MDPI

*Article*

# The Droserasin 1 PSI: A membrane-interacting antimicrobial peptide from the carnivorous plant *Drosera capensis*

**Marc A. Sprague-Piercy[1,†], Jan C. Bierma[1,†], Marquise G. Crosby[1], Brooke P. Carpenter[2], Gemma R. Takahashi[1], Joana Paulino[3], Ivan Hung[3], Rongfu Zhang[3], John E. Kelly[2], Natalia Kozlyuk[2], Xi Chen[2], Carter T. Butts[4], and Rachel W. Martin[1,2,*]**

[1] Department of Molecular Biology and Biochemistry, University of California, Irvine
[2] Department of Chemistry, University of California, Irvine
[3] National High Magnetic Field Laboratory, Tallahassee, Florida
[4] Departments of Sociology, Statistics, and Electrical Engineering and Computer Science, University of California, Irvine
* Correspondence: rwmartin@uci.edu
† These authors contributed equally to this work.

Version July 7, 2020 submitted to Biomolecules

**Abstract:** The Droserasins, aspartic proteases from the carnivorous plant *Drosera capensis*, contain a 100-residue plant-specific insert (PSI) that is post-translationally cleaved and independently acts as an antimicrobial peptide. PSIs are of interest not only for their inhibition of microbial growth, but also because they modify the size of lipid vesicles and strongly interact with biological membranes. PSIs may therefore be useful for modulating lipid systems in NMR studies of membrane proteins. Here we present the expression and biophysical characterization of the Droserasin 1 PSI (D1 PSI.) This peptide is monomeric in solution and maintains its primarily $\alpha$-helical secondary structure over a wide range of temperatures and pH values, even under conditions where its three disulfide bonds are reduced. Vesicle fusion assays indicate that the D1 PSI strongly interacts with bacterial and fungal lipids at pH 5 and lower, consistent with the physiological pH of *D. capensis* mucilage. It binds lipids with a variety of head groups, highlighting its versatility as a potential stabilizer for lipid nanodiscs. Solid-state NMR spectra collected at a field strength of 36 T, using a unique series-connected hybrid magnet, indicate that the peptide is folded and strongly bound to the membrane. Molecular dynamics simulations indicate that the peptide is stable as either a monomer or a dimer in a lipid bilayer. Both the monomer and the dimer allow the passage of water through the membrane, albeit at different rates.

**Keywords:** antimicrobial peptide; membrane protein; lipid-protein interactions; solid-state NMR; *Drosera capensis*; carnivorous plant

## 1. Introduction

Plant carnivory has independently evolved multiple times from defense mechanisms involving the jasmonate pathway [1,2]. In most cases, the leaves are modified for capturing prey, using a variety of mechanisms including pitfall traps, the unique snap trap mechanism of the Venus flytrap, and sticky flypaper traps, among others [3]. All of these plants must perform their digestion without any mechanical breakdown of prey tissues, such as mastication. In species with flypaper traps, such as the *Drosera*, digestion generally occurs in an exposed environment over a prolonged period of time, without the benefit of a pitcher or closed trap. These plants catch prey in the sticky polysaccharide mucilage of their leaf tentacles, which then wrap around the meal to increase contact with the digestive

mucilage [4]. The digestion is thus exposed to variable physical conditions from changes in weather and has increased risk of opportunistic microbial growth from bacteria and fungi that compete for nutrients from the captured prey and potentially infect the plant tissue, causing disease [5]. In the Droseracaea, the evolution of carnivory was accompanied by the loss of many genes common to other plants, and the concomitant expansion of genes specifically related to carnivory [6]. For these reasons, carnivorous plants are a potential source of novel and useful antimicrobial peptides as well as digestive enzymes. Here, we focus on an antimicrobial peptide discovered from the genome of the Cape sundew, *Drosera capensis*. This plant is relatively easily cultivated and has been the target of genome sequencing [7] and enzyme discovery [8–10] efforts.

Analysis of putative digestive enzyme sequences from *D. capensis* revealed several aspartic proteases that contain a segment of about 100 residues called a plant specific insert (PSI). PSIs are mostly-helical domains that are often cleaved off during maturation and act as independent proteins [11]. Structurally, PSIs are categorized as saposin-like proteins, a protein family whose members have substantial sequence diversity but share a strongly conserved, compact tertiary fold, usually stabilized by three disulfide bonds [12]. Although the *S. tuberosum* PSI is monomeric in solution [13], it was crystallized as a dimer [11] and appears to oligomerize upon interaction with anionic membranes [14,15]. Membrane permeabilization can induce cytotoxicity and is a common antimicrobial defense. PSI-containing aspartic proteases from *D. capensis* may have been recruited for digestive function due to their ability to digest prey proteins as well as inhibit microbial growth during digestion. These functions may be separate or synergistic: PSIs in the digestive mucilage could simply serve to suppress the growth of pathogens, or may also function to make insect lipids more available for digestion.

In other plants, aspartic proteases containing PSIs are implicated in stress responses, senescence, and pathogen responses [16–18]. Their overall function involves interacting with lipid membranes in a variety of ways, including membrane localization, increasing the availability of membrane lipids for enzymatic processing, and permeabilizing membranes [19]. For example, in the cardoon, the PSI from the aspartic protease cardosin A plays a role in vacuole localization [20,21] and induces vesicle leakage below pH 5.5, whether or not it is attached to the parent aspartic protease [22]. Similarly, the well-studied PSI from the potato *Solanum tuberosum* disrupts lipid vesicles and bilayers [11] and has been shown to exhibit antimicrobial activity against plant and human pathogens [23]. Here we report the expression and characterization of recombinantly expressed Droserasin 1 PSI (D1 PSI) from *D. capensis*. Our results show that the D1 PSI forms a compact, stable structure and is indeed capable of disrupting and permeabilizing membranes and inhibiting microbial growth.

## 2. Materials and Methods

### *Sequence Alignment and Clustering*

All sequences from the *Drosera capensis* genome [7] and the *Dionaea muscipula* transcriptome [24] that were previously annotated as coding for MEROPS A1 aspartic proteases using the MAKER-P (v2.31.8) pipeline [25] and a BLAST search against SwissProt (downloaded 8/30/15) and InterProScan [26] were examined for the presence of a PSI. Sequence alignment with the previously-characterized PSI from the *Arabidopsis thaliana* protease APA1_ARATH was used for quality control; proteins that did not contain a full-length PSI were not selected for modeling or further analysis. ClustalOmega was used to produce sequence alignments [27] for putative aspartic protease PSIs. The following settings were used: gap open penalty = 10.0, gap extension penalty = 0.05, hydrophilic residues = GPSNDQERK, and a BLOSUM weight matrix. This resulted in six complete PSIs from *D. capensis* and two from *D. muscipula*. Three previously-characterized aspartic proteases from other plants are also included as reference sequences [28–30]. The resulting PSI sequences were clustered by sequence similarity.

*Structure Prediction*

Structures for the PSIs were predicted using the Robetta [31] implementation of Rosetta [32]. This software uses a combination of comparative modeling and all-atom refinement based on a simplified forcefield, yielding five structures for each protein. For the PSIs, both open and closed conformations were observed in the models: we selected the lowest-energy representative of each type of structure for each PSI. The open conformation was observed as part of a dimer structure, which was also employed for subsequent modeling of the D1 PSI as described below. This process was performed for the *D. capensis* and *D. muscipula* PSIs, as well as a reference PSIs from *Arabidopsis thaliana*. The PDB files corresponding to the predicted structures for the PSIs reported in this manuscript are available in the Supplementary Information. Protein structure figures were generated using UCSF Chimera [33] and VMD [34].

## *2.1. Gene construction, expression, and purification*

Plasmids containing the DNA sequence of the *D. capensis* Droserasin 1 PSI (D1 PSI) genes were purchased from Integrated DNA Technologies (San Diego, CA). Each gene was flanked by regions containing restriction sites for NcoI and XhoI and contained an N-terminal 6x His tag and a TEV protease cleavage sequence (ENLYFQG). The gene was amplified using oligonucleotide primers purchased from Integrated DNA Technologies (Coralville, IA), and the resulting gene product was cloned into pET28a(+) vector (Novagen, Darmstadt, Germany). D1 PSI was overexpressed in SHuffle T7 *Escherichia coli* (New England Biolabs, Ipswich, MA) using an autoinduction protocol [35] at 25 °C with the modification of adding 50 $\mu$M IPTG [36]. Cells were allowed to grow for at least 24 hours, lysed via sonication, heated at 70 °C for 20 minutes to precipitate most *E. coli* proteins, and cell debris was removed by centrifugation. His-TEV-D1 PSI was purified on a Ni-charged Bio-Scale Mini Profinity IMAC Cartridges (Bio-Rad, Hercules, CA) where bound protein was washed with low imidazole wash buffer followed by washing with 40% isopropanol wash buffer mixture and 40% DMSO wash buffer mixture to remove potentially bound lipids before elution. The His tag was removed with the use of a His-tagged TEV protease (produced in-house), over one week, (the time required to obtain a sufficient yield of His tag-free protein), followed by reapplication to Ni-charged Bio-Scale Mini Profinity IMAC Cartridges (Bio-Rad, Hercules, CA) to remove His-tagged TEV protease and uncleaved His-TEV-D1 PSI. The final purification step consisted of applying the sample to a HiLoad 16/600 Superdex 75 pg gel filtration column (GE, Pittsburgh, PA) in 10 mM phosphate buffer. D1 PSI was dialyzed into 10 mM phosphate, 0.05% $NaN_3$, pH 6.9, for all experiments unless otherwise stated. The mass of the protein was confirmed by electrospray mass spectrometry. For $^{13}C$, $^{15}N$ labeled protein samples, protein was expressed using an optimized high-cell-density IPTG-induction minimal media protocol [37]. Purification was performed in the same manner as for natural abundance protein.

## *2.2. Circular dichroism*

D1 PSI was diluted to 0.125 mg/mL with either 10 mM succinate, acetate, MES, or phosphate buffer at pH 4, 5, 6, and 7, respectively, for the collection of full circular dichroism (CD) spectra between 190 and 260 nm in a 10 mm quartz cuvette. Three accumulations were collected and no smoothing function was applied to the collected data. Additional spectra were taken under the same conditions but where the D1 PSI was reduced in 1 mM DTT before dilution into the final buffer. Measurements were taken on a J-810 spectropolarimeter (JASCO, Easton, MD) equipped with a thermal controller.

