DISS. ETH NO. 27959

Engineering of Artificial Cells with a Focus on Membrane Proteins

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
Noah Kaspar Ritzmann
M.Sc. Biotechnology, ETH Zurich

born on 10.05.1990
citizen of Basel-Stadt

accepted on the recommendation of
Professor Dr. Daniel Müller
Professor Dr. Sebastian Hiller
Professor Dr. Dimitrios Fotiadis

2021

One Eternity Later
SpongeBob, episode *Squid on Strike*, Season 2, 2001

Table of Contents

Zusammenfassung

Das Leben auf der Erde wurde in den mehr als 4 Milliarden Jahren der Evolution mit einer Vielzahl an einzigartigen Fähigkeiten ausgestattet. Durch den ständigen selektiven Druck, sich an spezifische Nischen anzupassen, wird sogar von den einfachsten Formen des Lebens ein hohes Maß an Komplexität gefordert. Daher sind bis heute selbst einzelne biologische Zellen zu komplex, um diese vollständig zu verstehen, obwohl sie die kleinste funktionelle Einheit des Lebens darstellen. Dennoch, scheinbar unbeeindruckt von dieser überwältigenden Komplexität, sind im Bereich der synthetischen Biologie Ansätze entwickelt worden, die darauf abzielen, künstliche Zellen von Grund auf neu zu erschaffen. Die Beweggründe dazu sind vielfältig, konzentrieren sich aber vor allem auf zwei Hauptziele. Während sich manche Forscher darauf fokussieren, künstliche Zellen zu erschaffen, um die minimalen Anforderungen zu verstehen, die für die Entstehung des Lebens erforderlich sind, streben andere danach, natürliche Funktionen von biologischen Zellen nachzuahmen und zu erweitern, um neue Anwendungen im Gesundheitswesen und in der Biotechnologie zu entwickeln. Das Ziel dieser Arbeit entspricht dem letzteren Ansatz.

Obwohl viele unterschiedliche Ansätze existieren, um künstliche Zellen herzustellen, konzentriert sich diese Arbeit auf die Anwendung von Membranproteinen als funktionelle Bausteine. Membranproteine sind, wie der Name schon besagt, Proteine, die mit Membranen assoziiert sind. Die Membranen machen den Unterschied zwischen einer Zelle und einer losen Ansammlung von Biomolekülen aus. Biologische Membranen trennen das Zellinnere von der Umgebung ab und bilden dadurch Kompartimente, die den Zellinhalt einschliessen, während unerwünschte Moleküle ausgeschlossen werden. Diese räumliche Trennung ist überlebenswichtig für Zellen. Trotzdem sind sie auf den ständigen Austausch mit ihrer Umwelt angewiesen, um Nährstoffe aufzunehmen und mit benachbarten Zellen zu kommunizieren. Um dies zu gewährleisten, bauen Zellen Proteine in ihre Membranen ein. Proteine sind nanoskopisch kleine Maschinen, welche so gut wie jede Funktion innerhalb von Zellen erfüllen. Dies erledigen sie entweder als lösliche Proteine eingeschlossen durch die Zellmembran oder als Membranproteine, welche in die Membran eingebettet sind.

Aufgrund ihrer unglaublichen Vielfalt sind Proteine interessante Bausteine für die Herstellung von künstlichen Zellen. Dabei ist ein naheliegender Ansatz, um einzelne Bausteine zu einer künstlichen Zelle zusammenzusetzen und miteinander zu kombinieren, derjenige von natürlichen Zellen: ein Kompartiment gebildet durch eine Membran, das lösliche Proteine umschliesst und mit Membranproteinen ausgestattet ist. Obwohl diese Idee zunächst simpel erscheinen mag, ist deren Umsetzung anspruchsvoll. Zudem ist die Herstellung von komplexeren künstlichen Zellen begrenzt durch die heute verfügbaren Methoden und Bausteine. Dabei ist der funktionelle Einbau von Membranproteinen besonders anspruchsvoll und deshalb oftmals der limitierende Faktor. Deshalb ist das Hauptziel dieser Arbeit, neue Ansätze zu erkunden, um gegenwärtige Hindernisse in diesem Bestreben zu überwinden.

Dazu wurden Membranproteine im Hinblick auf ihre Faltung und den Membraneinbau untersucht, um neue Lösungen für die heutigen Limitationen in der Herstellung künstlicher Zellen zu finden. Darüber hinaus wurde eine Methode entwickelt, um den Einbau von Membranproteinen in künstliche Membranen besser kontrollieren zu können, welche danach zur Herstellung einfacher künstlicher Zellen verwendet wurde. Abschliessend wurden komplexere künstliche Zellen hergestellt, die aufzeigen, dass grosses Potential in der Kombination verschiedener Herstellungsmethoden liegt.

Um die komplexen Kräfte zu untersuchen, welche Membranproteine zusammenhalten, wurden LamB und BamA, beides Proteine der Äusseren Membrane von *Escherichia Coli* untersucht. Dazu wurde Rasterkraftmikroskopie basierte Einzelmolekül-Kraftspektroskopie verwendet, welche die mechanische Entfaltung und Extraktion einzelner Proteine aus ihrer Membranumgebung ermöglicht. Dadurch werden die individuellen strukturellen Segmente identifiziert, welche die jeweiligen Proteine stabilisieren. Durch das Aufbringen einer mechanischen Zugkraft wurden die beiden Proteine sequentiell entfaltet. Die Ergebnisse decken sich mit den Resultaten aus vorhergehenden Einzelmolekül-Kraftspektroskopie Experimenten an Proteinen der Äusseren Membran und deuten auf ein gemeinsames Entfaltungsverhalten hin.
BamA ist besonders interessant, da es die Hauptkomponente des β-Barrel Assembly Machinery (BAM)-Komplexes ist, der dafür zuständig ist, Proteine in die Äussere Membran von *E. coli* einzubauen. Dabei kann ein vertieftes Verständnis des natürlichen Proteineinbauprozesses dazu beitragen, besser zu verstehen, wie Proteine in künstliche Membranen eingebaut werden können. Deshalb wurden verschiedene Faktoren wie die Membran oder verschiedene mutierte BamA Varianten verglichen, um die jeweiligen Auswirkungen auf die dynamische β-Barrel Region von BamA zu untersuchen. Zudem wurde BamA in der Gegenwart von Darobactin untersucht, einem kürzlich entdeckten Antibiotikum, das BamA inhibiert. Durch den Einsatz von dynamischer Kraftspektroskopie wurden die mechanischen, energetischen und kinetischen Eigenschaften einzelner Struktursegmente, die BamA stabilisieren, charakterisiert. Die Ergebnisse deuten auf eine funktionell relevante Rolle der Linker-Region zwischen der Membran- und der löslichen Domäne von BamA hin und identifizieren zudem asymmetrische mechanische Eigenschaften in der β-Barrel Region.
Zusätzlich zu den Untersuchungen der natürlichen Faltungs- und Einbauprozesse von Membranproteinen wurden neue Membranproteine entwickelt, um die Kontrolle über Membranproteineinbau *in vitro* zu verbessern. Herkömmliche Proteine nehmen in der künstlichen Membran normalerweise zufällige Orientierungen an, da sie während der Aufreinigung und Rekonstitution ihre Orientierung nicht beibehalten. In vielen Fällen führt dies zu funktionsunfähigen künstlichen Zellen. Die präsentierten Ergebnisse zeigen jedoch, dass die Orientierung von Membranproteinen mit löslichen Fusionsproteinen kontrolliert werden kann. Dies blockiert eine der beiden möglichen Orientierungen und liefert somit eine neue Methode, um die Herstellung künstlicher Zellen besser kontrollieren zu können. Um die Funktionalität des Ansatzes zu demonstrieren, wurde Proteorhodopsin verwendet. Da es sich dabei um eine lichtgetriebene Protonenpumpe handelt, dienen die vorgestellten Proteinkonstrukte auch dazu, Energie bereitzustellen. Durch das Beleuchten dieser einfachen künstlichen Zellen wandelt Proteorhodopsin Lichtenergie in einen Protonengradienten über die Membran um. Die Richtung des Protonengradienten kann dabei durch die Proteorhodopsin-Orientierung in der Membran gesteuert werden. Die vorgestellten Proteinkonstrukte dienen dazu, zukünftige künstliche Zellen mit Energie zu versorgen und dabei weitere energieabhängige Bausteine anzutreiben.
In einer weiteren Studie wurde eine andere lichtbetriebene Protonenpumpe, Xenorhodopsin, in komplexeren künstlichen Zellen verwendet, um pH-abhängige Prozesse zu steuern. Dazu wurden die Protonenpumpen in Bakterien exprimiert, die in größere künstliche Kompartimente eingekapselt wurden, wo sie als künstliche Organellen funktionierten. Wenn diese beleuchtet wurden, veränderten die Protonenpumpen in der Bakterienmembran den pH-Wert ihrer Umgebung, was dazu genutzt wurde, die Assoziation oder die Dissoziation von pH-abhängigen DNA-

Konstrukten zu kontrollieren. Dadurch wurde ein künstliches, vereinfachtes Zytoskelett hergestellt, dessen Zusammenbau von einer externen Lichtquelle ausgelöst werden kann. In einem abschließenden Experiment wird ein komplexes System demonstriert, welches durch Beleuchtung mit einer Lichtquelle ermöglicht, die Morphologie künstlicher Kompartimente zu verformen. Die vorgestellten Umsetzungen künstlicher Zellen demonstrieren neue Möglichkeiten zur Herstellung komplexer Systeme und liefern neue Bausteine für weitere Ansätze im Bereich der künstlichen Zellentwicklung.

Summary

More than 4 billion years of evolution have equipped life on earth with a broad range of unique properties and abilities. Forced by the constant selective pressure to adapt to specific niches, life requires high levels of complexity, even in its simplest forms. As of today, individual biological cells are too complex to be fully understood, even though they represent the smallest functional unit of life. Nevertheless, seemingly unimpressed by this overwhelming complexity, approaches have emerged in the field of synthetic biology to assemble artificial cells from scratch. The motivations for this undertaking are diverse, but mainly revolve around two main goals. The first branch seeks to create artificial cells to understand the minimal requirements necessary for the emergence of life, whereas the other branch aims to mimic and extend natural functionality to develop novel applications in healthcare and biotechnology. The aim of this thesis is in line with the second branch.

While many different approaches exist to assemble artificial cells, this thesis focusses on the application of membrane proteins as functional building blocks. Membrane proteins, as indicated by their name, are proteins that are associated with membranes. The membranes, in turn, make the difference between a cell and a loose collection of biomolecules. By separating the inside of the cell from the surrounding environment, biological membranes form compartments, which contain the cellular content, while excluding undesired molecules. Although this spatial separation is essential for survival, cells need constant exchange with their environment in order to import nutrients and communicate with other cells. To this end, cells embed proteins into their membranes.

Proteins are nanoscopic machines, which perform virtually every function within cells, either in the form of soluble proteins enclosed by the membrane compartment or in the form of membrane proteins, embedded in the membrane boundary. Due to their incredible diversity, proteins are interesting building blocks for artificial cells. Thereby, the most obvious building plan to assemble and combine the individual building blocks to an artificial cell is that provided by natural cells: a compartment formed by a membrane, which encloses soluble proteins and is functionalized with membrane proteins. However, while the idea is simple, the implementation is challenging and with the methods and building blocks available today, the achievable complexity of artificial cells remains limited. The functional incorporation of membrane proteins in particular is the bottleneck of many approaches. The main aim of this thesis, therefore, is to provide new ways to overcome the current obstacles in this endeavour.

Hence, in order to provide new means to overcome the current limitations in the field of artificial cell engineering, membrane proteins have been investigated in the context of folding and insertion into the membrane environment. Moreover, a method has been developed, to improve the control over membrane protein insertion into artificial membranes, which was used to assemble simplistic artificial cells. Finally, more complex artificial cells were implemented to demonstrate the potential of combining different engineering approaches.

To study the intricate network of forces stabilizing membrane proteins, the outer membrane proteins (OMPs) LamB and BamA from *Escherichia Coli*., were investigated using atomic force microscopy (AFM)-based single-molecule force spectroscopy (SMFS). This allowed to mechanically unfold and extract single proteins from their membrane environment to identify the individual structural segments stabilizing the proteins. Upon applying a mechanical pulling force, the OMPs sequentially unfolded individual β-hairpins, which is in line with previous SMFS studies performed on β-barrel OMPs and substantiates a common unfolding behaviour. BamA is of particular interest, as it is the

main component of the β-barrel assembly machinery (BAM) complex, which is responsible to fold and insert OMPs into the *E. coli* outer membrane. Improving the understanding of the natural membrane protein insertion process, might help to improve our efforts to insert membrane proteins into artificial membranes. Hence, different factors modulating BamA, such as the membrane environment or specific mutations, were compared in order to investigate the respective effects on the dynamic BamA β-barrel region. Additionally, BamA was studied in presence of darobactin, a novel antibiotic compound, which was found to inhibit BamA. By employing SMFS in the dynamic mode, the mechanical, energetic and kinetic properties of individual structural segments stabilizing BamA were characterized. The results propose a functionally relevant role for the linker region between the BamA β-barrel and its soluble domains, and further identify asymmetric mechanical properties within the β-barrel.

In addition to the study of natural membrane protein folding and insertion processes, new membrane proteins were designed to gain control over *in vitro* membrane insertion processes. Without this approach reconstituted membrane proteins typically adopt random orientations, as the processes of membrane protein purification and reconstitution into artificial membranes does not maintain their original orientation. In many cases this renders the resulting artificial cell non-functional. However, the presented results demonstrate, that their final orientation can be controlled by attaching a soluble fusion protein to one terminus of the membrane protein of interest, which prevents one of the two possible orientations. The presented approach to control membrane protein orientation thus provides a new tool to better control the assembly of artificial cells. Since proteorhodopsin, a light-driven proton pump, was used to demonstrate the functionality of the approach, the presented protein constructs also serve as energizing building blocks for artificial cells. Upon illumination of the resulting simplistic artificial cells, proteorhodopsin converts the energy of light into a proton gradient across the membrane. The directionality of the proton gradient can be controlled by choosing the desired proteorhodopsin orientation in the membrane. In combination with downstream energy-dependent building blocks, the presented protein constructs will hopefully be useful to energize more complex artificial cells in the future.

In a final study, another light-driven proton pump, xenorhodopsin, was used in more complex artificial cells, to control pH-dependent processes. To this end, the proton pumps were expressed in bacteria, which were encapsulated as artificial organelles in larger artificial compartments. There, the proton pumps in the bacterial membrane changed the pH of their environment upon illumination, which controlled the assembly and disassembly of pH-dependent DNA constructs. With this system, an artificial cytoskeleton-mimic was built up, the assembly of which could be triggered externally with light. In a final experiment, a complex interdependent system is demonstrated, which allows to modulate the morphology of artificial compartments upon illumination. The presented implementations of artificial cells demonstrate new means to assemble complex systems and provide new building blocks for further artificial cell engineering efforts.

1 Introduction

1.1 The Concept of Artificial Cells

Inspired by the incredible properties of living cells, the field of synthetic biology aims to engineer artificial cells to bring about novel applications in healthcare and biotechnology. For this purpose, it is necessary to break down the biological complexity into components that are simpler to understand and quantify. One prominent component shared among all types of living cells is the membrane, which separates the inner from the outer environment. This creates a spatial compartment, which encloses the cellular content. In order to functionalize this compartment, cells embed proteins into their membrane. Those membrane proteins perform many essential functions, like sensing the cellular environment[1], importing nutrients[2], and transferring signals across the membrane[3]. Living cells further contain a large number of diverse soluble proteins enclosed within their membrane boundary.

Many artificial cell engineering approaches follow the natural model and employ artificial or natural membranes to enclose their soluble content. Furthermore, like in living cells, artificial compartments are often functionalized with proteins, to utilize their diverse properties. In order to assemble the individual components into artificial cells, membrane compartments are typically combined with membrane proteins that are embedded in the membrane envelope and with soluble proteins, encapsulated within the compartment. This functionally connects the incorporated components to mimic complex functions of living cells or to equip the compartment with novel properties, unseen in nature. With this approach, notable examples of artificial cells have been developed so far, for applications such as targeted drug delivery[4], nanoreactors[5] and the conversion of light energy into chemical energy[6] (Figure 1). However, the assembly process of artificial cells from individual components is a challenging endeavor and the achievable complexity remains limited to this date.

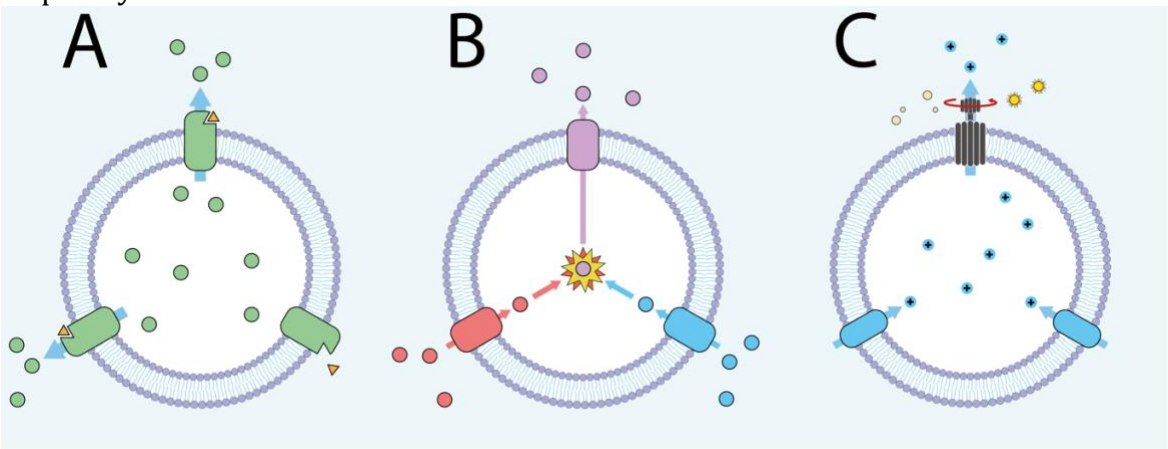


Figure 1. Examples of Artificial Cells. (A) Schematic representation of an artificial cell, designed to export encapsulated compounds (green spheres) in presence of a specific substrate (orange triangles). This might be applied to achieve targeted drug release. (B) Schematic representation of a nanoreactor, which imports and colocalizes educts (red and blue spheres) to accelerate the reaction to the product (purple sphere), which can be exported subsequently. (C) Schematic representation of an artificial cell, that converts the energy of light into chemical energy.

1.2 Forming a Compartment

The assembly of artificial cells, requires spatial separation from the environment, to form functional units. Inspirations for compartmentalization strategies can be found in great numbers in nature[7]. Virus capsids for example, consist of repeating protein units,

assembled to larger hollow structures and enclose the viral genetic material[8]. Similar supramolecular protein assemblies have been found to form microcompartments in bacteria, where they colocalize enzymes with their substrate to promote specific metabolic processes[9]. Moreover, cellular organelles form specialized sub-compartments within cells, to not only separate components of the cell from the outside, but also from each other.
While artificial counterparts to some of those natural compartments have been developed[10–12], most engineering approaches are focused on the assembly of artificial cells. Similar to living cells, which have evolved different types of architecture, various types of artificial cells have emerged, mostly due to varying assembly strategies. Some types of artificial cells consist of water-in-oil droplets, produced on microfluidic chips or stabilized in hydrogels[13]. Other approaches lead to compartments, which are separated from the environment by artificial membranes. Impressive results have been achieved with synthetic block copolymers, which form hollow spherical polymer-membranes[5,14]. They are generally thicker than natural membranes, which greatly improves their mechanical properties[15]. However, the increased thickness renders them incompatible with the functional incorporation of some membrane proteins. Hence, the building blocks most commonly used to separate the content of engineered cells from their environment are lipids, which are the closest thing to living cells, since lipids form the basis of natural cell membranes.

1.2.1 Biological Membranes

The membrane that surrounds a natural cell is called plasma membrane and is made up of lipids but also contains proteins and carbohydrates[16]. It serves many purposes like fending harmful molecules, selectively importing nutrients and transferring signals for cell-cell communication[17]. The foundation for the assembly of most biological membranes is laid by phospholipids, which form a bilayer due to their amphipathic properties[18], meaning they are partially polar and apolar (Figure 2A). The phospholipid head contains a negatively charged phosphate group, which can establish electrostatic interactions with the polar water molecules. However, the apolar tail domains disrupt the hydrogen bonding network formed by the surrounding water molecules, which is energetically unfavorable. Thus, driven by the hydrophobic effect[18], the phospholipids self-assemble to form bilayers, consisting of two stacked phospholipid layers, to protect their apolar tails from the aqueous solution. The orientations of the individual layers within the bilayer are mirrored, so that the phospholipid tails are protected from the surrounding aqueous solution by the upper and the lower layer of phospholipid head groups (Figure 2B). The bilayer closes to form a hollow compartment, which encapsulates the cellular content (Figure 2C).
The phospholipid bilayer is semipermeable. The size and charge of molecules determines whether they can penetrate the membrane or not. Generally, small hydrophobic molecules can effectively diffuse through the membrane, while the bilayer is poorly permeable for water molecules due to the hydrophobic core. Despite their barrier-like properties, plasma membranes are far from rigid but are two-dimensional fluids as components in the membrane can diffuse along its plane[19]. The chemical properties of the fatty acids in the hydrophobic tails of the lipid determine the fluidity of the membrane. Unsaturated fatty acids reduce the packing of lipids due to sterical hinderance and hence fluidice the membrane. Due to those fluid-like properties, plasma membranes can adapt to cell shape changes, to changes in the environment and can self-heal after mechanical deformation or even punctuation[20–22]. However, the membrane fluidity is not uniform

across the whole cell surface. Some components in the membrane can diffuse freely, while other components diffuse through the membrane assembled into larger entities[23–25].
Besides phospholipids, repeating carbohydrate units, called oligosaccharides contribute to biological membranes. They are covalently linked to the extracellular side of lipids or membrane proteins, forming glycolipids[26] or glycoproteins[27] respectively. They are involved in cell-cell communication and adhesion and are relevant markers for the immune system[27]. The third group of macromolecules contributing to biological membranes are membrane proteins, which equip the plasma membrane with essential properties, required for cell functionality. However, in order to understand membrane proteins and their relevance to artificial cells, we have to understand what proteins are in general and for what purpose they are maintained by cells.

