

Impact of novel detergents on membrane protein structural studies

Abstract

Detergents are widely used for membrane protein solubilization, purification and structural analysis. Conventional detergents are limited in their ability to stabilize membrane proteins in aqueous solution, and thus over the past two decades much effort has been made to design novel detergents with improved properties. Here we provide details of 23 different classes of novel detergents with particular focus on OGNG, LMNG and GDN, which have been widely applied to the successful determination of a wide range of membrane protein structures. These three detergents have distinct properties and this is reflected in their preferential use in different structural approaches; *in surfo* crystallization (OGNG), lipid cubic phase (LCP) crystallization (LMNG) and cryo-EM (LMNG and GDN). This review provides insights into the biophysical basis of the different behaviors of these detergents by analysis of the stability of protein-detergent complexes (PDCs) in terms of *detergent-detergent* and *detergent-protein* interactions. The review also [discusses recently developed variants of OGNG, LMNG and GDN as well as several other classes of amphiphiles in the context of these interactions](#). Finally, a detailed perspective of new detergent design is provided as a basis for expanding the applicability of small detergents in membrane protein structural study.

1. Introduction

Membrane proteins, encoded by approximately 25% of the human genome, perform a variety of vital functions such as material transport, cell-cell recognition, signal transduction, and cell adhesion and mobility.[#] Membrane proteins are implicated in various diseases including cancers and viral infections, and are associated with neurodegenerative diseases such as Alzheimer and Parkinson diseases. As a result, more than 50% of pharmaceuticals target these membrane-associated bio-macromolecules. High resolution membrane protein structures are key to a detailed understanding of their mechanism of action and successful drug discovery and development. However, as of July 31, 2021, ~1300 unique membrane protein structures have been determined (Fig. 1), a tiny proportion of the total structures deposited in the Protein Data Bank (PDB), highlighting the challenges of membrane protein structural study. X-ray crystallography, single particle cryo-electron microscopy (cryo-EM) and nuclear magnetic resonance (NMR) spectroscopy remain the principle techniques for protein structure determination (Fig 2). X-ray crystallography has been most widely used for this purpose since the structural characterization of myoglobin in 1958. NMR spectroscopy for protein structure determination dramatically increased in 1990s, but no further increase has been observed since then. Recent progress in cryo-EM technology has witnessed a notable increase in use of this method for protein structure elucidation and this is particularly true for membrane proteins. Cryo-EM is responsible for 78% of the 99 G protein-coupled

receptor (GPCR) structures determined this year (Jan-Jul 2021).

Membrane proteins have a high propensity to aggregate and/or denature in aqueous solution. Their dual hydrophobic-hydrophilic domains mean that it is difficult to maintain them in a soluble and stable state outside biological membranes. It is necessary to both extract the membrane protein from the lipid bilayer and replace it with another amphiphilic system for membrane protein structural study. Diverse amphiphilic systems have been developed to provide a suitable environment for membrane protein study. The conventional system is a micellar assembly formed by small amphiphilic molecules containing hydrophilic head and hydrophobic tail groups. These amphiphilic molecules, named detergents/surfactants, are not only able to dismantle cell membranes, but can self-assemble around the hydrophobic domains of membrane proteins. Thus, detergents are widely used to extract membrane proteins from the native membranes and to stabilize the resulting protein-detergent complexes (PDCs) in aqueous solution. Conventional detergents with a canonical structure such as DDM, OG, LDAO and C₈E₄ are seriously limited in their ability to stabilize membrane proteins, particularly eukaryotic proteins and large membrane protein complexes (Fig. 3). Over the past two decades, other amphiphilic systems have been invented, including amphiphilic polymers (e.g., amphipols (APols) and styrene-maleic acid copolymer (SMA)) and amphiphilic peptides/proteins (e.g., membrane scaffold proteins (MSPs) and saposin A). The polymeric materials (i.e., SMA, MSP and saposin A) are capable of forming disc-shaped nano-assemblies which incorporate membrane proteins surrounded by a patch of lipid bilayer, leading to styrene-maleic acid lipid particles (SMALPs), MSP-based nanodiscs (NDs) and saposin-lipid-protein complexes (Salipro), respectively (Fig. 4). Some special detergents such as 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPS) and 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) are known to form bicelles when mixed with lipid vesicles (e.g., 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC)) in specific ratios (Figs. 3 and 4). These bicelles thus are similar to MSP-based NDs, SMALP, and Salipro in that they accommodate a patch of membranes in the interior. In contrast, both detergents and Apols tend to directly interact with membrane protein surfaces without incorporation of a lipid patch, although small numbers of relatively tightly bound membrane lipids can remain associated through detergent extraction and isolation and if appropriate exchange into Apols. Due to their small size, typically 50 to 200 detergent molecules are required to associate with a membrane protein, while a single or a small number of Apols is sufficient to cover the whole hydrophobic surface of a membrane protein.

Amphiphilic systems incorporating a patch of lipid bilayers in the interior, as exemplified by bicelles, MSP-based NDs, and SMALP, are generally superior to detergent micelles in preserving the native conformation of membrane proteins. However, these nano-assemblies have limited applications for membrane proteins study. For example, although being widely applicable for protein crystallization, bicelles are not suitable for protein extraction. The MSP-based NDs and Salipro were shown to be useful

for cryo-EM structural study, but these are also unsuitable for protein extraction and have little application for protein crystallization. In contrast, detergent micelles can be used in all steps, encompassing protein solubilization, purification and structure determination by crystallization, cryo-EM and NMR spectroscopy. Furthermore, proteins need to be extracted and isolated first in detergents, prior to reconstitution into MSP-based NDs, Salipro and bicelle systems. As a result, despite recent breakthroughs in the development of new amphiphilic systems, detergent micelles are still the most widely used membrane mimetic.

Detergent properties and efficacies can be optimized *via* relatively simple structural modification based on detergent design principles. Such optimization is significantly more challenging for polymer/protein-based amphiphilic systems. Thus, since the first structural determination of the photosynthetic reaction centre of *Rhodospseudomonas viridis* using LDAO in 1985, much effort has been focused on the development of novel detergents. Examples include hemi-fluorinated surfactants (HFSs), facial amphiphiles (FAs), neopentyl glycol (NG) amphiphiles [glucose neopentyl glycols (GNGs) and maltose neopentyl glycols (MNGs)], and rigid hydrophobic-bearing detergents [e.g., chobimalt, glyco-diosgenin (GDN)] (Fig. 5). Some recently described detergents contain a rigid core in the central region, as in maltose-bis(hydroxymethyl)phenol amphiphiles (MBPs) [M. Ehsan, *ACS Chem Bio*, 2021] and 1,3,5-triazine-based maltosides (TEMs). New detergents with non-canonical hydrophilic and hydrophobic group have also been developed, as exemplified with penta-saccharide amphiphiles (PSEs), dendronic trimaltoside amphiphiles (DTMs), vitamin E-derived glycosides (VEGs) and oligoglycerol detergents (OGDs). Of the novel classes of detergents covered here, the NG and GDN class detergents are particularly notable as these have been widely and highly successfully adopted as new agents for membrane protein structural studies. Thus, this review starts with detailed descriptions of the molecular design and self-assembly behaviors of amphiphiles and their membrane protein applications, particularly focusing on OGNG, LMNG (commercial names of GNG-3 and MNG-3, respectively), and GDN. In addition, the current review includes upgraded versions of NG/GDN class that are better than the original detergents for protein stability. The review also covers several structurally distinct amphiphiles that hold significant potential for future membrane protein structural study. Each set of detergents is discussed in terms of their ability to form *detergent-detergent* or *detergent-protein* interactions, key to protein stability.

2. Detergent structures and properties of NG/GDN class

2.1 Detergent structures and synthetic accessibility

The NG class amphiphiles commonly have two alkyl chains and two carbohydrate head groups connected with each other *via* an NG unit in the central region. These amphiphiles are categorized into two classes, GNGs and MNGs, with glucoside or maltoside head groups, respectively. The resulting

NG detergents have a dimeric architecture compared to conventional detergents. For example, two representative NG amphiphiles, octyl glucose neopentyl glycol (OGNG) and lauryl glucose neopentyl glycol (LMNG), are dimers of OG and DDM, respectively (Figure 5). One important feature of the NG detergents is the presence of a quaternary carbon between the head and tail groups, central to the chemical linkage of the head and tail groups (Figure 5c,d). This rigidifying carbon center at the hydrophilic-hydrophobic interface likely affects detergent self-assembly behaviors and detergent efficacy for membrane protein stabilization. The NG class amphiphiles can be prepared using diethylmalonate as a starting material *via* a protocol comprising four synthetic steps: dialkylation, LiAlH_4 reduction, β -selective glycosylation and deprotection. Overall synthetic yields are in the range of 70 to 80%. The short synthetic protocol and high overall yield indicate ready accessibility to this class, facilitating widespread use in membrane protein research. Dozens of NG detergents are commercially available from Anatrace including OGNG and LMNG.

GDN contains a steroid unit, called diosgenin, and branched dimaltoside as the hydrophobic and hydrophilic groups, respectively (Figure 5f). This detergent is a synthetic version of digitonin, a steroid saponin obtained from the purple foxglove plant *Digitalis purpurea*. GDN and digitonin differ in the absence/presence of substituents on the steroid ring and carbohydrate structure of the head group. The head group of digitonin comprises two galactoses, two glucoses and a xylose, while GDN contains a branched maltoside head group (Figs. 3d,5f). The steroid hydrophobic groups of GDN and digitonin are distinct from typical conventional detergents and NG class detergents. Due to this unique structural feature of their hydrophobic group, GDN and digitonin are often used in the study of particularly challenging membrane proteins that are unstable in other classes of detergents (*vide infra*). Merits of GDN compared to digitonin include high homogeneity, no batch-to-batch variation, good water-solubility (>10%) and low toxicity. However, compared to the NG class detergents, GDN is synthetically less accessible; a seven-step synthetic protocol is required for preparation of this agent, with an overall yield of ~35%. Thus, synthetic variants with enhanced accessibility, attained by structural modification, would further increase the utility of this class for membrane protein structural study.