## *2.3. Fluorescence spectroscopy*

UV fluorescence measurements were made on D1 PSI at a concentration of 0.2 mg/mL under the same buffer conditions as for CD spectra including unreduced and DTT-reduced protein for full emission spectra, with an excitation wavelength of 280 nm. Spectra were taken using a Cary Eclipse Fluorescence Spectrophotometer (Agilent, Santa Clara, CA)

### 2.4. Characterization of oligomeric state

0.5 mg/mL of purified D1-PSI with 6x His-tag were incubated in 50 mM ammonium formate for pH 3.5, 50 mM ammonium acetate for pH 4, 5, and 6 or 50 mM ammonium bicarbonate for pH 7 and 8. Intact protein samples were diluted 5-fold in water and then run on the SYNAPT G2-Si Mass Spectrometer (Waters, Milford, MA) via direct injection in a 80/20% 0.1% formic acid/ACN running buffer. Samples were ionized by electrospray ionization with a capilary voltage of 2.7 kV and then separated by the T-Wave ion mobility collision cell (Waters, Milford, MA). The ion series generated was then deconvoluted using the MaxEnt1 algorithm supplied by the MassLynx (Waters Milford, MA) software package to calculate the mass of the proteins at each pH.

### 2.5. Antimicrobial assay

Antimicrobial activity of D1 PSI was measured by growing *Pichia pastoris* in standard yeast extract, peptone, dextrose media (YPD) with the addition of 50 mM sodium phosphate and adjusted to pH 4, 5, 6, or 7, with or without 25 $\mu$M D1 PSI, peptide concentration was selected to match previously reported experiments[23]. Two replicates were performed at each pH value. Cells were grown in 15 mL culture tubes in a final culture volume of 1 mL. Cultures were inoculated with a pregrown culture, resulting in a starting optical density (OD) of 0.005 at 600 nm, and were allowed to grow for 48 hours at 30 °C. After growth total cell yield was estimated by measuring the OD at 600 nm.

### 2.6. Vesicle fusion assay

Lipids used for experiments were extracted from either *E. coli* or *P. pastoris* using the Bligh and Dyer method [38]. The lipids were then dried under a stream of nitrogen gas and polar lipids were extracted using a cold acetone precipitation [39]. The *E. coli* or *P. pastoris* polar lipids were solubilized by repeated heating and cooling, between 42 °C and room temperature, in either 10 mM succinate pH 4, 10 mM acetate pH 5, 10 mM MES pH 6, or 10 mM phosphate pH 7. The solution was then run through a mini extruder equipped with a 100 nm polycarbonate membrane (Avanti Polar Lipids, Alabaster, AL) to create large unilamellar vesicles (LUV) of approximately 100 nm in diameter. Vesicle size was monitored over time at the different pH with or without 50 $\mu$M D1 PSI using dynamic light scattering on a Zetasizer Nano ZS (Malvern Instruments, Malvern, U.K.).

### 2.7. Lipid interaction quantification

1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) sodium salt (POPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine sodium salt (POPS), and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphate sodium salt (POPA) were purchased from Avanti Polar Lipids (Alabaster, AL). A solution of 10 mM sodium acetate pH 4.5 containing 75 $\mu$M of POPC, POPE, POPG, POPS, and POPA each was divided into two aliquots. One aliquot had His-tagged D1 PSI added to a final concentration of 0.04 mg/mL. Both solutions were sonicated for an hour at room temperature and applied to a Ni-charged Bio-Scale Mini Profinity IMAC Cartridges (Bio-Rad, Hercules, CA). The aliquot lacking D1 PSI had the flow-through collected, while the elution was collected from the aliquot containing D1 PSI. Lipids were extracted from the flow-through and elution respectively using the Bligh and Dyer method [38]. A series of ten-fold dilutions were made for each and the resulting samples were run on a Xevo G2-XS QTof spectrometer (Waters, Milford, MA) in positive mode, with an in-line ACUITY UPLC BEH C18 Column (Waters, Milford, MA), where lipids were eluted with a water/isopropanol gradient containing ammonium formate. Only the sample concentrations within the linear range of ion intensities were used. Ion intensities were tabulated for each lipid species and normalized to the total ion count to estimate relative lipid proportions.

### 2.8. Solid-state NMR

Samples were prepared by evaporating away chloroform from 5 mg of POPC and POPA, respectively, under a stream of nitrogen gas. Lipids were sonicated in 0.5 mL methanol to redissolve lipids then 0.5 mL of water was added followed by 2 mg of $^{13}$C, $^{15}$N labeled D1 PSI dissolved in water. Protein lipid mixture was briefly sonicated, flash frozen and lyophilized. Lyophilized sample was then hydrated with 10 $\mu$L 10 mM acetate, 0.025% sodium azide, pH 4.5 buffer. The resulting sample was cycled between 42 °C and room temperature ten times. Spectra were taken at the National High Magnetic Field Lab (Tallahassee, FL) using the 40 mm bore Series Connected Hybrid (SCH) magnet system, currently the highest-field NMR magnet [40]. Two-dimensional $^{13}$C-$^{13}$C cross polarization [41] dipolar assisted rotational resonance [42] (CP DARR) spectra were obtained at 36 T, with a 2 mm CPMAS probe tuned to frequencies of $^{1}$H, $^{13}$C, $^{15}$N, with ferroshims, at a temperature of 10 °C, a 100 ms mixing time, and a MAS rate of 24.4 kHz [40].

### 2.9. Molecular Modeling and Analysis

Predicted monomer and dimer structures of D1 PSI as described above were modeled within lipid bilayers using atomistic molecular dynamics (MD) simulations. Open monomer and PSI dimer structures were first adjusted by adding disulfide bonds based on homology to the potato (*S. tuberosum*) PSI, with protonation states corrected for pH 5 using PROPKA3 [43]. For each respective structure, a POPC membrane patch was prepared using the VMD membrane plugin; for the monomer structure, a 100 Å square patch was used, with a 150 Å patch employed for the PSI dimer. The membrane patch was then centered within a TIP3P water box [44] of dimension 80 Å normal to the patch, the PSI was added within the membrane center, and 0.1M NaCl was added. (All structure preparation was performed using VMD [34].) The prepared system was then equilibrated as follows. Initially, with all atoms other than those of the lipid tails were held fixed, the system was minimized for 5,000 iterations and simulated at 300K for 0.5 ns. (All simulations were performed in NAMD [45] using the CHARMM36 force field [46] under periodic boundary conditions in an $NpT$ ensemble at 1 atm pressure with Nosé-Hoover Langevin piston pressure control [47,48] and Langevin temperature control (damping coefficient 1/ps).) Water and lipid were then released (with the protein held fixed), and the system was then minimized for 1,000 iterations and simulated for an additional 0.5 ns at 300K; during this time, water was excluded from the lipid layer using the script of [49]. Following this, all atoms were released, and the system was equilibrated for 0.5 ns at 310 K. The final state of this trajectory was then used as the initial state for a production run of 100 ns at 310 K with frames collected at 10 ps intervals. This protocol was employed for both monomer and dimer structures.

Following simulation, water flow through PSI-induced channels was measured by counting transitions of water molecules through the membrane during each 10 ps observation window using a custom R [50] script using the `bio3d` library [51]. Underlying flow rates were estimated by Bayesian inference using an autoregressive latent rate model defined as

$$\lambda_i \sim t_\nu(\rho\lambda_{i-1}, \sigma) Y_i \sim \text{NegBin}(\lambda_i, \phi),$$

where $\lambda_i$ is the log expected flow rate in window $i$ (in units of counts per window), $Y_i$ is the number of water molecules transitioning in window $i$, $\nu$ and $\sigma$ are the degree of freedom and scale parameters for drift in the latent rate function, $\rho$ is an autocorrelation parameter, and $\phi$ is the overdispersion parameter of the negative binomial distribution. (Student's $t$ and negative binomial distributions were respectively employed because of the extremely bursty character of the flow process.) Weakly informative priors were used for all parameters, based on physically plausible limits for the processes in question; specifically, we take $1/\phi \sim \text{HalfCauchy}(0, 5)$, $\sigma \sim \text{HalfCauchy}(0, 5)$, $\nu \sim \text{Gamma}(2, 0.1)$, $\rho \sim \text{Normal}(0, 1)$, and $\lambda_1 \sim \text{Cauchy}(0, 5)$. Posterior inference was performed using No-U-Turn Hamiltonian Monte Carlo sampling [52] for 2,000 draws using four independent Markov chains; simulation was performed using the `rstan` package [53]. In addition to rate estimation, marginal

distributions of transition counts over all observation windows were also obtained; these were compared to the distributions arising from Poisson distributions with equivalent expected values to assess overdispersion.

## 3. Results and Discussion

### *3.1. D. capensis and related plants contain several aspartic proteases with PSIs*

The genome of *D. capensis* contains at least six aspartic proteases with moderate sequence identity to mammalian pepsin (Droserasins 1-6) [7]. These proteases are classified as belonging to the MEROPS A1 class [54], which also includes pepsin and the nepethesins found in the digestive fluid of pitcher plants of the related genus *Nepenthes* [55,56]. The PSIs (without the catalytic domains) from *D. capensis* and *D. muscipula* were clustered by protein sequence similarity. Three well-characterized PSI sequences from other plants were also included for reference: *C. cardunculus* Cardosins A and B (Uniprot IDs CARDA_CYNCA and CARDB_CYNCA, respectively) and *A. thaliana* APA1 (Uniprot ID APA1_ARATH) [28–30]. Previous studies have shown that recombinantly expressed APA1_ARATH is maximally efficient at pH 5.3, and has a highly specific cleavage profile with respect to the insulin $\beta$-chain [57]. The full-length droserasins share important functional sequence features with the vacuolar protein APA1_ARATH and the cardosins, including the active site residues, the disulfide bonding pattern, and the PSI. Protein sequence alignments comparing sequences of the PSIs described here can be found in Supplementary Figure S1. All the PSI sequences share a relatively high degree of sequence identity, but the two from *C. cardunculus* are much more similar to each other than to the *D. capensis*, *D. muscipula*, and *A. thaliana* sequences. The latter group further clusters into two subtrees, with three representatives from *D. capensis* and one from *D. muscipula* in each group. The Droserasin 1 PSI (D1 PSI) was chosen for further modeling and experimental characterization as a representative of the group that is less related to the previously characterized APA1_ARATH.

