Physics Department

# Cell-Free Synthetic Biology: From Biomolecular Switches to Polymersomes and Bacteriophages

Dissertation by

**Kilian Vogele**

Technische Universität München

Technische Universität München
Fakultät für Physik
Physik synthetischer Biosysteme

# Cell-Free Synthetic Biology: From Biomolecular Switches to Polymersomes and Bacteriophages

Kilian Vogele

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r): Prof. Dr. Karen Alim

Prüfer der Dissertation:

1. Prof. Dr. Friedrich C. Simmel

2. Prof. Dr. Gil G. Westmeyer

Die Dissertation wurde am 02.07.2020 bei der Technischen Universität München eingereicht und durch die Fakultät der Physik am 18.08.2020 angenommen.

# Contents

# List of Publications

"Production of Phage Therapeutics and Formulations: Innovative Approaches"
Maia Merabishvili, Jean-Paul Pirnay, **Kilian Vogele**, Danish J. Malik, Phage Therapy: A Practical Approach, Springer, Cham, 3-41 (2019) [1]

"Genetically Encoded Membranes for Bottom-Up Biology"
**Kilian Vogele**, Tobias Pirzer and Friedrich C. Simmel, ChemSystemsChem, Volume 1, e1900016 (2019) [2]

"Periodic Operation of a Dynamic DNA Origami Structure Utilizing the Hydrophilic-Hydrophobic Phase-Transition of Stimulus-Sensitive Polypeptides"
Marisa A. Goetzfried, **Kilian Vogele**, Andrea Mückl, Marcus Kaiser, Nolan B. Holland, Friedrich C. Simmel, Tobias Pirzer, Small, Volume 15, 45-1903541 (2019) [3]

"*In Vesiculo* Synthesis of Peptide Membrane Precursors for Autonomous Vesicle Growth"
**Kilian Vogele**, Thomas Frank, Lukas Gasser, Marisa A. Goetzfried, Mathias W. Hackl, Stephan A. Sieber, Friedrich C. Simmel, Tobias Pirzer JoVE, 148, e59831 (2019) [4]

"Enhanced Efficiency of an Enzyme Cascade on DNA-Activated Silica Surfaces"
**Kilian Vogele**, Jonathan List, Friedrich C. Simmel, Tobias Pirzer, Langmuir, Volume 34, 49 14780-14786 (2018) [5]

"Towards Synthetic Cells Using Peptide-Based Reaction Compartments"
**Kilian Vogele**, Thomas Frank, Lukas Gasser, Marisa A. Goetzfried, Mathias W. Hackl, Stephan A. Sieber, Friedrich C. Simmel and Tobias Pirzer, Nature Communications, Volume 9, 3862 (2018) [6]

"Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch"
**Kilian Vogele**, Jonathan List, Günther Pardatscher, Nolan B. Holland, Friedrich C. Simmel and Tobias Pirzer, ACS Nano, 10, 11377-11384 (2016) [7]

# Abstract

Life as we know it is structured. Through the spontaneous formation of biological building blocks, such as lipids or even proteins in higher organized structures, the units of life, the cells, are formed. The principle behind this is self-assembly, encoded in the molecular structure and guided by energy minimization. Based on this principle structures ranging from nanoscale to macroscopic dimensions can be built.

In this thesis, different biopolymers, like DNA and peptides are used to create structures based on self-assembly. The first investigated system is a structure designed according to the DNA-origami technique, which is capable of guiding electromagnetic energy below the optical diffraction limit. This is achieved by assembling gold nanoparticles on a DNA-origami structure in close proximity to each other as an active plasmonic waveguide. By modification of the central gold nanoparticle, with a thermomechanical switch based on an elastin-like polypeptide, it is also possible to alter the efficacy of the energy transfer.

In the second part, the self-assembly of peptides and proteins in a cell-free transcription-translation system is investigated. As a first system, a synthetic cell is assembled consisting of an amphiphilic elastin-like polypeptide membrane, which is capable of the formation of vesicles. Inside the vesicle, a variety of different active processes can be encapsulated, like RNA transcription or protein expression. By combining the gene encoding for the amphiphilic elastin-like polypeptide inside the vesicle growth of the membrane can also be achieved. As a proof of concept, the cell-free transcription-translation system is used to generate bacteriophages from their genome. It is shown that the DNA of the T7 bacteriophages is replicated and the expressed proteins self-assemble into fully functional bacteriophages. This cell-free system is then further used as a platform to show that the expression of a natural genome can be altered by the external addition of a short DNA strand. Additionally, it is shown that the cell-free system can be expanded to express a huge variety of bacteriophages, as well as serve as system to identify hypothetical proteins in the bacteriophage genome using mass spectrometry. In this work, we show the potential of the cell-free transcription-translation system for accelerating bacteriophage research and thus is likely to facilitate the development of treatments with bacteriophages to contribute in solving the emerging antibiotic crisis.

# Zusammenfassung

Das Leben, wie wir es kennen, ist strukturiert aufgebaut. Durch die spontane Formation von biologischen Bausteinen, wie Lipiden oder auch Proteinen in höher organisierte Strukturen, werden die Einheiten des Lebens, die Zellen, gebildet. Das Prinzip hinter diesem Effekt ist die Selbstassemblierung, implementiert in die molekulare Struktur und geführt durch die Minimierung der freien Energie. Mit diesem Prinzip können Strukturen von der Nanoskala bis hin zu makroskopischen Dimensionen gebaut werden.

In dieser Arbeit werden Biopolymere wie DNA und Proteine verwendet, um durch Selbstassemblierung verschiedene Strukturen herzustellen. Das erste untersuchte System ist eine Struktur, die mit der DNA-Origami Technik hergestellt worden ist und elektromagnetische Energie leiten kann. Dies wurde mit Goldnanopartikeln bewerkstelligt, die auf einer DNA-Origami-Struktur so angeordnet worden sind, dass sie einen aktiven plasmonischen Wellenleiter bilden. Durch die Ausstattung des zentralen Goldnanopartikels mit einem Elastin-ähnlichem Polypeptid konnte gezeigt werden, dass die Effizienz des Energietransfers verändert werden kann.

Im zweiten Teil wird die Selbstassemblierung von Peptiden und Proteinen in einem zellfreien transkriptions-translations System untersucht. Das erste hier untersuchte System ist eine synthetische Zelle, welche aus amphiphilen Elastin-ähnlichen Polypeptiden besteht, welche Vesikeln formen. Es werden eine Reihe von verschiedenen aktiven Prozessen in diese Vesikel eingeschlossen, wie ein Transkriptionssystem für RNA und ein Proteinexpressionssystem. Mit der Kombination der genetischen Information für das amphiphile Peptid und dem zellfreien Extrakt ist es auch möglich, Wachstum der Membran zu zeigen. Weiterhin ist mit dem zellfreien transkriptions-translations System möglich, dass Bakteriophagen mit Hilfe ihres Genoms hergestellt werden können. Das zellfreie System wurde auch als Plattform genutzt, um das Expressionsmuster von einem natürlichen Bakteriophagengenom im Zellextrakt durch die externe Hinzugabe von einem kurzen DNA Strang zu ändern. Zusätzlich konnte gezeigt werden, dass das zellfreie System Erweiterungsmöglichkeiten bietet, um eine Vielzahl an verschiedenen Bakteriophagen zu exprimieren, aber auch als System verwendet werden kann, mit welchem hypothetische Bakteriophagenproteine mit Massenspektrometrie identifiziert werden können. Damit wurde das Potenzial des Zellextrakts für die mögliche Anwendungen in der Bakteriophagenforschung gezeigt und wir können somit hoffentlich dazu

beitragen, dass die Bakteriophagentherapie schneller zum Patienten gebracht werden kann und somit einen Beitrag zur Lösung der aufkommenden Antibiotikakrise geleistet wird.

# 1  Emergence of Life

What is life? As simple as this question may sound, there is still no generally accepted definition of life.[8] There is an intuitive understanding what it means to be alive, but to narrow this down to a universal definition is problematic, especially in synthetic biology, which explores the border between the inorganic and the living world.[9] Several disciplines in the life sciences would require a list of typical characteristics of life, such as astrobiologists in their search for alternative life forms [8] or origin of life researchers in their attempt to understand how life on earth might have emerged in order to recreate similar conditions in experiments.[10] But also for synthetic biology, such a definition would be required in the search for a minimal system which would be still seen alive, or for the construction of a living system completely from scratch.[11] The problem with such a fixed definition is that there are no conclusive theories in biology, as it is the case for physics. Nevertheless, there are some common characteristics that an organism must fulfill to be considered as alive:[8, 12]

**Respond to stimuli** whereat various forms of responses and stimuli exist, from following a gradient of nutrients to sense the light in the environment to increase the amount of sunlight for photosynthesis.

**Metabolism** as the transformation of nutrients to either energy (catabolism) to fuel cellular processes or into other building blocks to create and maintain the cell (anabolism).

**Homeostasis** is maintaining an equilibrium of an open dynamical system through internal processes that are separated from the environment.

**Growth** as the ability to increase the cell or organism itself, by its part and in total.

**Adaption/evolution** as the ability to change and evolve over time as a function of the environment.

**Organization** of the cell itself in hierarchical structures, which especially include compartmentalization as the organism needs to have a membrane that encapsulates an interior where most of the biochemical processes take place.

**Reproduction** to create progeny, either by division from one organism or sexually with a partner.

To condense these characteristics several scientists have attempted to establish a definition, depending on the scope and focus of their research. One of the most popular ones is the working definition of the NASA based on a panel around Gerald Joyce: "Life is a self-sustaining chemical system capable of Darwinian evolution."[13] If the focus is more set towards the capability of self-organization Stuart Kauffman argued that life emerges in a phase transition.[14] This self-organization has emerged from a collectively autocatalytic set of polymers which are capable of molecular reproduction and completing at least one thermodynamic work cycle.[15] For a more thermodynamic scope of the definition, the physicist Erwin Schrödinger made an attempt to define life in his lecture series "What is life". He proposed that life is a well-ordered pattern of molecules that decrease their internal entropy by increasing the entropy of their environment.[16]

One of the even more discussed questions than 'What is life?', is how the first organisms emerged and how possible first self-reproducing systems may have looked like. As the origin of life on earth took probably place 4.3 to 3.8 billion years ago [17] the exact circumstances of the events cannot be traced back anymore. Therefore the research is more focused on recreating plausible scenarios on how life may have emerged on the early earth.

A promising candidate for a first self-replicating system is RNA, which resulted in the hypothesis of the RNA world.[18] The advantage of RNA in comparison to other polymers like DNA is that RNA can store information as well as be catalytically active like in ribozymes.[19] Whereas this higher reactivity of the RNA also limits the stability as an information storage polymer as it is more prone to hydrolysis. This is especially problematic for the storage if the geological time scales for the accumulation of the building blocks of the RNA are taken into account.[20] A possible pathway for the successful synthesis of the RNA monomers has been explored by the Sutherland group as they showed in a prebiotic plausible way the synthesis of activated pyrimidine nucleotides.[21] Moreover, they demonstrated with a prebiotic plausible scheme based on hydrogen cyanide that the precursors of pyrimidine nucleotides can also form precursors of lipids and amino acids.[22] This would be a piece of evidence that life on earth may be formed on common chemistry. Find-

ings in meteorites also suggested that the building blocks of RNA can be formed extraterrestrially, like adenine and guanine.[23]

Besides the synthesis of the monomers, another important step is the emergence of polymers with catalytic function, especially self-replication. This process can be divided into two major steps, the template-directed polymerization of the RNA and the ligation of RNA molecules to form larger structures for more complex tasks.[24] This combination is necessary as the currently available ribozymes which can perform template-directed polymerization of RNA monomers are longer as their own processivity.[25] With RNA-joining ribozymes, it was possible to construct from a total of four components a cross-catalytic system which was able to undergo self-sustained replication with a doubling time of about 1 hour.[26] By combing such a system with an RNA-template-directed RNA polymerase a self-replicating system could be constructed.

Despite the fact that current life uses RNA, there might be a pre-RNA-world at which also other molecules were used, like peptide nucleic acid (PNA)[27] or threose nucleic acid (TNA).[28] Problematic here is that even if it was the case that other polymers have been also used in the origin of life this information got lost due to the evolution towards the modern cell.

Another approach to recreate life is the attempt to create a protocell in synthetic biology. Here, the major difference to the origin of life research is that there are no limitations to prebiotic plausible conditions. Therefore also more advanced enzymes and already evolved subsystems of modern cells can be utilized and combined.

The creation of protocells is driven by two main motivations, the curiosity of advancing science and the opportunity for practical implementations. Here, the main focus would be on medical applications like drug-delivering devices triggered by external signals or diagnostic tools. Another field of usage could also be multifunctional biosensing.[11]

Building such a protocell from scratch by combining several non-living molecules to react with each other and to self-assemble into a living system would be a *bottom-up* approach. For that realization several subsystems which mimic certain aspects of the aforementioned criteria of life were recreated.[30] The main driving force behind most of the experimental realizations are spontaneous self-assembly as lipid-based membranes and template-directed assembly of polymers like DNA replication, tran-
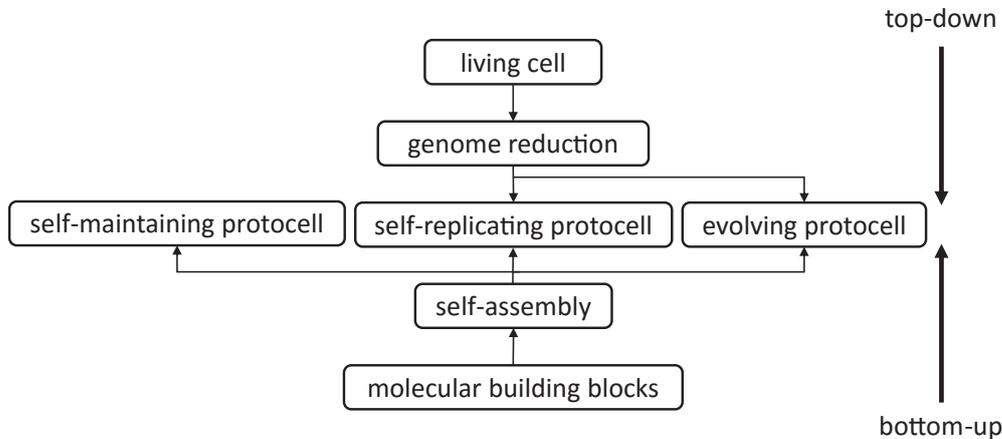
Figure 1: The two major approaches for building a protocell are the *top-down* and the *bottom-up*. Here, three potential intermediate steps can be achieved, a self-maintaining protocell, a replicating but not evolving, and a fully evolvable protocell. Adopted from [29].

scription and translation.[31] One of the major achievements was the utilization of cell-free expression systems, like crude *E.coli* cell lysates [32, 33] or the PURE system which is the purified expression machinery of a bacterial cell.[34] With combining this cell-free system several milestones towards an artificial living entities were reached, like the expression of membrane proteins in an encapsulated cell-free system.[35] It was also shown that in such an encapsulated cell-free system the constituents of the membrane can be synthesized by its encoded genes.[6, 36] A further example is the DNA replication of a genome encoding its own proteins in a cell-free system.[37, 38] Even evolution was shown in a translation system with RNA replication. In repeated dilution cycles RNA sequences were generated which were more resistant against parasitic RNA.[39] So far the *bottom-up* approach is limited by the mostly missing interaction of the subsystems and the loss of functions of certain parts over time which would make further research necessary.

Applying a different principle in comparison to *bottom-up* would be the *top-down* approach, as shown in Figure 1. Here a fully functional living cell is used where step-wise genes are removed to find a minimal functional system to just reproduce and maintain itself.[40, 41] As soon as such an artificial system is reached the process of evolution can be used to optimize this system and to adapt it to certain environments by designed experiments.

In this thesis, by applying the principles of self-assembly to different biopolymers, a range of structures were built like DNA-Origami structures, peptide protocells mimicking certain aspects of life, and fully functional *in vitro* synthesized bacteriophages. The different self-assembled structures are discussed in the following chapters whereat the work is mainly based on the following publications:

"Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch"
**Kilian Vogele**, Jonathan List, Günther Pardatscher, Nolan B. Holland, Friedrich C. Simmel and Tobias Pirzer, ACS Nano, 10, 11377-11384 (2016)

"Genetically Encoded Membranes for Bottom-Up Biology"
**Kilian Vogele**, Tobias Pirzer and Friedrich C. Simmel, ChemSystemsChem, Volume 1, e1900016 (2019)

"Towards Synthetic Cells Using Peptide-Based Reaction Compartments"
**Kilian Vogele**, Thomas Frank, Lukas Gasser, Marisa A. Goetzfried, Mathias W. Hackl, Stephan A. Sieber, Friedrich C. Simmel and Tobias Pirzer, Nature Communications, Volume 9, 3862 (2018)

"Target Specific Silencing of Genes in Cell-Free Systems for Controlling and Modification of Bacteriophages"
**Kilian Vogele**, Elisabeth Falgenhauer, Sophie von Schönberg, Friedrich C. Simmel and Tobias Pirzer, in preparation

"A Cell-Free Platform for Characterizing and Producing Therapeutic Bacteriophages Against Biohazardous and Antibiotic-Resistant Bacteria"
Quirin Emslander, **Kilian Vogele**, Peter Braun, Markus Joppich, Miriam Abele, Chen Meng, Christina Ludwig, Friedrich C. Simmel and Gil G. Westmeyer, in submission

The author of this thesis has made significant contributions to these publications, from contributions to the development of ideas, experimental designs and implementation, evaluation of the data and preparation of the papers.

The organization of this thesis is as follows: After the introduction, the second chapter gives an overview of the physics behind self-assembly, about the different possible biological building blocks, and about the structures which can be formed. In the third chapter, these principles are applied to build with the DNA-origami technique and gold nano-particles functionalized with DNA and DNA-peptide hybrids a switchable plasmonic waveguide. Chapter four is a review of the attempts of *bottom-up* biology to create protocells, especially focused on encapsulation. Followed by a chapter about the creation of such a genetically encoded system which can express - based on an *in vitro* cell-free system - its own membrane components which are in this case amphiphilic elastin-like polypeptides. In this system, we were also able to show growth of the self-assembled peptide vesicles. Utilizing this cell-free system further in chapter 7 proteins of the T7 bacteriophage have been expressed and fully self-assembled in a one-pot reaction. By suppressing the gene for the major capsid protein the DNA replication of the T7 bacteriophage has also been enhanced without altering the genome of the bacteriophage itself. The cell-free expression of phages was then further analyzed in chapter 8 where different bacteriophages were successfully expressed independently from their host. Additionally, it was also shown that the cell-free platform can be used as a tool to further characterize bacteriophages with mass-spectrometry to detect beforehand only hypothesized proteins. In the final chapter 9, the results are reviewed and a further outlook of the possible application of the findings towards the production of bacteriophage toward medical usage is given.

# 2  Fundamentals

## 2.1  Self-assembly of Biological Structures

Self-assembly is used ubiquitously in nature like in the formation of membranes,
[42] the folding of a polypeptide chain into a protein,[43] the capsid assembly of a
virus,[44] but also in inorganic materials like crystals [45] or in artificial systems like
structures built with the DNA-Origami technique.[46] In principle, the formation
of these systems have in common that they start spontaneously with a reversible
association of building blocks (tectons) into higher-ordered structures. The self-
assembly process is driven by the minimization of the free energy of the systems,
whereat the tectons usually interact via weak interactions with each other without
external directions. But this lowest state of energy and maximized entropy doesn't
lead to a uniform distribution of the building blocks as the interaction with the
surrounding molecules also have to be taken into account. This increase of order
of the resulting structure in comparison to the building blocks is also a distinct
feature of self-assembly. These principles can then be used to build arbitrary shapes
by altering the form, the number, the direction, and the range of the interaction
within the tecton. The major advantage of such build structures is that they all
form at the same time in contrast to the conventional approaches at which every
building block has to be placed separately.[47]

## 2.2  Thermodynamics of Self-assembly

The basics of the thermodynamic description in equilibrium of a system capable of
self-assembly (based on [48] and [49]) are building blocks that tend to aggregate.
The goal is to predict the portion of molecules which will form aggregates containing
a finite number of molecules. The volume fraction of a building block in a solution
with $N$ building blocks will be denoted as $X_N$. Therefore the total volume fraction
of the building blocks in the solvent is $\sum_N X_N = \phi$. In thermodynamic equilibrium,
the chemical potential $\mu$ of the dissolved building blocks must be the same in all
different coexisting aggregates. This can then be expressed as

$$\mu = \mu_N^0 + \frac{k_B T}{N} \ln \frac{X_N}{N}. \tag{1}$$

Whereat $\mu_N^0$ is the free energy change which is needed to take into account when a
free building block from the bulk is added to the aggregate with $N$ particles. The

second term is the contribution of the translational entropy of the aggregate. By solving for the volume fraction $X_N$ of the aggregate with $N$ building blocks we get

$$X_N = N \cdot \exp\left[\frac{N(\mu - \mu_N^0)}{k_B T}\right]. \tag{2}$$

Since in the thermodynamic equilibrium $\mu_N = \mu = \mu_1 = const$ is valid the volume fraction of the aggregate $X_N$ in respect to the monomers can be calculated by utilizing $\mu = \mu_1$ together with equation 1 which gives for the volume fraction of the aggregate in the solute with the size $N$ in terms of the fraction of the monomers of the building block $X_1$ the following expression:

$$X_N = N \left(X_1 \cdot \exp\left[\frac{(\mu_1 - \mu_N^0)}{k_B T}\right]\right)^N \tag{3}$$

This equation implies that if $\mu_N^0 > \mu_1^0$ most of the building blocks will be in a monomeric form. Therefore a condition for forming aggregates is that $\mu_N^0 < \mu_1^0$. For a finite size of the aggregates $\mu_N^0$ has to have a minimum at a finite value since for infinite size $\mu_N^0$ decreases progressively.

With the assumption that the monomer-monomer bond energy is $-\alpha \cdot k_B T$ the free energy of $\mu_N^0$ can be calculated whereat $\alpha$ is a positive constant which is given by the strength of the intermolecular interaction. To take into account that only the building blocks present on the interface to the solvent are unbund $\mu_N^0$ can be calculated with

$$\mu_N^0 = \mu_\infty^0 + \alpha \cdot k_B T \cdot N^{-P} \tag{4}$$

whereat $P$ is a number which depends on the dimensionality of the aggregate. Based on the generalized functional form of $\mu_N^0$ the concentration at which aggregates will form can be expressed by

$$X_N = N \left(X_1 \cdot \exp\left[\alpha(1 - 1/N^P)\right]\right)^N \approx N \left(X_1 \cdot \exp(\alpha)\right)^N. \tag{5}$$

With the conservation of mass $\sum_N X_N = \phi$ the monomer concentration at which aggregates will form, the so called critical aggregation concentration (CAC), can be expressed by:

$$(X_1)_{crit} = CAC \approx e^{-\alpha} \tag{6}$$

For sufficient small numbers of building blocks where $X_1 \cdot e^\alpha$ is smaller than one, $X_1 > X_2 > X_3 > ...$ is true for every $\alpha$ which implies that the majority of the building blocks are in a monomeric state. As $X_1$ approaches $e^{-\alpha}$ a further addition

of building blocks results in the formation of more aggregates whereat the monomer concentrations stays approximately constant. For a parameter $P = 1/3$ like for spheres a phase separation will take place to one aggregate of infinite size. This is true for simple building blocks like hydrophobic polymers. For more complex polymers, especially amphiphilic compounds $\mu_N^0$ has a minimum at a finite value of N.
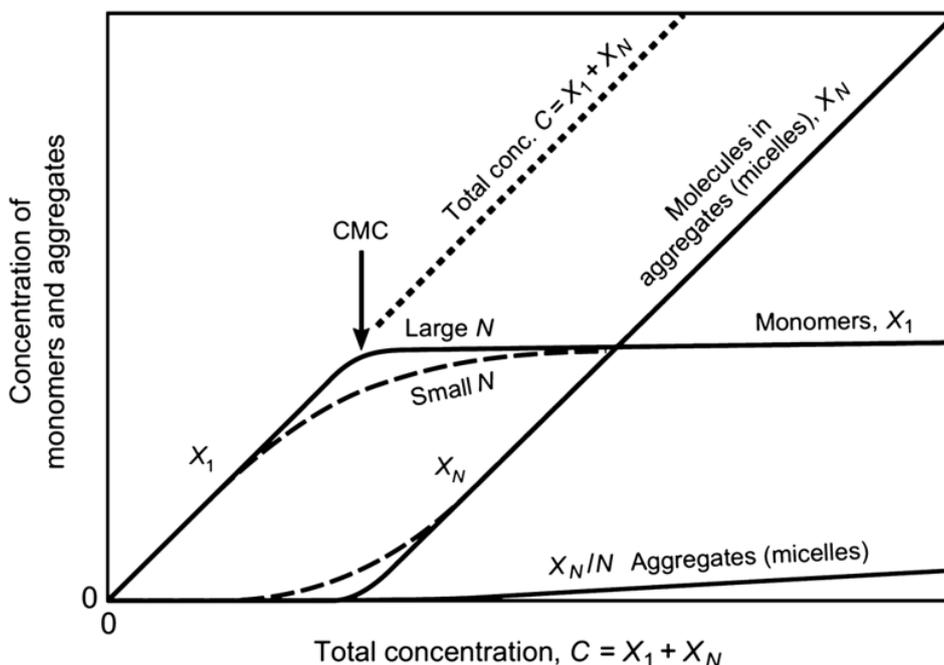


Figure 2: Concentration of the aggregation state of monomers in an aqueous solution as a function of the total concentration. The transition at the CMC is sharper for larger N. Reprinted by permission from Elsevier Inc.[48], copyright (2011).

**Forces in self-assembled systems:** The strength of typical interactions in self-assembled systems needs to be weak and reversible. Therefore the free energy of such bonds is in the range of $k_B T$. Molecular forces like charge-charge interaction are usually too strong, so therefore weak interactions like the dispersion force, hydrogen bonds, and entropic forces like the hydrophobic effect are utilized by self-assembled systems.

The dispersion force usually refers to all electrostatic interactions between neutrally charged molecules and decreases rapidly with the distance. This force plays a major roll in systems with biomolecules as a lot of them have a high static dipole moment

and these biomolecules can also be polarized which induces a dipole moment. The hydrogen bonds are interactions that occur between an acceptor with a high electronegativity in close proximity to a hydrogen atom which is bound to a donor. A requirement for the donor is that it contains an unbound electron pair which is usually the case for nitrogen, oxygen, sulfur, or fluorine. In general, the interactions described by hydrogen bonds are a mixture of the dipole interaction and quantum-mechanical effects. The hydrophobic effect is the clustering of nonpolar molecules in a polar medium like water. The water molecules have to arrange themselves around the hydrophobic molecule which results in a solvent cage. This increases the order of the water molecules which would result in a higher entropy. As now the hydrophobic molecules cluster together a smaller amount of solvent cages are required which result in a total increase of the entropy of the system.[50]

## 2.3 Biological Building Blocks

For successful self-assembly, the tectons can be designed accordingly to two different design principles. For the positive design strategy, the affinity of the building blocks is maximized to form into the desired structure and to minimize therefore the total free energy of the system with respect to other designs. Problematic here is that alternative structures with the same tectons can have an even lower free energy state and wrongly formed aggregates could emerge. In contrast to that stands the negative design strategy which aims to maximize the free energy distance between the desired structure and possible alternative assemblies. For successful self-assembly usually, a mixture of both design strategies brings reasonable results.[51, 52]

### 2.3.1 Lipids

Lipids are biomolecules which have a high solubility in nonpolar solvents, among others this includes fatty acids, phospholipids or sphingolipids.[53] They are used for energy storage, signaling and as the key component of the cellular membrane.[54] They all have in common that they have an amphiphilic structure which made them a good building block for self-assembly.[55]
One of the simplest molecules in the class of lipids are fatty acids. In general, a fatty acid is a class of molecules with an aliphatic chain of varying length and a carboxylic acid group. They can either be unsaturated, which means with at least

one or more double bonds between the carbon atoms of the aliphatic chain, or saturated. Most of the naturally occurring fatty acids have an even number of carbon atoms which are ranging from 4 to 28. Based on the hydrophilic head group and the hydrophobic tail they can be used as a building block for the self-assembly of micelles and vesicles.[56]

In cells, these fatty acids are used for energy storage. In the cytoplasm, they are usually stored in the form of triglyceride which is composed of three fatty acid molecules. Another important role is the one as the main component of membranes since they are precursors of phospholipids which are a major component of the cellular membrane.[56] In the context of the origin of life also the fatty acids themselves are a plausible candidate for the first membranes. Besides plausible prebiotic synthesis ways, fatty acid membranes have dynamical properties which are beneficial for growth and permeability. For this the exchange of the outer and the inner leaflet of the membrane is essential. For fatty acids, this exchange can occur in subsecond time scale.[57] All these lipids have in common that they can form based on their geometry of the head group and tail a range of different self-assembled structures, like micelles and vesicles.[48]

### 2.3.2 Peptides

Peptides are biological polymers of amino acids linked by an amide bond whereat their difference to proteins is their length of less than 50 amino acids. They have a huge variety of function in nature like as signaling molecules, as a defense against other cells (like antimicrobial peptides) or as a capsid of a virus.[58] As by tuning their interaction with each other by altering the amino acid sequence they can be used as a well designable tectons for self-assembly.[52]

**Elastin-like polypeptides.** Elastin-like polypeptides (ELP) are repeats of the synthetic pentapeptides which was inspired by the amino acid sequence of tropoelastin.[59] ELP consists of a repetitive pentameric amino acid sequence (Gly-$\alpha$-Gly-Val-Pro) where $\alpha$ can be every natural amino acid except for proline.[60] Characteristic for ELP is the distinct phase transition behavior, characterized by changing from a soluble state to an aggregated state when the transition temperature ($T_t$) is exceeded like it is shown in Figure 3. This transition temperature depends on the environmental conditions like the salt concentration, the pH, the ELP concentration and also the primary structure of the ELP.[60]
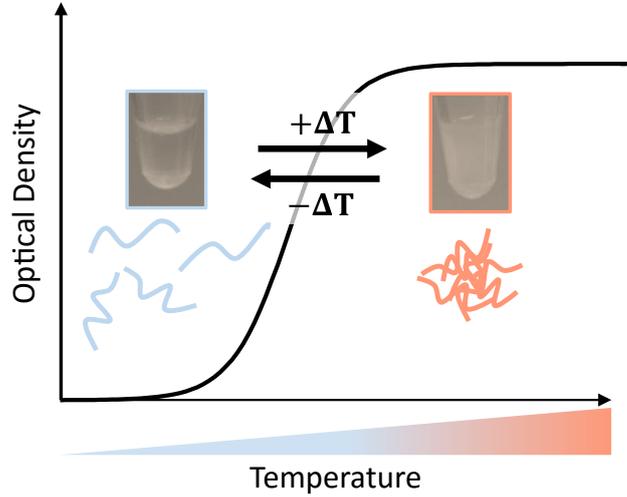
Figure 3: Schematic of the change of the optical density, based on the aggregation of the ELP depending on the transition temperature.

For tuning the $T_t$ of the ELP the primary amino acid sequence at the second position can be altered. The hydrophobicity of this amino acid determines the $T_t$ as it is shown in Figure 4. Beside the primary structure the concentration as well as the length of the ELP has a significant influence on the $T_t$. Phenomenological there is relation between the concentration, the length and the transition temperature:

$$T_t \propto \frac{1}{L} \cdot \ln \left( \frac{1}{C} \right) \tag{7}$$

where L is the chain length and C the concentration of the ELP.[62]

The transition temperature is also dependent on the salt concentration whereat the relation between them depends on the position of the salts in the Hofmeister series. For kosmotropic anions like $Cl^-$ or $H_2PO_4^-$ the relation is simply linear with

$$T(c) = T_0 + c[M]. \tag{8}$$

Where $T_0$ is the transition temperature without salt, c a constant with units of temperature/molarity and [M] is the concentration of the salt.[63] For a rough estimation the constant c can be set to $-14\frac{°C}{M}$ which was measured for the GVGVP ELP.[64] In contrast for the chaotropic anions like $SCN^-$ or $Br^-$ the transition temperature need an additional term the Langmuir binding isotherm:

24
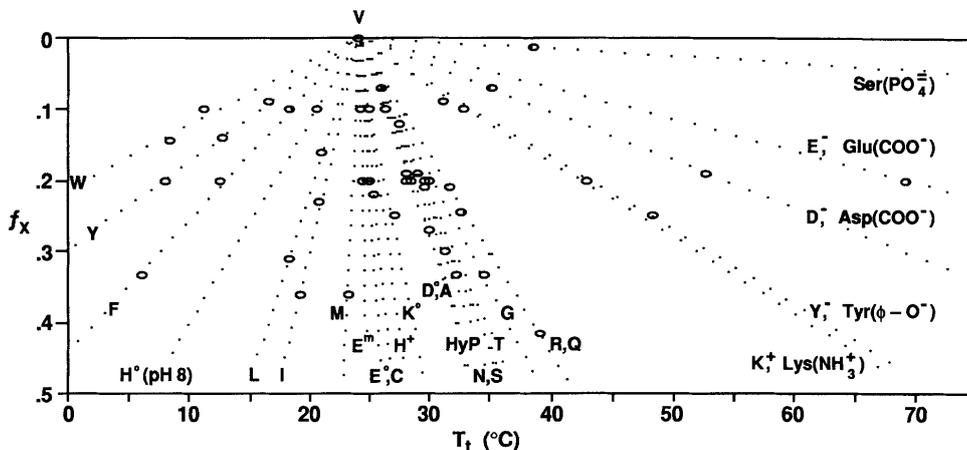
Figure 4: Inverse transition temperature of elastin-like polypeptides with the sequence poly[$f_v$(VPGVG),$f_x$(VPGXG)]. The transition temperature is plotted as a function of the mole fraction $f_x$ of the guest residue X for different amino acid. The letters stands for the single letter code of the amino acids. Reprinted by permission from Wiley [61] copyright (1992).

$$T(c) = T_0 + c[M] + \frac{B_{max} \cdot K_A[M]}{1 + K_A[M]} \tag{9}$$

where $K_A$ is the equilibrium association constant. Since the isotherm is unitless the constant $B_{max}$ is added which has a unit of temperature.[63]

From a structural point, the transition of the temperature above the $T_t$ is associated with the formation of $\beta$-spiral structures.[65] Due to the difficulty to measure the secondary structure with NMR due to their size and problems with the crystallization [65] there are however controversies which structure the ELP form in their aggregated state.[66] As they collapse the ELP reduce their volume. As for macroscopic crosslinked ELP gels, the volume change can be up to 20 times in respect to the volume below the $T_t$ [67] as for microscopic micelles structures of crosslinked ELP, the diameter change is only 35 %.[68] It was also shown that for a collapsed ELP the contour length is 20 % reduced.[69] These data differ between the used systems, especially as it is hypothesized that only at least three ELP can collapse to the aggregated state with the distinct temperature transition behavior. In contrast to other self-assembled building blocks like lipids, the water content in

25

the aggregated phase is with roughly 60 % [70] to 63 % [69] relatively high.[71]

### 2.3.3  Proteins

Proteins are long biomolecules consisting of an amino acid chain. They are responsible for several tasks in cells, like catalytic reactions, providing structure, the responsibility for transport between the interior and the exterior, and for the response to stimuli. To fulfill all these functions certain secondary structure motif is used, like the $\alpha$-helix and the $\beta$-sheet. These motifs can then also be self-assemble into higher-ordered structures like the coiled-coil motif as in $\alpha$-kreatin.

The folding of the proteins is mainly due to the self-assembly of the polypeptide chain. But as several interactions, like hydrogen bonds, hydrophobic forces, disulfide bonds, salt bridges, and other electrostatic interactions are creating a complicated energy landscape, which makes the prediction of the final tertiary structure based on the amino-acid sequence challenging. Due to kinetic traps, the proteins might not fold into its minimum free energy state. In such cases, chaperons may support the correct folding. These correctly folded proteins can also further self-assemble into higher-ordered structures as the tubulin polymerization of mikrotubuli or the viral capsid assembly of the tobacco mosaic virus.[56] This shows that the principles of self-assembly are used on several hierarchical levels to design all kinds of different structures.

### 2.3.4  DNA

Deoxyribonucleic acid (DNA) is the major molecule responsible for the storage of information in cells. Usually, DNA is in a double helix form and carries the genetic information for the expression of proteins. DNA is a polymer consisting of four different nucleotides: adenosine, guanine, cytosine, and thymine. The double helix which is formed by the DNA is mainly stabilized by two forces the hydrogen bonds between the obverse nucleotides and the base-stacking interaction between the aromatic rings of the nucleotides above and below. For ordinary Watson-Crick base-pairing adenosine pairs with thymine and guanine pairs with cytosine.[56] Based on that self-assembled structures can be designed like with the DNA-origami technique.[46]

26

## 2.4 Self-assembled Structures

From the free energy of the building blocks the volume fraction of the formed aggregates could be calculated (based on [48] and [49]). As for self-assembly certain order is required the building block need to fulfill special requirements, whereas the most prominent class are amphiphilic molecules. These molecules have a hydrophilic head, which is in contact with the water interface and a hydrophobic tail, which is energetically unfavorable in contacting water. Based on the geometry of the molecules there is a most favorable structure, based on the head-group area $a_0$ the hydrophobic tail length $l_c$ in a fully extended straight configuration and the volume of $v$ of the hydrophobic tail. The optimum head group area is defined by the minimum of the free energy of the distance of the head groups, which consists of repulsion based on the electrostatic, or other interaction and an increasing internal attraction. Based on that, the dimensionless packing parameter can be defined:

$$pa = v/a_0 l_c \tag{10}$$

Based on that packing parameter the type of the most favourable structure can be determined, like spherical micelles ($pa \approx 1/3$), cylindrical micelles ($pa \approx 1/2$), bilayers or vesicles ($pa \approx 1$) or inverted structures ($pa > 1$) as shown in Figure 5.

### 2.4.1 Vesicles

Vesicles are defined according to IUPAC as "Closed structure formed by amphiphilic molecules that contain a solvent (usually water)."[72] Based on the optimized ratio of the packing parameter as defined in equation 10 around one the building blocks tend to form bilayers. To determine the size distribution of the bilayers in a solution above the CAC the free energy of the monomer has to be determined. This free energy consists of two parts one representing the energy of the bond $\mu_\infty^0$, which is a constant term and one for the fraction of the edge which is in contact with the solvent, distributed to every building block with $\alpha \cdot k_B T/\sqrt{N}$. This results in:

$$\mu_N^0 = \mu_\infty^0 + \frac{\alpha \cdot k_B T}{\sqrt{N}} \tag{11}$$

Combining this equation with the expression of the volume fraction $X_N$ form equation 5 results to:

$$X_N \approx N \left( X_1 \cdot \exp(\alpha) \right)^N \cdot \exp(-\alpha N^{1/2}) \tag{12}$$

| Lipid | Critical packing parameter $v/a_0 l_c$ | Critical packing shape | Structures formed |
|---|---|---|---|
| Single-chained lipids (surfactants) with large head-group areas: *SDS in low salt* | $< 1/3$ | Cone | Spherical micelles |
| Single-chained lipids with small head-group areas: *SDS and CTAB in high salt, nonionic lipids* | $1/3$-$1/2$ | Truncated cone | Cylindrical micelles |
| Double-chained lipids with large head-group areas, fluid chains: *Phosphatidyl choline (lecithin), phosphatidyl serine, phosphatidyl glycerol, phosphatidyl inositol, phosphatidic acid, sphingomyelin, DGDG[a], dihexadecyl phosphate, dialkyl dimethyl ammonium salts* | $1/2$-$1$ | Truncated cone | Flexible bilayers, vesicles |
| Double-chained lipids with small head-group areas, anionic lipids in high salt, saturated frozen chains: *phosphatidyl ethanolamine, phosphatidyl serine + Ca²⁺* | $\sim 1$ | Cylinder | Planar bilayers |
| Double-chained lipids with small head-group areas, nonionic lipids, poly *(cis)* unsaturated chains, high *T*: *unsat. phosphatidyl ethanolamine, cardiolipin + Ca²⁺ phosphatidic acid + Ca²⁺ cholesterol, MGDG[b]* | $> 1$ | Inverted truncated cone or wedge | Inverted micelles |
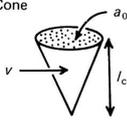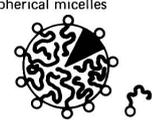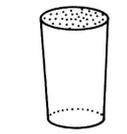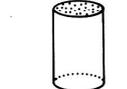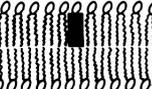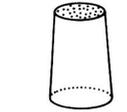
Figure 5: Packing parameter of different building blocks and the resulting structure of the aggregates. Reprinted by permission from Elsevier Inc.[48], copyright (2011).

From this expression, every added building block above the CMC would join an infinite sized sheet. Taken into account that the ends of the sheets can join thought bending the extra energy of the edges can be eliminated. For that a corresponding bending energy $\Delta E \propto -1/R^2$ per unit area of the bilayer a finite size distribution of vesicles can be derived, where R is the radius of the vesicle. From geometrical considerations the radius of such a vesicles, with a slightly truncated cone ($v/\alpha_0 l_c < 1$) can be calculated by:[73]

$$R \approx \frac{l_c}{1 - v/\alpha_0 l_c} \tag{13}$$

28

which is the critical radius which a vesicle would form, without the introduction of packing stress on the building blocks.