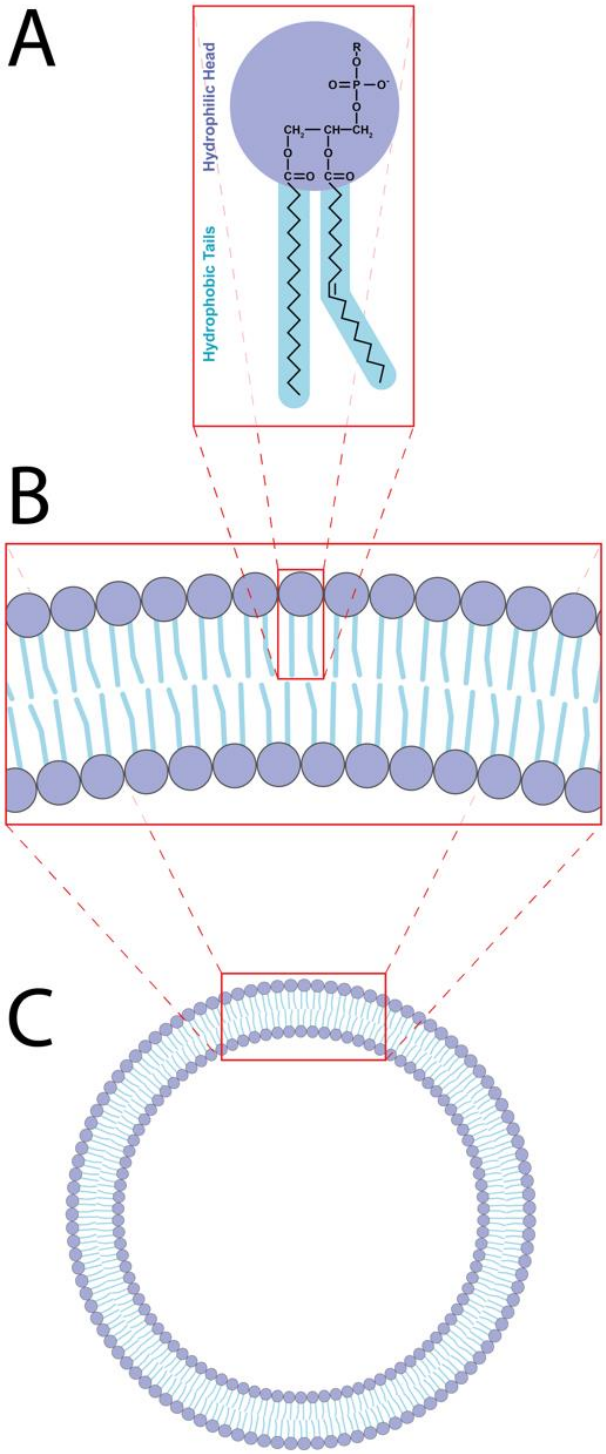


Figure 2. The Structure of Phospholipid Membranes (A) Structure of a phospholipid. The hydrophilic moiety is depicted in purple, the hydrophobic moiety in light blue. (B) Representation of phospholipids assembled to form a bilayer. (C) The phospholipid bilayer, rounded up to form a hollow spherical compartment.

1.3 Proteins as Functional Modules in Artificial Cells

Proteins are nanoscopic machines, which participate in virtually every process within living cells and therefore perform an incredibly diverse range of functions. Enzymes, for example, are proteins, which accelerate chemical reactions, necessary for most metabolic processes in cells. Structural proteins are involved in cell motility and cell shape maintenance by forming a cytoskeleton. Other proteins are responsible for immune

responses, cargo transport, cell adhesion or cell-cell communication to name just a few examples. Each group of proteins is further subdivided into subgroups, consisting of many individual members, which often perform one function only. Reaction cascades within cells are therefore usually performed by a series of proteins, each specialized to perform one step only. Consequentially, evolution has developed an immense number of proteins.
Due to their diverse properties, proteins are highly interesting building blocks for equipping artificial cells with specific functionalities. Although artificial alternatives have been developed[28], they are not equivalent substitutes for natural proteins. Hence, most artificial cells are functionalized with natural proteins, which can be further modified using genetic engineering to precisely match the intended use.

1.3.1 The Structure of Proteins

Proteins are large biomolecules, consisting of long amino acid chains. All proteinogenic amino acids share structural features, such as the central carbon atom Cα, which is connected to an amino group, a carboxyl group and a side chain (Figure 3A). The side chain is variable among amino acids, resulting in a total of 20 different amino acids found in almost all organisms. The side chains and thus also the amino acids differ in size, shape, charge and polarity, which affects the properties of the protein they are a part of. To form a protein, amino acids are chemically linked into a linear chain, called a polypeptide. The covalent bond connecting the amino group of one amino acid to the carboxyl group of the next amino acid is called peptide bond. The end of the linear amino acid chain, with the free amino group is referred to as N-terminus, the opposite end containing the free carboxyl group is referred to as C-terminus respectively (Figure 3B).
Peptide bonds exist in trans or cis conformation with respect to Cα, of which trans is much more abundant due to steric hindrance preventing the cis conformation[29]. However, the bonds between Cα and the adjacent amino group as well as between Cα and the carbonyl group are permitted to rotate. Thus, each amino acid can influence the overall structure of the protein along mainly two degrees of freedom. Since most proteins consist of hundreds of amino acids, this quickly adds up to a virtually endless number of possible structures a polypeptide can adopt. However, most proteins fold into unique structures, known as native conformations, which are specific to the functions performed by the proteins upon completion of their folding processes. For example, the active site of an enzyme must be precisely tailored to the respective substrate in order to exclusively bind to it and to prevent off-target interactions.
Protein structures are divided into four levels: primary, secondary, tertiary and quaternary structure. The primary structure refers to the amino acid sequence making up the polypeptide chain (Figure 3C), which is determined by the gene encoding the protein. Secondary structures are local structures formed by hydrogen bonding between carbonyl and amino groups of amino acids, which often display regularly repeating elements. The most common secondary structures are α-helices (Figure 3D), β-sheets (Figure 3E), as well as loops and turns. The formation of these structural elements or rather the absence of alternative secondary structures is partially explained by the limited range of compatible torsion angles in the protein backbone[30]. The tertiary structure refers to the relative spatial arrangement of secondary structural elements, which gives the protein its overall shape (Figure 3F). Tertiary structures are stabilized by a multitude of non-local interactions among amino acid side chains, which are mainly non-covalent. One exception is the formation of covalent disulfide bonds, which lock two parts of the polypeptide to each other. Another important factor stabilizing many tertiary

structures is the formation of a hydrophobic core, where amino acids with hydrophobic side chains assemble to prevent interactions with water molecules surrounding the protein. Tertiary structures are further influenced by external factors, such as the plasma membrane or the presence of specific substrates. Finally, the quaternary structure refers to a structure formed by several proteins assembled in a protein complex (Figure 3G). A protein, which is part of a supramolecular complex is called a protein subunit. However, not all proteins form complexes. Thus, not all proteins display a quaternary structure.
Moreover, proteins are not static, but undergo constant conformational changes, allowing them to shift between several states, which is necessary for most proteins to perform their function[31,32]. The interaction with other biomolecules or proteins often triggers large conformational changes in the protein. Simultaneously smaller deviations are constantly induced by thermal vibrations. Due to the highly complex nature of protein dynamics, movements within proteins cover a broad range of timescales, from sub-angstrom vibrations of covalent bonds in the femtoseconds, to collective movements of whole protein domains at microseconds timescale[33].

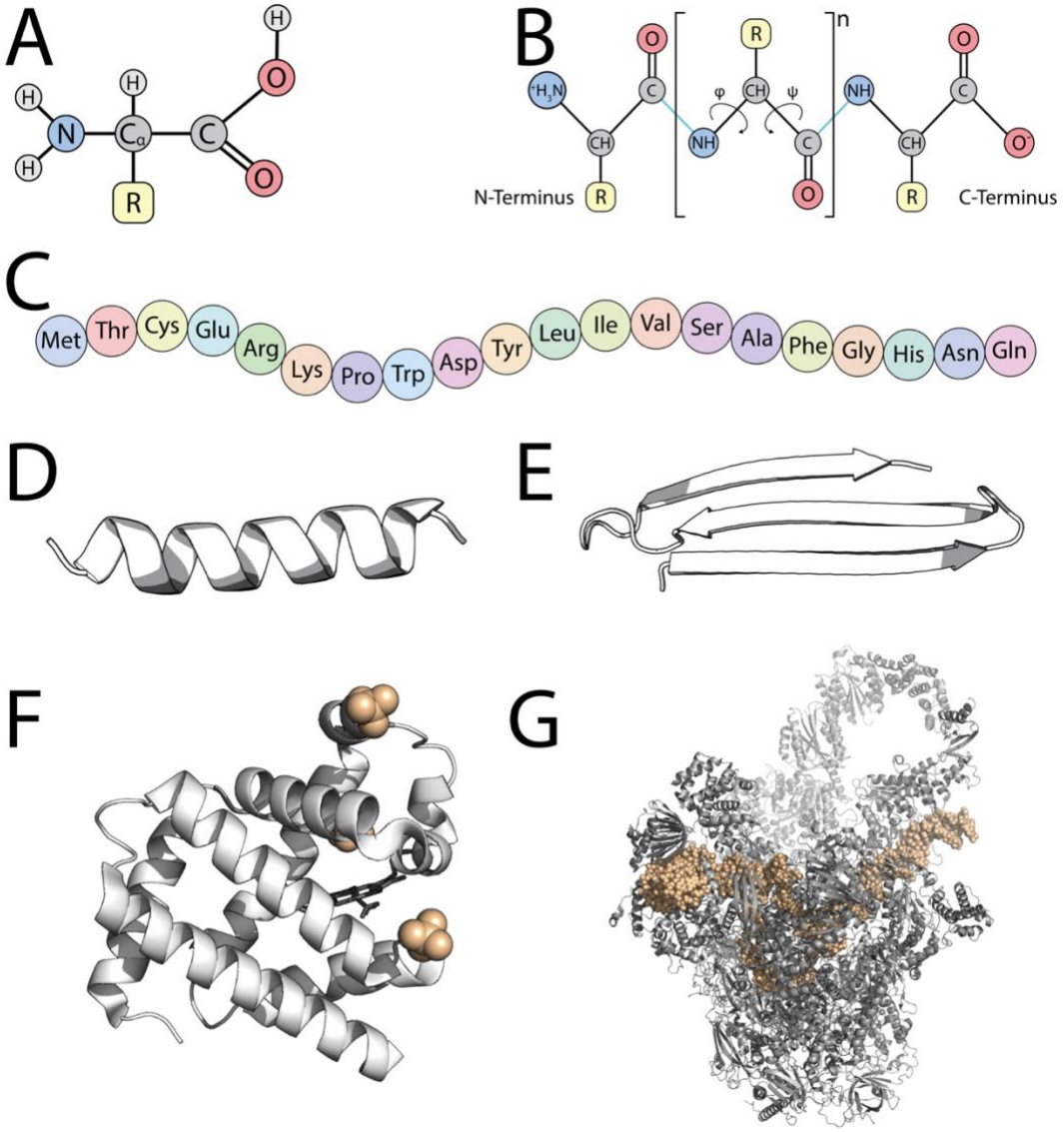


Figure 3. Amino Acids, Polypeptides and Proteins – From Primary to Quaternary Structure. (A) Structure of a proteinogenic amino acid. The nitrogen atom of the amino group is indicated in blue, the oxygen atoms of the carboxyl group in red. The variable side chain is depicted in yellow. (B) Structure of a polypeptide. The N-

terminus with the free amino group is shown on the left, the C-terminus exposing the free carboxyl group on the right. The peptide bonds are indicated in light blue. The two main torsion angles are indicated as φ and ψ in the central amino acid unit, which represents the array of amino acids, spanning the polypeptide chain between the two termini. (C) Primary structure of an exemplary polypeptide chain, containing all 20 proteinogenic amino acids, depicted in the three-letter code. (D-E) Cartoon representations of secondary structural elements α-helix (D) and β-sheets (E). (F) Tertiary structure of human myoglobin (PDB: 3RGK)[42]. (G) Quaternary structure of the human RNA polymerase, a multiprotein complex containing 27 subunits (PDB: 5IY8)[43].

1.3.2 Diversity of Protein Architectures

A common type of protein architecture is represented by globular proteins, which adopt a globe-like shape as implied by their name. Their linear polypeptide chain collapses into a somewhat spherical shape to protect hydrophobic residues in the core of the protein[18]. This exposes the hydrophilic residues on the protein surface, which renders globular proteins water soluble. Globular proteins can act as enzymes[34], messengers[35] or regulatory proteins[36]. Another type of protein architecture is adopted by fibrous proteins, which have a lower solubility than globular proteins and form long and thin structures. They consist of repetitive units assembled into supramolecular complexes crosslinked to each other and primarily perform structural roles to support cells biomechanically[37,38]. The third type of protein architecture is defined by intrinsically disordered proteins, which range from completely unstructured to partially structured[39]. They can convert into a structured state upon binding to another molecule[40] or remain flexible and act as hinge-like linkers in larger protein complexes[41]. The fourth common type of protein architecture is represented by membrane proteins, which will be discussed in detail in the following chapter.

1.3.3 Membrane Proteins

Membrane proteins are a class of proteins, which associate with biological membranes and equip them with essential properties, required for cell functionality. They are categorized into peripheral and integral membrane proteins. Peripheral membrane proteins are merely transiently associated with one leaflet of the phospholipid bilayer via electrostatic interactions, by burying hydrophobic loop regions into the membrane or by aligning amphipathic domains with the interface between phospholipid head- and tail-groups[44,45]. Alternatively, they can interact with other, more permanently associated membrane proteins[46], or bind covalently to a lipid, which they insert into the membrane to establish anchorage[47–49]. Integral membrane proteins are permanently associated with the membrane during their lifetime. They can only be separated from the membrane by either disturbing membrane or protein integrity. While integral monotopic membrane proteins are permanently attached to one leaflet of the membrane and do not span across the bilayer, integral bitopic membrane proteins span the membrane once and integral polytopic membrane proteins span the bilayer multiple times[45,50]. The latter two are referred to as transmembrane proteins.

Transmembrane proteins evolved to adapt to the very specific nature of the membrane environment. In order to be stably integrated into the membrane, transmembrane proteins expose apolar amino acid residues at the interface with the apolar phospholipid tails, resulting in a belt-shaped hydrophobic region[51,52]. Another feature specific to transmembrane proteins is the higher-than-average abundance of aromatic residues found at the interface between the membrane and the aqueous solution, which potentially favors interactions with lipid head groups[51].

The absence of water molecules in the hydrophobic membrane core, requires the protein backbone to form hydrogen-bonds with itself, which is only fully satisfied in the α-helical (Figure 4A) or β-sheet (Figure 4B) secondary protein structures[51], leading to the two main structural classes of transmembrane proteins: α-helix bundles and β-barrel proteins (Figure 4C)[53]. α-helical membrane proteins are far more abundant than β-barrels. They are present in most types of biological membranes and it was found that 20-30% of all genes are encoding the typical α-helix-bundle motif[54]. To traverse the hydrophobic membrane core with a width of 3 nm, transmembrane α-helices are typically comprised of 20 amino acids or more[55] whereas β-strands within membrane spanning β-barrel proteins generally contain 10 or more amino acids to fully traverse the hydrophobic core of the membrane[56]. In contrast to α-helical membrane proteins, transmembrane β-barrels are only found in outer membranes of Gram-negative bacteria, or organelles like mitochondria and chloroplasts and are thus far less abundant[54,56].

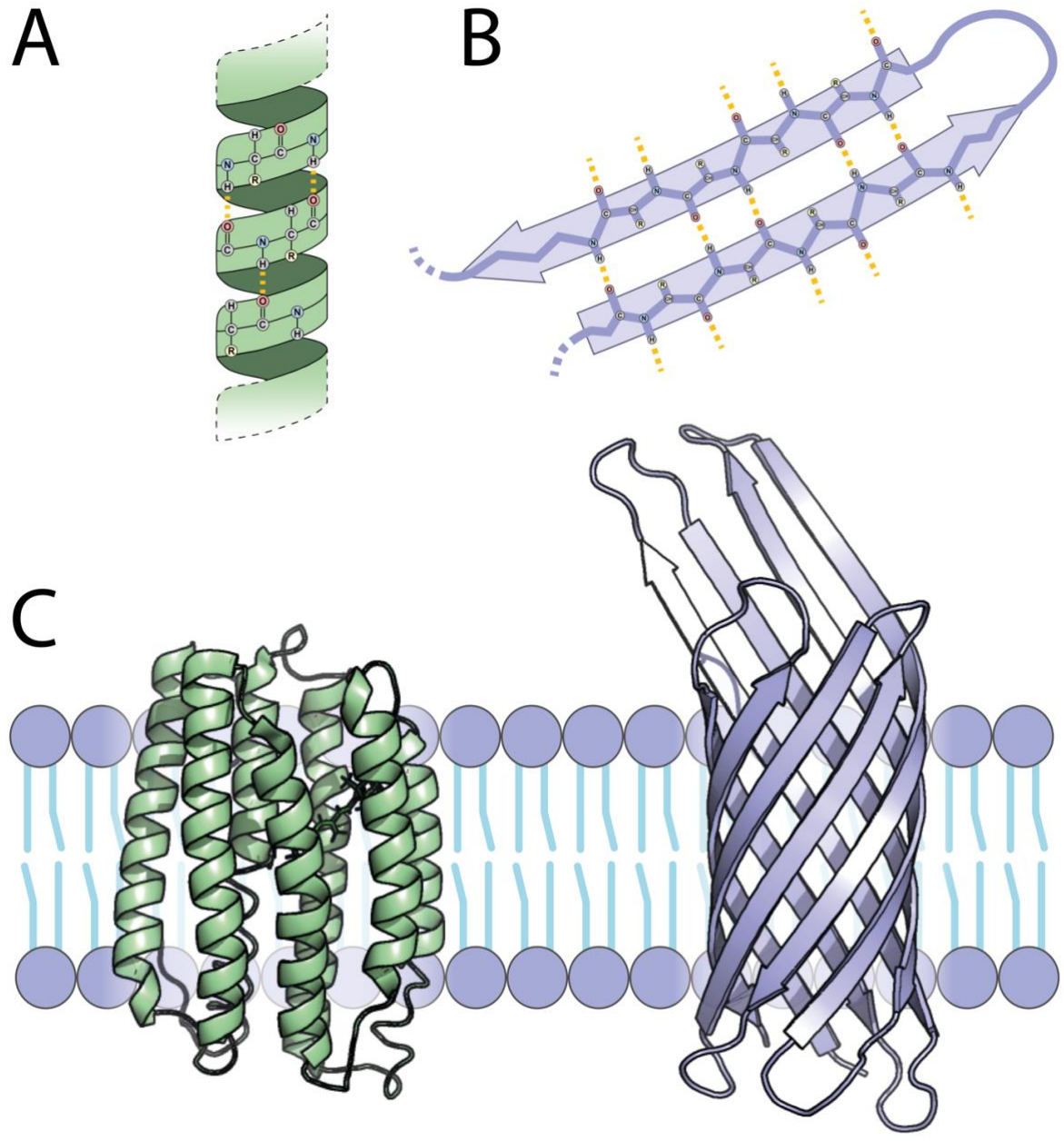


Figure 4. α-helical and β-barrel Membrane Proteins. (A-B) Cartoon representations of a transmembrane α-helix in (A) and of transmembrane β-sheets in (B). The yellow dotted lines indicate backbone hydrogen bonds. (C) Representation of a phospholipid bilayer, containing two transmembrane proteins. On the left in green: green-light absorbing proteorhodopsin (PDB: 2L6X)[57], a polytopic α-helical membrane protein. On the right in blue: outer membrane protein X (PDB: 1QJ8)[58], a polytopic β-barrel membrane protein.

1.3.4 Membrane Protein Functions

Compared to soluble proteins, membrane proteins display rather limited structural variability, due to the specific requirements of the membrane environment. Nevertheless, they perform vastly different tasks to functionally connect cells to their surroundings. Among the most abundant functions of membrane proteins is the translocation of substrates, such as ions[59], small molecules[2] or even macromolecules, *e.g.*, other proteins[60,61] with high specificity across the membrane. The transport processes are either passive, from high to low concentration, or active, against the concentration gradient, which requires an energy source.

