

CD81 extracted in SMALP nanodiscs comprises two distinct protein populations within a lipid environment enriched with negatively charged headgroups

Hoor Ayub, Michelle Clare, Ivana Milic, Nikola P. Chmel, Heike Böning, Andrew Devitt, Thomas Krey, Roslyn M. Bill, Alice J. Rothnie



PII: S0005-2736(20)30262-5

DOI: <https://doi.org/10.1016/j.bbamem.2020.183419>

Reference: BBAMEM 183419

To appear in: *BBA - Biomembranes*

Received date: 19 December 2019

Revised date: 5 June 2020

Accepted date: 7 July 2020

Please cite this article as: H. Ayub, M. Clare, I. Milic, et al., CD81 extracted in SMALP nanodiscs comprises two distinct protein populations within a lipid environment enriched with negatively charged headgroups, *BBA - Biomembranes* (2020), <https://doi.org/10.1016/j.bbamem.2020.183419>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

CD81 extracted in SMALP nanodiscs comprises two distinct protein populations within a lipid environment enriched with negatively charged headgroups

Hoor Ayub^a, Michelle Clare^a, Ivana Milic^a, Nikola P. Chmel^b, Heike Böning^c, Andrew Devitt^a, Thomas Krey^{c,d,e}, Roslyn M. Bill^a & Alice J. Rothnie^a

^a Life & Health Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK

^b Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

^c German Center for Infection Research (DZIF) Hannover-Braunschweig Site and Institute of Virology, Hannover Medical School, 30625 Hannover, Germany

^d Center of Structural and Cell Biology in Medicine, Institute of Biochemistry, University of Luebeck, Luebeck, Germany

^e Centre for Structural Systems Biology (CSSB), Hamburg, Germany

Corresponding authors: Dr Alice J. Rothnie & Professor M. Roslyn Bill, School of Life & Health Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK. a.rothnie@aston.ac.uk, r.m.bill@aston.ac.uk

Abbreviations:

CD	circular dichroism
DLS	dynamic light scattering
DDM	dodecylmaltoside
EC2	extracellular loop 2
ELISA	enzyme linked immune absorbant assay
EM	electron microscopy
HCV	hepatitis C virus
HRP	horseradish peroxidase
LEL	large extracellular loop
mAb	monoclonal antibody
MS	mass spectrometry
Ni-NTA	nickel nitrilotriacetic acid
OG	octyl glucoside
OPD	<i>o</i> -Phenylenediamine dihydrochloride
PBS	phosphate-buffered saline
PBST	phosphate-buffered saline + Tween 20 (0.05% v/v)
<i>P. pastoris</i>	<i>Pichia pastoris</i>
SEL	small extracellular loop
SD	Standard deviation
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SEM	Standard error of the mean
SMA	styrene maleic acid co-polymer
SMALP	styrene maleic acid lipid particle

HIGHLIGHTS

- CD81 expressed in *Pichia pastoris* can be solubilised and purified using SMA polymer.
- SMALP-encapsulated CD81 retains native folded structure.
- Expression and buffer conditions can be optimized to improve protein quality.
- The lipid environment surrounding CD81 is enriched with negatively charged lipids.

Journal Pre-proof

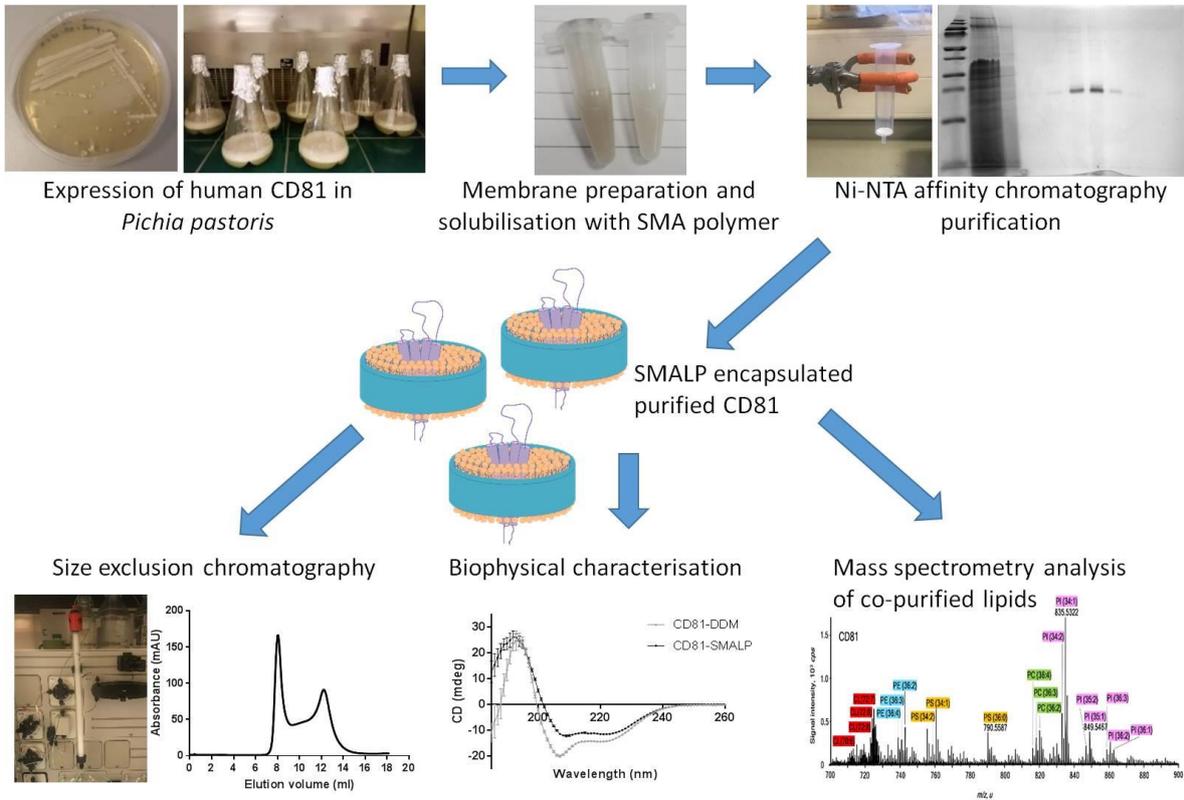
ABSTRACT

Tetraspanins exert a wide range of cellular functions of broad medical importance. Despite this, their biophysical characteristics are incompletely understood. Only two high-resolution structures of full-length tetraspanins have been solved. One is that of human CD81, which is involved in the infectivity of human pathogens including influenza, HIV, the malarial plasmodium parasite and hepatitis C virus (HCV). The CD81 crystal structure identifies a cholesterol-binding pocket, which has been suggested to be important in the regulation of tetraspanin function. Here we investigate the use of styrene-maleic anhydride co-polymers (SMA) for the solubilisation and purification of CD81 within a lipid environment. When CD81 was expressed in the yeast *Pichia pastoris*, it could be solubilised and purified using SMA2000. This SMALP-encapsulated CD81 retained its native folded structure, as determined by the binding of two conformation-sensitive anti-CD81 antibodies. Analysis by size exclusion chromatography revealed two distinct populations of CD81, only one of which bound the HCV glycoprotein, E2. Optimization of expression and buffer conditions increased the proportion of E2-binding competent CD81 protein. Mass spectrometry analysis indicated that the lipid environment surrounding CD81 is enriched with negatively charged lipids. These results establish a platform to study the influence of protein-lipid interactions in tetraspanin biology.

KEYWORDS

Tetraspanin, SMALP, membrane protein, solubilisation, purification, stability.

GRAPHICAL ABSTRACT



1. INTRODUCTION

CD81 is a member of the tetraspanin family of membrane proteins that have broad biological and medical importance. Tetraspanins engage in a wide range of incompletely-understood molecular interactions, exerting biological functions in cell-cell adhesion, cell proliferation, the immune system and during fertilization. They coordinate the trafficking of molecules into tetraspanin-enriched membrane microdomains thereby affecting cell signaling, morphology, motility and fusion [1]. Several tetraspanins have well-documented roles in infectious disease processes [2], while the expression of others correlates with tumour stage, type and patient outcome by affecting cell growth, morphology, invasion and metastasis [3].

There are 34 tetraspanin family members in mammals, of which 33 have been identified in humans [1]. Tetraspanin proteins have four transmembrane domains (TM1-4), intracellular N- and C-termini and two extracellular domains, one small (known as EC1 or SEL) and one large (typically 100 residues; known as EC2 or LEL). The family's signature motif is the presence of four or more cysteine residues in LEL; two in highly-conserved 'CCG' motifs [1]. Central to tetraspanin function are their extensive associations with an array of signalling proteins [1]; tetraspanin homodimers have been proposed to be the building blocks for the assembly of multicomponent tetraspanin protein complexes [4, 5]. The 6 Å cryo-electron microscopy structure of a uroplakin tetraspanin revealed a rod-shaped structural morphology consisting of four TM helical bundles bound to a single TM helix partner [6].

Only two high-resolution structures are available of full-length tetraspanins: human CD81 [7] and human CD9 [8]. The 2.95 Å crystal structure of CD81 confirms the presence of the conserved CCG motif and two disulfide bridges [7]. CD81 is involved in cell-cell adhesion, cell proliferation, the immune response, fertilization and the infectivity of several important human pathogens including influenza, HIV, the malarial plasmodium parasite and hepatitis C virus (HCV) [9]. It forms diverse, associations with other proteins in cell membranes: the oligomeric status of CD81 within these complexes and the mechanistic detail of how they exert their biological function remain incompletely understood. The CD81 crystal structure identifies a cholesterol-binding pocket; cholesterol binding appears to modulate CD81 activity in cells and has been suggested as a potential mechanism for the regulation of tetraspanin function [7] prompting a need to understand the role of the membrane lipid environment in CD81 folding and function.

Biological membranes are complex mixtures of membrane proteins and lipids. To study the function of an individual membrane protein often requires purification of that protein, however purification of membrane proteins can be challenging. This is largely due to their location within the lipid bilayer. First the protein needs to be extracted from the bilayer, and then the hydrophobic regions of the protein, which would normally interact with lipid acyl chains, need to be stabilised in some way to allow the protein to be soluble in water and not just aggregate. Typically this has been achieved with the use of mild detergents, such as octyl glucoside (OG) or dodecylmaltoside (DDM). These detergents

disrupt the membrane bilayer and then form a micelle around the hydrophobic parts of the protein. Detergents have proven successful for the study of many proteins, including the X-ray crystal structure of CD81 [7]. However one of the downsides of detergents is that during solubilisation they strip away the native lipid environment of a membrane protein. It is well established that specific lipid environments are important not just for stability of many proteins but also for their function [10-12], including CD81. In recent years an alternative approach to detergents has been developed using styrene maleic acid co-polymer (SMA), which extracts membrane proteins with their surrounding lipids [13-15]. The SMA polymer inserts into a membrane and extracts a disc of lipid bilayer with the polymer wrapping around the outer edge [16]. This nanodisc-like structure is termed an SMA lipid particle (SMALP), and the approach allows for the extraction and purification of membrane proteins without ever removing them from their natural lipid environment [13, 17-19]. Given the proposed importance of lipids for the function of CD81, in this study we aimed to investigate the use of SMA polymers for the solubilisation and purification of the tetraspanin CD81.

2. MATERIALS & METHODS

2.1 MATERIALS

Pichia pastoris wild type strain X-33 (Invitrogen) was electroporated to transform with hCD81 expression plasmid pPICZB containing recombinant and C-terminus His₆ tagged CD81 [20]. SMA 2000 co-polymer (2:1 Cray Valley) and SZ25010 co-polymer (3:1 Polyscope) were obtained as styrene maleic anhydride and hydrolysed to styrene maleic acid forms as described previously [21]. Detergent n-dodecyl-B-D-maltoside (DDM), biotin, zeocin, yeast extract, peptone, yeast nitrogen base and agar antibiotic were purchased from ThermoFisher Scientific. Ni²⁺-NTA resin for IMAC purification was purchased from Qiagen and Bio-Rad Econo-column empty chromatography column was used. Mini EDTA-free Protease Inhibitor Cocktail Tablets from Roche Applied Science. Anti-mouse HRP-conjugated secondary antibody was purchased from Cell Signaling Technology. Ammonium formate was purchased from Sigma Aldrich. Methanol, chloroform and ultra pure water, all LC-MS grade, were purchased from Fisher, UK. Anti-CD81 antibodies 2s131, 1s337 and 1s135 were generated as described previously [22] and hybridoma supernatants were used for all ELISA experiments.

2.2 EXPRESSION OF CD81

P. pastoris transformed with CD81 were streaked on to YPD-Zeocin agar plates (1%(w/v) yeast extract, 2%(w/v) peptone, 2%(w/v) glucose, 2%(w/v) agar, 100 mg/ml Zeocin) from a -80°C glycerol stock. The plates were incubated at 30 °C for three days or until single colonies were visible.

For large scale recombinant CD81 expression, large scale yeast growth was performed using large baffled shake flasks. One colony of *P. pastoris* from YPD agar plates was added to a 250 ml baffled flask containing 50 ml BMGY (1%(w/v) yeast extract; 2%(w/v) peptone, 100 mM potassium phosphate pH 6.0, 1.34%(w/v) YNB, 4 x 10⁻⁵%(w/v) biotin, 1%(v/v) glycerol) and incubated at 30 °C and 220 rpm for overnight growth. 5 ml of this seed culture were added into 200 ml of fresh BMGY media in separate 1 L baffled flask and grown for 22 - 24 h at 30 °C and 220 rpm. To induce recombinant CD81 expression, cell density was monitored until an OD₆₀₀ of 1 was reached. Cells were then pelleted by centrifugation at 5000 g for 10 min. The cells were resuspended in 500 ml of BMMY media (1%(w/v) yeast extract; 2%(w/v) peptone, 100 mM potassium phosphate pH 6.0, 1.34%(w/v) YNB, 4 x 10⁻⁵%(w/v) biotin, 0.5%(v/v) methanol) in a 2 L baffled shake flask. Culture was grown for 22 h, supplemented with 5 ml of absolute methanol and then grown for a further 22 h. Cells were harvested by centrifugation (5000 g for 20 min at 4 °C) and stored at -80 °C until further use.

2.3 MEMBRANE PREPARATION

The cells from large scale *P. pastoris* growth were lysed using an Emulsiflex-C3 cell disrupter (Avestin) to isolate membranes. The cell pellet was mixed with ice-cold breaking buffer (5%(v/v) glycerol, 2 mM EDTA, 100 mM NaCl, 50 mM NaH₂PO₄, 50 mM Na₂HPO₄, pH 7.4) supplemented with protease inhibitors, at a ratio of 3ml buffer per gram of cell pellet. The cells were passed through the cell lyser 5 times (for ≈15 min) at 20,000 to 25,000 psi. Any unbroken cells and cell debris were removed by centrifugation (10,000 g, for 30 min at 4 °C). The supernatant was collected and ultracentrifuged (100,000 g, for 1 h at 4 °C). The pellet containing yeast membranes was collected, resuspended in buffer A (20mM HEPES, 50mM sodium chloride, 10%(v/v) glycerol, pH 7.0) at a concentration of 160 mg/ml (wet pellet weight) and homogenised using a glass homogeniser. Membrane fractions were stored at 4 °C for immediate use or at -80 °C for future analysis.

2.4 SOLUBILISATION & PURIFICATION

For initial solubilisation trials CD81 expressing *P. pastoris* membranes, at 160 mg/ml (wet pellet weight) were diluted four fold with Tris-HCl purification buffer(20 mM Tris-HCl, 150 mM sodium chloride, pH 8.0), and incubated with varying concentrations of SMA 2000 polymer (0-5% w/v), SZ25010 polymer (2.5% w/v) or DDM (2% w/v), for time points ranging from 15 min to 4 hours at room temperature, with gentle shaking. Turbidity was monitored by measuring the OD₆₀₀ of the sample. Insoluble material was sedimented by ultracentrifugation (100,000 g, 20 min at 4 °C) to yield supernatant containing solubilised CD81. Samples of the solubilised protein within the supernatant and insoluble protein within the pellet were analysed by Western blotting, using an anti-CD81 primary antibody (2s131) at a dilution of 1:200, followed by an anti-mouse HRP-conjugated secondary antibody. Importantly no reducing agents were used in the sample buffer as this has been shown to prevent binding of the antibody [22].

For downstream purification CD81 expressing *P. pastoris* membranes, at 160 mg/ml (wet pellet weight) were diluted four fold with either Tris-HCl purification buffer(20 mM Tris-HCl, 150 mM sodium chloride, pH 8.0) or HEPES purification buffer (20 mM HEPES, 200 mM sodium chloride, 10% glycerol, pH 8.0) and incubated with 2.5%(w/v) of the SMA polymer for 1 h at room temperature with mild agitation, before harvesting the solubilized protein by ultracentrifugation (100,000 g, 20 min at 4 °C).

The solubilised protein was incubated with Ni-NTA agarose resin overnight at 4°C with mild agitation. Initial trials utilized 100µl resin per ml of solubilised protein, whilst optimized conditions increased the volume of resin 5-fold. The solution was poured into an empty column and the flow-through was collected. Resin was washed with 50 bed volumes of purification buffer containing 20 mM imidazole, a second wash with 20 bed volumes of purification buffer containing 40 mM imidazole and a final wash with 1 bed volume of purification buffer containing 60 mM imidazole. The bound CD81 was eluted with 3 bed volumes of purification buffer containing 300 mM imidazole. The elution fractions were

pooled and concentrated using a centrifugal concentrator (Vivaspin 20 kDa cut-off, Sartorius).

