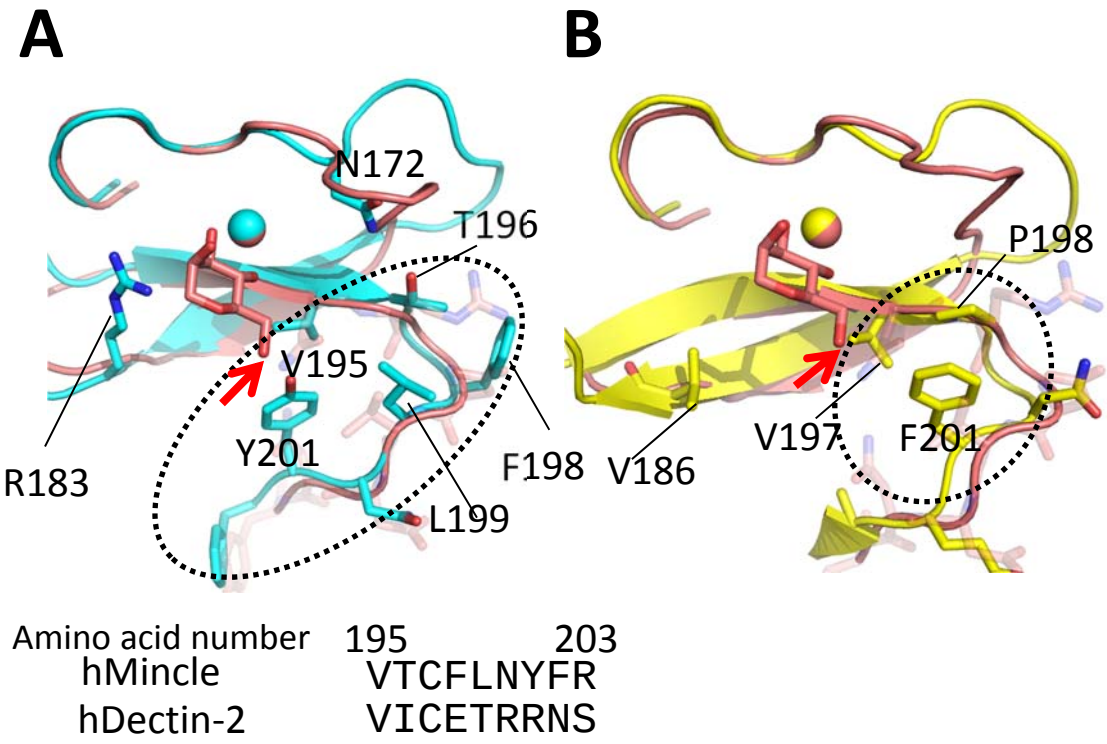
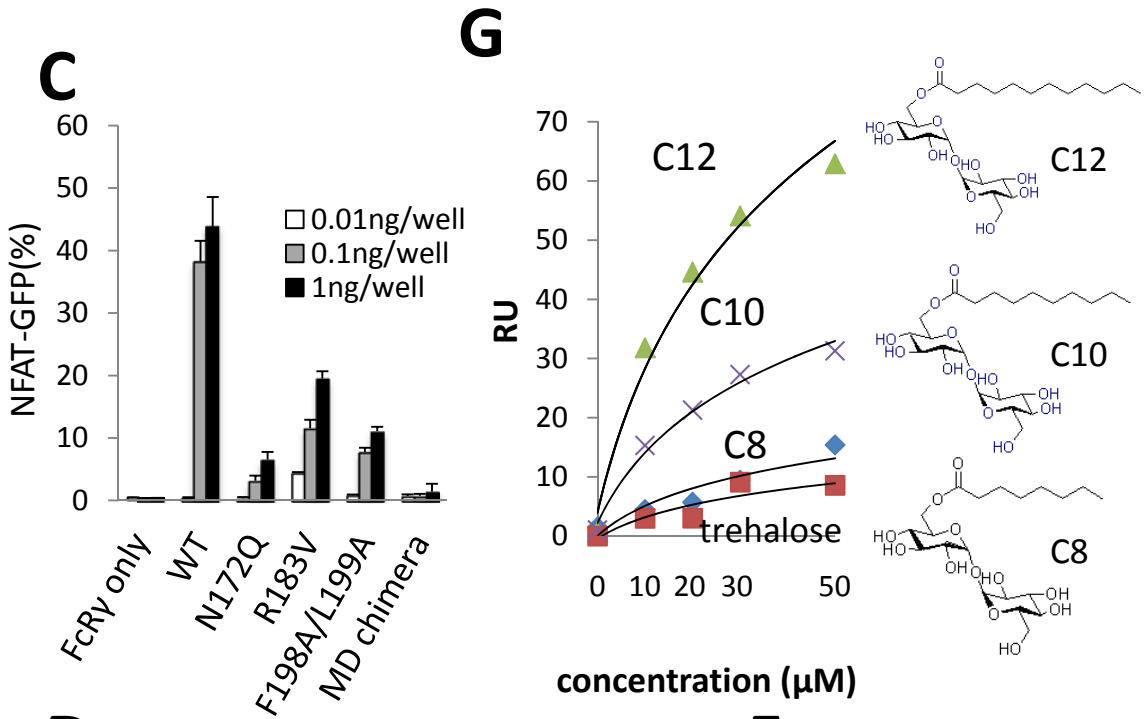

A
D175
E173
N176
D196
β9
N195
β8
MCL
B
N171
E169
N172
D194
β7
N193
β6
Mincle
C
N310
E308
E315
N316
D317
DC-SIGNR
D
citric acid
1
E
3-OH
4-OH
F
hIgG
hMincle WT(EPN)
hMincle EPN→QPD
hMincle EPN→EPD
OD 450 nm
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
0.03
0.1
0.3
1
Fusion-Ig protein (μg/ml)
G
hIgG
hMCL WT(EPD)
hMCL EPD→EPN
hMCL EPD→QPD
OD 450 nm
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
0.3
1
3
10
Fusion-Ig protein (μg/ml)

Fig.3



Amino acid number	195	203
hMincle	VTCFLNYFR	
hDectin-2	VICETRRNS	



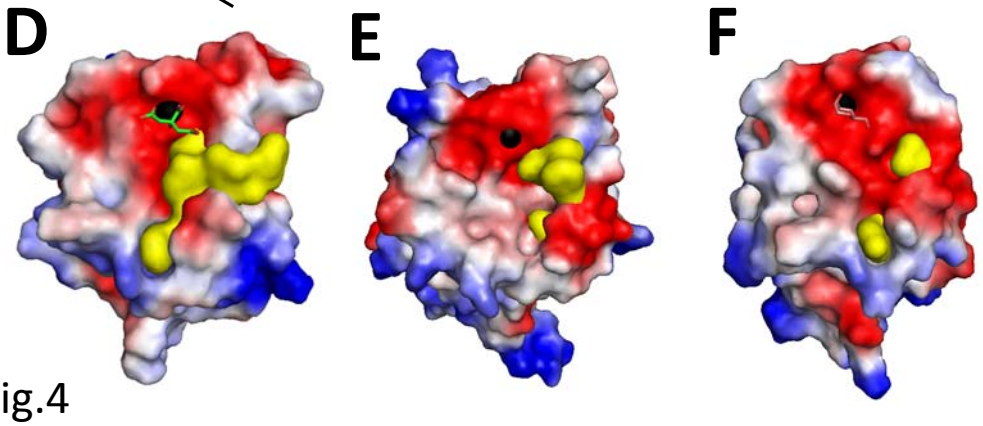


Fig.4