Published in final edited form as:

Cell. 2016 November 3; 167(4): 1041–1051.e11. doi:10.1016/j.cell.2016.09.056.

Crystal structure of a full-length human tetraspanin reveals a cholesterol-binding pocket

Brandon Zimmerman[1], **Brendan Kelly**[2], **Brian J. McMillan**[1], **Tom C.M. Seegar**[1], **Ron O. Dror**[2], **Andrew C. Kruse**[3], and **Stephen C. Blacklow**[4,*]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA; Dana Farber Cancer Institute, Boston, MA 02215, USA

[2]Departments of Computer Science and of Molecular and Cellular Physiology, and Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305

[3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

[4]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA; Dana Farber Cancer Institute, Boston, MA 02215, USA; Brigham and Women's Hospital, Boston, MA 02115, USA

SUMMARY

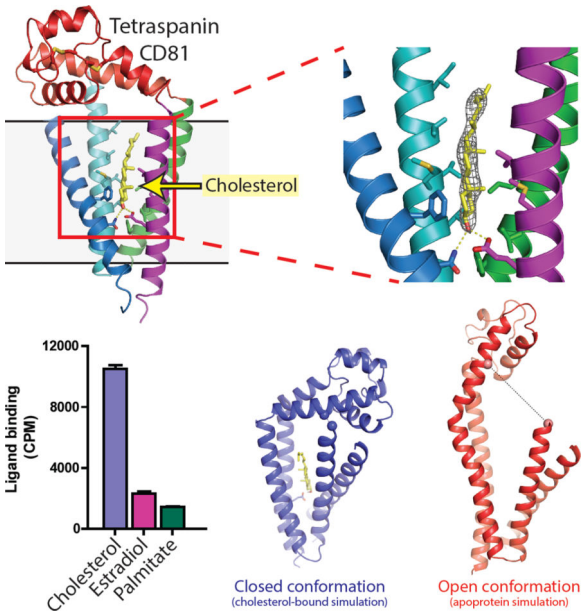
Tetraspanins comprise a diverse family of four-pass transmembrane proteins that play critical roles in the immune, reproductive, genitourinary, and auditory systems. Despite their pervasive roles in human physiology, little is known about the structure of tetraspanins or the molecular mechanisms underlying their various functions. Here we report the crystal structure of a full-length tetraspanin, human CD81. The transmembrane segments of CD81 pack as two largely separated pairs of helices, capped by the large extracellular loop (EC2) at the outer membrane leaflet. The two pairs of helices converge at the inner leaflet to create an intramembrane pocket with additional electron density corresponding to a bound cholesterol molecule within the cavity. Molecular dynamics simulations identify an additional conformation in which EC2 separates substantially from the transmembrane domain. Cholesterol binding appears to modulate CD81 activity in cells, suggesting a potential mechanism for regulation of tetraspanin function.

Graphical abstract

Correspondence to ACK or SCB: andrew_kruse@hms.harvard.edu or stephen_blacklow@hms.harvard.edu.
*Lead Contact

Author Contributions

The overall project was designed and developed by BZ, TCMS, ACK, and SCB. Molecular cloning, protein purification and crystallization experiments were conducted by BZ. Data collection was performed by BZ and ACK. Data processing, phase calculation, and structure refinement were carried out by BZ, BJM and ACK. BK and ROD performed and analyzed molecular dynamics simulations. The manuscript was written by BZ and SCB with input from BK, ROD, and ACK.



Introduction

Tetraspanins comprise a large family of four-pass transmembrane proteins arising evolutionarily in protists (Huang et al., 2005). In early unicellular eukaryotes, tetraspanins are believed to play a role in the dynamic regulation of membrane morphology, an activity thought to be subsequently co-opted for cell-cell interactions. As a result, there is an evolutionary link between the emergence of tetraspanins and the development of multicellularity (Huang et al., 2005). The development of specialized cell-cell interactions and new cell types likely selected for the duplication and differentiation of tetraspanins that evolved to effect specific functions. Thus, primitive fungi have a single tetraspanin, while the size of the tetraspanin family has grown to ten representatives in the sea urchin, *Strongylocentrotus purpuratus*, seventeen in the tunicate, *Ciona intestinalis* and 33 in *Homo sapiens* (Garcia-Espana et al., 2008).

Despite their common ancestry, the vertebrate tetraspanins have acquired a variety of discrete and important biological functions, which highlight their critical, though underappreciated, role in mammalian physiology (Hemler, 2005). Studies of tetraspanin knockouts in mice and other organisms have identified essential functions for tetraspanins in the immune, reproductive, genitourinary, and auditory systems. For instance, mice lacking CD151/tspan24 have abnormalities in hemostasis and lymphocytes that respond abnormally to mitogenic stimulation (Le Naour et al., 2000; Wright et al., 2004). Patients with frameshift mutations in CD151 present with a similar bleeding disorder, along with hereditary nephritis, deafness and epidermolysis bullosa (Karamatic Crew et al., 2004). Additionally, CD63/tspan30 knockout mice exhibit altered water homeostasis with increased urinary flow and water intake (Schroder et al., 2009), and CD9/tspan29 null mice are sterile due to failure of sperm-egg fusion (Le Naour et al., 2000).

The biological effects of many tetraspanins appear to be attributable to their activities as modulators of central signal transduction pathways. There is accumulating evidence that the six tetraspanins in the C8 subclass regulate Notch signaling by promoting ADAM10 trafficking and enzymatic maturation in both flies and mammals (Dornier et al., 2012; Haining et al., 2012; Jouannet et al., 2016; Noy et al., 2016). A number of other tetraspanins appear to exert their effects by influencing integrin signaling, either indirectly by modulating responses of integrins to their ligands (van Spriel et al., 2012; Wee et al., 2015), or by directly and stably associating with particular integrin heterodimers (Yauch et al., 1998). Moreover, knockout of CD37 in mice predisposes them to development of B-cell lymphoma, a phenotype attributed to unrestrained IL-6 signaling in the absence of CD37 (de Winde et al., 2016).

CD81/TAPA-1 (Target of Antiproliferative Antibody-1)/tspan28 was first identified as the target of an antibody discovered in a screen for the inhibition of proliferation in a lymphoma cell line (Oren et al., 1990). CD81, together with ME491/CD63 and CD37, became the founding members of the tetraspanin (TM4SF) protein family in mammals (Hemler, 2005; Huang et al., 2005). CD81 is also among the most well characterized tetraspanins because of its essential role in B-cell biology (Cherukuri et al., 2004b; Mattila et al., 2013), forming a multi-protein complex with CD19, CD21/CR2, and CD225 to regulate B cell receptor function. The importance of CD81 in B cell function is further highlighted by the recent report of a common variable immunodeficiency patient who has a germ line mutation that produces an alternatively spliced, truncated form of CD81 (van Zelm et al., 2010), which appears to sequester CD19 intracellularly (Vences-Catalan et al., 2015). CD81 is also a host factor that interacts with the Hepatitis C virus E2 envelope protein and is required for efficient HCV entry (Bradbury et al., 1992; Pileri et al., 1998).

Tetraspanins are the largest family of transmembrane proteins in mammals for which detailed structural information about the intramembrane domain remains unavailable. This lack of information, together with a limited understanding about the structural relationship between the intramembrane domain and the extracellular region of the protein, has greatly hindered efforts to understand the molecular mechanisms of action of these proteins. The tetraspanins are predicted to contain a number of shared structural features, including intracellular N- and C-termini, small (EC1) and large (EC2) extracellular loops, four transmembrane regions (TMs), and a short intracellular loop between TM2 and TM3 (Yanez-Mo et al., 2009). All tetraspanins also possess intracellular cysteine residues, which are typically palmitoylated (Yanez-Mo et al., 2009). These cysteines and their palmitoylation are required for efficient interaction with certain associated proteins and the formation of tetraspanin microdomains known as the tetraspanin web (Hopf et al., 2012), thought to be generated by tetraspanin-tetraspanin interactions as well as by heterologous interactions with other associated proteins in “signaling hubs” (Levy and Shoham, 2005). Though structures have been reported for the extracellular EC2 regions of *Schistosoma mansoni* TSP-2 and human CD81, they provide limited insight into other family members because of the poor conservation of the EC2 region, and offer no information about the intramembrane portion of the protein, which is the most highly conserved part of the molecule (Stipp et al., 2003).

We report here the structure of human CD81 (PDB: 5TCX), as a representative example of a full-length tetraspanin. The protein contains a bound cholesterol molecule in a large intramembrane pocket between two largely independent pairs of transmembrane helices capped by the EC2 domain, and specific cholesterol binding to this site is observed *in vitro*. The EC2 domain appears to more readily adopt an "open" conformation in molecular dynamics simulations when cholesterol is not present in its binding site, consistent with the modulation of tetraspanin function by mutations that compromise cholesterol binding. Together, the structural, computational and biochemical studies suggest a model for tetraspanin function, as well as a route to modulating tetraspanin activity as a therapeutic strategy in a variety of different diseases.

RESULTS

Crystal structure of CD81

To define the overall architecture of an intact tetraspanin, elucidate interdomain relationships between the EC2 and the intramembrane region, and gain insight into tetraspanin-ligand interactions and the tetraspanin web, we purified full-length CD81 from insect cells (Figures S1A and SA1B), crystallized the protein using the lipidic cubic phase method, and determined its structure. As is common for integral membrane protein crystals, the diffraction pattern was anisotropic with strong scattering along two reciprocal space axes and weaker scattering along the third. Prior to refinement we performed ellipsoidal truncation with resolution limits of 5.5 Å along the a* axis, 2.95 Å along the b* axis, and 2.95 Å along the c* axis (Table S1). To obtain phase information, we utilized a fragment-based iterative molecular replacement approach (Kruse et al., 2013). Briefly, we first located EC2 of CD81 by molecular replacement using a previously reported structure of the CD81 EC2 region (PDB, 1IV5, chain A) as a search model (Kitadokoro et al., 2002) (Figure S1C). After placing this domain, we used four more rounds of iterative molecular replacement using a polyalanine alpha helix search model to locate the four alpha helices comprising the transmembrane domain. This model was then used as a starting point for model building and refinement (Figure S2).

The overall structure of CD81 resembles a waffle cone in which the EC2 domain covers an intramembrane cavity bounded by the four transmembrane helices (Figure 1A). No electron density is visible for the small extracellular loop (EC1), suggesting this region is disordered. The overall fold of the four transmembrane helices does not resemble that of any other integral membrane protein of known structure. The transmembrane region to be two largely separated pairs of antiparallel helices: one pair comprises TM1/TM2 and the other TM3/TM4. The two pairs of helices only converge close to the cytoplasmic side of the membrane through contacts between TM2 and TM3. The central cavity bounded by the four transmembrane helices and the bottom face of EC2 encloses a total volume of 3300 $Å^3$ (Figure 1B).