Alternatively CD81 membranes were solubilised with 1% (w/v) DDM and 0.1% cholesterol hemisuccinate in 250 mM NaCl, 20 mM HEPES pH 7.4, 20 mM imidazole, supplemented with 10% (v/v) glycerol, at 4°C for 2 hours before ultracentrifugation (100,000 g, 60 min at 4 °C). DDM solubilised protein was mixed with Ni-NTA resin (at a ratio of 1:20) for 2 hours at 4°C. The solution was poured into a chromatography column and the flow-through was collected. Resin was washed with 20 column volumes of buffer containing 100 mM NaCl, 20mM HEPES pH 7.4, 1% glycerol, 0.1% (w/v) DDM, 0.01% CHS and 40mM Imidazole. The bound CD81 was eluted with 5 column volumes using the same buffer but with 300 mM imidazole. The elution fractions were pooled and concentrated using a centrifugal concentrator (Vivaspin 50 kDa cut-off, Sartorius).

Purified protein was quantified using an SDS-PAGE based assay with bovine serum albumin as a standard [23].

Purified CD81-SMALP was further analysed and purified through size exclusion chromatography (SEC) using AKTA Pure system (GE Healthcare) using a Superdex increase 200 10/300 GL SEC column.

2.5 EXPRESSION & PURIFICATION OF HCV E2 GLYCOPROTEIN

A full-length streptavidin-tagged HCV E2 ectodomain (sE2) of the clinical isolate UKN2b_2.8 was expressed in *Drosophila* S2 cells as previously described[24, 25]. Briefly, *Drosophila* S2 cells were transfected, amplified, and induced with 4 μ M CdCl₂ at a density of $\sim 8 \times 10^6$ cells/ml for 6-9 days for large-scale production. Proteins were purified from the supernatant by affinity chromatography using a Strep-Tactin Superflow column (IBA, Goettingen, Germany) followed by size exclusion chromatography (SEC) using a Superdex200 column (GE Healthcare, Uppsala, Sweden). Pure monomeric protein was concentrated to ~ 15 mg/ml for use in ELISA.

2.6 ELISA ASSAYS

For anti-CD81 binding ELISA, 50 μ l of purified protein (100 μ g/ml) were added to a 96-well plate (Immulon II ELISA plate Bunc) in triplicate and incubated overnight at 4 °C. Unbound protein was removed by washing three times with PBS (200 μ l/well). Plates were blocked with 100 μ l of 2%(w/v) bovine serum albumin in PBS per well for 20 min at room temperature to reduce any nonspecific binding and washed three times with PBS. Samples were incubated with 50 μ l primary anti-CD81 antibodies 1s337, 1s135 or 2s337 [22], diluted 1:2 in PBS-Tween (0.05% v/v). After three PBS washes (200 μ l/well), samples were incubated with HRP-conjugated anti-mouse secondary antibody (Sigma-Aldrich) at a dilution 1: 1000 in PBS-Tween for 1 h at room temperature. Three final washes were performed (200 μ l/well) followed by the addition of SIGMAFAST™ OPD Tablets solution (Sigma-Aldrich) (100 μ l/well)

and incubated for 15 to 20 min at room temperature. The stop reagent (1 M sulphuric acid from Sigma-Aldrich – 100 µl/well) was used, once the solution started changing colour from colourless to pale orange. The absorbance was measured on a Fusion plate reader (Perkin-Elmer) at 492 nm.

An ELISA was also used to measure binding of HCV E2 glycoprotein to CD81. CD81 protein samples were bound to plates, blocked and washed as described above. Following this samples were incubated with 50 µl of 85 µg/ml purified streptavidin-tagged E2 protein for 1 h at room temperature incubation before three PBS washes. 50 µl primary anti-strep antibody (Progen #910STR) diluted 1:5000 in PBST and incubated at room temperature for 1 h with gentle shaking. After three PBS washes (200 µl/well), samples were incubated with HRP-conjugated anti-mouse secondary antibody (Sigma-Aldrich) at a dilution 1: 1000 in PBST. Wells were washed and developed as described above.

2.7 DYNAMIC LIGHT SCATTERING (DLS)

50 µl of CD81-SMALP (0.05mg/ml) or CD81-DDM (0.09mg/ml) were used for the DLS measurement. The data was collected using a Malvern Instruments ZetasizerNano S (633 nm). Measurements were taken in disposable ultra-micro UV cuvettes (BrandTech Scientific) at 20 °C with 300 s equilibration time, where automated parameters were used. Each measurement was repeated seven times.

2.8 CIRCULAR DICHROISM SPECTROSCOPY (CD)

For circular dichroism (CD) analysis, the purified protein was buffer exchanged into 20 mM sodium phosphate buffer for CD81-SMALP. Sodium phosphate buffer was supplemented with 0.1% DDM and 0.01% cholesterol hemisuccinate for CD81-DDM purified protein. This was because high background signal was observed with the HEPES/NaCl buffer whereas minimum background noise was observed with sodium phosphate buffer. Data was collected between 260 and 180 nm (far-UV wavelength region) at 0.2 nm intervals at 20°C using a Jasco J-1500 CD spectrometer. A 1 mm path length quartz cuvette (Starna UK) containing 200 µl of CD81-SMALP (0.05mg/ml) or CD81-DDM (0.09mg/ml) was used for each analysis and 18 technical replicates were performed for each sample. For thermal melt analysis, temperature was increased from 25 - 90°C, with increments of 5°C. Lipid only SMALPs, using 1,2-dimyristoyl-sn-glycero-3-phosphocholine lipid, were used for background signal correction to subtract possible SMA co-polymer and lipid noise from the CD spectrum. The amount required as a control for CD analysis was calculated using the absorbance peak at 260 nm.

Structural CD data analysis was performed using Dichroweb[26, 27] using CDSSTR, Contin-LL and Selcon3 algorithms[28-34].

2.9 PROTEIN AGGREGATION ASSAY

The aggregation of purified protein following heating at a range of temperatures from 20 °C to 90 °C was assayed. 100 µl of purified protein samples (30 µg /ml) were heated for 10 min at selected temperature points followed by centrifugation at 10,000 g for 10 min to remove aggregation. Supernatant was collected and 20 µl sample from each were loaded on 12% SDS-PAGE gel following staining with Coomassie stain. The gel bands were analysed by densitometry and amount of remaining protein in each temperature treatment was normalised with the 4 °C temperature treatment reading.

2.10 NEGATIVE STAIN ELECTRON MICROSCOPY (EM)

EM grids (copper grid with a formvar-carbon support film from Agar Scientific) were glow discharged for 2 min to distribute negative charge on the surface before for sample adsorption. 10 µl of purified protein (2µg/ml) was placed on each grid and incubated for 2 min at room temperature. Excess solution was blotted off and grids were washed three times with distilled water before gently drying using Whatman filter paper (Agar Scientific). 3 µl drop of 2% (w/v) uranyl acetate stain was placed on each grid and incubated for 4 min following three washed with distilled water. The grids were visualised using a JEOL 2100Plus super high resolution TEM-STEM at room temperature and images were acquired at 60000x.

2.11 LIPID EXTRACTION

Lipids were extracted following a previously described method using methyl-*tert*-butyl ether [35] with slight modifications. Briefly, 50 µl of sample was mixed with 375 µl of methanol in a 2 ml polypropylene tube (Eppendorf, UK) and vortexed for 10 sec. Methyl-*tert*-butyl ether (1250 µl) was added and sample was vortexed for an additional 10 sec. Lipids were extracted for 1 h at 4°C on a rotary shaker. Phase separation was induced by addition of 375 µl of water (MS grade, Fisher, UK) followed by 10 sec vortexing and 10 min extraction on rotary shaker at 4 °C. Sample was centrifuged at 1,000 xg for 10 min, upper phase was collected and dried under vacuum (Vacuum concentrator, Eppendofr, Germany). Samples were stored at -20 °C prior to MS analysis.

2.12 MASS SPECTROMETRY

Lipid extracts corresponding to 150 µl of sample were dissolved in an electrospray (ESI) solution containing 5 mM ammonium formate in a mixture of methanol and chloroform (2:1, v:v) and analysed by mass spectrometry (MS). Each sample was directly infused into the 5600 TripleTof mass spectrometer (AB Sciex, UK) at a flow rate of 500 nl/min using a syringe pump (Harvard Apparatus, USA) and PicoTip ESI emitters (New Objective, Germany). Data were acquired in both positive (2.4 kV) and negative ion modes (-2.2 kV). MS and

MS/MS spectra were acquired in an information dependent mode (IDA) for 15 min. One IDA cycle consisted of a MS survey scan (m/z range from 400 to 1250 positive ion mode, and m/z 200 to 1250 in negative ion mode) followed by consecutive CID fragmentations (10 V collision energy) of the six most abundant ions in MS survey scan. Acquired ions were temporarily excluded from MS/MS acquisition for 300 sec.

2.13 DATA ANALYSIS

Western blots and SDS-PAGE gels were analysed by densitometry using the software ImageJ.

Mass spectrometry data visualization and analysis was done using Peak View software (version 2.2, AB sciex, UK). Ion list with precursor m/z values averaged over the acquisition time was exported to the Lipid maps search. All positive identifications were manually inspected for the presence of lipid-specific fragment ions or neutral losses.

3. RESULTS

3.1 SOLUBILISATION OF CD81

Recombinant CD81 was expressed in *Pichia pastoris* as described previously [20]. Membranes were extracted from these cells and the solubilisation of these membranes by two different SMA polymers, SMA 2000 or SZ25010 at 2.5 % (w/v), was investigated and compared to the conventional detergent DDM (1% (w/v)). SMA 2000 and SZ25010 are both co-polymers of styrene and maleic acid but they differ in their ratio of styrene:maleic acid and average molecular weight. SMA 2000 has a 2:1 ratio of styrene:maleic acid with an average molecular weight of 7.5kDa, whereas SZ25010 has a 3:1 ratio, and average molecular weight of 10kDa. Both have been shown previously to be effective at solubilising membrane proteins [36]. As shown in Figure 1A & B, the addition of either SMA polymer or DDM resulted in a dramatic drop in turbidity within 1 hour, whereas no change in turbidity was observed with the negative control (PBS). To investigate if this general breakup of the membrane correlated with solubilisation of CD81 specifically, the sample was subjected to ultracentrifugation, and the solubilised protein in the supernatant (S) and insoluble protein in the pellet (P) collected and analysed by Western blot. In Figure 1C it can be seen that DDM solubilisation yields a relatively intense band for CD81 in the soluble fraction. However, there is still some CD81 which remains insoluble, and overall DDM gives a solubilisation efficiency of approximately 70%. For both SMA 2000 and SZ25010 the band in the soluble fraction appears less intense than for DDM, and somewhat smeared. This finding supports previous suggestions that the presence of excess free SMA polymer interferes with binding to the antibody. However if the insoluble pellet fractions are examined (which do not contain the excess SMA polymer) they are of similar intensity to the pellet fractions obtained with DDM, suggesting that a substantial amount of the CD81 has been solubilised. In fact when the pellet fractions are compared to those obtained with a buffer negative control, it is apparent that both SMA 2000 and SZ25010 solubilised approximately 60% of the total CD81 (Figure 1D). Therefore going forward the amount of CD81 remaining insoluble in the pellet was always measured and compared to the negative PBS control. Having established that SMA polymers were capable of effectively extracting CD81 from *P. pastoris* membranes, the next step was to optimise the conditions for this extraction. As no difference was observed in the abilities of SMA 2000 and SZ25010 to extract CD81 (Figure 1), and previous reports with other proteins suggested SZ25010 gives lower stability and yield in downstream purification, it was decided to proceed solely with SMA 2000, which had previously been shown to be the best polymer to date [36]. The first variable investigated was solubilisation time from 15 min to 4 hours. As shown in Figure 2B the membrane turbidity dropped rapidly and after just 15 mins remained constant. However, although this bulk membrane disruption occurred quickly, the specific solubilisation of CD81 was much slower (Figure 2A & B), reaching its highest levels after 4 hours. Secondly the concentration of SMA 2000 for optimal solubilisation was investigated, from 0.5% (w/v) to 5% (w/v). Figure 2D shows that the optical density decreased substantially with just 0.5%

(w/v) SMA 2000, and increasing the concentration made only small improvements to this. However, this was not replicated in CD81 specific solubilisation (Figure 2C), which was only approximately 20% efficient with 0.5% (w/v) SMA 2000, and gradually increased as the concentration of SMA was increased. Thus it is clear that the changes in turbidity do not reflect the solubilisation of CD81. CD81 solubilisation by SMA 2000 was greatest after 4 hours at room temperature, and the efficiency could be improved by increasing the concentration of SMA 2000 to 5% (w/v) SMA. Although these results show it may be possible to increase the yield of CD81-SMALPs by increasing the concentration of SMA or solubilisation time, the effects of excess SMA [37, 38], or keeping CD81 at room temperature for longer periods would need to be evaluated. For these reasons, it was decided going forward to use standard conditions of 2.5% (w/v) SMA 2000 for 1 hour at room temperature, which achieved a solubilisation efficiency of approximately 60%.

3.2 PURIFICATION OF CD81

Having established that CD81 could be successfully solubilised by SMA 2000, the next step was to purify the protein. An initial Ni-NTA affinity purification (Supplementary Figure 1) followed a published protocol for SMALPs [21]. Whilst CD81-SMALP was found to bind well initially to the Ni²⁺ resin, with little protein in the flow through (Supplementary Figure 1B), CD81 started to come off the column during the washes, and in the elution fractions it co-purified with a number of contaminants (Supplementary Figure 1A). Mass spectrometry analysis of these contaminating bands identified plasma membrane H⁺ ATPase and alcohol dehydrogenase as particularly prominent contaminants (Supplementary Figure 2). Therefore the purification procedure was optimized by i) increasing the amount of Ni-NTA resin used, which seemed to improve the binding of CD81 and ii) increasing the volume of washing steps to remove contaminating bands. This resulted in a much more pure preparation of CD81 as shown in Figure 3A. The structural integrity of this purified CD81 was tested by an ELISA assay using conformation sensitive anti-CD81 antibodies 1s337 and 1s135, which are known to recognize non-linear epitopes within the large extracellular loop of CD81, alongside an antibody which recognizes a linear peptide sequence 2s131 [22], as well as binding to the Hepatitis C virus (HCV) E2 glycoprotein. As can be seen in Figure 3B all three antibodies and the E2 glycoprotein bound well to the purified CD81, showing that the purified CD81 retained its folded structure. For downstream comparisons, CD81 was also purified using DDM at comparable yields (Supplementary Table 2), and this CD81-DDM also bound the conformation sensitive antibodies and HCV E2 glycoprotein (Supplementary Figure 3).

3.3 BIOPHYSICAL COMPARISON OF CD81-SMALP AND CD81-DDM

The biophysical properties of purified CD81-SMALP were then analysed and compared to those of CD81-DDM. Dynamic light scattering (DLS) was used to investigate the size of the

particles. CD81-SMALPs had a diameter of approximately 10 nm which agrees well with previously reported studies using SMALPs [36, 39]. CD81-DDM particles were smaller at approximately 5 nm, but this also agrees well with previously reported DDM micelle size [36]. To examine the folding of the protein, CD spectroscopy was used. It should be noted that for this CD spectroscopy analysis CD81 samples were exchanged into a phosphate buffer lacking glycerol and NaCl, because of high light absorbance of Cl⁻ ions in the UV range. Figure 4 shows that both CD81-SMALP and CD81-DDM display two negative peaks at 222 nm and 208 nm and a positive peak at 190 nm in the CD spectrum. This indicates alpha-helical secondary structure, which suggests that protein in both samples is folded and abundant in alpha-helical secondary structure. However, there is a clear difference in the shape of the CD81-DDM curve compared to CD81-SMALP, particularly at the 208nm peak. After Dichroweb CD data analysis and taking the average alpha helical abundance from three secondary structure databases (CDSSTR, Contin-LL and Selcon3 – Supplementary Table 1), the CD81-SMALP appears to comprise 71% alpha helical structure and around 28% of the protein was classified as unstructured. This is in contrast to CD81-DDM where approximately 58% of the protein was alpha helical and 37% protein was classified as unstructured. This suggests that SMA solubilised and purified CD81 was better able to retain its correct structural conformation, since a previously-published CD81 crystal structure (PDB ID: 5TCX) indicated that approximately 68% of the protein is in alpha helical secondary structure [7].

The thermostability of purified CD81 was investigated next, using three different approaches. Firstly, CD spectroscopy was used to monitor protein secondary structure stability over a temperature range from 20°C to 90°C (Figure 5 A & B). At temperatures up to 40°C, minor unfolding was observed, with CD81-SMALP possibly being a little more stable than CD81-DDM. However at temperatures above 40°C the ellipticity reduced at each temperature increment. Despite the initial spectral differences observed between CD81-SMALP and CD81-DDM (Figure 4B), little difference was seen between them in resisting heat treatment (Figure 5C). The second approach examined protein aggregation upon heating (Figure 5 D & E). After heating, samples were centrifuged to remove large aggregates before visualising by Western blot. CD81-SMALP appeared unaffected up to temperatures of 40-50°C. However at temperatures of 60°C and higher, larger molecular weight aggregates are visible (Figure 5D). In contrast CD81-DDM showed little variation even at high temperatures (Figure 5E). Finally thermostability was monitored by binding to a conformationally sensitive antibody (Figure 5F). CD81-SMALP showed very little change in binding at increasing temperatures up until 60°C, suggesting intact conformation of CD81. At temperatures of 70°C and above, a gradual decrease in antibody binding was observed, likely due to heat-induced unfolding of the extracellular loop 2 epitope. In contrast CD81-DDM antibody binding decreased following incubation at 20°C and continued to decrease at each increased temperature, indicating lower protein stability in detergent micelles at this important extracellular loop than when CD81 is encapsulated in SMALPs. Both the aggregation and

antibody binding assays utilised CD81-DDM in its optimized buffer composition containing both glycerol and NaCl.