The role of vesicles in biological systems is quite diverse, as the complete cell membrane of L-form bacteria,[74] as containers for nutrients storage and processing and as transporters of substances inside and outside of the cells.[54] Despite their manifold tasks the principle behind the formation is always based on the same principle self-assembly.[55]

### 2.4.2 Bacteriophages

Bacteriophages (phages) are the most common biological particles on earth with an estimated number of more than $10^{32}$. Phages are viruses which are specialized in bacteria as their host.[75]

**A brief history of phages and their therapeutic usage.** Phages were first described in 1915 by Frederik Twort, whereas the antimicrobial use of phages was already postulated two years later by Felix d'Herelle. In 1919, the first documented clinical use of phages occurred at the Hôpital des Enfants-Malades in Paris, where phages were successfully used to treat 4 pediatric cases of bacterial dysentery. This was followed by larger studies with phages as therapeutics, such as a trial against cholera. With 73 patients in the experimental group versus 118 patients in the control group, d'Herelle observed a 90 % reduction in mortality with 74 fatal results in the control group and only 5 in the experimental group, which were treated with phages.[76] Due to the lack of knowledge of the underlying biological processes and errors in application and storage, the results of many other studies were rather mixed. With the spread of antibiotics as an alternative to the treatment of bacterial infections, bacteriophage therapy disappeared from western medicine.[77] Much of the further development in phage therapy has taken place in the states of the former Soviet Union since the 1930s. Over the years, the Eliava Institute in Georgia, in particular, has built up a large collection of more than 830 medically used phages.[78] The largest study to date on the use of bacteriophages as a therapeutic agent has also been carried out there. For this purpose, over 30.000 children were treated prophylactically with a cocktail of phages against Shigella. The study participants were divided into a placebo group and an experimental group. After 109 days it

was shown that the patients treated with phages were 74 % less likely to develop dysentery.[79] At the Eliava Institute, phages are still successfully used today to treat acute and chronic infections.[80]

**Phage biology.** Because of the high abundance and diversity of phages, there is a huge variety of different groups, which are mainly classified by their morphology, their genome and their host. In general, phages can have two distinct life cycles: the lysogenic and the lytic. The lytic cycle is the most important one for therapeutic usage since in this cycle the phage replicates in the host bacterium and releases them through lysis, which leads to cell death. Besides the lytic life cycle, which all phage have in common, the lysogenic life cycle is optional for a majority of phages.[81]
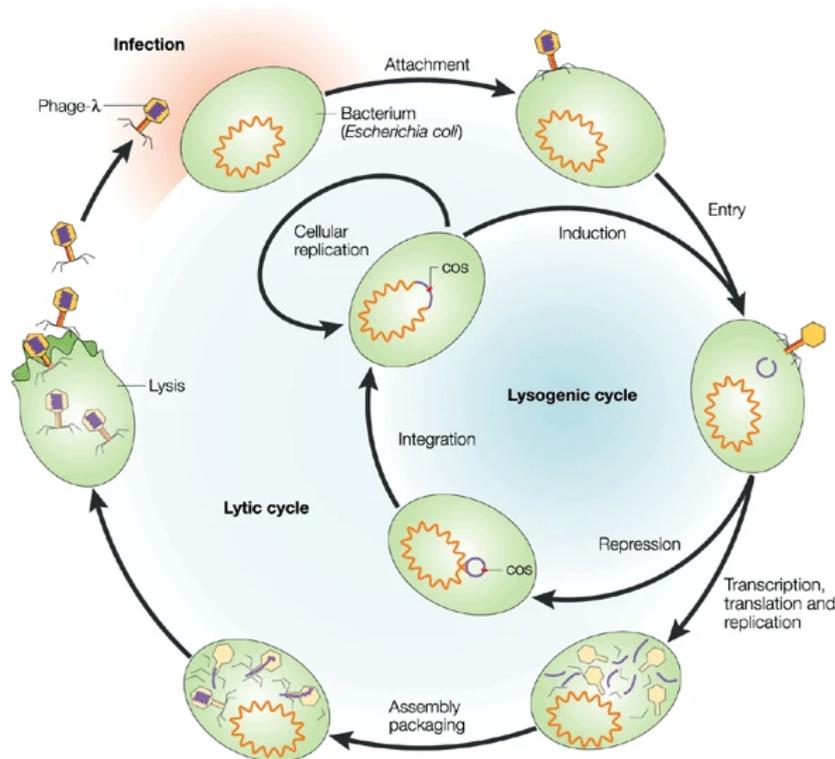


Figure 6: The lysogenic and the lytic life cycle of the temperate $\lambda$ bacteriophage. Reprinted by permission from Springer Nature [81] copyright (2003).

The life cycle of a bacteriophage, with the optional lysogenic life cycle, is depicted

in Figure 6. The reproduction of a lytic phage can be separated into five phases:

**Adsorption:** Based on the highly specific interaction of the bacterial cell wall and the receptor proteins attached to the phage capsid structure the corresponding host is identified. These specific receptors on the cell wall can be proteins, lipopolysaccharides, or even the flagellum.

**Injection:** If a corresponding host is identified the nucleic acid of the phage is ejected into the bacteria cell. Due to the high variety of phages, the genome of the phage can be single-stranded RNA, doubled-stranded RNA, single-stranded DNA, or double-stranded DNA.

**Production:** During this time the genes on the phage genome gets expressed, usually in a well-defined order and also the genome of the phage is replicated.

**Maturation:** In this phase the capsid protein of the phage self-assemble themselves into the capsid structure as well as the packing of the genome into the phage.

**Release:** For the release of the phages several strategies are used. The most common one is the cell lysis as the membrane of the host gets degraded and finally disrupted. This process is usually done by the so-called endolysin which are degrading the peptidoglycan in the cell wall. Another strategy is the continuous release of phages by secretion, which would be the case for filamentous phages. After the release these new phages can then start this cycle all over again.

For the lysogenic life cycle, the phage integrates temporarily into the genome of the host bacterium with the help of viral integrases and transposase. In most cases, they are separated from the genome again by an external stimulus and then return to the lytic infection cycle. However, other fragments of the bacterium's genome can also be transferred, such as antibiotic resistance and pathogenicity factors. For this reason, only strictly lytic phages are used for medical applications.[82]

**Phage structure.** A model system for the classification of phages according to their structure are the seven virulent Coliphages of the T-series. The family of the T1 and T5 is the *Siphoviridae* for the T3 and T7 it is the *Podoviridae* and T2, T4, and T6 belong to the *Myoviridae*. They all have a linear double-stranded DNA genome, with varying length and terminal redundancies.[83] The even-numbered phages of the T-series have instead of cytosine a modified 5-Hydroxymethylcytosine base as protection against restriction enzymes in the host.[84]
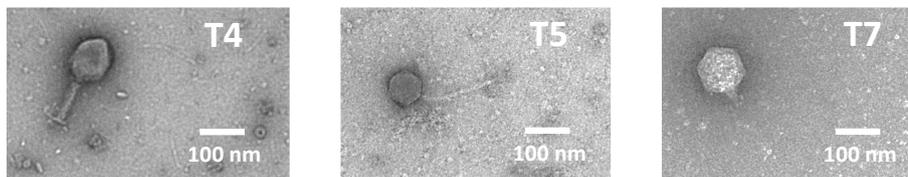


Figure 7: Morphology of different phages of the T-series. Transmission electron microscope image of the T4 phage (left), T5 phage (middle) and T7 phage (right). Scale bar: 100 nm.

Despite the huge variation in shape, as seen in Figure 7 generalized principles can be found for the structure of the capsid as it relays on the self-assembly of a small number of different asymmetric building blocks. The general structure of most capsid is a polyhedron, constructed by employing symmetry-principles to generate a large capsid with only a small number of genes.[85]

To arrange the asymmetric building blocks in such a way that a symmetric structure can be constructed, three of the building blocks have to be placed into a triangular conformation to build an icosahedron. In such a triangle the building blocks have six equivalent contact points AE, three other equivalent contact points BC, and three similar contact points DD, as shown in Figure 8. As the same contact points are used several times a symmetrical structure can be built from these triangular grids.[85] An icosahedron can be constructed consisting of 20 equilateral triangles with 12 vertices. As one triangle consist of 3 building blocks the smallest possible icosahedron consists of 60 subunits. To increase the size of the icosahedral capsid new smaller triangles have to be implemented into the original triangles from the icosahedron. As the structure has to be symmetric only a certain number of new triangles can be added to the icosahedron structure. These newly added triangles are forming hexamers, whereas on the vertices pentamers are formed.[86] The number of newly constructed hexamers $H$ can be calculated with the triangulation number
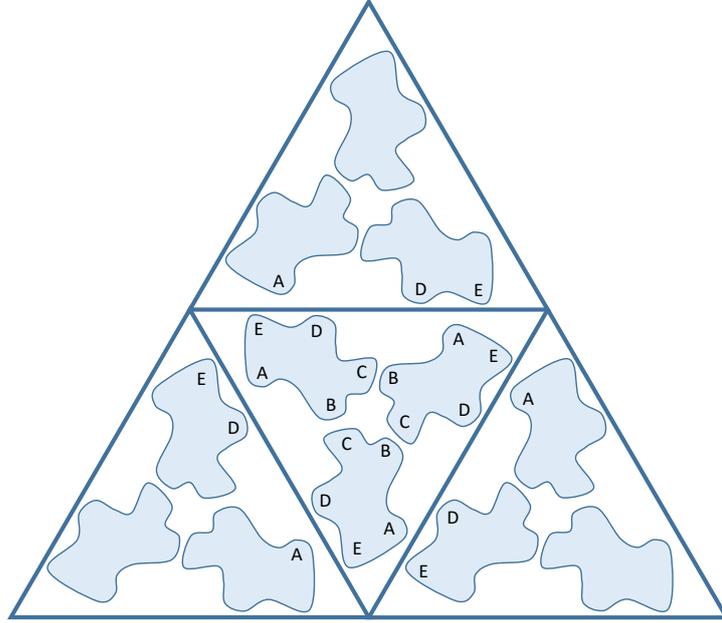
Figure 8: Illustration of the triangular grid constructed from asymmetric building blocks to build an icosahedron like structure. The possible contact points of the structure are denoted by the letters A-E. Adopted from [50].

T, which is the count of how often the original triangle of the icosahedron is divided, with the formula $H = 10 \cdot (T - 1)$. The triangulation number can be calculated by the integers $h \geq 1$ and $k \geq 0$. $h$ is the number of steps taken in a straight line from one pentamer to the closest point of the neighboring pentamer, whereas $k$ is the number of steps taken in a straight line from the point which was arrived with $h$ steps before towards the neighboring pentamer by turning $60°$. Therefore the triangulation number T is defined as[87]:

$$T = h^2 + h \cdot k + k^2 \tag{14}$$

There are several examples of such icosahedral structures among phages, like the PhiX174 bacteriophages from the family of the *Microviridae* with T=1, the Q$\beta$ bacteriophages from the family of the *Leviviridae* with T=3 and the T7 bacteriophages from the family of the *Podoviridae* with T=7.[88]

### 2.4.3 DNA-Origami

Structure prediction of self-assemblies of proteins based on their amino acid sequence is still challenging due to the huge variety of interactions. In contrast to that DNA would be a way easier predictable building block for the design of structures, as it mainly relys on Watson-Crick base-pairing. As an especially robust technique for the building of all kinds of arbitrary shapes is the DNA-origami technique. The two main components are a long single-stranded DNA strand the so-called scaffold, usually from the M13 phage, and many short single-stranded DNA strands, the so-called staples. The staple strands are responsible to shape the scaffold in the designed structure. One part of the staple strand is therefore complementary to a certain region of the scaffold and another part of the staple strand is complementary to another distant region. By binding to these two regions by complementary base-pairing the scaffold is brought to shape. When this is systematically performed any desired structure, like rectangular, smilies,[46] blocks [89] or even tunable arms can be designed.[90]

These self-assembled nano-sized objects can then also further be utilized as a bread-board for other structures, like enzymes,[91, 5] gold nanoparticles,[92, 7] or dyes for super-resolution microscopy.[93] By utilizing also different approaches switchable structures can be designed, by altering the salt concentrations,[94] the electric field,[90] or with the combination of switchable peptides such as ELP.[3] With all that different proof of principles, the DNA-origami technique is a simple and robust implementation of self-assembly of artificial designs in the lab.

# 3 Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch

The text and figures in this chapter are adapted from the following publication:

"Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch"

**Kilian Vogele**, Jonathan List, Günther Pardatscher, Nolan B. Holland, Friedrich C. Simmel and Tobias Pirzer, ACS Nano, 10, 11377-11384 (2016) Copyright 2020 American Chemical Society

In order to provide a structure better suited to this thesis, certain sections have been merged and rearranged or renamed.

In this chapter nanoscale plasmonic waveguides composed of gold nanoparticles and DNA-Origami structures are assembled. With such a waveguide structure electromagnetic energy can be guided below the optical diffraction limit. Typically, such waveguide structure are mainly fabricated by lithographic methods and, mostly operated using electronic effects or near-field optical techniques as the signal feed-in and readout. For a self-assembled structure in bulk, such techniques are problematic, therefore a simple spectroscopic excitation readout is facilitated here.
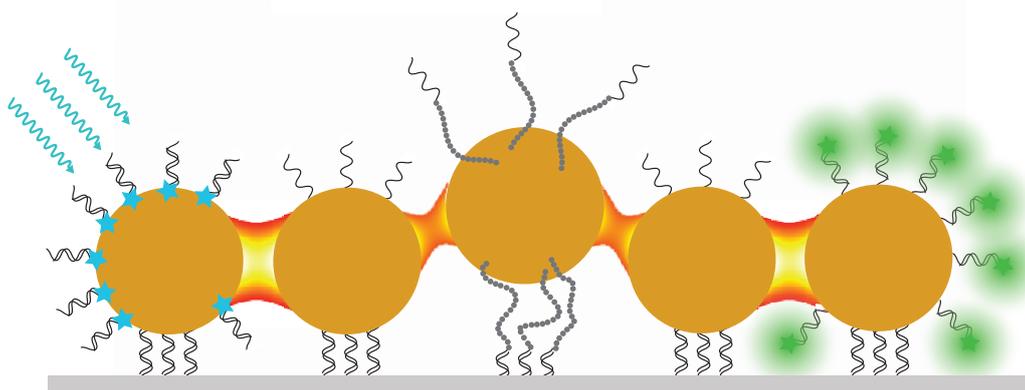


Figure 9: Schema of the plasmonic waveguide system with a donor dye (turquoise) and an acceptor dye (green) at each end of the waveguide and a gold nanoparticle modified with a thermoresponsive peptide-DNA hybrid as a switch in the middle.

The gold nanoparticles at the end of the waveguide are equipped with fluorescent dyes, whereat the first particle is modified with a donor dye and the last particle is modified with an acceptor dye, which is used to detect energy transfer. The distance between both dye-labeled nanoparticles is beyond the multicolor FRET range and below the Abbé limit. With the combination of a thermoresponsive peptide, as shown in Figure 9, a switchable waveguide is assembled.[7]

## 3.1 Introduction

Nanoparticle plasmons are collective electromagnetic excitations of mobile charges in metals confined to nanoscale volumes. Since they can be used for the collection and manipulation of optical signals below the diffraction limit, they attract growing interest.[95, 96, 97, 98, 99, 100, 101] In contrast to bulk metal surfaces, single nanoparticles can be experimentally and theoretically treated as plasmonic atoms that enable the well-defined creation of plasmonic molecules or nanostructures.[102] In close proximity the local surface plasmon resonances (LSPRs) of nanoparticles couple and enable optical properties such as energy transfer. The most popular material used for plasmonic nanostructures is gold, which shows on one hand strong energy dissipation for LSPR coupling and in consequence a low efficiency in energy transfer.[103, 104, 105] On the other hand, AuNPs can be easily functionalized with biomolecules such as DNA and peptides, and are experimentally well characterized, their interactions are predictable by theory, and hence they are generally well suited to create prototypes for photonic circuits.

In order to benefit from these advantages, AuNPs have to be precisely positioned and arranged. In the last years metallic nanoparticles have been successfully assembled into plasmonic waveguides for electromagnetic energy transfer. Nanoparticles were arranged mainly via lithographic methods, which provide the required precision,[102, 106] but at the same time the use of lithography pins the waveguides onto their solid supports. Also the local detection of energy transfer by optical near-field techniques limits the versatility of these waveguides.

Alternatively the DNA origami technique has already been proven suitable for controlling crucial parameters for LSPR and to create plasmonic nanostructures such as nanoantennas,[107, 108] nanoparticle heterodimers,[109] 3D plasmonic metamolecules,[110] and other plasmonically active structures.[111, 112, 92, 113, 114] However, the fabrication of larger, heterogeneous, and thus more complicated plas-

monic nanostructures remains challenging. In particular, for their utilization in a biological context, for instance to create sensory devices for nanomedical applications, manageable methods have to be developed. A possible strategy is one-pot self-assembly, which does not require external interference but still achieves sufficient yield and precise alignment.

In this work, we present a plasmonic waveguide created by the assembly of AuNPs that allows the transfer of excitation energy over a distance of approximately 50 nm. Our waveguide system utilizes a single-layered rectangular DNA origami structure as a molecular circuit board with five AuNPs lined up to achieve a coupling of the LSPR within the particles (Figure 10). In contrast to previous studies, fluorescent dyes are used for both excitation and visualization of plasmonic energy transfer. In order to achieve a high yield, the required alignment accuracy, and specific particle addressability, we developed a solid-phase-based assembly strategy using silica microparticles.

## 3.2   Results and Discussion

Our waveguide experiments were carried out exclusively in aqueous bulk solution utilizing conventional fluorescence spectrometers. In the visible spectrum quenching of fluorescent dyes by AuNPs and thus energy transfer from dyes to particles is most efficient at very short distances.[103, 115, 116, 117] For a successful excitation of plasmons a spectral overlap between the emission of the dye and the LSPR is required. We thus used fluorescein (FAM), which has a suitable overlap with the waveguide resonance, as a "feed-in" for the waveguide (see Supporting Information section 9.2.1). It is located in direct proximity to the gold surface by hybridization via DNA linker L1 as indicated in Figure 10a.

The actual transport of electromagnetic energy along the AuNP chain progresses by near-field coupling of LSPR between adjacent AuNPs. Detection of successful waveguide operation is enabled by surface plasmon coupled emission (SPCE) of an acceptor dye (Atto 532), a fluorescence enhancement due to the presence of plasmons, at the terminal end of the waveguide.[103][106] SPCE most efficiently occurs at the farther distance from the gold surface,[115, 117] and therefore the acceptor dye is attached to the terminal AuNP with DNA linker L2 acting as a spacer.

Since the spectra of both dyes overlap with the plasmon resonance spectrum, FAM
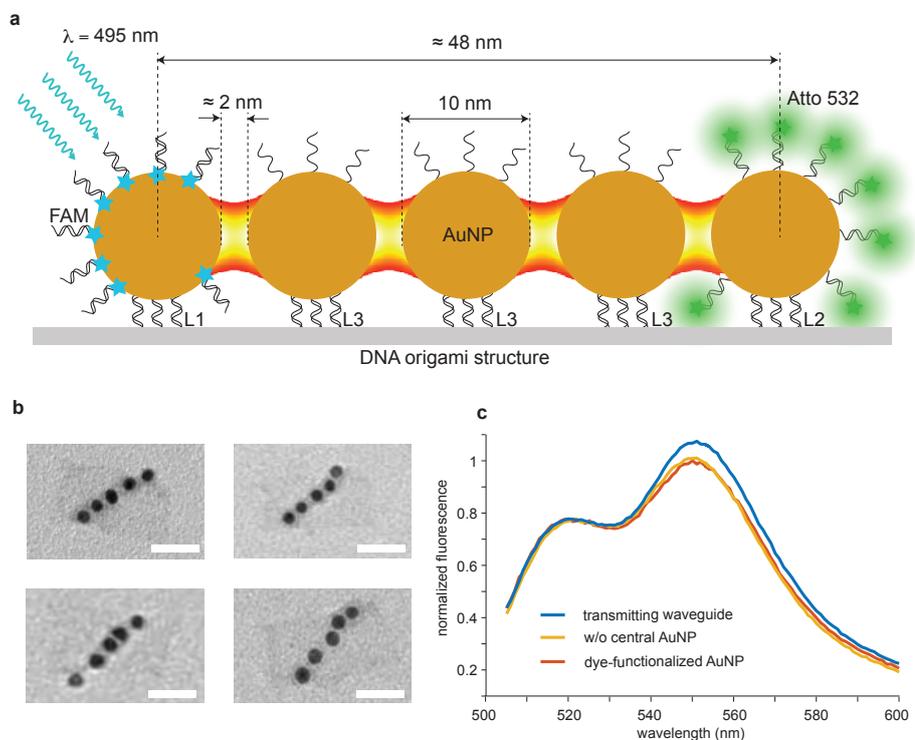
Figure 10: Plasmonic waveguide system. (a) Schematic of the waveguide on a DNA origami rectangle showing the linear arrangement of five gold nanoparticles. The waveguide is excited by donor fluorophores (FAM) on the left side and transfers energy via plasmonic coupling to acceptor fluorophores (Atto 532) on the right, which serve as a read-out. The plasmon coupling between the particles is color-coded, with yellow indicating high electrical field enhancement (good coupling) and red indicating lower electrical fields. (b) Typical TEM images of a plasmonic waveguide assembled on the DNA rectangle. Scale bars: 50 nm. (c) Combined emission spectra of FAM and Atto 532 for a sample containing single dye-functionalized AuNPs (red), for a nontransmitting waveguide that lacks the central AuNP (yellow), and for a fully assembled waveguide (blue). The spectra were normalized to the maximum FAM emission signal at $\lambda = 520$ nm.

and Atto 532 get directly excited at 490 nm and are partially quenched but still detectable. In Figure 10c, the spectrum of a fully assembled waveguide system is compared with the spectrum of a waveguide containing a defect in the middle (i.e., the central AuNP is missing); the latter is identical to the spectrum of a

38

single AuNP functionalized with FAM and Atto 532, respectively. With all five AuNPs present, the Atto 532 emission peak at 550 nm is increased with respect to the defect-containing system, indicating a successful transfer of energy from FAM to Atto 532 via the waveguide and a transfer effect of about 10 %. A possible fluorescence enhancement of the Atto dye mediated by a spectral red-shift of the transmitting waveguide can be can ruled out. Since the spectral overlap of Atto 532 and AuNPs rather decreases than increases with a red-shift, it would instead result in a reduced transmittance (see also Supporting Information section 9.2.1).

Apart from matching the emission and absorption spectra, the precise alignment of AuNP on the DNA structure is the most crucial point in achieving efficient through-waveguide energy transfer that can be detected in bulk. Since the interactions of the plasmons within the particle chain scale exponentially with distance,[118, 119] we employed the DNA origami technique to achieve the necessary proximity and nanometer accuracy.[120, 121, 122, 123, 124] A twist-corrected DNA origami rectangle (90 nm × 60 nm) was equipped with five binding sites with three anchor strands each.[92] In order to maximize LSPR, a interparticle gap of approximately 2 nm was chosen in the DNA origami design. With standard functionalization procedures, it is challenging to achieve precise alignment and high yield at the same time.[114, 125, 126, 127, 128] For instance, when using closely spaced binding sites, single AuNPs often bind to two neighboring binding positions simultaneously. Furthermore, a large excess of AuNPs has to be used to prevent interstructure cross-linking, followed by laborious purification steps. In order to overcome these issues, we developed a convenient functionalization protocol based on a solid-phase assembly process. An overview of the complete functionalization process is given in Figure 11 (see also Supporting Information section 9.2.1). In our approach, we utilize polydisperse silica particles (about 0.3-8 $\mu$m in size) as a solid support for DNA origami structures. In order to adsorb the DNA structures onto the negatively charged silica surface, they are gently mixed with suspended silica particles in the presence of 60 mM $MgCl_2$. Subsequently, DNA-coated AuNPs are added to the particles and gently mixed for at least 2 min. The maximum functionalization yield per site of 99.7 % is obtained at an excess of three nanoparticles per binding site and an incubation time of only 2 min. For longer incubation times the binding yield slightly decreases (see Figure 11b), probably as a result of abrasion during mixing. As anchor strands we preferentially used poly(dA) and poly(dT) sequences, respectively, as they do not exhibit any secondary structure and therefore provide

higher binding yields. The functionalization yields were determined using AFM imaging (see also Supporting Information section 9.2.1). The protocol works best for single-layered DNA origami structures, where it is irrelevant in which orientation the structures adsorb onto the silica surface. Even for structures with two DNA origami layers DNA anchor strands are still accessible due to threading through the origami mesh.[129] Figure 11c shows examples for AuNP-functionalized origami cross [130, 131] and rectangle structures.

For the fabrication of a plasmonic waveguide with three different types of gold nanoparticles attached, namely, the excited AuNP, the transmitting AuNPs, and the emitting AuNP (Figure 10a), DNA anchor positions with mutually orthogonal sequences were used (see Methods section). With less repetitive anchor sequences we obtained a slightly reduced assembly yield of 56 % for the complete waveguide using TEM (Supporting Information Figure 12) and gap sizes of $2.1 \pm 1.5$ nm, which are influenced by variances in AuNP diameter and shape (see also Supporting Information section 9.2.1). The transfer effect through the whole waveguide is only approximately 10 %, which can be rationalized by the strong energy dissipation in the AuNPs, the considerable variance of the gap sizes, and the slightly reduced assembly yield.

As an example of a more complex plasmonic assembly, we next introduced a reversible nanomechanical switch into the waveguide. We functionalized the central AuNP with thermoresponsive elastin-like peptides (ELPs) to make the transmittance conditional on an external stimulus. ELPs are synthetic peptides that show a fully reversible hydrophobic collapse when the temperature is increased above a certain transition temperature $T_t$.[132, 70, 69, 62] $T_t$ typically depends on amino acid sequence, peptide concentration, solvent conditions, etc. Here we used ELPs with the repetitive sequence $(GVGVP)_{40}$ with a cysteine at the C-terminus (abbreviated as V40C) and coupled them to oligomeric DNA linkers L4 (poly(dA)) via copper-based azide-alkyne cycloaddition. Utilizing the cysteine, the diblock copolymers L4-V40C were attached to AuNPs via gold-thiol bonds. Subsequently, the ELP-AuNP conjugates were positioned in the center of the waveguide via linker L4. The yield for the assembly of the ELP-enhanced waveguides -now each containing four distinct AuNP binding sites- slightly dropped to 45 % (see Supporting Information Figure 13).

The mechanism of the temperature-driven nanomechanical switch is illustrated in
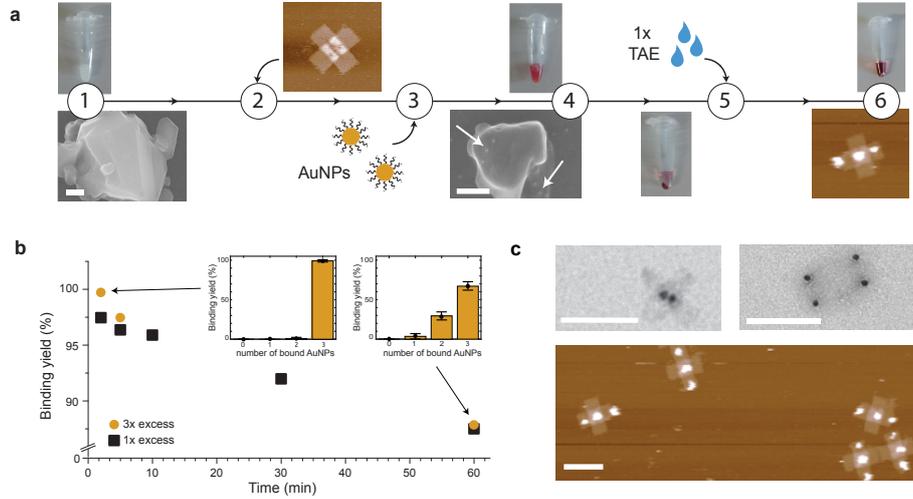
Figure 11: Silica-based functionalization protocol of DNA origami structures. (a) The schematic of the process illustrates the developed silica protocol: Initially, silica particles are suspended in TAE with 60 mM $MgCl_2$ (1) followed by mixing with the desired DNA origami structures (2). After the addition of DNA-functionalized AuNP (3) the suspension is incubated while shaking. After successful functionalization (4) the DNA structures are centrifuged and eluted by replacing the supernatant with TAE (5, 6). On silica adsorbed structures are indicated by white arrows. The assembled AuNP-DNA structures are separated from the silica particles by centrifugation and are located in the supernatant. The images of the silica particles are acquired by SEM, and DNA origami structures are visualized using AFM. Scale bars: 100 nm. (b) The binding yield per binding site decreases slightly for longer incubation times, probably due to abrasion effects. Insets show the binomial distributions of the binding yield for incubation times of 2 and 60 min for analyzed cross-shaped structures with three binding sites. The given uncertainties are standard deviations. Black dots indicate the theoretical binomial distributions. (c) AFM image (top) of fully equipped cross-shaped structures and TEM images (bottom) of completely functionalized cross-shaped and rectangular structures using the silica protocol. Scale bars: 100 nm.

Figure 12a. At a temperature below $T_t$ the ELPs are highly solvated and act as extended tethers for the central AuNP, which reduces the plasmonic coupling between the particles. In their collapsed state above $T_t$, the ELPs occupy a smaller volume and thus act as a shorter tether. As a consequence, the central AuNP is

localized more firmly between the adjacent particles, reducing the distance between the AuNPs and, hence, increasing the transmittance. The temperature dependence of the dyes was eliminated by normalization with a blank signal (see Supporting Information section 9.2.1). The fluorescence spectrum of the switchable waveguide is expected to be intermediate between the defect-containing system and the transmitting waveguide from Figure 10c. The aforementioned red-shift can again be ruled out, but since the permittivity in close proximity to the central AuNP is changed due to the switching between the hydrophilic and hydrophobic state, its influence cannot be excluded.

In order to prove that the altered transmittance originates from a controllable temperature effect of the ELP, we utilized the salt dependence of the transition temperature $T_t$. In TAE (40 mM Tris, 20 mM acetate, 1 mM EDTA) with 12.5 mM $MgCl_2$ the $T_t$ was determined to be 22.0 °C, whereas after the addition of 300 mM NaCl the $T_t$ shifted to 18.6 °C, which demonstrates the so-called $T_t$ mechanism of ELPs.[70] Figure 12b shows the temperature-dependent transmittance of the switchable waveguide tuned by the addition of NaCl. A waveguide without ELP attached shows no temperature effect. Compared to their bulk counterpart (see Supporting Information Figure 5) the measured $T_t$ values for the waveguide are lower, which is caused by a higher local concentration and less free volume.[70, 62] The evident change in the slope of the transition is observed only in experiments with the assembled waveguide. For ELP alone and L4-V40C-AuNP conjugates the addition of NaCl does not alter the slope (see Supporting Information Figures 5, 6, 7). It is known that the slope of the transition often changes with $T_t$,[133] where lower $T_t$ values show a steeper slope presumably due to a higher cooperativity. Since the transition of ELPs is fully reversible, we could also demonstrate repeated switching of transmittance by temperature cycling between 15 and 35 °C with a temperature gradient of 1 °C/min (see Figure 12c and Supporting Information Figure 4).

## 3.3 Conclusion

In summary, we have introduced a solid phase-based technique that allowed us to functionalize DNA nanostructures with nanoparticles with high efficacy and speed. Decoration of DNA nanostructures with gold nanoparticles required less than 20 min with a yield of almost 100 % and a reduced consumption of materials. The extraordinary yield of the functionalization method allowed us to create plasmonic waveg-
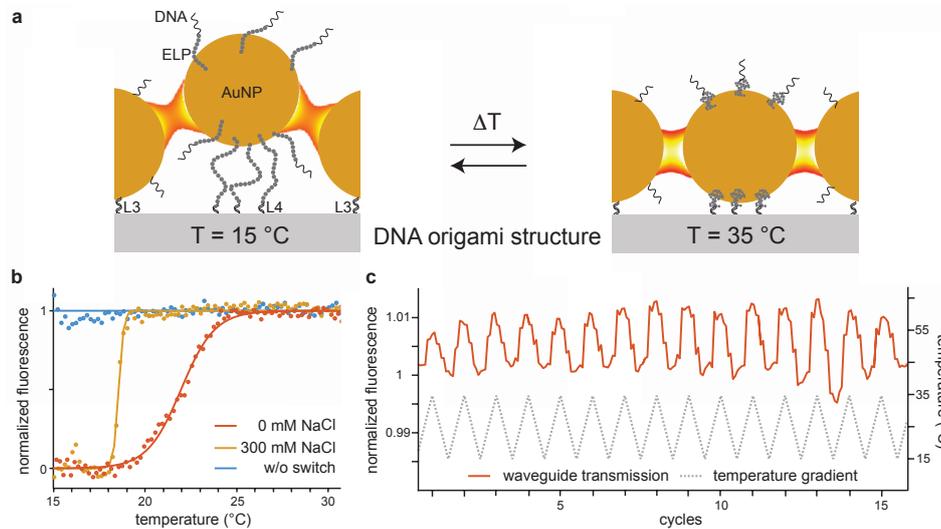
Figure 12: Thermoresponsive switching of a plasmonic waveguide. (a) The illustration shows the operating principle of the ELP switch. With elongated ELPs at 15 °C the central AuNP is less localized between adjacent AuNPs (left) than for 35 °C (above $T_t$) with collapsed ELPs (right). The plasmon coupling between the particles is color-coded, with yellow indicating high electrical field enhancement (good coupling) and red indicating lower electrical fields. (b) Fluorescence spectroscopy curves measured using a waveguide without implemented switch (blue dots) as a control and a switchable waveguide in TAE with 12.5 mM $MgCl_2$ without NaCl (red dots) and with 300 mM NaCl (yellow dots). All curves were normalized using a blank signal with dye-functionalized AuNPs. The specific $T_t$ values were determined from the inclination point of sigmoidal fits (solid lines). For the control a linear fit was used. (c) Fluorescence signal of a waveguide in TAE with 12.5 mM $MgCl_2$ without extra NaCl (red) that is repeatedly switched between its two states by temperature cycling between 15 and 35 °C (gray). All traces were normalized to the maximum emission of the dye-functionalized AuNPs.

uides composed of AuNP chains on DNA origami platforms that could be excited via donor and read out via acceptor fluorophores using a standard bulk spectrometer.

We further developed an active plasmonic waveguide including elastin-like peptide components that could be thermomechanically regulated and tuned by changing the ionic conditions. It is also conceivable to use more advanced stimuli such as

43

light, magnetism, or biochemical signals.[70] While our self-assembled, switchable plasmonic waveguides cannot compete with the response time of solid-state devices, they are excellently suited for biological applications, which require remote excitation and sensing. In future work, the electric field polarization along the waveguide could be utilized to alter bond energy levels analogous to Stark spectroscopy. Chemical reactions or conformations of proteins located at the waveguide's terminal end could be probed in their natural environment and without sophisticated equipment. The required resonance wavelengths could be adjusted by using nanoparticles with different composition and diameter. It is also readily conceivable to extend our approach to the creation of AuNP-based light-harvesting systems.

# 4 Genetically Encoded Membranes for Bottom-Up Biology

The text and figures in this chapter are taken from the following publication:

In order to provide a structure better suited to this thesis, certain sections have been merged and rearranged or renamed.

In this chapter, several approaches of synthetic biology to create a self-replicating cell-mimicking systems are discussed. In the *bottom-up* approach an artificial cell is build from scratch. Here compartmentalization of all the necessary components is essential, as shown in figure 13. To enable the growth of the compartment the
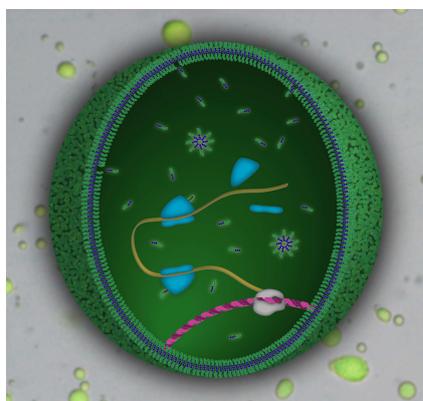


Figure 13: Illustration of a peptide based self-replicating cell-mimicking system, with fluorescence giant peptide vesicles in the background. The artwork is provided by Marisa A. Goetzfried.

building blocks can be either externally provided or generated *in situ* through internal processes. Of particular interest for an autonomously self-replicating system is the genetically encoding of the necessary proteins, which are required to produce the building blocks. Here the challenges associated with two different approaches, namely the generation of phospholipid via genetic and enzymatic processes and the expression of peptide-based membrane are discussed.[2]

## 4.1  Introduction

Artificial living systems are often conceived as compartmentalized chemical systems that are able to grow and divide, and at the same time replicate and pass on genetic information - which in turn would convey the potential for Darwinian evolution.[134] The creation of such systems necessarily involves the realization and study of out of equilibrium chemical reaction networks that control molecular self-assembly and self-organization processes, which are central topics of interest in systems chemistry. A key challenge in this context is the encapsulation of the systems and the coupling and coordination of their internal dynamics (gene replication and metabolism) with the dynamics of the compartment resulting in growth and division. A variety of theoretical models have been formulated that attempted to capture the basic requirements for self-replication of protocellular systems that would give rise to life-like dynamics (Figure 14). Ganti's well-known chemoton model consisted of three coupled self-reproducing reaction cycles - a metabolic cycle coupled to a gene replication and a membrane formation cycle.[135] Other schemes have been developed, in which membrane formation was driven by osmotic effects and the external addition of monomers.[136, 137] Various researchers have worked on different aspects of protocell dynamics in the past [138, 139, 11] such as gene expression and replication of genetic polymers,[140, 141, 142, 143, 144, 145, 37] reconstitution of metabolic processes,[146, 33] or growth and division of membranous vesicles.[147, 148, 149, 36, 150] However, the creation of systems, in which these different dynamics consistently play together is extremely challenging. In particular, the production of membrane components in situ is a challenge, and most of the systems studied so far were either static (non-growing), or fed externally. In this perspective, we will discuss various approaches towards compartmentalization of artificial cellular systems, with a focus on attempts to encode membrane production genetically.

## 4.2  Compartmentalization

### Membrane Components

Modern cells are encapsulated by phospholipid bilayers, which are densely packed with other non-lipid components such as transmembrane proteins and pores, peptidoglycans, etc. In fact, only about 50 % of the mass of membranes is actually
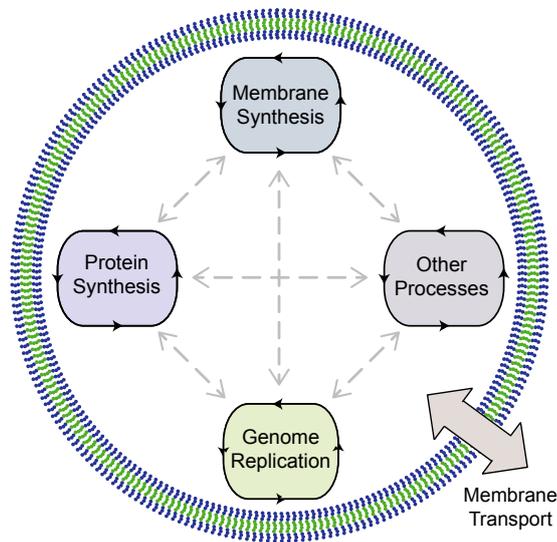
Figure 14: Illustration of an abstract model for a synthetic cell involving several linked reaction/replication cycles. Growth of cells is associated with genome replication, an increase in membrane area and internal volume/ contents. Next to membrane production this requires additional metabolic functions (in particular energy metabolism), and potentially also mechanical functions and the possibility to communicate with the environment. Depending on the nature of the reaction cycles (and its degree of autonomy), the cell has to be provided with different sets components from the outside.

contributed by phospholipids.[151] While there has been considerable recent interest in membrane-free compartmentalization via liquid-liquid phase separation or coacervation,[140, 152, 153, 154, 155] most of the research on protocells and artificial cells has previously focused on bilayers created from amphiphilic molecules.

Apart from phospholipids, other amphiphiles were considered, among them prebiotically more plausible fatty acids,[156] natural as well as synthetic peptides,[70, 157] or non-biological components such as block copolymers or other amphiphiles created by chemical means.[143, 150]

These types of membranes differ in various physicochemical aspects, which are important for the overall behavior of the artificial cells created from them. For instance, fatty acid vesicles are more dynamic than phospholipids (e. g., when considering "flip-flop" events of membrane monomers), which also makes them leaky. This is an advantage for the exchange of molecules with the environment, but pre-

47

vents the establishment of chemical gradients, which would be important from a bioenergetic perspective.[158] Also membranes made from peptides or polymers differ widely in dynamics, permeability and mechanical stability.