Evolutionary conservation of tetraspanin structure

To confirm the relative positions of the transmembrane helices and assess our assignments of the residues in this region, we performed an evolutionary coupling analysis, which compares

homologous sequences to determine amino acid residues that are correlated through evolution (Marks et al., 2012). The correlation map for CD81 indicates that strong evolutionary couplings occur between helices one and two, between helices three and four, and within the EC2 domain, as predicted by the X-ray structure. We extended this analysis to additional tetraspanin proteins dispersed throughout phylogeny to examine whether this unusual fold is evolutionarily conserved, or a specific feature of CD81 (Figure 2). CD81 (Figure 2A), hypothetical protein FGSG_08695 from the fungus, *Fusarium graminearum* PH-1 (Figure 2B), tetraspanin 3A from the fruit fly, *Drosophila melanogaster* (Figure 2C) and AX4 tetraspanin family protein from the eukaryotic slime mold, *Dictyostelium discoideum* (Figure 2D) all exhibit the same pattern of evolutionary coupling, providing support for the inference that the fold seen in the CD81 structure is conserved among all tetraspanins.

A more detailed analysis of sequence conservation patterns for CD81 homologs across species (Figures 3A and S3A) and for tetraspanin paralogs within humans (Figures 3B and S3B) highlights the tight evolutionary constraints within the transmembrane region and the greater variability of the second extracellular loop by comparison. Strikingly, highly conserved residues within EC2 are located on extracellular helices one and two at their points of contact with the junction between EC1 and the first two transmembrane segments, accounting for the “closed” conformation of the EC2 cap. Contacts between the EC2 and the transmembrane region include hydrophobic interactions between L35 of TM1 and V146 of EC2, and of F56 of TM2 with F126 of EC2 (Figure S3C).

Analysis of intramembrane binding pocket

The most striking feature of the CD81 structure is the large, hydrophobic pocket bounded by the four transmembrane helices and the EC2 cap. We observe unexpected additional F_o-F_c electron density within this pocket. On the basis of the shape of the electron density, the presence of cholesterol in the crystallization mix, and the chemical features of the pocket, we tentatively identified the additional density as a bound cholesterol molecule (Figures 4 and S4, related to Figure 4). Within 4 Å of the bound cholesterol are a number of hydrophobic and aromatic residues, including F21 of TM1, I64, V68, V71, M72 and V75 of TM2, F94, L98 and L101 of TM3, and V212, and M216 of TM4 (Figure 4A). N18 and E219, two polar residues belonging to TM1 and TM4, respectively, form hydrogen bonds to the cholesterol hydroxyl group. Analysis of the CD81 sequences from all 37 available homologs reveals that these 13 amino acids are nearly 100% conserved (Figure S4). Comparison of the CD81 sequence with its 32 human paralogs shows N18 is conserved in 27 of the 33 human tetraspanins. However, E219 is only present in CD81 and tspan10, with the majority (64%) of mammalian tetraspanins having a polar glutamate or glutamine residue (82%) on the preceding turn of the helix, and a glycine residue (64%) at this position instead.

CD81 binds specifically to cholesterol

To determine directly whether CD81 specifically binds cholesterol, we assessed the cholesterol-binding ability of wild type CD81 immunopurified from HEK293T cells using a radioactive binding assay, and compared it with the Beta-lactam binding receptor BlaR, a

negative control four-pass transmembrane protein prepared similarly (Figure 4B). Wild type CD81 immunoprecipitates recover approximately 15-fold more cholesterol than immunoprecipitates from untransfected control cells or from cells expressing BlaR. To determine whether cholesterol binding by CD81 relies on specific interactions seen in the structure of CD81, we mutated residue E219 to either alanine or glutamine, as it is a critical polar contact with the cholesterol hydroxyl in our structure (the hydrogen bond between E219 and cholesterol is formed 90% of the time in the simulations described in the following section). The E219A and E219Q point mutants of CD81 recover about 50% less cholesterol than wildtype (Figure 4C), whereas a G26F/G30F double mutant on the external face of helix one does not detectably affect cholesterol recovery (Figure S4). Finally, to address whether CD81 and the transmembrane pocket are capable of binding other lipids, we explored the binding of radioactive estradiol and palmitate using the same radioligand binding assay. The recovery of both estradiol and palmitate is dramatically reduced in comparison to cholesterol (Figure 4D). Moreover, the amount of bound palmitate is unaffected by the E219Q mutation (though it does appear that the amount of bound estradiol is reduced somewhat upon introduction of the E219Q mutation, suggesting that it might have very weak affinity for the cholesterol binding pocket, consistent with the fact that it is a cholesterol derivative).

Molecular dynamics simulations identify an open conformation of EC2

We performed molecular dynamics simulations of CD81 in a hydrated lipid bilayer, both with and without cholesterol bound in the intramembrane pocket. In three of nine simulations of the apoprotein (*i.e.* with cholesterol removed), EC2 transitioned to an “open” conformation, in which it disengaged from TM1 and TM2 (Figure 5A). EC2 remained in this open conformation for the remainder of these three simulations. After opening, EC2 is flexible and dynamic relative to the TM domain (Figure 5C); however, the same fully open conformation is observed in all three simulations that display opening.

The opening motion involves a substantial straightening of TM3 and TM4. During EC2 opening, a salt bridge between D196 on EC2 and K201 on TM4 that stabilizes the closed state breaks, leading to extension and straightening of TM4. A new interaction between K116 and D117 stabilizes the extended form of TM3 observed in the open state (Figure 5B).

Interestingly, cholesterol-bound simulations consistently maintained a closed conformation, similar to the crystal structure, in which EC2 remained in contact with TM1 and TM2 (Figure 5A). In two of the nine simulations that we initiated with cholesterol bound, however, cholesterol dissociated from the binding pocket into the membrane, exiting through the gap between TM1 and TM4. In one of these simulations, EC2 transitioned to the fully open conformation after cholesterol dissociated, and remained in that conformation for the rest of the simulation. The fact that we only observed opening when cholesterol was absent from the binding pocket suggests that the presence of cholesterol may stabilize the closed conformation, and that the absence of cholesterol may favor opening.

TM1 and TM2 undergo substantial motion in simulations both in the presence and in the absence of cholesterol (Figure S5). The intracellular end of TM1 remains in close contact with TM4 in simulations of the apoprotein, stabilized by a hydrogen bond between the side

chains of N18 and E219. When cholesterol is bound, it competes for interaction with E219, often leading the intracellular end of TM1 to separate from TM4.

Cholesterol binding regulates CD81 function

Numerous reports have highlighted the importance of CD81 in CD19 export to the cell surface, and a CD81 truncation in a human patient results in a combined variable immunodeficiency phenotype and intracellular retention of CD19 in cell-based assays (van Zelm et al., 2010; Vences-Catalan et al., 2015). Here, we used a flow cytometry assay to measure the amount of CD19 at the surface of transfected cells, and the ability of CD81 to increase the amount of CD19 detected at the cell surface. In the absence of added CD81, 293T cells transfected with FLAG-tagged CD19 show a minimal increase in surface staining compared to untransfected cells. Upon co-transfection of wild-type CD81, we see a ten-fold increase in surface staining compared to CD19 with control vector (Figure 6A). Moreover, when cholesterol binding is compromised by either the E219A or E219Q mutation, the amount of CD19 surface staining increases further by an additional 50% when comparable amounts of CD81 protein are expressed and present at the cell surface, whereas no effect is seen when the G26F/G30F mutant is tested (Figure 6B).

DISCUSSION

We have solved the crystal structure of a full-length human tetraspanin and defined an unknown intramembrane binding pocket for tetraspanins. The discovery of the pocket with cholesterol bound was then used to design experiments for interrogation of cholesterol binding and its potential role in tetraspanin function.

Our structure of CD81 revealed a monomeric form of a tetraspanin, which contrasts with the dimeric structure of the isolated EC2 fragment. The putative dimerization interface seen in the structure of the isolated EC2 fragment is located on its bottom face in the full length protein within 3.5 Å of the TM1/TM2 bundle, indicating that the dimer is likely a non-native interaction driven by lattice packing effects in the absence of the transmembrane regions of the protein.

Early modeling work proposed that the four TM helices would form a tightly associated four-helix bundle (Seigneuret, 2006), yet our structure reveals the transmembrane region to fold as two largely separated pairs of antiparallel helices, a conclusion supported by evolutionary coupling analysis (Figure 2). The overall similarity seen among diverse family members argues strongly that all tetraspanin proteins possess this transmembrane fold. The intramembrane pocket seen in the structure between TM1/4 and TM2/3 is likely to be accessible only through lateral diffusion within the membrane plane, because entrance from the extracellular space is precluded by the presence of EC2 above the pocket.

Much attention has been previously given to the concept of a tetraspanin web, where tetraspanins form homooligomers or heterooligomers with other tetraspanins to form higher order complexes and protein-rich microdomains in the cell membrane (Charrin et al., 2003a; Horvath et al., 1998; Levy and Shoham, 2005; Rubinstein et al., 2013). Recently, the existence of the tetraspanin web has come under more scrutiny, as early experiments done

with detergents did not effectively disrupt secondary interactions (Dornier et al., 2012; Zuidscherwoude et al., 2015). In our CD81 crystals, individual monomers in the lattice pack such that adjacent subunits in the membrane plane lie in antiparallel orientations (Figure S6), an arrangement that is necessarily non-physiological. Though the absence of lateral homotypic packing interactions among adjacent subunits does not in itself exclude the possibility of higher order CD81 complexes or the potential for assembly of a tetraspanin web, there is also evidence among the uroplakin tetraspanins that tetraspanin monomers do not interact with one another. Uroplakins Ia and Ib, in complex with their accessory proteins uroplakin II and IIIa, form hexameric lattices. Despite the close proximity of the six tetraspanin molecules in the lattice, the lattice is entirely bridged by interactions between the non-tetraspanin partners, uroplakin II and IIIa (Min et al., 2006). A recent study using super-resolution microscopy also shows that tetraspanins, including CD81, lie in closer proximity to their non-tetraspanin binding partners than to other tetraspanins in the membrane, again suggesting that tetraspanins need not be constitutively oligomeric in the cellular milieu (Zuidscherwoude et al., 2015). An important caveat from our structure is the fact that four intracellular cysteine residues were mutated to prevent disulfide crosslinking during purification. These sites are expected to be palmitoylated, and palmitoylation of tetraspanins has been implicated for at least some tetraspanin-tetraspanin interactions (Yang et al., 2002) and tetraspanin functions (Cherukuri et al., 2004a). However, it now seems at least equally likely that other factors are more responsible for the creation of tetraspanin-enriched microdomains than direct interactions between tetraspanins themselves.