3.4 SIZE EXCLUSION CHROMATOGRAPHY OF CD81-SMALP

Purified CD81-SMALP was concentrated and loaded on a size exclusion chromatography column. As shown in Figure 6A, this resulted in two major elution peaks as measured by the absorbance at 280 nm. Peak 1 was very close to the void volume of the column (≈ 8 ml) suggesting the presence of larger complexes or potential protein or SMALP aggregates. Peak 2 eluted at ≈ 13 ml column elution volume suggesting extraction of smaller molecules (potential isolation of individual CD81 SMALPs) in this region. When these fractions were analysed by SDS-PAGE (Figure 6B), it was clear that both contained CD81, at fairly similar concentrations, although Peak 1 also contained higher molecular weight bands. Interestingly, when fractions from both peaks were assayed for their ability to bind the HCV E2 glycoprotein, only those from peak 2 showed binding (Figure 1D).

As Peak 2 on the gel filtration trace appeared to contain functional, non-aggregated CD81, efforts were made to try to increase the proportion of CD81 in this peak. Consequently *P. pastoris* growth conditions were changed: lower cell biomass was taken forward for large scale growth in methanol medium ($OD_{600} = 1$) instead of inducing expression at high cell density ($OD_{600} > 5$). Although this produced less CD81 (0.43 mg/l culture) than the standard higher biomass protocol (1.06 mg/l culture), a higher proportion was seen in peak 2 (Figure 7 & Supplementary Table 2). Having identified that this change improved the quality of the protein, further optimization steps were made. HEPES buffer was used for solubilisation, purification and gel filtration work, supplemented with 200 mM sodium chloride and 10% glycerol, instead of the standard Tris/sodium chloride buffer. HEPES is often considered to buffer hydrogen ions more effectively than Tris, thus preventing pH-related SMALP or protein aggregation (although increasing the concentration of Tris could have been another approach), and glycerol is often reported to aid stability of membrane proteins. Finally, CD81-SMALP elution fractions from affinity purification were concentrated less before analysing on the SEC to minimise aggregation due to heavy concentration of the protein. A like-for-like comparison of the results obtained before and after optimisation are shown in Figure 7. After applying these changes, an increase in Peak 2 was observed in the SEC spectrum (Figure 7C). SDS-PAGE analysis of the SEC fractions shows considerably less protein in the aggregated fractions (Peak 1) and more in ≈ 13 ml elution fractions (Peak 2) (Figure 7D).

3.5 BIOPHYSICAL CHARACTERIZATION OF THE TWO SIZE EXCLUSION PEAKS

To further investigate the nature of the proteins present in each of the two peaks observed with SEC, the two samples were characterized by DLS, CD and EM. CD spectra of both peaks

from SEC overlaid perfectly with the sample straight off the affinity purification column, showing that both peak 1 and peak 2 contained fully folded CD81 (Figure 8A). However DLS analysis of the same samples showed that the particles within peak 1 were larger than those within peak 2, with an average diameter of approximately 20 nm. When examined by negative stain EM, the particles within peak 1 also appeared larger than those in peak 2 (Supplementary Figure 5).

3.6 ANALYSIS OF CD81 LIPID ENVIRONMENT

As SMALPs extract membrane proteins within their lipid surrounding, we used a shotgun lipidomics approach to investigate changes in the lipid environment of purified CD81-SMALP. We compared the lipid profiles of total membrane extracts from *Pichia pastoris* overexpressing CD81 (Figures 9A and 10A), the SMA solubilised membranes (Figures 9B and 10B) and purified CD81-SMALPs (Figure 9C and 10C). In total, we were able to detect lipids across eight different lipid classes, namely phosphatidylcholines, phosphatidylethanolamines, phosphatidic acids, phosphatidylserines, phosphatidylinositols, triacylglycerols, sphingomyelin and cardiolipins.

The positive ion mode mass spectrum of total membranes (Figure 9A) is dominated by phosphatidylcholine species. The phosphatidylcholine lipids have relatively long polyunsaturated fatty acid chains of 38 carbon atoms and five to eight double bonds and phosphatidylcholine (38:5) appears as the most intense lipid ion. Phosphatidylethanolamine lipids were detected with significantly lower signal intensities, lower complexity, shorter polyunsaturated fatty acid chains and lower unsaturation levels. A single sphingomyelin species, sphingomyelin (d36:3) was observed with relatively strong signal intensity. Finally, a few lipids from a triacylglycerol family, eg triacylglycerol (54:6) were also detected. Negative ion mode spectra (Figure 10A) suggest phosphatidylinositol as the most dominant lipid class detected. Phosphatidylinositol lipids were detected with polyunsaturated fatty acid chains of predominantly 34C atoms and up to two double bonds. Phosphatidylserine and cardiolipin were also observed in the total membrane extracts of *Pichia pastoris* overexpressing CD81.

Mass spectra of SMA solubilised membranes revealed some differences in measured lipid profiles compared to the total membranes. We observed the total loss of sphingomyelin and triacylglycerol lipids (Figure 9B). Although there was no change in the overall complexity of phosphatidylcholine lipids, it is apparent that phosphatidylcholine lipids with shorter polyunsaturated fatty acid chains and lower unsaturation levels, e.g. phosphatidylcholine(36:2) were now the most dominant ones. Phosphatidylethanolamine and phosphatidylserine profiles appear more diverse in SMA-solubilised membranes compared to total membranes. However, the complexity and relative intensities of negatively-charged lipids, in particular phosphatidylinositol and cardiolipin stayed unaffected by the SMA membrane solubilisation.

Significant changes in phosphatidylcholine and PE species were observed upon purification of CD81-SMALP from SMA solubilised membranes. Although phosphatidylcholine lipids with reduced complexity were still present, they were no longer the dominant species in the positive ion mode, while PE lipids were almost absent (Figure 9C). Surprisingly, phosphatidic acid and phosphatidylinositol lipids dominated even the positive ion spectrum of CD81 purified nanodiscs. A single phosphatidic acid lipid, namely phosphatidic acid(O-36:2)/phosphatidic acid(P-36:1) was detected as the most intense signal in a form of an alkyl ether. Corresponding tandem mass spectrum showed the prominent neutral losses of water and CO₂, characteristic for carboxylic acids. Similarly, a phosphatidylinositol alkyl ether lipid such as phosphatidylinositol(O-34:2)/phosphatidylinositol(P-34:1) was detected as the second most dominant signal. The negative ion mode measurements revealed that phosphatidylinositol and cardiolipin diversity and relative abundances remained conserved even upon purification of CD81-SMALP, while phosphatidylserine heterogeneity was reduced.

4. DISCUSSION

In this study we have shown that recombinant CD81 expressed in *P. pastoris* yeast can be effectively solubilised and purified using SMA polymer. It was previously shown that another member of the tetraspanin family, TSPAN7, could also be extracted with SMA from the yeast *Saccharomyces cerevisiae*. However this was not the case for several other tetraspanins expressed in *S. cerevisiae* [40]. Interestingly, we found that monitoring membrane turbidity was not a good measure for CD81 solubilisation. Membrane turbidity is quick and easy to measure and has been used previously to monitor kinetics of solubilisation [37, 41]. However we found that turbidity decreased much more quickly and at lower concentrations of SMA than were needed for CD81 solubilisation (Figure 2). This contrasts with a previous report where SMA was used to solubilise the ABC transporter MRP4/ABBC4 (multidrug resistance protein 4) from insect cell membranes, where the changes in turbidity were mirrored by protein specific solubilisation [42]. Secondly the speed of CD81 solubilisation was much slower than that observed for MRP4 [42], possibly related to the differences in lipid composition of yeast membranes compared to insect cells [43]. Thus this highlights the importance of measuring protein-specific solubilisation efficiency and that optimal conditions will likely be both protein and expression-system dependent.

As has been observed previously for many other proteins, purification of CD81 using SMA showed comparable yields to that obtained using conventional detergents but gave a cleaner purification [17, 36, 44]. CD spectroscopy also showed subtle differences between SMA and DDM purified CD81, suggesting that SMA encapsulated CD81 retained more of the native secondary structure. It should be noted that this was a side-by-side comparison under comparable conditions, in buffer that did not contain NaCl or glycerol, which have been found to stabilize the DDM purified protein. This may explain why the DDM sample was less folded than the previously published crystal structure which was also obtained using DDM [7]. However it highlights that SMALPs can offer improved stability without the need for special handling or buffer conditions. Interestingly, SMA did not offer any improvement over DDM for CD81 thermostability as measured by CD spectroscopy or when using a gel based aggregation assay, which contrasts with many previous reports [13, 36, 44]. Reasons for this are not totally clear, perhaps the secondary structure of CD81 is inherently quite stable, and it has been suggested previously that the presence of detergent can help to decrease aggregation for some proteins[36]. Differences between DDM and SMA purified CD81 were observed when using a binding assay for a conformation-sensitive antibody. This antibody is known to bind to a non-linear epitope within the large extracellular loop, which is important for function of CD81[22]. Thus SMA encapsulation stabilised the 3-dimensional folded structure of CD81, even if very little change in secondary structure occurred. Notably this antibody binding assay utilized CD81-DDM within its optimised buffer conditions containing both glycerol and NaCl, so this difference is not simply caused by using suboptimal conditions for the detergent solubilised protein.

SEC analysis of SMA purified CD81 showed a large peak in or close to the void volume in addition to a second peak that eluted later. Presumably the void peak contains some form

of larger assembly or aggregated species, however it is notable that both peaks contain fully folded CD81, thus the void peak does not contain denatured, aggregated protein. CD81 itself is known to form dimers and larger assemblies [4, 5] so it is possible that this peak contains larger assemblies of CD81. Notably, the CD81 crystal structure revealed a monomeric form of a tetraspanin, which contrasts with the crystal structure of the extracellular LEL domain of human CD81 (88 of 236 residues). This latter structure confirms the presence of the conserved CCG motif, two disulfide bridges and a potential dimerisation interface [45]. The DLS analysis of the void peak shows that it is approximately twice the size of a standard SMALP. One possibility is that CD81 has oligomerised and the SMALP formed around the oligomer is larger to accommodate this. Such flexibility of size has been observed previously [46]. Alternatively it has also been shown previously that SMALP discs themselves can dimerise [47]. It was interesting that the intensity of the A_{280} signal for the void peak was significantly higher than that of the second peak, and yet SDS-PAGE analysis showed almost equal concentrations of CD81 in both peaks. Therefore perhaps the species present in the void peak was also scattering light. Importantly only the CD81 in the second peak bound to the HCV E2 glycoprotein. Given that the CD spectroscopy showed the CD81 in this peak was fully folded, it is unlikely that the protein has become sufficiently unstructured such that the E2 binding site is not formed. It is more likely that access to the binding site is sterically blocked in some way. Whether this is due to oligomerisation of CD81 itself, or of the SMALP complex, or caused by a conformational change of the protein remains to be determined. Going forward it will be interesting to investigate in more detail the nature of the CD81-polymer complexes in the two peaks, especially given that the proportion of CD81 present in the second peak could be improved by changing expression and buffer conditions.

Lastly, we examined the lipids associated with SMA-purified CD81. Importantly, we first analysed the lipid species present in whole membranes from *P. pastoris* expressing CD81, and SMA solubilised *P. pastoris* membranes. In total, lipids from seven families were detected in total membrane extracts, such as phosphatidylcholine, phosphatidylethanolamine, sphingomyelin, triacylglycerol, phosphatidylserine, phosphatidylinositol and cardiolipin. These results are in good correlation with previously published work [48, 49]. Ejsing *et al.* investigated the total lipidome of *Saccharomyces cerevisiae* in one of the most comprehensive yeast lipid profiling studies published to date [48]. Although here we report the membrane lipidome of *P. pastoris*, we have detected the presence of lipids from the same lipid classes as observed with *S. cerevisiae* [48]. Furthermore, it appears that in both *P. pastoris* and *S. cerevisiae* phosphatidylcholine, phosphatidylethanolamine and phosphatidylinositol lipids are the most dominant ones. Grillitsch *et al.* reported that the most abundant phospholipid class detected in *P. pastoris* plasma membranes was phosphatidylethanolamine, followed by phosphatidylcholine, phosphatidylserine, phosphatidic acid, phosphatidylinositol and cardiolipin [49]. Our results detected no phosphatidylglycerol class of lipids in the *P. pastoris* membrane, in agreement with both previous studies in yeast [48, 49]. Further, sterol lipids, primarily ergosterol as one of the major constituents of yeast membranes was not identified. Its abundance decreases

up to 8-fold in *P. pastoris* compared to *S. cerevisiae* [50]. Thus, future work should include design of more targeted lipidomics approaches for the detection of sterols in membranes of yeast. It is important to note that the lipid fingerprint of yeast membranes depends on growth conditions such as temperature, carbon source and oxygen supply, and these could contribute to the quantitative difference observed in between studies [51-53]. Furthermore, it was previously shown that SMA solubilisation is not preferential for any particular lipid subset [13-15, 54]. Nevertheless, in SMA-solubilised membranes we have observed loss of the signals for sphingomyelin (d36:3) and triacylglycerol, higher diversity of phosphatidylethanolamine and phosphatidylserine lipids, and the change in signal intensities of phosphatidylcholine lipids while their heterogeneity remained the same. The most dramatic change in lipid fingerprint was observed in CD81-SMALP. This was reflected in an apparent decrease in positively charged phosphatidylethanolamine and phosphatidylserine lipids, and enrichment in negatively charged phosphatidylserine, phosphatidylacid and phosphatidylinositol. Both phosphatidylserine and phosphatidylinositol play a crucial role in generating a negative surface plasma membrane potential that is essential for lipid binding of proteins with pleckstrin homology domains [55]. Phosphatidylinositol lipids are present in relatively low abundance in cells and tissues, but in a high concentration in plasma membranes. They serve as precursors of several signalling lipids and play a crucial role in signal transduction. Carloni *et al.* showed that transfection of human cells with CD81 stimulated phosphatidylinositol-4-kinase activity at the plasma membrane, increasing the levels of phosphatidylinositol-4-phosphate which affected protein recruitment, cellular signalling and proliferation [56].

It should be noted that CD81 is a human protein that was expressed recombinantly in *P. pastoris*, so it is not within a fully-native lipid environment. Going forward, it would be of interest to investigate whether similar lipid profiles are observed with CD81 expressed in mammalian cells. It would also be interesting to reconstitute purified CD81 into liposomes of defined lipids, as identified in this study, and examine the effect of the lipids on CD81 function, structure and stability. Our findings set the scene for understanding the role of protein-lipid interactions in the biology of CD81 and of the tetraspanin family more broadly.

CREDIT AUTHOR CONTRIBUTION

Hoor Ayub: Conceptualization, Methodology, Validation, Analysis, Investigation, Writing – original draft. **Michelle Clare:** Methodology, Investigation, Writing – review & editing. **Ivana Milic:** Analysis, Investigation, Resources, Writing – original draft. **Nikola P. Chmel:** Analysis, Investigation, Resources, Writing – review & editing. **Heike Boning:** Resources, Investigation. **Andrew Devitt:** Supervision, Funding Acquisition, Writing – review & editing. **Thomas Krey:** Resources, Supervision, Funding Acquisition, Writing – review & editing. **Roslyn M. Bill:** Conceptualization, Resources, Supervision, Funding Acquisition, Writing – original draft. **Alice J. Rothnie:** Conceptualization, Methodology, Supervision, Writing – original draft.

ACKNOWLEDGEMENTS

We thank Dr Saskia Bakker, from the electron microscopy facility at the University of Warwick, for expert technical help with the negative stain electron microscopy. The underlying data for this study can be found at the Aston Explorer Data Repository (DOI link to be finalised).

FUNDING SOURCES

Our work was supported by grants from the Biotechnology & Biosciences Research Council (BBSRC); RMB and MC through BB/N007417/1, IM and AD through BB/S00324X/1 and BB/M006298/1. HA was supported by an Aston University School of Life & Health Sciences Doctoral Studentship, and TK received funding from the German Center of Infection Research (DZIF).

REFERENCES

- [1] M.E. Hemler, Tetraspanin functions and associated microdomains, *Nature reviews. Molecular cell biology*, 6 (2005) 801-811.
- [2] M.E. Hemler, Targeting of tetraspanin proteins-potential benefits and strategies, *Nature reviews. Drug discovery*, 7 (2008) 747-758.
- [3] M.E. Hemler, Tetraspanin proteins promote multiple cancer stages, *Nature reviews. Cancer*, 14 (2014) 49-60.
- [4] N. Bonander, M. Jamshad, D. Oberthur, M. Clare, J. Barwell, K. Hu, M.J. Farquhar, Z. Stamatakis, H.J. Harris, K. Dierks, T.R. Dafforn, C. Betzel, J.A. McKeating, R.M. Bill, Production, purification and characterization of recombinant, full-length human claudin-1, *PLoS one*, 8 (2013) e64517.
- [5] O.V. Kovalenko, X. Yang, T.V. Kolesnikova, M.E. Hemler, Evidence for specific tetraspanin homodimers: inhibition of palmitoylation makes cysteine residues available for cross-linking, *The Biochemical journal*, 377 (2004) 407-417.