2.1 Detergent self-assembly property

Two NG detergents, LMNG and OGNG, have been best characterized in terms of their self-association behavior in aqueous solution and thus our discussion focuses on these two detergents. Micelles formed by OGNG and LMNG are substantially different from those formed by the conventional analogs (DDM and OG, respectively). First, OGNG and LMNG have critical micelle concentrations (CMCs) significantly lower than OG and DDM, respectively, mainly due to the presence of two alkyl chains. OGNG and LMNG have CMCs of 1.0 and 0.01 mM, 25 and 170 times smaller than those of OG (25 mM) and DDM (0.17 mM), respectively, as determined by dye absorption method (Table 1). A similar

trend was found in the comparison of the CMC of DMNG (~0.036 mM) and its conventional equivalent (DM; ~1.8 mM). The reduced CMCs are a natural consequence of the increased hydrophobicity of their detergent lipophilic groups. In the case of LMNG, a lower CMC of 0.04 μ M obtained by analytical ultracentrifugation (AUC), has been reported.

OGNG and LMNG also form much larger micelles to their conventional equivalents, as estimated by the hydrodynamic radius (R_h). When measured at a detergent concentration of 1.0 wt%, OGNG forms micelles with a R_h of 4.4 nm, while OG micelles have smaller sizes in the range of 1.5 to 2.3 nm (Table 1). Large micelles are also observed for LMNG; the R_h of LMNG micelles was 10.6 nm at 1.0 wt%, substantially larger than DDM micelles with a R_h of 3.5 nm at the same conditions. The large aggregates formed by OGNG and LMNG imply that their self-assembly architectures differ from micellar aggregates formed by DDM and OG. Breyton *et al* investigated the self-assembly structure of LMNG using small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS). This study indicates that LMNG forms rod-shaped self-assemblies with a diameter of 2.1 nm, markedly different from the spherical or ellipsoidal micelles formed by DDM. In addition, consistent with our results, the self-assembly size of LMNG increases with increasing detergent concentration, while the assembly size of DDM varies little with detergent concentration (Fig. 6). Increasing LMNG aggregate size with increasing detergent concentration can be explained by the fact that additional LMNG molecules added into solution incorporate into both ends of the rod-shaped assemblies in preference to creating additional self-assemblies. The difference between the self-assembly architectures of LMNG and DDM (rod-shaped vs spherical) can be explained by the critical packing parameter - the relative area occupied by the detergent head and tail group. As the quaternary carbon center of LMNG is located close to the detergent head group, the mean area of the hydrophobic group is proportionally greater than that of the branched dimaltoside head group in comparison to the equivalent group in DDM, resulting in a higher packing parameter and a rod shaped self-assembly. OGNG also showed an increase in micelle size with increasing detergent concentration. The R_h of OGNG micelles steadily increases from 3.5 to 4.4 nm when OGNG concentration is increased from 0.4 to 1.0 wt%, indicating that this detergent also forms rod-shaped self-assemblies (Fig. 6c). The smaller size of the OGNG self-assemblies can be partly explained by the presence of the branched diglucoside head group, smaller than the branched dimaltoside in LMNG.

GDN and digitonin have similar lipophilic groups, but the former has a CMC approximately ten times lower than the latter (0.018 vs 0.25-0.50 mM), likely due to the better segregation of the hydrophobic and hydrophilic groups for GDN (Table 1). Of note, digitonin contains two hydroxyl groups in the lipophilic regions, which reduces the hydrophobicity of this natural detergent. The hydroxyl group at position C15 is likely to have a significant contribution to the higher CMC of digitonin as it would substantially reduce the hydrophobic interactions between the detergent molecules in aqueous solution.

GDN forms micelles with a R_h of 4.4 nm at 1.0 wt% (Table 1). The sizes of GDN self-assemblies vary little with changing detergent concentration, a similar behavior observed for DDM and digitonin (Fig. 6). Thus this detergent and digitonin probably form spherical or ellipsoidal micelles like DDM (Fig. 6d). The overall larger micelles formed by GDN compared to DDM (4.4 vs 3.5 nm) is likely the result of the large hydrophobic group, resulting in an increased packing parameter.

3. Use of OGNG, LMNG and GDN in membrane protein study

3.1 GNGs (OGNG)

OGNG (GNG-3) is efficient at extracting membrane proteins from the membranes. This detergent yielded a higher solubilization efficiency for the light harvest complex I-reaction centre (LHI-RC) complexes of *Rhodobacter capsulatus* than DDM (90 vs 80%). OGNG was as good as DDM for stabilizing LHI-RC complexes, but less effective at stabilizing *Arabidopsis thaliana* boron transporter 1 (AtBOR1), *Aquifex aeolicus* leucine transporter (LeuT), *Salmonella typhimurium* melibiose permease (MelB_{Si}), and human beta2 adrenergic receptor (β_2 AR). Despite this limitation, however, this glucoside has found its use in membrane protein structural study. As of 07.31.2021, 3D structures of a total of 15 membrane proteins have been obtained using OGNG based on the information obtained from an online data bank “membrane proteins of known 3D structure” (Table S1). Of note, this number includes multiple structures of the same membrane protein. Due to the high solubilisation efficiency OGNG has often been used for protein extraction. Of the 15 membrane protein structures obtained using OGNG, nine of the proteins were extracted, purified and structurally determined *via* crystallization using this glucoside (Figure 7a). This data suggests that OGNG can be used for every step of protein structural characterization: membrane protein solubilisation, purification and crystallization.

As indicated by the data analysis, this glucoside has contributed to the high resolution structure determination of relatively stable membrane proteins such as channels (9) and enzymes (3) (Figure 7b). Channel structures obtained using OGNG are categorized into three classes: ion channels [(two-pore domain potassium channel K_{2p}2.1 (TREK-1; 6CQ6) and trimeric intracellular cation (TRIC) channel (6IYU)], transport receptor potential (TRP) channels [TRPV2 transient receptor potential channel (6BWJ)] and aquaporins [AQP2 and AQP7 (4NEF)]. In addition, OGNG was used to the high resolution crystal structures of three enzymes: CAAX protease ZMPSTE24 (4AW6), Na⁺-translocating pyrophosphatase (PPase) (4AV3), and superoxide oxidase (5OC0). Notably, the high resolution structures of all these proteins were determined *via* vapor diffusion method, with the single exception of sugar transporter protein 10 (6H7D) determined *via* lipid cubic phase (LCP) crystallization (Figure 7c). This data suggests that OGNG is suitable for *in surfo* protein crystallization, likely due to its small PDC formation. The small PDC formation of OGNG relative to DDM was well illustrated by SEC profiles of a nuclear membrane zinc metalloprotease (ZMPSTE24) in these detergents. At first glance,

the data obtained appears to be inconsistent with the larger micelles formed by OGNG than DDM. However, it is important to note that micelle sizes formed by detergents do not necessarily correlate with their PDC sizes, although this is typical for conventional detergents with single head and tail groups. In the presence of membrane proteins, detergent molecules associate with the hydrophobic protein surfaces and thus detergent geometry, a main determinant for detergent micelle size, is not a major factor in PDC size. Rather, the length and conformation of the detergent head and tail groups can play dominant roles in the PDC size. Consequently, OGNG can form smaller PDCs than DDM, despite forming larger micelles formation.

3.2 MNGs (LMNG, DMNG, and Cymal-5-NG)

3.2.1 LMNG interactions with membrane proteins

Due to the large micelles formed by LMNG compared to DDM, there has been a question about the heterogeneity of PDCs formed by this detergent, originating from various amount of bound detergents. This question was specifically addressed by studying a few membrane proteins solubilized in LMNG. Size exclusion chromatography-light scattering (SEC-LS) analysis of *E. coli* outer membrane ferrichrome-iron transporter (FhuA)/LMNG complexes indicates that LMNG is bound to this membrane protein at 1.2 g detergent per g of protein, similar to that observed for DDM. A similar result was obtained with *Bacillus* multidrug resistance ATP (BmrA)/LMNG complexes investigated by MALDI-TOF. According to detergent belt modeling that assumes cylindrical micelle formation surrounding a given membrane protein, the BmrA/detergent complexes are calculated to have detergent belts with radii of 28 and 24 Å for DDM and LMNG, respectively. It is notable that LMNG, like OGNG as discussed above, is predicted to form smaller PDCs than DDM despite the tendency to form substantially larger micelles. Indeed, BmrA/LMNG study showed that the number of LMNG molecules bound to the protein is less than twice that of DDM, when the PDCs were purified by immobilized metal affinity chromatography (IMAC) (328 vs 138) or SEC (399 vs 167). More importantly, the PDCs formed by LMNG with the four membrane proteins [BmrA, *S. pneumoniae* NOX protein (SpNOX), FhuA, and bacteriorhodopsin from *Halobacterium* (bR)] remain homogeneous over the course of protein purification. This result is interesting because different trends were observed in previous studies. A hemi-fluorinated surfactant (HFS) with three glucose units forms homogeneous micelles and generates homogeneous complexes with bR and cytochrome *b_of* complexes, while a HFS with a single glucose unit forms long rod-shaped assemblies and produces heterogeneous PDCs. These results of LMNG and HFSs indicate that the size and homogeneity of detergent self-assemblies are not necessarily correlated to those of their PDCs. Due to the formation of homogeneous PDCs with a reasonable size, despite the formation of large rod-shaped self-assemblies, LMNG can be favorably used for membrane protein crystallization *via* vapor diffusion and/or cryo-EM protein structural study (*vide infra*).