The relationship between cholesterol and tetraspanins has been discussed in numerous studies, often in the context of the tetraspanin web connecting lipid rafts to other membrane signaling domains. A study examining the role of CD82 in the actin cytoskeleton of T-lymphocytes revealed that removal of cholesterol altered the cellular distribution of CD82, and furthermore disrupted all CD82-dependent signaling events (Delaguillaumie et al., 2004). Early work demonstrated a physical link between cholesterol and CD81 as well as CD9 and CD82 using photoactivatable cholesterol crosslinking, demonstrating the close proximity of cholesterol and tetraspanins within the membrane (Charrin et al., 2003b). Perhaps the most compelling evidence highlighting the central role of cholesterol to tetraspanin biology comes from a study showing that cholesterol in the host cell membrane was required for both maintenance of a CD81 monoclonal antibody epitope and for CD81-dependent infection by *Plasmodium falciparum* sporozoites, but not by sporozoites that were CD81-independent (Silvie et al., 2006). Our results reveal that CD81 binds cholesterol and that E219 is an important residue in the binding pocket, offering a potential mechanism for how tetraspanins might detect cholesterol or other membrane lipids (whereas CD81 and tspan10 both possess a polar residue in this position, most other tetraspanins have a polar residue one helical turn earlier, with other cholesterol-interacting residues highly conserved throughout evolution).

Our molecular dynamics simulations suggest that CD81 may exist in both closed and open conformations, and that the equilibrium between closed and open conformations may be shifted toward the closed state when cholesterol is bound. Mutations that interfere with cholesterol binding by CD81 result in enhanced delivery of its binding partner (CD19) to the cell surface. On the basis of these observations, we speculate that the open conformation

may bind partner proteins more tightly, and that the closed state, favored when cholesterol is bound, disfavors partner binding, and hence, protein export (Figure 6C). Modulation of tetraspanin conformation in response to differences in cholesterol concentration could explain how tetraspanins regulate the subcellular localization of their partner proteins by loading and unloading cargo in response to variation in lipid composition among different membrane compartments.

Accumulating evidence that tetraspanins such as CD81 have growth promoting activity in certain human cancers has led to their emergence as potential therapeutic targets (Hemler, 2014). Indeed, monoclonal antibodies directed at the tetraspanin26/CD37 have moved into phase I or phase II clinical trials for the treatment of chronic lymphocytic leukemia and non-Hodgkin lymphoma (Deckert et al., 2013; Zhao et al., 2007). The unexpected discovery of the intramembrane ligand binding pocket represents a potential targetable site for modulating tetraspanin function using small molecules, and provides a route forward for elucidation of the molecular mechanism of action of this enigmatic class of important but poorly understood proteins.

STAR METHODS

CONTACT FOR RESOURCE AND REAGENT SHARING

For additional information about reagents and resources, contact the Lead Contact, Stephen Blacklow, at Stephen_blacklow@hms.harvard.edu.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

CD81 protein for crystallographic studies was expressed in *Spodoptera frugiperda* ovarian tissue (SF9) cells grown in grown in ESF 921 Insect Cell Culture Medium, Protein-Free media at 27 °C. Protein was harvested 42 hours after baculov irus infection and isolated from the Sf9 membrane fraction.

Radioactivity experiments were performed with protein isolated from HEK293T cells. HEK293T cells were grown in DMEM media supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. HEK293T cells were also used in the flow cytometry assays.

METHOD

DETAILS Expression and purification

The sequence encoding full-length human CD81 was assembled using synthetic gene blocks (gBlocks, Integrated DNA Technologies) and inserted into the baculovirus transfer vector pVL1392 with an amino-terminal FLAG epitope tag followed by a 3C protease cleavage site. Four intracellular cysteine residues at positions 6, 9, 227 and 228 were mutated to serine to prevent disulfide crosslinking during purification. This protein was expressed in *Sf9* insect cells using the BestBac system (Expression Systems) according to the manufacturer's instructions with Fugene HD as the transfection reagent. Infection was performed when cells reached a density of 4×10^6 cells/ml, and flasks were then shaken at 27 °C for 42 hours prior to harvest. Cells were harvested by centrifugation and frozen at

−80 °C until purification. After frozen cell paste was thawed, cells were lysed by osmotic shock in 20 mM HEPES pH 7.4, 2 mM magnesium chloride, 2 mg/ml iodoacetamide (Sigma Aldrich) and 1:100,000 (v:v) benzonase nuclease (Sigma Aldrich). Lysed cells were centrifuged at 18,000 RPM in a Sorvall RC 5C Plus centrifuge with an SS-34 rotor for 15 minutes. CD81 was then extracted from the pellet using a glass dounce tissue grinder to homogenize lysed cells in a solubilization buffer consisting of 250 mM NaCl, 20 mM HEPES pH 7.5, 10% (v/v) glycerol, 1% (w/v) n-Dodecyl-β-D-Maltoside (DDM - Anagrade; Anatrace), 0.1% (w/v) cholesterol hemisuccinate (CHS; Steraloids) and 2 mg/ml iodoacetamide. Samples were stirred for 2 hours at 4 °C, then centrifuged at 20,000 RPM for 20 min. Supernatant containing solubilized protein was filtered through a glass microfiber prefilter and loaded by gravity flow onto 3 mL of M2 anti-FLAG antibody affinity resin (Sigma Aldrich). The resin was then washed with 50 ml of buffer containing 100 mM NaCl, 20 mM HEPES pH 7.4, 1% glycerol, 0.1% DDM, and 0.01% CHS. Bound protein was eluted in the same buffer supplemented with 0.2 mg/mL FLAG peptide. 3C protease was added (1:100 w:w) and incubated with CD81 at 4 °C overnight. CD81 was further purified by size exclusion chromatography (SEC) on a Sephadex S200 column (GE Healthcare) in buffer containing 0.1% DDM, 0.01% CHS, 100 mM NaCl, and 20 mM HEPES pH 7.4. CD81 was biochemically pure but consistently eluted as two broad peaks during SEC. The SEC-purified protein was concentrated to 30 – 40 mg/mL and flash frozen with liquid nitrogen in aliquots of 8 μL. Samples were stored at −80°C until use for crystallography. Purity and monodispersity of crystallographic samples was evaluated by SDS-PAGE and analytical SEC, respectively (Figure S1).

Crystallography and data collection

Purified CD81 was reconstituted into lipidic cubic phase by mixing with a pre-made 10:1 (w:w) mix of monoolein (Hampton Research) with cholesterol (Sigma Aldrich) at a ratio of 1.5:1.0 lipid:protein by mass, using the coupled syringe reconstitution method (Caffrey and Cherezov, 2009). All samples were mixed at least 100 times. The resulting phase was dispensed in 40 – 50 nL drops onto a glass plate, and overlaid with 650 nL of precipitant solution using a Gryphon LCP robot (Art Robbins Instruments). Crystals grew in precipitant solution containing 35 – 45% PEG 300, 100 – 400 mM dibasic sodium citrate, 0.1 M Tris pH 8. Initial crystallization hits grew slowly, with crystals appearing after more than three weeks. Crystals were harvested using mesh loops and stored in liquid nitrogen until data collection.

Data collection was performed at Advanced Photon Source GM/CA beamline 23ID-B. An initial grid raster with 80 × 30 μm beam dimensions was performed, followed by a sub-raster using a 20 μm beam to locate crystals in the loop. Additional rastering was performed using a 10 μm beam diameter to optimally position crystals for data collection. Data collection used a 10 μm beam and diffraction images were collected in 1 degree oscillations at a wavelength of 1.033 Å. The final data set for CD81 was compiled by merging data from five crystals using XDS (Kabsch, 2010). Data are summarized in Table S1, using the format of Harrison and colleagues (Corbett et al., 2010).

Visual inspection of diffraction frames revealed strong anisotropy, with weaker diffraction in the reciprocal *a** axis and stronger diffraction in the other two directions. We thus performed ellipsoidal truncation on the merged dataset using the UCLA anisotropy diffraction server (Strong et al., 2006). Resolution limits along reciprocal space axes were chosen based on an F/sigF > 2 criterion, giving resolution limits of 5.50 Å, 2.95 Å and 2.95 Å along the reciprocal space *a**, *b** and *c** axes, respectively. Anisotropic B scaling was also applied using the server to compensate for differences in intensity along each reciprocal space axis.

Phasing and structure refinement

To obtain phases, molecular replacement was performed in Phenix with Phaser using chain A from a CD81 ectodomain x-ray structure, PDB ID 1IV5 (Kitadokoro et al., 2002) (Figure S1C). A subsequent molecular replacement search was conducted using a polyalanine helix as a search model, with the ectodomain included as a fixed partial model. This procedure, which located one transmembrane helix, was performed three more times, in order to locate all four transmembrane helices. For refinement, reciprocal space optimization of XYZ coordinates and individual atomic B-factor parameters was performed with standard Phenix restraints in phenix.refine (Afonine et al., 2012); optimization for X-ray/stereochemistry weight and X-ray/ADP weight was also performed. Extensive additional model building was carried out manually in Coot (Emsley and Cowtan, 2004) and refined following each round of model building. Near the end of refinement, custom geometric restraints were placed on the two disulfide bonds present in the EC2. As a control for register assignment, the structure was built and register assigned using two independent approaches. First, the model was manually built and register assigned by inspection of electron density in combination with consideration of evolutionary conservation of residues interacting within the transmembrane domains (Marks et al., 2012). In parallel, sequence register was independently assigned automatically with phenix.autobuild, which resulted in the same register assignment. Representative composite omit map density is shown in Figure S2. Cholesterol was manually placed into an F_o-F_c difference map. Following refinement, structure quality was assessed using MolProbity (Chen et al., 2010), and figures were prepared in PyMOL (Schrodinger, 2010). All crystallographic data processing, refinement, and analysis software was compiled and supported by the SBGrid Consortium (Morin et al., 2013).

Evolutionary coupling analysis

For evolutionary coupling analysis, the EVcouplings option on www.evfold.org was used (Marks et al., 2012). The sequence of the indicated tetraspanins was used as a template and default search settings were used, except that alpha helical TM domain was set to “Yes”. The top 90 evolutionary coupling pairs were chosen for display.

Sequence conservation analysis

For conservation analysis, a multiple sequence alignment aligning human CD81 to its most similar 150 homologs was analyzed using the Consurf server (Ashkenazy et al., 2010; Landau et al., 2005). The multiple sequence alignment was generated using a protein sequence BLAST on the NCBI database. For conservation analysis among the 33 human

tetraspanins, the sequences were aligned using Clustal Omega and then analyzed using the Consurf server (Sievers et al., 2011).