- [6] G. Min, H. Wang, T.T. Sun, X.P. Kong, Structural basis for tetraspanin functions as revealed by the cryo-EM structure of uroplakin complexes at 6-Å resolution, *The Journal of cell biology*, 173 (2006) 975-983.
- [7] B. Zimmerman, B. Kelly, B.J. McMillan, T.C.M. Seegar, R.O. Dror, A.C. Kruse, S.C. Blacklow, Crystal Structure of a Full-Length Human Tetraspanin Reveals a Cholesterol-Binding Pocket, *Cell*, 167 (2016) 1041-1051 e1011.
- [8] R. Umeda, Y. Satouh, M. Takemoto, Y. Nakada-Nakura, K. Liu, T. Yokoyama, M. Shirouzu, S. Iwata, N. Nomura, K. Sato, M. Ikawa, T. Nishizawa, O. Nureki, Structural insights into tetraspanin CD9 function, *Nat Commun*, 11 (2020) 1606.
- [9] L. Feneant, S. Levy, L. Cocquerel, CD81 and hepatitis C virus (HCV) infection, *Viruses*, 6 (2014) 535-572.
- [10] R. Dawaliby, C. Trubbia, C. Delporte, M. Masureel, P. Van Antwerpen, B.K. Kobilka, C. Govaerts, Allosteric regulation of G protein-coupled receptor activity by phospholipids, *Nat Chem Biol*, 12 (2016) 35-39.
- [11] M.E. Zoghbi, R.S. Cooper, G.A. Altenberg, The Lipid Bilayer Modulates the Structure and Function of an ATP-binding Cassette Exporter, *J Biol Chem*, 291 (2016) 4453-4461.
- [12] A. Rothnie, D. Theron, L. Soceneantu, C. Martin, M. Traikia, G. Berridge, C.F. Higgins, P.F. Devaux, R. Callaghan, The importance of cholesterol in maintenance of P-glycoprotein activity and its membrane perturbing influence, *Eur Biophys J*, 30 (2001) 430-442.
- [13] J.M. Dorr, M.C. Koorengel, M. Schafer, A.V. Prokofyev, S. Scheidelaar, E.A. van der Cruisen, T.R. Dafforn, M. Baldus, J.A. Killian, Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K⁺ channel: the power of native nanodiscs, *Proc Natl Acad Sci U S A*, 111 (2014) 18607-18612.
- [14] A.C.K. Teo, S.C. Lee, N.L. Pollock, Z. Stroud, S. Hall, A. Thakker, A.R. Pitt, T.R. Dafforn, C.M. Spickett, D.I. Roper, Analysis of SMALP co-extracted phospholipids shows distinct membrane environments for three classes of bacterial membrane protein, *Sci Rep*, 9 (2019) 1813.
- [15] I. Prabudiansyah, I. Kusters, A. Caforio, A.J. Driessen, Characterization of the annular lipid shell of the Sec translocon, *Biochim Biophys Acta*, 1848 (2015) 2050-2056.
- [16] M. Jamshad, V. Grimard, I. Idini, T.J. Knowles, M.R. Dowle, N. Schofield, P. Sridhar, Y.P. Lin, R. Finka, M. Wheatley, O.R.T. Thomas, R.E. Palmer, M. Overduin, C. Govaerts, J.M. Ruysschaert, K.J. Edler, T.R. Dafforn, Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins, *Nano Research*, 8 (2015) 774-789.
- [17] S. Gulati, M. Jamshad, T.J. Knowles, K.A. Morrison, R. Downing, N. Cant, R. Collins, J.B. Koenderink, R.C. Ford, M. Overduin, I.D. Kerr, T.R. Dafforn, A.J. Rothnie, Detergent-free purification of ABC (ATP-binding-cassette) transporters, *Biochem J*, 461 (2014) 269-278.
- [18] M. Jamshad, J. Charlton, Y.P. Lin, S.J. Routledge, Z. Bawa, T.J. Knowles, M. Overduin, N. Dekker, T.R. Dafforn, R.M. Bill, D.R. Poyner, M. Wheatley, G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent, *Biosci Rep*, 35 (2015).
- [19] N.L. Pollock, S.C. Lee, J.H. Patel, A.A. Gulamhussein, A.J. Rothnie, Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer, *Biochim Biophys Acta*, 1860 (2018) 809-817.
- [20] M. Jamshad, S. Rajesh, Z. Stamataki, J.A. McKeating, T. Dafforn, M. Overduin, R.M. Bill, Structural characterization of recombinant human CD81 produced in *Pichia pastoris*, *Protein Expr Purif*, 57 (2008) 206-216.
- [21] A.J. Rothnie, Detergent-Free Membrane Protein Purification, *Methods Mol Biol*, 1432 (2016) 261-267.
- [22] J. Grove, K. Hu, M.J. Farquhar, M. Goodall, L. Walker, M. Jamshad, H.E. Drummer, R.M. Bill, P. Balfe, J.A. McKeating, A new panel of epitope mapped monoclonal antibodies recognising the prototypical tetraspanin CD81, *Wellcome Open Res*, 2 (2017) 82.

- [23] A. Rothnie, J. Storm, J. Campbell, K.J. Linton, I.D. Kerr, R. Callaghan, The topography of transmembrane segment six is altered during the catalytic cycle of P-glycoprotein, *J Biol Chem*, 279 (2004) 34913-34921.
- [24] T. Krey, J. d'Alayer, C.M. Kikuti, A. Saulnier, L. Damier-Piolle, I. Petitpas, D.X. Johansson, R.G. Tawar, B. Baron, B. Robert, P. England, M.A. Persson, A. Martin, F.A. Rey, The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule, *PLoS Pathog*, 6 (2010) e1000762.
- [25] A.W. Tarr, P. Lafaye, L. Meredith, L. Damier-Piolle, R.A. Urbanowicz, A. Meola, J.L. Jestin, R.J. Brown, J.A. McKeating, F.A. Rey, J.K. Ball, T. Krey, An alpaca nanobody inhibits hepatitis C virus entry and cell-to-cell transmission, *Hepatology*, 58 (2013) 932-939.
- [26] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res*, 32 (2004) W668-673.
- [27] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, *Biopolymers*, 89 (2008) 392-400.
- [28] N. Sreerama, R.W. Woody, A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal Biochem*, 209 (1993) 32-44.
- [29] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy, *Protein Sci*, 8 (1999) 370-380.
- [30] S.W. Provencher, J. Glockner, Estimation of globular protein secondary structure from circular dichroism, *Biochemistry*, 20 (1981) 33-37.
- [31] I.H. van Stokkum, H.J. Spoelder, M. Bloemendal, R. van Grondelle, F.C. Groen, Estimation of protein secondary structure and error analysis from circular dichroism spectra, *Anal Biochem*, 191 (1990) 110-118.
- [32] L.A. Compton, W.C. Johnson, Jr., Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication, *Anal Biochem*, 155 (1986) 155-167.
- [33] P. Manavalan, W.C. Johnson, Jr., Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra, *Anal Biochem*, 167 (1987) 76-85.
- [34] N. Sreerama, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Anal Biochem*, 287 (2000) 252-260.
- [35] V. Matyash, G. Liebisch, T.V. Kurzchalia, A. Shevchenko, D. Schwudke, Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics, *Journal of lipid research*, 49 (2008) 1137-1146.
- [36] K.A. Morrison, A. Akram, A. Mathews, Z.A. Khan, J.H. Patel, C. Zhou, D.J. Hardy, C. Moore-Kelly, R. Patel, V. Odiba, T.J. Knowles, M.U. Javed, N.P. Chmel, T.R. Dafforn, A.J. Rothnie, Membrane protein extraction and purification using styrene-maleic acid (SMA) copolymer: effect of variations in polymer structure, *Biochem J*, 473 (2016) 4349-4360.
- [37] B. Bersch, J.M. Dorr, A. Hessel, J.A. Killian, P. Schanda, Proton-Detected Solid-State NMR Spectroscopy of a Zinc Diffusion Facilitator Protein in Native Nanodiscs, *Angew Chem Int Ed Engl*, 56 (2017) 2508-2512.
- [38] S. Rehan, V.P. Jaakola, Expression, purification and functional characterization of human equilibrative nucleoside transporter subtype-1 (hENT1) protein from Sf9 insect cells, *Protein Expr Purif*, 114 (2015) 99-107.
- [39] D.J. Swainsbury, S. Scheidelaar, R. van Grondelle, J.A. Killian, M.R. Jones, Bacterial reaction centers purified with styrene maleic acid copolymer retain native membrane functional properties and display enhanced stability, *Angew Chem Int Ed Engl*, 53 (2014) 11803-11807.
- [40] K. Skaar, H.J. Korza, M. Tarry, P. Sekyrova, M. Hogbom, Expression and Subcellular Distribution of GFP-Tagged Human Tetraspanin Proteins in *Saccharomyces cerevisiae*, *PLoS One*, 10 (2015) e0134041.

- [41] C. Logez, M. Damian, C. Legros, C. Dupre, M. Guery, S. Mary, R. Wagner, C. M'Kadmi, O. Nosjean, B. Fould, J. Marie, J.A. Fehrentz, J. Martinez, G. Ferry, J.A. Boutin, J.L. Baneres, Detergent-free Isolation of Functional G Protein-Coupled Receptors into Nanometric Lipid Particles, *Biochemistry*, 55 (2016) 38-48.
- [42] D. Hardy, R.M. Bill, A.J. Rothnie, A. Jawhari, Stabilization of Human Multidrug Resistance Protein 4 (MRP4/ABCC4) Using Novel Solubilization Agents, *SLAS Discov*, (2019) 2472555219867074.
- [43] J.J. Dominguez Pardo, J.M. Dorr, A. Iyer, R.C. Cox, S. Scheidelaar, M.C. Koorengel, V. Subramaniam, J.A. Killian, Solubilization of lipids and lipid phases by the styrene-maleic acid copolymer, *Eur Biophys J*, 46 (2017) 91-101.
- [44] Y. Liu, E. Moura, J.M. Dorr, S. Scheidelaar, M. Heger, M.R. Egmond, J.A. Killian, T. Mohammadi, E. Breukink, *Bacillus subtilis* MraY in detergent-free system of nanodiscs wrapped by styrene-maleic acid copolymers, *PLoS One*, 13 (2018) e0206692.
- [45] K. Kitadokoro, M. Ponassi, G. Galli, R. Petracca, F. Falugi, G. Grandi, M. Bolognesi, Subunit association and conformational flexibility in the head subdomain of human CD81 large extracellular loop., *Biological Chemistry*, 383 (2002) 1447-1452.
- [46] C. Sun, S. Benlekbir, P. Venkatakrisnan, Y. Wang, S. Hong, J. Hosler, E. Tajkhorshid, J.L. Rubinstein, R.B. Gennis, Structure of the alternative complex III in a supercomplex with cytochrome oxidase, *Nature*, 557 (2018) 123-126.
- [47] V. Postis, S. Rawson, J.K. Mitchell, S.C. Lee, R.A. Parslow, T.R. Dafforn, S.A. Baldwin, S.P. Muench, The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy, *Biochim Biophys Acta*, 1848 (2015) 496-501.
- [48] C.S. Ejsing, J.L. Sampaio, V. Surendranath, E. Duchoslav, K. Ekroos, R.W. Klemm, K. Simons, A. Shevchenko, Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry, *Proceedings of the National Academy of Sciences*, 106 (2009) 2136-2141.
- [49] K. Grillitsch, P. Tarazona, L. Klug, T. Wriessnegger, G. Zellnig, E. Leitner, I. Feussner, G. Daum, Isolation and characterization of the plasma membrane from the yeast *Pichia pastoris*, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838 (2014) 1889-1897.
- [50] E. Zinser, F. Paltauf, G. Daum, Sterol composition of yeast organelle membranes and subcellular distribution of enzymes involved in sterol metabolism, *J Bacteriol*, 175 (1993) 2853-2858.
- [51] A.M. Galea, A.J. Brown, Special relationship between sterols and oxygen: were sterols an adaptation to aerobic life?, *Free radical biology & medicine*, 47 (2009) 880-889.
- [52] I. Schuiki, M. Schnabl, T. Czabany, C. Hrastnik, G. Daum, Phosphatidylethanolamine synthesized by four different pathways is supplied to the plasma membrane of the yeast *Saccharomyces cerevisiae*, *Biochim Biophys Acta*, 1801 (2010) 480-486.
- [53] M. Inan, M.M. Meagher, Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of *Pichia pastoris*, *Journal of bioscience and bioengineering*, 92 (2001) 585-589.
- [54] I.A. Smirnova, D. Sjostrand, F. Li, M. Bjorck, J. Schafer, H. Ostbye, M. Hogbom, C. von Ballmoos, G.C. Lander, P. Adelroth, P. Brzezinski, Isolation of yeast complex IV in native lipid nanodiscs, *Biochim Biophys Acta*, 1858 (2016) 2984-2992.
- [55] R.V. Stahelin, Lipid binding domains: more than simple lipid effectors, *J Lipid Res*, 50 Suppl (2009) S299-S304.
- [56] V. Carloni, A. Mazzocca, K.S. Ravichandran, Tetraspanin CD81 is linked to ERK/MAPKinase signaling by Shc in liver tumor cells, *Oncogene*, 23 (2004) 1566-1574.

Figure 1

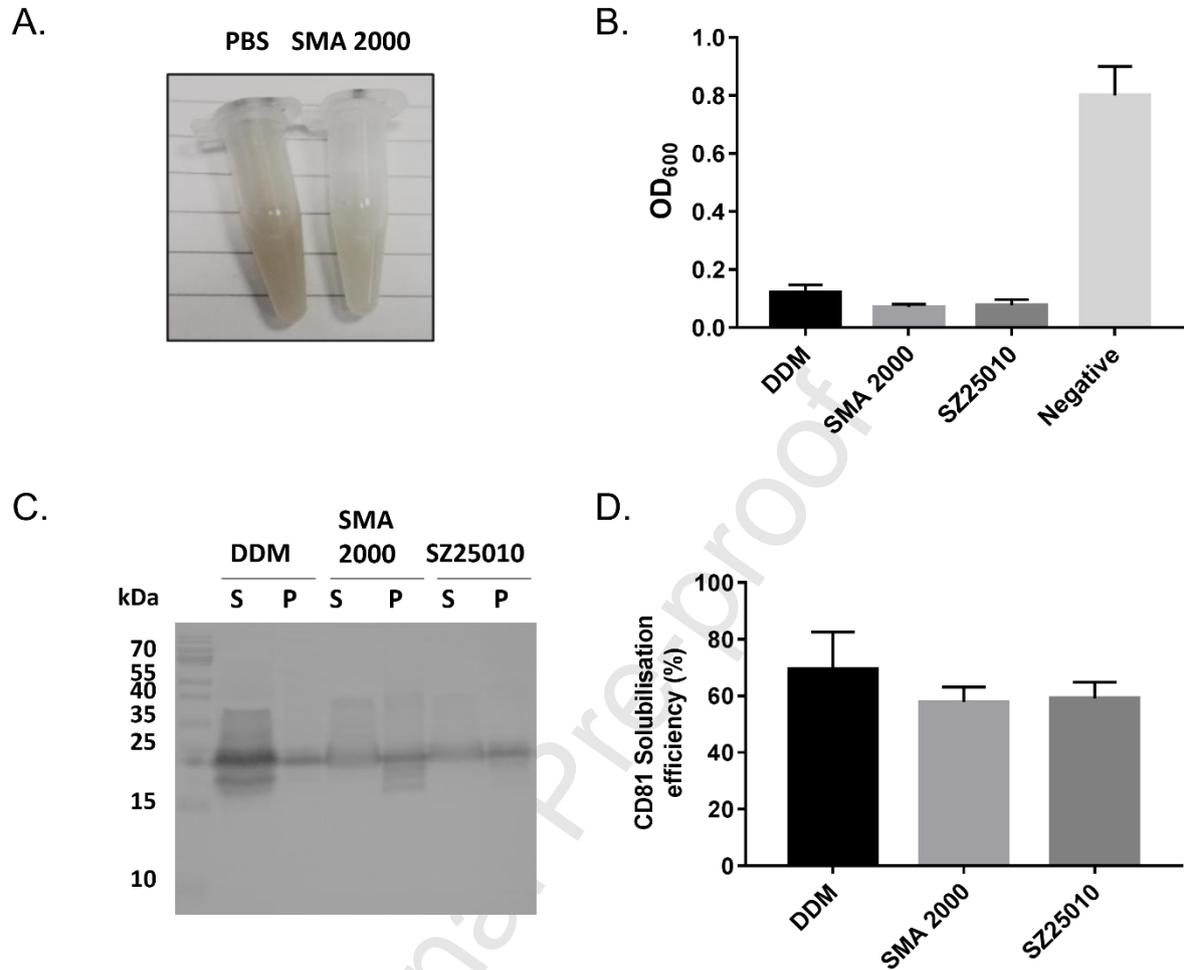


Figure 1. CD81 can be solubilised using SMA polymers or conventional detergent. Membrane preparations from CD81 expressing *Pichia pastoris* (30 mg/ml wet pellet weight) were solubilised with 2% (w/v) DDM, 2.5% (w/v) SMA polymer or PBS (negative control) for 1 hour at room temperature. A) Image showing the change in appearance as SMA clarifies the membrane suspension compared to PBS control. B) Optical density measurements of the samples measured at 600 nm. Data are mean \pm SD, n=3. C) Following solubilisation, samples were centrifuged at 100,000g for 20 min at 4°C. Solubilised protein in the supernatant (S) was harvested. Insoluble material in the pellet (P) was resuspended in an equal volume of purification buffer supplemented with 2% (w/v) SDS. Samples were analysed by Western blot using an anti-CD81 primary antibody (mAb 2s131). D) Average CD81 solubilisation efficiency. Data are mean \pm SD, n=3.

