Micelle-Catalyzed Domain Swapping in the Cytoplasmic Domain of *E. coli* Rhomboid Protease

Jason Ka-Cheong Kwok

Thesis Submitted to the

Faculty of Graduate and Postdoctoral Studies

University of Ottawa

In partial fulfillment of the requirements for the

M.Sc. degree in the

Ottawa-Carleton Chemistry Institute

Abstract

Domain swapping is a mechanism of protein oligomerization whereby subunits exchange identical structural elements with each other. This type of interaction is important in a number of biologically important processes, including the regulation of enzyme activity, the modulation of molecular recognition and in the development of some protein deposition diseases. Domain swapping can be promoted by exposure to chemical denaturants, high temperatures, or redox reagents, although these factors can deviate greatly from conditions encountered *in vivo.* Moreover, in some cases these conditions can alter the monomer-oligomer equilibrium or even promote the formation of alternate domain-swapped oligomers. In contrast, we have found that it is possible for detergent micelles to be used as a catalyst in the domain-swapping interaction involving the N-terminal cytoplasmic domain of *E. coli* GlpG rhomboid protease. Our results show that hexadecylphosphocholine micelles can catalyze domain swapping of NGlpG by lowering the Gibbs free energy of the kinetic barrier by 12 kcal/mol, while preserving the equilibrium populations of monomer and dimer. Micelle charge and size were found to be important for this catalysis, which involved the formation of a partially unfolded micelle-bound intermediate containing significant secondary structure. Overall, the results from this work reveal how detergent micelles can affect the energy landscape of domain swapping for NGlpG, and provide insight into the potential ability of local irregularities in lipid membrane environments to play a role in domain swapping *in vivo*.

Acknowledgements

First and foremost, I would like to thank Dr. Natalie Goto for being an exceptional supervisor. Throughout my time in this lab, you provided endless guidance, advice, and feedback. Under your guidance, I developed my critical thinking and presentation skills. Also, with your help I've come to appreciate the importance of seeing not only the trees, but also the forest. I'm certain everything I've learned and experienced in these past few years will help me in the future.

I'd like to thank all past and present members of the Goto lab. I thank Dr. Allison Sherratt and Dr. Houman Ghasriani for paving the way for progress on the Rhomboid project. Thank you, Allison, for taking the time to provide valuable tips on how to acquire good quality kinetic data for domain swapping of NGlpG. I would also like to thank Alex Foo and Saud Ayed for being great colleagues. The lab techniques and clever tricks you guys showed me often made things easier. I'm also grateful for the help that Tabussom Qureshi and Laura McLeod both provided when I first started working in the lab. Tab, thanks for helping me acquire NMR data when I first started working in the lab. Laura, thanks for showing me how to use the CD spectropolarimeter to obtain high quality CD spectra. I would also like to thank Roxana Filip for assisting me with the backbone amide hydrogen deuterium exchange project. Besides the Goto lab, I thank the Keillor lab for the use of their CD spectropolarimeter, and the Scaiano and Moon labs for the use of their plate readers.

Last but not least, I would like to thank my parents and my sisters for their continual support and encouragement.

Statement of Contribution

All data presented in the thesis was acquired and analyzed by me, except for the van't Hoff analysis for domain swapping of NGlpG in the absence of Fos16 micelles, and dissociation constants and pseudo-first order rate constants for domain swapping in the absence and presence of LPPG, DDM, and urea, which were acquired by Dr. Allison Sherratt.

Table of Contents

List of Abbreviations

BCA	Bicinchoninic Acid
CD	Circular Dichroism
CI2	Chymotrypsin Inhibitor 2
Ckshs1	Human Cell Cycle Protein
CMC	Critical Micelle Concentration
CV-N	Cyanovirin-N
DDM	n-Dodecyl-β-D-maltoside
ecGlpG	*Escherichia coli* rhomboid protease
E. coli	*Escherichia coli*
Fos	Phosphocholine
Fos16	Hexadecylphosphocholine
FPLC	Fast Protein Liquid Chromatography
GlpG	Rhomboid Protease
hnNOE	Heteronuclear Nuclear Overhauser Effect
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani
LPPG	1-palmitoyl-2-hydroxy-sn-glycero-3-[phosphor-rac-(1-glycerol)]
MRE	Mean Residue Ellipticity
MW	Molecular weight
NGlpG	N-terminal cytoplasmic domain of *Escherichia coli* rhomboid protease
Ni-NTA	Nickel-Nitriloacetic acid
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
p13suc1	Suppressor of Cyclin Dependent Kinase 1
PF	Protection Factor
PrP	Prion Protein
SDS-PAGE	Sodium Dodecylsulphate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
T_m	Transition Midpoint
TROSY	Transverse Relaxation-Optimized Spectroscopy
2D ^{1}H-^{5}N HSQC	Two-dimensional ^{1}H-^{15}N Heteronuclear Single Quantum Coherence

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Domain Swapping

Domain swapping is an unusual mechanism of protein oligomerization whereby protein subunits exchange identical structural elements with each other (1). Each subunit of the oligomer, called a protomer, is structurally identical to the monomer, but is made up of more than one polypeptide. Swapped elements can range in size from a single alpha helix or beta strand to a globular domain, with the swapped element typically coming from the N- or C-terminus. Regions of the domain-swapped oligomer that connect protomers are referred to as hinge loops, and are often in a turn conformation in the monomer to allow intramolecular interactions with the rest of the polypeptide, but are in an extended state in the domain-swapped oligomer. While domain-swapping interactions are relatively rare, crystal structures of domain-swapped states have been determined for more than 60 proteins to date (1).

Although the structures of many domain-swapped proteins have been determined, questions have been raised as to whether this oligomerization mechanism is biologically relevant, as some domain-swapped states have been obtained under physiological conditions, while others can only be obtained by exposure to conditions that may not be biologically relevant. For example, the x-ray crystal structure of the domain-swapped dimer of cyanovirin-N was solved at low pH, yet cyanovirin-N would not be expected to encounter low pH conditions *in vivo* (2). In contrast, some proteins have been shown to undergo domain-swapping interactions *in vivo*, such as T7 helicase, phosphoenolpyruvate mutase, T4 endonuclease, and BS-RNase, suggesting that domain swapping interactions can have biological relevance (3–5). For example BS-RNase can form a domain-swapped dimer that, unlike the monomer, exhibits allosteric ribonuclease activity and is necessary for antitumour activity (3, 6, 7). In addition, there are also examples of proteins whose domain-swapped

forms have not been observed *in vivo,* but can still form domain-swapped oligomers in solution conditions that could be found in physiological systems. For example, RNase A can form a domain-swapped dimer at pH 6.5 and 37°C (3, 8), and its domain-swapped state exhibits higher enzyme activity on RNA than does its monomeric state (9). Similarly, diphtheria toxin forms a domain-swapped dimer at low pH that resembles the acidic conditions in endosomes, and has also been crystallized in its domain-swapped state in complex with its receptor (3, 10). Altogether, these examples show that domain swapping can occur *in vivo,* and in some cases can regulate enzyme activity or molecular recognition.

1.2 Domain Swapping and Deposition Diseases

Domain swapping is also implicated in the development of some protein deposition diseases that are caused by the formation of amyloid or non-amyloid fibrils (11). All amyloid fibrils comprise proteins with beta strands that stack perpendicularly to the fibril axis in an arrangement referred to as a cross-beta spine. Domain swapping has been proposed to be involved in the pathological self-assembly of amyloid fibrils by cystatin C, β2-microglobulin, prion protein (PrP). The cross-beta spines of amyloid fibrils formed from these domain swapping proteins comprise hinge loops that adopt beta strand conformations in the domain-swapped oligomer. Alternatively, some domain swapping interactions appear to play a role in the formation of non-amyloid polymers, such as the z-mutant of α1-antitrypsin that is associated with serpinopathies (12).

The first protein that was shown to form amyloid fibrils through domain swapping was cystatin C (Fig. 1.1). The crystal structure of its domain-swapped dimer was solved for the full length protein and for N-terminally truncated variants produced by limited proteolysis (13, 14). These structures showed that the hinge loops in both domain-swapped

dimers form beta sheets, and that the dimers could stack on top of each other such that hydrogen bonds formed between beta sheets of protomers from adjoining dimers (Fig. 1.1) (15). More evidence suggesting that domain swapping interactions are involved in cystatin C fibril formation was provided by experiments showing inhibition by disulphide bridged mutations that prevent domain swapping (16).

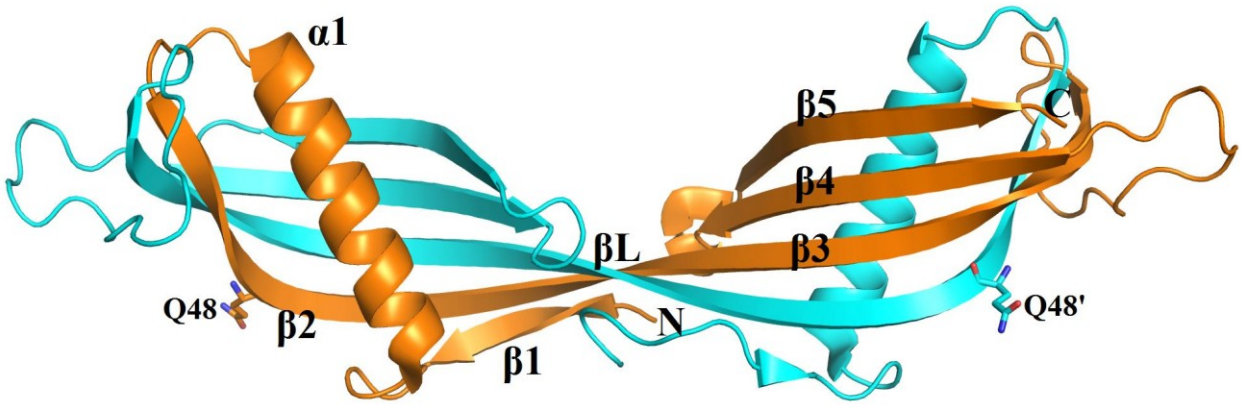


Figure 1.1: Crystal structure of the domain-swapped dimer of full-length human cystatin C (PDB ID 1TIJ) (15). One subunit is colored in orange, and the other is colored in blue. The secondary structural elements of the orange subunit are labeled. The swapped elements include α1, β1, and β2. The hinge loop is formed from loop βL, which connects β2 and β3. Amyloid fibrils have been proposed to form from stacking of cystatin C dimers, such that backbone amide hydrogen bonds form between β5 strands from different dimers and side chain hydrogen bonds form between Q48 (Q48 in the orange subunit and Q48' in the blue subunit) of β2 from different dimers (15).

β2-microglobulin (Fig. 1.2) and hamster PrP can each self-assemble into amyloid fibrils through domain swapping (17, 18). Interestingly, each of these proteins possesses a disulfide bond between its swapped element and the other polypeptide in the protomer, and domain swapping for both of these proteins has been proposed to be induced by redox processes. Both β2-microglobulin and hamster PrP are thought to form amyloid fibrils through open-ended domain swapping, with the hinge loops associating to form cross-beta-

spines. Reduction of the disulfide bond involving the swapped element converts the protein into an open intermediate state that subsequently undergoes re-oxidation to form either dimers or higher-order oligomers by open-ended domain swapping.

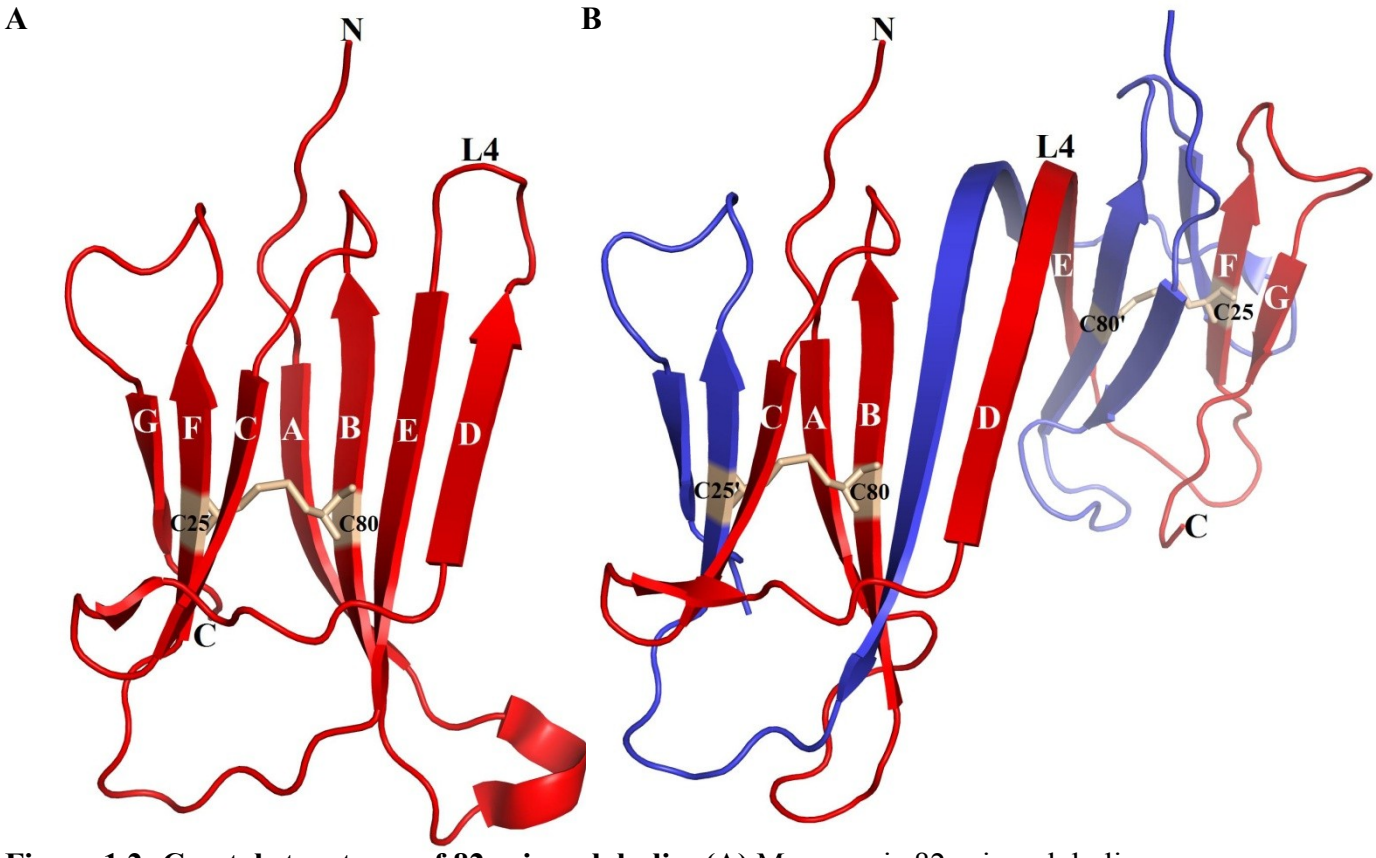


Figure 1.2: Crystal structures of β2-microglobulin. (A) Monomeric β2-microglobulin (PDB ID 1LDS) (19) possesses an intramolecular disulfide bridge (beige sticks) between C25 of βB and C80 of βF. Under reducing conditions, the disulfide bridge is disrupted, thereby freeing βF, along with βE and βG, for exchange with another subunit. **(B)** Domain-swapped dimer of β2-microglobulin (PDB ID 3LOW) (17). One subunit is colored in red, and the other is colored in blue. The swapped elements and the hinge loop of the red subunit are labeled in white. The swapped elements include βE, βF and βG, and the hinge loop comprises loop L4. Intermolecular disulfide bridges between C25 and C80' and between C25' and C80 are shown as beige sticks.

Given the roles of domain swapping in the regulation of protein function and the nucleation of some protein deposition diseases, there has been increasing interest in

identifying physiologically relevant conditions that can enhance the propensity for a protein to form domain-swapped oligomers. Linked to this is the potential impact of these conditions on the free energy landscape and mechanism of domain swapping, an issue that I seek to address in this thesis.

1.3 Factors Affecting Domain Swapping Propensities

To understand the propensity for a protein to undergo domain swapping, it is helpful to consider the relative free energies, and hence stabilities, of the monomer and the domain-swapped oligomer. The monomer and the domain-swapped oligomer are usually similar in their stabilities because both states possess nearly identical structures (3). From a thermodynamic perspective, domain swapping may be conducive to nucleation events that precede the formation of amyloid fibrils. The nucleation event involves the slow formation of a dimer, followed by growth of the fibril. Due to its thermodynamic stability, a domain-swapped dimer has the potential to serve as a long-lived nucleus for fibril growth (20).

The stability of a domain-swapped oligomer relative to that of the monomer is measured by the change in Gibbs free energy, $\Delta G°$, which depends on the enthalpy and entropy changes ($\Delta H°$ and $\Delta S°$, respectively) according to:

$$\Delta G° = \Delta H° - T\Delta S°$$

Since the interactions between the swapped element and the rest of the protein are identical in both the monomer and the domain-swapped oligomer, the exchange of the swapped element should not affect the enthalpy of the system. However, enthalpy changes can occur if novel interactions form between hinge loops or between protomers in the domain-swapped oligomer (3). The entropy is generally expected to decrease upon domain

swapping, due to the significant drop in the number of translational and rotational degrees of freedom resulting from the association of two polypeptide chains (3).

In addition to these fundamental considerations, results from a number of studies have shown that the structure of the monomer has significant influence on domain swapping propensity. Specifically, the hinge loop length, hinge loop sequence, and potential for cross-talk with residues outside of the hinge loop can influence the free energy of domain swapping (21). For example, shortening the hinge loop in some structures promotes domain swapping, since shorter hinge loops impose strain in the monomer that could be relieved by domain swapping, as was observed for p13suc1, ckshs2, and staphylococcal nuclease (21–24). In other cases, domain swapping can be promoted by lengthening the hinge loop. For example, the insertion of a polyglutamine sequence in the hinge loop of chymotrypsin inhibitor 2 (CI2) gave rise to domain-swapped dimers and higher-order oligomers (25). In this case, it was proposed that the insertion increased the flexibility of the hinge loop, thereby favoring domain swapping. This idea was also supported by experiments showing that polyalanine and polyglycine insertions could further increase the propensity for CI2 to domain swap.