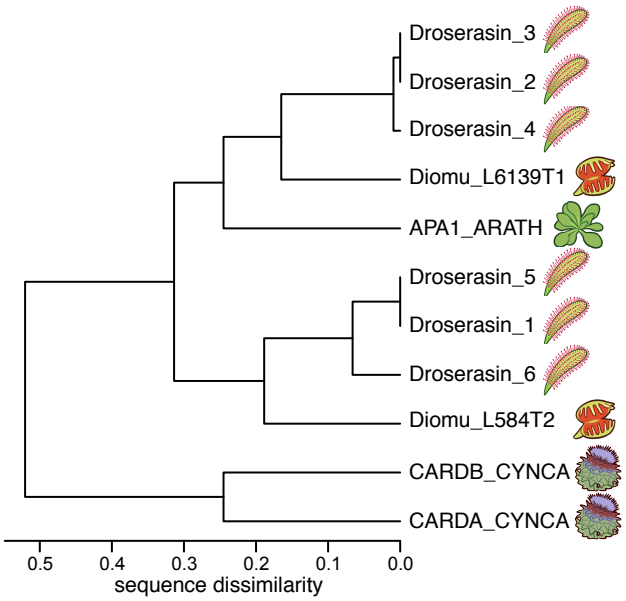


**Figure 1.** PSIs from *D. capensis* and *D. muscipula* clustered according to protein sequence similarity. Reverence sequences from *Arabidopsis thaliana* and *Cynara cardunculus* are also included for comparison.

Saposin-like proteins, including plant aspartic protease PSIs, are very stable, in part due to the presence of disulfide bonds that lock the tertiary structure into place [58]. This fold has two characteristic conformations, a compact closed form (Figure 2A) and an extended open form (Figure 2B) , both of which can be adopted by human saposin C [59,60]. The crystal structure of the *S. tuberosum*

PSI captures the open structure, observed as a domain-swapped dimer [11]. The corresponding model for the D1 PSI is shown in Figure 2C. The open conformation is proposed to be responsible for membrane-interacting activity due its increased exposure of hydrophobic residues compared to the closed conformation; in Sec. 3.7 we show via MD simulation that the open conformation of D1 PSI is indeed compatible with embedding in lipid bilayers. As we show, dimers formed from open-conformation monomers are also lipid-compatible, however; they have a distinct stabilization mechanism. The experiments and simulations that follow were performed in order to characterize the biophysical and membrane-interacting properties of the D1 PSI as a first step toward understanding its mode of antimicrobial activity.

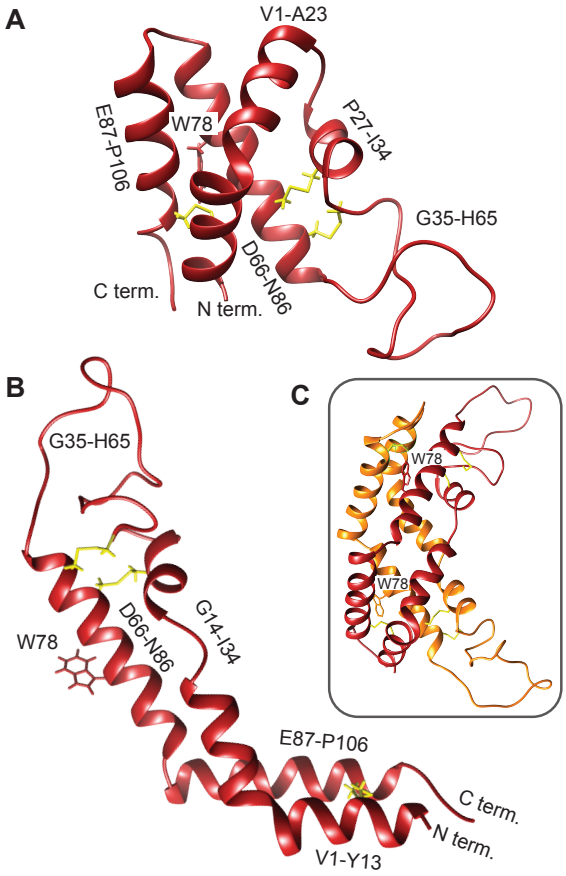


**Figure 2.** Comparative models of mature D1 PSI were predicted using the Robetta server [31,32]. **A** Closed conformation. **B** Open conformation. **C** Predicted dimer in solution.

### *3.2. D. capensis D1 PSI is highly stable*

In order to test the thermal stability of D1 PSI, experiments were performed with protein recombinantly expressed in *E. coli*. Typical yields of 20 mg/L of bacterial growth were achieved. Purification and characterization data are shown in Supplementary Figures S2-S3. CD spectra were collected under different pH and temperature conditions. Regardless of the pH, the CD spectra collected at 20, 55, or 90 °C show very little variance (Supplementary Figure S4), indicating that at least the secondary structure of this protein is extremely thermostable. Furthermore, after pretreating the PSI with DTT to reduce the disulfide bonds, only a marginal reduction of signal is observed, indicating that the secondary structure is perturbed very little even after its three presumptive disulfide bonds have been reduced.

To further probe the response of the D1 PSI to pH, temperature, and reducing agent, intrinsic tryptophan fluorescence spectroscopy was employed. This technique can provide information on

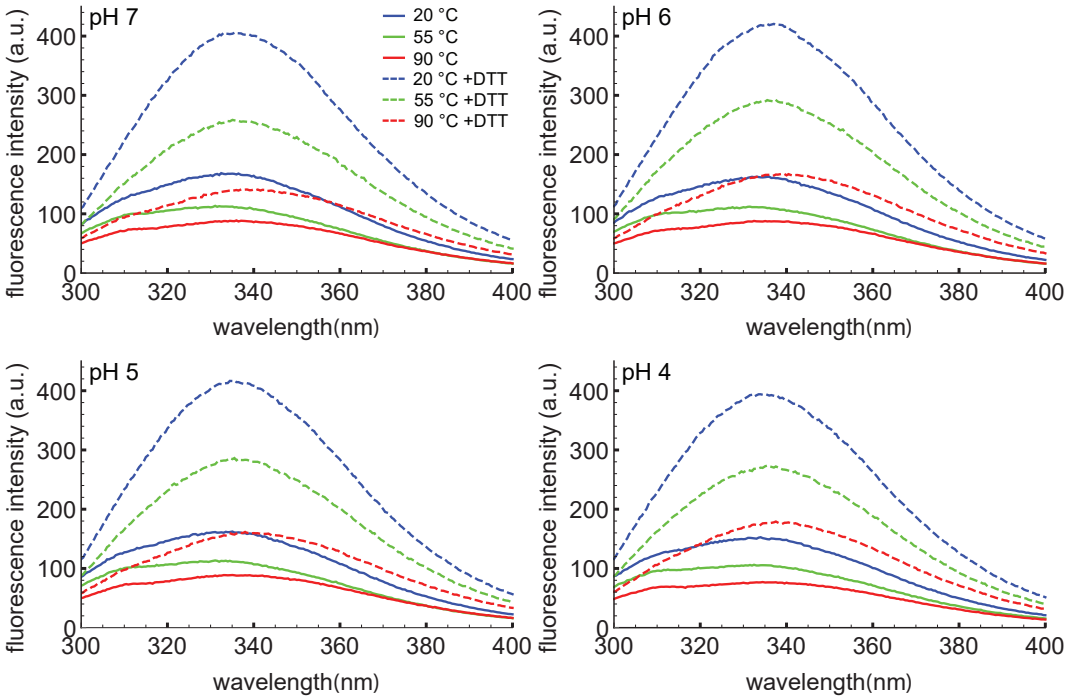


**Figure 3.** Intrinsic tryptophan fluorescence spectra of D1 PSI at pH 4, 5, 6, and 7. Spectra were collected at 20, 55 and 90 °C in the presence (dashed lines) or absence (solid lines) of DTT. The excitation wavelength was 280 nm. Under non-reducing conditions the emission maxima are at approximately 334 nm regardless of pH or temperature. Under reducing conditions the emission maxima for most of the spectra shift slightly to about 337 nm, indicating slightly more exposure of the single tryptophan. Reducing the protein also leads to an increase in the emission intensity, likely due to the reduction of a nearby disulfide bond that was quenching the tryptophan.

the local environment of the tryptophan as a more hydrophobic environment leads to a blue-shifted emission relative to a polar environment. The D1 PSI only has one tryptophan, allowing that particular position to be probed. Here, the intrinsic fluorescence was measured under the same conditions used for CD experiments. Under non-reducing conditions the wavelength of maximal emission is approximately 334 nm, regardless of temperature or pH, indicating that the single tryptophan (W78) is moderately exposed but insensitive to both pH and temperature. The emission intensity does decrease as a function of temperature, a known phenomenon in proteins [61]. When the D1 PSI has been reduced, there appears to be a slight shift in the emission maxima to approximately 337 nm, and more strikingly, the emission intensity increased relevant to the non-reduced PSI at equivalent temperature. A likely explanation could be that in the oxidized form, the tryptophan is next to a disulfide bond that is capable of quenching it, as in the closed conformation shown in Figure 2A but not the open monomer shown in Figure 2B, where the single Trp is solvent-exposed. Reducing the adjacent disulfide would release the quenching [62].