Radioactive lipid binding assay

[1,2-^{3}H]cholesterol (45 Ci/mmol), 2,4,6,7-^{3}H(N)-estradiol (94 Ci/mmol), and 9,10-^{3}H(N)-palmitic acid (30 Ci/mmol) were purchased from PerkinElmer Life Sciences. FLAG-tagged CD81 was subcloned into the mammalian expression vector pcDNA3.1 using NheI and BamHI restriction sites and mutations were made using gBlocks from IDT. FLAG-BlaR was constructed using PCR, and subcloned into the mammalian expression vector pcDNA3.1 using the same restriction sites. HEK293T cells were transfected with empty vector, FLAG-tagged CD81 (wildtype, E219A, E219Q), or FLAG-tagged BlaR. Cells were lysed 36 hours later with 20 mM HEPES pH 7.4, 2 mM $MgCl_2$, 1 μl benzonaze and 2 mg/ml iodoacetamide. Lysed cells were harvested by centrifugation at 15,000 RPM for 10 min at 4 °C. Pellets were resuspended in 1% DDM, 0.1% CHS, 20 mM HEPES pH 7.4 and 250 mM NaCl and solubilized for 2 hours. The samples were clarified by centrifugation at 15,000 RPM for 10 min at 4 °C. The soluble fraction was then incubated with 15 μl of anti-M2 FLAG beads per sample for 1 hour. Beads were recovered and washed three times with 0.1% DDM, 0.01% CHS, 20 mM HEPES pH 7.4 and 100 mM NaCl. FLAG-CD81 conjugated beads were then incubated with 1 μCi of ^{3}H-lipid for 1 hour. Beads were washed three times used a spin filter column and remaining ^{3}H-lipid was counted using a Beckman Coulter LS6500 scintillation counter.

Molecular dynamics simulation – System setup

Simulations of CD81 were based on the cholesterol-bound crystal structure described in this manuscript. The receptor was simulated in two distinct conditions: (1) the cholesterol-bound crystal structure; (2) the same structure with cholesterol removed.

Prime (Schrödinger, Inc.) was used to model in missing side chains and the missing residues E86, S87 and 41–54 of EC1. The crystallized protein construct with cysteine mutations at S6, S9, S227 and S228 was employed during simulation. *S*-(2-amino-2-oxoethyl)-cysteines 80 and 89 were returned to natural cysteines. Hydrogen atoms were added, and protein chain termini were capped with the neutral groups acetyl and methylamide. Titratable residues were left in their dominant protonation state at pH 7.0. An alternative side-chain orientation of K11 was used to prevent early formation of potentially artefactual interactions with E219 that was observed in initial simulations. For apoprotein simulation conditions, cholesterol was removed.

The prepared protein structures were aligned on the transmembrane helices to the *z*-axis, and internal waters added with Dowser (Hermans, 2003; Zhang and Hermans, 1996) (in addition to internal waters resolved in the crystal structure). The structures were then inserted into a pre-equilibrated palmitoyl-oleoyl-phosphatidylcholine (POPC) bilayer, and solvated with 0.15 M NaCl in explicitly represented water, then neutralized by removing chloride ions. Final system dimensions were approximately 80 × 70 × 96 Å^{3}, including about 124 lipids, 29 sodium ions, 25 chloride ions, and 10600 water molecules.

MD simulation protocol

We used the CHARMM36 parameter set for protein molecules, lipid molecules, cholesterol and salt ions, and the CHARMM TIP3P model for water; protein parameters incorporated CMAP terms (Best et al., 2012a; Best et al., 2012b; Huang and MacKerell, 2013; Klauda et al., 2010; MacKerell et al., 1998).

Simulations were performed on GPUs using the CUDA version of PMEMD (Particle Mesh Ewald Molecular Dynamics) in Amber14 (D.A. Case, 2015; Le Grand et al., 2013; Salomon-Ferrer et al., 2013).

Prepared systems were minimized, then equilibrated as follows: The system was heated using the Langevin thermostat from 0 to 100 K in the NVT ensemble over 12.5 ps with harmonic restraints of 10.0 kcal·mol^{-1}·Å^{-2} on the non-hydrogen atoms of lipid, protein and ligand, and initial velocities sampled from the Boltzmann distribution. The system was then heated to 310 K over 125 ps in the NPT ensemble with semi-isotropic pressure coupling and a pressure of one bar. Further equilibration was performed at 310 K with harmonic restraints on the protein and ligand starting at 5.0 kcal·mol^{-1}·Å^{-2} and reduced by 1.0 kcal·mol^{-1}·Å^{-2} in a stepwise fashion every 2 ns, for a total of 10 ns of additional restrained equilibration.

We performed nine simulations of cholesterol-bound CD81, and another nine simulations of unliganded CD81, for a total of 18 simulations. These simulations were conducted in the NPT ensemble at 310 K and 1 bar, using a Langevin thermostat and Monte Carlo barostat. In each of these simulations, we performed 5 ns of unrestrained equilibration followed by a production run of 1.0–2.5 μs. Simulations initiated with cholesterol bound totaled 13.7 μs, and simulations initiated with cholesterol removed totaled 14.8 μs.

Simulations used periodic boundary conditions, and a time step of 2.5 fs. Bond lengths to hydrogen atoms were constrained using SHAKE. Non-bonded interactions were cut off at 9.0 Å, and long-range electrostatic interactions were computed using the particle mesh Ewald (PME) method with an Ewald coefficient β of approximately 0.31 Å and B-spline interpolation of order 4. The FFT grid size was approximately 80 × 64 × 128.

MD simulation analysis

Trajectory snapshots were saved every 100 ps during production simulations. Trajectory analysis was performed using VMD (Humphrey et al., 1996) and CPPTRAJ (Roe and Cheatham, 2013), and visualization was performed using VMD. Trajectories were aligned to the crystal structure, along TM3 and TM4.

To determine how frequently the hydrogen bond between the hydroxyl group of cholesterol and the carboxylate side-chain of E219 was formed in simulation, we calculated the percentage of time the distance between the cholesterol hydroxyl hydrogen and either carboxylate oxygen of E219 was 3.3 Å or less during simulations.

To calculate a *p*-value associated with the alternative hypothesis that the absence of cholesterol in the binding pocket favors opening, we proceed as follows. We define the *p*-value as the probability that we would have observed data that favors the alternative

hypothesis at least as strongly as the data we observed under the null hypothesis that opening is equally likely when simulating CD81 with or without cholesterol in the binding pocket. Given that we observed no opening events when cholesterol was bound and four opening events when it was not, we calculate the probability, under the null hypothesis, of observing no opening events when cholesterol was bound and four or more opening events when it was not. Because we don't have an a priori estimate of the probability of opening in a given simulation, we maximize the calculated *p*-value over all possible opening probabilities—in other words, we report the largest of a family of possible *p*-values.

More specifically, we assume that, under the null hypothesis, the number of opening events observed in t μs of simulation of closed-state CD81 is given by a Poisson distribution with mean qt, where q is a proportionality or rate constant. The value of q is unknown but assumed to be the same in simulations with or without cholesterol bound. We calculate, as a function of q, the probability of observing no opening events when cholesterol is bound (in 11.04 μs of simulation of closed, cholesterol-bound CD81) and four or more opening events when cholesterol is not bound (in 12.55 μs of simulation of closed CD81 with no cholesterol bound, including the relevant portion of simulations in which cholesterol dissociated). Our calculated p-value, 0.027, is the maximum value of this probability over all possible values of q ($p = .03$; see Methods).

Flow cytometry assay

HEK293T cells were transfected with empty vector, Flag-tagged CD19 or ProtC-tagged CD81 (wildtype, E219A, E219Q). Cells were washed with PBS 36 hours post-transfection and resuspended with 20mM HEPES pH7.4, 150mM NaCl, and 0.1% BSA. Cells were incubated for 15min with Alexa 488-anti-CD19 (Molecular probes) and APC-anti-CD81 (Biolegend) per manufacturers recommendation. Cells were washed and plated into a 96 well U-bottom dish and sorted on a BD Accuri C6 flow cytometer. Experiments were done in triplicate or quadruplicate at least 4 independent times.

QUANTIFICATION AND STATISTICAL ANALYSIS

Bar graphs display mean ± standard deviation. p values were calculated by one-way ANOVA followed by post-hoc Bonferroni tests where applicable using GraphPad Prism. All radioactivity experiments were performed three independent times with duplicate measurements (figure 4C). Flow cytometry experiments were performed four independent times with triplicate measurements (figure 6A and 6B). Statistical analyses relevant to the structural model are included in Table S1, related to Figure 1.

DATA AND SOFTWARE AVAILABILITY

The structure reported in the paper is deposited in the PDB under code: 5TCX. Graphpad Prism 7.0 is available from an institutional license at Harvard Medical School. The EVfold server is a freely available online tool created by the Marks lab at Harvard Medical School and the Sander lab at Memorial Sloan-Kettering Cancer Center. All software used in structure determination (XDS, Phenix, Pymol and Coot) are accessible via the Structural Biology Grid (SBGrid) at Harvard Medical School.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank beamline staff at APS GM/CA beamline 23ID-B for their superb technical assistance and support. Financial support for this work was provided by NIH grant NCI 5 RO1 CA092433 to SCB and 1DP5 OD021345 to ACK. BZ was supported by a CIHR postdoctoral fellowship, and TCMS was supported by training grant 2T32 HL007627.

References

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012; 68:352–367. [PubMed: 22505256]

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic acids research. 2010; 38:W529–W533. [PubMed: 20478830]

Best RB, Mittal J, Feig M, MacKerell AD Jr. Inclusion of many-body effects in the additive CHARMM protein CMAP potential results in enhanced cooperativity of alpha-helix and beta-hairpin formation. Biophysical journal. 2012a; 103:1045–1051. [PubMed: 23009854]

Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, Mackerell AD Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. Journal of chemical theory and computation. 2012b; 8:3257–3273. [PubMed: 23341755]

Bradbury LE, Kansas GS, Levy S, Evans RL, Tedder TF. The CD19/CD21 signal transducing complex of human B lymphocytes includes the target of antiproliferative antibody-1 and Leu-13 molecules. Journal of immunology. 1992; 149:2841–2850.

Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. Nat Protoc. 2009; 4:706–731. [PubMed: 19390528]

Charrin S, Manie S, Billard M, Ashman L, Gerlier D, Boucheix C, Rubinstein E. Multiple levels of interactions within the tetraspanin web. Biochemical and biophysical research communications. 2003a; 304:107–112. [PubMed: 12705892]

Charrin S, Manie S, Thiele C, Billard M, Gerlier D, Boucheix C, Rubinstein E. A physical and functional link between cholesterol and tetraspanins. European journal of immunology. 2003b; 33:2479–2489. [PubMed: 12938224]

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

Cherukuri A, Carter RH, Brooks S, Bornmann W, Finn R, Dowd CS, Pierce SK. B cell signaling is regulated by induced palmitoylation of CD81. The Journal of biological chemistry. 2004a; 279:31973–31982. [PubMed: 15161911]

Cherukuri A, Shoham T, Sohn HW, Levy S, Brooks S, Carter R, Pierce SK. The tetraspanin CD81 is necessary for partitioning of coligated CD19/CD21-B cell antigen receptor complexes into signaling-active lipid rafts. Journal of immunology. 2004b; 172:370–380.