Besides the hinge loop length, specific amino acid residues in the hinge loop can influence domain swapping propensity. In particular, proline can impart strain in either the monomer or the domain-swapped oligomer, depending on its position in the hinge loop relative to the other structural elements (22). For example, in the case of p13suc1, it was determined that P90 and P92, two proline residues in the hinge loop, imposed strain in the monomer and the domain-swapped dimer, respectively (22). Substitution of either proline with alanine was found to alleviate the strain, thereby shifting the equilibrium toward the

state with less strain. Additionally, domain swapping of p13suc1, Ckshs1, CV-N, and Protein L was promoted by substitution of their hinge loop residues with proline (21–23, 26, 27).

Domain swapping propensity may also be influenced by cross-talk between hinge loop residues and regions outside the hinge loop (21, 28). Since the hinge loop is the only region of the protein that changes in conformation during domain swapping, any remote influence on the relative stabilities of the monomer and domain-swapped oligomer should be communicated through the hinge loop. Based on this premise, Schymkowitz et al have proposed that the domain swapping propensity of p13suc1 depends on the sensing of hinge loop strain by residues elsewhere in the protein, as evidenced by a shift in the monomer-dimer equilibrium that resulted from alanine substitutions in the beta sheets flanking the hinge region (28). The mutated residues included Glu 86 and Ile 94, which are adjacent to the hinge loop, and also residues that contribute to the phosphate binding site, namely, Arg 39 and Arg 99 (Fig. 1.3). Also, phosphate binding was found to shift the equilibrium toward the monomer. To test whether the hinge loop had a role in sensing ligand binding and translating it into a change in the monomer-dimer equilibrium, the authors mutated proline residues in the hinge loop and measured the equilibrium constant for dimerization before and after the addition of ligand. For these mutants, no difference in the monomer-dimer equilibrium was observed before and after ligand binding, confirming that the hinge loop proline residues are indeed necessary for translating ligand binding at remote sites into a shift in the monomer-dimer equilibrium.

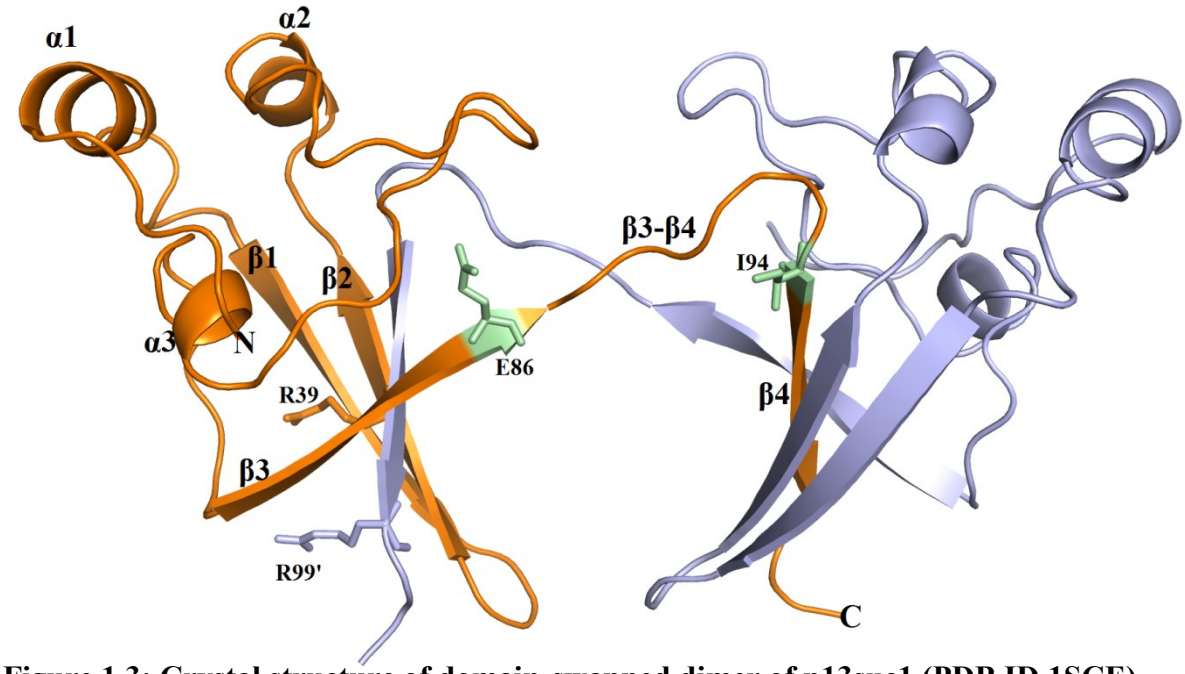


Figure 1.3: Crystal structure of domain-swapped dimer of p13suc1 (PDB ID 1SCE) (29). One subunit is colored in orange, and the other subunit is colored in blue. The secondary structure elements of the orange subunit are labeled. The swapped element comprises β4 and the hinge loop consists of β3-β4. E86 and I94, both shown as green sticks in the orange subunit, are proposed to sense hinge loop strain since mutation of either residue to alanine alters the monomer-dimer equilibrium. The phosphate binding site includes R39 and R99, which can also sense hinge loop strain. In the domain-swapped dimer, R39 (orange stick) and R99' (blue stick) come from different subunits. Binding of phosphate shifts the monomer-dimer equilibrium toward the monomer.

1.4 Domain Swapping Mechanisms

Although the structural similarity between the monomer versus the protomer in a domain-swapped oligomer gives rise to similar stabilities, several intramolecular interactions between the swapped element and the rest of monomer must be disrupted for exchange to occur; therefore, large activation energies are often associated with domain swapping. As such, it has been suggested for many domain-swapping interactions that the protein must first unfold before it can undergo domain swapping, although the extent of unfolding that is required seems to depend on the protein (21). In cases where the swapped element cannot fold independently, partial or complete unfolding of the protein must occur prior to

association, with this interaction being required for refolding to occur. Furthermore, the extent of unfolding required also depends on the location of the swapped element. Specifically, when the swapped element does not make significant contributions to the hydrophobic core, only partial unfolding is required, as was proposed for domain swapping by barnase, which swaps two alpha helices at its N-terminus (30). By contrast, if the swapped element forms an integral part of the protein hydrophobic core, then complete unfolding of the protein is thought to be required, and the association must occur between fully denatured polypeptides that concomitantly associate and refold. This mechanism for domain swapping was proposed for p13suc1 since its swapped element is a beta strand that forms part of a beta sheet located in the hydrophobic core of the folding nucleus (22). Complete unfolding was also proposed in the mechanism for domain swapping of the N-terminal domain of CD2 because its protomers are composed of beta sheets formed from highly intertwined intra- and intermolecular beta strands (31).

1.5 Domain Swapping Conditions

Since domain swapping often involves protein unfolding and refolding, denaturing conditions, or reagents that disrupt interactions between a swapped element and the rest of the protein, should facilitate the exchange of structural elements. For example, the formation of domain-swapped dimers of Fis1 and Stefin A could be facilitated with the use of denaturing conditions, such as high concentrations of chemical denaturants or high temperatures (32, 33). However, these denaturing conditions do not resemble physiological environments. Moreover, as amyloid and non-amyloid fibrils of domain-swapped proteins continue to be characterized, it is important to be aware of the conditions that are used to induce fibril formation because the use of non-physiological conditions could give rise to

alternate aggregates that do not actually exist *in vivo*. For example, the formation of amyloid fibrils of hamster PrP was induced by exposing samples to reducing agent, chemical denaturants, and salt, promoting the *in vitro* conversion of monomeric hamster PrP into an open intermediate. The slow removal of these additives is required for subsequent formation of higher order assemblies (18). Although the disulfide bond is important for the infectivity of PrP, no biological factors or conditions have been proposed to provide the redox environment necessary to promote the formation of the open intermediate (1, 18). Also, the high concentrations of guanidine hydrochloride and salt result in conditions that do not resemble those encountered in physiological systems. Therefore, it is still not known whether the PrP amyloid fibrils produced in this manner are related to the aggregates associated with Creutzfeldt-Jakob disease (1).

The potential problem of non-physiological domain swapping forced by denaturing *in vitro* conditions is perhaps best characterized in a study of the Z-mutant (G324L) of α1-antitrypsin that is associated with one of the serpinopathies (12). This mutant gives rise to a non-amyloid polymer that accumulates in hepatocytes, thereby causing liver cirrhosis. Studies support a model for polymerization involving open-ended domain swapping, where the reactive center loop from one subunit interacts with a β-sheet from another subunit (34). In 2008, Yamasaki and co-workers solved the crystal structure of a domain-swapped dimer of another serpin, antithrombin, and proposed a different model for the polymerization based on the structure (35). According to this model, polymerization still involves open-ended domain swapping, but the swapped element was proposed to be a beta hairpin consisting of the reactive center loop and β5 of β-sheet A, rather than just the reactive center loop itself. This model was later rejected based on the results from experiments using synthetic peptides, limited proteolysis, monoclonal antibodies, and ion mobility-mass spectrometry to

investigate the structure of the polymeric intermediate (12). It is believed that the use of guanidine to promote antithrombin polymerization resulted in the formation of the alternate polymer, which could not be recognized by the monoclonal antibody that binds only the pathological polymers of α1-antitrypsin. On the other hand, the polymer that was recognized by the antibody was consistent with the structure involving the swap of just the reactive center loop.

1.6 *E. coli* Rhomboid Protease and its N-terminal Cytoplasmic Domain (NGlpG)

Given the influence of local environment in triggering domain swapping events, particularly in disease-related proteins, it would be useful to identify conditions of physiological relevance that have the potential to induce this type of interactions. Not only would this improve understanding of mechanisms of pathological protein misfolding events, but it could also form the basis for the development of new strategies for inhibition of these interactions. My thesis has focussed on the ability of a membrane-mimetic detergent micelle to alter the free energy landscape for domain swapping. The system that I have used for these studies is the cytoplasmic domain from the *E. coli* GlpG rhomboid protease (ecGlpG).

Previously, the Goto lab discovered that the N-terminal cytoplasmic domain of *E. coli* rhomboid protease (NGlpG), an intramembrane serine protease, could form a domain-swapped dimer (36). Moreover, domain swapping of NGlpG was found to be catalyzed by hexadecylphosphocholine (Fos16) detergent micelles. Since detergent micelles can resemble local irregularities in lipid bilayers, they may have physiological relevance as factors that induce domain swapping *in vivo*. By studying the effect of Fos16 detergent micelles on the kinetics and thermodynamics of domain swapping of NGlpG, insights into the mechanism of action of similar biological factors on domain swapping can be obtained.

The domain studied in this thesis comes from the rhomboid family of intramembrane serine proteases, a group of proteins involved in a wide range of biological functions, including cell signalling, host cell invasion by apicomplexan parasites, and mitochondrial remodeling (37). These proteases each possess at least six transmembrane helices that form a hydrophilic cavity where the serine-histidine catalytic dyad is located (38). In addition to the transmembrane segments, many rhomboid proteases also possess extramembranous domains. In the case of the *E. coli* GlpG rhomboid protease, this domain is NGlpG, an N-terminal cytoplasmic domain composed of a three-stranded anti-parallel β sheet flanked by two α helices (Fig. 1.4) (39). The functional relevance of NGlpG remains unknown; its removal from full-length GlpG does not alter the ability of this rhomboid to cleave model substrates (36). However, it is not yet known whether NGlpG plays a role in *E. coli* GlpG cleavage of native substrates, since no biologically relevant target has yet been identified (40). Interestingly, the activity of the human rhomboid RHBDL2 on thrombomodulin has been shown to require the cytoplasmic domain (41), raising the possibility that NGlpG could play a similar role.

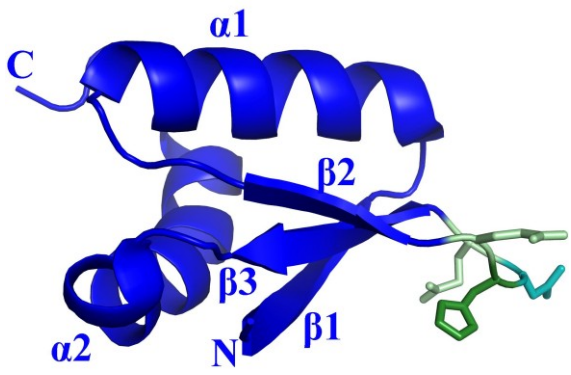


Figure 1.4: Solution NMR structure of the lowest energy conformer of monomeric NGlpG (PDB ID 2LEP) (39)**.** The swapped elements include α2 and β3, and the hinge loop comprises the β2-β3 turn. The hinge loop residues Q31 (light green), H32 (dark green), N33 (turquoise), and Q34 (light green) are shown as sticks.

During the preparation of samples for structure determination of the isolated NGlpG, the domain was found to elute from the size exclusion chromatography column in both monomeric and dimeric states (36). However, when either of the states was isolated and re-applied to the column at room temperature, no conversion to the other state was observed, suggesting that these two states were kinetically trapped. Consistent with this, it was found that the isolated monomer or dimer would convert to the other state if exposed to elevated temperatures. By measuring rate constants for exchange at several temperatures in the 45°C-60°C range, the activation energy for dimerization was found to be ~80 kcal/mol.. The large activation energy suggests that the NGlpG dimer is formed from a domain-swapping interaction.

To explore this possibility, 2D ^{1}H-^{15}N HSQC spectra were acquired on monomeric and dimeric NGlpG and average amide chemical shifts differences calculated for each residue (36). The largest chemical shift changes were found in residues H32, Q34, S35, and D36 localizing to a turn between the second and third beta strand (Fig. 1.4). No other residue showed a significant change in chemical shift, indicating that the local chemical environment of the monomer and dimer was identical for all residues outside this turn. This pattern of chemical shift changes was also highly consistent with a domain-swapping interaction. To confirm the domain-swap, the dimer interface of NGlpG was characterized by recording a NOESY experiment selective for intermolecular contacts (42). This experiment showed intermolecular NOE signals involving α2 and β3, suggesting that these secondary structure elements comprise the swapped element, with the β2-β3 turn serving as the hinge loop.

To obtain atomic resolution detail to this domain-swapped dimer, the solution NMR structure of the dimer was determined (36). However, prior to final refinement of this structure an x-ray crystal structure of dimeric NGlpG became available, showing good

agreement with our calculations (Fig. 1.5). In the case of the crystal structure, the region between β2 and β3 was resolved in a β-strand conformation, forming novel intermolecular and intramolecular hydrogen bonds that are not present in the monomer. In the case of the NMR structure, only very low β-strand propensity was predicted for backbone secondary shifts, suggesting that this additional β-strand structure may not be very stable in solution. The extended nature of the structure in this region made it difficult to precisely define its conformation from NOEs. Consequently, it is not known whether the hinge loops are structured in solution.

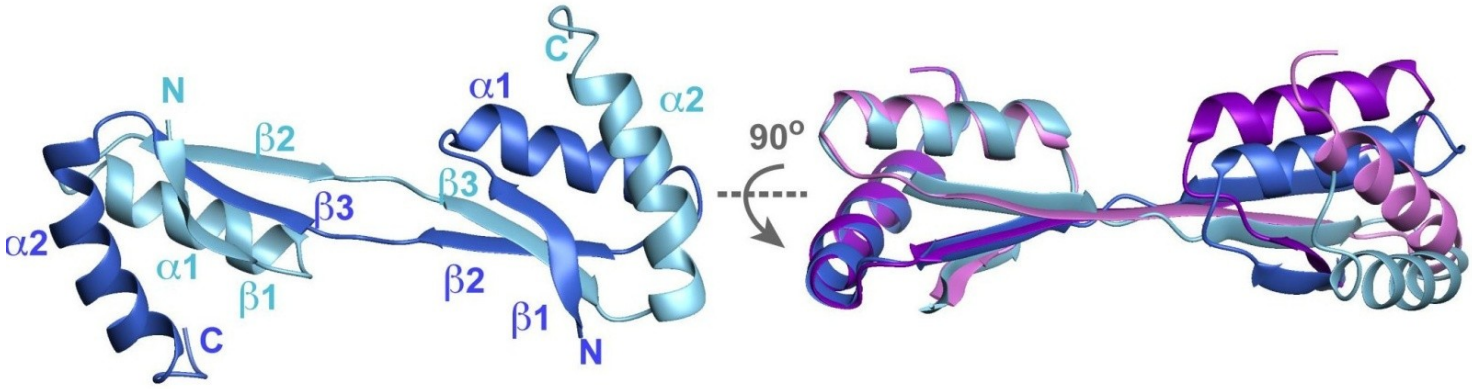


Figure 1.5: Dimer structure of NGlpG. Lowest energy conformer of the solution NMR structure of dimeric NGlpG shown on the left, with one subunit colored in dark blue and the other subunit colored in light blue. Superposition of the solution NMR structure with the x-ray structure (PDB ID 4HDD) (40) of dimeric NGlpG (purple) shown on the right.

An understanding of the structure, dynamics, and interactions of the hinge loops will provide clues as to what factors might affect the relative stabilities of monomeric and dimeric NGlpG, and is therefore one of the objectives of this thesis. In particular, heteronuclear $\{^1H\}$-^{15}N NOEs can be used to assess the conformational flexibility of the hinge loops relative to the rest of the structure (43). In addition, backbone amide hydrogen exchange experiments can be done to provide a measure of the stability of backbone hydrogen bonds

(44), and to determine whether any differences exist between the domain-swapped and monomeric states. Regions of the protein that show changes in hydrogen bond stability outside of the hinge loop region will help to account for the observed propensity for domain swapping by NGlpG.