**Self-Assembly of Membranes and Compartments**

Membrane formation by amphiphiles is an extensively studied self-assembly phenomenon, which is driven by both entropic and enthalpic effects.[159, 48] This involves the hydrophobic effect that leads to the aggregation of the hydrophobic parts of the amphiphiles as well as attractive interactions between their polar headgroups and the surrounding aqueous solvent. Above a critical aggregation concentration (cac) the amphiphiles - depending on the geometry of the molecules - assemble into micelles, vesicles, bilayer sheets or tubular structures. Importantly, the contributing interactions are rather weak, and monomers and superstructures coexist in a dynamic equilibrium.[48] The resulting dynamic nature of membranes formed by amphiphiles is an important property for the realization of compartments that are able to grow and divide, which distinguishes them from more rigid structures such as capsids and protein cages.[160]

**Techniques for Vesicle Formation and Encapsulation**

Before one can start to investigate vesicle growth and division in artificial cellular systems, an initial population of vesicles has to be generated. In the lab, several methods are available for the formation of vesicular compartments from amphiphiles, which can be roughly divided into solvent displacement and rehydration methods.[161]

Solvent displacement methods proceed through dissolution of the amphiphiles in an appropriate organic solvent, followed by direct transfer into an aqueous phase or by the intermediate formation of water-in-oil emulsion droplets. Simple solvent injection, in which amphiphilic molecules in an organic phase are added to aqueous solution under vigorous stirring, typically results in small unilamellar vesicles (SUVs). By contrast, in emulsion phase transfer methods water-in-oil droplets are generated and then added to an oil-water biphasic system, in which the oil phase also contains amphiphiles. Migration of the droplets through the oil-water interface results in the formation of a water-oil-water (wow) double emulsion.[35]

Emulsion phase transfer can also be performed in microfluidic systems, resulting in relatively monodisperse SUVs or giant unilamellar vesicles (GUVs).[162] While solvent displacement methods afford quite efficient encapsulation, they have the disadvantage of organic solvent remnants between the leaves of the formed bilayers. This problem does not arise in rehydration methods, which usually lead to SUVs and multilamellar vesicles (MLVs) with high polydispersity. Polydispersity can be reduced, however, by using extrusion methods. Electroformation can be used to produce GUVs with relatively low polydispersity [163] and proceeds from lipid films formed on metal (Pt) electrodes, which are subjected to an alternating electrical current.[164, 165] Electroformation has also been shown to work for polymersomes,[166] but the method is relatively limited by the low ionic strength required of the rehydration solution. In gel assisted rehydration, a hydrogel is coated with amphiphilic molecules and rehydrated with buffers of any ionic strength, resulting in GUVs with relatively low polydispersity. Notably, rehydration methods are consistent, in principle, with wet/dry cycles, which also play a role in various origin of life scenarios.[167, 168, 169, 170]

## 4.3   Growth and Division of Cell-Sized Compartments

Due to their dynamic nature, micelles and membranes formed by amphiphiles are amenable to growth, division and fusion. Growth necessarily requires the incorporation of amphiphilic molecules into an existing membrane, which in the context of artificial cellular structures has been previously accomplished either by external feeding, or by employing chemical reactions that generate amphiphiles from precursors.

### Growth by External Feeding and Fusion

Fatty acid (FA) micelles and vesicles have been extensively studied in the context of prebiotically plausible compartmentalization.[140, 148, 156, 171, 172, 173, 42, 174, 175] Fatty acid assemblies are highly dynamic and they quickly exchange fatty acid monomers with their environment. It has also been demonstrated that fatty acid vesicles can grow by feeding them with fatty acids or micelles.[174, 176] In a kinetic study it has been shown that growth of vesicles by micelle uptake is much faster than de novo formation of vesicles.[177] In contrast to FA membranes, phospholipid

(PL) membranes - more similar to modern cell membranes - are quite stable and do not spontaneously take up monomers from their environment. It could be shown, however, that it is possible to induce PL vesicle growth through fusion, which can be supported by proteins (e. g., SNARE proteins in modern cells [178, 179]), or other bridging molecules (such as DNA [180]) that bring two vesicles into close proximity. Fusion has also been shown to spontaneously occur between vesicles with oppositely charged membranes,[181, 182, 183, 184] and it has been successfully induced by applying pH changes.[185]

Growth of vesicles can be promoted by physical means such as electric fields or osmotic pressure differences between the interior and exterior solution. In the context of artificial cells, osmotic pressure differences arise as a result of biopolymer synthesis, which traps the counterions of the newly formed polyelectrolytes inside of the compartments.[176]

**Growth Driven by Chemical Reactions**

Apart from feeding scenarios, a wide variety of chemical approaches have been developed to drive vesicle growth, in particular in the context of abiological membrane forming molecules (for a review cf. Ref. [150]). The basic concept here is to generate membrane-forming amphiphiles through a chemical reaction. In early work, oleic acid/oleate vesicles have been shown to grow via autocatalytic hydrolysis of oleic acid esters or oleic acid anhydride.[186] Click chemistry has been used to join two single-chain amphiphiles (only forming micelles) to create a double-tailed amphiphile that is capable of creating or inserting into membranes.[175] Similar approaches were developed using chemical ligation or condensation reactions.[187, 188, 189, 190]

There also have been various attempts to reconstitute parts of the natural phospholipid synthesis pathway in the context of artificial cell membranes.[167, 191, 192, 193, 194, 195] Phospholipid synthesis involves membrane-inserted enzymes such as sn-glycerol-3-phosphate acyltransferase (GPAT, PlsB in *E.coli*) and lysophosphatidic acid acyltransferase (LPAAT, PlsC in *E.coli*), which convert glycerol-3-phosphate (G3P) into phosphatidic acid (PA), which can be further transformed into phosphocholine (PC) and phosphoethanolamine (PE) lipids. Using PlsB, PlsC and the water-soluble enzyme FadD10, which catalyzes the adenylation of long-chain fatty

acids into fatty acid adenylates, membrane expansion of liposomes has been demonstrated via cell-free phospholipid production starting from oleic acid.[195] Recently, a minimal biochemical route towards the creation of phospholipids was demonstrated also involving FadD10, in which FA adenylates were shown to spontaneously react with amine-functionalized lysolipids to form phospholipids.[196] Enzymatic reactions are of particular interest as they can also be encoded, in principle, into the genome of the artificial cells (see below).

**Budding and Fission**

Cell division is another essential prerequisite for a continuous compartmental replication cycle. In the absence of a sophisticated cell division machinery, purely physical processes may facilitate the division or budding of daughter compartments from vesicles that have grown large through one of the processes described above.[197, 198, 199, 200] Shape transitions of fatty acid or phospholipid vesicles can be induced thermally,[201, 202] or due to mechanical processes such as shearing.[147, 203] For instance, heating can change the excess area (the difference in the areas of the outer and inner bilayer leaflet), and thus generate shape transitions that lead to budding.[201, 204] Fission can also be induced isothermally through a change in osmotic pressure.[147]

In the case of mixed phospholipid/fatty acid vesicles, fission has been demonstrated to occur upon application of gentle shear flows.[175] Extrusion of vesicles through membranes with small pores is another physical approach to enforce fission,[174] which is also used in the lab to generate vesicle populations with a narrow size distribution. Both thermal and mechanical processes can potentially occur in the context of nonequilibrium cycles, e. g., during phoretic motion in gradients occurring in thermal diffusion columns.[205, 206]

Remarkably, proliferation via blebbing, tubulation and vesiculation has also been observed in the context of cell-wall deficient "L-form" bacteria.[207][208] This type of proliferation does not require the FtsZ-dependent cell division machinery and has therefore been speculated to represent a primordial form of bacterial cell division.[74, 209]

## 4.4 Genetic Encoding of Membrane Components

One of the central challenges for the realization of autonomously growing artificial cells that are reminiscent of modern biological cells is the production of the compartment boundaries by the artificial cell itself. This could, in principle, also facilitate a tight coupling of the genetic and the compartmental replication cycle as envisioned in some of the protocell scenarios described above.

### Genetic Encoding of Lipid Metabolism

The most obvious approach to achieve this would be to encode the complete phospholipid anabolism on the genome of the artificial cell. The first stage of phospholipid synthesis is the synthesis of fatty acids, which in most bacteria is catalyzed by the multienzyme complex fatty acid synthase II (FAS II). FAS II converts acetyl-CoA and malonyl-CoA into long chain fatty acids, where acetyl-CoA and malonyl-CoA themselves are derived from pyruvate, which is a key product of glycolysis. Fatty acids are further converted into zwitterionic phospholipids such as phosphatidylethanolamine (PE) or anionic glycerophospholipids and cardiolipin.[210] Cell-free expression of all enzymes involved in phospholipid synthesis and their proper (co)operation would be quite challenging for the simple artificial cells that have been created thus far. A less ambitious goal therefore is the reconstitution of just parts of (or alternatives to) the phospholipid pathway, and various research groups have attempted to do this in the past (cf. also [167, 191, 192, 193, 194, 196]). In an early study, Murtas et al. realized the production of the membrane-bound enzymes GPAT and LPAAT (see above) using cell-free protein synthesis inside of liposomes.[211] A remarkable step towards genetically encoded PL synthesis was then taken by Danelon and co-workers [36] who demonstrated *in vesiculo* synthesis of GPAT and LPAAT and six other PL pathway enzymes using the PURE [34] cell-free gene expression system (Figure 15a). This allowed them to synthesize both PE and PG lipids *in situ* starting from externally added acyl-CoA and glycerol-3-phosphate as precursors. It could be shown that an amount of lipids was newly synthesized that would be sufficient for an increase in membrane area of about 1 %. Devaraj and co-workers recently demonstrated that also the enzyme FadD10 that is not naturally involved in phospholipid biosynthesis can be expressed in the PURE system and leads to *de novo* phospholipid membrane formation when lipid-forming precursors are added.[196]

**Peptide-Based Artificial Cells**

An entirely different approach towards genetically encoded membrane-formation was taken by several other groups,[212, 213, 214] including ours,[6] that essentially circumvented the complexities of fatty acid and phospholipid biosynthesis. As noted above, rather than using fatty acids or phospholipids it is possible to generate membranes and vesicles also with other amphiphiles. This is particularly true of amphiphilic peptides, as they can, in principle, be genetically encoded and synthesized by cell-free transcription/translation reactions.

An interesting class of peptides in this context are the so-called elastin-like polypeptides (ELP).[70, 215, 216] ELP have the common amino acid sequence $(GXGVP)_n$, where X can be any amino acid except proline. The variable X can be used to tune the properties of the peptides, e. g., whether they are more hydrophobic or hydrophilic. Using this property, Vogele and co-workers created an amphiphilic elastin-like polypeptide composed of a hydrophilic glutamic acid-rich first block (mainly X=E) and a hydrophobic phenylalanine-rich second block (mainly X=F), which had been shown to be capable of self-assembling into small peptide vesicles. Importantly, it was shown that a cell-free gene expression system based on an *E.coli* cell extract [217] could be functionally enclosed inside of the peptide vesicles. When co-encapsulating a genetic template for the membrane-forming ELPs, new peptides could be generated *in situ* (Figure 15b), which led to a measurable vesicle growth. From the experiments, a more than four-fold increase in total membrane area was derived.

One advantage of this approach is the fact that it does not require any lipid metabolism at all, and only peptides have to be generated by the cell-free expression reaction rather than the metabolic enzymes. Using peptide membranes, artificial cells exclusively based on nucleic acids and proteins thus appear feasible. At this point it is not clear, however, whether peptide membranes could sustain a proton gradient, and thus also allow for ATP generation via oxidative phosphorylation.

## 4.5 Challenges for Genetically Encoded Growth

Even though the recent advances in genetically encoded *in situ* membrane production seem promising, a number of formidable challenges still have to be overcome in order to establish autonomously growing and dividing artificial cellular systems. For instance, all the examples described in the previous section relied on cell-free
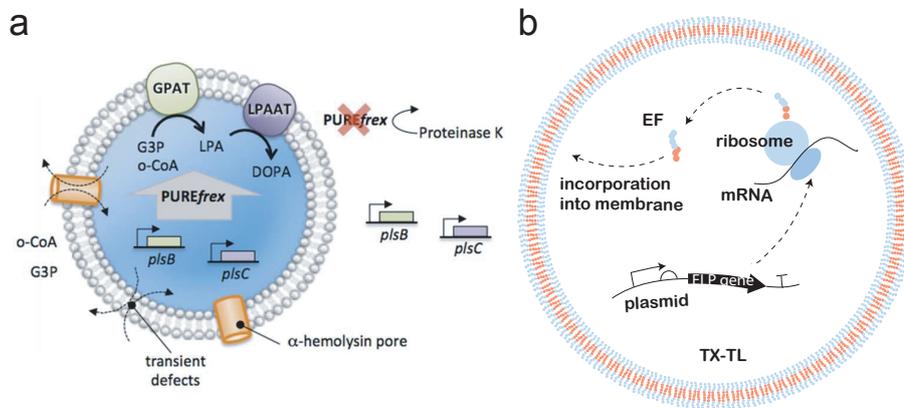
Figure 15: Recent examples for genetically encoded *in vesiculo* production of membrane components. (a) Lipid biosynthesis inside liposomes. Through expression of membrane proteins GPAT and LPAAT utilizing PUREfrex at the inside, the production of the lipids LPA and DOPA has been succeeded. (Reprinted from Ref. [36] under license CC BY 4.0). (b) Production of the peptide membrane constituent utilizing a crude bacterial cell extract. After expression the peptide incorporated into the membrane and facilitated vesicle growth. Adapted from Ref. [6] under license CC BY 4.0).

gene expression systems, they had no metabolic activity and they did not have the capability to replicate their own genome.

Following different lines of argument, researchers have previously estimated the number of genes required for minimal cells somewhere in the range of 150-200.[218, 219] Lipid synthesis or potentially peptide membrane synthesis (as described above) would only represent a minor fraction of "what is needed" for self-sustained growth, while the transcription-translation machinery would constitute the largest part. For instance, the PURE gene expression system contains 36 macromolecular components,[34, 220] which do not yet include ribosomes and tRNAs, however. Ribosomes and tRNAs would require a set of more than 80 additional genes.[218] Nevertheless, it has been shown possible to assemble ribosomes in a cell-free context,[221, 222] which ultimately could result in the creation of self-regenerating expression systems. If, in practice, only smaller gene sets can be realized in the context of artificial cells, one will have to make the choice, which compounds to provide externally to compensate for degradation processes and dilution by growth - this could

be achieved, e. g., via fusion with vesicles loaded with these essential components. Sustained growth would also have to be supported by appropriate metabolic processes. As for biological organisms, various degrees of metabolic autonomy are conceivable. Most systems realized so far simply use up what is supplied initially (i. e., during encapsulation), which obviously is incompatible with continuously growing compartments. To ensure a continuous supply with chemical energy and precursor molecules from the outside, either permeation or transport through the membrane has to be realized, or feeding and fusion with appropriately loaded vesicles. Such an "auxotrophic" life style currently is the most realistic approach for the creation of self-replicating artificial cells.

There have been many attempts to reconstitute metabolic processes in artificial cellular compartments, e. g., systems in which bacteriorhodopsin or photosystem II was coupled to ATP generation by ATP synthase,[223] or systems, in which chloroplasts were encapsulated as artificial "organelles" to perform photosynthesis.[224] It is important to note that bioenergetic processes such as oxidative phosphorylation, ATP synthesis or photosynthesis are all based on enzymes incorporated into phospholipid membranes. For this reason it is not clear whether and how an autonomous energy metabolism could be implemented in self-replicating compartments encapsulated by other types of membranes.

Other major challenges for the realization of replicating artificial cells are the replication of the cell's genome,[141, 37] and, moreover, the coupling of genome and compartmental replication. Modern cell division is an extremely complex and highly regulated process, which would require a large number of additional molecular components to be expressed by the artificial cells. In the absence of a complex cell division machinery, one would have to rely on vesicle shape transitions leading to budding or blebbing, in analogy to L-form proliferation mentioned above. In this case, one would not be able to guarantee a proper distribution of the genome copies to the daughter cells. One possible way to prevent genome loss in this case could be the use of a larger number of genome copies per artificial cell.

It might not be necessary to implement a full gene set for replicative compartments right away, however. When there is sufficient coupling of genome and compartmental replication, artificial cells will compete with each other for externally supplied resources, i. e., small molecule precursors and also other macromolecular components essential for gene expression or metabolism as discussed above. This could potentially already lead to some kind of evolutionary optimization of the systems,

in which the fraction of the faster replicating or "fitter" compartments would increase over time. One could then try to gradually increase the complexity of the artificial cells by externally supplying additional biochemical functions - these could be provided either in the form of autonomously self-replicating genetic templates (such as plasmids) or via recombinant integration into the existing genome.

## Conclusion

The realization of growing and dividing artificial cell systems would be an immense technical achievement, which not only would be of fundamental interest, but could also lead to technological applications.

From a general perspective, the realization of growing and dividing compartments with genetically encoded membranes will provide insights into fundamental aspects of the coupled cyclic dynamics governing self-replicating chemical systems, which could then be studied experimentally in a systematic way. It would also provide an instructive laboratory example for the transition from non-living to living matter. As the schemes discussed above are mainly based on modern biochemical components, they would not provide direct insight into processes leading to the emergence of life on earth, however.

Self-replicating vesicles could also be of interest as autonomously evolving bioreactors. For instance, one could attempt to couple the production of a molecule of interest to the replication success of the reactor by making components of the replication machinery dependent on the presence of this compound via appropriate regulatory mechanisms. This should result in an evolutionary optimization process, in which the better performing bioreactors would also replicate faster.

There may be other situations, in which replication of artificial cells could be desired - for instance, in the context of artificial differentiation, in which certain cell types would replicate and grow into artificial tissues depending on cues from the environment. Also in the context of environmental or medical sensing, detection of certain compounds (or cells) could trigger proliferation of artificial cells for bioremediation or defense, in analogy to something like an immune response.

# 5 Towards Synthetic Cells Using Peptide-Based Reaction Compartments

The text and figures in this chapter are taken from the following publication:

"Towards Synthetic Cells Using Peptide-Based Reaction Compartments"

**Kilian Vogele**, Thomas Frank, Lukas Gasser, Marisa A. Goetzfried, Mathias W. Hackl, Stephan A. Sieber, Friedrich C. Simmel and Tobias Pirzer, Nature Communications, Volume 9, 3862 (2018) under license CC BY 4.0

In order to provide a structure better suited to this thesis, certain sections have been merged and rearranged or renamed.

In the previous chapter, different concepts for a self-replicating cell-mimicking system has been discussed, whereas in this chapter the experimental realization of a peptide-based system is in the foreground.
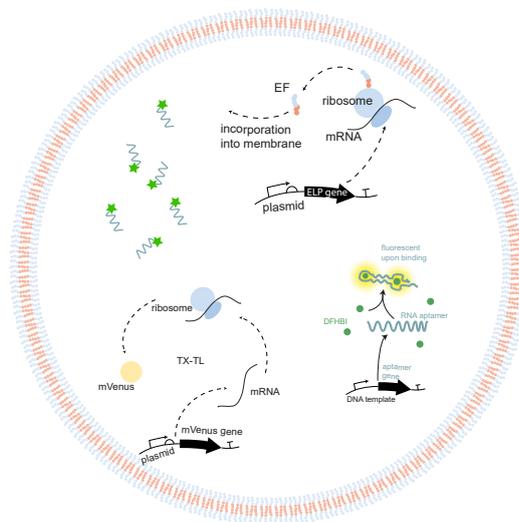


Figure 16: Illustration of the peptide vesicles based on the amphiphilic elastin-like peptides, with different internal subsystems for the detection of internal processes, like protein expression with a cell-free system, transcription of an RNA aptamer or the production of the membrane constituents themselves.

In living cells the information of the membrane constituents is encoded in the genome. In an artificially generated system capable of the production of their

own membrane building blocks a similar approach has to been taken. However, the implementation of the natural biochemical reaction networks to synthesize the membrane building blocks *in vitro* is notoriously difficult for lipids. To circumvent this an amphiphilic elastin-like peptide (ELP) is designed as a membrane molecule, which is capable to self-assemble into vesicular structures of about 200 nm diameter. With the combination of a cell-free transcription-translation system together with the DNA template inside the peptide vesicles the growth can be genetically encoded within. As a first proof of principle, the *in vesiculo* production of a functioning fluorescent RNA aptamer and a fluorescent protein is shown. Furthermore, it is shown that the membrane peptide itself is expressed and also the autonomous vesicle growth due to the incorporation of the newly produced ELP into the membrane, as shown in Figure 16.[6]

## 5.1 Introduction

Life is based on the complex interaction of numerous molecular components, which self-assemble and self-organize into higher ordered structures. Inspired by natural systems, several aspects of living cells have already been recapitulated *in vitro*, such as DNA replication, protein expression or compartmentalization of molecular reactions.[225, 33] Compartmentalized protein expression and DNA replication inside vesicles have been studied as well.[143, 226] For bottom up approaches towards the creation of protocellular compartments, these aspects need to be coupled and coordinated.

For instance, DNA amplification was coupled with vesicle self-reproduction,[143] or the genetically encoded synthesis of phospholipids inside lipid vesicles from precursor molecules such as acyl-coenzyme A and glycerol-3-phosphate.[36] Here, mainly phospholipids or fatty acids are used for compartmentalization.[36, 177, 144, 227] However, in principle also other suitable amphiphilic building blocks can serve as membrane material, such as peptides or synthetic block copolymers. For instance, Huber et al.[212] used a specifically designed amphiphilic elastin-like peptide (ELP) to form vesicles inside *E.coli*. The genetic template for the production of such peptides can be enclosed inside an *in vitro* reaction compartment and therefore directly linked to the expression machinery. In order to implement this in a cell-free context,[228, 229] transcription-translation systems based on purified components, e.g., the PURE system,[144, 34] or on bacterial cell extracts, e.g., the TX-TL

system,[230, 231] can be used. These systems employ the multicomponent bacterial translation machinery, to express proteins from externally added DNA in a one-pot reaction. For the transcription the T7 polymerase can be used as well as the present constitutive based transcription system in the case of the *E.coli* TX-TL system.

Such *in vitro* systems have been successfully used to implement and to study gene expression, gene circuits, expression and DNA self-replication.[34, 35, 232, 233, 37] For instance, cell-free systems were encapsulated into phospholipid vesicles to facilitate compartmentalized protein expression for several hours.[35]

In the present work, we encapsulate the TX-TL system in peptidosomes made of amphiphilic elastin-like peptides (ELP). These peptides can be easily expressed in cell-free systems [228, 229] and thus they simplify the synthesis of the membrane material in comparison to lipid synthesis. We show that biomolecules can be easily enclosed in ELP-based vesicles, and we further demonstrate *in vesiculo* transcription of an RNA aptamer and translation of a fluorescent protein. We finally, succeed in the expression of the membrane-constituting peptides inside the vesicles themselves and demonstrate their incorporation into the membrane and thus inherent vesicle growth.

## 5.2   Results

### The synthetic cell model

The molecular building block of our membrane was a synthetic peptide derived from the protein tropoelastin, which commonly comprises the sequence motif $(G\alpha GVP)_n$, where $\alpha$ can be any natural amino acid except for proline and n is the number of pentapeptide repeats. ELPs are stimulus-sensitive peptides and undergo a fully reversible phase transition from a hydrophilic to a hydrophobic state when the sample temperature exceeds the specific transition temperature $T_t$.[70] The latter depends on several parameters such as the amino acid used for $\alpha$, concentration, salt conditions, pH, etc. Unlike other hydrophobic molecules such as lipids, ELPs in a hydrophobic coacervate state can still exhibit a water content of about 63 % by weight.[234]

In order to create an amphiphilic peptide capable of membrane formation, we pro-
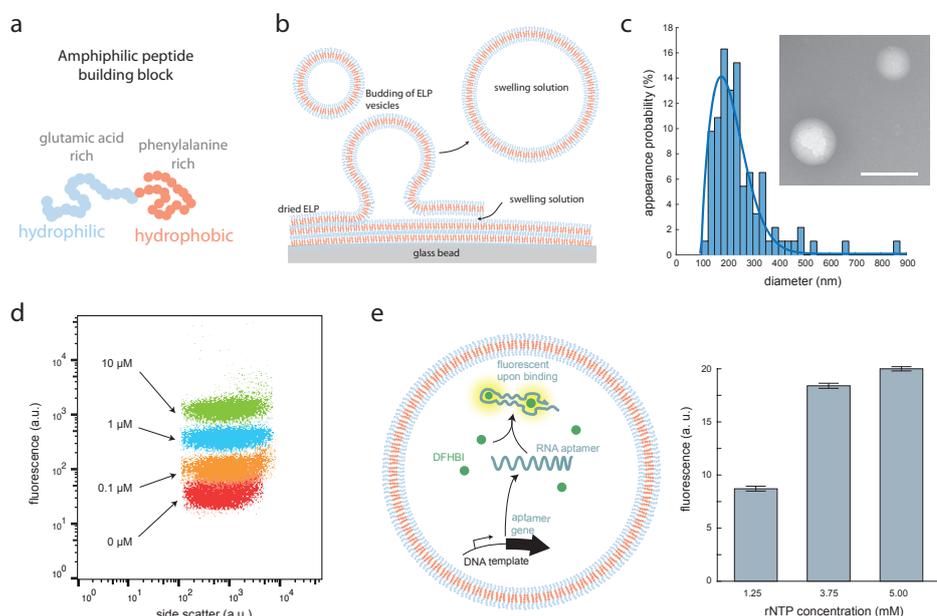
Figure 17: ELP vesicle formation and transcription. (a) The peptide building block is an amphiphilic ELP with a hydrophilic glutamic acid-rich domain and a hydrophobic phenylalanine-rich domain. (b) For vesicle formation dried ELPs are rehydrated from glass beads. (c) Size distribution of the produced vesicles obtained from TEM measurements with diameter of 176 nm. The data are described using a Weibull probability distribution (solid line). The inset shows a typical TEM image. Scale bar: 200 nm. (d) Co-localization of fluorescently labeled DNA (with indicated concentrations) into the vesicles measured by flow cytometry. (e) Left: Illustration of the transcription of an RNA aptamer inside of an ELP vesicle. The aptamer binds to DFHBI, which then fluoresces. Right: Plateau value means of the DFHBI fluorescence after 50 min of transcription for various rNTP concentrations. The given error bars indicate the sample standard deviation of the measured plateau values.

duced a diblock copolymer with the sequence MGH-GVGVP((GEGVP)$_4$(GVGVP))$_4$ ((GFGVP)4$_4$(GVGVP))$_3$(GFGVP)$_4$-GWP abbreviated as EF. At physiological pH, this peptide has a charged hydrophilic E-rich block (mainly GEGVP pentapeptides) with a specific T$_t$, E below sample temperature T and a hydrophobic F-rich block (mainly GFGVP pentapeptides) with a specific T$_t$, F above T (Fig. 17a). The peptide was expressed using *E.coli* cells carrying a plasmid coding for EF and pu-

rified using inverse transition cycling [235] at pH 2 and 7 (see 9.1). The purity of the protein was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with Roti-blue staining (Supplementary Fig. 16) and the concentration was determined using spectroscopic methods (see 9.1). For controlled formation of vesicles, the peptides were dissolved in a chloroform-methanol mixture together with spherical glass beads (see 9.1). We assume that the vesicle formation process is similar to the liposome formation process (Fig. 17b).[144] After fast evaporation of the organic solvent the glass beads are coated with EF. Due to the addition of the swelling solution (initially 1x phosphate-buffered saline (PBS)) the EF film rehydrates through budding and the swelling solution is encapsulated. Using dynamic light scattering (DLS) we determined the diameter distributions for 110, 180, 220, and 440 pM EF. The corresponding peak values were 87, 178, 220, and 250 nm with dispersion (sample standard deviation) values of 47, 67, 111, and 415 nm (Supplementary Table 4). For further experiments we used 180 pM EF since it resulted in the lowest relative dispersion (dispersion divided by the peak value). Using transmission electron microscopy (TEM) we determined the peak value of the diameter distribution for 180 pM EF (Fig. 17c) to 176 nm with a dispersion of 68 nm, which is in good agreement with the DLS data (Supplementary Fig. 17). The size distribution of the TEM and DLS data can be described by a Weibull extremal probability distribution (Supplementary Note 9.2.2),[236] and all given dispersion values are determined from a Weibull fit to the data. Membrane formation and its stability depend on the osmotic pressure, the critical aggregation concentration and the chemical potentials in general. Therefore, the samples were not diluted if applicable or purified.

We also verified the size stability of the vesicles over 18 h using DLS with a mean of 179 nm and a standard deviation of 10 nm (Supplementary Fig. 18). We found consistent mean diameters in DLS measurements from five repetitive rehydration experiments indicating no stochastic influence on the vesicle formation process with a student-t corrected standard deviation of 12.4 nm for a P-value of 0.95. For a swelling solution containing 1x PBS, the mean membrane thickness was roughly determined to be 4.9 nm with a standard deviation of 0.5 nm using TEM (Supplementary Table 5 and Supplementary Fig. 23), which is in good agreement with Huber et al. who used an amphiphilic ELP of similar size in a GFP-ELP fusion protein. Furthermore, we demonstrated vesicle stability at various NaCl membrane gradients (Supplementary Fig. 20 and 21). In order to show encapsulation using

the glass beads method, we utilized fluorescently-labeled DNA in the swelling solution. The resulting vesicle sample was diluted and subsequently measured with a flow cytometer. The fluorescence intensities of the vesicles scaled with the DNA concentrations used (Fig. 17d and Supplementary Fig. 22). In a control experiment without vesicles the fluorescence determined for a sample of FAM-labeled DNA was reduced due to dilution of the DNA, but it also scaled with the concentration (Supplementary Fig. 35).

**Transcription in vesicles**

We next studied the efficacy of *in vitro* transcription reactions encapsulated inside of the peptide vesicles. To this end, we transcribed the fluorogenic dBroccoli RNA aptamer [237] in the presence of its cognate fluorophore DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone) and monitored the increasing fluorescence of the vesicles over time. dBroccoli is the dimeric version of the Broccoli aptamer, which exhibits robust folding, even in low magnesium concentrations, and it's optimized for usage in living cells. DFHBI is a small non-fluorescent molecule that gets into a highly fluorescent state upon binding to its aptamers such as Broccoli. The encapsulated transcription mix (TX) consisted of ribonucleoside tri-phosphates (rNTPs), electrolytes, DFHBI (dimethyl sulfoxide (DMSO) stock), T7 RNA polymerase, and the DNA template for the dBroccoli aptamer. Encapsulation was performed with the glass beads method described above, whereby 10 % DMSO in solution did not affect vesicle formation. To suppress transcription outside of the vesicles DNase I was added to the outer solution after formation of the vesicles, which digested any non-encapsulated DNA template. As a negative control, the DNase I was added before the vesicle formation was done.

As expected, transcription inside the vesicles led to an increase in the fluorescence signal, which reached its maximum after $\approx 50$ min. The plateau phase was most probably caused by the depletion of resources such as the rNTPs, the formation of pyrophosphates or the exhaustion of the polymerase. To test the hypothesis of resource depletion the rNTP concentration was altered. The measured intensity values always reached their plateau phase after the same time but depending on the rNTP concentration the maximum fluorescence level increased (Fig. 17e and Supplementary Fig. 24). This increase was not linear with the rNTP concentration,

and hence the rNTP depletion was not solely responsible for the limitation of the transcription reaction.

**Protein expression in peptide vesicles**

In order to demonstrate a compartmentalized transcription-translation process, we expressed the fluorescent protein mVenus (Fig. 18a) and YPet (Supplemental Figs. 25 and 26) inside of the vesicles. Encapsulation again was carried out using the glass beads method. Upon gene expression, mVenus fluorescence increased and reached a plateau phase after 180 min (Fig. 18 and Supplementary Fig. 25). Expression of proteins occurring outside of the vesicles was suppressed using the antibiotic kanamycin, which blocks the 30S-subunit of the ribosome and thus prevents translation. As an alternative, we used EDTA to chelate magnesium ions in the buffer and thereby also suppress gene expression (Supplementary Fig. 26). Neither the presence of kanamycin nor EDTA outside of the vesicles compromised expression of mVenus inside of the peptidosomes, which indicates that the peptide membrane is not permeable for these small molecules on the time-scale of our experiments. When kanamycin or EDTA was encapsulated together with the expression mix as a control, protein expression was successfully suppressed (Fig. 18b red curve and Supplementary Figs. 25 and 26). The TX-TL system also contains cofactors, such as $NAD^+$ or FAD, which change their autofluorescence upon reduction to NADH or $FADH_2$ and thus create an additional change in the fluorescence signal. Therefore, we used a background correction with a sample containing the cell extract, but without plasmid. By assuming a Poisson distribution we can calculate the probability to find the TX-TL components inside the vesicles (Supplementary Note 9.2.2). For concentrations of about 1 $\mu$M and above (e.g., the proteins) the probability is close to 100 %, whereas for the plasmid (50 nM) the probability is 35 %.

**Vesicle growth caused by compartmentalized peptide synthesis**

As a proof of principle, we measured the growth of the peptide vesicles, when monomers are added to the outer solution. In DLS measurements we could measure a size change from 192 nm (dispersion 74 nm) to 234 nm (dispersion 101 nm) after the addition of 50 $\mu$M ELP (Supplementary Fig. 19).
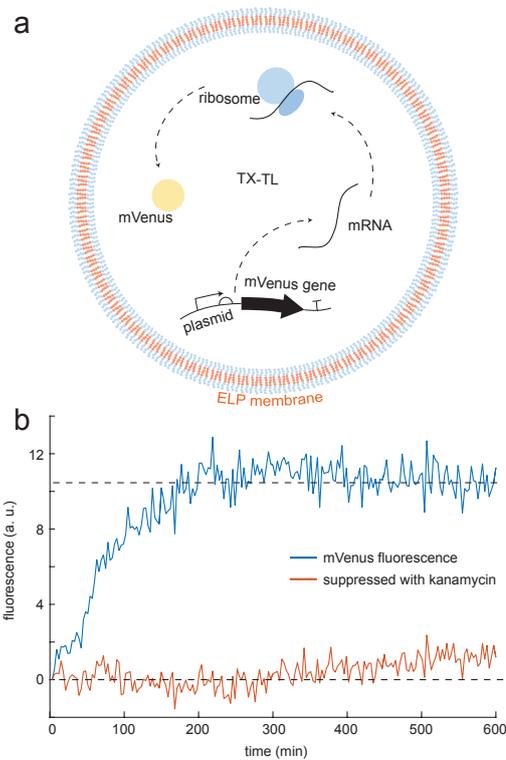
Figure 18: Protein expression inside ELP vesicles. (a) Illustration of the expression of mVenus inside an ELP vesicle containing bacterial cell extract TX-TL. (b) Time-dependent fluorescence of expressed mVenus (blue) and fluorescence of kanamycin-suppressed expression (red). The dashed lines are a guide to the eye.

Given the capacity for protein expression within the peptidosomes, we finally proceeded to synthesize the membrane-constituting peptide itself inside of the vesicles (Fig. 19a). To verify peptide expression inside the vesicles containing the TX-TL system, we equipped EF with a His-tag. After vesicle formation and EF expression we analyzed this sample for His-tagged peptides. External expression was suppressed using kanamycin. Using western blotting the peptide was clearly identified after 240 min of expression, whereas in the initial sample (before expression) no His-tagged peptide was found (Fig. 19b bottom and Supplementary Fig. 33). The peptide band was shifted upward with respect to the expected position at about 18 kDa, which is a well-known effect for elastin-like polypeptides.[238, 239] For clarification we confirmed the expected weight of our non-tagged EF construct by mass spectrometry. We found a peptide mass of 18180 Da (Fig. 19b top and Supplemen-

tary Fig. 36), which is the exact ELP mass reduced by the mass of the methionine at the beginning of the peptide sequence, which most probably has been removed through posttranslational modification.[240] The two peaks next to the parent mass most likely represent the peptide mass with $Na^+$ and the peptide mass with acetonitrile.

In the next step, we investigated the incorporation of the internally generated EF peptides into the vesicle bilayer and thus the growth of the membrane from within. The vesicles were produced as before and kanamycin was utilized to suppress outside expression. After 8 h of incubation, flow cytometry measurements indeed showed an increase of the forward scattering signal. This indicates a growth in vesicle size with respect to a reference sample where kanamycin was present inside the vesicles (Supplementary Fig. 34).

Using TEM we statistically verified the relative size change of the vesicles (Supplementary Figs. 27 and 28). The freshly prepared vesicles, which were able to grow, were divided into two batches. One was immediately flash-frozen to suppress peptide expression, whereas the second was incubated at 29 °C for 240 min and then flash-frozen to stop expression. At the beginning of the expression (before incubation of the sample) the peak of the size distribution was found at a diameter of 157 nm with a dispersion of 104 nm, while peptide synthesis for 240 min resulted in a diameter of 330 nm and a dispersion of 83 nm (Fig. 19c, Supplementary Fig. 31, and Supplementary Table 6). As mentioned before only 35 % of the vesicles should contain a plasmid, and are therefore able to express EF and to grow. Figure 19c shows no indication of a not growing subpopulation at t = 240 min. We assume that the vesicles exchange membrane peptides, which makes the whole population grow; perhaps this also indicates the existence of flip-flop between the leaflets. As a negative control, the hydrophilic ELP $(GVGVP)_{40}$ (further denoted as V40) was expressed to keep the load on the expression system similar. Neither flow cytometry nor TEM measurements (Supplementary Figs. 29 and 30) showed a measurable change in vesicle size in this case, which is in agreement with the fact that V40 is not able to incorporate into the membrane. The peak value determined from these TEM measurements was 149 nm with a dispersion of 102 nm at the beginning of V40 expression and 145 nm with a dispersion of 128 nm after 240 min.

To further examine vesicle growth, we utilized a Förster resonance energy transfer (FRET) assay to monitor the incorporation of internally expressed EF into the membrane. To this end, we prepared two batches of EF, which were either modified
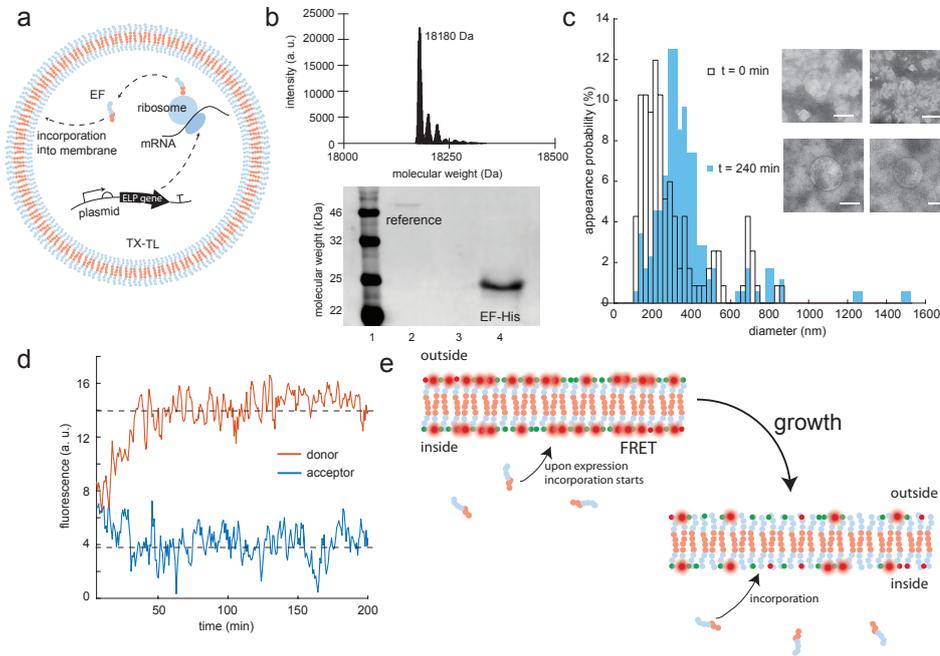
Figure 19: Expression of ELP and vesicle growth. (a) Illustration of the expression of EF inside a vesicle using TX-TL. (b) Top: Mass spectroscopy of full length EF. Bottom: Western blot of EF-His expressed in TX-TL within vesicles after 0 min (lane 3), where no EF-His is detected and after 240 min (lane 4). The reference is a histidine-tagged helicase (lane 2). (c) Size distribution of vesicles at the beginning of EF expression (t = 0 min) with a peak value of 149 nm and after 240 min with a peak value of 330 nm using TEM. Inset: Typical TEM images at t = 0 min (top) and t = 240 min (bottom). Scale bar: 200 nm. (d) Donor fluorescence (red) of dye-labeled EF located in the membrane and acceptor fluorescence (blue) of a labeled EF. The dashed lines are a guide to the eye. (e) Illustration of the FRET assay used. The vesicles are formed using Cy5-EF and Cy3-EF. Upon expression of unlabeled EF and its incorporation into the membrane the mean distance between the FRET pairs rises and the donor signal increases.

with the fluorophore Cy3 or the fluorophore Cy5 via copper catalyzed azide-alkyne Huisgen cycloaddition (see 9.1). Vesicles were then formed using a 1:1 mixture of Cy5-EF and Cy3-EF. As a result, the dyes were randomly and homogeneously distributed within the vesicle membrane after the formation with the glass beads method. As Cy3 and Cy5 constitute a FRET pair, excitation of the Cy3 fluo-

66

rophore, therefore, led to fluorescence emission of the Cy5 acceptor via FRET.

In a bulk experiment, we found that expression of non-labeled EF within the vesicles led to a decrease in the acceptor signal accompanied by an increase in donor fluorescence. This indicated an increasing average distance between the dyes inside the membrane (Fig. 19d, e) and thus an incorporation of new EF. The vesicles were produced as before and kanamycin was utilized to suppress outside expression. In control measurements with kanamycin inside the vesicles, the FRET signal stayed constant (Supplementary Fig. 32). This clearly demonstrates that EF peptides expressed in the interior of peptide vesicles incorporate into the membrane and cause vesicle growth.