Corbett KD, Yip CK, Ee LS, Walz T, Amon A, Harrison SC. The monopolin complex crosslinks kinetochore components to regulate chromosome-microtubule attachments. Cell. 2010; 142:556–567. [PubMed: 20723757]

D.A. Case, JTB.; Betz, RM.; Cerutti, DS.; Cheatham, TE., III; Darden, TA.; Duke, RE.; Giese, TJ.; Gohlke, H.; Goetz, AW.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, TS.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, KM.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, HT.; Omelyan, I.; Onufriev, A.; Roe, DR.; Roitberg, A.; Salomon-Ferrer, R.;

Simmerling, CL.; Smith, W.; Swails, J.; Walker, RC.; Wang, J.; Wolf, RM.; Wu, X.; York, DM.; Kollman, PA. Amber. Vol. 14. AMBER; 2015. 2015

de Winde CM, Veenbergen S, Young KH, Xu-Monette ZY, Wang XX, Xia Y, Jabbar KJ, van den Brand M, van der Schaaf A, Elfrink S, et al. Tetraspanin CD37 protects against the development of B cell lymphoma. J Clin Invest. 2016; 126:653–666. [PubMed: 26784544]

Deckert J, Park PU, Chicklas S, Yi Y, Li M, Lai KC, Mayo MF, Carrigan CN, Erickson HK, Pinkas J, et al. A novel anti-CD37 antibody-drug conjugate with multiple anti-tumor mechanisms for the treatment of B-cell malignancies. Blood. 2013; 122:3500–3510. [PubMed: 24002446]

Delaguillaumie A, Harriague J, Kohanna S, Bismuth G, Rubinstein E, Seigneuret M, Conjeaud H. Tetraspanin CD82 controls the association of cholesterol-dependent microdomains with the actin cytoskeleton in T lymphocytes: relevance to co-stimulation. Journal of cell science. 2004; 117:5269–5282. [PubMed: 15454569]

Dornier E, Coumailleau F, Ottavi JF, Moretti J, Boucheix C, Mauduit P, Schweisguth F, Rubinstein E. TspanC8 tetraspanins regulate ADAM10/Kuzbanian trafficking and promote Notch activation in flies and mammals. J Cell Biol. 2012; 199:481–496. [PubMed: 23091066]

Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

Garcia-Espana A, Chung PJ, Sarkar IN, Stiner E, Sun TT, Desalle R. Appearance of new tetraspanin genes during vertebrate evolution. Genomics. 2008; 91:326–334. [PubMed: 18291621]

Haining EJ, Yang J, Bailey RL, Khan K, Collier R, Tsai S, Watson SP, Frampton J, Garcia P, Tomlinson MG. The TspanC8 subgroup of tetraspanins interacts with A disintegrin and metalloprotease 10 (ADAM10) and regulates its maturation and cell surface expression. J Biol Chem. 2012; 287:39753–39765. [PubMed: 23035126]

Hemler ME. Tetraspanin functions and associated microdomains. Nature reviews Molecular cell biology. 2005; 6:801–811. [PubMed: 16314869]

Hemler ME. Tetraspanin proteins promote multiple cancer stages. Nature reviews Cancer. 2014; 14:49–60. [PubMed: 24505619]

Hermans J, Xia X, Zhang L, Cavanaugh D, Dowser D. DOWSER program. 2003

Hopf TA, Colwell LJ, Sheridan R, Rost B, Sander C, Marks DS. Three-dimensional structures of membrane proteins from genomic sequencing. Cell. 2012; 149:1607–1621. [PubMed: 22579045]

Horvath G, Serru V, Clay D, Billard M, Boucheix C, Rubinstein E. CD19 is linked to the integrin-associated tetraspans CD9, CD81, and CD82. The Journal of biological chemistry. 1998; 273:30537–30543. [PubMed: 9804823]

Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. Journal of computational chemistry. 2013; 34:2135–2145. [PubMed: 23832629]

Huang S, Yuan S, Dong M, Su J, Yu C, Shen Y, Xie X, Yu Y, Yu X, Chen S, et al. The phylogenetic analysis of tetraspanins projects the evolution of cell-cell interactions from unicellular to multicellular organisms. Genomics. 2005; 86:674–684. [PubMed: 16242907]

Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of molecular graphics. 1996; 14:33–38. 27–38. [PubMed: 8744570]

Jouannet S, Saint-Pol J, Fernandez L, Nguyen V, Charrin S, Boucheix C, Brou C, Milhiet PE, Rubinstein E. TspanC8 tetraspanins differentially regulate the cleavage of ADAM10 substrates, Notch activation and ADAM10 membrane compartmentalization. Cell Mol Life Sci. 2016; 73:1895–1915. [PubMed: 26686862]

Kabsch W. Xds. Acta crystallographica Section D. Biological crystallography. 2010; 66:125–132. [PubMed: 20124692]

Karamatic Crew V, Burton N, Kagan A, Green CA, Levene C, Flinter F, Brady RL, Daniels G, Anstee DJ. CD151, the first member of the tetraspanin (TM4) superfamily detected on erythrocytes, is essential for the correct assembly of human basement membranes in kidney and skin. Blood. 2004; 104:2217–2223. [PubMed: 15265795]

Kitadokoro K, Ponassi M, Galli G, Petracca R, Falugi F, Grandi G, Bolognesi M. Subunit association and conformational flexibility in the head subdomain of human CD81 large extracellular loop. Biological chemistry. 2002; 383:1447–1452. [PubMed: 12437138]

Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD Jr, Pastor RW. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. The journal of physical chemistry B. 2010; 114:7830–7843. [PubMed: 20496934]

Kruse AC, Manglik A, Kobilka BK, Weis WI. Applications of molecular replacement to G protein-coupled receptors. Acta crystallographica Section D. Biological crystallography. 2013; 69:2287–2292. [PubMed: 24189241]

Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic acids research. 2005; 33:W299–W302. [PubMed: 15980475]

Le Grand S, Götz AW, Walker RC. SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. Computer Physics Communications. 2013; 184:374–380.

Le Naour F, Rubinstein E, Jasmin C, Prenant M, Boucheix C. Severely reduced female fertility in CD9-deficient mice. Science. 2000; 287:319–321. [PubMed: 10634790]

Levy S, Shoham T. The tetraspanin web modulates immune-signalling complexes. Nature reviews Immunology. 2005; 5:136–148.

MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. The journal of physical chemistry B. 1998; 102:3586–3616. [PubMed: 24889800]

Marks DS, Hopf TA, Sander C. Protein structure prediction from sequence variation. Nature biotechnology. 2012; 30:1072–1080.

Mattila PK, Feest C, Depoil D, Treanor B, Montaner B, Otipoby KL, Carter R, Justement LB, Bruckbauer A, Batista FD. The actin and tetraspanin networks organize receptor nanoclusters to regulate B cell receptor-mediated signaling. Immunity. 2013; 38:461–474. [PubMed: 23499492]

Min G, Wang H, Sun TT, Kong XP. Structural basis for tetraspanin functions as revealed by the cryo-EM structure of uroplakin complexes at 6-A resolution. J Cell Biol. 2006; 173:975–983. [PubMed: 16785325]

Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, Sliz P. Collaboration gets the most out of software. Elife. 2013; 2:e01456. [PubMed: 24040512]

Noy PJ, Yang J, Reyat JS, Matthews AL, Charlton AE, Furmston J, Rogers DA, Rainger GE, Tomlinson MG. TspanC8 Tetraspanins and A Disintegrin and Metalloprotease 10 (ADAM10) Interact via Their Extracellular Regions: EVIDENCE FOR DISTINCT BINDING MECHANISMS FOR DIFFERENT TspanC8 PROTEINS. J Biol Chem. 2016; 291:3145–3157. [PubMed: 26668317]

Oren R, Takahashi S, Doss C, Levy R, Levy S. TAPA-1, the target of an antiproliferative antibody, defines a new family of transmembrane proteins. Molecular and cellular biology. 1990; 10:4007–4015. [PubMed: 1695320]

Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, Weiner AJ, Houghton M, Rosa D, Grandi G, et al. Binding of hepatitis C virus to CD81. Science. 1998; 282:938–941. [PubMed: 9794763]

Roe DR, Cheatham TE 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. Journal of chemical theory and computation. 2013; 9:3084–3095. [PubMed: 26583988]

Rubinstein, E.; Charrin, S.; Tomlinson, M. Organisation of the Tetraspanin Web. In: Berditchevski, F.; Rubinstein, E., editors. Tetraspanins. Netherlands: Springer; 2013. p. 47-90.

Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. Journal of chemical theory and computation. 2013; 9:3878–3888. [PubMed: 26592383]

Schroder J, Lullmann-Rauch R, Himmerkus N, Pleines I, Nieswandt B, Orinska Z, Koch-Nolte F, Schroder B, Bleich M, Saftig P. Deficiency of the tetraspanin CD63 associated with kidney pathology but normal lysosomal function. Molecular and cellular biology. 2009; 29:1083–1094. [PubMed: 19075008]

Schrodinger L. The PyMOL Molecular Graphics System, Version 1.3r1. 2010

Seigneuret M. Complete predicted three-dimensional structure of the facilitator transmembrane protein and hepatitis C virus receptor CD81: conserved and variable structural domains in the tetraspanin superfamily. Biophys J. 2006; 90:212–227. [PubMed: 16352525]

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. 2011; 7:539. [PubMed: 21988835]

Silvie O, Charrin S, Billard M, Franetich JF, Clark KL, van Gemert GJ, Sauerwein RW, Dautry F, Boucheix C, Mazier D, et al. Cholesterol contributes to the organization of tetraspanin-enriched microdomains and to CD81-dependent infection by malaria sporozoites. Journal of cell science. 2006; 119:1992–2002. [PubMed: 16687736]

Stipp CS, Kolesnikova TV, Hemler ME. Functional domains in tetraspanin proteins. Trends in biochemical sciences. 2003; 28:106–112. [PubMed: 12575999]

Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, Eisenberg D. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103:8060–8065. [PubMed: 16690741]

van Spriel AB, de Keijzer S, van der Schaaf A, Gartlan KH, Sofi M, Light A, Linssen PC, Boezeman JB, Zuidscherwoude M, Reinieren-Beeren I, et al. The tetraspanin CD37 orchestrates the alpha(4)beta(1) integrin-Akt signaling axis and supports long-lived plasma cell survival. Sci Signal. 2012; 5:ra82. [PubMed: 23150881]

van Zelm MC, Smet J, Adams B, Mascart F, Schandene L, Janssen F, Ferster A, Kuo CC, Levy S, van Dongen JJ, et al. CD81 gene defect in humans disrupts CD19 complex formation and leads to antibody deficiency. J Clin Invest. 2010; 120:1265–1274. [PubMed: 20237408]

Vences-Catalan F, Kuo CC, Sagi Y, Chen H, Kela-Madar N, van Zelm MC, van Dongen JJ, Levy S. A mutation in the human tetraspanin CD81 gene is expressed as a truncated protein but does not enable CD19 maturation and cell surface expression. J Clin Immunol. 2015; 35:254–263. [PubMed: 25739915]

Wee JL, Schulze KE, Jones EL, Yeung L, Cheng Q, Pereira CF, Costin A, Ramm G, van Spriel AB, Hickey MJ, et al. Tetraspanin CD37 Regulates beta2 Integrin-Mediated Adhesion and Migration in Neutrophils. Journal of immunology. 2015; 195:5770–5779.