Channel proteins transport water or ions passively along the concentration gradient by facilitated diffusion[62]. That is, they form a tube-like opening spanning the membrane, with an internal geometry to exclusively permeate specific substrates[63]. Some channels are permanently open while others are gated by chemical or electrical signals[64]. Larger molecules cannot pass through the narrow openings in channel proteins and thus need to be translocated by carrier proteins, which have specific substrate binding sites instead of continuous channels across the membrane[65]. Upon substrate binding, carrier proteins undergo conformational changes to expose and release the substrate on the opposite site of the membrane[66]. Transport is achieved by repeated transitions between those conformational states. Some carrier proteins can actively transport substrates uphill (against the concentration gradient), by coupling it to the downhill transport (along the concentration gradient) of another substrate[67]. While uniporters transport only one substrate, symporters transport two substrates along the same direction and antiporters transport two substrates in opposing directions.

Aside from coupled carriers, uphill transport of substrate can be achieved by proteins which actively pump substrate across the membrane upon energy consumption. Thereby this energy can be supplied by the hydrolysis of adenosine triphosphate (ATP)[68] or by converting the energy of light, as found in microbial rhodopsins, which are found in microorganisms belonging to all the three domains of life[69]. They are involved in sensing light (phototaxis)[70] and conversion of light to an ion gradient across the membrane they reside in[71]. They consist of seven transmembrane α-helices and a retinal molecule, which isomerizes upon absorbing a photon of a specific wavelength[72]. This induces a series of conformational rearrangements within the protein and ultimately results in the net flux of one ion across the membrane per absorbed photon[73,74]. The most prominent microbial rhodopsin is bacteriorhodopsin (BR), which was discovered in 1971 (ref. [75]). It is found in the cell membrane of *Halobacterium salinarum*, where it is densely packed to form two dimensional crystalline arrangements. The light absorbing properties of BR result in a strong purple coloration, which is why the membrane is typically referred to as purple membrane. BR exports protons upon illumination, which establishes a proton gradient across the membrane, which can be coupled to energize transport processes or converted to chemical energy in the form of ATP[76].

Similar to substrate transport processes, cells need to transfer signals across the plasma membrane to communicate with neighboring cells and to sense the external environment. Signal transduction typically involves binding of an extracellular ligand to

a transmembrane receptor, which causes conformational rearrangements within the receptor[77]. This triggers an intracellular signaling cascade, which ultimately affects the behavior of the cell[3]. Therefore, membrane receptors allow cells to sense and react to extracellular stimuli such as hormones[78], nutrients[79] or neurotransmitters[80].
Besides substrate transport and signal transduction across the plasma membrane, membrane proteins are involved in many other processes, such as cell adhesion[81] or enzymatic reactions[82]. Their essential roles in many vital processes, paired with the facilitated accessibility on the cell surface renders membrane proteins relevant drug targets[83]. As a result, ≈60% of all protein drugs target membrane proteins[84], mainly ion channels and G-protein coupled receptors (GPCRs)[85].

1.3.5 Biogenesis of Membrane Proteins

Membrane proteins have to be inserted into the membrane to adopt a functional state. The insertion process relies on a series of complex steps and involves multiple helper proteins, such as chaperones, translocons and insertases. Furthermore, α-helical membrane proteins undergo a different pathway than β-barrel proteins to arrive in their respective membrane-environments in a correctly folded and functional state. Hence, to exemplify both processes, this chapter discusses protein biogenesis and membrane protein insertion into the inner membrane (IM) and the outer membrane (OM) of *Escherichia Coli* (Figure 5).
Biogenesis of all proteins, soluble and membrane-associated, starts at the genetic level in the cytosol, which is the space enclosed within the IM. Upon transcription of a gene encoding a protein into the form of messenger RNA (mRNA), it is passed to the ribosome, a complex comprised of proteins and ribosomal RNA (rRNA)[86]. The ribosome then translates the information contained on the mRNA to assemble individual amino acids into the nascent polypeptide chain. If proteins are targeted towards a membrane or the periplasm (aqueous space between IM and OM), they are not directly folded in the cytosol upon exiting the ribosome, but remain in an unfolded state to successfully traverse and/or insert into a membrane. Due to their amphipathic properties, membrane proteins in aqueous solution are prone to precipitation and hence have to be carefully guided to their destination by a concerted cascade of events[87]. This process is initiated when the first few amino acids of the nascent chain exit the ribosome. The presence or absence of an N-terminal signal sequence dictates whether the nascent polypeptide chain will be transferred to the SecYEG translocon, which is a protein complex in the IM that assists either the folding of proteins into the inner membrane or their translocation into the periplasm[87]. Depending on the hydrophobicity of the signal sequence, the polypeptide chain is either post- or co-translationally directed to the translocon (Figure 5)[88].
α-helical IM proteins are primarily targeted to the signal recognition particle (SRP)-dependent pathway, which is usually co-translational[89,90]. The SRP is a protein-RNA complex that binds the ribosome close to the peptide tunnel, where the nascent chain emerges[91]. There the SRP scans the newly translated nascent chain for specific hydrophobicity patterns present in most IM proteins, but also for signal sequences of some secretory proteins and membrane associated proteins[92,93]. If the specific pattern is recognized, and SRP is bound to the ribosome-associated nascent chain, the ribosome-SRP complex is delivered to FtsY, a SecYEG-bound SRP receptor[94]. Upon contact with FtsY, the translating ribosome binds to the SecYEG translocon and continuous translation inserts the nascent chain into the translocon. A flexible opening in the SecY subunit of SecYEG acts as a lateral gate, and releases the hydrophobic α-helices of IM proteins into

the membrane, where they protein adopts a functional state[87]. Less complex membrane proteins can be inserted into the IM by the insertase YidC[95], independent of SecYEG.
Alternatively, proteins can reach the membrane through the post-translational pathway, in which the nascent chain is synthesized in the cytosol prior to contact with translocons or insertases. Soluble chaperones, such as trigger factor[96] and SecB[97] protect the nascent chain from precipitation and deliver it to the SecYEG translocon in a translocation competent state. There it is pushed into the translocon by SecA in a process driven by ATP-hydrolysis[98]. Proteins targeted to the periplasm or the OM are guided through the translocon and released into the periplasmic space, while their N-terminal signal sequence is cleaved off by a signal peptidase during the translocation process[99]. Soluble periplasmic proteins arrive in the periplasm in an unfolded and inactive state and need to be guided towards adopting their correctly folded and functional state by periplasmic chaperones[100]. Outer membrane proteins (OMPs) need to further traverse the periplasmic space to arrive at the OM. Therefore, soluble chaperones, such as Skp and SurA bind to the nascent chain upon translocation to prevent aggregation and to maintain folding competence[100,101]. The protein is then delivered to the OM in an unfolded, chaperone-stabilized state, where it is inserted and folded into the membrane by the β-barrel assembly machinery (BAM) complex.
The BAM complex is an insertase, found in the OM of *E. coli*, where it facilitates the folding of integral β-barrel OMPs into the membrane. The complex is comprised of five subunits, BamA-BamE, of which BamA is the central component[102–104]. BamA consists of five N-terminal soluble domains, termed polypeptide transport (POTRA) domains and a C-terminal transmembrane domain[105,106]. The membrane embedded region consists of a 16 β-stranded barrel, which establishes a loose connection between the first and the last β-strand[107]. The β-barrel seam is thus structurally more dynamic than that of other β-barrel OMPs. Moreover, the hydrophobic belt of BamA at the β-barrel seam is too narrow to completely span across the OM.
Although to date, the molecular mechanism of BAM-mediated OMP insertion is not fully understood, most models agree on the relevance of the atypical β-barrel seam[108]. The BamA-assisted model proposes, that the narrow BamA β-barrel seam locally disturbs the membrane integrity due to a hydrophobic mismatch, which facilitates the insertion of partially folded OMPs into the membrane[109]. In contrast, the budding model suggests the formation of a transient hybrid β-barrel, consisting of BamA and OMP-substrate, which releases the OMP through the β-barrel seam into the membrane. In this model, the dynamic and mechanically flexible BamA β-barrel seam functions as a lateral gate[110]. Interestingly, whereas sealing the lateral gate by linking the first to the last β-strand inhibits BamA *in vivo*[110], BamA-mediated OMP insertion *in vitro* remains unaffected[111]. However, despite contradicting findings, recent structures of folding intermediates support the role of the β-barrel seam as a lateral gate[112–114].
Compared to the SecYEG translocon in the IM, the BAM complex in the OM has no access to energy sources of the cell, such as ATP or membrane potentials. Consequentially, all steps past IM translocation, namely binding to periplasmic chaperones, transfer to BamA and folding into the OM need to be energetically favorable. Thus, the free energy of the pre-proteins is proposed to strictly decrease along the journey from the IM to and into the OM, acting as an energy sink[115]. However, to date, the exact mechanism underlying this process remains unknown.

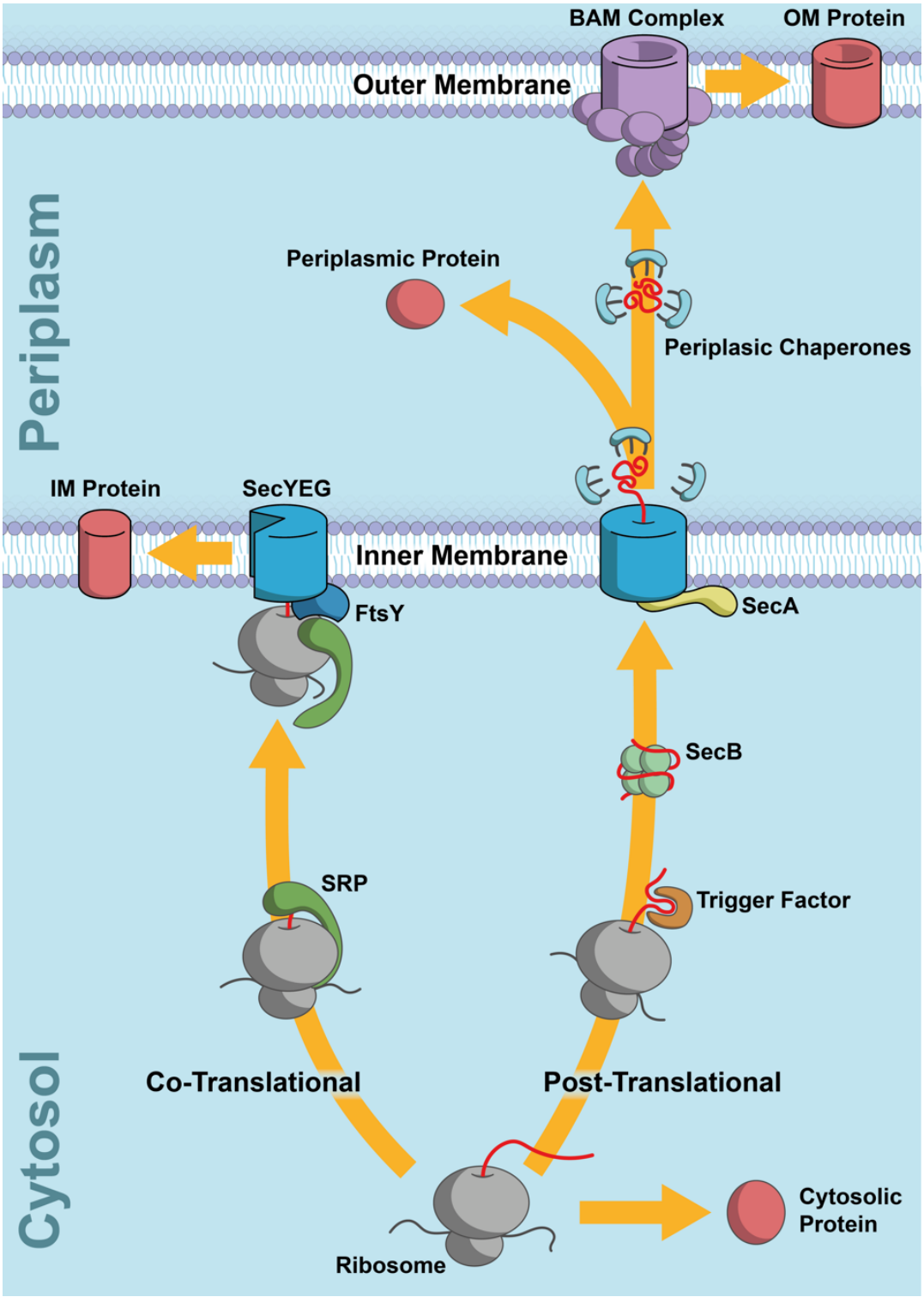


Figure 5. Protein Biogenesis. Overview over the processes involved in protein biogenesis in the cytosol, the IM, the periplasm and the OM. The ribosome is represented in grey, translating the nascent polypeptide chain (red string) upon reading the mRNA (black). Soluble cytosolic proteins fold directly after translation to adopt their native conformation. Membrane proteins and periplasmic proteins undergo either the co-translational (left) or the post-translational (right) pathway. SecYEG (blue) inserts proteins into the IM or translocates them to the periplasm. Periplasmic soluble proteins adopt their final structure in the periplasm. OMPs are transported to the BAM complex (purple) by periplasmic chaperones (light blue), where they are inserted into the OM.

1.4 The Assembly of Artificial Cells

1.4.1 Protein Expression and Purification

To integrate proteins into artificial cells, it is required to produce them in sufficient amounts and isolate them from the surrounding cellular components. To this end, different types of cell-based expression systems, such as bacterial, yeast, insect or mammalian cells, have been highly optimized to yield the maximal amount of the protein of interest[116]. While different types of expression systems stand out with different properties, genetically altered *E. coli*, optimized for protein overexpression, are most commonly used. The handling of *E. coli* is comparably simple, due to efficient tools to introduce genetic modifications, exponential growth and possibilities for long-term storage by freezing. After providing the bacteria with plasmid DNA encoding the protein of interest, protein expression can be initiated and tightly regulated.

After successful overexpression of the protein of interest, the bacteria are harvested by centrifugation and subsequently lysed to access and extract their content (Figure 6). Further centrifugation steps allow the separation of soluble and membrane fraction. Latter has to be solubilized for the purification of membrane proteins, as will be discussed below. Isolating the protein of interest from the remaining cellular content is achieved with the help of an affinity tag, which is a sequence of amino acids fused to the protein of interest, designed to bind specifically to an immobilized counterpart[117]. Already on the plasmid level, the DNA sequence coding for the affinity tag is fused to the gene encoding the protein, so that protein and tag are expressed as one continuous polypeptide chain. Upon binding of the tag, which immobilizes the protein of interest, the other proteins are removed by washing of the sample. Subsequent elution of the tagged protein allows efficient isolation from the remaining cellular components. Although decent purity can be achieved in this way, many applications require higher sample purity[118]. Thus, additional steps can be performed, to further increase the quality of the sample e.g., by separating the remaining proteins by charge, size or surface hydrophobicity and by removing the affinity tag by proteolytic cleavage[119].

The purification of membrane proteins is generally more laborious than that of soluble proteins, mainly because membrane proteins have to be solubilized and extracted from the membrane to be isolated from other membrane-associated components[120]. Moreover, during the extraction and purification, the hydrophobic belt of the membrane protein has to be carefully protected from the surrounding aqueous solution to prevent denaturation and aggregation. Due to their capabilities of disrupting lipid bilayers and stabilizing exposed hydrophobic surface regions of membrane proteins, detergents are well suited for this task[121]. Detergents, consist of polar and apolar parts[122] and usually have structures similar to lipids, yet they comprise relevant differences. While the lipid head-groups cover approximately the same area as spanned by their tails, detergent head-groups are bulkier than their tail domains, resulting in a cone-like shape as compared to the cylindrical shape adopted by lipids[123]. This directly affects the shape of the supramolecular structures formed by the hydrophobic effect. Instead of planar bilayers, detergents in aqueous solution assemble into spherical structures called micelles, in which the apolar tail domains are protected from the surrounding aqueous solution. However, micelles are dynamic structures and in constant exchange with detergent monomers in the surrounding solution. Detergents can also self-insert into and dissociate from the plasma membrane, where they intercalate the lipid molecules and disrupt the highly ordered bilayer structure. Hence, at sufficiently high concentration, detergents completely solubilize the bilayer, which results in a ternary solution comprised of detergents, lipids and membrane proteins[124,125]. Under optimal conditions, the detergent

molecules form a torus-shaped structure around the hydrophobic belt of the membrane protein of interest, which solubilizes and simultaneously preserves it in the native conformation[126]. After successful solubilization, membrane proteins can be purified as explained for soluble proteins (*e.g.*, with affinity tags), as long as the detergent concentration is constantly kept in an optimal range[127].

1.4.2 Detergent-Mediated Membrane Protein Reconstitution

Whereas purified soluble proteins can be readily encapsulated into artificial compartments, membrane proteins have to be carefully embedded into their new membrane environment. This process of reinserting membrane proteins into artificial membranes is referred to as membrane protein reconstitution (Figure 6). It is particularly relevant for membrane proteins involved in transport and signaling processes across the membrane, which is often referred to as vectorial activity[74]. In the absence of compartmentalization, *e.g.*, in a detergent-stabilized state, transport processes result in no net flux and can therefore not be measured.

In theory, membrane protein reconstitution into artificial membranes can be regarded as the reversed counterpart to membrane solubilization. Briefly, lipids are added to the detergent-stabilized proteins and the detergents are removed thereafter in order to force the protein into a newly formed lipid bilayer. Although there are different means of membrane protein reconstitution, detergent-mediated reconstitution is the most commonly used strategy[128]. The first step in the reconstitution process is to form artificial membranes. To this end, lipid or polymer vesicles, which consist of bilayers rounded up to form a hollow spherical compartment, similar to an empty cell, are prepared and mixed with detergents and membrane proteins. This destabilizes the vesicles to fully or partially disintegrate and allows dynamic exchange of detergents, lipids and proteins. Upon removal of detergent molecules, the lipid molecules can reassemble into vesicle structures and under optimal conditions, the membrane proteins end up embedded in the newly formed vesicle membrane[129]. The resulting structures are referred to as proteoliposomes.

During the protein reconstitution process, many conditions need to be carefully optimized[130] to successfully embed membrane proteins into their new membrane environment. If the chosen conditions are incompatible with the membrane proteins, they might adopt misfolded states and become functionally impaired. Therefore, after membrane protein reconstitution it is important to either experimentally verify membrane protein functionality or to probe whether they adopted a correctly folded state. The latter can be achieved by employing atomic force microscopy.

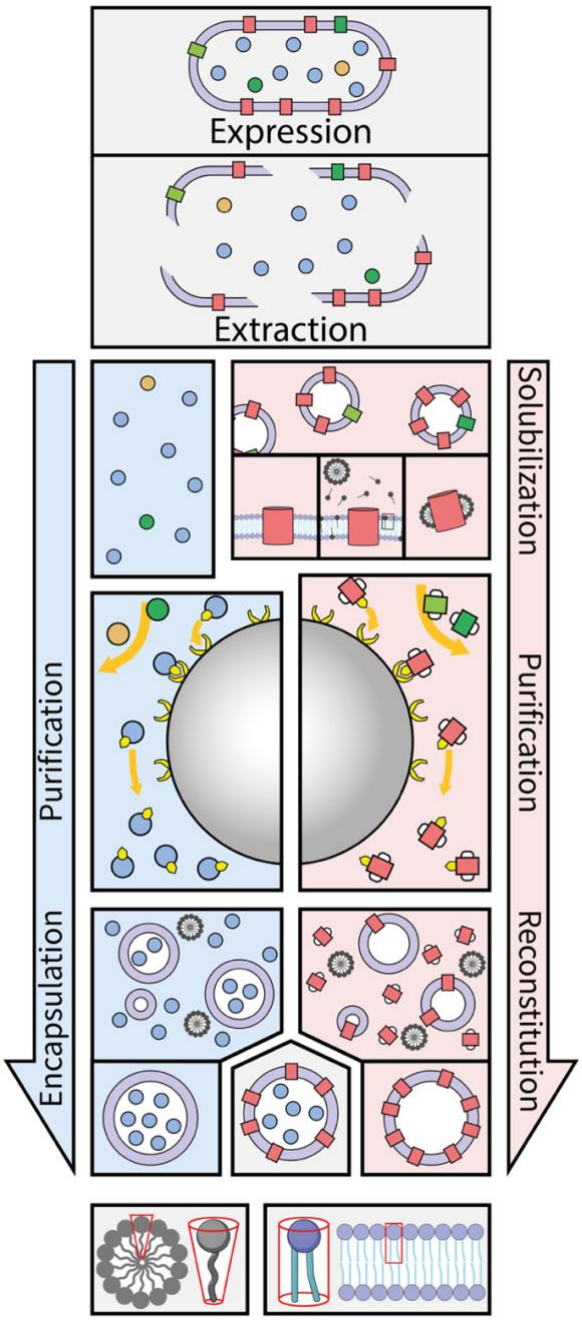


Figure 6. Protein Purification and Reconstitution. Overview of the production, purification and reconstitution processes of soluble and membrane proteins. From the top: simplified representation of an *E. coli* bacterium overexpressing membrane proteins (red) and soluble proteins (blue), together with other proteins (contaminants shown in different colors). The bacterial cells are lysed to access their content. The soluble fraction (blue background) is separated from the membrane fraction (red background). The membrane fraction is solubilized by detergents for further purification. The protein of interest is immobilized (*e.g.*, on beads, as indicated by the large grey sphere) via an affinity tag. Untagged proteins are removed by washing. The tagged protein is then eluted from the beads. Purified soluble and membrane proteins can be encapsulated or reconstituted in lipid vesicles respectively. The two grey bottom panels highlight the geometric differences between detergents (left) and lipids (right) and their effects on the shape of supramolecular assemblies. The cone-shaped detergents assemble into spherical micelles, the cylindrical lipids assemble into bilayers.

