

Supporting Information

Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs

Michael T. Marty, Kin Kuan Hoi, Joseph Gault, and Carol V. Robinson*

anie_201508289_sm_miscellaneous_information.pdf

Contents

Supplemental Methods

Materials. All materials were from Sigma Aldrich (St. Louis, MO) unless otherwise stated. All chromatography columns were acquired from GE Healthcare (Piscataway, NJ). Concentration was preformed using an Amicon 100 kDa molecular weight cutoff concentrator. Powdered lipids 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) or 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were acquired from Avanti Polar Lipids (Alabaster, AL). Lipids were dissolved in chloroform and stored at -20 $^{\circ}$ C. Concentration of lipid was determined using phosphate analysis.^[1] Prior to reconstitution in Nanodiscs, lipids were dried under nitrogen gas and stored in a vacuum desiccator overnight. Lipid films were solubilized by addition of 0.1 M sodium cholate solution followed by vortexing, sonication in a hot water bath, and freezing.

Protein Expression and Purification. All protein and Nanodisc manipulation and purification was performed at 4 °C unless otherwise stated. Membrane scaffold proteins, MSP1D1 and MSP1E3D1 were expressed and purified as previously described using a HisPrep FF 16/10 column.^[2] Following purification, overnight cleavage with TEV protease followed by reverse immobilized metal affinity chromatography (IMAC) was used to remove the polyhistadine tag. Cleaved products are referred to as MSP1D1(−) and MSP1E3D1(−). Cleaved MSP was dialyzed in buffer (20 mM Tris and 0.15 M NaCl) overnight before being concentrated, flash frozen, and stored at -80 °C.

AqpZ and AmtB were prepared as previously described.^[3] Briefly, 10X His – MBP – TEV – AmtB and AqpZ – TEV – GFP – 6X His were expressed in *E. coli*, and membranes were isolated by ultracentrifugation. Membrane proteins were extracted

from the membrane using overnight incubation in 1% octyl-glucoside (OG) dissolved in loading buffer (50 mM Tris, 0.2 M NaCl, 20 mM imidazole, 10% glycerol, 0.025% DDM and 5 mM BME). Following filtration, membrane protein extract was loaded on a 5 mL His-Trap HP column in loading buffer, washed with loading buffer, and eluted with elution buffer (50 mM Tris, 0.2 M NaCl, 0.5 M imidazole, 10% glycerol, 0.025% DDM and 5 mM BME). The eluent was then concentrated and purified by size exclusion chromatography (SEC) using a Superdex 200 16/600 column equilibrated in loading buffer. The purified membrane protein oligomers were concentrated, flash frozen in liquid nitrogen, and stored at -80 °C.

Nanodisc Preparation and Purification. Nanodiscs were prepared as previously described.^[2, 4] Briefly, MSP and purified membrane proteins were thawed on ice and combined with cholate-solubilized lipids. The ratio of membrane protein:MSP was 1:20, and the ratio of lipid:cholate was 1:2. Ratios of MSP:lipid were 1:55 POPC:MSP1D1(−), 1:70 DMPC:MSP1D1(−), and 1:120 POPC:MSP1E3D1(−). Additional cholate or buffer was added to bring the final cholate concentration to 23 mM. For DMPC, the assembly was performed at room temperature (20-25 °C). For POPC, the assembly was performed at 4 °C. The assembly mixture was incubated for 2 hours before addition of Amberlite XAD-2 hydrophobic beads at roughly 1:1 v/v ratio, which removed the detergent and drove the Nanodisc assembly.

Following overnight incubation with Amberlite XAD-2 beads, the reconstituted Nanodiscs were filtered and loaded onto a 1 mL His-Trap HP column in buffer A (40 mM Tris, 0.3 M NaCl, 20 mM imidazole) to remove excess empty Nanodiscs. Nanodiscs were washed with 5% buffer B (40 mM Tris, 0.3 M NaCl, 0.5 M imidazole)

before elution with 100% buffer B. The sample was then concentrated, filtered, and injected on a Superose 6 Increase 10/300 SEC column equilibrated in buffer A.

The peak fractions purified by SEC were incubated overnight with TEV protease in a 1:100 w/w ratio with 5 mM BME to cleave the respective MBP and GFP tags. The cleaved membrane protein Nanodiscs were purified by reverse IMAC in buffer A, concentrated, filtered, and injected on a Superose 6 Increase 10/300 column equilibrated in 0.2 M ammonium acetate.

