UNIVERSITY OF CALIFORNIA SAN DIEGO

Applications of Lipid Assemblies in Artificial Cell Development

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Chemistry

by

Ahanjit Bhattacharya

Committee in charge:

Professor Neal K Devaraj, Chair
Professor Stacey Brydges
Professor Simpson Joseph
Professor Ulrich F Müller
Professor Roger K Sunahara

2020

The Dissertation of Ahanjit Bhattacharya is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego

2020

DEDICATION

To the loving memory of my late grandfather

৺Sri Jagadis Banerjee (1938-2018)

Who taught me the values of

perseverance, and love for knowledge

EPIGRAPH

Mūkaṃ karoti vācālam

Paṅguṃ laṅghayate girim **|**

Yat kṛpā tamahaṃ

Vande Paramānanda Mādhavam **||**

You make the mute speak with eloquence, and the lame scale mountains by Your Divine Grace.

I bow to You, O Ever-Blissful Lord!

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AMP	adenosine monophosphate
ATP	adenosine triphosphate
β-ME	2-mercaptoethanol
BODIPY	boron dipyrromethene
BSA	bovine serum albumin
Cac	critical aggregation concentration
CcO	cytochrome c oxidase
CD	circular dichroism
Cmc	critical micelle concentration
CoA	coenzyme A
DAGK	diacylglycerol kinase
DDA	dodecanoic acid
DDM	*n*-dodecyl-*β*-D-maltopyranoside
DHPE	1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine
DLS	dynamic light scattering
DMEM	Dulbecco's modified Eagle's medium
DMF	*N*, *N*-dimethylformamide
DMSO	dimethyl sulfoxide
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DSC	differential scanning calorimetry
DTT	dithiothreitol
EDP	electron density profile
EDTA	ethylenediamine tetraacetic acid
ELSD	evaporative light scattering detector
FAA	fatty acyl adenylate
FRAP	fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
GOA	*N*-oleoyl *β*-D-galactopyranosylamine
GFP	green fluorescent protein

GUV	giant unilamellar vesicle
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HL	histidine ligation
HPLC	high-performance liquid chromatography
HPTS	8-hydroxypyrene-1,3,6-trisulfonate (trisodium salt)
IPTG	isopropyl *β*-D-1-thiogalactopyranoside
MeOH	methanol
MS	mass spectrometry
MWCO	molecular weight cutoff
NCL	Native Chemical Ligation
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
OG	*n*-octyl *β*-D-glucospyranoside
OTG	*S*-oleoyl β-D-1-thiogalactopyranoside
PBS	phosphate-buffered saline
PC	phosphocholine
PDMS	poly(dimethylsiloxane)
PMSF	phenylmethysulfonyl fluoride
POPC	1-palmitoyl-2-dioleoyl-*sn*-glycero-3-phosphocholine
PURE	protein synthesis using recombinant elements
rt	room temperature
SAXS	small-angle X-ray scattering
SEM	scanning electron microscopy
TCEP	tris(2-carboxyethyl) phosphine
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TGA	thermogravimetric analysis
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)aminomethane
TX-TL	transcription-translation
XRD	X-ray diffraction
XRR	X-ray reflectivity

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNWOLEDGEMENTS

When I was in my school, I first read about the Stanley Miller Experiment, where he synthesized few key building blocks of life by passing electricity through a mixture of atmospheric gases likely to have been present in the atmosphere of a primitive Earth. I always found this story highly fascinating and developed a great curiosity to know what defines life and what makes it different from non-living matter. It is this curiosity which motivated me to come to UC San Diego and work on artificial cells. Thank you Prof. Neal Devaraj for the great opportunity you gave me to explore the definition of life. It was amazing to work under a charismatic scientific leader like you. Thank you for always being encouraging when I came with new ideas and giving me the freedom to test it out. It truly helped to build my confidence. Thank you for your constructive criticism and attention to details. That truly helped me to organize my thoughts and think about science critically. Thank you for pushing me sometimes to 'go a little bit further' and step outside of comfort zone. It truly helped me to test my mettle. Thank you for teaching me not to lose hope when papers or grants got rejected. Those experiences will help me to be strong in the future and guide me about how to make a come-back and re-strategize in the event of a failure. Thank you for advising me to "knock others' doors" when stuck at some research problems. I found so many collaborations (big or small) ever since I started to do so. Thank you for your unorthodox approaches to scientific problems and looking at the 'big picture'. These lessons will be immensely valuable for me to fulfil my future academic aspirations.

I am fortunate to have a thesis committee with the professors who were extremely inspiring and encouraging to me. Thank you Prof. Uli Müller for your valuable advice and guidance all these years. I truly cherish the three weeks I worked in your lab as a rotation student. It may be a short period, but the experience left a deep impression on me about how to plan and execute research

meticulously. I immensely enjoyed the intellectually stimulating conversations of the origins of life with you. Thank you Prof. Simpson Joseph for your constant encouragement, especially during the early years of my PhD. There have been several times when I was stuck on a research problem and you shared your experience and knowledge with great patience to help me out. I was very happy to collaborate with your lab on the MURI projects. Thank you Prof. Roger Sunahara for being such an inspiring person. I had a few stimulating scientific conversations with you without even realizing that hours had passed. It was very motivating for a young person like me to watch you do the experiments yourself to help me answer my questions! Thank you Prof. Stacey Brydges for your valuable advice on improving my presentation skills.

Thank you Prof. Sunil Sinha for steering me into new directions in science. Had I not been involved in collaboration with your lab, a vast scientific landscape would probably have been unfamiliar to me. Through interaction with you I was able to appreciate the beauty and power of structural methods in chemistry and biology and I aspire to continue on that path in the future. I have definitely learnt much more effectively about X-ray physics just by discussing with you than reading textbooks and articles. I am always deeply inspired by your dedication for research and attention for fine details. I want to recall the time that you stayed with us at the beamtime in Argonne all day through 1 AM and we had to remind you to get some rest finally! Examples like this surely sets a benchmark about how much diligence we should put in my work in the coming days. Thank you Prof. Gourisankar Ghosh for inspiring me to be a good scientist. I will never forget the valuable advice you gave me on finding true purpose of life as a scientist. Thank you Prof. Mike Burkart for letting me collaborate with your lab on fatty acid synthases. Thank you Prof. Francesco Paesani for your amiable presence. It is amazing to see how caring you are about every student in the department and how well you appreciate their every small achievement. I felt

every word of appreciation from you as a big pat on my back. There is so much to learn from you about how to be a role model to students as a professor. Thank you Prof. Itay Budin for your wonderful work during your PhD which formed the basis of so many projects in our lab. It is always very refreshing discussing science with you.

I am extremely thankful to work for more than five years in a lab with a strong culture of cooperation and collaboration. My first introduction to the research at Devaraj Lab was done by Roberto Brea. Thank you Roberto for being a wonderful mentor and colleague! I admire your meticulous approach to experiments and attention to details. I admire your kindness, patience, and eagerness to teach things in the 'right way'. It was always fun to chat with you on 'random' topics in science. Thank you Henrike Niederholtmeyer for being a great colleague in the lab and an equally great friend outside the lab. Through our scientific partnership, I could learn so many things from you, not only biochemistry, but the overall approach to research too. You are truly inspiring! Thank you Brandon Cisneros for being an amazing colleague and friend for five years. I really appreciate your honest feedback to me on many matters. I share your enthusiasm for science, history, philosophy, and any other geeky stuff. I appreciate your highly unorthodox approach to everything in life. I am confident you will make a difference in this world. Thank you Andrew Rudd for being a great colleague and friend and for helping me in the lab in so many ways. I admire your passion for science and your determined nature. Thank you Christian Cole for being one of the 'founders' of the Devaraj Lab. You are a great mentor and thanks for your help on so many occasions, especially when I needed it the most. Along with you, everyone in the 'old cohort' – Michael Hardy, Seth Alexander, Eric Zhou, Haoxing Wu, and Fabian Ehret were fantastic lab citizens. Thank you Kira for being a great friend and enthusiastic coworker. I highly admire your I(q) and TEMperament. Thank you Mai Johnson for your accidental discovery that led us to work

together on a whole paper! Your enthusiasm is highly ‘contagious’ and it was a great pleasure to work with you and share the office as well. Thank you Luping Liu for being a great colleague and friend. The Mono Lake trip for our paper will be a highly memorable experience for me. Thank you Andrés Seoane for being great coworker and friend. Arguing on any topic has never been more fun than with you. Thank you Christy Cho and Satyam Khanal for being wonderful mentees and also for working with me on several papers. Thank you Kayla Busby, Dongyang Zhang, Shuaijiang Jin, Hetika Vora, Ember Tota, Judith Flores, Youngjun Lee, and Jorge Jiminez for wonderful memories both within and outside the lab. Finally, my sincere thanks goes to Rupak Bhattacharya and Jing-Jin Song for your fantastic efforts and collaboration on the X-ray projects.

PhD is not only about laboratory research and academic rigor. The life outside this realm is equally important in shaping an individual on this long journey. Staying far away from my country, family, and culture has never been easy, yet I never felt alone or isolated. Words will be insufficient to express how much grateful I am to people in my life, both family and friends who supported me all these years. My biggest support has been my little sister Kiran who has been with me through every moment of troughs and crests, even if she is eight thousand miles away. I want to thank my parents and grandparents for their constant warmth and encouraging words that I should make this long ‘absence’ worthy of it. It is difficult to appreciate how much sacrifice one’s family makes being far apart for so many years. I am thankful to my only relatives in USA - uncle Tamal Roy, aunt Archita Roy, and cousins Trinava and Anish who live in New Jersey for inviting me to their place few times during my PhD.

Now I will spend a dedicated paragraph on the people who made my time at San Diego highly memorable, if not one of the best phases of my life. I am blessed to be in close company with a wonderful group of friends without whom life would not have been the same. They are not

only extremely talented individuals with a remarkably intellectual side, but also they were always there to share every bit of joy or sadness. I will always cherish the great camaraderie between Sridip, Shouvik and Brato through weekend parties, road trips, hiking, watching and playing sports and every other activity however big or small. Through our mutual contacts, we became friends with Diptarka-da, Sumit, Siddhartha-da, Rashmi, Athira, Madhukar, Arijit, Abhik-da, Soumen-da, Antu, Nirjhar, Bodhisattwa, and Rourav. I will deeply cherish our Friday gatherings at 9226A Regents Rd. and later at 9647 Easter Way for all my life. Also, all the fantastic trips I made with you will be extremely special to me. Finally, I will give special thanks to Netai Mama, Anindita Mami and Orion for welcoming us with a warm heart and sharing so much love and affection.

Notes about the chapters

Chapter 1 is adapted partly adapted from materials published in the *Perspective* articles Bhattacharya, A.; Brea, R. J. & Devaraj, N. K. De novo vesicle formation and growth: an integrative approach to artificial cells. *Chem. Sci*. 2017, 8, 7912-7922 and Bhattacharya, A. & Devaraj, N. K. Tailoring the Shape and Size of Artificial Cells. *ACS Nano*, 2019, 13, 7396-7401. The dissertation author was a co-first author on the first manuscript and first author on the later.

Chapter 2 is adapted almost entirely from the materials published in Brea, R. J., Bhattacharya, A. & Devaraj, N. K. Spontaneous Phospholipid Membrane Formation by Histidine Ligation. *Synlett*, 2017, 28, 108-112. The dissertation author was a co-first author on this manuscript.

Chapter 3 is adapted majorly from the materials published in Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A Minimal Biochemical Route towards de novo Formation of Synthetic Phospholipid Membranes. *Nat. Commun*. 2019, 10, 300. There is a significant amount of unpublished material as well, especially those related to FadD2 coauthored with Cho, C. J.;

Brea, R. J. & Devaraj, N. K. and those on sphingomyelin coauthored with Devaraj, N. K. The dissertation author was the primary investigator for works described in this chapter.

Chapter 4 is adapted almost entirely from the materials published in Brea, R. J., Bhattacharya, A., Bhattacharya, R., Song, J., Sinha, S. K. & Devaraj, N. K. Highly Stable Artificial Cells from Galactopyranose-derived Single Chain Amphiphiles. *J. Am. Chem. Soc.*, 2018, 140, 17356-17360. The dissertation author was a co-first author on this manuscript.

Chapter 5 is adapted majorly from materials published in Bhattacharya, A., Brea, R. J., Bhattacharya, R., Song, J., Sinha, S. K. & Devaraj, N. K. Single-chain β-D-Glycopyranosylamides of Unsaturated Fatty Acids: Self-assembly Properties and Applications to Artificial Cell Development. *J. Phys. Chem. B.*, 2019, 123, 3711-3720. The dissertation author is the primary investigator on these works. This chapter also contains materials from a manuscript Johnson, M., Bhattacharya, A., Brea, R. J., Podolsky, K. A. & Devaraj, N. K. Temperature-dependent reversible Morphological Transitions in *N*-oleoyl β-D-galactopyranosylamine. This manuscript is currently under review in *J. Phys. Chem. B* and the dissertation author is a co-first author.

Chapter 6 is adapted majorly from materials which appear in Bhattacharya, A., Niederholtmeyer, H., Podolsky, K. A., R. J., Bhattacharya, R., Song, J., Brea, R. J., Hsien-Tsai, C., Sinha, S. K. & Devaraj, N. K. Lipid Sponge Droplets as Programmable Synthetic Organelles, *bioRxiv*, 2020, DOI: https://doi.org/10.1101/2020.01.20.913350. This manuscript is currently submitted to *PNAS*. The dissertation author is the co-first author in this work. This chapter also contains small amount of unpublished data on lectin-induced droplet structuration coauthored with Devaraj, N. K

VITA

Education

PhD in Chemistry (2014-2020)

University of California San Diego

Integrated MSc. in Chemistry (2009-2014)

Indian Institute of Technology Kharagpur

Publications

1. **A. Bhattacharya***, H. Niederholtmeyer*, K. A. Podolsky, R. Bhattacharya, C. H. Tsai, R. J. Brea, S. K. Sinha, N. K. Devaraj. Lipid Sponge Droplets as Programmable Synthetic Organelles. *bioRxiv*, **2020**, DOI: https://doi.org/10.1101/2020.01.20.913350 (*equal contribution, submitted to *PNAS*)
2. L. Liu, Y. Zou, **A. Bhattacharya**, D. Zhang, S. Q. Lang, K. N. Houk, N. K. Devaraj. Enzyme-free synthesis of natural phospholipids in water. *bioRxiv*, **2019**, DOI: https://doi.org/10.1101/746156. (in revision for *Nature Chemistry*)
3. **A. Bhattacharya**, N. K. Devaraj. Tailoring the Shape and Size of Artificial Cells. (Perspective) *ACS Nano*, **2019**, 13, 7396-7401.
4. **A. Bhattacharya**, R. J. Brea, J. Song, R. Bhattacharya, S. K. Sinha, N. K. Devaraj. Single-chain β-D-Glycopyranosylamides of Unsaturated Fatty Acids: Self-assembly Properties and Applications to Artificial Cell Development. *J. Phys. Chem. B*, **2019**, 123, 3711-3720.
5. **A. Bhattacharya**, R. J. Brea, H. Niederholtmeyer, N. K. Devaraj. A Minimal Biochemical Route towards de novo Formation of Synthetic Phospholipid Membranes. *Nat. Commun.* **2019**, 10, 300.
6. R. J. Brea*, **A. Bhattacharya***, R. Bhattacharya*, J. Song, S. K. Sinha, N. K. Devaraj. Highly Stable Artificial Cells from Galactopyranose-derived Single Chain Amphiphiles. *J. Am. Chem. Soc.* **2018**, 140, 17356-17360. (*equal contribution)
7. T. Enomoto*, R. J. Brea*, **A. Bhattacharya**, N. K. Devaraj. In Situ Lipid Membrane Formation Triggered by an Intramolecular Photoinduced Electron Transfer. *Langmuir*, **2018**, **34, 750-755.** (*equal contribution)
8. **A. Bhattacharya***, R. J. Brea*, N. K. Devaraj. De novo vesicle formation and growth: an integrative approach to artificial cells. (Perspective) *Chem. Sci*. **2017**, 8, 7912-7922. (*equal contribution)
9. C. Chen, P. Kattanguru, O. A. Tomashenko, R. Karpowicz, G. Siemiaszko, **A. Bhattacharya**, V. Calasans, Y. Six. Synthesis of functionalised azepanes and piperidines from bicyclic halogenated aminocyclopropane derivatives. *Org. Biomol. Chem.* **2017**, 15, 5364-5372.
10. R. J. Brea*, **A. Bhattacharya***, N. K. Devaraj. Spontaneous Phospholipid Membrane Formation by Histidine Ligation. *Synlett*, **2017**, 28, 108-112. (*equal contribution)

11. P. S. Addy, **A. Bhattacharya**, S. M. Mandal, A. Basak. Label assisted laser desorption/ionization mass spectrometry (LA-LDI-MS): an emerging technique for rapid detection of ubiquitous *cis*-1,2-diol functionality. *RSC Adv*., **2014**, 4, 46555.

Awards and scholarships

1. *Syncell2020 Travel Scholarship* from the International Conference on Engineering Synthetic Cells & Organelles committee.
2. *Luna Fung Scholarship 2019* from the Friends of International Center, University of California San Diego for scientific research achievements.
3. *Teddy Traylor Award 2017-18* from the Department of Chemistry and Biochemistry, University of California San Diego for scientific research achievements.
4. *Prime Minister of India Award 2014* for the Best Academic Performance in the IIT Kharagpur among all departments.
5. *Khorana Program for Scholars Award 2012* from the Indo-US Science and Technology Forum (IUSSTF) for research internship (May-July 2012) at the University of Minnesota, Twin Cities.
6. *Kishore Vaigyanik Protsahan Yojana (KVPY) Fellowship* awarded by the Department of Science and Technology, Government of India during 2008-2014.
7. *Gold Medal of Merit in Indian National Biology Olympiad* awarded by the Department of Science and Technology (DST), Govt. of India in 2009.

Presentations (Conference talks)

1. "Single-chain Synthetic Glycolipids: Self-assembly Properties and Applications to Artificial Cell Development" at the Biological Chemistry Graduate Student and Postdoctoral Fellow Symposium in 258th ACS National Meeting (25-29 August 2019) at San Diego, CA.

2. "Enzymatic *de novo* Synthesis of Artificial Phospholipid Membranes" at the Synthetic Biology: Engineering Evolution and Design 2018 (SEED18) Conference (3-7 June 2018) held at Scottsdale. AZ.

3. "Enzymatic Strategies of *de novo* Phospholipid Membrane Formation" at the Biomembrane Synthesis, Structure, Mechanics and Dynamics Symposium in 255th ACS National Meeting (18-22 March 2018) at New Orleans, LA.

4. "Linking Gene Expression with *de novo* Phospholipid Membrane Formation" at the Self Assembly of Synthetic and Biological Surfactants Symposium in 254th ACS National Meeting (20-24 August 2017) at Washington DC.

5. "Spontaneous Phospholipid Membrane Formation by Histidine Ligation" at the Biomembrane Symposium 253rd ACS National Meeting (2-6 April 2017) at San Francisco, CA.

ABSTRACT OF THE DISSERTATION

Applications of Lipid Assemblies in Artificial Cell Development

by

Ahanjit Bhattacharya

Doctor of Philosophy in Chemistry

University of California San Diego, 2020

Professor Neal K Devaraj, Chair

It is of great fundamental interest to develop experimental model systems to reconstruct design principles underlying formation of living cells and shed light on how life emerged on Earth and possibly elsewhere. An ambitious strategy is the bottom-up approach, which aims to systematically control the assembly of basic building blocks with defined functionality to construct a life-like entity. This dissertation will explore the roles of miscellaneous lipid assemblies as building blocks of artificial cellular compartments.

The first part (Chapters 2 and 3) of this dissertation shows how lipid environments facilitate chemoselective reactions between amphiphilic partners through physical partitioning. Through examples of histidine ligation, and acyl phosphate chemistry we show functionalized micelle-

forming amphiphiles react to form amidophospholipids which self-assemble into micron-sized vesicles.

The second part (Chapter 3) of the dissertation seeks to understand the origins of cellular phospholipid synthesis pathways by repurposing soluble fatty acid activating enzymes (fatty acyl adenylate ligase and fatty acyl CoA ligase) to synthesize phospholipids. Such simplified biochemical pathways may provide a hint at how lipids were likely synthesized in a minimal protocell without the necessity of transmembrane protein enzymes.

The third part (Chapters 4 and 5) of the dissertation discusses the vesicle formation from a novel class of single-chain amphiphiles derived from galactopyranosyl head groups and unsaturated fatty acid tails. The geometric parameters and thermodynamic properties of these amphiphilic assemblies are characterized by a host of physical techniques. The vesicles are further shown to sustain model biochemical reactions.

The final part (Chapters 6) of the dissertation describes the formation of sponge phase droplets from fatty acyl galactopyranosylamides and non-ionic detergents. The droplets contain a dense bicontinuous network of bilayers and nanometric aqueous channels, which facilitates partitioning of molecules based on their size, polarity, and specific binding motifs. The sequestration of biomolecules can be programmed by doping the droplets with suitably functionalized amphiphiles. The droplets can harbor functional soluble and transmembrane proteins, allowing for the co-localization and concentration of enzymes and substrates to enhance reaction rates. Droplets protect bound proteins from proteases, and these interactions can be engineered to be reversible and optically controlled.

Chapter 1 | General Introduction to Artificial Cells

1.1 Synthetic Organic Chemistry and Synthetic Biology

In the middle of the 19th century, a trend started in chemistry, namely, *to make new organic molecules*. The trend was set by Friedrich Wöhler in 1828, when he famously synthesized the organic substance urea by heating the inorganic substance ammonium cyanate and thereby putting an end to the idea of *vitalism*. Soon, Hermann Kolbe followed the same path and synthesized acetic acid from carbon disulfide in 1845 (Interestingly it was Kolbe who first used the term 'synthesis' for describing chemical transformations). Marcellin Berthelot - a prolific chemist of this era was the first to show that organic molecules which do not exist in nature can be made inside a laboratory. These seminal early works built the foundations of synthetic organic chemistry, ironically even before the concepts of fundamental particles, chemical bonds and quantum theory existed. The zeal for synthesizing bigger and complex molecules continued and within 100 years or so, some astonishing feats were achieved. In 1957, John Sheehan reported the synthesis of the then wonder drug Penicillin V.[1] In 1971, Robert Merrifield published the synthesis of a full functional protein ribonuclease A.[2] In 1972, Woodward and Eschenmoser completed the humongous task of synthesis of Vitamin B_{12} in collaboration with 91 postdoctoral scholars and 12 doctoral students. Har Gobind Khorana synthesized the complete gene of the tyrosine suppressor tRNA in 1979.[3] With all these remarkable achievements it seemed likely that it may be possible to build up further in biological complexity by harnessing the power of chemistry. When combined with engineering and materials science, a formidable force is generated. Thus, the 21st century is seeing a highly ambitious effort, namely, *to build life-like entities in a laboratory*. This pursuit of building a so called artificial cell or minimal cell from basic building blocks forms one of the pillars of bottom-up synthetic biology.

1.2 Synthesizing Life

The idea of synthetic life was being considered (at least theoretically) as early as the middle of the 20th century. Alexander Oparin, the father of 'Primordial Soup' Theory asserted that there is no fundamental difference between non-living matter and living organisms. In the 1944 classic "*What is Life?*" Erwin Schrödinger described life in the qualitative sense of entropy and non-equilibrium. In 1971, the theoretical biologist Tibor Gánti put forward his abstract Chemoton Model to describe the chemical origins of life. This model states that a living system (evolutionary unit) should display three fundamental features – lipid compartment, metabolism, and self-replication. The metabolic and self-replication functionalities constitutes an autocatalytic subsystem. In 2001, three eminent origins of life scientists Jack W, Szostak, David Bartel and Pier Luigi Luisi published a highly influential article in *Nature* entitled "Synthesizing Life".[4] Here they posited that a living system is a union of two kinds of replicating systems: (i) informational substance (genome) (ii) compartment material (lipid membrane). This suggestion has an important distinction from the Chemoton Model that here replication of the lipid compartment has been considered as well. Over the last two decades, various models of protocells have attempted to couple growth of the compartment with metabolic reactions, albeit with variable degree of success. Indeed, there are more examples where these two issues have been considered independently.

1.3 A Lipid World

Self-assembly is central to the very organization of life. Every living cell is bound by a highly dynamic and semi-permeable cell membrane, which is mainly composed of amphiphilic phospholipid molecules. In aqueous media, phospholipids spontaneously self-assemble into membrane-bound structures called vesicles. This process is largely driven by hydrophobic

interactions between the non-polar fatty acyl chains of the phospholipid molecules, which segregates them from the aqueous bulk solution (Figure 1.1).

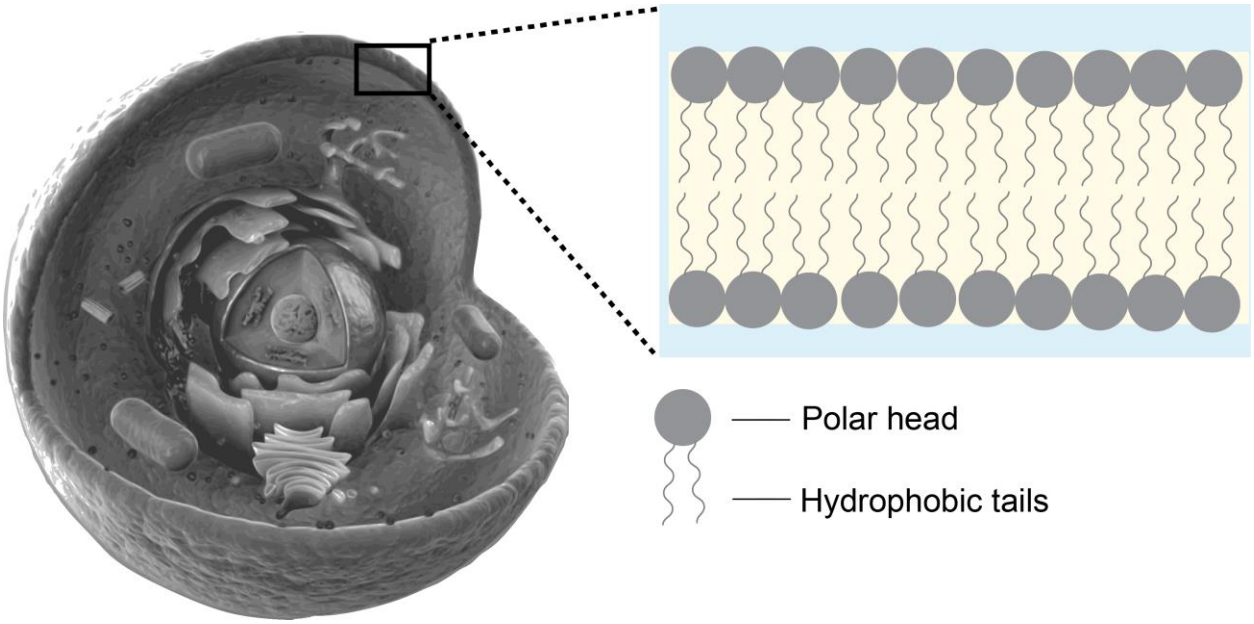


Figure 1.1. Cellular compartments are formed by amphiphilic lipid molecules. Phospholipids are composed of a polar head region and hydrophobic tail region which spontaneously self-assemble in water to form a bilayer membrane.

However, membranes are also known to form from simpler single-chain amphiphiles such as fatty acids. In the context of origins of life, single-chain amphiphiles are thought to have played a founding role in forming the earliest membranes. There is a significant debate over whether protocellular compartments were formed first or whether informational/catalytic molecules (for example, RNA) were formed first. Although the debate is mostly in favor of an RNA World view, an alternate scenario called Lipid World[5] has been considered. This view considers that even though lipids and amphiphiles are typically considered as inert structural molecules, they might have played a more significant role via physical catalysis. Lipid assemblies may have served as sites where molecules were concentrated from a dilute primordial soup and therefore enabled building of chemical reaction networks. Moreover, the Graded Autocatalysis Replication Domain (GARD) Model put forward by Doron Lancet considers that lipid assemblies carried *compositional*

information which could be propagated and therefore served as earliest evolutionary units. Therefore, in the development of artificial cells, both structural and functional aspects of lipid assembles need consideration.

1.4 On shape and form

Appreciation for shapes and sizes of living cells began as early as the invention of the microscope by Anton von Leeuwenhoek, and perhaps preceded any understanding of their function. Over time, it was realized that there is a strong correlation between the structure and function of living cells and the phrase "function presupposes structure" was considered to be axiomatic in biology.[6] So, it comes without surprise that in the turn of the 21st century, scientists will have similar appreciation for shape and form while designing an artificial cell. Lipid molecules spontaneously self-assemble to form spectacular-looking structures such as one shown in Figure 1.2 that resembles organelles in cell. Unfortunately, the outcomes of lipid self-assemblies such as the one presented here are mostly random and specific interventions or considerations will be necessary to achieve a well-defined structure deterministically.

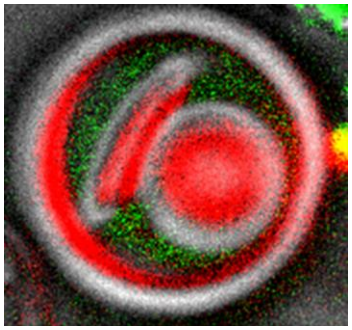


Figure 1.2. A random outcome of lipid self-assembly.

Therefore, the question arises, how can artificial cells be built with specific shapes and sizes? Let us consider a dichotomy first. Some objects have a given shape and size because they take the shape of the container in which they were molded, while other objects are inherently programmed to have so. Which of these two comes first, particularly in the context of artificial

cell design? Bottom-up synthetic biologists attempt to address this question by building cell-mimetic systems that obey either of these two paradigms. The first approach is to build artificial cellular compartments through mechanical manipulation of the building blocks, in a manner analogous to casting, forging, or milling objects in a manufacturing industry. This can be thought of as "container first" strategy. The alternate and perhaps less explored approach is based on the informational code inherent in the building blocks that allows them to be assembled in a predisposed manner. This can be thought of as "code first" strategy.

1.5 On growth and proliferation

To mimic the basic features of living cells, growth and proliferation of the artificial cellular compartment is necessary. Growth of membranes composed of single-chain amphiphiles like fatty acids has been demonstrated simply by addition of fresh monomers. However, owing to the less dynamic nature of phospholipid membranes, *in situ* generation of membrane-forming material is required for their growth. This can be accomplished either by a membrane-embedded catalyst or by a soluble catalyst encapsulated within the vesicle lumen.

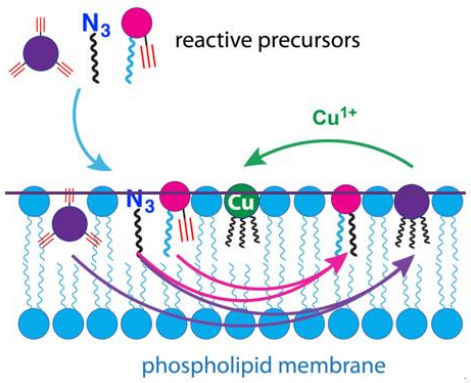


Figure 1.3. Self-reproducing system that can drive the repeated synthesis and growth of phospholipid membranes by using a CuAAC reaction. The regeneration of membrane-bound autocatalysts continually induces the formation of triazole phospholipids, mimicking natural membrane generation.

Simple autocatalytic reactions have been utilized in the past to drive the growth of membrane-bound vesicles. For instance, the growth of oleic acid vesicles has been shown to be driven by the autocatalytic hydrolysis of oleic anhydride droplets. Seeking a catalytic strategy to

obtain vesicle self-reproduction, a self-replicating catalyst, which can continuously drive phospholipid formation by click reaction over many generations has been described (Figure 1.3).[7] In principle, such a system should be able to sustain efficient membrane reproduction if the precursors are replenished. Design of an artificial cell needs to take this idea one step further, namely, to replicate genetic material along with membrane growth. This dissertation deals with approaches designed to develop lipid compartments with robust structure and dynamic behavior.

Notes on chapter

Chapter 1 is adapted partly adapted from materials published in the *Perspective* articles Bhattacharya, A.; Brea, R. J. & Devaraj, N. K. De novo vesicle formation and growth: an integrative approach to artificial cells. *Chem. Sci*. 2017, 8, 7912-7922 and Bhattacharya, A. & Devaraj, N. K. Tailoring the Shape and Size of Artificial Cells. *ACS Nano*, 2019, 13, 7396-7401. The dissertation author was a co-first author on the first manuscript and first author on the later.

1.6 References

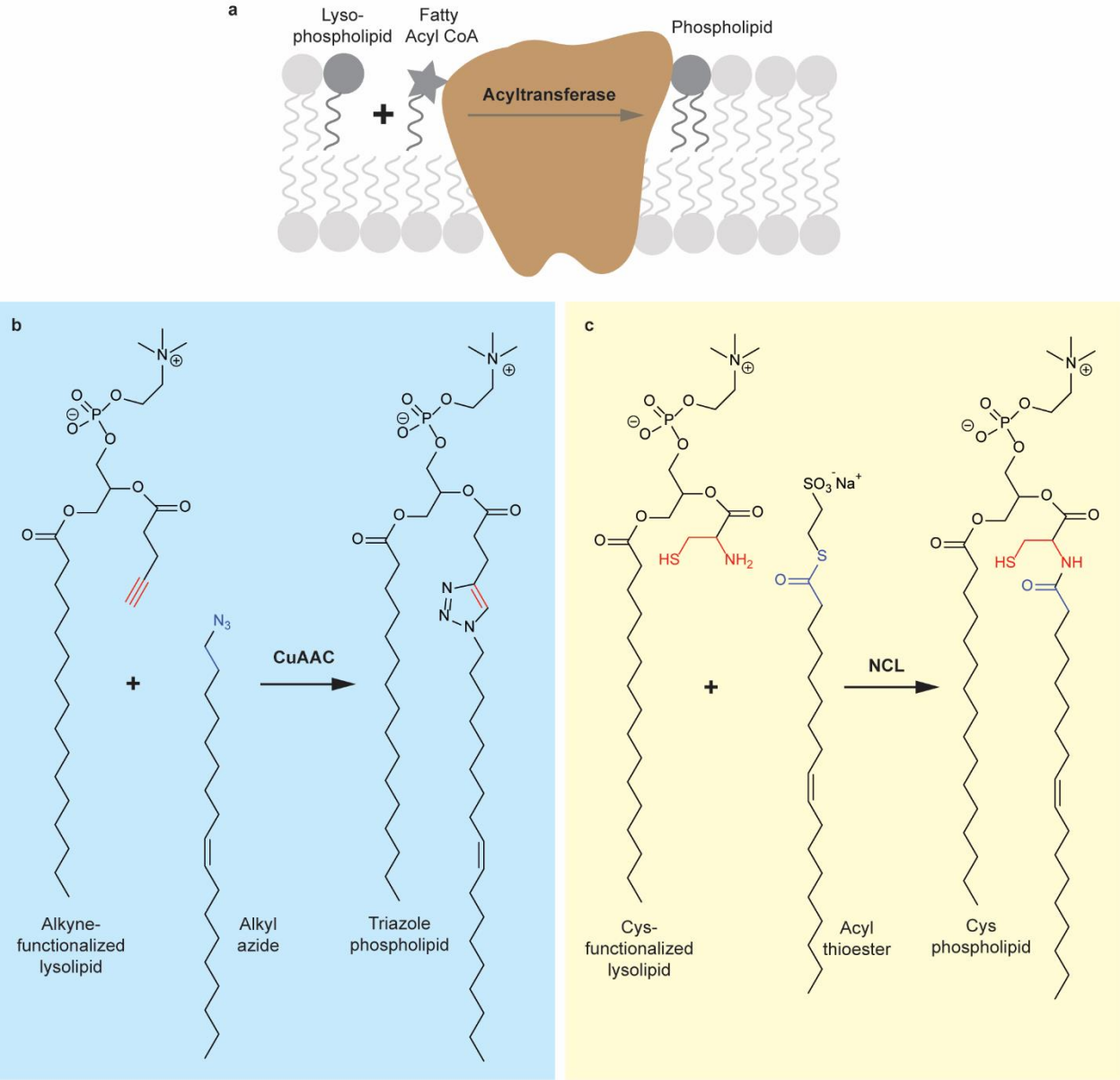
1. Sheehan, J. C. & Henbry-Logan, K. R. The total synthesis of penicillin V. *J. Am. Chem. Soc.* **79,** 1262–1263 (1957).
2. Gutte, B. & Merrifield, R. B. The synthesis of ribonuclease A. *J. Biol. Chem.* **246,** 1922–1941 (1971).
3. Khorana, H. G. Total synthesis of a gene. *Science* **203,** 614–625 (1979).
4. Szostak, J. W., Bartel, D. & Luisi, P. L.. Synthesizing life. *Nature* **409,** 387–390 (2001).
5. Segré, D., Ben-Eli, D., Deamer, D. W. & Lancet, D. The Lipid World. *Orig. Life Evol. Biosph.* **31,** 119–145 (2001).
6. Thompson, D. W. On Growth and Form. *Cambridge at the University Press*, (1942) pp 290.
7. Hardy, M. D., Yang, J., Selimkhanov, J., Cole, C. M., Tsimring, L. & Devaraj, N. K. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Nat. Acad. Sci*.**112**, 8187-8192 (2015).

Chapter 2 | Chemoselective Synthesis of Membrane-forming Phospholipids by Histidine Ligation

2.1 Introduction

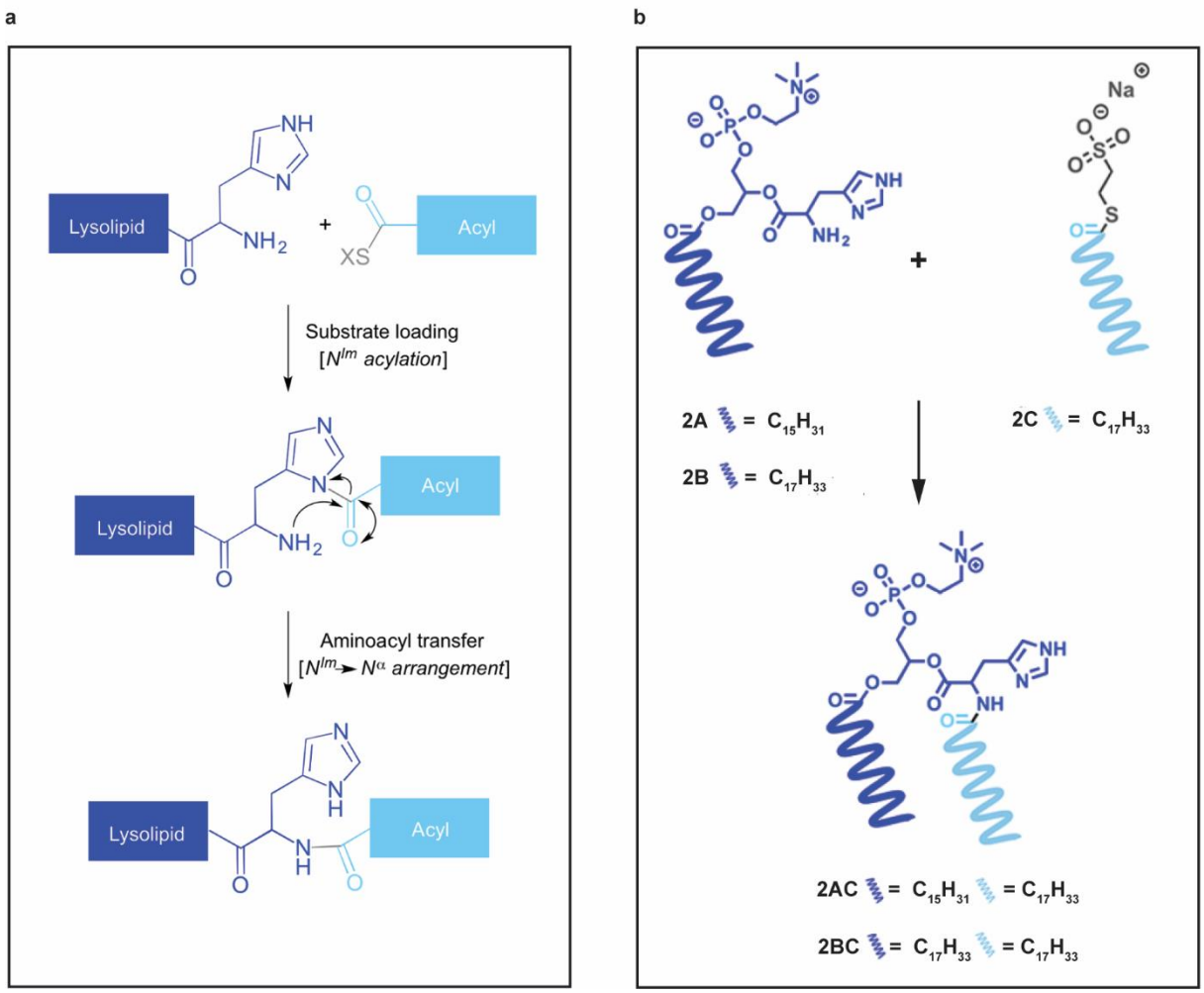
Living organisms carry out the *de novo* synthesis of lipid membranes by enzymatic acyl transfer reactions that couple the single-chain amphiphiles lysophospholipids and fatty acyl CoA's into membrane-forming phospholipids (Scheme 2.1a).[1] However, the acyltransferase enzymes are transmembrane proteins which can be challenging to integrate in an artificial cellular system. Therefore, efficient synthetic methodologies mimicking this biochemical coupling step offer a general route towards artificial membrane-forming lipids. Over the past decade, Devaraj group has described several biomimetic strategies for synthesis of membrane-forming phospholipids. In these methods, a functionalized lysophospholipid is reacted with another functionalized lipid fragment via chemoselective ligation reactions such as copper-catalyzed azide-alkyne cycloaddition (CuAAC)[2] (Scheme 2.1b) and native chemical ligation (NCL) (Scheme 2.1c).[3] The reactions proceeded extremely rapidly and quantitatively to yield functionalized phospholipids that self-assembled to form membrane-bound vesicles. The extremely high selectivity and reaction rates can be attributed to the fact that the reactive partners are present in micellar or bilayer membrane environments, that increases the effective concentrations by several orders of magnitude. These previous studies inspired us to reason that amphiphilic assemblies could offer a general means to drive otherwise inefficient reactions to high conversion.

Coupling between N-terminal histidine and C-terminal thioester, also known as Histidine Ligation (HL) has been described as a method for ligating two peptide fragments.[4,5] The proposed mechanism of the HL involves a two-step process starting with an acyl-exchange step between a



Scheme 2.1. Biomimetic synthesis of membrane-forming phospholipids using chemoselective ligation strategies. a. Schematic representation of coupling reaction between lysophospholipid and fatty acyl CoA in biological membranes (Land's Cycle). Synthesis of membrane-forming phospholipids based on b. click reaction. c. Native Chemical Ligation.

long-chain acyl thioester and the N^{Im} moiety of a histidine residue, followed by an intramolecular nucleophilic attack by the α-amino group of the histidine ($N^{Im} \rightarrow N^{\alpha}$ acyl rearrangement) to form the final amide bond (Scheme 2.2a). Unlike NCL, HL is a less popular strategy for orthogonal peptide ligation because of its low efficiency.[4] Here we demonstrate that the chemoselective HL reaction can be utilized to drive the biomimetic synthesis of amidophospholipids from histidine-



Scheme 2.2. Phospholipid formation by Histidine Ligation. a. Proposed mechanism for chemoselective amide bond formation by Histidine Ligation. b. Synthesis of membrane-forming phospholipids by Histidine Ligation.

functionalized lysolipids and fatty acyl thioesters in high efficiency (Scheme 2.2b). The phospholipids can spontaneously self-assemble to generate micron-sized vesicles *in situ* and encapsulate molecules of various sizes.

2.2 Synthesis of histidine-functionalized phospholipids and *in situ* vesicle formation

We initially synthesized two amphiphilic precursors: a histidyl derivative of the lysolipid 1-palmitoyl-*sn*-glycero-3-phosphocholine (**2A**) and a sodium 2-mercaptoethanesulfonate (MESNA) oleoyl thioester[3] (**2C**) (Scheme 2.2b). Precursors **2A** and **2C** are water-soluble amphiphiles, forming micelles of approximately 8.9 nm and 3.8 nm in diameter, respectively as

estimated by dynamic light scattering (DLS) measurements. Phospholipid synthesis by HL took place at mild conditions and low millimolar concentration ranges. At pH 7.5 (HEPES buffer), precursors **2A** and **2C** coupled over 24 h to afford the histidine-functionalized phospholipid **2AC**, that resembles the natural phospholipid 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) (Scheme 2.2b). In a similar manner, histidine-functionalized lysolipid **2B** and oleoyl thioester **2C** allow efficient formation of the phospholipid **2BC**, a mimic of the natural phospholipid 1-oleoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (DOPC) (Scheme 2.2b) and the phospholipid formation was analyzed using HPLC-MS measurements (Figure 2.1a). The phospholipids **2AC** and **2BC** readily formed membrane-bound vesicles *in situ*. Vesicular structures were identified by fluorescence microscopy using the membrane staining dye Texas Red® DHPE (Figure 2.1b). We observed the reaction between **2B** and **2C** using phase contrast microscopy to follow the formation of membranes. Initially, the solution was clear as expected. Within an hour or so, small spherical droplets appeared which grew bigger. With further progress of reaction, vesicles with distinguishable membranes appeared. By 24 h, vesicles of various shapes, sizes and lamellarities could be identified in the reaction mixture (Figure 2.1c).

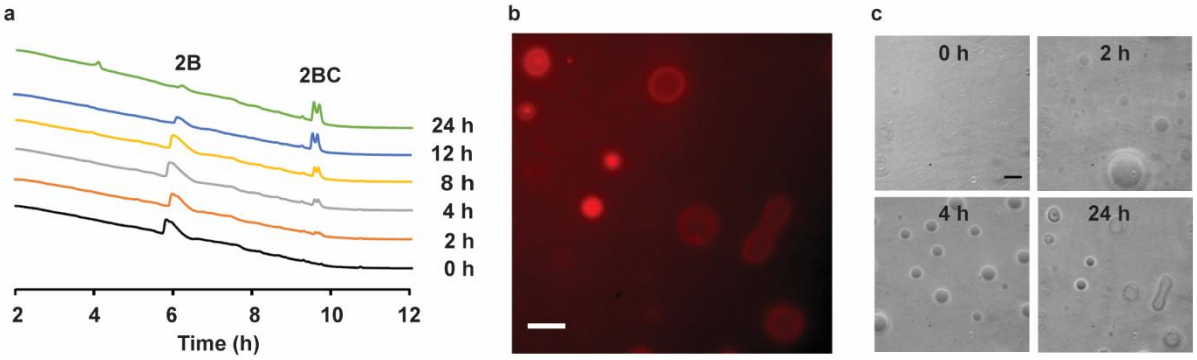


Figure 2.1. Formation of histidine-functionalized phospholipids and self-assembly into vesicles. a. HPLC chromatograms (210 nm) corresponding to the time course of formation of the phospholipid **2BC** from precursors **2B** and **2C**. b. Vesicles formed by hydrating a film of the purified **2AC** and stained with Texas Red-DHPE. Scale bar: 5 µm. c. Time course of *in situ* vesicle formation from **2BC** formed by the reaction between **2B** and **2C**. Scale bar: 10 µm.

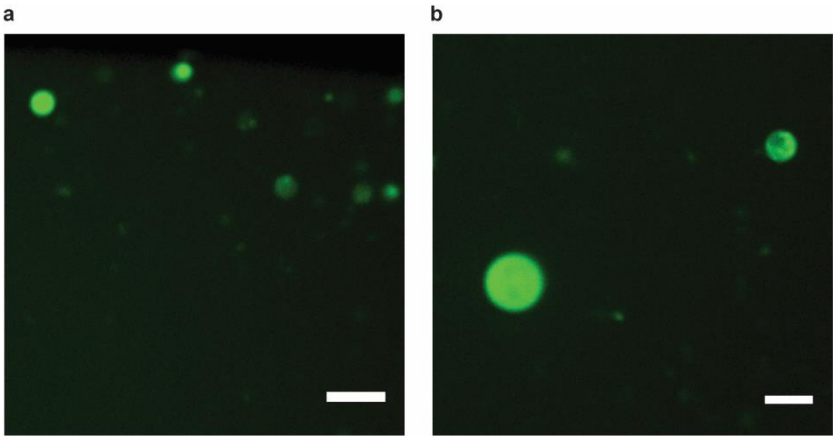


Figure 2.2. Encapsulation of fluorescent molecules in vesicles formed from the histidine-functionalized phospholipid **2AC**. a. HPTS and b. EGFP. Scale bars: 5 μm.

2.3 Encapsulation of molecules

The encapsulation ability of the phospholipid vesicles was determined by hydration of thin lipid films in the presence of a polar fluorescent dye, 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS), followed by removal of non-encapsulated dye by spin-filtration and vesicle characterization using fluorescence microscopy (Figure 2.2a). To check for orthogonality and biocompatibility of the lipid-forming HL reaction, EGFP was spontaneously encapsulated *in situ* by addition of the lipid precursors to a solution containing EGFP. After 24 h of reaction and removal of non-encapsulated EGFP by spin filtration, the lipid-containing solution was examined by fluorescence microscopy, and stable vesicles containing EGFP were observed (Figure 2.2b). The compatibility of lipid formation with biological molecules could lead to applications involving the use of these vesicles as bioreactors.

2.4 Conclusions

In summary, we have developed a highly efficient and chemoselective approach based on HL reaction to readily prepare a new class of amidophospholipids, which self-assemble *in situ* to form synthetic membranes. The selectivity, high reactivity and the biocompatibility of this

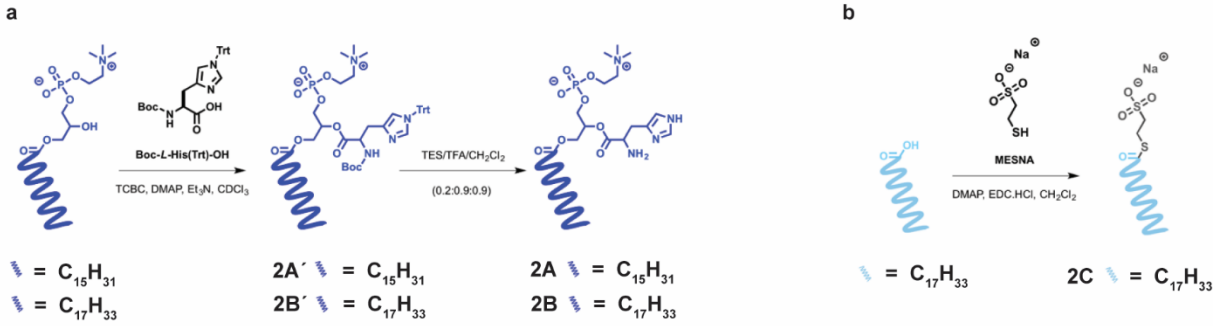
methodology are key features that make it a novel tool for the efficient encapsulation of biomolecules for bottom-up synthetic biology applications.