1.4.3 AFM-Based SMFS of Membrane Proteins

The atomic force microscope (AFM) is a valuable tool for the study of membrane proteins, especially in the context of folding and membrane insertion. It was developed in 1986 by Gerd Binnig, Calvin Quate, and Christopher Gerber and is based on the working principles of the scanning tunneling microscope (STM)[131,132]. However, unlike the STM, the AFM is not dependent on conductive surfaces, but is capable of analyzing insulating materials. The AFM is capable to operate in liquid environments, which renders the AFM a suitable tool for the study of biological samples under physiological conditions[133].

The working principle of the AFM is centered around a micro-cantilever with a sharp stylus, which is brought into direct physical contact with the sample (Figure 7A). Upon interaction with the sample, the cantilever deflects either towards or away from the sample, depending whether the interactions are attractive or repulsive. The deflection of the cantilever is measured by reflecting a laser beam from the backside of the cantilever

via a mirror onto a segmented photodetector. By employing piezo-electric actuators, the position of the cantilever relative to the sample can be precisely controlled in all three dimensions. Thus, by moving the cantilever along the sample, the AFM can record a topograph of the sample by raster-scanning the sample surface. Thereby, the cantilever deflection is monitored and the information is transferred to ultimately control the piezo-electric actuators. This closes a feedback loop, as controlling the distance between cantilever and sample directly affects the cantilever deflection. This feedback system can be employed to maintain a setpoint deflection, by adapting the height of the cantilever. As a result, both the piezo-height and the cantilever deflection at each sampled point are recorded to generate the final topograph of the sample surface with sub-nanometer resolution.

The most striking difference between the AFM and other microscopy methods, is the ability to mechanically interact with the sample. Additionally, upon calibration[134] the applied force can be calculated from the cantilever deflection, by approximating the cantilever as a Hookean spring. This allows to record high resolution topographs while simultaneously mapping physical, chemical and biological interactions[135], to study cell adhesion[136] and interactions between viruses and cells[137] to name just a few examples.

Moreover, the high sensitivity of the AFM allows to study interactions between the cantilever-stylus and single molecules. In particular, single molecule force spectroscopy (SMFS) studies the biophysical properties of single proteins by mechanically unfolding them[138]. In order to study proteins with SMFS, the protein needs to be immobilized. Conveniently, membrane proteins can be directly investigated by adsorbing the membrane in which they are embedded in to an atomically flat mica support[139]. There, the membrane forms a supported bilayer, that can be readily investigated with the AFM[140]. Typically for AFM studies, membrane proteins are purified and reconstituted into artificial membranes. However, some native membranes are suitable for AFM studies too, such as the purple membrane from *Halobacterium salinarum*[141], or protein-enriched outer membrane vesicles from *E. coli*[142].

Membrane patches containing protein-rich areas are localized by raster scanning the sample. Then, the cantilever tip is pushed onto the sample with a non-destructive force to establish an unspecific interaction between the cantilever tip and an individual membrane protein. This interaction is sufficient to fully extract the protein from the membrane upon retraction of the cantilever. During the extraction process, first the terminus attached to the cantilever-tip is fully extended, which is followed by the sequential unfolding of individual structural segments of the membrane protein until it is fully extracted (Figure 7B). During the unfolding process, the height of the z-piezo as well as the applied forces are recorded. The distance covered by the z-piezo is corrected by accounting for vertical tip-displacement caused by the cantilever deflection, resulting in the actual tip-sample-separation. For representation, typically, the applied force is plotted *versus* the vertical tip-sample-separation in a force-distance (FD) curve (Figure 7B).

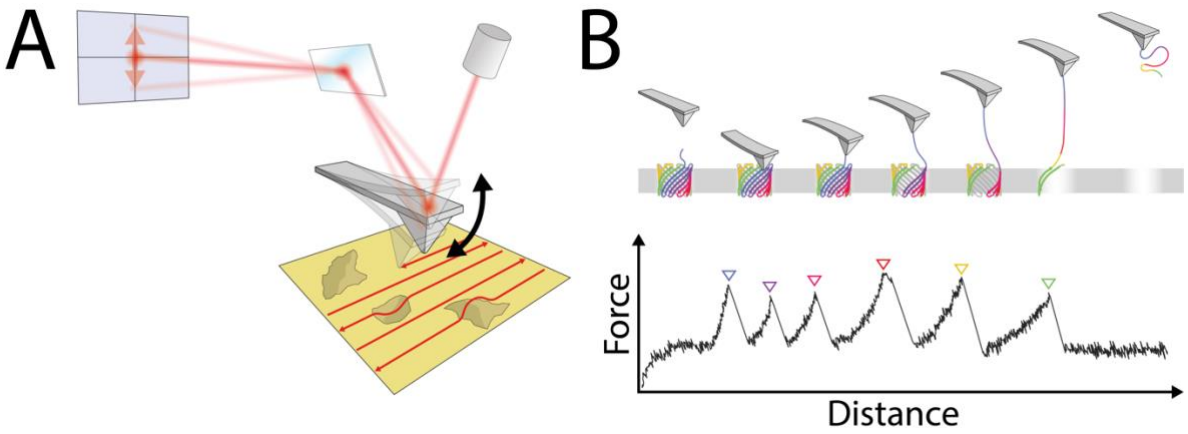


Figure 7. The Atomic Force Microscope and Single-Molecule Force Spectroscopy. (A) Schematic representation of the working principles of the AFM. The AFM stylus is brought in contact with the underlying sample, which causes the cantilever to deflect. The deflection is amplified by a laser beam, reflected from the back of the cantilever via a mirror onto a segmented photodiode. A topograph of the sample is recorded by raster-scanning the sample surface. (B) Top: cartoon representing the stepwise unfolding of a single membrane protein by AFM based SMFS. The cantilever-tip is pushed onto the surface to establish an interaction with the protein and is subsequently retracted to sequentially unfold stable structural segments from the membrane. The forces required to unfold structural segments are recorded and typically plotted *versus* the distance between the apex of the cantilever stylus and the sample surface in an FD-curve. Bottom: exemplary FD-curve, displaying force peaks, each corresponding to the unfolding of an individual structural segment. The saw-tooth like pattern, is characteristic to the unfolding of membrane proteins.

So far a variety of α-helical and β-barrel membrane proteins have been mechanically unfolded with SMFS[143–146]. Their unfolding typically results in FD curves displaying a characteristic saw-tooth like pattern with each force peak corresponding to the unfolding of an individual stable structural segment. Due to variations in structural and mechanical properties, different proteins display unique unfolding patterns. To identify the common unfolding behavior of a protein-type, many proteins of the same type are unfolded and their FD curves are superimposed. For a meaningful interpretation of the acquired data, it is important to correlate the recurring force peaks with structural segments of the protein. To this end the worm-like chain (WLC) model is fitted to each force peak in order to describe the stretching behavior of the unfolded polypeptide chain. This reveals the contour length of each force peak, that is the length of the unfolded part of the protein tethered between the cantilever and the sample at full extension. The contour lengths of individual force peaks in nm can be converted to amino acids, assuming an average length of 0.36 nm per amino acid. This precisely maps force peaks and thus the mechanical stability to secondary structural elements[147]. Therefore, the effects of mutations[148] or external factors on the mechanical stability of secondary structural elements can be investigated[147]. Thus, it has been studied how the pH[149], small molecules[150] or the temperature[151] modulate the intricate networks of interactions stabilizing membrane proteins.

Furthermore, SMFS can be operated in the dynamic mode, referred to as dynamic force spectroscopy (DFS). In DFS the protein is mechanically unfolded at varying speeds to probe the dependence on the loading rate (*i.e.*, force applied over time). Upon determining the most probable rupture force and loading rate of each structural segment, parameters describing the free-energy landscape can be approximated by applying the Bell-Evans model[152–154]. This contours the free-energy minimum hosting the folded state of the protein and maps kinetic, energetic and mechanical properties to individual structural segments. More precisely, the width of the free-energy well, the height of the

free-energy barrier stabilizing the folded state against unfolding, the unfolding rate at equilibrium (*i.e.*, no force applied) and the mechanical flexibility can be extracted for each structural segment. DFS is thus suitable to complement static structural data and to study how different factors shape the free-energy landscape of proteins. It was used to study receptor-ligand bonds[155], the effects of substrate binding to the lactose permease LacY[156] and how the energy landscape of the GPCR Par1[157] is affected upon binding of different ligands.

Although, protein unfolding via classic SMFS and DFS describes the folded state of membrane proteins, it does not elucidate how membrane proteins are initially inserted and folded into the membrane. The complex process of membrane protein folding and insertion *in vivo* cannot be replaced by simply pushing the membrane protein back into the membrane with the AFM. Nevertheless, the folding of structural segments can be indirectly studied by means of SMFS. To this end, the protein is partially or fully extracted from the membrane as in conventional protein unfolding experiments. However, after the initial unfolding process, the cantilever-tip is reapproached into close proximity of the surface for a specific amount of time. Given the right conditions, the protein or parts of the protein refold into the membrane. A second retraction of the cantilever unfolds the refolded structural segments. Careful comparison of the initial unfolding curve to the FD curve corresponding to the unfolding of the refolded protein allows classification of the folded state adopted by the protein into partially or fully folded, misfolded or unfolded. Similar to conventional SMFS unfolding experiments, the full potential of re-folding experiments can be tapped by comparing different experimental conditions. Single molecule refolding experiments have been conducted to study the influence of periplasmic chaperones on OMP folding and membrane insertion[158] and to investigate the IM insertase YidC[159]. A detailed hands-on introduction to SMFS unfolding and refolding experiments is provided in chapter 2.

1.4.4 Bottom-Up *versus* Top-Down Approaches

As described above, membrane proteins need to be extracted from cell membranes, purified and reconstituted into artificial membranes, in order to be employed in artificial cells. There, the folded state of the protein can be investigated for example by directly measuring the activity or with AFM-based SMFS. Even though all of those methods have been used for decades to understand membrane protein functionality, detergent-mediated reconstitution in particular has re-gained popularity in recent years in the field of artificial cell engineering, due to advances in synthetic biology[160]. Important factors for this development, were strong advances in the fields of DNA and protein engineering. The process from *in silico* DNA design to protein expression *in vivo* has been streamlined, which, in theory, allows to tailor proteins to our needs, produce them with high yields and purity and finally reconstitute and combine them in novel arrangements into artificial membranes to create artificial cells. This would allow to couple the activity of different proteins to assemble complex reaction cascades. By further encapsulating soluble proteins inside the artificial cells, the combinations are virtually endless. Unfortunately, in practice, detergent-mediated membrane protein purification and reconstitution often depends on meticulous fine-tuning of many inter-dependent parameters and needs adaptation for each membrane protein, which is why it is often regarded as black magic[161]. In fact, the process of protein purification and reconstitution often poses a major bottleneck, even in the study of individual membrane proteins. Hence, detergent mediated co-reconstitution (Figure 8A) of multiple membrane protein species, although possible[162], becomes increasingly complex with every additional

component, as optimal parameters are rarely overlapping between two or more reconstitution protocols. Additional complicating factors are the control over stoichiometric ratios between reconstituted membrane proteins and membrane protein orientation. Latter is essential for functional integration of vectorial membrane proteins. However, approaches exist to assemble increasingly complex systems from the bottom up, like the fusion of vesicles after reconstitution (Figure 8B) or the sequential reconstitution of different membrane proteins (Figure 8C)[163]. An alternative, intriguing approach is the reconstitution of membrane protein insertases, which might be exploited to insert and assemble different membrane proteins in a highly controlled manner (Figure 8D)[164].

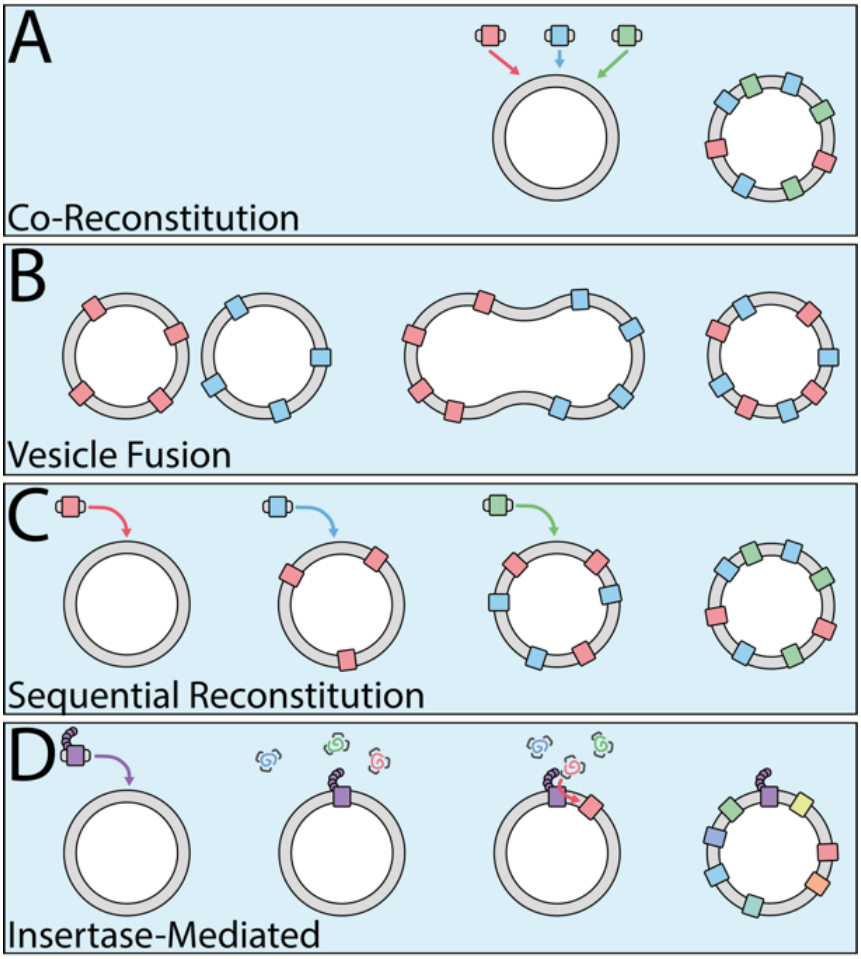


Figure 8. Bottom-Up Approaches for the Assembly of Multicomponent Artificial Cells. Schematic representations of different bottom-up strategies to assemble artificial cells. (A) Co-reconstitution of multiple membrane proteins into the same vesicle. (B) Separated reconstitution of multiple membrane proteins into individual vesicles, followed by their fusion. (C) Sequential reconstitution of multiple membrane proteins into the same vesicle. (D) Detergent-mediated insertion of an insertase into a vesicle, followed by insertase-mediated insertion of multiple chaperone stabilized membrane proteins into the same vesicle.

The construction of synthetic systems from individual components, such as reconstituted membrane proteins in artificial membrane compartments, is referred to as bottom-up assembly. Whereas bottom-up assembled systems can be constructed to precisely match the initial design, it remains challenging to reach high levels of complexity. An alternative strategy is to modify existing biological systems, which are inherently complex. This approach is referred to as top-down engineering and is the counterpart to bottom-up assembly[165]. In the top-down approach, natural cells are manipulated to fulfill new

functions. This is generally achieved by means of genetic engineering, which reprograms the cells. Upon altering their genetic information, the expression of specific proteins can be up or down regulated or completely knocked out. Alternatively, genes from other organisms or altered genes coding for engineered proteins can be introduced, to equip cells with completely new functions. Due to the highly efficient apparatus responsible for *in vivo* protein biogenesis, complex multicomponent systems are more feasible by employing the top-down approach as compared to the bottom-up approach. However, the machinery within living cells is too complex to fully understand, which often results in limited control over the engineered system with unforeseen side effects[166]. Moreover, not all functions are compatible with the host cell.
In summary, bottom-up as well as top-down approaches have inherent advantages and disadvantages. Therefore, hybrid approaches have emerged recently, merging bottom-up and top-down strategies with the intention to combine the strengths of the individual approaches[167–169].

1.5 Aim of the Thesis

This thesis explores membrane protein folding and insertion into natural and artificial membranes by means of insertases and detergents, with the goal to advance the methodology to assemble artificial cells from the bottom up. To this end AFM-based SMFS was employed to investigate the insertase BamA, the main component of the BAM complex. Moreover, two different light-driven proton pumps were modified and applied as energizing modules in simplistic artificial cells and in more complex systems. The presented protein constructs provide a tool to control the membrane protein orientation upon reconstitution, a common limitation in the process of artificial cell engineering. Furthermore, encapsulating *E. coli*, enriched with light-driven proton-pumps, in larger artificial cells, demonstrates the potential of combining bottom-up and top-down approaches. Although, the presented artificial cells are proof-of-concept, the demonstrated methods and protein-modules have a variety of potential applications and will hopefully be beneficial for future advances in the field.

1.6 References

1. Tetsch, L. & Jung, K. The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. *Mol. Microbiol.* **73**, 982–991 (2009).
2. Kimanius, D., Lindahl, E. & Andersson, M. Uptake dynamics in the Lactose permease (LacY) membrane protein transporter. *Sci. Rep.* **8**, 1–12 (2018).
3. Tuteja, N. Signaling through G protein coupled receptors. *Plant Signal. Behav.* **4**, 942–947 (2009).
4. Onaca, O., Enea, R., Hughes, D. W. & Meier, W. Stimuli-responsive polymersomes as nanocarriers for drug and gene delivery. *Macromol. Biosci.* **9**, 129–139 (2009).
5. Nardin, C., Thoeni, S., Widmer, J., Winterhalter, M. & Meier, W. Nanoreactors based on (polymerized) ABA-triblock copolymer vesicles. *Chem. Commun.* **15**, 1433–1434 (2000).
6. Lee, K. Y. *et al.* Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nat. Biotechnol.* **36**, 530–535 (2018).
7. Chen, A. H. & Silver, P. A. Designing biological compartmentalization. *Trends Cell Biol.* **22**, 662–670 (2012).
8. Perlmutter, J. D. & Hagan, M. F. Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* **66**, 217–239 (2015).
9. Kerfeld, C. A., Heinhorst, S. & Cannon, G. C. Bacterial microcompartments. *Annu. Rev. Microbiol.* **64**, 391–408 (2010).
10. Catti, L., Zhang, Q. & Tiefenbacher, K. Advantages of Catalysis in Self-Assembled Molecular Capsules. *Chem. - A Eur. J.* **22**, 9060–9066 (2016).
11. Matsuura, K., Ueno, G. & Fujita, S. Self-assembled artificial viral capsid decorated with gold nanoparticles. *Polym. J.* **47**, 146–151 (2015).
12. Godoy-Gallardo, M., York-Duran, M. J. & Hosta-Rigau, L. Recent Progress in Micro/Nanoreactors toward the Creation of Artificial Organelles. *Adv. Healthc. Mater.* **7**, 1–35 (2018).
13. Bayoumi, M., Bayley, H., Maglia, G. & Sapra, K. T. Multi-compartment encapsulation of communicating droplets and droplet networks in hydrogel as a model for artificial cells. *Sci. Rep.* **7**, 1–11 (2017).
14. Zhang, X. & Zhang, P. Polymersomes in Nanomedicine - A Review. *Curr. Nanosci.* **13**, 124–129 (2017).
15. Discher, B. M. *et al.* Polymersomes: Tough vesicles made from diblock copolymers. *Science* **284**, 1143–1146 (1999).
16. Bretscher, M. S. Membrane Structure : Some General Principles. *Science* **181**, 622–629 (1973).
17. Grecco, H. E., Schmick, M. & Bastiaens, P. I. H. Signaling from the living plasma membrane. *Cell* **144**, 897–909 (2011).
18. Tanford, C. The Hydrophobic Effect and the Organization of Living Matter. *Science* **200**, 1012–1018 (1978).
19. Singer, S. J. & Nicolson, G. L. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* **175**, 263–265 (1972).
20. Figard, L. & Sokac, A. M. A membrane reservoir at the cell surface: unfolding the plasma membrane to fuel cell shape change. *Bioarchitecture* **4**, 39–46 (2014).
21. Carlton, J. G., Jones, H. & Eggert, U. S. Membrane and organelle dynamics during cell division. *Nat. Rev. Mol. Cell Biol.* **21**, 151–166 (2020).
22. Boucher, E. & Mandato, C. A. Plasma membrane and cytoskeleton dynamics during

