BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

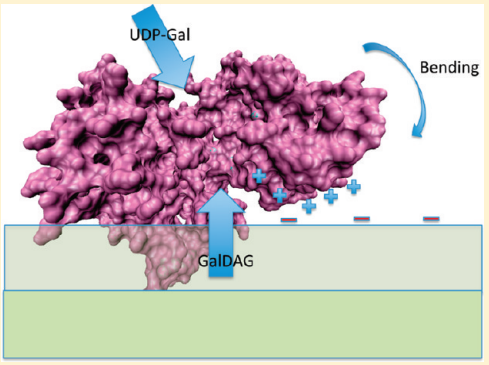
pubs.acs.org/biochemistry

Lipid Interacting Regions in Phosphate Stress Glycosyltransferase atDGD2 from *Arabidopsis thaliana*

Scarlett Szpryngiel, Changrong Ge, Irina Iakovleva, Alexander Georgiev, Jesper Lind, Åke Wieslander, and Lena Mäler*

Department of Biochemistry and Biophysics, Center for Biomembrane Research, The Arrhenius Laboratory, Stockholm University, 10691 Stockholm, Sweden

ABSTRACT: Membrane lipid glycosyltransferases (GTs) in plants are enzymes that regulate the levels of the non-bilayer prone monogalactosyldiacylglycerol (GalDAG) and the bilayer-forming digalactosyldiacylglycerol (GalGalDAG). The relative amounts of these lipids affect membrane properties such as curvature and lateral stress. During phosphate shortage, phosphate is rescued by replacing phospholipids with GalGalDAG. The glycolsyltransferase enzyme in *Arabidopsis thaliana* responsible for this, atDGD2, senses the bilayer properties and interacts with the membrane in a monotopic manner. To understand the parameters that govern this interaction, we have identified several possible lipid-interacting sites in the protein and studied these by biophysical techniques. We have developed a multivariate discrimination algorithm that correctly predicts the regions in the protein that interact with lipids, and the interactions were confirmed by a variety of biophysical techniques. We show by bioinformatic methods and circular dichroism (CD), fluorescence, and NMR spectroscopic techniques that two regions are prone to interact with lipids in a surface-charge dependent way. Both of these regions contain Trp residues, but here charge appears to be the dominating feature governing the interaction. The sequence corresponding to residues 227−245 in the protein is seen to be able to adapt its structure according to the surface-charge density of a bilayer. All results indicate that this region interacts specifically with lipid molecules and that a second region in the protein, corresponding to residues 130−148, also interacts with the bilayer. On the basis of this, and sequence charge features in the immediate environment of S227−245, a response model for the interaction of atDGD2 with the membrane bilayer interface is proposed.



Plants depend on a complex molecular machinery to maintain a functional cell membrane during stress conditions. The polar lipids that are most abundant in chloroplasts (thylakoids) are glycosylated and lack phosphate head groups, in contrast to the phosphatidylcholine-rich bilayers seen in other eukaryotic organelle membranes. These lipids are monogalactosyl-diacylglycerol (GalDAG: 1,2-diacyl-3-*O*-(β-D-galactopyranosyl)-*sn*-glycerol) and digalactosyl-diacylglycerol (GalGalDAG: 1,2-diacyl-3-*O*-[α-D-galactopyranosyl-(1→6)-*O*-β-D-galactopyranosyl]-*sn*-glycerol).[1] GalDAG lipids, containing only one sugar moiety, are prone to form non-bilayer structures, due to their relatively small headgroup size, while GalGalDAG forms bilayers. The amount of these lipids in the membrane affects the spontaneous curvature of the bilayer and therefore also physical properties such as the lateral stress. At least five different enzymes are responsible for producing GalDAGs and GalGalDAGs in higher plants under normal, stress, and various developmental conditions, and the relative amount of the lipids produced is ultimately controlled by these enzymes and their interactions with the local bilayer environment. It has been indicated that the physical properties of the membrane seem to control the activity of the enzymes and hence dictates the production of lipids necessary for maintaining a viable membrane under the prevailing conditions,[2] analogous to certain bacteria.[3] Hence, lipid recognition and membrane interaction of the enzymes are key elements for understanding the control in lipid production.

GalDAG consists of two main parts: a lipid precursor, *sn*-1,2-diacylglycerol (DAG), that is synthesized in the endoplasmic reticulum and plastids and then transported to the chloroplasts, and a galactose headgroup originating from UDP-galactose, which is attached by GalDAG synthases (MGDs). GalDAG synthases are lipid galactosyltransferases (GTs) that together with GalGalDAG synthases are responsible for the trafficking and incorporation of lipids into the thylakoid membranes. GalGalDAG synthases (atDGD1 and atDGD2) attach yet another sugar moiety to the GalDAG lipids, resulting in GalGalDAG. atDGD1 is the main ("default") enzyme for this process in the green tissues of *Arabidopsis thaliana* (leaves, etc.), whereas atDGD2 seems to be more important in the root tissue, triggering GalGalDAG production during specific growth conditions such as phosphate or draught stress, but without GalDAG accumulation as in the thylakoids. As part of a rescue mechanism,

Received: February 1, 2011
Revised: April 20, 2011
Published: April 20, 2011

phosphate shortage stress results in breakdown of phospholipids in the thylakoid and root membranes and a necessary substitution of them with another bilayer-forming lipid, that is, GalGalDAG, in order to save phosphate for other central metabolic purposes.[4−8]

atDGD2 in *A. thaliana* was first discovered by Dörmann et al in 1999, based on similarity searches with BLASTP for the atDGD1 sequence.[9] The existence of the protein was then verified experimentally.[6] The complete sequence consists of 473 amino acid residues, with a total molecular mass of 54 kDa. The majority of glycosyltransferases adopt single or double Rossmann fold domains, and many lack transmembrane segments.[10−12] There is no crystal or NMR structure available for atDGD2, although there is a modeled structure, based on sequence similarity and fold recognition with others GTs, which shows that the protein adopts a two-domain (GT-B) Rossman fold.[13] The protein does not contain any known transmembrane segments but is instead believed to have a monotopic interaction with the lipid bilayer interface. It is however unknown which part(s) of the enzyme is responsible for lipid interactions.

There are several ways a protein may interact monotopically with lipids in a biological membrane. One mechanism involves anchoring with the aid of one or more amphipathic helices, often in combination with additional electrostatic interactions by charged (cationic) amino acids. Certain amino acid residues have been shown to be more abundant than others in these regions; especially tryptophan is often mentioned as important for membrane interaction for both transmembrane and monotopic membrane proteins.[14−16] We have previously shown that the structurally related, prokaryotic lipid GT alMGS may dock into the membrane interface by an amphipathic helix enriched in Lys/Arg pairs, while it has been suggested that membrane interaction of MurG, a key enzyme in bacterial cell wall synthesis, is facilitated by surface-located Trp residues.[12] Tryptophans are unusually abundant in atDGD2 (with a total number of eight tryptophans in the sequence) and also in the other *A. thaliana* MGD and DGD enzymes.[17] These Trp residues are believed to be surface-exposed, and regions in the protein containing most of the Trps have also been shown to be extremely important for acitivity.[13] If one of the tryptophans in these regions is exchanged for Phe or Ala, the activity of the enzyme drops to almost zero. The role of the Trp residues is not elucidated but could point to membrane interaction or substrate (lipid) recognition/binding. In transmembrane segments, both cationic and tryptophan residues are crucial for the interactions with and orientations in the lipid bilayer.[18,19] In the eukaryotic organelle membranes, transmembrane segment length and amino acid sequence features seem to vary as adaptions to the organelle lipid bilayers.[20] For monotopic GTs, the exact localization and internal properties of the bilayer-interacting segments will determine how the enzyme is oriented in the interface, to be able to properly reach both its bilayer acceptor.

In this study, we wished to elucidate the mechanism by which the atDGD2 protein interacts with a bilayer, in order to better understand orientation and regulation of the enzyme. Many segments in atDGD2 can be predicted to interact with the bilayer interface, and several of these were previously identified as crucial for enzymatic activity.[13] In order to elucidate which part(s) of this enzyme is responsible for membrane interaction, as well as to see if it is possible to map out any substrate interacting areas, we used a combination of techniques. The segments were identified using the MPEx algorithm, based on hydropathy,[21] and their potential membrane interaction properties were further examined by a novel multivariate comparison of polypeptide sequence properties. These segments may express some property differences from the full-length enzyme, but the main characteristics should be the same. We have used fluorescence and NMR spectroscopy as well as a binding assay based on GFP constructs to study the extent of membrane interaction of four selected segments. We have also performed a structural study, with circular dichroism (CD) spectroscopy, followed by a NMR structure determination of one of the segments identified as crucial for interacting with lipids in the bilayer. Taken together, our results lead to a response model for how atDGD2 interacts with the membrane bilayer interface.

EXPERIMENTAL PROCEDURES

Materials. Five peptides derived from atDGD2 corresponding to the sequences 11−29, 46−64, 130−148, 169−187, and 227−245 in the full-length protein were purchased from NeoMPS (Strasbourg, France). 1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-*rac*-glycerol) (POPG), deuterated 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC-d_{54}), and 1,2-dihexanoyl-*sn*-glycero-3-phosphocoline (DHPC-d_{22}) were obtained from Avanti Polar Lipids (Alabaster, AL, USA). 1,2-Dimyristoyl-*sn*-glycero-3-[phosphor-*rac*-(1-glycerol)] (DMPG-d_{54}) was from Larodan AB (Malmö, Sweden), and dodecylphosphocholine (DPC-d_{38}) was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA).

Preparation of Vesicles for CD and Fluorescence Measurements. Samples containing vesicles were prepared by the extruder technique. Vesicles with zwitterionic lipids were prepared with POPC, and vesicles containing negative headgroup charge were prepared by substituting 10, 20, or 30 mol% of POPC for POPG. The lipids were dissolved in chloroform and dried under nitrogen gas until a uniform lipid film was attained. It was then dissolved in 50 mM phosphate buffer, pH 5.7. The sample was vortexed extensively to produce large multilamellar vesicles (LMVs). The LMVs then went through a freeze−thaw cycle in liquid nitrogen and a 40 °C water bath. This was done five times to reduce the lamellarity and achieve large unilamellar vesicles (LUVs). The solution was then extruded through double 100 nm polycarbonate Nuclepore membranes 21 times to achieve a uniform sample and thereafter diluted to a final lipid concentration of 1 mM.[22] The vesicles were stored at +5 °C between measurements for a maximum of one week. The size of the vesicles was estimated by dynamic light scattering (DLS). DLS experiments were conducted on a CGS-3 Compact Goniometer from ALV-GmbH using a 5 mm diameter tube. Three measurements were averaged and the hydrodynamic radius was calculated with the manufacturer software (data not shown) to verify a homogeneous vesicle size. Peptide was added to the final vesicles to a concentration of 50 μM (i.e., peptide/lipid 1:20 mol/mol) and then incubated at 25 °C for 1 h before measurements.

Construction of Vectors for GFP-atDGD2 Segments. DNA oligonucleotide pairs encoding the desired segments with overlapping ends encoding restrictions sites were synthesized by Eurofins MWG Operon (Germany). Four oligonucleotides coding for the same segment (0.15 nM each) were mixed in a buffer containing 20 mM TrisCl, pH 8.0 and 10 mM $MgCl_2$, heated for 5 min at 95 °C and then cooled to room temperature.

The assembled oligonucleotide pairs were then ligated into *Xma*I and *Hind*III sites of the modified pET43.1b-GFP(+) vector in which GFP-ORF was introduced between NdeI and SacII sites. The linker region between GFP and the DGD2 segments contains 27 amino acids (SAGKETAAAKFERQHMDSPPPT GLVPR). Finally, these constructs were transformed into *Escherichia coli* (BL21 Star(DE3) pLysS) for recombinant protein expression.