3.2.2 MNG application for membrane protein study

LMNG is highly effective at stabilizing diverse membrane proteins. This maltoside was shown to be superior to DDM in stabilizing a number of membrane proteins including G protein-coupled receptors (GPCRs) [e.g., β 2AR-T4L, muscarinic M3 acetylcholine receptor (M3AchR), and mouse μ -opioid receptor (MOR)], bacterial transporters (e.g., MelB_{St}, LeuT, BmrA, and FhuA), a prokaryotic analogous of the eukaryotic NADPH oxidases (SpNOX) and a prokaryotic respiratory complexes (succinate:quinone oxidoreductase (SQR)). The enhanced protein-stabilizing effect of LMNG compared to DDM was observed with almost every membrane protein tested. One exceptional case is uric acid-xanthine transporter from *Aspergillus nidulans* (UapA), where LMNG was inferior to DDM for protein stability as assessed *via* fluorescence size exclusion chromatography (FSEC) (NBMs, 2017). In addition to the enhanced efficacy for protein stabilization, LMNG exhibits high extraction efficiency. For example, this NG agent extracted LeuT, β 2AR and LHI-RC complexes from the membrane, as efficiently as DDM. As a result, this maltoside can be widely used over the course of membrane protein study: membrane protein solubilization, purification and downstream characterizations (*vide infra*).

From the structural information on the online data bank, we found that LMNG has been used for solubilization, purification, and/or structural determination of 247 membrane proteins over the past ten years (from 2011 to 2021.07), approximately 13 % of total membrane protein structures (1936) reported during the same period (Table S2). Of the 247 membrane protein structures obtained using LMNG, 38 membrane proteins were solubilized and purified from the native membranes using LMNG, and then exchanged into other amphiphilic systems such as NDs or Apols for protein structure determination (Figure 7a). Some membrane proteins (42) were initially extracted using a conventional detergent (mostly DDM) from the native membranes and then exchanged into LMNG for protein structure determination. LMNG has also been used for extraction, purification and structure determination of a further 91 proteins. It is important to note that only a few detergents have such diverse applications in membrane protein study. LMNG has been particularly effective for structural study of GPCRs. Over the last 10 years, 121 GPCR structures were reported where this maltoside was used for receptor solubilization, purification and/or structural determination (Figure 7b). In addition to GPCRs, LMNG has been widely used for structural study of other classes of membrane proteins (*i.e.*, non-GPCRs) : channels (50) [mechanosensitive (3), ion-selective (20), transient receptor potential (TRP) (16), gap junctions (5), others (6)], transporters (40) [ATP-binding cassette (ABC) (14), exchange transporters (3), multi-drug efflux (2), major facilitator superfamily (MSF) (2), solute carrier (SLC) (5)], enzymes (25) [e.g., ATPase (3), cellulose synthase (2), oxidoreductase (3), translocase and insertase (10)], electron transport chain complexes [complex I (4), II (1), and IV (1), and respirasome (1)], amino acid/polyamine/organocation (APC) superfamily (5), flagellar motor proteins (3), novel receptors (2), and virus-coat proteins (2) (Figure 7). Collectively, LMNG has contributed to the structural

determinations of 126 non-GPCRs.

Analysis of the total 247 membrane protein structures obtained using LMNG, including both 121 GPCR and 126 non-GPCR structures, indicates that cryo-EM was most frequently used (175 structures; 71%) as a structure determination method, followed by LCP crystallization (58 structures; 23%) and *in surfo* crystallization (13 structures; 5%) (Figure 7c). When cryo-EM method was utilized, LMNG-purified proteins were directly used for the structure determination in the most cases (121 protein cases) with a small number transferred to another amphiphilic system at the vitrification step (16 protein cases). A similar trend was found for the data analysis of 121 GPCR structures obtained using LMNG. Cryo-EM and LCP crystallization account for 59 % (71) and 39% (47) of these structures, respectively (Figure 8), while vapor diffusion method (*i.e.*, *in surfo* crystallization) and NMR spectroscopy account for only two and one GPCR structures, respectively ([M₁ human muscarinic acetylcholine receptor (M₁AchR; 6WJC) and metabotropic glutamate receptor 5 (mGlu5; 6N4X)]; β_2 AR [6KR8]). This data indicates that *in surfo* crystallization and NMR spectroscopy are less successful methods for GPCR structural study using LMNG. Interestingly, the method used for GPCR structure determination tends to depend on the GPCR class (A, B, and the others (C, E and F)). LCP crystallization has been most favorably used for class A GPCRs (Figure 8), accounting for 43 (54%) of the 80 class A GPCR structures obtained using LMNG. The other class A GPCR structures were determined mainly *via* cryo-EM method (35 structures; 44%) where the LMNG-purified receptors were used directly. Interestingly, a different trend was found for 31 class B GPCR structures obtained using LMNG. All these structures were determined *via* cryo-EM where LMNG was used for receptor purification *and* vitrification. The only exception is the glucagon like peptide-1 receptor (GLP-1R) (6LN2), whose structure was determined *via* LCP crystallization. As for the 126 non-GPCR structures obtained using LMNG, cryo-EM is dominant (105 structures; 83%), followed by LCP method (11 structures; 9%) and *in surfo* crystallization (10 structures; 8%). When these protein structures were determined *via* cryo-EM method, LMNG was often exchanged into another amphiphilic system at the vitrification step. Other amphiphilic systems used for this purpose include Apols, NDs, digitonin, GDN, CHAPS and poly (maleic anhydride-*alt*-1-decene) substituted with 3-(dimethylamino)-1-propylamine (PMAL-C8). Although being the least favorable among three structure determination methods for non-GPCR structures obtained using LMNG, *in surfo* crystallization has generated high resolution crystal structures of several non-GPCRs such as the cytochrome *b6f* complex] the TatC core of the twin-arginine protein transporter, an ABC transporter lipid-linked oligosaccharide flippase, a borate efflux transporter, a plant voltage-gated two-pore channel, and FhuA. These findings indicate that LMNG is suitable for *in-surfo* protein crystallization. LMNG is thus compatible with all three main membrane protein structure determination methods: *in surfo* crystallization, LCP crystallization and cryo-EM (Fig. 7c).

Other MNG amphiphiles, decyl maltose neopentyl glycol (DMNG) and Cymal-5-NG, have been also

used for membrane protein structural study (Table S3). DMNG has been used for protein solubilization, purification, and/or structure determination (Fig. 7a). This decyl version of LMNG has facilitated the structure elucidation of 14 membrane proteins (Fig. 7b): 7 channels [e.g., lysosomal K^+ channel (5VRE), TRPV2 transient receptor potential channel (6BWM), SWELL1 (LRRC8A) volume-regulated anion channel (6DJB)], 4 ABC transporters [e.g., MacAB-TolC ABC-type tripartite multidrug efflux pump (5NIK), ABCG5/ABCG8 sterol transporter (5DO7), cystic fibrosis transmembrane conductance regulator (CFTR) (6D3R)], 3 enzymes [endoplasmic reticulum (ER)-associated protein degradation (ERAD) protein Hrd1 ubiquitin ligase complex (6VJY) and ATG9A transmembrane protein of the core autophagy machinery (7JLO)]. Interestingly, this short chain version of LMNG has not yet used for GPCR structural study. Similar to the trend found for LMNG use in protein structural study, cryo-EM (10) was most frequently employed for membrane protein structural study using DMNG, followed by *in surfo* crystallization (4) (Figure 7c). When the protein structures were determined via cryo-EM, DMNG-purified proteins were often transferred into another amphiphilic system (NDs, digitonin, or Apos). In the case of Cymal-5-NG, two membrane proteins [WRB/CAML/TRC40 insertase (6SO5) and TRPA1 transient receptor potential channel (6V9X)] were extracted and purified in this cyclic version of LMNG and then exchanged into PMAL-C8 for vitrification. The utility of DMNG and Cymal-5-NG is unsurprising given the widespread use of their conventional equivalents (DM and Cymal-7, respectively) in membrane protein structural study.

3.3 GDN

Similar to LMNG, GDN is markedly effective at stabilizing membrane proteins. In the initial study, compared to DDM, this steroidal detergent conferred notably enhanced stability to a few transporters [LeuT, MelB_S, the murine cytidine-5'-monophosphate-sialic acid (CMP-Sia)] and GPCRs [β_2 AR and δ -opioid receptor (δ OR)]. As exemplified with δ OR, GDN can confer greater stability to a target membrane protein than LMNG, indicating that this detergent has particular utility for unstable protein complexes. This steroid detergent has been widely used for protein purification, as expected from the excellent efficacy of this agent for protein stabilization. The online structural data for membrane proteins showed that GDN has been used for solubilization, purification, and/or structure determination of 157 membrane proteins over the past 5 years (2017 to 2021.07) (Table S4). This number constitutes ~13% of total membrane protein structures (1229) reported over the same period. The majority of these membrane protein structures (120; 76%) were determined after initial protein solubilization using a conventional detergent (mostly DDM; 75), LMNG (46), or others (e.g., digitonin), followed by detergent exchange into GDN (Figure 7a). Use of GDN for protein solubilization, purification and structure determination resulted in a further 21 protein structures. GDN is effective for structural study of GPCRs as well as non-GPCRs; this detergent has facilitated structure elucidation of 47 GPCRs and 110 non-GPCR including 53 channels, 23 transporters and 23 enzymes (Figure 7b). Thus, GDN is

suitable for structural study of diverse classes of membrane proteins, in contrast to LMNG which appears to be most suitable for structural analysis of GPCRs (121 of 247 structures; 49%).