single-cell wound healing. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853**, 2649–2661 (2015).
23. Lingwood, D. & Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **327**, 46–50 (2010).
24. de la Serna, J. B., Schütz, G. J., Eggeling, C. & Cebecauer, M. There is no simple model of the plasma membrane organization. *Trends Biochem. Sci.* **14**, 1–17 (1989).
25. Simons, K. & Vaz, W. L. C. Model systems, lipid rafts, and cell membranes. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 269–295 (2004).
26. Yamakawa, T. & Nagai, Y. Glycolipids at the cell surface and their biological functions. *Annu. Rev. Anal. Chem.* **3**, 128–131 (1978).
27. Paulson, J. C. Glycoproteins: what are the sugar chains for? *Trends Biochem. Sci.* **14**, 272–276 (1989).
28. Göpfrich, K. *et al.* Large-Conductance Transmembrane Porin Made from DNA Origami. *ACS Nano* **10**, 8207–8214 (2016).
29. LaPlanche, L. A. & Rogers, M. T. cis and trans Configurations of the Peptide Bond in N-Monosubstituted Amides by Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **86**, 337–341 (1964).
30. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
31. McCammon, J. A. Protein dynamics. *Reports Prog. Phys.* **47**, 1–46 (1984).
32. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **254**, 1598–1603 (1991).
33. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
34. Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. Structure and possible catalytic residues of Taka-amylase A. *J. Biochem.* **95**, 697–702 (1984).
35. Weiss, M. A. The Structure and Function of Insulin: Decoding the TR Transition. **6729**, 1–16 (2012).
36. Bell, C. E. & Lewis, M. The Lac repressor: A second generation of structural and functional studies. *Curr. Opin. Struct. Biol.* **11**, 19–25 (2001).
37. Cox, R. W., Grant, R. A. & Rhone, R. W. The structure and assembly of collagen fibrils. *J. R. Microsc. Soc.* **87**, 123–142 (1967).
38. Taubenberger, A. *et al.* Revealing Early Steps of α2-β1 Integrin-mediated Adhesion to Collagen Type I by Using Single-Cell Force Spectroscopy. *Mol. Biol. Cell* **18**, 1634–1644 (2007).
39. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
40. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
41. Huang, Q., Li, M., Lai, L. & Liu, Z. Allostery of multidomain proteins with disordered linkers. *Curr. Opin. Struct. Biol.* **62**, 175–182 (2020).
42. Hubbard, S. R., Hendrickson, W. A., Lambright, D. G. & Boxer, S. G. X-ray crystal structure of a recombinant human myoglobin mutant at 2·8 Å resolution. *J. Mol. Biol.* **213**, 215–218 (1990).
43. He, Y. *et al.* Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
44. Whited, A. M. & Johs, A. The interactions of peripheral membrane proteins with

biological membranes. *Chem. Phys. Lipids* **192**, 51–59 (2015).
45. Allen, K. N., Entova, S., Ray, L. C. & Imperiali, B. Monotopic Membrane Proteins Join the Fold. *Trends Biochem. Sci.* **44**, 7–20 (2019).
46. Marsh, D., Horváth, L. I., Swamy, M. J., Mantripragada, S. & Kleinschmidt, J. H. Interaction of membrane-spanning proteins with peripheral and lipid-anchored membrane proteins: Perspectives from protein-lipid interactions. *Mol. Membr. Biol.* **19**, 247–255 (2002).
47. Ikezawa, H. Glycosylphosphatidylinositol (GPI)-Anchored Proteins. *Biol. Pharm. Bull.* **25**, 409–417 (2002).
48. Ferguson, M. A. J., Homans, S. W., Dwek, R. A. & Rademacher, T. W. Glycosyl-phosphatidylinositol moiety that anchors Trypanosoma brucei variant surface glycoprotein to the membrane. *Science* **239**, 753–759 (1988).
49. Braun, V. & Rehn, K. Chemical Characterization, Spatial Distribution and Function of a Lipoprotein (Murein-Lipoprotein) of the E. coli Cell Wall: The Specific Effect of Trypsin on the Membrane Structure. *Eur. J. Biochem.* **10**, 426–438 (1969).
50. Blobel, G., Walter, P. & Gilmore, R. Intracellular protein topogenesis. *Prog. Clin. Biol. Res.* **168**, 3–10 (1984).
51. Von Heijne, G. The membrane protein universe: What's out there and why bother? *J. Intern. Med.* **261**, 543–557 (2007).
52. Ulmschneider, M. B. & Sansom, M. S. P. Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta - Biomembr.* **1512**, 1–14 (2001).
53. Von Heijne, G. Principles of membrane protein assembly and structure. *Prog. Biophys. Mol. Biol.* **66**, 113–139 (1996).
54. Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
55. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).
56. Wimley, W. C. Toward genomic identification of β-barrel membrane proteins: Composition and architecture of known structures. *Protein Sci.* **11**, 301–312 (2002).
57. Reckel, S. *et al.* Solution NMR structure of proteorhodopsin. *Angew. Chemie - Int. Ed.* **50**, 11942–11946 (2011).
58. Vogt, J. & Schulz, G. E. The structure of the outer membrane protein OmpX from Escherichia coil reveals possible mechanisms of virulence. *Structure* **7**, 1301–1309 (1999).
59. Kelkar, D. A. & Chattopadhyay, A. The gramicidin ion channel: A model membrane protein. *Biochim. Biophys. Acta - Biomembr.* **1768**, 2011–2025 (2007).
60. Higgins, C. F. ABC transporters: physiology, structure and mechanism – an overview. *Res. Microbiol.* **152**, 205–210 (2001).
61. Manting, E. H., Van Der Does, C., Remigy, H., Engel, A. & Driessen, A. J. M. SecYEG assembles into a tetramer to form the active protein translocation channel. *EMBO J.* **19**, 852–861 (2000).
62. Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).
63. Doyle, D. A. *et al.* The structure of the potassium channel: Molecular basis of K+ conduction and selectivity. *Science* **280**, 69–77 (1998).
64. Yellen, G. The voltage-gated K+ channels and their relatives. *Nature* **419**, 35–42

(2002).
65. Quistgaard, E. M., Löw, C., Guettou, F. & Nordlund, P. Understanding transport by the major facilitator superfamily (MFS): Structures pave the way. *Nat. Rev. Mol. Cell Biol.* **17**, 123–132 (2016).
66. Yan, N. Structural advances for the major facilitator superfamily (MFS) transporters. *Trends Biochem. Sci.* **38**, 151–159 (2013).
67. Henderson, R. K., Fendler, K. & Poolman, B. Coupling efficiency of secondary active transporters. *Curr. Opin. Biotechnol.* **58**, 62–71 (2019).
68. Morth, J. P. *et al.* A structural overview of the plasma membrane Na+,K +-ATPase and H+-ATPase ion pumps. *Nat. Rev. Mol. Cell Biol.* **12**, 60–70 (2011).
69. Sharma, A. K., Spudich, J. L. & Doolittle, W. F. Microbial rhodopsins: functional versatility and genetic mobility. *Trends Microbiol.* **14**, 463–469 (2006).
70. Spudich, J. L., Sineshchekov, O. A. & Govorunova, E. G. Mechanism divergence in microbial rhodopsins. *Biochim. Biophys. Acta - Bioenerg.* **1837**, 546–552 (2014).
71. Beja, O., Spudich, E. N., Spudich, J. L., Marion, L. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786–789 (2001).
72. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899–911 (1999).
73. Hirschi, S., Kalbermatter, D., Ucurum, Z., Lemmin, T. & Fotiadis, D. Cryo-EM structure and dynamics of the green-light absorbing proteorhodopsin. *Nat. Commun.* **12**, 1–10 (2021).
74. Subramanlam, S. & Henderson, R. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* **406**, 653–657 (2000).
75. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like Protein from the Purple Membrane of Halobacterium halobium. *Nature* **233**, 149–152 (1971).
76. Matsuno-Yagi, A. & Mukohata, Y. Two possible roles of bacteriorhodopsin; a comparative study of strains of Halobacterium Halobium differing in pigmentation. *Biochem. Biophys. Res. Commun.* **78**, 237–243 (1977).
77. Maruyama, I. N. Activation of transmembrane cell-surface receptors via a common mechanism? The 'rotation model'. *BioEssays* **37**, 959–967 (2015).
78. Dufau, M. L. The luteinizing hormone receptor. *Annu. Rev. Physiol.* **60**, 461–496 (1998).
79. Husted, A. S., Trauelsen, M., Rudenko, O., Hjorth, S. A. & Schwartz, T. W. GPCR-Mediated Signaling of Metabolites. *Cell Metab.* **25**, 777–796 (2017).
80. Betke, K. M., Wells, C. A. & Hamm, H. E. GPCR mediated regulation of synaptic transmission. *Prog. Neurobiol.* **96**, 304–321 (2012).
81. Strohmeyer, N., Bharadwaj, M., Costell, M., Fässler, R. & Müller, D. J. Fibronectin-bound α5β1 integrins sense load and signal to reinforce adhesion in less than a second. *Nat. Mater.* **16**, 1262–1270 (2017).
82. Castresana, J., Lübben, M., Saraste, M. & Higgins, D. G. Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen. *EMBO J.* **13**, 2516–2525 (1994).
83. Yin, H. & Flynn, A. D. Drugging Membrane Protein Interactions. *Annu. Rev. Biomed. Eng.* **18**, 51–76 (2016).
84. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there ? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
85. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).

86. Ramakrishnan, V. Ribosome Structure and the Mechanism of Translation. *Cell* **108**, 557–572 (2002).
87. Oswald, J., Njenga, R., Natriashvili, A., Sarmah, P. & Koch, H. G. The Dynamic SecYEG Translocon. *Front. Mol. Biosci.* **8**, 1–30 (2021).
88. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *Trends Biochem. Sci.* **31**, 563–571 (2006).
89. Ulbrandt, N. D., Newitt, J. A. & Bernstein, H. D. The E. coli Signal Recognition Particle is Required for the Insertion of a Subset of Inner Membrane Proteins. *Cell* **88**, 187–196 (1997).
90. Koch, H. G., Moser, M. & Müller, M. Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. *Rev. Physiol. Biochem. Pharmacol.* **146**, 55–94 (2003).
91. Wild, K., Halic, M., Sinning, I. & Beckmann, R. SRP meets the ribosome. *Nat. Struct. Mol. Biol.* **11**, 1049–1053 (2004).
92. Schibich, D. *et al.* Global profiling of SRP interaction with nascent polypeptides. *Nature* **536**, 219–223 (2016).
93. Huber, D. *et al.* Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. *J. Bacteriol.* **187**, 2983–2991 (2005).
94. Saraogi, I. & Shan, S. ou. Co-translational protein targeting to the bacterial membrane. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1433–1441 (2014).
95. Samuelson, J. C. *et al.* YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637–641 (2000).
96. Saio, T., Guan, X., Paolo, R., Economou, A. & Kalodimos, C. G. Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science* **344**, 597–610 (2014).
97. Knoblauch, N. T. M. *et al.* Substrate specificity of the SecB chaperone. *J. Biol. Chem.* **274**, 34219–34225 (1999).
98. Economou, A. & Wickner, W. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* **78**, 835–843 (1994).
99. Carlos, J. L. *et al.* The role of the membrane-spanning domain of type I signal peptidases in substrate cleavage site selection. *J. Biol. Chem.* **275**, 38813–38822 (2000).
100. Goemans, C., Denoncin, K. & Collet, J. F. Folding mechanisms of periplasmic proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1517–1528 (2014).
101. Thoma, J., Burmann, B. M., Hiller, S. & Müller, D. J. Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **22**, 795–802 (2015).
102. Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262–265 (2003).
103. Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol.* **164**, 19–24 (2004).
104. Wu, T. *et al.* Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121**, 235–245 (2005).
105. Sánchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M. & Valencia, A. POTRA: A

conserved domain in the FtsQ family and a class of β-barrel outer membrane proteins. *Trends in Biochemical Sciences* vol. 28 523–526 (2003).
106. Gentle, I. E., Burri, L. & Lithgow, T. Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58**, 1216–1225 (2005).
107. Noinaj, N. *et al.* Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–390 (2013).
108. Konovalova, A., Kahne, D. E. & Silhavy, T. J. Outer Membrane Biogenesis. *Annu. Rev. Microbiol.* **71**, 539–556 (2017).
109. Fleming, K. G. A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 1–6 (2015).
110. Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C. & Buchanan, S. K. Lateral opening and exit pore formation are required for BamA function. *Structure* **22**, 1055–1062 (2014).
111. Doerner, P. A. & Sousa, M. C. Extreme Dynamics in the BamA β-Barrel Seam. *Biochemistry* **56**, 3142–3149 (2017).
112. Lee, J. *et al.* Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *Elife* **8**, 1–20 (2019).
113. Tomasek, D. *et al.* Structure of a nascent membrane protein as it folds on the BAM complex. *Nature* **583**, 473–478 (2020).
114. Höhr, A. I. C. *et al.* Membrane protein insertion through a mitochondrial β-barrel gate. *Science* **359**, 1–12 (2018).
115. Moon, C. P., Zaccai, N. R., Fleming, P. J., Gessmann, D. & Fleming, K. G. Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proc. Natl. Acad. Sci.* **110**, 4285–4290 (2013).
116. Tate, C. G. *et al.* Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim. Biophys. Acta - Biomembr.* **1610**, 141–153 (2003).
117. Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J. & Tan, S. Comparison of affinity tags for protein purification. *Protein Expr. Purif.* **41**, 98–105 (2005).
118. Ilari, A. & Savino, C. Protein structure determination by X-ray crystallography. *Bioinformatics* **452**, 63–87 (2008).
119. Walls, D. & Loughran, S. T. *Strategies for the purification of membrane proteins*. *Methods in Molecular Biology* vol. 1485 (2017).
120. Tate, C. G. Practical Considerations of Membrane Protein Instability during Purification and Crystallisation. *Methods Mol. Biol.* **601**, 187–203 (2010).
121. Seddon, A. M., Curnow, P. & Booth, P. J. Membrane proteins, lipids and detergents: Not just a soap opera. *Biochim. Biophys. Acta - Biomembr.* **1666**, 105–117 (2004).
122. Helenius, A., Mccaslin, D. R., Fries, E. & Tanford, C. Properties of Detergents. **56**, 734–749 (1979).
123. Israelachvili, J. N., Marcelja, S. & Horn, R. G. Physical principles of membrane organization. *Q. Rev. Biophys.* **13**, 121–200 (1980).
124. Helenius, A. & Simons, K. Solubilization of Membranes by Detergents. *Biochim. Biophys. Acta* **415**, 29–79 (1975).
125. Ilgü, H. *et al.* Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents. *Biophys. J.* **106**, 1660–1670

(2014).
126. le Maire, M., Champeil, P. & Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta* **1508**, 86–111 (2000).
127. Privé, G. G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388–397 (2007).
128. Rigaud, J.-L. & Levy, D. Reconstitution of Membrane Proteins into Liposomes. *Methods Enzymol.* **372**, 65–86 (2003).
129. Rigaud, J. L., Pitard, B. & Levy, D. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *BBA - Bioenerg.* **1231**, 223–246 (1995).
130. Goers, R. *et al.* Optimized reconstitution of membrane proteins into synthetic membranes. *Commun. Chem.* **1**, 1–10 (2018).
131. Binning, G., Rohrer, H., Gerber, C. & Weibel, E. Surface Studies by Scanning Tunneling Microscopy. *Am. Phys. Soc.* **49**, 57–61 (1982).
132. Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930–934 (1986).
133. Meyer, G. & Amer, N. M. Novel optical approach to atomic force microscopy. *Appl. Phys. Lett.* **53**, 1045–1047 (1988).
134. Burnham, N. A. *et al.* Comparison of calibration methods for atomic-force microscopy cantilevers. *Nanotechnology* **14**, 1–6 (2003).
135. Dufrêne, Y. F., Martínez-Martín, D., Medalsy, I., Alsteens, D. & Müller, D. J. Multiparametric imaging of biological systems by force-distance curve-based AFM. *Nat. Methods* **10**, 847–854 (2013).
136. Helenius, J., Heisenberg, C. P., Gaub, H. E. & Muller, D. J. Single-cell force spectroscopy. *J. Cell Sci.* **121**, 1785–1791 (2008).
137. Alsteens, D. *et al.* Nanomechanical mapping of first binding steps of a virus to animal cells. *Nat. Nanotechnol.* **12**, 177–183 (2017).
138. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112 (1997).
139. Müller, D. J. *et al.* Single-molecule studies of membrane proteins. *Curr. Opin. Struct. Biol.* **16**, 489–495 (2006).
140. Müller, D. J., Amrein, M. & Engel, A. Adsorption of biological molecules to a solid support for scanning probe microscopy. *J. Struct. Biol.* **119**, 172–188 (1997).
141. Müller, D. J. *et al.* Atomic force microscopy of native purple membrane. *Biochim. Biophys. Acta - Bioenerg.* **1460**, 27–38 (2000).
142. Thoma, J. *et al.* Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. *Commun. Biol.* **1**, 1–9 (2018).
143. Oesterhelt, F. *et al.* Unfolding pathways of individual bacteriorhodopsins. *Science* **288**, 143–146 (2000).
144. Thoma, J., Bosshart, P., Pfreundschuh, M. & Müller, D. J. Out but not in: The large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. *Structure* **20**, 2185–2190 (2012).
145. Serdiuk, T., Sugihara, J., Mari, S. A., Kaback, H. R. & Müller, D. J. Observing a lipid-dependent alteration in single lactose permeases. *Structure* **23**, 754–761 (2015).
146. Sapra, K. T. *et al.* One β hairpin after the other: Exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG. *Angew. Chemie - Int. Ed.* **48**,

8306–8308 (2009).
147. Bippes, C. A. & Muller, D. J. High-resolution atomic force microscopy and spectroscopy of native membrane proteins. *Reports Prog. Phys.* **74**, 1–43 (2011).
148. Sapra, K. T., Balasubramanian, G. P., Labudde, D., Bowie, J. U. & Muller, D. J. Point Mutations in Membrane Proteins Reshape Energy Landscape and Populate Different Unfolding Pathways. *J. Mol. Biol.* **376**, 1076–1090 (2008).
149. Kedrov, A., Krieg, M., Ziegler, C., Kuhlbrandt, W. & Muller, D. J. Locating ligand binding and activation of a single antiporter. *EMBO Rep.* **6**, 668–674 (2005).
150. Kedrov, A. *et al.* Probing the Interactions of Carboxy-atractyloside and Atractyloside with the Yeast Mitochondrial ADP/ATP Carrier. *Structure* **18**, 39–46 (2010).
151. Janovjak, H., Kessler, M., Oesterhelt, D., Gaub, H. & Müller, D. J. Unfolding pathways of native bacteriorhodopsin depend on temperature. *EMBO J.* **22**, 5220–5229 (2003).
152. Bell, G. I. Models for the Specific Adhesion of Cells to Cells. **200**, 618–627 (1978).
153. Evans, E. Probing the Relation Between Force - Lifetime - and Chemistry in Single Molecular Bonds. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 105–128 (2001).
154. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–1555 (1997).
155. Merkel, R., Nassoy, P., Leung, A., Ritchie, K. & Evans, E. Energy landscapes of receptor - ligand bonds explored with dynamic force spectroscopy. **397**, 50–53 (1999).
156. Serdiuk, T. *et al.* Substrate-induced changes in the structural properties of LacY. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1571–1580 (2014).
157. Spoerri, P. M. *et al.* Conformational Plasticity of Human Protease-Activated Receptor 1 upon Antagonist- and Agonist-Binding. *Structure* **27**, 1517–1526 (2019).
158. Thoma, J., Burmann, B. M., Hiller, S. & Müller, D. J. Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **advance on**, 795–802 (2015).
159. Serdiuk, T. *et al.* YidC assists the stepwise and stochastic folding of membrane proteins. *Nat. Chem. Biol.* **12**, 911–917 (2016).
160. Ciancaglini, P. *et al.* Proteoliposomes in nanobiotechnology. *Biophys. Rev.* **4**, 67–81 (2012).
161. Rigaud, J. & Levy, D. Reconstiution of Membrane Proteins into Liposomes. *Methods Enzymol.* **372**, (2003).
162. Biner, O., Fedor, J. G., Yin, Z. & Hirst, J. Bottom-Up Construction of a Minimal System for Cellular Respiration and Energy Regeneration. *ACS Synth. Biol.* **9**, 1450–1459 (2020).
163. Weiss, M. *et al.* Sequential bottom-up assembly of mechanically stabilized synthetic cells by microfluidics. *Nat. Mater.* **17**, 89–95 (2018).
164. Matsubayashi, H., Kuruma, Y. & Ueda, T. Cell-Free Synthesis of SecYEG Translocon as the Fundamental Protein Transport Machinery. *Orig. Life Evol. Biosph.* **44**, 331–334 (2014).
165. Roberts, M. A. J., Cranenburgh, R. M., Stevens, M. P. & Oyston, P. C. F. Synthetic biology: Biology by design. *Microbiology* **159**, 1219–1220 (2013).
166. Frei, T. *et al.* Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.* **11**, 1–14 (2020).
167. Elani, Y. Interfacing Living and Synthetic Cells as an Emerging Frontier in Synthetic Biology. *Angew. Chemie - Int. Ed.* **60**, 5602–5611 (2021).
168. Schmid, Y. R. F., Scheller, L., Buchmann, S. & Dittrich, P. S. Calcium-Mediated

Liposome Fusion to Engineer Giant Lipid Vesicles with Cytosolic Proteins and Reconstituted Mammalian Proteins. *Adv. Biosyst.* **4**, 1–9 (2020).
169. Einfalt, T. *et al.* Bioinspired Molecular Factories with Architecture and In Vivo Functionalities as Cell Mimics. *Adv. Sci.* **7**, 1–13 (2020).

2 Mechanical Unfolding and Refolding of Single Membrane Proteins by Atomic Force Microscopy

Noah Ritzmann and Johannes Thoma

Published in: Perez C., Maier T. (eds) Expression, Purification, and Structural Biology of Membrane Proteins. *Methods in Molecular Biology*, vol 2127, pp 359-372. Humana, New York, NY. February 29, 2020
Doi: 10.1007/978-1-0716-0373-4_23

2.1 Abstract

Atomic force microscopy (AFM)-based single-molecule force spectroscopy allows direct physical manipulation of single membrane proteins under near-physiological conditions. It can be applied to study mechanical properties and molecular interactions as well as unfolding and folding pathways of membrane proteins. Here, we describe the basic procedure to study membrane proteins by single-molecule force spectroscopy and discuss general requirements of the experimental setup as well as common pitfalls typically encountered when working with membrane proteins in AFM.