Expression and Purification of GFP-atDGD2 Segments. Expression and purification of GFP fusions was performed as described before.[13] Briefly, *E. coli* cells harboring the GFP-fused segments were harvested and resuspended in lysis buffer (100 mM TrisCl, pH 7.4, 50 mM NaCl, 50 mM KAc, 10% glycerol) and then were broken by French Press. The suspension was subsequently cleared by ultracentrifugation at 15000*g* for 15 min. HisTrapHP column (GE Lifescience) was used in the purification step, and the GFP-fused segments were eluted with elution buffer (100 mM Tris-Cl pH 8.0, 250 mM NaCl, 250 mM imidazol, 10% glycerol). Purified peptides were stored at 4 °C until use after addition of NaN_3 (0.1% final conc).

Vesicle Binding Assay for atDGD2 Segments. The binding assay in vitro with vesicles was also performed as described before.[13] Solutions of sucrose-loaded large unilamellar vesicles (LUVs) with the desired lipid composition were produced by passing the solution through two layers of polycarbonate membranes with 100 nm pore size using a Mini Extruder (Avanti Polar Lipids) and stored at +5 °C until use but not more than a week. The binding of the various atDGD2-GFP fusion peptides to sucrose-loaded vesicles was measured using the ultracentrifugation technique described previously.[23] The sucrose-loaded vesicles (0.5 mM total lipid) were mixed with the purified peptides at a peptide/lipid ratio of 1/500 in a total volume of 100 μL of vesicle solution for 15 min at RT and then centrifuged at 100000*g* for 1 h at room temperature. GFP fluorescence in the supernatant and pellet was measured with an excitation wavelength of 390 nm, emission wavelength of 530 nm, and a 495 nm cutoff filter using a SpectraMaxGemini EM (Molecular Devices, California). The amount of proteins and lipids were quantified both in the supernatant and pellet to evaluate the percentage of peptides bound to vesicles.

Fluorescence Spectroscopy. Intrinsic tryptophan fluorescence was used to study the local environment of the synthetic peptides dissolved in various solvents. All fluorescence measurements were performed on a FluoroLog spectrometer (HORIBA Jobin Yvon) at 25 °C. The excitation wavelength was 280 nm and the emission was observed between 300 and 450 nm. In the quenching studies, a stock solution of 1 M acrylamide was used to quench the fluorescence of the tryptophans. The acrylamide solution was titrated into the cuvette and mixed gently before measuring. The incubating time between measurements was 3 min to allow for complete diffusion and solubilization of the quencher. The resulting acrylamide concentration ranged between 4 and 72 mM. In order to calculate the quenching constant, the Stern-Volmer equation was used:

$$\frac{F_0}{F} = 1 + K_{SV}[Q] \qquad (1)$$

where F_0 and F are the fluorescence intensities without and with added quencher, K_{SV} is the quenching constant, and $[Q]$ is the concentration of added quencher.[24,25]

CD Spectroscopy. The CD measurements were done with a Chirascan circular dichroism spectrometer from Applied Photophysics. The temperature was adjusted to 25 °C by a Quantum Northwest TC 125 temperature controller. The signal was detected between 190 and 260 nm using a 0.5 step resolution. A 0.1 cm quartz cuvette was used, and 10 spectra were collected and averaged. Background spectra of the buffer solution or the various LUV solutions were recorded for all samples, and these spectra were subtracted from the spectra recorded for the peptides. The CD spectra were analyzed with Dichroweb[26,27] using the CONTIN method.[28]

NMR Spectroscopy. In all NMR experiments, a peptide concentration of 500 μM was used, and the pH value of the sample was adjusted with 50 mM phosphate buffer to 5.7. All experiments were carried out at 25 °C. The homonuclear two-dimensional (2D) experiments for structure analysis were measured on a Bruker Avance NMR spectrometer operating at a 500 MHz ^{1}H frequency and equipped with a cryo-probe. The peptide was dissolved in 50 mM DPC-d_{38} with 10% of D_2O added for the field/frequency lock. DLS measurements confirmed a homogeneous micelle size and interaction of the peptide with the micelles. TOCSY[29] spectra with mixing times of 30, 60, and 90 ms and NOESY[30] spectra with mixing times of 100 and 300 ms were recorded. 32 scans were acquired and 2048 × 512 data points collected in the direct and indirect acquisition dimensions. Processing of the NMR data was done with TopSpin 2.1 (Bruker BioSpin) and the assignment was done with Sparky.[31]

The translational diffusion experiments were measured on a Bruker Avance NMR spectrometer operating at a 600 MHz ^{1}H frequency. For the diffusion experiments, zwitterionic and negatively charged bicelles (20 mol% PG) were used. The zwitterionic bicelles were prepared by mixing chain-deuterated DMPC with chain-deuterated DHPC (300 mM lipids in total, $q = 0.15$). This mixture was vortexed in several steps until a clear solution was obtained. Partly anionic bicelles were prepared in the same way, but 20 mol% of the DMPC was exchanged for DMPG. The samples were prepared in D_2O to remove the water signal and to reduce overlap in the aromatic region of the peptide NMR spectra. The pulsed field gradient spin−echo (PFGS) experiment with bipolar gradient pulses[32] was used in order to estimate the translational diffusion constants. In the PFGS experiment, the gradient strength was increased in 32 steps between 5 and 95% gradient strength. Five experiments with 32 or 8 scans were acquired for samples with and without peptide, respectively. Three experiments to measure the diffusion of water were also measured for each sample condition in order to calculate the effect of bicelles and peptides on the viscosity in the sample. For the measurements of diffusion constants for bicelles or peptides, δ and Δ were set to 2 and 300 ms, respectively. For the water diffusion experiments, δ and Δ were 6 and 90 ms. The obtained signal intensities of the spectra measured for each sample were then fitted to the Stejskal−Tanner equation:[33]

$$\frac{I}{I_0} = e^{-\gamma^2 g^2 D \delta^2 (\Delta - \delta/3)} \qquad (2)$$

where I and I_0 is the measured intensities without and with an applied gradient pulse, γ is the gyromagnetic ratio of the studied nucleus, g is the gradient strength, D is the diffusion constant, δ is the gradient pulse length, and Δ is the total refocusing time. Correction for possible nonlinearity in the gradients, according to Damberg et al, was done to improve the accuracy of the diffusion constants.[34] The calculated diffusion constants for each sample were then averaged over all measurements. The diffusion

constants of the peptides together with diffusion constants for the DMPC lipids in the bicelles can be used to estimate the amount of peptide that is bound to the bicelles. This can be done by assuming that all DMPC molecules participate in bicelle formation, and assuming that the diffusion coefficient for DMPC is a measure of the bicelle translational movement. If a two-state exchange process is assumed between the bicelle-bound peptide and the peptide in solution, it is possible to estimate the extent of interaction by using the following relation:[35]

$$xD_{\mathrm{DMPC/DMPG}} + (1-x)D_{\mathrm{free}} = D_{\mathrm{bound}} \qquad (3)$$

where $D_{\mathrm{DMPC/DMPG}}$ is the diffusion coefficient of the bicelle, D_{free} is the diffusion coefficient of the peptide in buffer, D_{bound} is the diffusion coefficient of the peptide in the bicelle-solution, and x is the amount of peptide that is bound to the bicelle.

Structure Calculation. The structural calculation for S227−245 in 50 mM DPC micelles was performed with the CYANA software[36] and was based on the chemical shift assignment and 162 proton−proton distance constraints derived from the NOESY with a mixing time of 100 ms. Distances were determined from NOESY cross-peak volumes by the use of CALIBA[37] in CYANA and optimized manually at later stages in the calculation. The chemical shift index method was used for determination of H^{α} secondary chemical shifts and H^{N} average chemical shifts.[38] A total of 100 structures were calculated and 23 of the structures were selected based on the constraint violations to give the final ensemble. PROCHECK-NMR[39] was used to analyze the quality of these structures.

Databases and Web Tools. The complete atDGD2 sequence (UniProt ID: Q8W1S1) was analyzed by the Membrane Protein Explorer algorithm (http://blanco.biomol.uci.edu/mpex).[21,40] The interfacial (IF) hydrophobicity scale[41] was used and the window size was 19 amino acid residues.

The National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov), Uniprot (Universal Protein Resource) database (http://www.uniprot.org), and PDB database (http://www.pdb.org) were used to procure GT-B family protein sequences in this study. The TargetP 1.1 server,[42,43] available at the Center for Biological Sequence Analysis (CBS; http://www.cbs.dtu.dk/index.shtml), was used to identify chloroplast transit peptides, mitochondrial targeting peptides, or secretory pathway signal peptides sequences, which were removed in the final model building. Potential membrane-interacting surfaces in atDGD2 was searched using the MAPAS prediction facility.[44] The hydrophobic moment ("amphipathicity") was calculated according to Eisenberg[45] (MOMENT at http://fold.doe-mbi.ucla.edu/services). All protein sequences of the data set are available as fasta files upon request.

Multivariate Data Analysis. Partial least-squares projections to latent structures discriminant analysis (PLS-DA) is a frequently used classification method and was developed to find the relationships between a matrix X and a matrix Y.[46] Usually, matrix Y is composed of "dummy" variables; hence a value of 1 is given to one class and 0 to the other class. In the present study, PLS-DA was based on the assumption that polypeptide sequences (GTs) belonging to the same class (membrane-associated or not) have common features and therefore are likely to behave similarly in the analysis. Amino acid features in each polypeptide (GT in this case) can be described by a number of measured and computed parameters, such as hydration potential, isoelectric point, molecular mass, etc. However, to narrow down the number of parameters describing each amino acid, so-called z scales were used in the present study. These z scales are derived from a multiproperty matrix comprising 29 physical-chemical experimental parameters for the 20 coded amino acids with principal component analysis. They can be tentatively interpreted as z1 for "hydrophobicity" z2 for "bulk of side-chain," and z3 for "electronic properties".[47] Each amino acid residue in a polypeptide sequence was represented by three z-scales (Figure 1) and then the periodic properties in a polypeptide sequence were calculated by autocross-covariance (ACC), which combines auto covariances between the same z-scale (property) in each position, and cross covariances, between two different z-scales in each position.[48] The auto covariances with lags = 1, 2, ... l are given by

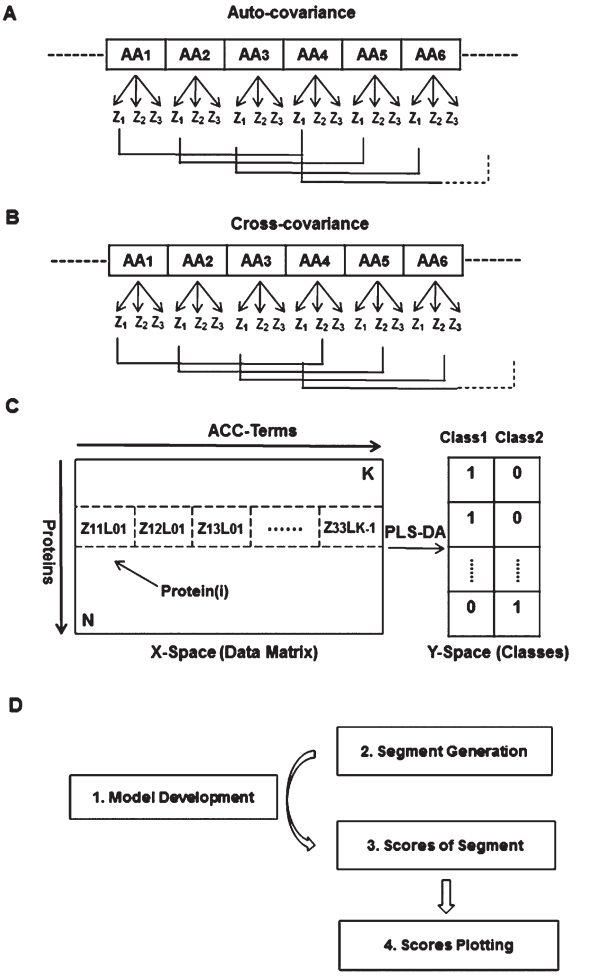


Figure 1. Schematic illustration for mutivariate data analysis rationale and steps performed in establishing a MVDA model. Protein sequences from the membrane-associated GTs and soluble GTs classes (from Table 1), are translated into three z scales (z1, aa hydrophobicity/hydrophilicity; z2, volume; and z3, polarizability) and used for (A) calculation of autocovariance and (B) cross-covariance variables ("ACC"), and (C) collection into data matrices from which partial-least-squares discriminant analyses (PLS-DA) was applied to correlate variables in X space and Y space. (D) The steps used in the present study for predicting membrane bound segments in atDGD2 (c.f. prediction by multivariate data analysis model in Experimental Procedures).