2.5 Experimental methods

2.5.1 General considerations

Commercially available 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (Lyso C_{16} PC-OH) and 1-oleoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (Lyso C_{18} PC-OH) were used as obtained from Avanti® Polar Lipids. N-Boc-*L*-His(Trt)-OH, 2,4,6-trichlorobenzoyl chloride (TCBC), triethylamine (Et_3N), 4-dimethylaminopyridine (DMAP), trifluoroacetic acid (TFA), triethylsilane (TES), oleic acid, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC.HCl), sodium 2-mercaptoethanesulfonate (MESNA), 4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid (HEPES) sodium salt and 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS) were obtained from Sigma-Aldrich. Texas Red 1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine, triethylammonium salt (Texas Red DHPE) was obtained from Life Technologies. Deuterated chloroform ($CDCl_3$), methanol (CD_3OD) and dimethyl sulfoxide (d_6-DMSO) were obtained from Cambridge Isotope Laboratories. All reagents obtained from commercial suppliers were used without further purification unless otherwise noted. Analytical thin-layer chromatography was performed on E. Merck silica gel 60 F_{254} plates. Compounds, which were not UV active, were visualized by dipping the plates in a ninhydrin or potassium permanganate solution and heating. Silica gel flash chromatography was performed using E. Merck silica gel (type 60SDS, 230-400 mesh). Solvent mixtures for chromatography are reported as v/v ratios. HPLC analysis was carried out on an Eclipse Plus C8 analytical column with *Phase A*/*Phase B* gradients [*Phase A*: H_2O with 0.1% formic acid; *Phase B*: MeOH with 0.1% formic acid]. HPLC purification was carried out on Zorbax SB-C18 semipreparative column with *Phase*

A/*Phase B* gradients [*Phase A*: H_2O with 0.1% formic acid; *Phase B*: MeOH with 0.1% formic acid]. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on a Varian VX-500 MHz or Jeol Delta ECA-500 MHz spectrometers, and were referenced relative to residual proton resonances in $CDCl_3$ (at δ7.24 ppm), CD_3OD (at δ4.87 or 3.31 ppm) or d_6-DMSO (at δ2.50 ppm). Chemical shifts were reported in parts per million (ppm, δ) relative to tetramethylsilane (δ 0.00). ^{1}H NMR splitting patterns are assigned as singlet (s), doublet (d), triplet (t), quartet (q) or pentuplet (p). All first-order splitting patterns were designated on the basis of the appearance of the multiplet. Splitting patterns that could not be readily interpreted are designated as multiplet (m) or broad (br). Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Varian VX-500 MHz or Jeol Delta ECA-500 MHz spectrometers, and were referenced relative to residual proton resonances in $CDCl_3$ (at δ77.23 ppm), CD_3OD (at δ49.15 ppm) or d_6-DMSO (at δ39.51 ppm). Electrospray Ionization-Time of Flight (ESI-TOF) spectra were obtained on an Agilent 6230 Accurate-Mass TOF-MS mass spectrometer.



Scheme 2.3. Synthetic schemes for His-functionalized lysolipids and oleoyl MESNA thioester.

2.5.2 General procedure for synthesis of histidine-functionalized lysolipids

A solution of lysolipid (Lyso C_{16} PC-OH or Lyso $C_{18:1}$ PC-OH, 1 eqvt.), *N*-Boc-L-His(Trt)-OH (2.5 eqvt.), DMAP (5 eqvt.) and Et_3N (3.5 eqvt.) in dry chloroform was stirred at rt for 10 min. Then, TCBC (6.5 eqvt.) was added. After 12 h stirring at rt, H_2O was added to quench the acyl chloride, and the solvent was removed under reduced pressure to give a pale yellow solid. The

corresponding residue was dissolved in MeOH, filtered using a 0.2 µm syringe-driven filter, and the crude solution was purified by HPLC to obtain **2A´** or **2B´** (Scheme 2.3a). For deprotection of the Boc-group, a solution of **2A´** or **2B´** in 2 mL of TFA/CH_2Cl_2/TES (0.9:0.9:0.2) was stirred at rt for 45 min. The residue obtained after solvent removal was purified by HPLC (Scheme 2.3a).

2.5.3. Synthesis of MESNA oleoyl thioester (2C)

A solution of oleic acid (189.2 mg, 670.0 µmol) in CH_2Cl_2 (5 mL) was stirred at 0 ºC for 10 min, and then DMAP (7.4 mg, 60.9 µmol) and EDC.HCl (128.4 mg, 670.0 µmol) were successively added. After 10 min stirring at 0 ºC, MESNA (100.0 mg, 609.1 µmol) was added. After 5 h stirring at rt, the mixture was extracted with H_2O (2×3 mL) and the combined aqueous phases were washed with EtOAc (3 mL). After evaporation of H_2O under reduced pressure, the residue was washed with CH_3CN (5 mL), and then filtered to obtain **2C** (Scheme 2.3b).

2.5.4 General procedure for synthesis of His-functionalized phospholipids

2A or **2B** (1 eqvt.) and **2C** (1 eqvt.) were dissolved in 100 mM HEPES buffer pH 7.5 and stirred under N_2 at rt. After 24 h, the crude mixture was purified by HPLC to obtain the phospholipids **2AC** or **2BC** (Scheme 2.2b).

2.5.5 Compound characterization data

2A´: white foam (68.3 mg, 73%). ^{1}H NMR ($CDCl_3$): 7.45-7.11 (m, 12H, 12×CH_{Ar}), 7.11-6.89 (m, 4H, 4×CH_{Ar}), 6.55 (s, 1H, 1×CH_{Ar}), 6.22 (m, 1H, 1×NH), 5.04-4.83 (m, 1H, 1×CH), 4.49-4.13 (m, 3H, 1×CH+1×CH_2), 4.11-3.82 (m, 3H, 1.5×CH_2), 3.80-3.62 (m, 1H, 0.5×CH_2), 3.59-3.33 (m, 4H, 2×CH_2), 3.25 (s, 9H, 3×CH_3), 3.03-2.79 (m, 2H, 1×CH_2), 2.23-1.78 (m, 2H, 1×CH_2), 1.32 (s, 9H, 3×CH_3), 1.25-0.94 (m, 24H, 12×CH_2), 0.90-0.65 (m, 3H, 1×CH_3). ^{13}C NMR ($CDCL_3$): 173.2, 171.4, 155.9, 144.3, 138.9, 136.3, 128.3, 128.3, 128.2, 120.0, 79.5, 75.5, 72.3, 66.6, 64.0,

62.5, 59.5, 54.7, 53.9, 34.0, 32.1, 30.0, 29.9, 29.9, 29.8, 29.8, 29.7, 29.5, 29.4, 29.2, 28.6, 24.9, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{54}H_{80}N_4O_{10}P$ ($[MH]^+$) 975.5607, found 975.5603.

2B´: white foam (72.3 mg, 75%). ^{1}H NMR ($CDCl_3$): 7.52-7.12 (m, 12H, 12×CH_{Ar}), 7.11-6.84 (m, 4H, 4×CH_{Ar}), 6.55 (s, 1H, 1×CH_{Ar}), 6.22 (d, *J* = 7.4 Hz, 1H, 1×NH), 5.47-5.13 (m, 2H, 2×CH), 5.07-4.82 (m, 1H, 1×CH), 4.61-4.13 (m, 3H, 1×CH+1×CH_2), 4.10-3.81 (m, 3H, 1.5×CH_2), 3.80-3.53 (m, 1H, 0.5×CH_2), 3.46-3.08 (m, 11H, 1×CH_2+3×CH_3), 3.06-2.78 (m, 2H, 1×CH_2), 2.26-1.74 (m, 8H, 4×CH+2×CH_2), 1.59-0.96 (29H, 3×CH_3+10×CH_2), 0.80 (t, *J* = 5.8 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$): 173.1, 171.4, 155.8, 144.3, 138.9, 136.3, 130.1, 130.1, 129.9, 128.3, 128.2, 120.0, 79.5, 75.5, 72.3, 66.6, 64.0, 62.5, 59.5, 54.7, 53.9, 33.9, 32.0, 29.9, 29.9, 29.7, 29.5, 29.4, 29.4, 29.3, 29.2, 28.6, 28.5, 27.4, 27.3, 24.8, 22.9, 14.3. HRMS (ESI-TOF) calculated for $C_{56}H_{82}N_4O_{10}P$ ($[MH]^+$) 1001.5763, found 1001.5756.

2A: colorless film (7.0 mg, 62%). ^{1}H NMR (CD_3OD): 7.89-7.61 (m, 1H, 1×CH_{Ar}), 7.20-6.90 (m, 1H, 1×CH_{Ar}), 5.42-5.15 (m, 1H, 1×CH), 4.54-3.85 (m, 7H, 3×CH_2+1×CH), 3.71-3.51 (m, 2H, 1×CH_2), 3.38-3.34 (m, 2H, 1×CH_2), 3.23 (s, 9H, 3×CH_3), 2.47-2.20 (m, 2H, 1×CH_2), 1.71-1.48 (m, 2H, 1×CH_2), 1.44-1.16 (m, 24H, 12×CH_2), 0.90 (t, *J* = 6.6 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 174.9, 168.1, 136.9, 134.0, 117.2, 74.2, 67.8, 64.9, 63.1, 60.5, 54.7, 54.5, 34.9, 34.7, 33.1, 30.8, 30.8, 30.7, 30.7, 30.7, 30.6, 30.6, 30.5, 30.4, 30.2, 29.7, 29.6, 25.9, 23.7, 14.4. MS. HRMS (ESI-TOF) calculated for $C_{30}H_{58}N_4O_8P$ ($[MH]^+$) 633.3987, found 633.3986.

2B: colorless film (8.3 mg, 73%). ^{1}H NMR (CD_3OD): 7.85-7.62 (m, 1H, 1×CH_{Ar}), 7.17-6.89 (m, 1H, 1×CH_{Ar}), 5.45-5.15 (m, 3H, 3×CH), 4.47-3.94 (m, 7H, 3×CH_2+1×CH), 3.70-3.56 (m, 2H, 1×CH_2), 3.38-3.32 (m, 2H, 1×CH_2), 3.22 (s, 9H, 3×CH_3), 2.46-2.21 (m, 2H, 1×CH_2), 2.16-1.90 (m, 4H, 2×CH_2), 1.69-1.49 (m, 2H, 1×CH_2), 1.49-1.15 (m, 20H, 10×CH_2), 0.89 (t, *J* = 6.7 Hz, 3H, 1× CH_3). ^{13}C NMR (CD_3OD): 174.8, 167.8, 137.0, 134.0, 130.9, 130.7, 117.6, 74.2, 67.4,

66.2, 64.9, 63.1, 60.5, 54.7, 34.9, 34.7, 33.1, 30.8, 30.8, 30.6, 30.5, 30.4, 30.2, 30.2, 29.6, 29.5, 28.1, 25.9, 23.8, 14.5. HRMS (ESI-TOF) calculated for $C_{32}H_{60}N_4O_8P$ ($[MH]^+$) 659.4143, found 659.4139.

2C: white solid (194.7 mg, 75%). ^{1}H NMR (d_6-DMSO): 5.36-5.27 (m, 2H, 2×CH), 3.05-2.99 (m, 2H, 1×CH_2), 2.60-2.51 (m, 4H, 2×CH_2), 2.02-1.92 (m, 4H, 2×CH_2), 1.58-1.49 (m, 2H, 1×CH_2), 1.34-1.18 (m, 20H, 10×CH_2), 0.85 (t, J = 6.9 Hz, 3H, 1×CH_3). ^{13}C NMR (d_6-DMSO): 198.7, 129.8, 129.7, 51.0, 43.4, 31.4, 29.2, 29.1, 28.9, 28.8, 28.7, 28.6, 28.5, 28.3, 26.7, 26.6, 25.1, 24.4, 22.2, 14.1. HRMS (ESI-TOF) calculated for $C_{20}H_{38}NaO_4S_2$ ($[MH]^+$) 429.2104, found 429.2105.

2AC: colorless oil (1.21 mg, 43%). HRMS (ESI-TOF) calculated for $C_{48}H_{90}N_4O_9P$ ($[MH]^+$) 897.6440, found 897.6437.

2BC: colorless oil (1.43 mg, 51%). HRMS (ESI-TOF) calculated for $C_{50}H_{92}N_4O_9P$ ($[MH]^+$) 923.6596, found 923.6594.

2.5.6 *De novo* phospholipid synthesis and vesicle formation

12.5 µL (10 mM stock solution) of lysolipid (**2A** or **2B**) and 15 µL (10 mM stock solution) of thioester **2C** were added to 22.5 µL of HEPES buffer (100 mM, pH 7.5) at rt, and the mixture was briefly agitated. The resulting solution was added to a glass microscope slide and covered with a glass coverslip supported with vacuum grease. The sample was then monitored by phase contrast microscopy in order to analyze the *in situ* phospholipid (**2AC** or **2BC**) vesicle formation.

2.5.7 Encapsulation Experiments

*Encapsulation of HPTS***.** 10 µL of a 10 mM solution of phospholipid (**2AC** or **2BC**) in MeOH:$CHCl_3$ (1:1) were added to a glass vial, placed under a steady flow of N_2 and dried for 10 min to prepare a lipid film. Then, 100 µL of 0.1 mM HPTS aqueous solution were added to the

lipid film and briefly vortexed. The solution was tumbled at rt for 30 min. Afterward, the resulting cloudy solution was diluted with an additional 200 µL of H_2O and transferred to a 100 kDa molecular weight cut-off (MWCO) centrifugal membrane filter and centrifuged for (3 min × 5) at 10,000 rcf (Eppendorf 5415C).

In situ encapsulation of EGFP. 25.0 µL (10 mM stock solution) of lysolipid (**2A** or **2B**) and 30 µL (10 mM stock solution) of thioester **2C** were added to 45.0 µL of HEPES buffer (100 mM, pH 7.5). Then, 2.0 µL of 220 µM solution of EGFP were added and the solution was kept at rt for 24 h. The resulting solution was diluted with an additional 200 µL of 45 mM HEPES buffer pH 7.5, transferred to a 100 kDa MWCO centrifugal membrane filter and centrifuged for (3 min × 5) at 10,000 rcf (Eppendorf 5415C). The vesicles were imaged on a spinning disc confocal microscope (488 nm laser) to observe encapsulation of EGFP in the *in situ* formed vesicles.

Notes on chapter

Chapter 2 is adapted almost entirely from the materials published in Brea, R. J., Bhattacharya, A. & Devaraj, N. K. Spontaneous Phospholipid Membrane Formation by Histidine Ligation. *Synlett*, 2017, 28, 108-112. The dissertation author was a co-first author on this manuscript.

2.6 References

1. Wang, B. & Tontonoz, P. Phospholipid Remodeling in Physiology and Disease. *Annu. Rev. Physiol.* **81,** 165–188 (2019).
2. Budin, I. & Devaraj, N. K. Membrane assembly driven by a biomimetic coupling reaction. *J. Am. Chem. Soc.* **134,** 751–753 (2012).
3. Brea, R. J., Cole, C. M. & Devaraj, N. K. In Situ Vesicle Formation by Native Chemical Ligation. *Angew. Chem. Int. Ed.* **53**, 14102–14105 (2014).
4. Zhang, L. & Tam, J. P. Orthogonal coupling of unprotected peptide segments through histidyl amino terminus. *Tetrahedron Lett.* **38,** 3–6 (1997).
5. Tam, J. P., Xu, J. & Eom, K. D. Methods and strategies of peptide ligation. *Biopolym. Pept. Sci. Sect.* **60,** 194–205 (2001).

Chapter 3 | Minimal Biochemical Routes Towards Synthesis of Membrane-forming Phospholipids

3.1 Introduction

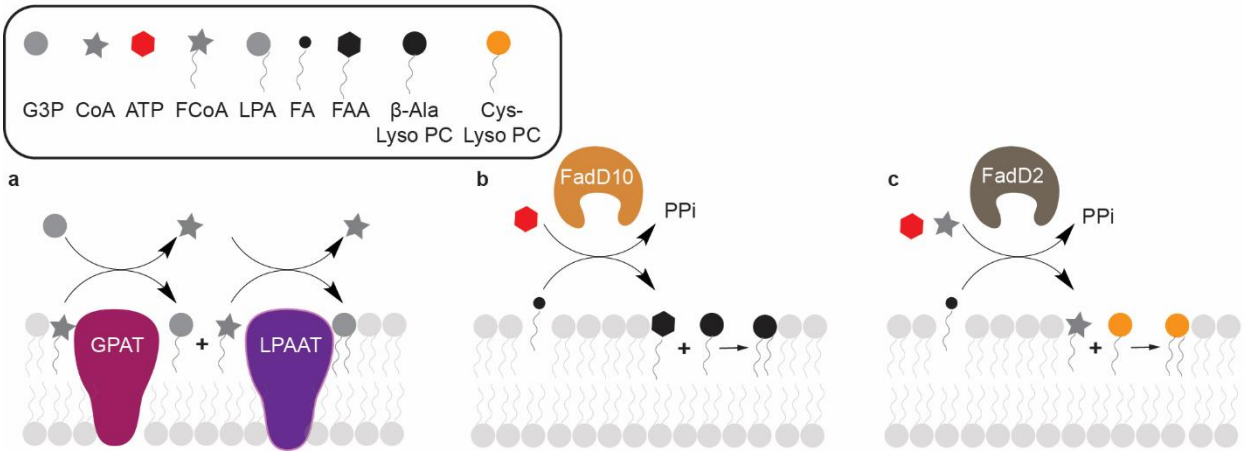
Phospholipids are the primary constituents of cell membranes. In living organisms, phospholipids are generated enzymatically by the reaction of a polar head group with long-chain acyl derivatives (Scheme 3.1a). These key steps rely on integral membrane proteins, such as acyltransferases, which require pre-existing membranes for proper folding and function.[1] This mechanism implies that all biological membranes must arise from pre-existing membranes.[2] However, the principle of biological membrane continuity presents a challenge for explaining how phospholipid membranes were generated *de novo* before the current membrane-dependent enzymes and mechanisms for phospholipid synthesis developed. We reasoned that designing a minimal route for enzymatic *de novo* phospholipid synthesis could help us understand how early cellular membrane synthesizing machinery evolved.[3–5] Since present-day integral membrane proteins cannot carry out true *de novo* phospholipid formation, a method by which soluble enzymes could facilitate the synthesis of membrane-forming phospholipids is required. It can also provide simplified strategies for generating membrane compartments in synthetic cells,[6,7] enable the development of tools for reconstituting membrane proteins,[8,9] and facilitate strategies for synthesizing structured lipids.[10]

As there are no analogous reactions in biology, we sought to design unique lipid synthesizing strategies by repurposing soluble fatty acid activating enzymes. In these methods, we enzymatically activate fatty acids to a high-energy intermediate, which can thereafter undergo non-enzymatic coupling with a suitably functionalized lysolipid to form a phospholipid. We used the following two mycobacterial ligases for phospholipid formation:

(i) FadD10 – a fatty acyl adenylate ligase. FadD10 catalyzes the generation of fatty acyl adenylates (FAAs) from fatty acid, Mg^{2+} and ATP precursors. FAAs react spontaneously and chemoselectively with amine-functionalized lysolipids to form amidophospholipids (Scheme 3.1b, Figure 3.2a).

(ii) FadD2 – a fatty acyl CoA ligase. FadD2 activates fatty acids to the corresponding acyl CoA (FCoA) using ATP, CoA, and Mg^{2+}. The FCoAs can undergo chemoselective coupling with cysteine-functionalized lysolipids at the thioester linkage to form amidophospholipids (Scheme 3.1c).

Our results demonstrate that pathways radically different from those taking place in living cells can be developed for synthesizing membrane-forming lipids.



Scheme 3.1. Natural and artificial pathways of phospholipid synthesis. a. Natural (Kennedy or de novo) pathway b. fatty acyl adenylate ligase (FadD10)-mediated pathway c. fatty acyl CoA ligase (FadD2)-mediated pathway. Abbreviations: G3P – glycerol-3-phosphate, FCoA – fatty acyl coenzyme A, LPA – lysophosphatidic acid, FA – fatty acid, FAA – fatty acyl adenylate, Ppi – inorganic pyrophosphate, GPAT – glycerol-3-phosphate acyltransferase, LPAAT – lysophosphatidic acid acyltransferase.

3.2 Reactivity of fatty acyl adenylates (FAA)

Acyl adenylates (AAs) are metabolic intermediates that are subset of a broader class of molecules called acyl phosphates. Acyl phosphates are found in diverse biochemical pathways ranging from fatty acid metabolism, bioluminescence, tRNA aminoacylation, protein acetylation,

and nitrogen fixation (Figure 3.1). AAs typically undergo enzymatic coupling with thiol nucleophiles, such as coenzyme A (CoA).[11] Acyl phosphates can also react non-enzymatically with primary amines in aqueous media.[12,13]

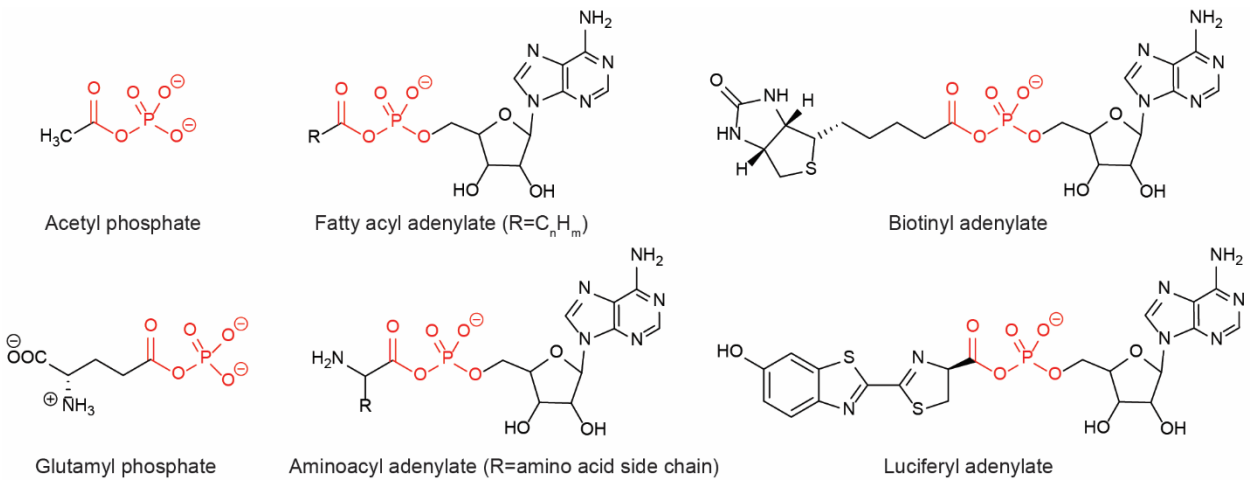


Figure 3.1. Chemical structures of common acyl phosphates found in biochemistry.

In order to evaluate the scope and applicability of the proposed lipid synthesis pathway, we first explored the reactivity patterns of FAAs. We synthesized dodecanoyl-AMP **3A1**[14] as a model FAA, and found that it was fairly stable to hydrolysis at 37 ^{o}C in HEPES buffer, pH 7.5 in the absence or presence (10 mM) of Mg^{2+}, and over time scales relevant to our subsequent experiments (Figure 3.2b). We observed that **3A1** showed negligible reactivity towards the hydroxy groups of the naturally occurring lysolipids 1-oleoyl-2-hydroxy-*sn*-glycero-3-phospho-choline (Lyso $C_{18:1}$ PC) and 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (Lyso $C_{16:0}$ PC-OH). Remarkably, when we mixed **3A1** with the corresponding amine-functionalized lysolipids **3B1** or **3B2** (Figure 3.1a) in HEPES buffer, pH 7.5 at 37 ^{o}C, the solution became turbid after 20 minutes, and we observed the formation of a large population of vesicles (Figure 3.1c) which were further imaged by TEM (Figure 3.2d). The progress of the reaction between **3A1** and **3B1** was analyzed over time using liquid chromatography-mass spectrometry (LC-MS) measurements (Figure 3.2e). The second order rate constant for the reaction was determined to be 87.0 ± 9.1 M^{-}

$^{1}\cdot s^{-1}$ (standard deviation from n = 3 replicates). We obtained the nuclear magnetic resonance (NMR) spectra from purified phospholipids **3A1B1** and **3A1B2.** These data verified that acylation occurred at the primary amine functionality of **3B1** and **3B2**, respectively. We carried out differential scanning calorimetry (DSC) on an aqueous dispersion of **3A1B1** and did not observe any peaks above 1 °C, suggesting that the lipid is in fluid phase at rt or above.

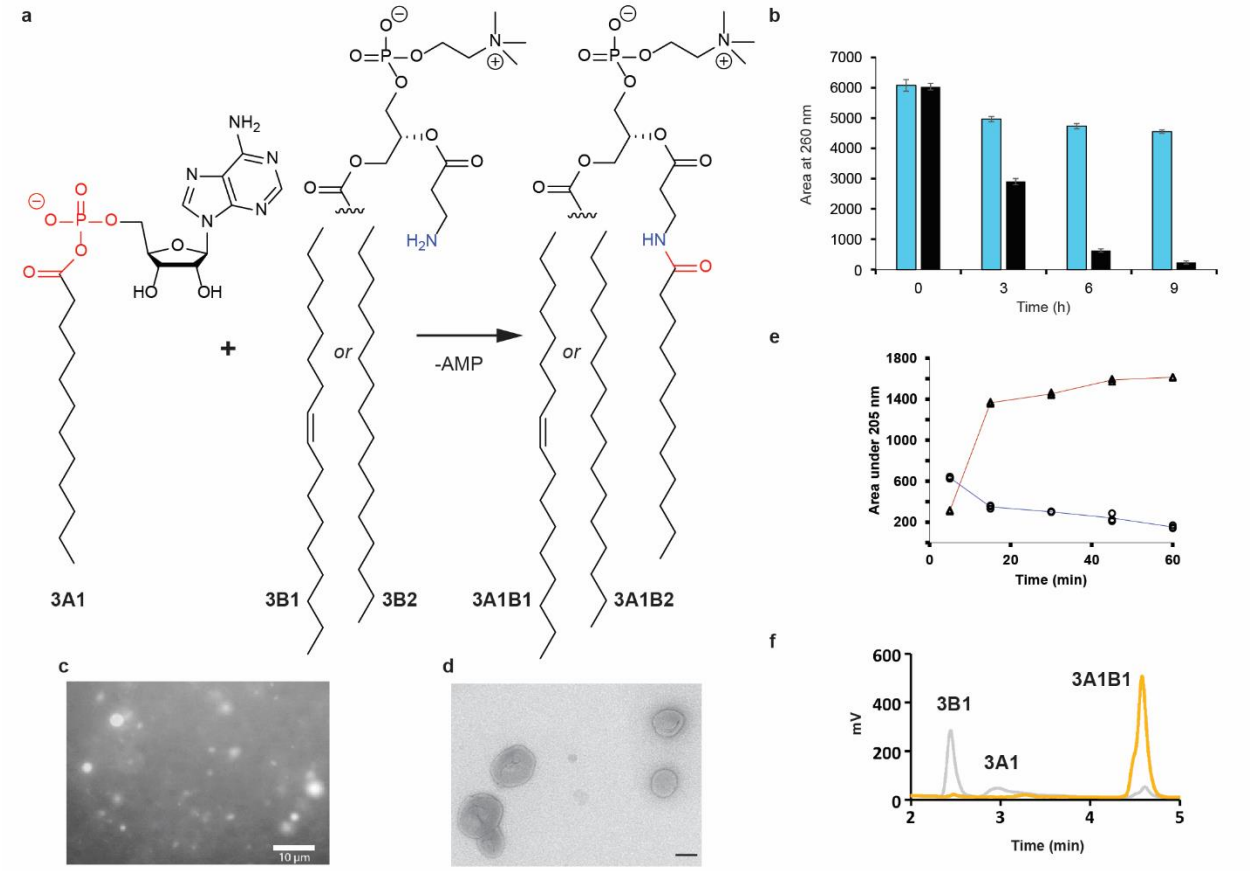


Figure 3.2. Formation of phospholipids based on acyl phosphate chemistry. a. Reaction between FAA and amine-functionalized lysolipids leading to the formation of amidophospholipids. b. Stability of **3A1** over time in the absence or presence (10 mM) of Mg^{2+}. The cyan bars represent integrated HPLC peak areas (260 nm) for the 1 in the absence of Mg^{2+}, whereas the black bars represent integrated HPLC peak areas (260 nm) for **3A1** in the presence of 10 mM Mg^{2+}. The error bars represent standard deviation (n = 3). c-d. *In situ* formation of phospholipid **3A1B1** vesicles by reaction of **3A1** and lysolipid **3B1** visualized by Texas Red-DHPE staining (c) and negative-staining TEM (d) (Scale bars: c – 10 µm, d – 100 nm). e. Kinetics of phospholipid **3A1B1** formation by the reaction of **3A1** with lysolipid **3B1**. Integrated HPLC peak areas (205 nm) were used to monitor the progress of the reaction. The error bars represent standard deviation (n = 3). f. HPLC/ELSD traces monitoring the selective formation of phospholipid **3A1B1** by reaction of **3A1** and lysolipid **3B1** in the presence of 50 mM lysine.

Because amines are ubiquitous in biology, we asked if FAAs have any selectivity for amine-functionalized lysolipids over simple water-soluble amines. When the reaction between **3A1** and **3B1** was conducted in the presence of 50 mM lysine in HEPES buffer, we found that coupling occurred with excellent selectivity between the FAA and the lysolipid. We did not observe significant side products that corresponded to a reaction with the primary amine groups in lysine (Figure 3.2f). We found similar selectivity when the coupling reaction between **3A1** and **3B1** was carried out in 100 mM Tris buffer, pH 8.0. We followed the reaction kinetics between **3A1** and a non-amphiphilic primary amine, Fmoc-*L*-Lys-OH, by HPLC-MS. The second order rate constant for the reaction was obtained to be 0.0033 ± 0.0004 $M^{-1} \cdot s^{-1}$ (error value represents standard deviation from $n = 3$ replicates), almost 30,000-fold slower than the reaction with the lysolipid **3B1**. We attribute this high selectivity of phospholipid synthesis to hydrophobic interactions between the alkyl chains of **3A1** and **3B1**, which are brought into close proximity, likely within mixed micelles or formed lipid membranes. These reactivity patterns indicate that FAAs could serve as an activated acyl precursor in a more complex biochemically relevant media.

3.3. Synthesis of sphingolipids using acyl phosphate chemistry

We also considered natural membrane forming precursors containing $–NH_2$ functionality such as lyso-Sphingomyelin d18:1 (lyso SM) as a model lysolipid and carried out the reaction with oleoyl-AMP (**3A2**) under similar conditions to form the naturally occurring 18:1 Sphingomyelin (**3A2S1**) (Figure 3.3a). Full conversion (as determined by HPLC-ELSD-MS) was observed in this case as well (Figure 3.3b) along with spontaneous vesicle formation. However, it was found that, the reaction was slower than the earlier example – typically required 5-6 hours to go to completion. This was not quite unexpected since the $–NH_2$ group is situated on a secondary carbon and hence sterically more hindered. However, a significant improvement in the reaction rate was observed

when Mg^{2+} or Ca^{2+} was added. The reaction was complete within 30 min without any noticeable hydrolysis.

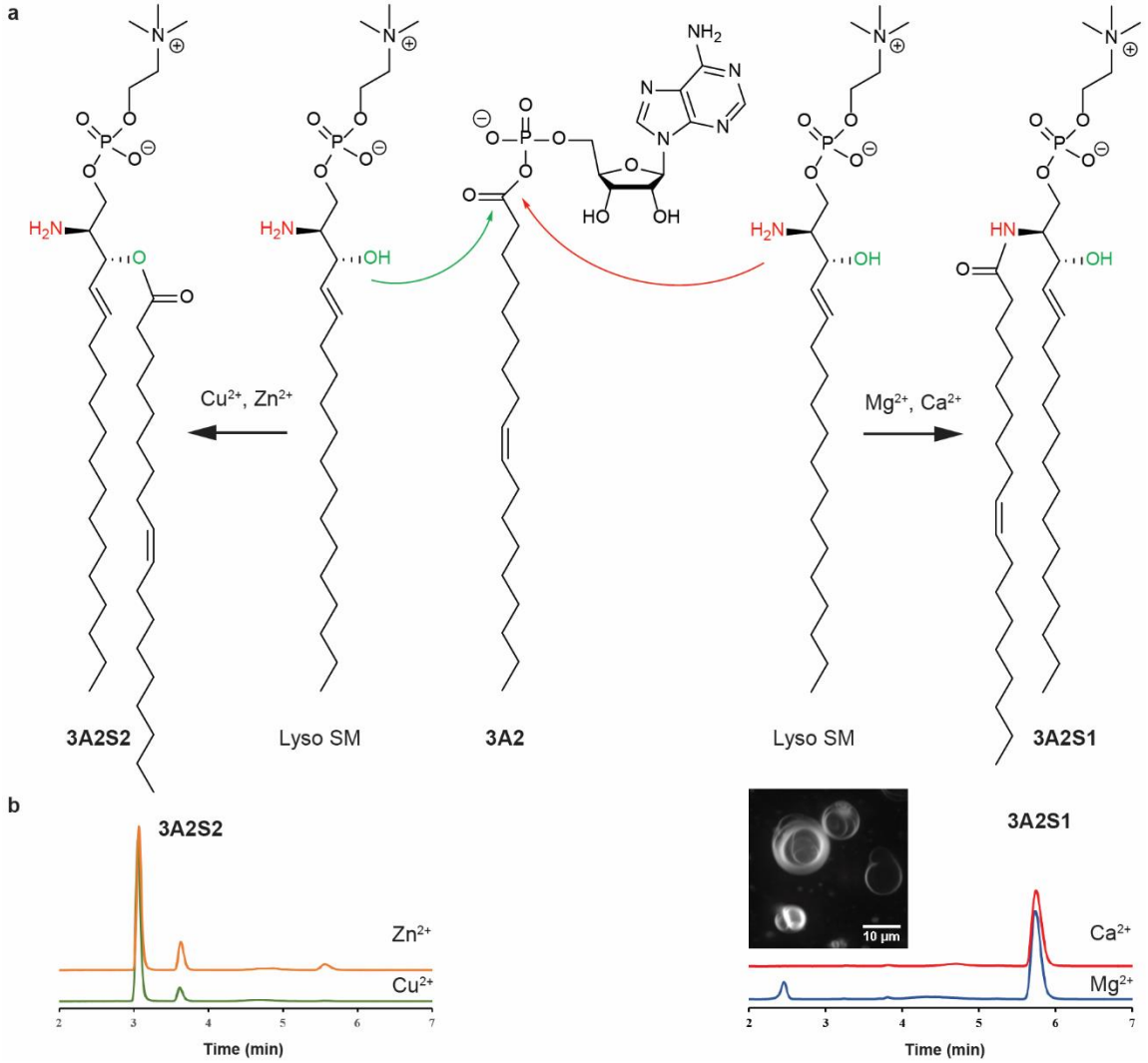


Figure 3.3. Formation of sphingomyelins using acyl phosphate chemistry. a. Reaction schemes showing formation of *N*-acylated and *O*-acylated products in presence of hard and soft metal ions respectively. b. Corresponding HPLC-ELSD chromatograms showing lipid formation. The inset image shows vesicles of **3A2S1**. Scale bar: 10 µm.

. Interestingly, Zn^{2+} and Cu^{2+} promoted mostly *O*-acylation under the same conditions (**3A2S2**) (Figure 3.3a-b). The evidence of *O*-acylation can be found from the (i) position of the product peak – **3A2S2** elutes significantly earlier compared to **3A2S1** as expected from the charged amine in the former (ii) mass spectrometric peaks for **3A2S2** gives as strong M/2 peak also expected from an amine-containing species. It is remarkable that when Cu^{2+} (as $CuSO_4$) was added

to the reaction, the color change from pale blue to intense blue, suggesting coordination between -NH_2 group of lyso-SM and Cu(II). It is also noteworthy that, the *O*-acylated product produced insoluble aggregates in the reaction mixture and vesicles were not observed to form. This is likely because of strong coordination with Zn^{2+} and Cu^{2+} that causes aggregation.

3.4 *De novo* phospholipid formation mediated by FadD10

We next aimed to identify an appropriate enzyme capable of generating FAAs such as **3A1** from fatty acid and ATP precursors. In biological systems, fatty acids are activated to FAAs by various fatty acyl CoA ligases (FACLs) and fatty acyl adenylate ligases (FAALs).[11] While FACLs are common, the FAA intermediate remains tightly bound to the active site to facilitate its subsequent reaction with CoA, which limited their attractiveness as candidate enzymes. We sought an enzyme that could form FAAs that could subsequently react with amine-functionalized lysolipids. We selected FadD10,[15] a recently characterized FAAL involved in the biosynthesis of a putative lipopeptide virulence factor in *Mycobacterium tuberculosis.*[16] FadD10 is a soluble enzyme that converts long-chain saturated fatty acids (typically C_{12}-C_{16}) into corresponding FAAs in presence of ATP and Mg^{2+}. Since FadD10 displays an 'open' conformation of its active site, it does not have a high binding affinity to the FAA product.[15] As such, we hypothesized the newly synthesized FAAs will be free to diffuse away and react with amine-functionalized lysolipids to form the corresponding amidophospholipids.

We expressed N-terminal His_6-tagged FadD10 in *E. coli* and purified it according to a published procedure.[15] In a typical phospholipid synthesis reaction, FadD10 was incubated with ATP, $MgCl_2$, lysolipid **3B1** and sodium dodecanoate in HEPES buffer at 37 ^oC. The enzymatically formed FAA intermediate **3A1** was detected by HPLC-MS, even after FadD10 was separated from the reaction mixture by spin filtration. These results indicate that **3A1** was released from the

enzyme's active site. Approximately 30 min after mixing FadD10 and substrates, small vesicles could be detected by optical microscopy. These vesicles gradually transformed into larger vesicles of various sizes and lamellarity (Figure 3.4a).

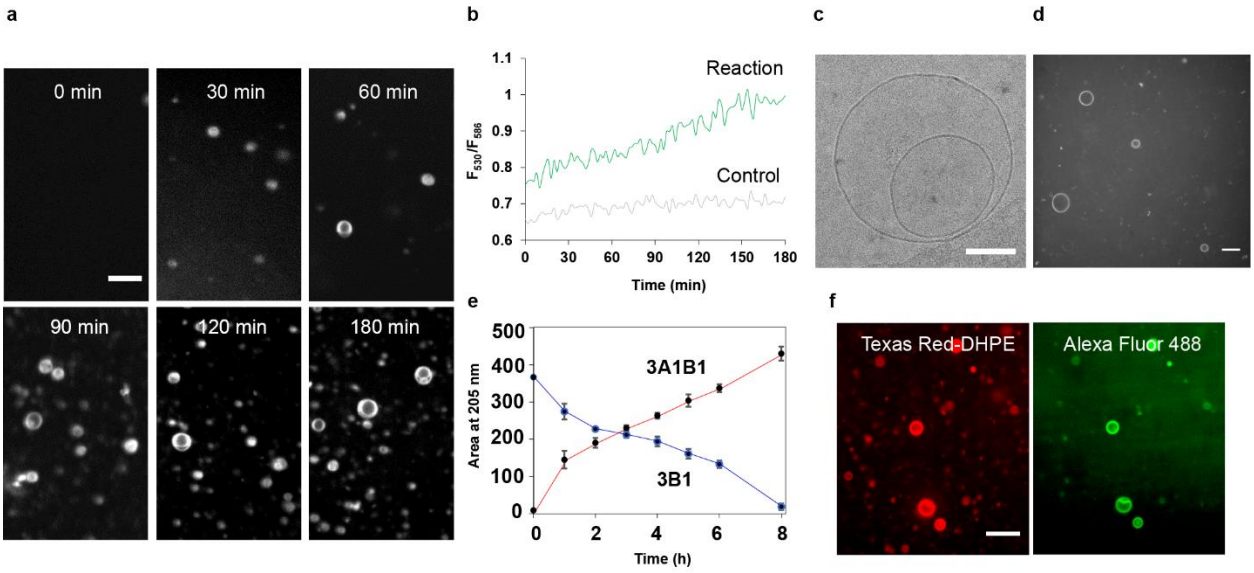


Figure 3.4. FadD10-mediated phospholipid membrane formation. a. Time series of spinning disk confocal microscopy images depicting *de novo* phospholipid **3A1B1** vesicle formation resulting from the incubation of an aqueous solution of dodecanoic acid, lysolipid **3B1**, ATP, $MgCl_2$, FadD10 and 0.1 mol% Texas Red® DHPE at 37 ºC. Scale bar: 10 µm. b. Assaying *de novo* membrane formation using FRET between 0.2 mol% of each of NBD-DHPE (*donor*; λ_{ex}: 430 nm, λ_{em}: 530 nm) and Rhodamine-DHPE (*acceptor*; λ_{ex}: 560 nm, λ_{em}: 586 nm). The *green* trace shows a gradual increase in the F_{530}/F_{586} corresponding to *de novo* membrane growth arising from phospholipid synthesis (***Reaction***). The *grey* trace corresponds to a control condition where ATP is substituted with GTP and no phospholipid synthesis took place (***Control***). c. Cryogenic-transmission electron microscopy (cryo-TEM) image of a *de novo* formed phospholipid vesicle showing the presence of membranes. Scale bar: 100 nm. d. FadD10 assisted formation of the anionic phosphoglycerol head group bearing phospholipid **3A1G1**. Scale bar: 10 µm. e. Kinetics of the consumption of lysolipid **3B1** and formation of phospholipid **3A1B1** at 37 ºC. Integrated HPLC peak areas (205 nm) were used to monitor the progress of the reaction. The error bars represent standard deviation (n = 3). f. Spinning disk confocal microscopy images of phospholipid **3A1B1** vesicles formed with Alexa Fluor 488 labeled FadD10 and 0.1 mol% Texas Red DHPE, showing association of the enzyme with the membrane. Scale bar: 10 µm.

We followed the progress of *de novo* formation of phospholipid membranes using a well-established Fluorescence Resonance Energy Transfer (FRET) assay involving fluorescently labeled phospholipid probes.[17] We observed a linear increase in the ratio of the fluorescence intensities of the donor dye NBD-DHPE (λ_{ex}: 430 nm, λ_{em}: 530 nm) and the acceptor dye Rhodamine-DHPE (λ_{ex}: 430 nm, λ_{em}: 586 nm), suggesting that membrane growth is taking place during phospholipid synthesis (Figure 3.4b). In a control experiment, where GTP was substituted

for ATP, a nearly constant ratio of the donor and acceptor fluorescence intensities was observed, consistent with the lack of phospholipid formation.

In addition, we performed several control experiments where one of the key reaction components was omitted or replaced by an unreactive substitute to confirm that vesicle formation occurred due to phospholipid generation. Cryo-electron microscopy further verified the presence of membranes (Figure 3.4c). In order to test the broader applicability of our method, we carried out the reaction in presence of the amine-functionalized lysophosphatidylglycerol **3G**. We obtained the bilayer forming phospholipid **3A1G**, which self-assembled to form vesicles as well (Figure 3.4d). We analyzed the formation of phospholipid **3A1B1** using HPLC-MS and found that the precursors were almost completely consumed in about 8 hours (Figure 3.4e).

Interestingly, when we used Alexa Fluor 488 labeled FadD10, we observed that the enzyme was significantly associated with the *de novo* formed phospholipid **3A1B1** membranes (Figure 3.4f). We isolated the vesicles from the reaction mixture by spin filtration and subsequent SDS-PAGE analysis of the vesicle fraction showed the presence of FadD10. During vesicle formation, FadD10 associates with the membranes possibly through electrostatic interaction with the charged amphiphilic precursors. Given that the theoretical isoelectric point (pI) of His_6-FadD10 is 5.4 (ExPASy), it is likely to bear a negative charge in the pH range of our experiments and undergo electrostatic interactions with the positively charged amine-functionalized lysolipid to become encapsulated during *de novo* formation of phospholipid vesicles.

3.5 *De novo* phospholipid formation mediated by FadD2

In a previous work by Brea *et al.*, it was shown that cysteine-functionalized lysolipids can be coupled to water-soluble fatty acyl thioesters by Native Chemical Ligation (NCL)[18] under mild aqueous conditions to generate phospholipids which self-assembled to form membranes. In

subsequent works, this strategy was applied to remodel artificial membranes and reconstitute functional transmembrane proteins. In biochemistry, fatty acids are commonly activated as thioesters of coenzyme A (CoA) by the action of enzymes called CoA ligases. We hypothesized that an enzymatically generated fatty acyl CoA thioester could be non-enzymatically coupled to a cysteine-functionalized lysolipid (Figure 3.5a). We expressed and purified N-terminal His_6-tagged FadD2 from *E. coli* based on a previously published procedure.[19] In order to test the CoA ligase activity of the enzyme, we incubated FadD2 with sodium dodecanoate, $MgCl_2$, ATP, and CoA and detected the formation of dodecanoyl-CoA (Figure 3.5b) by HPLC-MS (m/z = 948.26 for $[M-H]^-$). When the cysteine-functionalized lysolipid **3C1** was included in the reaction medium, we observed facile formation of the amidophospholipid **3A1C1**. Interestingly, when excess sodium dodecanoate and CoA are added, the free thiol group on **3A1C1** undergoes acylation to generate the three-tailed phospholipid. We also tested several saturated (C_{10}, C_{14}, C_{16}) and unsaturated ($C_{14:1}$, $C_{16:1}$, $C_{18:1}$) fatty acids and detected product formation. Thus the substrate scope of FadD2 is broader than that of FadD10.

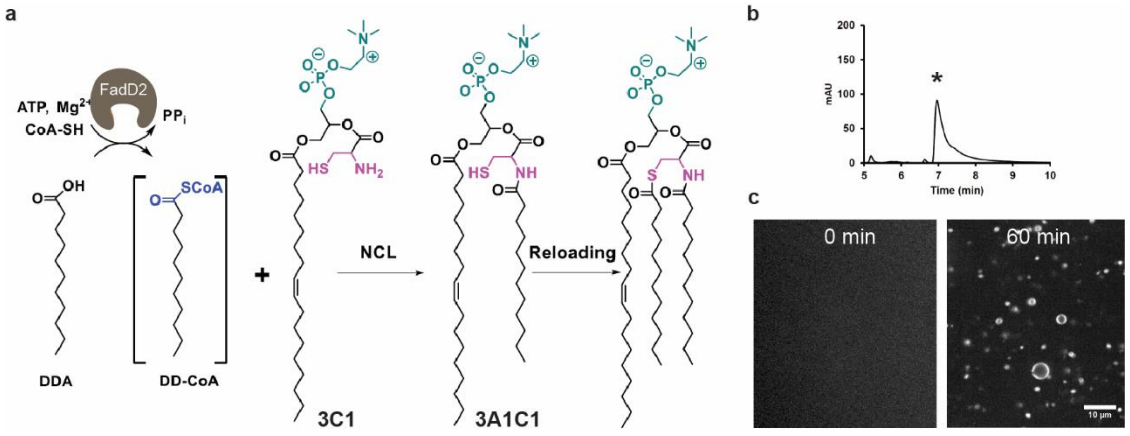


Figure 3.5. FadD2 mediated phospholipid synthesis. a. Reaction scheme corresponding to Cys-phospholipid formation. b. HPLC chromatogram (260 nm) corresponding to dodecanoyl-CoA generated by FadD2. c. Observation of in situ phospholipid membrane formation. Scale bar: 10 µm.

Initially, we sought to determine the optimum conditions for synthesis of phospholipids. We incubated lysolipid **3C1**, sodium dodecanoate, $MgCl_2$, ATP, CoA, and FadD2 at various pH's

(7.0, 7.5, 8.0, 8.5) and measured the extent of phospholipid formation over a fixed time by HPLC-ELSD-MS. Below pH 7.5, the conversion was low likely due to poor solubility of the fatty acid and weak nucleophilicity of the thiol and amine groups on cysteine. An increase in the degree of conversion could be observed with increase in pH, however, increased pH also led to instability of the phospholipid product due to ester hydrolysis at the *sn2* position. So, we carried out subsequent experiments at pH 8.0 for optimal yield and stability of the phospholipid product. The reaction mixture was clear upon mixing all components and after about an hour, micron-sized vesicles could be detected (Figure 3.5c). It is important to state that the sizes of the vesicles obtained were smaller compared to those obtained using FadD10. Also, the number of micron-sized vesicles were much fewer.

3.6 Microfluidics experiments

Since FadD10 could facilitate the *de novo* formation of phospholipids, we determined if vesicles encapsulating FadD10 could generate additional phospholipids if we continuously supplied reactive precursors. We expected to observe membrane growth and the formation of new vesicles occur if lipid synthesis is efficient. Since continual feeding of substrates requires maintaining non-equilibrium steady-state conditions for extended periods of time, we utilized a microfluidic chip[20] (Figure 3.6a) to trap giant vesicles (Figure 3.6b) composed of phospholipid **3A1B1** encapsulating Alexa Fluor 488 labeled FadD10 and ATP. We then continuously flowed reactive precursors (lysolipid **3B2**, sodium dodecanoate, ATP, and $MgCl_2$ in HEPES buffer) for 12 hours with simultaneous imaging (Figure 3.6a). In all experiments, we carefully matched the osmolarities of the initial vesicle dispersion and the flow solution using an osmometer in order to minimize non-specific effects on membrane morphology[21,22] arising from osmotic mismatch. During the first few hours, the vesicles underwent drastic morphological transformations followed

by growth (Figure 3.6c). Some vesicles exhibited membrane growth that promoted division into smaller vesicles, while maintaining their internal content, as observed using the fluorescence signal from labeled FadD10 (Figure 3.6d). This mode of division resembles that in L-form bacteria, where excess membrane synthesis increases the cellular surface-area-to-volume ratio to produce proliferation and budding events.[23] We would like to point out that the amphiphilic precursors likely enhanced the permeability of the membranes to polar solutes such as ATP due to formation of local non-lamellar phases[24,25] or transient defects.[26–28]

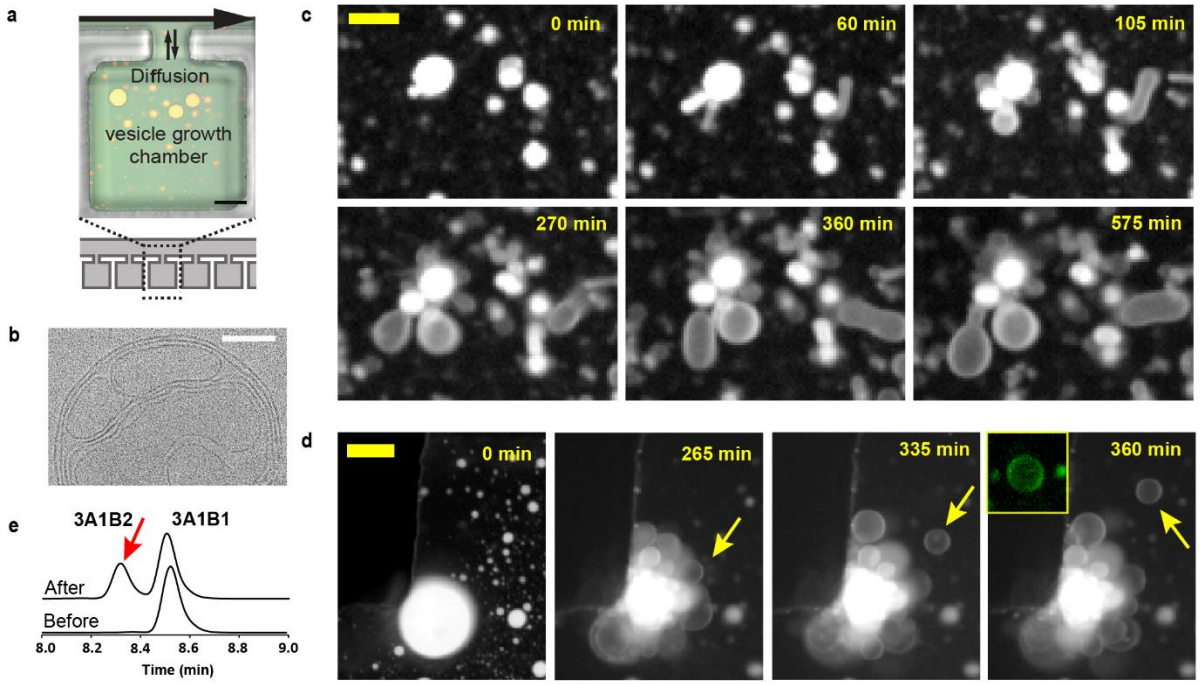


Figure 3.6. Proliferation of giant vesicles on a microfluidic device. a. Schematic representation of a microfluidic chip utilized to entrap phospholipid **3A1B1** giant vesicles encapsulating Alexa Fluor 488 labeled FadD10 and ATP. Reactive precursors (dodecanoic acid, lysolipid **3B2**, ATP and $MgCl_2$) were continuously flowed with simultaneous spinning disk confocal microscopy. Scale bar: 50 µm. b. Cryo-transmission electron microscopy (cryo-TEM) image of a large multilamellar vesicle prepared as per the vesicles used for the microfluidics experiments. Note the presence of abundant internal membranes. Scale bar: 100 nm. c. Fluorescence microscopy images (Texas Red channel) corresponding to vesicle growth over time. Scale bar: 10 µm. See Supplementary Movie 3. d. Fluorescence microscopy images (Texas Red® channel) corresponding to vesicle division. The yellow arrows indicate the formation and departure of a daughter vesicle. Scale bar: 20 µm. Inset depicts the daughter vesicle in the Alexa Fluor 488 channel. In c and d, the images are scaled logarithmically to enhance visibility of internal membranous structures. e. HPLC-MS experiment demonstrating formation of phospholipid **3A1B2** (indicated in red arrow) upon addition of lysolipid **3B2** and other precursors to phospholipid **3A1B1** giant vesicles encapsulating FadD10.