2.2 Introduction

Following the first appearance in 1986 [1], over the last decades atomic force microscopy (AFM) has evolved into an exceptional tool to study biological membranes and membrane proteins [2]. The latter is mainly owed to the ability to operate AFM in liquid environments, thereby allowing membranes to be studied under near-physiological conditions. In order to be studied by AFM, biological membranes are immobilized on a flat surface to form supported bilayers of 5–10 nm in height. Contouring the membrane surface with an atomically sharp tip at the free end of a microcantilever allows imaging of membrane topographies at sub-nanometer resolution. AFM has generated valuable insight into the molecular details of membrane proteins and their assemblies in biological membranes [3–6].

However, AFM also makes the direct physical manipulation of single membrane proteins possible. Single-molecule force spectroscopy (SMFS) allows probing the force-response of individual macromolecules under mechanical stress and can be used to study the unfolding and folding behavior of membrane proteins [7, 8]. In order to mechanically unfold a membrane protein, the cantilever is first pushed onto the membrane surface applying a non-destructive force to facilitate the unspecific attachment of one terminus of the membrane protein to the tip of the cantilever by physisorption. Subsequent retraction induces the stepwise unfolding of the protein starting from the terminus which is tethered to the cantilever [9]. During retraction the deflection of the cantilever (correlated to the force required to unfold the protein) is recorded together with the retraction distance. Corrected for the cantilever deflection, this distance yields the effective length of the stretched molecule tethered between the tip of the cantilever and the sample surface. Plotting the unfolding force against this tip-sample-separation results in so-called force distance (FD) curves. FD curves recorded during unfolding of membrane proteins typically feature sawtooth- shaped series of unfolding force peaks. These characteristic unfolding fingerprint patterns, which are determined by the interaction-network stabilizing the structural segments of a protein, are unique for each membrane protein. The unfolding fingerprint patterns can therefore be used to identify proteins [10], structural alterations within proteins, different conformational states [11, 12], or to distinguish folded and misfolded states of membrane proteins [13, 14].

SMFS can yield information not only on the unfolding pathways of membrane proteins but also on their folding behavior [15, 16]. To this end, a mechanically unfolded membrane protein is brought into close proximity of the membrane surface in a relaxed state. While still bound to the cantilever, the unfolded protein can reinsert into the lipid bilayer to adopt a folded state, which is then probed by unfolding the protein again [17]. The FD curves resulting from the two successive unfolding processes allow direct comparison of the conformations the protein adopted before and after refolding. This way SMFS helped to reveal to what extent membrane proteins can fold in a self-guided process [14, 18] and how external factors such as molecular chaperones influence the folding pathways of membrane proteins [19, 20].

2.3 Materials

1. Membrane protein sample: Ideal samples for force spectroscopy are proteoliposomes containing the membrane protein of interest embedded in a lipid bilayer at high purity and density. These can either be native membranes, which are naturally rich in a certain protein, or bilayers reconstituted from purified components. If possible, one of the protein's termini should be elongated to ensure efficient attachment to the cantilever during SMFS (see Notes 1 and 2).

2. Buffer: SMFS experiments are typically performed under physiological buffer conditions. However, the buffer should also facilitate firm adsorption of proteoliposomes to mica supports, which might require optimization. A good starting point is the storage buffer the proteoliposomes are supplied in since this buffer should ensure sample stability. Other buffers frequently used in SMFS include sodium or potassium-based PBS (20 mM (Na/K)P_i, pH 7.5, 100–150 mM (Na/K)Cl) and Tris buffers (20 mM Tris–HCl, pH 8, 100–300 mM NaCl) (see Note 6). All buffers should be freshly prepared from analytical grade reagents and ultrapure deionized water (using stock solutions is not recommended) in freshly cleaned glass bottles (see Note 4).
3. AFM: SMFS of membrane proteins can be performed using most commercial AFMs equipped with a 3-axis piezo positioning system. It should support an xy range of >15 x 15 μm and a z range of >1 μm. Moreover, the capability of designing and recording multi-segment force-distance curves is required. To ensure optimal stability during the measurements, the AFM should be placed on an active damping table and acoustically shielded (see Notes 3 and 4).
4. Cantilevers: Soft silicon nitride cantilevers with spring constants in the range 0.01–0.1 N/m and a tip radius <10 nm are the cantilevers of choice and offer a good balance between the ability to image the sample topography prior to and good force resolution during SMFS.
5. Mica supports: Muscovite mica is widely used to immobilize biological membranes for AFM studies due to its atomically flat and negatively charged surface. To prepare supports a mica disk with a diameter of 5 mm and a thickness of ~0.5 mm is glued either to a microscopy glass slide or to a Teflon foil covered metal disk (depending on the sample holder of the AFM) using two-component epoxy glue (see Note 5). It is important that the entire surface of the mica disk is covered in an even thin layer of glue and no air is entrapped in the glue.
6. Fluid cell: It is important that the AFM is equipped with a fluid cell, which can either be a closed design or an open design holding a liquid volume of ≥1 mL. Using a fluid cell ensures maximal stability of the setup during the measurement, minimizes thermal drift and allows measurements to be maintained for longer durations, especially when using a closed design.
7. Cleaning agents: 1% solution of household detergent in a spray bottle, laboratory-grade ethanol and ultrapure deionized water in wash bottles.
8. Compressed air or nitrogen to dry equipment.
9. Scotch tape or other adhesive tape for cleaving mica supports.

2.4 Methods

All preparations and experiments can be performed at room temperature.

3.1 Sample Adsorption and Imaging

1. Wash supports thoroughly with detergent solution, rinse several times iteratively with ethanol and water, dry using compressed air or nitrogen.
2. Cleave the top layer of the mica. To this end push a stripe of adhesive tape onto the mica surface and detach the top layer by pulling the tape off. Inspect the detached mica layer on the tape to ensure it was removed across the entire area of the support. Inspect the freshly cleaved support surface for visible edges. If necessary, repeat until an even and smooth surface is obtained.
3. Add 50 μL of proteoliposome suspension at a concentration of 10 μg/mL to the freshly cleaved mica and allow to adsorb for ~15 min (see Note 6 for optimal conditions). Cover

the support with a glass dish while adsorbing to prevent contamination with dust particles.

4. To wash out unbound material, remove 45 μL of sample solution then add 45 μL of fresh buffer and repeat seven times. Avoid direct contact with the mica surface when pipetting.

5. Transfer the sample to the AFM, fill the fluid cell with fresh buffer and engage the cantilever. Allow to equilibrate for at least 15 min or until the system is stabilized and thermal drift of the cantilever is no longer observed.

6. Approach the cantilever to the sample surface. Choose approach parameters conservatively using only minimal target force and approach velocity to prevent damaging the tip upon surface contact.

7. Record a topography of the sample at low to medium magnification and low resolution in contact mode (scan area ≥10 x 10 μm, 256 x 256 pixels). The topography can be recorded at rather high imaging velocity (up to five lines per second). Try to apply as low imaging force as possible in order not to damage the sample. The topography should give an initial impression of the sample quality. Adsorbed proteoliposomes should be visible as separated individual single-layered membrane patches of 5–10 nm in height. Areas of empty mica between membrane patches should be smooth and free of particulates (see Notes 5–9).

8. Select an area with a clean mica surface devoid of proteoliposomes and record a few force distance curves. Approach and subsequently retract the AFM cantilever without pausing using a constant velocity of 500 nm/s (see Note 10).

9. Assess the quality of the FD curves (Fig. 1, see Note 13). FD curves should not deviate from zero force level in the non-contact regime and increase linearly with piezo-movement in the contact regime. If no hysteresis and no strong adhesion are observed, use one of the recorded FD curves to determine the cantilever's deflection sensitivity in the linear regime of the contact region of the FD curve.

10. Retract the cantilever from the surface for at least 1 μm and calibrate the cantilever's spring constant using the thermal noise method [21].

11. Re-approach the cantilever to the sample surface, select areas of interest and record topographies at increased magnification. It is not necessary to obtain very high resolution; however, the image quality should be sufficient to discern protein-rich regions within membrane patches from regions containing mainly empty lipid (Fig. 2).

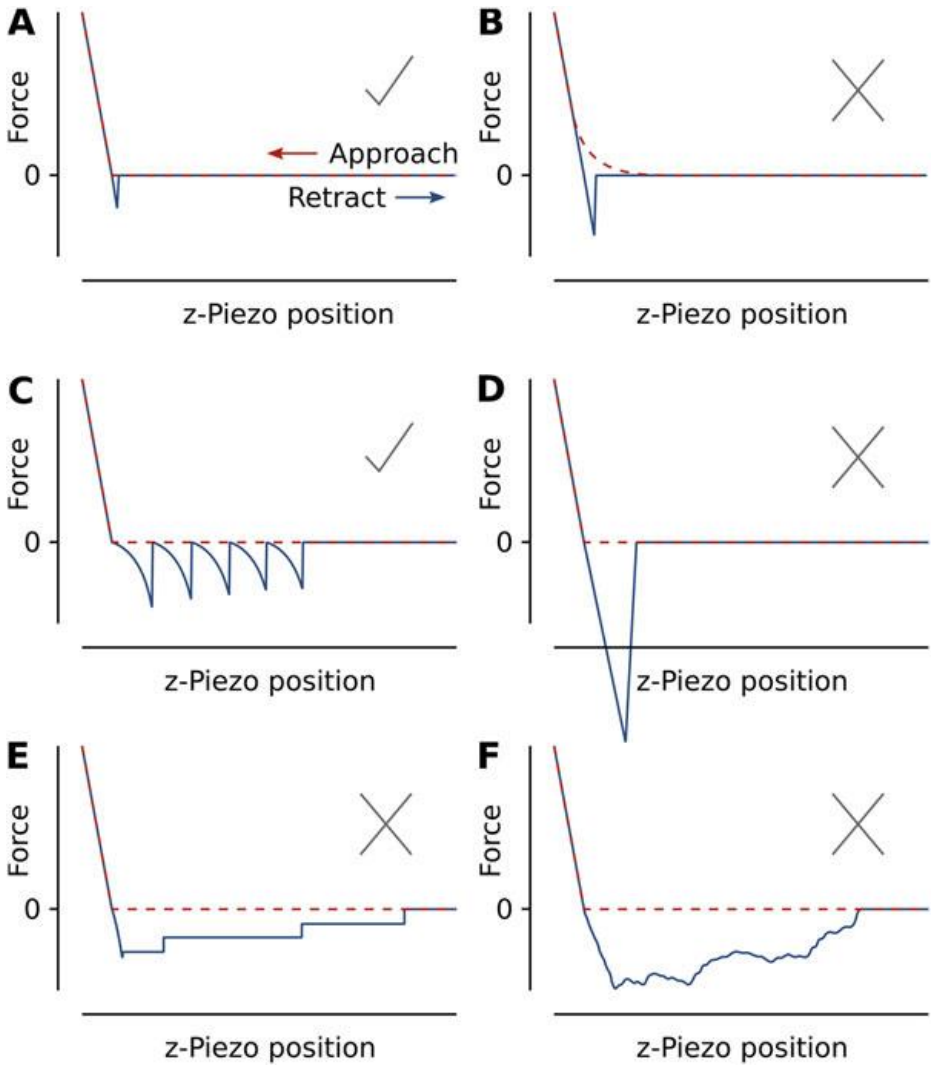


Fig. 1 Typical shapes of FD curves. (a) Ideal FD curves as measured on empty mica showing a sharp bend at the transition between non-contact and contact regime. No hysteresis is observed between the FD curves recorded during the approach (red) and following retract segment (blue). Upon retraction, only a weak surface adhesion event is registered. (b) A smooth bend at the transition between non-contact and contact regime in the approach curve is indicative of debris contaminating the cantilever. (c) Typical sawtooth-shaped series of force peaks as it occurs upon the stepwise unfolding of a membrane protein. (d) Pronounced force peaks with linear force increase resulting from strong surface adhesion can be caused by unfavorable tip geometries. (e) Elongated force plateaus caused by membrane tethers, which often occur when using blunt cantilevers with a large tip radius. (f) Highly irregular and irreproducible force patterns resembling a mountain range indicate sample degradation

3.2 Mechanical Unfolding of Membrane Proteins

1. Select a protein-rich region within a membrane patch with an area of min. 100 x 100 nm and set up a grid of consecutive measurement points with ~10 nm spacing (Fig. 2).

2. For each measurement point, repeat an approach-pause-retract cycle to record FD curves with the following settings: Approach and retract the AFM cantilever with a constant velocity of 500 nm/s (see Note 11). Use a target force of 1 nN for the approach and pause at constant force for 0.5 s (see Note 12). Set the retraction distance ≥2 times the contour length of the studied protein (see Note 13).

3. Assess the quality of initial FD curves (Fig. 1, see Note 14). If no irregular force patterns are observed in these FD curves, pass through the point grid repeating the approach-

pause- retract cycle for every measurement point and save all the recorded FD curves. The point grid can be sampled several times.

4. Interrupt the measurement every 1–2 h to record a new topography. Ensure that the membrane patch is still intact and that the measurement area is set on the correct sample position. Reposition the point grid if sample drift occurred.

5. Keep recording FD curves until the repetitively occurring characteristic unfolding fingerprint pattern of the studied membrane protein can be clearly recognized (see Note 15).

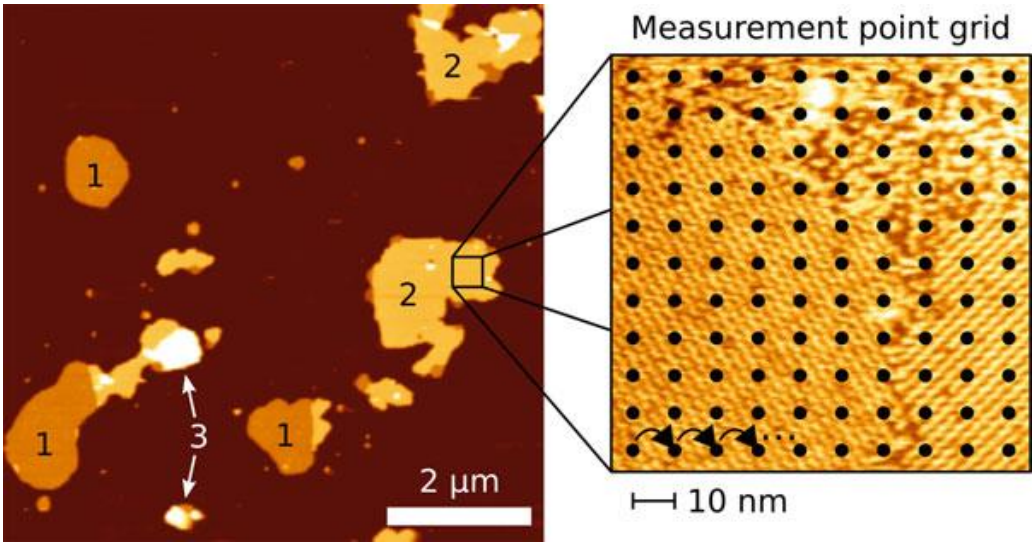


Fig. 2 Exemplary AFM topography. Proteoliposomes of FhuA reconstituted in *E. coli* polar lipids adsorbed to mica in buffer (20 mM Tris–HCl, pH 8, 150 mM NaCl). The topography shows membrane patches containing sparsely distributed membrane proteins (1) and membrane proteins at high density (2). Features protruding far from the surface indicate intact proteoliposomes, which did not break open upon adsorption (3). Areas containing densely packed membrane proteins provide ideal conditions for mechanical unfolding or refolding experiments and should be chosen to set the grid of consecutive measurement points

3.3 Refolding of Membrane Proteins

In contrast to the previously described unfolding of membrane proteins, setting up refolding experiments requires pre-existing knowledge of the characteristic unfolding fingerprint pattern of a protein, which can be established in unfolding experiments. Only then the fraction of the protein that is to be refolded can be selected based on the unfolding pattern. Moreover, since the throughput of refolding experiments is substantially lower, prior knowledge of the unfolding fingerprint pattern simplifies analysis of the resulting data. Otherwise the setup follows the same principles as mechanical unfolding experiments.

1. Select a protein-rich region within a membrane patch with an area of min. 100 x 100 nm and apply a grid of measurement points with ~10 nm spacing.

2. For each measurement point set up an approach-pause-retract-approach- pause-retract cycle to record FD curves with the following settings: Approach and retract the AFM cantilever with a constant velocity of 500 nm/s. Use a target force of 1 nN for the first approach and pause at constant force for 0.5 s. The retraction distance of the first retract depends on the unfolding fingerprint pattern of the studied membrane protein. Set the retraction distance depending on the length of the protein segments that are to be unfolded and subsequently refolded (see Note 16). Thereby take into account that the piezo travels a certain distance while the cantilever is in contact with the sample and deflected (Fig. 3). The distance of the second approach segment depends on the length of

the first retract segment relative to the sample surface and should be chosen to end 5–10 nm above the contact point with the sample surface (see Note 17). Set the second pause at constant height for 1 s and the second retraction distance to ≥2 times the contour length of the studied membrane protein (Fig. 3).

3. Interrupt the measurement every 1–2 h to record a new topography. Ensure that the membrane patch is still intact and that the measurement area is still set on the correct sample position. Reposition the point grid if sample drift occurred.

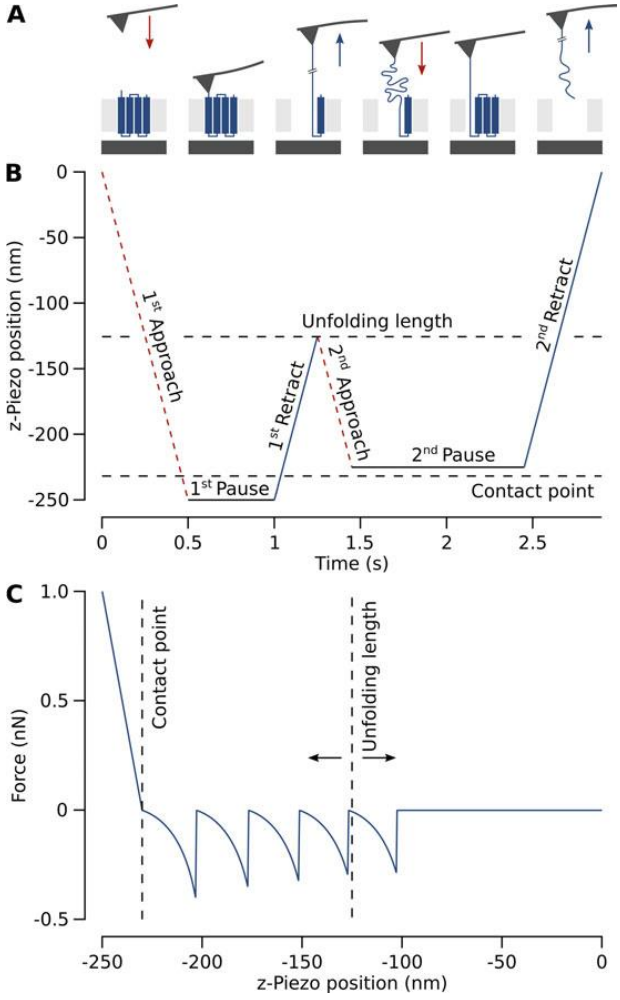


Fig. 3 Refolding setup. (a) Experimental sequence for protein refolding experiments. First, the cantilever is approached to and then pushed onto the membrane surface to facilitate the attachment of the membrane protein. Subsequent retraction of the cantilever induces unfolding of a predefined fraction of the protein. The unfolded protein is then held in close proximity of the membrane surface, upon which the unfolded polypeptide can reinsert into the membrane. The resulting fold is probed by unfolding the protein again. (b) Exemplary time sequence showing movement of the z-piezo during the experiment. The cantilever is approached with a velocity of 500 nm/s. A target force of 1 nN is maintained during the first pause segment for 0.5 s. Note that during the first pause segment, the cantilever is in contact with the sample surface. The cantilever is then retracted for 125 nm and re-approached for 100 nm, taking into account 20 nm the cantilever travels while in contact with the surface (see c) this places the cantilever 5 nm above the surface contact point. The cantilever is then held at this distance for 1 s and subsequently retracted for 225 nm. (c) Typical sawtooth-shaped series of force peaks as it occurs upon the stepwise unfolding of a membrane protein. Based on the unfolding pattern of the protein, the distance of the first retract segment can be chosen to unfold a predefined length of the protein (or the entire protein)

3.4 Data Processing

1. Correct FD curves for cantilever deflection by subtracting the cantilever deflection from the z-piezo position to obtain the actual tip-sample separation (see Note 18). Correct for force offset by setting the non-contact regime of the FD curve to zero force. To this end use 20% of the data points in the non-contact regime of the FD curve most distant from the contact regime. Correct for distance offset by setting the contact regime of the FD curve to zero distance (see Note 19). For representation FD curves are typically oriented to display unfolding force peaks as positive forces. Flip FD curves if necessary.

2. Full unfolding events are typically registered only in a small fraction (<1/1000) of all recorded FD curves, while in the majority of the approach-retract cycles no protein adhered to the cantilever tip, which results in FD curves showing no significant force peaks. FD curves corresponding to protein unfolding events can be selected by applying an automated coarse filtering step to the data of the retract segment. To this end define a

force threshold (>100 pN), which substantially exceeds the noise level of the FD curves and a distance threshold corresponding to >75% of the contour length of the membrane protein being unfolded (Fig. 4, see Note 20). For refolding experiments, set the distance threshold to >75% of the length of the first retract segment.

3. Inspect all resulting FD curves in order to select for FD curves showing the typical saw-tooth-shaped unfolding pattern corresponding to the stepwise unfolding of a membrane protein (Fig. 1) and sort out FD curves showing irregular force patterns.