Based on the popularity of digitonin in cryo-EM structural studies, GDN has recently received attention as a synthetic drop-in substitute for digitonin. The popular use of GDN for cryo-EM structural study can be found in our data obtained from the online membrane protein data bank. This structure determination method was used for 156 cases of the 157 membrane protein structures obtained using GDN (Figure 7c). The single exception is the structure of the Na_v1.7 sodium channel voltage-sensor domain with bound ProTx (6N4I) determined via *in surfo* crystallization. Similar to LMNG, membrane proteins solubilized/purified in GDN were sometimes exchanged into NDs (13) or digitonin (2) for cryo-EM study. It is noteworthy that the use of GDN for membrane protein structural study is rapidly growing (Figure 9). While no protein structure was obtained using GDN before 2017, GDN facilitated two membrane protein structures in 2017: activated GLP-1 receptor – G protein complex (5VAI) and the dimeric F_o region of mitochondrial ATP synthase (6B2Z). In 2018, a further four protein structures were obtained using GDN: the mechanically activated ion channel Piezo1 (6BPZ), TRPV3 channel (6DVZ), TRPM2 (6DRK), and Na_v1.4 sodium channel in complex with the β1 subunit (6AGF). The cryo-EM structure of Na_v1.4 protein contained clear electron density for GDN, inserted in the intracellular gate of the pore domain. A marked increase of GDN use was witnessed during 2019 and 2020, with a total of 35 and 63 membrane protein structures obtained using this detergent, respectively. Representatives include the mouse Piezo2 mechanosensitive ion channel (6KG7), the human TRPM2 channel (6PUO), and human ATP8A1-CSC50a complex (6K7G). The number of membrane protein structures obtained using GDN had already reached 53 this year as of July 31, 2021 (Figure 9). This data implies wide utility for GDN in future membrane protein structural efforts. In contrast to its popularity in cryo-EM structural study, GDN has yet to prove particularly useful for other protein structure determination methods. Caffrey, *et al* showed that the cubic phase formed by mono-olein was maintained in the presence of 20 mM GDN, indicating compatibility of this detergent with the LCP system. Thus, GDN has potential for LCP-crystallization of membrane proteins.

Two novel detergents (GDN and LMNG) have contributed significantly to membrane protein structural study. Use of GDN, LMNG, and a GDN/LMNG mixture have contributed to the determination of a total of 365 membrane protein structures (208, 118 and 39 protein structures, respectively) over the last 10 years. This number corresponds to about 20% of total membrane protein structures reported during the same period (Figure 9). When we consider membrane proteins structures reported over the last two years (from 2020 to July 2021), these two detergents have contributed to the determination of 223 membrane protein structures, accounting for more than one third of the total structures (37%) reported over the same period. LMNG has been more widely used for membrane protein structural study than GDN as can be seen in the number of total membrane protein structures obtained using each detergent

over the past 10 years (247 vs 157), but GDN became more widely used than LMNG for membrane protein structural study in 2021 (53 vs 40). A similar trend is found when we consider detergent use only for protein structure determination, thus excluding detergent use in protein solubilization and/or purification (Figure 9 inset). In 2020, LMNG and GDN were used in structure determination of 66 and 52 membrane proteins, corresponding to 17% and 13% of the total membrane protein structures determined over this period. In 2021, the number of membrane proteins structures determined by direct use of GDN is almost twice that determined using LMNG (49 (23%) vs 26 (12%)). Recent advances in cryo-EM technology are responsible for the increased use of GDN in membrane protein structural study.

3. Detergent structure-efficacy relationship of OGNG, LMNG and GDN

3.1 NG class detergents (OGNG/LMNG/DMNG)

OGNG and LMNG are dimeric forms of OG and DDM, respectively, and are significantly more effective than the respective conventional equivalent for membrane protein stabilization. This could be explained by enhanced interactions of these detergents with membrane proteins. Due to the presence of the two alkyl chains, this class of detergents (GNGs and MNGs) interacts more strongly with the hydrophobic surface of membrane proteins than the conventional equivalents with a single alkyl chain. MD simulation analysis with wild type adenosine A_{2A} receptor (WT-A_{2A}R) showed that *detergent-protein* interaction energy was increased by 17% in the case of LMNG compared to DDM. The NG class detergents (e.g., OGNG and LMNG) are also likely to strongly interact with other NG molecules in the PDC environments. This is suggested by the significantly lower CMCs of OGNG and LMNG compared to OG and DDM, respectively. The MD simulation showed that *detergent-detergent* interaction energy was increased by 15% when LMNG instead of DDM was used for interaction with WT-A_{2A}R. The enhanced *detergent-protein* and *detergent-detergent* interactions would result in tight packing of detergent micelles around membrane protein surfaces. The simulation study also suggested that the branched dimaltoside head group of LMNG favors formation of sustained intramolecular hydrogen bonds between the two maltose groups. In the resulting conformation, the branched head group of LMNG is able to form bifurcated hydrogen bonds with polar residues of transmembrane helices and loops in WT-A_{2A}R. These additional interactions between detergent molecules and membrane protein surfaces are favorable for further stabilizing and rigidifying the receptor. OGNG is likely to have weaker *detergent-protein* and *detergent-detergent* interactions than LMNG, mainly due to the presence of the short alkyl chain (*i.e.*, reduced hydrophobicity). While LMNG micelles are tightly packed around the membrane protein surfaces, consequently, OGNG micelles would be more loosely associated with membrane proteins, thereby reducing membrane protein stability. This is consistent with MD simulations of these two detergents in complex with two GPCRs (WT-A_{2A}R and WT-β₂AR) [Biochemistry 2020, 59, 2125–2134], revealing that OGNG undergoes frequent molecular tumbling within detergent micelles surrounding the receptors, not observed for LMNG. This molecular tumbling

is likely due to the relatively weak interaction of OGNG with other detergent molecules as well as with the protein surfaces. The 100-fold higher CMC of OGNG compared to that of LMNG indicates a relatively weaker tendency of OGNG to self-associate, thus supporting the relatively weak molecular interactions between detergent molecules in PDCs.

Strong *detergent-detergent* interactions are not only necessary to prevent protein aggregation, but also play a favorable role in minimizing protein denaturation. Thus, the dimeric architecture of the NG class detergents which allow strong interactions can be advantageous for stabilizing membrane proteins, as can be seen with LMNG. However, the same feature of NG detergents increases *detergent-protein* interaction, as described above. While being favorable for preventing protein aggregation, strong *detergent-protein* interactions could be detrimental for membrane protein stability as it induces protein denaturation, particularly for denaturation-sensitive proteins. The anionic detergent, SDS is a good example of a detergent which forms strong *detergent-protein* interactions and as a result causes disruption of the hydrophobic interactions holding a protein together. Thus, NG detergents should be favorable for stabilizing membrane proteins with a marked tendency to aggregate, but may be sub-optimal for stabilization of denaturation-prone membrane proteins. In order to stabilize a wide range of membrane proteins with diverse tendencies to aggregate and denature, detergents need to have minimal interaction with the membrane proteins, but are required to make strong interaction with other detergent molecules. These favorable molecular interactions for membrane protein stability are exemplified by the lipid molecules in cell membranes.

3.2 GDN

This steroidal detergent is structurally distinct from conventional detergents as well as the NG class agents. GDN contains a multiple fused ring (*i.e.*, steroid unit) in the lipophilic group and thus has a high hydrophobic density. The hydrophobic group of this detergent is very rigid, while those of conventional and NG class detergents are flexible. These features of GDN, high hydrophobic density and rigidity of the detergent hydrophobic group, are favorable for achieving enhanced *detergent-detergent* interaction in a PDC environment. Resultantly, GDN has the potential to form stronger *detergent-detergent* interactions than LMNG. In addition, this detergent contains a polar spiroketal unit at the terminus of the hydrophobic group, a feature common with digitonin and distinct from other steroid detergents (*e.g.*, chobimalt). This unique structural feature seems to be favorably associated with the successful use of GDN/digitonin in cryo-EM structural studies. The presence of the polar functional group in the terminal region significantly reduces detergent interactions with the hydrophobic surfaces of membrane proteins. As exemplified by CHAPSO or fluorinated surfactants (FSs), introduction of a lipophobic group into the detergent lipophilic region is often detrimental for protein stability, presumably as the result of reduced *detergent-protein* interaction. Consequently, membrane proteins solubilized in these detergents tend to aggregate over time. To circumvent this problem, the reduced *detergent-protein* interaction

needs to be accompanied by enhanced *detergent-detergent* interaction. This is the case for the GDN/digitonin where the spiroketal unit in the lipophilic group modulates *detergent-protein* interaction, while its multiple-fused ring with its rigid architecture strengthens *detergent-detergent* interaction. Thus, we envision that these detergents are highly effective at stabilization of both denaturation- and aggregation-prone membrane proteins. Based on the current discussion, the diosgenin hydrophobic group rather than the hydrophilic group is mainly responsible for the favorable behaviors of GDN with respect to membrane protein structural study. Compared to GDN/digitonin, other steroid detergents (e.g., chobimalt) have an alkyl chain at the termini of the hydrophobic groups, similar to the single alkyl chain found in conventional detergents. Thus, these steroidal detergents are similar to OG and DDM in terms of *protein-detergent* interactions although they are likely to have significantly improved *detergent-detergent* interactions.