In a subsequent experiment, we replaced the amine-functionalized lysolipid **3B2** with Lyso $C_{16:0}$ PC-OH, a lysolipid unreactive towards **3A1**. During the first 5 hours, the vesicles underwent minor structural deformations without significant changes in size. The vesicles then started to solubilize likely due to surfactant action.[29,30] The chambers were depleted of vesicles over the next 3 hours. In a separate experiment, we simulated the conditions in the microfluidic device by slowly adding reactive precursors (lysolipid **3B2**, sodium dodecanoate, ATP, and $MgCl_2$ in HEPES buffer) to the giant vesicles encapsulating FadD10 and ATP using a syringe pump and observed the formation of new phospholipid **3A1B2** over time by HPLC-MS (Figure 3.6e). These results demonstrate that the vesicle growth and division events we observed in the microfluidic device were the result of the synthesis of new phospholipid, and not simply due to incorporation of amphiphilic precursors to the membranes.

3.7 FadD10 expression and phospholipid formation in PURE system

Functional synthetic cells require coupled gene expression and membrane formation to enable genetic circuits and sensors to effect cell membrane growth. We therefore linked the expression of FadD10 to lipid synthesis using a cell-free protein expression system (Figure 3.7a). In a typical reaction, plasmid DNA encoding FadD10 was first added to the PURExpress System, a minimal recombinant transcription/translation system. We confirmed FadD10 expression by SDS-PAGE (Figure 3.7b). Then, we added lipid precursors and incubated the mixture at 37 ^{o}C for 3-6 hours. HPLC-ELSD traces of the reaction mixture indicated that phospholipid **3A1B1** was formed, and no lysolipid **3B1** was detectable (Figure 3.7c). This result demonstrates that FAAs can be continuously generated in a complex medium such as PURE system and subsequently react with an amine-functionalized lysolipid to form the corresponding phospholipid. We observed a large number of vesicles in the reaction mixture using microscopy (Figure 3.7d). When we omitted

FadD10 DNA, no phospholipid synthesis occurred (Figure 3.7c), and no vesicles were observed. When we replicated the experiment using a plasmid encoding sfGFP fused to FadD10, we found that the enzyme was spontaneously associated with the *de novo* formed membranes (Figure 3.7e). These results suggest that expression of FadD10 will enable the coupling of genotype with membrane growth in synthetic cells.

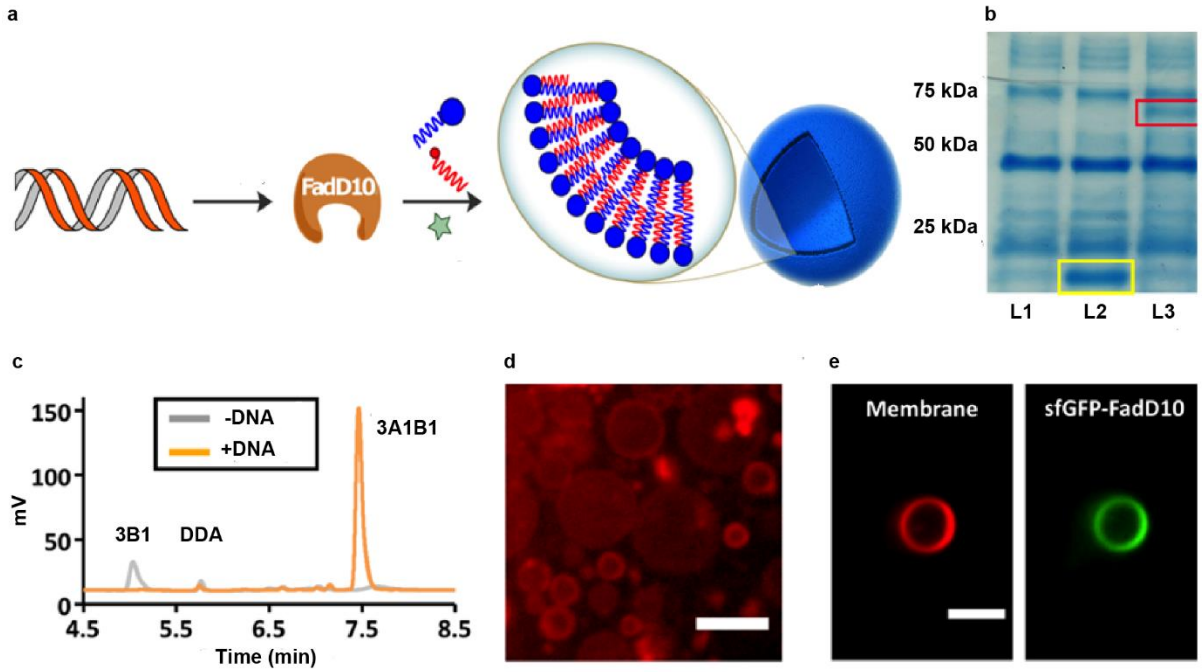


Figure 3.7. FadD10 expression and phospholipid synthesis in PURE system. a. Schematic representation of the cell-free expression of FadD10 and subsequent assembly of the *de novo* synthesized phospholipid into vesicles in the presence of appropriate reactive precursors. b. SDS-PAGE analysis of the expression of FadD10 in the *PURExpress*® System. Lane **L1**: No DNA; Lane **L2**: DHFR DNA; Lane **L3**: FadD10 DNA. (**c**) HPLC/ELSD traces monitoring the formation of phospholipid **3A1B1** by incubation of PURE System with an aqueous solution of dodecanoic acid, lysolipid **3B1**, ATP and $MgCl_2$ at 37 ºC in the absence (grey line) or presence (orange line) of plasmid DNA coding for FadD10. d. Spinning disk confocal microscopy of the *in situ* formed phospholipid vesicles in the PURE System driven by FadD10 expression. Membranes were stained using 0.1 mol% Texas Red DHPE dye. Scale bar: 5 µm. e. Localization of sfGFP-FadD10 to the membrane of the vesicles formed upon addition of the plasmid encoding the former into PURE system. External proteins were digested by Proteinase K. Scale bar: 5 µm.

3.8 One-pot synthesis of a membrane-forming lipid in PURE system mediated by FadD2

In the previous section, we described the linking of FadD10 expression with phospholipid formation in PURE system. One major limitation of that approach was that protein expression and phospholipid synthesis were carried out in separate steps. Ideally, it will be desirable to perform

both steps in the simultaneously in a 'one-pot' fashion (Figure 3.8a). We added the DNA encoding FadD2, lysolipid **3C1**, DDA, ATP, CoA, and TCEP in PURE system and did not observe any expression of protein and synthesis of phospholipid. Through thorough analysis, we identified that (i) Cys-modified lysolipids like **3C1** are hydrolytically unstable at the *sn2* ester linkage in undiluted PURE system (ii) when ATP is added without adding an equivalent of Mg^{2+}, TX-TL is inhibited (iii) addition of extra Na^{+} ions should be avoided (iv) linear DNA template is better for protein expression than plasmid template.

To avoid the issue of hydrolysis of the Cys-lysolipid, we synthesized a Cys-modified lysosphingomyelin (**3C2**). This lysolipid is stable to hydrolysis because the Cys group is attached via amide bond. We also prepared the corresponding Cys-sphingolipid **3A1C2** (Figure 3.8b) and found that it forms vesicles in aqueous media by optical microscopy and TEM (Figure 3.8c). This lipid was found to be fluid at room temperature as it did not show any peak in DSC above 1 °C. Following this, we tested the one-pot condition of FadD2 expression and lipid formation and indeed observed formation of the lipid **3A1C2** by HPLC-ELSD-MS in good conversion (Figure 3.8d). We also observed micron-sized vesicles in the same reaction mixture (Figure 3.8e). In comparison, we only obtained clear solution when no FadD2 was expressed. Finally, we tested if an additional protein can be co-expressed and membrane-formation can be achieved. We chose sfGFP-NT lysenin as a model protein for this experiment. NT-lysenin is a truncated form of the sphingomyelin-interacting protein lysenin that has been shown to selectively bind to sphingomyelins over other sphingolipid species. When one-pot condition was tested with co-expression of FadD2 and sfGFP-NT lysenin, we observed vesicles with the fluorescent protein bound to the membranes (Figure 3.8f). We tested additional conditions where we carried out sfGFP-NT lysenin expression in presence of vesicles from purified **3A1C2** and observed

localization of the protein to the membrane. When DOPC vesicles were used, no protein localization was observed suggesting that the expressed sfGFP-NT lysenin is not interacting with the membrane non-specifically.

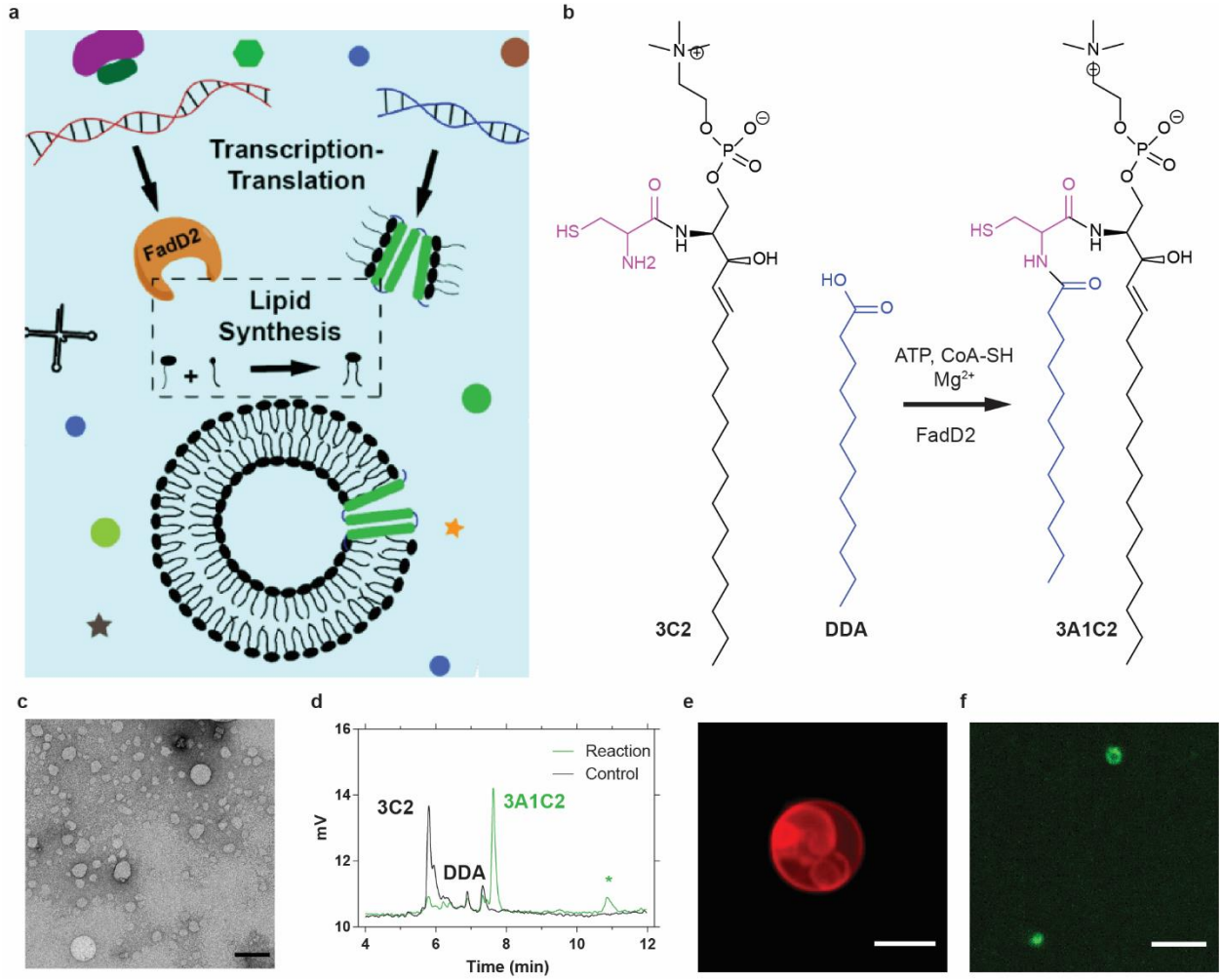


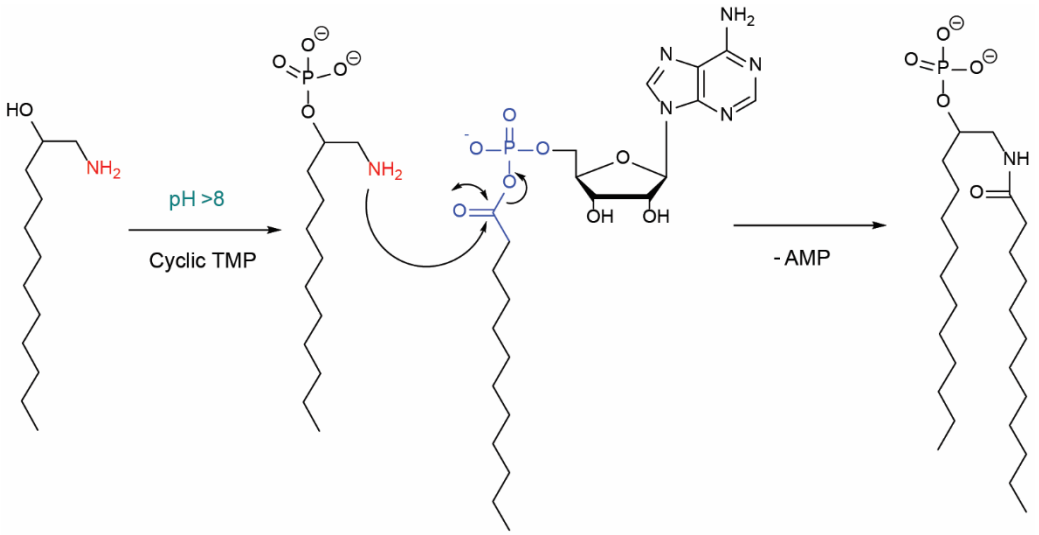
Figure 3.8. One-pot synthesis of membrane-forming lipids in PURE system mediated by FadD2. a. Schematic diagram outlining the concept of one-pot expression of the fatty acid activating enzyme FadD2 and additional protein and carrying out synthesis of membrane-forming lipids b. Reaction scheme corresponding to the formation of Cys-sphingomyelin c. Negative-staining TEM image showing vesicles formed from purified **3A1C2**. Scale bar: 250 nm. d. HPLC-ELSD chromatograms showing the formation of **3A1C2** under one-pot condition. The asterisk indicates traces of the three-tailed lipid formed by an additional acylation step on the free -SH group of **3A1C2**. e. A vesicle formed from **3A1C2** under one-pot condition in PURE system. Scale bar: 5 µm. f. Vesicles formed by one-pot condition showing sfGFP-NT lysenin localized to the membranes. Scale bar: 5 µm.

3.9 Summary and outlook

In summary, we have developed two efficient and minimal biochemical routes to synthesize phospholipids by repurposing the activity of fatty acid activating enzymes. Our strategy of using

soluble enzyme to promote phospholipid formation has a distinct advantage over previously described methods using integral membrane proteins.[27,31,32] The latter are difficult to reconstitute and require pre-existing membranes for maintaining proper structure and function. A method that relies on a soluble enzyme could be more straightforward and effective in linking genetic circuits with membrane formation in artificial cells.

Finally, our results also shed light on possible mechanisms leading to the origins of modern cellular membranes. Although it is believed that the earliest protocell membranes were composed of simple single-chain amphiphiles derived from geochemical processes,[33] there is still no consensus on how complex lipid synthesis arose during evolution. It has been hypothesized that phospholipid synthesis may have arisen in the RNA-world.[4] If so, then a ribozyme capable of driving phospholipid synthesis would be required.[5] If ribozymes were to closely mimic the function of natural lipid synthesizing acyltransferases, they would most likely be membrane-bound to interact effectively with their lipid substrates. This possibility seems unlikely, since RNAs are highly charged and polar, which makes them unlikely to embed catalytic centers within lipid membranes.[34] Our approach suggests an alternative scenario. Early ribozymes may have evolved to activate single-chain precursors such as fatty acids into corresponding acyl adenylates. The coupling of such activated precursors to other single-chain amphiphiles may have occurred spontaneously, either in micelles or within membranes. Indeed, a ribozyme capable of synthesizing acyl phosphates has previously been identified,[35] and the prebiotic plausibility of a single-chain phosphorylated amino amphiphile has been suggested (Scheme 3.2).[36] If fatty acyl adenylate (or CoA) synthesizing ribozymes could be selected through directed evolution and used to drive the *de novo* formation of membrane-forming lipids, it would suggest that catalysis by a soluble macromolecule could have played important role in the evolution of complex cellular membranes.



Scheme 3.2. Proposed scheme for reaction between a prebiotically plausible single chain amino-amphiphile[6] and a fatty acyl adenylate leading to the formation of a two-chain phospholipid-like molecule. TMP stands for trimetaphosphate.

The use of soluble enzymes to efficiently generate membrane-forming lipids *de novo* could have numerous applications. There is tremendous interest in the development of synthetic cells,[37–39] and advanced synthetic cells will undoubtedly require simplified methods to generate and maintain phospholipid membranes.[40,41] Our system provides a means to link gene expression and lipid formation in a synthetic cell. We also envision that the methodology of enzymatic *de novo* phospholipid membrane formation will enable synthesis of specific lipids in a cellular milieu, by expressing FadD10 and providing the reactive precursors. Finally, our method provides a route to synthesize proteoliposomes, which could have applications in biomaterials design and the reconstitution of disease-relevant membrane bound proteins.[8,9]

3.10 Experimental methods

3.10.1 General considerations

Commercially available 1-oleoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (Lyso $C_{18:1}$ PC-OH), 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (Lyso C_{16} PC-OH), 1-oleoyl-2-hydroxy-*sn*-glycero-3-phospho-(1´-rac-glycerol) (sodium salt) (Lyso $C_{18:1}$ PG-OH) and 1,2-

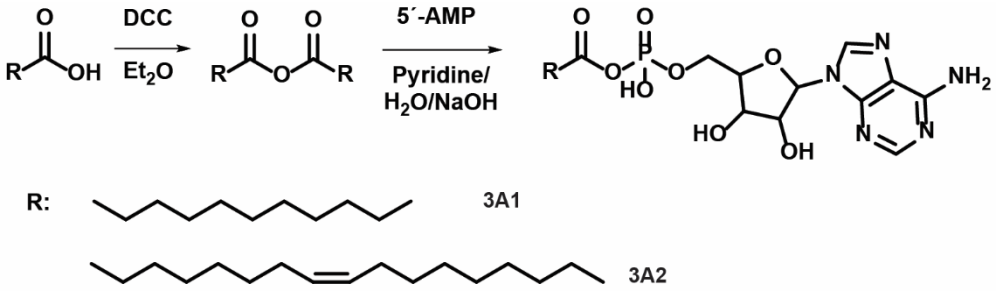
dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) were used as obtained from Avanti® Polar Lipids. Adenosine monophosphate monohydrate (AMP.H_2O), sodium dodecanoate (DDA), *N*-Boc-β-Ala-OH, *N*,*N*´-dicyclohexylcarbodiimide (DCC), 4-dimethyl-aminopyridine (DMAP), 2,2-dimethoxypropane (DMP), trifluoroacetic acid (TFA), HEPES sodium salt and 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris) were obtained from Sigma-Aldrich. Texas Red® DHPE and Alexa Fluor® 488 Succinimidyl (NHS) ester were obtained from Life Technologies. 2´,3´-*O*-(2,4,6-Trinitrophenyl) adenosine5´-triphosphate tetra(triethylammonium) salt (TNP-ATP) and lyso-sphingomyelin were obtained from Cayman Chemicals. NBD-DHPE and Rhodamine B-DHPE were obtained from Biotium. Deuterated chloroform ($CDCl_3$), methanol (CD_3OD) and dimethyl sulfoxide (d_6-DMSO) were obtained from Cambridge Isotope Laboratories.

^{1}H NMR spectra were recorded on a Varian VX-500 MHz or Jeol Delta ECA-500 MHz spectrometers, and were referenced relative to residual proton resonances in $CDCl_3$ (at δ 7.24 ppm), CD_3OD (at δ 4.87 or 3.31 ppm) or d_6-DMSO (at δ 2.50 ppm). Chemical shifts were reported in parts per million (ppm, δ) relative to tetramethylsilane (δ 0.00). ^{1}H NMR splitting patterns are assigned as singlet (s), doublet (d), triplet (t), quartet (q) or pentuplet (p). All first-order splitting patterns were designated based on the appearance of the multiplet. Splitting patterns that could not be readily interpreted are designated as multiplet (m) or broad (br). ^{13}C NMR spectra were recorded on a Varian VX-500 MHz or Jeol Delta ECA-500 MHz spectrometers, and were referenced relative to residual proton resonances in $CDCl_3$ (at δ 77.23 ppm), CD_3OD (at δ 49.15 ppm) or d_6-DMSO (at δ 39.51 ppm).

UV-Vis measurements for miscellaneous purposes were recorded on a Thermo Scientific NanoDrop 2000C spectrophotometer. Osmolarities were measured using Advanced Micro Osmometer (Advanced Instruments, Model 3300). Negative staining TEM were carried out on FEI

Tecnai^{TM} Spirit G^2 BioTWIN microscope operating at 80 kV. Cryo-EM was carried out on an FEI Tecnai^{TM} G^2 Polara microscope operating at 300 kV. Microscopy was carried out using an Olympus BX51 optical microscope and a Carl Zeiss Axio Observer Z1 spinning-disk confocal microscope. Differential scanning calorimetry (DSC) measurements were carried out on a Microcal VP-Capillary DSC instrument (Ward Lab, The Scripps Research Institute).

3.10.2 General procedure for synthesis of fatty acyl adenylates



Scheme 3.3. Synthesis of fatty acyl adenylates.

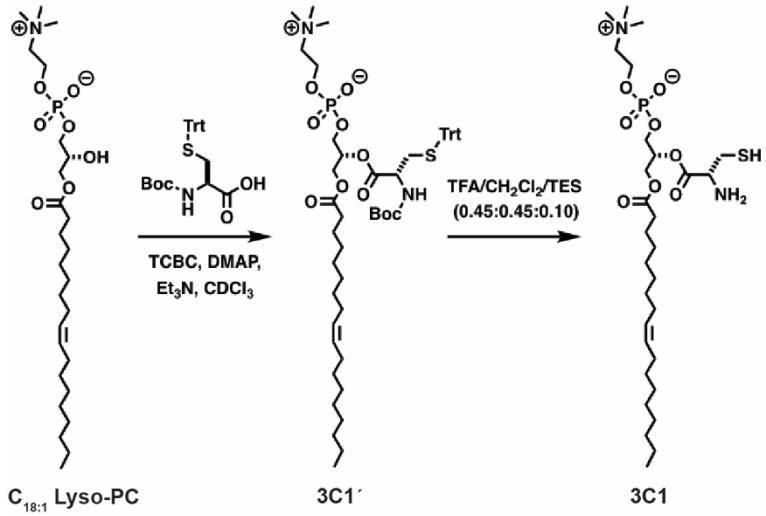
The fatty acid (1 mmol) was dissolved in anhydrous Et_2O (7 mL) and stirred for 10 min at rt. Then, a 1 M solution of DCC (1 mmol) in CH_2Cl_2 previously diluted with anhydrous Et_2O (4 mL volume) was added dropwise during 10 min. The resulting white suspension was stirred for 12 h at rt. The mixture was filtered, and the filtrate was evaporated *in vacuo* to afford a white solid. Dodecanoic anhydride was dried under high vacuum for 3 h and used without further purification.

5´-AMP.H_2O (0.4 mmol) was dissolved in a solution of 8 mL of 50% H_2O:pyridine (v/v) containing NaOH (20 mg, 0.50 mmol). Then, a solution of dodecanoic anhydride in 8 mL of THF was added in three portions. A white suspension resulted immediately, and stirring was continued at rt. After 15 min, the suspension had slowly thinned to an almost clear solution. Et_2O (20 mL) was then added and the two-phase system stirred vigorously for 2 min. Then, the ether phase was removed. This procedure was repeated three times. Distilled H_2O (3 mL) was added to the resulting whitish gel, and the pH of this suspension was brought to ~3 with HCl (5%). The gel turned into

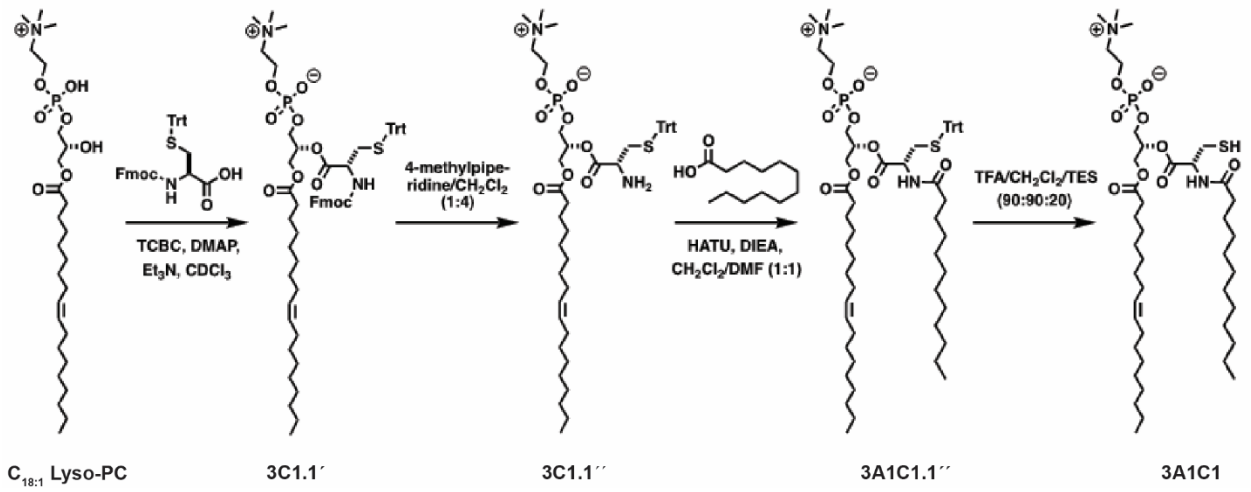
a white suspension and was extracted with Et_2O (3×15 mL) by stirring vigorously and separating the top organic layer. The white solid was transferred to a Büchner funnel and the remaining aqueous solution filtered. Successive washings with acetone (3×5 mL), and Et_2O (2×5 mL), followed by drying *in vacuo*, afforded fatty acyl adenylates as white powder.

3.10.3 General procedure for synthesis of amino acid esters of lysolipids (PC head group)

A solution of lyso $C_{18:1}$ PC-OH or lyso C16:0 PC-OH (1 eqvt.), *N*-Boc-β-Ala-OH or *N*-Boc-L-Cys (Trt) or *N*-Fmoc-L-Cys (Trt) (2.5 eqvt.), DMAP (6 eqvt.) and Et_3N (3.4 eqvt.) in dry chloroform was stirred at rt for 10 min. Then, TCBC (6.5 eqvt.) was added. After 12 h stirring at rt, H_2O (50 µL) was added to quench the acid chloride, and the solvent was removed under reduced pressure to give a pale yellow solid. The corresponding residue was dissolved in MeOH and purified by HPLC to afford the protected amino acid esters of lysolipids (**3B1´**, **3B2´, 3C1´, 3C1.1´**). Deprotection of the Boc-group is carried out in TFA/CH_2Cl_2 and the Fmoc groups are deprotected by 1:4 4-methylpiperidine/CH_2Cl_2. The functionalized lysolipids are obtained as formate salts after HPLC purification.

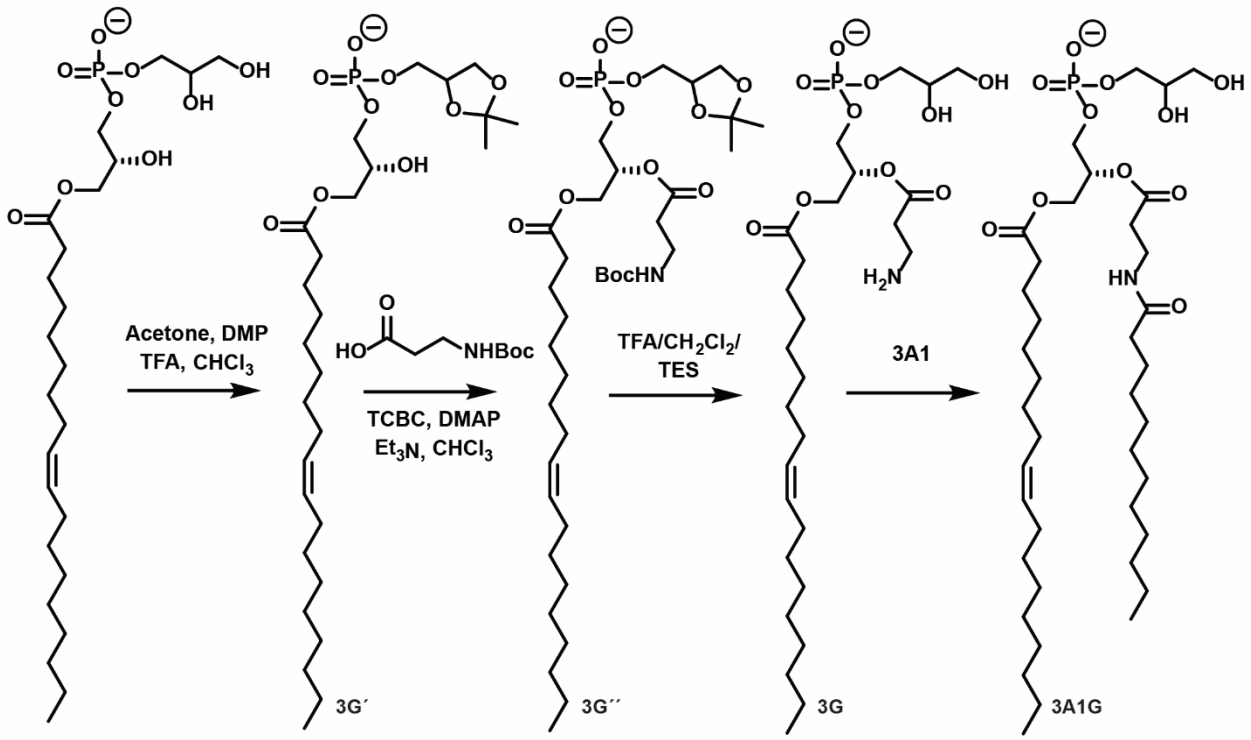


Scheme 3.4. Synthesis of Cys-functionalized lysolipids (PC head group)



Scheme 3.5. Synthesis of Cys-functionalized phospholipids (PC head group).

3.10.4 Procedure for synthesis of β-Ala lyso-PG



Scheme 3.6. Synthesis of β-Ala-functionalized phospholipids (PG head group).

A suspension of 1-oleoyl-2-hydroxy-*sn*-glycero-3-phospho-(1´-*rac*-glycerol) (sodium salt) (Lyso $C_{18:1}$ PG-OH, 50.0 mg, 93.9 µmol) in acetone (650 µL) was subsequently treated with 2,2-dimethoxypropane (DMP, 40 mL) and a catalytic amount of TFA (25 µL). After 5 min stirring at

rt, $CHCl_3$ (40 mL) was added until the lysolipid was fully dissolved. The reaction mixture was stirred for 60 h at rt. After evaporation of solvent under reduced pressure, the residue was dried *in vacuo* for 30 min to remove the TFA. The product was then re-dissolved in dry $CHCl_3$ (10 mL), and the solvent was removed by rotary evaporation. This was repeated a total of three times to remove trace TFA in the oily residue. The corresponding residue was dried *in vacuo* for 5 h to give a pale yellow oil, which was used without further purification.

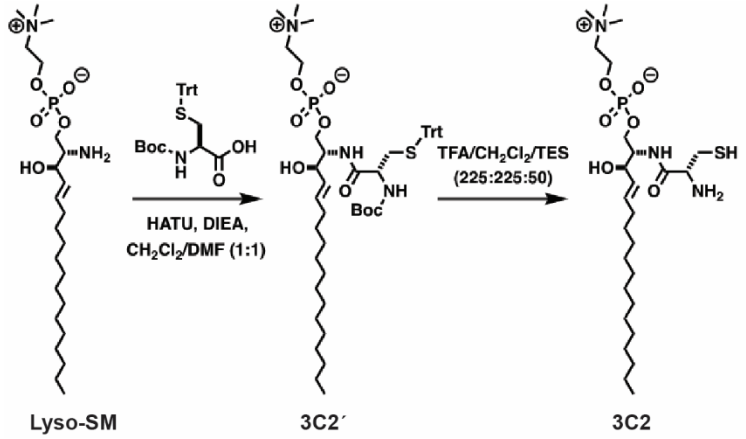
A solution of acetonide-protected 1-oleoyl-2-hydroxy-*sn*-glycero 3phosphoglycerol (**3G´**, 40.0 mg, 69.9 µmol), *N*-Boc-β-Ala-OH (33.1 mg, 174.7 µmol), DMAP (51.2 mg, 419.4 µmol) and Et_3N (34 µL) in chloroform (2.5 mL) was stirred for 10 min at rt. Then, 2,4,6-trichlorobenzoyl chloride (TCBC, 71 µL) was added and the reaction was stirred for 12 h at rt. Afterwards, H_2O (75 µL) was added to the reaction mixture to quench the acid chloride, and the solvent was removed by rotary evaporation to give a yellow oil. The crude was dissolved in 500 µL of $CHCl_3$ and purified by flash column chromatography (0-10% MeOH in $CHCl_3$) to afford **3G´´**.

Deprotection of **3G´´** (10.0 mg, 13.45 µmol) was carried out in 1 mL of TFA/CH_2Cl_2/TES (0.45:0.45:0.1) by stirring at rt for 30 min. The crude was purified by HPLC.

3.10.5 General procedure for synthesis of β-Ala phospholipids

An aqueous solution of β-Ala lysolipid (**3B1, 3B2, 3G**) was added to a 10 mM solution of **3A1** (1 eqvt.) in 100 mM HEPES pH 7.5 buffer, and stirred at rt. After 2 h, the mixture was concentrated and dried *in vacuo*. The resulting white residue was dissolved in MeOH (800 µL) and filtered, and the crude was purified by HPLC, affording the corresponding amidophospholipids.

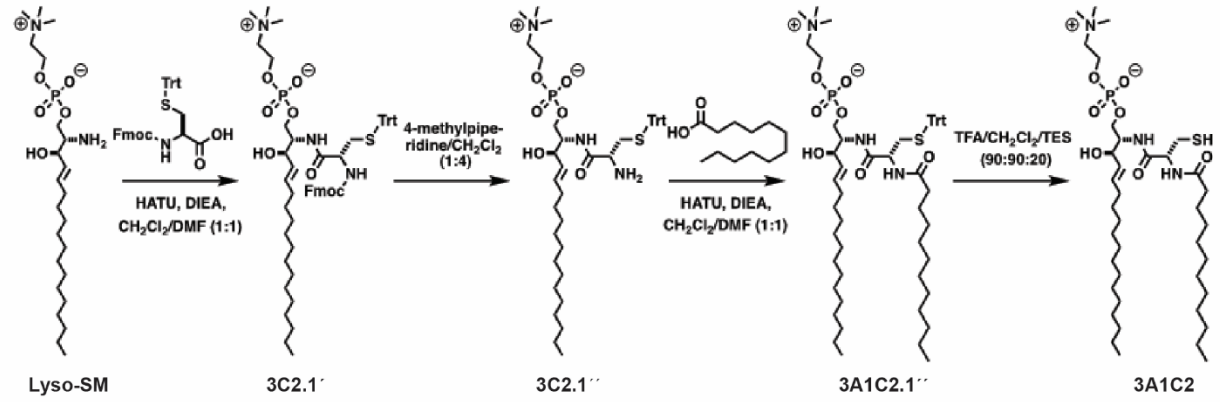
3.10.6 Procedure for synthesis of Cys-lyso sphingomyelin



Scheme 3.7. Synthesis of Cys-functionalized lysosphingomyelin.

A solution of *N*-Boc-*L*-Cys(Trt)-OH (6.0 mg, 12.9 μmol) in CH_2Cl_2/DMF (1:1) (250 μL) was stirred at 0 ºC for 10 min, and then HATU (5.4 mg, 14.2 μmol) and DIEA (9.0 μL, 51.7 μmol) were successively added. After 10 min stirring at 0 ºC, lyso-sphingomyelin (d18:1) (Lyso $d_{18:1}$ SM-NH_2, 6.0 mg, 12.9 μmol) was added. After 1 h stirring at rt, the mixture was concentrated under reduced pressure. The crude was purified by HPLC to obtain **3C2´**. Deprotection was carried out in 500 μL of TFA/CH_2Cl_2/TES (9:9:2) by stirring at rt for 30 min. The crude was purified by HPLC to afford **3C2**.

3.10.7 Procedure for synthesis of Cys-sphingomyelin



Scheme 3.8. Synthesis of Cys-functionalized sphingomyelin.

A solution of *N*-Fmoc-*L*-Cys(Trt)-OH (6.3 mg, 10.8 µmol) in CH_2Cl_2/DMF (1:1) (250 µL) was stirred at 0 ºC for 10 min, and then HATU (4.5 mg, 11.8 µmol) and DIEA (7.5 µL, 43.0 µmol) were successively added. After 10 min stirring at 0 ºC, lyso-sphingomyelin (d18:1) (Lyso $d_{18:1}$ SM-NH_2, 5.0 mg, 10.8 µmol) was added. After 1 h stirring at rt, the mixture was concentrated under reduced pressure and the residue was purified by HPLC to afford **3C2.1´**. Fmoc group was removed in 125 µL of 4-methylpiperidine/CH_2Cl_2 (1:4) by stirring at rt for 30 min to afford **3C2.1´´**. A solution of dodecanoic acid (1.0 mg, 4.9 µmol) in CH_2Cl_2/DMF (1:1) (250 µL) was stirred at 0 ºC for 10 min, and then HATU (2.1 mg, 5.4 µmol) and DIEA (3.4 µL, 19.8 µmol) were successively added. After 10 min stirring at 0 ºC, lyso-Sphingomyelin(d18:1)-[*L*-Cys(Trt)] (**3C2.1´´**, 4.0 mg, 4.9 µmol) was added. After 1 h stirring at rt, the mixture was concentrated under reduced pressure and the crude was purified by HPLC to afford **3A1C2.1´´**. The trityl group was removed in 200 µL TFA/CH_2Cl_2/TES (90:90:20) and the product **3A1C2** was HPLC-purified.

3.10.8 Compound characterization data

3A1: white solid (125.2 mg, 60%). ^{1}H NMR (d_6-DMSO): 8.55 (s, 1H, 1×CH_{Ar}), 8.30 (s, 1H, 1×CH_{Ar}), 5.94 (d, *J* = 5.6 Hz, 1H, 1×CH), 4.57 (t, *J* = 5.3 Hz, 1H, 1×CH), 4.24-3.98 (m, 4H, 2×CH+1×CH_2), 2.33 (t, *J* = 7.3 Hz, 2H, 1×CH_2), 1.55-1.35 (m, 2H, 1×CH_2), 1.31-1.09 (m, 16H, 8×CH_2), 0.88 (t, *J* = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (d_6-DMSO): 163.3, 152.8, 148.8, 148.5, 140.8, 118.7, 87.3, 83.6, 73.8, 70.4, 66.0, 34.6, 31.3, 29.0, 29.0, 28.9, 28.7, 28.7, 28.3, 24.1, 22.1, 14.0. HRMS (ESI-TOF) calculated for $C_{22}H_{37}N_5O_8P$ ($[MH]^+$) 530.2374, found 530.2374.

3A2: white solid (150.1 mg, 62%). ^{1}H NMR (d_6-DMSO): 8.57 (s, 1H, 1×CHAr), 8.31 (s, 1H, 1×CHAr), 5.93 (d, J = 5.8 Hz, 1H, 1×CH), 5.35-5.25 (m, 2H, 2×CH), 4.58 (t, J = 5.3 Hz, 1H, 1×CH), 4.22-4.15 (m, 1H, 1×CH), 4.14-4.01 (m, 3H, 1×CH+1×CH_2), 2.33 (t, J = 7.3 Hz, 2H, 1×CH_2), 2.03-1.87 (m, 4H, 2×CH_2), 1.51-1.40 (m, 2H, 1×CH_2), 1.33-1.15 (m, 20H, 10×CH_2), 0.84

(t, J = 6.9 Hz, 3H, 1×CH_3). ^{13}C NMR (d_6-DMSO): 169.5, 169.4, 152.9, 148.9, 140.9, 129.7, 129.7, 118.7, 87.3, 83.7, 73.9, 70.6, 66.1, 34.7, 31.4, 29.2, 29.1, 28.9, 28.8, 28.7, 28.7, 28.6, 28.4, 26.7, 26.6, 24.2, 22.2, 14.0. HRMS (ESI-TOF) calculated for $C_{28}H_{45}N_5O_8P^-$ ([M-H^+]) 610.3011, found: 610.3009.

3B1´: colorless film (14 mg, 76%). ^{1}H NMR ($CDCl_3$): 5.48 (br s, 1H, 1×NH), 5.41-5.27 (m, 2H, 2×CH), 5.26-5.16 (m, 1H, 1×CH), 4.46-4.24 (m, 3H, 1.5×CH_2), 4.22-4.10 (m, 1H, 0.5×CH_2), 4.09-3.93 (m, 2H, 1×CH_2), 3.90-3.75 (m, 2H, 1×CH_2), 3.51-3-18 (m, 2H, 1×CH_2), 3.36 (s, 9H, 3×CH_3), 2.67-2.43 (m, 2H, 1× CH_2), 2.29 (t, *J* = 7.6 Hz, 2H, 1×CH_2), 2.13-1.89 (m, 4H, 1×CH_2), 1.65-1.51 (m, 2H, 1×CH_2), 1.42 (s, 9H, 3×CH_3), 1.37-1.22 (m, 20H, 10×CH_2), 0.88 (t, *J* = 6.8 Hz, 3H, 1t×CH_3). ^{13}C NMR ($CDCl_3$): 173.7, 171.8, 156.1, 130.2, 129.9, 79.4, 71.3, 66.7, 63.9, 62.8, 59.4, 54.8, 36.5, 35.1, 34.2, 32.1, 29.9, 29.9, 29.7, 29.5, 29.5, 29.4, 29.3, 29.3, 28.6, 27.4, 27.3, 25.0, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{34}H_{66}N_2O_{10}P$ ([MH]$^+$) 693.4450, found 693.4446.

3B2´: white foam (49.8 mg, 78%). ^{1}H NMR ($CDCl_3$): 6.15 (br s, 1H, 0.3H, 0.3 × NH), 5.41 (br s, 0.7H, 0.7 × NH), 5.21 (m, 1H, 1 × CH), 4.40-4.26 (m, 3H, 1.5× CH_2), 4.17-4.08 (m, 1H, 0.5× CH_2), 4.07-3.94 (m, 2H, 1 × CH_2), 3.85-3.75 (m, 2H, 1 × CH_2), 3.38-3-31 (m, 2H, 1 × CH_2), 3.30 (s, 9H, 3 × CH_3), 2.59-2.46 (m, 2H, 1 × CH_2), 2.27 (t, *J* = 7.6 Hz, 2H, 1 × CH_2), 1.61-1.50 (m, 2H, 1 × CH_2), 1.40 (s, 9H, 3 × CH_3), 1.30-1.20 (m, 24H, 12 × CH_2), 0.86 (t, *J* = 6.8 Hz, 3H, 1 × CH_3). ^{13}C NMR ($CDCl_3$): 173.8, 171.8, 156.1, 79.4, 70.9, 66.3, 64.2, 62.6, 59.8, 54.5, 36.4, 34.9, 34.2, 32.1, 29.9, 29.9, 29.8, 29.7, 29.5, 29.5, 29.3, 28.6, 25.0, 22.9, 14.3. HRMS (ESI-TOF) calculated for $C_{32}H_{63}N_2O_{10}PNa$ ([M+Na]$^+$) 689.4113, found 689.4114.

3B1: colorless film (8.4 mg, 96%). ^{1}H NMR ($CDCl_3$): 8.60 (br s, 2H, 1×NH_2), 5.40-5.27 (m, 2H, 2×CH), 5.26-5.06 (m, 1H, 1×CH), 4.56-3.57 (m, 8H, 4×CH_2), 3.47-3.02 (m, 2H, 1×CH_2), 3.27

(s, 9H, 3×CH_3), 2.94-2.64 (m, 2H, 1×CH_2), 2.30 (t, *J* = 7.8 Hz, 2H, 1×CH_2), 2.11-1.89 (m, 4H, 2×CH_2), 1.69-1.48 (m, 2H, 1×CH_2), 1.47-1.15 (m, 20H, 10×CH_2), 0.88 (t, *J* = 6.6 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$): 173.6, 170.3, 130.2, 129.8, 72.4, 70.7, 66.4, 62.1, 59.6, 54.4, 35.4, 34.1, 32.7, 32.1, 29.9, 29.9, 29.7, 29.5, 29.5, 29.4, 29.3, 29.3, 27.4, 27.4, 25.0, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{29}H_{58}N_2O_8P$ ($[MH]^+$) 593.3925, found 593.3921.

3B2: colorless film (7.2 mg, 80%). ^{1}H NMR (CD_3OD): 5.19 (m, 1H, 1×CH), 4.37-4.32 (dd, 1H, J_1 = 3.9 Hz, J_2 = 12.1 Hz, 0.5×CH_2), 4.26-4.16 (m, 3H, 1.5×CH_2), 4.14-4.06 (m, 1H, 0.5 ×CH_2), 4.03-3.95 (m, 1H, 0.5×CH_2), 3.61-3.56 (m, 2H, 1×CH_2), 3.22-3.13 (m, 2H, 1×CH_2), 3.17 (s, 9H, 3×CH_3), 2.82-2.65 (m, 2H, 1×CH_2), 2.28 (t, *J* = 7.4 Hz, 2H, 1×CH_2), 1.59-1.50 (m, 2H, 1×CH_2), 1.30-1.20 (m, 24H, 12×CH_2), 0.85 (t, *J* = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 175.1, 171.5, 73.4, 67.5, 65.2, 63.3, 60.6, 54.8, 36.4, 34.9, 33.2, 32.8, 31.0, 31.0, 31.0, 31.0, 30.9, 30.9, 30.8, 30.7, 30.6, 30.4, 26.1, 23.9, 14.6. HRMS (ESI-TOF) calculated for $C_{27}H_{56}N_2O_8P$ ($[MH]^+$) 567.3769, found 567.3770.

3G´: ^{1}H NMR ($CDCl_3$): 8.28 (br s, 1H, 1×OH), 5.45-5.20 (m, 2H, 2×CH), 4.33-4.21 (m, 1H, 1×CH), 4.19-4.10 (m, 1H, 0.5×CH_2), 4.09-3.96 (m, 4H, 2×CH_2), 3.95-3.83 (m, 3H, 1.5×CH_2), 3.77-3.66 (m, 1H, 1×CH), 2.29 (t, *J* = 7.5 Hz, 2H, 1×CH_2), 2.08-1.90 (m, 4H, 2×CH_2), 1.67-1.49 (m, 2H, 1×CH_2), 1.38 (s, 3H, 1×CH_3), 1.32 (s, 3H, 1×CH_3), 1.30-1.17 (m, 20H, 10×CH_2), 0.85 (t, *J* = 6.9 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$): 174.3, 130.2, 129.9, 110.1, 74.5, 69.0, 67.8, 67.2, 66.1, 64.4, 34.2, 32.1, 30.0, 29.9, 29.7, 29.5, 29.5, 29.5, 29.4, 29.4, 27.4, 27.4, 26.8, 25.3, 25.0, 22.9, 14.3. HRMS (ESI-TOF) calculated for $C_{27}H_{50}O_9P$ ($[M]^-$) 549.3198, found 549.3193.

3G´´: pale yellow oil (43.1 mg, 83%). ^{1}H NMR ($CDCl_3$): 6.66 (d, *J* = 7.3 Hz, 1H, 1×NH), 5.37-5.31 (m, 2H, 2×CH), 5.30-5.21 (m, 1H, 1×CH), 4.43-3.62 (m, 9H, 4×CH_2+1×CH), 3.43-3.20

(m, 2H, 1×CH_2), 2.60-2.43 (m, 2H, 1×CH_2), 2.26 (t, *J* = 7.6 Hz, 2H, 1×CH_2), 2.00 (q, 4H, *J* = 6.8 Hz, 2×CH_2), 1.66-1.49 (m, 2H, 1×CH_2), 1.40-1.35 (m, 15H, 5×CH_3), 1.33-1.22 (m, 20H, 10×CH_2), 0.87 (t, *J* = 7.0 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$): 173.5, 169.9, 157.1, 130.1, 129.8, 106.7, 81.6, 74.8, 71.2, 66.6, 66.2, 64.6, 62.3, 40.1, 34.1, 32.0, 29.9, 29.9, 29.7, 29.5, 29.4, 29.3, 29.2, 28.5, 28.4, 27.4, 27.3, 26.8, 25.4, 24.9, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{35}H_{64}NO_{12}PNa$ ($[M+Na]^+$) 744.4058, found 744.4056.