### *3.3. D1 PSI is monomeric in solution over a wide pH range*

There is experimental evidence that PSI of *Solanum tuberosum* forms dimers under low pH conditions [15], leading to the hypothesis that the activation of *D. capensis* D1 PSI requires dimerization at low pH values. To test this hypothesis, the presence of monomeric and dimeric D1 PSI was measured by intact protein mass spectrometry at a range of pH values, from 3 to 8 (Supplementary Figure S5). A dimeric form of D1 PSI would be expected to predominate at pH 3.5 and 4 but not at pH 7 and above if there is pH-dependent dimerization. For each pH value measured, D1 PSI was found to exist predominately in the monomeric form. Notably even at low pH values (pH 3.5 and 4) the monomeric form of the protein predominates. Also, the relative amounts of monomer to dimer do not change much as pH changes. Taken together we conclude that the oligomeric state of D1 PSI in aqueous solution is primarily monomeric, and is not strongly pH-dependent. This suggests that either the active form of D1 PSI is the monomeric state or that lipid bilayer interactions are required for dimerization.

### *3.4. D1 PSI enables vesicle fusion at acidic pH*

Based on previous studies in the literature, the simulations described in Section 3.7, and preliminary results consistent with inhibition of microbial growth at pH 5 (Supplementary Figure S6), we hypothesize that the PSI permeabilizes the membrane. We therefore used a vesicle fusion assay to test whether and how the D1 PSI interacts with membranes. Dynamic light scattering (DLS) was used to monitor increasing vesicle size due to fusion from membrane disruption. First, 100 nm large unilamellar vesicles (LUVs) were prepared using *E. coli* polar lipid extract. At all pH conditions tested, LUVs were stable in size over time as shown in Fig. 4A. Upon addition of D1 PSI the distribution of LUV size begin to shift to larger sizes, but only at pH 4 and 5: the pH 6 and 7 samples do not change in size. DLS measurements of the D1 PSI alone were also recorded at each pH. For each sample, only one peak was present, corresponding to a size of about 2 nm, which was stable over time. When LUVs were made using yeast polar lipid extract instead, again the LUVs are stable over time at each pH (Fig. 4B). When D1 PSI is added to the LUVs, only at pH 4 and 5 does the size distribution increase, similar to the results for the *E. coli* polar lipids. These findings are consistent with the antimicrobial assay in that D1 PSI interacts with membranes in a pH-dependent manner where it is only active at acidic pH. Another noteworthy observation is that with LUVs composed of *E. coli* polar lipids the rate of vesicle fusion was slow and gradual while in the case of the yeast polar lipids the change was very rapid at pH 4. Not only is the rate faster in this case, but for the pH 4 yeast polar lipid condition, smaller peaks appear at the latest time points. A potential hypothesis to explain the presence of these smaller particles after extended time is that it is possible that the D1 PSI is actually able to extract some of the lipids from the vesicle into small lipoprotein particles, consistent with the role of some saposins as surfactants.

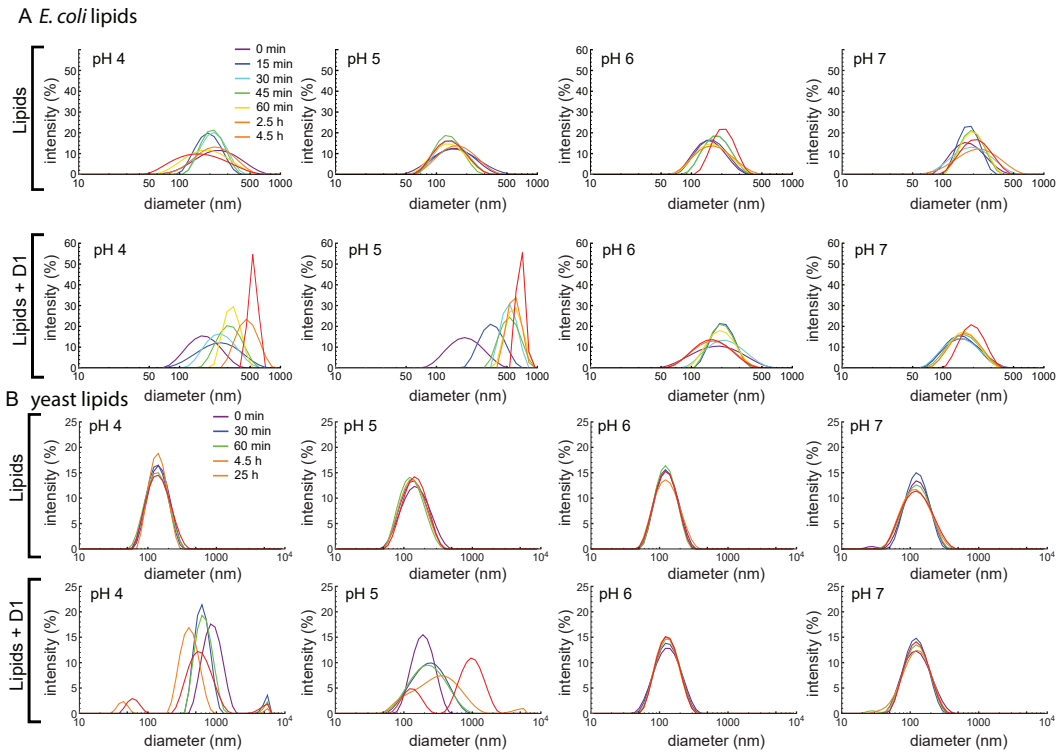


**Figure 4.** **(A)** LUVs of 100 nm were made using an *E. coli* polar lipid extract. LUVs were made in buffered solutions at pH 4, 5, 6, and 7. Vesicle size was monitored over time using DLS, either with (bottom) or without (top) D1 PSI. LUVs made with lipids alone are stable in size for all pH values, but upon addition of D1 PSI they fuse into larger vesicles at acidic pH. **(B)** LUVs of 100 nm were made using a yeast polar lipid extract. LUVs were made in buffered solutions at pH 4, 5, 6, and 7 and size was monitored using DLS over time, either without (top) or with (bottom) D1 PSI. Again, LUVs are stable in size for all pH values, but upon addition of D1 PSI they fuse into larger vesicles at acidic pH.

### *3.5. D1 PSI is able to interact with lipids having diverse head groups*

After showing that D1 PSI was able to interact with membrane lipids from a natural source that contains a variety of lipid species, we sought to gain insight into whether the D1 PSI preferentially binds specific polar lipids. The *S. tuberosum* PSI was found to interact with both anionic and neutral lipids, although its specific mode of membrane interaction does depend on the head groups, with more membrane disruption observed in negatively charged membranes [14]. Furthermore, the presence of cholesterol inhibits membrane fusion activity, making the PSI non-toxic to animal cells [13]. First, the affinity of D1 PSI for different phospholipid head groups was tested. Five phospholipid species containing different head groups, but the same acyl groups (a singly unsaturated 18-carbon acyl chain and saturated 16-carbon acyl chain), were tested for their ability to associate with the D1 PSI. A solution was prepared containing both the neutrally-charged POPC and POPE as well as negatively-charged POPG, POPS, and POPA. One aliquot of this lipid solution had His-tagged D1 PSI added and was mixed, allowing time for lipid-protein interactions to occur. Then the lipid-protein mixture was applied to $Ni^{2+}$ resin where the PSI would bind, bringing with it any associated lipids, while unbound lipids were washed away.

The PSI was eluted and any lipids that associated with it were quantified by MS and compared to the composition of the original lipid solution (Fig. 5A). As shown in Fig. 5B there is no significant difference in relative lipid composition between the original lipid solution and the lipids that co-purified with the PSI. However, it is possible the kinetics of association may be different depending on the lipid composition as hinted with the vesicle fusion assay (Fig. 4). During membrane association the PSI must first interact with the head groups, before interacting with the acyl chains buried in the membrane. It is possible that specific head groups more strongly interact with the PSI, promoting initial association, followed by contact with the acyl chains, at which point PSI-acyl chain interactions predominate. The lack of sensitivity to the lipid head group makes the D1 PSI a promising candidate for making lipoprotein nanoparticles of different sizes for NMR studies of other membrane proteins, as previously demonstrated for saposin A [63–65].

### *3.6. Solid-state NMR shows that D1 PSI is ordered and strongly bound to the membrane*

Solid-state NMR (ssNMR) is often used to solve the structures of membrane-associated proteins whose complexes with lipids are too large to tumble isotropically in solution. Because ssNMR makes use of magic angle spinning (MAS) to average out chemical shift anisotropy and dipolar coupling interactions it does not suffer the size limitations of solution-state NMR. This allows for the use of vesicles or bicelles as a membrane system to study membrane proteins [66]. Because of its similarity to saposins, the D1 PSI may be useful for stabilizing lipid nanodiscs, which are composed of a small section of lipid bilayer that is encircled by a membrane scaffold protein [67]. For ssNMR studies, $^{13}C$, $^{15}N$ isotopically labeled D1 PSI was expressed, purified and mixed with a mixture of 1:1 POPC and POPA at pH 4.5. The sample was sent to the National High Magnetic Field Laboratory in Tallahassee, FL to obtain spectra.

Our NMR investigations of this protein begin with 2D $^{13}C$-$^{13}C$ correlation spectra collected using the dipolar-assisted rotational resonance (DARR) experiment [42,68]. Early experiments using nano- or microcrystalline preparations of small, globular model proteins such as BPTI [69], ubiquitin [70], GB1 [71], and the $\alpha$-spectrin SH3 domain [72] demonstrated that this type of homonuclear correlation can be used to provide partial assignments. In the usual procedure, well-resolved resonances with distinctive chemical shift values are identified, followed by mapping of spin systems via a "walk" among the proximal $^{13}C$ atoms of the sidechain. Full assignments can then be obtained using further 2D, 3D and 4D experiments [73,74], followed by the measurement of through-space distance restraints (often using these same simple homonuclear correlations), and finally, structure determination [72,75,76]. This methodology is also fully applicable to membrane proteins [77–79], most readily in cases where the protein adopts a well-ordered conformation in the membrane. The $^{13}C$-$^{13}C$ CP DARR spectrum (Fig. 6) shows that there are clearly-defined peaks that are reasonably well-dispersed, demonstrating