4. The following detailed analysis of the resulting filtered dataset strongly depends on the type of experiment. Recently, several computational tools were developed to largely automate routine operations in the analysis and processing of force spectroscopy data, such as the alignment of FD curves, the automated fitting of force peaks, or the identification of unfolding pathways [22–24]. However, in particular the analysis of data obtained in refolding experiments remains a specialized task which needs to be adapted to the membrane protein under investigation.

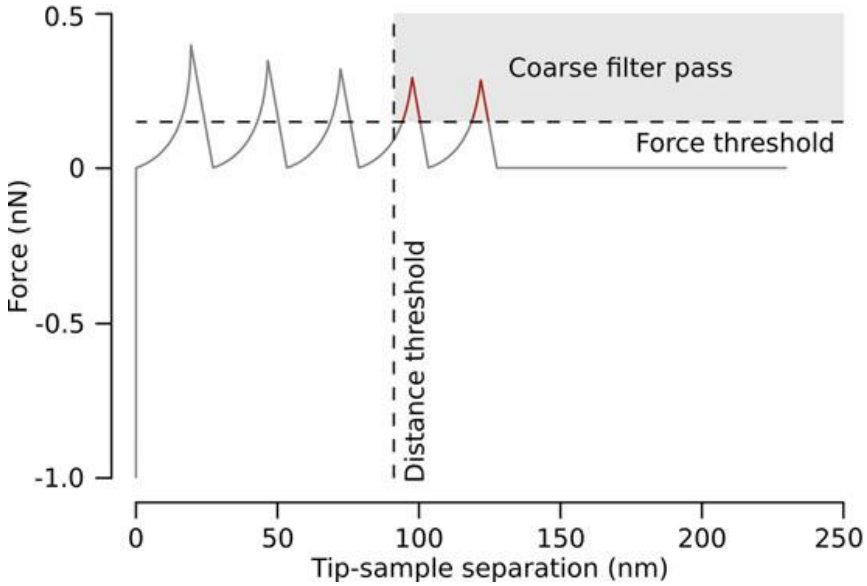


Fig. 4 Transformation and coarse filtering. FD curves are corrected for cantilever deflection shifted to zero force and zero distance. In order to select FD curves corresponding to the full unfolding of a membrane protein, a force threshold and a distance threshold corresponding to >75% of the contour length of the membrane protein being unfolded are defined (grey area). FD curves featuring force peaks within this region (red) pass the coarse filtering step and are taken into account for further analysis

4 Notes

1. Elongation of a terminus can also help to determine from which terminus the protein was unfolded in SMFS experiments since it strongly increases the probability of unfolding from the elongated terminus. Alternatively, this can be achieved through sequence alterations in the protein, e.g., through proteolytic cleavage or presence/absence of a disulfide bridge connecting two segments of the protein. Each of these alterations should result in a shift of either the entire or a part of the unfolding fingerprint pattern of the protein, which will allow determination of the unfolding direction.

2. Purple membrane from *H. salinarium* containing Bacteriorhodopsin at high density represents an ideal reference sample for newcomers, since it is commercially available (e.g., from Cube Biotech, Monheim am Rhein, Germany), easy to handle, and well-studied by SMFS, allowing direct comparison of the obtained results [9, 25].

3. The described method follows the procedure as carried out on a commercially available NanoWizard AFM (JPK Instruments, Berlin, Germany). However, it should be readily adaptable to most AFMs independent of model and manufacturer, under the premise that a basic knowledge of the working principles of AFM is existing and instrument-specific routine procedures such as cantilever mounting, laser alignment, and cantilever calibration can be performed.
4. All parts of the AFM, all tools used to handle cantilevers, fluid cells, etc., as well as all glassware used to prepare buffers must be meticulously clean. This is best achieved by lathering all surfaces extensively using detergent solution, followed by rinsing them several times alternately with ethanol and deionized water and drying them in a stream of nitrogen or compressed air.
5. Ready-to-use mica discs are commercially available in various diameters (e.g., from Electron Microscopy Sciences, Hatfield, USA); however, they can also be prepared from mica sheets using a punch and die set.
6. A proteoliposome concentration of ~10 µg/mL should result in dense but separated membrane patches upon adsorption to mica for many membrane protein samples. Decrease the concentration and adsorption duration if the density is too high and overlapping membrane patches are observed. Increase the concentration and adsorption duration if only few membrane patches are observed. Adding divalent ions can also help to improve adsorption (add $MgCl_2$ or $CaCl_2$ at concentrations of up to 20 mM to the adsorption buffer).
7. Typically, the majority of proteoliposomes will break open to form supported bilayers when adsorbed to the mica surface. However, in some cases a large fraction remains intact. These unbroken proteoliposomes are then visible as bulky blobs, which protrude highly from the surface. If this is the case, one can make use of the AFMs ability to physically interact with the sample and repeatedly image the proteoliposomes in contact mode, iteratively increasing the imaging force until open bilayers become visible (careful, might contaminate tip). If this procedure remains ineffective, another approach can be breaking the proteoliposomes by sonication prior to adsorption using a bath sonicator. To this end several samples should be prepared with varying sonication time and power, in order to find a condition where the proteoliposomes are destabilized enough to form open bilayers on the mica surface but not fragmented. Increased osmotic pressure by using buffers with a reduced salt content during adsorption can aid the formation of open bilayers as well.
8. If large numbers of small particles are observed, clean all equipment thoroughly and ensure all buffers are free from contaminants. However, also the sample itself may contain particulate contaminations such as protein/lipid aggregates, other precipitates, or dust particles. If contaminants are soluble, collect the proteoliposomes by centrifugation (rcf > 20000 x g for >30 min), remove the supernatant, and resuspend the membrane pellet in fresh buffer. If contaminants are insoluble, collect the precipitate by centrifugation (rcf < 5000 x g for <10 min), then recover the supernatant containing the proteoliposomes.
9. Proteoliposome preparations sometimes contain large populations of vesicles, which are too small to perform appropriate SMFS experiments. In such case it can help to repeatedly collect the proteoliposomes by centrifugation, remove the supernatant and resuspend the membrane pellet in fresh buffer. Thereby lower centrifugal forces and shorter centrifugation times help to select for larger proteoliposomes.

10. Prior to calibration, the instrument-specific set point (typically in V) corresponding to a force of 1 nN is not known but can be estimated from previous experience with cantilevers of the same type.
11. A velocity of 500 nm/s represents a good starting value, but of course the velocity can be varied. Choose the sampling rate for approach and retract segments to a value resulting in 8–10 data points per nm (*e.g.*, 4096 Hz at 500 nm/s). Adjust the sampling rate accordingly to obtain an equal number of data points per distance if the velocity is changed.
12. The duration of the pause segment can be increased if a contact time of 0.5 s results in a very low probability of attaching a polypeptide to the cantilever (unfolding events registered in <1/1000 of FD curves).
13. Calculate the contour length of the protein, which is the length at maximum extension, based on the protein sequence using a length of 0.36 nm per amino acid.
14. Replace the cantilever and/or prepare a fresh sample if multiple FD curves show irregular force patterns. Recording FD curves with a non-ideal setup rarely results in interpretable data and unnecessarily complicates downstream data processing.
15. Beware that a single protein species can yet result in more than one characteristic unfolding pattern, for example, if not all proteins are unfolded from the same terminus or if proteins are probed in different conformational states.
16. The protein can either be unfolded partially with the final segments remaining embedded in the lipid bilayer or fully. For partial unfolding, choose a retraction distance not exceeding the position of the last unfolding force peak. For full unfolding choose a retraction distance clearly exceeding the length of the fully stretched protein.
17. Decrease the approach distance used for the second approach if contact with the sample surface is observed (visible as a force increase close to the sample surface).
18. After correction, the contact regime of the FD curve should resemble a vertical line. If not, check whether the deflection sensitivity was determined correctly.
19. Data processing software included with commercially available AFMs should support these operations, otherwise they can be adapted from the procedure described by Bosshart et al. [26].
20. Check whether the AFM control software allows online-filtering of FD curves already during data acquisition. This way only FD curves passing the coarse-filtering step will be saved, which can substantially reduce the post-processing time.

Acknowledgment
We thank Nico Strohmeyer for proofreading the manuscript. Johannes Thoma was supported by a long-term European Molecular Biology Organization (EMBO) fellowship (ALTF 413-2018).

2.5 References

1. Binnig G, Quate C, Gerber C (1986) Atomic force microscope. Phys Rev Lett 56:930–934

2. Bippes CA, Muller DJ (2011) High-resolution atomic force microscopy and spectroscopy of native membrane proteins. Reports Prog Phys 74:086601

3. Butt HJ, Downing KH, Hansma PK (1990) Imaging the membrane protein bacteriorhodopsin with the atomic force microscope. Biophys J 58:1473–1480

4. Fotiadis D, Liang Y, Filipek S et al (2003) Atomic-force microscopy: rhodopsin dimers in native disc membranes. Nature 421:127–128

5. Hoogenboom BW, Suda K, Engel A, Fotiadis D (2007) The supramolecular assemblies of voltage-dependent anion channels in the native membrane. J Mol Biol 370:246–255

6. Shibata M, Yamashita H, Uchihashi T et al (2010) High-speed atomic force microscopy shows dynamic molecular processes in Photoactivated bacteriorhodopsin. Nat Nanotechnol 5:208–212

7. Kedrov A, Janovjak H, Sapra KT, Müller DJ (2007) Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu Rev Biophys Biomol Struct 36:233–260

8. Engel A, Gaub HE (2008) Structure and mechanics of membrane proteins. Annu Rev Biochem 77:127–148

9. Oesterhelt F, Oesterhelt D, Pfeiffer M et al (2000) Unfolding pathways of individual bacteriorhodopsins. Science 288:143–146

10. Thoma J, Manioglu S, Kalbermatter D et al (2018) Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. Commun Biol 1:23

11. Kedrov A, Ziegler C, Muller DJ (2006) Differentiating ligand and inhibitor interactions of a single antiporter. J Mol Biol 362:925–932

12. Bippes CA, Ge L, Meury M et al (2013) Peptide transporter DtpA has two alternate conformations, one of which is promoted by inhibitor binding. Proc Natl Acad Sci U S A 110:E3978–E3986

13. Serdiuk T, Sugihara J, Mari SA et al (2015) Observing a lipid-dependent alteration in single lactose permeases. Structure 23:754–761

14. Thoma J, Bosshart P, Pfreundschuh M, Müller DJ (2012) Out but not in: the large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. Structure 20:2185–2190

15. Kessler M, Gottschalk KE, Janovjak H et al (2006) Bacteriorhodopsin folds into the membrane against an external force. J Mol Biol 357:644–654

16. Yu H, Siewny MGW, Edwards DT et al (2017) Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. Science 355:945–950

17. Kedrov A, Ziegler C, Janovjak H et al (2004) Controlled unfolding and refolding of a single sodium-proton antiporter using atomic force microscopy. J Mol Biol 340:1143–1152

18. Bosshart PD, Iordanov I, Garzon-Coral C et al (2012) The transmembrane protein KpOmpA anchoring the outer membrane of Klebsiella Pneumoniae unfolds and refolds in response to tensile load. Structure 20:121–127

19. Thoma J, Burmann BM, Hiller S, Müller DJ (2015) Impact of Holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat Struct Mol Biol 22:795–802

20. Serdiuk T, Steudle A, Mari SA et al (2019) Insertion and folding pathways of single membrane proteins guided by translocases and Insertases. Sci Adv 5:eaau6824

21. Butt H-J, Jaschke M (1995) Calculation of thermal noise in atomic force microscopy. Nanotechnology 6:1–7

22. Bosshart PD, Frederix PLTM, Engel A (2012) Reference-free alignment and sorting of single-molecule force spectroscopy data. Biophys J 102:2202–2211
23. Thoma J (2017) Unfolding and folding pathways of Escherichia Coli outer membrane proteins. PhD thesis (ETH Zürich). https://doi. org/10.3929/ethz-b-000214458
24. Galvanetto N, Perissinotto A, Pedroni A, Torre V (2018) Fodis: software for protein unfolding analysis. Biophys J 114:1264–1266
25. Müller D, Heymann J, Oesterhelt F et al (2000) Atomic force microscopy of native purple membrane. Biochim Biophys Acta 1460:27–38
26. Bosshart PD, Casagrande F, Frederix PLTM et al (2008) High-throughput single-molecule force spectroscopy for membrane proteins. Nanotechnology 19:384014

3 Maltoporin LamB unfolds beta-hairpins along mechanical stress-dependent unfolding pathways

Johannes Thoma,[1,‡] Noah Ritzmann,[1,‡] Dominik Wolf,[1] Estefania Mulvihill,[1] Sebastian Hiller,[2] and Daniel J. Müller[1,*]

[1]Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH), 4058 Basel, Switzerland
[2]Biozentrum, Universität Basel, 4056 Basel, Switzerland

[‡]Contributed equally
[*]Correspondence: Daniel J. Müller, Dr.

Author Contributions: J.T., N.R., and D.J.M. discussed and designed the project. J.T. and N.R. engineered, expressed, purified and reconstituted the LamB constructs. J.T. and N.R. performed SMFS experiments. J.T., N.R., D.W. and D.J.M analyzed the data. E.M. recorded high-resolution AFM images. All authors wrote and discussed the paper.

Published in: *Structure*, 25, 1139–1144, July 5, 2017
Doi: 10.1016/j.str.2017.05.010

Summary

Upon mechanically pulling either terminal, end beta-barrel forming outer membrane proteins stepwise unfold beta-strands or -hairpins until the entire beta-barrel has been extracted from the membrane. So far, this unique unfolding pathway has been described for beta-barrels comprised of 8, 14, or 22 beta-strands. To substantiate these findings, we here mechanically unfold the 18-stranded beta-barrel outer membrane protein LamB from *Escherichia coli*. We find that its mechanical unfolding pathway is shaped by the stepwise unfolding of beta-hairpins. However, we also observe that beta-hairpins can unfold in groups. Thereby, beta-hairpins unfolding at higher pulling forces show a higher probability to unfold collectively whereas beta-hairpins unfolding at lower forces tend to unfold individually. This result suggests that the collective unfolding of beta-hairpins resembles a far-from-equilibrium process, whereas the unfolding of individual beta-hairpins describes a closer-to-equilibrium process. Our findings support a direct link between outer membrane protein structure and unfolding pathway and contribute to a better understanding of their unfolding in response to mechanical stress.

Introduction

Transmembrane beta-barrels represent a small and unique group of membrane proteins involved in basic cellular functions ranging from solute transport to signal transduction (Koebnik et al., 2000). In contrast to alpha-helical membrane proteins, which are found in the majority of cellular membranes, beta-barrels are exclusively found in outer membranes of Gram-negative bacteria, mitochondria and chloroplasts (Wimley, 2003). Despite showing considerable differences in size, beta-barrel membrane proteins share a common architecture. Generally an even number of antiparallel beta-strands spans the membrane and closes to form a beta-barrel *via* the first and last beta-strand (Koebnik et al., 2000; Schulz, 2002). For bacterial outer membrane proteins this beta-barrel architecture has been shown to be exceptionally stable against heat, chaotropic salts, detergents, and proteolysis (Bonhivers et al., 2001; Phale et al., 1998). Bulk experiments performed in the presence of denaturing chaotropic agents, detergents or temperature describe beta-barrel membrane proteins to unfold *via* two or three intermediates (Huysmans et al., 2010; Otzen and Andersen, 2013; Tamm et al., 2004). The experiments also suggest that during folding small beta-barrel proteins comprising 8 beta-strands in their denatured state first interact with the lipid membrane, where they tilt and insert to establish their native fold (Huysmans et al., 2010; Otzen and Andersen, 2013; Tamm et al., 2004).

Atomic force microscopy (AFM)-based single-molecule force spectroscopy (SMFS) provides a unique approach to characterize the mechanical unfolding and refolding of membrane proteins (Borgia et al., 2008; Engel and Gaub, 2008; Žoldák and Rief, 2013). To this end SMFS applies a mechanical pulling force to one terminal end of the protein embedded in a synthetic or native membrane. At sufficiently high forces the membrane protein unfolds sequentially in a series of steps (Oesterhelt et al., 2000). This stepwise unfolding process, which is characteristic for each membrane protein (Bippes and Muller, 2011; Engel and Gaub, 2008; Kedrov et al., 2007), is enforced by the spatial confinement of the membrane (Hensen and Mueller, 2013). Particularly, applying a mechanical pulling force to either terminal end causes a beta-barrel membrane protein to stepwise unfold individual beta-strands or -hairpins (Bosshart et al., 2012a; Sapra et al., 2009; Thoma et al., 2012). Such mechanical stress may be expected when bacteria adhering to surfaces are exposed to hydrodynamic flow or undergo severe mechanical deformations (Otto et al., 2001). It has also been shown that beta-barrel proteins once unfolded by mechanical force can insert and fold back into the lipid membrane (Bosshart et al., 2012a; Damaghi et al., 2011). Thereby the unfolded polypeptide can insert one beta-hairpin after the other until folding of the beta-barrel has been completed. However, compared to smaller beta-barrel membrane proteins, larger ones cannot refold into the membrane and are prone to misfolding (Thoma et al., 2012). In the latter case periplasmic chaperones can prevent misfolding and support the stepwise insertion and folding of beta-hairpins towards the native membrane protein structure (Thoma et al., 2015). Thus, studying the mechanical unfolding of beta-barrel proteins can provide complementary mechanistic insight into their folding behavior (Bosshart et al., 2012a; Damaghi et al., 2011; Horne and Radford, 2016; Thoma et al., 2015). In previous studies we used AFM-based SMFS to investigate the mechanical unfolding of bacterial outer membrane proteins having different sizes: i) the small outer membrane protein OmpA from Klebsiella pneumoniae comprising 8 beta-strands (Bosshart et al., 2012a), ii) the intermediate sized outer membrane protein OmpG from *Escherichia coli* comprising 14 beta-strands (Sapra et al., 2009), and iii) the large outer membrane protein FhuA from *E. coli* comprising 22 beta-strands (Thoma et al., 2012). The mechanical unfolding pathway of each of these beta-barrel proteins is

characterized by the stepwise unfolding of beta-hairpins and sometimes of single beta-strands. Thereby individual beta-hairpins and beta-strands can either unfold individually or groupwise. However, so far it is not entirely understood what causes beta-hairpins and beta-strands to unfold individually or in groups and how they shape the unfolding pathways of transmembrane beta-barrel proteins.
Here we attempt to gain a deeper understanding of the mechanical unfolding pathways of outer membrane proteins observed so far and apply SMFS to unfold another beta-barrel membrane protein: maltoporin (LamB) from *E. coli*. LamB forms homo-trimeric pores in the bacterial outer membrane, which are specific for the uptake of maltose (Luckey and Nikaido, 1980). The structure of each LamB monomer comprises 18 antiparallel beta-strands, which are connected by short periplasmic turns and extended loops on the extracellular side (Schirmer et al., 1995). Thus, LamB is distinguished from beta-barrel proteins previously studied by SMFS in size, structure and oligomeric assembly. In agreement with previously unfolded beta-barrel proteins we find that the unfolding pathway of LamB is shaped by beta-hairpins. However, by analyzing the unfolding pathways of single LamB molecules we observe that the propensity of beta-hairpins to unfold groupwise is higher if the externally applied force is high, and lower if the applied force is lower. This indicates a direct force-dependency of the unfolding pathways, which beta-barrel membrane proteins undergo when exposed to mechanical stress.

3.1 Results

Unfolding of LamB by AFM-based SMFS requires the attachment of a terminal end to the AFM stylus. Pushing the AFM stylus onto the membrane protein facilitates the nonspecific attachment of its polypeptide to the stylus (Müller and Engel, 2007). This nonspecific attachment is transient but sufficiently strong to mechanically pull the terminus and to unfold and extract the protein from the membrane. To enhance the probability to non-specifically attach the N-terminal end of LamB we elongated it by 15 amino acids (aa), which included a His_6-tag for purification of the protein (Figure S1). This LamB construct was expressed, purified, and reconstituted into 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) proteoliposomes (Experimental Procedures). High-resolution AFM of the proteoliposomes revealed densely packed LamB trimers, which either exposed the periplasmic or extracellular surface (Figure 1). For SMFS the stylus of the AFM cantilever was pushed onto these LamB assemblies to non-specifically attach one of the elongated N-termini (Figure 2A). Upon subsequent retraction of the AFM stylus we recorded the cantilever deflection and the distance travelled by the stylus in a so-called force-distance (FD) curve. In ≈ 0.02 % of all cases (n = 591'380) FD curves showed a distinct series of up to nine force peaks within a distance of up to 120 nm (Figure 2B). This distance corresponds to the contour length of the fully unfolded and stretched LamB polypeptide (Experimental Procedures). Each of the force peaks, which reoccurred at certain positions most of the FD curves, represents one unfolding step of the membrane protein. However, not all FD curves contained the same number of unfolding steps, but could miss a single or several force peaks (Figure 2B). These force peaks thus recorded the variable unfolding steps along the unfolding pathways of single maltoporins.