3G: colorless oil (5.2 mg, 67%). ^{1}H NMR (MeOD): 5.44-5.30 (m, 2H, 2×CH), 5.29-5.19 (m, 1H, 1×CH), 4.44-4.34 (m, 1H, 0.5×CH_2), 4.33-4.22 (m, 1H, 0.5×CH_2), 4.21-4.12 (m, 1H, 0.5×CH_2), 4.09-3.98 (m, 1H, 0.5×CH_2), 3.95-3.81 (m, 2H, 1×CH_2), 3.80-3.72 (m, 1H, 1×CH), 3.67-3.50 (m, 2H, 1×CH_2), 3.23 (t, *J* = 6.2 Hz, 2H, 1×CH_2), 2.93-2.66 (m, 2H, 1×CH_2), 2.34 (t, *J* = 7.5 Hz, 2H, 1×CH_2), 2.12-1.93 (m, 4H, 2×CH_2), 1.69-1.52 (m, 2H, 1×CH_2), 1.46-1.19 (m, 20H, 10×CH_2), 0.90 (t, *J* = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (MeOD): 175.0, 171.3, 130.9, 130.8, 73.4, 72.5, 67.7, 64.9, 63.8, 63.2, 36.3, 34.8, 33.1, 32.8, 30.9, 30.8, 30.6, 30.5, 30.4, 30.3, 30.2, 30.2, 28.1, 26.0, 23.8, 14.5. HRMS (ESI-TOF) calculated for $C_{27}H_{53}NO_{10}P$ ($[MH]^+$) 582.3402, found 582.3403.

3A1B1: colorless film (5.2 mg, 80%). ^{1}H NMR ($CDCl_3$): 6.98 (br s, 1H, 1×NH), 5.43-5.27 (m, 2H, 1×CH), 5.26-5.16 (m, 1H, 1×CH), 4.46-4.21 (m, 3H, 1.5×CH_2), 4.21-4.16 (m, 1H, 0.5×CH_2), 4.10-3.88 (m, 2H, 1×CH_2), 3.86-3.70 (m, 2H, 1×CH_2), 3.55-3.40 (m, 2H, 1×CH_2), 3.33 (s, 9H, 3×CH_3), 2.67-2.42 (m, 2H, 1×CH_2), 2.28 (t, *J* = 7.7 Hz, 2H, 1×CH_2), 2.17 (t, *J* = 7.3 Hz, 2H, 1×CH_2), 2.07-1.92 (m, 4H, 2×CH_2), 1.67-1.48 (m, 4H, 2×CH_2), 1.38-1.15 (m, 36H, 18×CH_2), 0.87 (t, *J* = 6.8 Hz, 6H, 2×CH_3). ^{13}C NMR ($CDCl_3$): 174.0, 173.9, 171.9, 130.2, 129.8, 71.3, 66.6, 64.1, 62.6, 59.4, 54.7, 36.6, 35.3, 34.7, 34.2, 32.1, 32.1, 29.9, 29.9, 29.9, 29.8, 29.8, 29.7, 29.7,

29.6, 29.6, 29.5, 29.4, 29.4, 29.3, 27.4, 27.4, 26.0, 25.0, 22.9, 22.9, 14.3. HRMS (ESI-TOF) calculated for $C_{41}H_{79}N_2O_9P$ ($[MH]^+$) 775.5596, found 775.5594.

3A1B2: colorless film (5.3 mg, 94%). ^{1}H NMR ($CDCl_3$): 7.00 (br s, 1H, 1×NH), 5.35-5.09 (m, 1H, 1×CH), 4.45-4.24 (m, 3H, 1.5×CH_2), 4.20-3.93 (m, 3H, 1.5×CH_2), 3.91-3.77 (m, 2H, 1×CH_2), 3.53-3.41 (m, 2H, 1×CH_2), 3.34 (s, 9H, 3×CH_3), 2.69-2.42 (m, 2H, 1×CH_2), 2.27 (t, J = 7.7 Hz, 2H, 1×CH_2), 2.15 (t, J = 7.5 Hz, 2H, 1×CH_2), 1.69-1.38 (m, 4H, 2×CH_2), 1.37-1.06 (m, 40H, 20×CH_2), 0.85 (t, J = 6.7 Hz, 6H, 2×CH_3). ^{13}C NMR ($CDCl_3$): 174.0, 173.9, 171.9, 71.2, 66.5, 64.4, 62.5, 59.8, 54.8, 36.7, 35.3, 34.7, 34.3, 32.1, 29.9, 29.9, 29.9, 29.9, 29.8, 29.8, 29.7, 29.6, 29.6, 29.5, 29.4, 26.0, 25.1, 22.9, 14.4. HRMS (ESI-TOF) calculated for $C_{39}H_{78}N_2O_9P$ ($[M+H]^+$) 749.5439, found 749.5437.

3C1´: white solid (79.1 mg, 84%). ^{1}H NMR ($CDCl_3$): 7.38 (d, J = 7.5 Hz, 6H, 6×CH_{Ar}), 7.33-7.26 (m, 6H, 6×CH_{Ar}), 7.25-7.19 (m, 3H, 3×CH_{Ar}), 5.40-5.29 (m, 2H, 2×CH), 5.27-5.13 (m, 1H, 1×CH), 5.07 (d, J = 9.0 Hz, 1H, 1×NH), 4.41-3.88 (m, 7H, 3×CH_2+1×CH), 3.77-3.57 (m, 2H, 1×CH_2), 3.25 (s, 9H, 3×CH_3), 2.74-2.48 (m, 2H, 1×CH_2), 2.31-2.09 (m, 2H, 1×CH_2), 2.07-1.93 (m, 4H, 2×CH_2), 1.61-1.43 (m, 2H, 1×CH_2), 1.42 (s, 9H, 3×CH_3), 1.31-1.17 (m, 20H, 10×CH_2), 0.88 (t, J = 7.0 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$): 173.5, 170.4, 163.9, 155.3, 144.4, 130.1, 129.9, 129.7, 129.6, 128.2, 128.2, 127.1, 80.1, 72.3, 67.2, 66.5, 63.8, 62.7, 59.4, 54.6, 52.7, 34.1, 34.1, 34.0, 32.0, 29.9, 29.9, 29.7, 29.5, 29.5, 29.4, 29.4, 29.3, 29.2, 29.2, 28.5, 28.5, 27.4, 27.3, 24.9, 24.8, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{53}H_{80}N_2O_{10}PS$ ($[MH]^+$) 967.5266, found 967.5269.

3C1: colorless film (5.1 mg, 79%). ^{1}H NMR (CD_3OD): 5.46-5.26 (m, 3H, 3×CH), 4.49-4.35 (m, 1H, 1×CH), 4.34-4.19 (m, 3H, 1.5×CH_2), 4.16-3.99 (m, 3H, 1.5×CH_2), 3.72-3.59 (m, 2H, 1×CH_2), 3.29-3.26 (m, 1H, 0.5×CH_2), 3.23 (s, 9H, 3×CH_3), 3.19-3.00 (m, 1H, 0.5×CH_2), 2.35 (t,

J = 6.7 Hz, 2H, 1×CH_2), 2.09-1.90 (m, 4H, 2×CH_2), 1.69-1.53 (m, 2H, 1×CH_2), 1.41-1.22 (m, 20H, 10×CH_2), 0.90 (t, *J* = 6.7 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 174.9, 172.3, 131.6, 130.8, 73.7, 67.4, 64.9, 63.4, 63.2, 60.6, 54.6, 34.9, 33.7, 33.1, 30.8, 30.8, 30.7, 30.6, 30.5, 30.4, 30.3, 30.3, 30.2, 28.1, 26.0, 23.8, 14.5. HRMS (ESI-TOF) calculated for $C_{29}H_{58}N_2O_8PS$ ($[MH]^+$) 625.3651, found 625.3647.

3C1.1´: white solid (40.9 mg, 78%). MS (ESI-TOF) [m/z (%)]: 1111 ($[M +Na]^+$, 100), 1089 ($[MH]^+$, 72).

3C1.1´´: colorless film (3.4 mg, 84%). MS (ESI-TOF) [m/z (%)]: 889 ($[M +Na]^+$, 60), 867 ($[MH]^+$, 100).

3A1C1.1´´: colorless film (1.8 mg, 76%). MS (ESI-TOF) [m/z (%)]: 1071 ($[M +Na]^+$, 44), 1049 ($[MH]^+$, 100).

3A1C1: colorless film (0.8 mg, 85%). ^{1}H NMR ($CDCl_3$): 6.95 (d, *J* = 7.9 Hz, 1H, 1×NH), 5.43-5.18 (m, 3H, 3×CH), 4.88 (dd, J_1 = 8.3 Hz, J_2 = 4.5 Hz, 1H, 1×CH), 4.47-4.28 (m, 3H, 1.5×CH_2), 4.26-4.00 (m, 3H, 1.5×CH_2),3.99-3.74 (m, 2H, 1×CH_2), 3.37 (s, 9H, 3×CH_3), 3.05 (ddd, J_1 = 13.3 Hz, J_2 = 8.5 Hz, J_3 = 4.3 Hz, 1H, 0.5×CH_2), 2.91 (ddd, J_1 = 13.3 Hz, J_2 = 8.5 Hz, J_3 = 4.3 Hz, 1H, 0.5×CH_2), 2.38-2.17 (m, 4H, 2×CH_2), 2.05-1.97 (m, 4H, 2×CH_2), 1.93 (br s, 1H, 1×SH), 1.69-1.48 (m, 4H, 2×CH_2), 1.37-1.16 (m, 36H, 18×CH_2), 0.88 (t, *J* = 6.8 Hz, 6H, 2×CH_3). ^{13}C NMR ($CDCl_3$): 173.7, 173.6, 169.9, 130.2, 129.9, 72.5, 66.6, 64.2, 62.5, 59.7, 54.8, 53.6, 36.5, 34.2, 32.1, 29.9, 29.9, 29.8, 29.8, 29.7, 29.7, 29.6, 29.5, 29.5, 29.4, 29.4, 29.3, 27.4, 27.3, 25.8, 25.0, 22.8, 14.3. HRMS (ESI-TOF) calculated for $C_{41}H_{80}N_2O_9PS$ ($[MH]^+$) 807.5317, found 807.5311.

3C2´: colorless film (9.7 mg, 83%). ^{1}H NMR (CD_3OD): 7.44-7.36 (m, 6H, 6×CH_{Ar}), 7.35-7.28 (m, 6H, 6×CH_{Ar}), 7.27-7.21 (m, 3H, 3×CH_{Ar}), 5.69-5.55 (m, 1H, 1×CH), 5.47-5.31 (m, 1H,

1×CH), 4.34-4.17 (m, 2H, 1×CH_2), 4.13-4.04 (m, 1H, 1×CH), 4.03-3.96 (m, 2H, 1×CH_2), 3.95-3.88 (m, 1H, 1×CH), 3.87-3.81 (m, 1H, 1×CH), 3.65-3.53 (m, 2H, 1×CH_2), 3.19 (s, 9H, 3×CH_3), 2.55-2.47 (m, 1H, 0.5×CH_2), 2.46-2.39 (m, 1H, 0.5×CH_2), 1.90-1.77 (m, 2H, 1×CH_2), 1.45 (s, 9H, 3×CH_3), 1.51-1.37 (m, 22H, 11×CH_2), 0.89 (t, J = 7.0 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 172.5, 157.3, 146.0, 134.9, 130.8, 130.7, 129.1, 128.0, 80.7, 72.1, 68.0, 67.5, 65.5, 60.4, 55.4 and 55.2, 54.7, 36.0, 33.4, 33.1, 30.9, 30.8, 30.8, 30.7, 30.5, 30.4, 30.3, 28.8, 23.8, 14.5. MS (ESI-TOF) [m/z (%)]: 932 ($[M+Na]^+$, 16), 910 ($[MH]^+$, 100).

3C2: colorless film (2.2 mg, 78%). ^{1}H NMR (CD_3OD): 5.85-5.70 (m, 1H, 1×CH), 5.56-5.43 (m, 1H, 1×CH), 4.35-4.22 (m, 2H, 1×CH_2), 4.15-4.01 (m, 3H, 1×CH+1×CH_2), 3.98-3.88 (m, 1H, 1×CH), 3.73-3.59 (m, 2H, 1×CH_2), 3.49-3.41 (m, 1H, 1×CH), 3.36-3.33 (m, 1H, 0.5×CH_2), 3.28-3.26 (m, 1H, 0.5×CH_2), 3.23 (s, 9H, 3×CH_3), 2.11-1.98 (m, 2H, 1×CH_2), 1.48-1.22 (m, 22H, 11×CH_2), 0.90 (t, J = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 172.5, 169.5*, 135.6, 130.9, 72.6, 67.4, 65.9, 60.5, 55.8, 54.6, 54.0, 33.6, 33.1, 30.9, 30.9, 30.9, 30.8, 30.6, 30.5, 30.5, 23.8, 14.5. MS (ESI-TOF) [m/z (%)]: 568 ($[MH]^+$, 100). (* Carbon peak corresponding to formic acid since **3C2** was obtained as a formate salt).

3C2.1´: white solid (10.1 mg, 91%). MS (ESI-TOF) [m/z (%)]: 1054 ($[M+Na]^+$, 23), 1032 ($[MH]^+$, 100).

3C2.1´´: colorless film (5.3 mg, 84%). MS (ESI-TOF) [m/z (%)]: 832 ($[M+Na]^+$, 37), 810 ($[MH]^+$, 100).

3A1C2.1´´: colorless film (3.9 mg, 84%). MS (ESI-TOF) [m/z (%)]:1014 ($[M+Na]^+$, 21), 992 ($[MH]^+$, 100).

3A1C2: colorless film (1.3 mg, 87%). ^{1}H NMR (CD_3OD): 5.73 (dt, J_1 = 15.3 Hz, J_2 = 6.6 Hz, 1H, 1×CH), 5.46 (ddt, J_1 = 15.3 Hz, J_2 = 7.6 Hz, J_3 = 1.5 Hz, 1H, 1×CH), 4.51 (dd, J_1 = 7.7

Hz, J_2 = 4.9 Hz, 1H, 1×CH), 4.36-4.20 (m, 2H, 1×CH+0.5×CH_2), 4.18-4.04 (m, 2H, 1×CH_2), 4.01-3.90 (m, 2H, 1×CH+0.5×CH_2), 3.64 (t, J = 4.7 Hz, 2H, 1×CH_2), 3.22 (s, 9H, 3×CH_3), 2.85 (dd, J_1 = 13.9 Hz, J_2 = 4.9 Hz, 1H, 0.5×CH_2), 2.74 (dd, J_1 = 13.8 Hz, J_2 = 7.7 Hz, 1H, 0.5×CH_2), 2.36-2.21 (m, 2H, 1×CH_2), 2.10-1.97 (m, 2H, 1×CH_2), 1.69-1.55 (m, 2H, 1×CH_2), 1.45-1.21 (m, 38H, 19×CH_2), 0.90 (t, J = 6.9 Hz, 6H, 2×CH_3). ^{13}C NMR (CD_3OD): 174.9, 170.6, 134.0, 129.5, 70.8, 66.0, 64.1, 59.1, 55.4, 54.1, 53.2, 35.5, 32.1, 31.7, 29.5, 29.4, 29.4, 29.4, 29.4, 29.4, 29.3, 29.1, 29.1, 29.0, 28.9, 26.2, 25.6, 22.4, 13.1. HRMS (ESI-TOF) calculated for $C_{38}H_{77}N_3O_7PS$ ($[MH]^+$) 750.5210, found 750.5214.

3.10.9 Expression and purification of FadD10

The *fadD10* (*Rv0099*)-pDEST17 plasmid was transformed into Novagen BL21(DE3) competent *E. coli* cells and grown overnight at 37 ºC in Luria-Bertani (LB) broth containing 0.1 mg/mL of ampicillin. Afterward, 1 mL of the overnight culture was used to inoculate 1 L of freshly autoclaved LB medium containing 0.1 mg/mL of ampicillin. The rest of the overnight culture was stored as 25% glycerol stocks at -80 ºC. The culture was grown at 37 ºC in a shaker-incubator till the OD_{600} reached about 0.7. Over-expression of FadD10 was induced by addition of 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG). The cells were then grown for 18-20 h at 18 ºC, after which the cells were harvested by centrifuging at 6,000 rcf for 20 min at 4 ºC. The pellet was resuspended by vortexing in 10 mL of lysis buffer containing 25 mM Tris pH 8.0, 0.5 M NaCl, 2 mM β-mercaptoethanol, 1 mg/mL lysozyme and a cocktail of protease inhibitors (SigmaFast®). Following cell lysis by an ultrasonicator probe, debris were removed by centrifuging (13,000 rcf, 20 min, 4 ºC). The supernatant was incubated in a gravity column with Ni^{2+}-nitrilotriacetate (Ni-NTA) agarose resin pre-equilibrated with 10 mM imidazole on a shaker table for 3 h at 4 ºC. Next, the flow-through was discarded and the resin was washed with 10 mL each of 20 mM imidazole

and 50 mM imidazole solutions. Elution was performed with 250 mM imidazole and the collected fractions were stored in aliquots containing 10% glycerol each (final concentration: 6.75 mg/mL) and stored at -80 ºC.

3.10.10 Expression and purification of FadD2

The pET28b-*fadD2* (*Rv0270*) and pGro7 (Takara) plasmids were co-transformed into BL21(DE3) competent *E. coli* cells and grown overnight at 37 ºC in Luria-Bertani (LB) broth containing kanamycin and chloramphenicol. Afterward, 1 mL of the overnight culture was used to inoculate 1 L of freshly autoclaved LB medium containing 0.05 mg/mL of kanamycin and 0.1 mg/mL of chloramphenicol. The culture was grown at 37 ºC in a shaker-incubator till the OD_{600} reached ~0.7. Over-expression of FadD2 was induced by addition of 1 mM IPTG. The cells were then grown for 18 h at 18 ºC, after which the cells were harvested by centrifuging at 6,000 rcf for 20 min at 4 ºC. The pellet was resuspended by vortexing in 10 mL of lysis buffer containing 50 mM Na-phosphate pH 8.0, 10 mM imidazole pH 8.0, 0.5 M NaCl, 2 mM β-ME, Mg-ATP (1 mM) 1 mg/mL lysozyme and PMSF (1 mM). Following cell lysis by an ultrasonicator probe, debris were removed by centrifuging (10,000 rcf, 30 min, 4 ºC). The supernatant was incubated in a gravity column with Ni^{2+}-nitrilotriacetate (Ni-NTA) agarose resin for 3 h at 4 ºC. Next, the flow-through was discarded and the resin was washed with 20 mM (10 mL), 35 mM (5 mL), 40 mM (2 mL) and 50 mM (2 mM) imidazole solutions buffered with 50 mM Na-phosphate and containing 1 mM Mg-ATP. Finally, His_6-tagged FadD2 was eluted with 250 mM imidazole and collected in 1 mL fractions and the pure fractions were pooled and concentrated using 30 kDa MWCO centrifuge filters at 4 ºC. They were stored in small aliquots containing 10% glycerol each (final concentration: 2.4 mg/mL) and stored at -80 ºC.

3.10.11 *De novo* phospholipid formation mediated by FadD10

In a typical *de novo* phospholipid synthesis reaction, to a 2.0 µL of lysolipid **3B1** solution (10 mM stock in H_2O) were successively added 2 µL HEPES (1 M), 1.8 µL of $MgCl_2$ (100 mM stock solution), 0.6 µL of ATP (100 mM stock solution) and 9.6 µL of H_2O. Next, 3.3-13 µg of FadD10 (from a 10% glycerol stock, 6.75 mg/mL) were added, and the solution was mixed by gentle tapping. Afterward, 2.0 µL of dodecanoic acid (as 10 mM sodium dodecanoate stock solution) were added. The reaction mixture was kept incubated in a 37 ºC water bath. Small aliquots (~2 µL) were taken out at various time points and placed on a glass slide for microscopic observations. We found that the pH range 7.5-8.5 is optimum for vesicle formation. Below pH 7.5, vesicle formation was impaired due to aggregation caused by poor solubility of dodecanoic acid. Above pH 8.5, the amino-lysolipids and phospholipids had reduced stability due to hydrolysis at the *sn2* position. We carried out *in situ* vesicle formation experiments in the concentration range 0.1-1.0 mM of the amphiphilic precursors.

3.10.12 FRET assay for studying *de novo* phospholipid membrane formation

A phospholipid synthesis reaction was set up using 0.5 mM lysolipid **2**, 0.5 mM sodium dodecanoate, 7.5 mM $MgCl_2$, 2.5 mM ATP, and FadD10 in HEPES buffer (pH 8.0, 100 mM) in presence of 0.5 µM of each of the FRET dyes NBD-DHPE (donor; λ_{em}: 530 nm) and Rhodamine-DHPE (acceptor; λ_{em}: 586 nm). The reaction mixture was placed in a 384-well plate at 37 ºC. 430 nm (bandwidth: 7.5 nm) was used as the wavelength for excitation and emission was monitored at 530 nm (bandwidth: 7.5 nm) and 586 nm (bandwidth: 7.5 nm). The ratio of the fluorescence intensities at 530 nm and 586 nm calculated and plotted with time. In a control experiment, ATP was substituted with GTP and fluorescence measurements were carried out using same parameters.

3.10.13 Microfluidics experiments

Giant vesicles were prepared by adding the following components in the given order: 20 µL lysolipid **3B1** (10 mM in 50 mM HEPES pH 7.5 buffer), 9 µL $MgCl_2$ (100 mM), 3 µL ATP (100 mM), 20 µL ddH_2O, 20 µL 100 mM HEPES pH 7.5 buffer, 6.5 µL FadD10 (6.75 mg/mL), 1.5 µL Alexa Fluor® 488 labelled FadD10 (14 µM), 0.5 µL Texas Red® DHPE (100 µM in EtOH) and 20 µL FAA **3A1** (10 mM in 50 mM HEPES pH 7.5 buffer). The reaction mixture was placed in a glass vial and stirred at rt for 2 h, obtaining phospholipid **3A1B1** vesicles encapsulating FadD10. We also carried out vesicle formation following the described procedure in presence of 0.1 mM TNP-ATP, a surrogate of ATP that binds to ATP-binding proteins resulting in marked increase in fluorescence intensity. We observed that the fluorescence intensity is significantly higher inside the vesicles compared to the background, suggesting that FadD10 is encapsulated highly efficiently in functional form.

Long-term vesicle growth was observed in a microfluidic device fabricated from polydimethylsiloxane (PDMS) and plasma bonded to cover glass. Pre-formed vesicles were vacuum-loaded into chambers that were 230 µm wide, 230 µm long and 80 µm high. The chip was then connected to a reservoir of precursor solution (40 µM lysolipid **3B2**, 40 µM sodium dodecanoate, 3 mM ATP, 9 mM $MgCl_2$, 40 mM HEPES pH 7.5 buffer) and placed in a temperature-controlled microscope box at 37 ºC. In each experiment, the osmolarities of the vesicle solution and the flow solution were measured and adjusted by adding glycerol accordingly prior to loading using an osmometer. Precursor solution flow was started immediately after the first image of the time-lapse movies were acquired. Flow was regulated to approximately 20 µL/h using air pressure. Precursors flowing in the main channel of the device entered vesicle chambers by diffusion through connections that were 50 µm wide, 60 µm long and 50 µm high. This

architecture allows observation of vesicles under shear-free conditions without convective flow while ensuring a constant supply of precursors, which are expected to completely equilibrate within 10 min. Alternatively, in the case of control experiments, lysolipid **3B2** was replaced by 1-palmitoyl-2-hydroxy-s*n*-glycero-3-phosphocholine (Lyso $C_{16:0}$ PC-OH).

3.10.14 PURE system experiments with FadD10

In a typical 10 µL protein expression reaction, the following components were added in the given order: 4 µL of *Solution A* (containing amino acids, energy factors, etc.), 3 µL *Solution B* (containing ribosomes, aminoacyl tRNA synthetases, etc), 0.2 µL (4 U) Murine RNase Inhibitor (New England Biolabs), *x* µL nuclease free H_2O, *y* µL DNA. Initially, protein expression was optimized using different amounts of FadD10 DNA: 50 ng, 100 ng, 150 ng and 200 ng. A *negative control* (no DNA) and a *positive control* (50 ng DHFR DNA) experiment were also carried out at the same time. All reaction mixtures were incubated in a 37 ºC water bath for 3 h and then analyzed by SDS-PAGE. 2.5 µL of each reaction mixture was loaded onto a precast 4-20% polyacrylamide gel (Mini Protean, Bio-Rad). The gel was run for 50 min at 130 V and then stained with Coomassie Blue (Instant Blue®, CBS Scientific) for 1 h. The stained gel was washed with distilled water and imaged with a commercial scanner. Expression of FadD10 was observed as expected. It was found that 50 ng DNA was sufficient for a 5 µL protein expression reaction for FadD10. Increasing the amount of DNA did not lead to any significant change in the protein expression. Expression of sfGFP-FadD10 was carried out in a similar manner and monitored using spectrofluorimetrically.

For phospholipid synthesis reactions, FadD10 was first expressed in PURExpress® System with the total reaction volume as 10 µL. A mixture containing precursors for lipid synthesis was prepared by adding 1.25 µL lysolipid **3B1** (5 mM), 0.25 µL sodium dodecanoate (25 mM), 1.0 µL ATP (100 mM), 2.0 µL $MgCl_2$ (100 mM) and 10.5 µL 100 mM HEPES pH 8.0 buffer. Then, it

was added to the PURExpress® System solution, and the resulting reaction mixture was further incubated at 37 ºC for 3 h with tumbling. Afterward, 25 µL of MeOH was added to the reaction mixture, which cause formation of a white precipitate. The precipitate was separated by brief centrifugation and the supernatant was collected and injected into HPLC-ELSD-MS. Formation of phospholipid **3A1B1** was confirmed from the retention time and verified by mass spectrometry. In separate experiments, it was also found that higher concentrations of phospholipid **3A1B1** (tested up to 1 mM) can be synthesized, albeit with longer incubation times (up to 12 h). Alternatively, in a control experiment, where FadD10 DNA was omitted in the first step, no phospholipid formation was observed.

Synthesis of phospholipid **3A1B1** (in the concentration range 0.25-1.0 mM) was carried out in PURExpress® System as described before using FadD10 and sfGFP-FadD10 plasmids. Vesicle formation can be improved and aggregate formation can be minimized by tumbling, occasional mild vortexing and tapping the reaction tube. Longer incubation time (24 h) was required when lipid synthesis (0.5 mM) was carried out using sfGFP-FadD10 plasmid. The membranes were stained using 0.1 mol% Texas Red® DHPE. An aliquot of the corresponding sample was placed on a glass slide and observed using fluorescence microscopy. In a control experiment where no DNA was added in the beginning, vesicles were not observed.

Notes on chapter

Chapter 3 is adapted majorly from the materials published in Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A Minimal Biochemical Route towards de novo Formation of Synthetic Phospholipid Membranes. *Nat. Commun*. 2019, 10, 300. There is a significant amount of unpublished material as well, especially those related to FadD2 coauthored with Cho, C. J.;

Brea, R. J. & Devaraj, N. K. and those on sphingomyelin coauthored with Devaraj, N. K. The dissertation author was the primary investigator for works described in this chapter.

We thank James Sacchettini (Texas A&M University) and Anthony Baughn (University of Minnesota) for generously providing us the plasmid for FadD10 and FadD2 respectively. We also thank Prof. Jeff Hasty, Andriy Didovyk, Omar Din and Ryan Johnson at UC San Diego for their collaboration with the microfluidic chip fabrication and experimental setup.

3.11 References

1. Yao, J. & Rock, C. O. Phosphatidic acid synthesis in bacteria. *Biochim. Biophys. Acta* **1831,** 495–502 (2013).

2. Blobel, G. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **77,** 1496–1500 (1980).

3. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409,** 387–390 (2001).

4. Blain, J. C. & Szostak, J. W. Progress toward synthetic cells. *Annu. Rev. Biochem.* **83,** 615–640 (2014).

5. Budin, I. & Szostak, J. W. Physical effects underlying the transition from primitive to modern cell membranes. *Proc. Natl. Acad. Sci.* **108,** 5249–5254 (2011).

6. Budin, I. & Devaraj, N. K. Membrane assembly driven by a biomimetic coupling reaction. *J. Am. Chem. Soc.* **134,** 751–753 (2012).

7. Brea, R. J., Cole, C. M. & Devaraj, N. K. In Situ Vesicle Formation by Native Chemical Ligation. *Angew. Chem. Int. Ed.* **53**, 14102-14105 (2014)..

8. Cole, C. M., Brea, R. J., Kim, Y. H., Hardy, M. D., Yang, J. & Devaraj, N. K. Spontaneous reconstitution of functional transmembrane proteins during bioorthogonal phospholipid membrane synthesis. *Angew. Chem. Int. Ed.* **54,** 12738–12742 (2015).

9. Brea, R. J., Cole, C. M., Lyda, B. R., Ye, L., Prosser, R. S., Sunahara, R. K. & Devaraj, N. K. In Situ Reconstitution of the Adenosine A_{2A} Receptor in Spontaneously Formed Synthetic Liposomes. *J. Am. Chem. Soc.* **139,** 3607–3610 (2017).

10. Iwasaki, Y. & Yamane, T. Enzymatic synthesis of structured lipids. *Adv. Biochem. Eng. Biotechnol.* **90**, 151-171 (2004).

11. Schmelz, S. & Naismith, J. H. Adenylate-forming enzymes. *Curr. Opin. Struct. Biol.* **19,** 666–671 (2009).

12. Dhiman, R. S., Opinska, L. G. & Kluger, R. Biomimetic peptide bond formation in water with aminoacyl phosphate esters. *Org. Biomol. Chem.* **9,** 5645–5647 (2011).

13. Kluger, R. & Cameron, L. L. Activation of acyl phosphate monoesters by lanthanide Ions: enhanced reactivity of benzoyl methyl phosphate. *J. Am. Chem. Soc.* **124,** 3303–3308 (2002).

14. Schall, O. F., Suzuki, I., Murray, C. L., Gordon, J. I. & Gokel, G. W. Characterization of acyl adenyl anhydrides: differences in the hydrolytic rates of fatty acyl-AMP and aminoacyl-AMP derivatives. *J. Org. Chem.* **63,** 8661–8667 (1998).

15. Liu, Z., Ioerger, T. R., Wang, F. & Sacchettini, J. C. Structures of Mycobacterium tuberculosis FadD10 protein reveal a new type of adenylate-forming enzyme. *J. Biol. Chem.* **288,** 18473–18483 (2013).

16. Wang, F., Langley, R., Gulten, G., Wang, L. & Sacchettini, J. C. Identification of a type III thioesterase reveals the function of an operon crucial for Mtb virulence. *Chem. Biol.* **14,** 543–551 (2007).

17. Hanczyc, M. M. & Szostak, J. W. Experimental Models of Primitive Cellular Compartments: Encapsulation, Growth, and Division. *Science* **302,** 618–622 (2003).

18. Brea, R. J., Cole, C. M. & Devaraj, N. K. In situ vesicle formation by native chemical ligation. *Angew. Chem. Int. Ed.* **53,** 14102–5 (2014).

19. Rosen, B. C., Dillon, N. A., Peterson, N. D., Minato, Y. & Baughn, A. D. Long-Chain Fatty Acyl Coenzyme A Ligase FadD2 Mediates Intrinsic Pyrazinamide Resistance in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother*. **61,** 1–13 (2017).

20. Kolnik, M., Tsimring, L. S. & Hasty, J. Vacuum-assisted cell loading enables shear-free mammalian microfluidic culture. *Lab Chip* **12,** 4732–4737 (2012).

21. Zhu, T. F. & Szostak, J. W. Coupled growth and division of model protocell membranes. *J. Am. Chem. Soc.* **131,** 5705–5713 (2009).

22. Zhu, T. F. & Szostak, J. W. Exploding vesicles. *J. Syst. Chem.* **2,** 4 (2011).

23. Mercier, R., Kawai, Y. & Errington, J. Excess membrane synthesis drives a primitive mode of cell proliferation. *Cell* **152,** 997–1007 (2013).

24. Richard Burack, W., Dibble, A. R. G., Allietta, M. M. & Biltonen, R. L. Changes in vesicle morphology induced by lateral phase separation modulate phospholipase A2 activity. *Biochemistry* **36,** 10551–10557 (1997).

25. Schroder, A. P., Crassous, J. J., Marques, C. M. & Olsson, U. Rapid confocal imaging of vesicle-to-sponge phase droplet transition in dilute dispersions of the C10E3 surfactant. *Sci.*

Rep. **9,** 2292 (2019).

26. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chem. Int. Ed.* **51,** 3114–3118 (2012).

27. Scott, A., Noga, M. J., de Graaf, P., Westerlaken, I., Yildirim, E. & Danelon, C. Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One* **11,** 1–23 (2016).

28. Jin, L., Kamat, N. P., Jena, S. & Szostak, J. W. Fatty Acid/Phospholipid Blended Membranes: A Potential Intermediate State in Protocellular Evolution. *Small* **14,** 1704077 (2018).

29. Tomita, T., Sugawara, T. & Wakamoto, Y. Multitude of morphological dynamics of giant multilamellar vesicles in regulated nonequilibrium environments. *Langmuir* **27,** 10106–10112 (2011).

30. Shoji, Y., Igarashi, T., Nomura, H., Eitoku, T. & Katayama, K. Liposome solubilization induced by surfactant molecules in a microchip. *Anal. Sci.* **28,** 339 (2012).

31. Stano, P. & Luisi, P. L. Achievements and open questions in the self-reproduction of vesicles and synthetic minimal cells. *Chem. Commun.* **46,** 3639–3653 (2010).

32. Exterkate, M., Caforio, A., Stuart, M. C. A. & Driessen, A. J. M. Growing Membranes in Vitro by Continuous Phospholipid Biosynthesis from Free Fatty Acids. *ACS Synth. Biol.* **7,** 153–165 (2018).

33. Segré, D., Ben-Eli, D., Deamer, D. W. & Lancet, D. The lipid world. *Orig. Life Evol. Biosph.* **31,** 119–145 (2001).

34. Khvorova, A., Kwak, Y.-G., Tamkun, M., Majerfeld, I. & Yarus, M. RNAs that bind and change the permeability of phospholipid membranes. *Proc. Natl. Acad. Sci.* **96,** 10649–10654 (1999).

35. Kumar, R. K. & Yarus, M. RNA-catalyzed amino acid activation. *Biochemistry* **40,** 6998–7004 (2001).

36. Mullen, L. B. & Sutherland, J. D. Formation of potentially prebiotic amphiphiles by reaction of *β*-hydroxy-*n*-alkylamines with cyclotriphosphate. *Angew. Chem. Int. Ed.* **46,** 4166–4168 (2007).

37. Buddingh, B. C. & van Hest, J. C. M. Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Acc. Chem. Res.* **50,** 769–777 (2017).

38. Engelhart, A. E., Adamala, K. P. & Szostak, J. W. A simple physical mechanism enables

homeostasis in primitive cells. *Nat. Chem.* **8,** 448–453 (2016).

39. Adamala, K. & Szostak, J. W. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* **5,** 495–501 (2013).

40. Brea, R. J., Hardy, M. D. & Devaraj, N. K. Towards self-assembled hybrid artificial cells: novel bottom-up approaches to functional synthetic membranes. *Chem. Eur. J.* **21,** 12564–12570 (2015).

41. Kurihara, K., Tamura, M., Shohda, K., Toyota, T., Suzuki, K. & Sugawara, T. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* **3,** 775–781 (2011).

Chapter 4 | Artificial Cell Model based on Vesicles from Galactopyranose-derived Single-Chain Amphiphiles

4.1 Introduction

In Chapters 2 and 3 we described the syntheses and applications of synthetic phospholipids in forming cell-mimetic vesicle compartments. Although dialkyl amphiphiles such as phospholipids better mimic the membrane-bound architecture of cells, there is a significant interest in the use of membrane-forming single-chain amphiphiles (SCAs) in designing protocellular models.[1–4] SCAs offer several advantages, including simplicity of molecular design and ease of synthesis.[5] Importantly, prior studies have demonstrated that vesicles from SCAs are more permeable to charged small molecule precursors, enabling internal polymerization of macromolecules such as oligonucleotides.[5,6] However, fatty acids, the most common SCAs used to form vesicles, suffer from several drawbacks.[7,8] They have a narrow pH range over which they form stable vesicles.[9] Moreover, vesicles formed from fatty acids are not tolerant of divalent cations, such as magnesium and calcium, which are essential in all living system.[10] SCAs also have limited compatibility with various biochemical processes. For instance, fatty acids are known to inhibit nucleic acid polymerases,[11,12] making DNA/RNA amplification or transcription reactions difficult to perform in such vesicles. Additionally, integral membrane proteins have not been demonstrated to function within membranes consisting solely of SCAs. It would thus be of considerable value to develop membranes consisting of SCAs where these limitations could be overcome. Here we report that *S*-oleoyl β-D-1-thiogalactopyranose (**OTG**), a newly synthesized class of water-soluble SCA, is capable of spontaneously forming highly stable micron-sized vesicles (Figure 4.1a). The uncharged glycolipid can self-assemble into vesicles at sub-millimolar

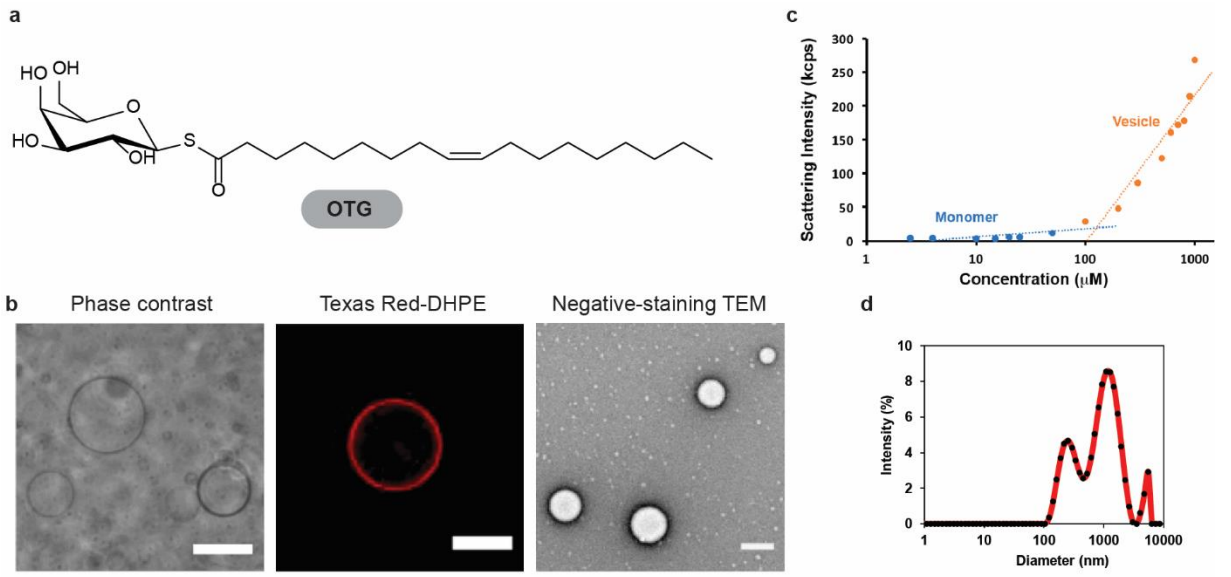


Figure 4.1. Vesicle formation from **OTG**. a. Chemical structure of *S*-oleoyl β-D-1-thio-galactopyranose (**OTG**). b. Phase contrast, fluorescence and negative-staining TEM images of **OTG** vesicles (scale bars: 10 μm, 5 μm, 500 nm). c. Concentration-dependent variation of scattering intensity in DLS. This plot is used to estimate the cac. d. A representative DLS intensity profile corresponding to a dispersion 5 mM OTG in water.

concentrations over a wide pH range and in the presence of free divalent metal cations. Moreover, the resulting vesicles are compatible with the conditions necessary for biomolecules such as proteins and nucleic acids to function. Thus, the thiogalactopyranose-derived SCA provides a simple model system to explore the characteristics of simple membranes, as well as host confined biochemical reactions.

4.2 Vesicle formation from OTG

We initially synthesized **OTG** by standard coupling of 1-thio-β-galactopyranose with oleic acid. HPLC purification was used to isolate the glycolipid, which was characterized by NMR spectroscopy and HPLC-ELSD-MS. Our initial intention was to use this lipid as a detergent, as we expected the amphiphile would form micelles due to the large head group and single alkyl tail. To our surprise, a thin film of the purified **OTG**, when hydrated, readily formed membrane-bound vesicles. **OTG** vesicles were initially identified by phase-contrast microscopy (Figure 4.1b) and fluorescence microscopy using the membrane-staining dye Texas Red® DHPE (Figure 4.1b).

Confirmation that the structures observed were vesicles was shown by transmission electron microscopy (TEM) (Figure 4.1b). The critical aggregation (or vesicle) concentration (cac or cvc) of **OTG** was determined to be 127 μM using a method based on dynamic light-scattering (DLS) measurements (Figure 4.2c).[13] A typical DLS intensity distribution showed the presence of a broad size distribution (approximately 0.1-10 μm) in a typical dispersion of vesicles (Figure 4.2d). The compound was stable to hydrolysis at 37 °C at pH 7 over a period of more than 24 h.

As noted, SCAs used in model protocells present several disadvantages regarding their self-assembly into vesicles.[14–19] For instance, fatty acids assemble into bilayer membranes only when the working pH is near the pK_a of the bilayer-associated acid (pH ~ 7-9).[14,15] Higher and lower pH favor the spontaneous formation of micelles and oil droplets, respectively. Additionally, divalent metal cations present in the hydration media drives the complete destruction of the membrane-bound vesicles by causing fatty acid precipitation.[16–19] These phenomena are due to the carboxylic acid polar head group, which can be protonated/deprotonated and can bind metal ions. In contrast, **OTG** amphiphiles lack an ionizable headgroup and would not be expected to bind metal cations tightly. We confirmed this by testing the effect of pH on the formation of **OTG** vesicles. Hydration of thin films of **OTG** at 37 °C using several aqueous solutions at different pH (2, 5, 7 and 10) showed that glycolipid-based micron-sized vesicles were formed under all these conditions. We also explored the effect of free divalent metal ions on the self-assembly of **OTG** into vesicles. Formation of vesicles at 37 °C was shown for hydrated thin films of **OTG** in aqueous solutions containing different divalent metal cations (1-10 mM Mg^{2+} or Ca^{2+}).

4.3 Encapsulation of fluorescent molecules

The efficient encapsulation ability of the **OTG** vesicles was demonstrated by inclusion of several polar fluorescent dyes such as HPTS and sulfo-Cy5-azide in the hydration media followed

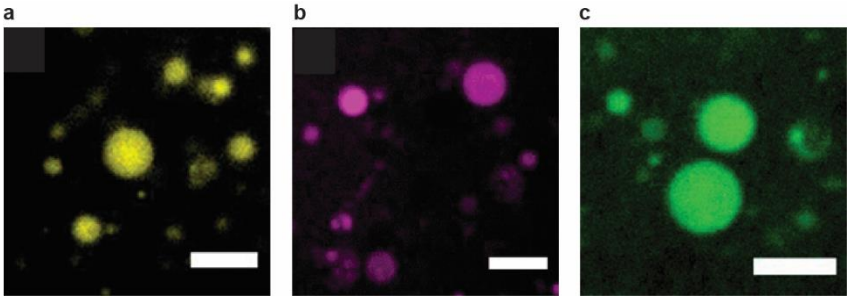


Figure 4.3. Encapsulation of fluorescent molecules in **OTG** vesicles. a. HPTS b. sulfo-Cy5-azide c. GFP. All scale bars: 5 μm.

followed by vesicle characterization using spinning-disk confocal fluorescence microscopy (Figure 4.3a-b). It was also possible to encapsulate GFP (Figure 4.3c). The compatibility of **OTG** vesicle self-assembly with soluble biological molecules lead us to consider the use of these vesicles as compartments for enzymatic reactions.

4.4 Characterization of OTG multilayers using X-ray methods

We have found that, when deposited from solvent on a solid substrate such as silicon, **OTG** spontaneously self-assembles into a multilamellar structure of stacked bilayers. This enabled us to determine the electron density profile (EDP) of the lipid bilayers via X-ray diffraction (XRD) analysis (Figure 2). Relative humidity (RHs) of the samples were controlled using a saturated salt solution. We used different saturated salt solutions, which provided three different RHs (K_2SO_4 ~ 98%, KNO_3 ~ 92% and NaCl ~ 75%),[14] keeping the temperature fixed at 25 °C. Figure 4.4a represents the typical XRD intensity profile of the sample at 75% RH. We observed a distinct set of Bragg peaks in the XRD profile, which indicated long-range ordering in the deposited **OTG** lipid multilayers. To quantify this further, we calculated the relative EDP of the stacked layers from the XRD intensity profiles data. Figure 4.4b shows relative electron density (ρ) profile as a function of the distance *z*. Here the center of each lipid bilayer is considered as the $z = 0$ plane for obtaining a symmetric profile. From this figure, we estimated the distance between the head groups

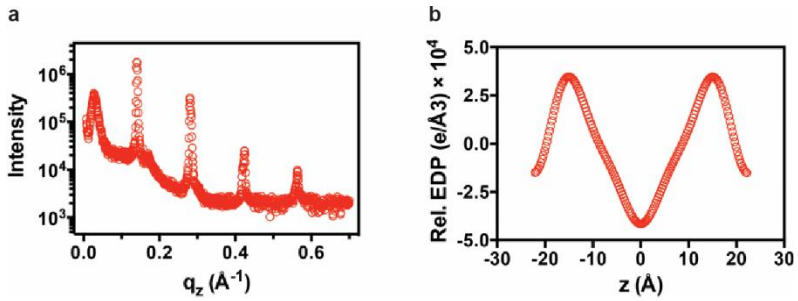


Figure 4.4. Structural characterization of **OTG** multilayer film. a. X-ray diffraction (XRD) intensity profile of an **OTG** multilayer film at 25 °C with 75% relative humidity. b. Relative electron density profile (EDP) of an **OTG** multilayer film at 75% relative humidity and 25 °C.

of the lipid bilayer and used it to calculate the thickness of the interstitial water layer between the head groups. The water layer is on the order of 5 Å at 25 °C and its thickness increases at higher relative humidity, likely due to swelling of the membrane. We also estimated the membrane thickness (in terms of the head-to-head distance of the lipid bilayer) to be 30.02 Å at 25 °C. The EDP also indicated a decreased chain overlap in the central region inside the bilayer.

4.5 Reconstitution of cytochrome c oxidase (CcO) in functional form

Having characterized the **OTG** membranes, we sought to explore if they are capable of mimicking critical features of biological membranes. Transmembrane proteins are essential membrane components for all cells, regulating transport and signaling. So far, they have not been functionally reconstituted into previously analyzed SCA-based membranes. We tested if functional transmembrane proteins could be incorporated into **OTG** membranes during vesicle formation. Cytochrome c oxidase (CcO) was chosen as a test protein because of the existence of well-known methods for assaying its function.[15–18] We initially exchanged the detergent *n*-octyl-β-D-glucopyranoside (OGP) for *n*-dodecyl-β-D-maltoside (DDM) present in commercially available bovine heart CcO preparations by spin filtration. OGP was used as the solubilizing agent for CcO because of its high critical micelle concentration (cmc) and low aggregation number,[19] which facilitates the removal of excess detergent in the subsequent steps. We dissolved a lipid film of **OTG** in excess OGP to obtain a clear solution into which OGP-solubilized CcO was added.

Subsequent treatment with Bio-Beads® SM-2 and overnight dialysis enabled complete removal of the OGP. Successful detergent elimination was confirmed by HPLC-ELSD-MS (Figure 4.5a). It is useful to mention here that direct exchange of DDM with **OTG** using the Biobead method was unsuccessful as these two amphiphiles have very similar cmc's and both were desorbed at nearly equal extent.

We then determined if the CcO reconstituted in the **OTG** membranes (Figure 4.5b) retained its catalytic function. When ferrocytochrome c, the substrate of CcO, was added to the CcO/**OTG**

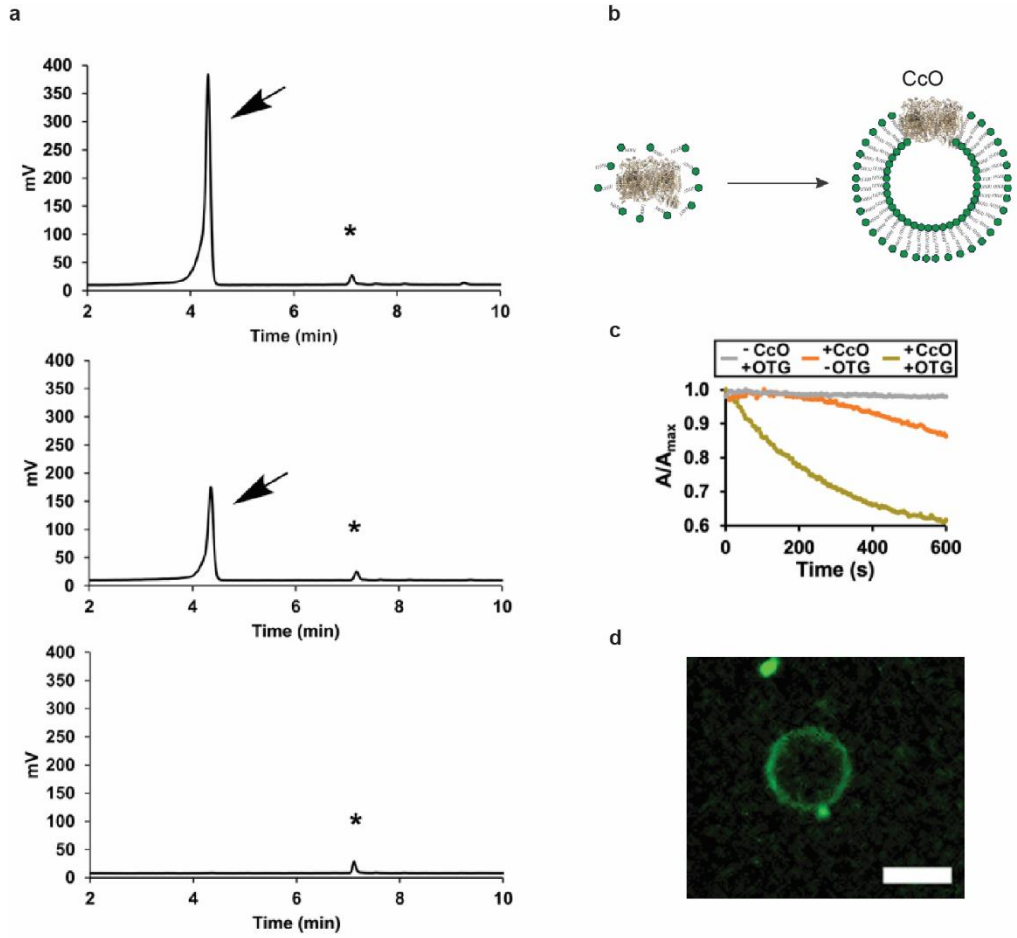


Figure 4.5. Reconstitution of cytochrome c oxidase (CcO) in functional form. a. HPLC-ELSD analysis of detergent removal by Bio-Beads™ SM-2 resin and dialysis. *Arrow* indicates the OGP peak and the *asterisk* indicates the **OTG** peak. b. Schematic representation of spontaneous reconstitution of CcO during vesicle formation. c. Normalized absorbance at 550nm of ferrocytochrome c in **OTG** vesicles with (*gold*) and without (*gray*) embedded CcO. Activity of the isolated CcO (*orange*) is also shown. d. Spinning-disk confocal image of a spontaneously reconstituted CcO (labeled with Alexa Fluor 488)/**OTG** proteoliposome. Scale bar denotes 5 µm.

vesicles, we observed a steady decrease in the absorbance at 550 nm over time (Figure 4.5c), indicating the formation of the oxidized product, ferricytochrome c. A control experiment where ferrocytochrome c was added to **OTG** vesicles lacking CcO, showed negligible change in the absorbance at 550 nm over same time span. A small decrease in absorbance at 550 nm was found in the presence of the isolated CcO (CcO in OGP only). Finally, we were able to visualize fluorescently labeled CcO reconstituted in **OTG** membranes by microscopy (Figure 4.5d).