$$\mathrm{ACC}_{j,l} = \sum_{i}^{n-l} \frac{z_{j,i} \times z_{j,i+l}}{n-l} \qquad (4)$$

dx.doi.org/10.1021/bi200162f |Biochemistry 2011, 50, 4451–4466

where the index j is used for the z scales ($j = 1, 2, 3$), index i is the amino acid position ($i = 1, 2, \ldots n$), and n is the number of amino acids in the polypeptide sequence. Cross covariances between two different properties, j and k, are calculated from

$$\mathrm{ACC}_{j\neq k,l} = \sum_{i}^{n-l} \frac{z_{j,i} \times z_{k,i+l}}{n-l} \quad (5)$$

From eqs 4 and 5, the ACC terms were calculated for each polypeptide and a data matrix X was created (see Figure 1 for a schematic description).

In this way, identified polypeptide sequences (GTs) in the present study were divided into two classes, that is, 24 membrane-associated and 26 soluble proteins (Table 1), and then each polypeptide sequence was translated into a series of variables calculated by the aforementioned autocross-covariance (ACC). The final result is a multivariate data matrix (matrix X) with (24 + 26) rows and a number of columns containing variables (ACC values). PLS-DA was performed to find a model that could separate these sequences according to whether they are soluble or membrane-bound, based on their periodic physical properties, or ACC terms. Cross validation (CV) was performed to evaluate the predictive ability of the obtained models. In short, all objects (polypeptides) in matrix X are divided into a number of groups and by developing a number of parallel models from reduced data in which data from one of the groups was deleted, a new data matrix, Y, was predicted from an updated model based on the remaining groups (objects). A $Q_{2\mathrm{cum}}$ (goodness of prediction) value was calculated that described how much of the variance in the Y matrix that could be predicted by the developed model. To obtain a perfect score of 1, all objects should be predicted back to the exact position given by the Y matrix. A $Q_{2\mathrm{cum}}$ value larger than 0.1 corresponds to 95% significance of the model.[49] Response permutation was also performed to evaluate the statistical significance of the calculated $Q_{2\mathrm{cum}}$ value. The SIMCA-P+ software (version 11.0, Umetrics AB, Umeå, Sweden) was used for the PLS-DA calculation.

Prediction by Multivariate Data Analysis Model. The multivariate data analysis in our study was performed in four steps (Figure 1D):

(i) A model was built by using partial least-squares discriminant analysis (PLS-DA) (described above) to examine the subcellular localization (membrane associated or not) of the full-length enzyme sequences in the data set.

(ii) In order to use this model to determine (predict) regions of the protein that possess membrane-binding capacity, we generated a set of peptides from each protein in the data set corresponding to a sliding window with user-defined length.

(iii) A score was calculated by using the model obtained in (i) to predict the membrane interaction properties (associated or not) of each peptide generated in (ii), and then

(iv) All the scores obtained from all the peptides (in one protein) were plotted against the protein sequence.

RESULTS

Multivariate Discriminant Analysis of Soluble and Membrane-Bound GTs. The two-Rossmann-fold-domain organization is common for many glycosyltransferases (GTs), both soluble and membrane-bound. Of the latter, bilayer interface-associated ones lacking the typical hydrophobic transmembrane segment are difficult to discriminate from the soluble ones, because both these types contain peptide segments of intermediate hydrophobicity, which may be of surface or internal localization. To better find and explore the possible regions that are important for membrane interaction, we examined atDGD2 with a multivariate discriminant analysis where models were made that distinguish (many) soluble from membrane-associated (monotopic) GTs. The proteins used in the analysis and their classification according to being soluble or membrane-associated are listed in Table 1. Principally, all combinations of side-chain hydrophobicity (Z1), volume (Z2), and "charge" (Z3) (i.e., polarizability) for all amino acid residues within a sliding window are analyzed for position-specified correlations, for all sequences in the two classes (see Experimental Procedures). A good separation of glycosyltransferases of the GT-B fold family was achieved with a PLS-DA (partial least-squares projections to latent structures discriminant analysis)[50] in terms of being membrane-associated or soluble. Cross validation was applied to evaluate the models, and by varying the sliding window length (formally "lag"), an optimal length of 22 amino acid residues was found. The score plot obtained from the model (with window length of 22 amino acid residues) reveals two well separated groups (Figure 2), and the model has three significant PLS components and a $Q_{2\mathrm{cum}}$ value (model quality parameter where 1.0 is maximum) of 0.56, that is, a good model. To examine the most important features that separated soluble GTs from membrane-bound, an analysis of the distribution of variables that contributed to the separation was made. The highest correlations were Z33L20 (reads as Z3 × Z3 for Lag 20, i.e., between positions 1 and 21), Z12L13, Z31L13, Z21L19, Z11L01, Z33L14, Z33L10, Z13L07, Z31L02, Z33L02, Z33L06 which are all positively correlated with membrane-associated GTs. Hence, the variable Z33L20 can be interpreted as a recurrent correlation between polarizability/charge at sequence position 1 with polarizability/charge at position 21 (distance between the two sites is Lag 20), and Z12L13 can be interpreted as a correlation between the hydrophobicity/hydrophilicity of site 1 and the size of the side-chain of residue 14. Of the aforementioned top variables, Z3 (polarizability/charge) is the most common variable correlating with either Z1 or Z3 in different positions alongside the sequence, which indicates that charged and hydrophobic amino acids are important for the membrane association process among these monotopic GTs. In our experience, short Lags, like 01, 02, 06, and 07 above, are often associated with amphipathicity, that is, properties at certain positions ("sidedness") in helices (data not shown).

Prediction of Membrane Bound Segments by MVDA-Model and MPEx. The MPEx analysis of sequence segments revealed nine potential membrane interface-binding segments in atDGD2 (http://blanco.biomol.uci.edu) as employed before,[13] of which the five most pronounced contained at least one Trp residue (Figure 3A). However, from the comparison here, it is obvious that these segments seem not to occupy the same positions in six other interface-associated GTs, where sequences were aligned with ClustalW (Figure 3A). Still, note the fairly good alignment of the secondary structure elements for these established three-dimensional (3D) structures. To improve the MVDA model and better discriminate between internal segments and "true" membrane-interacting (surface-exposed) sequences, we screened for the criteria found to be important for discriminating between soluble and membrane-bound GTs

Table 1. Experimentally Localized GTs Used in the MVDA Analysis

organism	protein name	SwissProt Code
	Soluble GTs[a]	
Escherichia coli	alpha,alpha-trehalose-phosphate synthase [UDP-forming]	OTSA_ECOLI
Escherichia coli	maltodextrin phosphorylase	PHSM_ECOLI
Escherichia coli	glycogen synthase	GLGA_ECO57
Escherichia coli	UDP-*N*-acetylglucosamine 2-epimerase	WECB_ECOLI
Candida albicans	trehalose-6-phosphate phosphatase	Q5AI14_CANAL
Enterobacteria phage T4	DNA beta-glucosyltransferase	GSTB_BPT4
Bacillus subtilis	glycogen synthase	GLGA_BACSU
Aspergillus niger	alpha,alpha-trehalose-phosphate synthase	TPSA_ASPNG
Arabidopsis thaliana	alpha,alpha-trehalose-phosphate synthase [UDP-forming]	TPS1_ARATH
Amycolatopsis orientalis	glycosyltransferase GtfD	Q9AFC7_AMYOR
Amycolatopsis orientalis	glycosyltransferase GtfA	P96558_AMYOR
Rhizobium radiobacter	glycogen synthase 1	GLGA1_RHIRD
Pyrococcus abyssi	GlgA glycogen synthase	Q9V2J8_PYRAB
Oryctolagus cuniculus	glycogen phosphorylase, muscle form	PYGM_RABIT
Enterobacteria phage T4	DNA alpha-glucosyltransferase	GSTA_BPT4
Streptomyces fradiae	glycosyl transferase	Q9RPA7_STRFR
Arabidopsis thaliana	probable hydroquinone glucosyltransferase	HQGT_ARATH
Amycolatopsis orientalis	glycosyltransferase GtfB	P96559_AMYOR
Chlamydia trachomatis	protein CT_858	Y858_CHLTR
Streptomyces antibioticus	oleandomycin glycosyltransferase	OLED_STRAT
Streptomyces antibioticus	oleandomycin glycosyltransferase	Q3HTL7_STRAT
Medicago truncatula	triterpene UDP-glucosyl transferase UGT71G1	Q5IFH7_MEDTR
Vitis vinifera	anthocyanidin 3-*O*-glucosyltransferase 2	UFOG_VITVI
Corynebacterium glutamicum	D-inositol-3-phosphate glycosyltransferase	MSHA_CORGL
Medicago truncatula	(iso)flavonoid glycosyltransferase	A6XNC5_MEDTR
Halothermothrix orenii	sucrose-phosphate synthase	B8CZ51_HALOH
	Membrane GTs[a]	
Synechocystis sp.	digalactosyldiacylglycerol synthase (Slr508)	P73948_SYNY3
Mycobacterium smegmatis	phosphatidylinositol alpha-mannosyltransferase	PIMA_MYCS2
Streptomyces viridochromogenes	putative glycosyltransferase	Q93KV2_STRVR
Streptococcus pneumoniae	glycosyl transferase CpoA	O06452_STRPN
Streptococcus pneumoniae	glycosyl transferase, group 1	Q97QX1_STRPN
Staphylococcus aureus	processive diacylglycerol glucosyltransferase	UGTP_STAAC
Pseudomonas paucimobilis	beta-1,4-glucuronosyltransferase	Q9APE1_PSEPA
Pseudomonas aeruginosa	rhamnosyl transferase 2	Q51560_PSEAE
Neisseria meningitidis	undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase	MURG_NEIMA
Cucumis sativus	monogalactosyldiacylglycerol synthase, chloroplastic	MGDG_CUCSA
Borrelia burgdorferi	monogalactosyldiacylglycerol synthase (BB0454)	O51410_BORBU
Bacillus subtilis	processive diacylglycerol glucosyltransferase	UGTP_BACSU
Avena sativa	UDP-glucose:sterol glucosyltransferase	O22678_AVESA
Arabidopsis thaliana	putative uncharacterized protein F7J8_200	Q9LFB4_ARATH
Arabidopsis thaliana	sulfolipid synthase	Q8S4F6_ARATH
Arabidopsis thaliana	UDP-glucose:sterol glucosyltransferase	O23649_ARATH
Arabidopsis thaliana	monogalactosyldiacylglycerol synthase 3, chloroplastic	MGDG3_ARATH
Arabidopsis thaliana	digalactosyldiacylglycerol synthase 2, chloroplastic	DGDG2_ARATH
Arabidopsis thaliana	digalactosyldiacylglycerol synthase 1, chloroplastic	DGDG1_ARATH
Acholeplasma laidlawii	1,2-diacylglycerol-3-glucose (1−2)-glucosyltransferase	Q8KQL6_ACHLA
Acholeplasma laidlawii	1,2-diacylglycerol 3-glucosyltransferase	Q93P60_ACHLA
Escherichia coli	ADP-heptose--LPS heptosyltransferase 2	RFAF_ECOLI
Escherichia coli	lipopolysaccharide heptosyltransferase-1	Q8FC98_ECOL6
Escherichia coli	undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase	MURG_ECOLI

[a] 24 membrane-associated and 26 soluble GTs in the GT-B fold family with experimentally determined localization properties are included.