One of the most unique aspects of NGlpG domain swapping was the observation that rates of interconversion between monomer and dimer could be significantly accelerated by the addition of detergent micelles (36). In particular, NMR spectroscopy and SEC experiments both showed that the micellar state of hexadecylphosphocholine (Fos16), which has a critical micelle concentration (CMC) of 13 μM as reported by Anatrace, could accelerate domain swapping without changing the apparent affinity for dimer formation. This is the first known example of micelle-catalyzed domain swapping, with no other demonstrations of catalysis for domain swapping yet described. However, several questions still remained from this study. Specifically the kinetics and thermodynamics of the micelle-accelerated domain swapping of NGlpG had not been quantitatively measured. In addition, the properties of the micelle (e.g. size, charge) that are important for this phenomenon had not been well characterized. This information should provide insight into the mechanism of micelle-catalyzed domain swapping, and potential physiological factors that could induce domain swapping. Therefore, one of the goals of this thesis is to address these questions, providing a more detailed mechanistic description of micelle-catalyzed domain swapping for NGlpG.

1.7 Thesis Goals

NGlpG undergoes a domain-swapping interaction that can be catalyzed by Fos16 micelles. Critical to our understanding of this process is the solution state structure of the

dimer, and information about the energetics, physical requirements and intermediates involved in this process. To address these issues, my thesis has focussed on these experimental objectives:

- Obtain NMR data required to refine the solution NMR structure of dimeric NGlpG.
- Quantitate the amount of energy that is lost from the activation barrier to domain swapping when the reaction is catalyzed by Fos16 micelles
- Determine the properties of the micelles that are important for the acceleration of domain swapping.
- Characterize the structural properties of the micelle-bound intermediate formed during micelle-accelerated domain swapping.

The results of this study are expected to provide insights into how detergent micelles alter folding energy landscapes to accelerate domain swapping. Since detergent micelles resemble certain lipid membrane environments, such as local lipid irregularities or phase boundaries of lipid rafts, the micelle properties that accelerate domain swapping may have physiological relevance. Also, an understanding of the mechanism by which detergent micelles accelerate domain swapping might provide insights into how lipid-like environments might stabilize intermediates that form domain-swapped dimers or oligomers of higher order assemblies.

Since this work involves the use of solution NMR and CD to characterize the structures of the dimeric and micelle-bound intermediate state of NGlpG, some of the basic theory behind these biophysical techniques will be introduced in the following section.

1.8 Theoretical Background for NMR and CD Experiments

1.8.1 Investigating Protein Dynamics on the Picosecond to Nanosecond Timescale by Solution NMR

Protein dynamics can be classified by the timescales within which particular protein motions occur (43). On the picosecond to nanosecond timescale, side chain motions, loop movements, and molecular rotation are most relevant. Nuclear spin relaxation measurements on this timescale can therefore be used to evaluate the flexibility of loop regions relative to the rest of the protein. In this work, one of these types of measurements, namely the heteronuclear nuclear Overhauser enhancement (hnNOE) was performed on monomeric and dimeric NGlpG to study the flexibility of the hinge loop in NGlpG relative to the rest of the protein. The $^{15}N\{^{1}H\}$ NOE arises from magnetization transfer from a ^{1}H to ^{15}N nucleus due to dipole-dipole interactions (43). The transfer of magnetization depends on the distance between the spins and the extent of motion experienced by the spins (45). In the case of an amide proton that is directly bonded to an ^{15}N atom, the distance between the two spins is fixed at a distance that allows strong dipole-dipole interactions between them. The efficiency of magnetization transfer that gives rise to the hnNOE can be modulated by motions of the N-H bond if they are faster than the timescale of molecular reorientation rates; in the case of proteins in aqueous solutions, this is the picosecond to nanosecond timescale. Internal motions that occur on these timescales can give rise to large and negative steady-state hnNOE values, while smaller extents of internal motion or loop flexibility give rise to positive hnNOE values.

Experimentally, steady-state hnNOE values are obtained by measuring NMR signal intensities in a ^{1}H-^{15}N HSQC spectrum acquired after irradiation of protons and taken as a

ratio to those from a control experiment run without irradiation (45). For proteins that possess little motion on the nanosecond to picosecond timescale, magnetization transfer is less efficient and the ratio is close to 1. For amides that possess more motion on this timescale, magnetization transfer is more efficient, and the resulting peak intensities and ratios are decreased, with the maximal effect observed for protein motions that are faster than 0.1 ns.

1.8.2 Hydrogen Exchange Experiments

NMR spectroscopy was also employed to measure rates of backbone amide hydrogen exchange with solvent deuterium (44). Hydrogen exchange occurs across timescales ranging from milliseconds to hours and weeks (46). Hydrogen exchange experiments can be performed to determine the stabilities of backbone amide hydrogen bonds. Solvent-exposed hydrogen bonds are generally less stable and therefore more likely to be disrupted, giving rise to a state that would allow exchange with solvent deuterium. By contrast, hydrogen bonds in parts of the protein that are buried in the hydrophobic core are generally less susceptible to disruptions, being stabilized by multiple interactions. Nonetheless, local conformational fluctuations will still be possible that temporarily disrupt the hydrogen bond and allow exchange with the solvent such that:

$$NH_{closed} \underset{k_{close}}{\overset{k_{open}}{\rightleftharpoons}} NH_{Open} \xrightarrow{k_{ex}} ND$$

where NH_{closed} is an amide that is engaged in a hydrogen bond, NH_{open} is an amide whose hydrogen bond has been disrupted, and ND is the amide after its hydrogen is exchanged with a deuteron from the D_2O solvent (44). k_{open} is the rate constant for local fluctuations that

disrupt the amide hydrogen bond, k_{close} is the rate constant for local fluctuations that allow the hydrogen bond to re-form, and k_{ex} is the rate constant for hydrogen exchange with the solvent. The propensity for local opening fluctuations that disrupt amide hydrogen bonds is given by the equilibrium constant K_o , where

$$K_o = \frac{k_{open}}{k_{close}}$$

In practice, the hydrogen exchange rate is characterized by an observed rate constant, k_{obs}, which is measured experimentally (44). The extent to which an amide hydrogen is protected from exchange is given by the protection, PF, which is obtained by dividing k_{obs} by the intrinsic rate constant,k_{int} , as follows:

$$PF = \frac{k_{obs}}{k_{int}}$$

k_{int} is the hydrogen exchange rate constant for an amide hydrogen that is completely exposed to the solvent and is not engaged in a hydrogen bond. k_{int} is sensitive to local sequence context, but can be can be calculated with an online tool called SPHERE (http://www.fccc.edu/research/labs/roder/sphere/). An amide hydrogen with a higher protection factor is more protected from exchange than an amide hydrogen with a lower protection factor.

If an amide is mainly engaged in a hydrogen bond at equilibrium, then k_{obs} is related to k_{open}, k_{close}, and k_{ex}, as follows (44):

$$k_{obs} = \frac{k_{open}k_{ex}}{(k_{close} + k_{ex})}$$

Depending on the magnitude of k_{close} versus that of k_{ex}, hydrogen exchange can be classified in one of two regimes (44). One regime is termed EX2, which is typically encountered when

the pH is less than 8. Under this condition, the hydrogen exchange rate is pH-dependent, and k_{close} is greater than k_{ex}, allowing the observed rate constant to be simplified as.

$$k_{obs} = \frac{k_{open}k_{ex}}{k_{close}} = K_o k_{ex}$$

In this case, K_o is the inverse of PF. That is,

$$PF = \frac{1}{K_o}$$

The other regime is termed EX1, which is encountered at high pH. Under this condition, the hydrogen exchange rate is pH independent, and k_{ex} is greater than $k_{close.}$, such that the opening and closing fluctuations limit the rate of hydrogen exchange.

$$k_{obs} = k_{open}$$

In this work, the EX2 regime will be used to study hydrogen exchange, as the pH of the buffer will be less than 8.

Hydrogen exchange rate constants can be obtained by acquiring a series of ^{1}H-^{15}N HSQC spectra on a protein just after introduction into a deuterated buffer. As each residue exchanges its amide hydrogen with a deuteron from the D_2O solvent, the intensity of its corresponding peak in the HSQC spectrum will decrease, since this NMR experiment does not detect ^{2}H. For each residue that can be seen in the HSQC spectrum, the peak intensity reduction as a function of time can be fit to a first order exponential decay to give the exchange constant. These are then used to calculate protection factors, allowing the identification of regions of high and low local conformational stability.

1.8.3 Circular Dichroism (CD) Spectroscopy

CD spectroscopy is a type of absorption spectroscopy useful for characterizing the secondary or tertiary structure of proteins (47). CD is the differential absorption of left and right circularly polarized light by chiral molecules, and is measured in units of ellipticity. For proteins, the arrangements of peptide bonds in secondary structures gives rise to chirality (48). To facilitate the comparison of CD results, the ellipticity is often converted to units of mean residue ellipticity (MRE), which normalizes the ellipticity by peptide bond concentration, and represents the average ellipticity for any residue within the protein sequence.

In this work, CD spectroscopy was used to estimate secondary structure content for monomeric and dimeric NGlpG. Each class of secondary structure gives rise to a characteristic CD spectrum. Specifically, an alpha helix gives rise to a spectrum with characteristic minima at 222 nm and 208 nm, while the spectrum of an anti-parallel beta sheet has a minimum at 218 nm and a maximum at 195nm (48). This allows secondary structure content to be estimated through the use of deconvolution algorithms with a reference library of spectra from proteins of known structure (48, 49). Typically, a linear combination of spectra from these proteins from the reference library is used to reproduce the experimental spectrum. The success of these methods tends to improve when reference libraries contain a large database of proteins, ideally from a similar structure class to the protein being analyzed (50).

In addition to secondary structure determination, CD spectroscopy can also be used to assess the thermal stability of a protein, which is quantified by the transition midpoint of the melt, T_m, and the enthalpy change for unfolding, ΔH_m (51). This type of analysis can be done on a protein of interest by monitoring the evolution of its ellipticity at a single wavelength as

a function of temperature. In the simple case of a monomer with only one melt transition, the thermal unfolding curve consists of a pre-transition baseline, a transition with an inflection point, and a post-transition baseline (51). The inflection point represents the transition midpoint of the melt, T_m, and the width of the transition is inversely related to ΔH_m. Since the entropy change, ΔS_m, at the transition midpoint is the ratio of ΔH_m to T_m, its magnitude can be inferred by comparing the width of the transition to how far the inflection point lies along the temperature axis (51, 52).

Chapter 2: Materials and Methods

2.1 NGlpG Expression

E. coli BL21(DE3) cells (Novagen) were transformed with a pET-30a vector encoding NGlpG (residues 1-61 of ecGlpG) followed by a Leu-Glu sequence and a C-terminal hexahistadine tag (36), as previously described (53). A single transformed colony was inoculated into 100 mL of LB media or M9 minimal media (47.9 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 0.1% (w/v) D-glucose, 100 µM $CaCl_2$) containing NH_4Cl, 0.01% (v/v) MEM vitamin (Gibco) or LB media, and 50 µg/mL kanamycin. The culture was incubated overnight at 37°C with shaking at 220 rpm. Following overnight incubation, the culture was added to 1 L of fresh LB media or M9 minimal media containing ^{15}N-labelled NH_4Cl, 0.01% (v/v) MEM vitamin (Gibco) or LB media, and 50 µg/mL kanamycin, and incubated again at 37°C and 220 rpm. To monitor culture growth, the optical density of the culture was measured at 600 nm using an Ultrospec 2100 pro UV/Visible (Biochrom). The culture was grown to an optical density of 0.5-0.7, and then NGlpG expression was induced by addition of 1 mL of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture. The culture was incubated for another 4 hours at 37°C, and then cells were harvested by centrifugation for 10 minutes at 4°C and 4,800xg using an Avanti J-E centrifuge (Beckman Coulter). Following centrifugation, cell pellets were removed and stored at -20°C until purification. Prior to induction of NGlpG expression, a 500 µl aliquot was removed for expression analysis by sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE). The aliquot was centrifuged for 2 minutes at 16,000xg using a 5415D Centrifuge (Eppendorf) to pellet the cells. The cell pellets were then resuspended in SDS-PAGE loading buffer (50 mM Tris-Cl at pH 6.8, 10% (w/v) glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue, and 2% (w/v) β-mercaptoethanol).

2.2 Purification of NGlpG

Purification of NGlpG was performed as previously described (36). Briefly, a cell pellet from 1 L of cell culture was thawed on ice and resuspended in 25 mL of harvest buffer (25 mM phosphate buffer at pH 6.5, 150 mM NaCl, and 10 mM imidazole). One EDTA-free protease inhibitor tablet (Roche) containing 4-(2-aminoethyl)-benzenesulfonyl fluoride was dissolved in the suspension. The suspension was sonicated with 1 second on and 1 second off pulses for 1 minute at 50% power on ice using a Sonic Dismembranator Model 500 (Fisher Scientific). The lysate was incubated on ice for 10 minutes, followed by a second round of sonication. Cell debris were removed from the soluble fraction by centrifugation for 20 minutes at 4°C and 16,000xg using an Avanti J-E centrifuge (Beckman Coulter). The soluble fraction was applied to a column containing 3 mL of nickel-nitrilotriacetic acid (Ni-NTA) resin (Novagen) that had been equilibrated with harvest buffer. The flow-through was collected and re-applied to the column to maximize the yield of NGlpG. The column was washed with 20 mL of harvest buffer, and then NGlpG was eluted with 3 x 10 mL of elution buffer (25 mM phosphate buffer at pH 6.5, 150 mM NaCl, and 250 mM imidazole). 50 µL aliquots were taken from the flow-through, wash, and elution fractions and mixed with 50 µL of SDS-PAGE loading buffer.

Nickel-affinity purified NGlpG was concentrated to ~1.5 mL in an Amicon Ultra centrifugal filter device (Millipore) with a 3 kDa molecular weight cut-off (Millipore) by centrifugation at 4^{o}C and 4,000xg using a Spinchron 15R centrifuge (Beckman Coulter). The concentrated sample was subjected to size exclusion chromatography (SEC) using a Superdex-75 10/300 GL column on an AKTA FPLC (GE Healthcare). NGlpG was eluted in filtered and degassed FPLC buffer (25 mM phosphate buffer, 150 mM NaCl, 100 µM EDTA,

pH 6.5) at a flow rate of 0.5 mL/min, and fractions corresponding to peaks representing monomeric or dimeric NGlpG were collected and pooled.

2.3 SDS-PAGE

SDS-PAGE analysis was performed to verify expression of NGlpG and to assess its purity following purification. Polyacrylamide gels were cast with a 5% stacking phase (5% (w/v) 37.5:1 Acrylamide/Bis-acrylamide, 0.12 M Tris pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.1% (v/v) TEMED) and a 15% resolving phase (15% (w/v) 37.5:1 Acrylamide/Bis-acrylamide, 0.38 M Tris pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.04% (v/v) TEMED). The gel was placed in a Mini-PROTEAN electrophoresis cell (Biorad), to which Tris-Glycine SDS running buffer (25 mM Tris, 190 mM glycine, and 3.5 mM SDS) was added. Samples in SDS loading buffer were heated at 100°C for 5 minutes, and then 10 μL of each sample was loaded. 2.5 μL of an EZ-Run Prestained Rec Protein Ladder (Fisher) was also run to allow determination of sample molecular weights. Gels were run at 180 V for 55 minutes, stained with Coomassie blue stain (50% (v/v) MeOH, 0.1% (w/v) Coomassie brilliant blue, 10% (v/v) acetic acid), and destained with destaining solution (40% (v/v) MeOH, 10% (v/v) acetic acid).

2.4 Determination of Protein Concentration

The protein concentrations of samples were determined using the BCA Protein Assay (Thermo). A series of standards was prepared from a 2 mg/mL stock of BSA (Thermo). 50 μL of each standard or sample was mixed with 1 mL of BCA reagent containing reagent A (200 mM sodium carbonate, 120 mM sodium bicarbonate, 30 mM bicinchoninic acid, 10

mM sodium tartarate , and 0.1 M sodium hydroxide) (Thermo) and reagent B (4% cupric sulfate) in a 50:1 (v/v) ratio (Thermo). Samples were incubated at 37°C for 30 minutes, and the absorbance at 562 nm measured at room temperature using an Ultrospec 2100 pro UV/Visible (Biochrom).

2.5 Circular Dichroism Spectroscopy

All CD spectroscopy experiments were performed on a Jasco J-815 CD instrument equipped with a PFD-425S Peltier temperature controller. All samples were placed in a 0.1 cm quartz cuvette (Hellma). Prior to each set of experiments, the cuvette was cleaned by soaking it in nitric acid overnight, and extensively washing with water the following morning. The cuvette was then rinsed with 70% (v/v) ethanol and dried under a stream of argon gas. The detergent Fos16 (Affymetrix) was dissolved in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, pH 6.5 on the same day that CD spectra were acquired.

Observed ellipticities were converted into units of mean residue elliptcity (MRE) according to:

$$\Theta_{MRE} = \frac{\Theta_{obs}}{NCl} \qquad \text{[eq.1]}$$

where Θ_{MRE} is the mean residue ellipticity, Θ_{obs} is the observed ellipticity, N is the number of backbone peptide bonds in the protein, C is the concentration of the protein in dmol/cm^3, and l is the path length of the cuvette in centimeters (48).

A ~50 μM stock of monomeric NGlpG was prepared, its concentration verified by BCA assay, and stored at 4^{o}C until the day CD spectroscopy was performed. For each set of experiments, CD spectra were acquired on:

1) 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, pH 6.5

2) 20 mM Fos16 in the same buffer as 1).

3) 5-7 μM NGlpG in the same buffer as 1).

4) 5-7 μM NGlpG in the same solution as 2).

The NGlpG concentration was kept at ~5 - 7 μM to minimize the tendency for dimerization during the experiment. (The extent of dimerization should be <1 % given a K_d = 0.4 mM). Spectra were acquired at 45^{o}C from 200 nm to 250 nm, with a scan speed of 20 nm/minute, a data pitch of 0.5 nm, a response time of 8 seconds, a bandwidth of 1.00 nm, and 8 accumulations.