4.6 Rolling circle amplification (RCA) of DNA inside vesicles

A possible route toward creating advanced artificial cells would be the coupling of simple vesicular systems with amplification and evolution of nucleic acids. Rolling circle amplification (RCA) is an isothermal DNA amplification technique where a polymerase adds nucleotides unidirectionally to short primers annealed to a single-stranded circular DNA template.[20] The robustness of this technique make it an attractive choice for demonstrating nucleic acid replication within vesicles generated by glycolipids (Figure 4.6a). We encapsulated a single-stranded circular M13mp18 DNA template and the RCA machinery [primers, deoxynucleoside triphosphates (dNTPs) and ϕ29 DNA polymerase] into **OTG** vesicles. The resulting vesicles were treated with proteinase K to remove any residual RCA machinery in the surrounding solution and therefore prevent DNA amplification outside the vesicle. At the start of the experiment, addition of EvaGreen®, a dye used to detect double-stranded DNA, into an aliquot of the reaction mixture showed that the level of double-stranded DNA in vesicles was initially below the level of detection (Figure 4.6b). After 21 h of RCA at 29 °C, the addition of the staining dye revealed high fluorescence levels within the vesicles, indicating that DNA amplification occurred (Figure 4.6b). Agarose gel electrophoresis confirmed that DNA had been amplified in the vesicles by RCA (Figure 4.6c). Moreover, we observed by phase-contrast microscopy that the vesicles were stable

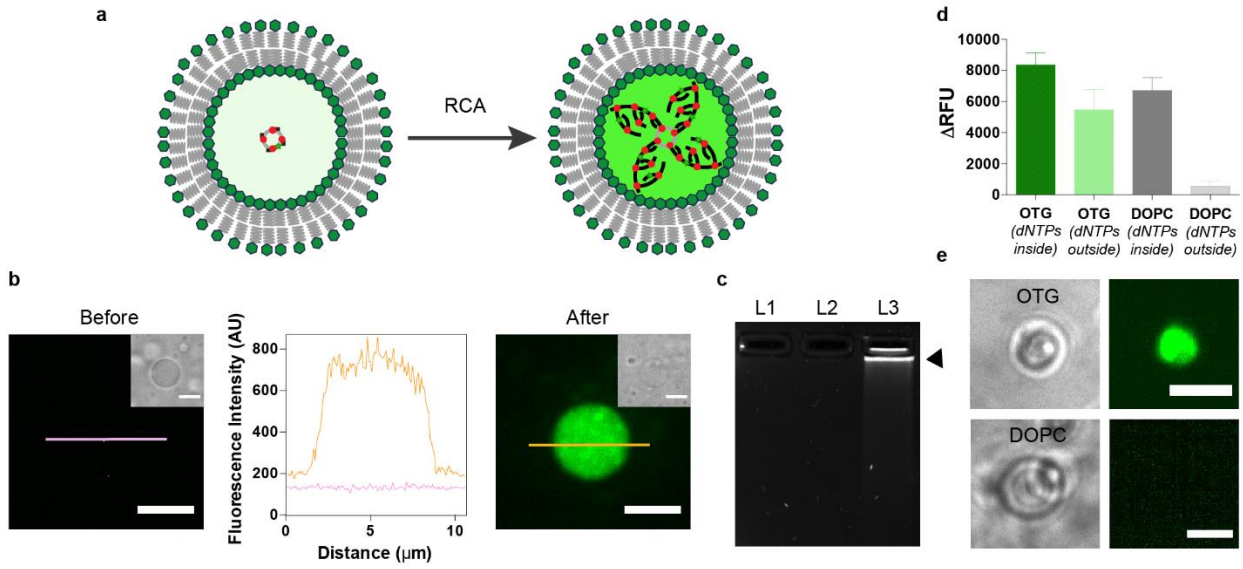


Figure 4.6. Rolling circle amplification of DNA in **OTG** vesicles. a. Schematic representation of the RCA.in **OTG**-based vesicles, showing the efficient amplification of DNA. b. Fluorescence microscopy images of representative vesicles *before* (left) and *after* (right) RCA. Insets show the phase-contrast microscopy images of each step, observing that the vesicles were stable during the RCA process. The *pink* and *orange* lines represent the section used to make the plot profile histogram (center), which shows the fluorescence intensity in a typical DNA-stained vesicle. EvaGreen was used as the DNA staining dye. Scale bars: 5μm. c. 1.0% agarose gel stained with GelRed®, verifying amplification of DNA by RCA in **OTG** vesicles. Lane L1: RCA reaction mixture minus φ29 polymerase encapsulated in **OTG** vesicles (*Control sample*); Lane L2: RCA reaction mixture pre-treated with proteinase K encapsulated in **OTG** vesicles (*Control sample*); Lane L3: RCA reaction mixture encapsulated in **OTG** vesicles followed by proteinase K treatment to digest the unencapsulated φ29 polymerase and other proteins (*Reaction sample*). The arrow-head denotes the amplified DNA band. Prior to loading on gel, all samples were cleaned up by phenol:chloroform:isoamyl alcohol (25:24:1 v/v/v) extraction. d. Increment of fluorescence (in RFUs) resulting from 24 h of RCA in **OTG** and DOPC vesicles adding dNTPs *inside* or *outside*. Error bars indicate standard deviation (SD; $n = 3$). e. Microscopy images of representative **OTG** vesicles and DOPC vesicles in bright-field (*left*) and fluorescence (*right*) channels after 24 h of RCA adding dNTPs outside. The amplification was possible only in the case of using **OTG** vesicles. The images were captured after adding EvaGreen. Scale bars: 5μm.

under the RCA conditions and **OTG** was hydrolytically stable (determined by HPLC-MS) over the times scales relevant to the corresponding experiments (Figure 4.6b).

Previous seminal studies have established that SCAs can have increased permeability to charged small molecules such as nucleotides when compared to two-chain lipids such as phospholipids.[8,21] Such permeability can facilitate internal biochemical reactions within synthetic cells, such as oligonucleotide polymerization, even when precursors are supplied externally. Thus, we were interested if **OTG** vesicles were more permeable to externally added NTPs compared to phospholipid vesicles. Additional experiments showed that **OTG** vesicles have significantly

higher permeability to dNTPs compared to that for two-chain phospholipid vesicles (Figure 4.6d-e). For instance, we observed that the **OTG** vesicles could sustain RCA reaction when the dNTPs were solely supplied from the outside, while 1,2-dioleoyl-*sn*-3-phosphocholine (DOPC) vesicles could not under identical conditions (Figure 4.6d-e). Fluorescence imaging indicated that **OTG** vesicles retained newly formed oligonucleotides. Therefore, **OTG** vesicles could be utilized as a semi-permeable artificial cell model where transport of charged solutes can take place without the necessity of special transporters or pore-forming proteins.

4.7 Experimental methods

4.7. 1 Synthesis of *S*-Oleoyl β-D-1-thiogalactopyranose (OTG)

A solution of oleic acid (25.9 mg, 91.6 μmol) in DMF (916 μL) was stirred at 0 ºC for 10 min, and then HATU (38.3 mg, 100.8 μmol) and DIEA (17.6 μL, 100.8 μmol) were successively added. After 10 min stirring at 0 ºC, 1-thio-β-galactopyranose sodium salt (20.0 mg, 91.6 μmol) was added. After 1 h stirring at rt, the mixture was concentrated under reduced pressure. The corresponding residue was dissolved in MeOH (500 μL), filtered using a 0.2 μm syringe-driven filter, and the crude solution was purified by HPLC, affording 39.8 mg of **OTG** as a white solid (94%). ^{1}H NMR ($CDCl_3$, 500.13 MHz, δ): 5.45-5.23 (m, 2H, 2×CH), 5.08 (d, *J* = 9.8 Hz, 1H, 1×CH), 4.38-3.96 (m, 3H, 3×OH), 3.94-3.61 (m, 6H, 4×CH+1×CH_2) 3.42-3.09 (br s, 1H, 1×OH), 2.63 (t, *J* = 7.6 Hz, 2H, 1×CH_2), 2.15-1.81 (m, 6H, 3×CH_2), 1.78-1.59 (m, 2H, 1×CH_2), 1.41-1.13 (m, 18H, 9×CH_2), 0.88 (t, *J* = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR ($CDCl_3$, 125.77 MHz, δ): 197.4, 130.2, 129.8, 82.6, 78.3, 75.1, 69.7, 69.5, 62.1, 44.7, 32.1, 29.9, 29.9, 29.7, 29.5, 29.5, 29.4, 29.3, 29.2, 27.4, 27.3, 25.3, 22.9, 14.3. MS (ESI-TOF) [m/z (%)]: 483 ($[M+Na]^+$, 100), 461 ($[MH]^+$, 20]. HRMS (ESI-TOF) calculated for [$C_{24}H_{44}O_6SNa$] ($[M+Na]^+$) 483.2751, found 483.2748.

4.7.2 GUV preparation by electroformation method

10.0 µL of a 10 mM solution of **OTG** in MeOH:$CHCl_3$ (1:1) were placed on the conductive side of an indium tin oxide (ITO) slide. The slide was then dried in vacuo for 1 h, after which a rubber O-ring was placed around the dried lipid film and filled with 260.0 µL of 150 mM sucrose solution. A second ITO slide was placed onto of the first and everything assembled within a Vesicle Prep Pro® device. The giant unilamellar vesicles were electroformed with the following parameters: Frequency 5 Hz, Amplitude 3 V, Temperature 55 ºC, Rise time 10 min, Main time 120 min, Fall time 10 min. Afterward, 5 µL of the vesicle solution were mixed with 0.1 µL of a 100 µM solution of Texas Red® DHPE in EtOH, placed on a clean glass slide and imaged.

4.7.3 Estimation of critical vesicle concentration (cvc)

100.0 µL of aqueous solutions (from 1 mM to 2.5 µM) of **OTG** were analyzed by Dynamic Light Scattering (DLS) in order to determine the critical vesicle concentration (cvc's) according to previously described methods.[3] This technique allowed us to measure the cvc at 25 °C, since the scattered light intensity (measured at 90°) is dependent of molecular weight and size of the particle and, hence, it increases when monomers start to aggregate in solution. Light-scattering intensity was very low and constant when the concentration was less than 50 µM, denoting the presence of monomers in solution. After sequential concentration increase, light-scattering intensity increased, and the cvc was determined to be 127 µM by intersecting the two straight lines corresponding to monomer and vesicle regions.

4.7.4 X-ray experiments

OTG lipid multilayers were deposited from a 8.0 mg/ml of **OTG** in $CHCl_3$/TFE (1:1) on freshly cleaned [MeOH/deionized (DI) H_2O] hydrophilic silicon [100] substrate. The deposited dried films were then put in high vacuum for more than 12 h to remove trapped solvents. A slow

and steady humidity incubation for more than 48 h at 45 ºC and 100% RH produced a well-oriented smectic lipid multilayer membrane with long range correlation among consecutive layers. Subsequently, samples were stored in sealed boxes at room temperature and under relative humidity of ~ 98 % using saturated salt solution vapor. X-ray diffraction (XRD) measurements were carried out at our in-house X-ray diffraction setup consisting of a four circle goniometer and a Cu Kα (8.04 keV, λ ~ 1.54 Å) source. A special sealed humidity chamber was used to control precisely the relative humidity at the sample. Using three different saturated salt solutions [K_2SO_4, KNO_3 and NaCl], the relative humidity was kept constant (at 98%, 85% and 75%, respectively) over the period of the experiment. For the representation of the typical XRD intensity profile of the OTG sample, we have plotted integrated intensity as a function of transferred wave vector (q_z),

$$q_z = \frac{4\pi}{\lambda} \sin\theta \tag{4.1}$$

where θ is the angle of incidence and λ stands for the wavelength of the incident X-ray. To quantify further, we have calculated the relative electron density profile (EDP) of the stacked layers from the XRD intensity profiles data. Ideally, one should convert the relative to an absolute EDP using some scaling instrumental parameters. These constants can be achieved by comparing the experimental EDP with standard known EDP at absolute scale, using the following equation

$$\rho_{relative}(z) = \frac{2}{d} \sum_{n=-\infty}^{\infty} \nu_n \, |F_n| \cos\left(\frac{2\pi n z}{d}\right) \tag{4.2}$$

where d represents the periodicity of the stacked multilayer, n is the Bragg peak order number, ν_n is the phase factor, and $F_n = n\sqrt{I_n}$ where I_n is the integrated intensity of the n-th order peak and the factor n include the Lorentz correction factor apart from a multiplicative constant. The phase's v_n were determined by measuring XRD peaks under three different hydrations and utilizing the swelling method.

4.7.5 CcO reconstitution

Reduction of ferricytochrome c. 1.84 mg of reddish brown powder of ferricytochrome c were dissolved in 190 µL PBS (1X). Then, 10 µL of sodium ascorbate (200 mM) were added and kept at 4 ^{o}C overnight, observing that the color changed to rose-red. The solution was washed with PBS (20X volume) using a 10 kDa MWCO spin filter to remove residual ascorbate. The final concentration of ferrocytochrome c was estimated as 0.75 mM and the A_{550}/A_{565} ratio was 14.98.

Fluorescent labeling of CcO. 60 µg of CcO (bovine heart) were exchanged with $NaHCO_3$ buffer (100 mM, pH 8.3) containing 0.4 mM *n*-dodecyl-*β*-maltoside (DDM). Then, the solution was treated with 10 equivalents of Alexa Fluor® 488 NHS ester (dissolved in anhydrous DMSO), and the mixture was stirred at room temperature for 2 h. Afterwards, the excess dye was removed, and DDM was exchanged with *n*-octyl-*β*-D-glucopyranoside (OGP) by spin filtration.

Activity assay of CcO. 0.3 mg of **OTG** lipid film deposited on the walls of a glass vial were hydrated with a solution (650 µL) containing sodium phosphate buffer (5 mM, pH 6.5) and OGP (30 mM) such that the final concentration of **OTG** was ~1 mM. Then, ~25 µg CcO (10 µL in 30 mM OGP) were added and allowed to sit for 10 min. Afterward, the clear solution was added to 34.3 mg of pre-washed Bio-Beads™ SM-2 and tumbled at room temperature. Within 30 min, turbidity was detected, and the dispersion was transferred to a dialysis chamber (20 kDa MWCO). Dialysis was carried out against 45 mL of sodium phosphate (5 mM, pH 6.5) with two buffer changes. The resulting vesicle dispersion was sonicated for 15 min. Then, 400 µL of the vesicle dispersion were taken in a UV-Vis cuvette and 10 µL of ferrocytochrome c (0.75 mM) were added and rapidly mixed by tapping. The absorbance at 550 nm was monitored spectrophotometrically at every 5 s interval using a Nanodrop 2000c spectrophotometer.

4.7.6 Rolling circle amplification (RCA) of DNA

Reaction conditions and microscopic observations. 60 µL of a 5 mM solution of **OTG** (in 1:1 MeOH:$CHCl_3$) were evaporated under a steady stream of Ar gas to produce a thin lipid film. The film was hydrated using a solution containing φ29 polymerase buffer [Final composition: 50 mM HEPES pH 7.5, 10 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 1 mM TCEP.HCl], dNTPs (Thermo Scientific, 5 mM each), Exo-resistant random primers (Thermo Scientific, final concentration: 29.2 µM), thermostable inorganic pyrophosphatase (New England Biolabs, 3.5 U), bovine serum albumin (New England Biolabs, 2.5 µg), M13mp18 single stranded DNA (New England Biolabs, 0.25 ng), NxGen® φ29 polymerase (Lucigen, 50 U), and nuclease free H_2O such that the final volume is 30 µL. Then, 1 µg of proteinase K (New England Biolabs) was added to digest any un-encapsulated protein. The resulting mixture corresponds to the *reaction sample*. In one of the *control* samples, the φ29 polymerase was substituted with its storage buffer as described by the manufacturer. In another *control* sample, the RCA reaction mixture was treated with 1 µg of proteinase K for 30 min at room temperature prior to hydration. All samples (reaction and controls) were placed at 29 ^oC. The samples were checked by confocal microscopy at various time points by taking out 4 µL and adding 0.2 µL EvaGreen® (20X).

Agarose Gel Electrophoresis. After 24 h of incubation at 29 ^oC, equal volume of phenol:chloroform:isoamyl alcohol (25:24:1 v/v/v) was added into the vesicle dispersion and mixed well by vigorous vortexing. The aqueous layer was collected and loaded onto 1% agarose gel using Purple Gel Loading Dye (New England Biolabs). The gel was run at 115 V for 60 min and the bands were then visualized on Gel Doc™ (Bio-Rad) based on GelRed® dye.

Comparison of RCA in OTG vesicles versus DOPC vesicles. 60 µL of a 5 mM solution of **OTG** (in 1:1 MeOH:$CHCl_3$) were evaporated under a steady stream of Ar gas to produce a thin

lipid film. The film was hydrated using a solution containing φ29 polymerase buffer [Final composition: 50 mM HEPES pH 7.5, 10 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 1 mM TCEP.HCl], Exo-resistant random primers (Thermo Scientific, final concentration: 29.2 μM), thermostable inorganic pyrophosphatase (New England Biolabs, 3.5 U), bovine serum albumin (New England Biolabs, 2.5 μg), M13mp18 single stranded DNA (New England Biolabs, 10 ng), NxGen® φ29 polymerase (Lucigen, 30 U), and nuclease free H_2O such that the final volume is 30 μL. Then, 2.5 μg of proteinase K (New England Biolabs) was added and incubated for 30 min to digest any un-encapsulated protein. After this, 1.5 μL of dNTP mix (5 mM each) is added to the vesicle dispersion and the reaction is incubated at 31 ºC for 24 h. The reaction corresponding to DOPC vesicles was prepared in an identical way except that the molar concentration of DOPC was half of that of OTG. 5 μL of each sample were diluted with 14 μL H_2O and 1 μL of Eva Green dye (5X) and the fluorescence were measured on a plate reader (λ_{ex}: 485 nm, λ_{em}: 530 nm).

After the 24 h incubation, equal volume of phenol:chloroform:isoamyl alcohol (25:24:1 v/v/v) was added into the vesicle dispersions and mixed well by vigorous vortexing. The aqueous layer was collected and loaded onto 0.7% agarose gel. The gel was run at 110 V for 40 min and the bands were then visualized on Gel Doc™ based on GelRed® dye.

Notes on chapter

Chapter 4 is adapted almost entirely from the materials published in Brea, R. J., Bhattacharya, A., Bhattacharya, R., Song, J., Sinha, S. K. & Devaraj, N. K. Highly Stable Artificial Cells from Galactopyranose-derived Single Chain Amphiphiles. *J. Am. Chem. Soc.*, 2018, 140, 17356-17360. The dissertation author was a co-first author on this manuscript.

4.8 References

1. Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. Experimental Models of Primitive Cellular Compartments: Encapsulation, Growth, and Division. *Science* **302,** 618–622

(2003).

2. Rendón, A., Carton, D. G.; Sot, J., Garcia-Pacios, M., Montes, L.R., Valle, M., Arrondo, J. L., Goni, F. M. & Ruiz-Mirazo, K. Model systems of precursor cellular membranes: Long-chain alcohols stabilize spontaneously formed oleic acid vesicles. *Biophys. J.* **102,** 278–286 (2012).

3. Walde, P., Wessicken, M., Rädler, U., Berclaz, N., Conde-Frieboes, K. & Luisi, P. L. Preparation and characterization of vesicles from mono-n-alkyl phosphates and phosphonates. *J. Phys. Chem. B* **101,** 7390–7397 (1997).

4. Du, N., Song, R., Li, H., Song, S., Zhang, R. & Hou, W. A Nonconventional Model of Protocell-like Vesicles: Anionic Clay Surface-Mediated Formation from a Single-Tailed Amphiphile. *Langmuir* **31,** 12579–12586 (2015).

5. Mansy, S. S. Model protocells from single-chain lipids. *Int. J. Mol. Sci.* **10,** 835–843 (2009).

6. Mansy, S. S. Membrane transport in primitive cells. *Cold Spring Harbor Perspect. Biol*. **2**, a002188 (2010).

7. Hargreaves, W. R. & Deamer, D. W. Liposomes from ionic, single-chain amphiphiles. *Biochemistry* **17,** 3759–3768 (1978).

8. Chen, I. A., Salehi-Ashtiani, K. & Szostak, J. W. RNA catalysis in model protocell vesicles. *J. Am. Chem. Soc.* **127,** 13213–13219 (2005).

9. Jin, L., Engelhart, A. E., Adamala, K. P. & Szostak, J. W. Preparation, purification, and use of fatty acid-containing liposomes. *J. Vis. Exp.* **2018,** e57324 (2018).

10. Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342,** 1098–1100 (2013).

11. Mizushina, Y., Yoshida, S., Matsukage, A. & Sakaguchi, K. The inhibitory action of fatty acids on DNA polymerase *β*. *Biochim. Biophys. Acta, Gen. Subj*. **1336**, 509-521 (1997).

12. Mansy, S. S. & Szostak, J. W. Thermostability of model protocell membranes. *Proc. Natl. Acad. Sci.* **105,** 13351–13355 (2008).

13. Budin, I., Prwyes, N., Zhang, N. & Szostak, J. W. Chain-length heterogeneity allows for the assembly of fatty acid vesicles in dilute solutions. *Biophys. J.* **107,** 1582–1590 (2014).

14. Ma, Y., Ghosh, S. K., Bera, S., Jiang, Z., Tristram-Nagle, S., Lurio, L. B. & Sinha, S. K. Cholesterol Partition and Condensing Effect in Phase-Separated Ternary Mixture Lipid Multilayers. *Biophys. J.* **110,** 1355–1366 (2016).

15. Cole, C. M., Brea, R. J., Kim, Y. H., Hardy, M. D., Yang, J. & Devaraj, N. K. Spontaneous

Reconstitution of Functional Transmembrane Proteins During Bioorthogonal Phospholipid Membrane Synthesis. *Angew. Chem. Int. Ed.* **54**, 12738-12742 (2015).

16. Vik, S. B. & Capaldi, R. A. Lipid Requirements for Cytochrome c Oxidase Activity. *Biochemistry* **16,** 5755–5759 (1977).

17. Yoshikawa, S., Muramoto, K. & Shinzawa-Itoh, K. The O_2 reduction and proton pumping gate mechanism of bovine heart cytochrome c oxidase. *Biochim. Biophys. Acta - Bioenerg.* **1807,** 1279–1286 (2011).

18. Yoshikawa, S., Muramoto, K., Shinzawa-Itoh, K. & Mochizuki, M. Structural studies on bovine heart cytochrome c oxidase. *Biochim. Biophys. Acta - Bioenerg.* **1817,** 579–589 (2012).

19. Stubbs, G. W. & Litman, B. J. Effect of Alterations in the Amphipathic Microenvironment on the Conformational Stability of Bovine Opsin. 1. Mechanism of Solubilization of Disk Membranes by the Nonionic Detergent, Octyl Glucoside. *Biochemistry* **17,** 215–219 (1978).

20. Ali, M. M., Li, F., Zhang, Z., Zhang, K., Kang, D.-K., Ankrum, J. A., Le, X. C. & Zhao, W. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem. Soc. Rev.* **43,** 3324–3341 (2014).

21. Mansy, S. S., Schrum, J. P., Krishnamurthy, M., Tobe, S., Treco, D. A. & Szostak, J. W. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454**, 122–125 (2008).

Chapter 5 | Self-Assembly and Thermodynamic Properties of Single-Chain β-D-Glycopyranosylamides of Miscellaneous Fatty Acids

5.1 Introduction

Amphiphilic molecules self-assemble in aqueous medium to form supramolecular aggregates belonging to various lyotropic mesophases. Molecular geometry and functional groups on an amphiphile are the primary determinants of what kind of aggregate it will form. Some of the mesophases well documented in the literature are micellar, lamellar, cubic, hexagonal, and inverse micellar. In this chapter we will discuss how subtle differences in molecular structures of single-chain glycolipid amphiphiles have dramatic manifestations on the self-assembled aggregates.

In nature, glycolipids constitute an important class of amphiphiles which exhibit a wide variety of structural polymorphism depending on the molecular architecture.[1,2] These amphiphiles are characterized by a monosaccharide or oligosaccharide head group and a hydrophobic tail often originating from condensation of fatty acids. Glycolipids play important roles in maintaining the stability of cellular membranes, as well as serving as markers for cellular recognition. Despite their ubiquity, the structure-function relationship of glycolipids with respect to membrane formation is relatively less explored and requires detailed biophysical studies. In the previous chapter, we discussed the formation of vesicles from the single-chain amphiphile *S*-oleoyl β-D-thiogalactopyranose (**OTG**) in water. Motivated by the unexpected discovery, we decided to study other amphiphiles having similar molecular design. We synthesized a series of glycopyranosylamides of a variety of fatty acids and characterized their self-assembly and thermodynamic properties using a host of physical techniques. In particular, we carried out the study of (i) galactopyranosylamides and (ii) glucopyranosylamides of unsaturated fatty acids in most extensive detail (Figure 5.1).

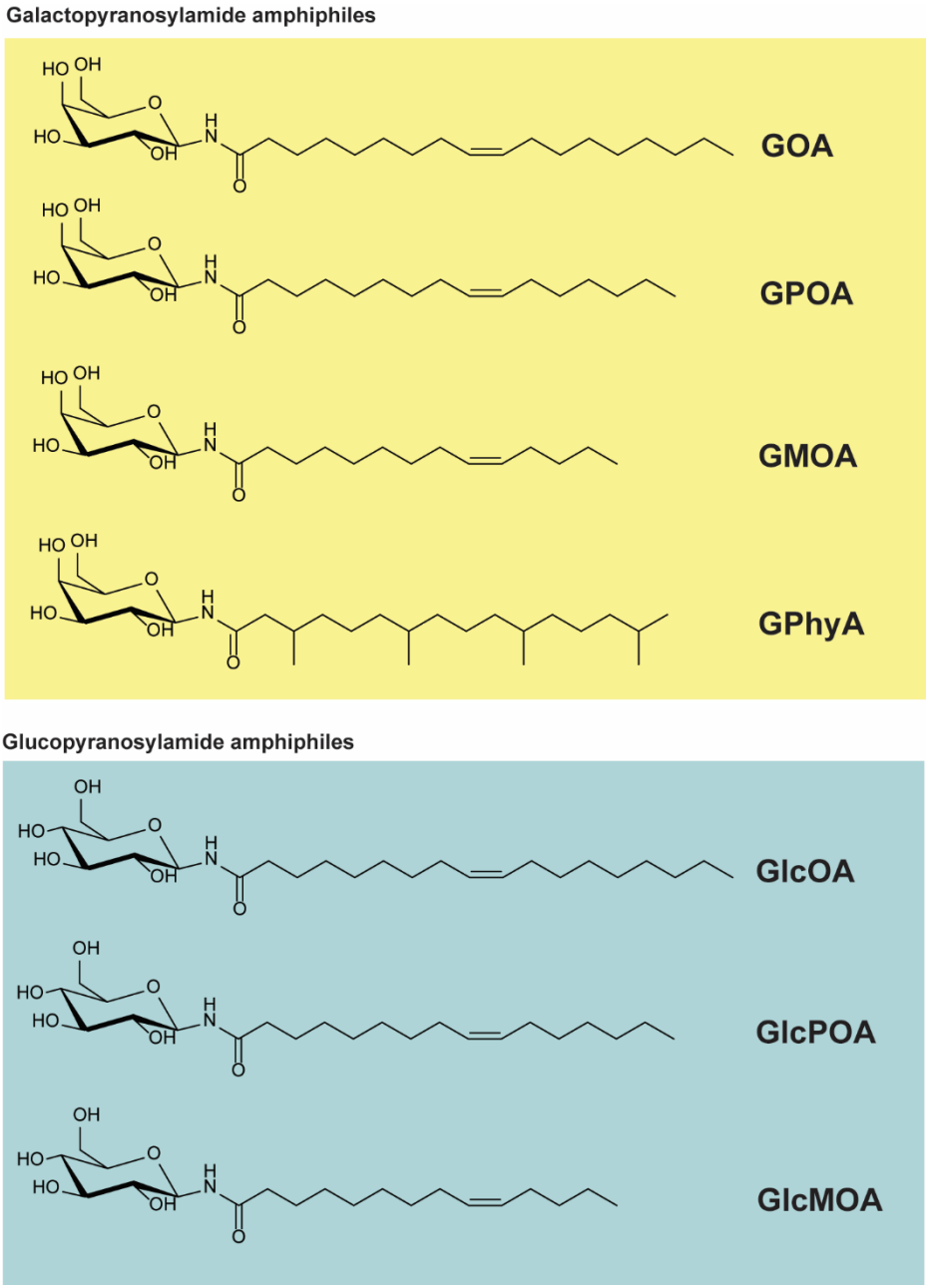


Figure 5.1. Chemical structures of the galactopyranosylamides and glucopyranosylamides most commonly used in this study.

5.2 Self-assembly behavior of β-D-glycopyranosylamide amphiphiles

Similar to the observation made in the previous chapter, we observed the formation of vesicles from *N*-oleoyl β-D-galactopyranosylamine (**GOA**) in water. Then we asked if galactopyranose amphiphiles with shorter unsaturated hydrophobic tails can also form membrane

bound vesicles. To address this question, we condensed the naturally occurring monounsaturated fatty acids palmitoleic acid (16:1, ω7) and myristoleic acid (14:1, ω5) with β-D-galactopyranosylamine and obtained the corresponding amides (**GPOA** and **GMOA**, respectively). Interestingly, we observed that both **GPOA** and **GMOA** self-assemble to form vesicles in water. Using a method based on the fluorescence properties of the solvatochromic dye Laurdan, we estimated the critical aggregation concentrations (cac) for all three amphiphiles and obtained the values 0.079 mM, 0.173 mM and 0.972 mM for **GOA**, **GPOA** and **GMOA**, respectively (Figure 5.2, Table 5.1). The trend is consistent with a previous study[3] on the dependence of cac on alkyl chain length. We visualized the formation of micron-sized vesicles stained with Texas Red DHPE from each of these amphiphiles by optical microscopy (Figure 5.2). We also synthesized the galactopyranosylamides with saturated fatty acid chains (palmitic, stearic), however, because of their very high melting points, it was not convenient to study them. Interestingly, we observed that the corresponding amphiphile (**GPhyA**) derived from the branched chain phytanic acid formed vesicles in water. We further characterized the vesicular structures formed from **GOA** by negative staining TEM and observed spherical compartments of the approximate size range 0.1-1 μm (Figure 5.3a).

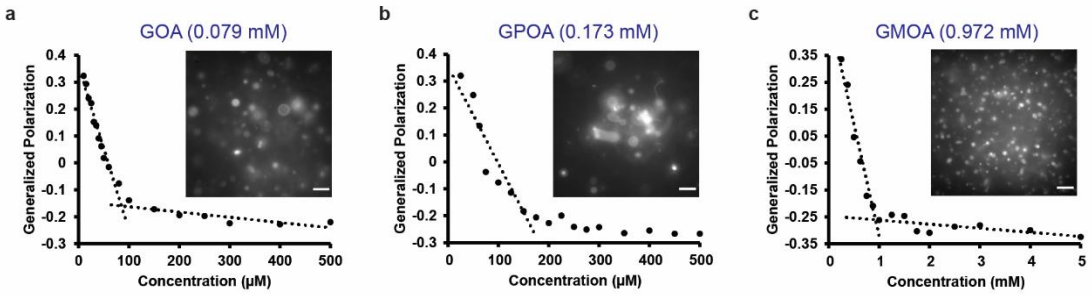


Figure 5.2. Critical aggregation concentration (cac) of β-D-galactopyranosylamides. Representative plots of cac determination using a method based on Laurdan fluorescence are shown for a. **GOA** b. **GPOA** c. **GMOA**. Inset images show corresponding vesicles membranes stained with 0.1 mol% Texas Red-DHPE. All scale bars: 10 μm.

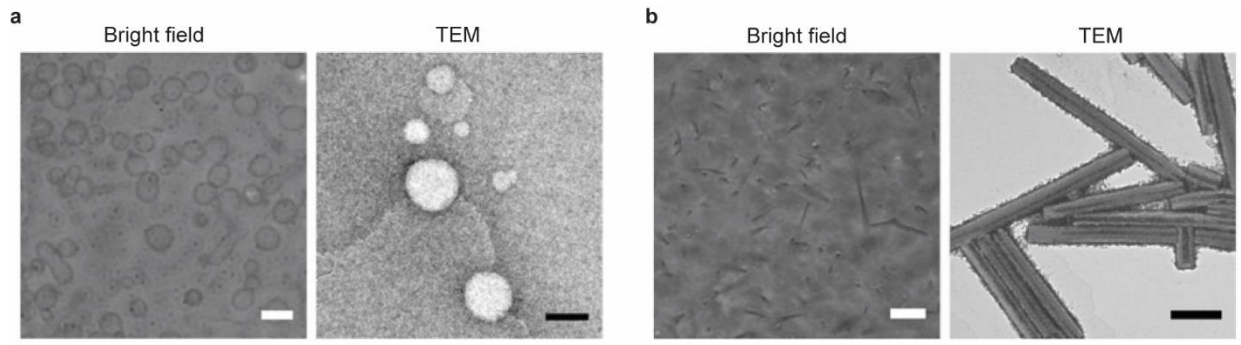


Figure 5.3. Comparison of self-assembled structures formed by **GOA** and **GlcOA**. Comparison of bright field (phase contrast) and TEM (negative staining) images of a. **GOA** vesicles (scale bars: 5 μm and 100 nm respectively) b. **GlcOA** nanotubes (scale bars: 5 μm and 1 μm respectively).

Next, we carried out the synthesis of *N*-oleoyl β-D-glucopyranosylamine (**GlcOA**). It was interesting to observe that **GlcOA** exclusively formed micron-sized nanotubes in aqueous medium (Figure 5.3b). TEM corroborated the nanotubular architecture for the corresponding self-assembled glycolipid structures (Figure 5.3b). The nanotubes possess well-defined hollow cylinders with inner diameters of ~70 nm and outer diameters of ~220 nm. Our observations are consistent with the previous report by Shimizu and coworkers.[4] Furthermore, we observed that the palmitoleoyl and myristoleoyl derivatives (**GlcPOA** and **GlcMOA**, respectively) also self assembles to form micron-sized nanotubes in water. It is interesting to note that how variation in the monosaccharide head group can have drastic manifestations on the supramolecular assemblies of amphiphiles and opens up the possibilities for further exploration.

5.3 X-ray studies of multilayer stacks on a solid substrate

The X-ray reflectivity (XRR) intensity profiles with respect to q_z for **GOA**, **GMOA**, **GPOA** and **GlcOA** are presented in Figure 5.4. The presence of equidistant Bragg peaks indicates the formation of multilamellar structures by each of these amphiphiles. Further, to obtain the molecular picture of these assemblies, as described in the Experimental Methods section, electron density profiles (EDPs) were constructed from intensities of the Bragg reflections using inverse Fourier transformation. Due to the symmetric nature of the bilayer leaflet, the $z = 0$ indicates the

horizontal plane at the center of the bilayer. From the EDPs of different amphiphiles, we estimated how the molecules align along the bilayer normal axis. We summarized the *d*-spacing, and head-to-head distances of the lipids in Table 5.1.

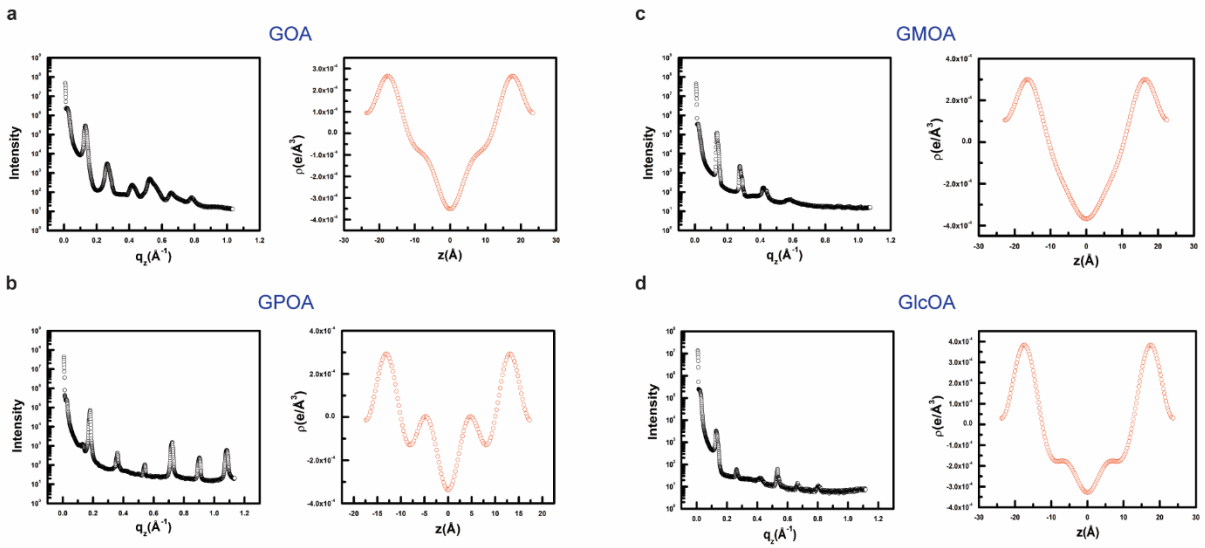


Figure 5.4. Characterization of the glycolipid amphiphiles by X-ray reflectivity (XRR). Multilayer films of a. **GOA**, b. **GPOA**, c. **GMOA**, and d. **GlcOA** are analyzed at room temperature (~22 ºC) and 98% relative humidity (RH). The XRD intensity profiles are displayed on the *left* and the relative electron density profiles (EDP) on the *right*.

The EDPs indicate that the folding of the alkyl chains varies across the four samples. In **GOA**, we observed a shoulder at 7.36 Å from the equatorial plane (Figure 5.4a). In case of **GPOA**, the effect of chain folding was observed to be more pronounced, giving rise to a distinct shoulder peak at 4.67 Å from the equatorial plane (Figure 5.4b). In contrast, we did not observe any shoulder for **GMOA** (Figure 5.4c). This is likely due to least splay in its alkyl chain, resulting in a nearly straight configuration of the latter. In qualitative terms, the EDP of **GlcOA** multilayers was observed to be similar to that of **GOA** suggesting that in a multilayer structure, the arrangement of the oleoyl chains is similar in these two lipids (Figure 5.4d). The observed differences (vesicles vs. nanorods) in self-assembly properties of **GOA** and **GlcOA** in aqueous medium are possibly stemming from the nature of intermolecular hydrogen bonding interactions between the head-groups.

Table 5.1. Summary of the physical parameters of the various glycolipid amphiphiles described in this work.

Parameter	GOA	GPOA	GMOA	GlcOA
cac[a]	0.079 mM	0.173 mM	0.972 mM	--
d-spacing (XRR)	48.30 Å	34.88 Å	46.86 Å	47.53 Å
d-spacing (SAXS)				
50 ºC	45.86 Å	40.75 Å[a]	36.11 Å[b]	--
30 ºC	45.86 Å, 37.40 Å	34.40 Å	36.21 Å[b]	46.76 Å[c]
4 ºC	--	34.60 Å	35.97 Å[b]	--
head-to-head distance	35.25 Å	26.16 Å	32.64 Å	35.00 Å
shoulder in the alkyl region	7.36 Å	4.67 Å	--	6.77 Å
T_m (DSC)	37 ºC	34 ºC	30 ºC	60 ºC

[a] cac values are reported as the mean obtained from three independent measurements
[b] values calculated from fitting curve to Gaussian distribution
[c] SAXS performed at room temperature

5.4 Small-angle X-ray scattering (SAXS) studies on glycolipid dispersions

We carried out SAXS studies on aqueous dispersions of the glycolipid amphiphiles using synchrotron radiation source. The results of the experiments are summarized in Table 5.1 and some of the representative plots are given in Figure 5.5. In case of **GOA** and **GPOA**, transitions between different lamellar phases could be observed as evident from the shifts in the peak positions. Interestingly, in case of **GMOA**, broad peaks with nearly identical maxima were obtained at all temperatures, suggesting a highly disordered molecular packing. It may be possible that, transitions between fluid and gel phases is slow in case of **GMOA** and we could not observe any significant difference over the time scale (~1 h) of the experiment. We discuss in more detail the phase behavior or **GOA** at temperatures below room temperature in more detail in section **5.9**. In addition, we also carried out SAXS studies on dispersion of **GlcOA** (3 mM) tubes at room temperature and observed that the monomers are packed with a lamellar periodicity of 46.76 Å.

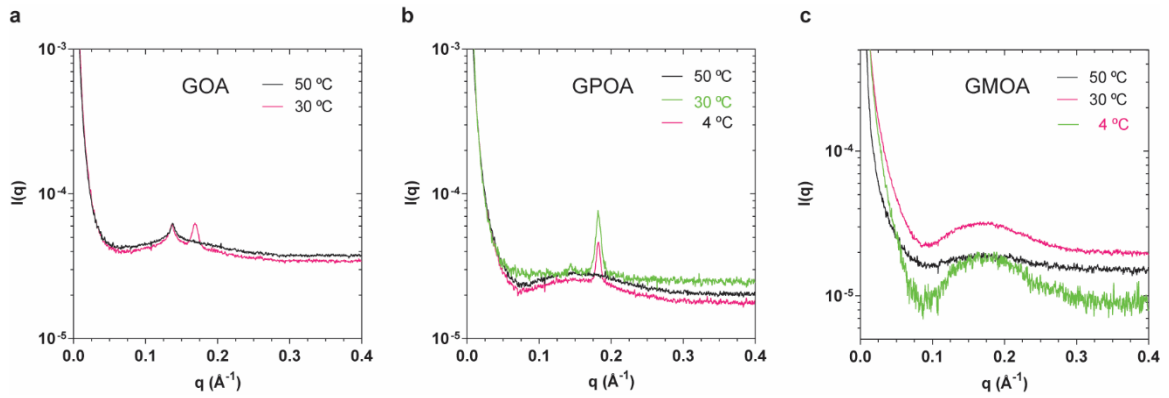


Figure 5.5. SAXS studies on multilamellar dispersions of miscellaneous galactopyranosylamide amphiphiles at varying temperatures. Representative SAXS intensity profiles for 4 mM dispersions of a. **GOA** (50 °C, 30 °C) b. **GPOA** (50 °C, 30 °C, 4 °C) c. **GMOA** (50 °C, 30 °C, 4 °C) in water.

5.5 Phase transition behavior of glycolipid amphiphiles

Transition between various lyotropic phases are defining thermodynamic events for all amphiphiles. We sought to determine the behavior of the amphiphiles used in our study as a function of temperature. We analyzed aqueous dispersions of different single-chain glycolipids (**GOA**, **GPOA**, **GMOA, GlcOA**, **GlcPOA**, and **GlcMOA**) using differential scanning calorimetry (DSC) over a broad range of temperature (1-90 ºC) and obtained the corresponding thermograms (Figure 5.6). All amphiphiles exhibited an asymmetric endothermic main transition peak the maxima of which can be interpreted as the chain melting temperature (T_m) (Figure 5.6, Table 5.1). In addition, **GOA** exhibited a sharp, relatively narrow, and intense endothermic peak around 25 ºC, preceding the peak corresponding to main transition at 37 ºC (Figure 5.6). We propose that the aforementioned peak corresponds to a pre-transition event, likely arising from a rearrangement from a subgel phase to a bilayer gel phase as it had been observed in case of several bilayer membrane forming lipids.[5–7] We investigate in further detail on the origin of this peak in Section **5.9**. Overall, the β-D-glucopyranosylamides exhibited significantly higher transition temperatures as compared to the corresponding β-D-galactopyranosylamides (Figure 5.6). These differences likely arise from the more extensive intermolecular hydrogen bonding between the head groups of

glucose units in the former class. Interestingly, the T_m values among the β-D-galactopyranosylamides varies as **GOA** (37 ºC)>**GPOA** (34 ºC)>**GMOA** (30 ºC), whereas among β-D-glucopyranosylamides, the trend is reverse as **GlcOA** (60 ºC)<**GlcPOA** (61 ºC)<**GlcMOA**

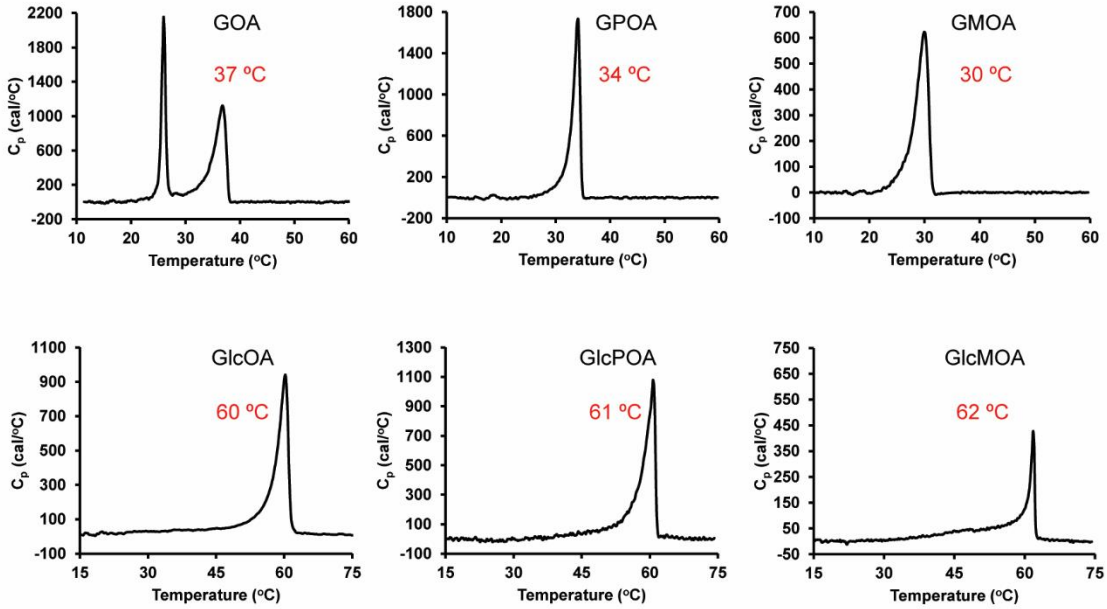


Figure 5.6. Characterization of the phase transition behavior by differential scanning calorimetry (DSC). DSC thermograms of aqueous dispersions of **GOA** (0.5 mM), **GPOA** (1 mM), **GMOA** (2 mM), **GlcOA** (0.5 mM), **GlcPOA** (0.5 mM), and **GlcMOA** (1.5 mM) at scan rate 30 ºC/h.

(62 ºC). These trends suggest that, the combined effect of the alkyl chain length and position of the *cis*-double bond have different manifestations on the phase transition behavior in two kinds of self-assembled structures (vesicles and tubes), likely because of the way the fatty acid chains are packed.

5.6 Effects of Tween 80 on GOA membranes

We next sought to determine the effects of non-ionic detergents on the phase behavior and fluidity of **GOA** membranes. Various non-ionic detergents are used in liposomal preparations for drug delivery to tune the fluidity and permeability properties of the membranes.[8] In particular, we studied the effects of the commercially available non-ionic surfactant Tween 80 on **GOA** membranes using DSC (Figure 5.7a). Tween 80 is typically used in liposomal drug delivery

systems as an "edge activator" or fluidity enhancer.[9–11] We obtained DSC thermograms of mixtures of **GOA** and Tween 80 at various molar ratios (0-50 mol% of Tween 80). The main transition peak broadened with increasing concentration of Tween 80 and at 35 mol% it disappeared. Interestingly, the lower temperature peak decreased in intensity but retained its position and sharp feature until it completely disappeared at 35 mol% Tween 80. We observed that Tween 80 significantly facilitated the hydration of **GOA** films into vesicles and improved its solubility in buffered solutions even when present at very low concentrations (2.5 mol%). This is likely because Tween 80 attenuates intermolecular hydrogen bonding between the galactose head groups and alters the order of water structure near the membrane surface. To support the data obtained from DSC, we observed the **GOA** vesicles formed with varying percentages of Tween 80 by microscopy (Figure 5.7b). As a general trend, we observed that with increasing percentage of Tween 80 the membranes were more fluid and dynamic (Figure 5.7c) and the vesicles were smaller in size.

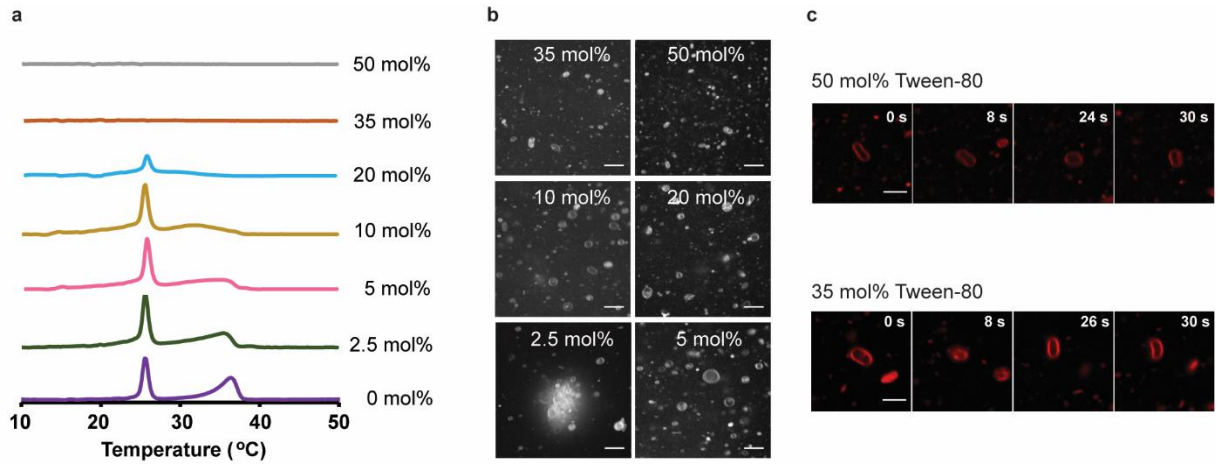


Figure 5.7. Effects of various percentages of Tween 80 on **GOA** self-assembly studied by DSC and fluorescence microscopy. a. DSC thermograms of aqueous dispersions of a 0.5 mM **GOA** aqueous solution containing 0, 2.5, 5, 10, 20, 35, and 50 mol% of Tween 80 at scan rates of 30 °C/h. b. Representative fluorescence microscopy images of **GOA** vesicles containing different percentages of Tween 80. Scale bars: 10 µm. c. Dynamic behavior of **GOA** vesicles containing high percentages of Tween 80 as visualized by spinning disk confocal microscopy. Scale bars: 5 µm. In b and c, membranes are stained with Texas Red-DHPE.

5.7 Encapsulation of molecules of various sizes

For practical applications, it is highly important to know if the vesicles derived from β-D-galactopyranosylamide glycolipids can entrap molecules of biological relevance. Initially, we encountered some practical difficulties to encapsulate molecules in **GOA** vesicles because of its relatively high main transition temperature (~37 ºC) and poor hydration. However, we observed that inclusion of small quantities (6.25 mol%) of Tween 80 in the hydration buffer significantly improved the hydration and subsequent isolation of the vesicles. We encapsulated fluorescent or fluorescently-labeled molecules of various sizes, such as the polar fluorescent dye HPTS (pyranine), fluorescein isothiocyanate (FITC)-labeled dextran (3-5 kDa), FAM-labeled DNA oligonucleotide (dN_{20}) and super-folder green fluorescent protein (sfGFP) in **GOA** vesicles as observed by microscopy (Figure 5.8a i-iv). These results suggest that **GOA** vesicles can encapsulate and retain molecules relevant to biological systems.