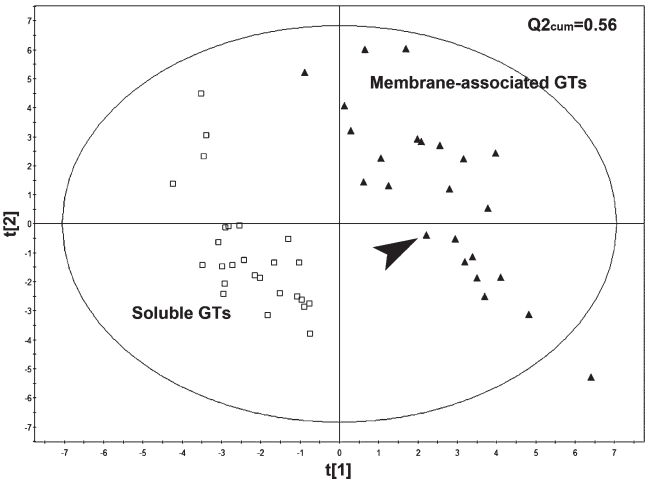


Figure 2. Separation of membrane-associated glycosyltransferases from soluble glycosyltransferases (GTs) based on amino acid properties (*z* variables). Score plot from the PLS-DA analysis of all ACC terms within a sliding window of 22 amino acid residues (Lag 21). Both the membrane-associated and soluble GTs form clear clusters separately, with three significant PLS components separating the two groups and a Q_{2cum} value for the model of 0.56. The first two significant components (score vectors t[1] and t[2]) are shown here. The score vectors t[1] or t[2] can be thought of as a new variables, which reflect the information in the original X-variables (calculated by ACC, c.f. Methods) that is of relevance for modeling and predicting the response variable in Y space (1 or 0, c.f. Methods). The set of sequences are shown in Table 1. atDGD2 is indicated by an arrow.

(above). The amino acid sequence of atDGD2 was divided into a large number of overlapping peptide sequences from the N-terminus to the C-terminus, using a sliding window length of 29 amino acid residues. Each of these segments in atDGD2 was predicted against the MVDA two-class model (above) to find their class belongings, and from this scores were obtained and plotted against the full atDGD2 sequence (Figure 3B). These visualize the probability of membrane association — a higher value indicates a membrane-interacting region. From Figure 3B, there are four regions with significantly higher values, which also overlap with the results predicted by MPEx for interface association (Figure 3A) (denoted S130−148, S81−99, S169−187, and S227−245), while the segments S11−29 and S46−64, shown earlier to be important for atDGD2 enzyme activity,[13] only appeared as potential membrane-interacting sequences in the MPEx analysis but not in the MVDA analysis here (Figure 3B). In addition, a segment downstream of S227−245, that is, 240−260, is predicted to have a high hydrophobic moment ("amphipathicity"; ranging 0.61−0.54), typical for bilayer surface-associated helices.[45]

On the basis of the MVDA and MPEx analyses, the four (out of these six) sequences that contained a Trp residue and sufficient ΔG values were chosen for further biophysical studies to examine their membrane interaction properties (i.e., S11−29, S46−64, S130−148, and S227−245). Sequence S81−99 does not contain a Trp residue (and has a small ΔG value), and S169−187 (large ΔG) was too hydrophobic to be able to study experimentally in isolation, and these were therefore omitted in this study. S169−187 could not be dissolved in either buffer or lipid vesicles, and it has recently been shown that this peptide lyses liposomes.[13] The aggregation propensities and bilayer-disrupting properties of this segment, the latter lethal for a living cell, indicate that it may be dependent on the rest of the protein for proper structure and hence appears not to be a likely candidate for a monotopic (interface) interaction when alone.

Binding of GFP-Fused Segments to Lipid Bilayers. To quantitatively evaluate the potential lipid binding properties of these four predicted segments in more detail, all of them were cloned as GFP fusions, which may also illustrate their abilities to keep a larger protein moiety at a membrane surface. Since properties related to charge was found to be important in the MVDA analysis (above), vesicle binding assays (liposome pull-down) were performed which enabled us to test the binding properties of these segments to lipid vesicles containing either 10 or 30% phosphatidylglycerol (PG) head-groups or phosphatidylcholine (PC) head-groups only (Figure 4), analogous to binding experiments with GFP-free peptides and bicelles; see below. Neither S11−29 nor S46−64 GFP fusions were observed to bind strongly to any of the vesicles (less than 5% bound) as compared to GFP control protein (Figure 4). S130−148 and S227−245 GFP fusions, on the other hand, exhibited a stronger binding to the PG-containing vesicles, with an increased fraction of bound fusion as the PG content increased. We also observed that these peptides bound somewhat less to vesicles containing PC-lipids only, in agreement with the bioinformatic analysis above indicating that charge was important. Furthermore, there was around 20% self-aggregation for these two GFP hybrids due to their hydrophobicity in vitro in the absence of lipids, which might indicate these two segments are more likely membrane-bound in vivo.

Fluorescence Spectroscopy. Potential contacts between synthetic peptides corresponding to the sequences S11−29, S46−64, S130−148, and S227−245, and bilayer interior regions were analyzed by fluorescence measurements in 50 mM phosphate buffer, and zwitterionic (PC) vesicles or anionic vesicles containing 10, 20, and 30 mol% PG in PC. Emission wavelength values and Stern-Volmer quenching constants are shown in Table 2. Both quenching constants and wavelength shifts indicate that S46−64, S130−148, and S227−245 all have some degree of interaction with the bilayer. In zwitterionic vesicles, the emission wavelengths are similar to those observed in buffer, while upon introduction of anionic headgroup charge (PG lipids), the emission wavelengths gradually shift to lower values with an increasing amount of PG for all segments, except for S11−29. For this segment, no blue-shift is observed. The blue-shift is indicative of a less hydrophilic environment for the tryptophans, that is, burial into the bilayer interior of the vesicles, and hence this interaction is increased with a higher amount of negatively charged lipids present in the bilayer for three of the peptides. S11−29 on the other hand does not seem to interact with any bilayer, which is also consistent with the binding assay and MVDA analysis, respectively.

CD Spectroscopy. Secondary structures of the peptides were investigated by CD measurements on the same samples (under the same conditions) that were used in the fluorescence experiments. The results from CD measurements in buffer and zwitterionic vesicles show a random coil conformation for all of the studied peptides (Figure 5). Introducing anionic lipid charge in the bilayer surface results in various structural effects for the different peptides. S11−29 and S46−64 do not alter their structure significantly, regardless of the extent of lipid surface charge. The differences in the relative intensities of the spectra are most likely related to solubility problems for the fragments,

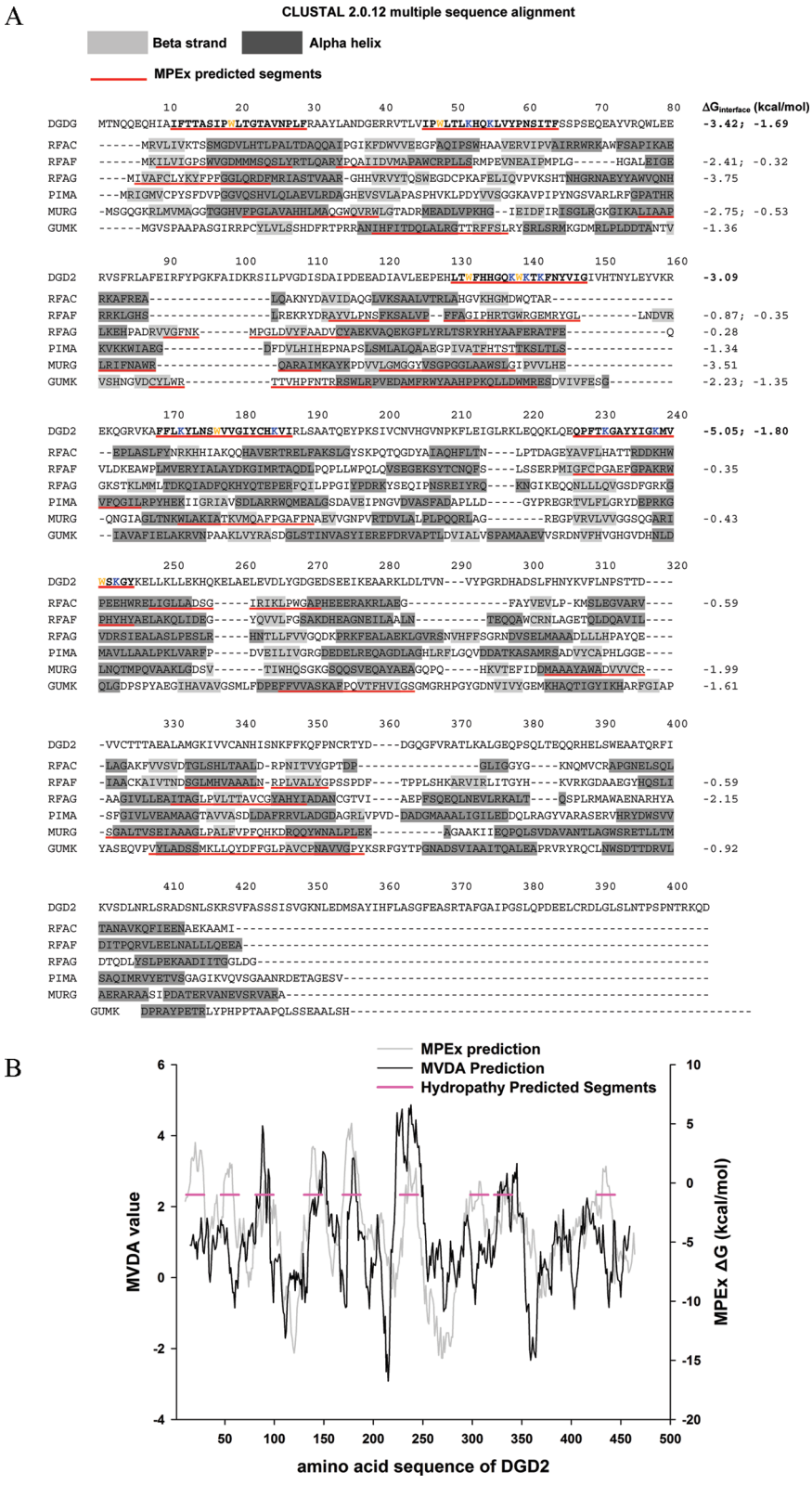


Figure 3. Predicted membrane-associated segments in atDGD2. Panel (A) shows MPex results for atDGD2 and six membrane-interacting GTs with established 3D structures. Underlined sequence segments indicate sequence with $\Delta G_{interface}$ values sufficient for interface interaction; the latter are indicated to the right of the sequences. Positively charged amino acid residues in atDGD2 are blue. Only atDGD2 sequence segments investigated in this study are visualized, and the remaining (low $\Delta G_{interface}$) are indicated in the bottom panel. Panel (B) shows the results of the analysis by using the MVDA model compared with the MPEx prediction. The gray curve indicates the ΔG value (kcal/mol) from the MPEx analysis, while the black curve shows the predicted score values for atDGD2 peptides (22 amino acid residue window length). Regions indicated by a purple line are the most likely membrane-bound segments from the MPEx analysis. Triangles define the five selected 19 amino acid-long segments based on MPEx prediction that were analyzed in the present study.

dx.doi.org/10.1021/bi200162f |*Biochemistry* 2011, 50, 4451–4466

and hence it is difficult to make quantitative estimations of the amount of structure. Nevertheless, the appearance of the spectra for these two fragments are very similar in the different solvents. S130−148 shows a slight shift in wavelength minimum in vesicles containing 10 or 20 mol% PG and is suddenly transformed into a α-helical structure when the anionic charge is increased to 30 mol% (Figure 5). The signal intensity is however quite low and the peptide appears to aggregate to some extent in all of these conditions. Hence, it is also in this case difficult to make quantitative estimates of the amount of structure.