Figure 2

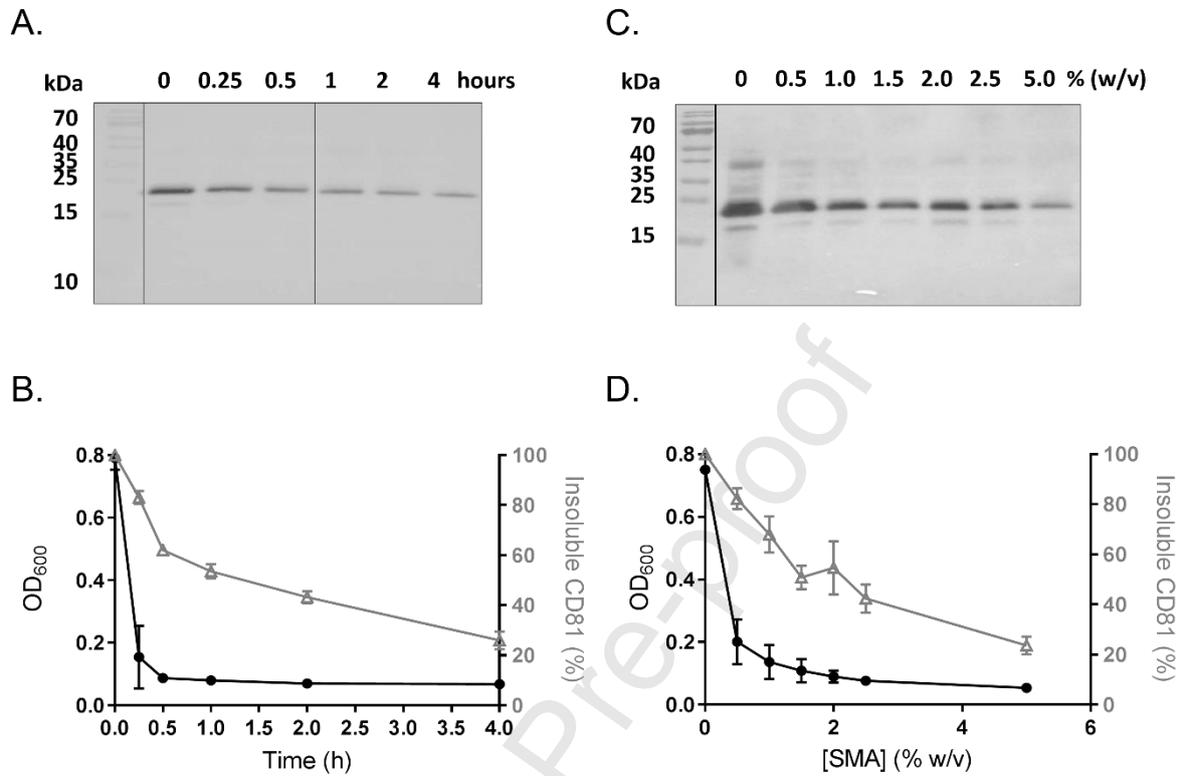
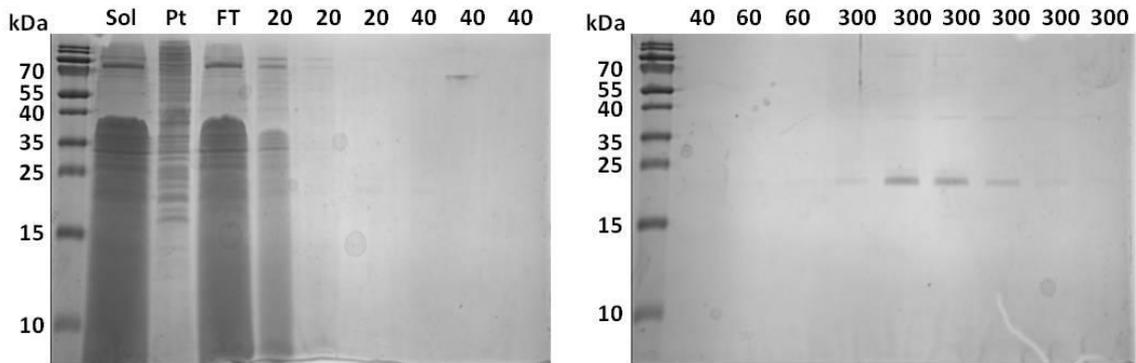


Figure 2. Solubilisation of CD81 by SMA 2000 is slower than the break-up of the total membrane. A) Representative Western blot showing the amount of CD81 remaining insoluble (in the pellet) after various lengths of incubation with 2.5% (w/v) SMA 2000 at room temperature. B) Average measurements of solubilisation at different time points. Optical density measurements at 600nm (solid black circles). Percentage insoluble CD81 determined by densitometric analysis of Western blots as shown in A (open grey triangles). Data are mean±SD, n=3. C) Representative Western blot showing the amount of CD81 remaining insoluble after incubation with varying concentrations of SMA 2000 for 1 hour at room temperature. D) Average measurements of solubilisation using different SMA 2000 concentrations. Optical density measurements at 600nm (solid black circles). Percentage insoluble CD81 determined by densitometric analysis of Western blots as shown in C (open grey triangles). Data are mean±SD, n=2.

Figure 3

A.



B.

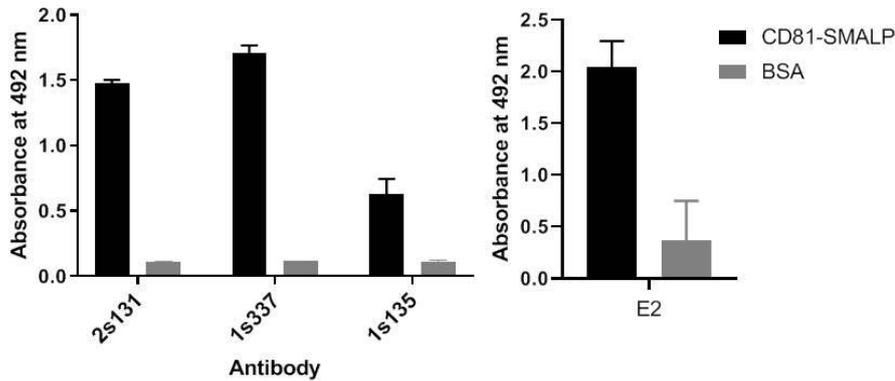


Figure 3. Purification of SMALP-encapsulated CD81 yields a functionally-folded protein. A)

The Ni-NTA affinity purification procedure was modified to increase the amount of resin used, and increase the volume of washes. Samples of each step of the purification analysed by SDS-PAGE and stained with Instant Blue. Sol is the total solubilised protein, Pt is the insoluble protein, FT is flow-through and numbers show the concentration in mM of imidazole in the wash buffers. B) Elution fractions containing purified CD81 were pooled together and their folded state assessed by an ELISA assay using conformation-sensitive anti-CD81 antibodies; 1s337 and 1s135, alongside the linear peptide antibody 2s131. Bovine serum albumin was used as a negative control. Primary antibodies were detected with an anti-mouse-HRP secondary antibody and visualised by SIGMAFAST™ OPD tablet. Optical density reading at 492 nm wavelength was measured. Data are mean±SD, n=2. Purified CD81 was also analysed for binding to the E2 glycoprotein that acts as CD81 receptor ligand by an ELISA assay, n=3.

Figure 4

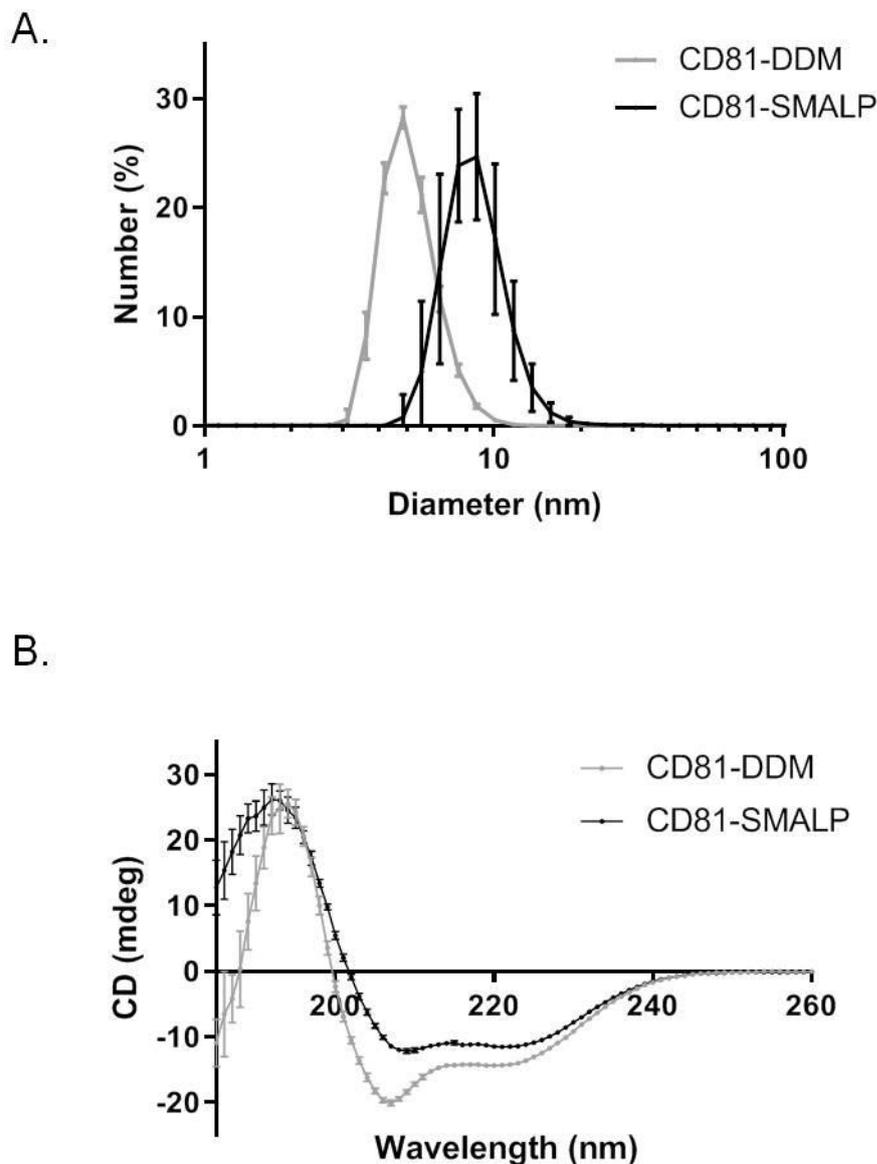


Figure 4. Biophysical characterisation of purified CD81. A) DLS analysis of purified CD81 either encapsulated within SMALPs (black) or within DDM micelles (grey). Obtained using a Malvern Instruments Zetasizer Nano S (633 nm) with a 1.0 cm path length disposable cuvette. Seven technical replicates were performed for each sample, $n=2$. B) CD spectra of purified CD81 either encapsulated within SMALPs (black) or within DDM micelles (grey). Purified CD81-SMALP and CD81-DDM was buffer exchanged into 20 mM sodium phosphate buffer pH 8 at concentrations of 0.05 mg/ml and 0.09 mg/ml, respectively. 200 μ l of each sample volume was used in a 1 mm path length cuvette and measured using a Jasco J-1500 instrument. Maximum absorbance of 260 nm and minimum absorbance of 180 nm was used for CD detection at 0.2 nm intervals at 20 $^{\circ}$ C; 18 technical replicates were performed for each sample, where $n=2$.

Figure 5

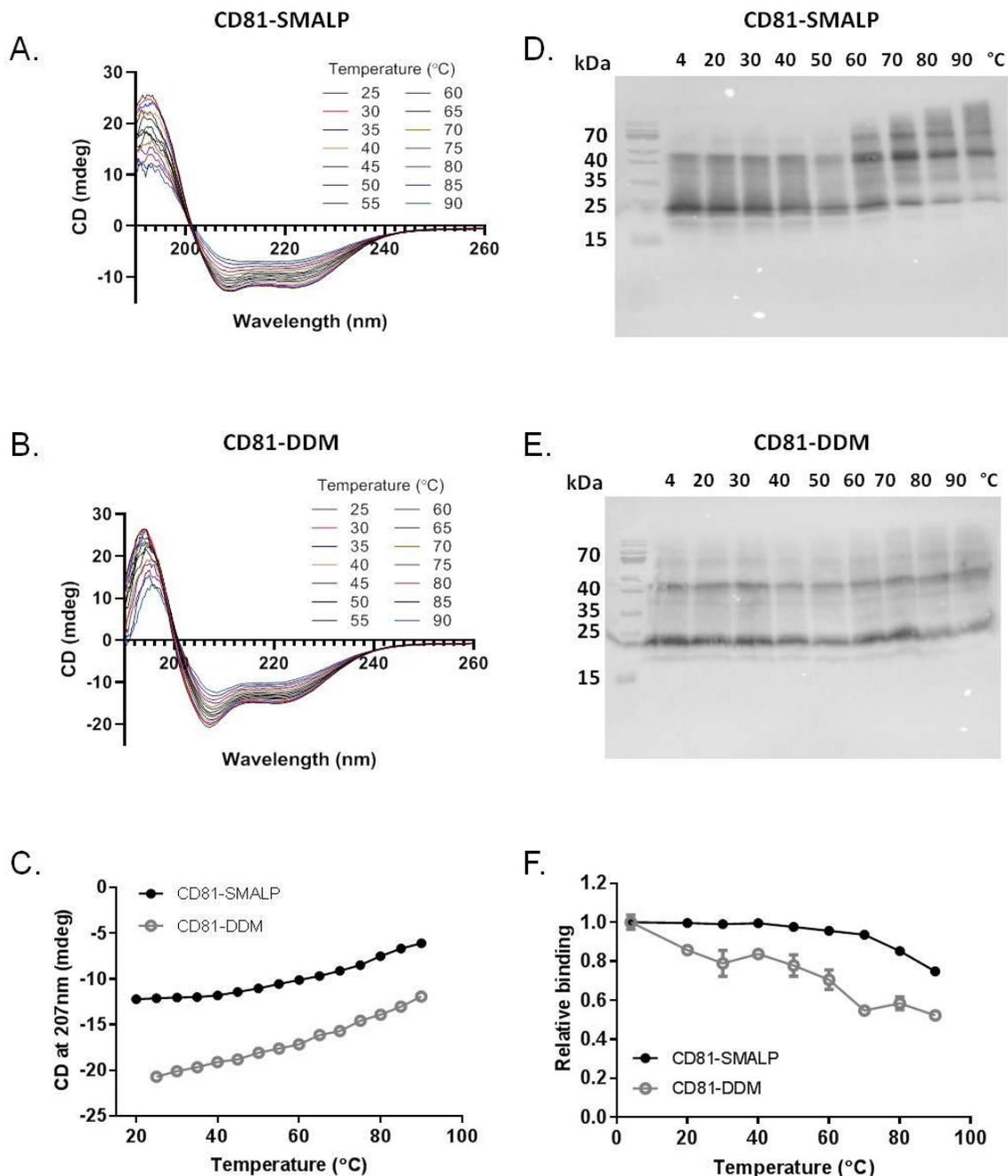


Figure 5. Thermostability of purified CD81. Effect of temperature on secondary structure measured by CD spectroscopy over a range of different temperatures for A) CD81-SMALP and B) CD81-DDM. Temperatures ranged from 25 °C to 90°C with 5° C increment. Eighteen technical replicates were taken at each condition, n=2. C) The temperature dependence on the CD signal at 207nm for both CD81-SMALP (black closed circles) and CD81-DDM (grey open circles). D & E) Effect of temperature on protein aggregation analysed by Western blot using an anti-CD81 antibody (mAb 2s131) for D) CD81-SMALP and E) CD81-DDM. F) Effect of increasing temperature on the binding of CD81 to a conformation sensitive anti-CD81 (mAb 1s337) antibody, using an ELISA assay. Black solid circles are CD81-SMALP, open grey squares are CD81-DDM. Data are mean±SEM, n=3.