5.8 Activity of β-galactosidase encapsulated in GOA vesicles

Construction of cell-mimetics necessitates encapsulation of proteins and other macromolecules in functional form. As a proof of concept demonstration, we chose the 540 kDa (tetramer) enzyme β-galactosidase as a model protein (Figure 5.8b). At first, we hydrated a thin film of **GOA** with a buffer containing β-galactosidase. Following this, we digested the unencapsulated β-galactosidase with proteinase K. Then, we added a fluorogenic substrate 4-methylumbelliferyl β-D-galactopyranoside (MUG) and observed a continuous linear increase in the fluorescence from 4-methylumbelliferone (λ_{ex}: 375 nm, λ_{em}: 475 nm) – the fluorescent product generated from MUG by β-galactosidase. In order to verify that the observed activity is only from the encapsulated enzyme, we carried out a control experiment where proteinase K was added in the hydration buffer from the beginning. Upon addition of MUG, we did not observe any increase

in the fluorescence signal, suggesting that proteinase K digestion is effective and there is no residual activity from the cleaved fragments.

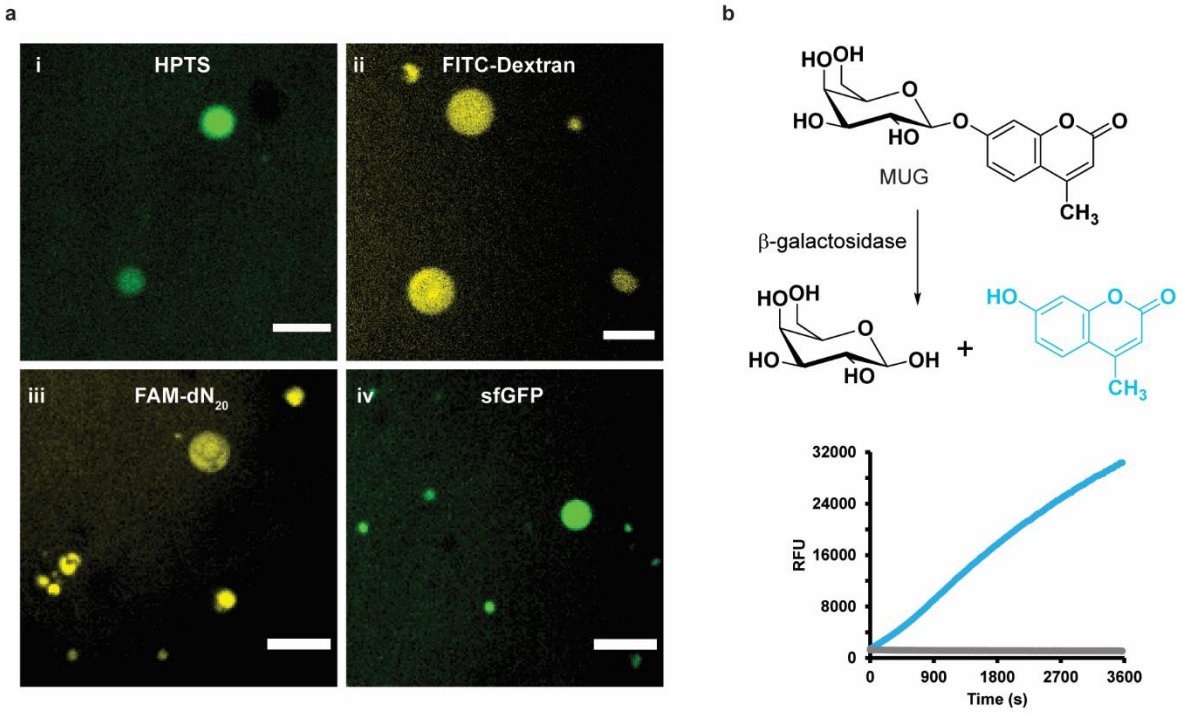


Figure 5.8. Encapsulation of various molecules in GOA vesicles. a. Spinning-disk confocal microscopy images of **GOA** vesicles encapsulating i. HPTS (scale bar: 10 μm), ii. FITC-Dextran (scale bar: 5 μm), iii. 5′-FAM dN_{20} oligonucleotide (scale bar: 10 μm) and iv. sfGFP (scale bar: 10 μm). b. Activity of β-galactosidase encapsulated in **GOA** vesicles monitored fluorimetrically. 6.25 mol% Tween 80 was used in all vesicle preparations.

5.9 Temperature dependent morphological transformation of GOA

5.9.1 Microscopic observations

With these results, we were aiming to use **GOA**/Tween 80 vesicles for delivering doxorubicin to HeLa cells. However, while attempting to carry out these experiments, we obtained an unexpected result. We observed that if the vesicles are stored below room temperature for prolonged period, then they transform into fibers consisting mostly of flat sheets and needle-like structures (Figure 5.9a). TEM, SEM and cryo-SEM further confirmed these morphologies (Figure 5.9b-g). We also followed the vesicle-to-fiber and fiber-to-vesicles in real time using phase-contrast microscopy.

5.9.2 Physical basis of morphological transformations

Interconversion between various supramolecular morphologies in response to external stimuli has been extensively documented in the literature including temperature-dependent transformation between vesicles and tubular or other high aspect ratio morphologies.[12,13]

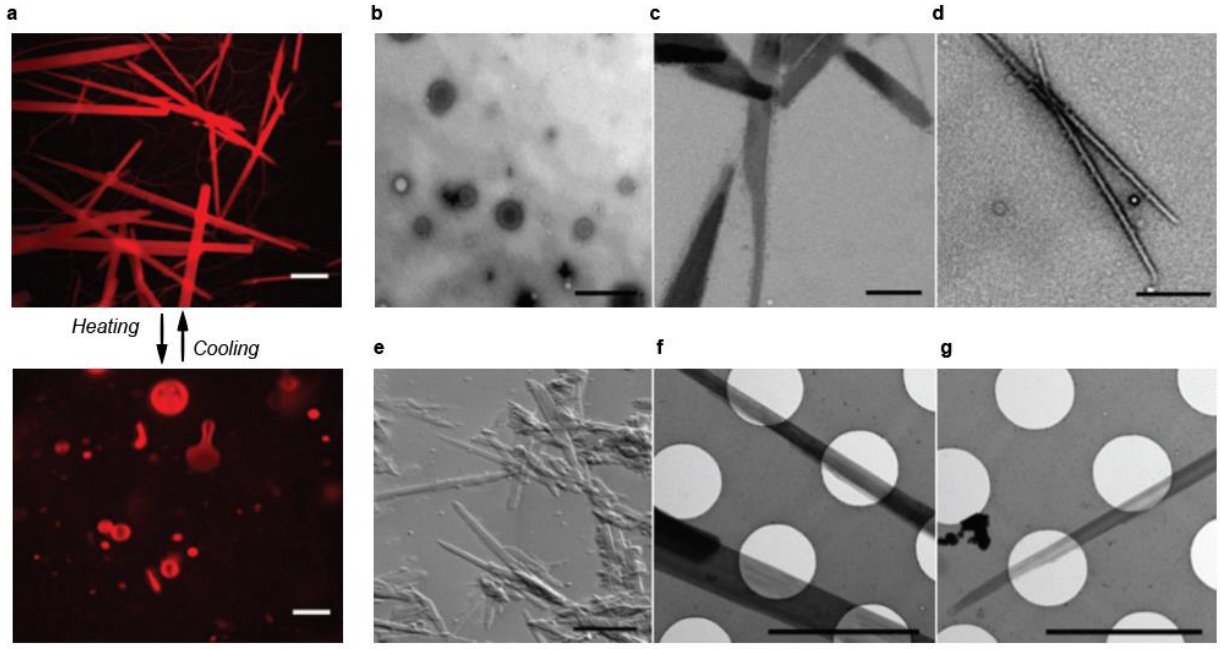


Figure 5.9. Microscopic observation of self-assembled aggregates of **GOA** at different temperatures. a. Fluorescence microscopy images showing reversible transformation between fiber and vesicle morphologies stained with Nile Red. Negative-staining TEM images of b. vesicles c. flat sheets d. needle-like morphologies. e. SEM images of **GOA** fibers, cryo-TEM images showing f. flat sheet and g. needle-like morphologies. Scale bars: a – 10 μm, b – 0.5 μm, c – 1.5 nm, d – 0.5 μm, d, e, f – 5 μm.

Among the SCAs, those bearing a glucose head group are perhaps the most well-studied for temperature dependent transformations.[4,14,15] Similar observations were reported for galactose-containing bolaamphiphiles.[16] Our observation with **GOA** is the first report on the temperature-dependent morphological transformation of a galactose-containing single-chain amphiphile. Circular dichroism (CD) spectroscopy can provide experimental evidence that formation of fibers is directed by chiral molecular packing.[13] At 25 °C, CD spectrum of the **GOA** sample containing fiber-like structures showed a negative peak centered approximately at 220 nm (Figure 5.10a). When the temperature of the sample was raised to 37 °C, the CD signal diminished and disappeared

at 60 °C. The hydrocarbon chains are packed in a fluid state at higher temperatures in vesicle morphologies resulting in loss of CD signal. These results indicated the presence of chiral molecular packing in the fiber-like structures but not in the vesicles.

To study the mesophase behavior of the vesicles and fibers, we employed SAXS using synchrotron radiation source. A dispersion of **GOA** fibers loaded in a glass capillary gave Bragg reflections (n = 1, 2), suggesting the presence of lamellar ordering (Figure 5.10b). From the intensity profile, the d-spacing was calculated to be 37.40 Å. This distance is slightly higher than the previously calculated head-to-head distance of 35.25 Å (for **GOA** bilayers), suggesting that the **GOA** molecules are only weakly hydrated in the fibers. When the same capillary was heated above main phase transition temperature, Bragg reflections (n = 1, 2) were obtained at higher lamellar repeat distances (Figure 5.10b) and the intensity was decreased. The d-spacing was calculated to be 45.53 Å, as expected from a more disordered molecular packing in fluid phase vesicular morphology.

To better understand the thermodynamic nature of this transformations, we carried out DSC studies on multilamellar dispersions of **GOA**. When the sample block was incubated at 4 °C for 9-10 h, DSC thermograms revealed the presence of two peaks – an intense, sharp, and symmetrical peak centered at 25 °C, and a broad, strong, asymmetrical peak with maximum at 37 °C (Figure 5.10c-I). As the holding time is shortened to ~3 h, the first peak is diminished while the second peak remained unchanged (Figure 5.10c-II). When the samples were incubated throughout at 25 °C, the first peak disappeared (Figure 5.10c-III). Thus, incubation of a multilamellar vesicular dispersion of **GOA** at low temperature for prolonged time showed how the vesicles transformed to fibers *in situ* and displayed the respective DSC peaks. However, when the samples were incubated at higher temperatures, the vesicles partly retained their fluid nature and no conversion

to fibers was observed. We further corroborated these results by microscopic observations of the left-over dispersions and observed fibers in the first two conditions, but only vesicles in the final condition. We attributed the peak at 37 °C to be the main transition peak (gel to fluid), and the

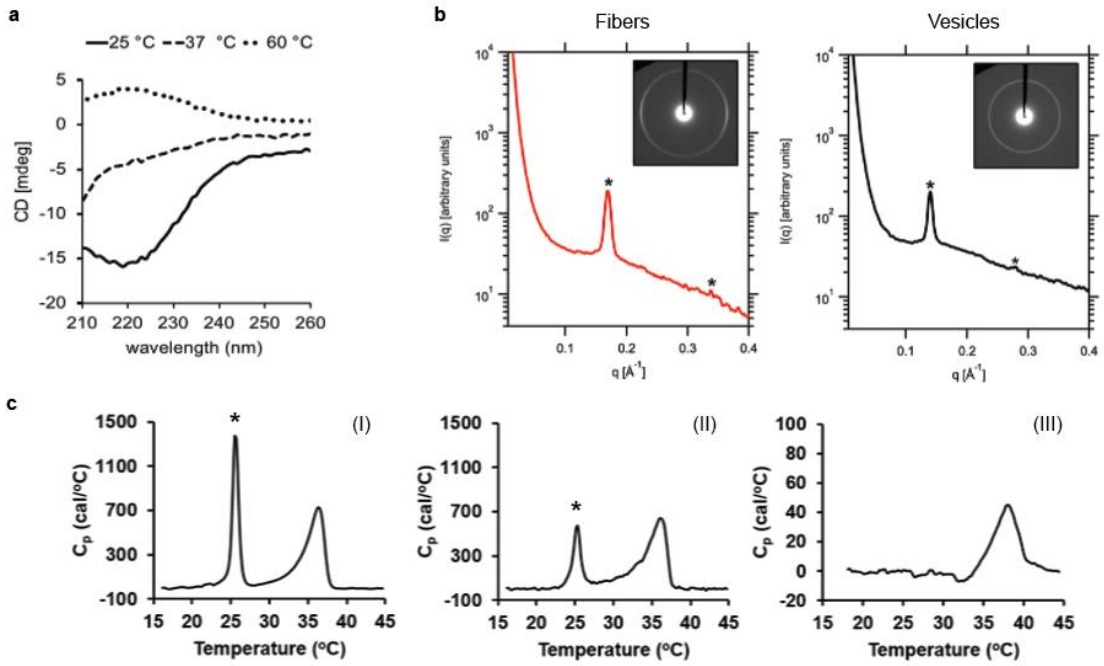


Figure 5.10. Physical basis of morphological transformations of **GOA**. a. CD spectra of a **GOA** dispersion (fibers and vesicles) at varying temperatures, b. SAXS intensity profiles of dispersions of **GOA** fibers and vesicles, c. DSC thermograms obtained when the sample is held for (I) 9-10 h (at 4 °C) (II) 2-3 h (at 4 °C) (III) at 25 °C.

sharp peak at 25 °C as the one corresponding to the fiber-to-vesicle transition. Based on the sharpness of the DSC peak, we reason that the fiber-to-vesicle transformation is a cooperative process and entropy-driven as suggested in previous reports.[17]

On a molecular level, we believe that strong directional intermolecular interactions among the amide groups near the head-group region play a key role in the transformation between vesicles to fibrous structures similar to what has been reported with other amphiphilic systems.[16,18] Interestingly, we observed that vesicles derived from *S*-oleoyl β-D-thiogalactopyranose (**OTG**) do not transform to fibers even after one week of incubation at 4 °C, further highlighting the significance of the amide bond for such transformation to take place. Also, **GPhyA** vesicles (Figure 5.1) do not undergo a similar morphological transformation even when kept at 4 °C for

several days, suggesting that a rigid segment (such as unsaturated bond) in the hydrophobic chain is necessary.

5.10 Conclusion and outlook

In this chapter, we have described in detail the self-assembly of a collection of single-chain β-D-glycopyranosylamide amphiphiles. We have shown that amphiphiles bearing a galactose headgroup are capable of self-assembly into vesicles when the fatty acid chain is unsaturated or branched. Although lamellar phases have been observed in aqueous dispersions of various double-chain glycolipids,[1,19,20] reports of stable vesicle formation from single-chain glycolipids are rare. Moreover, it is likely that membranes derived from amphiphiles having various fatty acid chains will have different fluidity and permeability properties and such properties may be exploited for biological membrane-mimetic applications. The observation that the identity of the monosaccharide group drastically influences the self-assembly properties opens avenues for exploration of the properties of single-chain amphiphiles derived from other monosaccharides. Since the effect of stereochemical variations on structure and assembly is drastic in glycolipids compared to glycerophospholipids, it is highly likely that new properties could be discovered from various permutations of glycolipids. We briefly studied the *N*-oleoyl derivatives of β-D-mannopyranosylamine and β-L-fucopyranosylamine, and found that they form various fibrous aggregates, although the precise nature of those aggregates will require detailed studies.

Development of novel glycolipid amphiphiles may be beneficial for finding potential applications. We would like to point out that the syntheses of various β-D-glycopyranosylamides are straightforward and can be scaled up easily. This is largely because the β-D-glycopyranosylamine precursors are readily accessible from the corresponding sugars[21,22] and various fatty acids are available from natural sources. Single-chain glycolipids being non-ionic

amphiphiles, are likely to be non-toxic to the cells[23] than cationic lipids. Hence, they may be used in the development of vesicular drug delivery systems such as niosomes[8,24] and transferosomes.[25] Finally, the temperature-dependent morphological transformations of **GOA** may find applications in controlled released of molecular payload.

5.11 Experimental methods

5.11.1 General considerations

Commercially available β-D-galactopyranosylamine and β-D-glucopyranosylamine were used as obtained from Carbosynth LLC. Oleic acid (OA), O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluorophosphate (HATU), *N*,*N*-diisopropylethylamine (DIPEA), *N,N*-dimethylformamide (DMF), 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS), Tween 80, β-galactosidase (*E. coli* overproducer, Roche) and 4-methylumbelliferyl β-D-galactopyranoside (MUG) were obtained from Sigma-Aldrich. Palmitoleic acid (POA) and myristoleic acid (MOA) were obtained from Cayman Chemicals. Proteinase K was obtained from New England Biolabs. Texas Red 1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine, triethylammonium salt (Texas Red DHPE) was purchased from Biotium. *N*,*N*-Dimethyl-6-dodecanoyl-2-naphthylamine (Laurdan) was obtain from Chemodex. Nuclear magnetic resonance (^{1}H and ^{13}C NMR) spectra were acquired on a Varian VX-500 MHz instrument. Transmission electron microscopy (TEM) images were recorded on a FEI Tecnai^{TM} Sphera and FEI Tecnai^{TM} Spirit G2 120 kV microscopes. Circular dichroism (CD) spectra were recorded on an Aviv CD Spectrometer. Differential scanning calorimetry (DSC) measurements were carried out on a Microcal VP-Capillary DSC instrument (Ward Lab, The Scripps Research Institute). Microscopy was carried out using an Olympus BX51 optical microscope and a Carl Zeiss Axio Observer Z1 spinning-disk confocal microscope.

5.11.2 General procedure for synthesis of fatty acyl glycopyranosylamides

Fatty acid (1 eqvt.) was dissolved in DMF and stirred at 0 °C. After this, HATU (1.2 eqvt.) and DIPEA (1.2 eqvt.) were added respectively. After stirring for about 10 min, □-D-glycopyranosylamine (1 eqvt) was added. Stirring was continued for 1 h at rt, and the organic solvents were removed *in vacuo*. The yellow residue was dissolved in MeOH, filtered, and purified using HPLC.

5.11.3 Compound characterization

GOA: white solid (36.8 mg, 91%). ^{1}H NMR (CD_3OD): 5.44-5.25 (m, 2H, 2×CH), 4.84-4.83 (m, 1H, 1×CH), 3.88 (d, J = 3.1 Hz, 1H, 1×CH), 3.74-3.63 (m, 2H, 2×CH), 3.61-3.47 (m, 3H, 1×CH+1×CH_2), 2.34-2.16 (m, 2H, 1×CH_2), 2.13-1.91 (m, 4H, 2×CH_2), 1.69-1.55 (m, 2H, 1×CH_2), 1.41-1.24 (m, 20H, 10×CH_2), 0.90 (t, J = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 177.3, 130.9, 130.8, 81.5, 78.2, 75.8, 71.3, 70.5, 62.5, 37.2, 33.1, 30.9, 30.9, 30.6, 30.5, 30.4, 30.3, 30.3, 30.2, 28.2, 28.1, 26.6, 23.7, 14.4. HRMS (ESI-TOF) calculated for [$C_{24}H_{45}NO_6Na$] ([M+Na]$^+$) 466.3139, found 466.3139.

GPOA: white solid (20.7 mg, 89%). ^{1}H NMR (CD_3OD): 5.42-5.27 (m, 2H, 2×CH), 4.85 (d, J = 8.6 Hz, 1H, 1×CH), 3.88 (dd, J_1 = 3.1 Hz, J_2 = 1.1 Hz, 1H, 1×CH), 3.73-3.63 (m, 2H, 2×CH), 3.60-3.47 (m, 3H, 1×CH+1×CH_2), 2.34-2.17 (m, 2H, 1×CH_2), 2.10-1.98 (m, 4H, 2×CH_2), 1.71-1.57 (m, 2H, 1×CH_2), 1.40-1.24 (m, 16H, 8×CH_2), 0.91 (t, J = 6.9 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 177.3, 130.9, 130.8, 81.5, 78.2, 75.8, 71.3, 70.4, 62.5, 37.2, 33.0, 30.9, 30.4, 30.4, 30.3, 30.1, 28.2, 26.7, 23.8, 14.5. HRMS (ESI-TOF) calculated for [$C_{22}H_{41}NO_6Na$] ([M+Na]$^+$) 438.2826, found 438.2821.

GMOA: white solid (8 mg, 74%). ^{1}H NMR (CD_3OD): 5.42-5.29 (m, 2H, 2×CH), 4.85 (d, J = 8.6 Hz, 1H, 1×CH), 3.88 (dd, J_1 = 3.1 Hz, J_2 = 1.1 Hz, 1H, 1×CH), 3.73-3.64 (m, 2H, 2×CH),

3.60-3.48 (m, 3H, 1×CH+1×CH_2), 2.33-2.19 (m, 2H, 1×CH_2), 2.11-1.98 (m, 4H, 2×CH_2), 1.71-1.56 (m, 2H, 1×CH_2), 1.41-1.29 (m, 12H, 6×CH_2), 0.91 (t, J = 7.1 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 177.3, 130.8, 130.8, 81.5, 78.2, 75.8, 71.3, 70.4, 62.5, 37.2, 33.2, 30.9, 30.4, 30.4, 30.2, 28.1, 27.9, 26.7, 23.4, 14.4. HRMS (ESI-TOF) calculated for [$C_{20}H_{37}NO_6Na$] ([M+Na]$^+$) 410.2513, found 410.2508.

GlcOA: white solid (34.3 mg, 85%). ^{1}H NMR (CD_3OD): 5.41-5.26 (m, 2H, 2×CH), 4.89 (d, J = 9.1 Hz, 1H, 1×CH), 3.82 (dd, J_1 = 12.0 Hz, J_2 = 2.1 Hz, 1H, 1×CH), 3.65 (dd, J_1 = 12.0 Hz, J_2 = 5.1 Hz, 1H, 1×CH), 3.39 (t, J = 8.8 Hz, 1H, 1×CH), 3.37-3.32 (m, 1H, 1×CH), 3.30-3.20 (m, 2H, 1×CH_2), 2.24 (td, J_1 = 7.5 Hz, J_2 = 4.1 Hz, 2H, 1×CH_2), 2.10-1.96 (m, 4H, 2×CH_2), 1.72-1.52 (m, 2H, 1×CH_2), 1.39-1.24 (m, 20H, 10×CH_2), 0.90 (t, J = 6.8 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 177.3, 130.9, 130.8, 81.0, 79.6, 79.0, 73.9, 71.4, 62.6, 37.1, 33.1, 30.9, 30.9, 30.7, 30.5, 30.4, 30.4, 30.4, 30.3, 28.2, 28.1, 26.6, 23.8, 14.5. HRMS (ESI-TOF) calculated for [$C_{24}H_{44}NO_6$] ([M-H]$^-$) 442.3174, found 442.3175.

GlcPOA: white solid (3.9 mg, 42%). ^{1}H NMR (CD_3OD): 5.43-5.26 (m, 2H, 2×CH), 4.89 (d, J = 9.1 Hz, 1H, 1×CH), 3.82 (dd, J_1 = 12.0 Hz, J_2 = 2.2 Hz, 1H, 1×CH), 3.65 (dd, J_1 = 11.9 Hz, J_2 = 5.1 Hz, 1H, 1×CH), 3.39 (t, J = 8.8 Hz, 1H, 1×CH), 3.35-3.32 (m, 1H, 0.5×CH_2), 3.30-3.26 (m, 1H, 0.5×CH_2), 3.23 (t, J = 9.1 Hz, 1H, 1×CH), 2.32-2.16 (m, 2H, 1×CH_2), 2.10-1.98 (m, 4H, 2×CH_2), 1.70-1.53 (m, 2H, 1×CH_2), 1.40-1.26 (m, 16H, 8×CH_2), 0.91 (t, J = 7.0 Hz, 3H, 1×CH_3).^{13}C NMR (CD_3OD): 177.3, 130.9, 130.8, 81.0, 79.6, 79.0, 73.9, 71.3, 62.6, 37.2, 33.0, 30.9, 30.9, 30.4, 30.4, 30.2, 30.1, 28.2, 26.7, 23.8, 14.5. HRMS (ESI-TOF) calculated for [$C_{22}H_{41}NO_6Na$]$^+$ ([M+Na]$^+$) 438.2826, found 438.2823.

GlcMOA: white solid (12.9 mg, 73%). ^{1}H NMR (CD_3OD): 5.42-5.25 (m, 2H, 2 × CH), 4.89 (d, J = 9.1 Hz, 1H, 1×CH), 3.82 (dd, J_1 = 12.0 Hz, J_2 = 2.2 Hz, 1H, 1×CH), 3.65 (dd, J_1 =

11.9 Hz, J_2 = 5.1 Hz, 1H, 1×CH), 3.39 (t, J = 8.8 Hz, 1H, 1×CH_3), 3.36-3.32 (m, 1H, 0.5×CH_2), 3.30-3.26 (m, 1H, 0.5×CH_2), 3.23 (t, J = 9.1 Hz, 1H, 1×CH), 2.32-2.16 (m, 2H, 1×CH_2), 2.12-1.94 (m, 4H, 2×CH_2), 1.71-1.52 (m, 2H, 1×CH_2), 1.38-1.29 (m, 12H, 6×CH_2), 0.91 (t, J = 7.1 Hz, 3H, 1×CH_3). ^{13}C NMR (CD_3OD): 177.3, 130.8, 130.8, 81.0, 79.6, 79.0, 73.9, 71.3, 62.6, 37.1, 33.1, 30.8, 30.4, 30.4, 30.2, 28.1, 27.9, 26.6, 23.4, 14.4. HRMS (ESI-TOF) calculated for $[C_{20}H_{37}NO_6Na]^+$ ($[M+Na]^+$) 410.2513, found 410.2509.

5.11.4 Critical aggregation concentration (cac) estimation

Multilamellar dispersions of glycolipids were prepared in Milli-Q H_2O by hydration of a thin film followed by sonication. Afterwards, various solutions (20 µL each) were prepared by dilution of the concentrated dispersions. The samples were kept at 37 ºC for 1 h, following which 0.25 µL of Laurdan (100 µM in EtOH) were added to each and mixed by gentle tapping. Then, the samples were transferred to a 384 well plate and analyzed on a Tecan Infinite Plate Reader at 37 ºC. The samples were excited at 364 nm and emission spectra acquired over 430-500 nm. Generalized polarization (GP) was calculated as follows:

$$GP = \frac{I_{440} - I_{490}}{I_{440} + I_{490}} \quad (5.1)$$

where I_{440} and I_{490} stands for the fluorescence intensities at the wavelengths 440 nm and 490 nm, respectively. The values of GP were plotted against the lipid concentrations for each dilution.

5.11.5 TEM imaging

Copper grids (formvar/carbon-coated, 400 mesh copper) were glow discharged at 20 mA for 1.5 min. After this, 3.5 µL of a dispersion of the glycolipid (**GOA** and **GlcOA**) in H_2O were added to the grid surface and allowed to sit for about 10 seconds. It was washed with 10 drops of ultrapure H_2O and stained subsequently with 3 drops of 1% uranyl acetate. The staining was carried out for about 10 seconds before blotting with filter paper. Samples were then imaged by TEM.

5.11.6 Cryo-EM imaging

Fibers were prepared from 3 mM **GOA** as described previously. Immediately before grid preparation, the fiber sample was pipetted onto plasma-cleaned 200-mesh Quantifoil R 2/2 copper grids (Quantifoil). Using a Vitrobot EM grid plunger (FEI), excess buffer was blotted at room temperature and 95% humidity and the grids were plunge-frozen in liquid ethane maintained at about -180 °C. The grids were stored in liquid nitrogen until use. Fiber samples were imaged on a 120 kV Libra TEM (Zeiss) fitted with a 2k×2k charge coupled device (CCD) camera.

5.11.7 Scanning electron microscopy (SEM) imaging

A dispersion of **GOA** (0.5 mM) fibers was dropped on a glass slide and freeze-dried. Later they were sputter coated with 5 nm of iridium and imaged at 5 kV on FEI Apreo SEM.

5.11.8 X-ray studies on multilayered lipid stacks

A solution (10 mM) of the glycolipid (**GOA**, **GPOA**, **GMOA** or **GlcOA**) in methanol was used to drop coat the film on freshly cleaned silicon [100] substrate. Methanol was evaporated at controlled rate to obtain a uniform coated film over 10×10 mm of surface. The films were then put in high vacuum (10^{-3} Torr) for more than 12 h to remove trapped solvents. To obtain a well-oriented lipid multilayer, the dried films were then rehydrated under 100% relative humidity (RH) at 50 ºC. A slow and steady humidity incubation for more than 24 h produced a well-oriented smectic lipid multilayer membrane with long range correlation among consecutive layers. Subsequently, samples were stored in sealed boxes at room temperature and under relative humidity of ~98 % using saturated salt (K_2SO_4) solution vapor. X-ray measurements were carried out at our in-house X-ray diffraction setup (D8 Discover, Bruker) consisting of a four-circle goniometer and a Cu Kα (8.04 keV, λ ~ 1.54 Å) source. A special sealed humidity chamber was used to control precisely the relative humidity at the sample.

Intensity profiles. For the representation of the typical intensity profiles of the glycolipid samples, we have plotted integrated intensity as a function of transferred wave vector (q_z),

$$q_z = \frac{4\pi}{\lambda}\sin\theta \tag{5.2}$$

where θ is the grazing angle of incidence and λ stands for the wavelength of the incident X-ray.

Relative electron density profiles (EDP). We obtained a set of Bragg peaks from the XRD profile, indicating long-range ordering in the glycolipid multilayers deposited on the substrate. To quantify further, we determined the relative electron density profile (EDP) of the stacked multilayers. The EDP can be can generated from the XRD intensity profile using the following equation:

$$\rho_{relative}(z) = \frac{2}{d}\sum_n \nu_n |F_n| \cos\left(\frac{2\pi n z}{d}\right) \tag{5.3}$$

where d represents the d-spacing of the stacked multilayer, n is the Bragg peak order number, ν_n is the phase factor, and $F_n = n\sqrt{I_n}$ where I_n represents integrated intensity of the n-th order peak and the factor n include the Lorentz correction factor apart from a multiplicative constant. In practice, we considered the first 6 Bragg peaks from the intensity profile. The phase factors ν_n were taken from those obtained from similar lipid multilayers utilizing the swelling method.[26] Since the EDPs of bilayers are assumed to be centrosymmetric, the phase's ν_n can only be of combination of +1 or -1. The relative EDPs are indeterminate to the extent of an overall scale factor and a shift arising from $n = 0$ Fourier coefficient in Eqn. 5.3. The corresponding EDP data allowed us to estimate the membrane thickness (head to head distance of the lipid bilayer) for the glycolipids under consideration.

5.11.9 Small-angle X-ray scattering (SAXS)

SAXS studies were carried out at the beamline 12-ID-C at the Advanced Photon Source (APS), Argonne National Laboratory (ANL). The samples were placed in capillary tubes with 0.01

mm walls and 15 mm outer diameter (15-SG, Charles Supper Company). SAXS measurements were performed with monochromatic photon flux of 1×10^{13} photons s^{-1} cm^{-2}. A 1×1 mm^2 focused beam of incident photon energy at 18 keV (wavelength = 0.69 Å) was used with exposure times of 1 s. The SAXS data were collected by a CCD detector (2048×2048 pixels) and the sample-to-detector distance was 3825 mm. One-dimensional SAXS profiles were derived by azimuthal averaging of the two-dimensional scattering data with proper background subtraction using the Igor Pro macro Nika. Finally, the scattering intensity profiles were obtained as a function of the scattering vector (q).

5.11.10 Differential scanning calorimetry (DSC) measurements

GOA, **GPOA**, **GMOA**, **GlcOA, GlcPOA** and **GlcMOA** dispersions were prepared in Milli-Q H_2O at concentrations of 0.5 mM, 1.0 mM, 2.0 mM, 0.5 mM, 0.5 mM and 1.5 mM, respectively. The volume of sample used for analyses were 450 µL each. The scan rate for each sample was 30 ºC/h. The *gain* settings were set to "high". Background subtraction and processing of the thermograms were carried out using Microcal Origin Software provided by the manufacturer.

5.11.11. Encapsulation experiments

A thin lipid film was prepared on the walls of a glass vial by evaporating a solution of **GOA** in $CHCl_3$:MeOH (2:1). Hydration of the lipid film was carried out by vortexing with 60 µL of a solution containing (i) 1 mM HPTS, 0.2 mM Tween 80, and 50 mM Tris (pH 8.0) (ii) 2 mg/mL FITC-labeled Dextran (3-5 kDa, Sigma Aldrich), 0.2 mM Tween 80, and 50 mM Tris (pH 8.0) (iii) 50 µM 5′-FAM dN_{20} oligonucleotide (Sequence: 5′-FAM-TAATACGACTCACTATAGGG-3', Integrated DNA Technologies), 0.2 mM Tween 80, and 50 mM Tris (pH 8.0) (iv) 8.33 µM His_6-sfGFP, 0.2 mM Tween 80, and 50 mM HEPES-K (pH 7.6). The final concentration of **GOA**

in the dispersions were 3.0 mM. The dispersion was further tumbled at 37 ºC. The unencapsulated molecules were removed using a spin filter (300 kDa MWCO, Pall Corporation). A small volume of the retentate was examined by spinning-disk confocal microscopy.

For encapsulation of β-galactosidase, hydration of **GOA** film (3 mM final concentration) was carried out by vortexing with 60 µL of a solution containing 50 U/mL β-galactosidase, 0.2 mM Tween 80, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 25 mM β-mercaptoethanol in 100 mM HEPES-K (pH 7.6). The dispersion was tumbled at 37 ºC for 1.5 h. Following this, 2 µL proteinase K (20 mg/mL) was added and tumbling continued overnight at 37 ºC. In the control experiment, equal amount of proteinase K was added to the hydration buffer in the beginning. 20 µL of the vesicle dispersion was taken and 0.2 µL of the fluorogenic substrate 4-methylumbelliferyl β-D-galactopyranoside (50 mM in DMSO) was added. The dispersions were immediately placed in a 384 well plate and fluorescence (λ_{ex}: 375 nm, bandwidth: 20 nm; λ_{em}: 475 nm, bandwidth: 20 nm) was monitored every 30 s at 30 ºC for 1 h.

5.11.12 Circular dichroism (CD) spectroscopy

GOA was suspended in deionized water at a concentration of 3.0 mM. Then, 200 µL of the sample was loaded into a quartz cuvette (1 mm light path) and scanned from 260 nm to 190 nm at every 1 nm interval. The sample was measured at 25, 37, and 60 ºC. The background consisted of deionized water and an average of five individual scans were measured. Presented spectra represent averages of ten individual scans with background subtraction.

Notes on chapter

Chapter 5 is adapted majorly from materials published in Bhattacharya, A., Brea, R. J., Bhattacharya, R., Song, J., Sinha, S. K. & Devaraj, N. K. Single-chain β-D-Glycopyranosylamides of Unsaturated Fatty Acids: Self-assembly Properties and Applications to Artificial Cell

Development. *J. Phys. Chem. B.*, 2019, 123, 3711-3720. The dissertation author is the primary investigator on these works. This chapter also contains materials from a manuscript Johnson, M., Bhattacharya, A., Brea, R. J., Podolsky, K. A. & Devaraj, N. K. Temperature-dependent reversible Morphological Transitions in *N*-oleoyl β-D-galactopyranosylamine. This manuscript is currently under review in *J. Phys. Chem. B* and the dissertation author is a co-first author.

We acknowledge Gabriel Ozorowski (TSRI, Ward Lab) for collaboration with the DSC measurements and Henrike Niederholtmeyer for providing sfGFP, Marcus Jager and Christian M. Cole (Kelly Lab, The Scripps Research Institute) for their assistance with CD Spectroscopy.

5.12 References

1. Hato, M., Minamikawa, H., Tamada, K., Baba, T. & Tanabe, Y. Self-assembly of synthetic glycolipid/water systems. *Adv. Colloid Interface Sci.* **80,** 233–270 (1999).

2. Schneider, M. F., Zantl, R., Gege, C., Schmidt, R. R., Rappolt, M. & Tanaka, M. Hydrophilic/hydrophobic balance determines morphology of glycolipids with oligolactose headgroups. *Biophys. J.* **84,** 306–313 (2003).

3. Budin, I., Prwyes, N., Zhang, N. & Szostak, J. W. Chain-length heterogeneity allows for the assembly of fatty acid vesicles in dilute solutions. *Biophys. J.* **107,** 1582–1590 (2014).

4. Kamiya, S., Minamikawa, H., Jung, J., Yang, B., Masuda, M. & Shimizu, T. Molecular structure of gluco-pyranosylamide lipid and nanotube morphology. *Langmuir* **21,** 743–750 (2005).

5. Chen, S. C., Sturtevant, J. M. & Gaffney, B. J. Scanning calorimetric evidence for a third phase transition in phosphatidylcholine bilayers. *Proc. Natl. Acad. Sci. USA* **77,** 5060–3 (1980).

6. Ruocco, M. J. & Shipley, G. G. Characterization of the sub-transition of hydrated dipalmitoyl-phosphatidylcholine bilayers. *Biochim. Biophys. Acta* **691,** 309–320 (1982).

7. Saveyn, P., Van Der Meeren, P., Zackrisson, M., Narayanan, T. & Olsson, U. Subgel transition in diluted vesicular DODAB dispersions. *Soft Matter* **5,** 1735–1742 (2009).

8. Kumar, G. P. & Rajeshwarrao, P. Nonionic surfactant vesicular systems for effective drug delivery - an overview. *Acta Pharm. Sin. B* **1,** 208–219 (2011).

9. El Zaafarany, G. M., Awad, G. A. S., Holayel, S. M. & Mortada, N. D. Role of edge

activators and surface charge in developing ultradeformable vesicles with enhanced skin delivery. *Int. J. Pharm.* **397,** 164–172 (2010).

10. Khan, M. A., Pandit, J., Sultana, Y., Sultana, S., Ali, A., Aqil, M. & Chauhan, M. Novel carbopol-based transfersomal gel of 5-fluorouracil for skin cancer treatment: In vitro characterization and in vivo study. *Drug Deliv.* **22,** 795–802 (2015).

11. Ahad, A., Al-Saleh, A. A., Al-Mohizea, A. M., Al-Jenoobi, F. I., Raish, M., Yassin, A. E. B. & Alam, M. A. Formulation and characterization of Phospholipon 90 G and tween 80 based transfersomes for transdermal delivery of eprosartan mesylate. *Pharm. Dev. Technol.* **23,** 787–793 (2018).

12. Fujima, T., Frusawa, H., Minamikawa, H., Ito, K. & Shimizu, T. Elastic precursor of the transformation from glycolipid nanotube to vesicle. *J. Phys. Condens. Matter* **18,** 3089–3096 (2006).

13. Schnur, J. M., Ratna, B. R., Selinger, J. V., Singh, A., Jyothi, G. & Eashwaran, K. R. K. Diacetylenic lipid tubules: Experimental evidence for a chiral molecular architecture. *Science* **264**, 945-947 (1994)

14. Shimizu, T., Minamikawa, H., Kosigo., M., Aoyagi, M., Kameta, N.,Ding, W. & Masuda, M. Self-organized nanotube materials and their application in bioengineering. *Polym. J.* **46,** 831–858 (2014).

15. Kameta, N., Matsuzawa, T., Yaoi, K., Fukuda, J. & Masuda, M. Glycolipid-based nanostructures with thermal-phase transition behavior functioning as solubilizers and refolding accelerators for protein aggregates. *Soft Matter* **13,** 3084–3090 (2017).

16. Masuda, M. & Shimizu, T. Lipid nanotubes and microtubes: Experimental evidence for unsymmetrical monolayer membrane formation from unsymmetrical bolaamphiphiles. *Langmuir* **20,** 5969–5977 (2004).

17. Smulders, M. M. J., Nieuwenhuizen, M. M. L., Greef, T. F. A., Schoot, P., Schenning, A. P. H. J. & Meijer, E. How to distinguish isodesmic from cooperative supramolecular polymerisation. *Chem. Eur. J.* **16**, 362-367 (2010).

18. Jung, J. H., John, G., Yoshida, K. & Shimizu, T. Self-assembling structures of long-chain phenyl glucoside influenced by the introduction of double bonds. *J. Am. Chem. Soc.* **124,** 10674–10675 (2002).

19. Kitamoto, D., Morita, T., Fukuoka, T., Konishi, M. a. & Imura, T. Self-assembling properties of glycolipid biosurfactants and their potential applications. *Curr. Opin. Colloid Interface Sci.* **14,** 315–328 (2009).

20. Chemical, T. acid) belong to glycolipids ,. 763–766 (1987).

21. Lubineau, A., Aug, J. & Drouillat, B. Improved synthesis of glycosylamines and a straightforward preparation of N-acylglycosylamines as carbohydrate-based detergents. **266,** 211–219 (1995).

22. Sagar Reddy, G. V., Rao, G. V., Subramanyam, R. V. K. & Iyengar, D. S. A New Novel and Practical One Pot Methodology for Conversion of Alcohols to Amines. *Synth. Commun.* **30,** 2233–2237 (2000).

23. Ernst, R. & Arditti, J. Biological effects of surfactants, IV. Effects of non-ionics and amphoterics on HeLa cells. *Toxicology* **15,** 233–242 (1980).

24. Bartelds, R., Nematollahi, M. H., Pols, T., Stuart, M. C. A., Pardakhty, A., Asadikaram, G. & Poolman, B. Niosomes, an alternative for liposomal delivery. *PLoS One* **13,** 1–18 (2018).

25. Rai, S., Pandey, V. & Rai, G. Transfersomes as versatile and flexible nano-vesicular carriers in skin cancer therapy: the state of the art. *Nano Rev. Exp.* **8,** 1325708 (2017).

26. Ma, Y., Ghosh, S. K., Dilena, D. A., Bera, S., Lurio, L. B., Parikh, A. N. & Sinha, S.K. Cholesterol Partition and Condensing Effect in Phase-Separated Ternary Mixture Lipid Multilayers. *Biophys. J.* **110,** 1355–1366 (2016).

Chapter 6 | Lipid Sponge Droplets as Programmable Synthetic Organelles

6.1 Introduction

6.1.1 Need for novel model systems to mimic membrane-rich organelles

In the previous chapters we discussed the importance of vesicular structures as mimics of membrane-bound compartmentalized structure of cells and many of its constituent organelles. Several methodologies have been developed to sequester specific biochemical processes to defined locations and to emulate functions of natural cellular organelles.[1] For instance, mimics of chloroplasts[2,3] and nuclei[4] have been described using vesicles as model systems. However, vesicle models do not capture the essential features of many membrane-rich intracellular organelles and structures such as the endoplasmic reticulum (ER), Golgi apparatus, and inner mitochondrial membrane. These organelles contain extremely high densities of tightly packed membrane networks. In eukaryotic cells, the ER is often the organelle with the greatest surface area and may contain 50-60% of the total membrane.[5] Lipid membranes of the ER form a vast convoluted network interspersed by aqueous channels. Due to the presence of numerous hydrophobic and hydrophilic volumes, the ER hosts an enormous diversity of soluble and membrane-bound proteins at high concentrations and serve as biochemical factories of cells. Vesicle models cannot adequately mimic various nuances of endomembrane systems, such as occurrence of interconnected bilayers, highly convoluted topology, and occurrence of triply periodic minimal surfaces. Further, encapsulation of high concentrations (~100 mg/mL) of molecules within vesicles is often inefficient.

Coacervates (or liquid-liquid phase separated droplets) formed from oppositely charged polyions or immiscible polymers can ameliorate some of the limitations existing with vesicle models. For instance, coacervates can spontaneously encapsulate molecules at concentrations

approaching those inside cells. Unlike vesicles, they lack a physical barrier and they are open to the surrounding environment, thus allowing facile transport of molecules. Their ability to spontaneously enrich molecules from the environment has led to enhanced reaction rates for ribozymes,[6] improved transcription,[7] and acceleration of enzymatic reactions.[8] However, coacervates lack a membrane architecture, therefore they cannot host functional membrane proteins.

In this chapter, we will describe how we combined the best features of vesicle and coacervate systems into a novel bicontinuous sponge phase droplets systems and mimicked basic functions of membrane-rich organelles.

6.1.2 Bicontinuous lipidic mesophases

Many lipids self-assemble into bicontinuous mesophases such as cubic and sponge (also referred to as L_3 phase, molten cubic phase, or lipid coacervates). These phases consist of three dimensional domain of bilayer networks intersected by network of aqueous channels. The word 'bicontinuous' refers to the property that which can be traversed in any direction in both the hydrophilic and hydrophobic domains. Bicontinuous lipidic materials possess properties that resemble highly bilayer-rich organelles such as the ER. Although these lipidic structures are often considered exotic, extensive sponge-like morphology of smooth endoplasmic reticulum has been documented in early electron microscopy studies by Don Fawcett and Vittorio Luzzati. Cubic structures have been observed in plant prolamellar bodies, inner mitochondrial membranes of the stressed amoeba *Chaos carolinensis*, and fat digestion intermediates. Bicontinuous lipid phases offer enormous membrane surface areas, the capacity to harbor both hydrophobic and hydrophilic molecules, and the ability to sequester soluble and membrane proteins.[9] Different kinds of cubic phases have been studied in detail and have been widely applied as a matrix for membrane protein

crystallization.[10] On the other hand, reports of sponge phases are relatively fewer. They have been observed in dispersions of non-ionic surfactants[11], alkylammonium salts[12], Aerosol OT[13], gemini surfactants[14], glycolipid biosurfactants[15], and monoglycerides[16]. Sponge phases frequently form a single bulk phase, exist only within a narrow range of concentration, pH, and temperature, and are highly sensitive to additives.

6.1.3 Sponge phase from GOA and IGEPAL: An unexpected finding

In Chapters 4 and 5, we described the formation of lamellar membranes from single-chain amphiphiles derived from galactopyranose and unsaturated fatty acids. On course of those studies, we discovered serendipitously that *N*-oleoyl β-D-galactopyranosylamine (**GOA**, **Fig. 1a**) forms micron-sized spherical droplets in aqueous media in presence of the commercially available non-ionic octylphenoxypolyethoxyethanol surfactant IGEPAL® CA-630 (IGEPAL, **Fig. 1a**). Using many physical and biochemical techniques we found that the droplets belong to sponge phase that is stable and compatible with a wide range of biological reaction systems. In the following sections, we describe that uptake of specific biomolecules such as nucleic acids and proteins into the droplets can be predictably programmed by doping sponge phase droplets with small quantities of amphiphiles bearing affinity groups. The lipid bilayer-rich nature of the droplets allows spontaneous reconstitution of functional transmembrane proteins. Co-localization and concentration of an enzyme and its substrate in droplets enhanced the rate of a biochemical reaction, and that reversible light-controlled sequestration and release of a protein can be used to regulate an enzymatic reaction. Lipid sponge droplets thus offer a unique platform for mimicking both membrane- and solution-associated organelle functions.

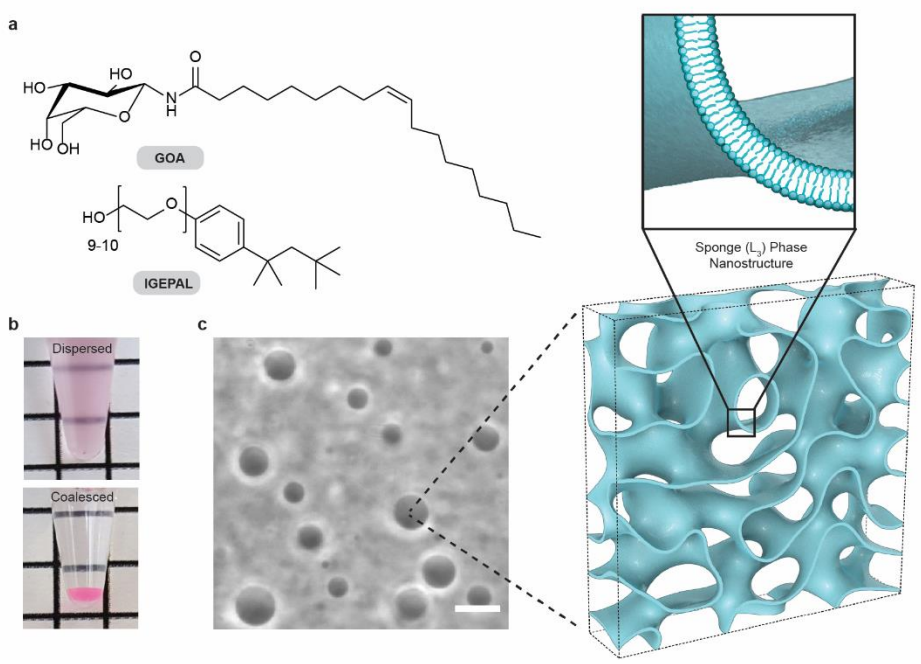


Figure 6.1. Formation of lipid sponge droplets. a. Chemical structures of *N*-oleoyl β-D-galactopyranosylamine (**GOA**) and octylphenoxypolyethoxyethanol (IGEPAL). b. A dispersed sample of droplets merges into a single coalesced phase at the bottom of the tube under gravity. A small quantity of the water soluble dye Rhodamine B was added to the dispersion to demarcate the coalesced phase. c. Phase-contrast image of a typical droplet dispersion and illustration of the droplets' porous, bilayer-rich nanostructure. Illustration was based on a computationally generated model of a bicontinuous structure[54]. Scale bar: 10 μm.

6.2 Preparation of lipid sponge droplets and study of physical properties

When a thin film of **GOA** deposited on the walls of a glass vial is hydrated with a solution containing the non-ionic detergent IGEPAL (Figure 6.1a), a uniformly turbid dispersion is produced. On centrifugation or long standing, the dispersion coalesces into an optically isotropic single dense phase at the bottom of the tube (Figure 6.1b). By phase contrast microscopy, the dispersion was found to consist of a polydisperse population of micron-sized spherical droplets with distinct optical contrast compared to the surrounding medium (Figure 6.1c). The spherical shape suggested to us that the droplets are liquid in nature. The coalesced phase can be resuspended by gentle agitation or vortexing. We used high-performance liquid chromatography and mass spectrometry (HPLC-MS) to determine the compositions of the supernatant and the coalesced

phases and found that both amphiphiles had partitioned almost entirely (~99%) into the coalesced phase.

We constructed a phase diagram by varying the relative proportions of **GOA** and IGEPAL and found that droplet formation takes place at molar ratios of approximately 1:1 to 1:4 (IGEPAL:**GOA**) occupying a large portion of the phase space (Figure 6.2a). Outside this region, micellar solutions (at higher relative IGEPAL concentrations) or lamellar vesicles (at lower relative IGEPAL concentrations) are obtained. We acquired the fluorescence emission spectrum of the solvatochromic dye Laurdan incorporated in the droplets and obtained a maximum around 490 nm (Figure 6.2b). This data suggests that the droplets have a water-accessible environment where the dye can undergo dipolar relaxation typical of fluid lipid membranes. In comparison, in oil-in-water droplet systems such as oleic acid, Laurdan exhibits a maximum peak at 440 nm. When thermogravimetric analysis (TGA) was carried out on the coalesced phase (obtained from 4:3 GOA:IGEPAL by molar ratio), we observed a gradual loss of about 67% mass over the range 30-100 °C (Figure 6.2c). This lost mass is likely from water present in the coalesced phase.