Peak fractions were pooled, and the concentration was measured using the absorbance at 280 nm (A280). If necessary, samples were concentrated to between 1 and 5 AU prior to mass spectrometry (MS) analysis. In some cases, residual sodium adducts were observed by mass spectrometry, specifically on the low *m/z* lipid peaks. For these samples, an additional round of purification using a centrifugal buffer exchange column (Micro Bio-Spin 6, Bio-Rad, Hercules, CA) equilibrated in 0.2 M ammonium acetate was performed. Concentrated membrane protein Nanodiscs in ammonium acetate were generally used fresh but could also be flash frozen and stored at -80 °C.

Following purification, Nanodisc complexes were analyzed by SDS-PAGE to ensure homogeneity and incorporation of membrane protein and MSP. Nanodisc complexes were also characterized by SEC using a Superdex 200 Increase 3.2/300 column. Stokes diameters were calibrated using five gel filtration standards: carbonic anhydrase, albumin, alcohol dehydrogenase, apoferritin, and thyroglobulin. MSP1E3D1(−) Nanodiscs had an average Stokes diameter of 12.0±1.0 nm, which is in excellent agreement with the known 12.1 nm Stokes diameter of POPC MSP1E3D1 Nanodiscs. MSP1D1(−) Nanodiscs had an average Stokes diameter of 12.2±0.9 nm, which is larger than the 9.7 nm previously measured for POPC

MSP1D1 Nanodiscs.^[5] This may be due to the stacking of Nanodiscs into dimeric structures, which has been observed previously by electron microscopy.^[6] The Stokes diameter is too small for the formation of larger disc complexes with more than two MSP molecules.^[7] Most importantly, the consistent similarities between mass spectra from the MSP1E3D1(−) and MSP1D1(−) indicate that this interaction does not influence the gas-phase behavior of the complexes.

Three independent replicate samples were prepared for each type of Nanodisc for each membrane protein complex. Each of these was characterized and analyzed independently to establish reproducibility.

Mass Spectrometry. Mass spectrometry was performed using a modified Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany) optimized for analysis of high *m/z* ions. Modifications included a lower frequency RF applied to the quadrupole, higher pressures in the HCD cell, and modified electronics for extended frequency range detection on the Orbitrap.^[8] Ionization was performed in positive mode using a static nanospray source with gold coated capillary needles prepared using a Model P-97 (Sutter Instrument Co.) capillary puller.^[9] A small amount of backing pressure was used on some samples to improve the nanospray. The capillary voltage was set at 1.4 kV, and the ion transfer capillary was set to 200 °C. Although this temperature seems excessive, evaporative cooling^[10] and the short time spent in the transfer capillary likely limit the thermal energy actually imparted to the complex. No additional insource collision energy was applied. The S-lens RF potential was set to maximum to bias towards higher *m/z*, and the quadrupole was set to select 2,500-20,000 *m/z*. Native mode was used to trap ions in the HCD cell.

Collisional activation was applied by adjusting the voltage applied in the HCD cell. This was ramped automatically from 0 to 190 V in 10 V steps with a 2 min collection at each step. Argon pressure in the HCD cell was set at 1.5×10^{-9} mbar. The C-trap entrance lens was set to 6.0 V. The automatic gain control was set to 1×10⁶ with a maximum inject time of 20 ms. Scans were collected from 2,500-20,000 *m/z* at a resolution of 17,500 at *m/z*=200 (a transient time of 64 ms) with 10 microscans summed into a single scan. Calibration was performed using CsI clusters.

The noise level parameter was set to 3, which is lower than the default of 4.68, to avoid automatic removal of the baseline that forms naturally for Nanodisc complexes^[11] as a result of ions overlapping in m/z . All spectra were summed over the two minute acquisition but with no additional processing.

Analysis of Mass Spectrometry Data. Data was converted from Thermo raw files to mzML using ProteoWizard.^[12] Scans over each two minute collection window were summed and converted into a single text file using Python and pymzml.^[13] Deconvolution was performed using UniDec software, which is available for download at http://unidec.chem.ox.ac.uk.^[14] Minor baseline subtraction was performed prior to deconvolution using a curved background as previously described.^[15] No other smoothing, linearization, or processing of the data was performed. Deconvolution was performed similarly to as detailed previously for empty Nanodisc complexes.^[14] The Bayesian priors were smoothed for the charge state distribution and the mass distribution, which used the mass of a lipid as the expected mass difference. The charge state was limited to between 1 and 25, and masses were limited to below 220 kDa. No significant peaks were observed outside

this range. A split Gaussian/Lorentzian peak shape was used with a full width at half maximum (FWHM) of 5 Th. Masses were generally sampled by interpolation at 10 Da intervals. Because the FWHM of the high-mass peaks in the zero-charge mass spectra was generally around 100 Da, 10 Da sampling was sufficient to accurately define the peak while minimizing computation time. Exact masses used for mass defect analysis were determined using a 1 Da interval. Mass accuracy was judged by comparing the deconvolved mass of MSP with the mass predicted from the sequence. We observed that the measured masses were generally within one bin (either 10 or 1 Da depending on the resolution of the deconvolution) of the expected masses.