S227−245 seems to change its structural behavior to the greatest extent among all the studied peptides. The structure of this peptide is also unaffected by the presence of zwitterionic vesicles but as the anionic content is increased the peptide transforms into an α-helix which then undergoes another structural change and ends up in a somewhat more undefined, but structured, state (Figure 5). This mixture probably contains a large portion of α-helix but also some β-strand structure. Here, the CONTIN-LL algorithm[26,27] in DichroWeb[28] was used in an attempt to quantitatively analyze the S227−245 CD data, but the relatively low signal intensities, again most likely due to solubility problems, make the numeric results relatively unreliable. The results show an increase in both α-helical (from less than 5 to around 20%) and β-strand content (from around 20 to 40%), whereas the unstructured portion of the peptide decreased (from 50 to 20%) when comparing samples with 0 mol% to 30 mol% PG. Although the numbers are lower than expected, we clearly see that this peptide undergoes a significant structural rearrangement when altering the amount of PG in the LUVs, mostly evidenced by the shift in the overall appearance of the spectra. In summary, S130−148 and S227−245 undergo structural changes when interacting with model bilayers, while S11−29 and S46−64 do not. This is in perfect agreement with the binding studies, which indicated that the latter two do not interact with model bilayers.

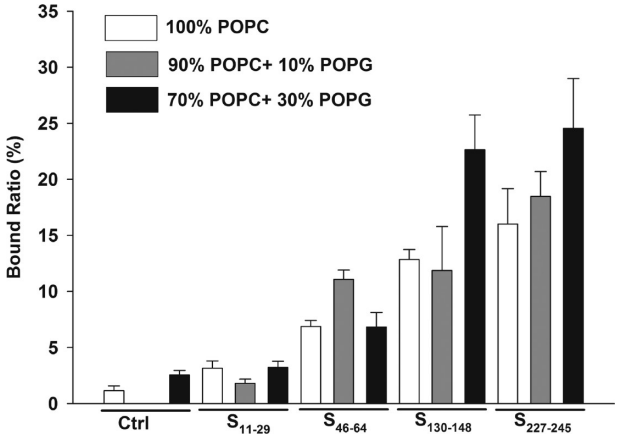


Figure 4. Binding assay of GFP-DGD2 fusion segments to vesicles. Relative binding of recombinant GFP (control), GFP-S11−29, GFP-S46−64, GFP-S130−148, and GFP-S227−245 to vesicles with varying lipid composition. POPC is the matrix lipid, which was stepwise exchanged with 10 mol% or 30 mol% POPG. The relative amount of binding is shown in arbitrary units. The error bars in the graph show the standard error of the mean ($n \geq 3$).

In Silico Studies. To investigate the propensities of the selected peptides to form amphipathic helices simple helical wheel calculations were performed at (http://cti.itc.virginia.edu/~cmg/). Only S227−245 shows signs of amphipathicity when displayed in this manner (Figure 6A). The alghoritm AmphipaSeeK (http://npsa-pbil.ibcp.fr/)[16] predicts amphipathic, so-called in-plane membrane anchors (IPM anchors). When running the complete atDGD2 sequence, there is only one region predicted as an IPM anchor (score > 0.0, Figure 6B), the sequence neighboring S227−245. We believe that this region may be important for membrane interaction due to its many cationic charges and amphipathic nature. In the modeled structure by Ge et al.,[13] this region is an α-helix. Also, the algorithm Agadir[51] was used to predict the helical content in the peptides. S227−245 is predicted by Agadir to contain 23% helix in solution, whereas the other peptides have values of 2% (S11−29), 16% (S46−64), and 5% (S130−148). Once again S227−245 clearly has different properties than the other segments, although Agadir only makes predictions in water solution and does not take into account any effects due to a lipid environment.

NMR Spectroscopy and Structure Calculation. Since all methods showed that S227−245 interacts with lipids in a surface-charge dependent manner, a solution structure of this sequence was generated from standard homonuclear ^{1}H NMR data. The vesicles used for the fluorescence and CD measurements are however too large and would result in severe line broadening of all the peaks in the NMR spectra. Therefore, the peptide was initially dissolved in a solution containing isotropic bicelles (q = 0.15, that is, the ratio of DMPC to DHPC). Also in this solvent, large line-broadening effects were observed, yielding spectra of poor quality. In our experience, this is the result of an interaction between the peptide and lipids that leads to a significant reduction of mobility.[52−54] The peptide was instead dissolved in a detergent (dodecylphosphocholine (DPC)) solution that gave much better NMR spectra. Controls were made to make sure that the spectra in the detergent solution were similar to the ones in the bicelle solution, and also that the CD spectrum for this new sample condition looked the same as the previously measured CD spectra in bicelles. Around 95% of all ^{1}H chemical

Table 2. Fluorescence Results [a]

	S11−29[b]		S46−64[b]		S130−148[b]		S227−245[b]	
solvent	λ_{max} (nm)	K_{SV} (M^{-1})	λ_{max} (nm)	K_{SV} (M^{-1})	λ_{max} (nm)	K_{SV} (M^{-1})	λ_{max} (nm)	K_{SV} (M^{-1})
buffer	354	34	356	34.2	356	18.6	356	25.1
0 mol% PG	356	35.6	355	24.8	354	20	355	21.7
10 mol% PG	356	26.7	356	26.3	352	11.2	353	14.0
20 mol% PG	354	18.4	350	14.5	346	7.0	350	11.2
30 mol% PG	354	23.0	348	8.8	344	10.3	346	8.7

[a] Quenching constants (K_{SV}) and maximum emission wavelength (λ_{max}) for the four synthetic segments. [b] The peptide/lipid ratio was in all cases 1:20.

dx.doi.org/10.1021/bi200162f |*Biochemistry* 2011, 50, 4451–4466

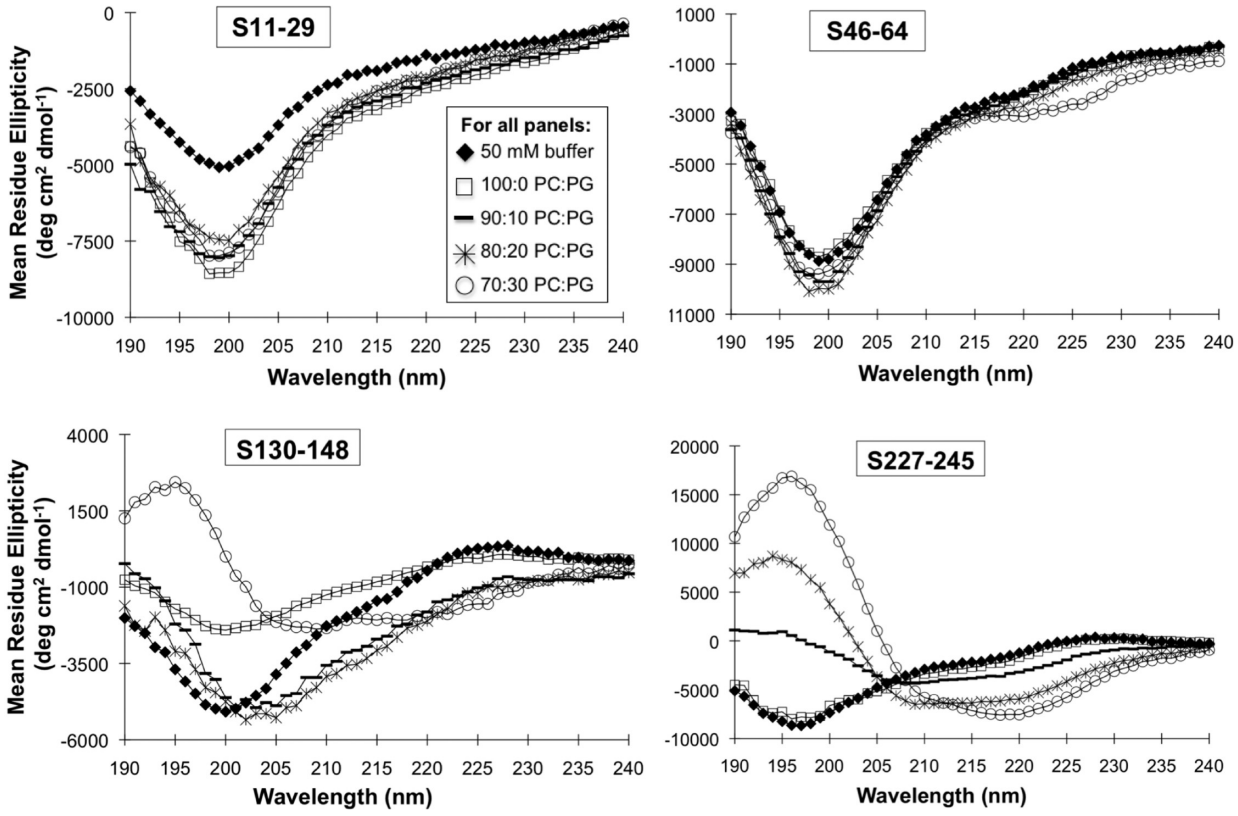


Figure 5. CD results for the synthetic sequences derived from atDGD2 corresponding to S11−29, S46−64, S130−148, and S227−245. The solvents are indicated in the inset. The temperature was in all cases 25 °C.

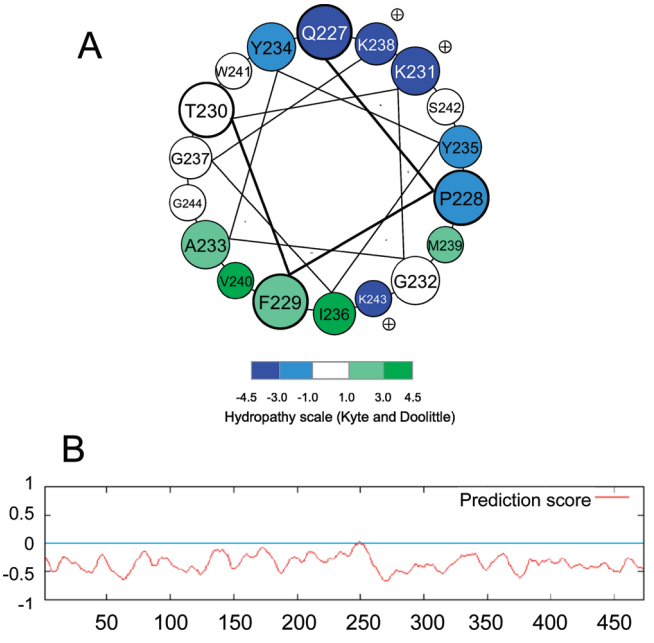


Figure 6. The S227−245 sequence as seen in an α-helical wheel representation (A). The coloring is based upon hydropathy where the most hydrophilic residues are colored blue and the positively charged residues are highlighted with ⊕ symbols. In (B) the results from the AmphipaSeek algorithm is shown.

shifts (including side-chain protons) were assigned. The calculated H^{α} secondary shifts for S227−245 (Figure 7A) clearly indicate a helical structure. This is also supported by the presence of strong sequential amide-amide cross-peaks, medium-range NOE signals (i to i+3) and the absence of long-range NOEs (Figure 7C). A region of a NOESY spectrum displaying strong H^N-H^N cross-peaks characteristic for helical structure is shown in Figure 7D.