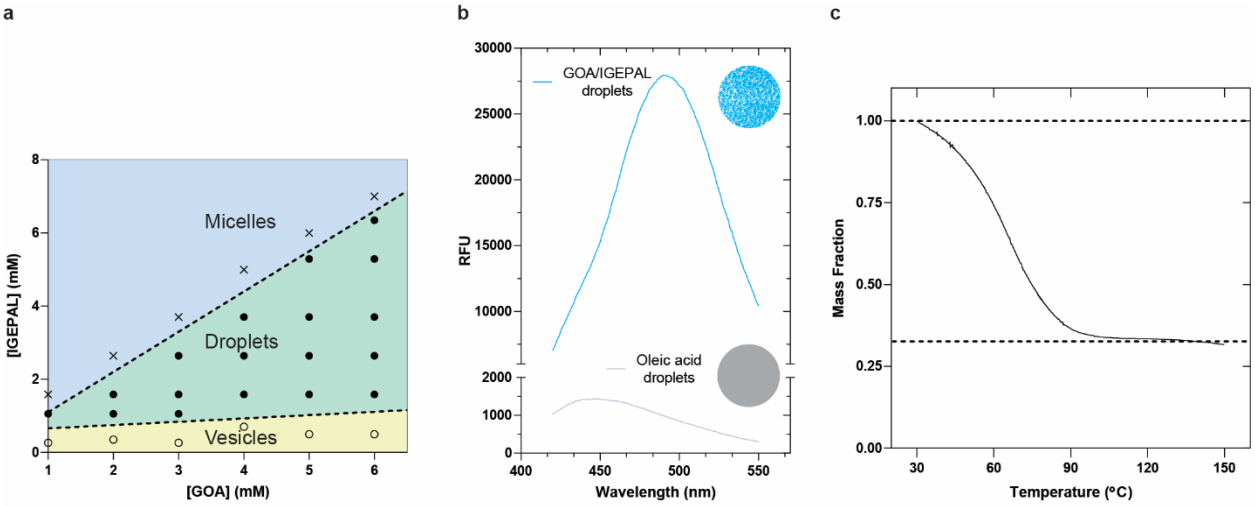


Figure 6.2. Lipid sponge droplets are water-rich. a. Phase diagram showing the relative molar compositions of GOA and IGEPAL in 1X PBS over which vesicles (*yellow*), droplets (*green*), and micelles (*blue*) are obtained. b. Fluorescence spectra of laurdan (0.3 mol%) incorporated in sponge phase droplets and oleic acid oil droplets. c. TGA trace obtained when a GOA:IGEPAL (4:3 molar ratio) coalesced phase was heated. The dotted lines are meant to provide a guide to the eye.

In addition to IGEPAL, we found that GOA forms droplets with other nonionic octylphenoxypolyethoxyethanol detergents such as Triton X-114 and Triton X-100. Single-chain galactolipids with shorter unsaturated fatty acid chains such as **GPOA** and **GMOA** were also found to form droplets with the same class of non-ionic surfactants. In addition, **GPhyA** – the single-chain glycolipid with saturated branched chain fatty acid was also found to form droplets with the same class of surfactants.

6.3 Small-angle X-ray scattering on droplet dispersions

We carried out small-angle X-ray scattering (SAXS) investigation on droplet dispersions using a synchrotron radiation source. When the intensity [$I(q)$] profile from a 4:3 GOA:IGEPAL droplet dispersion was plotted against the scattering vector (q), a single broad peak was obtained (Figure 6.3a). The peak is fitted to a Gaussian distribution and a maximum is obtained at q_{max} = 0.153±0.0002 Å^{-1}. In comparison, SAXS on a dispersion of GOA vesicles gave rise to an intensity profile consisting of a sharp peak (Figure 6.3a) consistent with a lamellar (L_α) phase.[17] The characteristic length ($d = 2\pi/q_{max}$) for the sponge phase (from 4:3 GOA:IGEPAL) was calculated to be 4.11 nm. This length can be interpreted as the average channel diameter of the sponge phase.[16] We found that the nature of the peak and the position of the maxima of the SAXS profile remains nearly unchanged in the coalesced phase, suggesting that the structural characteristics are retained (Figure 6.3b). For various compositions of the droplets, the intensity profile showed a decay behavior as $I(q) \approx q^{-2}$ consistent with a local bilayer structure (Figure 6.3c).[18–20] Such features are typical of a lipidic sponge (L_3) phase (Figure 6.1c) similar to that reported with various surfactant systems.[16,19] We further carried out SAXS studies on various combinations and ratios of glycolipids and octylphenoxypolyethoxyethanol surfactants to characterize the sponge phase.

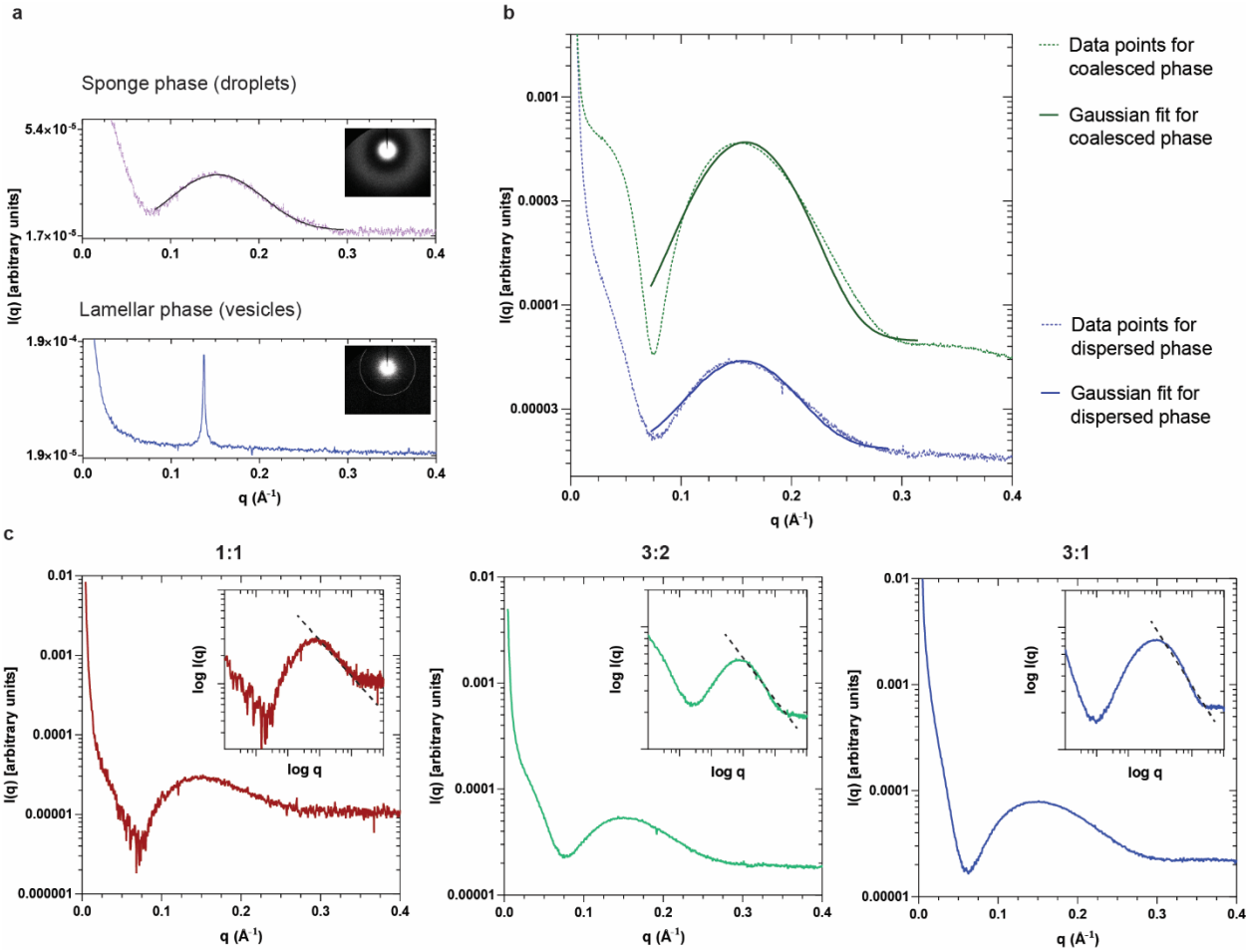


Figure 6.3. SAXS investigation of **GOA**/IGEPAL droplets. a. Top: synchrotron small-angle X-ray scattering (SAXS) intensity profile (dotted) from a droplet dispersion (4:3 **GOA**:IGEPAL by molar ratio). The solid black line represents a Gaussian fit to the data to obtain the position of maximum. Bottom: synchrotron SAXS intensity profile from a dispersion of GOA vesicles. The inset images show the SAXS patterns. b. When the SAXS intensity profiles of dispersed and coalesced phases from 3:2 **GOA**:IGEPAL (by molar ratio) are compared, the nature of the profiles and position of maxima are observed to be practically unchanged. The solid lines correspond to the Gaussian distribution fit to the sponge phase peak data points. For dispersed phase, q_{max} = 0.153±0.0002 Å^{-1}, for coalesced phase q_{max} = 0.156±0.0002 Å^{-1}. c. SAXS intensity profiles from GOA:IGEPAL droplet dispersions in the molar ratio 1:1, 3:2, 3:1 shows a $I(q) \approx q^{-2}$ decay, suggesting a local bilayer structure. The corresponding log-log plot is shown in the inset. The slopes are calculated to be 1.90, 1.96, and 2.18 respectively.

6.4 Cryogenic electron microscopy of sponge phase

We further characterized the droplets by cryogenic transmission electron microscopy (Cryo-TEM) and found that they possess a disordered network of lamellar structures, consistent with the characteristics of a sponge phase (Figure 6.4a).[21] In contrast, Cryo-TEM analysis of a sample of **GOA** vesicles showed distinct membrane bound structures of various lamellarities (Figure 6.4a). We imaged the droplets by freeze-fracture cryogenic scanning electron microscopy

(Cryo-SEM) and observed a porous morphology (Figure 6.4b) corresponding to a three-dimensional network of lipid bilayers.[12,13,15]

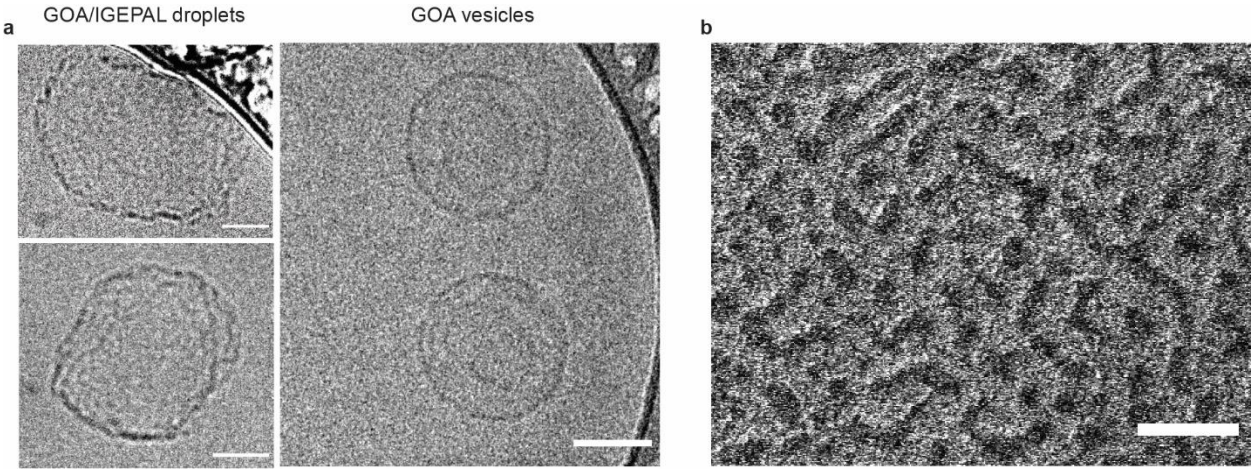


Figure 6.4. Cryogenic electronic microscopy of sponge phase. a. Cryo-TEM of **GOA**/IGEPAL droplets compared to **GOA** vesicles. Scale bars: 50 nm. b. Freeze-fracture cryo-SEM of **GOA**/IGEPAL droplet showing sponge phase morphology. Scale bar: 1 µm.

6.5 Proposed mechanism of sponge phase formation

Previously, it has been proposed that transformations of lamellar to sponge phases are mediated by bending of bilayers followed by formation of interconnections between them.[22] **GOA** and similar galactolipids alone form lamellar phases, which are composed of locally flat stiff bilayer membranes. We speculate that *t*-octylphenoxypolyethoxyethanol surfactants, with their flat aromatic ring and short bulky aliphatic segment, insert between the **GOA** molecules and induce a negative curvature. Additionally, the detergents may make the bilayers more fluid by lowering the bending rigidity. These combined effects would allow adjacent bilayers to fuse and form a three-dimensional sponge-like network.

We made a few key observations that may allow us to elucidate in the future the precise mechanism of sponge phase formation:

i. Only the single-chain glycopyranosylamides consisting of a galactose head-group form droplets. Amphiphiles with glucose, mannose, and fucose head-groups did not form droplets.

ii. For droplet formation, a galactolipid should have a fatty acid chain which occupies a large volume. This includes unsaturated fatty acids (oleic, palmitoleic, myristoleic) and multiply branched-chain fatty acid (phytanic). Glycolipids with saturated chains (palmitic, stearic) were not found to form droplets.

iii. For droplet formation, galactolipids bearing an amide linkage between the head-group and the fatty acid tail are necessary. For comparison, OTG was not found to form droplets.

iv. The aromatic ring of the *t*-octylphenoxypolyethoxyethanol surfactants is necessary for sponge phase formation. For instance, fully hydrogenated Triton X-100 does not cause droplet formation. However, *n*-nonylphenoxypolyethoxyethanol forms droplets with GOA, suggesting that variations in the aliphatic segment can be accommodated.

v. *t*-Octylphenoxyethanol does not form droplets in combination with GOA, suggesting that the short PEG head-group present on the *t*-octylphenoxypolyethoxyethanol surfactants plays important role in influencing hydrogen bonding interaction between GOA molecules. On the other hand, Tergitol^{TM} 15-S-40, which has a much longer PEG chain (n = 40) do not facilitate droplet formation, suggesting that an optimum length of PEG chain is necessary.

vi. Tween detergents, octyl glucoside, DDM, and SDS do not form sponge phase with galactopyranosylamides.

6.6 Robustness of droplet formation

We found that stable droplets can be formed over a wide range of pH, ionic strength, buffer composition, and dilution. We observed droplet formation over the pH range 2-10, high ionic strength (up to 1 M NaCl), high concentrations of Mg^{2+} (100 mM) and Ca^{2+} (10 mM). The droplets also formed in complex media like sea water, cell culture media, and cell-free transcription/translation (TX-TL) reactions. Due to lack of an ionizable head group, glycolipid sponge droplets assemble more robustly compared to other droplet systems based on single chain amphiphiles, such as fatty acids,[23] which can exist only within a relatively narrow range of pH conditions and ionic environments.

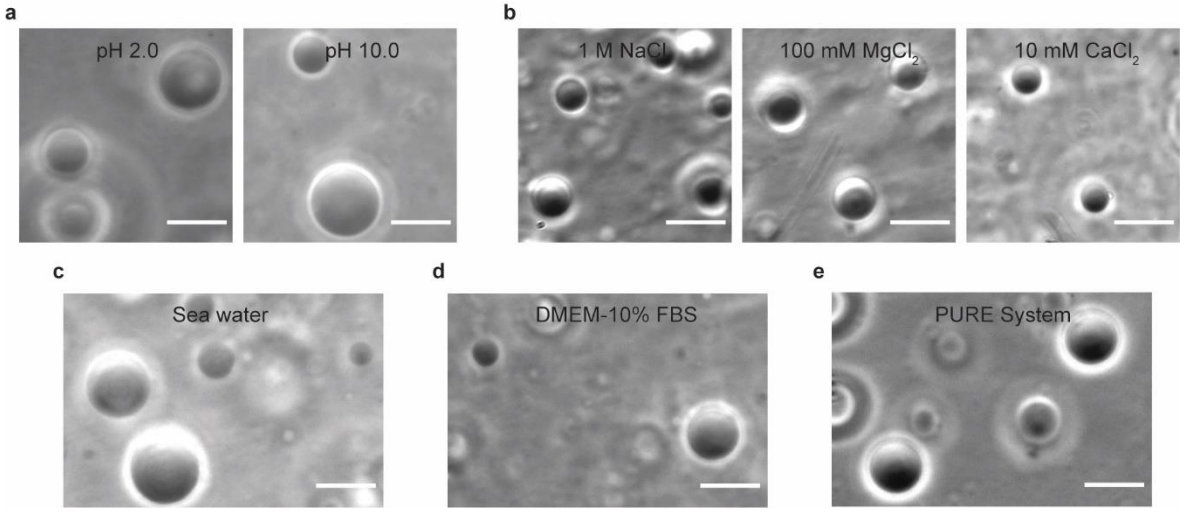


Figure 6.5. Robustness of droplet formation. Formation of droplets from **GOA** and IGEPAL at a. pH 2.0 (Na-glycine buffer), pH 10.0 (Na-glycine buffer) b. 1 M NaCl, 100 mM $MgCl_2$, 10 mM $CaCl_2$ c. Sea water d. DMEM-10% fetal bovine serum (FBS) e. PURE system (cell free transcription-translation medium). All scale bars: 10 μm.

6.7 Partitioning properties of droplets

6.7.1. Small molecules and fluorescent dyes

By virtue of their extensive bilayer network and high water content, lipid sponge droplets can provide a unique milieu for harboring molecules spanning from hydrophobic to hydrophilic in nature. At first, we studied the partitioning behavior of a variety of molecules into the sponge phase

using HPLC-MS. We measured the concentration of an analyte in the supernatant and compared it with the total concentration to estimate the percentage of partitioning into the sponge phase (Table 6.1). Polar lipid molecules such as cholesterol and 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) were 100% partitioned. Water soluble fluorescent dye Rhodamine B was 81.4% partitioned while the more polar dye Sulforhodamine B was 47.3% partitioned.

Table 6.1. Summary of extent of partitioning of molecules (<1 kDa) into sponge phase. In each of these experiments, 5 µg of a molecule of interest was tested for partitioning in 80 µg of droplet material (composed of 4:3 **GOA**:IGEPAL by molar ratio) in 1X PBS. Using HPLC, the concentration of the molecule of interest in the supernatant was measured and then compared with the total to estimate the extent of partitioning. We define "capacity" as the mass of a molecule of interested partitioned per unit mass of droplet-forming amphiphiles.

Partitioning of various molecules into lipid sponge phase

Molecule	Molar mass (g mol^{-1})	Charge	Percentage partitioned	Capacity (mg g^{-1})
Rhodamine B	479.02	0	81.4±5.2	50.9
Rhodamine 6G	479.02	+2	70.2±2.3	43.9
Sulforhodamine B	558.67	-1	47.3±1.9	29.6
7-Diethylamino-4-methylcoumarin	231.29	+1	80.6±1.3	50.4
7-Amino-4-methylcoumarin	175.18	0	73.7±5.5	46.1
7-Hydroxycoumarin (pH 7.4)	162.14	-1	6.3±0.9	3.9
Fluorescein (pH 6.0)	332.21	-1	43.7±1.2	27.3
Fluorescein (pH 8.0)		-2	13.2±3.4	8.3
Doxorubicin (pH 7.4)	543.52	+1	24.7±7.2	15.4
Cholesterol	386.65	0	100±0	62.5
POPC	760.09	0	100±0	62.5
Adenosine	267.24	0	<1	--
5´-N-Ethylcarboxamidoadenosine	308.29	0	<1	--

Next we used fluorescence microscopy to visualize the partitioning behavior of several lipophilic, and water soluble fluorescent molecules (Figure 6.6a). Highly sulfonated dyes such as Alexa Fluor 488, and sulfo-Cy5-azide showed little or no enrichment (Figure 6.6a, vii-viii). This suggests that dyes like Alexa Fluor 488 can be used as a fluorescent label for macromolecules to study their localization into the droplets in an unambiguous manner. Non-sulfonated xanthene dyes such as fluorescein and 4´,5´-dibromofluorescein showed different degrees of partitioning depending on pH and hence their ionization state. For instance, at lower pH (~6.0) the phenolic -OH of

fluorescein remains mostly protonated, resulting in greater partitioning (43.7%) to the sponge phase. At higher pH (~8.0), the phenolic -OH is ionized, leading to significantly less partitioning (13.2%) to the sponge phase. Based on the above studies, we conclude that overall charge, hydrophobicity/hydrophilicity, and structure of the molecules play an important role in determining the partitioning behavior.

6.7.2 Galactophilic proteins

Galactose and its derivatives play a crucial role in biology as affinity or recognition groups for protein binding on the surfaces of cells or organelles. **GOA** offers ample β-galactopyranosyl moieties throughout the bulk of the sponge droplets. We hypothesized that proteins with affinity for galactose residues would spontaneously partition into the droplets. We chose to test the well-studied galactophilic lectin LecA (PA-IL) from the pathogenic bacterium *Pseudomonas aeruginosa*, where LecA promotes biofilm formation via cross-linking between bacterial cell-surface galactose residues.[24] We added Alexa Fluor 488-labeled LecA (1 μM final concentration) to droplets (formed from 3 mM **GOA** and 1.8 mM IGEPAL) and observed a high extent of partitioning into the droplets (Figure 6.6b-i). Using fluorometric analysis of the supernatant, we estimated that 97.3% of the protein partitioned into the droplets. We carried out control experiments where isopropyl β-D-1-thiogalactopyranoside (IPTG) or phenyl β-D-galactopyranoside were added as competing ligands for LecA and significantly less partitioning of the Alexa Fluor 488-labeled LecA was observed by microscopy, thus supporting our hypothesis that sequestration was due to a specific interaction of the lectin with GOA. Next we tested the galactophilic lectin galectin 3, which is known to bind primarily to β-galactoside residues and plays an important role in cell adhesion, macrophage activation, and apoptosis.[25] As expected, we observed highly efficient partitioning of Alexa Fluor 488-labeled galectin-3 into the droplets

(Figure 6.6b-ii). Apart from galactophilic lectins, we found that a GFP fusion of the *Escherichia coli lac* operon repressor protein (sfGFP-LacI) spontaneously partitions into the droplets (Figure 6.6b-iii). LacI is known to bind to galactopyranoside effector ligands primarily via hydroxyl groups (O2, O3, O4, O6) of the galactose moiety on the latter. In addition, apolar substituents (if any) on the C1 (anomeric) position of galactose group interact with a hydrophobic surface in the binding site, thereby increasing the binding affinity.[26] We reason that the presence of the galactose group and hydrophobic oleoyl chain on **GOA** contribute to its binding to sfGFP-LacI.

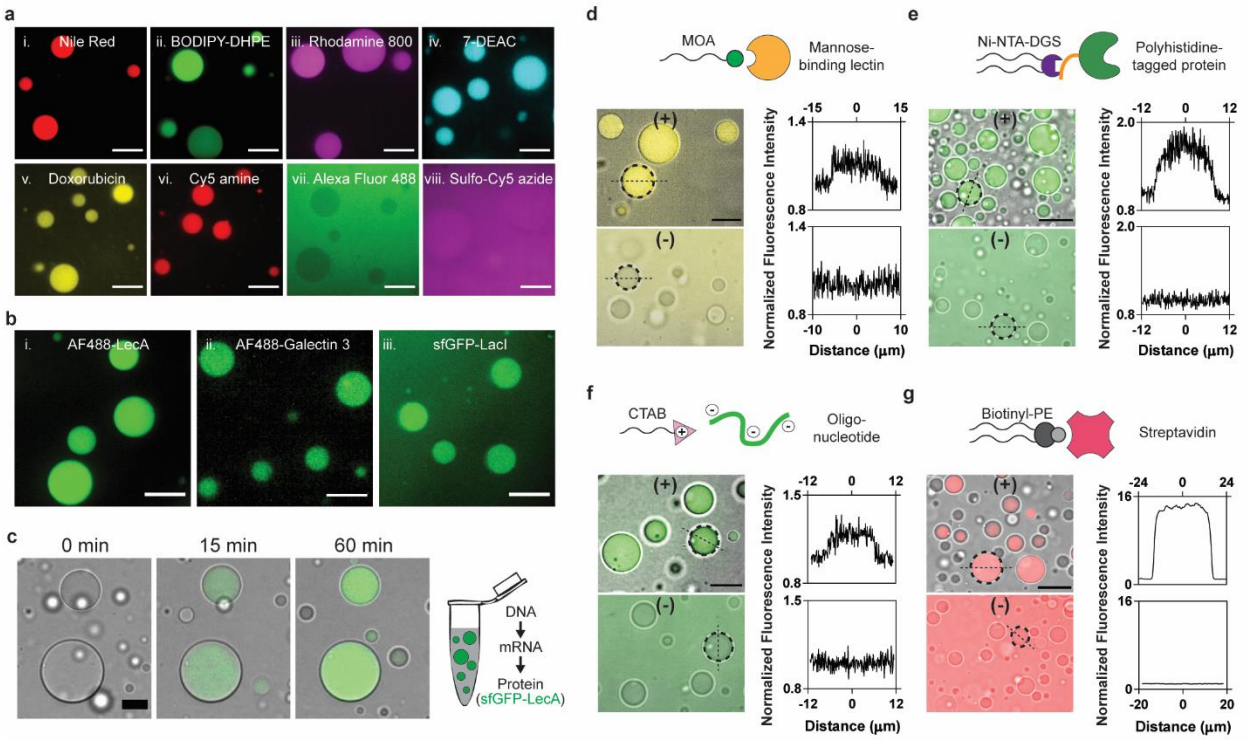


Figure 6.6. Partitioning of various molecules into lipid sponge droplets. a. Partitioning of i) Nile Red ii) BODIPY FL-DHPE iii) Rhodamine 800 iv) 7-diethylaminocoumarin v) Doxorubicin vi) Cy5-amine vii) Alexa Fluor 488 and viii) sulfo-Cy5-azide into lipid sponge droplets. Scale bars: 10 μm. b. Partitioning of galactophilic proteins i) Alexa Fluor 488-LecA ii) Alexa Fluor 488-Galectin 3 and iii) sfGFP-LacI into lipid sponge droplets. Scale bars: 10 μm. c. Time lapse images showing gradual incorporation of sfGFP-LecA expressed in PURE system into lipid sponge droplets. Scale bar: 15 μm. d. Selective partitioning of FITC-Concanavalin A into droplets doped with *N*-oleoyl β-D-mannosylamine (MOA). Scale bar: 15 μm. e. Selective partitioning of sfGFP-His_6 into droplets doped with Ni-NTA-DGS. Scale bar: 20 μm. f. Selective partitioning of a FAM-labeled DNA oligonucleotide into droplets doped with CTAB. Scale bar: 15 μm. g. Selective partitioning of Alexa Fluor 568-labeled streptavidin into droplets doped with biotinyl-PE. Scale bar: 35 μm. In d-g, binding (doped) droplets are represented by (+) while non-binding droplets are represented by (-). Plots of the fluorescence intensity profile along a straight line drawn through the center of the marked droplets (broken circle) are shown on the right. On the *x*-axes, 0 corresponds to the center of the droplets.

6.7.3 Partitioning of galactophilic proteins expressed by cell-free TX-TL

To evaluate the compatibility of droplets and their sequestration properties with complex biochemical reactions, we formed droplets in a recombinant TX-TL system (PURExpress). The PURE TX-TL system consists of almost 30 enzymes, 56 tRNAs, ribosomes, and a highly optimized mixture of precursors and buffer components.[27] Even when droplets were present at high densities (reaction contained 12 mM **GOA** and 6.7 mM IGEPAL), efficient protein synthesis took place and we observed gradual sequestration of galactophilic proteins during their synthesis (Figure 6.6c). While LecA and LacI fused with sfGFP spontaneously partitioned into the droplets upon synthesis, sfGFP alone did not. Cell-free protein synthesis in the presence of droplets shows that droplets are highly biocompatible and will likely support many different biochemical reactions. Further, we can utilize this methodology as a rapid route to introduce functional proteins into the droplets to carry out specific biochemical transformations.

6.7.4 Programming partitioning behavior

Living cells and sub-cellular organelles contain specific receptors to recognize macromolecular binding partners or signaling molecules. For example, lysosomal enzymes synthesized in the rough ER are modified with mannose-6-phosphate for targeting them to lysosomes.[28] Since our droplets are composed of amphiphilic species, we asked if such targeting mechanisms could be mimicked by doping the droplets with amphiphiles bearing affinity handles. At first, we doped the droplets with 11.3 mol% *N*-oleoyl β-D-mannopyranosylamine (MOA) and added fluorescein isothiocyanate (FITC) labeled lectin Concanavalin A, which has specificity for terminal mannopyranosyl residues. As expected, we observed that the droplets doped with MOA had a 1.14-fold higher fluorescence signal compared to the background while those without did not show increased fluorescence (Figure 6.6d). Encouraged by this result, we doped the droplets

with 1.6 mol% 1,2-dioleoyl-*sn*-glycero-3-[(*N*-(5-amino-1-carboxypentyl) iminodiacetic acid) succinimidyl] (nickel salt) (Ni-NTA-DGS) – a headgroup modified phospholipid that binds to polyhistidine-tagged proteins with high affinity and specificity. When we added sfGFP-His_6, we observed highly efficient localization using confocal microscopy. In absence of Ni-NTA-DGS, no localization of sfGFP-His_6 was observed (Figure 6.6e).

Next, we doped the droplets with 10 mol% of the cationic amphiphile cetyltrimethylammonium bromide (CTAB) and found that it facilitated the sequestration of a fluorescently labeled oligonucleotide (5´-FAM dN_{20}) (Figure 6.6f). We measured the fluorescence signal in the droplets to be 1.2-fold higher than the background when 2 µM 5´-FAM dN_{20} was added to CTAB-doped droplets, while no increase was observed in non-doped droplets. We performed SAXS on the droplets doped with CTAB and found that sponge phase characteristics were retained. We also observed efficient enrichment of DNA oligonucleotide when the droplets were doped with 2-hydroxydodecylamine. Doping with a biotinylated phospholipid (biotinyl-PE) allowed us to recruit fluorescently labeled Streptavidin into the sponge phase, where the protein was concentrated more than 10-fold from solution (Figure 6.6g). Recruitment of streptavidin suggests that additional biotinylated molecules may be co-sequestered when bound to streptavidin tetramers.

6.7.5 Quantification of partitioning of sfGFP-His_6

To quantify how much sfGFP-His_6 partitioned into Ni-NTA-DGS-doped droplets, we analyzed droplet supernatants fluorometrically. Droplets sequestered more than 90% of the protein when we titrated sfGFP-His_6 between 1 nM and 10 µM (Figure 6.7a). At 10 µM total sfGFP-His_6 concentration, we estimate the protein concentration to be almost 1 mM in the droplet phase, based on droplet formation experiments at comparable **GOA** and IGEPAL concentrations in large

volumes, where the droplets phase occupied approximately 1% of the total volume. Our calculations indicate that poly-His-tagged proteins in the droplet phase can therefore easily reach the concentrations of even the most abundant proteins in cells (for instance, the translation elongation factor EF-Tu is present at approximately 100 µM in *E. coli*).[5] The capacity of droplets for sequestering His-tagged protein obviously depends on the density of droplets and the amount of Ni-NTA-DGS, which explains why partitioning saturates as the protein concentration approaches the concentration of Ni-NTA-DGS (Figure 6.7b-c).

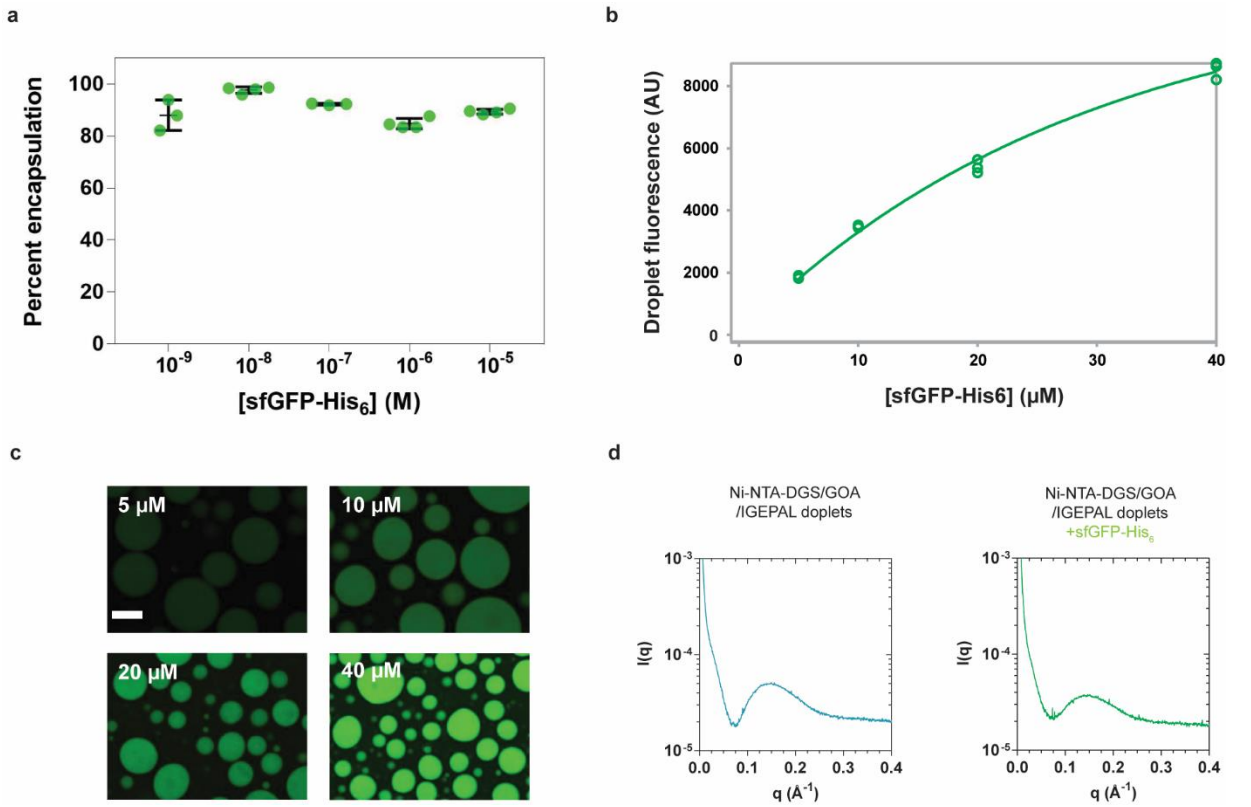


Figure 6.7. Partitioning of sfGFP-His$_6$ in Ni-NTA-DGS doped sponge droplets. a. Quantification of the extent of partitioning of sfGFP-His$_6$ into sponge phase droplets doped with 1.6 mol% Ni-NTA-DGS at various concentrations of the protein. Partitioning was estimated from fluorescence of the supernatant. b. Partitioning measured by confocal microscopy. Droplet fluorescence intensities at increasing sfGFP-His$_6$ concentrations were measured in three images and data fitted to an exponential saturation function. Fluorescence begins to saturate as protein concentration approaches total Ni-NTA-DGS concentration (83.3 µM). c. Representative images of droplet fluorescence displayed at identical contrast settings. Scale bar: 25 µm. d. SAXS intensity profiles of droplets doped with Ni-NTA-DGS in *absence* and *presence* of sfGFP-His$_6$.

To verify that small quantities of Ni-NTA-DGS do not alter the sponge phase structure of the droplets, we carried out SAXS on the doped droplets. We observed that the nature of the SAXS intensity profile is practically unchanged when the droplets are doped with Ni-NTA-DGS with sfGFP-His_6 added or not added (Figure 6.7d).

6.8 Mobility of molecular cargo in lipid sponge droplets

For reactions to occur in a synthetic organelle model, it is important that molecules are mobile within the compartment. As mentioned earlier, the spherical shape suggests that the droplets are liquid in nature. Additionally, we frequently observe fusion events between individual droplets that coalesce into a bigger droplet. To quantify the mobility of molecules inside droplets, we performed fluorescence recovery after photobleaching (FRAP) experiments. To study the mobility of hydrophobic molecules, we bleached the center of a droplet containing a fluorescently labeled phospholipid (Figure 6.8a) and the half-maximal recovery of fluorescence was observed within 6.8 s ($t_{1/2}$). To study the mobility of soluble proteins, we carried out FRAP on Ni-NTA-DGS-doped droplets binding a His_6-tagged fluorescent protein and measured a $t_{1/2}$ of 16 s. As controls, we compared protein diffusion in droplets to the mobility of the protein in aqueous water-in-oil emulsion droplets and bound to a Ni-NTA gel matrix, where, respectively, fluorescence recovery was much more rapid or did not occur over the duration of the experiment (Figure 6.8b). The FRAP results indicate that molecular cargo within the droplet phase diffuses slower than in a dilute aqueous solution but they are sufficiently mobile to allow dynamic molecular interactions to occur.

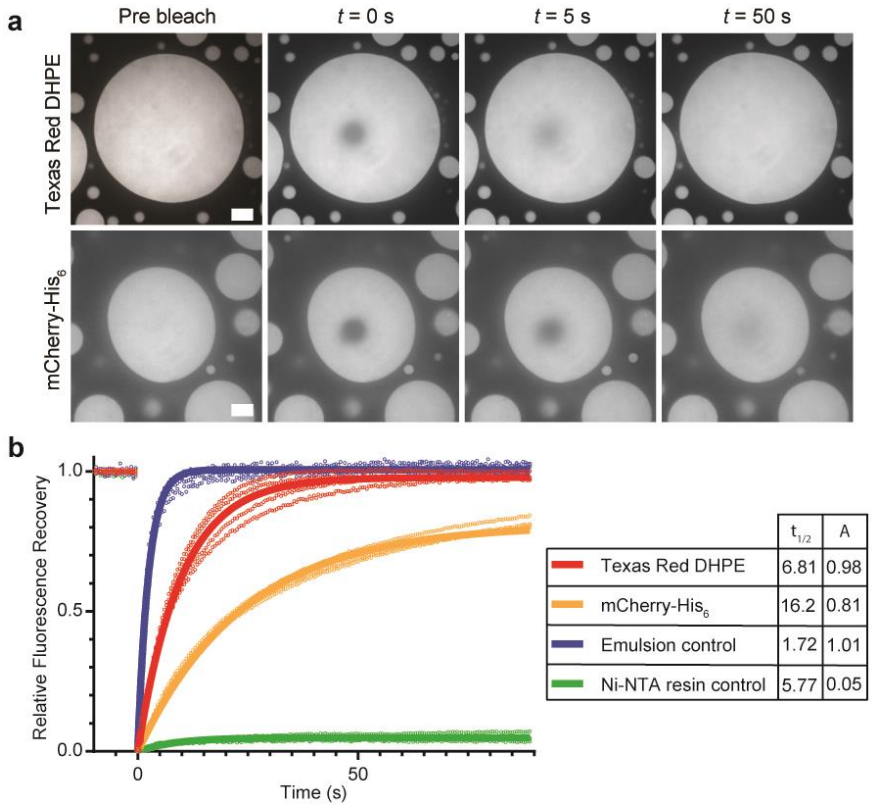


Figure 6.8. FRAP analysis to characterize diffusion of biomolecules in lipid sponge droplets. a. Representative fluorescence images of FRAP experiments on lipid sponge droplets containing the fluorescent phospholipid Texas Red-DHPE or fluorescent protein mCherry-His_6. Scale bars: 10 µm. b. Relative fluorescence recovery of Texas Red DHPE and mCherry-His_6 in droplets, compared to fluorescence recovery of mCherry-His_6 in water-in-oil emulsion droplets and bound to a solid agarose Ni-NTA resin. Shown are data points from at least three FRAP experiments. The fraction of recovery (plateau), A, and $t_{1/2}$ (time to half-maximal recovery, in seconds) were obtained by fitting the data (solid lines) as described in Section **6.15.10**.

6.9 LecA-induced spontaneous structuration of the droplets

Living cells and their constituent organelles undergo various shape transformations in response to stimuli. We observed that incorporation of large quantities of the galactophilic protein LecA within the droplets cause massive structural deformations to happen. The spherical droplets gradually transform to berry-like structures upon incorporating LecA (Figure 6.9a). A likely explanation is that LecA is a tetrameric protein and can bind to four galactose groups. This can cause cross-linking of **GOA** molecules to happen within the sponge phase and hence the structuration events. Occasionally, we observed the splitting of a droplet into daughter droplets induced by lectin binding (Figure 6.9b). Interestingly, the structured droplets reverted to the

original spherical shape along with release of the lectin when a non-amphiphilic competitive ligand such as phenyl β-D-galactopyranoside is added in excess (Figure 6.9c).

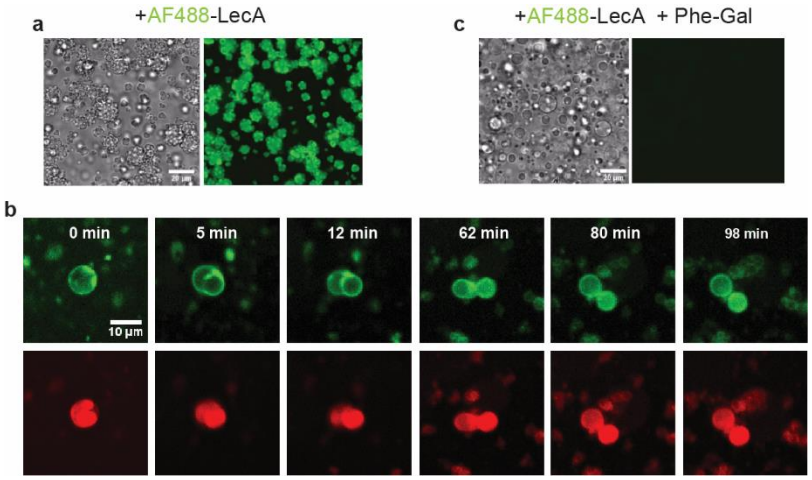


Figure 6.9. LecA-induced spontaneous structuration of the droplets. a. Droplets transform to berry-like morphology upon excess LecA incorporation. Scale bar: 20 µm. b. Splitting of a droplet into daughter droplets. Top panel corresponds to Alexa Fluor 488 channel (LecA) and bottom panel corresponds to Texas Red-DHPE (lipid) channel. Scale bar: 10 µm. c. Structured droplets revert to spherical shape when Phe-Gal is added. Scale bar: 20 µm.

6.10 Reconstitution of functional transmembrane proteins

Membrane proteins are encoded by a large proportion (~30%) of the genome and they fulfill essential cellular functions.[9] However, they are challenging to employ in systems built from the bottom up, limiting the range of biochemical transformations that can be implemented in cell-free reactions. Since the droplets are composed of biocompatible lipid bilayers, we were interested in testing if spontaneous reconstitution of transmembrane proteins was possible. As a proof of concept demonstration, we chose commercially available cytochrome c oxidase (CcO) solubilized in *n*-dodecyl-*β*-D-maltoside (DDM) as a model integral membrane protein. Initially, we exchanged DDM with IGEPAL using spin filtration. Next, we prepared the droplets by hydrating a **GOA** film in the presence of IGEPAL-solubilized CcO. When ferrocytochrome c was added to the droplets, we observed a steady decrease in absorbance at 550 nm indicating the enzyme-catalyzed formation of ferricytochrome c (Figure 6.10a). In comparison, when the activity was measured with IGEPAL-solubilized CcO only, a two-fold lower reaction rate was observed. The bilayer

membrane environment within the droplets likely allowed proper folding and membrane insertion of CcO in comparison to solubilization in IGEPAL only. In addition, it may be possible that the lipid environment of the droplets led to increased solubility of molecular O_2, the electron acceptor in the reaction, and hence an increased reaction rate. Finally, for direct observation, we labeled CcO with Alexa Fluor 488, and observed spontaneous partitioning into the droplets (Figure 6.10a).

Membrane proteins of the endomembrane system participate in a variety of transformations of lipid substrates. The ER contains the biosynthetic machinery for the production of phospholipids and steroids.[29] Comparably packed with membrane bilayer networks and with their ability to sequester membrane proteins and hydrophobic molecules, sponge droplets may offer an excellent environment for reactions involving lipid substrates. To test our hypothesis, we expressed bacterial (*E. coli*) diacylglycerol kinase (DAGK), a 13.2 kDa protein containing three transmembrane helices, in a PURE TX-TL reaction in the presence of droplets. DAGK catalyzes the phosphorylation of diacylglycerols and monoacylglycerols. The droplets were doped with 9.1 mol% 1-decanoyl-*rac*-glycerol, a substrate for DAGK, and additional ATP and $MgCl_2$ were added in the reaction mixture to support phosphorylation activity. After 3 h, the expected product decanoyl lysophosphatidate could be detected in nearly quantitative conversion as evidenced by the mass spectrometric signals (Figure 6.10b). To visualize the partitioning of the protein into the droplets, we expressed sfGFP-fused DAGK in the TX-TL system and observed fluorescence signal in the droplets indicating localization of the expressed protein (Figure 6.10b). Our results demonstrate that droplets can facilitate an important enzymatic step in the biosynthesis of phospholipids.

In addition, we also tested the membrane proteins adenosine A_{2A} receptor (Alexa Fluor 488-labeled) and proteorhodopsin and observed that they are excellently partitioned to the droplets.

In the future, lipid sponge droplets, with their high membrane density and large lipid-aqueous interface, could provide a platform for reconstituting more complex biochemical pathways involving multiple transmembrane proteins. We believe there is potential for the reconstitution of lipid biosynthesis pathways to emulate and characterize functions of the smooth ER.

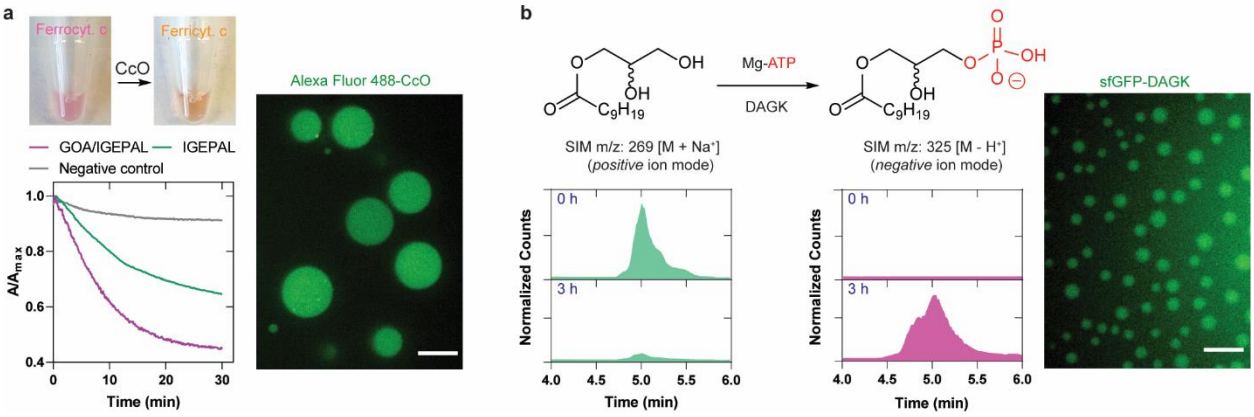


Figure 6.10. Reconstitution of functional transmembrane proteins in lipid sponge droplets. a. Cytochrome c oxidase (CcO) reaction. Addition of ferrocytochrome c to a droplet dispersion containing CcO leads to a gradual decrease in the absorbance at 550 nm indicating the formation of ferricytochrome c. Alexa Fluor 488-labeled CcO partitions into the droplets. Scale bar: 10 µm. b. *E. coli* diacylglycerol kinase (DAGK) reaction in droplets. DAGK is expressed in PURE system in the presence of droplets containing the substrate 1-decanoyl-*rac*-glycerol. Selective ion monitoring (SIM) chromatogram shows transformation to the phosphorylated product. Similarly expressed fluorescent fusion protein sfGFP-DAGK partitions into the droplets. Scale bar: 5 µm.

6.11 Rate enhancement of enzymatic reactions in the droplets

Living cells use organelles to co-localize molecules to facilitate reactions and minimize unwanted side reactions. Using compartmentalization to enhance biochemical reaction rates in living cells has been demonstrated by synthetic biologists and metabolic engineers. For instance, confinement of enzymatic pathways and key intermediates in mitochondria of engineered yeast has been shown to increase metabolite titer significantly as compared to cytoplasmic pathways.[30] Similar strategies have been employed to enhance the yield of fatty acid-derived metabolites in yeast peroxisomes.[31] We asked if colocalization of an enzyme and its substrate in droplets would enhance the rate of a biochemical reaction. First, we bound His_6-tagged cathepsin K, a lysosomal cysteine protease, within Ni-NTA-DGS doped droplets (Figure 6.11a). Next, we added a

fluorogenic dipeptide substrate benzyloxycarbonyl-L-leucyl-L-arginine 7-amido-4-methyl-coumarin (Z-LR-AMC), which gets sequestered within droplets by virtue of its partial hydrophobic nature (Figure 6.11a). We monitored the progress of the cathepsin K catalyzed hydrolysis reaction by measuring the steadily increasing fluorescence signal from the by-product 7-amino-4-methyl-coumarin (AMC) (Figure 6.11b). We observed that the progress of reaction was markedly higher when compared to the bulk phase reaction in the absence of droplets. We performed additional control reactions where we (i) omitted Ni-NTA-DGS from the droplets (non-binding droplets), and (ii) used IGEPAL only, and in each case, we found that the increase in AMC fluorescence signal was lower than the condition where both enzyme and substrate were co-localized in the droplets. We further corroborated and quantified these results by measuring the amount of AMC generated using HPLC-MS based on standard calibration curves. We estimated that after 6 hours, 84.7% of Z-LR-AMC was hydrolyzed in the binding droplets (Figure 6.11c). In the presence of non-binding droplets or IGEPAL alone, 25.7% and 18.6% of Z-LR-AMC was hydrolyzed respectively. Only 2.5% of Z-LR-AMC was hydrolyzed in the bulk condition. A likely explanation for the higher yield of product in the non-binding droplets and IGEPAL only conditions compared to the bulk without additives is that the amphiphiles facilitated dispersal of the substrate in the reaction media. Nonetheless, the product yield is enhanced greatly by programming the droplets to sequester cathepsin K and by concentrating and co-localizing enzyme and substrate. We believe that the droplets will offer a general scaffold for studying the effects of colocalization and confinement on enzymatic reactions.