Kendrick mass analysis was performed using the Kendrick Mass Tools module in UniDec. Deconvolved mass distributions from HCD/CID spectra from 50- 190V were summed to yield an average mass distribution. We defined a normalized mass defect as: $\frac{M}{M_l}-Floor\left(\frac{M}{M_l}\right)$ $\frac{M}{M_l}$), where *M* is the measured mass and *M_I* is the mass of the lipid species. To generate a square plot, a matrix of intensity values was created, where each row was specified by the zero-charge mass rounded down to the nearest integer multiple of the lipid mass, defined as: $M_l * Floor \left(\frac{M}{M}\right)$ $\frac{m}{M_l}$). Each column was specified by the normalized mass defect as defined above. Intensity values corresponded to the intensity of the zero-charge mass spectra for each *M*. To calculate the average mass defect for the high-mass peaks, the matrix of intensity values was summed along each column for all mass values greater than 90 kDa. The average of triplicate sample replicates was normalized and fit to a Gaussian to determine the mass and standard deviation of the mass defect (Table S1). Manual inspection of the mass defects of individual spectra showed very good agreement with the averages shown in Figure 3 and S4. Moreover, low mass peaks for MSP

and membrane protein monomers are in excellent agreement with predicted mass defects.^[2-3]

Lipid distributions were extracted from the zero-charge mass spectrum as the maximum intensity within a window of 200 Da from the predicted mass for the membrane protein oligomer plus 0-120 lipids. Average lipid distributions were calculated for each protein by averaging all nine lipid distributions (three replicates of three types of Nanodisc). These average distributions were fit to overlapping Gaussian distributions using a least squares approach. A pairwise F-test was used to test models with 1, 2, 3, or 4 Gaussian distributions plus a constant background. AmtB showed significance for up to three independent distributions, and AqpZ showed significance for up to four. To keep both models consistent, we used three Gaussian distributions for both. We report the mean and standard deviation of the Gaussian distributions.

Molecular Dynamics. Molecular dynamics of AmtB and AqpZ were performed using $VMD^{[16]}$ and NAMD.^[17] NAMD and VMD were developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. A 96 $\AA \times 96$ Å patch of POPC membrane was created by the VMD Membrane Plugin using CHARMM27.^[18] AmtB trimer^[19] and Aqp $Z^{[20]}$ tetramer were generated from crystal structure (PDB ID: 2NPE and 1RC2 respectively) and parameterized using the VMD Automatic PSF Generation Plugin. Membrane protein complexes were aligned manually in the membrane. All lipids with at least one atom within 0.5 Å of the protein were deleted. The system was solvated and ionized with 0.15 M NaCl.

MD was performed using the CHARMM27 force field at 300 K and 1 bar using the same simulation setup and equilibration protocol as previously described.^[21] The first 20 ns of the simulations were discarded to allow sufficient equilibration of the system. For analysis, the simulation was sampled at 0.1 ns intervals from 20 ns to 100 ns.

To classify interactions between lipids and the membrane protein, we defined the minimum distance metric between two sets of atoms as the shortest pairwise distance between any atom in the first set, a subset of atoms of a single lipid, and any atom in the second set, a subset of protein atoms. Thus, the minimum distance is determined for each lipid molecule separately and depends entirely on the shortest contact with the protein. Lipids in the annular belt were classified using heavy atom (non-hydrogen) subset in both lipids and the protein, but similar numbers were obtained using all atoms. The lipid head group was defined to include the fatty acid carboxyl, the glycerol, and the phosphatidylcholine moieties. Minimum distance for lipid head groups was also calculated using heavy atoms in the protein and the lipids. For lipid anions, distances were measured between the two singly-bonded oxygen atoms on the lipid phosphate and nitrogen atoms on the side chains of lysine or arginine as well as on the N-terminus. For lipid cations, distances were measured between the methyl carbons on the choline and carboxyl oxygen atoms on the side chains of glutamate or aspartate residues or on the C-terminus.