On the basis of these NOEs, a solution structure was calculated. An ensemble of the 23 structures with the lowest energies is presented in Figure 8A. Analysis of this ensemble resulted in low target function values and small distance violations (Table 3). A PROCHECK-NMR analysis shows clear α-helical hydrogen bond patterns for amino acid residues Tyr4-Met13 in the peptide sequence, that is, Tyr230-Met239 in the full-length sequence. As expected, alignment of the backbone atoms in this region also give rise to lower root-mean-square deviation (rmsd) values than the rest of the sequence, confirming a high precision in the structures of the ensemble. Ramachandran plot statistics confirm a right-handed α-helical conformation and that all residues are located in allowed regions. Comparing the H^N chemical shifts to H^N average chemical shifts (Figure 7B) demonstrates apparent features corresponding to an amphipathic helix in which every fourth amino acid residue is potentially located in a hydrophobic environment and repeated after a set of three residues situated in a hydrophilic environment. Examining the location of side-chains in the calculated NMR structure reveals that all of the hydrophobic side-chains are located on one side of the structure, while two of the positively charged are on the other side (Figure 8B). All these structural NMR data are in good agreement with results obtained from the CD experiments.

Diffusion Experiments. NMR diffusion experiments were conducted for synthetic S227−245 in bicelle solutions for

dx.doi.org/10.1021/bi200162f |*Biochemistry* 2011, 50, 4451–4466

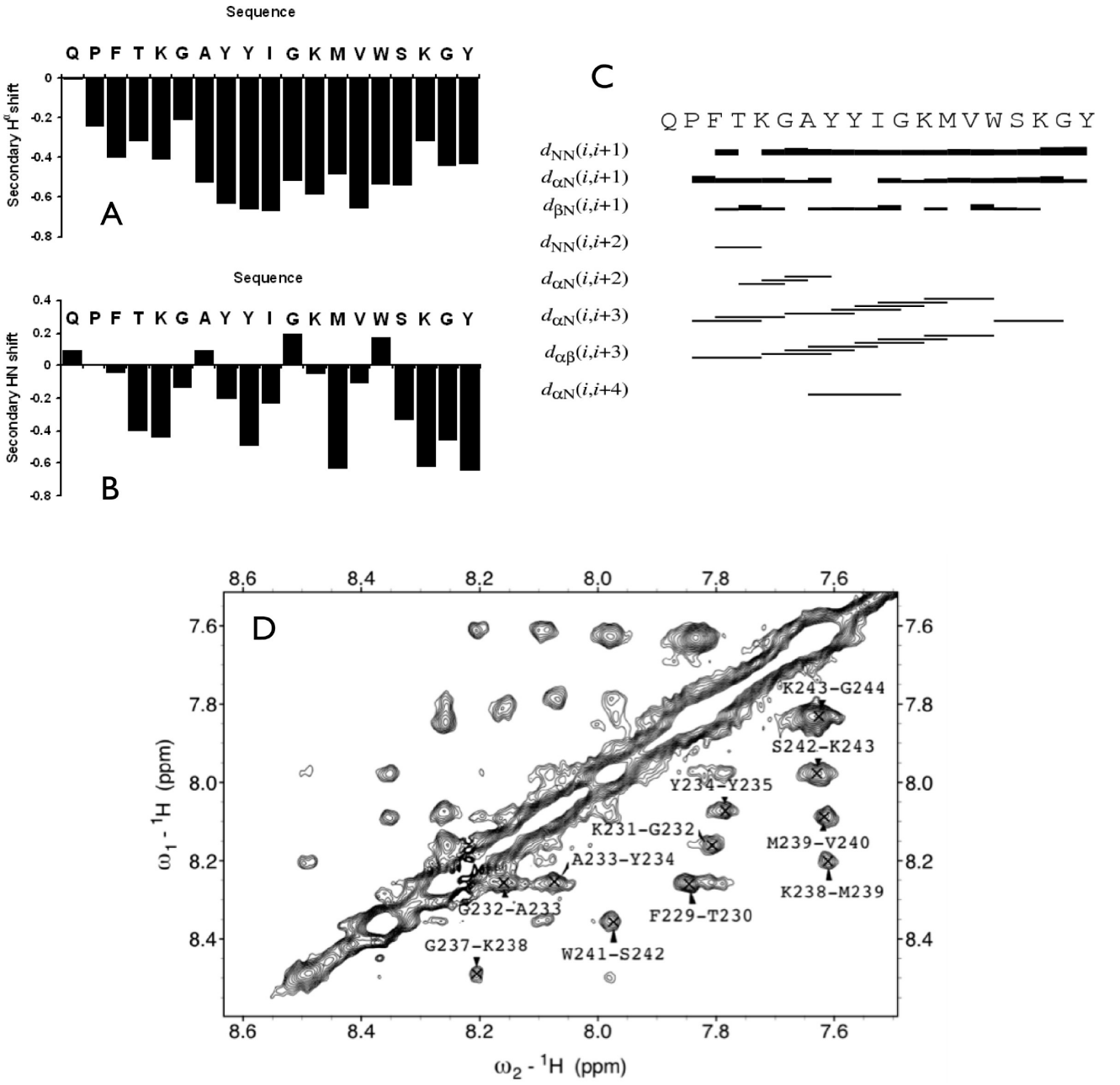


Figure 7. NMR structural data for S227−245 in DPC-d_{38}. The difference between the observed H^{α} chemical shift and random coil values are shown in (A) and the corresponding difference for H^{N} chemical shifts are shown in (B). A summary of sequential and medium-range NOE cross-peaks is shown I (C). In (D) the H^{N}-H^{N} region from the 2D NOESY with a mixing time of 300 ms is shown. The assignment for the most intense cross-peaks are indicated and labeled according to the sequence of the full-length protein.

estimating the amount of peptide that binds to phospholipid bicelles (Table 4). From the diffusion coefficient in buffer, 19.9×10^{-11} m^2 s^{-1}, and a calibrated version of Stoke−Einstein relationship,[55] we can estimate the molecular size of the peptide to be around 2300 Da, in excellent agreement with a peptide monomer (Mw = 2266 Da). For S227−245 in two different bicelle compositions (DMPC, or DMPC + DMPG), we see that the diffusion coefficient is altered and is more on the order of that of the lipid molecules. For the peptide in DMPC-containing bicelles, we observe that the diffusion coefficient of the peptide is 12% larger than that of the DMPC lipid, indicating than not all of the peptide is bound. We can estimate from eq 3 that this corresponds to around 93% of the peptide being associated with the bicelles. In PC/PG-bicelles, the diffusion of the peptide is even slower than that of the lipids, which indicates that all of the peptide is bound and that the bicelles with the peptide are of larger size than bicelles without peptide. In this case, the peptide is not only fully associated to the bicelles but appear to alter the composition to make them larger. One must keep in mind that the concentration of lipids is very high in these experiments (total PC/PG amount 100 mM corresponding to a peptide/lipid ratio 1:200) compared to experiments of LUV solutions, and therefore the amount of associated peptide may be different. Nevertheless, we conclude that in the presence of bicelles with 20 mol% PG all

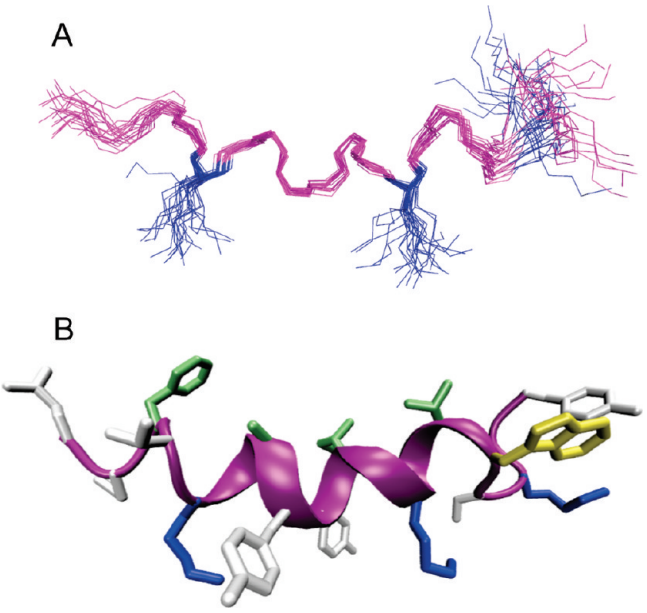


Figure 8. Solution NMR structure of the S227−245 segment in DPC micelles (A) represented by an ensemble of 23 structures, shown as lines. The structures are aligned using backbone atoms in residues 4−13. Lys residues are shown in blue. (B) One representative structure in cartoon mode showing the secondary structure of S227−245. Lys (blue), hydrophobic (green), Trp (yellow), and other (white) residues are depicted as sticks. The figures were made with (A) Pymol and (B) VMD.

Table 3. Structural Statistics for the Ensemble of 23 Structures of S227−245

no. of constraints	162
CYANA target function	0.016 ± 0.005
maximum distance violation	0.06 ± 0.006
backbone atom rmsd (Å)	
Residue 1−19	1.16
Residue 4−13	0.37
Ramachandran plot regions (%)	
most favored	78.6
allowed	17.4
generously allowed	4.0
disallowed	0.0

of S227−246 is bound, while somewhat less is bound to PC bicelles, again in agreement with the vesicle binding studies and fluorescence measurements and also with the observation that the secondary structure of the peptide depends on bilayer surface charge.

DISCUSSION

The atDGD2 gene is transcriptionally upregulated under certain conditions, such as phosphate starvation, leading to phospholipid replacement by increased synthesis of "new" GalGalDAG which partially replaces membrane phospholipids in both green and root tissues. The enzyme responds to the membrane bilayer charge properties and accordingly "adjusts" its activity. To understand which properties of the enzyme govern this interaction between the protein and the lipids, we have localized several sites (local regions) in the enzyme that appear to

Table 4. Translational Diffusion Data for the S227−245 Segment with and without the Presence of Zwitterionic or Anionic Bicelles

	D_{obs} ($\times 10^{-11}$ m^2/s)		
solvent	S227−245[a]	DMPC[b]	DHPC[b]
buffer	19.9	N/A	N/A
bicelles, zwitterionic	7.83 ± 0.16	6.93 ± 0.03	7.59 ± 0.01
bicelles, 20 mol% PG	6.21 ± 0.10	6.60 ± 0.02	7.34 ± 0.06

[a] On the basis of intensities of the aromatic peaks. [b] On the basis of the methyl peak of the aliphatic chain.

be of importance for either attaching the protein to the bilayer interface or for binding to lipid molecules in other capacities, such as substrate binding. We have investigated the bilayer interaction properties of the sequences corresponding to these sites by a variety of biophysical methods.