Figure 6

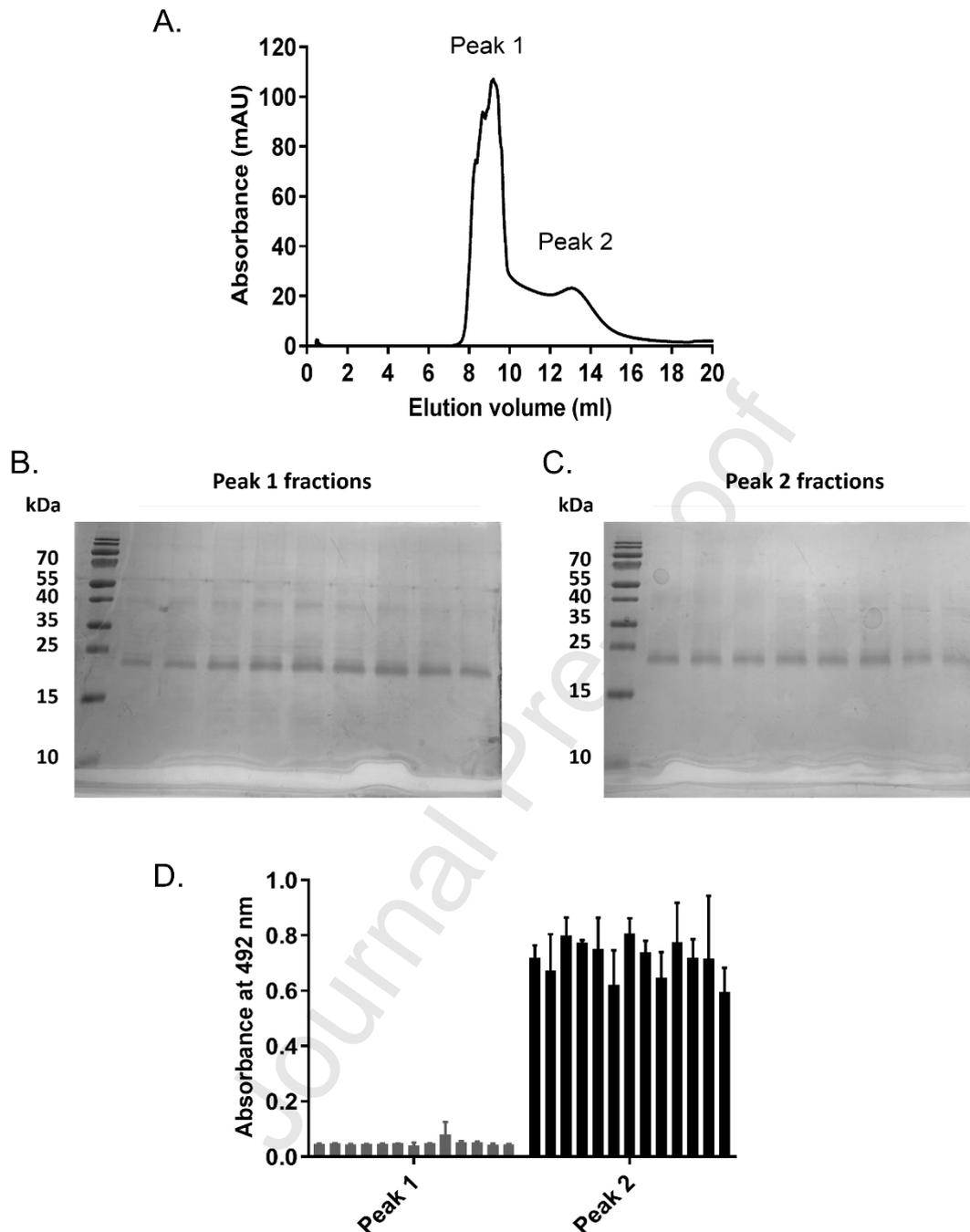


Figure 6. Size exclusion chromatography (SEC) analysis of CD81-SMALP reveals two distinct protein populations. A) Following Ni-NTA affinity chromatography elution fractions were pooled, concentrated and 500 μ l at 1 mg/ml was loaded on a Superdex 200 30/10 size exclusion column, run at 0.5ml/min. Protein absorbance monitored at 280nm. B & C) Fractions from the two main peaks obtained by SEC were run on SDS-PAGE and stained with Instant Blue. D) Fractions from the two SEC peaks were analysed for binding to the E2 glycoprotein that acts as CD81 receptor ligand by an ELISA assay, n=2.

Figure 7

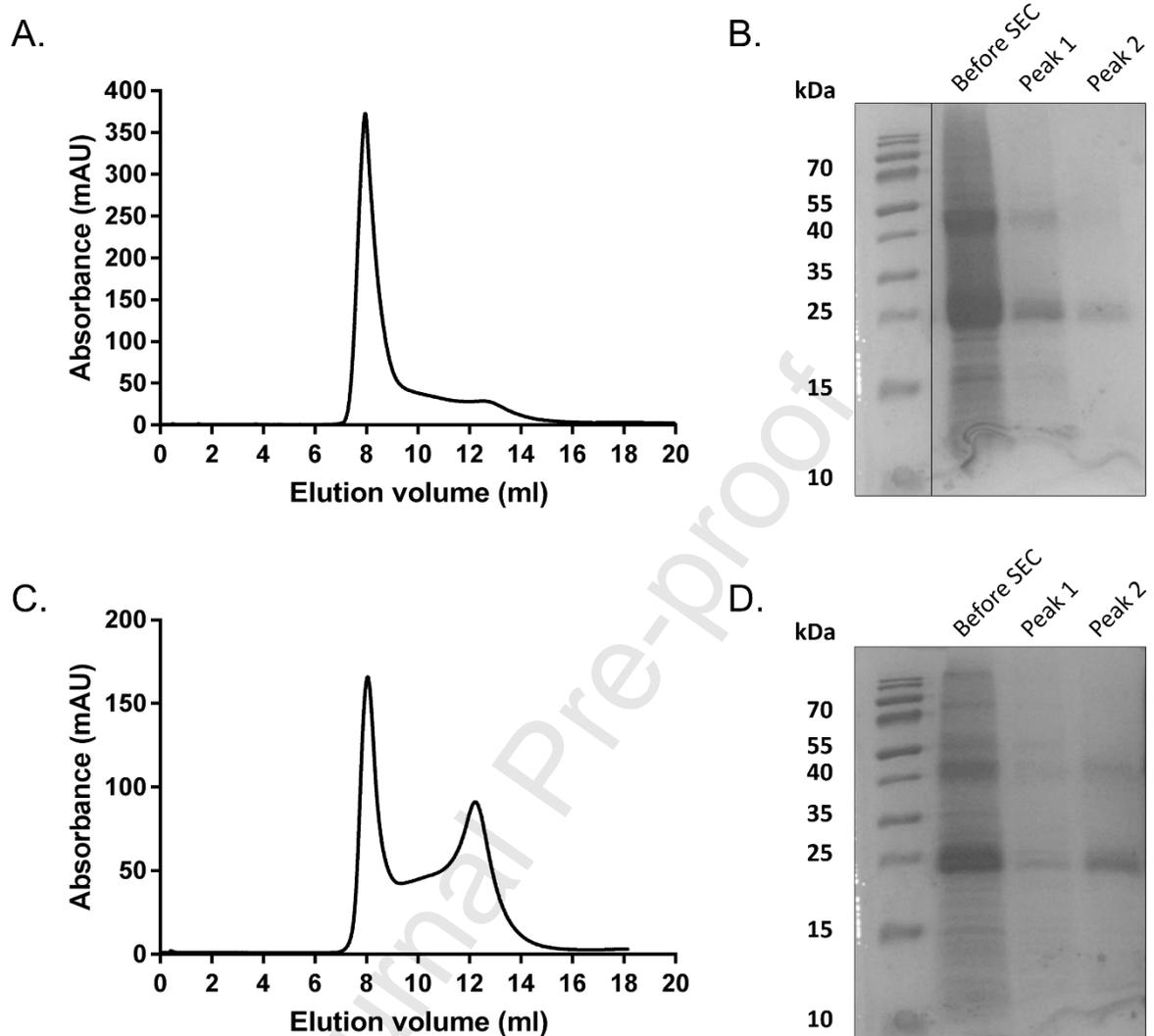


Figure 7. Optimization of the SEC conditions increases the proportion of E2-binding-competent CD81. A) SEC profile of purified CD81-SMALP obtained using standard conditions. B) SDS-PAGE of the CD81 sample loaded on the column, peak 1 and peak 2 from the SEC trace in A. C) SEC profile of purified CD81-SMALP obtained using modified expression and buffer conditions. D) SDS-PAGE of the CD81 sample loaded on the column, peak 1 and peak 2 from the SEC trace in C. Traces are representative of those seen in 3 repeats.

Figure 8

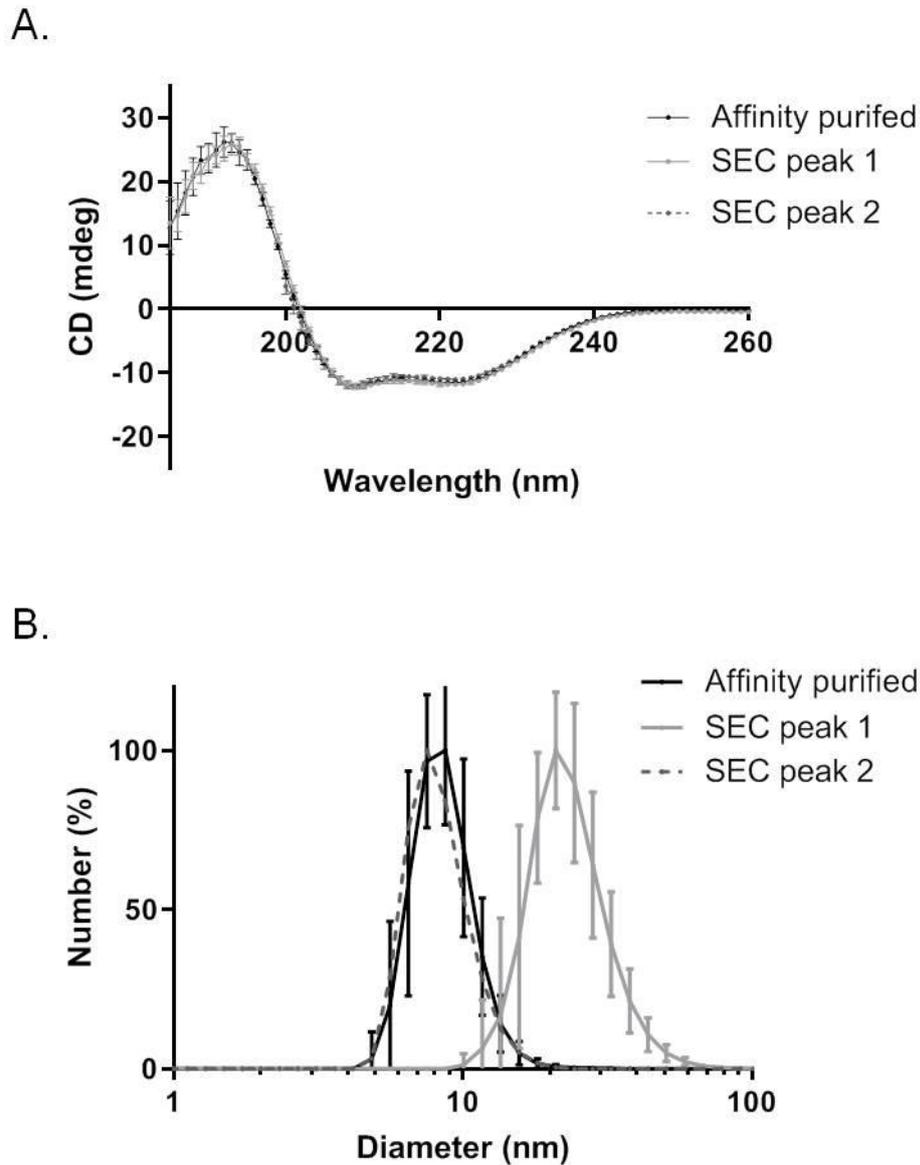


Figure 8. Biophysical characterisation of CD81 -SMALPs following size exclusion chromatography. A) CD spectra of affinity purified CD81 -SMALPs prior to SEC (black) , CD81-SMALPs in peak 1 following SEC (light grey) and from peak 2 following SEC (dark grey, dashed). 200 μ l of each sample volume was used in a 1 mm path length cuvette and measured using a Jasco J-1500 instrument. Maximum absorbance of 260 nm and minimum absorbance of 180 nm was used for CD detection at 0.2 nm intervals at 20 $^{\circ}$ C; 18 technical replicates were performed for each sample, where n=2. B) DLS analysis of purified CD81-SMALPs prior to SEC (black) , CD81-SMALPs in peak 1 following SEC (light grey) and from peak 2 following SEC (dark grey, dashed). Obtained using a Malvern Instruments Zetasizer Nano S (633 nm) with 1.0 cm path length disposable cuvette. Seven technical replicates were performed for each sample, n=2.

Figure 9

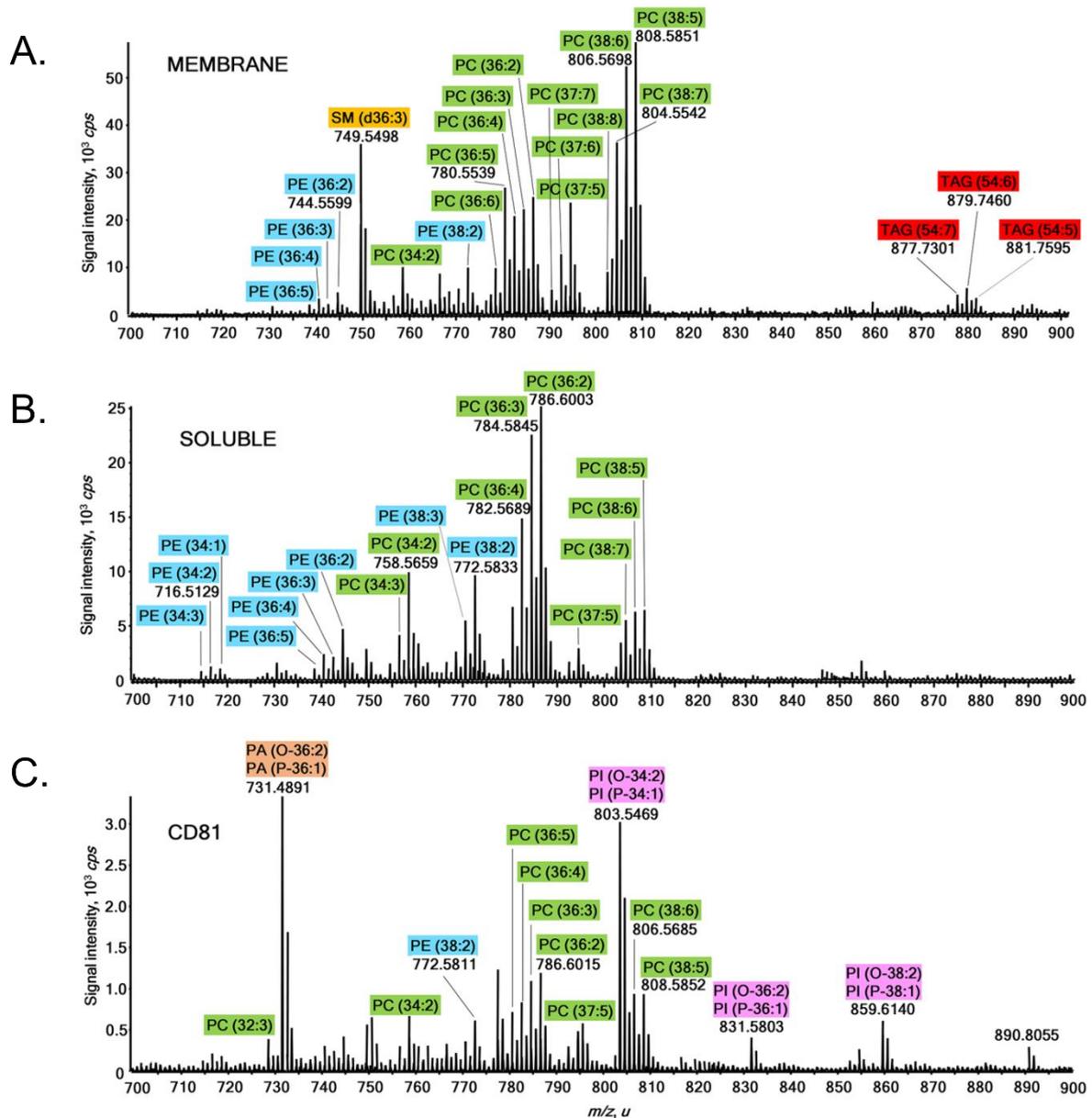


Figure 9. Positive ion mode MS spectra measured for total lipid extracts of total yeast membrane, SMA solubilised membrane and purified CD81 SMALPs. Identified glycerophosphatidylcholines are labelled in green (PC), glycerophosphatidyletanolamines in blue (PE), glycerophosphatidylinositols in magenta (PI), glycerophosphatidycacids in (PA) and sphingomyelin in orange (SM).

Figure 10

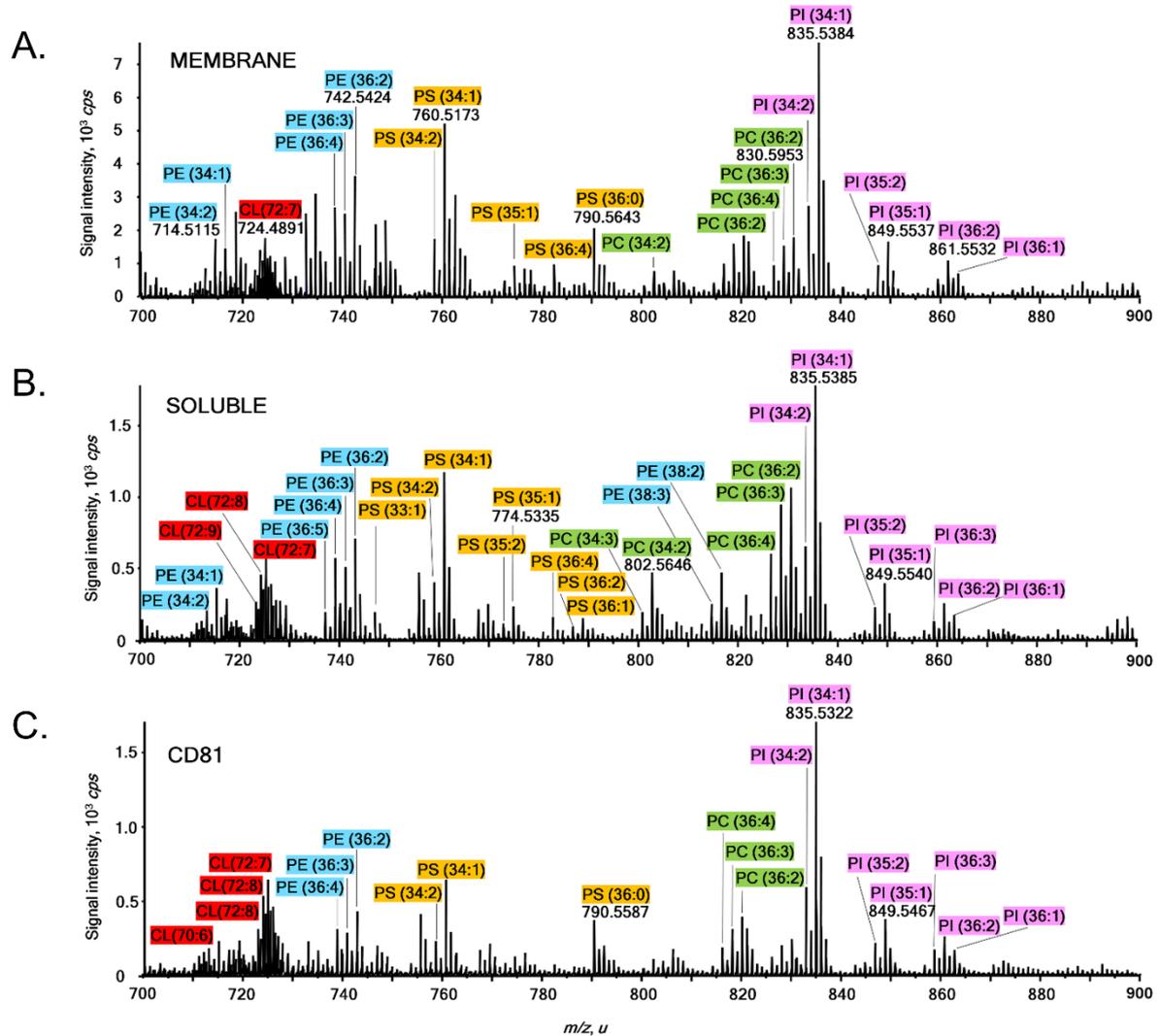


Figure 10. Negative ion mode MS spectra measured for total lipid extracts of total yeast membrane, SMA solubilised membrane and purified CD81 SMALPs. Identified glycerophosphatidylcholines are labelled in green (PC), glycerophosphatidyletanolamines in blue (PE), glycerophosphatidylinositols in magenta (PI), glycerophosphatidylserines in orange (PS) and cardiolipins in red (CL).