The minimum distance between each lipid and the membrane protein oligomer was calculated for each type of interaction in each frame. A histogram of all the minimum distances allowed an appropriate cut-off to be selected at a local minimum following the first main peak in the probability density function (See Figures S5). Cumulative density functions were also examined to confirm that there was a

plateau. The cut-offs were selected as 5.0 Å for the annular belt, 4.6 Å for head groups, 3.25 Å for lipid anionic interactions, and 4 Å for lipid cationic interactions. In each frame, the number of lipids meeting each cut-off was counted. The mean and standard deviation (Table S2) in the number of lipids of each class was calculated for all frames from 20 ns to 100 ns (Figure S6), and a histogram at 1 lipid resolution was generated (Figure 4 and S7).

Supplemental Tables

Table S1. Normalized mass defect of AmtB trimer and AqpZ tetramer with varying types of Nanodiscs. Mean and standard deviation of the mass defects were calculated by fitting the average mass defect of three replicates, all collision energies, and all masses greater than 90 kDa (Figure S4) to Gaussian distributions. Predicted mass defects were calculated using masses of 126,783 Da for the AmtB trimer, 98,904 for the AqpZ tetramer, 760.076 for POPC, and 677.933 for DMPC. AmtB and AqpZ masses were determined internally from the deconvolutions, and lipid masses were used as the average mass reported by Avanti Polar Lipids. Although the mass defects are systematically slightly higher than predicted, this is likely due to adduction of solvent or ions.^[22]

Table S2. The mean and standard deviation of the number of lipids determined by MD for the lipid annular belt, the annular belt considering only head groups, and the ionic contacts for frames collected at 0.1 ns intervals from 20 ns to 100 ns (Figure S6). Peak 1, 2, and 3 are listed as the mean and standard deviation of three overlapping Gaussian distributions fit to the MS lipid distributions (Figure 4).

Supplemental Figures

Figure S1. Representative mass spectrum and deconvolution of AqpZ Nanodiscs with POPC and MSP1D1(−). The mass spectrum at 80V CID (A) deconvolves to yield the mass distribution (B). Increasing the collision voltage causes the mass distribution to shift to lower mass species (C). The average mass distribution over all CID states (D) shows three distinct species (E): AqpZ trimer with a broad mass distribution including tens of bound lipids (*red*), AqpZ monomer (*green*), and MSP (*blue*). These three states are annotated in A, B, and D. Two separate lipid bound states of AqpZ trimer are shown in E are the full annular belt complex (*red left*) and the annular belt considering only head group interactions (*red right*).

Figure S2. Mass spectra of AmtB-POPC-MSP1D1(−), AmtB-POPC-MSP1E3D1(−), AmtB-DMPC-MSP1D1(−), AqpZ-POPC-MSP1D1(−), AqpZ-POPC-MSP1E3D1(−), and AqpZ-DMPC-MSP1D1(−) Nanodiscs collected from 50 V to 190 V.

Figure S3. Zero-charge mass spectra of AmtB and AqpZ Nanodiscs deconvolved from corresponding spectra shown in Figure S2.

Figure S4. Average mass defect values and lipid distribution in AmtB and AqpZ Nanodiscs. Mass defects for AmtB (*top left*) and AqpZ (*bottom left*) were averaged over all mass species greater than 90 kDa from three replicates for 15 spectra from 50-190V CID for Nanodiscs formed with POPC and MSP1D1(−) (*blue*), POPC and MSP1E3D1(−) (*blue dashed*), and DMPC and MSP1D1(−) (*red*). Vertical lines show predicted mass defect for oligomer complexes in POPC (*blue*) and DMPC (*red*). Extracted lipid distributions averaged from the same spectra are shown in the *top right* and *bottom right* respectively.

Figure S5. Protein-lipid minimum distance distributions and cutoffs extracted from MD. In each subplot, the top shows the normalized probability density function of minimum distances between a subset of atoms in each lipid and a subset of atoms in the protein for frames collected at 0.1 ns intervals from 20 ns to 100 ns. The bottom shows the corresponding cumulative density function averaged over the same frames. Specific subsets are: all non-hydrogen atoms in the lipid and non-hydrogen atoms in the protein (A); non-hydrogen atoms in the lipid head group and nonhydrogen atoms in the protein (B); choline methyl carbon atoms on the lipid and carboxylate oxygen atoms on the protein (C); and phosphate oxygen atoms on the lipid and nitrogen atoms on sidechains of lysine and arginine residues on the protein (D). The cutoffs for each bond distance are marked with vertical dotted lines at 5, 4.6, 4, and 3.25 Å respectively. Results for AqpZ and AmtB are shown in *blue* and *green* respectively.