To estimate the secondary structure content of NGlpG in the presence or absence of Fos16, the spectra were baseline-corrected and analyzed using CDPro (49), which includes three methods for secondary structure calculation: SELCON3, CDSSTR, and CONTIN (49, 50). Secondary structure estimation of NGlpG in the presence/absence of Fos16 was performed using all three methods and a reference set of 43 soluble proteins (SP43) (50). The output of each calculation was provided as percentages of six forms of secondary structure: regular helix, distorted helix, regular strand, distorted strand, turns, and unordered. For each output, the regular and distorted forms of each of the helix and strand classes of secondary structure were summed and outputs from the three calculations averaged.

2.6 Thermal Denaturation Experiments

Thermal denaturation experiments on monomeric NGlpG were monitored by CD. Before each experiment, a CD spectrum was acquired at room temperature as described in the previous section. Thermal unfolding of monomeric NGlpG was monitored from 20^{o}C to 100^{o}C at 222 nm, with a ramp rate of 1.0^{o}C/minute (in the absence of Fos16) or 2.0^{o}C/minute

(in the presence of Fos16), a sampling temperature of 0.2°C, an equilibration time of 0.2 minutes, a response time of 0.25 seconds, and a bandwidth of 1.00 nm. Refolding was then carried out from 100°C to 20°C using the same parameters that had been used for the unfolding experiment. Unfolding and refolding curves were found to overlap, confirming reversibility. Reversibility was also confirmed by comparing CD spectra of NGlpG before and after the cycle of thermal denaturation and refolding.

The thermal unfolding transition usually consists of a pre-transition baseline, a transition with an inflection point, and a post-transition baseline (51). To determine the transition midpoint of a melt (T_m), and the enthalpy change for thermal unfolding (ΔH_m), the thermal curve can be fit to the following equation:

$$\Theta_{obs} = \frac{(y_n + m_n T) + (y_d + m_d T)K}{1+K} \qquad \text{[eq. 2]}$$

where Θ_{obs} is the ellipticity at any temperature T, y_n and y_d are the y-intercepts of the native and unfolded states, respectively; m_n and m_d are the slopes of the pre- and post-melt transitions; and K is the equilibrium constant for the transition at temperature T, determined from the van't Hoff equation (54, 55):

$$K = \exp\left[\frac{\Delta H_m}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right)\right] \qquad \text{[eq. 3]}$$

where R is the gas constant.

Since the fit depends on parameters of both pre- and post-transition baselines, these equations cannot be used to determine T_m and ΔH_m from thermal unfolding curves that are missing either baseline. However in these cases, the T_m and ΔH_m for unfolding can still be determined by fitting the differential form of the thermal curve, as the equations used for the fit do not require baseline parameters (52). In this method, the average slope of the thermal unfolding curve ($\frac{d(\Theta_{obs})}{dT}$) can be fit to:

$$\frac{d(\Theta_{obs})}{dT} = Af(1-f)T^2 \qquad \text{[eq. 4]}$$

where A is a correction factor, and f is the fraction of denatured protein at temperature T calculated using (52):

$$f = \frac{K}{K+1} \qquad \text{[eq. 5]}$$

The average slope of the thermal unfolding curve was taken over 2 degree intervals.

2.7 Domain Swapping Rate Measurements

Monomeric NGlpG was prepared as a 0.26 mM stock and divided into 100 uL aliquots. The following detergents were tested for their effects on the rates of domain swapping: dodecyl maltoside (DDM), and decyl-, dodecyl- tridecyl- and hexadecyl-phosphocholines (all from Affymetrix) (Table A.1). All detergents were dissolved in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, pH 6.5 on the day the exchange assay was initiated. 5 μL of the detergent stock was mixed with each 100 μL aliquot of monomeric NGlpG so that the final protein concentration was 0.25 mM. Aliquots were incubated at 45^{o}C in a Mastercyler Personal thermal cycler (Eppendorf), and an aliquot was removed at various time points and placed on ice for ~5 minutes to quench the monomer-dimer interconversion. Subsequently, 100 μL of the quenched aliquot was subjected to SEC at room temperature using a Superdex-75 10/300 GL column on an AKTA FPLC (GE Healthcare), and the peaks on the SEC profile integrated using Unicorn FPLC software to determine the relative populations of monomer and dimer.

The relative populations of monomer and dimer were plotted as a function of time,

and the interconversion was found to follow first-order kinetics for a reversible system consisting of monomer (M) and dimer (D) states (36):

$$M + M \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} D$$

Starting from the monomeric state, the loss of monomer over time could be described by the following integrated rate equation:

$$f^M = f_{eq}^M + \left(f_0^M - f_{eq}^M\right)e^{-k_{ex}t} \quad \text{[eq. 6]}$$

where f^M is the fraction of monomeric NGlpG at any time t, f_{eq}^M is the fraction of monomeric NGlpG at equilibrium, f_0^M is the initial fraction of monomeric NGlpG, and k_{ex} is an observed rate constant for monomer-dimer exchange and is the sum of the forward and rate constants, as follows:

$$k_{ex} = 2k_1 + k_{-1} \quad \text{[eq. 7]}$$

where k_1 and k_{-1} are first order rate constants for dimerization and dimer dissociation, respectively.

The forward and reverse rate constants could be determined by combining eq. 7 with the equilibrium-constant expression for dimer dissociation,

$$K_d = \frac{k_{-1}}{k_1} \quad \text{[eq. 8]}$$

where K_d is the equilibrium constant for dimer dissociation.

For all detergents, $k_{ex}$was measured over a range of micelle concentrations, and a linear relation was established between k_{ex} and micelle concentration, as follows:

$$k_{ex=}k_2[micelle] \quad \text{[eq. 9]}$$

where k_2 is a second order rate constant for the reversible monomer-dimer exchange and is equal to the slope of the linear relation.

k_2 is the sum of second order forward rate constant, $k_{2,f}$, and the second order reverse rate constant, $k_{2,r}$ for dimerization and dimer dissociation, respectively.

$$k_2 = k_{2,f} + k_{2,r} \quad \text{[eq. 10]}$$

2.8 Thermodynamics of Fos16-Catlayzed Domain Swapping

The activation enthalpy and activation entropy for Fos16-accelerated dimer association were determined by fitting k_2 for the forward or reverse reactions obtained at 21°C, 30°C, 35°C, and 45°C to the Eyring equation:

$$\ln\left(\frac{k_2}{T}\right) = -\frac{\Delta\ddagger H^\circ}{RT} + \frac{\Delta\ddagger S^\circ}{R} + \frac{k_B}{h} \quad \text{[eq. 11]}$$

where k_2 is the second order rate constant for domain swapping, T is the temperature, $\Delta\ddagger H^\circ$ is the activation enthalpy, $\Delta\ddagger S^\circ$ is the activation entropy, R is the gas constant, k_B is Boltzmann's constant, and h is Planck's constant. The activation enthalpy and entropy were used to calculate the activation Gibbs free energy, using the Gibbs free energy equation:

$$\Delta\ddagger G^\circ = \Delta\ddagger H^\circ - T\Delta\ddagger S^\circ \quad \text{[eq. 12]}$$

Similarly, the enthalpy and entropy changes for Fos16-catalyzed domain swapping were determined by fitting the dissociation constants, K_d, obtained at 21°C, 30°C, 35°C, and 45°C to the van't Hoff equation:

$$lnK_d = -\frac{\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R} \quad \text{[eq. 13]}$$

where ΔH° and ΔS° are the enthalpy and entropy changes for dimer dissociation.

The enthalpy and entropy changes were used to determine the change in Gibbs free energy, ΔG°, as follows:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad \text{[eq. 14]}$$

2.9 NMR Spectroscopy

All NMR experiments were performed on a Varian Inova 500 spectrometer equipped with a triple resonance probe. All spectra were processed using NMRPipe (56) and analyzed in NMRViewJ (57)

2.10 Backbone Amide Hydrogen-Deuterium Exchange

$^{2}H/^{1}H$ exchange experiments were initiated by buffer exchange of a sample of NGlpG in a monomeric or dimeric state. A 500 μL sample of 1-2 mM ^{15}N-labeled NGlpG in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, pH 6.5 was applied to a NAP-5 column (GE Healthcare). The sample was then eluted in 1 mL of the same buffer made in 99.9% D_2O, such that the final protein concentration was 0.5-1mM.

Acquisition of ^{1}H-^{15}N HSQC spectra started approximately 10-15 minutes after buffer exchange was completed. The spectra were acquired at 298 K, with 11 (3) experiments 45 minutes in duration, followed by 6 (13) experiments 90 minutes in duration for the monomer (dimer). The decay of integrated peak volumes was fit using non-linear least-squares fitting to extract proton exchange constants. Amide proton protection factors were calculated by dividing intrinsic exchange rate constants estimated by the program SPHERE (http://www.fccc.edu/research/labs/roder/sphere/) with experimentally determined proton exchange constants (58, 59).

2.11 Heteronuclear steady-state $\{^{1}H\}^{15}N$ NOE

Heteronuclear steady-state $\{^{1}H\}^{15}N$ NOEs for the ^{15}N-labeled monomer or dimer were acquired at 500 MHz. TROSY-based methods were used to increase the sensitivity and

resolution of NOE measurements (60). The experiment included 3 s of presaturation preceded by a recycle delay of 3 s for the NOE experiment, and a 6-s recycle delay for the reference experiment (61, 62). Duplicate NOE and control spectra were recorded in an interleaved manner, using 64 transients, and spectral widths of 7017 Hz with 1024 complex points in F2 (^{1}H) and 1580.7 Hz with 100 complex points in F1 (^{15}N). Steady-state NOEs were calculated from the ratio of peak intensities from spectra with and without proton saturation, and errors were estimated from variations measured in duplicate spectra.

2.12 Determination of K_d for binding of Fos16 micelles to monomeric NGlpG

A 46 μM stock of monomeric NGlpG was prepared in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, pH 6.5. A fresh 12 mM stock solution of Fos16 was prepared and used to make a series of duplicate samples containing ~5 μM NGlpG and 0 – 10 mM Fos16. Binding of NGlpG to Fos16 micelles was monitored by CD, with spectra being acquired as described above. The MRE at 222 nm, which changed significantly upon micelle binding, was divided by the MRE at 217 nm, which showed little change upon micelle binding, and the ratio was used to determine the fraction (f_b) of NGlpG bound:

$$f_b = \frac{r - r_0}{r_{max} - r_0} \qquad \text{[eq. 15]}$$

where $r = \frac{\Theta_{222}}{\Theta_{217}}$, r_0 is the ratio of MREs in the absence of micelles, and r_{max} is the ratio of MREs when NGlpG is completely bound to the micelles.

A binding isotherm was obtained by plotting f_b as a function of the Fos16 micelle concentration, which was calculated as follows:

$$[Micelle] = \frac{[Fos16] - cmc}{N} \qquad \text{[eq. 16]}$$

Where cmc is the critical micelle concentration, (13 μM), and N is the aggregation number (178).

Since the concentration of NGlpG used was similar to the equilibrium constant for dissociation of monomeric NGlpG from Fos16, K_d, the dissociation constant was determined by fitting the isotherm to the following quadratic equation (63):

$$f_b = \frac{([NGlpG]+[Micelle]+K_d)-\sqrt{([NGLpG]+[Micelle]+K_d)^2-4[NGLpG][Micelle]}}{2[NGlpG]} \qquad \text{[eq. 17]}$$

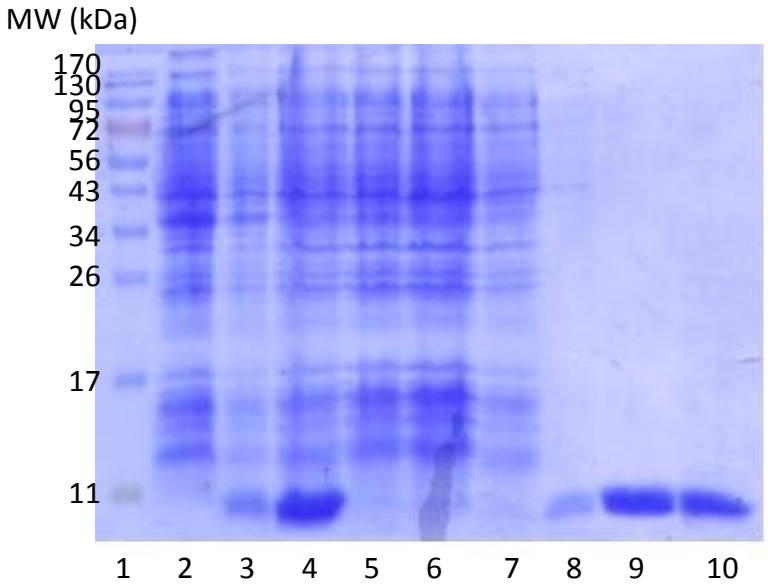
Chapter 3: Results

3.1 Protein Expression and Purification

Expression and purification of monomeric and dimeric NGlpG was done using established methods and monitored by SDS-PAGE and SEC (39). Overexpression of NGlpG in *E. coli* cells was confirmed by the presence of an intense band on an SDS-PAGE gel at ~8 kDa, the approximate molecular weight of monomeric NGlpG (36) (Fig. 3.1A, Lane 3). Following cell lysis by sonication, His-tagged NGlpG was purified by affinity chromatography and then by SEC.

For affinity chromatography, the cell lysate containing NGlpG (Fig. 3.1, Lane 4) was applied to a nickel affinity column. Binding of NGlpG to the nickel affinity column was confirmed by the absence of NGlpG in the flow-through (Fig. 3.1A, Lanes 5 and 6). Subsequently, the bound NGlpG was washed with buffer containing 10 mM imidazole to remove unwanted proteins. Most NGlpG had remained bound to the column during the wash, as confirmed by the presence of only a faint band at ~8 kDa in lane 7 (Fig. 3.1A). Other bands in lane 7 are also weak in intensity, indicating that most contaminating proteins were not bound to the column but had instead already been removed in the flow-through (Compare lanes 5, 6 and 7). Lastly, NGlpG was eluted from the column with buffer containing 250 mM imidazole, as revealed by the presence of a single, intense band in lane 8 at ~8 kDa (Fig. 3.1A).

A



B

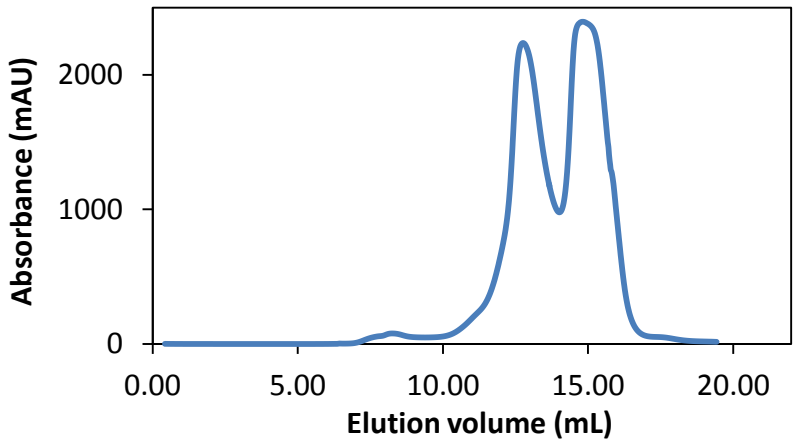


Figure 3.1: Expression and purification of NGlpG. A) Coomassie-stained SDS-PAGE gel. Lane 1: Molecular weight ladder; Lane 2: Whole cell lysate from pre-induction; Lane 3: Whole cell lysate ~4 hours after induction with IPTG; Lane 4: Soluble fraction from lysate applied to nickel column; Lanes 5 and 6: First and second flow-through from nickel column; Lane 7: Wash with 10 mM imidazole; Lane 8: Elution with 250 mM imidazole; Lane 9: SEC fractions corresponding to the peak for monomeric NGlpG; Lane 10: SEC fractions corresponding to the peak for dimeric NGlpG. The molecular weight of monomeric NGlpG is ~8 kDa (36). B) Example of SEC profile during purification of NGlpG using Superdex-75 10/300 GL column, monitored by absorbance at 280 nm. SEC was carried out in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, at pH 6.5. Two peaks at elution volumes of 12.5 mL and 15 mL correspond to monomeric and dimeric NGlpG, respectively (36).

Following elution from the nickel column, NGlpG was subjected to a second round of purification by SEC (Fig. 3.1B). NGlpG was applied to a Superdex-75 10/300 GL column, and the absorbance of the effluent was monitored at 280 nm. The SEC profile comprises two intense peaks at elution volumes of ~15 mL and ~12.5 mL, corresponding to the molecular weights of monomeric and dimeric NGlpG, respectively (36). The SEC fractions corresponding to both the lower and higher molecular weight species appear as single, intense bands at ~8 kDa on an SDS-PAGE gel, indicating that the higher molecular weight species is a non-covalent, homodimer of NGlpG.

3.2 NGlpG is Stable in the Temperature Range over which Monomer-Dimer Exchange was Conducted.

At room temperature, NGlpG converts slowly into a domain-swapped dimer, taking several days to reach equilibrium (36). The slow conversion makes it difficult to study the kinetics and thermodynamics of the interaction. To facilitate the measurement of rate and equilibrium constants, domain swapping was accelerated by conducting monomer-dimer exchange assays at higher temperatures (45^{o}C, 50^{o}C, 55^{o}C, and 60^{o}C) (36). However, at elevated temperatures proteins are more susceptible to denaturation. To see whether NGlpG remains stable from 45^{o}C to 60^{o}C, its melting temperature (T_m) was determined from analysis of thermal unfolding and refolding curves obtained by CD spectroscopy (Fig. 3.2). The T_m is the temperature at which 50% of a protein population is thermally denatured and is represented as the midpoint of the transition between upper and lower baselines of a thermal unfolding/refolding curve. Since the thermal unfolding/refolding curve lacks an upper baseline in the temperature range of 20^{o}C -100^{o}C (Fig. 3.2A), the T_m was instead determined by fitting the differential unfolding curve by eq. 4 (Fig. 3.2B). The fraction of NGlpG

unfolded and the equilibrium constant at each temperature were determined using eq. 3 and eq. 5 and by estimating the values of ΔH_m and T_m. The fit yielded a T_m of ~90°C, and the fraction of NGlpG denatured at 45°C, 50°C, 55°C, and 60°C was estimated to be less than $1x10^{-6}$, $5x10^{-6}$, $3x10^{-5}$, and $1x10^{-4}$. These low fractions indicate that NGlpG essentially remains folded in the temperature range over which the monomer-dimer exchange was performed. Based on the fit, it was also possible to determine thermodynamic values for unfolding. For monomeric NGlpG, the enthalpy change for unfolding, ΔH_m, was found to be 70 ± 10 kcal/mol, and the entropy change for unfolding, ΔS_m at the melting point was found to be 193 ± 28 cal/ (K mol).