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

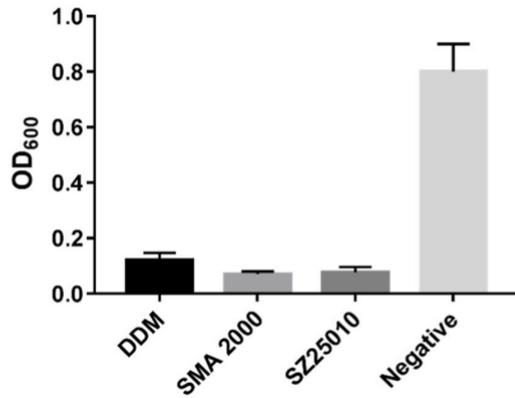
The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof

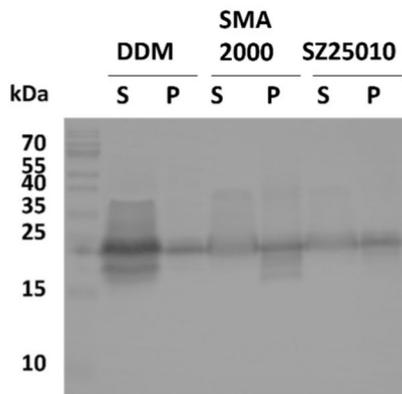
A.



B.



C.



D.

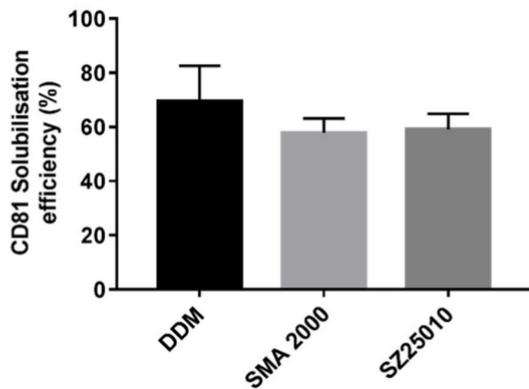
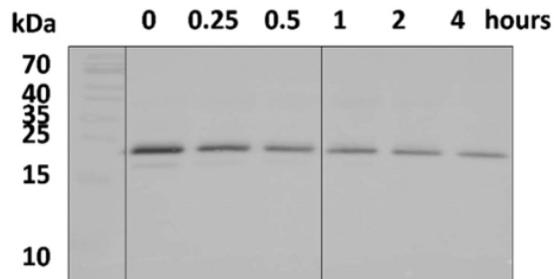
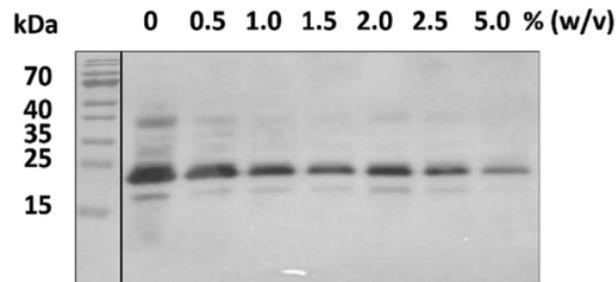


Figure 1

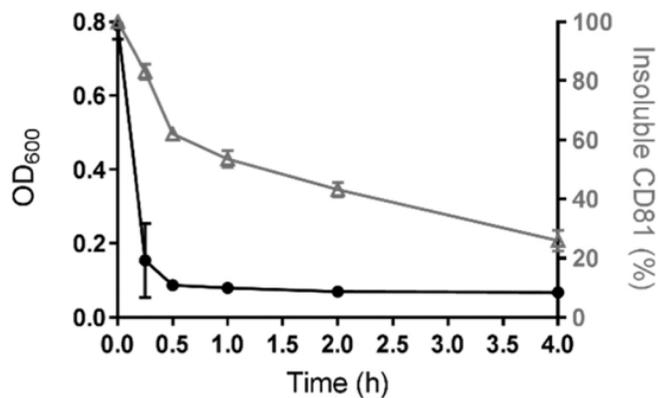
A.



C.



B.



D.

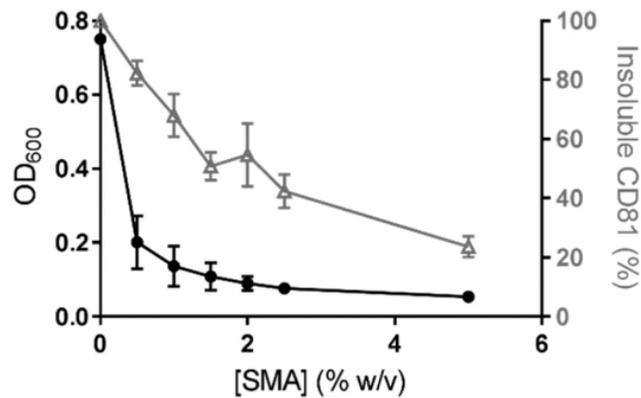
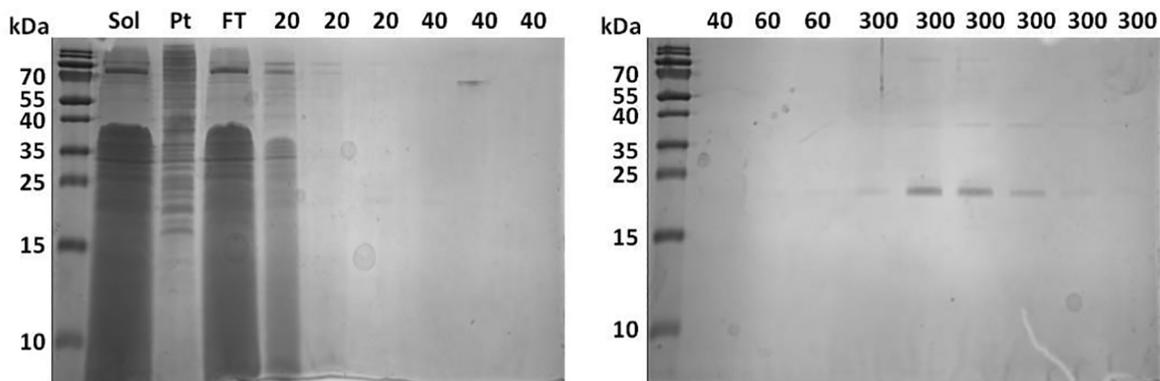


Figure 2

A.



B.

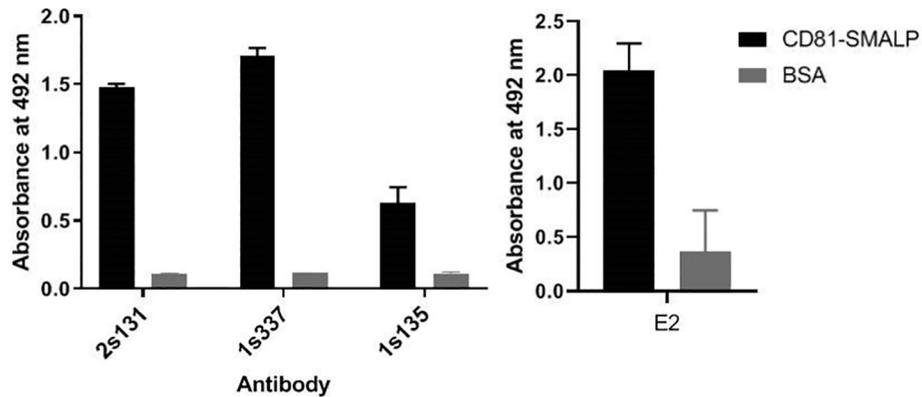


Figure 3

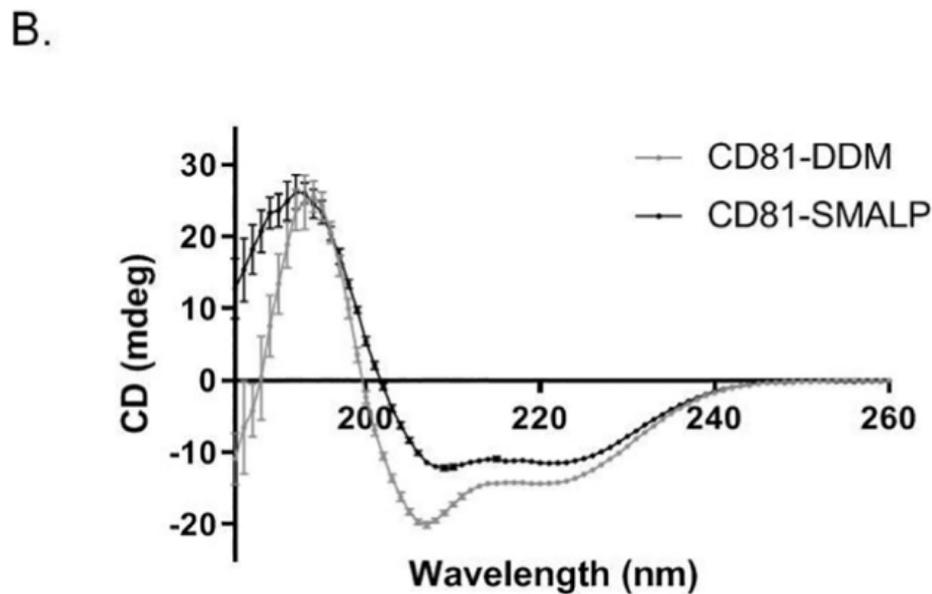
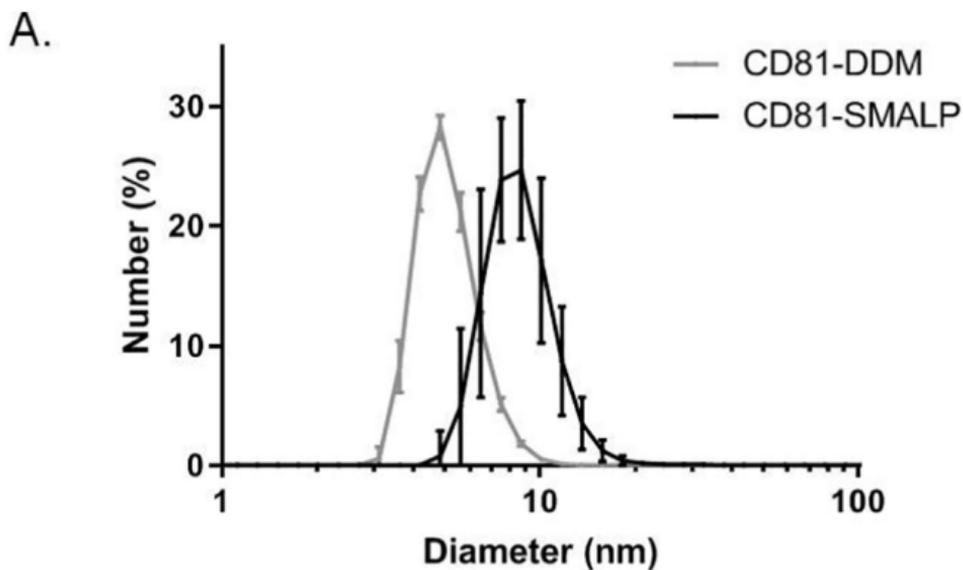


Figure 4

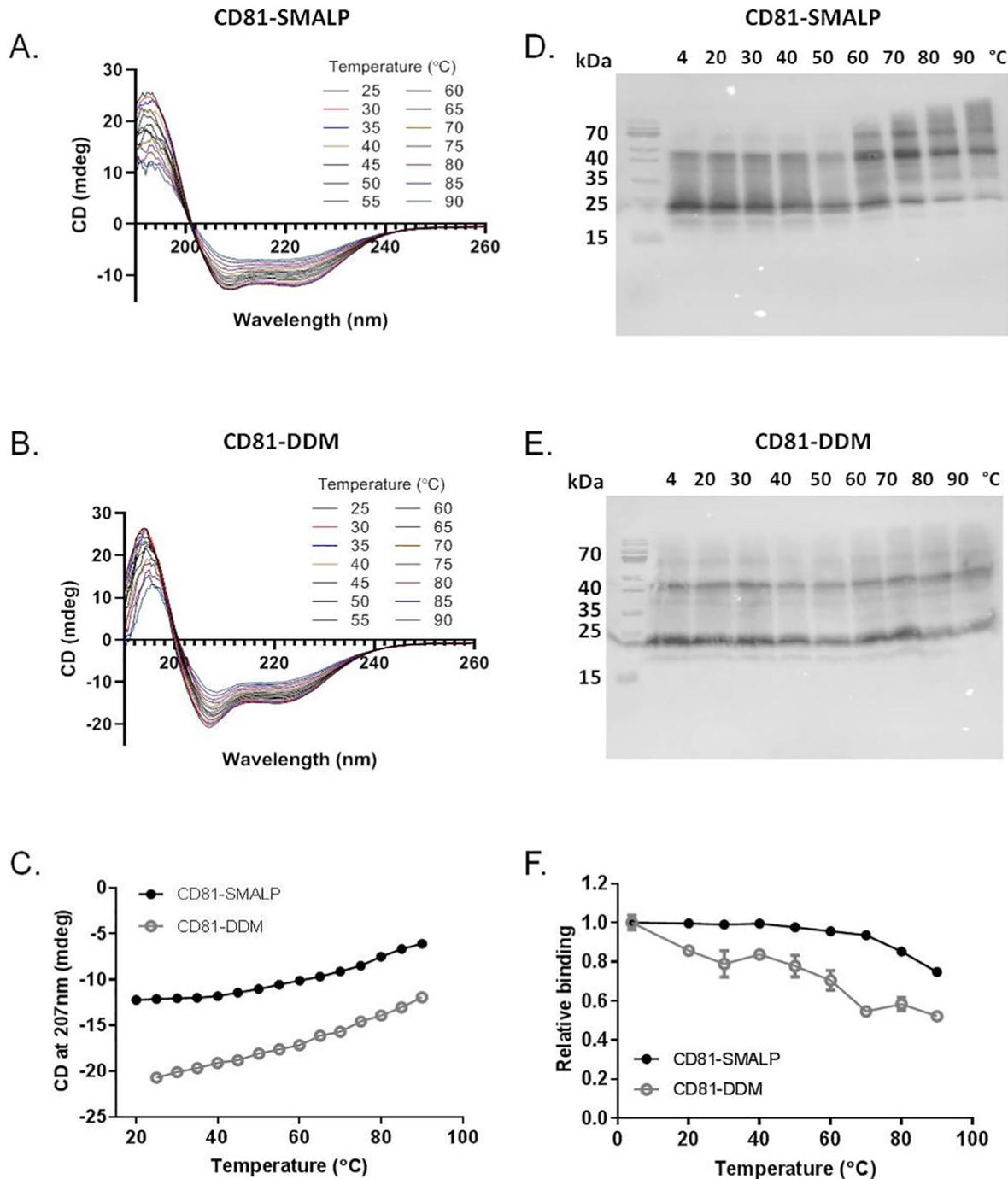
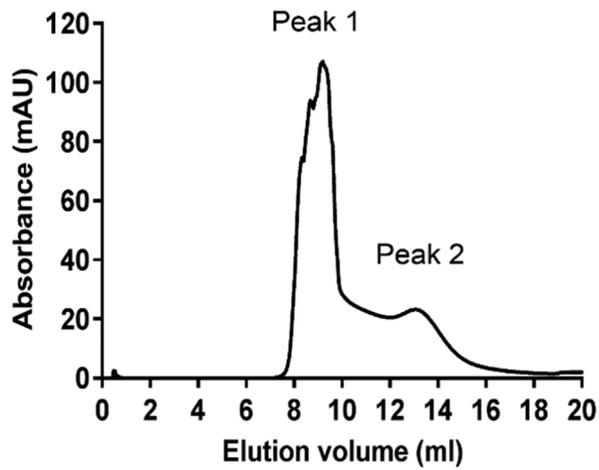


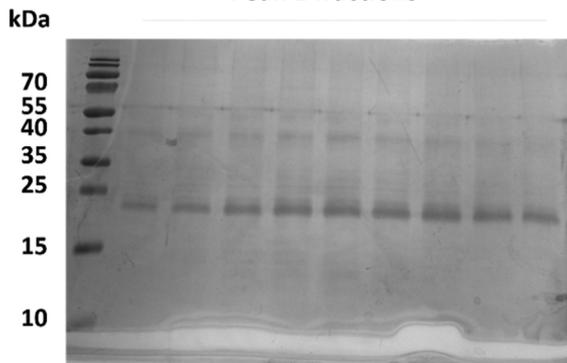
Figure 5

A.