Figure S6. Dynamics of the number of lipids in contact with the membrane protein oligomer in MD simulation for AmtB (*left*) and AqpZ (*right*). Traces are shown for lipids classified as belonging to the lipid annular belt (*purple*), the annular belt when considering only head groups (*blue*), and the ionic contacts (*cyan*) based on cutoffs defined in Figure S5.

Figure S7. Comparison of lipid distributions measured by MS with distributions predicted by MD. Lipid distributions extracted as function of collision voltage are shown in A and B for AmtB and AqpZ respectively with the corresponding distribution in number of lipids in the annular belt (*purple*), the number of lipids in the annular belt when considering only head groups (*blue*), and the number of ionic contacts (*cyan*) predicted by MD in C and D.

Supplemental References

- [1] P. S. Chen, T. Y. Toribara, H. Warner, *Anal. Chem.* **1956**, *28*, 1756-1758.
- [2] T. K. Ritchie, Y. V. Grinkova, T. H. Bayburt, I. G. Denisov, J. K. Zolnerciks, W. M. Atkins, S. G. Sligar, in *Methods Enzymol., Vol. Volume 464* (Ed.: D. Nejat), Academic Press, **2009**, pp. 211-231.
- [3] A. Laganowsky, E. Reading, T. M. Allison, M. B. Ulmschneider, M. T. Degiacomi, A. J. Baldwin, C. V. Robinson, *Nature* **2014**, *510*, 172-175.
- [4] T. H. Bayburt, Y. V. Grinkova, S. G. Sligar, *Nano Lett.* **2002**, *2*, 853-856.
- [5] I. G. Denisov, Y. V. Grinkova, A. A. Lazarides, S. G. Sligar, *J. Am. Chem. Soc.* **2004**, *126*, 3477-3487.
- [6] J. T. S. Hopper, Y. T.-C. Yu, D. Li, A. Raymond, M. Bostock, I. Liko, V. Mikhailov, A. Laganowsky, J. L. P. Benesch, M. Caffrey, D. Nietlispach, C. V. Robinson, *Nat. Meth.* **2013**, *10*, 1206-1208.
- [7] Y. V. Grinkova, I. G. Denisov, S. G. Sligar, *Protein Eng. Des. Sel.* **2010**, *23*, 843-848.
- [8] a) A. Dyachenko, G. Wang, M. Belov, A. Makarov, R. N. de Jong, E. T. J. van den Bremer, P. W. H. I. Parren, A. J. R. Heck, *Anal. Chem.* **2015**, *87*, 6095-6102; b) R. J. Rose, E. Damoc, E. Denisov, A. Makarov, A. J. R. Heck, *Nat. Meth.* **2012**, *9*, 1084- 1086.
- [9] H. Hernandez, C. V. Robinson, *Nat. Protoc.* **2007**, *2*, 715-726.
- [10] a) J. A. Silveira, K. L. Fort, D. Kim, K. A. Servage, N. A. Pierson, D. E. Clemmer, D. H. Russell, *J. Am. Chem. Soc.* **2013**, *135*, 19147-19153; b) S.-W. Lee, P. Freivogel, T. Schindler, J. L. Beauchamp, *J. Am. Chem. Soc.* **1998**, *120*, 11758-11765.
- [11] M. T. Marty, H. Zhang, W. Cui, M. L. Gross, S. G. Sligar, *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 269-277.
- [12] D. Kessner, M. Chambers, R. Burke, D. Agus, P. Mallick, *Bioinformatics* **2008**, *24*, 2534-2536.
- [13] T. Bald, J. Barth, A. Niehues, M. Specht, M. Hippler, C. Fufezan, *Bioinformatics* **2012**, *28*, 1052-1053.
- [14] M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch, C. V. Robinson, *Anal. Chem.* **2015**, *87*, 4370-4376.
- [15] N. Morgner, C. V. Robinson, *Anal. Chem.* **2012**, *84*, 2939-2948.
- [16] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33-38.
- [17] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781-1802.
- [18] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545-1614.
- [19] A. Javelle, D. Lupo, L. Zheng, X.-D. Li, F. K. Winkler, M. Merrick, *J. Biol. Chem.* **2006**, *281*, 39492-39498.
- [20] D. F. Savage, P. F. Egea, Y. Robles-Colmenares, J. D. O. C. Iii, R. M. Stroud, *PLoS Biol.* **2003**, *1*, e72.
- [21] C. Bechara, A. Nöll, N. Morgner, M. T. Degiacomi, R. Tampé, C. V. Robinson, *Nat. Chem.* **2015**, *7*, 255-262.
- [22] A. R. McKay, B. T. Ruotolo, L. L. Ilag, C. V. Robinson, *J. Am. Chem. Soc.* **2006**, *128*, 11433-11442.