From the measured vesicle diameters, we could estimate the relative number of peptides expressed inside the vesicles. The membrane volume at times 0 min and 240 min after start of the expression reaction can be calculated using the volume of a single ELP $V_{\mathrm{ELP}}$ and the number of peptides $N_0$ and $N_{240}$ at these times. The relative volume increase of the membrane $\xi = N_{240} \cdot V_{\mathrm{ELP}}/N_0 \cdot V_{\mathrm{ELP}} = N_{240}/N_0$ is related to the vesicle radii at times 0 min and 240 min via the expression $\xi = R_{240}/R_0$ (Supplementary Note 9.2.2). From our experiments, we found $\xi = 4.4$, which means a 4.4-fold increase of the initially present number of ELP due to peptide production inside the vesicles.

## 5.3 Discussion

In conclusion, our results demonstrate that peptide vesicles are promising candidates for the generation of artificial cell-like compartments. Their fabrication is relatively straightforward and the encapsulation of biochemical reaction mixtures is - at the moment - only limited for low-concentrated molecules by the vesicle size. We successfully showed transcription of an RNA aptamer and the expression of fluorescent proteins inside our peptide vesicles. Most importantly, we demonstrated vesicle growth through expression of the membrane peptide in vesiculo and its incorporation into the membrane.

It is conceivable that in future work peptide vesicle growth and also replication of the encapsulated genetic templates could be coupled, which would be a major step towards the generation of self-replicating protocellular compartments.

# 6 Target Specific Silencing of Genes in Cell-Free Systems for Controlling and Modification of Bacteriophages

The text and figures in this chapter are in preparation for a publication:

"Target Specific Silencing of Genes in Cell-Free Systems for Controlling and
Modification of Bacteriophages"

**Kilian Vogele**, Elisabeth Falgenhauer, Sophie von Schönberg, Friedrich C.
Simmel and Tobias Pirzer, in preparation

In order to provide a structure better suited to this thesis, certain sections have
been merged and rearranged or renamed.

In the previous chapter, the cell-free transcription-translation system was used to
express a peptide which is capable to self-assemble into vesicular structures. In the
following a different building block for self-assembly in cell-free systems is investigated, bacteriophage capsid proteins.

The open nature of the cell-free system allowed a diverse methodology for rapid
prototyping. Controlling the ribosome binding site (RBS) of genes is a widely
used approach to alter the expression of proteins by nucleic acid-based logic gates
to develop sensor-based systems. Here a highly sequence-specific and orthogonal
DNA-based suppression system is developed to alter the expression pattern in a
similar way. The single-stranded suppressor DNA binds to its complementary messenger RNA (mRNA) and blocks the ribosome binding site of the gene of interest.
The achieved reduction of gene expression is approximately 50-fold. By applying
this system to alter the replication of the native T7 phage genome by suppressing
the major capsid protein the genome replication in the cell-free system is improved.
Simultaneously the titer of the cell-free produced phages is reduced significantly.
This system is able to extend the toolbox of controlling *in vitro* translation and
gives the options to alter the expression pattern of native genomes for prototyping.

## 6.1 Introduction

Cell-free transcription-translation has become a wide spread tool in synthetic biology for rapid prototyping and as model platform. Employing its open nature, a variety of applications have been developed, such as metabolic networks, genetic circuits and oscillators. These examples come with different feedback systems, and therefore require a variety of regulatory elements and promoters with diverse strengths. Moreover, in the last years *in vitro* expression has been used to mimic living systems as well. A variety of created *bottom-up* systems already display essential properties of life, such as encapsulation of an expression system in various compartments,[35, 232] the enclosed synthesis of the membrane constitutes [36] and the growth of such compartments.[6]

Another essential aspect of a living system is the replication of its genome, e.g. DNA- or RNA-based, and hence the ability for Darwinian evolution.[12] To implement this, very often phage-based systems have been employed.[241, 242] One of the first prominent *in vitro* attempts is the Spiegelman experiment,[243] where the Q$\beta$ bacteriophage's RNA polymerase replicates its own genome. The main goal is to achieve a system, which is not only capable of genome replication, but which also encodes its complete replication machinery. Two of the main obstacles are the genome length and the appearance of so called parasitic sequences, which are faster replicated than the replicase encoding sequences. For the latter a promising solution is a combination of encapsulation and sorting.[244] For instance, by using emulsion droplets to enclose the genome and the replicase of the Q$\beta$ phage genome, to couple replication with an intermediate sorting step for a few generations of serial dilutions. This system, however, is still not capable of replicating longer genomes. As an alternative, the replication machinery of the phi29 phage can be utilized, which was successfully used to replicate templates up to more than 100 kilobases (kb).[37] For comparison, a minimal protein-based cell, which is capable of reproduction a set of about 150 genes would be required, which translates into a minimal genome with a length of about 113 kb.[245]

Most importantly, phages have been utilized to develop a variety of different tools for synthetic biology, such as the T7 *in vitro* transcription systems,[246] the T5 exonuclease for Gibson assembly,[247] and the phage assisted continuous evolution (PACE) approach.[248] And since phages are the most abundant biological particles on earth, they still hold a huge potential for functional tools; despite the fact that

the T4 phages, which are one of the best characterized model phages, still encode hypothetical proteins.[249]

Beside fundamental questions about their proteins, the modification of bacteriophages can be very useful to improve their therapeutic value.[250] Here, the potential applications are tail fiber engineering to alter the phage host range,[251] the suppression of the lysogenic life cycle,[252] or the improvement of biofilm degradation.[253] But phage genome engineering lacks efficacy due to the limited time of phages inside their host bacteria. Therefore, conventional techniques such as homologous recombination requires screening of up to $10^4$ to $10^{10}$ phages.[254] More advanced techniques employing control systems such as CRISPR-based systems can reduce these numbers.[255] But essential genes still needed to be expressed which is also a problems with toxic genes. Another constrain is that all these techniques are involving electroporation, which can't be utilized with all bacteria strains. This also limits the attempts to modify bacteriophages to get insights in the function of the corresponding proteins.

To solve this limitation we developed an *in vitro* system with which we were able to show that we can suppress sequence specific genes in the T7 phage. We targeted the gene encoding the major capsid protein, which is an essential component of the phage capsid. As a result, the titer of the phages produced cell-free decreased, whereas the phage genome replication increased 8-fold. For the suppression a single stranded DNA complementary to the ribosomal binding site (RBS) and the following downstream sequence was added to the cell-free system. By blocking the RBS of the mRNA we show that the expression end levels of YPet can be also successfully suppressed. By this technique we were able to suppress gene expression in general as well as in full native genomes to alter the outcome of the expression program.

## 6.2   Results and Discussion

**Suppression of cell-free expression of a gene encoding a fluorescent protein.**

RNAi is a prominent example of suppressing gene expression in nature.[256] But the stability of RNA in cell-free systems is quite limited. Therefore, we used complementary suppressor DNA (supDNA) to alter the RBS access of the ribosome to block the translation of the mRNA (Fig. 20a). In a first step we characterized

the capabilities of supDNA through the suppression of YPet translation, which was encoded on a model plasmid; the cell-free system used was derived from *E.coli*. The expression of YPet at 29 °C was monitored by its fluorescence in a plate reader. The total length of the supDNA was 60 bp (base pairs), which were complementary to the RBS and the downstream sequence. We used supDNA in concentrations of 0 µM, 0.05 µM, 0.5 µM, 5 µM and 10 µM and compared the mean end-level fluorescence between 13 h and 14 h. The gene is either controlled by a T7 promoter (Fig. 20b) or a constitutive promoter (J23106) (Fig. 20c). As expected the suppression strength increased with increasing supDNA concentration, whereas 0.05 µM did not show any effect when compared to the control. For both promoters suppression was highest at 10 µM, 4-fold for the constitutive promoter and 50-fold for the T7 promoter. Since supDNA and the ribosome are competing for the RBS, it is possible to gradually tune the suppression strength of the targeted gene by changing the concentration of the supDNA. A rationale for the different results for the two promoters is the transcription velocities of the RNA polymerases used. The native RNA polymerase of *E.coli* transcribes 20-90 nucleotides per second,[257] whereas the T7 polymerase can transcribe up to 240 nucleotides per second.[258] The ribosome initiation rate and the RNA production rate in *E.coli* are adjusted by evolution to minimize the number of free ribosomes. In contrast, the T7 RNA polymerase is optimized to produce higher levels of phage RNA. This creates a situation in which the ratio between available ribosomes and mRNA level is imbalanced in favor of the mRNA. Hence, more supDNA can bind to the RBS than it is blocked by the ribosome.

**Gene suppression in a native phage genome**

For the prediction of the function of a specific gene, commonly a knockout variant of the genome is generated; but for this it is necessary to engineer the phage genome. Although the techniques used for mammalian cells developed quite fast, for bacteriophages laborious screening is still necessary.[254] However, using a cell-free system only the native phage genome is required, and it has been already shown that a variety of *E.coli* phages can be produced *in vitro*.[230, 233] Here, we used the T7 phage and suppressed the translation of its major capsid protein (Fig. 22a) using supDNA. The gene of interest is controlled by a T7 promoter and terminated
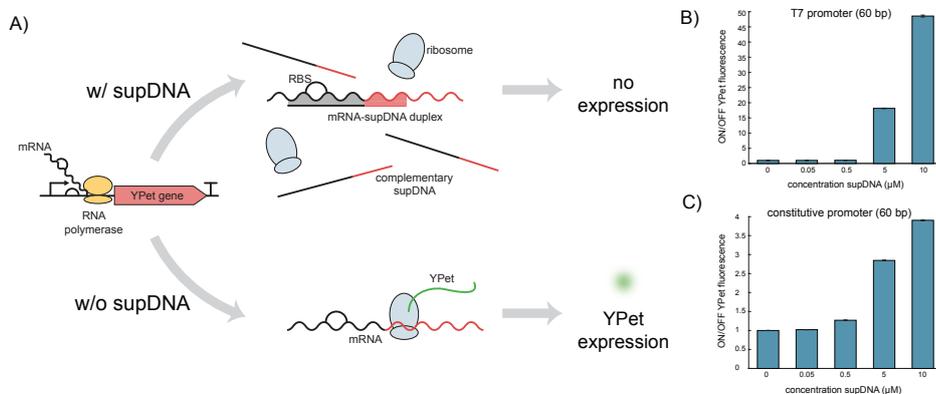
Figure 20: Suppression by the addition of complementary DNA. (a) Schematic of the principle of the suppression by the complementary DNA by inhibiting the binding of the ribosome to the RBS of the gene of interest. As a first step the RNA polymerase (yellow) transcribes the mRNA (wavelike line) containing the RBS. In the presents of the DNA (straight line) which is complementary to the RBS region the ribosome (blue) can't bind and translation is blocked. (b) The ratio of the suppression by complementary DNA at varying concentrations of the added DNA of a YPet gene controlled under a T7 promoter in a cell-free *E.coli* system. The bars representing the mean of the on/off ratio of the end levels between 13-14 h of the fluorescence measured with a plate reader. Biological triplicates were measured. c) The ratio of the suppression by the complementary DNA at varying concentrations of the added DNA of a YPet gene under a constitute promoter (J23106) in a cell-free *E.coli* system. The bars representing the mean of the on/off ratio of the end levels between 13-14 h of the fluorescence measured with a plate reader. Biological triplicates were measured.

by a T7 terminator (T7-T phi), and therefore it is a suitable candidate; additionally, the major capsid protein is not encoded in a polycistronic region of the genome. Phage expression was measured after 4 h at 29 °C using a plaque assay, which resulted in phage concentrations of about $3 \times 10^9$ PFU per mL to $1.3 \times 10^{10}$ PFU per mL; additionally transmission electron microscopy (TEM) was used to check the formation of the phages (Fig. 22b). In the negative control, either without DNA or without the cell extract no phages were detected.

Upon the addition of nucleoside triphosphate (dNTP) to the sample containing the cell-free system and the phage genome, replication of the T7 DNA could be detected
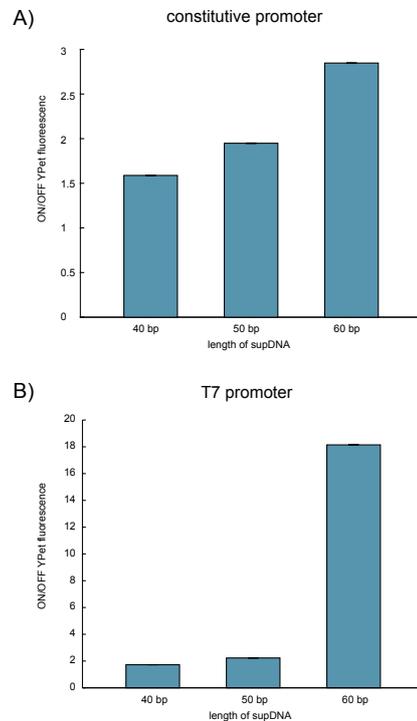
Figure 21: Suppression by complementary DNA with varying length of the complementary DNA. (a) The ratio of the suppression by complementary DNA at varying length of the complementary DNA at a concentration of 5 $\mu$M on a YPet gene under a constitute promoter (J23106) in a cell-free *E.coli* system. The bars representing the mean of the on/off ratio of the end levels between 13-14 h of the fluorescence measured with a plate reader. Biological triplicates were measured. (b) The ratio of the suppression by complementary DNA at varying length of the complementary DNA at a concentration of 5 $\mu$M on a YPet gene under a T7 promoter in a cell-free *E.coli* system. The bars representing the mean of the on/off ratio of the end levels between 13-14 h of the fluorescence measured with a plate reader. Biological triplicates were measured.

after 4h of incubation by its encoded proteins; the detection was carried out using agarose gel electrophoresis (Fig. 22c). The visible bands at about 1 kb to 1.5 kb resulted from the cell-free system and the bands below 0.5 kb are residual DNA fragments from the phage DNA preparation. To quantify the replication yield of the T7 genome with and without supDNA qPCR was utilized. Two identical sam-

ples were prepared, one with supDNA of 60 bp length complementary to the RBS and a reference sample without. In the qPCR, the region between the bases 8426 and 9007 was targeted by the qPCR primers; the calibration curve for the DNA concentration calculation was measured beforehand using purified phage DNA. In the control sample with replication but without supDNA the DNA concentration showed a 4-fold increase, whereas the sample with supDNA the DNA replication increased more than 30-fold (Fig. 22d). We also tested how the suppression of the major capsid protein affected production of fully assembled phages and found, that the phage titer roughly decreased by a factor of four (Fig. 22e). These results show that supDNA can be used to successfully alter gene expression levels of a native phage. The major capsid protein is essential for the assembly of the phage capsid and therefore its suppression reduced the number of produced phages. The simultaneous increase of replicated DNA can be explained through the higher amount of available DNA for replication, since it couldn't be packaged into phage capsids and simultaneously resources are drawn from the capsid expression.

**Serial dilution of the phage genome in fresh cell extract**

The replication capability of the T7 genome could be utilized in approaches for the creation of artificial cells. For that, the phage genome has to be able to replicate itself at least several times. We next tested the application of supDNA in serial dilution experiments similar to the Spiegelman experiment.[243] As shown before supDNA improves DNA replication and thus the major capsid protein was suppressed in the course of the serial dilution steps. The phage DNA and the cell-free system were incubated at 29 °C for 4 h and diluted afterwards by transferring 3 % (vol %) of the sample into freshly prepared cell extract. After that it was again incubated at 29 °C for 4 h (Fig. 23a). Before every dilution step the DNA concentration was quantified using qPCR as mentioned before; the replication of the full genome was again measured using an agarose gel. As expected the DNA concentration increased due to replication in each generation, whereat over several generation the increase per generation is decreasing.

To investigate the influence of encapsulation on replication, the phage genome and the cell-free system were enclosed in emulsion droplets formed by manual shaking. After a 4 h incubation at 29 °C the droplets were broken and the contents pooled. For DNA quantification qPCR and agarose gel were used, and 3 % of this mixture
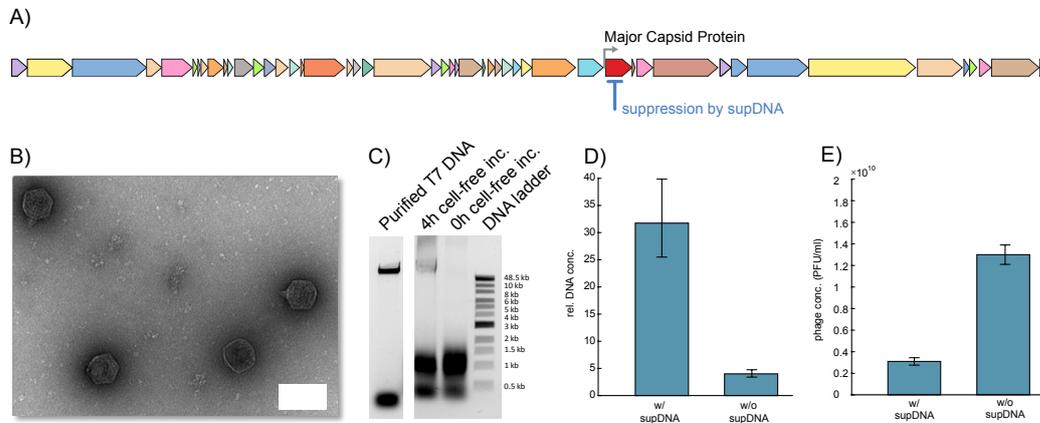
Figure 22: Suppression in the gene encoding the major capsid protein of the T7 phage. (a) Schematic drawing of the full genome of the T7 phage with its encoded genes, as well as the T7 promoter controlling the major capsid protein and the binding position of the complementary DNA. (b) Transmission electron microscope images of the T7 phage after the expression in the cell-free system after 4 h at 29 °C. (c) Gel electrophoresis image of the T7 phage DNA incubated in the cell-free system. From left to right the samples are the column purified DNA of the T7 phage extracted from a beforehand prepared phage stock, the second sample is the cell extract with the phage DNA incubated for 4 h at 29 °C, the third sample is the same system but without the incubation and the fourth sample is the Quick-Load 1 kb Extend DNA Ladder from NEB. (d) The relative DNA concentration of the phage DNA incubated in the cell extract with the addition of complementary DNA and without. The concentration was measured with qPCR. (e) The concentration of the phages in PFU/ml measured with a plaque assay from a sample with and without the complementary DNA.

was diluted as mentioned before and again encapsulated. The global trend is similar to the measurements in bulk, but instead of a 10-fold increase due to replication in bulk, the encapsulated samples showed a 100-fold increase of DNA in the first two generations. After four generations the DNA concentrations in droplets were still 100-fold larger than the bulk concentrations. The plaque assays carried out after each generation showed a reduction of active phages from generation to generation (Fig. 23 insets) in both experiments, bulk and droplets. The measured titer for generations 2 to 4 in these two experiments were almost the same, except for the

first generation were in droplets nearly one order of magnitude less active phages were produced.
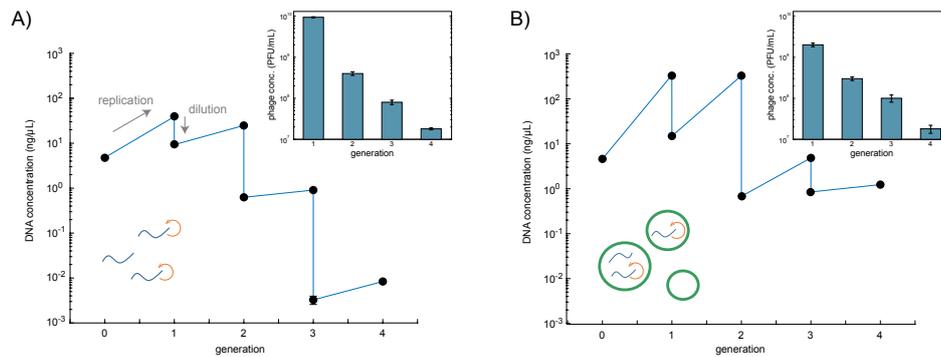


Figure 23: Propagation of serial transfer of phage DNA in fresh cell extra. For that the phage DNA was incubated for 4 h at 29 °C in the cell extract and afterwards 3 % of the volume was transferred into a fresh solution of cell extract. This was repeated for four generation. The DNA concentration of each sample was measured by qPCR at the beginning of the incubation and after 4 h of incubation. Additional the concentration of fully assembled phages was determined by a plaque assay (insert). (a) incubation of the samples in bulk (b) incubation of the cell extract during the incubation in droplets. The droplets were produced by mixing the cell extract with FC-40 oil with 2 % (w/w) PFPE/PEG-surfactant and subsequent shaking. For the propagation to the next generation the droplets were broken with perfluorooctanol and pooled again for the next generation. Biological triplicates were measured.

The global reduction of replicated DNA from generation to generation could be caused by two main reasons. First, because of the accumulation of mutations in genes required for DNA replication, to the point where the translated proteins are less or not functional anymore. Second, the appearance of short parasitic DNA sequences, which are replicated faster than the full genome. From generation to generation, these parasitic sequences are losing the genes encoding essential proteins for DNA replication. When a phage is produced in a bacterium parasitic sequences get lost after lysis or result in less active or non-active phages, which will diminish due to evolutionary pressure. However, in serial dilution experiments there is less evolutionary pressure, and thus parasitic sequences can accumulate until they poison the system. As a consequence, the replication success of the full phage genome is

reduced. What speaks for this hypothesis is that the decay in DNA concentration over the four generations is less for the encapsulated replication compared to bulk, which is more prone to spreading of the parasitic DNA sequences, as similar effects were shown with the $Q\beta$ replication system.[244, 242]

## 6.3   Conclusion

With the recent findings we were able to show that the suppression by the addition of a DNA primer, which is complementary to the RBS is possible. This can be used as a tool for rapid prototyping in cell-free systems. The complementary DNA system is highly sequence specific as it relies on the complementary base pairing and is therefore also orthogonal, as only a length of 60 bp is providing sufficient suppression. As the expression end-levels can be continuously be tuned by the concentration of the complementary DNA also different promoter strengths in gene networks can be screened. This can therefore allow faster tuning of such systems *in vitro* as the design and use of the complementary DNA technique is easy and fast. Whereas it is limited to *in vitro* systems as the complementary DNA has to be added externally.

From these capability also another field of application arose, the suppression in native genomes, like phages to alter their behavior. As the genome engineering of phages is rather laborious and also due to the huge variety of phages, this rapid tool to suppress sequence specific genes can improve the function prediction of newly found phages genes drastically. As this technique relays on the suppression of the RBS still polycistronic genes can be a limitation. With our systems we were able to increase the capability to replicate long DNA *in vitro* which can be used in further attempts to create a protocell. To conclude the complementary DNA technique gives a huge variety of possible further applications and expand the possibilities of cell-free expression as a rapid prototyping toolbox.

# 7 A Cell-Free Platform for Characterizing and Producing Therapeutic Bacteriophages Against Biohazardous and Antibiotic-Resistant Bacteria

The text and figures in this chapter are in preparation for a publication:

"A Cell-Free Platform for Characterizing and Producing Therapeutic Bacteriophages Against Biohazardous and Antibiotic-Resistant Bacteria"

Quirin Emslander, **Kilian Vogele**, Peter Braun, Markus Joppich, Miriam Abele, Chen Meng, Christina Ludwig, Friedrich C. Simmel and Gil G. Westmeyer, in submission

In order to provide a structure better suited to this thesis, certain sections have been merged and rearranged or renamed.

In the previous chapter, the T7 bacteriophages were expressed *in vitro* as a model system to study the possibilities of the cell-free system as a platform for phage production and research. As bacteriophages can also be used as a therapeutic against multidrug-resistant bacteria, which are emerging rapidly, the interest from researchers and regulatory authorities has been increased. However, due to the lack of safe production standards and insufficient phage characterization bacteriophage therapy is currently impeded.

Therefore, the cell-free system is further developed to a host-independent platform for the production of gram-positive and gram-negative bacteria. In a one-pot reaction phages against potentially antibiotic-resistant bacteria such as enterohemorrhagic *E.coli* (EAEC), *Yersinia pestis*, or gram-positive *Bacillus subtilis* are produced in medically relevant concentrations. With high-resolution mass spectrometry the expression of 40 hypothetical proteins from two different phages has also been discovered. Additionally, with time-resolved mass spectrometry genes in the T7 genome are identified, which are expressed exceptionally late during the phage life cycle. This shows that the cell-free system is a useful platform for the acceleration of phage research, as well as a basis for the safe and scalable production of clinical-grade therapeutic bacteriophages.

## 7.1  Introduction

The growing number of multidrug-resistant bacteria (MDR) has been classified as one of the main global health threats by the WHO.[259, 260] It is estimated that costs of 2.2 billion dollars in the USA in 2014 alone, [261] as well as 33,000 deaths per year in Europe, are caused by MDR.[262]

MDR is particularly severe in gram-negative bacteria, with most cases reported for *E.coli*, but classical hospital germs such as gram-negative *S. aureus* are also increasingly resistant. Additionally, in some cases, the treatment of these infections with antibiotics is discouraged because of antibiotics induced side effects.[262, 260, 263, 264, 265]

Due to the increasing necessity of antibiotic alternatives, Belgium and France, as the first countries in the EU, have recently allowed the use of naturally occurring, well-characterized bacteriophages for personalized bacteriophage (short "phage") therapy under the legal framework of the magistral preparation,[266] as well as the use of genetically engineered phages in exceptional cases.[252] A rapid administration of characterized phages may also be crucial against pathogens such as *Yersinia pestis*, which are classified as potential bioterrorism agents and still contribute to naturally occurring outbreaks of epidemics in Africa with high morbidity rates.[267] Unfortunately, the current production processes for both natural and engineered phages pose considerable biosafety concerns and are often quite inefficient and unreliable. In both cases, a pure culture of the pathogenic host is required for phage amplification and engineering.[252, 268] This process has a high risk of prophage contamination, [269] which is known to induce virulence factors in bacteria.[270, 271] Moreover, the standard cultivation of phages with bacteria is time-consuming, generating relatively low titers and requiring the corresponding lab security level for the pathogen. This makes it challenging to parallelize the production of effective phage cocktails that typically contain up to 16 phages.[272] Recent double-blind patient trials have also revealed the instability and correspondingly low shelf half-life of prepared phage cocktails, emphasizing the importance of a fast and reliable production process.[268] Additionally, the production of engineered phages is limited by the necessity of modifying phage genomes within living bacteria, e.g., by CRISPR-Cas, or using yeast artificial chromosomes, which limits the efficiency.[251, 249]

Apart from the elaborate production process, the poor understanding of the molecular relationship between phages and their host can lead to erratic treatment out-

comes.[273, 274] Even though the existence of phages has been known for over a century,[275] the characterization of the phage proteome is still patchy. Based on the analysis of several hundred phage genomes, recent studies determined over 60 % of the annotated gene products as "hypothetical proteins" [251, 249] without any experimental validation. Even for intensively studied model phages such as T7, 38.5 % of the annotated proteome is merely hypothetical, although its proteome consists of only 57 predicted proteins. Thus, it is unclear whether and when these proteins are expressed and how they contribute to phage assembly and the modification of bacterial physiology.[276, 277]

Cell-free production systems, derived, for instance, from *E.coli* extracts, could address all of the above-mentioned challenges: they can (i) produce phages efficiently and safely at high titers, (ii) facilitate novel methodologies for phage engineering, and (iii) enable a better characterization of phage proteomes and function, which is important also for improved phage therapies. Cell-free expression systems have recently been employed for the biosynthesis of glycoproteins,[278] the incorporation of non-canonical amino acids,[278] and as components of minimal cells.[279, 6] With respect to phage production, until now cell-free systems were restricted to the generation of model phages with gram-negative hosts.[280] Most notably, in seminal work, Noireaux and coworkers used their cell-free system to produce the *E.coli* phages T7, phiX174, and MS2.[145] However, neither the cell-free generation of clinically relevant bacteriophages nor a comprehensive characterization of phage proteomes has been achieved so far.

We have therefore introduced innovations in several key aspects: rapid production, transient phage bioengineering, and the characterization of a broad spectrum of phage in a host-independent manner, including therapeutic phage against potentially antibiotic-resistant and potentially bioweaponizable bacteria (Figure 24). This approach has also enabled us to systematically monitor self-assembly and phage expression profiles using time-resolved mass spectrometry to identify dozens of previously hypothetical proteins, which has a significant impact on phage biology and future phage engineering.
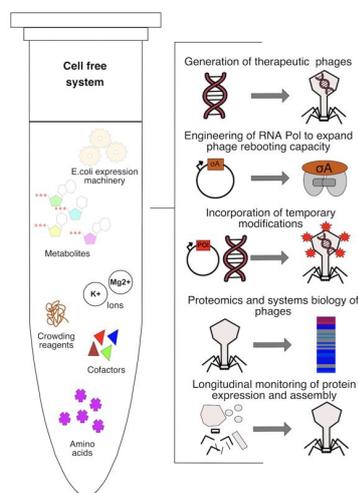
Figure 24: Schematic illustration of a cell-free toolbox for basic and therapeutic phage research. The cell-free system consists of an endogenous *E.coli* molecular expression machinery complemented with metabolites, amino acids, and buffers promoting phage assembly. An optimized transcription-translation system with additional host factors and crowding reagents can manufacture phages targeting both gram-positive and gram-negative bacteria. The open nature of this system enables novel forms of bioengineering (transient engineering) by which the incorporated modification is not encoded on the phage genome. Phages generated in this manner can be analyzed via mass spectrometry without purification to also characterize non-structural phage proteins that may be lost upon purification of phages from bacterial hosts. Cell-free phage production also enables the monitoring of protein expression in a time-resolved manner. Phage engineering can also be performed completely *in vitro*, either by stable genome engineering or by temporary modifications at the protein level.

## 7.2 Results

**Host-independent rebooting of therapeutic phages**

To achieve fast, on-demand production of phages, we optimized the metabolite and energy solutions, e.g., PEG and ATP concentration, of our cell-free expression system to generate high titers ($1.2 \times 10^{10}$ PFU/ml) of DNA (T7) and RNA phages (MS2) from their genomes from a minimum volume of 10 $\mu$l within 2 to 4 hours (Fig. 25a, Fig. 25b). We then tested, under these optimized conditions, the production

of therapeutically relevant phages such as phage CLB-P3 against the pathogenic EAEC strain and phage PhiA1122 against *Y.pestis*. As shown in Figure 25c it was indeed possible to produce clinically relevant titers (on the order of $10^8$ PFU/ml) of both therapeutic phages within 4 hours.

**Rebooting of phages targeting gram-positive bacteria and transient engineering of single-use phages**

To demonstrate the generalizability of our approach to bacterial hosts genetically substantially different from *E.coli*, we next aimed at assembling the phages Phi29 and Goe1 (vB-BsuP-Goe1) targeting *B.subtilis* as part of the phylum Firmicutes. We reasoned that the major limitation of the cell-free production of phages, which had so far been restricted to phages against gram-negative bacteria, is the transcription of foreign host-factors such as promoters. As the early genes of a phage genome are usually transcribed by the host's RNA polymerase, one would at least need to modify the *E.coli* RNA polymerase by incorporating the *B.subtilis* housekeeping sigma factor sigA to reboot these phages (Figure 25a). Indeed, assembly of the *B.subtilis*-targeting phages was only successful when the sigA host factor was co-expressed (Figure 25b,c).

Next, we sought to exploit this modularity of the cell-free system to enable phage bioengineering at the protein level without changing the phage genome. To demonstrate this approach, we co-expressed variants of the T7 capsid protein G10B, including a polyhistidine-tag and the split luciferase HiBit.[281] With co-expression of the his-tagged capsid protein, phages could be purified by affinity chromatography at nearly two log-fold higher titers, indicating that the G10B-His was integrated into the outer phage structure and that the modified phages were still effective against the host bacteria (Figure 25d). We also co-expressed HiBit with a His-tag fused to G10B, purified the resulting phages, and then incubated them with *E.coli*. After washing the bacteria, a substantial luminescence signal was detected, as compared to phages produced by co-expression of his-tagged HiBit without the capsid protein G10B or wildtype T7. These data show that the engineered phages could attach to the host bacteria. Importantly, the bioluminescence signal disappeared after one phage replication cycle, showing that the phage modification at the protein level was only transient.
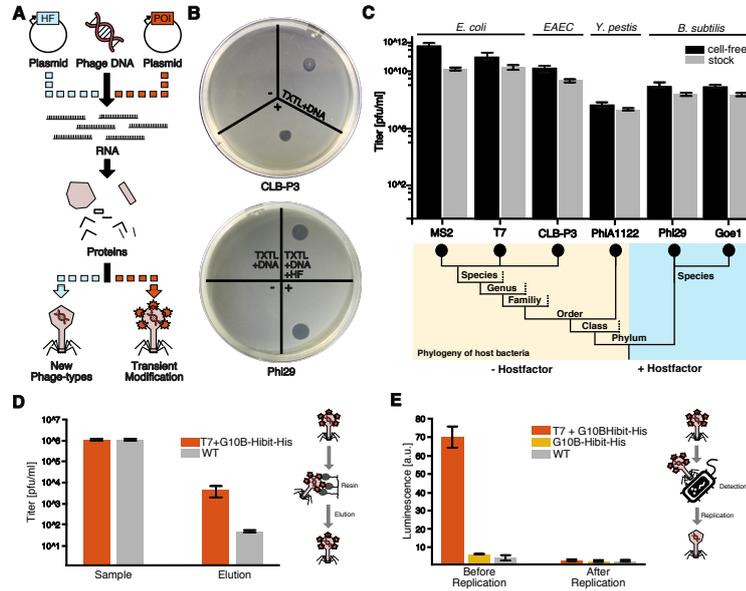
Figure 25: Cell-free production of phages against gram-negative and gram-positive bacteria, including therapeutic phages. (a) Illustration of phage rebooting showing how the addition of phage DNA enables transcription, translation, and self-assembly. If an appropriate host factor (HF) such as sigA is co-expressed (blue), also phages targeting bacteria distant from *E.coli* can be produced. Complementarily, structural proteins of interest (POI, red) can be co-expressed for co-assembly into the phage structure to engineer phages with transiently modified properties. (b) Determination of successful phage synthesis of CLB-P3 (EAEC) and Phi29 (*B.subtilis*). The spot-on assay includes the phage stock, to show lytic ability as a positive control (+) as well as the cell-free system with (+) phage DNA and without (-). Phage Phi29 can only be assembled with the HF. (c) Rebooting of different phages achieves titers sufficient for therapy ($>10^6$ PFU/ml), e.g., CLB-P3 targeting EAEC and PhiA1122 targeting *Y.pestis*. Black and grey bars correspond to phage titers obtained by cell-free phage production and bacterium-dependent cultivation respectively. (d) Titers of T7 phages with co-expression of a modified capsid protein (G10B) fused to split luciferase (HiBit) and a polyhistidine-tag (G10B-HiBit-His) before (sample, adjusted to $10^6$ PFU/ml for both conditions) and after affinity purification (elution) compared to wildtype T7 phages (WT). (e) Bioluminescent signal obtained from bacteria after incubation with phages produced with co-expression of G10B-HiBit-His or controls (HiBit-His without G10B or WT). Error bars in c and d represent the standard deviation of the mean (n=3).

**Total phage proteome identification within the cell-free system**

We subsequently set out to compare the protein composition of phages derived from bacteria-dependent amplification (phage stock) with cell-free synthesized phages assembled by high-resolution mass spectrometry (MS). As can be seen in Figure 26a, more proteins were identified in the cell-free approach than in the phage stock samples. The proteome coverage derived from the cell-free system was 72 % (41 of 57 T7 proteins), with 35-42 % (20-24 proteins) of the proteome represented by hypothetical proteins. Thus, almost all of the previously characterized and at least half of the hypothetical proteins (12 proteins) were validated at the protein level. From the protein expression data, it can be seen that predominantly non-structural proteins (based on the protein uniprot ID and their correlated function), which are present in the bacterial cytosol or membrane, exhibited much lower abundance, in the phage stock samples as compared with the cell-free samples. Several of the hypothetical proteins from phage T7 were found to be conserved in other clinically relevant phages targeting *K. pneumoniae* and *Y.pestis*, e.g., Yersinia phage PhiA1122 (NCBI ID: NP 848312.1) and Klebsiella phage (phage vB-KpnP-KpV767 GenBank ID: AOZ65500), suggesting that these proteins play a universal role in phage biology.

**Bottom-up phage characterization**

Motivated by our success in validating hypothetical proteins based on an existing, well-characterized genome, we next tested whether the cell-free approach can be applied to yet uncharacterized phages. To this end, we de-novo sequenced CLB-P3 phages targeting EAEC using Next-Generation Sequencing (using both Nanopore Sequencing with a MinION device (Oxford Nanopore Technologies, UK) and Illumina Sequencing with a MiSeq Instrument (Illumina Technologies, USA)). We then performed a hybrid genome assembly from MinION and MiSeq reads and analyzed the proteins from cell-free expression by tandem mass spectrometry.
As shown in Figure 26b for the CLB-P3 phage, the majority of proteins were identified in both samples, but with 72 % the proteome coverage was much higher in the cell-free system compared to the phage stock samples (approx. 59 %). In total 68 from 87 predicted proteins were identified by MS. Using our genome annotation tool, only 38 of these 68 proteins could be directly assigned functional names. The remaining 30 are thus labeled as hypothetical. The corresponding heatmap (Fig-
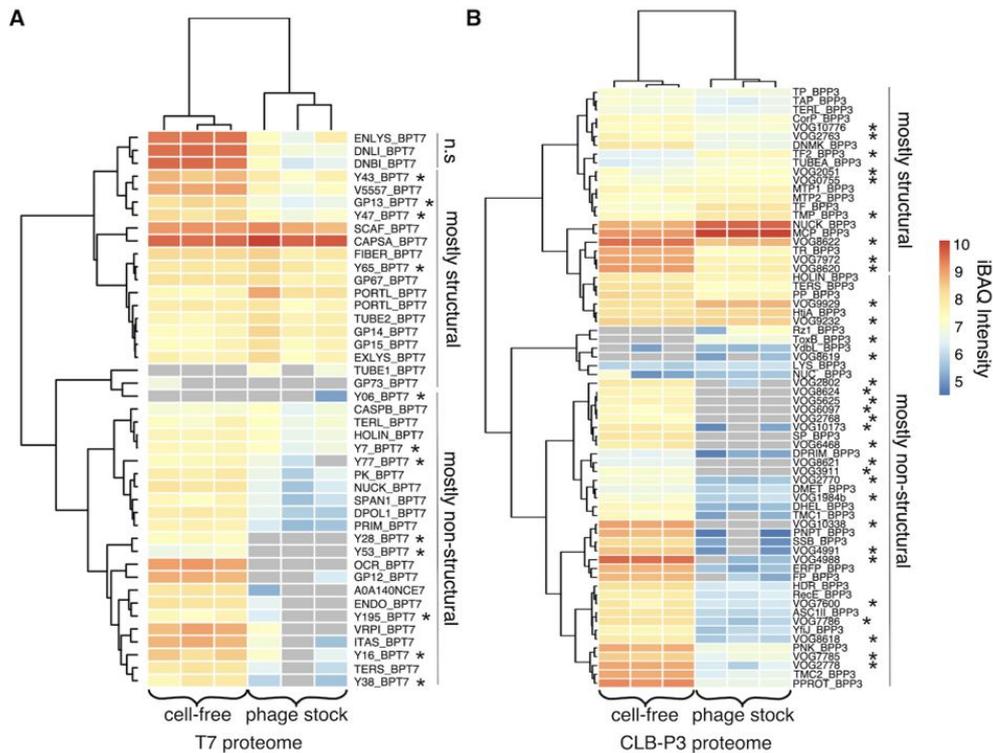
Figure 26: Highly resolved phage proteomes obtained from cell-free systems reveal hypothetical proteins. Phage proteomes expressed in a cell-free system were analyzed by mass spectrometry. The heatmaps from (a) T7 and (b) CLB-P3 proteomes obtained from phage stock and cell-free assembled samples show significant divergence in mass spectrometric intensity between certain protein groups (iBAQ = intensity-based absolute quantification method). Less abundant proteins in the phage stock proteome clustered mostly annotated as non-structural (n.s.) proteins, e.g., not contributing to phage capsid, base-plate, and tail fiber. Stars indicate all hypothetical proteins validated in this study.

86

ure 26b) exhibits a similar distribution of the precursor signal as the T7 heatmap. The intensities of the soluble factors are lower in the phage stock sample than compared with the cell-free system, whereas structural proteins were present with high abundance in both sample types. Based on the protein ID (pVOGs [282]) of validated proteins it was possible to track their existence in several hundred phages targeting gram-negative- and positive bacteria. These include potentially therapeutic phages like the Shigella phage psf-2, the Salmonella phage Jersey, the Klebsiella phage Sushi, and the Cronobacter phage phiES15. But also phages potentially relevant for the food industry were identified, such as Lactococcus phage TP901-1 and Pseudomonas phage phiPSA1.