First, we evaluated the possible sites in the proteins most likely to be involved in lipid interactions. Comparisons with six other membrane-interacting GTs of very similar structures showed that such sites are not conserved, either in amino acid sequence or localization in the proteins (Figure 3). Therefore, we have developed a multivariate discriminant analysis method to first classify GTs as soluble or membrane-bound and then to identify membrane interacting regions in the latter proteins. By this method, we identified segments S81−99, S130−148, S169−187, and S227−245 as membrane-interacting sequences, while the sequences S11−29 and S46−64 only appeared as potential membrane-interacting sequences in the MPEx analysis (cf. Figure 3). No structure induction, and only limited lipid interaction, is observed for S11−49 or S46−64 peptide segments, which is in perfect agreement with the predictions from the MVDA analysis. Notably, these two sequences show up as potential bilayer interface-interacting sequences when predicted from hydropathy calculations (Figure 3A), which indicates that more factors need to be considered in order to correctly predict lipid interaction in GTs. Hence, we developed a computational approach that accurately predicts the regions in atDGD2, which are confirmed by biophysical methods and experiments to interact with lipids in model bilayers.

Notably, all but one of the identified sites in the protein contain Trp residues, which appear to be localized on or close to the surface of the protein. Figure 9A shows a model structure of the protein,[13] in which the Trp residues are displayed. Trp residues are known to be enriched in bilayer interface regions in many membrane proteins[15,56] and at the surface of membrane interacting proteins.[57] Previous studies of atDGD2 have shown that all of the identified Trps are important for activity but that the one in S130−148 was likely to be of particular importance for membrane anchoring.[13] Replacing Trp139 by Ala resulted in a dramatic decrease in liposome binding.[13] Mutating Trp241 in S227−245, which was also found to be a lipid-interacting segment, was shown not to affect binding, indicating that this Trp is not crucial for lipid interaction. Our data show that both of these sequences have the possibility to adopt secondary structure in the presence of phospholipids. CD data show that S130−148 may form some structure when the anionic lipid content of the vesicles is raised high enough. S227−245, on the other hand, shows even more evident structural inducement already at a low amount of charge. Since zwitterionic lipid vesicles are not enough to induce any structure in any of the peptides, we conclude that

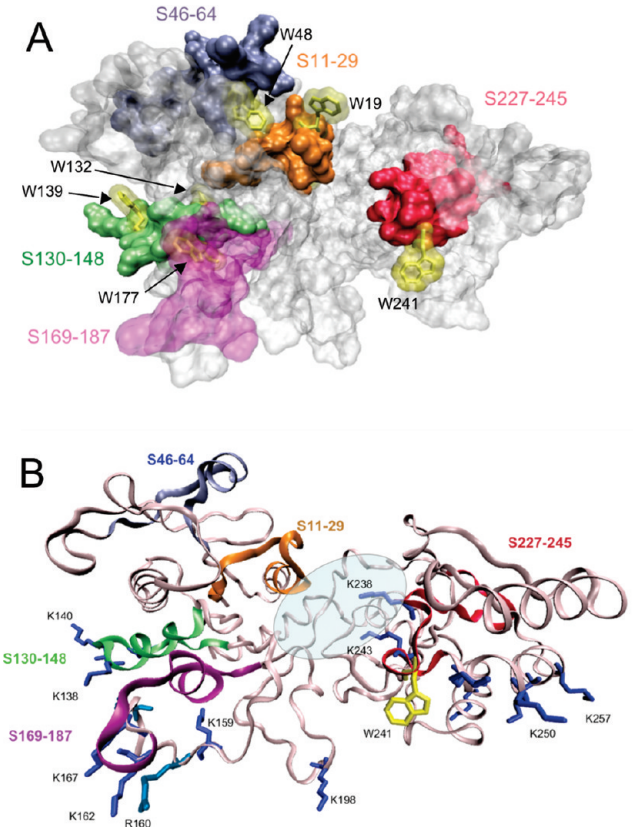


Figure 9. Model of the atDGD2 structure based on fold recognition and alignment performed with SWISS-MODEL, shown in surface mode with the tryptophans indicated in yellow (A). The regions corresponding to the sequences studied here are indicated in orange (S11−29), light blue (S46−64), green (S130−148), magenta (S169−187), and red (S227−245). (B) Ribbon-model of atDGD2. Side-chains in some of the positively charged amino acid residues are indicated by blue color. The shaded area indicates the substrate-binding region.

anionic lipid charge is a prerequisite for this to happen in both sequences.

The binding assay and fluorescence experiments give strong evidence that S130−148 and S227−245 interact with anionic lipids and that the interaction is strengthened the more anionic charge there is on the membrane surface. Diffusion NMR experiments further confirm these observations for S227−245. Lys, Phe, and Trp are the most overrepresented amino acids in peripherally membrane-interacting segments.[16] S227−245 contains one Trp, one Phe, and three Lys but no negative charges (Figure 3A) and further point at S227−245 taking part in either peripheral membrane anchoring or lipid recognition. The negative membrane surface charge of the chloroplasts,[58,59] together with our findings, point at a mostly electrostatic interaction between this segment of the protein and the lipid bilayer. It has recently been shown that Lys243 is crucial for activity and replacing this by Ala results in almost complete abolishment of activity. This is in agreement with our findings that this segment of the protein interacts with lipids in a charge-dependent way. When examining the location of this segment in the protein (Figure 9B), one clearly sees that this sequence appears to be a part of the structure that is only partially exposed. Lys243 was in a previous study selected for mutagenesis due to its closeness to the UDP-Gal substrate in the active site.[13] This indicates that it is not part of the membrane anchoring of the bulk atDGD2 protein, but more likely involved in substrate recognition and binding. Previous studies have also shown that by increasing the PG fraction in vesicles in the same manner as in the present study, increases both affinity (decrease K_M) and turnover (increase V_{max}) for the soluble UDP-Gal substrate during GalGalDAG synthesis in vitro.[13] Our results indicate that this sequence segment itself has the capability to form different secondary structure depending on the surface charge of the bilayer, from α-helical at modest PG content to more β-sheet like at higher PG content. This is a well-known feature for amphipathic sequences and indicates a chameleon-like lipid interaction, which may well be important for recognizing lipid composition in a bilayer. In order to further examine the structural propensities of this sequence, we turned to solution NMR.

The NMR measurements confirm the induced α-helical structure seen in CD spectra and predicted by in silico methods (Figures 5 and 6). The Trp residue of this segment is located in the N-terminal part of the sequence, and it is not clear what its exact orientation is. Lys231 and Lys238 are well-defined and clearly point in the opposite direction from the hydrophobic residues of the segment, as expected in an amphipathic helix. Lys243 is somewhat diffuse in orientation in comparison to the other lysine residues since the terminals of the fragment are less structured, which may indicate flexibility. As this Lys residue seems to be close to UDP-Gal,[13] this indicates that the sequence has to be able to alter its conformation and that the shown dynamic flexibility is part of the mechanism for how enzyme activity can be enhanced by anionic phospholipids, by stimulating substrate binding. On the basis of the model structure of atDGD2, this residue appears to be close in space to the proposed active site in the protein. Hence, the possibility for conformational changes in this region appears to be important. Intriguingly, a similar distribution of beta strand and α helix organization predicted for this segment is present in all the other established 3D structures of the membrane-interacting GTs in Figure 3A. Furthermore, the change in secondary structure for S227−245 upon increased anionic lipid charge is analogous to the structural changes for a peptide segment in the phosphatidylinositol-phosphate binding site in the structure of cytoskeleton protein gelsolin, upon lipid binding.[60]

Several other important features for the S227−245 segment are also evident. First, in PimA, a monotopic lipid GT in mycobacteria of double Rossmann fold, a loop in the active site next to the soluble donor substrate is linked to a membrane-binding amphipathic helix. The latter seems to influence enzyme activity in a lipid environment as indicated from mutant studies.[61] The S227−245 segment in atDGD2 is immediately followed downstream by an amphipathic helix ending at amino acid residue 260 (Figure 9B), of the highest hydrophobic moment[45] in the entire atDGD2 sequence (cf. Figure 3B legend). This is also a likely segment for interaction with anionic phospholipids due to its high content of plus-charged amino acids and potentially for affecting the K_m and V_{max} for the UDP-Gal substrate (cf. above), due to the immediate link with S227−245 making pulling/pushing movements possible. Second, in PimA, one Arg and one Lys residue interacting with its soluble substrate GDP-Mannose[58] is located in a segment corresponding to S227−245 (Figure 3A), cf. K231 and K243 in atDGD2 here. Third, the helix just next to (above) the amphipathic one in atDGD2 (i.e., 268−285; see Figure 9B) is strongly anionic and should be repelled from an anionic bilayer

interface. However, it has been shown that acidic amino acid residues become protonated by increased fractions of anionic lipids due to decreasing interfacial pH.[62] Consequently, more anionic lipids will facilitate an approach of the C-terminal domain of atDGD2 to the bilayer interface, where it can reach its acceptor substrate GalDAG, due to increased interactions with the cationic amino acid residues and decreased repulsion from the anionic ones at the atDGD2 bottom face (cf. Figure 9B). Hence, in vivo the newly transcribed/translated atDGD2 enzyme will bind more strongly to and be more active in local membrane regions with higher contents of anionic phospholipids. Because of the strongly stimulated synthesis of GalGalDAG, the increasing fraction of this lipid in the membrane will yield a decreasing fraction of membrane phospholipids (saving phosphate), and where the smaller fraction of anionic species will lower atDGD2 activity, eventually reaching a new balance corresponding to the lipid-dependent activity profile of atDGD2.

In summary, our studies show that peptide segment S227−245 has the ability to adopt an α-helical structure in the presence of lipids and that it has a quite strong interaction with the bilayer of vesicles and bicelles. The interaction is mostly mediated by positively charged amino acid residues, although the role of the tryptophan cannot be neglected since they are often found in the membrane-interacting regions of membrane-anchored proteins. We observe that this region of the protein has, under certain conditions, the propensity to form an amphiphilic helical structure. This region of the enzyme, flanking the active site, seems to potentially be part of the mechanism for stimulation of enzyme activity by anionic lipids. Among the other regions of atDGD2 studied here, we can accurately predict their lipid interacting properties, as evaluated by a range of biophysical techniques, with our MVDA algorithm. We also see that the S130−148 segment, a region that has earlier been implicated in bilayer anchoring, interacts with lipids and has the possibility to adopt secondary structure in a lipid-dependent manner.

Accession Codes

The chemical shifts for S227−245 in DPC solution have been deposited with the BMRB (www.bmrb.wisc.edu) under accession code 17356. The coordinates for the solution structure of S227−245 in DPC solution have been deposited within the PDB under accession code 2L7C (www.pdb.org).

AUTHOR INFORMATION

Corresponding Author

*E-mail: lena.maler@dbb.su.se. Phone: +46 8 162448. Fax: +46 8 155597.

Funding Sources

This work was supported by the Swedish Science Research Council, the Carl Trygger Foundation, the Magnus Bergvall Foundation, and the EU-MCRTN Biocontrol.

ABBREVIATIONS USED

GalDAG, monogalactosyl-diacylglycerol (1,2-diacyl-3-*O*-(ss-D-galactopyranosyl)-*sn*-glycerol); GalGalDAG, digalactosyl-diacylglycerol (1,2-diacyl-3-*O*-[α-D-galactopyranosyl-(1→6)-*O*-β-D-galactopyranosyl]-*sn*-glycerol); GT, glycosyltransferase; MGD, monoglycosyldiacylglycerol synthase; DGD, diglycosyldiacylglycerol synthase; atDGD1, digalactosyldiacylglycerol synthase1 from *Arabidopsis thaliana*; atDGD2, digalactosyldiacylglycerol synthase2 from *Arabidopsis thaliana*; PG, phosphatidylglycerol; GFP, green fluorescent protein; PC, phosphatidylcholine; NMR, nuclear magnetic resonance; POPC, 1-palmitoyl-2-oleoyl-phosphatidylcholine; POPG, 1-palmitoyl-2-oleoyl-phospatidylglycerol; DLS, dynamic light scattering; CD, circular dichroism; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DPC, *n*-dodecyl phosphocholine; LUV, large unilamellar vesicle; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy

REFERENCES

(1) Siebertz, H. P., Heinz, E., Linscheid, M., Joyard, J., and Douce, R. (1979) Characterization of lipids from chloroplast envelopes. *Eur. J. Biochem. 101*, 429–438.