4. Variants of OGNG, LMNG and GDN

4.1 OGNG variants

Due to the limited efficacy of OGNG for protein stabilization, a number of variants of this glucoside have been invented. Initial variants contain different linkages in the lipophilic region, introduced to conjugate the alkyl chains to the central NG unit. Ether and amide linkages are utilized to generate GNG-1 and GNG-5, respectively (Figure 10a). GNG-1 was generally better than OGNG at stabilizing membrane proteins, but the difference was relatively small. GNG-5 was comparable to OGNG at stabilizing LHI-RC complexes. These variants having two identical alkyl chains, referred to as symmetric GNGs, were overall comparable to or slightly inferior to DDM for protein stability. The limited efficacy of these symmetric GNGs for protein stabilization is mainly attributed to the short alkyl chain length that matches poorly with typical membrane protein dimensions. The short chain length of the GNGs is also unfavorable for effective *detergent-detergent* interaction in PDCs. Thus, it is necessary to increase the alkyl chain length of GNGs in order to enhance protein stabilization efficacy. However, due to the presence of the relatively small glucoside head group, an increase in alkyl chain length resulted in limited water-solubility of the detergents. For example, GNG-4, containing two alkyl chains longer than those of OGNG by just one carbon unit (CH_2), gave only 1.0 wt% water-solubility (Figure 4). As the presence of the glucoside head group is essential for small PDC formation, it is important to increase alkyl chain length without changing this head group. We solved this issue by designing P-GNGs having two different alkyl chains. In the P-GNGs, one alkyl chain is short and the other chain is long and thus these two chains were designated as pendant and main chains, respectively (Fig. 10b). The pendant chains varied from C1 (methyl) to C3 (propyl), while the alkyl chains in the range of C11 to C14 were used as main chains. When tested with several membrane proteins including two GPCRs ($\beta_2\text{AR}$ and MOR), these P-GNGs were superior to OGNG and DDM in stabilizing the tested membrane proteins. Detergent efficacy of the P-GNGs for protein stabilization tended to increase with the length

of pendant and/or main chains. Consequently, the P-GNGs with C3 pendant and C13/C14 main chain (*i.e.*, GNG-3,13 and GNG-3,14) were the most effective for protein stabilization. It is notable that those P-GNGs were as effective as LMNG at stabilizing LeuT and MOR, in spite of having a glucoside head group. The pendant chain placed at the interface between the hydrophobic and hydrophilic groups fills the empty spaces in the micelle interiors, without disturbing micellar architecture. Thus, the presence of the pendant chain increases hydrophobic interaction between detergent alkyl chains in the interior of micelles formed by the P-GNGs. In addition, the pendant chains of the P-GNGs are unlikely associated with *protein-detergent* interaction as these are located far from the membrane protein surfaces. In other words, only the detergent main chains directly interact with membrane proteins, similar to the conventional detergents, OG and DDM. Thus, the introduction of the pendant chains here effectively enhances *detergent-detergent* interaction without a concomitant increase in *detergent-protein* interaction. This is an ideal feature for a detergent that is effective for both denaturation-sensitive and aggregation-prone membrane proteins, as described above. In terms of solubilisation efficiency, the P-GNGs were comparable to OGNG and DDM, as supported by analysis of extraction of the bacterial transporter MelB.

We have also developed hydrophilic variants of OGNG. Like OGNG, neopentyl glycol-derived triglucosides (NDTs) have two identical alkyl chains, but with a head group altered from a branched diglucoside to a branched tri-glucoside. Using the pentaerythritol linker, we readily introduced two identical alkyl chains in the range of C10 to C12 and the three glucose units into the detergent scaffold (Fig. 10c). Of the NDTs, NDT-C11 was most effective at protein stability and superior to DDM for stabilizing multiple membrane proteins tested, including MelB and LeuT. A preliminary result showed that this detergent was also effective at stabilizing β_2 AR and thus has potential for GPCR structural study (unpublished). We also invented the tetra-glucoside versions of OGNG, designated malonate-derived tetra-glucosides (MTGs) (Fig. 10d). The MTGs with the ether linkages (*e.g.*, MTG-E12) were shown to be effective at stabilizing multiple membrane proteins compared to DDM and OGNG. Both the NDT and MTG detergents have a greater number of glucose units (three and four, respectively) than OGNG (two). Thus, the head groups of these detergents have enhanced hydrophilicity compared to that of OGNG, allowing increase of the alkyl chain length up to C12/C14 without causing any water-solubility issue. The increased alkyl chain length of these detergents (NDT-C11 and MTG-E12) is likely to be responsible for their superior efficacies for protein stabilization compared to OGNG.

We also utilized tandem structures of the NG unit in detergent invention. As an example, tandem malonate-based glucosides (TMG-As/Ts) contain two NG units connected to each other in a parallel fashion using a methylene/thioether linkage (Fig. 10e). Because of the presence of the alkyl spacer between the two NG units, the distance between the two alkyl chains was substantially increased in the TMG architecture compared to OGNG. These tandem detergents (*e.g.*, TMG-T13, TMG-A13) were

shown to display favorable behaviors for protein stability compared to DDM when tested with UapA, LeuT, β_2 AR, and MelB. As a further expansion, we also prepared a class of detergents, designated mesitylene-cored glucoside amphiphiles (MGAs), that contains three NG units in the structures (Fig. 10f). These NG units were connected to each other by a mesitylene core and the resulting variants bear six glucose units. A couple of detergents such as MGA-C13 and MGA-C14 displayed generally enhanced efficacy for stabilizing multiple membrane proteins such as LeuT, β_2 AR, and MelB. MGA-C13 formed small PDCs with β_2 AR, probably due to the presence of the glucoside head group, and retained the receptor stability more effectively than DDM in the detergent-free conditions. It is noteworthy that the OGNG variants mentioned above [hydrophobic (GNG-1/5, P-GNGs), hydrophilic (MTGs and NDTs) and tandem (TMG-As/Ts and MGAs)] were efficiently prepared and developed owing to the convenience of structural modification in the NG scaffold. These variants are the significantly upgraded versions of OGNG, a detergent that has proved effective for *in-surf*o crystallization of membrane proteins. In addition, these glucosides conferred enhanced stability to membrane proteins comparable to or better than LMNG, a molecule widely used for protein structural determination *via* LCP crystallization and cryo-EM study. Thus, these glucoside variants, particularly NDT-C11, GNG-3,13, and MGA-C13, will find use in membrane protein structural study in a near future.

4.2 LMNG variants

Based on the wide utility of LMNG in membrane protein structural study, LMNG variants could be more useful than OGNG variants. However, it would be harder to develop LMNG variant with an enhanced efficacy for protein stabilization as the original LMNG is already a significantly optimized detergent in this context. The first variants of LMNG contained different linkages in the lipophilic region, amide (MNG-1) and ether (MNG-2) (Figure 10g). These MNGs were rather inferior to LMNG (MNG-3) in membrane protein stability. The hydrophobic variants of LMNG with different alkyl chain length were invented, as exemplified with undecyl maltose neopentyl glycol (MNG-4; UDMNG) and DMNG (MNG-5) (Fig 4d). In addition, we reported a cyclohexyl version of LMNG, designated MNG-3-C9Cy (commercial name: Cymal-5-NG) (Figure 10h). This detergent was designed based on the fact that CYMAL-7, a cyclohexyl version of DDM, is frequently used in membrane protein purification and structural determination. The hemi-fluorinated versions of LMNG are also notable as they contain a distinctive feature from other HFSs. By incorporating a 3,5-difluorinated benzene ring at the end of the lipophilic group (Figure 10i), we prepared F4-MNG which was markedly effective at stabilizing the denaturation-sensitive *R. capsulatus* LHI-RC complex. In the presence of variant with a 3,5-trifluoromethyl (CF₃)-substituted benzene ring (F12-MNG) instead of the fluorine (F)-substituted ring, the LHI-RC complexes underwent significant protein aggregation. Thus, the behaviors of F4-MNG and F12-MNG are similar to those of the HFSs and FSs, respectively. As the hydrophobic group containing

the fluorinated benzene ring is commercially available, F4-MNG is highly accessible compared to other HFSs that require multiple steps for synthesis of their hemi-fluorinated chains.

All the MNG analogs described above have two identical alkyl groups like the original LMNG, and these symmetrical MNGs may not be optimal for molecular packing of detergent micelles around membrane proteins. As the central region of detergent micelles or the interfaces between detergent micelles and membrane proteins in PDCs are very congested with many alkyl chains, the presence of two alkyl chains of the same length could exert steric hindrance in this area. To relieve this energetically unfavorable interaction in the detergent micelle interiors or at detergent-protein interfaces, we introduced MNG amphiphiles with two different alkyl chains, designated A-MNGs ('A' stands for asymmetric), as exemplified by MNG-8,12, MNG-6,14 and MNG-2,18 (Fig 10j). As indicated in the numbers in detergent designation, these A-MNGs have the same number of carbon units as LMNG (*i.e.*, MNG-10,10) but these are unevenly distributed through the two alkyl chains. As expected, these detergents (MNG-8,12, MNG-6,14 and MNG-2,18) gave lower CMCs (8, 6, and 2 mM, respectively) than LMNG (10 mM), due to the reduced steric hindrance in the micelle core. More importantly, MNG-8,12 and MNG-6,14 were superior to LMNG for stabilization of multiple membrane proteins (LHI-RC, LeuT, β_2 AR and MOR), with MNG-8,12 most effective at stabilizing two GPCRs (β_2 AR and MOR). As LMNG is widely used for structural studies of GPCRs and channels, these A-MNGs (MNG-8,12 and MNG-6,14) have significant potential for structure determination of these types of membrane proteins via LCP and cryo-EM. For the preparation of tandem variants of LMNG, designated tandem neopentyl glycol maltosides (TNMs), the two NG units were connected to each other in series *via* two different linkers and then three alkyl chains and three maltose groups were introduced onto either side of the tandem NG unit (Figure 10k). TNM-C11S and TNM-C12L exhibited a marked stabilizing effect on LeuT and β_2 AR compared to DDM, respectively.