**Time-resolved analysis of protein expression and assembly**

Encouraged by the high proteome coverage, we next focused on comparing the time frame required for phage assembly with the temporal profile of phage protein expression. We monitored the protein expression from a 70 $\mu$L cell-free reaction over 10 time points (0-240 minutes) by assessing protein composition and T7 phage titer. Specifically, half of the removed volume at each time point (3 $\mu$L) was immediately mixed with LDS to stop the expression and assembly process (Figure 27a). From the remaining volume, the phage titer was determined via a plaque assay (Figure 27b). The mass spectrometric results show that most phage proteins are present after 40 minutes (Figure 27a), but fully assembled and functional phages first appear after 60 minutes, as determined via plaque assay (Figure 27b). Interestingly, four proteins are expressed particularly late, i.e., their highest concentration level is reached way after 40 minutes. These four proteins entail the two subunits of the terminase, (TERL and TERS, necessary for the loading of DNA into the phage capsid [283]), the cell lysis protein Holin (necessary for the timed release of the phage from the bacteria,[284, 285] and Y195, a protein which has not been experimentally validated or functionally characterized before. In combination with the plaque assay results (Figure 27b), which show the first assembled phages between 60-90 min, these proteins appear to be the limiting factors. On this basis, we aligned the protein expression patterns to the position of their coding sequence on the T7 genome. The graph in Figure 27c is in compliance with the current consecutive expression model of the T7 phage. It is even possible to analyze the kinetics of cell-free expression of every single protein with molecular precision (Figure 27d), exemplarily shown for

87
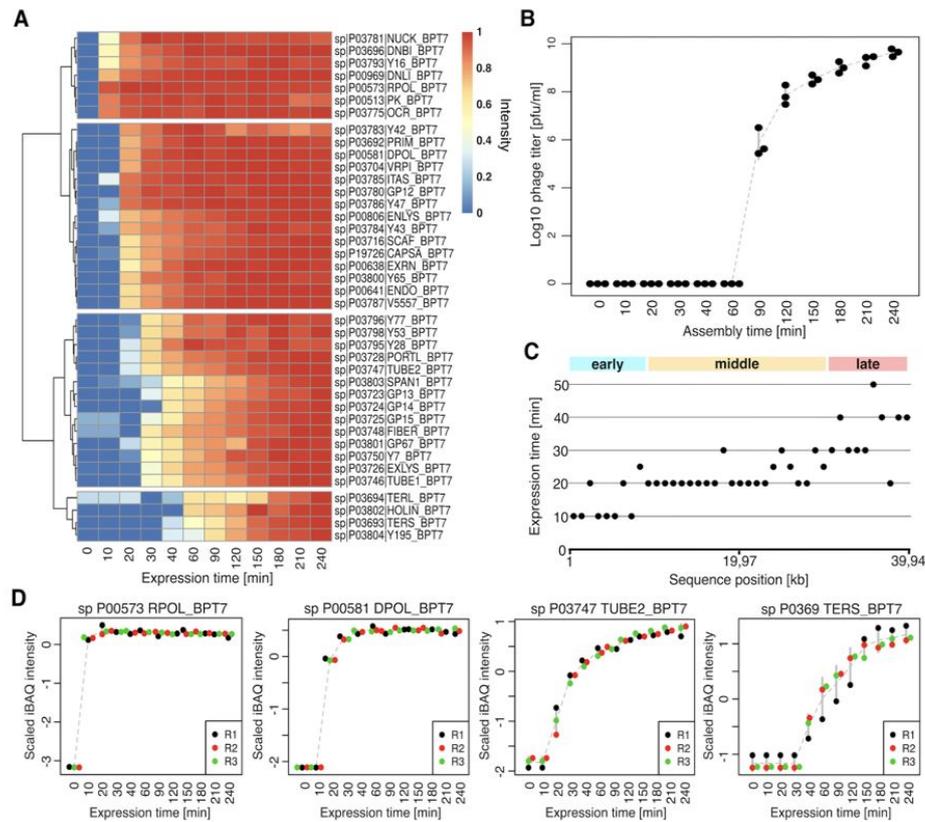
Figure 27: Time-resolved cell-free protein expression analysis of the T7 phage elucidates the self-assembly process. Protein expression and phage assembly were monitored over time from one single cell-free reaction over different time points for proteome characterization and titer determination. (a), The orchestrated expression of the T7 proteome was analyzed by mass spectrometry from 0 to 240 minutes within the cell-free system. The heatmap shows the relative abundance of proteins at different time points. (b) Phage titers during the 240 min phage assembly in a cell-free reaction determined with Plaque-Assay. (c) Mapping of protein expression on the phage genome accurately recapitulates early, middle, and late gene products in the synthetic system. The time-points of detected proteins crossing a relevant iMS intensity (50% max. MS intensity) were aligned to the position in the T7 genome. (d) Protein expression patterns can be differentiated with molecular precision. Early genes, e.g., DNA/RNA polymerase (P00581 and P00573) are already expressed after 10 min, whereas structurally relevant proteins, e.g., TUBE2 and late proteins such as TERS are expressed after 20 and 30 min. All data is obtained from three biological replicates. Error bars in (b) and (d) represent the standard deviation of the mean (n=3).

88

the T7 RNA polymerase (RPOL), the middle gene T7 DNA polymerase (DPOL), and the late genes tail tubular gp12 (TUBE2) and Terminase S (TER S).

## 7.3 Discussion

We have introduced an enhanced cell-free expression platform that facilitates the host-independent one-pot synthesis of clinically relevant and gram-positive bacteria targeting phages enabling the systematic identification of new phage proteins and accelerating phage engineering.

The titers that could be achieved for the therapeutic phages PhiA1122 and CLB-P3 targeting *Y.pestis* and EAEC, respectively, were substantially higher (fivefold) than those obtained by amplification in bacterial liquid culture (Figure 25b, c). This may be explained by the fact the cell-free system provides a stably buffered solution with high metabolite concentration and is unaffected by phage-induced bacterial defense mechanisms such as CRISPR or by cell death.[286]

We could furthermore substantially increase the diversity of bacteriophages that can be produced by cell-free expression. Our data show that co-expression of the primary *B.subtilis* RNA polymerase sigma factor sigA, which mediates the interaction of RNA polymerase with *B.subtilis*-specific promoter sequences,[287] enables cell-free production of the Phi29 and Goe1 phage targeting gram-positive bacteria (Figure 25b,c). This is in line with the finding by Whiteley et al. [288, 289] that the *B.subtilis* sigA factor can direct the *E.coli* RNA-polymerase to transcribe foreign promoters, even if the natural sigma70 cofactor is still active in the system. Our results suggest that the *E.coli* and *B.subtilis* RNAP core enzymes are sufficiently similar to allow for some promiscuity in the integration of the host factor. Our data also make it appear likely that additional phages for hosts distinct from *E.coli* could be generated by supplementing appropriate host factors (Figure 25c). With this approach, it should be possible to generate a variety of phages that tackle different kinds of MDRs for which the development of antibiotic resistance has become a major global health problem. This particularly applies to *E.coli* related bacteria, including *K.pneumoniae*, *E.faecium*, and *S.aureus* in the case of gram-positive bacteria.[262, 260]

Furthermore, we demonstrated an approach for transient phage engineering by temporarily integrating a luciferase and a polyhistidine tag in the capsid structure of the T7 phage. This form of phage bioengineering is modular and convenient as

it only requires the addition of plasmids without the need to modify the phage genome. The method thus also offers a high degree of biosafety because the non-genomic modification cannot be passed on to the offspring. This transient phage bioengineering at the protein level can thus become a powerful method to, e.g., systematically modify tail-fibers to respond to MDRs without risking the spread of genetically modified phages to the environment. In addition to its use for therapeutic and safe phage production, our cell-free expression system proved to be a reliable and powerful platform for the identification of phage proteins by mass spectrometry. In order to demonstrate this capability, we investigated the proteome of the well-characterized phage T7 as well as that of the previously uncharacterized phage CLB-P3. For both phages, the proteomics data generated from the cell-free samples was highly reproducible and enabled the identification of over 44 proteins with almost complete proteome coverage that had been previously categorized only as hypothetical (Figure 26b). By contrast, samples obtained via bacteria-dependent production displayed less identified proteins with a high variability of the obtained intensities and low peptide counts in this case. In these samples, many of the proteins could not be identified reproducibly (Figure 26b).

The cell-free system should also give preferred access to non-structural phage proteins as these are likely to be lost in conventional protocols that purify phages from bacterial cultures. In fact, phage stock samples enrich functioning phages including the structural phage proteins during PEG purification, e.g., tail fiber proteins, stem proteins, and base plate proteins. On the other hand, soluble factors necessary for phage amplification and host subjugation, e.g., polymerases, nuclease inhibitors, etc., are probably lost during buffer exchange, as they are not physically connected to the phage structure.

We thus analyzed the identified proteins with respect to these two categories and found that the set of proteins in both the cell-free and bacteria-derived samples of phage T7 could be assigned mainly to structural proteins as opposed to soluble factors not attached to the phage structure. By contrast, those proteins exhibiting lower MS intensities intensity were mostly related to proteins that are not part of the phage structure (Figure 26b). As the majority of the hypothetical proteins clustered together with already characterized non-structural T7 proteins, we assume that these proteins also belong to the class of non-structural proteins. A similar intensity distribution was found for the previously uncharacterized CLB-P3 phage.

90

We thus expected that those found in both types of preparations - as in the case of phage T7-proteins of the "core proteome" - are structural. Accordingly, we assume that proteins that are exclusively well-represented in samples from the cell-free system are mainly non-structural. It is likely that this identification method for phage proteins, exemplarily demonstrated for CLB-P3, will facilitate the further characterization of other phages. It can for instance be seen that a range of CLB-P3 proteins is well conserved in many bacteriophages, while others appear to be host-specific (Enterobacteria/Escherichia phage). It is interesting to see that CLB-P3 shares a large portion of proteins with Escherichia phage ADB-2 and Shigella phage pSf-2, highlighting its clinical potential.

We also observed that the proteome detected from the cell-free system itself (the *E.coli* proteome) is ten-times less concentrated than the actual bacterial cytosol, allowing the detection of less abundant phage proteins by MS analysis. Given that several of these proteins are highly conserved in other phages, it will be informative to test whether they exist and serve similar functions in those other phages as well. With our methodology, we also introduce a model system for phage protein expression analysis. When examining the time course of phage protein expression, the natural protein expression patterns for phage assembly is closely recapitulated *in vitro*. The *E.coli* RNA polymerase dependent proteins cluster separately from the T7 RNA polymerase dependent proteins (Figure 27a), and the T7 typical order of gene expression from early via middle to late genes is confirmed (Figure 27c).[290] However, the overall expression trajectory appears to be decelerated by about a factor of two to three compared to the *in vivo* case, where T7 phage replicates typically within 20 to 30 min.[291] This may be partially explained by the dilution of the bacterial cell extract with respect to the bacterial cytosol (by a factor of 10) which decreases reaction rates in the extract, potentially counteracted by the addition of the crowding agent polyethylene glycol.

Interestingly, we have discovered a cluster of proteins (TERL, TERS, HOLIN, and Y195) that express about 20 minutes later than the designated "late genes" (Figure 27a). This is an interesting finding as it suggests a particular regulatory mechanism for this group of genes and points to protein products that may be involved in phage assembly and release. In fact, it had been previously suggested that TERS and TERL are responsible for the loading of DNA into the phage capsid [283] and HOLIN for the disruption of the bacterial cell wall [285]. In compliance with the phage titer (Figure 27b), we determined that the first infectious phages are gener-

ated within a time frame of 60-90 minutes. This also indicates a crucial role of the previously hypothetical protein Y195 in finalizing the phage assembly or triggering the release process. In summary, based on a number of methodological advances, we have developed an optimized cell-free production system for bacteriophages that enables the rapid, reliable, and host-independent production of clinically relevant phages also against gram-positive bacteria. This versatile one-pot reaction also allows for the systematic analysis of phage proteomes and bioengineering of phages at the protein level, which will accelerate the biological development of tailor-made phage therapies against the growing number of pathogenic MDR bacteria and the emerging threat of bioterrorism.

# 8 Summary and Outlook

All the different structures built in this thesis have in common that they were created by self-assembly. The main advantage of employing this principle is the simultaneous creation of large amounts of the same construct in a one-pot reaction. This enables the production of nanoscale objects in higher amounts, in contrast to conventional methods, where each part is assembled through external control one after the other.

For the combination of the cell-free transcription-translation system with the expression of polymers capable of self-assembly, a variety of applications were shown. In the search for an artificial system that is capable of replication substantial progress was made, as growth by incorporation of the membrane building blocks of peptide vesicles was shown and the *in vitro* replication of the T7 phage DNA. As the genome of the T7 phage is rather large this system could have the potential of encoding a huge variety of proteins, besides the genes which are necessary for the replication itself. Both successes were achieved in one-pot reactions by the expression of the encoding genes in the cell-free system.

To develop this system further to a fully self-replicating protocell, these systems would have to be coupled. As a first step, the peptide sequences for the membrane of the vesicles would have to be encoded into the T7 phage genome, replacing the current capsid proteins. These biological entities would still rely on some external supply of nutrients and proteins for the expression machinery. Based on that system other necessary encoding genes would have to be cloned stepwise into this system to create a more and more independent artificial cell, until the point where with guided evolution experiments this biological entity could be adapted further independently.

Besides the more fundamental approach of self-assembly in the cell-free system, such as the creation of a protocell is the production of bacteriophages for therapeutic usage. Here the main advantage is the host-independence of the cell-free system. As with conventional methods for therapeutic phage production the pathogenic host is still required.[292] These difficulties concerning the safety and quality of the phage preparation are faced, which is one of the reasons why the potential of phages for fighting drug-resistant bacteria is still not fully developed. The major impurity in such preparations are possible prophages which can be present in the genome of the host-bacteria of the phages, bacterial host DNA packed into the capsid of the

phage but also endotoxins, which requires elaborate purification techniques.[271] A further limitation is also that only a minority of bacteria can be properly cultivated in the lab for phage production.[293] Another important step in the preparation of drugs is also the standardization of the production. As the conventional method for phage preparation for therapeutic usage is relying on the host-bacteria a standardization would be only possible for single phage therapeutic, which makes it a time-consuming and expensive process and therefore not interesting for a broader application.

To enable the potential of phage therapy a standardized host independent way for the production would be required. With the cell-free platform for host-independent expression of phages, a broader spectrum of patients can be reached. As the production of potentially therapeutically relevant phages was shown in this thesis, the next step towards the application of phages in patients would be the regulatory approval. In Belgium a pragmatical legal framework for phage therapy was introduced, the so-called magistral preparation.[294] As the first step of a certification process the phage and the DNA of the production-host are evaluated to exclude possible lysogenic phages. These can spread antibiotic resistances and pathogenicity factors and therefore must be excluded from therapeutic preparations.[271] After successful approval of the phage and production-host pair, the phage solution is tested for their endotoxin content, the sterility, and the pH value. If these values are all in the required range the therapeutic phages can be handed over to a pharmacist as they must prepare the therapeutic phages for the application in the patient. After these steps, the phages can be applied to patients. For the first applications in patients the cell-free system will be further developed and standardized so that in the future, phage therapy can be applied in a broader way to help fighting the emerging antibiotic crisis.

# 9 Appendix

## 9.1 Materials and Methods

Remark: Parts of this appendix were adapted from the methods section and supplementary information of the given references. Techniques which are used in the previous methods sections of the different papers are not repeated, if they are the same.

**Methods to Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch [7]**

**Matrials**

Fluorescently labeled DNA oligomers with PAGE purification were ordered from Biomers (Ulm, Germany). Other DNA was purchased from Eurofins MWG with HPLC purification. Gold nanoparticles were ordered form BBI Solutions. NHS-azide linker and alkyne-functionalized DNA were purchased from Baseclick (Munich, Germany). All other reagents and solvents were of analytical grade and were purchased from Sigma-Aldrich.

**DNA Origami Assembly**

About 200 short staple strands were used to fold a 7249 nucleotides (nt) long single-stranded "scaffold" strand from phage M13mp18 [123] into a twist-corrected rectangular origami structure and a twist-corrected cross-shaped structure, respectively. For both DNA origami structures a 3-fold excess of standard staples and 5-fold excess of extended staples were used for the implementation of the desired binding sites. As folding buffer (FB) we used $1\times$ TAE with 12.5 mM $MgCl_2$. For the folding of the structures an initial temperature of 70 °C was used, followed by a temperature ramp from 65 to 40 °C within about 1 h followed by a temperature jump to 20 °C. After the assembly of the structures the solution was purified using poly(ethylene glycol) (PEG) precipitation.[124] The sample was mixed with 15 % PEG-8000 dissolved in FB with 100 mM NaCl (further denoted as PEG solution) in a 1:1 (v/v) ratio followed by centrifugation at 16000 rcf for 25 min. The supernatant was removed, and the pellet redissolved in FB and afterward mixed with the PEG

solution at a final ratio of 1:1 (v/v). The last step was repeated three times. After the last centrifugation the supernatant was removed and the pellet redissolved in 1× FB. The concentration of the DNA origami structures was measured by absorption spectrometry (Nanophotometer IMPLEN vers. 7122 V2.3.1, Munich, Germany) using an extinction coefficient of $1.12 \times 10^8$ cm-1 M-1 and adjusted to 100 nM.

## Expression of Elastin-like Polypeptides

Polypeptides were prepared by overexpression in *E.coli*. The genes encoding the elastin-like polypeptide sequences were designed and synthesized using methods reported previously [69, 62] and ultimately were ligated in a pET20b expression vector (Novagen) and transformed into the BL21(DE3) strain of *E.coli*. As confirmed by DNA sequencing, the gene encoded the polypeptide sequence MSKGP(GVGVP)$_{40}$ GWPC (abbreviated as V40C). The ELP was expressed in a shaker at 37 °C in a 1 L culture of LBA media (10 g of peptone, 5 g of yeast extract, 5 g of NaCl, and 100 mg of ampicillin) induced with 240 mg of IPTG (isopropyl $\beta$-d-1-thiogalactopyranoside) when the optical density at 600 nm reached 0.8. Five hours after induction, the bacteria were isolated through centrifugation. The bacteria were lysed by sonication in phosphate-buffered saline (PBS, pH 7.4) supplemented with 2 mM EDTA. ELPs were purified through sequential cold (4 °C) and warm (40 °C) centrifugations. During cold centrifugations, the ELPs remained in the supernatant, while during warm centrifugations they phase-separated into the pellet, which was resuspended in PBS. After three cycles of centrifugation, the ELPs were prepared at 100 $\mu$M in PBS.

## Conjugation of ELPs to DNA

Initially, elastin-like peptides were activated with an azide group and conjugated to DNA via copper-based azide-alkyne Huisgen cycloaddition (further denoted as click-chemistry). The used NHS-azide linker (y-azidobutyric acid oxysuccinimide ester) was diluted in DMSO to a final concentration of 20 mM. ELPs were dissolved in 1× PBS (8 g/L NaCl, 2 g/L KCl, 1.42 g/L Na$_2$HPO$_4$, 0.27 g/L K$_2$HPO$_4$, pH 6.8-7.0). The peptides were mixed with a 2-fold excess of linker and incubated for 12 h at room temperature. To remove residual NHS-azide linker, an equal volume of 5 M NaCl was added, which lowers the transition temperature and precipitates all

96

ELP compounds. The sample was centrifuged at 16000 rcf for 10 min, and the supernatant was replaced by $1\times$ PBS at room temperature and incubated in a refrigerator at 4 °C for 10 min. After redissolving the ELPs the purification was repeated three times.

The coupling of V40C and DNA linker L4 was carried out by click-chemistry. The alkyne-modified L4 was mixed with V40C at a molar ratio of 1:4 (DNA:V40C); the sequence of L4 is poly(dA) with an alkyne group attached at the 5' end, alkyne-AAAAAAAAAAAAAAAAA. Afterward we added 1 mM TBTA (tris(benzyltriazolyl - methyl) amine), 10 mM TCEP (tris(2-carboxyethyl)phosphine hydrochloride), and 10 mM $CuSO_4$; all given concentrations are final concentrations. The mixture was incubated at 4 °C for 12 h. Remaining linker strands were removed using the purification procedure mentioned in the last paragraph.

**Atomic Force Microscopy**

For AFM imaging a sample volume of 5 $\mu$L with 5 nM DNA structures was deposited onto freshly cleaved muscovite mica (Plano GmbH, Wetzlar, Germany) followed by the addition of 60 $\mu$L of FB. We used the AC mode of the Cypher ES Environmental AFM (Asylum Research, Santa Barbara, CA, USA), Asylum Research's blueDrive for photothermal excitation [295] of the cantilever, and AR software version 13.06.82 implemented in IGOR Pro 6.34A. The used silicon cantilever (Olympus microcantilevers BL-AC40TS-C2) had a nominal force constant of 0.09 N/m and a nominal resonance frequency in water of 25 kHz. The used scan rates were between 1 and 4 Hz. Image processing was carried out using the aforementioned software and the plugin Scale Bar Tools for Microscopes for Java-based software ImageJ.

**Scanning Electron Microscopy**

For SEM imaging silicon wafers were used as a substrate and subsequently cleaned by rinsing with acetone, ethanol, and water. After drying under a gentle stream of nitrogen the silica suspension was deposited on the cleaned silicon surfaces. After 5 min the remaining fluid was removed and the wafer was dried in the oven at 50 °C for 30 min. Imaging was performed using an e-LiNE (Raith) system at an acceleration voltage of 5 kV and an aperture of 10 $\mu$m.

**Transmission Electron Microscopy**

Before staining a final concentration of 7 % DMSO was added to the DNA sample.[296] The DNA origami structures were adsorbed on glow-discharged Formvar-supported carbon-coated Cu400 TEM grids (FCF400-CU, Science Services, Munich, Germany) for 30 s and afterward positively stained using a 2 % aqueous uranyl formate solution with 25 mM sodium hydroxide for 60 s. Afterward the grid was washed with water and dried under vacuum for 30 min. Imaging was performed using a Philips CM100 transmission electron microscope at 100 kV. An AMT 4 megapixel CCD camera was used for acquiring images and performed at $18.000\times$ up to $28.500\times$ magnification. Image processing was carried out using the plugin Scale Bar Tools for Microscopes for Java-based software ImageJ.

**Preparation of Gold Nanoparticles Including DNA Coating**

For the preparation of the gold nanoparticles we adapted the protocol from Kuzyk et al..[92] Initially, the commercial AuNP suspension was concentrated. Then 50 mL of 10 nM AuNPs was dissolved in 20 mg of BSPP (bis(p-sulfonatophenyl)phenyl-phosphine dihydrate dipotassium salt) while stirring for 3 days at room temperature. Afterward a 5 M NaCl solution was added until the color of the sample changed to blue. The suspension was centrifuged at 1600 rcf at room temperature for 30 min, and the supernatant was discarded. The pellet was redissolved in 800 $\mu$L of BSPP solution (2.5 mM in H2O), and an equal volume of methanol was added afterward. This suspension was centrifuged at 1600 rcf at room temperature for 30 min. The supernatant was discarded and the pellet again redissolved in 800 $\mu$L of BSPP solution (2.5 mM in H2O). The concentration of AuNPs was determined by absorption spectroscopy (Nanophotometer IMPLEN vers. 7122 V2.3.1) at a wavelength of 518 nm and using an extinction coefficient of $8.5 \times 10^7$ cm-1 mol-1. In the second step, AuNPs were coated with thiolated DNA. We used three thiolated DNA linkers with the sequences thiol-5'-TTCTCTCTCTCTCTCTTTCT-3' (denoted as L1), thiol-5'-TTTTCTTTCCTTCCTTCCTT-3' (denoted as L2), and thiol-5'-TTCTTCTTCTTCTTCTTCTT-3' (denoted as L3) for the specific attachment of AuNPs. The following steps were used for all three types of nanoparticles. The DNA linker solution was mixed with TCEP (final concentration 10 mM) and incubated for at least 30 min. Afterward AuNPs were mixed with a 100-fold excess of the DNA linker followed by an addition of sodium citrate buffer in a final con-

centration of 20 mM and a pH adjustment to pH 3 by adding 1 M HCl. After 3 min HEPES buffer was added in a final concentration of 20 mM with pH 7. To remove the unbound oligonucleotides, 100 kDa MWCO centrifugal filters (Amicon Ultra, Merck Millipore) were used. The filters were filled with 400 $\mu$L of 0.5$\times$ TBE and centrifuged at 2000 rcf for 5 min. Afterward the filters were loaded with the sample and centrifuged at 10000 rcf for 10 min followed by an additional four centrifugation steps (10000 rcf, 10 min) with 400 $\mu$L of 0.5$\times$ TBE added to the filter before each step. To remove the AuNP-DNA solution from the filter, it was flipped, placed into a fresh tube, and centrifuged at 10000 rcf for 10 min. The filtration was repeated one more time.

We also prepared a fourth type of AuNP coated with the diblock copolymer L4-V40C. Here, the cysteine of the ELP was used to bind the copolymer to the AuNPs; the DNA linker L4 does not contain a thiol group. The ELP-DNA copolymer solution was mixed with TCEP (final concentration 10 mM) and incubated for at least 30 min to prevent the formation of disulfide bonds. Afterward AuNPs were mixed with a 100-fold excess of L4-V40C conjugate followed by the addition of sodium citrate buffer in a final concentration of 20 mM and a pH adjustment to pH 3 by adding 1 M HCl. After 30 min HEPES buffer was added in a final concentration of 20 mM with a pH of 7. Unbound L4-V40C conjugates were removed by centrifugation at 10000 rcf at 4 °C for 10 min. The supernatant (without AuNPs) was replaced by cooled water. This purification was repeated five times.

**Silica Protocol: Optimized Protocol for the Functionalization of DNA Origami Structures with AuNPs**

Silica particles were suspended in 1$\times$ FB and can be optionally cleaned from contaminants beforehand. In the next step a solution of DNA origami structures was mixed with the silica suspension at final concentrations of 5.7 nM DNA structures, 1$\times$ TAE, and 60 mM MgCl$_2$. If necessary a final concentration of up to 300 mM NaCl can be added. After 5 min of gentle shaking a suspension of AuNPs was mixed with this silica-DNA suspension and gently shaken for 2 min. If longer shaking times were used, the binding yield decreased slightly. Here, a 3-fold excess of AuNPs relative to the amount of binding sites at the DNA origami structures was used. A 1:1 ratio would slightly decrease the binding yield. Afterward the suspension was centrifuged at 20000 rcf for 3 min and the supernatant was replaced by 1$\times$

TAE to elute the DNA structures for 3 min, followed by another 3 min centrifugation at 20000 rcf. The supernatant, which contained the AuNP-functionalized DNA structures, was mixed with 10× FB to obtain a 1× buffered solution.

**Fluorescence Spectroscopy**

The fluorescence spectroscopy experiments were performed with a Cary Eclipse spectrometer (Agilent Technologies Deutschland GmbH, Böblingen, Germany). In all measurements both FAM and Atto 532 were simultaneously excited at $490 \pm 5$ nm. For each experiment 60 $\mu$L of sample was filled into fluorescence cuvettes (105.251-QS, Hellma GmbH and Co. KG, Müllheim, Germany), and 100 $\mu$L of hexadecane was placed on the top to prevent evaporation. Emission spectra were recorded at a constant temperature of 20 °C using a Peltier block. Temperature ramp experiments (Figure 12b) were carried out by increasing the operating temperature from 15 °C to 35 °C at 0.1 °C/min, and the emission was measured at $520 \pm 5$ and $550 \pm 5$ nm. The data were normalized by dividing the signal by a blank signal that contained two types of AuNPs (5.7 nM) with FAM and Atto 532, respectively attached. In Figure 12b we used the fully assembled waveguide and additionally suspended L4-V40-AuNP conjugates (5 nM) in FB as a control. The transition temperature was determined from the inflection point of a sigmoidal fit. For the cycling experiments (Figure 12c) the waveguide with ELPs was heated and cooled between a temperature of 15 and 35 °C at 0.5 °C/min. The emission was measured at $550 \pm 5$ nm and normalized as mentioned before using a working waveguide without implementing an ELP switch as a blank.

**Methods to Towards Synthetic Cells Using Peptide-Based Reaction Compartments [6]**

**Expression of elastin-like peptides**

For bacterial ELP expression, the peptide gene was cloned into a pET20b(+) expression vector (Novagen) and transformed into the BL21(DE3)pLysS strain of *E.coli*. As confirmed by Sanger sequencing, the gene encoded the polypeptide sequence MGHGVGVP((GEGVP)$_4$(GVGVP))$_4$((GFGVP)$_4$(GVGVP))$_3$(GFGVP)$_4$GWP (abbreviated as EF). ELP expression was performed in a culture flask shaker at 37 °C in a 1 L culture of LB (Luria/Miller) medium (10 g of tryptone, 5 g of yeast extract, 10 g of NaCl, and 100 mg of carbenicillin), induced with 240 mg of IPTG (isopropyl $\beta$-D-1-thiogalactopyranoside) when the optical density at 600 nm reached approximately 0.8. After 16 h of incubation at 16 °C, the bacteria were harvested through centrifugation. The bacteria were lysed by sonication in phosphate-buffered saline (PBS, pH 7.4) supplemented with lysozyme (1 mg/mL), 1 mM PMSF, 1 mM benzamidin and 0.5 U of DNase I. After lysis 2 mL of 10% (w/v) PEI was added per 1 L of original cell culture. The samples were incubated at 60 °C for 10 min and afterwards at 4 °C for 10 min, followed by a centrifugation at 16000 × g at 4 °C for 10 min.

The ELPs were purified through sequential centrifugations under acidic (pH 2) and neutral (pH 7) conditions. For the pH adjustment phosphoric acid and sodium hydroxide were used. During centrifugations at pH 7, the ELPs remained in the supernatant, while during centrifugations at pH 2 they phase-separated into the pellet, which was re-suspended in water. After three cycles of centrifugation, the ELPs were dissolved in water at a concentration of 700 $\mu$M. The concentration of the peptides was measured using absorption spectrometry (Nanophotometer IMPLEN vers. 7122 V2.3.1, Munich, Germany), assuming an extinction coefficient of 5500 M/cm at 280 nm.

**Glass beads method**

Two-hundred microliters of concentrated 1.1 mM ELP solution was mixed with 1250 $\mu$L of a 2:1 chloroform/methanol mixture, for fast evaporation. A total of 1.5 g of spherical glass beads (212 $\mu$m to 300 $\mu$m in size) were added to a round-bottom flask. Using a rotary evaporator the solvent was evaporated, resulting in

a peptide film on the glass beads. For further experiments 100 mg of the glass beads were mixed with 60 $\mu$L of the swelling solution containing the molecules to be encapsulated. After an incubation for 5 min at 25 °C, the vesicles had formed and the sample was centrifuged to sediment the glass beads. The vesicle solution was removed using a pipette.

**Western blotting**

Samples were mixed with 2× Laemmli buffer and heated to denature the peptide structure. SDS PAGE (12 %) was used for separation of the sample components. For further analysis, the peptides were fixed on a PVDF (polyvinylidene difluoride) membrane by transferring the content of the SDS gel to the membrane using a Semi-Dry blotter. The peptide-free areas of the membrane were blocked by incubation in a blocking solution containing bovine serum albumine (BSA). Afterwards, the membrane was rinsed several times with PBST (phosphate buffered saline with Tween 20) to remove residual BSA. The detection of the immobilized peptides was carried out by incubating the membrane with a specific anti-His antibody (6 ×-His Epitope tag antibody, mouse, purchased from Life Technologies GmbH, Darmstadt, Germany: catalog number MA1135, clone 4E3D10H2/E3) at 4 °C overnight. Residual antibodies were removed by washing with PBST. For visualization secondary antibodies (anti-mouse Alexa Fluor 680, goat, purchased from Life Technologies: catalog number A28183) were added onto the membrane and incubated for 1 h at room temperature. Residual secondary antibodies were removed through washing with PBST before the membrane was imaged using a fluorescent scanner (Typhoon Fla 9500, GE Healthcare Life Science). Both antibodies were used at dilutions of 1:1000.

**Transcription reaction (TX)**

The sequence of the RNA aptamer dBroccoli was GAGACGGTCGGGTCCATCTG-AGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTCA-GATGTCGAGTAGAGTGTGGGCTC.[237] The transcription solution contained 1 × RNAPol reaction buffer (40 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 2 mM spermidine), 125 mM KCl, 15 mM $MgCl_2$, 4 mM rNTP, 10 $\mu$M DFHBI, 200 nM DNA template, 4 U/$\mu$L T7 polymerase (NEB, M0251S),

0.5 U/$\mu$L RNase inhibitor murine (NEB, M0314S) and water. All measurements took place at 37 °C.

**Transcription translation reaction (TX-TL)**

For the generation of crude S30 cell extract a BL21-Rosetta 2(DE3) mid-log phase culture was bead-beaten with 0.1 mm glass beads in a Minilys homogenizer (Peqlab, Germany) as described in ref. [217]. The composite buffer contained 50 mM Hepes (pH 8), 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM coenzyme A, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM PEP, 1 mM DTT and 2 % PEG-8000. As an energy source in this buffer phosphoenolpyruvate (PEP) was utilized instead of 3-phosphoglyceric acid (3-PGA). All components were stored at -80 °C before usage. A single cell-free reaction consisted of 42 % (v/v) composite buffer, 25 % (v/v) DNA plus additives and 33 % (v/v) S30 cell extract. For ATP regeneration 13.3 mM maltose and 1 U of T7 RNA polymerase (NEB, M0251S) were added to the reaction mix.[33] All measurements took place at 29 °C with 50 nM of plasmid if it is not indicated differently.

**Dynamic light scattering**

For the DLS experiments the instrument DynaPro Nanostar (Wyatt technology corporation) was used. The buffers were sterile filtered before usage and the samples were measured in a disposable cuvette. For one distribution a set of 50 single measurements were performed for 2 s and averaged afterwards. The values were averaged and processed with the DYNAMICS software using a CONTIN-like algorithm.

**Transmission electron microscopy**

The vesicle solution was adsorbed on glow-discharged formvar-supported carbon-coated Cu400 TEM grids (FCF400-CU, Science Services, Munich, Germany) for 2 min, followed by a negative stain using a 2 % aqueous uranyl formate solution with 25 mM sodium hydroxide for 45 s. Afterwards the grid was dried and stored under vacuum for 30 min. Imaging was carried out using a Philips CM100 transmission electron microscope at 100 kV. For acquiring images an AMT 4 megapixel CCD

camera was used and imaging was performed at magnification between $\times 8900$ and $\times 15.500$. For image processing the plugin Scale Bar Tools for Microscopes for Java-based software ImageJ was used.

## Flow cytometry

The flow cytometer measurements were performed by using a CyFlow Cube 8 cytometer (Sysmex Partec GmbH, Germany) equipped with a blue laser emitting at 488 nm. The measured signals were the forward scattering signal (FSC), the side scattering signal (SSC) and the fluorescence signal, which was band-pass filtered at 536 nm $\pm$ 40 nm. The buffers were sterile filtered and degassed before usage. For a measurement 100 $\mu$L of the sample was diluted with 500 $\mu$L 1$\times$ PBS (8 g/L NaCl, 2 g/L KCl, 1.42 g/ L $Na_2HPO_4$, 0.27 g/L $K_2HPO_4$, pH 6.8-7.0) and measured immediately. The analysis of the data was performed with the FlowJo v10 software (FlowJo LLC, USA).

## Fluorescence measurements

Cell-free expression and transcription was characterized via plate reader measurements, with the corresponding filter sets for the fluorescence (BMG FLUOstar Optima) using 15 $\mu$L reaction volumes in 384-well plates.

## Fluorescence measurements

The vesicles were prepared according to the glass beads method, with a mixture of 100 $\mu$L of 1.1 mM Cy3 labeled ELPs and 100 $\mu$L of 1.1 mM Cy5 labeled ELPs. For rehydration the TX-TL solution was used with the plasmid containing the EF gene (50 nM). Expression of EF occurring outside of the vesicles was suppressed using the antibiotic kanamycin. For the reference, kanamycin was added before the vesicle formation. Cell-free expression was characterized via plate reader measurements in a bulk measurement, with the corresponding filter sets for the FRET dyes (BMG FLUOstar Optima) using 15 $\mu$L reaction volumes in 384-well plates.

## Mass spectrometry

Full-length protein mass spectrometry was performed on a Dionex Ultimate 3000 HPLC system coupled to a Thermo LTQ-FT Ultra mass spectrometer with electrospray ionization source (spray voltage 4.2 kV, tube lens 120 V, capillary voltage 48 V, sheath gas 60 arb, aux gas 10 arb, sweep gas off). In all, 2.5 $\mu$L of sample corresponding to 1.64 nmol of peptide were on-line separated using a BioBasic-4 column (Thermo; 150 mm $\times$ 1 mm, 5 $\mu$m) by applying a multistep gradient from 2 % to 20 % eluent B over 6 min; 20 % to 25 % B over 1 min and 25 % to 85 % B over 14 min (eluent A: water with 0.1 % (v/v) formic acid; eluent B: 90 % (v/v) water, 10 % (v/v) acetonitrile with 0.1 % (v/v) formic acid; flow: 0.2 mL/min). All solvents were of liquid chromatography-mass spectrometry grade. The mass spectrometer was operated in positive mode collecting full scans at R = 50000 from m/z 400 to m/z 2000. Collected data was deconvoluted using Thermo Xcalibur Xtract algorithm.

## Click chemistry

Initially, elastin-like peptides were activated with an azide group and then conjugated to dyes via copper-based azide-alkyne Huisgen cycloaddition (denoted as click chemistry). The used NHS-azide linker (y-azidobutyric acid oxysuccinimide ester) was diluted in DMSO to a final concentration of 20 mM. ELPs were dissolved in 1$\times$ PBS (8 g/L NaCl, 2 g/L KCl, 1.42 g/ L $Na_2HPO_4$, 0.27 g/L $K_2HPO_4$, pH 6.8-7.0). The peptides were mixed with a 2-fold excess of NHS-azide and incubated for 12 h at room temperature. To remove residual NHS-azide, the sample was loaded into a 10 kDa dialysis cassette and stored at 4 °C for 12 h. In the next step, the activated EF and dye were conjugated using the aforementioned click chemistry. The alkyne-modified dye was mixed with activated EF at a molar ratio of 1:1 (dye:ELP). Afterwards 1 mM TBTA (tris(benzyltriazolylmethyl)amine), 10 mM TCEP (tris(2-carboxyethyl)-phosphine hydrochloride), and 10 mM CuSO4 were added; all given concentrations are final concentrations. The mixture was incubated at 4 °C for 12 h. Remaining linker strands were removed by dialysis with a 10 kDa dialysis cassette, which was stored at 4 °C for 12 h.

## Methods to Target Specific Silencing of Genes in Cell-Free Systems for Controlling and Modification of Bacteriophages

### Phage assembly

Phages were assembled corresponding to Rustad et al.[280] with the following adjustments. Phage DNA was mixed with a TX-TL system based on cell extract, an energy solution and an amino acid solution as described in Sun et al.[217] with one amendment in Buffer B (6 mM Mg-glutamate, 100 mM K-glutamate, 3 mM DTT, 1.5 mM each amino acid except leucine, 1.25 mM leucine, 50 mM HEPES, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/ml tRNA, 0.26 mM CoA, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA, 4 % PEG-8000).[230] For 6 reactions a 13 $\mu$L, 2.5 $\mu$L PEG 8000 (64 % w/v), 4 $\mu$L dNTPs (25 mM), 0.8 $\mu$L ATP (500 mM), 37.5 $\mu$L. Buffer B, 2 $\mu$L GamS (150 $\mu$M), 28.5 $\mu$L TX-TL and 1.6$\mu$L DNA (10 nM) were mixed with nuclease free water to a final volume of 80 $\mu$L. All constituents were mixed (except DNA), chilled for 5 min on ice, followed by the addition of DNA. This 13 $\mu$L assembly mix was incubated for 4 h at 29 °C to generate the phages.

### qPCR

For the quantification of the amplification of the T7 DNA the samples were flash frozen in liquid nitrogen and stored at -80 °C until all samples were collected. The qPCR reactions were performed on a BioRad IQ5 instrument by cycling 1x 1 min 95 °C, 45x 30 s 95 °C and 15 s 60 °C, 1x melt curve 55–95 °C. The reactions were prepared with 7 $\mu$L of 1:100 diluted cell extract sample containing the DNA and LunaScript Universal MasterMix 2x (New England Biolabs) in white PCR stripes with flat lid (AB-1191, ThermoFisher). Three technical replicates were recorded for each sample. The cycle threshold (ct) values were calculated by the intersection of the fluorescence curve from the DNA dye at 20 % of the maximum intensity of the brightest sample. From the ct value the concentration was determined based on a beforehand measured calibration curve.

### Droplet generation

FC-40 oil (Sigma Aldrich, F9755) with 2 % (w/w) PFPE/PEG-surfactant (Raindance Technologies) were used to create the droplets by shaking. After the reaction, droplets were broken by adding perfluorooctanol (PFO; 370533, Sigma) to the droplets (five volumes of PFO to one volume of the aqueous drop-let contents).

### T7 Stock preparation

A single plaque of the T7 phage was picked to incubate at 37 °C and 250 rpm in a rotary shaker with the corresponding host-bacteria at an OD of 0.4 at 600 nm. The sample was incubated for approximately 2 h until the lysis cleared the solution. This was centrifuged at 5000 rcf at 4 °C for 5 minutes and the supernatant, including the phages was mixed with 10 % w/v PEG 8000 and 1 M NaCl for precipitation of the phages. After storage of the phages at 4 °C overnight the sample was centrifuged for 15 minutes at 4 °C at 7000 rcf. After discarding the supernatant, the precipitated phages were resuspended in phage buffer (1x PBS, 1mM MgCl2, 1mM MgSO4), followed by a filtration with a 0.45 $\mu$m sized filter. The resulting titer was measured by a plaque assay.