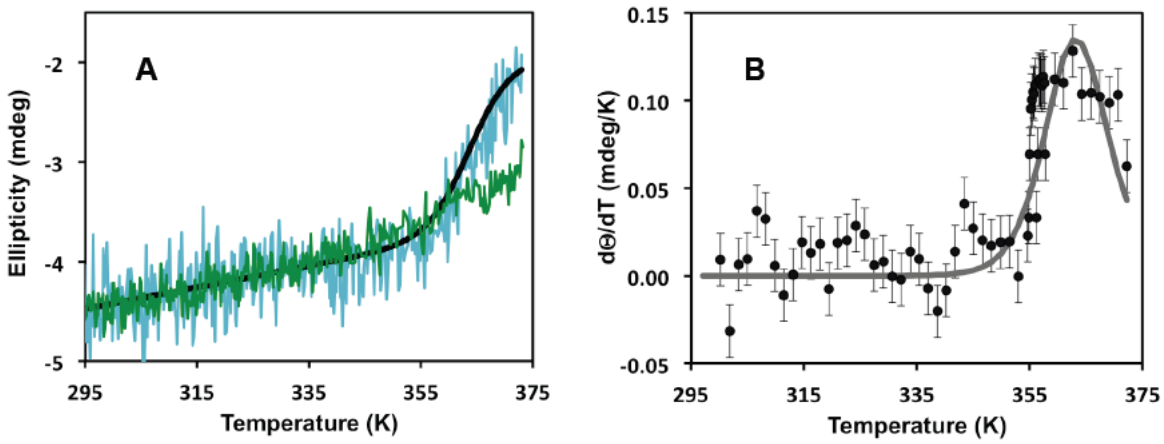


Figure 3.2: Thermal denaturation of monomeric NGlpG A) Thermal denaturation curves for 5 µM monomeric NGlpG in the absence (blue) and presence (green) of 2.5 mM Fos16. The ellipticity at 222 nm was monitored by CD spectroscopy. The buffer used for CD contained 25 mM phosphate buffer, 150 mM NaCl, 100 µM EDTA, at pH 6.5 B) Differential thermal unfolding curve for 5 µM monomeric NGlpG in the absence of Fos16. Differential ellipticities (dΦ/dT) were determined from the thermal unfolding curve for NGlpG in the absence of Fos16 in (A). A T_m of 90°C and a ΔH_m of 70±10 kcal/mol were obtained by fitting the differential unfolding curve to eq. 4 with a scaling factor of A= 4.1 x 10^{-6}. All error bars represent the standard deviation of the average dΦ/dT for 300-340 K temperature range. The parameters obtained from fitting of the differential curve were used with eq. 2 to fit the thermal denaturation curve in (A) for illustration purposes.

3.3 The Picosecond to Nanosecond Dynamics of the Hinge (31-34) Region in NGlpG are Similar to those in the Structured Domains.

The structures of monomeric and dimeric NGlpG have both been solved by solution NMR. The structure of the dimer was solved by defining the dimeric interface with distance constraints derived from intermolecular NOEs (42). However, few intermolecular NOEs were found between the hinge loops, thereby limiting the global precision of the NMR ensemble of the 20 lowest energy dimer structures. The scarcity of intermolecular NOEs between the hinge loops is expected, given the extended and solvent-exposed nature of the hinge loops. However, the low precision could also be due to enhanced motions of the hinge loops relative to the structured domains. Loop motions could occur on the picosecond to nanosecond timescale, which can be probed through NMR nuclear spin relaxation experiments (43). For this purpose, hnNOEs were measured for the monomeric and dimeric states of NGlpG (Fig. 3.3). The magnitude of the hnNOE is inversely related to the rate and extent of site-specific internal flexibility and loop motions on the ns to ps time scale. We found that the mean hnNNOE was 0.63 ± 0.13, and that the majority of residues, including those in the hinge region, are within one standard deviation of this mean. Since hnNOEs of NGlpG are similar throughout its sequence, the low precision of the NMR ensemble is likely not a consequence of enhanced loop motions but instead results from the scarcity of intermolecular NOEs in the hinge loops.

Based on a van't Hoff analysis previously done in the absence of micelles (36), the $\Delta S°$ for domain swapping of NGlpG was found to be -26.2 ± 1.4 cal/K/mol (Fig. 3.4, white data points). While the association of two monomer subunits decreases the entropy, it is also possible that $\Delta S°$ is influenced by any changes in the picosecond to nanosecond dynamics associated with domain swapping. To see whether domain swapping alters the picosecond to

nanosecond dynamics of NGlpG, hnNOEs were acquired on the monomer and compared with those for the dimer (Fig. 3.3) The mean residue hnNOE for the monomer was 0.68±0.13, which is identical within standard deviation to that for the dimer. Additionally, domain swapping does not significantly alter the dynamics of each residue, as indicated by the similar hnNOEs for respective residues in both the monomer and the dimer. Since the dynamics of NGlpG are unaffected by domain swapping, the ΔS° associated with domain swapping is mainly due to the association of two monomer subunits (36).

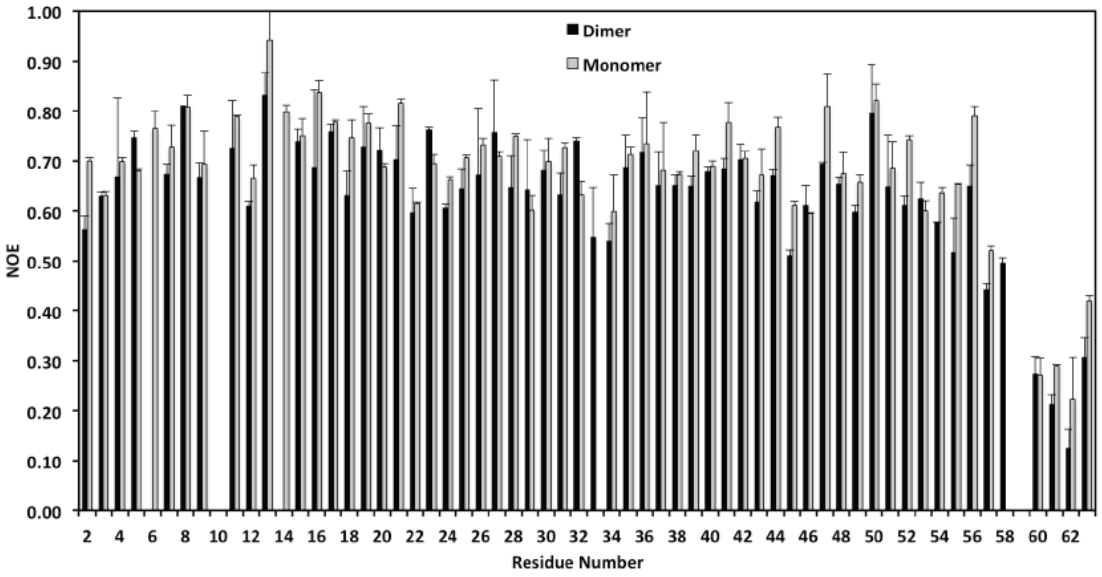


Figure 3.3: Steady-state heteronuclear {^{1}H}-^{15}N NOEs for monomeric and dimeric NGlpG. {^{1}H}-^{15}N NOEs were acquired at 500 MHz and 298 K for ^{15}N-labelled monomeric (grey) and dimeric (black) NGlpG in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, and pH 6.5. The height of each bar represents an average of {^{1}H}-^{15}N NOEs from duplicate acquisitions on the same sample. Each error bar represents the standard deviation of the average {^{1}H}-^{15}N NOE for the corresponding residue. The mean residue {^{1}H}-^{15}N NOEs for the monomer and dimer are 0.63±0.13 and 0.68±0.13, respectively.

3.4 Backbone Amide Hydrogen-Deuterium Exchange

The entropic penalty associated with domain swapping in the absence of additives is compensated by an enthalpy decrease ($\Delta H° = -15$ kcal/mol) (Fig. 3.4, white data points) (36), which is generally attributed to interactions that form exclusively within the dimer between hinge loop residues or between structured domains from different subunits. The structured domains of the NGlpG protomers are too far apart to engage in additional intermolecular interactions outside the domain swap, as confirmed by the absence of intermolecular NOEs between these regions (42). On the other hand, the side chains of Q31 and N33 from the extended hinge loops of the dimer appear to be close enough to engage in inter- and intra-molecular hydrogen bonds (Fig. 3.5) that are unlikely to occur in the monomer, whose hinge loop folds back on itself such that Q31 and N33 are unable to interact with each other (Fig. 1.4). These hydrogen bonds might account for part of the enthalpy decrease associated with domain swapping.

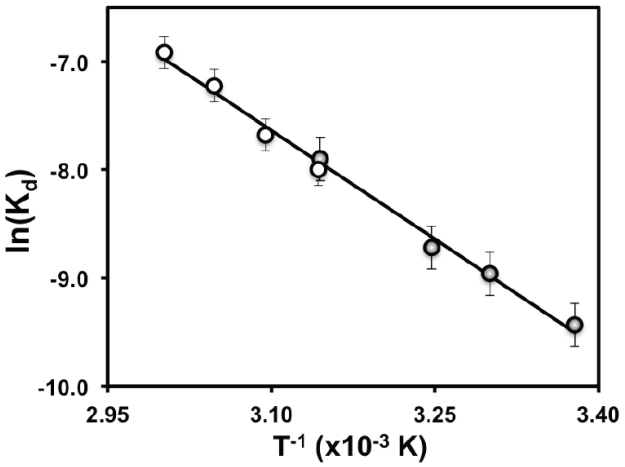


Figure 3.4: Van't Hoff analysis for domain swapping of NGlpG in the absence (white) and presence (grey) of Fos16. All measurements were performed in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, at pH 6.5. The plot shows the relation between K_d and temperature in the 318-333 K (294-318 K) range for domain swapping in the absence (presence) of Fos16. Each data point represents an average of K_d values obtained individually from at least two separate samples. Each error bar represents the standard deviation of the average $\ln(K)_d$. Linear regression to the van't Hoff equation gave $\Delta H°$ and $\Delta S°$ of -13.4±0.4 kcal/mol and -26.2±1.4 cal/K/mol, respectively.

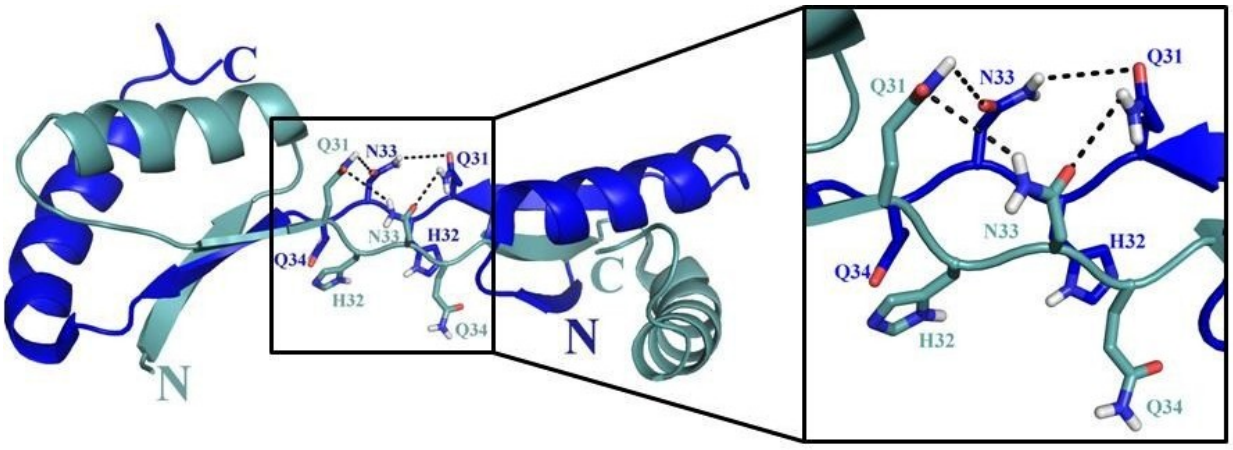
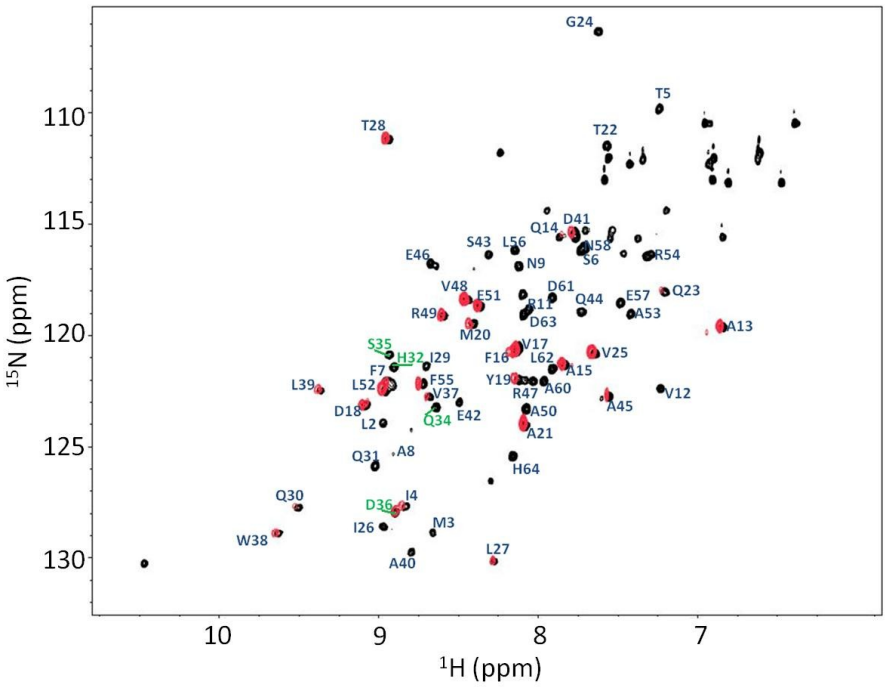


Figure 3.5: Solution NMR structure of dimeric NGlpG. The hinge loop residues Q31, H32, N33, and Q34 from both subunits are shown as sticks. Potential intra- and intermolecular hydrogen bonds are shown as black, dashed lines. (Inset) Closer view of the hinge loop residues and hydrogen bonds from both subunits.

To see how hydrogen bonding affects local conformational stability throughout NGlpG, hydrogen exchange experiments were performed on NGlpG to determine protection factors, which quantify the extent to which amide hydrogens are protected from exchange with the solvent (64). Higher protection factors indicate greater protection from exchange and hence greater local stability. Examples of HSQC acquired on dimeric NGlpG before and after exchange into deuterated buffer are shown in Figure 3.6A. Peaks for 26 backbone amides remain following exchange into D2O. These peaks correspond mainly to residues in the β sheet (ex: W38 and V37 of β3, T28 and Q30 of β2, I4 of β1) and to regions of the α helices that are less exposed to the solvent (ex: V48 and L52 of α2, F16 and M20 of α1). For each residue, the time-dependent decay in signal intensity was used to determine a rate constant for exchange, which was converted to a protection factor by dividing the observed rate constant by the intrinsic rate constant for exchange. As shown in Fig. 3.6Bii, protection factors in the dimer are highest in β3, the hydrophobic face of α1, and β2. Similar results were obtained for the monomer (Fig. 3.6Bi), in that the same three secondary structural elements possess high protection factors. However, a few differences in

A



B

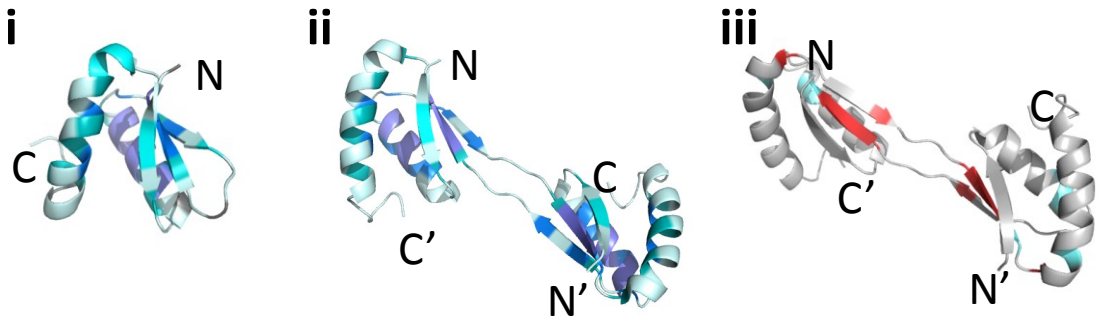


Figure 3.6: Measurement of hydrogen exchange protection factors. (A) 2D ^{1}H-^{15}N HSQC spectra at 25°C of 0.5-1 mM dimeric NGlpG in 25 mM phosphate buffer, 150 mM NaCl, 100 µM EDTA, at pH 6.5 before (black) and at the first time point (~10-15 minutes) following exchange into D_2O (red). Backbone amide peaks are labelled. Peaks corresponding to hinge loop residues are labelled in green. (B) Hydrogen exchange protection factors (PF) mapped onto the solution NMR structures of i) monomeric and ii) dimeric NGlpG. Protection factors greater than 7, 6, 5, and 4 are shown in purple, dark blue, teal and cyan, respectively. iii) Changes in protection factors due to domain swapping are mapped onto the structure of the dimer. Increases in protection factors greater than 1 are shown in red. Decreases in protection factors with magnitudes greater than 1 are shown in blue. The rest of the residues have protection factors that change by a magnitude of less than 1, or were too fast to be measured (grey).

protection factors can be observed between the monomer and the dimer. These differences localize to β3, Q30 of β2, and D41 of α2 (Fig. 3.6Biii), suggesting that domain swapping increases the protection factors and hence conformational stabilities in these regions. Since the protomers in the domain-swapped dimer are structurally identical to the monomer, the increased stabilities in these regions might be associated with hydrogen bond formation between Q31 and N33 in the dimer.