Wright MD, Geary SM, Fitter S, Moseley GW, Lau LM, Sheng KC, Apostolopoulos V, Stanley EG, Jackson DE, Ashman LK. Characterization of mice lacking the tetraspanin superfamily member CD151. Molecular and cellular biology. 2004; 24:5978–5988. [PubMed: 15199151]

Yanez-Mo M, Barreiro O, Gordon-Alonso M, Sala-Valdes M, Sanchez-Madrid F. Tetraspanin-enriched microdomains: a functional unit in cell plasma membranes. Trends in cell biology. 2009; 19:434–446. [PubMed: 19709882]

Yang X, Claas C, Kraeft SK, Chen LB, Wang Z, Kreidberg JA, Hemler ME. Palmitoylation of tetraspanin proteins: modulation of CD151 lateral interactions, subcellular distribution, and integrin-dependent cell morphology. Molecular biology of the cell. 2002; 13:767–781. [PubMed: 11907260]

Yauch RL, Berditchevski F, Harler MB, Reichner J, Hemler ME. Highly stoichiometric, stable, and specific association of integrin alpha3beta1 with CD151 provides a major link to phosphatidylinositol 4-kinase, and may regulate cell migration. Mol Biol Cell. 1998; 9:2751–2765. [PubMed: 9763442]

Zhang L, Hermans J. Hydrophilicity of cavities in proteins. Proteins. 1996; 24:433–438. [PubMed: 9162944]

Zhao X, Lapalombella R, Joshi T, Cheney C, Gowda A, Hayden-Ledbetter MS, Baum PR, Lin TS, Jarjoura D, Lehman A, et al. Targeting CD37-positive lymphoid malignancies with a novel engineered small modular immunopharmaceutical. Blood. 2007; 110:2569–2577. [PubMed: 17440052]

Zuidscherwoude M, Gottfert F, Dunlock VM, Figdor CG, van den Bogaart G, van Spriel AB. The tetraspanin web revisited by super-resolution microscopy. Sci Rep. 2015; 5:12201. [PubMed: 26183063]

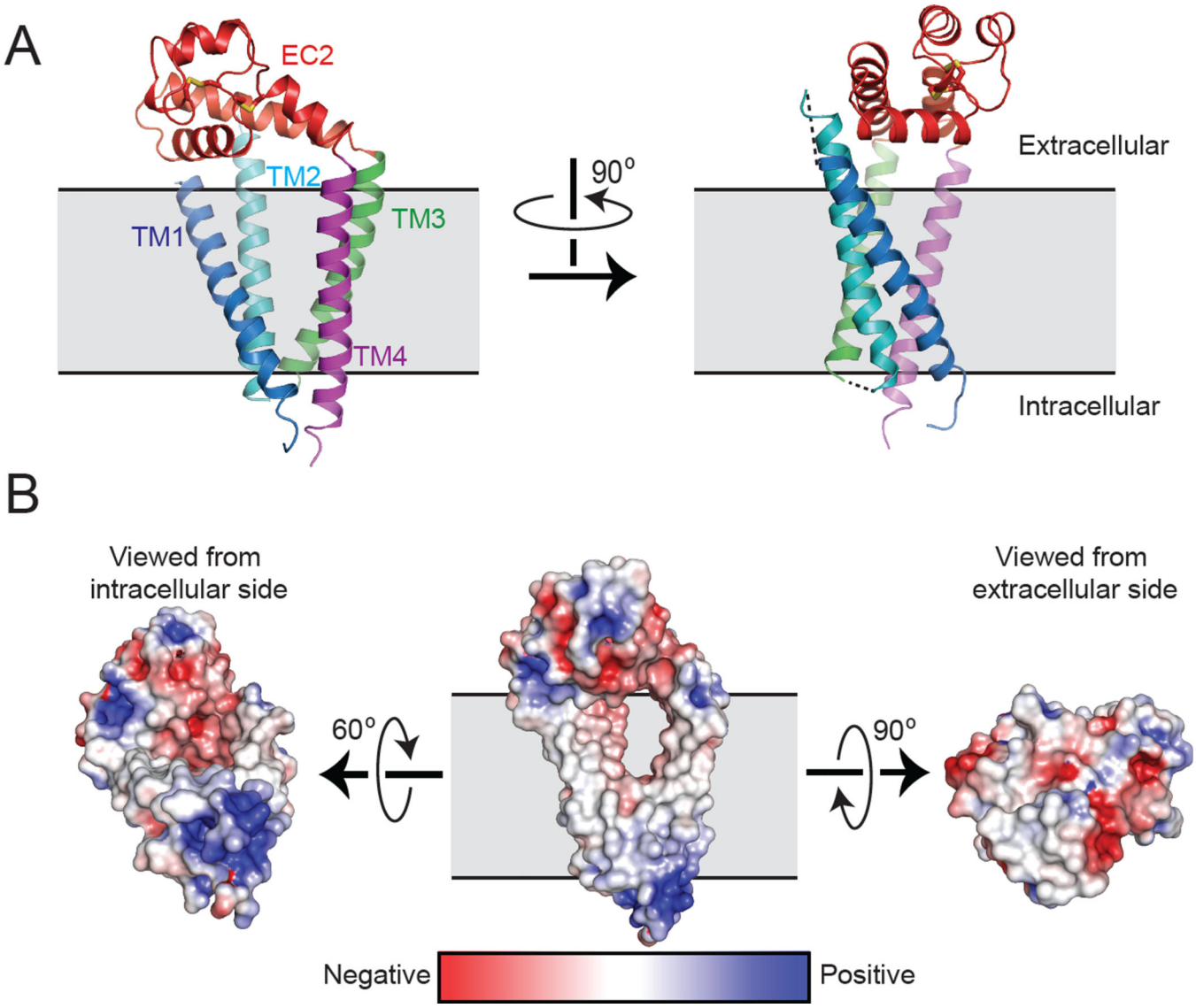


Figure 1.

Overall structure of human CD81. (A) Cartoon representation viewed parallel to the membrane plane. Helix one (TM1) is blue, helix two (TM2) cyan, helix three (TM3) green and helix four (TM4), magenta. The large extracellular region (EC2) between TM3 and TM4 is red. (B) Surface representation, colored by electrostatic surface potential on a sliding scale from blue (basic) to red (acidic). See also Figures S1, S2 and S6.

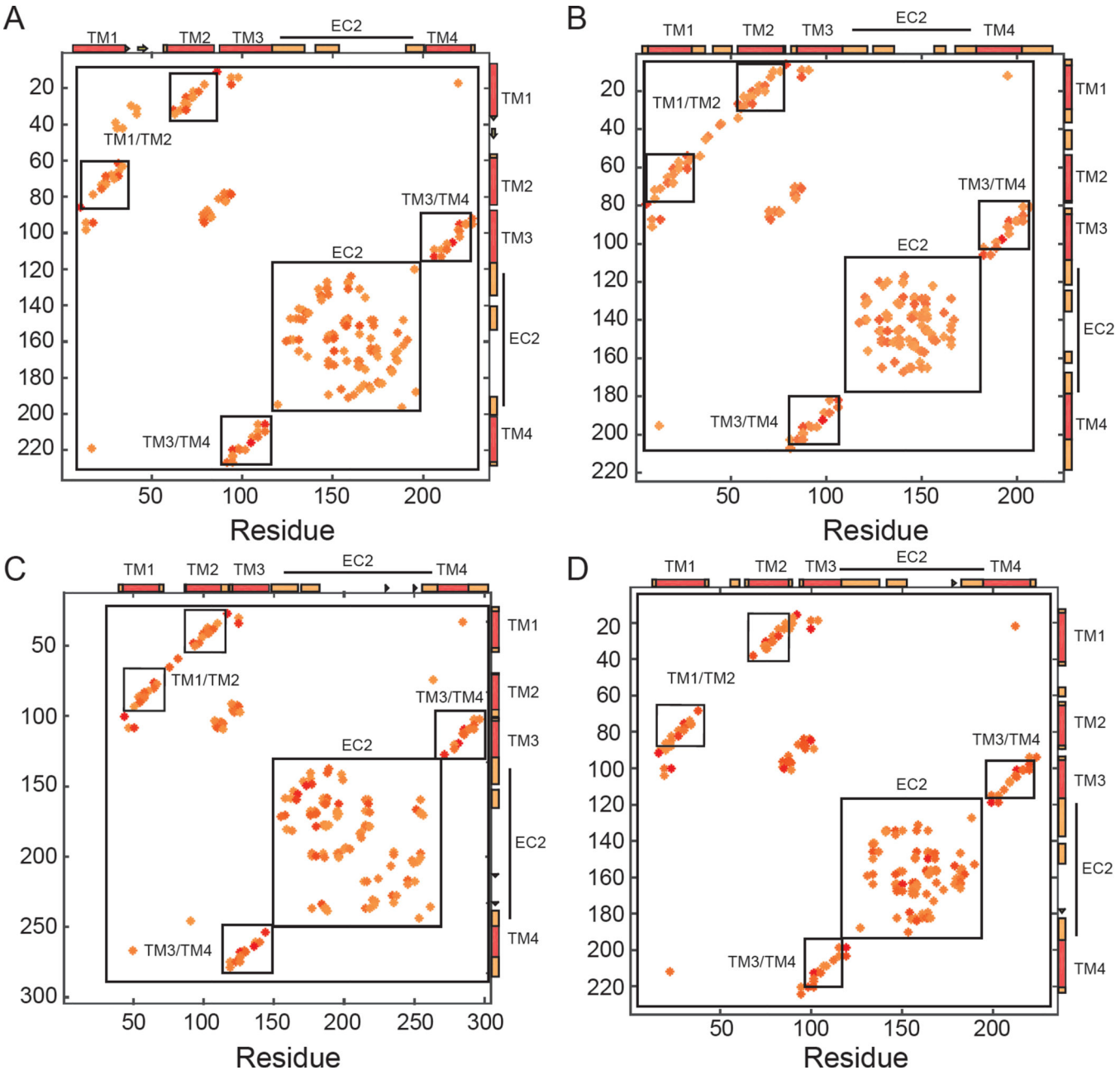


Figure 2.