### Plaque assay

For the plaque-assay 0.5 % agarose NZCYM medium was melted, split into 4 mL aliquots and stored in a water bath at 48 °C. Separately, the assembled phages were diluted $10^2$-$10^8$-fold in phage buffer (1x PBS, 1mM MgCl2, 1mM MgSO4). 100 $\mu$L of each dilution was mixed with an equal volume of corre-sponding host bacterium overnight culture. This mixture was added to the 0.5 % agarose NZCYM medium aliquots and poured on a 1 % NZCYM agar plate. After the suspension solidified at room temperature, the plates were incubated at 37 °C until plaques became visible.

## Methods to A Cell-Free Platform for Characterizing and Producing Therapeutic Bacteriophages Against Biohazardous and Antibiotic-Resistant Bacteria

### Phenol-chloroform extraction

The phage stock ($10^6$-$10^{10}$ PFU/mL) was mixed at a 1:1 ratio with Roti-Phenol/ Chloroform/Isoamyl alcohol (pH 7.5-8.0) in a 5PRIME Phase Lock GelTM tube (Quantabio, USA). The tubes were gently inverted and centrifuged for 5 minutes at 16000 g at room temperature. Afterward, two volumes of pure chloroform (Roth, Germany) were added to the upper phase of the tube, inverted and centrifuged as before. The supernatant was transferred to a separate Eppendorf tube, and 20 $\mu$L of 3M sodium acetate was added. The DNA was precipitated with 1 mL of -80 °C pure ethanol (Roth, Germany) for 1h at -80 °C. The DNA the sample was centrifuged (16000 x g, 30 min, 4 °C), and the supernatant was discarded. After adding 1 mL of -20 °C cold 70 % (v/v) ethanol (Roth, Germany) the DNA was pelleted (16000 x g, 5 min, 4 °C) and the supernatant was discarded again. To evaporate the remaining ethanol the sample was stored at RT for approximately 15 min. The DNA pellet was dissolved with 30 $\mu$L of nuclease-free water (Thermo Fisher Scientific, USA).

### Transient engineering

For the transient modification, the phages were assembled as before, except for the addition of 0.2 nM of a plasmid encoding for a fusion of the major capsid protein of the T7 phage G10B via a 3xGS Linker to HiBiT and a 6xHis-Tag. (MLGVASTVAA-SPEEASVTSTEETLTPAQEAARTRAANKARKEAELAAATAEQGSGSGSVSGW-RLFKKISHHHHHH).

### His-Tag Purification

The assembled phages were diluted to $10^6$ PFU/mL in 1x PBS and 20 mM imidazole directly after the assembly in the cell extract in a total volume of 800 $\mu$L. The phage suspension was then applied onto Ni-NTA Agarose beads (250 $\mu$L of Ni-NTA Agarose from Qiagen ID: 30210), which had been pre-equilibrated with a washing buffer containing 1x PBS and 20 mM imidazole. The column was subsequently washed with 6 column volumes (800 $\mu$L) of 1x PBS and 20 mM imidazole. The

phages were eluted with one column volume of 1x PBS and 250 mM imidazole, before the titer was measured.

**HiBit-Assay**

The cell-free reaction was carried out as described before with or without the addition of 0.2 nM of a plasmid encoding either G10B-HiBit-His or just HiBit-His without G10B. 200 nM of T7 RNA polymerase was added to conduct transcription. After the incubation of the cell-free reaction, 20 $\mu$L was added to 1 ml of the corresponding bacteria (*E.coli* DSMZ 613) with an OD adjusted to 0.8 and incubated for 5 min at room temperature. The bacteria were pelleted down by centrifugation at 7000 x g at room temperature and the supernatant was discarded. The cells were resuspended in 1x PBS. The HiBiT was detected by measuring the luminescence after mixing with the Nano-Glo HiBiT Extracellular Detection System (Promega N2420). To confirm that the modification is only temporary, the progeny of the phages were isolated from a plaque assay and adjusted to the same titer as reached in the cell-free system. The luminescence was measured again with the Nano-Glo HiBiT Extracellular Detection System.

**Nanopore sequencing**

Phage DNA was prepared with the SQK-RAD004 kit (Oxford Nanopore Technologies, UK) on an FLO-MIN106 flow cell (Oxford Nanopore, UK) according to the manufacturer's protocol.

**Illumina Sequencing**

The phage genome library (Ultra II DNA library prep for Illumina, NEB) and sequencing (MiSeq v3 600 cycle kit, Illumina) was performed with the Illumina MiSeq platform according to the manufacturer's protocol. The DNA was sheared with an E220 supersonicator (Covaris).

**Genome assembly and annotation of phage genomes**

The reads derived from Oxford Nanopore MinION sequencing were basecalled using guppy (ONT) version Version 2.3.7+e041753. For the short reads, the Illumina

Miseq derived FASTQ-files (Miseq reporter version 2.6.2.3) were used. Reads mapping to *E.coli* K-12 MG1655 (U00096.3) were filtered from both MinION and Miseq reads. Following this, a hybrid assembly from both MinION and Illumina reads was performed using MaSURCA 3.3.2 in auto mode for the assembly graph k-mer size. After assembly, a quality control of the assembled sequence was performed by re-aligning the sequence itself (using BLAST) while aligning the Illumina or MinION reads onto the sequence for a coverage plot. The BLAST result was visualized in a dot plot, from which it was verified that the assembled sequence was a duplication of the actual genome. Likewise, the aligned Illumina reads showed a secondary alignment position in the first/second copy of the actual genome. After deduplicating the assembled sequence, the same alignment check was performed to confirm the correct deduplication. These positions were also verified by Polymerase Chain Reaction (PCR), and the generated PCR products were confirmed by Sanger Sequencing. Open-reading frames (ORFs) were identified using PHANOTATE.[297] Found ORFs were then compared to representative members for each pVOG class to determine the most similar representative, which allows assigning a function to a specific ORF.[282] Particularly those ORFs which could be associated with a pVOG are likely to be conserved, since similar sequences had already been found in other bacteriophages. To compare the found ORFs and the pVOG sequences, pairwise global alignments, using the PAM250 substitution matrix with gap open/extend costs -12 and -1, were computed (GOTOH algorithm). If for any alignment a sequence identity of the ORF with more than 40 % was detected, the alignment was accepted and the ORF was assigned the respective pVOG. The remaining ORFs (15) were blasted against all bacteriophage protein sequences downloaded from Genbank (file gbphg1.seq.gz). Again, an alignment was accepted if more than 70 % of the ORF length aligned and 50 % of the pVOG sequence aligned, respectively. Overall, for 72 of the 87 ORF a pVOG assignment was found. 8 ORFs were assigned a known bacteriophage protein and 7 ORFs remained unassigned.

**Sample preparation for mass spectrometry**

For sample preparation, 3 $\mu$l cell-free phage- or 5 $\mu$L native sample (stock concentration > $10^6$ PFU/mL) were diluted with NuPAGE LDS sample buffer (ThermoFisher). The samples were reduced with 50 mM dithiothreitol (DTT) at 70 °C for 10 min and alkylated with 55 mM chloroacetamide (CAA) at room temperature

for 30 min in the dark. The samples were run on a NuPAGE 4-12 % Bis-Tris protein gel (ThermoFisher) for 5 min (200V, 500 mA, 1x MOPS buffer), 30 min fixated (40 % MeOH, 2 % AA), colored (1x RotiBlue, Carl Roth, Germany) and destained (25 % EtOH, 1 % AA). All samples were prepared via in-gel digest(Shevchenko et al., 2006). The collected supernatant (130 $\mu$L) was dried in a centrifugal evaporator (Centrivap Cold Trap -50, Labconco, USA). The sample was freshly suspended before MS measurement in 20 $\mu$L washing solution (2 % ACN, 0.1 FA), and 5 $\mu$L was injected into the mass spectrometer per measurement.

**Proteomic data acquisition**

For LC-MS/MS the Ultimate 3000 RSLCnano system was coupled to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific Bremen, Germany). For each analysis, a final injection amount of around 0.1 $\mu$g of peptides was delivered to a trap column (ReproSil-pur C18-AQ, 5 $\mu$m, Dr. Maisch, 20 mm × 75 $\mu$m, self-packed) at a flow rate of 5 $\mu$L/min in 100 % solvent A (0.1 % formic acid in HPLC grade water). After 10 min of loading, peptides were transferred to an analytical column (ReproSil Gold C18-AQ, 3 $\mu$m, Dr. Maisch, 400 mm × 75 $\mu$m, self-packed) and separated using a 50 min gradient from 4 % to 32 % of solvent B (0.1 % formic acid in acetonitrile and 5 % (v/v) DMSO) at 300 nL/min flow rate. Both nanoLC solvents contained 5 % (v/v) DMSO. The Fusion Lumos Tribrid Mass Spectrometer mass spectrometer was operated in the data-dependent acquisition and positive ionization mode. MS spectra (360–1300 m/z) were recorded at a resolution of 60,000 using an automatic gain control (AGC) target value of 4e5 and maximum injection time (maxIT) of 50 ms. After peptide fragmentation using higher-energy collision-induced dissociation (HCD), MS2 spectra of up to 20 precursor peptides were acquired at a resolution of 15000 with automatic gain control (AGC) target value of 5e4 and maximum injection time (maxIT) of 22 ms. The precursor isolation window width was set to 1.3 m/z and normalized collision energy to 30 %. Dynamic exclusion was enabled with 20 s exclusion time (mass tolerance ±10 ppm). MS/MS spectra were singly-charged, unassigned, or with charge states >6+ were excluded. Raw data processing and statistical analysis Peptide identification and quantification were performed using MaxQuant version 1.6.3.4.[298] and the reference proteomes *E.coli* BL21(DE3) (NCBI: 469008), phage CLB-P3 (de-novo assembled: P3.prot.fa) and phage T7 (Uniprot: UP000000840). A contaminant list

(a built-in option in MaxQuant) was also added. For all MaxQuant searches, typical default parameters were employed. Those included carbamidomethylation of cysteine as fixed modification and oxidation of methionine and N-terminal protein acetylation as variable modifications. Trypsin/P was specified as the proteolytic enzyme. The precursor tolerance was set to 4.5 ppm, and fragment ion tolerance to 20 ppm. Results were adjusted to a 1 % false discovery rate (FDR) on the peptide spectrum match (PSM) level employing a target-decoy approach using reversed protein sequences. The minimal peptide length was defined as 7 amino acids. Intensity-based absolute quantification (iBAQ) [299] was enabled. The iBAQ intensities represent an estimate of the absolute concentration (rank) of a given protein in a sample and were the applied quantification unit throughout this study.

**Heatmaps**

To examine how the phage proteins are differentially expressed in the cell-free system and phage stock, triplicates of each sample were measured using tandem-MS. All phage proteins were selected and further analyzed. The phage protein (RPOL BPT7) cannot be distinguished from *E.coli* protein (A0A140NCE7 ECOBD) and was included in the analysis as well. The logarithm transformed (base 10) iBAQ intensity was used to represent the protein abundance, missing values were replaced with the lowest measured iBAQ intensity. The Euclidean distance together with the Ward linkage method (no scaling) was used in the heatmap (Fig. 26b). Time-resolved heatmap and single protein analysis Three replicates were measured over the 12 time-points. The logarithm transformed (base 10) iBAQ intensity was used to represent the protein abundance, missing values were replaced with the constant value of 5.5, which is slightly lower than the lowest measured values. To remove the potential batch effect between the three replicates, the iBAQ intensities were z-score transformed within each replicate. Then the z-scored iBAQ values of all three replicates were visually checked using beeswarm plots (Fig. 27d, R package Beeswarms version 0.2.3). To cluster phage proteins according to their time-points of expression (Fig. 27A), the z-scored iBAQ values of the three replicates were combined using their mean values, which were further scaled to the range of 0 to 1. Manhattan distance and Ward's method were used to cluster the final transformed intensities. Three phage proteins (P03751GP73BPT7, P03778Y06BPT7, P03797Y38BPT7) were removed in this analysis because they did not show a clear

trend over the time course. All analyses were performed in R (version 3.6.1).

## 9.2    Supplementary Data

### 9.2.1    Supporting Information to Self-Assembled Active Plasmonic Waveguide with a Peptide-Based Thermomechanical Switch

**Fluorescence spectroscopy measurements**
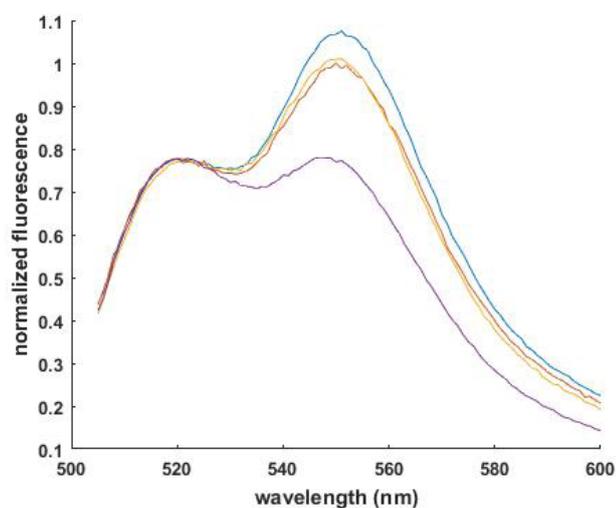


Supplementary Figure 1: Combined emission spectra of fluorescein (FAM) and Atto 532 for a sample containing only dyes (purple), single dye-functionalized AuNP (yellow), non-transmitting waveguides that lack the central AuNP (red) and for a fully assembled waveguide (blue). The spectra were normalized to the maximum FAM emission signal at $\lambda = 520$ nm.

**Computational results.** In calculations with MatLab-based program MNPBEM 13 the maximum absorption of a chain of five AuNP (single particle diameter 10 nm) in water was determined at 547 nm, whereas for the situation of a central defect in such a chain the calculated maximum absorption is at 542 nm. From these calculations no substantial spectral shifts are expected for the transmitting waveguide and the sample without the central AuNP.
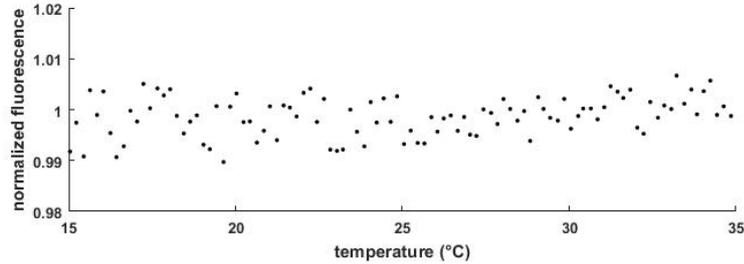
Supplementary Figure 2: Fluorescence spectroscopy data of a typical signal of the FAM dye at wavelength of 520 nm at different temperatures. The sample contains the fully assembled waveguide, with the dyes attached in FB. The curve was normalized using a blank signal, which contains dye functionalized AuNP.

**Other influences on LSPR transmission.** Potentially, LSPR transmission through the waveguide is not only influenced by the position of the central AuNP, but also by the local permittivity of the AuNP, which changes during the ELP transition. Whereas LSPR coupling changes exponentially with distance, the influence of the permittivity only scales with the inverse square-root,[300] and is thus expected to merely have a moderate effect.
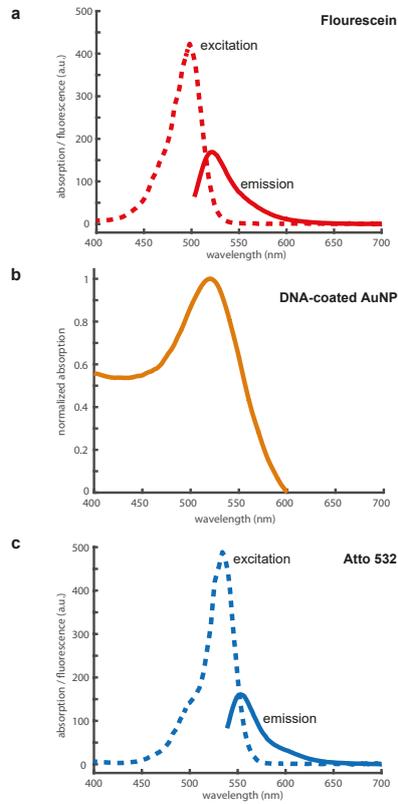
Supplementary Figure 3: Extinction/absorbance and fluorescence spectra. (a) Excitation (dashed) and emission (solid) spectra of fluorescein. (b) Absorption of single DNA-coated AuNP. (c) Excitation (dashed) and emission spectra of Atto 532 (solid).
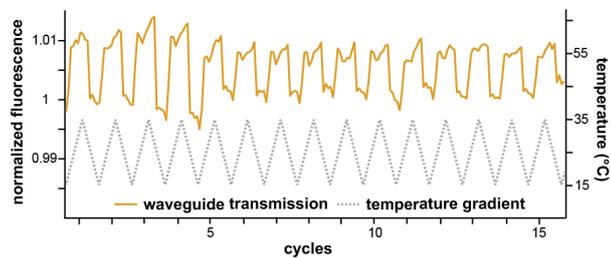


Supplementary Figure 4: Fluorescence signal of a waveguide in FB with extra 300 mM NaCl (orange) that is repeatedly switched between its two states by temperature cycling between 15 °C and 35 °C (grey). The curve was normalized using a blank signal, which contains dye-functionalized AuNP.

**Temperature transition of bulk ELP and L4-V$_{40}$C-AuNP**

To determine the transition temperature of the L4-V$_{40}$C conjugates attached to the AuNP the absorption spectra between 500 nm and 600 nm of the ELP-AuNP solution was investigated at different temperatures. When ELPs collapse the permittivity around the AuNP changes and therefore the maximal peak of the absorbance spectrum shifts to longer wavelengths. The inclination point of the transition curve is used as the transition temperature. T$_t$ of 10 nM L4-V$_{40}$0C-AuNP in FB was measured by changing the temperature from 15 °C to 50 °C by absorbance spectrometry. (J-815 CD-Spectrometer from JASCO Germany GmbH.)
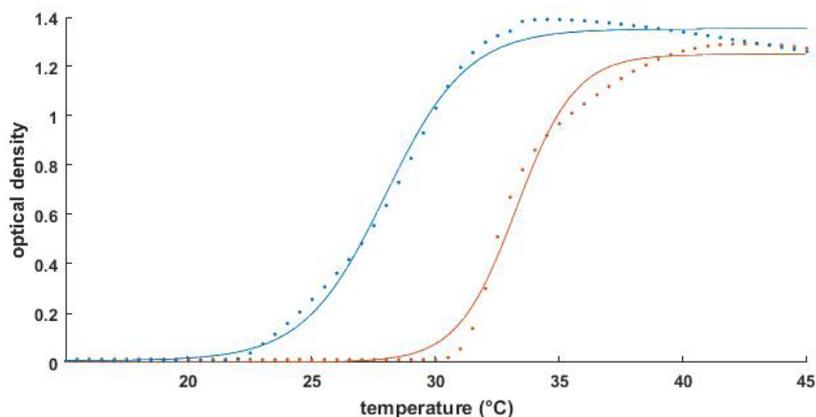


Supplementary Figure 5: Transition temperature of the ELP V$_{40}$C. To determine the transition temperature of the V$_{40}$C (75 $\mu$M) the optical density of the solution was measured at 350 nm. As buffered solution we used FB (red dots) and FB with 300 mM NaCl (blue dots). In order to prevent the formation of disulfide bridges TCEP was added at a final concentration of 10 mM. The transition temperatures were determined from the inflection points of sigmodal fits of the data (solid lines) resulting in T$_t$ = 33.3 °C (no NaCl, in red) and 28.0 °C (with 300 mM NaCl, in blue).
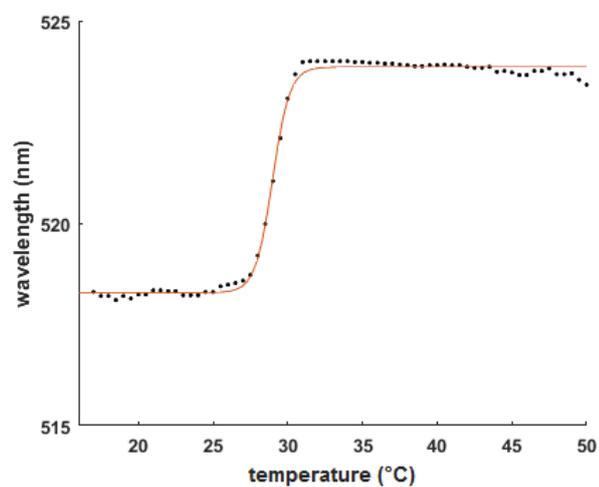
Supplementary Figure 6: Wavelengths of maximum absorbance for 10 nM L4-V$_{40}$C-AuNP in FB and extra 300 mM NaCl versus temperature (black dots). From the sigmoidal fit (red solid line) the transition temperature is determined to be 24.1 °C.
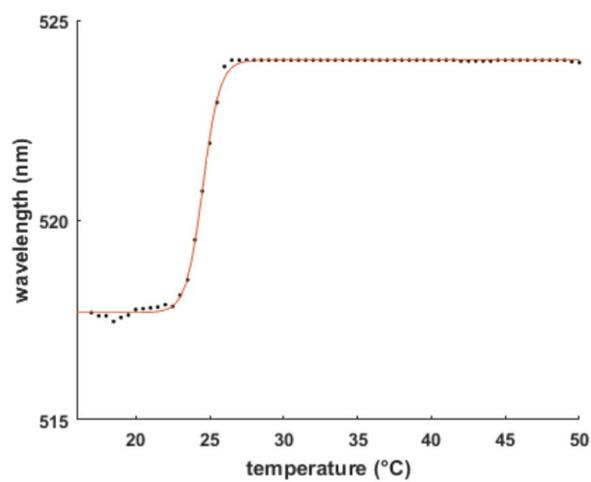


Supplementary Figure 7: Wavelengths of maximum absorbance for 10 nM L4-V$_{40}$C-AuNP in FB versus temperature (black dots). From the sigmoidal fit (red solid line) the transition temperature is determined to be 28.5 °C.

## Gold nanoparticle assembly using the silica protocol

### Adsorption and desorption of DNA origami structures on silica particles

DNA binds to negatively charged silica surface via salt bridges using divalent $Mg^{2+}$ ions. In order to find the right salt conditions a screening of the $MgCl^{2+}$ concentration was performed. Various samples containing 5.7 nM cross-shaped origami structures in TAE with various $MgCl^{2+}$ concentrations between 12.5 mM and 95 mM were prepared. After the adsorption onto the silica particles (see supplementary Figure 8 and Figure 11) the concentration of DNA in each supernatant (denoted as supernatant 1) was determined by UV/VIS spectrometry (NanoDrop 8000, Thermo Scientific, USA). As extinction coefficient for the DNA structures we used $1.12 \times 10^8$ Mol-1 cm-1. After elution from the silica particles the DNA concentration in each supernatant (denoted as supernatant 2) was determined. The result is shown in supplementary Figure 9. Supernatant 1 did not contain DNA structures when a $Mg^{2+}$ concentration below approx. 65 mM was used, which is apparently the limit for adsorption onto the utilized silica particles. Above 65 mM a large amount of DNA structures was found in supernatant 1. From the analysis of supernatant 2 after the elution with TAE it can be seen that in the range between 50 mM and 65 mM for the adsorption of DNA structures is best and the starting concentration of 5.7 nM is almost recovered (see supplementary Figure 9). To visualize the adsorbed DNA origami structures with attached AuNP we used SEM (see supplementary Figure 8). SEM imaging was performed as described in the Methods section.
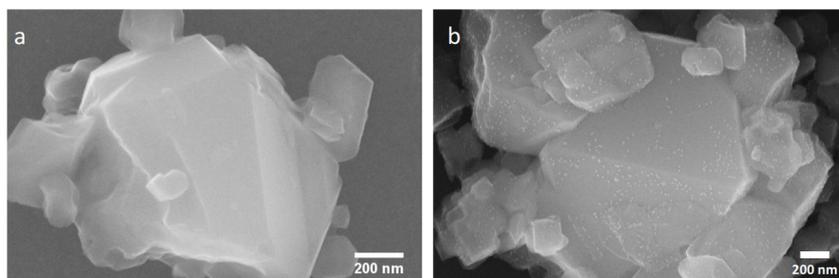


Supplementary Figure 8: SEM images. (a) Bare silica particle before functionalization. (b) Silica particle with attached DNA origami structures (not visible) with bound AuNP.
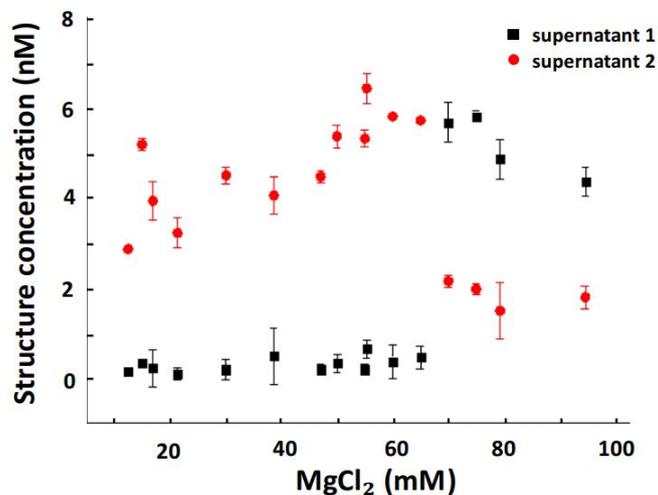
Supplementary Figure 9: Influence of $Mg^{2+}$ concentration. After adsorption supernatant 1 (black squares) did not contain DNA structures below 65 mM $MgCl_2$. Supernatant 2 (red circles) contained approx. the initial concentration of DNA structures when $MgCl^{2+}$ in the range between 50 mM and 65 mM was used.

**Analysis of the binding yield from AFM images.**

The binding yield achieved with the silica protocol was determined by counting cross-shaped structures in AFM images (supplementary Figure 10). We also determined the influence of used excess of AuNP and the incubation time (the adsorption time) on the binding yield. AFM imaging was performed as described in the Methods section. Some structures were not taken into consideration due to structure misfolding or only partial visibility at the edges of the images. The found DNA structures were categorized into five groups, 1. structures with three bound AuNP (denoted as $CRO_{3AuNP}$), 2. structures with two AuNP ($CRO_{2AuNP}$), 3. structures with one AuNP($CRO_{1AuNP}$), 4. empty structures ($CRO_{0AuNP}$) and 5. dismissed structures. Supplementary Figure 10 shows an AFM image with marked structures. The total binding yield $p_{binding}$ of a single AuNP to a designated binding position was calculated using

$$p_{binding} = \frac{3 \cdot CRO_{3AuNP} + 2 \cdot CRO_{2AuNP} + 1 \cdot CRO_{1AuNP}}{3 \cdot \sum_i CRO_{iAuNP}} \tag{15}$$

Additionally, we calculated the binomial distribution from the total binding yield with its standard deviation

$$\sigma = \sqrt{N p_{binding}(1 - p_{binding})} \qquad (16)$$

with $N = 3 \cdot \sum_i CRO_{iAuNP}$. Supplementary Figure 11 shows the distributions of the number of bound AuNP on one DNA structure for 1-, 2- and 3-fold excess of AuNP relative to the number of used binding positions (red bars with error bars indicating the calculated standard deviation). The black dots display the binomial distribution. For 1-fold and 2-fold excess about 92 % of the structures were fully equipped and for 3-fold 99.1 % were fully equipped. We also varied the incubation time for the attachment of the AuNP. We used incubation times of 2, 5, 10, 30 and 60 minutes for 1-fold excess and 2, 5 and 60 minutes for 3-fold excess. An overview of the counts is shown in Tables 1 and 2.
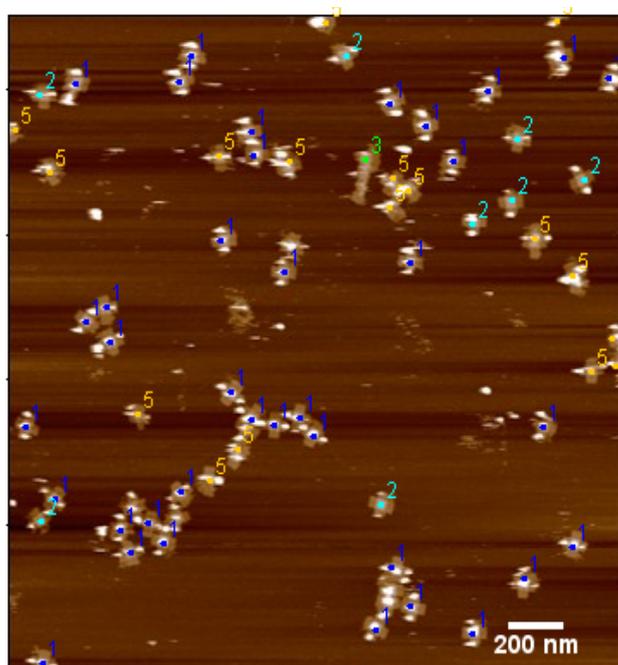


Supplementary Figure 10: Analysis of AFM images. Exemplary AFM image of AuNP-functionalized crossshaped DNA origami structures. Here, a 1-fold excess of AuNP was used.

| Time (min) | 2 | 5 | 10 | 30 | 60 |
|---|---|---|---|---|---|
| 0 occupied | 0 | 0 | 1 | 0 | 0 |
| 1 occupied | 0 | 2 | 1 | 3 | 7 |
| 2 occupied | 13 | 14 | 22 | 33 | 31 |
| 3 occupied | 157 | 86 | 174 | 126 | 82 |
| Total number of structures | 170 | 102 | 198 | 162 | 120 |

Supplementary Table 1: Binding statistics for Figure 11. Number of occupied binding sites (left column) per cross-shaped DNA origami structure for different incubation times and a 1:1 ratio of binding positions and AuNP.

| Time (min) | 2 | 5 | 60 |
|---|---|---|---|
| 0 occupied | 0 | 0 | 0 |
| 1 occupied | 0 | 0 | 3 |
| 2 occupied | 2 | 10 | 25 |
| 3 occupied | 223 | 212 | 57 |
| Total number of structures | 225 | 131 | 85 |

Supplementary Table 2: Binding statistics for Figure 11. Number of occupied binding sites (left column) per cross-shaped DNA origami structure for different incubation times and a 1:3 ratio of binding positions and AuNP.
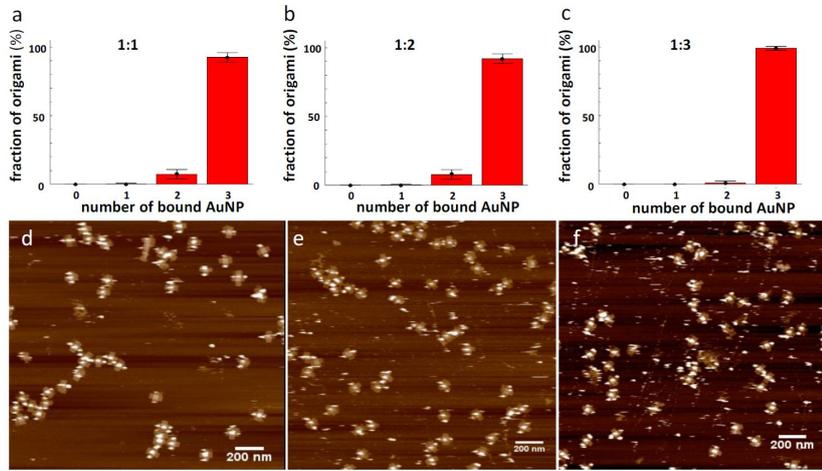
Supplementary Figure 11: Binding yield distribution. Occupied binding positions per structure in percent with (a) 1-fold excess of AuNP relative to the number of binding positions (b) 2-fold excess and (c) 3- fold excess. d-f, corresponding representative AFM images.

**AuNP binding yield of the waveguide**

To determine the binding yield of the AuNP on the waveguide structure several TEM images were analyzed. Here TEM images were used because the resolution of the AuNP in close proximity is better at the TEM. Imaging was performed as described in the methods section. Typical images used for analysis are shown in supplementary Figures 12 and 13. In both images fully functionalized waveguides are shown in the top row and waveguides with a single missing AuNP (the second main fraction) are shown in the bottom row. In supplementary Figure 12 the results for the waveguide without ELP switch are shown, whereas in supplementary Figure 13 the results for the waveguide including ELP switch are presented. In the analysis some structures were not taken into consideration due to structure misfolding and partial visibility at the edges of the images. The imaged twist-corrected rectangular origami structures (abbreviated as tcRRO) were categorized into seven groups, 1. structures with five bound AuNP (denoted as $tcRRO_{5AuNP}$), 2. with four AuNP ($tcRRO_{4AuNP}$), 3. with three AuNP ($tcRRO_{3AuNP}$), 4. with two AuNP ($tcRRO_{2AuNP}$), 5. with one AuNP ($tcRRO_{1AuNP}$), 6. empty structures ($trRRO_{0AuNP}$) and 7. dismissed structures.
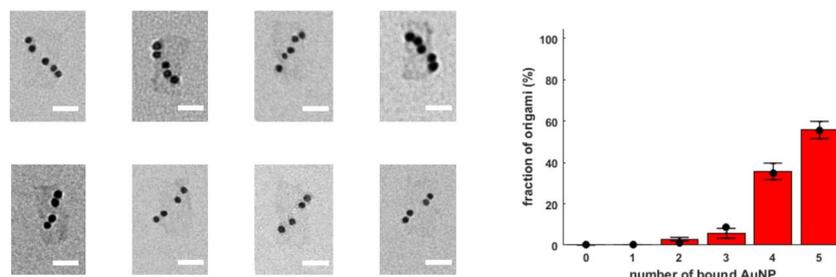
Supplementary Figure 12: Analysis of the waveguide without ELP switch. Left, TEM images of waveguides. The top row shows typical TEM images of the fully assembled waveguide and in the bottom row typical images of incompletely assembled waveguides. Scale bars, 50 nm. Right, corresponding binding yield distribution, with occupied binding positions per structure. The given uncertainties are calculated standard deviations. Black dots indicate the theoretical bionomial distribution.
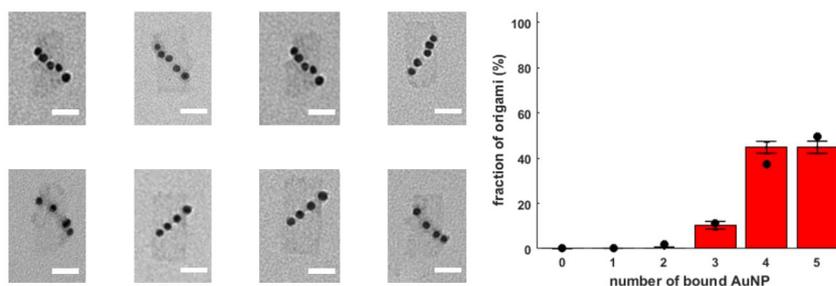


Supplementary Figure 13: Analysis of switchable waveguides. Left, TEM images of the waveguide with the ELP-AuNP as the central AuNP. The top row shows typical TEM images of the fully assembled waveguide and in the bottom row typical images of the incompletely assembled waveguide are shown. Scale bars, 50 nm. Right, corresponding binding yield distribution, with occupied binding positions per structure. The given uncertainties are calculated standard deviations. Black dots indicate the theoretical bionomial distribution.

| waveguide species | waveguide without ELP | waveguide with ELP |
|---|---|---|
| 0 occupied | 0 | 0 |
| 1 occupied | 0 | 0 |
| 2 occupied | 2 | 0 |
| 3 occupied | 4 | 6 |
| 4 occupied | 25 | 26 |
| 5 occupied | 39 | 26 |
| Total number of structures | 70 | 59 |

Supplementary Table 3: Binding statistics. Number of occupied binding sites (left column) per waveguide with and without attached ELP.

**Statistics of gap size for waveguides:** The average gap size was determined by measuring the edge-to-edge distance between two AuNPs. A total amount of 72 gaps were analyzed and resulted in 2.1 nm with a standard deviation of 1.5 nm.

**DNA origami Design and sequences**
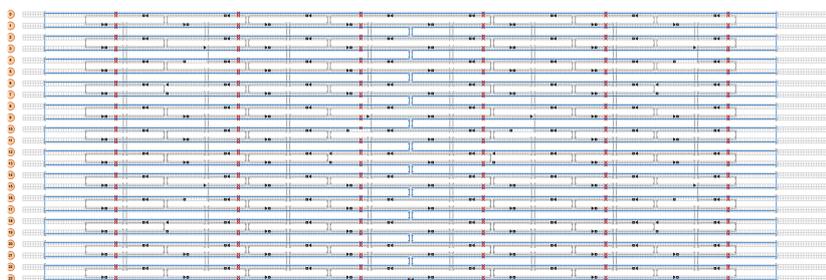


Supplementary Figure 14: Design 1: Twist corrected rectangular origami structure (tcRRO).
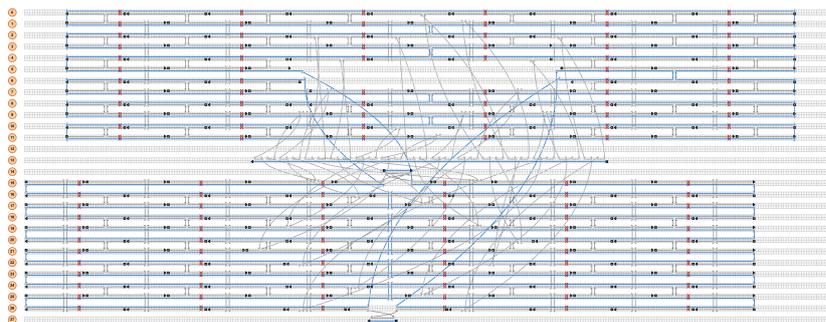
Supplementary Figure 15: Design 2: Twist corrected cross-shaped origami structure (CRO).