3.5 Determination of Rate Constants and Kinetic Order for Fos16-Accelerated Domain Swapping.

At 45^{o}C, domain swapping of 0.25 mM monomeric NGlpG reaches equilibrium in ~2 days (36). When the same concentration of monomeric NGlpG was incubated with 2.5 mM Fos16 at the same temperature, equilibrium was reached in less than 10 minutes (Fig. 3.7 A, B). The dimer had formed by domain swapping, as confirmed by a 2D ^{1}H-^{15}N correlation HSQC spectrum that had previously been acquired on NGlpG at 25^{o}C following a one hour incubation of the monomer with 2.5 mM Fos16 (36). Besides the monomer and the dimer, other states were not observed by SEC (Fig. 3.7A), indicating that Fos16 does not promote the formation of higher order aggregates or trap detergent-bound intermediates that are stable under the conditions used for SEC.

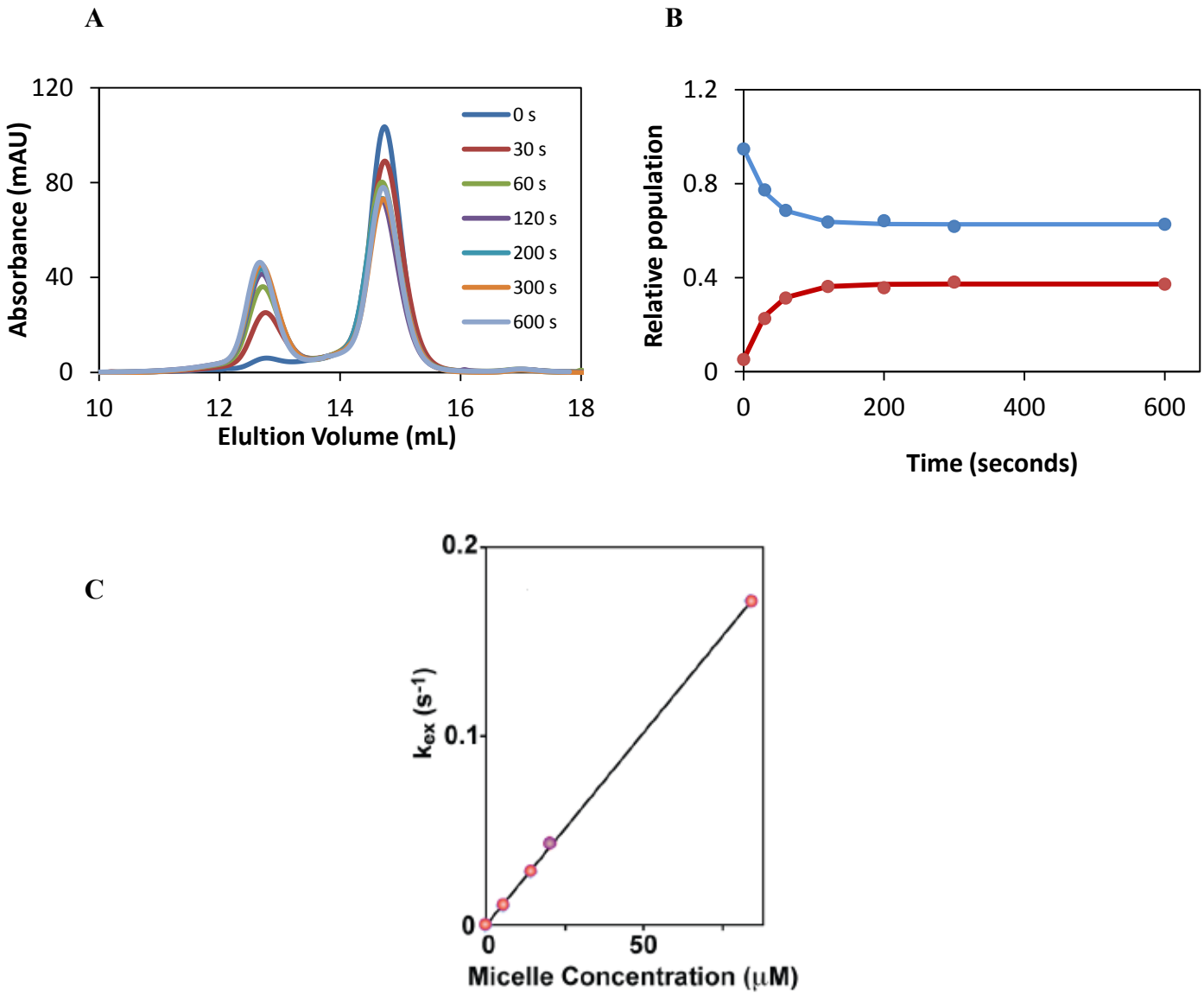


Figure 3.7: Dimerization kinetics of 0.25 mM NGlpG in the presence of 2.5 mM Fos16 (~14 μM micelles) at 318 K.
(A) Size exclusion chromatography profiles showing a decrease in monomeric NGlpG and increase in dimeric NGlpG at 0, 30, 60, 120, 200, 300, and 600 seconds. Elution volumes for the dimer and monomer are ~12.5 mL and ~15 mL, respectively. The buffer used for monomer-dimer exchange and for SEC contained 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, at pH 6.5. SEC was performed at room temperature. (B) Kinetic profile showing time-dependent changes in the relative monomer (blue data points, blue trace) and dimer (red data points, red trace) populations determined from peak areas in SEC profiles. (C) Demonstration of linear relationship between pseudo-first order rate constant, k_{ex}, for monomer-dimer exchange and concentration of Fos16 micelles. k_{ex} obtained at Fos16 micelle concentrations of 6 μM, 14 μM, and 84 μM were (9.3±1.1) x 10^{-3} s^{-1}, (2.88±0.03) x 10^{-2} s^{-1}, and 0.17 s^{-1}. The slope of the line is a second order rate constant, $k_{2,ex}$, equal to 2.02 $mM^{-1}s^{-1}$.

To quantify the rate and extent of domain swapping in the presence of Fos16, the peak areas from the SEC profiles were used to construct a kinetic profile representing the depletion of monomer over time (Fig. 3.7B). The kinetic profile was fit to an integrated rate equation for a first order, reversible two-state system, from which a first order rate constant of $0.034 \pm 0.010\ s^{-1}$ was obtained (Table 3.1). By comparison, the rate constant for domain swapping in the absence of Fos16 is $(2.0 \pm 0.4)\ x10^{-5}\ s^{-1}$; thus, this concentration of Fos16 enhances the rate of domain swapping by ~1700 fold. Using the relative populations of the monomer and dimer states at equilibrium, a dissociation constant of 0.35 ± 0.05 mM was obtained, which is equivalent to the K_d of 0.36 ± 0.01mM obtained in the absence of Fos16 (Table 3.1). Therefore, Fos16 increases the domain swapping rate at 45^oC without altering the extent of dimerization.

Table 3.1: Effect of additive on pseudo-first order rate constant for domain swapping at 318K.
First order rate constants were obtained for domain swapping of NGlpG in the presence of 2.5 mM additive (1M for urea).

Additive	**K_d (mM)**	**k_{ex} (s^{-1})**	**Fold acceleration**
None	0.36 ± 0.01	$(2.0 \pm 0.4)\ x10^{-5}$	1
Fos16	0.35 ± 0.05	0.034 ± 0.010	1700
LPPG	0.68 ± 0.03	0.044 ± 0.016	2200
DDM	0.48 ± 0.13	$(12 \pm 2)\ x10^{-5}$	6.0
Urea	0.71 ± 0.02	$(8.9 \pm 1.0)\ x10^{-5}$	4.5

To quantify the influence of Fos16 on the domain swapping rate, pseudo-first order rate constants were determined over a range of Fos16 concentrations (Fig. 3.7C). For any detergent, increasing the detergent concentration above its critical micelle concentration increases the number of micelles rather than the number of free detergent molecules in solution. For Fos16-accelerated domain swapping, the rate constants were found to increase linearly with detergent concentrations above the CMC, suggesting that the increase in rate was due to an increase in micelle concentration rather than the free detergent. Further, the linear relationship between k_{ex} and Fos16 concentration indicates that the rate is first order with respect to Fos16 micelles. The slope of this line yields a second order rate constant, k_2, of 2.0 $mM^{-1}s^{-1}$, which indicates that one micelle and one NGlpG monomer or dimer participate in the rate-limiting step of this process.

3.6 Determination of Thermodynamic Parameters for Fos16-Catalyzed Domain Swapping.

To determine the enthalpy and entropy changes, ΔH^o and ΔS^o, for domain swapping of NGlpG in the presence of Fos16, a van't Hoff analysis was carried out by measuring the K_d values for Fos16-accelerated domain at 21°C, 30°C, 35°C, and 45°C (Fig. 3.4, grey points). Based on the analysis, ΔH^o and ΔS^o were found to be -13.4 ± 0.4 kcal/mol and -26.2 ± 1.4 cal/K/mol, respectively. These values are identical to the corresponding ΔH^o and ΔS^o for domain swapping of NGlpG in the absence of Fos16 (Fig. 3.4, white points), suggesting that Fos16 micelles function as catalysts and therefore accelerate domain swapping of NGlpG without changing the free energy difference between the monomer and the dimer. In addition, the van't Hoff analysis yielded a linear relation between the K_d and temperature in

the 21°C to 45°C temperature range, indicating that unfolding and other temperature-dependent processes do not occur from 21°C to 45°C.

If Fos16 is a catalyst for domain swapping, it should accelerate domain swapping by lowering the free energy for activation. To determine the extent to which Fos16 lowers this barrier, an Eyring analysis was performed by measuring rate constants at 30^{o}C, 35^{o}C, 45^{o}C, and 60^{o}C and then fitting the data points to obtain the activation parameters Δ‡H°, Δ‡S°, and Δ‡G° (Fig. 3.8). Based on the analysis, Δ‡H°and Δ‡S°for Fos16-catalyzed domain swapping were found to be 20.1 ± 1.5 kcal/mol and 16.4 ± 5.0 cal/K/mol, which are both considerably less than the respective values of 73.2 ± 0.9 kcal/mol and 147.1 ± 2.8 cal/K/mol obtained for domain swapping in the absence of Fos16 micelles (Table 3.2). The lower activation enthalpy for domain swapping in the presence of Fos16 micelles might arise from stabilizing interactions that form between the micelle and NGlpG, while the lower activation entropy is indicative of the entropy cost that comes from the association of NGlpG with the micelle. Together, these contributions lower the free energy of activation, Δ‡G°, at 45^{o}C by ~12 kcal/mol.

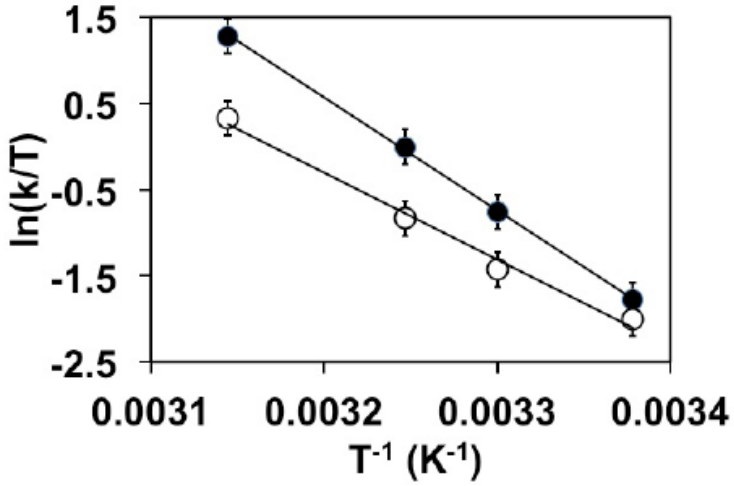


Figure 3.8: Eyring plot for NGlpG dimer association (white) and dissociation (black) in the presence of Fos16. All measurements were performed in 25 mM phosphate buffer, 150 mM NaCl, 100 μM EDTA, at pH 6.5. The rate constant *k* used in this analysis represents the second order forward or reverse rate constant in the 294-318 K range. Each data point represents an average obtained from two independent samples. The activation enthalpy and entropy for dimerization and dimer dissociation determined from the linear fit are listed in Table 3.2.

Table 3.2: Activation parameters for NGlpG monomer (M) association and dimer (D) dissociation of NGlpG in the absence and presence of Fos16.

	No Detergent		**Fos16**	
	2 M → D	D → 2 M	2 M → D	D → 2 M
$\Delta^{\ddagger}H^{o}$ (kcal/mol)	73.2 ± 0.9	85.0 ± 1.7	20.1 ± 1.5	26.1 ± 0.4
$\Delta^{\ddagger}S^{o}$ (cal/K/mol)	147.1 ± 2.8	185.7 ± 5.3	16.4 ± 5.0	37.3 ± 1.2
$\Delta^{\ddagger}G^{o}$ at 45 °C (kcal/mol)	26.5 ± 1.3	25.9 ± 2.4	14.9 ± 2.2	14.2 ± 0.5

3.7 Investigation of the Effects of Detergent Size and Charge on Domain Swapping

While the results of the Eyring analysis indicate that Fos16 micelles catalyze domain swapping, the micelle properties important for catalysis had not been investigated. For this purpose, the influences of two physical properties, namely micelle size and charge (Fig. 3.9, Table 3.1), were examined by measuring and comparing rate constants for domain swapping promoted by detergents with different chain lengths and head groups. As shown in Figure 3.9, increases in phosphocholine detergent alkyl chain length and hence micelle size were found to increase the second order rate constant for reversible monomer-dimer exchange. Negatively charged micelles were also found to accelerate domain swapping since 1-palmitoyl-2-hydroxy-sn-glycero-3-[phosphor-rac-(1-glycerol)] (LPPG) (Table A.1), an anionic detergent containing a 16-carbon acyl chain, enhanced domain swapping by 2200-fold (Table 3.1). However, the nonionic detergent DDM (Table A.1) only gave rise to a 6-fold acceleration, indicating that the presence of a charge in the detergent head group is necessary for increasing the domain swapping rate.

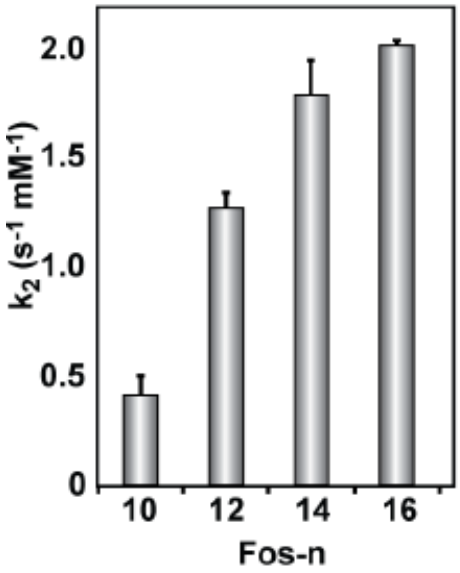


Figure 3.9: Correlation between alkyl chain length of phosphocholine detergents and second order rate constant for domain swapping at 318 K.
Second order rate constants were determined for reversible monomer-dimer exchange in the presence of phosphocholine detergent.

3.8 Characterization of the Micelle-Bound Intermediate State of NGlpG

Since Fos16 micelle-catalyzed domain swapping is first order with respect to NGlpG and Fos16 micelles (Fig. 3.7), a single subunit of monomeric NGlpG and a single Fos16 micelle participate in the transition state of the rate-limiting step for dimerization. If the association occurs before the rate-limiting step, it may be possible to isolate the micelle-bound state of NGlpG and to characterize the affinity of the association. To investigate this possibility, Fos16 micelles were titrated into monomeric NGlpG and the effect was monitored at 45^{o}C using CD spectroscopy. To ensure that changes in the CD signals were due to binding of monomeric NGlpG to Fos16, the concentration of monomeric NGlpG was kept low (~5 μM) to minimize the influence of dimer formation on the CD spectrum (less than 3% dimer at equilibrium predicted at this concentration). The CD spectrum of monomeric NGlpG changed upon addition of excess Fos16 (20 mM Fos16) (Fig. 3.10), with a significant increase in intensity at ~208 nm, suggesting a significant change in secondary structure. Secondary structure content prediction based on these spectra indicate that the

association of monomeric NGlpG with a Fos16 micelle results in a 10% gain of alpha helix content, concomitant with 7% and 3% decreases in beta strand and coil contents, respectively (Table 3.3).