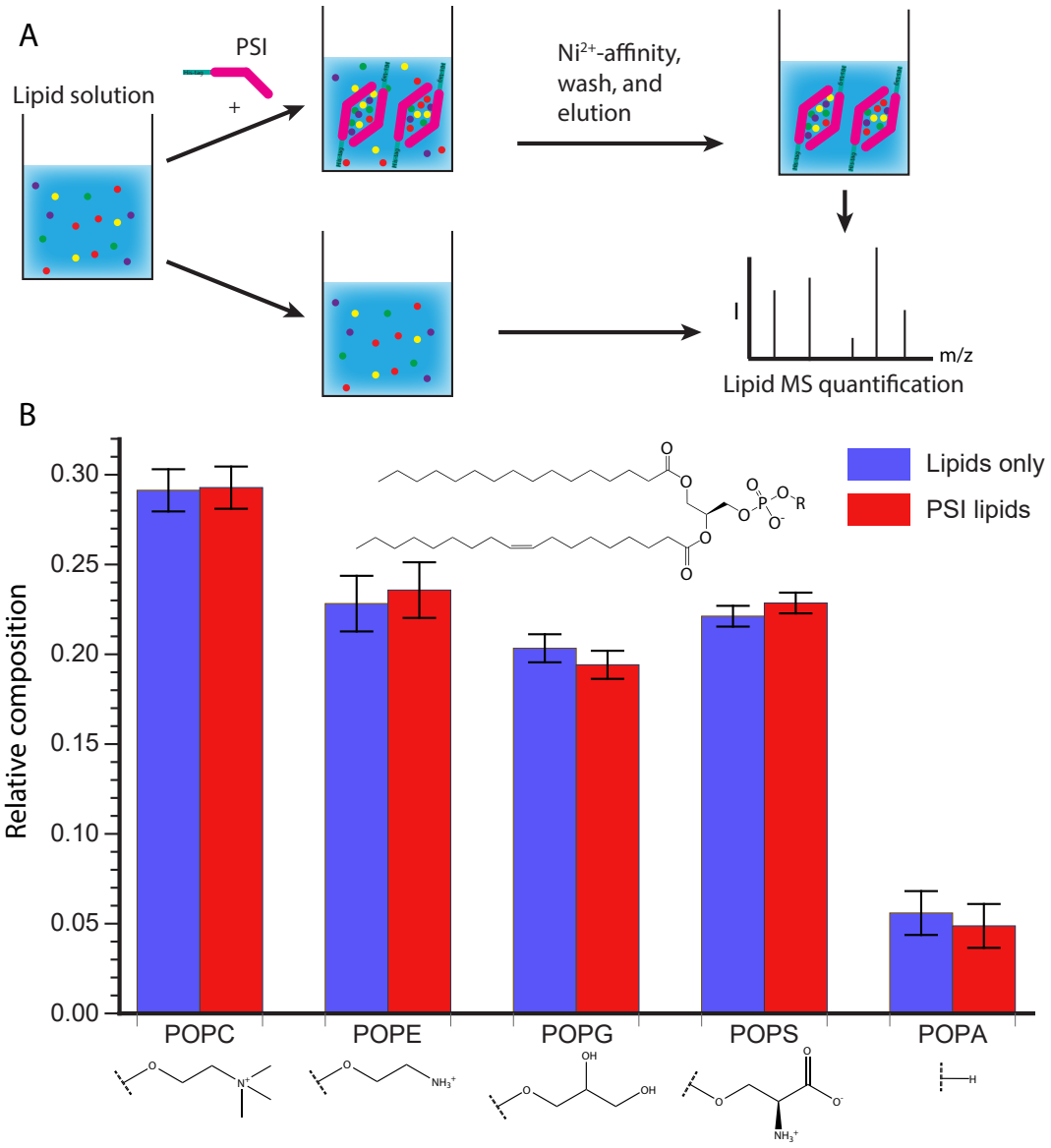


**Figure 5.** **(A)** Schematic representation of the experimental procedure. A lipid solution containing POPC, POPE, POPG, POPS, and POPA was prepared and split into two aliquots. One aliquot was treated with His-tagged D1 PSI and allowed to mix before purifying the His-tagged PSI and bound lipids from the unbound lipids. Lipids were then extracted from the D1 PSI and lipid solution and relative lipid species were quantified using mass spectrometry. **(B)** Results of lipid quantification. Blue represents lipids from the initial lipid mixture while red represents lipids that were bound to D1 PSI. The data indicate that there is no difference between the initial composition of lipids and the composition of lipids bound to the D1 PSI.

that the PSI is not overly mobile and is not structurally heterogeneous in the sample so as to distribute the signal over multiple chemical shifts. Additional DARR spectra were collected with mixing times of 50 and 400 ms (Supplementary Figure S7.) These spectra are similarly well-resolved, and comparing the three spectra shows the expected increase in the number of cross-peaks with increasing mixing time, as correlations between more distant spin pairs are observed.

These preliminary spectra indicate that this protein is amenable to structure determination by ssNMR. Based on common $^{13}$C chemical shifts we can tentatively assign some residues based on correlations between the C$\alpha$ and C$\beta$ carbons. We can tentatively assign 1 Ala, 3 Val, 1 Pro, and 4 Asn residues. This suggest that a backbone walk is possible to do with D1-PSI in a POPC and POPA membrane system. Carbon chemical shifts were compared to the average chemical shifts from all the proteins in the BioMagResBank, [80] gathered from http://www.bmrb.wisc.edu/ref_info/statful.htm as published on 5/14/2020. More experiments, including 3D experiments, will be needed for resonance complete assignments. Further sample preparation, such as optimization of lipid composition and protein-lipid ratios, is also needed to increase the signal of the sample and to explore how the D1 PSI responds to different lipids in the membrane system.

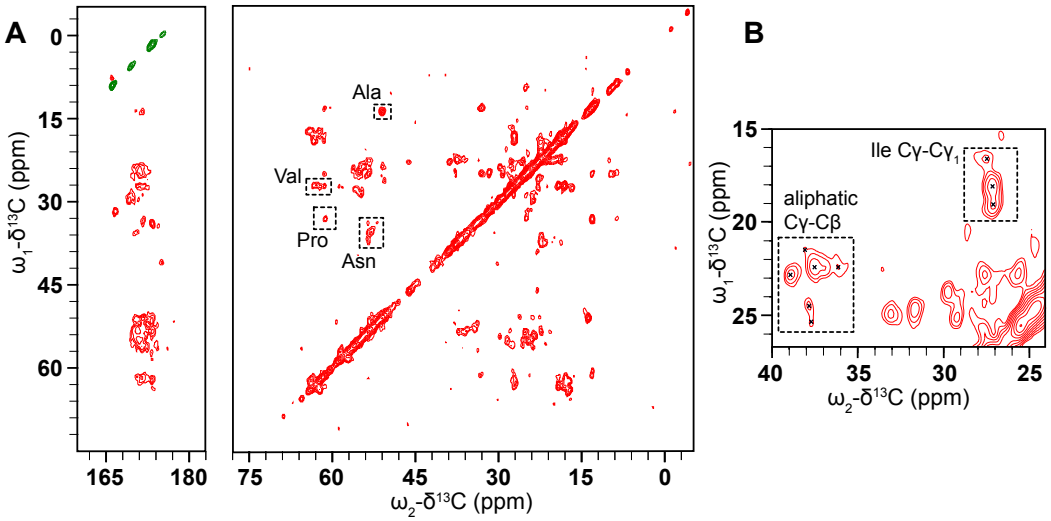


**Figure 6. (A)** CO (left) and aliphatic (right) regions of a $^{13}$C-$^{13}$C CP DARR spectrum of D1 PSI in a 1:1 POPC, POPA membrane system was taken at 10 °C and spinning at 24.4 kHz. The spectrum displays well-resolved off-diagonal peaks of reasonable intensity indicating that the PSI is suitable for ssNMR structure determination. Tentative assignments are labeled by residue type. **(B)** The inset region shows some representative cross-peaks in the aliphatic region.

### *3.7. Molecular modeling suggests potential stability of both monomeric and dimeric D1 PSI within membranes, and indicates that both induce permeability*

To further explore the potential interaction of the D1 PSI with membranes, in particular the question of whether the membrane-interacting form is likely to be monomeric or dimeric, we performed all-atom molecular dynamics (MD) simulations of the PSI within POPC bilayers. As both monomeric and dimeric conformations have been proposed to be biologically relevant for other PSIs, we examined both cases. After 100 ns (following initial stabilization and equilibration) at 310K, we see that both monomeric and dimeric forms retain stable – but quite distinct – conformations within the bilayer. Fig. 7B provides a schematic depiction of the monomeric case. The PSI assumes the open "L" conformation, with the terminal helices V1-Y13 and E87-P106 forming a partially solvated "pontoon" that sits parallel to the lipid surface and the central helix pair (residues G14-I34 and D66-N86) spanning the bilayer at an angle of roughly 120 degrees with respect to the "pontoon." On the opposite side, the protein is anchored by a mostly unstructured, hydrophilic loop region (G35-H65) that extends well

into the solvent (see Fig. 7C). This conformation appears to be stable, and indeed the "hinge" between the pontoon and the spanning helices would appear to allow the PSI to accommodate substantial deformation in membrane curvature or thickness without extensive conformational change.