## 9.2.2 Supporting Information to Towards Synthetic Cells Using Peptide-Based Reaction Compartments
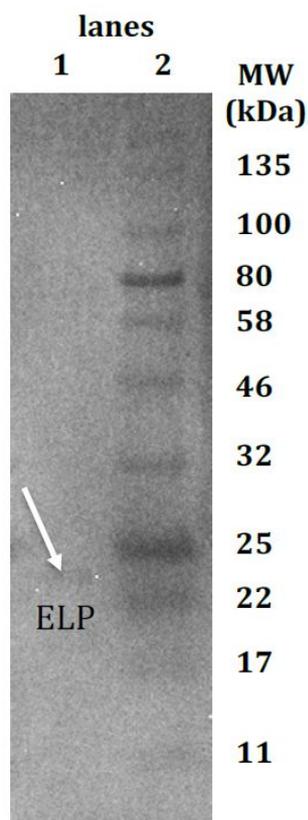


Supplementary Figure 16: SDS gel of purified EF. Lane 1 shows the purified EF peptide. Lane 2 is the Color Prestained Protein Standard ladder, Broad Range (11–245 kDa) from NEB. EF has a mass of about 18 kDa, but the peptide band is shifted upwards to about 23 kDa, which is a well-known effect for elastin-like peptides.
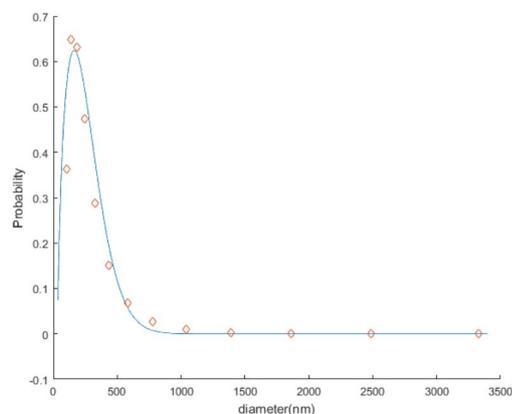
Supplementary Figure 17: Appearance probability of the vesicles measured with DLS. The vesicles were produced using the glass beads method, 180 pM EF and 1x PBS as the swelling solution. Original DLS data from a binned histogram (red diamonds) is shown with a peak value of 178 nm and a dispersion of 67 nm. The blue solid curve is a Weibull probability distribution fit.
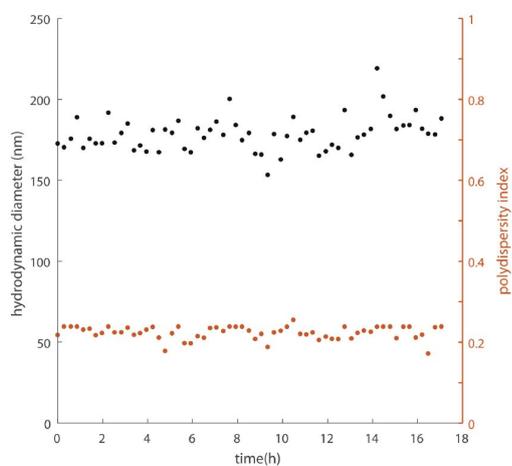


Supplementary Figure 18: Size stability in time. Measurement of the hydrodynamic diameter (black dots) and the polydispersity index (red dots) of the vesicles for 18 h using DLS. The vesicles were produced using the glass beads method and 1x PBS as the swelling solution. The mean of the diameter is 179 nm with a standard deviation of 10 nm.
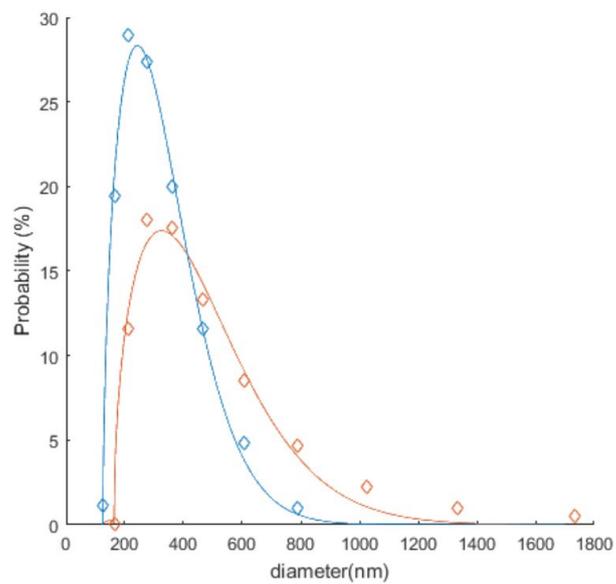
Supplementary Figure 19: Vesicle growth through the addition of EF from the outside. Appearance probability of the vesicles measured with DLS. The vesicles were produced using the glass beads method and 1x PBS as the swelling solution. The blue diamonds are original DLS data from a binned histogram. The solid blue curve is a Weibull probability distribution fit. The peak value of the diameter was found to be 192 nm with a dispersion of 74 nm. After the addition of 50 $\mu$M EF the peak value of the binned histogram (red diamonds) changed to 234 nm with a dispersion of 101 nm.
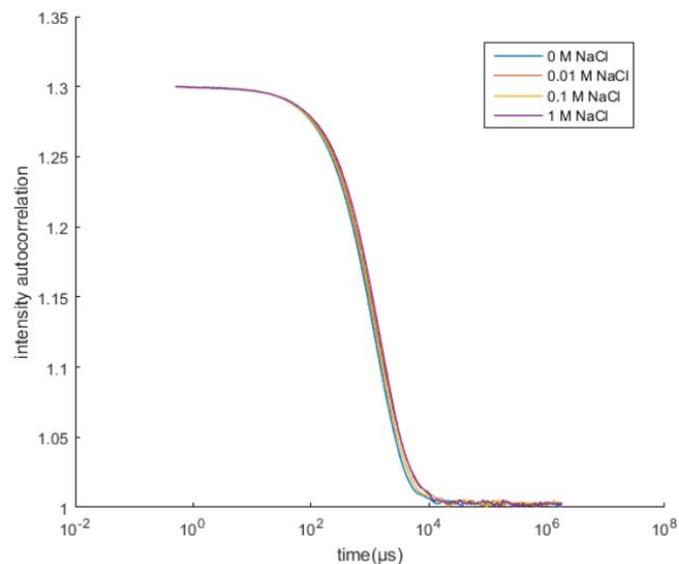
Supplementary Figure 20: Membrane stability in NaCl gradients. After vesicle formation using the glass beads method and 1x PBS we tested the vesicles size stability for NaCl concentration gradients between outside and inside. Therefore, we added NaCl to outside solution to get NaCl concentration differences of 0 M, 0.01 M, 0.1 M and 1 M. The correlation functions from DLS measurements show no variations for these NaCl concentrations. In TEM measurements we see spherical vesicles for all gradients except 1 M NaCl with some vesicles showing ellipsoid shapes.
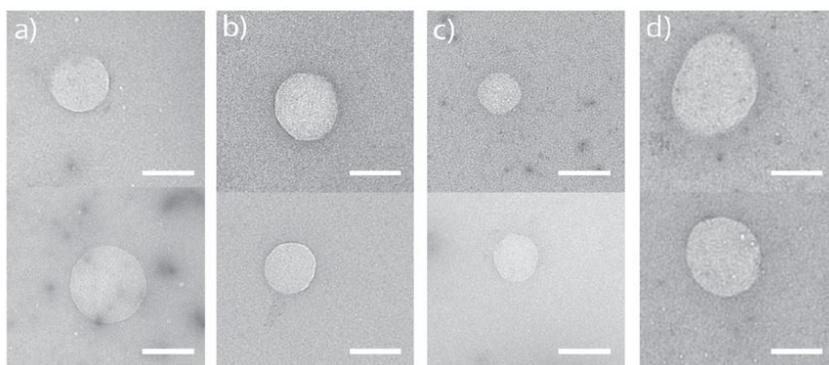


Supplementary Figure 21: Shape of EF vesicles at various NaCl membrane gradients. The concentration difference of NaCl between inside and outside the vesicles are: 0 M (blue), 0.01 M (red), 0.1 M (yellow) and 1 M (purple). Scale bar: 200 nm.
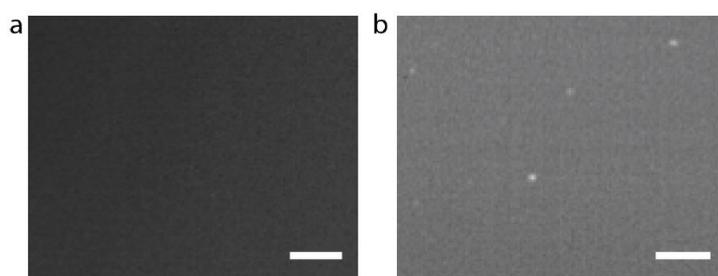
Supplementary Figure 22: Light microscopy of vesicles with encapsulated fluorophore-labeled DNA. As described in the main part EF vesicles were produced with encapsulated fluorophore-labeled DNA with 1x PBS. (a) Brightfield image of EF vesicles. Due to their size of about 200 nm and the diffraction limit there are no vesicles visible. (b) Fluorescence microscopy image of the same vesicles. Scale bar: 10 $\mu$m.
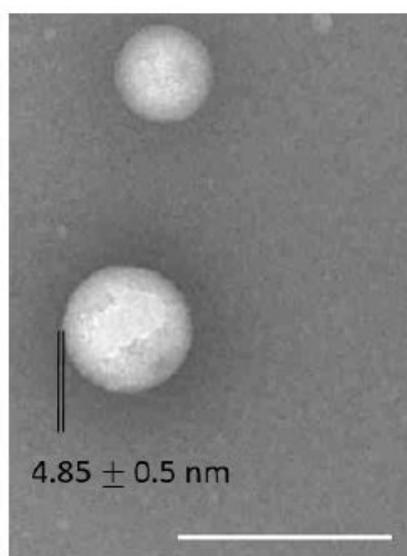


Supplementary Figure 23: Determination of membrane thickness. TEM image of EF vesicles used for the determination of the membrane thickness. The glass beads were rehydrated with 1x PBS. The thickness of the membrane was calculated for these vesicles to 4.9 nm with a standard deviation of 0.5 nm. Scale bar: 200 nm.
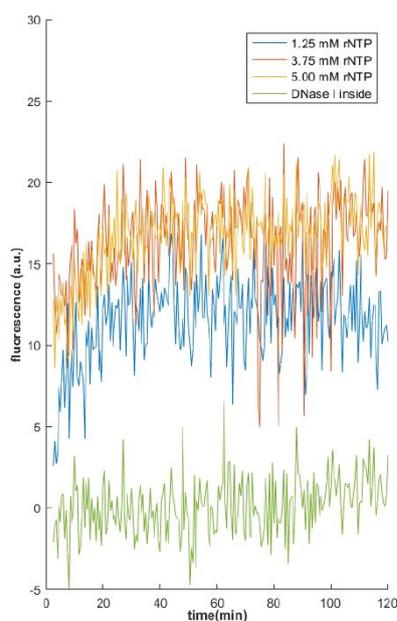
Supplementary Figure 24: TX in vesicles: transcription of the dBroccoli aptamer. Fluorescence time traces during the transcription of the dBroccoli aptamer inside the EF vesicles. Here, only transcription mixture TX was encapsulated using the glass beads method. After vesicle formation DNase I was added to prevent transcription and hence fluorescence at the outside. Excitation was carried out at 480 nm and the emission was measured at 520 nm. Three different rNTP concentrations were used, namely 1.25 mM (blue), 3.75 mM (red) and 5.00 mM (yellow). As a negative control DNase I was also incorporated into the vesicles with 1.25 mM rNTP (green). For signal correction a sample containing vesicles with transcription mixture TX but without DNA template was used.
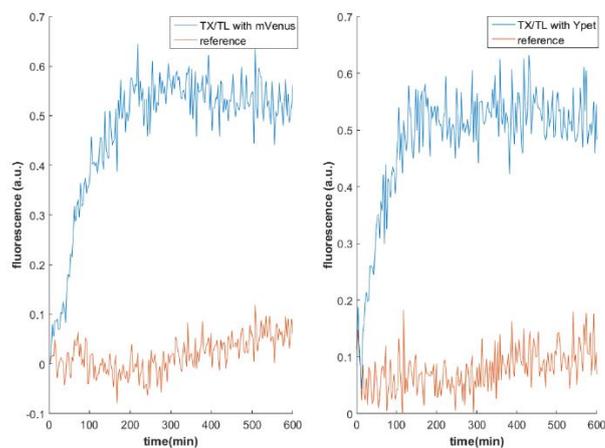
Supplementary Figure 25: *In vesiculo* translation vs. time. Fluorescence signals (excitation at 500 nm and emission at 520 nm) of two fluorescent proteins (blue curves), namely mVenus (left) and Ypet (right), which are expressed inside of EF vesicles utilizing TX-TL. For mVenus expression a T7 promoter was used and for Ypet expression a constitutive promoter from *E.coli* was used. To suppress expression at the outside an antibiotic was added to the solution surrounding the vesicles. For a reference the antibiotic was also encapsulated inside the vesicles and successfully suppressed protein expression (red curves). Both signals (blue and red) were background corrected using a sample containing EF vesicles with TX-TL, but no plasmids.
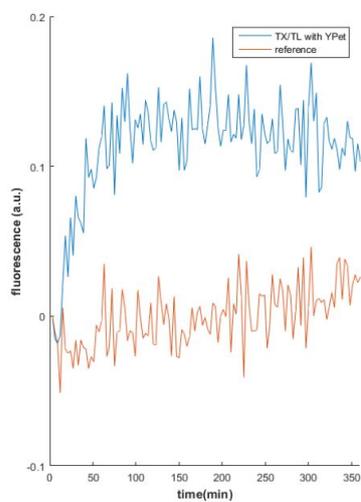
Supplementary Figure 26: YPet expression inside vesicles. Fluorescence signals (excitation at 500 nm and emission at 520 nm) of YPet (blue curves), namely, which are expressed inside of EF vesicles utilizing TX-TL. For Ypet expression a constitutive promoter from *E.coli* was used. To suppress expression at the outside EDTA was added to the solution surrounding the vesicles. For a reference EDTA was also encapsulated inside the vesicles and successfully suppressed protein expression (red curves). Both signals (blue and red) were background corrected using a sample containing EF vesicles with TX-TL, but no plasmids.
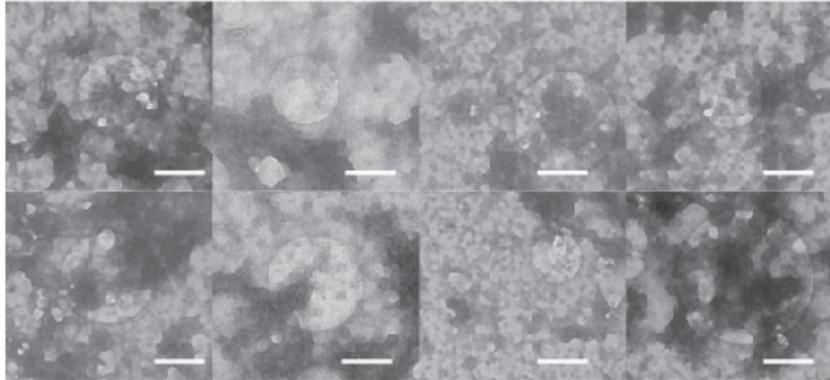
Supplementary Figure 27: EF expression in vesicles using TX-TL. TEM images of the EF vesicles after 0 min incubation time. The TX-TL system and the plasmid encoding EF are encapsulated with the glass beads method. After vesicle formation an antibiotic is added to the outer solution. The samples were not purified from the TX-TL (see main text). All shown subfigures are from the same TEM grid and sample. Scale bars: 200 nm.
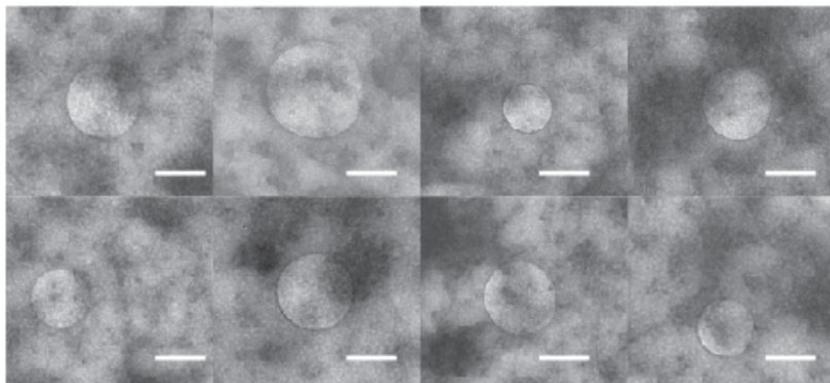


Supplementary Figure 28: EF expression in vesicles using TX-TL. TEM images of the EF vesicles after 240 min incubation at 29 °C. The TX-TL system and the plasmid encoding EF are encapsulated with the glass beads method. After vesicle formation an antibiotic is added to the outer solution. The samples were not purified from the TX-TL (see main text). All shown subfigures are from the same TEM grid and sample. Scale bars: 200 nm.
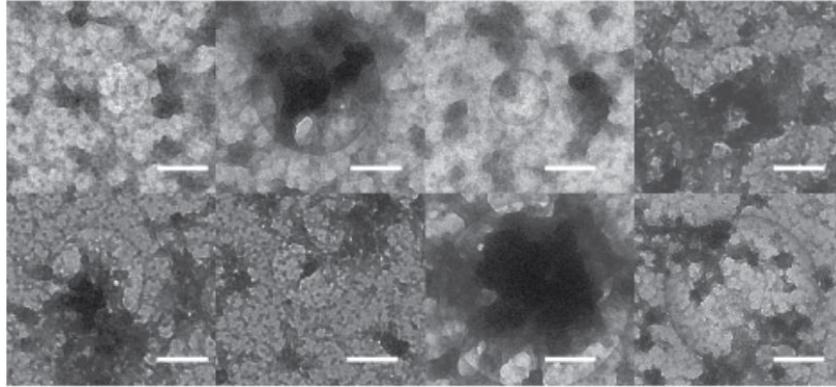
Supplementary Figure 29: V40 expression in vesicles using TX-TL. TEM images of the EF vesicles after 0 min. The TX-TL system and the plasmid encoding V40 are encapsulated with the glass beads method. After vesicle formation an antibiotic is added to the outer solution. The samples were not purified from the TX-TL. All shown subfigures are from the same TEM grid and sample. Scale bars: 200 nm.
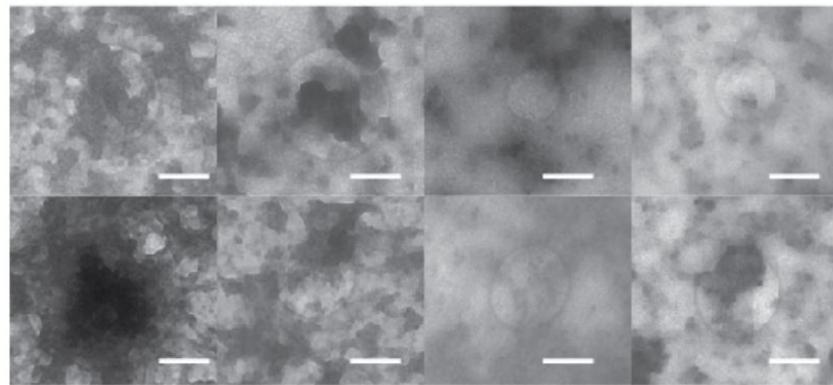


Supplementary Figure 30: V40 expression in vesicles using TX-TL. TEM images of the EF vesicles after 240 min incubation at 29 °C. The TX-TL system and the plasmid encoding V40 are encapsulated with the glass beads method. After vesicle formation an antibiotic is added to the outer solution. The samples were not purified from the TX-TL. All shown subfigures are from the same TEM grid and sample. Scale bars: 200 nm.
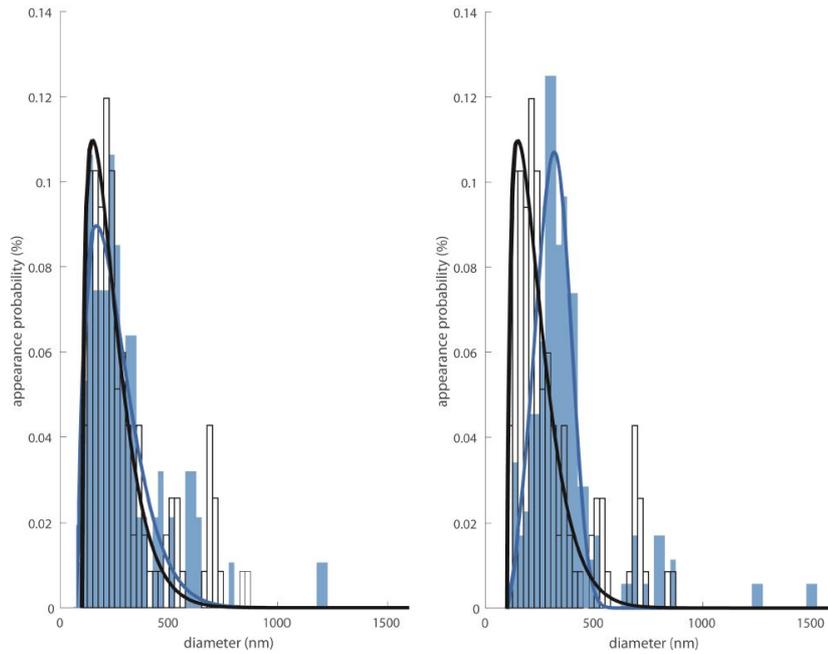
Supplementary Figure 31: Vesicle size. Size distributions of the vesicles measured by TEM, from about 100 vesicles. Left, Vesicles containing TX-TL and the V40 plasmid. The black framed histogram shows the size distribution at t = 0 min. The solid black line is a Weibull function fitted to the histogram resulting in a peak value of 149 nm. The blue histogram is the size distribution at t = 240 min. The solid blue line is a Weibull function fitted to the histogram resulting in a peak value of 145 nm. Right, Vesicles containing TX-TL and the EF plasmid. The black framed histogram shows the size distribution at t = 0 min. The solid black line is a Weibull function fitted to the histogram resulting in a peak value of 157 nm. The blue histogram is the size distribution at t = 240 min. The solid blue line is a Weibull function fitted to the histogram resulting in a peak value of 330 nm. For the statistical analysis about 100 vesicles were evaluated.

138
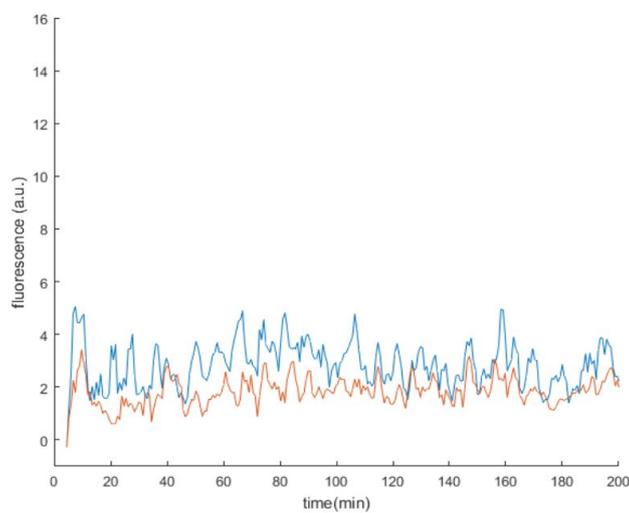
Supplementary Figure 32: Negative control of the FRET assay. Time trace of the fluorescence signal of the donor dye Cy3 (red) and the acceptor dye Cy5 (blue), which are attached to EF via copper-based azidealkyne Huisgen cycloaddition. The TX-TL mixture, the EF plasmid and kanamycin were encapsulated in vesicles made of Cy5-EF and Cy3-EF using the glass beads method. Both signals are background corrected with a sample containing vesicles made of Cy5-EF and Cy3-EF, but without TX-TL and plasmid.
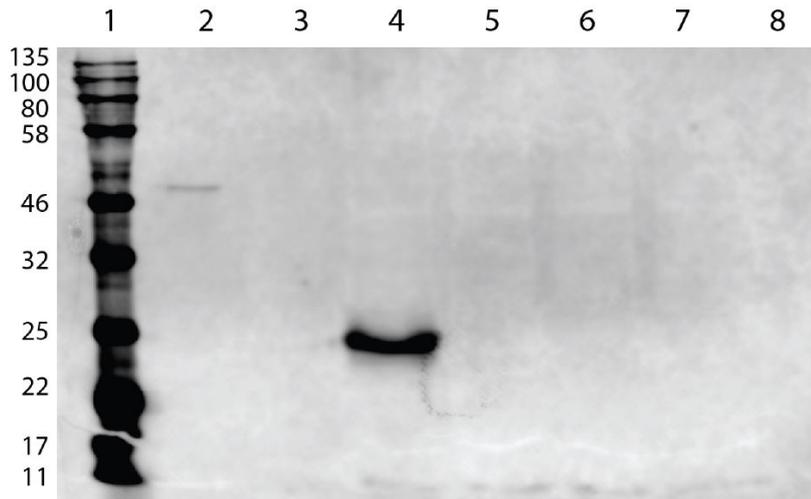
Supplementary Figure 33: Western blot of the peptides EF-His (6x histidin) expressed inside the vesicles. The first lane (from left) is the Color Prestained Protein Standard ladder, Broad Range (11–245 kDa) from NEB. The second lane is a positive control with a 67 kDa histidine-tagged helicase. For lane three a sample with vesicles containing TX-TL and the EF-His plasmid without incubation was used. For lane four a sample with vesicles containing TX-TL and the EF-His after an incubation time of 240 min at 29 °C was used. The sample for lane five is the same as for lane three, but with kanamycin inside the vesicles. The sample for lane six is the same as for lane four, but also with kanamycin inside the vesicles. For lane seven only TX-TL without any ELP was used. For lane eight a sample with peptide vesicles in 1x PBS was used. The used antibodies were purchased from Life Technologies GmbH, Darmstadt, Germany. Primary antibody: anti-His antibody (6x-His Epitope tag antibody ) (mouse, Lot: SI252938, Catalogue: MA1135). Secondary antibody: anti-mouse Alexa Fluor 680 (goat, Lot: RG233737A, Catalogue: A28183).
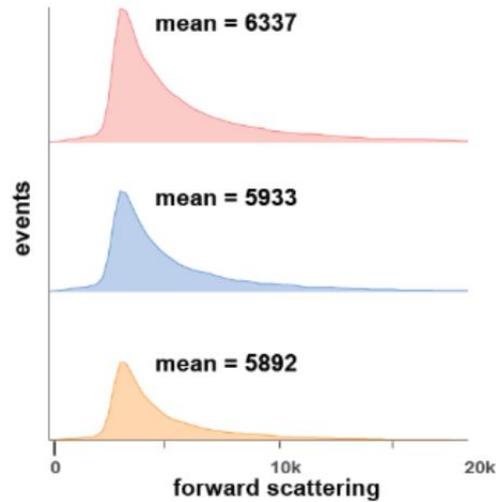
Supplementary Figure 34: Forward scattering of the EF vesicles using flow cytometry. Top, The red histogram shows the scattering of vesicles containing TX-TL and the EF plasmid for peptide expression. The sample was incubated for eight hours at 29 °C. To suppress expression outside the vesicles the antibiotic kanamycin was added to the external solution. The mean of the forward scattering was measured to be 6337, which indicates vesicle growth. Center, The blue histogram shows the scattering of vesicles containing TXTL and the V40 plasmid. It was also incubated for eight hours at 29 °C. Bottom, The yellow histogram shows the scattering of vesicles containing TX-TL, the EF plasmid and the antibiotic kanamycin to suppress peptide expression.
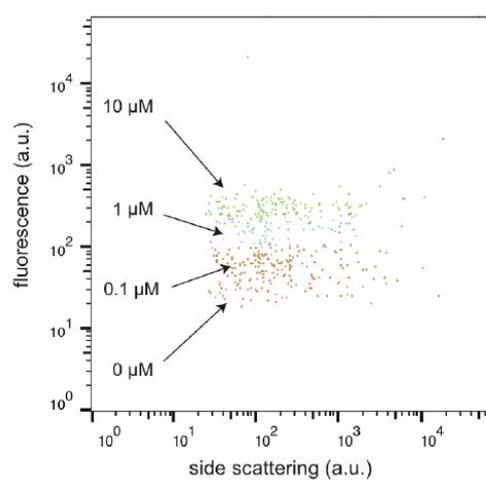
Supplementary Figure 35: Flow Cytometry reference. Flow cytometry measurement of the FAM-labeled DNA without EF with starting concentrations of 10 $\mu$M (green dots, 1.7 $\mu$M after dilution for the measurement), 1 $\mu$M (blue dots, 0.17 $\mu$M after dilution), 0.1 $\mu$M (orange dots, 0.02 $\mu$M after dilution) and the control without dyes (red dots).

**A**



**B**



rt= 21.15 -21.66 min

**C**
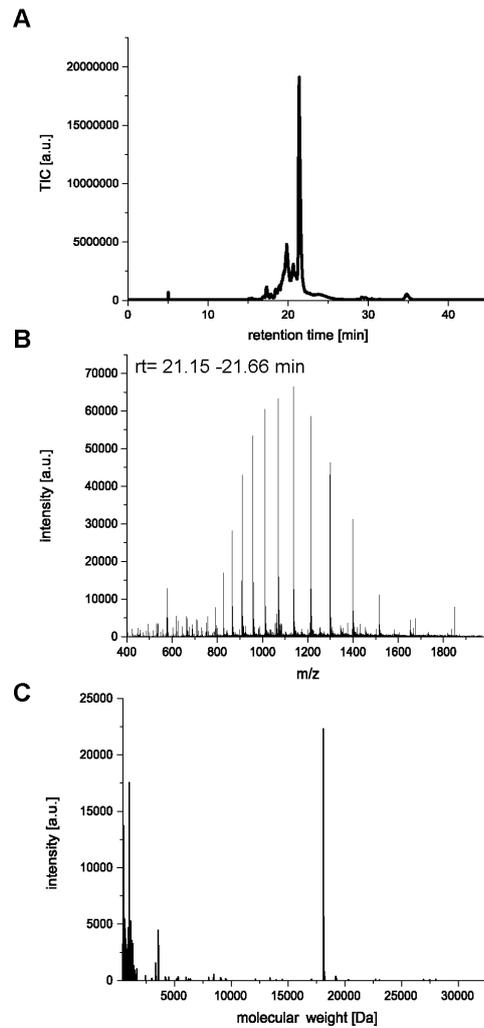


Supplementary Figure 36: Intact-protein mass spectrometry of EF. (a) Total ion current chromatogram. (b) Summed-up raw spectra of retention time 21.15 to 21.66 min. (c) Deconvoluted spectra of retention time 21.15 to 21.66 min. EF is the only prominent species.

**Weibull distribution**

To obtain parameters of the size distributions a Weibull distribution function was used with the following probability density function

$$f(d) = \frac{\delta}{\eta} \cdot \left(\frac{d - d_0}{\eta}\right)^{\delta - 1} \cdot \exp\left[-\left(\frac{d - d_0}{\eta}\right)^{\delta}\right] \tag{17}$$

with $d$ as the diameter of the vesicles and $\eta$, $\delta$ and $d_0$, as free parameters. The peak value where the distribution of the diameter reaches its maximum can be obtained with the equation:

$$d_{max} = d_0 + \eta \left(\frac{\delta - 1}{\delta}\right)^{\frac{1}{\delta}} \tag{18}$$

The expected value $E$ can be calculated using

$$E = d_0 + \eta \cdot \Gamma\left(\frac{1}{\delta} + 1\right) \tag{19}$$

and the dispersion (standard deviation) $D$ of the Weibull distribution is calculated through

$$D^2 = \eta^2 \left(\Gamma\left(\frac{2}{\delta} + 1\right) - \Gamma^2\left(\frac{1}{\delta} + 1\right)\right) \tag{20}$$

| EF conc. (pM) | dispersion D(nm) | expected value E(nm) | peak value $d_{max}$ (nm) |
|---|---|---|---|
| 110 | 47.0 | 115.5 | 86.8 |
| 180 | 67.0 | 207.3 | 177.9 |
| 220 | 110.5 | 236.3 | 219.6 |
| 440 | 414.6 | 587.5 | 249.7 |

Supplementary Table 4: Vesicle sizes for various EF concentrations. We used several EF concentrations and tested the concentration influence on the size of the vesicles using the glass beads method and 1x PBS as swelling solution. Supplementary Table 4 shows the different hydrodynamic radii determined from the Weibull fit of the vesicle size distribution, which was obtained by DLS measurements of different EF concentrations. For further experiments we used a concentration of 180 pM, since it worked best with TXTL.

| vesicle content | mean (nm) | Standard deviation (nm) |
|---|---|---|
| 1x PBS | 4.9 | 0.5 |
| t = 0 min TX-TL EF plasmid | 5.2 | 0.6 |
| t = 240 min TX-TL EF plasmid | 5.3 | 0.7 |
| t = 0 min TX-TL V40 plasmid | 5.6 | 0.8 |
| t = 240 min TX-TL V40 plasmid | 5.7 | 0.6 |

Supplementary Table 5: Membrane thickness of the EF vesicles. The membrane thickness of the EF vesicles was measured by analyzing 20 vesicles at a representative spot in the TEM images. For image analysis and processing the Java-based software ImageJ was used. The thickness was determined for five different samples.

| vesicle content | variance $D^2$ (nm$^2$) | dispersion D (nm) | expected value E (nm) | peak value $d_{max}$ (nm) | number of analyzed vesicles |
|---|---|---|---|---|---|
| 1x PBS | 4666 | 68 | 213 | 176 | 92 |
| t = 0 min TX-TL EF plasmid | 10866 | 104 | 239 | 157 | 171 |
| t = 240 min TX-TL EF plasmid | 6818 | 83 | 318 | 330 | 173 |
| t = 0 min TX-TL V40 plasmid | 10505 | 103 | 233 | 150 | 114 |
| t = 240 min TX-TL V40 plasmid | 16378 | 128 | 247 | 145 | 93 |

Supplementary Table 6: Weibull curve fitting parameters used in 31.

**Encapsulation efficiency of TX-TL components**

The low signal-to-noise ratio of the mVenus and YPet fluorescence signals in Fig. 18 resp. supplementary Fig. 25 can be rationalized by the small size of the vesicles resp. the encapsulation efficiency. We assume a Poisson distribution to derive the probability that at least one molecule of a kind is encapsulated inside of each vesicle. The probability that all of the N necessary components of the TX-TL system were encapsulated in a vesicle with volume V and radius R is therefore given by

$$P(R) = \prod_{n=1}^{N} [1 - e^{-C_n \cdot V(R)}] \tag{21}$$

where $C_n$ is the concentration (number density) of molecule n that is encapsulated. If we assume that the vesicles have a radius of R = 150 nm the concentration of the molecules needed to be in the low $\mu$M range to have an encapsulation probability close to 100 %, which is only valid for the proteins ($\mu$M) and the amino acids (mM). The plasmid, however, had a concentration of 50 nM and therefore had an encapsulation probability of only 35 %, resulting in a relatively low and noise fluorescence signal

**Number of peptides expressed inside the vesicles**

To estimate the relative number of peptides incorporated into the membrane, we used the following assumption:

$$\frac{N_{t=240min} \cdot V_{ELP}}{N_{t=0min} \cdot V_{ELP}} = \frac{V_{vesicleshellat t=240min}}{V_{vesicleshellat t=0min}} = \tag{22}$$

$$\frac{\frac{4}{3}\pi R^3_{t=240min} - \frac{4}{3}\pi(R_{t=240min} - a)^3}{\frac{4}{3}\pi R^3_{t=0min} - \frac{4}{3}\pi(R_{t=0min} - a)^3} = \tag{23}$$

$$\frac{R^3_{t=240min} - (R_{t=240min} - a)^3}{R^3_{t=0min} - (R_{t=0min} - a)^3} \tag{24}$$

where $N_t$ is the number of ELP incorporated into the membrane at time $t$ and the volume of a single peptide $V_{ELP}$. $R_t$ is the radius of the vesicles at time $t$ and $a$ is the membrane thickness. We further assumed that the volume of a single peptide stays constant and is independent of the vesicle radius, which is supported by the measurement of the membrane thickness, which indicate no swelling of the membrane:

$$\frac{N_{t=240min}}{N_{t=0min}} = \frac{3R^2_{t=240min}a - 3R_{t=240min}a^2 + a^3}{3R^2_{t=0min}a - 3R_{t=0min}a^2 + a^3} \tag{25}$$

This would result into a relative membrane volume increase of $\xi = 4.44$ for a membrane thickness $a$ of 5.33 nm, a radius of 157 nm at 0 min and a radius of 330 nm at 240 nm. If we assume that $a^n << R^n$ for every $n$ larger than 1 we can simplify the relative volume increase to:

$$\frac{N_{t=240min}}{N_{t=0min}} = \frac{R^2_{t=240min}}{R^2_{t=0min}} \tag{26}$$

For the same values of the radii we get for the relative membrane volume increase $\xi = 4.42$ which is close to the exact formula.

# References

[1] Gorski, A., Miedzybrodzki, R. & Borysowski, J. *Phage Therapy: A Practical Approach* (Springer, Cham, 2019).

[2] Vogele, K., Pirzer, T. & Simmel, F. C. Genetically encoded membranes for bottom–up biology. *ChemSystemsChem* **1** (2019).

[3] Goetzfried, M. A. *et al.* Periodic operation of a dynamic dna origami structure utilizing the hydrophilic-hydrophobic phase-transition of stimulus-sensitive polypeptides. *Small* **15**, e1903541 (2019).

[4] Vogele, K. *et al.* In vesiculo synthesis of peptide membrane precursors for autonomous vesicle growth. *Journal of visualized experiments : JoVE* (2019).

[5] Vogele, K., List, J., Simmel, F. C. & Pirzer, T. Enhanced efficiency of an enzyme cascade on dna-activated silica surfaces. *Langmuir* **34**, 14780–14786 (2018).

[6] Vogele, K. *et al.* Towards synthetic cells using peptide-based reaction compartments. *Nature Communications* **9**, 3862 (2018).

[7] Vogele, K. *et al.* Self-assembled active plasmonic waveguide with a peptide-based thermomechanical switch. *ACS Nano* **10**, 11377–11384 (2016).

[8] McKay, C. P. What is life—and how do we search for it in other worlds? *PLoS biology* **2**, e302 (2004).

[9] Xu, C., Hu, S. & Chen, X. Artificial cells: From basic science to applications. *Materials Today* **19**, 516–532 (2016).

[10] Zhuravlev, Y. N. & Avetisov, V. A. The definition of life in the context of its origin. *Biogeosciences* **3**, 281–291 (2006).

[11] Rasmussen, S. *Protocells: Bridging nonliving and living matter* (MIT Press, Cambridge, Mass., 2009).

[12] Koshland Jr., D. E. The seven pillars of life. *Science* **295**, 2215–2216 (2002).

[13] Benner, S. A. Defining life. *Astrobiology* **10**, 1021–1030 (2010).

[14] Kauffman, S. A. *The origins of order: Self-organization and selection in evolution* (Lightning Source UK. Ltd. and Oxford University Press, Milton Keynes UK and New York, 1993).

[15] Kauffman, S. A. Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. *Journal of Cybernetics* **1**, 71–96 (1971).

[16] Schrödinger, E. & Fischer, E. P. *Was ist Leben? Die lebende Zelle mit Augen des Physikers betrachtet*, vol. 1134 of *Serie Piper* (Piper, München and Berlin and Zürich, November 2015).

[17] Dodd, M. S. *et al.* Evidence for early life in earth's oldest hydrothermal vent precipitates. *Nature* **543**, 60–64 (2017).

[18] Neveu, M., Kim, H.-J. & Benner, S. A. The "strong" rna world hypothesis: Fifty years old. *Astrobiology* **13**, 391–403 (2013).

[19] Robertson, M. P. & Joyce, G. F. The origins of the rna world. *Cold Spring Harbor Perspectives in Biology* **4**, a003608–a003608 (2012).

[20] Levy, M. & Miller, S. L. The stability of the rna bases: Implications for the origin of life. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 7933–7938 (1998).

[21] Matthew W. Powner, Béatrice Gerland & John D. Sutherland. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).

[22] Bhavesh H. Patel, Claudia Percivalle, Dougal J. Ritson, Colm D. Duffy & John D. Sutherland. Common origins of rna, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chemistry* **7**, 301–307 (2015).

[23] Callahan, M. P. *et al.* Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13995–13998 (2011).

[24] Pressman, A., Blanco, C. & Chen, I. A. The rna world as a model system to study the origin of life. *Current Biology* **25**, R953–R963 (2015).

[25] Lawrence, M. S. & Bartel, D. P. Processivity of ribozyme-catalyzed rna polymerization. *Biochemistry* **42**, 8748–8755 (2003).

[26] Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an rna enzyme. *Science* **323**, 1229–1232 (2009).

[27] Nielsen, P. E. Peptide nucleic acids and the origin of life. *Chemistry & Biodiversity* **4**, 1996–2002 (2007).

[28] Orgel, L. Origin of life: Enhanced: A simpler nucleic acid. *Science* **290**, 1306–1307 (2000).

[29] Sole, R. V., Munteanu, A., Rodriguez-Caso, C. & Macía, J. Synthetic protocell biology: from reproduction to computation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **362**, 1727–1739 (2007).

[30] Dzieciol, A. J. & Mann, S. Designs for life: protocell models in the laboratory. *Chemical Society Reviews* **41**, 79–85 (2012).

[31] Blain, J. C. & Szostak, J. W. Progress toward synthetic cells. *Annual review of biochemistry* **83**, 615–640 (2014).

[32] Nirenberg, M. W. & Matthaei, J. H. The dependence of cell-free protein synthesis in e. coli upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America* **47**, 1588–1602 (1961).

[33] Caschera, F. & Noireaux, V. Synthesis of 2.3 mg/ml of protein with an all escherichia coli cell-free transcription-translation system. *Biochimie* **99**, 162–168 (2014).

[34] Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nature Biotechnology* **19**, 751–755 (2001).

[35] Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17669–17674 (2004).

[36] Scott, A. *et al.* Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLOS ONE* **11**, e0163058 (2016).

[37] van Nies, P. *et al.* Self-replication of dna by its encoded proteins in liposome-based synthetic cells. *Nature Communications* **9**, 1583 (2018).

[38] Libicher, K., Hornberger, R., Heymann, M. & Mutschler, H. In vitro self-replication and multicistronic expression of large synthetic genomes. *Nature Communications* **11**, 904 (2020).

[39] Ichihashi, N. *et al.* Darwinian evolution in a translation-coupled rna replication system within a cell-like compartment. *Nature Communications* **4**, 2494 (2013).

[40] Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).

[41] Sleator, R. D. Jcvi-syn3.0 – a synthetic genome stripped bare! *Bioengineered* **7**, 53–56 (2016).

[42] Monnard, P.-A. & Deamer, D. W. Membrane self-assembly processes: steps toward the first cellular life. *The Anatomical Record* **268**, 196–207 (2002).

[43] McManus, J. J., Charbonneau, P., Zaccarelli, E. & Asherie, N. The physics of protein self-assembly. *Current Opinion in Colloid & Interface Science* **22**, 73–79 (2016).

[44] Aksyuk, A. A. & Rossmann, M. G. Bacteriophage assembly. *Viruses* **3**, 172–203 (2011).

[45] Blake, A. J. *et al.* Inorganic crystal engineering using self-assembly of tailored building-blocks. *Coordination Chemistry Reviews* **183**, 117–138 (1999).

[46] Rothemund, P. W. K. Folding dna to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).

[47] Grzybowski, B. A., Wilmer, C. E., Kim, J., Browne, K. P. & Bishop, K. J. M. Self-assembly: From crystals to cells. *Soft Matter* **5**, 1110 (2009).

[48] Israelachvili, J. N. *Intermolecular and surface forces* (Elsevier, Amsterdam and Boston and Paris, 2011).

[49] Jones, R. A. L. *Soft condensed matter*, vol. 6, Ed. 2003 of *Oxford master series in physics Condensed matter physics* (Oxford University Press, Oxford, 2003).

[50] Sackmann, E. & Merkel, R. *Lehrbuch der Biophysik.* Master (Wiley-VCH, Weinheim, 2012).

[51] Mason, J. M., Müller, K. M. & Arndt, K. M. Positive aspects of negative design: simultaneous selection of specificity and interaction stability. *Biochemistry* **46**, 4804–4814 (2007).

[52] Bromley, E. H. C., Channon, K., Moutevelis, E. & Woolfson, D. N. Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. *ACS chemical biology* **3**, 38–50 (2008).

[53] Nic, M., Jirat, J., Kosata, B., Jenkins, A. & McNaught, A. *IUPAC Compendium of Chemical Terminology* (IUPAC, Research Triangle Park, NC, 2009).

[54] Muro, E., Atilla-Gokcumen, G. E. & Eggert, U. S. Lipids in cell biology: how can we understand them better? *Molecular biology of the cell* **25**, 1819–1823 (2014).