Evolutionary coupling map of tetraspanins. The top 90 amino acid evolutionary coupling pairs of (A) human CD81, (B) hypothetical protein FGSG_08695 from Fusarium graminearum PH-1, (C) tetraspanin 3A from Drosophila melanogaster, and (D) AX4 tetraspanin family protein from Dictyostelium discoideum. Hot spots include couplings between residues of TM1 and TM2, TM3 and TM4, the junction between TM2 and TM3, and intradomain coupling within EC2. Analysis performed using the EVFold server (http://www.evfold.org). See also Figure S2.

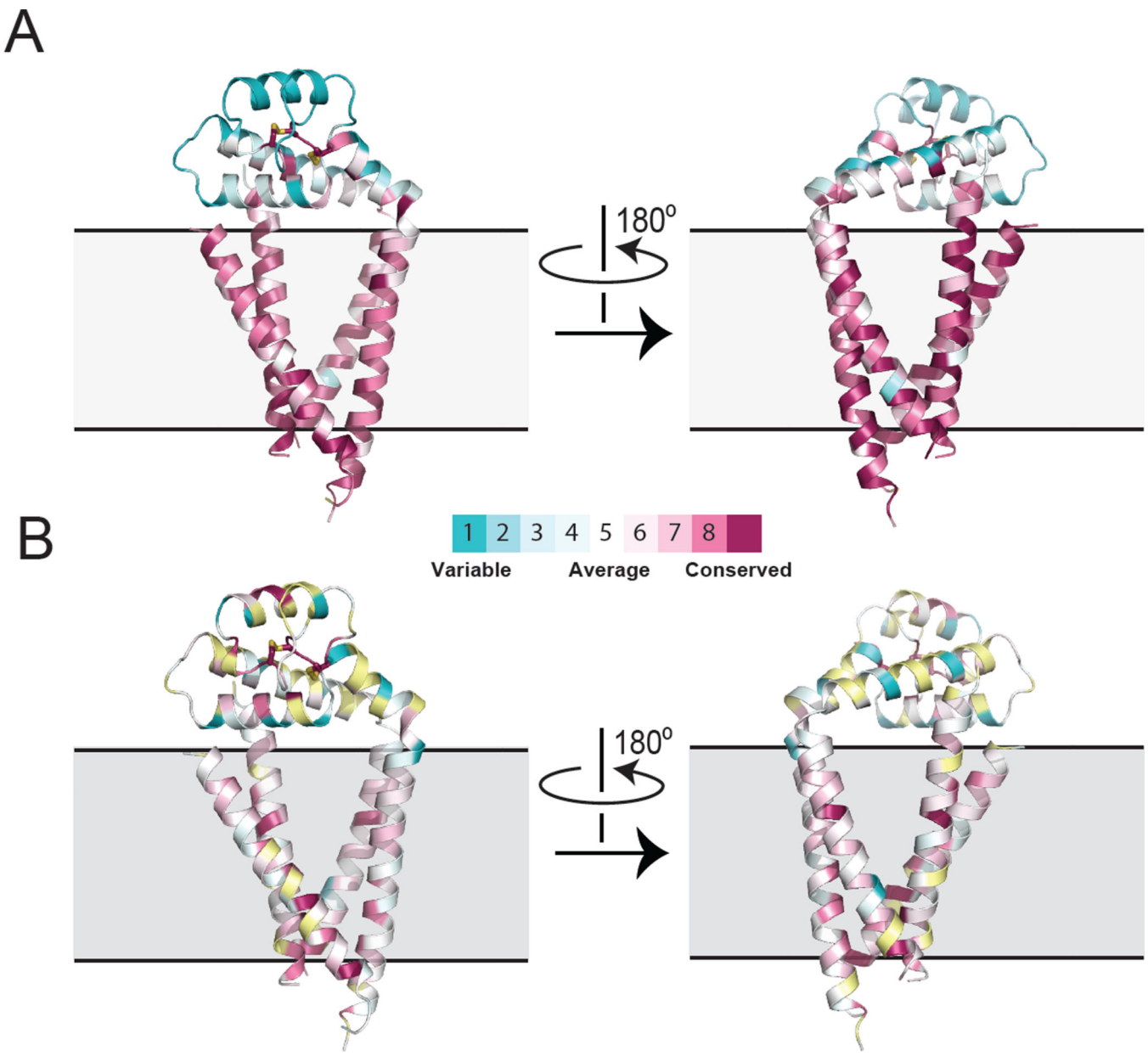


Figure 3.

Sequence conservation of CD81. Cartoon representation of CD81 versus the top 50 CD81-related sequences determined by Consurf (A) or CD81 versus the 32 human tetraspanin paralogs (B) colored on a sliding scale from teal (poorly conserved) to maroon (highly conserved) (Landau et al., 2005). Residues with insufficient information for analysis are yellow. The high degree of conservation of the transmembrane region contrasts with the high divergence at the surface of the extracellular domain. The large pocket within the membrane bounded by the ectodomain and the TM helices is almost 3300 $Å^3$ in volume. See also Figure S3.

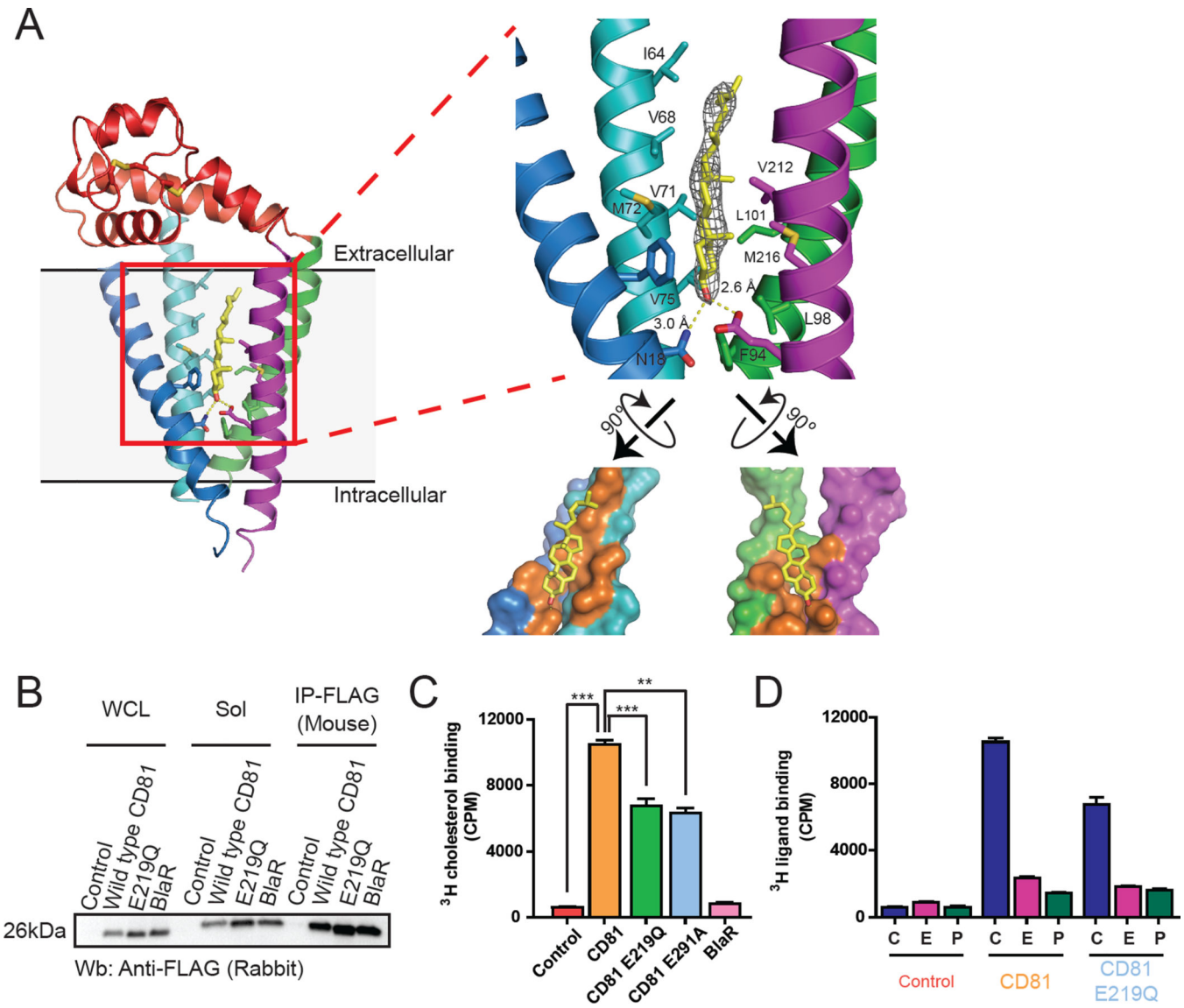


Figure 4.

CD81 binds cholesterol within its intramembrane cavity. (A) CD81-cholesterol interactions. CD81 residues within 4 Å of the bound cholesterol molecule are rendered as sticks, and labeled in the zoomed in view (right). An Fo-Fc omit map of electron density contoured at 2.0 σ is shown for the bound cholesterol. N18 and E219 coordinate the cholesterol hydroxyl group. Views of the pocket in surface representation are shown in open-book form projecting onto the TM1/2 bundle (left) and the TM3/4 bundle (right). CD81 residues at the ligand interface are colored orange. (B) CD81 wild type, CD81 mutant proteins (E219A, E219Q) and BlaR were immunopurified from HEK293T cells. Proteins captured on FLAG beads were used for radioactive cholesterol binding experiments. WCL, whole cell lysate, Sol, Solubilized protein, IP, immunoprecipitation. (C) Cholesterol binding by immunopurified proteins. 1,2-^{3}H-cholesterol was incubated with immunopurified FLAG-CD81 or FLAG-BlaR from 293T cells and bound cholesterol was measured. The figure represents three independent experiments performed in duplicate. Statistical analysis was performed using

ANOVA and a Bonferroni post-hoc test was performed comparing all columns. **, $p<0.01$, ***, $p<0.001$. (D) Specificity of lipid binding to CD81. Immunopurified FLAG-CD81 from 293T cells prepared as in (C) was incubated with 1,2-^{3}H-cholesterol [C], 2,4,6,7-^{3}H(N)-estradiol [E] and 9,10-^{3}H(N)-palmitic acid [P]. Bound ^{3}H-lipid was measured in a scintillation counter. The figure represents three independent experiments performed in duplicate. See also Figure S4.