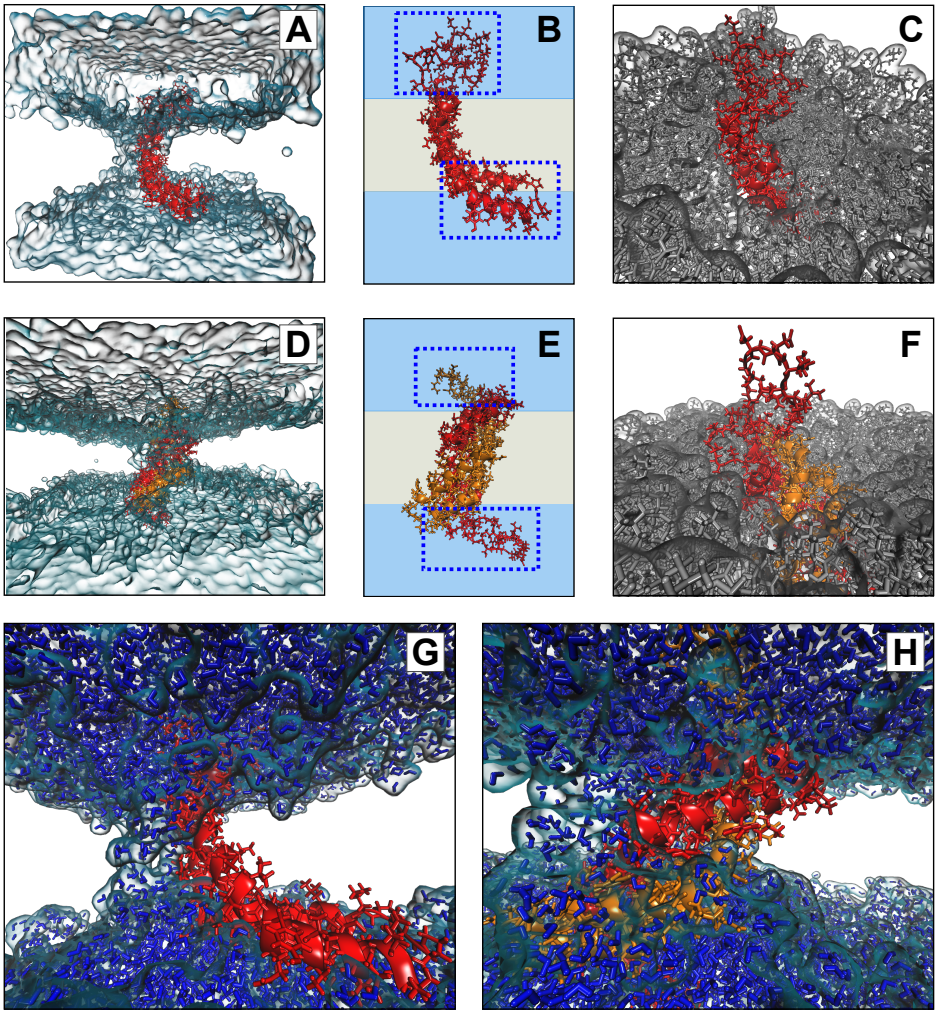


**Figure 7.** Conformations of monomeric (**(A)-(C)**) and dimeric (**(D)-(F)**) D1 PSI in POPC bilayer after 100ns. Both monomer and dimer assume stable but distinct conformations within the membrane (**(A)**,**(D)**), anchored by loop regions and (for the monomer) a helical "pontoon" as shown schematically in **(B)**,**(E)**; blue regions indicate solvent, whereas grey regions indicate lipid. Loops extend substantially into solvent in both cases (**(C)**,**(F)**), while a depression forms in the surrounding bilayer. Bilayer depressions are linked to channels along the hydrophilic regions of the protein (**(G)**,**(H)**), through which water diffuses.

In the case of the dimer, we observe a very different anchoring strategy. Fig. 7E shows a schematic of the D1 PSI within the bilayer. In this case, the PSI forms a symmetric homodimer that anchors to both sides of the bilayer by respective copies of the hydrophilic anchor loop (G35-H65). Interestingly, the hydrophobic residues of the central helices that are lipid-exposed in the monomeric case serve here as the core of the homodimer, remaining stable despite burial in a hydrophobic environment. By contrast, the nominally lipid-exposed residues about the core (which form the "back" of the open "L" in the monomer, and the pontoon) are largely hydrophilic. This does not appear to destabilize the dimer; instead, these residues appear to facilitate the maintenance of a water channel, as described below. As with the monomeric case, the anchor loop is aggressively solvated, with a water layer extending several nanometers beyond the lipid head groups (Fig. 7F). The highly flexible nature of the anchoring loops suggests a substantial entropic contribution to the free energy of the lipid-embedded

dimer (more so than for the monomer), possibly allowing the dimer to remain anchored within the bilayer at higher temperatures or at higher ionic concentrations.

Both monomeric and dimeric D1 PSI are observed to induce local changes in membrane conformation, creating "depressions" in the membrane surface (Fig. 7C,F) that are linked to transmembrane water channels (Fig. 7G,H). In the monomeric case, a single channel is formed that follows the hydrophilic residues on the "back" of the spanning helices, bridging to the hydrophilic residues on the sides and "bottom" of the pontoon (Fig. 7G). For the dimer, the larger available hydrophilic surface area tends to produce multiple locally solvated regions (Fig. 7H), supporting a broader channel structure. Examination of water transport across the membrane confirms that these apparent channels do indeed induce membrane permeability. Fig. 8A shows posterior estimates of water transport rates (molecules/ns) across both simulated MD trajectories. As suggested by the broader channel structure, the PSI dimer supports a higher and somewhat more consistent mean flow rate (apx 37 molecules/ns) than the monomer (apx 7 molecules/ns), though the former is still somewhat "bursty" and irregular. Flow through the monomer channel is extremely irregular, with infrequent bursts and substantial changes in the overall flow rate over longer time scales (the former being evident from the roughness of the estimated rate function). Additional confirmation of the bursty nature of water flow in both cases can be seen by comparing the marginal distribution of observed transition events over short (10 ps) intervals to an equivalent Poisson process, as shown in Fig. 8B. While the upper tail of the transition counts (i.e., numbers of water molecules observed to transition through the channel) is heavier than the Poisson in both cases, the departure is especially profound for the monomer trajectory. This may arise from the relative asymmetry of the channel structure (as shown in Fig. 7G), which may create "pools" of water that can only pass when the monomer assumes specific conformations. Whether or not this is the case, the substantially higher mean flow rates for dimeric D1 PSI suggest that enhanced concentrations of PSI within membranes will produce greater than linear increases in membrane permeability, potentially contributing to toxicity in biological settings.

## 4. Conclusions

In conclusion, the D1 PSI is a mostly $\alpha$-helical, highly thermostable peptide that maintains its secondary structure up to 90 °C, even under conditions where its three disulfide bonds are reduced. It causes vesicle fusion in lipid mixtures from both yeast and *E. coli*, but only at a pH $\leq$ 5, consistent with the physiological pH of *D. capensis* digestive mucilage. Mass spectrometry of bound lipids upon extraction from lipid mixtures indicate that the PSI interacts with a wide range of lipids, independent of the charge on the head groups. Solid-state NMR data indicate that it strongly interacts with the membrane in a bicelle mixture, consistent with its robust lipid interactions in vesicle fusion assays. Furthermore, these data suggest that the PSI adopts an ordered conformation when interacting with the membrane, making structural studies in the membrane-bound state feasible. A major question left to be resolved is the oligomeric state of the PSI in membranes. In aqueous solution, it is predominantly monomeric over a wide range of pH values from 3-8. However, MD simulations show that either the monomeric or dimeric state may interact with and permeabilize membranes. In both the monomeric an dimeric states, the PSI in its open conformation can span the membrane, inducing local changes in membrane curvature and allowing the passage of water molecules from one side to the other. Further experimental characterization is needed to elucidate the mechanism of membrane interaction, which may involve monomers, dimers, or even larger pore-forming complexes.

**Author Contributions:** M.A.S.P., J.C.B., M.G.C, G.R.T., B.P.C., N.K., and X.C. prepared protein samples, performed biophysical experiments and analyzed data. J.E.K, C.T.B., and R.W.M. performed protein sequence analysis. J.P., I.H., and R.Z. collected solid-state NMR data. M.A.S.P., J.C.B., and R.W.M. analyzed solid-state NMR data. C.T.B. conceptualized and performed molecular dynamics simulations and analyzed MD data. R.W.M designed the experimental studies. M.A.S.P., J.C.B., M.G.C, C.T.B., and R.W.M wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by NSF awards DMS-1361425 to C.T.B and R.W.M, and CHE-1308231 to R.W.M., ARO award W911NF-14-1-0552 to C.T.B, and the CIFAR Molecular Architecture of Life program. This

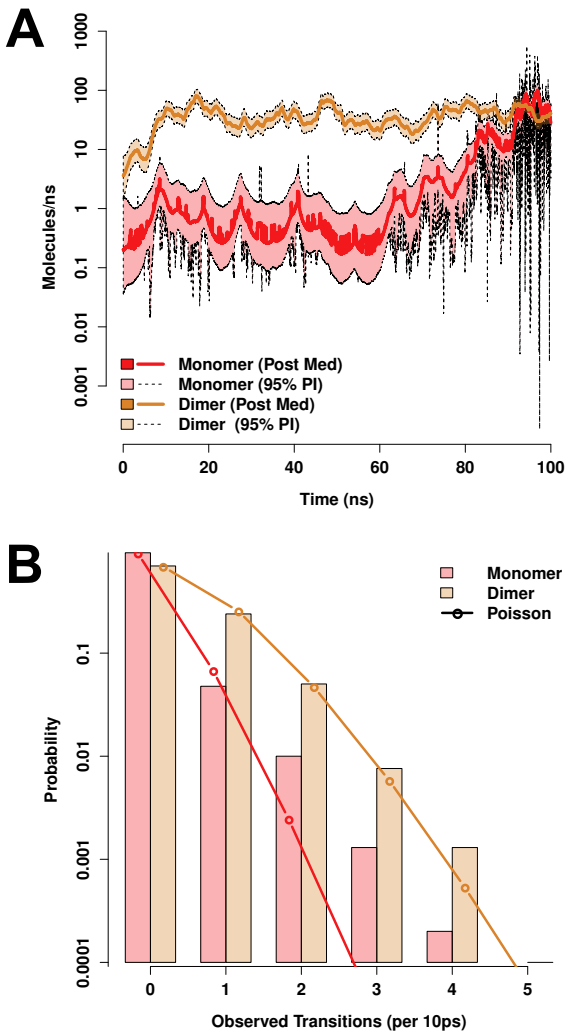


**Figure 8.** Both monomeric and dimeric D1 PSI conformations induce permeability in lipid bilayers, as revealed by MD trajectories. **(A)** Posterior median estimates (solid lines) and 95% posterior intervals (shaded areas) of trans-membrane water flow as a function of simulation time. Transport is extremely "bursty" in both conformations, with substantially larger fluctuations in mean rate for the PSI monomer trajectory. **(B)** Marginal distribution of transport events (cross-membrane water molecule transitions) per 10ns observation interval (bars). Observed distributions are heavier-tailed than Poisson distributions with equivalent expectation (lines), indicating high levels of "burstiness' even on short timescales.

work was made possible, in part, through access to the Genomics High Throughput Facility Shared Resource of the Cancer Center Support Grant (P30CA-062203) at the University of California, Irvine and NIH shared instrumentation grants 1S10RR025496-01, 1S10OD010794-01, and 1S10OD021718-01. The National High Magnetic Field Laboratory (NHMFL) is supported by the National Science Foundation through NSF/DMR-1644779 and the state of Florida. Development of the 36-T series connected hybrid magnet and NMR instrumentation was supported by NSF (DMR-1039938 and DMR-0603042) and NIH (BTRR 1P41 GM122698).