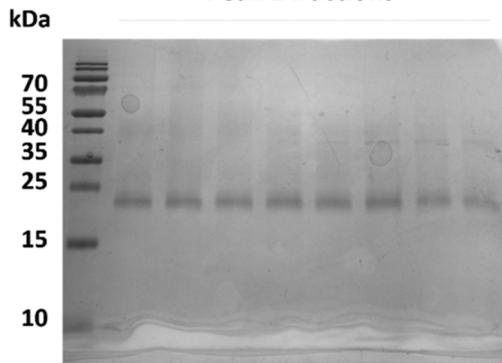
B.

Peak 1 fractions



C.

Peak 2 fractions



D.

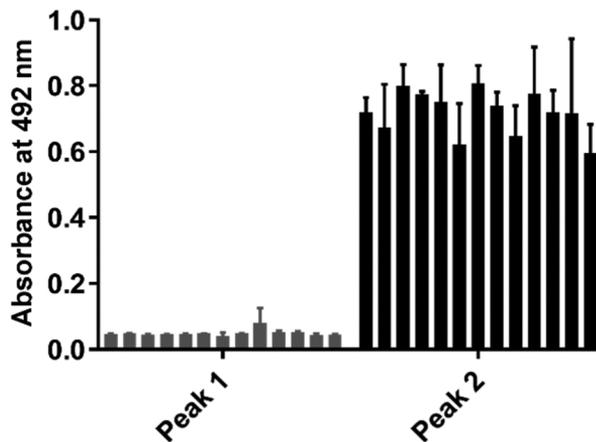


Figure 6

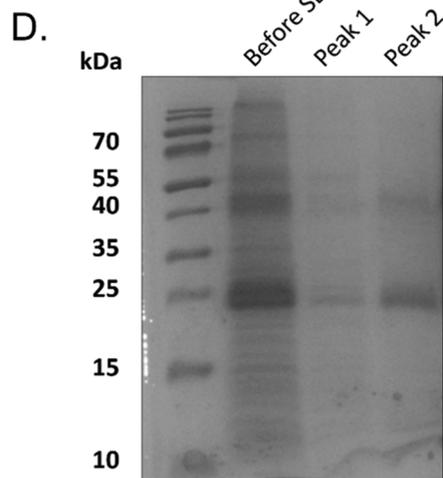
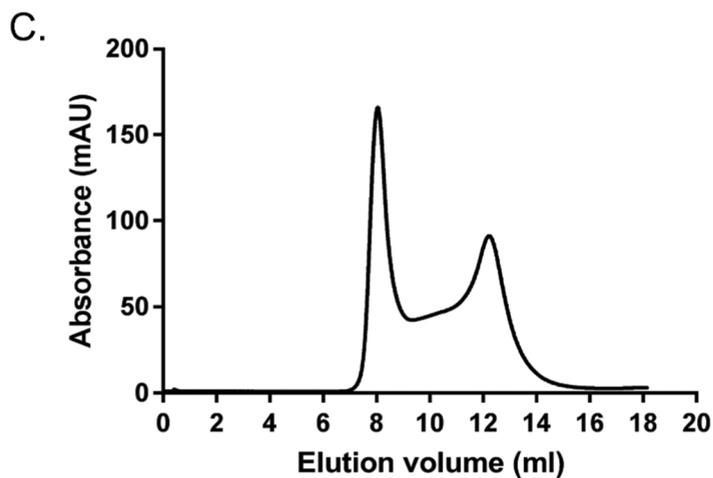
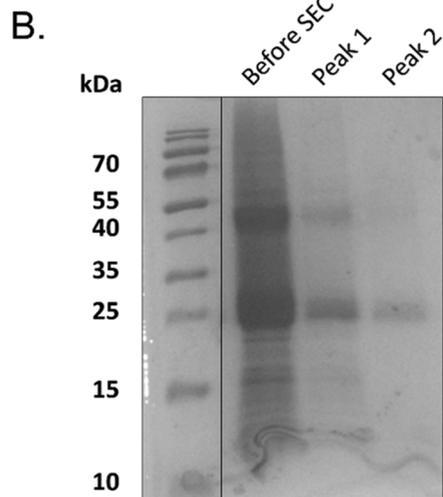
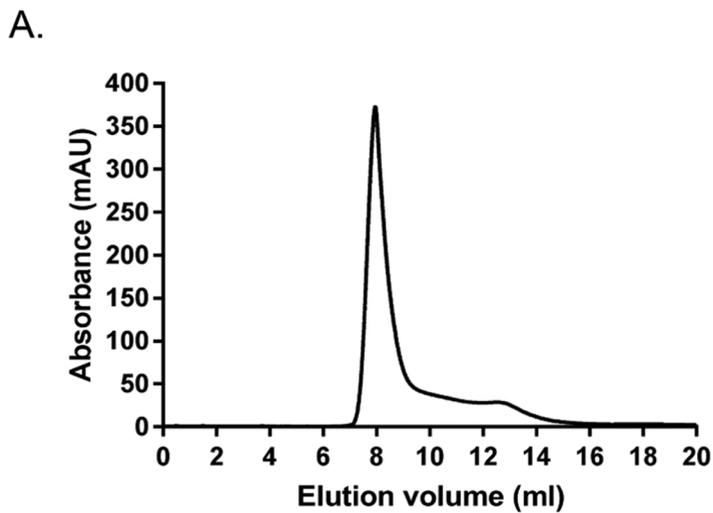
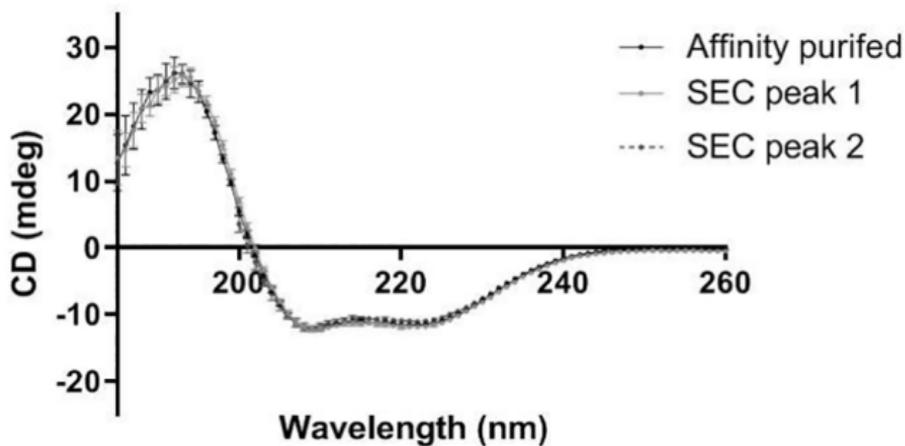


Figure 7

A.



B.

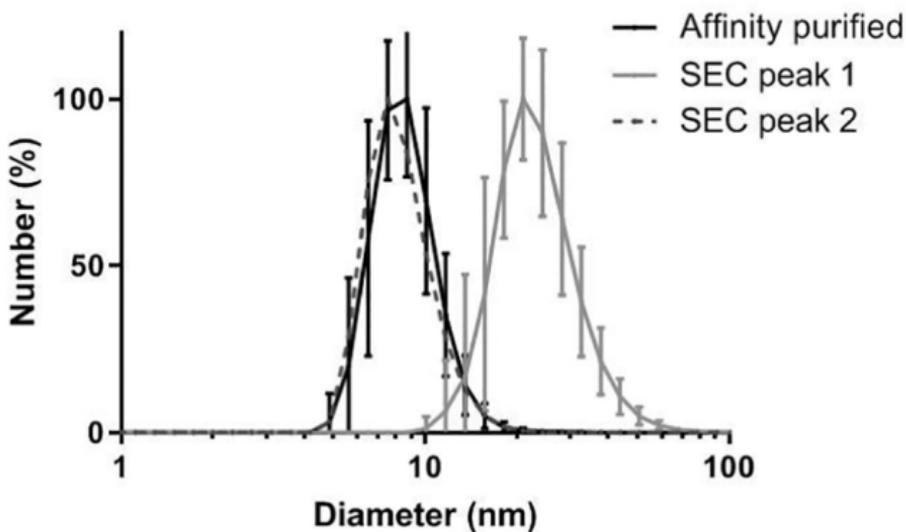


Figure 8

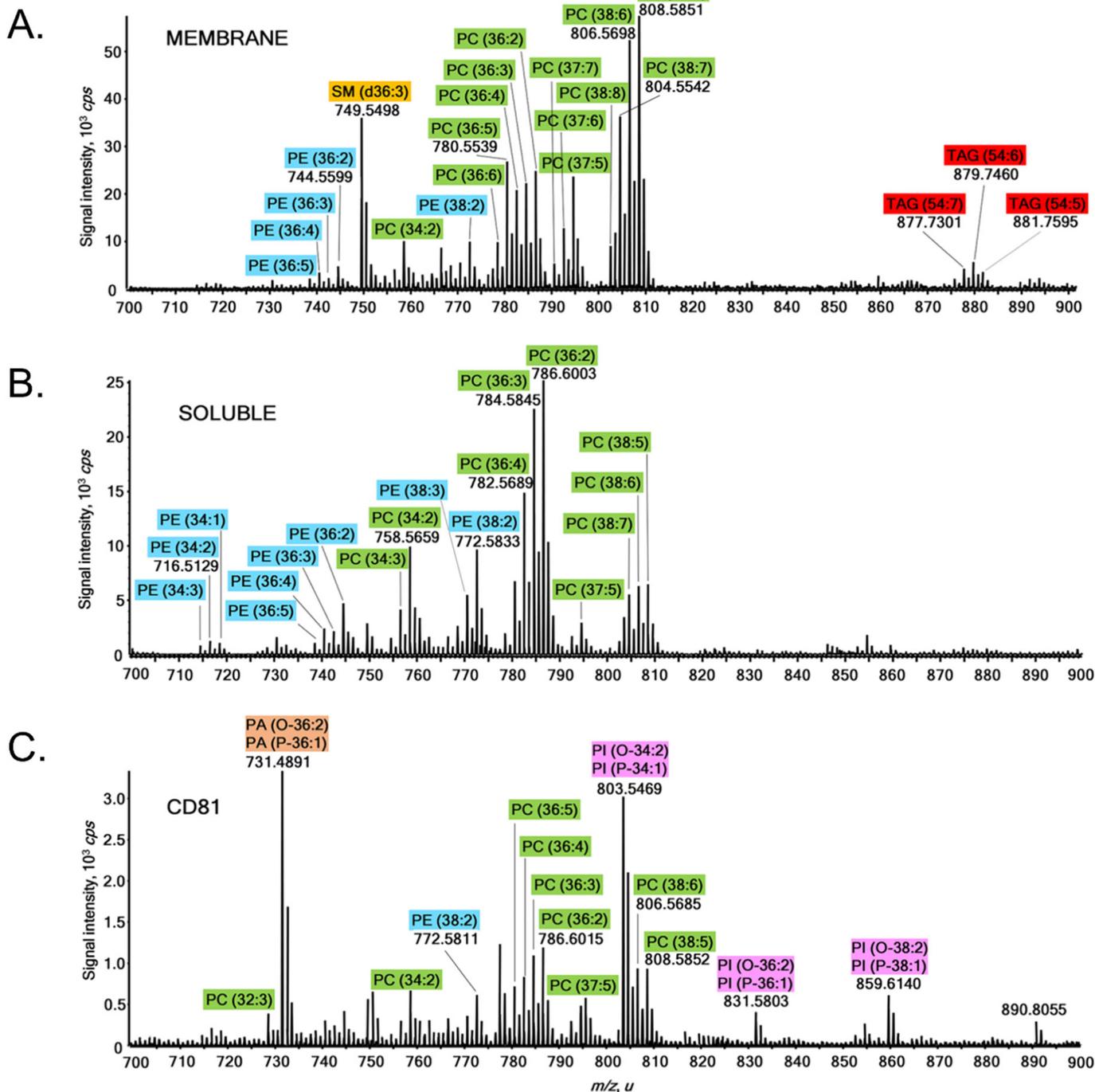
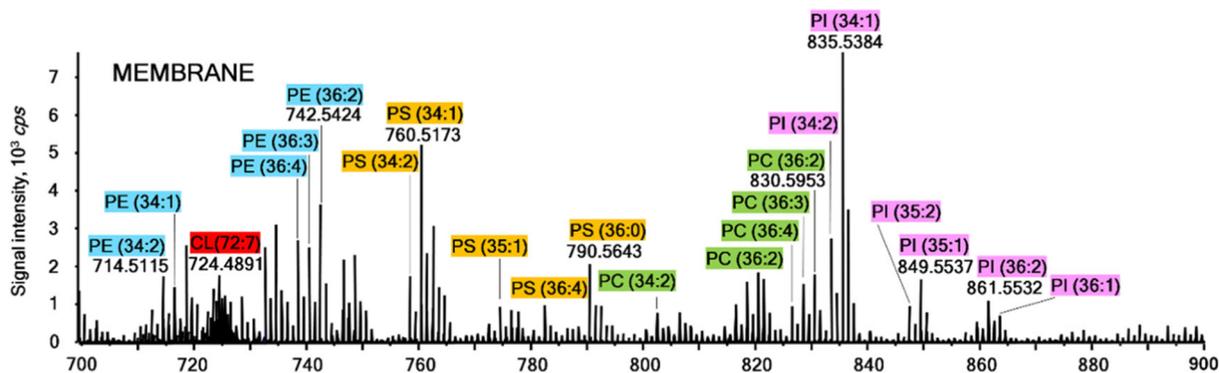
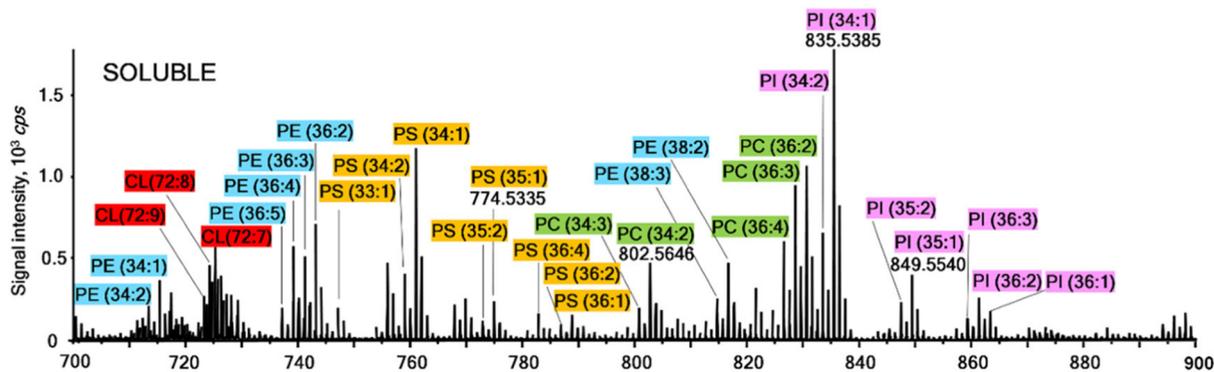


Figure 9

A.



B.



C.

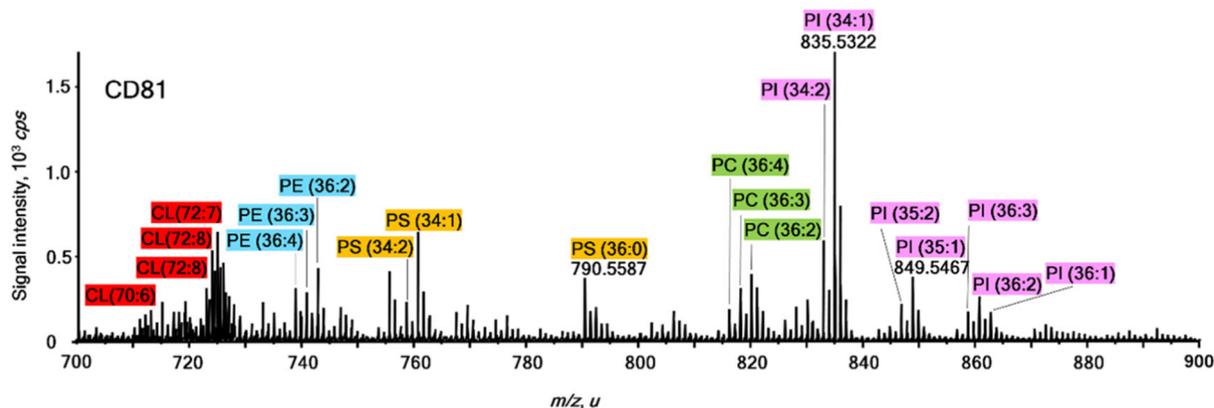


Figure 10