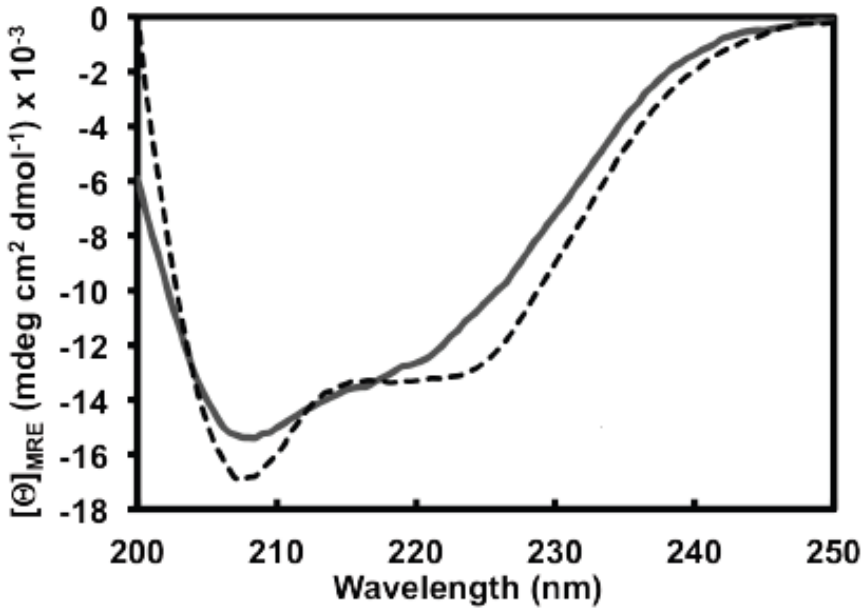


Figure 3.10: CD spectra of ~5 µM monomeric NGlpG in the absence (solid line) and presence (dashed line) of 20 mM Fos16 (~110 µM micelles) at 318 K. Buffers for CD spectroscopy contained 25 mM phosphate buffer, 150 mM NaCl, 100 µM EDTA, at pH 6.5.

Table 3.3: Secondary structure content of monomeric NGlpG in the absence and presence of 20 mM Fos16 (~110 mM micelles) at 318 K. Each value represents an average of mean percentages determined individually from two independent samples calculated from three algorithms (SELCON3, CDSTTR, and CONTIN).

	α	β	Turn	Random
NGlpG	35 ± 3	17 ± 1	19 ± 1	29 ± 1
NGlpG + Fos16	42 ± 2	10 ± 1	19 ± 1	26 ± 1

To assess the accuracy of the secondary structure content predictions, the secondary structure content of monomeric NGlpG, as predicted from its CD spectrum, was compared with values determined from the Define Secondary Structure of Proteins (DSSP) algorithm, which calculates secondary structures of proteins based on their atomic coordinates (65). The DSSP algorithm predicts that the solution NMR structure of NGlpG (PDB ID 2LEP) possesses 42% helix character and 21% beta sheet character (65). By comparison, the secondary structure of monomeric NGlpG, determined from analysis of its CD spectrum, consists of 35% α helix character and 17% β strand character (Table 3.3). Although analysis of the CD spectrum results in an underestimation of α helix and β strand character, the CD and DSSP methods both predict that NGlpG possesses twice as much α helix character as β strand character.

To verify the stoichiometry of the association and to determine the affinity of the interaction between monomeric NGlpG and Fos16 micelles, a binding isotherm was plotted using the molar ellipticity values from the titration and fit to eq. 17. The isotherm was found to consist of a single saturation point (Fig. 3.11), confirming that one NGlpG associates with only one Fos16 micelle. Moreover, a K_d of 2.8 μM was obtained from the quadratic fit of the isotherm, corresponding to a ΔG^o for binding of approximately -8 kcal/mol. This value is comparable in magnitude to the 12 kcal/mol stabilization of the transition state for domain swapping in the presence of Fos16 micelles, suggesting that catalysis of domain swapping is due predominantly to stabilizing interactions that form from the association of NGlpG with a Fos16 micelle.

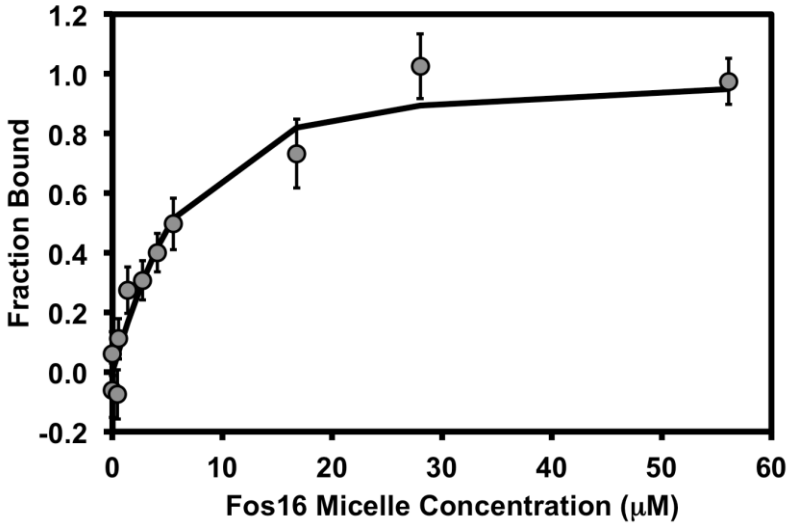


Figure 3.11: Titration of Fos16 micelles into ~5 μM monomeric NGlpG at 318 K. The fraction of NGlpG bound to Fos16 micelles was determined based on ratios of mean residue ellipticities at 222 nm to 217 nm. Each data point represents an average of fraction bound values determined from individual measurements on two separate samples. Each error bar represents a standard deviation of the average fraction bound. The binding isotherm was fit to eq. 17, from which a K_d of 2.8 μM was obtained.

3.9 Tertiary Interactions in the Micelle-Bound Intermediate

To assess the thermal stability of the micelle-bound intermediate, a thermal denaturation analysis was performed on monomeric NGlpG in the presence of 2.5 mM Fos16. As shown in Figure 3.2B, the thermal unfolding curve does not exhibit an upper baseline or a cooperative melting transition, suggesting that there is no cooperativity for thermal unfolding. This would indicate that the secondary structural elements of the micelle-bound intermediate are not stably associated with each other, but instead are bound to the micelle in a stable interaction that is not disrupted by the increase in temperature.

Chapter 4: Discussion

4.1 Insights into the Solution Structure of NGlpG

In this work, NMR data was obtained to determine if the solution structure of dimeric NGlpG is flexible in the hinge domain region. Although the solution structure of the NGlpG dimer is superimposable with the crystal structure, the hinge loops adopt a beta strand conformation in the crystal structure, whereas the hinge loops are not well-defined in the solution structure. This may be due to the extended nature of the hinge loop, which could limit the number of observed NOEs, or that the hinge loops are more dynamic in solution than in crystal form. As such, heteronuclear $\{^{1}H\}$-^{15}N NOEs were measured to see whether the low precision was caused by enhanced internal flexibility of the hinge loops relative to the protomers. The results show that heteronuclear NOEs in the hinge loop residues have magnitudes that are similar to those in the rest of the domain-swapped dimer, suggesting that the low precision of the ensemble does not result from enhanced internal flexibility of the hinge loops on the nanosecond to picosecond timescale. However, these residues might have enhanced motions on a longer timescale, such as those that are more typically seen in inter-domain motion or protein folding (43). This has been seen in the domain-swapped dimer of cyanovirin-N, since the protomers adopt different relative orientations in the solution and crystal structures, with enhanced conformational mobility of the hinge loop being detected on the microsecond time scale (26). Specifically, dimerization gave rise to line broadening of peaks corresponding to the hinge region in the NMR spectrum. However, no obvious line broadening is apparent in hinge loop residues in the HSQC spectrum of the dimer (Fig. 3.6A, black peaks), suggesting that there may not be significant inter-domain flexibility in the dimer.

To better understand the difference in stabilities of the monomer and the dimer we also performed backbone amide hydrogen exchange experiments monitored by solution NMR. The results show that the protection factors in β2 and β3 are higher in the dimer than in the monomer (Fig. 3.6B) reflecting an increase in local conformational stability. Given that the protomer structures in the dimer are structurally identical to the monomer, the increase in the stability of the hydrogen bonds between β2 and β3 may result from the formation of hydrogen bonds between the hinge loops in the dimer. For example, domain swapping of NGlpG appears to enhance the potential for hydrogen bond formation between side chains in the hinge loop (e.g. between Asn33 and Gln31) that would not be possible in the monomer (Fig. 3.5, Fig. 1.4). As has been observed in other cases (28), these changes in the hinge loop interactions may be sensed outside of the hinge region, such as in β2 and β3. These changes in stability, along with the loss of translational and rotational entropy that arises from the association of two polypeptides to form one complex, result in a domain-swapped dimer that is ~5 kcal/mol more stable than the monomer at 45°C in terms of Gibbs free energy.

4.2 Micelle Catalysis in NGlpG Domain Swapping

One of the most novel aspects of NGlpG domain swapping is the catalysis of this process that can be achieved through the use of zwitterionic detergent micelles. While a few proteins have been shown to undergo domain swapping that could be promoted by detergents, catalysis by micelles has never previously been demonstrated. For example, although domain swapping of Bax was induced by the non-ionic detergents Triton X-100 and Nonidet P40 (Table A.1), it was not shown whether dimerization was due to micelles or the free detergent, or if the equilibrium constant for dimerization was preserved (66, 67). In

another case, domain swapping of Bet v4 could be accelerated in the presence of SDS (Table A.1) (68); however, only sub-micellar concentrations were required. Furthermore, approximately 10% of Bet v4 formed aggregates rather than domain-swapped dimers under these conditions. Unlike these examples, domain swapping of NGlpG could be catalyzed by Fos16 micelles without undergoing aggregation.

Although not yet explored, it is possible that the use of micelles to accelerate domain swapping may be applicable to other proteins, particularly if they have already been shown to have some propensity to interact with micelles or lipid vesicles (32, 69–72). A promising candidate is saposin C (SapC), a protein that is involved in destabilization and fusion of acidic vesicles (71). Monomeric SapC forms a bundle of 5 α helices that binds to SDS micelles in an open conformation, with two of the α helices being displaced to expose the hydrophobic core that interacts with the micelle surface (70). Interestingly, the structure of the micelle-bound monomer is highly similar to the structure adopted by each subunit in a crystal structure of the SapC domain-swapped dimer (71). Based on this structure the authors proposed that SapC may have encountered endogenous yeast lipids during expression in *P. pastoris* to give rise to the domain-swapped structure captured in the crystal. It is noteworthy that the micelle-bound intermediate state of SapC is similar to that of NGlpG, in that both proteins undergo an increase in helix content and experience disruptions of tertiary interactions. These similarities raise the possibility that micelles can be generally useful for the catalysis of domain swapping, particularly if the protein of interest has some propensity to interact with lipid bilayers or micelles.

Another domain-swapping protein that has some propensity for membrane binding is the cytoplasmic domain of the membrane protein Fis1 (Fis1ΔTM), which forms a domain-swapped dimer that is necessary for activity (32). Slow exchange between oligomeric states

could be overcome by raising the temperature above ~45^{o}C, or by treatment with >4 M guanidinium, although these conditions gave rise to significantly different populations of monomeric and dimeric states. The ability of Fis1ΔTM to reversibly bind to lipid membranes (69) raises the possibility that detergent micelles could catalyze domain swapping while maintaining the free energy difference between monomeric and dimer states, a hypothesis that remains to be explored in the future. Overall, the potential for micelles to be used as catalysts of domain swapping has been an underappreciated phenomenon. The results of this thesis have the potential to expand the arsenal of tools to facilitate the study of domain-swapping interactions.

4.3 Mechanism for Micelle-Catalyzed Domain Swapping of NGlpG

During the characterization of NGlpG domain swapping kinetics, it was found that dimerization is first order with respect to NGlpG. Therefore, the rate-limiting step does not involve the association of monomeric subunits, but rather requires some high barrier activation event in a single monomer (36). This has been seen in other cases, for example domain-swapping dimerization of cyanovirin-N, where the rate-limiting step was proposed to involve unfolding (73). In the case of NGlpG it is not yet known whether complete unfolding is necessary for domain swapping. However, the free energy of unfolding can be estimated based on the thermal denaturation data obtained on monomeric NGlpG. ΔH_m and ΔS_m for unfolding of NGlpG at the transition midpoint (T_m = 90 ± 2°C) was found to be 70 ± 10 kcal/mol and 193 ± 28 cal/(mol K), respectively (Fig. 3.2). If it is assumed that ΔH_m and ΔS_m are temperature-independent then $\Delta G°$ for unfolding at 45°C would be 9 kcal/mol, corresponding to ~0.00008% of NGlpG in the unfolded state under these conditions. In this case, if unfolding is the rate-determining step then the activation barrier to unfolding should

be the same height as the barrier for domain swapping. However, entropy is temperature-dependent, while enthalpy varies with the heat capacity, a variable that also depends on temperature. Consequently, it is not possible to directly compare the enthalpy and entropy of unfolding at the melting temperature to the activation enthalpy and entropy of domain swapping determined at lower temperatures. Instead it would be necessary to determine $\Delta G°$ for unfolding at 45°C, for example through titration of NGlpG with chemical denaturants, such as guanidine hydrochloride (74). If the unfolded state is also similar to the transition state for domain swapping, then this free energy difference would be expected to be similar to the activation free energy for domain swapping. However this may not be the case, since most proteins have only small free energy differences between folded and unfolded states. Nonetheless, the transition state for unfolding may be the same as that for domain swapping, a possibility that could be explored in future experiments by measuring folding/unfolding rates for NGlpG, along with an Eyring analysis.

In the presence of micelles, domain swapping remains first order with respect to NGlpG, and also first order in micelle concentration, indicating that the transition state must involve both species (Fig 4.1). Based on our measurements of Fos16 micelle binding affinity, this interaction is quite favourable, with a $\Delta G°$ of -8 kcal/mol for complex formation. This is a benefit for micelle catalysis since high concentrations of the micelle catalyst are not required for rate enhancement (75), as we have observed in our experiments. In addition, binding is associated with conformational changes, both at the secondary and tertiary structure levels. Specifically, there is an increase in alpha helix content, along with decreases in β strand and random coil contents that occur upon binding of the monomer to Fos16 micelles (Table 3.3). It is likely that these secondary structure elements are not stably associated in a compact globular state since the thermal denaturation of NGlpG in the

presence of Fos16 micelles did not show evidence of a cooperative transition. Despite these structural characterizations, it is important to note that the secondary structure prediction was performed using a standard set consisting only of soluble proteins (50). However, the micelle-bound state may also exhibit characteristics of a membrane protein, as NGlpG interacts with the micelle in a partially unfolded state with a potentially exposed hydrophobic core. If so, this would make the secondary structure prediction less valid. Future studies should focus on characterizing the structure of the micelle-bound state using higher resolution techniques, such as NMR spectroscopy.

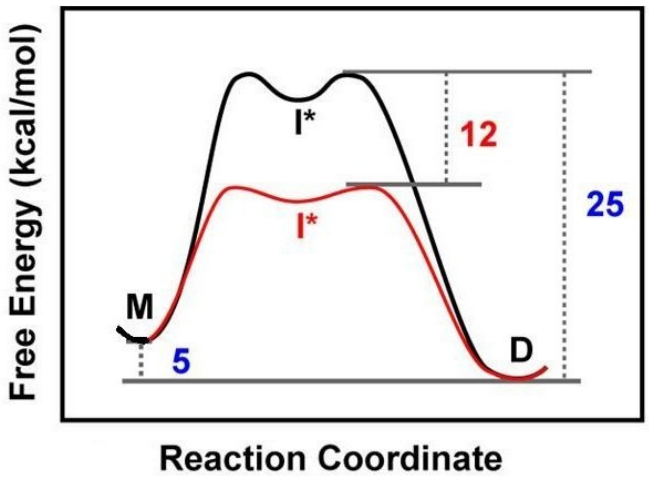


Fig. 4.1: Free energy landscape for domain swapping of NGlpG at 45°C in the absence (black) and presence (red) of Fos16 micelles. The free energy difference between the monomer (M) and dimer (D) is ~5 kcal/mol. Domain swapping involves a high-energy intermediate, I*, which is formed when the activation barrier (~25 kcal/mol) is overcome. In the presence of Fos16 micelles, the activation barrier for domain swapping is lowered by ~12 kcal/mol.

The results of our Eyring analysis in the presence and absence of Fos16 micelles show that the micelle may bind more tightly to the transition state than the intermediate state, since the decrease in activation energy (~12kcal/mol) is larger than the energy of binding the partially unfolded intermediate (~8 kcal/mol). The results of the Eyring analysis also reveal that rate acceleration comes from a significant decrease in the activation enthalpy for domain

swapping. Since micelle charge appears to influence the rate of micelle-accelerated domain swapping, it is likely that electrostatic interactions comprise some part of this stabilization. However, the large amount of entropy that is gained in the transition state for uncatalyzed domain swapping is significantly reduced by micelles. This is not unexpected since the association of two species will reduce the translational and rotational degrees of freedom available to the system. However, it is also possible that the higher order of the micelle-catalyzed transition state reflects a more compact state than the uncatalyzed pathway. Since detergent tail length affects the rate of domain swapping, it is possible that smaller micelles impose a more restrictive environment for the NGlpG transition state, thereby decreasing the activation entropy for micelle-accelerated domain swapping. Although the effects of detergent charge and size on the activation enthalpy and activation entropy were not directly tested in this study, this is an interesting aspect of micelle-catalyzed domain swapping that could be pursued in the future.

4.4 A Chaperonin-Like Mechanism of Catalysis

Structural analysis of the micelle-bound state of NGlpG showed a high degree of secondary structure that did not undergo a cooperative thermal transition, suggesting a partially unfolded state for the micelle-bound protein. This ability of the micelle to at least partly unfold NGlpG is reminiscent of the action of chaperonins, which also unfold kinetically trapped folding intermediates (76–78). In particular, FRET studies have shown that binding of misfolded intermediates to chaperonins such as GroEL/GroES results in disruption of non-native interactions (78, 79). Given the stabilities of monomeric and dimeric NGlpG, either of these two states can be considered kinetically trapped intermediates. Fos16

micelles help to free these trapped intermediates by disrupting tertiary interactions to form a partially unfolded state that is primed for domain swapping.