Due to their modular synthesis and synthetic convenience, more MNG variants can be readily designed and developed. For example, further variants of the A-MNGs are interesting as the previous A-MNGs (*e.g.*, MNG-8,12) were notably superior to LMNG for GPCR stability. While the previously reported A-MNGs are variants of LMNG with two C10 alkyl chains (*i.e.*, MNG-10,10), we can prepare asymmetric variants of the MNG with C11/C12 alkyl chains (*i.e.*, MNG-11,11/MNG-12,12). MNG-11,11 showed limited water-solubility (1.0 wt%), but our study showed that introduction of asymmetry in the detergent lipophilic region increases water—solubility. Thus, asymmetric versions of C11/C12 alkyl-chained MNG, as exemplified by MNG-10,12, MNG-9,13, and MNG-8,14; MNG-11,13, MNG-10,14, and MNG-9,15, are expected to be sufficiently water-soluble for membrane protein study. We anticipate that some of these A-MNGs might further enhance detergent efficacy for GPCR stabilization compared to MNG-8,12. Hydrophilic variants of LMNG are also interesting for structural study of membrane proteins as the detergent hydrophilic group often dictates the structural determination

methods and/or the target membrane protein classes. As mentioned above, small PDC-forming glucoside detergents such as OG and OGN are generally favorable for *in surfo* crystallization, whereas maltoside detergents with enhanced protein stabilization efficacy (e.g., DDM and LMNG) are more widely used for LCP-based crystallization or cryo-EM study. Zwitterionic detergents such as LDAO and dodecylphosphocholine (DPC) are more suitable for protein structural study *via* NMR spectroscopy (Fig. 3e). Thus, zwitterionic MNGs, dimeric forms of LDAO and DPC, can be useful for NMR-based membrane protein structural study. These detergents can be prepared by incorporating the head group of amine-oxide/phosphocholine into the MNG scaffold. We are currently investigating these versions of the MNG in terms of their effect on protein stability and their utility for NMR-based structural study. Although LMNG has been widely applicable for diverse classes of membrane protein structural study, this maltoside is rarely useful for structure determination of β -barrel proteins (e.g., porins) with only one example [Aerolysin (5JZT)] reported to date. It is known that detergents with an oligoethylene glycol head group such as C₈E₄ and C₁₂E₉ are favorable for structural study of this class of proteins. Thus, the oligoethylene glycol versions of LMNG could be useful when we are to determine 3D structures of challenging β -barrel proteins. These hydrophilic variants of LMNG are expected to be more effective at stabilizing membrane proteins than the respective conventional analog (e.g., LDAO, DPC, or C₈E₄) as OGN and LMNG are superior to OG and DDM in this regard, respectively.

4.3 GDN variants

As GDN contains a steroidal hydrophobic group (*i.e.*, diosgenin), there is little opportunity for development of hydrophobic variants of this detergent. Other steroid units such as cholesterol and cholestanol can be used as structural variants of diosgenin. However, these hydrophobic groups lack the spiroketal unit in the terminal region that plays a crucial role in modulating *detergent-protein* interaction in the PDCs. Thus, these variants are unlikely to be as useful as GDN for membrane protein structural study. In addition, while diosgenin consists of six cycloalkane rings in a flat architecture, cholesterol/cholestanol comprises four cycloalkane rings, an arrangement that is likely to be less effective than the former for favorable *detergent-detergent* interaction in the PDCs. Thus, we have focused on hydrophilic or linker variations as for development of new GDNs. For instance, we utilized a penta-saccharide head group to generate steroid-based penta-saccharide (SPS) detergents (e.g., SPS-4 and SPS-4L) based on the high efficacy of PSEs with the same head group for protein stabilization (Fig. 10I). However, these hydrophilic variants were less effective than GDN for LeuT and β_2 AR stability. The presence of the bulky head group seems to hinder favorable *detergent-detergent* interaction of these new molecules in PDCs. To resolve this issue, we recently developed GDN variants, named GDN-C2/C3, that contain a branched dimaltoside head group, structurally similar to that of the original GDN, but with different linkers between the head and tail groups [*under revision*]. Compared to the original GDN, these GDN variants have improved synthetic convenience (*i.e.*, short synthetic

steps and high overall yields) owing to the utilization of the more accessible linkers. The resulting GDNs (GDN-C2 and GDN-C3) were comparable to or better than the original GDN at stabilizing multiple model membrane proteins tested. For example, GDN-C3 was comparable to GDN for MelB stability, while this new detergent conferred an increased T_m of MOR by 2.6 °C compared to the original GDN. Since GDN has proved particularly successful for membrane protein structural analysis by cryo-EM, these GDN variants (e.g., GDN-C2/C3) represent excellent drop-in substitutes for the original GDN. The CMCs and R_h for the variants of OGNG, LMNG and GDN are summarized in Table 2, where the R_h values MNG-2 and MNG-3-C9Cy were measured for this work.

5. Structurally distinct amphiphiles

5.1. Carbohydrate-cored amphiphiles (MNAs, SIGs and TCGs/TCMs)

Recently, several classes of novel amphiphiles with distinctly different architectures from the NGs/GDN have been developed. Due to the distinct architecture of the NG/GDN classes, these amphiphiles are likely to provide unique behaviors toward membrane protein structural studies. Carbohydrate cored amphiphiles, whose representatives include mannitol-based amphiphiles (MNAs), scyllo-inositol glycosides (SIGs), trehalose-cored maltosides/glucosides (TCMs/TCGs), and [Hazrat Hussain, et al, *Chem Eur J* 2016; A. Sadaf, et al, *ACS Chem Bio* 2020; M. Das et al, *Org Biomol Chem.* 2018], contain different carbohydrates (mannitol, scyllo-inositol and trehalose, respectively). Due to the presence of multiple alcohol groups, these carbohydrates units within the core of the molecule effectively act as linkers for the introduction of the detergent alkyl chains and head groups. In the MNA class, two alkyl tails and four glucose head groups were attached to the six alcohol groups of the mannitol linker [Figure 11a]. The resulting MNAs contain the same number of the head and tail groups as the TMGs and MTGs (described above), but the distance between the two alkyl chains and/or the head groups arrangement is substantially smaller/bigger??. These differences result in a large variation in the molecular geometry of the MNAs compared to the TMGs/MTGs, leading to distinct molecular packing of the detergent micelles around membrane proteins. The MNA with a C12/C13 alkyl chain (MNA-C12/C13) was not only efficient at extracting BOR1 from the membrane, but also highly effective at stabilizing several membrane proteins; BOR1, LeuT, LHI-RC complex and β_2 AR [H. Hussain, et al, *Chem Eur J* 2016]. Notably, MNA-C12 and MNA-C13 were superior to LMNG at stabilizing BOR1. In addition, these MNA detergents produced smaller PDCs with β_2 AR than DDM. These results suggest significant potential of these MNAs for protein structural studies *via* cryo-EM. Indeed, the cryo-EM structure of the full length hemagglutinin-Fab complex was recently determined in MNA-C12 micelles [D. J. Benton, *PNAS*, **2018**, 115, 10112-10117]. Of note, MNA-C12 and MNG-C13 were used to form markedly more stable picodiscs with saposin A at neutral pH compared to DDM. The resulting MNA/Saposin picodiscs have promise for structural characterization of a range of membrane proteins [K.W. Kurgan, *Biochemistry*, 2021, 60, 1108-1119]. Thanks to the synthetic accessibility and modular

Commented [BB1]: This sentence doesn't really add anything.

Commented [BB2]: Could say a bit more here-what precisely are the variations in molecular geometry and molecular packing?

synthetic approach, the MNA architecture could be further expanded to incorporate branched alkyl chains instead of the straight alkyl chains, as exemplified with MNA-AC15 and MNA-AC16 [H. Hussain, *Analyst*, 2018] (Figure 11b). The branched alkyl chains would play a favorable role in membrane protein stability by filling the empty spaces in the detergent interiors, thereby increasing *detergent-detergent* interactions, as shown by the comparative study of P-GNGs with the GNGs. Indeed, the branched C15 alkyl-chained MNA (*i.e.*, MNA-AC15) was superior to the linear MNA-C13 for LeuT and MelB stability, but was suboptimal for β_2 AR stability, indicating that this detergent is most suitable for structural studies of membrane transporters. A further variation in length of the branched chain might be necessary to improve detergent efficacy for stabilization of GPCRs (*e.g.*, β_2 AR)

The SIGs contain a rigid unit of scyllo-inositol in the detergent core region, used as the linker to connect detergent head and tail groups. This carbohydrate has a unique structure, placing six hydroxyl groups around a cyclohexane ring all in the equatorial positions. Thus three alkyl chains and three glycosides (glucosides/maltosides) could be strategically introduced into either face of this ring to produce the amphiphilic molecules [Figure 11c]. The resulting SIGs possess the same number of head and tail groups as the TNMs, the tandem variants of LMNG, but the head and tail groups in the SIGs are in closer proximity to each other *via* the small and rigid linker (*i.e.*, scyllo-inositol) than the TNMs. This compact architecture generates high hydrophobic density in the hydrophilic-hydrophobic interfaces of detergent micelles, a favorable feature for effective *detergent-detergent* interactions associated with stable micelle and PDC formations. Of the various SIGs, a ones with C11/C12 alkyl chains and maltoside head groups, designated STM-11/12, conferred particularly marked stability on a range of membrane proteins including LHI-RC complexes, LeuT, MelB, and β_2 AR compared to DDM [A. Sadaf, et al, *ACS Chem Bio* 2020]. These two STMs were even better than LMNG in maintaining the long term stability of LeuT and the LHI-RC complex.