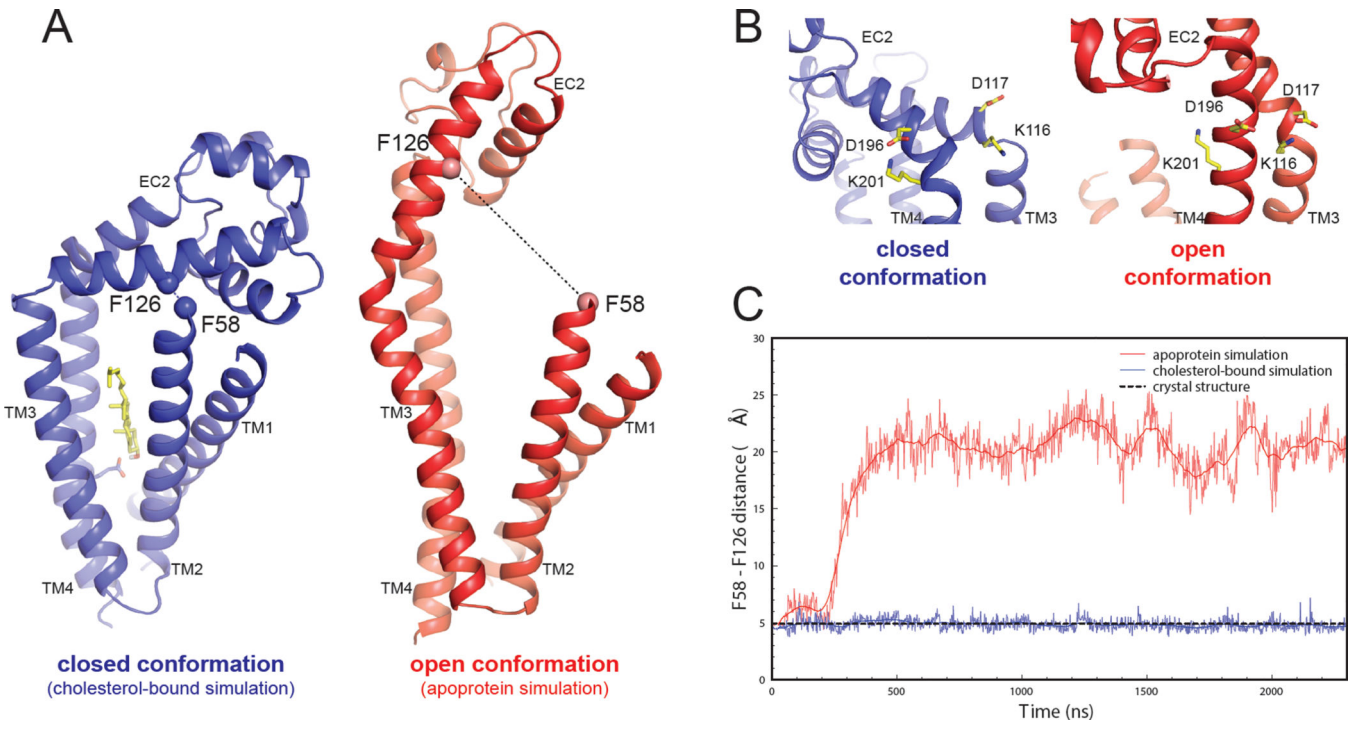


Figure 5.

Molecular dynamics simulations reveal an “open” conformation in which EC2 separates from the transmembrane domain. **A**, Closed (blue) and open (red) states of CD81, as observed in cholesterol-bound and apoprotein simulations, respectively. The open conformation is characterized by substantial domain separation and straightening of TM helices 3 and 4. **B**, The salt bridge from EC2 to TM4 (D196 – K201) stabilizes the closed conformation and breaks during opening, while a new electrostatic interaction K116 – D117 helps stabilize the open conformation. **C**, Interdomain distance (measured between alpha carbons of F58 and F126) as a function of time for an apoprotein simulation in which the domains separate and a cholesterolbound simulation in which they do not. Thin traces show values every 1 ns, and thick traces are smoothed. See also Figure S5.

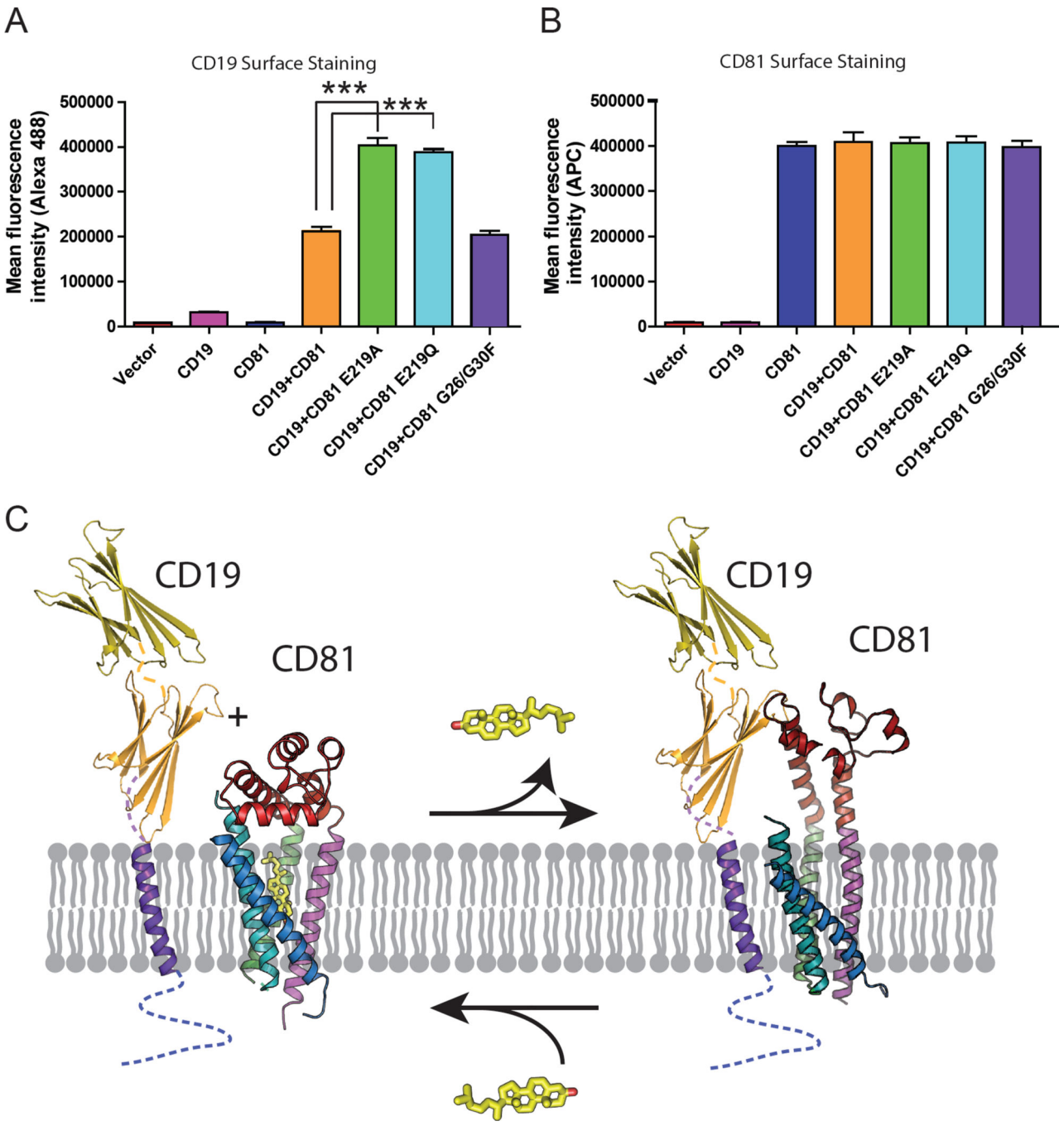


Figure 6.

Cholesterol binding regulates CD81-mediated export of CD19. HEK293T cells were transfected with cDNA encoding the indicated proteins and cell surface amounts of CD19 (A) and CD81 (B) were assessed by flow cytometry. Histograms shown represent four independent experiments done in triplicate. Statistical analysis was performed using ANOVA and a Bonferroni post-hoc test was performed comparing all columns. ***, $p<0.01$. (C) Proposed model for modulation of cargo binding in response to cholesterol. CD81 favors a closed conformation when cholesterol is bound (left), and more readily accesses an open

conformation, which allows more efficient export of its cargo CD19 (modeled in cartoon form), when not bound to cholesterol (right).

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Alexa Fluor® 488 anti-human CD19 Antibody	Biolegend	302219
APC anti-human CD81 (TAPA-1) Antibody	Biolegend	349510
ANTI-FLAG® M2 Affinity Gel	Sigma Aldrich	A2220
Chemicals, Peptides, and Recombinant Proteins		
[1,2-3H]-cholesterol	Perkin Elmer	NET139250UC
2,4,6,7-3H(N)-estradiol	Perkin Elmer	NET013250UC
9,10-3H(N)-palmitic acid	Perkin Elmer	NET043001MC
Benzonase recombinant nuclease	Sigma Aldrich	E1014
Iodoacetamide	Sigma Aldrich	I1149
In-Fusion® HD Cloning Plus	Clontech	638911
ESF 921 Insect Cell Culture Medium, Protein-Free	Expression systems	96-001
BestBac 2.0, v-cath/chiA Deleted Linearized Baculovirus DNA	Expression systems	91-002
Fugene HD	Promega	E2311
Lipofectamine 2000	Invitrogen	11668019
n-Dodecyl-β-D-Maltoside (Anagrade)	Anatrace	D310
5-Choesten-3β-ol Hemisuccinate	Steraloids	C6823-000
Monoolein	Hampton Research	HR2-435
Cholesterol	Sigma	C8667
PEG300	Hampton Research	HR2-517
Deposited Data		
CD81 Structure	This paper	PDB: 5TCX
Experimental Models: Cell Lines		
293T (ATCC® CRL-3216™)	ATCC	CRL-3216
Spodoptera frugiperda ovarian tissue (SF9)	Expression systems	94-001S
Recombinant DNA		
pVL1392	Expression systems	91-030
pCDNA3.1 (+) hygro	Invitrogen	V87020
gBlocks	Integrated DNA Technologies	N/A
Software and Algorithms		
GraphPad Prism 7.0		http://www.graphpad.com/scientific-software/prism/
EVFold conservation server	Hopf *et al.* Cell. 2012	http://evfold.org/evfold-web/evfold.do
BD Accuri C6 software	BD Accuri C6	https://www.bdbiosciences.com/instruments/accuri/features/software.jsp
UCLA-DOE LAB — Diffraction Anisotropy Server	Strong *et al.* PNAS. 2006.	https://services.mbi.ucla.edu/anisoscale/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SBGrid Consortium	Morin et al., 2013	https://sbgrid.org/software/
XDS	Kabsch, 2010	https://sbgrid.org/software/
Phenix (1.10_2155)	Afonine et al., 2012	https://sbgrid.org/software/
Coot	Emsley and Cowtan, 2004	https://sbgrid.org/software/
Pymol	Schrodinger, 2010	https://sbgrid.org/software/