The ability of micelles to act in a chaperonin-like manner is novel, and distinct from previous studies that used micelles to function as protein chaperones (80, 81). In those studies, micelles were used to bind proteins that were completely denatured by exposure to harsh conditions such as high temperatures or chemical denaturants. The designed micelles did not have any propensity to bind the folded state or induce unfolding on their own. Instead the primary function of these micelles was to bind and sequester the unfolded state to prevent aggregation. However, because the interaction with the unfolded state was strong, micelle removal through exposure to detergent-binding reagents (80), or enzymatic systems that disrupted the micelle structure (81), was required to allow polypeptide release under refolding conditions. That is distinct from micelle catalysis of NGlpG domain swapping, since detergent removal was not required to achieve acceleration of domain swapping. This was demonstrated by the NMR spectrum of NGlpG in the presence of micromolar quantities of detergent micelles, with the pure monomeric species giving rise to the dimeric species at an accelerated rate (36). If the micelle interaction with NGlpG was too strong, then the micelles would be expected to become saturated with protein, and unable to release protein for the domain-swapping interaction.

4.5 Relevance of Micelle-Catalyzed Domain Swapping to Fibril Formation

An intriguing possibility that is raised by the results from this thesis is that lipid-protein interactions may nucleate domain swapping in some protein deposition diseases. Many amyloid fibrils are formed from open-ended domain swapping, including fibrils of PrP, β2-microglobulin, α1-antitrypsin (12,17,18). Fibril formation usually occurs over

several years and involves a lag phase followed by a growth phase. The lag phase is associated with the formation of a nucleus that seeds the formation of aggregates. In the case of cystatin C, the nucleus is formed from its domain-swapped dimer (13). Since the results from my study show that domain swapping can be accelerated by detergent micelles, it is also possible that some regions of lipid membranes may promote nucleation. Specifically, local deformations in lipid bilayer structure that increase solvent exposure of the hydrophobic phase, as might be seen in regions of significant membrane curvature or by hydrophobic mismatch at phase boundaries of lipid rafts may have some propensity to bind proteins in a partially denaturing manner. As shown in this thesis, the energetic barrier for domain swapping could be lowered through this interaction, potentially nucleating the formation of amyloid fibril formation, and accelerating the development of amyloidosis. The electrostatic and steric properties of the micelles that are important for the catalysis of domain swapping of NGlpG could also be important for protein-lipid interactions that result in rate enhancement of fibril formation *in vivo*. Ultimately, the structural and energetic characterization of the micelle-bound state of NGlpG presented in this thesis has the potential to facilitate the rational design of therapeutics that inhibit fibril formation involving domain swapping oligomers in the future.

References

1. Gronenborn AM (2009) Protein acrobatics in pairs--dimerization via domain swapping. *Current Opinion in Structural Biology* 19:39–49.

2. Yang F, Bewley CA, Louis JM, Gustafson KR, Boyd MR, Gronenborn AM, Clore GM, Wloodawer, A (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *Journal of Molecular Biology* 288:403–12.

3. Liu Y, Eisenberg D (2002) 3D domain swapping: as domains continue to swap. *Protein Science* 11:1285–1299.

4. Mazzarella L,Capasso S, Demasi D, Di Lorenzo G, Mattia C, Zagari A. (1993) Bovine seminal ribonuclease: structure at 1.9 A resolution. *Acta Crystallogrica Section D Biological Crystallography* 49:389–402.

5. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* 101(6):589–600.

6. Piccoli R, Di Donato A, D'Alessio G (1988) Co-operativity in seminal ribonuclease function. Kinetic studies. *The Biochemical Journal* 253:329–36.

7. Cafaro V, De Lorenzo C, Piccoli R, Bracale A, Mastronicola MR, Di Donato A, D'Alessio G (1995) The antitumor action of seminal ribonuclease and its quaternary conformations. *FEBS Letters* 359:31–4.

8. Park C, Raines RT (2000) Dimer formation by a "monomeric" protein. *Protein Science* 9:2026–33.

9. Gotte G, Bertoldi M, Libonati M (1999) Structural versatility of bovine ribonuclease A. Distinct conformers of trimeric and tetrameric aggregates of the enzyme. *European Journal of Biochemistry* 265:680–7.

10. Louie G V, Yang W, Bowman ME, Choe S (1997) Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor. *Molecular Cell* 1:67–78.

11. Bennett MJ, Sawaya MR, Eisenberg D (2006) Deposition diseases and 3D domain swapping. *Structure* 14:811–24.

12. Ekeowa UI, Freeke J, Miranda E, Gooptu B, Bush MF, Pérez J, Teekman J, Robinson CV, Lomas DA (2010) Defining the mechanism of polymerization in the serpinopathies. *Proceedings of the National Academy of Sciences USA* 107:17146–51.

13. Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nature Structural Biology* 8:316–20.

14. Janowski R, Abrahamson M, Grubb A, Jaskolski M (2004) Domain swapping in N-truncated human cystatin C. *Journal of Molecular Biology* 341:151–60.

15. Janowski R, Kozak M, Abrahamson M, Grubb A, Jaskolski M (2005) 3D domain-swapped human cystatin C with amyloidlike intermolecular beta-sheets. *Proteins* 61:570–8.

16. Nilsson M, Wang X, Rodziewicz-Motowidlo S, Janowski R, Lindström V, Onnerfjord P, Westermark G, Grzonka Z, Jaskolski M, Grubb A. (2004) Prevention of domain swapping inhibits dimerization and amyloid fibril formation of cystatin C: use of engineered disulfide bridges, antibodies, and carboxymethylpapain to stabilize the monomeric form of cystatin C. *Journal of Biological Chemistry* 279:24236–45.

17. Liu C, Sawaya MR, Eisenberg D (2011) β2-Microglobulin Forms Three-Dimensional Domain-Swapped Amyloid Fibrils With Disulfide Linkages. *Nature Structural and Molecular Biology* 18:49–55.

18. Lee S, Eisenberg D (2003) Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction-oxidation process. *Nature Structural Biology* 10:725–30.

19. Trinh CH, Smith DP, Kalverda AP, Phillips SEV, Radford SE (2002) Crystal structure of monomeric human beta-2-microglobulin reveals clues to its amyloidogenic properties. *Proceedings of the National Academy of Sciences U S A* 99:9771–6.

20. Wolf MG, Gestel J, de Leeuw SW (2008) Modeling amyloid fibril formation: a free-energy approach. *Methods in Molecular Biology* 474:153–179.

21. Rousseau F, Schymkowitz JWH, Itzhaki LS (2003) The unfolding story of three-dimensional domain swapping. *Structure* 11:243–51.

22. Rousseau F, Schymkowitz JWH, Wilkinson HR, Itzhaki LS (2001) Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proceedings of the National Academy of Sciences USA* 98:5596-5601.

23. Seeliger MA, Schymkowitz JWH, Rousseau F, Wilkinson HR, Itzhaki LS (2002) Folding and association of the human cell cycle regulatory proteins ckshs1 and ckshs2. *Biochemistry* 41:1202–10.

24. Green SM, Gittis AG, Meeker AK, Lattman EE (1995) One-step evolution of a dimer from a monomeric protein. *Nature Structural Biology* 2:746–751.

25. Chen YW, Stott K, Perutz MF (1999) Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proceedings of the National Academy of Sciences USA* 96:1257–61.

26. Barrientos LG, Louis JM, Botos I, Mori T, Han Z, O'Keefe BR, Boyd MR, Wloodawer A, Gronenborn AM. (2002) The domain-swapped dimer of cyanovirin-N is in a metastable folded state: reconciliation of X-ray and NMR structures. *Structure* 10:673–86.

27. Kuhlman B, O'Neill JW, Kim DE, Zhan KYJ, Baker D (2001) Conversion of monomeric protein L to an obligate dimer by computational protein design. *Proceedings of the National Academy of Sciences USA.* 98:10687-91.

28. Schymkowitz JWH, Rousseau F, Wilkinson HR, Friedler A, Itzhaki LS (2001) Observation of signal transduction in three-dimensional domain swapping. *Nature Structural Biology* 8:888-92.

29. Bourne Y, Arvai AS, Bernstein SL, Watson MH, Reed SE, Endicott JE, Noble ME, Johnson LN, Tainer JA (1995) Crystal structure of the cell cycle-regulatory protein suc1 reveals a beta-hinge conformational switch. *Proceedings of the National Academy of Sciences USA* 92:10232–6.

30. Zegers I, Deswarte J, Wyns L (1999) Trimeric domain-swapped barnase. *Proceedings of the National Academy of Sciences USA* 96:818–822.

31. Hayes M V, Sessions RB, Brady RL, Clarke AR (1999) Engineered assembly of intertwined oligomers of an immunoglobulin chain. *Journal of Molecular Biology* 285:1857–67.

32. Lees JPB, Manlandro CM, Picton LK, Tan AZE, Casares S, Flanagan JM, Fleming KG, Hill RB (2012) A designed point mutant in Fis1 disrupts dimerization and mitochondrial fission. *Journal of Molecular Biology* 423:143–58.

33. Jerala R, Zerovnik E (1999) Accessing the global minimum conformation of stefin A dimer by annealing under partially denaturing conditions. *Journal of Molecular Biology* 291:1079–1089.

34. Lomas DA, Evans DL, Finch JT, Carrell RW (1992) The mechanism of Z α_1-antitrypsin accumulation in the liver. *Nature* 357:605-607.

35. Yamasaki M, Li W, Johnson DJD, Huntington JA (2008) Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 455:1255–1258.

36. Sherratt AR (2011) Beyond the active site of the bacterial rhomboid protease : Novel interactions at the membrane to modulate function. PhD Thesis, University of Ottawa.

37. Freeman M (2008) Rhomboid proteases and their biological functions. *Annual Review of Genetics* 42:191–210.

38. Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* 444:179–183.

39. Sherratt AR, Blais DR, Ghasriani H, Pezacki JP, Goto NK (2012) Activity-Based Protein Pro filing of the Escherichia coli GlpG Rhomboid Protein Delineates the Catalytic Core. *Biochemistry* 51:7794-7803.

40. Lazareno-Saez C, Arutyunova E, Coquelle N, Lemieux MJ (2013) Domain swapping in the cytoplasmic domain of the Escherichia coli rhomboid protease. *Journal of Molecular Biology* 425:1127–1142.

41. Lohi O, Urban S, Freeman M (2004) Diverse substrate recognition mechanisms for rhomboids; thrombomodulin is cleaved by Mammalian rhomboids. *Current Biology* 14:236–241.

42. Kwok J (2012) Structure Determination of a Dimer Formed by the Bacterial Rhomboid Cytoplasmic Domain. Honours Thesis, Univeristy of Ottawa.

43. Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta* 1814:942–968.

44. Dempsey CE (2001) Hydrogen exchange in peptides and proteins using NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* 39:135–170.

45. Morin S (2011) A practical guide to protein dynamics from ^{15}N spin relaxation in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* 59:245–262.

46. Gryk M, Jardetzky O (1997) Flexibility and Function of the Escherichia coli trp Repressor. *Biological NMR Spectroscopy*, eds Markley J, Opella S (Oxford University Press), pp 29–49.

47. Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* 1751:119–139.

48. Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*. 1:2876–2890.

49. Sreerama N, Woody RW (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical Biochemistry*. 209:32-44.

50. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry* 287:252–260.

51. Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nature Protocols* 1:2527–2535.

52. John DM, Weeks KM (2000) van't Hoff enthalpies without baselines. *Protein Science* 9:1416–1419.

53. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. (2002) Current protocols in molecular biology. New York: John Wiley & Sons; 2002.

54. Robertson AD, Swint L (1993) Thermodynamics of unfolding for turkey ovomucoid third domain: Thermal and chemical denaturation. *Protein Science* 2:2037–2049.

55. Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27:8063–8068.

56. Delaglio F, Grzeseik S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* 6:277–293.

57. Johnson BA, Blevins RA (1994) NMR View: A computer program for the visualization and analysis of NMR data. *Journal of Biomolecular NMR* 4:603–614.

58. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins: Structure, Function, and Genetics* 17: 75-86.

59. Zhang Y-Z. (1995) Protein and Peptide Structure and Interactions Studied by Hydrogen Exchange and NMR. Ph.D. Thesis (Structural Biology and Molecular Biophysics, University of Pennsylvania).

60. Rule GS, Hitchens TK (2006) *Fundamentals of Protein NMR Spectroscopy*. Springer.

61. Zhu G, Xia Y, Nicholson LK, Sze KH (2000) Protein dynamics measurements by TROSY-based NMR experiments. *Journal of Magnetic Resonance* 2000 143:423–426.

62. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* 33:5984–6003.

63. Pollard TD (2010) A guide to simple and informative binding assays. *Molecular Biology of the Cell* 21:4061–4067.

64. Dempsey CE (2001) Hydrogen exchange in peptides and proteins using NMR spectroscopy. *Progress in Nuclear Magnetic Resononance Spectroscopy* 39:135–170.

65. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.

66. Czabotar PE, Westphal D, Dewson G, Ma S, Hockings C, Fairlie WD, Lee EF, Yao S, Robin AY, Smith BJ, Huang DCS, Kluck RM, Adams JM, Colman PM (2013) Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. *Cell* 152: 519-531.

67. Hsu Y-T, Youle RJ (1997) Nonionic Detergents Induce Dimerization among Members of the Bcl-2 Family. *Journal of Biological Chemistry* 21:13829-13834.

68. Magler I, Nüss D, Hauser M, Ferreira F, Brandstetter H (2010) Molecular metamorphosis in polcalcin allergens by EF-hand rearrangements and domain swapping. *FEBS Journal*. 277: 2598-2610.

69. Wells RC, Hill RB (2011) The cytosolic domain of Fis1 binds and reversibly clusters lipid vesicles. *PLoS One* 6:1-12.

70. Hawkins CA, de Alba E, Tjandra N (2005) Solution structure of human saposin C in a detergent environment. *Journal of Molecular Biology* 346:1381-1392.

71. Rossmann M, Schultz-Heienbrok R, Behlke J, Remmel N, Alings C, Sandhoff K, Saenger W, Maier T (2008) Crystal structures of human saposins C and D: implications for lipid recognition and membrane interactions. *Structure* 16:809-817.

72. De Alba E, Weiler S, Tjandra N (2003) Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles *Biochemistry* 42:14729-14740.

73. Liu L, Byeon I-JL, Bahar I, Gronenborn AM (2012) Domain swapping proceeds via complete unfolding: a ^{19}F- and ^{1}H-NMR study of the Cyanovirin-N protein. *Journal of the American Chemical Society* 134:4229–4235.

74. Greenfield NJ (2006) Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. *Nature Protocols* 1:2733–2741.

75. Anslyn EV, Dougherty DA (2006). *Modern Physical Organic Chemistry*. University Science Books.

76. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Hartl FU (2013) Molecular chaperone functions in protein folding and proteostasis. *Annual Review of Biochemistry* 82:323-355.

77. Horwich AL (2013) Chaperonin-mediated protein folding. *Journal of Biological Chemistry* 288:23622-23632.

78. Lin Z, Madan D, Rye HS (2008) GroEL stimulates protein folding through forced unfolding. *Nature Structural & Molecular Biology* 15:303-311.

79. Sharma S, Chakraborty K, Müller BK, Astola N, Tang Y-C, Lamb DC, Hayer-Harti M, Harti, FU (2008) Monitoring protein conformation along the pathway of chaperonin-assisted folding *Cell* 133:142-153.

80. Khodarahmi R, Yazdanparast R (2005) Fluorimetric study of the artificial chaperone-assisted renaturation of carbonic anhydrase: A kinetic analysis. *International Journal of Biological Macromolecules* 36:191-197.

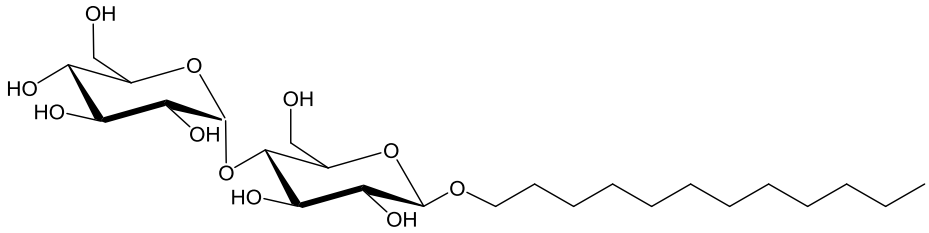
81. Morimoto N, Ogino N, Narita T, Akiyoshi K (2009) Enzyme-responsive artificial chaperone system with amphiphilic amylose primer. *Journal of Biotechnology* 140:246-249.

Appendix

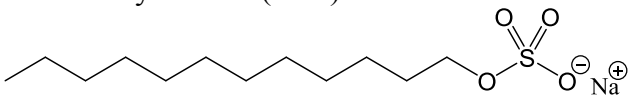
A.1 Tables

Table A.1: Chemical structures of detergents

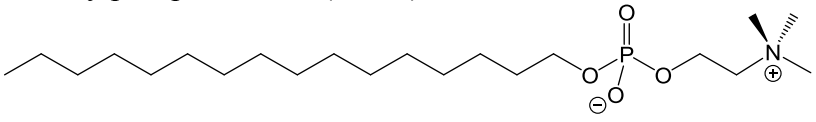
n-dodecyl-β-D-maltoside (DDM)



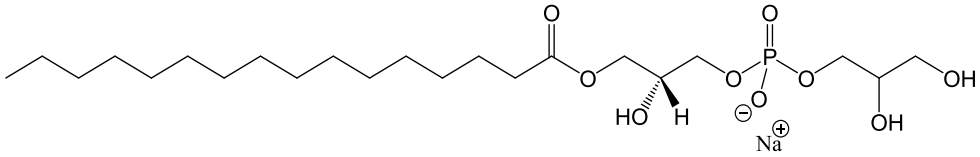
Sodium dodecyl sulfate (SDS)



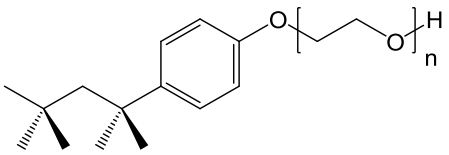
Hexadecylphosphocholine (Fos16)



1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LPPG)



Triton X-100



Octylphenoxypolyethoxyehtanol (Nonidet P40)