In the case of the TCMs/TCGs, the eight alcohol groups in a trehalose unit were utilized to introduce four alkyl chains on one face and four maltoside/diglucoside groups on the other face of the disaccharide ring (Figure 11d). The trehalose group was previously used as detergent head group for generation of dodecyl trehaloside detergents (DDTre) [H. Tao, et al, *Langmuir*, **2012**, 28, 1117311181]. The DDTre detergents have a single dodecyl chain attached to different locations of the trehaloside alcohol groups, different markedly from the TCMs/TCGs with their four alkyl chains projecting from same face of the core ring. The DDTre detergents, derivatives of DDM, were shown to be comparable to DDM for stabilization of *S. typhimurium* MsbA (ATP-binding cassette transporter) and nociceptin/orphanin FQ peptide receptor (ORL1; GPR). In contrast, all the tested TCMs/TCGs were shown to be markedly effective at stabilizing LHI-RC complexes and LeuT compared to DDM, while only the C10 alkyl-chained TCM (TCM-C10) was better than DDM in terms of β_2 AR stability [M. Das et al, *Org Biomol Chem*. 2018]. Some carbohydrate-cored amphiphiles, particularly MNA-C13, MNA-AC15, and STM-

11/12, have been shown to be more effective than LMNG for membrane protein stability, suggesting they have significant potential in membrane protein structural study.

5.2. Stereo-isomeric amphiphiles (BTMs, NBMs, and CPMs)

The carbohydrate-cored amphiphiles introduced above have defined stereo-chemistry in their core structures, while the NG class has no chirality center in the core unit. Thus, the relative directions of the detergent tail and head groups of these detergents were restricted by the stereochemistry of the carbohydrate core units (MNAs, SIGs, and TCMs) or by the structure of the NG linker (NG class). As this relative orientation of detergent head and tail groups is strongly associated with the molecular packing of detergent assembly around membrane proteins, a stereo-chemical variation in the detergent core unit is important to enhance detergent efficacy for membrane protein stabilization. The first systematic study of this type of feature was carried out with butane-1,2,3,4-tetraol-based maltosides (BTMs). This class contains two alkyl chains and two maltoside head groups like LMNG, but the BTMs have a butane-1,2,3,4-tetraol (BT) unit containing two variable chirality centers (two asymmetric carbons; C2 and C3) in the core region (Figure 11e), while the NG linker of LMNG has no chiral carbon. Accordingly, we could generate the three stereoisomers (A-BTMs, B-BTMs, and M-BTMs) with different configurations of two chirality carbons (2*S*,3*S*; 2*R*,3*R*; 2*S*,3*R* or 2*R*,3*S*, respectively), in contrast to the presence of a single stereoisomer for the carbohydrate-cored amphiphiles or the absence of a stereoisomer for the NG class detergents. When the three sets of BTMs were compared to each other in terms of membrane protein stability, the M-BTMs were superior to the other stereoisomers (A- and B-BTMs) for stabilization of the three tested membrane proteins (UapA, MelB and β_2 AR) [A. Sadaf, *ChemSci*, 2016]. A similar outcome associated with detergent stereochemistry was obtained for the NBMs. The norbonane core unit (*i.e.*, NB) in the NBMs contains the two variable chirality centers (C2 and C3) like the BT unit of the BTMs, but the former is more rigid than the latter due to its bicyclic structure [Figure 11f]. When two sets of stereoisomers (*exo* and *endo* forms) of the NBMs (X-NBMs and D-NBMs, respectively) were tested for membrane protein stability, the *exo*-forms of the NBMs (e.g., X-NBM-C11) were more effective than the *endo*-counterparts (e.g., D-NBM-C11) at stabilizing UapA, β_2 AR and the β_2 AR-G_s complex [M. Das, *J. Am. Chem. Soc.* 2017]. Furthermore, X-NBM-C11 was superior to LMNG for UapA stabilization and fully maintained the integrity of the β_2 AR-G_s complex over a three-week incubation at room temp^{??}. The combined results of the BTMs and the NBMs highlights the importance of detergent stereochemistry in membrane protein stability. A similar yet distinct stereo-chemical effect on protein stability was obtained from evaluation of the stereo-isomeric CPMs with membrane proteins [M. Das, *J. Am. Chem. Soc.* 2020]. This class contains a conformationally variable cyclopentane (CP; monocyclic) ring in the detergent core region, in contrast to the conformationally rigid NB unit (bicyclic) of the NBMs. In view of the stereochemistry of the detergent core units (*i.e.*, CP and NB), the *cis* and *trans* forms of the CPMs (designated CPM-Cs and

CPM-Ts, respectively) correspond to the *endo* and *exo* forms of the NBMs, respectively [Figure 11g]. When evaluated with several membrane proteins including LeuT, MelB, β_2 AR and MOR, CPM-C12 was better than CPM-T12 for protein stability. In contrast, for the NBMs, the *trans* forms (e.g., X-NBM-C11) was better than the *cis* forms (e.g., D-NBM-C11) for stabilization of the same set of membrane proteins. The conversion of the CP ring from *envelope* to *half-chair* conformation and the resulting change in the detergent hydrophobic length is likely to be responsible for the enhanced behavior of CPM-C12 compared to CPM-T12. Consequently, detergent efficacy for protein stabilization not only varies according to stereochemistry (*cis/trans*) of the detergent core unit, but also is further optimized by the conformational change in the core unit (*half-chair/envelope*). Notably, CPM-C12 was superior to LMNG in terms of stability of two GPCRs (β_2 AR and MOR), and effectively used for extraction and purification of the prokaryotic voltage-dependent potassium channel KvAP cloned from *Aeropyrum pernix* [M. Das, *J. Am. Chem. Soc.* 2020]. The collective findings on the BTMs, NBMs and CPMs indicate high potential of in particular M-BTM-C11, X-NBM-C11 and CPM-C12, as biochemical tools for membrane protein structural study and demonstrate the importance of the conformational variability as well as stereochemistry of the detergent core unit for optimal membrane protein stability.

5.2. Facial amphiphiles (FAs, TFAs and LFAs)

Bile acids such as cholic and deoxycholic acids have a rigid and hydrophobic scaffold similar to other steroid-based compounds such as diosgenin and cholesterol used for preparation of GDN and chobimalt, respectively. However, the bile acids are distinct from these other steroids, as they contain multiple hydroxyl groups which project only in one direction, and are thus effective scaffolds for generation of facial amphiphiles. Conventional detergents such as CHAPS and CHAPSO also contain a cholate-based scaffold, but these molecules are only weakly amphiphilic. Q. Zhang *et al* reported several facial amphiphiles derived from cholic acid [Q. Zhang, *et al*, *Angewandte Chemie*, **2007**, 46, 7023-7025; S. C. Lee, *et al*, *PNAS*, **2013**, E1203-E1211] (Figure 11h-j). As expected from a facial interaction with a membrane protein surface, only small numbers of these detergent molecules tended to associate with membrane proteins, resulting in small detergent-protein ratios in their PDCs. For example, the number of FA-1 molecules that associate with MsbA was estimated to be 37, significantly smaller than the number of the UDM counterpart (219) [Q. Zhang, *et al*, *Angewandte Chemie*, **2007**, 46, 7023-7025]. The formation of a small PDC for dimeric MsbA associated with the FAs, compared to MsbA + UDM/DDM was supported by analytical SEC [S. C. Lee, *et al*, *PNAS*, **2013**, E1203-E1211]. When evaluated with a sample of membrane proteins, two FAs (FA-3 and FA-5) with two maltoside and three glucoside head groups, respectively, were superior to OG at stabilization of bR from *Halobacterium halobium* and were better than DDM for stabilization of MsbA from *E. coli*. These facial amphiphiles were also comparable to or slightly better than DDM at stabilization of the human gap junction channel composed of connexin 26 (Cx26). The enhanced stability of MsbA and Cx26 attained using these

amphiphiles (FA-3 and FA-5) resulted in high resolution crystal structures of these membrane proteins [S. C. Lee, et al, *PNAS*, **2013**, E1203-E1211].

In the presence of membrane proteins, detergent faciality can be largely induced by the introduction a small alkyl chains as a detergent pendant. One class of detergents with this feature is the lithocholate-based facial amphiphiles (LFAs) that contain two different hydrophilic groups (a maltoside and a branched maltoside group) positioned at either end of the detergent molecule (Figure 11k). LFA-C4 bearing a short butyl pendant showed a remarkable efficacy for membrane protein stabilization compared to DDM/OG when tested with multiple membrane proteins (LeuT, bR, and β_2 AR) [M. Das, *Chem Eur J*, 2018]. In contrast, an analog with no alkyl pendant, designated lithocholate-based bola-amphiphile (LBA-C0), resulted in poor stability of these membrane proteins indicating that a facial amphiphile is a better architecture than a bola-amphiphile for membrane protein stability. Due to the presence of the butyl chain, LFA-C4 has enhanced ability to increase detergent facility by inducing facial arrangement of the detergent head and tail groups and resultantly allows facial interaction of this molecule with membrane protein surfaces. Furthermore, the butyl pendant chain of LFA-C4 has an additional role in tuning the micelle hydrophobic dimensions according to the hydrophobic dimensions of an individual target membrane protein. This customizable nature can be attained by adjusting the degree of chain overlap between the butyl chains of two LFA molecules aligned to span membrane protein surfaces. These two features of LFA-C4, protein-induced detergent faciality and target protein-dependent tenability of the degree of chain overlap, are responsible for the generally favorable behavior of LFA-C4 in terms of membrane protein stabilization. These features are impossible in both LBA-C0 and the FAs. LBA-C0 has no alkyl pendant and the hydrophobic lengths of the FAs (~ 13.4 Å) are too small to allow chain overlap when interacting with membrane proteins of variable hydrophobic widths (28–32 Å).