**Acknowledgments:** The authors acknowledge Dmitry Fishman of the UCI Laser Spectroscopy Labs for assistance with optical spectroscopy, Ben Katz and Felix Grun of the UCI Mass Spectrometry Facility for assistance with mass spectrometry, Tim Cross for helpful discussions and facilitating SCH use, and Veronika Cesar for preparing the plant diagrams in Fig. 1. R.W.M. is a CIFAR fellow.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| D1 | Droserasin 1 |
| DMSO | dimethylsulfoxide |
| DTT | dithiothreitol |
| IPTG | Isopropyl $\beta$-D-1-thiogalactopyranoside |
| PSI | plant-specific insert |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| POPC | 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine |
| POPE | 1-palmitoyl-2-oleoyl-sn-glycero-3- phosphoethanolamine |
| POPG | 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) |
| POPS | 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine |
| POPA | 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphate |
| LUV | large unilamellar vesicles |
| CD | circular dichroism |
| CP | cross-polarization |
| DARR | dipolar-assisted rotational resonance |
| MAS | magic angle spinning |
| MD | molecular dynamics |
| MS | mass spectrometry |
| NMR | nuclear magnetic resonance |

## References

1. Pavlovič, A.; Saganová, M. A novel insight into the cost–benefit model for the evolution of botanical carnivory. *Annals of Botany* **2015**, *115*, 1075–1092. doi:10.1093/aob/mcv050.
2. Paniw, M.; Gil-Cabeza, E.; Ojeda, F. Plant carnivory beyond bogs: reliance on prey feeding Drosophyllum lusitanicum (Drosophyllaceae) in dry Mediterranean heathland habitats. *Annals of Botany* **2017**, p. mcw247. doi:10.1093/aob/mcw247.
3. Ellison, A.M.; Gotelli, N.J. Energetics and the evolution of carnivorous plants—Darwin's 'most wonderful plants in the world'. *Journal of Experimental Botany* **2009**, *60*, 19–42.
4. Kokubun, T. Occurrence of myo-inositol and alkyl-substituted polysaccharide in the prey-trapping mucilage of Drosera capensis. *The Science of Nature* **2017**, *104*. doi:10.1007/s00114-017-1502-4.
5. Król, E.; Płachno, B.J.; Adamec, L.; Stolarz, M.; Dziubińska, H.; Trębacz, K. Quite a few reasons for calling carnivores 'the most wonderful plants in the world'. *Annals of Botany* **2011**, *109*, 47–64. doi:10.1093/aob/mcr249.
6. Palfalvi, G.; Hackl, T.; Terhoeven, N.; Shibata, T.F.; Nishiyama, T.; Ankenbrand, M.; Becker, D.; Förster, F.; Freund, M.; Iosip, A.; Kreuzer, I.; Saul, F.; Kamida, C.; Fukushima, K.; Shigenobu, S.; Tamada, Y.; Adamec, L.; Hoshi, Y.; Ueda, K.; Winkelmann, T.; Fuchs, J.; Schubert, I.; Schwacke, R.; Al-Rasheid, K.; Schultz, J.; Hasebe, M.; Hedrich, R. Genomes of the Venus flytrap and close relatives unveil the roots of plant carnivory. *Current Biology* **2020**, p. https://doi.org/10.1016/j.cub.2020.04.051.

7. Butts, C.T.; Bierma, J.C.; Martin, R.W. Novel proteases from the genome of the carnivorous plant *Drosera capensis*: structural prediction and comparative analysis. *Proteins: Structure, Function, and Bioinformatics* **2016**, *84*, 1517–1533.
8. Butts, C.T.; Zhang, X.; Kelly, J.E.; Roskamp, K.W.; Unhelkar, M.H.; Freites, J.A.; Tahir, S.; Martin, R.W. Sequence comparison, molecular modeling, and network analysis predict structural diversity in cysteine proteases from the Cape sundew, *Drosera capensis*. *Computational and Structural Biotechnology Journal* **2016**, *14*, 271–282.
9. Unhelkar, M.H.; Duong, V.T.; Enendu, K.N.; Kelly, J.E.; Tahir, S.; Butts, C.T.; Martin, R.W. Structure prediction and network analysis of chitinases from the Cape sundew, *Drosera capensis*. *Biochimica et Biophysica Acta* **2017**, *1861*, 636–643.
10. Duong, V.T.; Unhelkar, M.H.; Kelly, J.E.; Kim, S.H.; Butts, C.T.; Martin, R.W. Network analysis provides insight into active site flexibility in esterase/lipases from the carnivorous plant *Drosera capensis*. *Integrative Biology* **2018**, *10*, 768–779.
11. Bryksa, B.; Bhaumik, P.; Magracheva, E.; De Moura, D.; Kurylowicz, M.; Zdanov, A.; Dutcher, J.; Wlodawer, A.; Yada, R. Structure and mechanism of the saposin-like domain of a plant aspartic protease. *Journal of Biological Chemistry* **2011**, *286*, 28265–28275.
12. Bruhn, H. A short guided tour through functional and structural features of saposin-like proteins. *Biochemical Journal* **2005**, *389*, 249–257. doi:10.1042/bj20050051.
13. Muñoz, F.; Palomares-Jerez, M.F.; Daleo, G.; Villalaín, J.; Guevara, M.G. Cholesterol and membrane phospholipid compositions modulate the leakage capacity of the swaposin domain from a potato aspartic protease (StAsp-PSI). *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2011**, *1811*, 1038–1044. doi:10.1016/j.bbalip.2011.08.013.
14. Muñoz, F.F.; Palomares-Jerez, M.F.; Daleo, G.; Villalaín, J.; Guevara, M.G. Possible mechanism of structural transformations induced by StAsp-PSI in lipid membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2014**, *1838*, 339–347.
15. Bryksa, B.C.; Yada, R.Y. Protein structure insights into the bilayer interactions of the saposin-like domain of *Solanum tuberosum* aspartic protease. *Scientific Reports* **2017**, *7*, 16911.
16. D'Hondt, K.; Stack, S.; Gutteridge, S.; Vandekerckhove, J.; Krebbers, E.; Gal, S. Aspartic proteinase genes in the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*. *Plant Molecular Biology* **1997**, *33*, 187–192.
17. Xia, Y.; Suzuki, H.; Borevitz, J.; Blount, J.; Guo, Z.; Patel, K.; Dixon, R.A.; Lamb, C. An extracellular aspartic protease functions in *Arabidopsis* disease resistance signaling. *The EMBO Journal* **2004**, *23*, 980–988.
18. Simoes, I.; Faro, C. Structure and function of plant aspartic proteinases. *European Journal of Biochemistry* **2004**, *271*, 2067–2075.
19. White, P.C.; Cordeiro, M.C.; Arnold, D.; E., B.P.; Kay, J. Processing, activity, and Inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). *Journal of Biological Chemistry* **1999**, *274*, 16685–16693.
20. Duarte, P.; Pissarra, J.; Moore, I. Processing and tracking of a single isoform of the aspartic proteinase cardosin A on the vacuolar pathway. *Planta* **2008**, *227*, 1255–1268.
21. Pereira, C.; Pereira, S.; Satiat-Jeunemaitre, B.; Pissarra, J. Cardosin A contains two vacuolar sorting signals using different vacuolar routes in tobacco epidermal cells. *The Plant Journal* **2013**, *76*, 87–100.
22. Egas, C.; Lavoura, N.; Resende, R.; Brito, R.; E., P.; de Lima, M.; Faro, C. The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *Journal of Bioloical Chemistry* **2000**, *275*, 38190–38196.
23. Muñoz, F.F.; Mendieta, J.R.; Pagano, M.R.; Paggi, R.A.; Daleo, G.R.; Guevara, M.G. The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides* **2010**, *31*, 777–785. doi:10.1016/j.peptides.2010.02.001.
24. Risør, M.W.; Thomsen, L.R.; Sanggaard, K.W.; Nielsen, T.A.; Th/ogersen, I.B.; Lukassen, M.V.; Rossen, L.; Garcia-Ferrer, I.; Guevara, T.; Scavenius, C.; Meinjohanns, E.; Gomis-Rüth, F.X.; Enghild, J.J. Enzymatic and structural characterization of the major endopeptidase in the Venus flytrap digestion fluid. *The Journal of Biological Chemistry* **2016**, *291*, 2271–2287.
25. Campbell, M.; Law, M.; Holt, C.; Stein, J.; Moghe, G.; Hufnagel, D.; Lei, J.; Achawanantakun, R.; Jiao, D.; Lawrence, C.J.; Ware, D.; Shiu, S.H.; Childs, K.L.; Sun, Y.; Jiang, N.; .; Yandell, M. MAKER-P: A Tool-kit for

the Rapid Creation, Management, and Quality Control of Plant Genome Annotations. *Plant Physiology* **2013**, *164*, 513–524.