(2) Dörmann, P., and Benning, C. (2002) Galactolipids rule in seed plants. *Trends Plant Sci. 7*, 112–118.

(3) Vikström, S., Li, L., Karlsson, O. P., and Wieslander, Å. (1999) Key role of the diglucosyldiacylglycerol synthase for the nonbilayer-bilayer lipid balance of *Acholeplasma laidlawii* membranes. *Biochemistry 38*, 5511–5520.

(4) Andersson, M. X., Stridh, M. H., Larsson, K. E., Liljenberg, C., and Sandelius, A. S. (2003) Phosphate-deficient oat replaces a major portion of the plasma membrane phospholipids with the galactolipid digalactosyldiacylglycerol. *FEBS Lett. 537*, 128–132.

(5) Gigon, A., Matos, A. R., Laffray, D., Zuily-Fodil, Y., and Pham-Thi, A. T. (2004) Effect of drought stress on lipid metabolism in the leaves of *Arabidopsis thaliana* (Ecotype Columbia). *Ann. Bot. 94*, 345–351.

(6) Kelly, A. A., and Dörmann, P. (2002) DGD2, an Arabidopsis gene encoding a UDP-galactose-dependent digalactosyldiacylglycerol synthase is expressed during growth under phosphate-limiting conditions. *J. Biol. Chem. 277*, 1166–1173.

(7) Hartel, H., Dörmann, P., and Benning, C. (2000) DGD1-independent biosynthesis of extraplastidic galactolipids after phosphate deprivation in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A. 97*, 10649–10654.

(8) Tjellström, H., Hellgren, L. I., Wieslander, Å., and Sandelius, A. S. (2010) Lipid asymmetry in plant plasma membranes: phosphate deficiency-induced phospholipid replacement is restricted to the cytosolic leaflet. *FASEB J. 24*, 1128–1138.

(9) Dörmann, P., Balbo, I., and Benning, C. (1999) Arabidopsis galactolipid biosynthesis and lipid trafficking mediated by DGD1. *Science 284*, 2181–2184.

(10) Botte, C., Jeanneau, C., Snajdrova, L., Bastien, O., Imberty, A., Breton, C., and Marechal, E. (2005) Molecular modeling and site-directed mutagenesis of plant chloroplast monogalactosyldiacylglycerol synthase reveal critical residues for activity. *J. Biol. Chem. 280*, 34691–34701.

(11) Holzl, G., Leipelt, M., Ott, C., Zahringer, U., Lindner, B., Warnecke, D., and Heinz, E. (2005) Processive lipid galactosyl/glucosyltransferases from *Agrobacterium tumefaciens* and *Mesorhizobium Loti* display multiple specificities. *Glycobiology 15*, 874–886.

(12) Lind, J., Ramo, T., Klement, M. L., Barany-Wallje, E., Epand, R. M., Epand, R. F., Mäler, L., and Wieslander, Å. (2007) High cationic charge and bilayer interface-binding helices in a regulatory lipid glycosyltransferase. *Biochemistry 46*, 5664–5677.

(13) Ge, C., Georgiev, A., Öhman, A., Wieslander, Å., and Kelly, A. A. (2011) Tryptophan residues promote membrane association for a plant lipid glycosyltransferase involved in phosphate stress. *J. Biol. Chem. 286*, 6669–6684.

(14) Yau, W. M., Wimley, W. C., Gawrisch, K., and White, S. H. (1998) The preference of tryptophan for membrane interfaces. *Biochemistry 37*, 14713–14718.

(15) Granseth, E., von Heijne, G., and Elofsson, A. (2005) A study of the membrane-water interface region of membrane proteins. *J. Mol. Biol. 346*, 377–385.

(16) Sapay, N., Guermeur, Y., and Deleage, G. (2006) Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. *BMC Bioinf. 7*, 255.

(17) Kelly, A. A., Öhman, A., Sedoud, A., and Wieslander, Å. (2007) Membrane homeostasis mechanisms for plant glycosyl-transferases involved in galactolipid biosynthesis: biochemical and in-silico structure studies of key enzymes, in *Current Advances in the Biochemistry and Cell Biology of Plant Lipids* (Benning, C., and Olhlrogge, J., Eds.) 1st ed., Aardvark Global Publishing Company, LLC, Salt Lake City, UT.

(18) Nyholm, T. K., Ozdirekcan, S., and Killian, J. A. (2007) How protein transmembrane segments sense the lipid environment. *Biochemistry 46*, 1457–1465.

(19) Rath, A., Tulumello, D. V., and Deber, C. M. (2009) Peptide models of membrane protein folding. *Biochemistry 48*, 3036–3045.

(20) Sharpe, H. J., Stevens, T. J., and Munro, S. (2010) A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell 142*, 158–169.

(21) Snider, C., Jayasinghe, S., Hristova, K., and White, S. H. (2009) MPEx: A tool for exploring membrane proteins. *Protein Sci. 18*, 2624–2628.

(22) Mäler, L., and Gräslund, A. (2009) Artificial membrane models for the study of macromolecular delivery. *Methods Mol. Biol. 480*, 129–139.

(23) Buser, C. A., and McLaughlin, S. (1998) Ultracentrifugation technique for measuring the binding of peptides and proteins to sucrose-loaded phospholipid vesicles. *Methods Mol. Biol. 84*, 267–281.

(24) Eftink, M. R., and Ghiron, C. A. (1981) Fluorescence quenching studies with proteins. *Anal. Biochem. 114*, 199–227.

(25) Lakowicz, J. R., and Weber, G. (1973) Quenching of fluorescence by oxygen. Probe for structural fluctuations in macromolecules. *Biochemistry 12*, 4161–4170.

(26) Provencher, S. W., and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry 20*, 33–37.

(27) van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R., and Groen, F. C. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal. Biochem. 191*, 110–118.

(28) Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers 89*, 392–400.

(29) Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson. 53*, 521–528.

(30) Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun. 95*, 1–6.

(31) Goddard, T. D., and Kneller, D. G. *Sparky*, University of California, San Francisco, USA.

(32) Wu, D. H., Chen, A. D., and Johnson, C. S. (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J. Magn. Reson., Ser. A. 115*, 260–264.

(33) Stejskal, E. O., and Tanner, J. E. (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys. 42*, 288–292.

(34) Damberg, P., Jarvet, J., and Gräslund, A. (2001) Accurate measurement of translational diffusion coefficients: a practical method to account for nonlinear gradients. *J. Magn. Reson. 148*, 343–348.

(35) Andersson, A., Almqvist, J., Hagn, F., and Mäler, L. (2004) Diffusion and dynamics of penetratin in different membrane mimicking media. *Biochim. Biophys. Acta 1661*, 18–25.

(36) Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol. 278*, 353–378.

(37) Güntert, P., Braun, W., and Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol. 217*, 517–530.

(38) Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The Chemical Shift Index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry 31*, 1647–1651.

(39) Laskowski, R. A., Rullmann, A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477.

(40) Jaysinghe, S., Hristova, K., Wimley, W., Snider, C., and White, S. H. (2006) *Membrane Protein Explorer (MPEx)*, http://blanco.biomol.uci.edu/mpex.

(41) Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol. 3*, 842–848.

(42) Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol. 300*, 1005–1016.

(43) Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng. 10*, 1–6.

(44) Sharikov, Y., Walker, R. C., Greenberg, J., Kouznetsova, V., Nigam, S. K., Miller, M. A., Masliah, E., and Tsigelny, I. F. (2008) MAPAS: A tool for predicting membrane-contacting protein surfaces. *Nat. Methods 5*, 119.

(45) Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol. 179*, 125–142.

(46) Wold, S.; Eriksson, L. and Sjöström, M. Partial least squares projections to latent structures (PLS) in chemistry. In *The Encyclopedia of Computational Chemistry* (Schleyer, P., Allinger, N. L., Clark, T., Dasteiger, J., Kollman, P. A., Schaefer , H. F., III, Eds.) John Wiley & Sons: Chichester, 1998; pp 2006−2021.

(47) Hellberg, S., Sjöstrom, M., Skagerberg, B., and Wold, S. (1987) Peptide quantitative structure-activity relationships, a multivariate approach. *J. Med. Chem. 30*, 1126–1135.

(48) Wold, S., Jonsson, M., Sjöström, M., Sandberg, M., and Rännar, S. (1993) DNA and peptide sequences and chemical processes multivariately modelled by principle components analysis and partial least squares projections to latent structures. *Anal. Chim. Acta 277*, 239–253.

(49) Eriksson, L., Johansson, E., Kettaneh-Wold, N., and Wold, S. (2001) Multi- and megavariate data analysis principles and applications. Umetrics, Umeå.

(50) Rosén, M. L., Edman, M., Sjöström, M., and Wieslander, Å. (2004) Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *J. Biol. Chem. 279*, 38683–38692.

(51) Lacroix, E., Viguera, A. R., and Serrano, L. (1998) Elucidating the folding problem of α-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol. 284*, 173–191.

(52) Barany-Wallje, E., Andersson, A., Gräslund, A., and Mäler, L. (2006) Dynamics of transportan in bicelles is surface charge dependent. *J. Biomol. NMR 35* 137–147.

(53) Lind, J., Gräslund, A., and Mäler, L. (2006) Membrane interactions of dynorphins. *Biochemistry 45*, 15931–15940.

(54) Biverståhl, H., Andersson, A., Gräslund, A., and Mäler, L. (2004) NMR solution structure and membrane interaction of the N-terminal sequence (1−30) of the bovine prion protein. *Biochemistry 43*, 14940–14947.

(55) Danielsson, J., Jarvet, J., Damberg, P., and Gräslund, A. (2002) Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1−40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length. *Magn. Reson. Chem. 40*, S89–S97.

(56) Liu, W., and Caffrey, M. (2006) Interactions of tryptophan, tryptophan peptides, and tryptophan alkyl esters at curved membrane interfaces. *Biochemistry 45*, 11713–11726.

(57) Bhardwaj, N., Stahelin, R. V., Langlois, R. E., Cho, W., and Lu, H. (2006) Structural bioinformatics prediction of membrane-binding proteins. *J. Mol. Biol. 359*, 486–495.

(58) Dobrikova, A., Taneva, S. G., Busheva, M., Apostolova, E., and Petkanchin, I. (1997) Surface electric properties of thylakoid membranes from *Arabidopsis thaliana* mutants. *Biophys. Chem. 67*, 239–244.

(59) Block, M. A., Dorne, A. J., Joyard, J., and Douce, R. (1983) Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts. II. Biochemical characterization. *J. Biol. Chem. 258*, 13281–13286.

(60) Xian, W., and Janmey, P. A. (2002) Dissecting the gelsolin−polyphosphoinositide interaction and engineering of a polyphosphoinositide-sensitive gelsolin C-terminal half protein. *J. Mol. Biol. 322*, 755–771.

(61) Guerin, M. E., Schaeffer, F., Chaffotte, A., Gest, P., Giganti, D., Kordulakova, J., van der Woerd, M., Jackson, M., and Alzari, P. M. (2009) Substrate-induced conformational changes in the essential peripheral membrane-associated mannosyltransferase PimA from mycobacteria: implications for catalysis. *J. Biol. Chem. 284*, 21613–21625.

(62) Johnson, J. E., Xie, M., Singh, L. M., Edge, R., and Cornell, R. B. (2003) Both acidic and basic amino acids in an amphitropic enzyme, CTP:phosphocholine cytidylyltransferase, dictate its selectivity for anionic membranes. *J. Biol. Chem. 278*, 514–522.