[55] Mashaghi, S., Jadidi, T., Koenderink, G. & Mashaghi, A. Lipid nanotechnology. *International Journal of Molecular Sciences* **14**, 4242–4282 (2013).

[56] Berg, J. M., Tymoczko, J. L., Gatto, G. J. & Stryer, L. *Biochemistry* (W.H. Freeman & Company a Macmillan Education Imprint, New York, 2015).

[57] Schrum, J. P., Zhu, T. F. & Szostak, J. W. The origins of cellular life. *Cold Spring Harbor Perspectives in Biology* **2**, a002212 (2010).

[58] Nelson, D. L., Cox, M. M. & Lehninger, A. L. *Lehninger principles of biochemistry* (W.H. Freeman and Co, New York, 2005).

[59] Henze, R. & Urry, D. W. Dielectric relaxation studies demonstrate a peptide librational mode in the polypentapeptide of elastin. *Journal of the American Chemical Society* **107**, 2991–2993 (1985).

[60] Chockalingam, K., Blenner, M. & Banta, S. Design and application of stimulus-responsive peptide systems. *Protein engineering, design & selection : PEDS* **20**, 155–161 (2007).

[61] Urry, D. W. *et al.* Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* 1243–1250 (1992).

[62] Meyer, D. E. & Chilkoti, A. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules* **5**, 846–851 (2004).

[63] Cho, Y. *et al.* Effects of hofmeister anions on the phase transition temperature of elastin-like polypeptides. *The Journal of Physical Chemistry B* **112**, 13765–13771 (2008).

[64] Luan, C. H., Parker, T. M., Prasad, K. U. & Urry, D. W. Differential scanning calorimetry studies of nacl effect on the inverse temperature transition of some elastin-based polytetra-, polypenta-, and polynonapeptides. *Biopolymers* **1991**, 465–475 (1991).

[65] Serrano, V., Liu, W. & Franzen, S. An infrared spectroscopic study of the conformational transition of elastin-like polypeptides. *Biophysical Journal* **93**, 2429–2435 (2007).

[66] Lee, H., Kim, H. R., Larson, R. G. & Park, J. C. Effects of the size, shape, and structural transition of thermosensitive polypeptides on the stability of lipid bilayers and liposomes. *Macromolecules* **45**, 7304–7312 (2012).

[67] Trabbic-Carlson, K., Setton, L. A. & Chilkoti, A. Swelling and mechanical behaviors of chemically cross-linked hydrogels of elastin-like polypeptides. *Biomacromolecules* **4**, 572–580 (2003).

[68] Kracke, B. *et al.* Thermoswitchable nanoparticles based on elastin-like polypeptides. *Macromolecules* **48**, 5868–5877 (2015).

[69] Ghoorchian, A., Cole, J. T. & Holland, N. B. Thermoreversible micelle formation using a three-armed star elastin-like polypeptide. *Macromolecules* **43**, 4340–4345 (2010).

[70] Urry, D. W. Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *The Journal of Physical Chemistry B* **101**, 11007–11028 (1997).

[71] Marsh, D. Membrane water-penetration profiles from spin labels. *European biophysics journal : EBJ* **31**, 559–562 (2002).

[72] Slomkowski, S. *et al.* Terminology of polymers and polymerization processes in dispersed systems. *Pure and Applied Chemistry* **83**, 2229–2259 (2011).

[73] Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, Faraday Transactions 2* **72**, 1525 (1976).

[74] Errington, J. L-form bacteria, cell walls and the origins of life. *Open Biology* **3**, 120143 (2013).

[75] Lin, D. M., Koskella, B. & Lin, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World journal of gastrointestinal pharmacology and therapeutics* **8**, 162–173 (2017).

[76] Chanishvili, N. Phage therapy–history from twort and d'herelle through soviet experience to current approaches. *Advances in virus research* **83**, 3–40 (2012).

[77] Summers, W. C. The strange history of phage therapy. *Bacteriophage* **2**, 130–133 (2012).

[78] Vandamme, E. J. & Mortelmans, K. A century of bacteriophage research and applications: Impacts on biotechnology, health, ecology and the economy! *Journal of Chemical Technology & Biotechnology* **94**, 323–342 (2019).

[79] Sulakvelidze, A., Alavidze, Z. & Morris, J. G. Bacteriophage therapy. *Antimicrobial agents and chemotherapy* **45**, 649–659 (2001).

[80] Kutateladze, M. Experience of the eliava institute in bacteriophage therapy. *Virologica Sinica* **30**, 80–81 (2015).

[81] Campbell, A. The future of bacteriophage biology. *Nature Reviews Genetics* **4**, 471–477 (2003).

[82] Sulakvelidze, A. & Kutter, E. *Bacteriophages: Biology and applications* (CRC Press, Boca Raton, FL, 2005).

[83] Carter, J. B. & Saunders, V. A. *Virology: Principles and applications* (John Wiley & Sons, Chichester, England and Hoboken, NJ, 2007).

[84] WYATT, G. R. & COHEN, S. S. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *The Biochemical journal* **55**, 774–782 (1953).

[85] Caspar, D. L. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harbor symposia on quantitative biology* **27**, 1–24 (1962).

[86] Keef, T., Twarock, R. & Elsawy, K. M. Blueprints for viral capsids in the family of polyomaviridae. *Journal of theoretical biology* **253**, 808–816 (2008).

[87] Baker, T. S., Olson, N. H. & Fuller, S. D. Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiology and molecular biology reviews : MMBR* **64**, 237 (2000).

[88] Wurtz, M. Bacteriophage structure. *Electron Microscopy Reviews* **5**, 283–309 (1992).

[89] Douglas, S. M. *et al.* Self-assembly of dna into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).

[90] Kopperger, E. *et al.* A self-assembled nanoscale robotic arm controlled by electric fields. *Science* **359**, 296–301 (2018).

[91] Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable dna nanostructures. *Journal of the American Chemical Society* **134**, 5516–5519 (2012).

[92] Kuzyk, A. *et al.* Dna-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* **483**, 311–314 (2012).

[93] Jungmann, R. *et al.* Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on dna origami. *Nano Letters* **10**, 4756–4761 (2010).

[94] Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic dna devices and assemblies formed by shape-complementary, non-base pairing 3d components. *Science* **347**, 1446–1452 (2015).

[95] Baev, A., Prasad, P. N., Ågren, H., Samoć, M. & Wegener, M. Metaphotonics: An emerging field with opportunities and challenges. *Physics Reports* **594**, 1–60 (2015).

[96] Hentschel, M., Dregely, D., Vogelgesang, R., Giessen, H. & Liu, N. Plasmonic oligomers: the role of individual particles in collective behavior. *ACS Nano* **5**, 2042–2050 (2011).

[97] Lermusiaux, L., Sereda, A., Portier, B., Larquet, E. & Bidault, S. Reversible switching of the interparticle distance in dna-templated gold nanoparticle dimers. *ACS Nano* **6**, 10992–10998 (2012).

[98] Luk'yanchuk, B. *et al.* The fano resonance in plasmonic nanostructures and metamaterials. *Nature Materials* **9**, 707–715 (2010).

[99] Nordlander, P., Oubre, C., Prodan, E., Li, K. & Stockman, M. I. Plasmon hybridization in nanoparticle dimers. *Nano Letters* **4**, 899–903 (2004).

[100] Wang, H., Brandl, D. W., Nordlander, P. & Halas, N. J. Plasmonic nanostructures: artificial molecules. *Accounts of Chemical Research* **40**, 53–62 (2007).

[101] Zhang, J. *et al.* Single-mode waveguiding in bundles of self-assembled semiconductor nanowires. *Applied Physics Letters* **97**, 221915 (2010).

[102] Maier, S. A. *et al.* Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides. *Nature Materials* **2**, 229–232 (2003).

[103] Govorov, A. O., Lee, J. & Kotov, N. A. Theory of plasmon-enhanced förster energy transfer in optically excited semiconductor and metal nanoparticles. *Physical Review B* **76** (2007).

[104] Maier, S. A., Kik, P. G. & Atwater, H. A. Observation of coupled plasmon-polariton modes in au nanoparticle chain waveguides of different lengths: Estimation of waveguide loss. *Applied Physics Letters* **81**, 1714–1716 (2002).

[105] Barrow, S. J. *et al.* Electron energy loss spectroscopy investigation into symmetry in gold trimer and tetramer plasmonic nanoparticle structures. *ACS Nano* **10**, 8552–8563 (2016).

[106] Kuzyk, A. *et al.* Molecular coupling of light with plasmonic waveguides. *Optics Express* **15**, 9908–9917 (2007).

[107] Acuna, G. P. *et al.* Fluorescence enhancement at docking sites of dna-directed self-assembled nanoantennas. *Science* **338**, 506–510 (2012).

[108] Simoncelli, S. *et al.* Quantitative single-molecule surface-enhanced raman scattering by optothermal tuning of dna origami-assembled plasmonic nanoantennas. *ACS Nano* **10**, 9809–9815 (2016).

[109] Weller, L. *et al.* Gap-dependent coupling of ag–au nanoparticle heterodimers using dna origami-based self-assembly. *ACS Photonics* **3**, 1589–1595 (2016).

[110] Kuzyk, A. *et al.* Reconfigurable 3d plasmonic metamolecules. *Nature Materials* **13**, 862–866 (2014).

[111] Tan, S. J., Campolongo, M. J., Luo, D. & Cheng, W. Building plasmonic nanostructures with dna. *Nature Nanotechnology* **6**, 268–276 (2011).

[112] Klein, W. P. *et al.* Multiscaffold dna origami nanoparticle waveguides. *Nano Letters* **13**, 3850–3856 (2013).

[113] Thacker, V. V. *et al.* Dna origami based assembly of gold nanoparticle dimers for surface-enhanced raman scattering. *Nature Communications* **5**, 3448 (2014).

[114] Gür, F. N., Schwarz, F. W., Ye, J., Diez, S. & Schmidt, T. L. Toward self-assembled plasmonic devices: High-yield arrangement of gold nanoparticles on dna origami templates. *ACS Nano* **10**, 5374–5382 (2016).

[115] Anger, P., Bharadwaj, P. & Novotny, L. Enhancement and quenching of single-molecule fluorescence. *Physical Review Letters* **96**, 113002 (2006).

[116] Ghosh, D. & Chattopadhyay, N. Gold nanoparticles: Acceptors for efficient energy transfer from the photoexcited fluorophores. *Optics and Photonics Journal* **03**, 18–26 (2013).

[117] Swierczewska, M., Lee, S. & Chen, X. The design and application of fluorophore-gold nanoparticle activatable probes. *Physical Chemistry Chemical Physics* **13**, 9929–9941 (2011).

[118] Jain, P. K., Huang, W. & El-Sayed, M. A. On the universal scaling behavior of the distance decay of plasmon coupling in metal nanoparticle pairs: A plasmon ruler equation. *Nano Letters* **7**, 2080–2088 (2007).

[119] Sönnichsen, C., Reinhard, B. M., Liphardt, J. & Alivisatos, A. P. A molecular ruler based on plasmon coupling of single gold and silver nanoparticles. *Nature Biotechnology* **23**, 741–745 (2005).

[120] Acuna, G. P. *et al.* Distance dependence of single-fluorophore quenching by gold nanoparticles studied on dna origami. *ACS Nano* **6**, 3189–3195 (2012).

[121] Dutta, P. K. *et al.* Dna-directed artificial light-harvesting antenna. *Journal of the American Chemical Society* **133**, 11985–11993 (2011).

[122] Stein, I. H., Steinhauer, C. & Tinnefeld, P. Single-molecule four-color fret visualizes energy-transfer paths on dna origami. *Journal of the American Chemical Society* **133**, 4193–4195 (2011).

[123] Kick, B., Praetorius, F., Dietz, H. & Weuster-Botz, D. Efficient production of single-stranded phage dna as scaffolds for dna origami. *Nano Letters* **15**, 4672–4676 (2015).

[124] Stahl, E., Martin, T. G., Praetorius, F. & Dietz, H. Facile and scalable preparation of pure and dense dna origami solutions. *Angewandte Chemie* **53**, 12735–12740 (2014).

[125] Gates, E. P., Jensen, J. K., Harb, J. N. & Woolley, A. T. Optimizing gold nanoparticle seeding density on dna origami. *RSC Advances* **5**, 8134–8141 (2015).

[126] Ko, S. H. *et al.* High-speed, high-purity separation of gold nanoparticle-dna origami constructs using centrifugation. *Soft Matter* **10**, 7370–7378 (2014).

[127] Pilo-Pais, M., Goldberg, S., Samano, E., LaBean, T. H. & Finkelstein, G. Connecting the nanodots: programmable nanofabrication of fused metal shapes on dna templates. *Nano Letters* **11**, 3489–3492 (2011).

[128] Shaw, A., Benson, E. & Högberg, B. Purification of functionalized dna origami nanostructures. *ACS Nano* **9**, 4968–4975 (2015).

[129] Wu, N. *et al.* Molecular threading and tunable molecular recognition on dna origami nanostructures. *Journal of the American Chemical Society* **135**, 12172–12175 (2013).

[130] Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline two-dimensional dna-origami arrays. *Angewandte Chemie* **50**, 264–267 (2011).

[131] Aghebat Rafat, A., Pirzer, T., Scheible, M. B., Kostina, A. & Simmel, F. C. Surface-assisted large-scale ordering of dna origami tiles. *Angewandte Chemie* **53**, 7665–7668 (2014).

[132] Nath, N. & Chilkoti, A. Interfacial phase transition of an environmentally responsive elastin biopolymer adsorbed on functionalized gold nanoparticles studied by colloidal surface plasmon resonance. *Journal of the American Chemical Society* **123**, 8197–8202 (2001).

[133] Ghoorchian, A. & Holland, N. B. Molecular architecture influences the thermally induced aggregation behavior of elastin-like polypeptides. *Biomacromolecules* **12**, 4022–4029 (2011).

[134] Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chemical Society Reviews* **46**, 2543–2554 (2017).

[135] Gánti, T. *The principles of life* (Oxford Univ. Press, Oxford, 2003).

[136] Piedrafita, G., Ruiz-Mirazo, K., Monnard, P.-A., Cornish-Bowden, A. & Montero, F. Viability conditions for a compartmentalized protometabolic system: a semi-empirical approach. *PLOS ONE* **7**, e39480 (2012).

[137] Piedrafita, G., Monnard, P.-A., Mavelli, F. & Ruiz-Mirazo, K. Permeability-driven selection in a semi-empirical protocell model: the roots of prebiotic systems evolution. *Scientific Reports* **7**, 3141 (2017).

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

[139] Luisi, P. L., Ferri, F. & Stano, P. Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* **93**, 1–13 (2006).

[140] Walde, P., Goto, A., Monnard, P.-A., Wessicken, M. & Luisi, P. L. Oparin's reactions revisited: Enzymic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. *Journal of the American Chemical Society* **116**, 7541–7547 (1994).

[141] Kita, H. *et al.* Replication of genetic information with self-encoded replicase in liposomes. *ChemBioChem* **9**, 2403–2410 (2008).

[142] Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454**, 122–125 (2008).

[143] Kurihara, K. *et al.* Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated dna. *Nature Chemistry* **3**, 775–781 (2011).

[144] Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angewandte Chemie* **124**, 3168–3172 (2012).

[145] Shin, J., Jardine, P. & Noireaux, V. Genome replication, synthesis, and assembly of the bacteriophage t7 in a single cell-free reaction. *ACS Synthetic Biology* **1**, 408–413 (2012).

[146] Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Molecular Systems Biology* **4**, 220 (2008).

[147] Döbereiner, H. G., Käs, J., Noppl, D., Sprenger, I. & Sackmann, E. Budding and fission of vesicles. *Biophysical Journal* **65**, 1396–1403 (1993).

[148] Zhu, T. F. & Szostak, J. W. Coupled growth and division of model protocell membranes. *Journal of the American Chemical Society* **131**, 5705–5713 (2009).

[149] Terasawa, H., Nishimura, K., Suzuki, H., Matsuura, T. & Yomo, T. Coupling of the fusion and budding of giant phospholipid vesicles containing macromolecules. *Proceedings of the National Academy of Sciences* **109**, 5942–5947 (2012).

[150] Bhattacharya, A., Brea, R. J. & Devaraj, N. K. De novo vesicle formation and growth: an integrative approach to artificial cells. *Chemical Science* **8**, 7912–7922 (2017).

[151] Alberts, B. *et al.* *Molecular biology of the cell* (Garland Science Taylor and Francis Group, New York, NY, 2015).

[152] Li, M., Huang, X., Tang, T.-Y. D. & Mann, S. Synthetic cellularity based on non-lipid micro-compartments and protocell models. *Current Opinion in Chemical Biology* **22**, 1–11 (2014).

[153] Dora Tang, T.-Y. *et al.* Fatty acid membrane assembly on coacervate micro-droplets as a step towards a hybrid protocell model. *Nature Chemistry* **6**, 527–533 (2014).

[154] Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annual Review of Cell and Developmental Biology* **30**, 39–58 (2014).

[155] Douliez, J.-P. *et al.* Catanionic coacervate droplets as a surfactant-based membrane-free protocell model. *Angewandte Chemie* **56**, 13689–13693 (2017).

[156] Mansy, S. S. Model protocells from single-chain lipids. *International Journal of Molecular Sciences* **10**, 835–843 (2009).

[157] Bellomo, E. G., Wyrsta, M. D., Pakstis, L., Pochan, D. J. & Deming, T. J. Stimuli-responsive polypeptide vesicles by conformation-specific assembly. *Nature Materials* **3**, 244–248 (2004).

[158] Lane, N. & Martin, W. F. The origin of membrane bioenergetics. *Cell* **151**, 1406–1416 (2012).

[159] Israelachvili, J. N., Mitchell, D. & Ninham, B. W. Theory of self-assembly of lipid bilayers and vesicles. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **470**, 185–201 (1977).

[160] Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C. & Shively, J. M. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature reviews. Microbiology* **6**, 681–691 (2008).

[161] Rideau, E., Dimova, R., Schwille, P., Wurm, F. R. & Landfester, K. Liposomes and polymersomes: a comparative review towards cell mimicking. *Chemical Society Reviews* **47**, 8572–8610 (2018).

[162] Deshpande, S., Caspi, Y., Meijering, A. E. C. & Dekker, C. Octanol-assisted liposome assembly on chip. *Nature communications* **7**, 10447 (2016).

[163] Pereno, V. *et al.* Electroformation of giant unilamellar vesicles on stainless steel electrodes. *ACS omega* **2**, 994–1002 (2017).

[164] Angelova, M. I. & Dimitrov, D. S. Liposome electroformation. *Faraday Discussions of the Chemical Society* **81**, 303 (1986).

[165] Montes, L.-R., Alonso, A., Goñi, F. M. & Bagatolli, L. A. Giant unilamellar vesicles electroformed from native membranes and organic lipid mixtures under physiological conditions. *Biophysical Journal* **93**, 3548–3554 (2007).

[166] Kita-Tokarczyk, K., Grumelard, J., Haefele, T. & Meier, W. Block copolymer vesicles—using concepts from polymer chemistry to mimic biomembranes. *Polymer* **46**, 3540–3563 (2005).

[167] Deamer, D. W. & Gavino, V. Lysophosphatidylcholine acyltransferase: purification and applications in membrane studies. *Annals of the New York Academy of Sciences* **414**, 90–96 (1983).

[168] Mulkidjanian, A. Y., Bychkov, A. Y., Dibrova, D. V., Galperin, M. Y. & Koonin, E. V. Origin of first cells at terrestrial, anoxic geothermal fields. *Proceedings of the National Academy of Sciences* **109**, E821–30 (2012).

[169] Djokic, T., van Kranendonk, M. J., Campbell, K. A., Walter, M. R. & Ward, C. R. Earliest signs of life on land preserved in ca. 3.5 ga hot spring deposits. *Nature Communications* **8**, 15263 (2017).

[170] Becker, S. *et al.* Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nature Communications* **9**, 163 (2018).

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

[172] Apel, C. L., Deamer, D. W. & Mautner, M. N. Self-assembled vesicles of monocarboxylic acids and alcohols: Conditions for stability and for the encapsulation of biopolymers. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1559**, 1–9 (2002).

[173] Walde, P., Wick, R., Fresta, M., Mangone, A. & Luisi, P. L. Autopoietic self-reproduction of fatty acid vesicles. *Journal of the American Chemical Society* **116**, 11649–11654 (1994).

163

[174] Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* **302**, 618–622 (2003).

[175] Budin, I. & Szostak, J. W. Physical effects underlying the transition from primitive to modern cell membranes. *Proceedings of the National Academy of Sciences* **108**, 5249–5254 (2011).

[176] Chen, I. A., Roberts, R. W. & Szostak, J. W. The emergence of competition between model protocells. *Science* **305**, 1474–1476 (2004).

[177] Chen, I. A. & Szostak, J. W. A kinetic study of the growth of fatty acid vesicles. *Biophysical Journal* **87**, 988–998 (2004).

[178] Tamm, L. K., Crane, J. & Kiessling, V. Membrane fusion: a structural perspective on the interplay of lipids and proteins. *Current Opinion in Structural Biology* **13**, 453–466 (2003).

[179] Südhof, T. C. & Rothman, J. E. Membrane fusion: grappling with snare and sm proteins. *Science* **323**, 474–477 (2009).

[180] Chan, Y.-H. M., van Lengerich, B. & Boxer, S. G. Lipid-anchored dna mediates vesicle fusion as observed by lipid and content mixing. *Biointerphases* **3**, FA17 (2008).

[181] Pantazatos, D. P. & MacDonald, R. C. Directly observed membrane fusion between oppositely charged phospholipid bilayers. *The Journal of membrane biology* **170**, 27–38 (1999).

[182] Sunami, T. *et al.* Detection of association and fusion of giant vesicles using a fluorescence-activated cell sorter. *Langmuir* **26**, 15098–15103 (2010).

[183] Caschera, F. *et al.* Programmed vesicle fusion triggers gene expression. *Langmuir* **27**, 13082–13090 (2011).

[184] Lei, G. & MacDonald, R. C. Lipid bilayer vesicle fusion: Intermediates captured by high-speed microfluorescence spectroscopy. *Biophysical Journal* **85**, 1585–1599 (2003).

[185] Kurihara, K. *et al.* A recursive vesicle-based model protocell with a primitive model cell cycle. *Nature communications* **6**, 8352 (2015).

[186] Wick, R., Walde, P. & Luisi, P. L. Light microscopic investigations of the autocatalytic self-reproduction of giant vesicles. *Journal of the American Chemical Society* **117**, 1435–1436 (1995).

[187] Takakura, K., Toyota, T. & Sugawara, T. A novel system of self-reproducing giant vesicles. *Journal of the American Chemical Society* **125**, 8134–8140 (2003).

[188] Toyota, T. *et al.* Population study of sizes and components of self-reproducing giant multilamellar vesicles. *Langmuir* **24**, 3037–3044 (2008).

[189] Minkenberg, C. B. *et al.* Responsive vesicles from dynamic covalent surfactants. *Angewandte Chemie* **50**, 3421–3424 (2011).

[190] Brea, R. J., Cole, C. M. & Devaraj, N. K. In situ vesicle formation by native chemical ligation. *Angewandte Chemie* **53**, 14102–14105 (2014).

[191] Schmidli, P. K., Schurtenberger, P. & Luisi, P. L. Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes. *Journal of the American Chemical Society* **113**, 8127–8130 (1991).

[192] Wick, R. & Luisi, P. L. Enzyme-containing liposomes can endogenously produce membrane-constituting lipids. *Chemistry & Biology* **3**, 277–285 (1996).

[193] Kuruma, Y., Stano, P., Ueda, T. & Luisi, P. L. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochimica et biophysica acta* **1788**, 567–574 (2009).

[194] Liu, Z. *et al.* In vitro reconstitution and optimization of the entire pathway to convert glucose into fatty acid. *ACS Synthetic Biology* **6**, 701–709 (2017).

[195] 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 Synthetic Biology* **7**, 153–165 (2018).

[196] Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A minimal biochemical route towards de novo formation of synthetic phospholipid membranes. *Nature Communications* **10**, 300 (2019).

[197] Seifert, Berndl & Lipowsky. Shape transformations of vesicles: Phase diagram for spontaneous- curvature and bilayer-coupling models. *Physical review. A, Atomic, molecular, and optical physics* **44**, 1182–1202 (1991).

[198] Yanagisawa, M., Imai, M. & Taniguchi, T. Shape deformation of ternary vesicles coupled with phase separation. *Physical Review Letters* **100**, 148102 (2008).

[199] Budin, I. & Szostak, J. W. Expanding roles for diverse physical phenomena during the origin of life. *Annual Review of Biophysics* **39**, 245–263 (2010).

[200] Ichii, T., Suzuki, H. & Yomo, T. Micro-droplet model for recursive growth and division dynamics of the cell. *EPL (Europhysics Letters)* **96**, 48006 (2011).

[201] Käs, J. & Sackmann, E. Shape transitions and shape stability of giant phospholipid vesicles in pure water induced by area-to-volume changes. *Biophysical Journal* **60**, 825–844 (1991).

[202] Mansy, S. S. & Szostak, J. W. Thermostability of model protocell membranes. *Proceedings of the National Academy of Sciences* **105**, 13351–13355 (2008).

[203] Bar-Ziv & Moses. Instability and pearling states produced in tubular membranes by competition of curvature and tension. *Physical Review Letters* **73**, 1392–1395 (1994).

[204] Miao, Seifert, Wortis & Döbereiner. Budding transitions of fluid-bilayer vesicles: The effect of area-difference elasticity. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* **49**, 5389–5407 (1994).

[205] Möller, F. M., Kriegel, F., Kieß, M., Sojo, V. & Braun, D. Steep ph gradients and directed colloid transport in a microfluidic alkaline hydrothermal pore. *Angewandte Chemie* **56**, 2340–2344 (2017).

[206] Budin, I., Bruckner, R. J. & Szostak, J. W. Formation of protocell-like vesicles in a thermal diffusion column. *Journal of the American Chemical Society* **131**, 9628–9629 (2009).

[207] Leaver, M., Dominguez-Cuevas, P., Coxhead, J. M., Daniel, R. A. & Errington, J. Life without a wall or division machine in bacillus subtilis. *Nature* **457**, 849–853 (2009).

[208] Mercier, R., Kawai, Y. & Errington, J. General principles for the formation and proliferation of a wall-free (l-form) state in bacteria. *eLife* **3** (2014).

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

[210] Berg, J. M., Tymoczko, J. L. & Stryer, L. *Biochemistry* (Freeman, New York, 2002).

[211] Murtas, G., Kuruma, Y., Bianchini, P., Diaspro, A. & Luisi, P. L. Protein synthesis in liposomes with a minimal set of enzymes. *Biochemical and Biophysical Research Communications* **363**, 12–17 (2007).

[212] Huber, M. C. *et al.* Designer amphiphilic proteins as building blocks for the intracellular formation of organelle-like compartments. *Nature Materials* **14**, 125–132 (2015).

[213] Schreiber, A., Huber, M. C. & Schiller, S. M. Prebiotic protocell model based on dynamic protein membranes accommodating anabolic reactions. *Langmuir* **35**, 9593–9610 (2019).

[214] Martín, L., Castro, E., Ribeiro, A., Alonso, M. & Rodríguez-Cabello, J. C. Temperature-triggered self-assembly of elastin-like block co-recombinamers:the controlled formation of micelles and vesicles in an aqueous medium. *Biomacromolecules* **13**, 293–298 (2012).

[215] MacEwan, S. R. & Chilkoti, A. Elastin-like polypeptides: biomedical applications of tunable biopolymers. *Biopolymers* **94**, 60–77 (2010).

[216] Amiram, M., Quiroz, F. G., Callahan, D. J. & Chilkoti, A. A highly parallel method for synthesizing dna repeats enables the discovery of 'smart' protein polymers. *Nature Materials* **10**, 141–148 (2011).

[217] Sun, Z. Z. *et al.* Protocols for implementing an escherichia coli based tx-tl cell-free expression system for synthetic biology. *Journal of visualized experiments : JoVE* e50762 (2013).

[218] Forster, A. C. & Church, G. M. Towards synthesis of a minimal cell. *Molecular Systems Biology* **2**, 45 (2006).

[219] Gil, R., Silva, F. J., Peretó, J. & Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiology and molecular biology reviews : MMBR* **68**, 518–37 (2004).

[220] Lavickova, B. & Maerkl, S. J. A simple, robust, and low-cost method to produce the pure cell-free system. *ACS Synthetic Biology* **8**, 455–462 (2019).

[221] Jewett, M. C., Fritz, B. R., Timmerman, L. E. & Church, G. M. In vitro integration of ribosomal rna synthesis, ribosome assembly, and translation. *Molecular Systems Biology* **9**, 678 (2013).

[222] Caschera, F. *et al.* High-throughput optimization cycle of a cell-free ribosome assembly and protein synthesis system. *ACS Synthetic Biology* **7**, 2841–2853 (2018).

[223] Lee, K. Y. *et al.* Photosynthetic artificial organelles sustain and control atp-dependent reactions in a protocellular system. *Nature Biotechnology* **36**, 530–535 (2018).

[224] Pavan Kumar, B. V. V. S. *et al.* Chloroplast-containing coacervate micro-droplets as a step towards photosynthetically active membrane-free protocells. *Chemical Communications* **54**, 3594–3597 (2018).

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

[226] Martino, C. *et al.* Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures. *Angewandte Chemie* **51**, 6416–6420 (2012).

[227] Hardy, M. D. *et al.* Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8187–8192 (2015).

[228] Chu, H.-S. *et al.* Expression analysis of an elastin-like polypeptide (elp) in a cell-free protein synthesis system. *Enzyme and Microbial Technology* **46**, 87–91 (2010).

[229] Martin, R. W. *et al.* Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nature Communications* **9**, 1203 (2018).

[230] Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The all e. coli tx-tl toolbox 2.0: A platform for cell-free synthetic biology. *ACS Synthetic Biology* **5**, 344–355 (2016).

[231] Schwarz-Schilling, M., Aufinger, L., Mückl, A. & Simmel, F. C. Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets. *Integrative Biology* **8**, 564–570 (2016).

[232] Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. & Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nature Chemistry* **9**, 431–439 (2017).

[233] Rustad, M., Eastlund, A., Jardine, P. & Noireaux, V. Cell-free txtl synthesis of infectious bacteriophage t4 in a single test tube reaction. *Synthetic Biology* **3**, 253 (2018).

[234] Urry, D. W. Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. *Journal of Muscle Research and Cell Motility* **23**, 543–559 (2002).

[235] Meyer, D. E. & Chilkoti, A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature Biotechnology* **17**, 1112–1115 (1999).

[236] Tenchov, B. G., Yanev, T. K., Tihova, M. G. & Koynova, R. D. A probability concept about size distributions of sonicated lipid vesicles. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **816**, 122–130 (1985).

[237] Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an rna mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *Journal of the American Chemical Society* **136**, 16299–16308 (2014).

[238] McPherson, D. T., Xu, J. & Urry, D. W. Product purification by reversible phase transition following escherichia coli expression of genes encoding up to 251 repeats of the elastomeric pentapeptide gvgvp. *Protein Expression and Purification* **7**, 51–57 (1996).

[239] Meyer, D. E. & Chilkoti, A. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* **3**, 357–367 (2002).

[240] Ben-Bassat, A. *et al.* Processing of the initiation methionine from proteins: properties of the escherichia coli methionine aminopeptidase and its gene structure. *Journal of Bacteriology* **169**, 751–757 (1987).

[241] Sakatani, Y., Yomo, T. & Ichihashi, N. Self-replication of circular dna by a self-encoded dna polymerase through rolling-circle replication and recombination. *Scientific Reports* **8**, 13089 (2018).

[242] Bansho, Y., Furubayashi, T., Ichihashi, N. & Yomo, T. Host-parasite oscillation dynamics and evolution in a compartmentalized rna replication system. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 4045–4050 (2016).

[243] Mills, D. R., Peterson, R. L. & Spiegelman, S. An extracellular darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences* **58**, 217–224 (1967).

[244] Shigeyoshi Matsumura *et al.* Transient compartmentalization of rna replicators prevents extinction due to parasites. *Science* **354**, 1293–1296 (2016).

[245] Anthony C Forster & George M Church. Towards synthesis of a minimal cell. *Molecular Systems Biology* **2** (2006).

[246] Chamberlin, M., McGrath, J. & Waskell, L. New rna polymerase from escherichia coli infected with bacteriophage t7. *Nature* **228**, 227–231 (1970).

[247] Gibson, D. G. *et al.* Enzymatic assembly of dna molecules up to several hundred kilobases. *Nature methods* **6**, 343–345 (2009).

[248] Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).

[249] Hatoum-Aslan, A. Phage genetic engineering using crispr-cas systems. *Viruses* **10** (2018).

[250] Barbu, E. M., Cady, K. C. & Hubby, B. Phage therapy in the era of synthetic biology. *Cold Spring Harbor Perspectives in Biology* **8** (2016).

[251] Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering modular viral scaffolds for targeted bacterial population editing. *Cell systems* **1**, 187–196 (2015).

[252] Dedrick, R. M. *et al.* Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant mycobacterium abscessus. *Nature Medicine* **25**, 730–733 (2019).

[253] Lu, T. K. & Collins, J. J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11197–11202 (2007).

[254] Pires, D. P., Cleto, S., Sillankorva, S., Azeredo, J. & Lu, T. K. Genetically engineered phages: a review of advances over the last decade. *Microbiology and molecular biology reviews : MMBR* **80**, 523–543 (2016).

[255] Kiro, R., Shitrit, D. & Qimron, U. Efficient engineering of a bacteriophage genome using the type i-e crispr-cas system. *RNA biology* **11**, 42–44 (2014).

[256] Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by rnai in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 1443–1448 (2002).

[257] Malinen, A. M. *et al.* Active site opening and closure control translocation of multisubunit rna polymerase. *Nucleic acids research* **40**, 7442–7451 (2012).

[258] Bonner, G., Lafer, E. M. & Sousa, R. Characterization of a set of t7 rna polymerase active site mutants. *The Journal of biological chemistry* **269**, 25120–25128 (1994).

[259] Leung, E., Weil, D. E., Raviglione, M. & Nakatani, H. The who policy package to combat antimicrobial resistance. *Bulletin of the World Health Organization* **89**, 390–392 (2011).

[260] Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: The who priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet Infectious Diseases* **18**, 318–327 (2018).

[261] Thorpe, K. E., Joski, P. & Johnston, K. J. Antibiotic-resistant infection treatment costs have doubled since 2002, now exceeding $2 billion annually. *Health affairs (Project Hope)* **37**, 662–669 (2018).

[262] Cassini, A. *et al.* Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the eu and the european economic area in 2015: A population-level modelling analysis. *The Lancet Infectious Diseases* **19**, 56–66 (2019).

[263] Casey, P. G., Hill, C. & Gahan, C. G. E. coli o104:h4: social media and the characterization of an emerging pathogen. *Bioengineered bugs* **2**, 189–193 (2011).

[264] Kimmitt, P. T., Harwood, C. R. & Barer, M. R. Toxin gene expression by shiga toxin-producing escherichia coli: the role of antibiotics and the bacterial sos response. *Emerging infectious diseases* **6**, 458–465 (2000).

[265] Michael, M., Elliott, E. J., Ridley, G. F., Hodson, E. M. & Craig, J. C. Interventions for haemolytic uraemic syndrome and thrombotic thrombocytopenic purpura. *The Cochrane database of systematic reviews* CD003595 (2009).

[266] Pirnay, J.-P., Vos, D. & Verbeken, G. Clinical application of bacteriophages in europe. *Microbiology Australia* (2019).

[267] Chatterjee, M. *et al.* Antibiotic resistance in pseudomonas aeruginosa and alternative therapeutic options. *International journal of medical microbiology : IJMM* **306**, 48–58 (2016).

[268] Jault, P. *et al.* Efficacy and tolerability of a cocktail of bacteriophages to treat burn wounds infected by pseudomonas aeruginosa (phagoburn): A randomised, controlled, double-blind phase 1/2 trial. *The Lancet Infectious Diseases* **19**, 35–45 (2019).

[269] Rohde, C., Wittmann, J. & Kutter, E. Bacteriophages: A therapy concept against multi-drug-resistant bacteria. *Surgical infections* **19**, 737–744 (2018).

[270] Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D. Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the enterobacteriaceae family - a review. *Frontiers in microbiology* **8**, 1108 (2017).

[271] Rohde, C. *et al.* Expert opinion on three phage therapy related topics: Bacterial phage resistance, phage training and prophages in bacterial production strains. *Viruses* **10** (2018).

[272] Azeredo, J. & Sillankorva, S. *Bacteriophage Therapy*, vol. 1693 (Springer New York, New York, NY, 2018).

[273] Brüssow, H. What is needed for phage therapy to become a reality in western medicine? *Virology* **434**, 138–142 (2012).

[274] Lu, T. K. & Koeris, M. S. The next generation of bacteriophage therapy. *Current opinion in microbiology* **14**, 524–531 (2011).

[275] Twort, F. W. An investigation on the nature of ultra-microscopic viruses. *The Lancet* **186**, 1241–1243 (1915).

[276] Consortium, T. U. Uniprot: A worldwide hub of protein knowledge. *Nucleic acids research* **47**, D506–D515 (2019).

[277] Geer, L. Y. *et al.* The ncbi biosystems database. *Nucleic acids research* **38**, D492–6 (2010).

[278] Jaroentomeechai, T. *et al.* Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nature communications* **9**, 2686 (2018).

[279] Jia, H., Heymann, M., Bernhard, F., Schwille, P. & Kai, L. Cell-free protein synthesis in micro compartments: building a minimal cell from biobricks. *New biotechnology* **39**, 199–205 (2017).

[280] Rustad, M., Eastlund, A., Marshall, R., Jardine, P. & Noireaux, V. Synthesis of infectious bacteriophages in an e. coli-based cell-free expression system. *Journal of visualized experiments : JoVE* (2017).

[281] Oh-Hashi, K., Furuta, E., Fujimura, K. & Hirata, Y. Application of a novel hibit peptide tag for monitoring atf4 protein expression in neuro2a cells. *Biochemistry and biophysics reports* **12**, 40–45 (2017).

[282] Grazziotin, A. L., Koonin, E. V. & Kristensen, D. M. Prokaryotic virus orthologous groups (pvogs): a resource for comparative genomics and protein family annotation. *Nucleic acids research* **45**, D491–D498 (2017).

[283] Daudén, M. I. *et al.* Large terminase conformational change induced by connector binding in bacteriophage t7. *The Journal of biological chemistry* **288**, 16998–17007 (2013).

[284] Nguyen, H. M. & Kang, C. Lysis delay and burst shrinkage of coliphage t7 by deletion of terminator tphi reversed by deletion of early genes. *Journal of virology* **88**, 2107–2115 (2014).

[285] Wang, I. N., Smith, D. L. & Young, R. Holins: the protein clocks of bacteriophage infections. *Annual review of microbiology* **54**, 799–825 (2000).

[286] Garneau, J. E. *et al.* The crispr/cas bacterial immune system cleaves bacteriophage and plasmid dna. *Nature* **468**, 67–71 (2010).

[287] Haldenwang, W. G. The sigma factors of bacillus subtilis. *Microbiological reviews* **59**, 1–30 (1995).

[288] Davison, B. L., Murray, C. L. & Rabinowitz, J. C. Specificity of promoter site utilization in vitro by bacterial rna polymerases on bacillus phage phi 29 dna. transcription mapping with exonuclease iii. *The Journal of biological chemistry* **255**, 8819–8830 (1980).

[289] Whiteley, H. R. & Ernest Hemphill, H. The interchangeability of stimulatory factors isolated from three microbial rna polymerases. *Biochemical and Biophysical Research Communications* **41**, 647–654 (1970).

[290] McAllister, W. T. & Wu, H. L. Regulation of transcription of the late genes of bacteriophage t7. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 804–808 (1978).

[291] Heineman, R. H. & Bull, J. J. Testing optimality with experimental evolution: lysis time in a bacteriophage. *Evolution; international journal of organic evolution* **61**, 1695–1709 (2007).

[292] Jurač, K., Nabergoj, D. & Podgornik, A. Bacteriophage production processes. *Applied microbiology and biotechnology* **103**, 685–694 (2019).

[293] Stewart, E. J. Growing unculturable bacteria. *Journal of Bacteriology* **194**, 4151–4160 (2012).

[294] Pirnay, J.-P. *et al.* The magistral phage. *Viruses* **10** (2018).

174

[295] Labuda, A. *et al.* Comparison of photothermal and piezoacoustic excitation methods for frequency and phase modulation atomic force microscopy in liquid environments. *AIP Advances* **1**, 022136 (2011).

[296] Mallik, L. *et al.* Electron microscopic visualization of protein assemblies on flattened dna origami. *ACS Nano* **9**, 7133–7141 (2015).

[297] McNair, K., Zhou, C., Dinsdale, E. A., Souza, B. & Edwards, R. A. Phanotate: a novel approach to gene identification in phage genomes. *Bioinformatics (Oxford, England)* **35**, 4537–4542 (2019).

[298] Cox, J. & Mann, M. Maxquant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**, 1367–1372 (2008).

[299] Schwanhäusser, B. *et al.* Corrigendum: Global quantification of mammalian gene expression control. *Nature* **495**, 126–127 (2013).

[300] Hohenester, U. & Trügler, A. Mnpbem – a matlab toolbox for the simulation of plasmonic nanoparticles. *Computer Physics Communications* **183**, 370–381 (2012).

# 10 Acknowledgments