26. Quevillon, E.; Silventoinen, V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R. InterProScan: Protein Domains Identifier. *Nucleic Acids Research* **2005**, *33*, W116–W120.
27. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J.D.; Higgins, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011**, *7*, 539–539.
28. Silva, S.; F.X., M. Proteolysis of ovine caseins by cardosin A, an aspartic acid proteinase from *Cynara cardunculus* L. *Lait* **1998**, *78*, 513–519.
29. Silva, S.; Malcata, F. On the activity and specificity of cardosin B, a plant proteinase, on ovine caseins. *Food and Chemical Toxicology* **1999**, *67*, 373–378.
30. Takahashi, K.; Niwa, H.; Yokota, N.; Kubota, K.; Inoue, H. Widespread tissue expression of nepenthesin-like aspartic protease genes in Arabidopsis thaliana. *Plant Physiology and Biochemistry* **2008**, *46*, 724–729.
31. Kim, D.; Chivian, D.; Baker, D. Protein Structure Prediction and Analysis Using the Robetta Server. *Nucleic Acids Research* **2004**, *32*, W526–31.
32. Raman, S.; Vernon, R.; Thompson, J.; Tyka, M.; Sadreyev, R.; Pei, J.; Kim, D.; Kellogg, E.; DiMaio, F.; Lange, O.; Kinch, L.; Sheffler, W.; Kim, B.H.; Das, R.; Grishin, N.V.; Baker, D. Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins* **2009**, *77*, 89–99.
33. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25*, 1605–1612.
34. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38, 27–28.
35. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **2005**, *41*, 207–234.
36. Faust, G.; Stand, A.; Weuster-Botz, D. IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured *Escherichia coli*. *Engineering in Life Sciences* **2015**, *15*, 824–829. doi:10.1002/elsc.201500011.
37. Sivashanmugam, A.; Murray, V.; Cui, C.; Zhang, Y.; Wang, J.; Li, Q. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Science* **2009**, *18*, 936–948. doi:10.1002/pro.102.
38. Bligh, E.G.; Dyer, W.J. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **1959**, *37*, 911–917. doi:10.1139/o59-099.
39. Vandana, V.; Karuna, M.S.L.; Vijayalakshmi, P.; Prasad, R.B.N. A simple method to enrich phospholipid content in commercial soybean lecithin. *Journal of the American Oil Chemists' Society* **2001**, *78*, 555–556. doi:10.1007/s11746-001-0303-2.
40. Bird, M.D.; Brey, W.W.; Cross, T.A.; Dixon, I.R.; Griffin, A.; Hannahs, S.T.; Kynoch, J.; Litvak, I.M.; Schiano, J.L.; Toth, J. Commissioning of the 36 T Series-Connected Hybrid Magnet at the NHMFL. *IEEE Transactions on Applied Superconductivity* **2018**, *28*, 1–6. doi:10.1109/tasc.2017.2781727.
41. Pines, A.; Gibby, M.G.; Waugh, J.S. Proton-Enhanced Nuclear Induction Spectroscopy. A Method for High Resolution NMR of Dilute Spins in Solids. *Journal of Chemical Physics* **1972**, *56*, 1776–1777.
42. Takegoshi, K.; Nakamura, S.; Terao, T. 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* **2001**, *344*, 631–637.
43. Olsson, M.H.; Sondergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation* **2011**, *7*, 525–537.
44. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.
45. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26*, 1781–1802.

46. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.; Mittal, J.; Feig, M.; Mackerell, Jr, A.D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone $\phi$, $\psi$ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J Chem Theory Comput* **2012**, *8*, 3257–3273. doi:10.1021/ct300400x.
47. Martyna, G.J.; Tobias, D.J.; Klein, M.L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101*, 4177–4189.
48. Feller, S.E.; Zhang, Y.; Pastor, R.W.; Brooks, B.R. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics* **1995**, *103*, 4613–4621.
49. Aksimentiev, A. Script to Exclude Water from Lipid Bilayers. Software File, 2016.
50. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.
51. B.J., G.; A.P.C., R.; K.M., E.; J.A., M.; L.S.D., C. Bio3D: An R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695–2696.
52. Homan, M.D.; Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* **2014**, *15*, 1593–1623.
53. Stan Development Team. RStan: the R interface to Stan, 2020. R package version 2.19.3.
54. Rawlings, N.; Waller, M.; Barrett, A.; Bateman, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **2014**, *42*, D503–D509.
55. Takahashi, K.; Athauda, S.; Matsumoto, K.; Rajapakshe, S.; Kuribayashi, M.; Kojima, M.; Kubomura-Yoshida, N.; Iwamatsu, A.; Shibata, C.; Inoue, H. Nepenthesin, a unique member of a novel subfamily of aspartic proteinases: enzymatic and structural characteristics. *Current Protein and Peptide Science* **2005**, *6*, 513–525.
56. Buch, F.; Kaman, W.E.; Bikker, F.J.; Yilamujiang, A.; Mithöfe, A. Nepenthesin protease activity indicates digestive fluid dynamics in carnivorous *Nepenthes* plants. *PLoS ONE* **2015**, *10*, e0118853.
57. Mazorra-Manzano, M.A.; Tanaka, T.; Dee, D.R.; Yada, R.Y. Structure-function characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry* **2010**, *71*, 515–523.
58. Motta, M.; Tatti, M.; Salvioli, R. Autophagic dysfunction in Gaucher disease and its rescue by cathepsin B and D proteases. In *Autophagy: Cancer, Other Pathologies, Inflammation, Immunity, Infection, and Aging*; Elsevier, 2014; pp. 131–146. doi:10.1016/b978-0-12-405529-2.00009-3.
59. Ahn, V.; Leyko, P.; Alattia, J.; Chen, L.; Prive, G. Crystal structures of saposins A and C. *Protein Science* **2006**, *15*, 1849–1857.
60. Rossmann, M.; Schultz-Heienbrok, R.; Behlke, J.; Remmel, N.; Alings, C.; Sandhoff, K.; Saenger, W.; Maier, T. Crystal structures of human saposins C and D: Implications for lipid recognition and membrane interactions. *Structure* **2008**, *16*, 809–817.
61. Gally, J.; Edelman, G. The effect of temperature on the fluorescence of some aromatic amino acids and proteins. *Biochimica et Biophysica Acta* **1962**, *60*, 499–509. doi:10.1016/0006-3002(62)90869-7.
62. Hennecke, J.; Sillen, A.; Huber-Wunderlich, M.; Engelborghs, Y.; Glockshuber, R. Quenching of tryptophan fluorescence by the active-site disulfide bridge in the DsbA protein from Escherichia coli. *Biochemistry* **1997**, *36*, 6391–6400. doi:10.1021/bi963017w.
63. Frauenfeld, J.; Löving, R.; Armache, J.P.; Sonnen, A.F.P.; Guettou, F.; Moberg, P.; Zhu, L.; Jegerschöld, C.; Flayhan, A.; Briggs, J.A.G.; Garoff, H.; Löw, C.; Cheng, Y.; Nordlund, P. A saposin-lipoprotein nanoparticle system for membrane proteins. *Nature Methods* **2016**, *13*, 345–351.
64. Chien, C.T.H.; Helfinger, L.R.; Bostock, M.J.; Solt, A.; Tan, Y.L.; Nietlispach, D. An adaptable phospholipid membrane mimetic system for solution NMR studies of membrane proteins. *Journal of the American Chemical Society* **2017**, *139*, 14829–14832.
65. Flayhan, A.; Mertens, H.D.; Ural-Blimke, Y.; Molledo, M.M.; Svergun, D.I.; Löw, C. Saposin lipid nanoparticles: A highly versatile and modular tool for membrane protein research. *Structure* **2018**, *26*, 345–355.
66. Ladizhansky, V. Applications of solid-state NMR to membrane proteins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2017**, *1865*, 1577–1586. doi:10.1016/j.bbapap.2017.07.004.
67. Denisov, I.G.; Sligar, S.G. Nanodiscs for structural and functional studies of membrane proteins. *Nature Structural & Molecular Biology* **2016**, *23*, 481–486. doi:10.1038/nsmb.3195.
68. Zilm, K.W. Oral Presentation at the 40th Experimental NMR Conference. Orlando, FL, 1999.

69. McDermott, A.; Polenova, T.; Böckmann, A.; Zilm, K.W.; Paulson, E.K.; Martin, R.W.; Montelione, G.T. Partial NMR assignments for uniformly ($^{13}$C, $^{15}$N)-enriched BPTI in the solid state. *Journal of Biomolecular NMR* **2000**, *16*, 209–219.
70. Igumenova, T.I.; McDermott, A.E.; Zilm, K.W.; Martin, R.W.; Paulson, E.K.; Wand, A.J. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *Journal of the American Chemical Society* **2004**, *126*, 6720–6727.
71. Franks, W.T.; Zhou, D.H.; Wylie, B.J.; Money, B.G.; Graesser, D.T.; Frericks, H.L.; Sahota, G.; Rienstra, C.M. Magic-angle spinning solid-state NMR spectroscopy of the $\beta$1 immunoglobulin binding domain of protein G (GB1): $^{15}$N and $^{13}$C chemical shift assignments and conformational analysis. *Journal of the American Chemical Society* **2005**, *127*, 12291–12305.
72. Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **2002**, *420*, 99–102.
73. Lange, A.; Becker, S.; Seidel, K.; Giller, K.; Pongs, O.; Baldus, M. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angewandte Chemie* **2005**, *44*, 2089–2092.
74. Sperling, L.J.; Berthold, D.A.; Sasser, T.L.; Jeisy-Scott, V.; Rienstra, C.M. Assignment strategies for large proteins by magic-angle spinning NMR: the 21-kDa disulfide bond forming enzyme DsbA. *Journal of Molecular Biology* **2010**, *399*, 268–282.
75. Manolikas, T.; Herrmann, T.; Meier, B.H. Protein structure determination from $^{13}$C spin-diffusion solid-state NMR spectroscopy. *Journal of the American Chemical Society* **2008**, *130*, 3959–3966.
76. van der Wel, P.C.A. New applications of solid-state NMR in structural biology. *Emerging Topics in Life Sciences* **2018**, *2*, 57–67.
77. Shi, L.; Ahmed, M.A.; Zhang, W.; Whited, G.; Brown, L.S.; Ladizhansky, V. Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump–Structural insights. *Journal of Molecular Biology* **2009**, *386*, 1078–1093.
78. Shahid, S.A.; Bardiaux, B.; Franks, W.T.; Krabben, L.; Habeck, M.; van Rossum, B.J.; Linke, D. Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nature Methods* **2012**, *9*, 1212–1217.
79. Mandala, V.S.; Williams, J.K.; Hong, M. Structure and dynamics of membrane proteins from solid-state NMR. *Annual Review of Biophysics* **2018**, *47*, 201–222.
80. Ulrich, E.L.; Akutsu, H.; Doreleijers, J.F.; Harano, Y.; Ioannidis, Y.E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C.F.; Tolmie, D.E.; Wenger, R.K.; Yao, H.; Markley, J.L. BioMagResBank. *Nucleic Acids Research* **2008**, *36*, D402–D408.