The FAs and LFA-C4 have hydrophobic lengths of 13.4 and 17.9 Å, respectively, roughly half of the hydrophobic width of membrane proteins, thus two molecules of each detergent are needed to fully span the hydrophobic thickness of membrane proteins [Q. Zhang, et al, *Angewandte Chemie*, **2007**, 46, 7023–7025; M. Das, *Chem Eur J*, 2018]. A facial amphiphile having two steroid units connected by a short linker would have the combined hydrophobic length sufficient in a single detergent molecule to span the hydrophobic surfaces of membrane proteins by a single detergent molecule. By connecting two units of deoxycholic acid using a 1,3-diamine-functionalized alkyl linker, we constructed tandem facial amphiphiles (TFAs) with four maltoside head groups projecting in the same direction [Figure 10m]. Detergent hydrophobicity was varied through the introduction of different lengths of alkyl appendages (methyl (TFA-1), ethyl (TFA-2), and butyl chains (TFA-3)) onto the amide amine group of the TFAs. Micelles formed by TFA-1 with the two methyl appendages were estimated to contain only 6 detergent molecules, in marked contrast to DDM micelles containing 176 detergent molecules. TFA-1 was shown

to be superior to DDM, LDAO or OG at stabilizing bR, LHI-RC complexes, and cytochrome *bo3* ubiquinol oxidase (Cyt *bo3*), but was inferior to DDM for LeuT and β_2 AR stability [P. S. Chae, *J. Am. Chem. Soc.*, 2010]. The suboptimal property of TFA-1 for LeuT and β_2 AR stability could be due to the relatively high polarity and *trans-cis* isomerism associated with the amide functional groups present within the detergent hydrophobic surface. To minimize the polarity associated with the amide linkage and avoid *trans-cis* isomerism, ether linkage variants were invented by utilizing diol-functionalized linkers. The resulting TFAs bear two cholate units and six maltosides as the tail and head groups, respectively, designated TFM-Es (Figure xx). Different lengths of the diol-functionalized linkers were used to vary the hydrophobic length of these agents, important for stabilization of individual membrane proteins with different hydrophobic widths (28–32 Å). TFM-C3E, containing a propylene spacer between the two cholate units, showed improved protein stabilization efficacy compared to the amide version (*i.e.*, TFM-C3Am) and TFA-1 when tested with a few membrane proteins (LeuT, MelB and β_2 AR) [M. Das, *ChemBioChem*, 2018]. Interestingly, LeuT with the shortest hydrophobic width (~29.8 Å) of the tested proteins was most effectively stabilized by the TFM with the shortest lipophilic group (*i.e.*, TFM-C0E), while β_2 AR with the longest hydrophobic width (~31.8 Å) retained its stability most effectively in the longest lipophilic group-bearing TFM (*i.e.*, TFM-C5E). This result indicates that the hydrophobic length of a detergent molecule can be adjusted according to the hydrophobic dimensions of any given membrane protein in order to confer enhanced protein stability. Detergent efficacy of the TFMs for protein stabilization may be improved by further increasing detergent hydrophobicity *via* linkage variation. In this context, the connection of the two cholate-based scaffolds *via* a thioether linkage rather than an ether linkage or their connection without the introduction of a functional group are possibilities. The physical properties of the carbohydrate-cored, stereo-isomeric and facial amphiphiles introduced above such as CMC and micelle size are summarized in Table 3.

6. Future directions and challenges for new detergent design

NG class detergents (OGNG and LMNG) and GDN have significantly contributed to membrane protein structure determination *via* both crystallization and cryo-EM. The recent popularity of GDN/digitonin for cryo-EM study indicates that detergents with enhanced *detergent-detergent* rather than *detergent-protein* interactions are likely to be favorable for protein structural study. Thus, new detergent design should focus on this detergent-detergent feature. So far, enhanced *detergent-detergent* interaction has been achieved by either increasing detergent alkyl chain number as in the cases of LMNG and carbohydrate-cored amphiphiles, increasing the hydrophobic surfaces and rigidity of detergent molecules (*e.g.*, GDN), or *via* stereo-chemical variations of the detergent core unit as observed with the BTMs, NBMs and CPMs. In contrast, the facial amphiphiles (FAs, LFAs, and TFAs) have increased *detergent-protein* interactions, as supported by a study of FA-3-MsbA PDCs *via* electrospray ionization mass spectrometry (MS) [S. C. Lee, et al, *PNAS*, 2013, E1203-E1211].

An increase in the alkyl chain number and the hydrophobic surface/rigidity of detergent hydrophobic group, and variations in the stereochemistry of the detergent core unit effectively enhance the hydrophobic interactions between detergent molecules within micelles or PDCs. However, the hydrophobic interaction is inherently weak compared to other intermolecular interactions and stronger intermolecular interactions between detergent molecules such as aromatic-aromatic and hydrogen-bonding interactions have the potential to further enhance protein stability. Some recent detergents contain structural features that allow aromatic-aromatic interactions, as exemplified by the terphenyl-cored trimaltosides (TPMs) and penta-phenylene maltoside (PPM) (Fig. 12a,b). These amphiphilic agents bear multiple aromatic rings (terphenyl and penta-phenylene groups, respectively) in the lipophilic region and spectroscopic studies of these molecules indicated the presence of *aromatic-aromatic* interaction between detergent molecules in their micelles. These detergents, however, gave an only marginal improvement on protein stabilization compared to DDM. The relative polar nature of the aromatic system (terphenyl/penta-phenylene) compared to cyclohexane counterparts and the short length of this hydrophobic group are likely responsible for the limited enhancement of protein stability. Hydrogen bonding interaction between detergent molecules is more attractive as this is the strongest non-covalent interaction. A hydrogen bond-mediated intermolecular interaction was incorporated into β -peptides (e.g. BP-1) (Fig. 12c) and, as expected, these peptide detergents were shown to confer enhanced stability to various membrane proteins. However, these peptide-based detergents are not well suited to large scale preparation and tend to form water-insoluble aggregates. Implanting a hydrogen bonding motif into a more conventional detergent architecture is attractive as it is likely to increase the detergent practical utility for protein structural study and such molecules are likely to be easier to produce in large scale. Considering the strength of hydrogen bonding compared to other intermolecular forces, a small number of hydrogen bonds between detergent molecules will make a dramatic difference to membrane protein stability.

In contrast to their wide use in protein structure determination *via* LCP/*in-surf*o crystallization and cryo-EM, LMNG and GDN have seldom been used for NMR-based protein structural study. We found only one example of a solution NMR protein structure (β_2 AR; 6KR8) determined using LMNG in the PDB. Another study showed that the quality of NMR spectra of human α_{1B} -adrenoceptor (α_{1B} -AR) in LMNG micelles was better than that of the receptor in nanodiscs (NDs). There are no reports of OGNG and GDN being used for NMR-based structural studies, suggesting a lack of compatibility with this method. The carbohydrate head groups of the detergent tend to increase PDC size by recruiting more water molecules into the surfaces of detergent micelles, making the PDCs too large for NMR analysis. As described above, zwitterionic MNG variants with an amine oxide or phosphocholine head group should form smaller PDCs and thus may be more suitable for NMR based studies. However, zwitterionic head groups are known to be unfavorable for membrane protein stability. So design of detergents that

incorporate such groups and maintain membrane protein in a stable state will be challenging. The GNGs, MNGs and GDN are not suitable for native MS, a powerful technique to identify multiple protein isomeric forms, to determine the exact position of post-translational modifications, and to investigate lipid-protein interactions]. The photo-cleavable detergent (Azo) and OGDs were shown to be useful for membrane protein study *via* native MS (Figure 5m). As only a limited number of novel detergents has been explored for this application, development of additional photo-cleavable or oligoglycerol-based detergents would facilitate MS-based membrane protein structural study. In this context, it would be interesting to introduce a photo-reactive unit or glycerol-based head group into the NG class detergents or the GDN scaffold.

Along with GDN/digitonin, protein/polymer-supported nanodiscs (NDs) are widely used for cryo-EM structural study. Proteins (e.g., MSP and saposin A) and polymeric materials such as SMAs and diisobutylene/maleic acid (DIBMA) are capable of formation of NDs containing a membrane protein and a patch of lipids when added to biological membranes. Very recently, small diglucoside amphiphiles, dodecyl diglucoside (DDDg) and C8-fluorocarbon diglucoside (F₈ODG), showed the ability to form NDs when mixed with *E. coli* membranes (Fig. 12d,e). This result suggests that ND formation is not an exclusive property of large amphiphilic molecules such as polymers or proteins. In this context, we recently noted unique behavior of MNG-2,18 with MelB_{st}; this detergent fully maintained the soluble state of MelB_{st} even at 65 °C, while the other A-MNGs including MNG-8,12 and LMNG yielded little amount of soluble protein under the same conditions. We hypothesize that the formation of NDs containing MelB is responsible for the exceptionally favorable behavior of this detergent. A preliminary study showed that, when treated with DMPC vesicles, this detergent was capable of forming NDs with a diameter of 10 to 40 nm depending on detergent concentration (unpublished). Interestingly, DDDg and MNG-2,18 both have long alkyl chains compared to conventional detergents, a structural feature that seems favorable for ND formation with a patch of lipid molecules; the long alkyl-chained detergents tend to interact with the fatty acid chain of phospholipids in preference to the membrane protein surfaces. In addition, both detergents possess a branched carbohydrate head group (glucose and maltose for DDDg and MNG-2,18, respectively), which likely facilitates formation of discoidal assemblies (*i.e.*, NDs). These bulky hydrophilic groups should be effective at introducing large curvature into the detergent assembly formed at the periphery of the NDs. Thus, detergents with a large head group and a long alkyl chain have potential as ND-forming agents. As structural modifications can be readily introduced into the MNG architecture, detergent optimization on the basis of MNG-2,18 for effective ND formation is feasible. Development of ND-forming detergents would further expand the utility of small amphiphiles in membrane protein research.