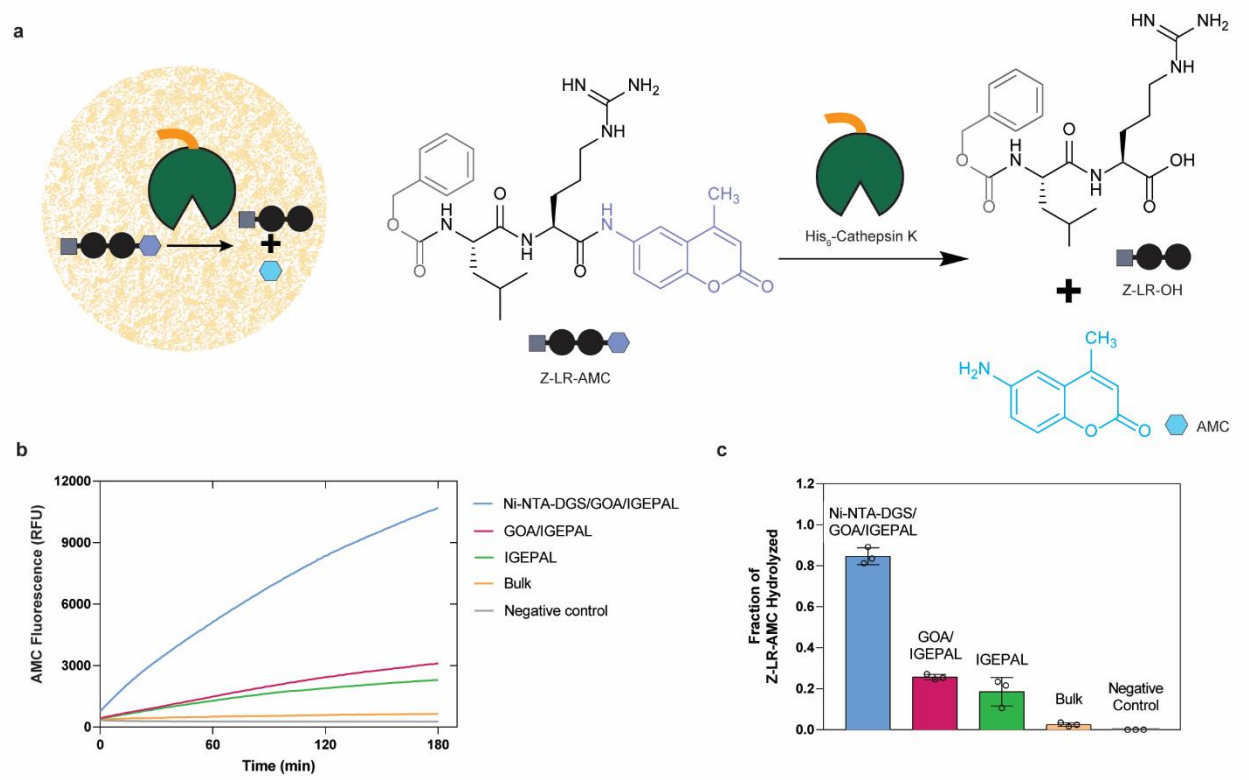


Figure 6.11. Rate enhancement of a model biochemical reaction due to co-localization of an enzyme and its substrate. a. Reaction scheme showing hydrolysis of a fluorogenic dipeptide Z-LR-AMC catalyzed by His_6-Cathepsin K. b. Progress of hydrolysis of Z-LR-AMC in droplets and corresponding controls followed by fluorometry. c. Quantification of the extent of hydrolysis of Z-LR-AMC in droplets and corresponding controls by HPLC-MS.

6.12 Light-controlled protein encapsulation

In living cells, proteins often shuttle dynamically between different subcellular locations in response to changes in the environment. For example, transcription factors shuttle between the nucleus and the cytoplasm, and cells dynamically alter the number of signaling receptors in the plasma membrane by endocytosis.[32] Furthermore, cells regulate the availability of signaling proteins by forming and dissolving membraneless compartments such as stress granules in response to environmental factors.[33] Sequestration of proteins into intracellular phase separated droplets has also been engineered to be responsive to light stimuli.[34] We aimed to engineer light-controlled and reversible protein sequestration into lipid sponge droplets. The photoreceptor phytochrome B (PhyB) from *Arabidopsis thaliana* switches between a binding and a non-binding conformation upon exposure to red (660 nm) and far-red (740 nm) wavelengths respectively. In

its binding conformation, PhyB binds to phytochrome interacting factors (PIFs) that can be fused to cargo proteins.[35] We used this system to control the capture and release of proteins by droplets with light. PhyB was produced with a biotin affinity handle, bound to Streptavidin and encapsulated in droplets doped with biotinyl-PE (Figure 6.12a). PhyB-loaded droplets encapsulated PIF-sfGFP-ssrA after exposure to daylight or 660 nm light. The protein was not enriched in droplets in 740 nm light or in the absence of PhyB (Figure 6.12b). Alternating exposure of droplets with red and far-red light reversibly switched the localization of PIF-sfGFP-ssrA between droplets and solution. Release and capture of the protein were rapid. Droplets lost 90% of GFP fluorescence within 2-3 min after switching illumination to 740 nm far-red light, and fluorescence signal from the solution in the vicinity of droplets increased immediately. When droplet samples were illuminated with 660 nm light, fluorescence signal from the solution decreased and fluorescence of droplets began to increase. Fluorescence of PIF-sfGFP-ssrA was observed to spread from the edges into the interior of droplets. In large droplets with diameters of more than 40 μm it took more than 30 min to reach a homogeneous intensity of GFP from droplet edge to interior (Figure 6.12b). The observed differences in the release and binding speeds can be explained by the differences in PIF-tagged protein concentration under the different illumination conditions in droplets and solution respectively. Sequestration into droplets is slower than release because the protein is at a relatively low concentration in solution, and sequestration into the droplets only occurs after an encounter of a PIF-tagged protein in solution and a droplet-bound PhyB. Light controlled localization changes were readily reversible and could be observed for more than 6 h. The observed fast release kinetics will be useful to rapidly make proteins available for reaction in solution.

6.13 Light-controlled sequestration protects proteins from degradation

Cells utilize a "nucleolar detention" strategy to regulate molecular networks by capturing and immobilizing essential cellular factors within the nucleolus away from their effector molecules.[36] We reasoned that light-controlled protein release and capture by droplets should allow us to control reactions in a similar manner. In contrast to the previous reactions that we engineered to occur within droplets, we focused on a reaction happening in solution, outside of droplets: ClpXP mediated protein degradation. ClpXP (from *E. coli*) is a large protein complex of approximately 600 kDa that consists of 6 ClpX and 14 ClpP subunits. ClpXP degrades ssrA-tagged protein substrates in a highly specific, processive and ATP-dependent manner and has a K_M of about 500 nM for ssrA-tagged sfGFP.[37] Since this protein complex has a cylindrical shape of 14.5 nm length and 11 nm diameter,[38] which are significantly larger than the average channel diameter of 4.1 nm, we expect it to be excluded from the droplets. Indeed, we observe that ClpP-msfGFP-His_6 is excluded from biotinyl-PE-doped droplets and cannot enter the droplets even when targeted inside using Ni-NTA-DGS (Figure 6.12a). Light-controlled changes in PIF-sfGFP-ssrA concentration in solution should therefore affect the rate of proteolytic degradation of the protein by ClpXP. As predicted, we found that in 660 nm illumination conditions degradation of PIF-sfGFP-ssrA was substantially slower than in 740 nm illumination conditions in the presence of PhyB-loaded droplets. We observed differences in degradation rate by fluorometric measurements of the reaction over time (Figure 6.12c). We further corroborated these data by microscopy of reaction samples, and immunoblotting. The light-dependent behavior allowed us to dynamically control the degradation rate by changing the illumination wavelength. Switching illumination between 660 nm and 740 nm led to pronounced and immediate changes in degradation rate (Figure 6.12c). Protection from ClpXP-mediated degradation required presence of PhyB-loaded droplets.

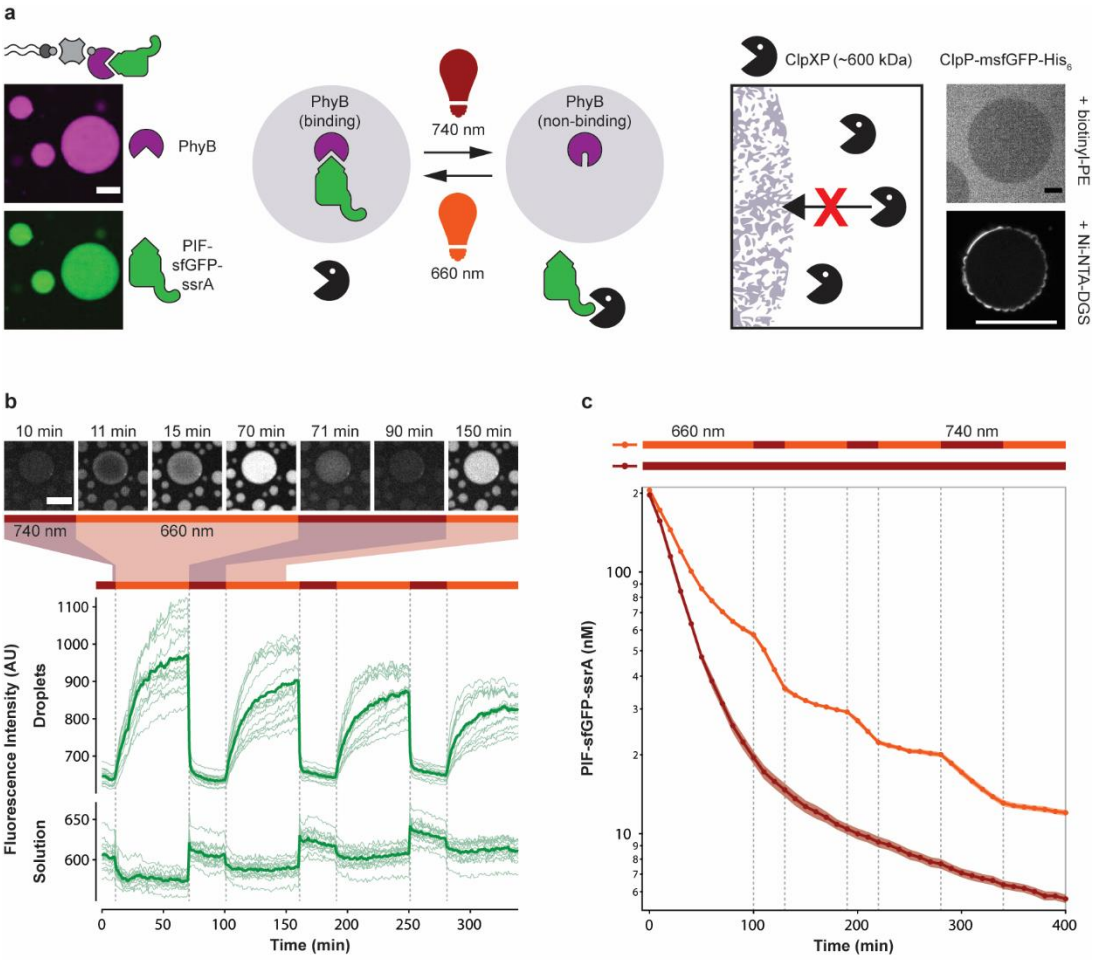


Figure 6.12. Light-dependent protein capture and release to dynamically control degradation rates. a. Design of the light-controlled protein degradation reaction. Left: Biotin-Streptavidin-mediated binding of light-responsive phytochrome B (PhyB) and capture of the protease substrate PIF-sfGFP-ssrA in 660 nm light. Right: Exclusion of ClpXP protease from lipid sponge droplets. ClpP-msfGFP-His_6 is excluded from droplets doped with biotinyl-PE and is unable to enter the droplets when targeted inside via Ni-NTA-DGS. b. Light-controlled release and capture of PIF-sfGFP-ssrA by PhyB-loaded droplets. Traces show fluorescence intensities of droplets and solution over time in response to changing illumination conditions. Bars above indicate 660 nm and 740 nm conditions. Intensities were extracted from a total of 15 droplets in 5 time-lapse movies (thin lines) and averaged (bold lines). Selected images from a time-lapse movie are shown on top. c. Light-controlled switching between protection and degradation of PIF-sfGFP-ssrA by ClpXP. Degradation of PIF-sfGFP-ssrA was monitored by reading the fluorescence of the reaction every 10 min. Reactions were performed in triplicates. Shown is the average (bold) and the standard deviation (shaded areas). Scale bars: 20 µm.

The Streptavidin-PhyB complex alone, without droplets, slowed degradation only marginally. Also, in the presence of droplets without PhyB, no protection or responses to wavelength changes were observed. In the absence of ClpXP, GFP signal decreased only marginally. As shown here for light-switchable protection of a sequestered protein from degradation in solution, light-

responsive droplets will enable precise spatiotemporal control of reactions both in solution and in the droplet phase.

6.14 Conclusions and outlook

Organelle-mimics. In summary, we describe the application of lipid sponge droplets as mimics of membrane-rich organelles and subcellular structures. With their nanoporous structure offering abundant hydrophobic-hydrophilic interfaces, the droplets spontaneously sequester molecules from the surrounding medium based on their physico-chemical properties. Further, the droplets are programmable in a modular fashion using amphiphiles with affinity handles that specifically target tagged proteins into the sponge phase at high concentrations. This is a significant advantage over vesicle-based artificial organelle models, which suffer from the limitation that they have no mechanism for efficient and spontaneous internal sequestration of macromolecules and solutes. Being composed of biocompatible building blocks, the lipid sponge droplets support a wide range of enzymatic reactions catalyzed by both transmembrane and soluble proteins and involving both water soluble and lipophilic substrates. Furthermore, we showed that a combination of colocalization, increased concentration, and the molecularly crowded environment can enhance enzymatic reactions in the droplets. The open and liquid nature of the droplets allows exchange with the surrounding environment, as well as reversible, light-responsive control over protein localization and concentration, which is difficult to engineer in a vesicle-based artificial organelle. Due to their versatile, self-assembled, and programmable nature lipid sponge droplets will find application as organelle-mimics to spatially organize biochemical reactions within artificial cellular systems and in cell-free systems.

Relevance to cell biologists. As of now, several aspects of cellular lipidic architecture remain poorly understood. For instance, what is the physical basis of vesicle production from

endoplasmic reticulum (ER) and how simple model systems can mimic such events? Also, what are the dynamical features of the interfaces and junctions in highly convoluted lipidic structures in cells? The endomembrane systems house a large diversity of membrane proteins for carrying out various metabolic transformations. A simple model system for studying how membrane proteins influence the phase behavior of lipids needs to be developed. Also, it needs to be understood if the proteins help to reinforce the relatively fragile lipid structures and maintain the sizes of interconnecting nanochannels. Sponge phase droplets could provide a simplified platform for studying diffusion and interaction of membrane proteins and their substrates in a meshwork of lipids. Interconnected lipid bilayers have often been reported in diseased or physiologically stressed cells. It is not clear if the underlying causes are biochemical or simply physical. Sponge droplets could be used to mimic such conditions and help us develop fundamental understanding of pathological states. ER is known to be site of processing and detoxification of various drug molecules. However, further understanding is needed regarding how drugs are partitioned to ER networks. Are there underlying biochemical mechanisms or do physical factors and lipid architecture play significant roles. We believe that simple model systems based on sponge phases and similar bicontinuous structures will help us develop better understanding of these matters.

Origins of life. Finally, alternative, non-vesicular lipid structures such as the one reported here could serve as additional models for protocellular evolution. Coacervate Theory was popularized by the likes of Alexander Oparin, J. B. S. Haldane, and Krishna Bahadur to describe possible origins of cellular life. It was suggested that coacervates derived from macromolecules were possibly progenitors of membrane-bound cells. Although this theory fell out of favor for some decades, there has been a renewed interest recently through artificial cell research. Fatty acid vesicles, which are popular protocell models, have been shown to transform to coacervates or other

bilayer-rich structures under certain conditions.[23,39] Interconversion or coexistence of such structures may have important implications in protocellular structural diversity, stability, and permeability and deserves further exploration.

6.15 Experimental methods

6.15.1 Droplet preparation

Typically, 9 µL of **GOA** from a 10 mM stock solution in methanol were added to 50 µL chloroform in a glass vial. When droplets were doped with amphiphiles bearing affinity handles, stock solutions were prepared in methanol and added into the chloroform as well. The organic solvents were evaporated under a slow stream of nitrogen gas and, while carefully rotating the glass vial, a thin film was generated at the bottom of the vial. The film was hydrated with a solution containing different amounts of IGEPAL (Sigma I8896, Lot#MKCC9036) within the droplet forming composition range (Figure 6.2a) as indicated for each experiment. The vial was vortex mixed until the lipid film was completely resuspended and the dispersion appeared turbid. To ensure complete hydration of the film, the sides and bottom of the vial are scratched with the pipette tip. As droplets form robustly at different **GOA** and IGEPAL concentrations and in different buffer compositions, the volume and composition of the rehydration solution was adjusted as indicated for each experiment.

6.15.2 Thermogravimetric analysis

A droplet dispersion was prepared in MilliQ H_2O from 3 mM **GOA** and 2.25 mM IGEPAL and allowed to coalesce into a single phase. The coalesced phase (2.351 mg) was carefully collected from the bottom of the tube and loaded onto a ceramic crucible. Thermogravimetric analysis was carried out on a Perkin-Elmer STA 6000 Simultaneous Thermal Analyzer. The

sample was heated at 10 °C/min over the range 30-350 °C. Nitrogen gas flow was maintained at 20 mL/min. We present the data over the range 30-150 °C.

6.15.3 Cryo-TEM

Sponge droplets were prepared from 3 mM **GOA** and 2.25 mM IGEPAL and sonicated with heating to ensure formation of small droplets. Immediately before grid preparation, the droplet sample was pipetted onto plasma-cleaned 300-mesh Quantifoil Multi-A carbon grids (Quantifoil). Using a Vitrobot EM grid plunger (FEI), excess buffer was blotted at room temperature and 95% humidity and the grids were plunge-frozen in liquid ethane maintained at about -180 °C. The grids were stored in liquid nitrogen until use. Droplet were imaged on a JEM-2100F microscope (JOEL) fitted with a OneView 4K charge-coupled device camera (Gatan).

6.15.4 Cryo-SEM

Approximately 3 µL of sample (dispersion of 3 mM **GOA** and 2.25 mM IGEPAL) was pipetted onto a 3 mm Al planchette and flash frozen in liquid ethane before loading into a Leica cryo-transfer shuttle for insertion into a Leica ACE600 cooled down to -130 °C. The sample was then freeze-fractured and freeze-etched for 6 minutes at -105 °C and coated with 7 nm of C/Pt and an additional 5 nm of C. The sample was then transferred to in a Hitachi S-4800 cFEG-SEM fitted with a Leica cryo-stage cooled down to -130 °C. Imaging occurred at 5 kV with an approximate working distance of 9.7 mm.

6.15.5 Small-angle X-ray scattering

For studying the sponge phase characteristics, we prepared a series of samples with controlled ratios between **GOA** and IGEPAL. The droplet dispersions were loaded in thin walled (0.01 mm) special glass capillary tubes of 15 mm outer diameter (15-SG, Charles Supper Company). Synchrotron X-ray scattering measurements were performed at the Advanced Photon

Source (APS), Argonne National Laboratory (ANL) at beamline 12-ID-C with monochromatic photon flux 1×10^{13} photons s^{-1} cm^{-2}. A focused beam of dimension 1×1 mm^2 with incident photon energy at 18 keV (wavelength = 0.69 Å) was employed with exposure times of 10 s to get a reasonable scattering signal from the sample. The SAXS data were collected by a CCD detector (2048×2048 pixels) and the sample-to-detector distance was 3825 mm. To obtain one-dimensional SAXS profiles, 2-dimensional scattering data were azimuthally averaged with proper background subtraction. Finally, the data were transformed into the profiles of scattering intensity as a function of scattering vector (q).

6.15.6 Sequestration of small molecules

All droplet dispersions for the sequestration experiments were prepared from 3 mM **GOA** and 2.25 mM IGEPAL such that the total mass of droplet forming amphiphiles was 80 µg. 5 µg of the molecule to be sequestered were added in each case. In case of sequestration studies of fluorescein, POPC, and cholesterol, 5 µL of 1 mg/mL solution in EtOH were added during **GOA** film formation. 20 µL of the cloudy droplet dispersion was collected in a 0.6 mL Eppendorf tube and placed in an incubator at 37 ^{o}C for 2 h. The tubes were centrifuged at 14,000 rcf for 2 min to ensure complete coalescence of the droplets. 2 µL of the supernatant was collected carefully and diluted with 98 µL 1:1 MeOH: H_2O. This was injected as whole into HPLC-MS and the peak area corresponding to the analyte was measured (S). The coalesced phase was resuspended by vigorous vortexing and 2 µL of the dispersion was diluted with 98 µL 1:1 MeOH: H_2O. This was injected as whole into HPLC-MS and the peak area corresponding to the analyte in consideration was measured (T). The sequestered fraction was calculated from the following equation:

$$Fraction\ Sequestered = 1 - S \div [S \times 0.1 + T \times 0.9] \quad (6.1)$$

6.15.7 Sequestration of LecA (PA-IL)

Fluorescent labeling of LecA. Terminal galactose-binding lectin LecA (or PA-IL) from *Pseudomonas aeruginosa* was purchased from Sigma-Aldrich. 0.31 mg of the solid powder (~80% protein) was dissolved in 200 µL 100 mM $NaHCO_3$, pH 8.3. 3.3 equivalents of Alexa Fluor 488 NHS ester (dissolved in anhydrous DMSO) was added to this solution, and the mixture was stirred at room temperature for 2 h. The excess dye was removed by a commercial dye removal kit (Pierce, Thermo Fisher Scientific). The eluent was further washed and exchanged with 1X PBS using a 10 kDa MWCO spin filter and stored at -20 ^oC as 10% glycerol stock containing 0.02% (*w*/*v*) NaN_3.

Microscopic observations. Droplets were prepared with 3 mM **GOA** and 1.33-2.25 mM IGEPAL. Sequestration of various concentrations of Alexa Fluor 488 labeled LecA (1.0 µM-0.25 nM) was studied by spinning disk confocal microscopy.

6.15.8 Reconstitution of CcO in droplets

Reduction of ferricytochrome c. 1.72 mg of reddish-brown powder of ferricytochrome c from bovine heart (Sigma) were dissolved in 190 µL PBS (1X). Then, 10 µL of sodium ascorbate (200 mM) were added. The mixture was kept at 4 ^oC overnight, observing that its color changed to rose-red. The residual ascorbate was removed by dialysis (mini dialysis device, 10 kDa MWCO, Slide-A-Lyzer^{TM}) against 10 mM sodium phosphate buffer (pH 7.0) followed by spin filtration using a 10 kDa MWCO spin filter with 6X volume. The final concentration of ferrocytochrome c was determined to be 1.11 mM and the A_{550}/A_{565} ratio was 17.43.

Detergent exchange of CcO. 5 µL of CcO from bovine heart (~5 mg/mL in 39 mM DDM, 25 mM Tris pH 7.8, 5 mM EDTA) was added to 197.5 µL (in 10 mM sodium phosphate buffer, pH 7.0) of 1.68 mM IGEPAL. 2.5 µL of this solution was dissolved in 97.5 µL MeOH and mixed by vigorous vortexing. The tube was centrifuged at maximum speed to remove any insoluble

material. The supernatant was carefully collected and run on HPLC-MS and the peaks corresponding to DDM and IGEPAL were seen distinctly. The remaining solution was transferred to a 100 kDa MWCO spin filter and thoroughly washed with (5×200 µL) of 1.68 mM IGEPAL (in 10 mM sodium phosphate buffer, pH 7.0). 1 µL of the final retentate (~35 µL) was added to 99 µL MeOH and run on HPLC-MS as described. The analysis verified that the DDM present in the commercially available CcO was completely exchanged with IGEPAL. Based on a calibration curve, the concentration of IGEPAL was estimated to be 18.8 mM.

Activity assay of CcO. A lipid film from 39 µL of 10 mM **GOA** was hydrated with a solution containing 377 µL sodium phosphate (10 mM, pH 7.0), 6 µL CcO (in 18.8 mM IGEPAL and 10 mM sodium phosphate pH 7.0), and 7 µL of IGEPAL (16.83 mM in 10 mM sodium phosphate pH 7.0). The turbid droplet dispersion was transferred to a cuvette and 15 µL of ferrocytochrome c was added. The dispersion was mixed by tapping and pipetting and placed in Nanodrop 2000c. Absorbance at 550 nm was measured every 5 s over 30 min and a steady decrease was observed. To compare the activity of CcO at the same concentration of IGEPAL in the absence of droplets, a sample was prepared in an identical manner and the activity was assayed. A negative control sample was prepared and assayed where no CcO was added. The A_{550} was recorded identically and a negligible decrease was observed.

Fluorescent labeling of CcO. 60 µg of commercially available CcO from bovine heart (Sigma) was exchanged with sodium bicarbonate buffer (100 mM, pH 8.3) containing 0.4 mM *n*-dodecyl-*β*-maltoside (DDM). Then, the solution was treated with 10 equivalents of Alexa Fluor 488 NHS ester (dissolved in anhydrous DMSO), and the mixture was stirred at room temperature for 2 h. Then, the excess dye was removed and DDM was exchanged with IGEPAL by spin filtration (100 kDa MWCO).

6.15.9 Activity of cathepsin K in droplets

General. Recombinant His_6-tagged human cathepsin K was purchased from EMD Millipore. The enzyme was stored as 18.5 µM aliquots in 100 mM NaOAc/HOAc buffer (pH 5.5) containing 1 mM TCEP at -80 °C. The protein was labeled with Alexa Fluor 488 and microscopy verified its sequestration into droplets doped with Ni-NTA-DGS. Benzyloxycarbonyl-L-leucyl-L-arginine 7-amido-4-methylcoumarin (Z-LR-AMC) was purchased from Peptides International and stored as 10.8 mM DMSO solution at -80 °C. It was estimated from HPLC that 75% of Z-LR-AMC gets partitioned to the droplets used in the subsequent experiments.

Fluorometric assay. Ni-NTA-DGS/**GOA**/IGEPAL (26.7 µM/0.96 mM/0.64 mM) droplets were prepared in 100 mM HEPES-Na containing 1.5 mM TCEP. His_6-cathepsin K was added to the droplet dispersion so that the enzyme concentration with respect to bulk was 2.3 nM. After incubating at room temperature for 10 min, Z-LR-AMC (dissolved in DMSO) was added such that the final concentration was 81 µM. Additional conditions were tested where one or more of the components were omitted: **GOA**/IGEPAL (0.96 mM/0.64 mM), IGEPAL (0.64 mM), bulk, and negative control (no amphiphiles and no enzyme). 20 µL of each sample was loaded onto a 384 well-plate and analyzed for fluorescence from 7-amino-4-methylcoumarin (λ_{ex}: 360 nm, λ_{em}: 460 nm; 10 nm bandwidth) every 1 min at room temperature in a Tecan Spark microplate reader.

HPLC-MS assay. Calibration curves were prepared for Z-LR-AMC and AMC with known quantities of the pure substances. Ni-NTA-DGS/**GOA**/IGEPAL (26.7 µM/0.96 mM/0.64 mM) droplets were prepared in 100 mM HEPES-Na containing 2.0 mM TCEP. His_6-cathepsin K was added to the droplet dispersion to a bulk concentration of 2.8 nM. Z-LR-AMC was added to a final concentration of 81 µM. **GOA**/IGEPAL (0.96 mM/0.64 mM), IGEPAL (0.64 mM), bulk, and negative control (no amphiphiles and no enzyme) samples were also prepared in identical buffer

and same quantity of enzyme (except for the negative control) and substrate were used. The reactions (20 µL) were incubated at room temperature for 6 h and quenched by addition of 80 µL of 1:1 MeOH:H_2O. The reactions were analyzed by HPLC-MS and the areas of the peaks under 330 nm chromatogram were used for calculation of yields. The experiments were repeated in triplicate.

6.15.10 FRAP experiments

For the analysis of lipid diffusion, droplets were formed with 3.6 mM **GOA**, 2.56 mM IGEPAL and 36 µM Texas Red DHPE in 1X PBS. For the analysis of protein diffusion, droplets were formed with 3 mM **GOA**, 2.25 mM IGEPAL and 83 µM Ni-NTA-DGS in 100 mM HEPES pH 8. Purified mCherry-His_6 protein (Antibodies-online Inc.) was added to a final concentration of 10 µM. As controls, we performed FRAP experiments on mCherry-His_6 bound to Ni-NTA agarose resin (HisPur, Thermo Fisher Scientific) and a water-in-oil emulsion of 90 µM mCherry-His_6 in mineral oil containing 0.25% (*v*/*v*) Span-80. FRAP experiments were performed with a Zeiss Spinning Disk microscope using a 63x 1.4 NA oil objective. A circular region of interest (ROI) with a diameter of 9.6 µm was bleached in large droplets (>50 µm) with a 5 s pulse of the 561 nm laser at 100% intensity. Samples were imaged every 500 ms with a 150 ms exposure at 50% intensity. Movies were analyzed in Fiji/ImageJ to extract average fluorescence over time of the bleached ROI, entire droplet and background. Background corrected fluorescence values were used to correct for possible photobleaching

$$f(t) = \frac{F_{ROI}(t)}{F_{Whole}(t)} \tag{6.2}$$

and then scaled between 0 and 1

$$F(t) = \frac{f(t)-f(0)}{f(-10s)-f(0)} \tag{6.3}$$

where f(0) is the intensity at time 0s, right after the bleach, and f(-10s) is the initial prebleach intensity.[40] Data were then fitted to

$$F(t) = A(1 - e^{-\frac{t}{\tau}}) \tag{6.4}$$

where A is the amplitude of recovery and τ is the time constant that allowed us to calculate the half-time to recovery $t_{1/2}$.

$$t_{1/2} = ln2 \cdot \tau \tag{6.5}$$

6.15.11 TX-TL in the presence of droplets

A thin film was prepared from 18 µL 10 mM **GOA** and rehydrated by vortexing with 9 µL rehydration solution (6 µL PURExpress Solution A, 0.375 µL Murine RNase Inhibitor (NEB), 1.5 µL 67.3 mM IGEPAL, 1.125 µL H_2O). For the reaction to start, 4.5 µL PURExpress Solution B and linear DNA templates were added for a final volume of 15 µL. DNA template concentrations in the TX-TL reaction were 5 nM for P_{T7}-sfGFP-lecA and P_{T7}-sfGFP, and 10 nM for P_{T7}-sfGFP-lacI. Linear DNA templates were assembled by PCR. Expression reactions were carried out at 33 °C, in droplets of 5 µL volume sandwiched between cover glass and gas permeable Lumox plates (Sarstedt) or in chambers of a microfluidic device, where fluorescence increase of individual droplets could be followed by time lapse imaging. For expression of DAGK, a lipid film was prepared from 15 µL of 10 mM **GOA** and 3 µL of 10 mM 1-decanoyl-*rac*-glycerol (Cayman Chemicals). Hydration was carried out with 30 µL of a solution containing 5 mM IGEPAL, 0.5 mM ATP, and 0.5 mM $MgCl_2$. Protein expression reaction was consisted of 8 µL PURExpress Solution A, 6 µL PURExpress Solution B, 0.5 µL Murine RNase Inhibitor (NEB), 1 µL DAGK linear DNA (28.8 nM final concentration), 1 µL H_2O, and 4 µL droplet dispersion (containing the substrates for DAGK). 10.25 µL of the reaction is taken out immediately and quenched with 4 µL

Na_2-EDTA (100 mM) and 90 µL methanol. The remainder is incubated for 3 h at 37 °C and quenched similarly. The initial and final time points were analyzed by HPLC-MS in selective ion mode for *m*/*z* 269 (positive ion), and *m*/*z* 325 (negative ion). For the expression of sfGFP-DAGK, 13.4 nM of DNA template was used.

6.15.12 Protein expression, purification and analysis

Proteins were expressed in *E. coli* BL21 (DE3) strain. ClpX-His_6 and ClpP-His_6 were gifts from Xingying Shi and Simpson Joseph (UC San Diego).[37] Biotinylated PhyB(1-651)-AviTag with phycocyanobilin was produced in *E. coli*.[35] Proteins were purified by Ni-NTA affinity chromatography. Purified proteins were buffer exchanged to storage buffer using spin filter devices or dialysis. Storage conditions were 200 mM potassium phosphate pH 7.2, 2 mM EDTA, 400 mM NaCl, 1 mM DTT, 1 mM NaN_3, 50% (*v*/*v*) glycerol at -20 °C for sfGFP-LacI-His_6, 1X PBS at -80 °C for sfGFP-His_6, 100 mM HEPES pH 8.0, 25 mM NaCl, 0.5 mM TCEP, 50% (*v*/*v*) glycerol at -20 °C for PhyB-AviTag, 100 mM HEPES pH 8.0, 25 mM NaCl, 5 mM DTT, 50% (*v*/*v*) glycerol at -20 °C for His_6-PIF-sfGFP-ssrA, and 50 mM Tris pH 7.5, 200 mM KCl, 25 mM $MgCl_2$, 0.67 mM IGEPAL, 0.5 mM TCEP, 20% (*v*/*v*) glycerol at -80 °C for ClpP-msfGFP-His_6. Protein concentrations were measured by BCA assay. PhyB-AviTag was verified by Zinc staining to visualize the bound chromophore phycocyanobilin by incubating an SDS-PAGE gel in 1 mM zinc acetate for 15 min followed by visualization under UV light. Photoresponsive switching was verified by recording absorbance spectra of purified PhyB protein after exposure to 660 nm or 740 nm light using a Tecan Spark microplate reader. For immunoblotting, protein samples were separated by SDS-PAGE and transferred to PVDF membranes. Antibodies used were GFP Antibody (B-2) (Santa Cruz Biotechnology) at 1:200 and 6x-His Tag Monoclonal Antibody (Invitrogen) at 1:1000. Bands were detected with a horseradish peroxidase coupled secondary

antibody and SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific).

6.15.13 Light-control of protein localization

Samples were illuminated using LEDs (Roithner LaserTechnik GmbH, LED660N-03 and LED740-series) in light insulated containers, where LEDs were placed in proximity to samples using a custom-built setup. In 740 nm light conditions, samples were continuously illuminated by 740 nm light. In 660 nm light conditions, samples were illuminated every 10 min with a 30 s light pulse and otherwise incubated in the dark.

6.15.14 ClpXP reactions

For light controlled protein degradation, droplets for one ClpXP reaction were prepared as follows, where the volumes were scaled according to the reactions performed in one experiment. Per reaction, a lipid film was formed from 7.2 µL 10 mM **GOA** and 0.8 µL 2.5 mM 18:1 Biotinyl-PE (Avanti). The lipid film was vortexed in 40 µL rehydration solution (1.35 mM IGEPAL in 100 mM HEPES pH 8.0) to form droplets. 40 µL of 2 µM Streptavidin and 2 µM PhyB-AviTag in reaction buffer [50 mM HEPES pH 8.0, 20 mM KCl, 5 mM $MgCl_2$, 1% (*v*/*v*) glycerol, 1% (*w*/*v*) BSA, 5 mM β-ME, 0.067 mM IGEPAL] were added and incubated with droplets for 15 min. Unbound PhyB and Streptavidin were then removed by washing with reaction buffer by pelleting droplets by centrifugation at 1500 rcf. During washing, droplets were concentrated in 10 µL reaction buffer, and the substrate PIF-sfGFP-ssrA was added for a final concentration of 250 nM in the reaction for binding to PhyB-loaded droplets and incubated for 15 min. The reaction was started by adding premixed ClpXP and ATP in reaction buffer (concentrations were 4 mM ATP, 50 nM $ClpX_6$, 200 nM $ClpP_{14}$ in a final reaction volume of 20 µL), and droplets were coalesced in the bottom of the reaction plate (384-well v-bottom plates, Promega) by centrifugation. The

reaction was performed at room temperature, monitored by reading fluorescence of PIF-sfGFP-ssrA in a Tecan Spark microplate reader (excitation 485 +/- 20 nm (filter), emission 520 nm +/- 20 (monochromator)) every ten minutes and placed in specific illumination conditions between reads. For analysis by western blot, samples were collected at different time points, quenched by adding Laemmli buffer and stored at -20 °C. For analysis by microscopy, samples were first placed in 740 nm from 5 min and then in 660 nm light for 30 min to compare droplet fluorescence.

Notes on chapter

Chapter 6 is adapted majorly from materials which appear in Bhattacharya, A., Niederholtmeyer, H., Podolsky, K. A., R. J., Bhattacharya, R., Song, J., Brea, R. J., Hsien-Tsai, C., Sinha, S. K. & Devaraj, N. K. Lipid Sponge Droplets as Programmable Synthetic Organelles, *bioRxiv*, 2020, DOI: https://doi.org/10.1101/2020.01.20.913350. This manuscript is currently submitted to *PNAS*. The dissertation author is the co-first author in this work. This chapter also contains small amount of unpublished data on lectin-induced droplet structuration coauthored with Devaraj, N. K.

We thank Soenke Seifert (Advanced Photon Source, Argonne National Laboratory) for guidance and valuable suggestions with the Synchrotron SAXS experiments, Simpson Joseph and Xinying Shi (UCSD) for ClpX and ClpP proteins, Sunahara Lab for adenosine A_{2A} receptor Maximilian Hörner and Wilfried Weber (University of Freiburg) for PhyB and PIF-tag plasmids, Florian Stenger and Axel Voigt (TU Dresden) for 3D models of a bicontinuous structure, and Veronica Falconieri Hays for illustration of the lipid sponge phase. TEM work was performed at the UC Irvine Materials Research Institute (IMRI). We acknowledge the use of the UCSD Cryo-Electron Microscopy Facility, which is supported by NIH grants to Timothy S. Baker and a gift from the Agouron Institute to UCSD. This work made use of the BioCryo facility of Northwestern

University's NU*ANCE* Center, which has received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205); the MRSEC program (NSF DMR-1720139) at the Materials Research Center; the International Institute for Nanotechnology (IIN); and the State of Illinois, through the IIN. It also made use of the CryoCluster equipment, which has received support from the MRI program (NSF DMR-1229693).

6.16 References

1. Simmel, F. C. Synthetic organelles. *Emerg. Top. Life Sci.* ETLS20190056 (2019).

2. Feng, X., Jia, Y., Cai, P., Fei, J. & Li, J. Coassembly of photosystem II and ATPase as Artificial chloroplast for light-driven ATP synthesis. *ACS Nano* **10,** 556–561 (2016).

3. Lee, K. Y., Park, S., Lee, K. H., Kim, S., Kim, H., Meroz, Y., Mahadevan, L., Jung, K., Ahn, T. K., Parker, K. K. & Shin, K. Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nat. Biotechnol.* **36,** 530–535 (2018).

4. Deng, N.-N., Yelleswarapu, M., Zheng, L. & Huck, W. T. S. Microfluidic Assembly of Monodisperse Vesosomes as Artificial Cell Models. *J. Am. Chem. Soc.* **139,** 587–590 (2017).

5. Milo, R., Phillips, R. & Orme, N. *Cell Biology by the Numbers* (Garland Science, New York, 2015).

6. Poudyal, R. R., Guth-Metzler, R. M., Veenis, A. J., Frankel, E. A., Keating, C. D. & Bevilacqua, P. C. Template-directed RNA polymerization and enhanced ribozyme catalysis inside membraneless compartments formed by coacervates. *Nat. Commun.* **10,** 490 (2019).

7. Sokolova, E., Spruijt, E., Hansen, M. M. K., Dubuc, E., Groen, J., Chokkalingam, V., Piruska, A., Heus, H. A. & Huck, W. T. S. Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proc. Natl. Acad. Sci.* **110,** 11692–11697 (2013).

8. Koga, S., Williams, D. S., Perriman, A. W. & Mann, S. Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat. Chem.* **3,** 720–724 (2011).

9. Conn, C. E. & Drummond, C. J. Nanostructured bicontinuous cubic lipid self-assembly materials as matrices for protein encapsulation. *Soft Matter* **9,** 3449–3464 (2013).

10. Caffrey, M. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Cryst.* **71,** 3–18 (2015).

11. Strey, R., Schomacker, R., Roux, D., Nallet, F. & Olsson, U. Dilute lamellar and L_3 phases in the binary water-C12E5 system. *J. Chem. Soc. Faraday Trans.* **86,** 2253–2261 (1990).

12. Strey, R., Jahn, W., Porte, G. & Bassereau, P. Freeze fracture electron microscopy of dilute lamellar and anomalous isotropic (L_3) phases. *Langmuir* **6,** 1635–1639 (1990).

13. Menger, F. M. & Sykes, B. M. Anatomy of a Coacervate. *Langmuir* **14,** 4131–4137 (1998).

14. Peresypkin, A. V. & Menger, F. M. Zwitterionic Geminis. Coacervate Formation from a Single Organic Compound. *Org. Lett.* **1,** 1347–1350 (1999).

15. Imura, T., Yanagishita, H. & Kitamoto, D. Coacervate formation from natural glycolipid: One acetyl group on the headgroup triggers coacervate-to-vesicle transition. *J. Am. Chem. Soc.* **126,** 10804–10805 (2004).

16. Angelov, B., Angelova, A., Mutafchieva, R., Lesieur, S., Vainio, U., Garamus, V., Jensen, G. V. & Pedersen, J. S. SAXS investigation of a cubic to a sponge (L_3) phase transition in self-assembled lipid nanocarriers. *Phys. Chem. Chem. Phys.* **13**, 3073–3081 (2011).

17. Pabst, G., Rappolt, M., Amenitsch, H. & Laggner, P. Structural information from multilamellar liposomes at full hydration: Full q-range fitting with high quality X-ray data. *Phys. Rev. E* **62,** 4000–4009 (2000).

18. Schaefer, D. W. Polymers, fractals, and ceramic materials. *Science* **243,** 1023–1027 (1989).

19. Maldonado, A., Urbach, W., Ober, R. & Langevin, D. Swelling behavior and local topology of an L_3 (sponge) phase . *Phys. Rev. E* **54,** 1774–1778 (1996).

20. Bhansali, S. H., Malik, S. H., Jarvis, S. M., Akartuna, I., Dabbs, D. M., Carbeck, J. M. & Aksay I. K. The stability of L_3 sponge phase in acidic solutions. *Langmuir* **22,** 4060–4064 (2006).

21. Valldeperas, M., Dabkowska, A. P., Pálsson, G. K., Rogers, S., Mahmoudi, N., Carnerup, A., Barauskas, J. & Nylander, T. Interfacial properties of lipid sponge-like nanoparticles and the role of stabilizer on particle structure and surface interactions. *Soft Matter* **15,** 2178–2189 (2019).

22. Porte, G. Lamellar phases and disordered phases of fluid bilayer membranes. *J. Phys.* **4,** 8649–8670 (1992).

23. Garenne, D., Beven, L., Navailles, L., Nallet, F., Dufourc, E. J. & Douliez, J-P. Sequestration of Proteins by Fatty Acid Coacervates for Their Encapsulation within Vesicles. *Angew. Chem. Int. Ed.* **55,** 13475–13479 (2016).

24. Winzer, K., Falconer, C., Garber, N. C., Diggle, S. P., Camara, M. & Williams, P. The *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL Are Controlled by Quorum Sensing

and by RpoS. *J. Bacteriol.* **182,** 6401–6411 (2000).

25. Johannes, L., Jacob, R. & Leffler, H. Galectins at a glance. *J. Cell Sci.* **131,** jcs208884 (2018).

26. Daber, R., Stayrook, S., Rosenberg, A. & Lewis, M. Structural Analysis of Lac Repressor Bound to Allosteric Effectors. *J. Mol. Biol.* **370,** 609–619 (2007).

27. Shimizu, Y., Inoue, A., Tomari, A., Suzuki, T., Yokogawa, T., Nishikawa, K. & Ueda, T. Cell-free translation reconstituted with purified components. *Nat. Biotechnol*. **19,** 751–755 (2001).

28. Dahms, N. M., Olson, L. J. & Kim, J. J. P. Strategies for carbohydrate recognition by the mannose 6-phosphate receptors. *Glycobiology* **18,** 664–678 (2008).

29. Schwarz, D. S. & Blower, M. D. The endoplasmic reticulum: Structure, function and response to cellular signaling. *Cell. Mol. Life Sci.* **73,** 79–94 (2016).

30. Avalos, J. L., Fink, G. R. & Stephanopoulos, G. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **31,** 335–341 (2013).

31. Zhou, Y. J., Buijs, N. A., Zhu, Z., Gómez, D. O., Boonsombuti, A., Siewers, V., & Nielsen, J. Harnessing Yeast Peroxisomes for Biosynthesis of Fatty-Acid-Derived Biofuels and Chemicals with Relieved Side-Pathway Competition. *J. Am. Chem. Soc.* **138,** 15368–15377 (2016).

32. Itzhak, D. N., Tyanova, S., Cox, J. & Borner, G. H. H. Global, quantitative and dynamic mapping of protein subcellular localization. *Elife* **5,** 1–36 (2016).

33. Hyman, A. A., Weber, C. A. t& Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **30,** 39–58 (2014).

34. Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E., & Brangwynne, C. P. Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168,** 159–171 (2017).

35. Beyer, H. M., Thomas, O. S., Riegel, N., Zurbriggen, M. D., Weber, W. & Hörner, M. Generic and reversible opto-trapping of biomolecules. *Acta Biomater.* **79,** 276–282 (2018).

36. Audas, T. E., Jacob, M. D. & Lee, S. The nucleolar detention pathway: A cellular strategy for regulating molecular networks. *Cell Cycle* **11,** 2059–2062 (2012).

37. Shi, X., Wu, T., Cole, C. M., Devaraj, N. K. & Joseph, S. Optimization of ClpXP activity and protein synthesis in an *E. coli* extract-based cell-free expression system. *Sci. Rep.* **8,** 3488 (2018).

38. Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S. A. & Raunser, S. Cryo-EM structure of the ClpXP protein degradation machinery. *Nat. Struct. Mol. Biol.* **26,** 946–954 (2019).

39. Cornell, C. E., Black, R. A., Xue, M., Litz, H. E., Ramsay, A., Gordon, M., Mileant, A., Cohen, Z. R., Williams, J. A., Lee, K. K., Drobny, G. P. & Keller, S. Prebiotic amino acids bind to and stabilize prebiotic fatty acid membranes. *Proc. Natl. Acad. Sci.* **116,** 17239–17244 (2019).

40. Kang, M., Andreani, M. & Kenworthy, A. K. Validation of normalizations, scaling, and photofading corrections for FRAP data analysis. *PLoS One* **10,** e0127966 (2015).

Chapter 7 | Outlook and Future Directions

This dissertation addresses a few subjective problems in the artificial cells area which are extremely broad in scope. Our focus was to understand the fundamental role of lipid assemblies in the basic design plan of cells. We offered a few plausible 'solutions' or 'partial solutions' to those problems. We emphasize that many alternate routes to those problems may be developed given their open-ended nature. Here we provide a brief summary of the key achievements, the existing limitations, and possible future directions.

7.1 Summary of results

Lipid environments facilitate chemical reactions through physical catalysis. We have shown through several examples that lipid assemblies partition amphiphilic molecules through physical interactions and therefore increases their effective concentrations by several folds of magnitude. Moreover, through this partitioning mechanism, 'unwanted' reactions with non-amphiphilic species are minimized. As a result, even the reactions which are inherently sluggish or are prone to side-reactions (such as hydrolysis) can be rendered highly efficient in lipid environments. We discussed this phenomenon with the examples of histidine ligation in Chapter 2, adenylate chemistry and Native Chemical Ligation in Chapter 3. In all these examples, the corresponding reactions between amphiphilic partners are rendered efficient by many folds due to the concentrating mechanism of lipid assemblies and protection from non-amphiphilic competing groups. In principle, any simple, robust, and chemoselective coupling reaction that takes place under mild aqueous condition could be used to generate phospholipids *in situ*. This opens the opportunity to build a diverse toolbox for generating membrane-forming lipids chemically. One limitation that needs to be addressed is that in the methods described, we formed an amide bond

which does not occur in natural glycerophospholipids. So, future efforts will be necessary to develop chemoselective non-enzymatic strategies to generate ester bond.

Membrane-forming lipids can be synthesized via biochemical pathways significantly different from the existing pathways. In Chapter 3 we discussed two minimal biochemical approaches to synthesize phospholipids. In both approaches we substituted an enzymatic acyl transfer step with a non-enzymatic ligation step. The chemoselectivity of the ligation reactions allowed us to achieve high specificity even in a complex biochemical medium such as PURE system. Analogous minimal routes may be developed using a ribozyme to carry out the fatty acid activation step, although selecting a ribozyme with such activity may be technically challenging. Such efforts will help bridging important gaps in the understanding of the evolution of membrane synthesis machinery. Our current approaches still rely on amphiphilic precursors, which are not ideal to use for demonstrating growth and division events as they themselves can have non-specific effects (even disruptive) on the membranes. Future efforts will be necessary to carry out phospholipid synthesis entirely from non-amphiphilic precursors. A plausible route is to employ fatty acid synthases (FAS) to generate a fatty acid or fatty acyl CoA *in situ* which can thereafter react with a functionalized non-amphiphilic headgroup to form a two-chain lipid. we envision that strategies of *de novo* lipid formation can potentially be applied to novel applications in living cells. Lipids remain among the less understood and explored classes of biomolecules. There are multiple disease conditions where aberrations in membrane fluidity, lipid metabolism and signaling pathways are implicated. It will be particularly interesting to explore if site-specific or stimuli-responsive lipid synthesis and/or remodeling in cells can ameliorate such conditions.

Membrane-forming lipids can be designed with unique structural scaffolds. The discovery of bilayer membrane-forming single-chain galactolipids described in Chapters 4 and 5 was a

serendipitous one. We synthesized a series of glycolipids with similar structural design and identified the governing 'rules' for membrane formation. In general, we observed that the glycolipids offer an enormous diversity of supramolecular aggregates even with tiny variations in chemical structures. The physical characterization studies described in these chapters may be useful for studying other novel amphiphilic systems in the future. We further demonstrated that these lipids are compatible with biochemical reactions and will find more applications in artificial cell construction.

Lipid sponges offer a simple chassis for programmable encapsulation of biomolecules. Study of bicontinuous sponge phases formed from single-chain galactolipids and non-ionic detergents as described in Chapter 6 was another outcome of a serendipitous discovery. We show that the lipid sponges behave in many ways as other liquid-liquid phase separated systems or coacervates in terms of sequestering molecules at high concentrations. However, unlike complex coacervates derived from oppositely charged polymers, lipid sponges offer more well-defined sequestration properties and programmability. Particularly, the ability of lipid sponges to host functional membrane proteins offer a huge benefit to reconstitute diverse biochemical pathways within them. A few shortcomings do exist with the lipid sponges. For instance, in case of membrane proteins (for example, ion channels, ATP synthase) which require a specific directional orientation for their proper function, cannot work in a desired way as the sponge phase do not have an 'inside' and 'out' property. Further, because of the porous structure of the lipid sponges, nucleases and proteases can destroy the inner macromolecular content if their sizes are smaller than the aqueous channel sizes. In spite of this disadvantage, lipid sponges can offer some novel insight into protocellular evolution and an interesting parallel can be drawn with marine sponges. Marine sponges are organisms with one of the most primitive and open body plans, yet they are

surviving for hundreds of millions of years on Earth. Lipid sponges similarly have an open structure, freely exchanging matter with the surroundings. This can be thought of as a disadvantage in some ways, especially considering that modern cells are bound by membrane with strict permeability regulations. However, this property of 'openness' may have proved beneficial in the early stages of evolution of protocellular life. Lipid sponges may have continuously sequestered molecules from a dilute primordial soup and enabled the establishment of reaction networks. They offer a structural plan where sophisticated mechanisms of molecular transport were not necessary, especially those necessary for transporting charged molecules. In modern cells, lipidic structures having bicontinuous features have been observed in endoplasmic reticulum and inner mitochondrial membranes and their role is not fully understood. The question remains whether those structures are evolutionary relics and their presence hint us at a deeper fundamental significance.

7.2 Open questions

There are several open questions that need to be addressed too. For instance, how can shape and size control be achieved simultaneously in an artificial cell undergoing growth and division? Although several artificial systems displaying growth and division-like events have been described, those proliferation events mostly take place non-deterministically and may be considered highly rudimentary compared to their biological counterparts. Proliferation of an artificial cellular container into two identical copies of itself is still an unachieved goal. We surmise that approaches radically different from those prevalent in nature may need to be considered to further this goal. Recently, the topic of division of active droplets into identical daughter droplets has been treated theoretically and it will be interesting to translate those ideas experimentally. Another long-sought goal in artificial cell research is generation of Turing instabilities, which

coupled to artificial cellular membrane growth could lead to controlled deformation of membranes and eventual division. We envision that an integrative approach for building artificial life will take us to newer avenues in science and technology and provide answers to many fundamental questions in biology.