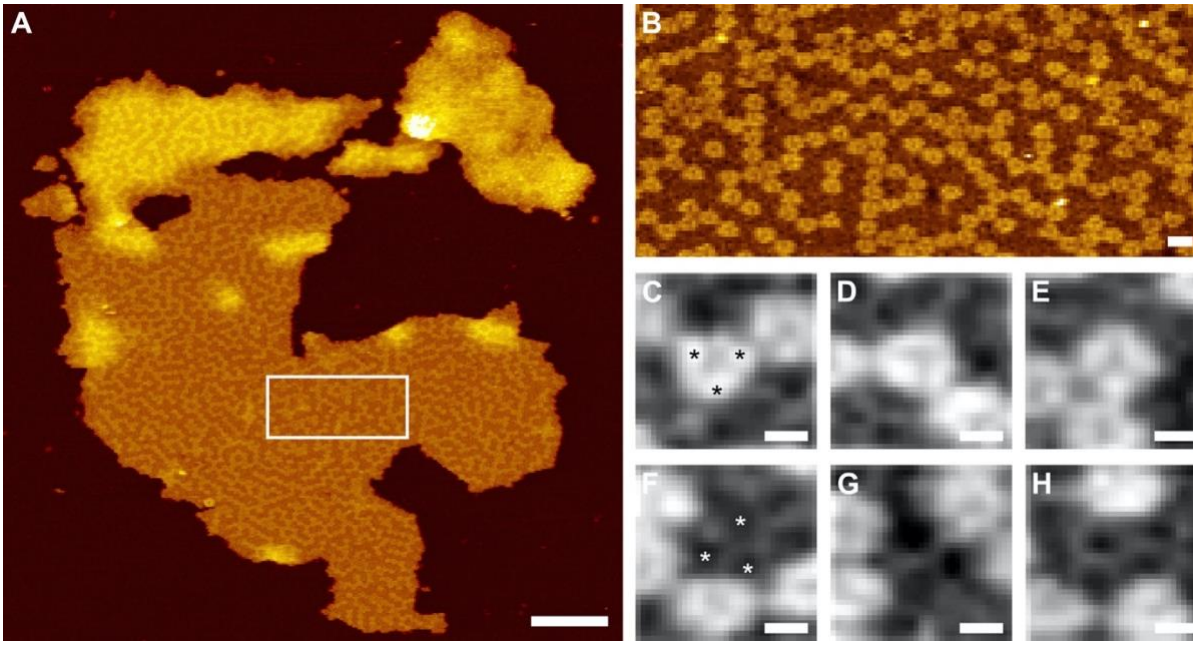


Figure 1. High-resolution AFM of Reconstituted LamB. (A) Overview AFM topograph of a proteoliposome containing LamB in a densely packed arrangement, adsorbed to mica. Upon adsorption to mica the proteoliposomes opened as single layered membrane patches. (B) High-resolution topograph of the area marked with a white rectangle in (A). Individual LamB trimers expose either their periplasmic (dark) or extracellular (bright) surface. (C-E) Enlarged view of the extracellular surface showing the extracellular loops (bright protrusions) protruding from the LamB trimer. (F-H) Enlarged view of the periplasmic surface showing the entrances (dark holes) of the three pores of the LamB trimer. Scale bars, 100 nm (A), 10 nm (B) and 2 nm (C-H). Pixels represent true pixels of the AFM topographs (raw data). Full color scales correspond to vertical scales of 15 nm (A) and 5 nm (B). Topographs were recorded in buffer solution at room temperature. See also Figure S1.

To reveal the common unfolding pattern of LamB the FD curves were aligned and superimposed (Figure 3A). This common pattern comprised nine dominant force peaks each corresponding to one unfolding step. To determine the contour length of the polypeptide unfolded and stretched in every of these force peaks, each peak in every FD curve was fitted applying the worm-like chain (WLC) model (Figure 3B) (Müller and Engel, 2007; Siggia et al., 1994). The contour lengths and unfolding forces of all force peaks were plotted and clustered using a density-based clustering method (Figure 3C) (Ester et al., 1996). For every cluster the mean contour length and the mean unfolding force was calculated (Table S1). The unfolding force peaks distributed at contour lengths ranging from 25 to 372 aa. The unfolding forces required to unfold LamB ranged from ≈ 550 pN for the first unfolding step to ≈ 100 pN for the last unfolding step. To assign the unfolding force peaks to the structural segments unfolded in each unfolding step we had to confirm experimentally that LamB was mechanically unfolded from the N-terminal end. We thus created a LamB construct, in which extracellular loop 6 was shortened by 23 aa, and unfolded this LamBΔL6 construct by SMFS (Figure S2). The resulting FD curves showed that all force peaks detected before the force peak at contour length ≈ 231 aa, remained unchanged compared to full-length LamB (Figure 2 and Figure S2). However, the force peaks detected at longer contour lengths were shifted to shorter contour lengths by ≈ 26 aa, which corresponded to the shortened loop 6 (aa positions 255 - 277). This result confirmed that we unfolded LamB by mechanically pulling its N-terminal end.

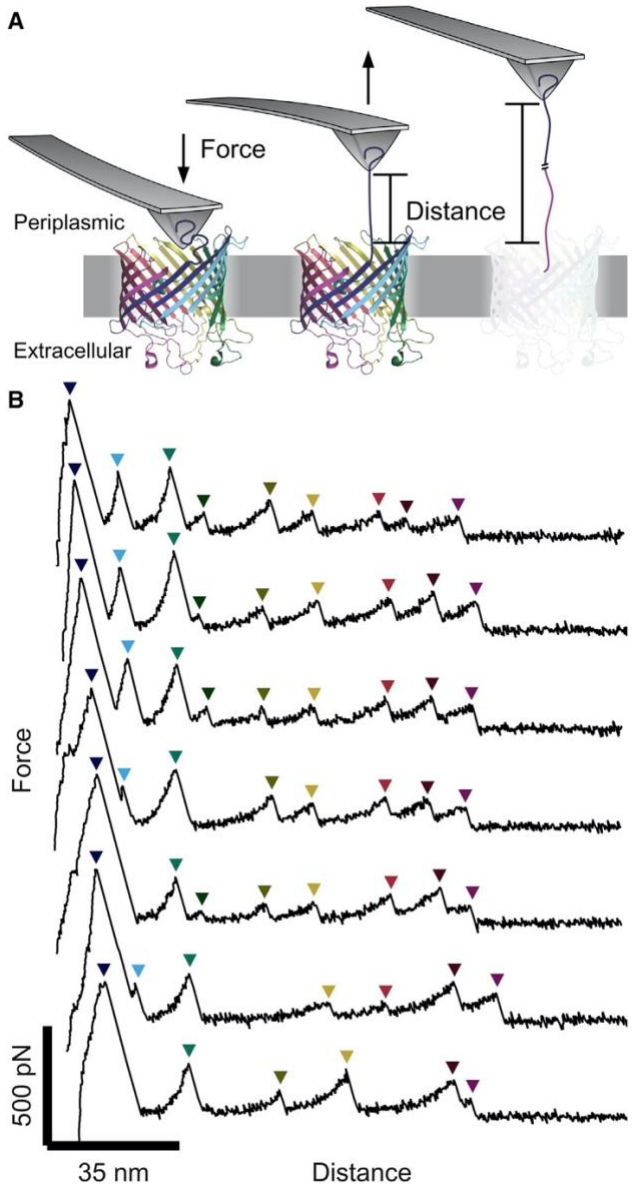


Figure 2. Mechanical Unfolding of LamB by SMFS. (A) Single LamB monomers are non-specifically attached to the stylus of an AFM cantilever by pushing the AFM stylus onto LamB. Retraction of the cantilever increases the distance and applies a mechanical force, which induces unfolding of LamB. The force required to unfold a single LamB monomer is recorded in force vs distance (FD) curves. The Lamb structure has been taken from PDB-code: 1MAL (Schirmer et al., 1995). (B) Exemplary FD curves showing the complete unfolding of individual LamB monomers. Colored arrows indicate the positions of force peaks, each of which represents one unfolding step. See also Figure S2.

After confirming that LamB was unfolded from the N-terminal end, we assigned the structural segments unfolded in each unfolding step. To localize the unfolding steps on the secondary structure of LamB the contour length of each unfolding force peak was subtracted from the N-terminal end (Figure 3D) (Müller and Engel, 2007). Each unfolding force peak was located at the beginning of an individual beta-hairpin. The unfolding steps of all nine beta-hairpins thus described the unfolding pathway of LamB. During this stepwise unfolding process, the partially unfolded beta-barrel remained folded in the membrane thereby forming unfolding intermediates along the unfolding pathway of LamB (Figure 4A).

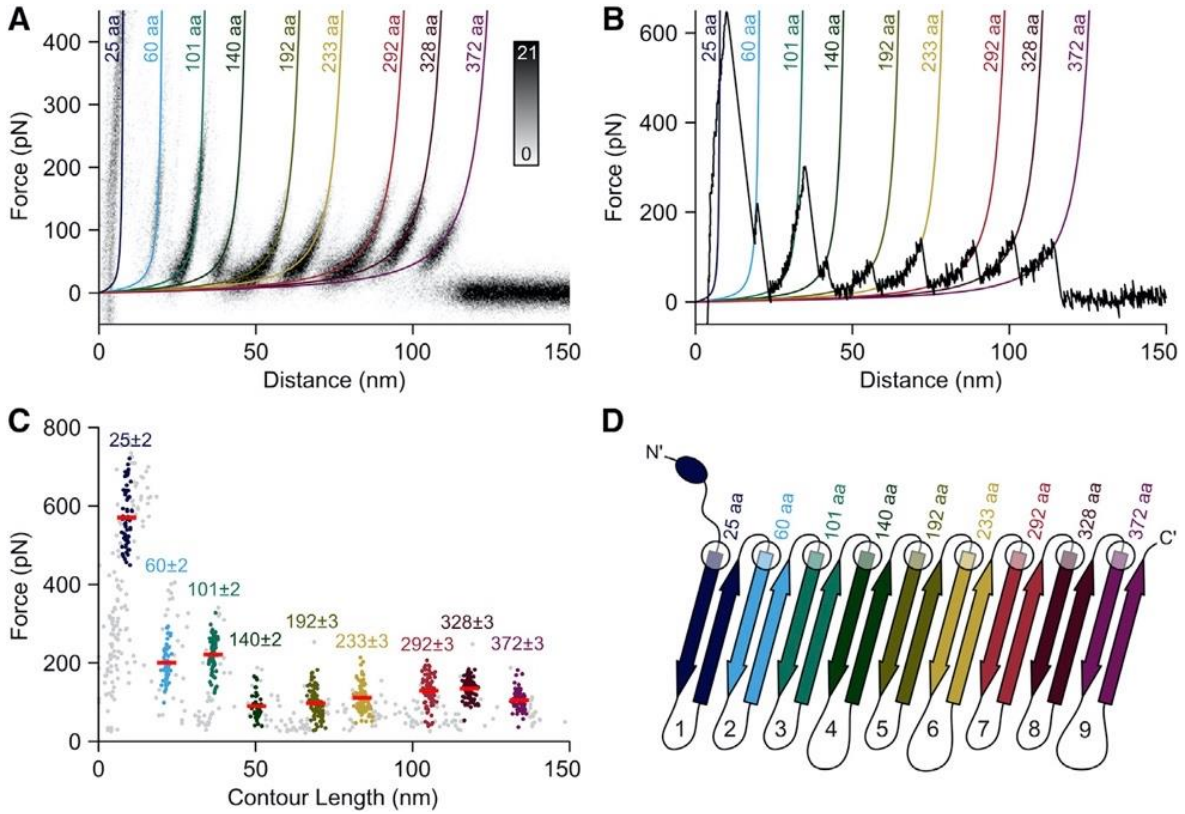


Figure 3. Mechanical Unfolding Steps of LamB. (A) Superimposition of 99 aligned FD curves, each recording the complete unfolding of a single LamB, highlights their common unfolding pattern. The gray scale bar of the density plot indicates the number of data points superimposing per bin. Colored lines are worm-like chain (WLC) curves approximating the mean contour length of each common unfolding force peak. Contour lengths are given in amino acids (aa). (B) Single FD curve showing full unfolding of LamB. Colored lines are WLC curves fitted to every force peak with contour lengths given at their ends in aa. (C) Unfolding force *vs* contour lengths of all force peaks detected in all FD curves. Data points of the same color belong to the same force peak class. Unclassified points are grey. Each force peak class indicates one unfolding step. Red lines indicate the mean force of each force peak class. Values above each force peak class provide the average contour length in aa ± SD. (D) Secondary structure cartoon of LamB showing the locations of predominant mechanical unfolding steps (circles). Numbers 1–9 indicate periplasmic loops as well as beta-hairpins 1–9. The secondary structure of Lamb has been taken from PDB-code: 1MAL (Schirmer et al., 1995). See also Figure S2 and Table S1.

The mechanical unfolding pathway of LamB was shaped by the stepwise unfolding of nine beta-hairpins. However, some FD curves recorded upon unfolding LamB missed individual force peaks, while all other peaks remained unchanged of their position (Figure 2B). To describe such variations of the unfolding pathways of individual maltoporins we analyzed how often an unfolding force peak was detected at certain contour lengths (Figure 4B). Among the 99 FD curves recorded 11% showed all nine unfolding steps of the nine beta-hairpins forming the LamB beta-barrel. The probabilities that 1, 2, 3, and more unfolding steps were missing were 26%, 31%, 16%, and 16%, respectively (Figure S3). To further describe the unfolding pathway of LamB, we analyzed the transition-frequency between the individual unfolding steps (Figure 4B). This frequency is visualized by the thickness of the lines connecting individual unfolding steps. The analysis revealed that the unfolding of the first beta-hairpin, which was detected in > 90% of the FD curves, was followed by a separate unfolding step of the second beta-hairpin in 62% of the FD curves. However, in 38% of all FD curves the first two or more beta-hairpins unfolded cooperatively in a single unfolding step. The same tendency was observed for the third and fourth beta-hairpin, which unfolded individually in 56% of the

FD curves but cooperatively in 44% of the FD curves. The probability of beta-hairpins to unfold pairwise decreased with the force, which had to be applied to induce their unfolding. Consequently, the remaining beta-hairpins 5–9, which unfolded at much lower force, predominantly unfolded in individual steps.

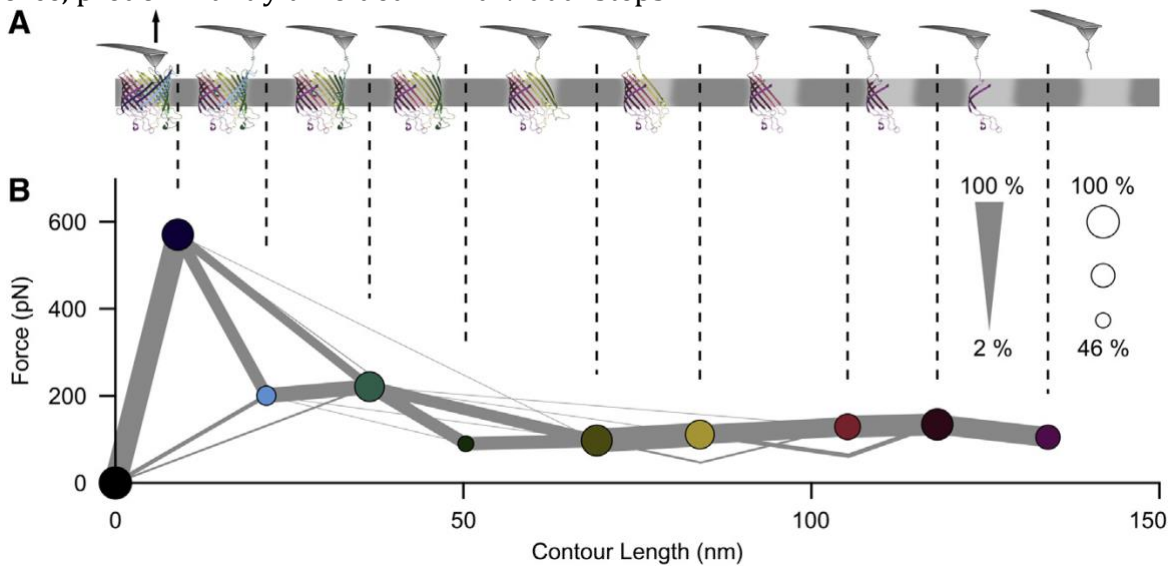


Figure 4. Mechanical Unfolding Intermediates and Unfolding Pathways of LamB. (A) Tertiary structure cartoon illustrating the unfolding intermediates of LamB. The unfolding intermediates, which progresses from the left to the right, are shaped by the partially unfolded beta-barrel remaining in the membrane while mechanically pulling the N-terminal end induces the stepwise unfolding of beta-hairpins. (B) Probabilities of unfolding steps shaping the unfolding pathway of LamB. Colored circles represent mean forces and mean contour lengths of the force peak classes found in Figure 3C. Grey lines indicate transitions between unfolding steps and the line thickness gives the probability of transiting between unfolding steps (see thickness indicator). Size of circles indicates the probability of occurrence of each force peak class (see reference circles). See also Figure S3, S4 and Table S1.

3.2 Discussion

In our SMFS experiments individual LamB monomers resisted pulling forces of up to ≈ 900 pN before starting unfolding. Previous experiments indicated, that typical forces required to unfold the transmembrane beta-barrel proteins OmpA, OmpG and FhuA range from 150 to 400 pN (Bosshart et al., 2012a; Sapra et al., 2009; Thoma et al., 2012), which are higher than those required to unfold alpha-helical transmembrane proteins. Alphahelical proteins such as bacteriorhodopsin, lactose permease (LacY), and human aquaporin-1 (hAQP1) typically unfold at forces ranging from 100 to 150 pN (Müller et al., 2003; Oesterhelt et al., 2000; Serdiuk et al., 2014). We thus conclude that LamB is an exceptionally stable outer membrane protein. However, unlike OmpA, OmpG and FhuA, which occur as monomers, LamB forms trimers. In this trimeric arrangement, the first beta-hairpin and the last beta-hairpin closing the beta-barrel are located at the interface formed between the three LamB monomers (Figure S4). The trimeric interface may thus stabilize LamB against mechanical unfolding. Whereas for OmpA the last beta-strand required the highest force to unfold, OmpG, FhuA, and LamB required the highest unfolding force to unfold the first beta-hairpin. The latter finding is in agreement with steered molecular dynamics simulations, which suggested that the highest force is required to open the beta-barrel of OmpG and to initiate the unfolding of the first beta-hairpin (Hensen and Mueller, 2013).

The unfolding pathway of LamB is shaped by the nine beta-hairpins of the transmembrane beta-barrel. Each of these beta-hairpins can unfold in a single step. Such

unfolding steps defined by individual beta-hairpins were observed before in the 14 betastranded beta-barrel of OmpG, which unfolded in seven steps (Sapra et al., 2009) as well as in the 22 beta-stranded beta-barrel of FhuA, which unfolded in 11 steps (Thoma et al., 2012). The four unfolding steps found for the eight beta-stranded beta-barrel of OmpA, however, described the unfolding of either single, pairs or triplets of beta-strands (Bosshart et al., 2012a). With LamB confirming the trend towards a hairpin-wise unfolding pattern of transmembrane beta-barrel proteins, the differing unfolding pattern of OmpA likely presents an outlier. Due to the small size of the beta-barrel of OmpA and the effective absence of a barrel lumen (Renault et al., 2009), this difference in unfolding behavior can be rationalized by interactions of sidechains protruding into the interior of the barrel (Tamm et al., 2004). Here we observe that the first four N-terminal beta-hairpins of LamB are prone to unfold pairwise with their adjacent hairpin. Among all beta-hairpins of LamB these beta-hairpins require the highest pulling force to unfold.
A force mechanically stressing a folded structure is considered to shift the structure out of equilibrium until it unfolds (Bell, 1978; Evans and Ritchie, 1997). The higher the externally applied force the more the system is shifted out of equilibrium. It may be thus concluded, that beta-hairpins unfolding at lower forces describe a closer-to-equilibrium process, whereas the pairwise unfolding of beta-hairpins represents a far-from-equilibrium process. Similarly, the absence of one or more unfolding steps was observed in the unfolding patterns of other transmembrane beta-barrel proteins. For example, only 31% of all OmpG and 15% of all FhuA molecules unfolded by SMFS showed the unfolding of every single beta-hairpin of the transmembrane beta-barrel (Sapra et al., 2009; Thoma et al., 2012). For LamB, we here show that only in 11% of all cases every beta-hairpin of the transmembrane beta-barrel unfolds in a separate step. However, so far such instances of cooperatively unfolding beta-hairpins could neither be localized to specific regions of the proteins nor be correlated to high unfolding forces. The hypothesis that the groupwise unfolding of beta-hairpins represents a far-from-equilibrium process compared to the individual unfolding of beta-hairpins should be experimentally further tested.
In conclusion, the mechanical unfolding pathway of LamB is shaped by the stepwise unfolding of beta-hairpins, which has been described for other outer membrane proteins having either smaller or larger beta-barrels. Our findings therefore strengthen the hypothesis that the secondary structure of transmembrane beta-barrel proteins determines the unfolding steps along the unfolding pathway. It remains to be elucidated whether this phenomenon can be further confirmed with other beta-barrel membrane proteins. We also found evidence that the tendency of a beta-hairpin to unfold individually or collectively with adjacent beta-hairpins depends on how far the folded structure is shifted out of equilibrium. Whereas this dependency, which has been demonstrated for transmembrane alpha-helical proteins (Janovjak et al., 2003), can be generalized for transmembrane beta-barrel proteins must be confirmed by complementary approaches. The existence of such a dependency would allow us to describe common folding and unfolding behaviors among membrane proteins having different secondary structure elements.

3.3 Experimental Procedures

Cloning. The lamB gene was amplified from genomic DNA from *E. coli* strain MG1655 using primers CCA TCA CCA TCC ATG GGG AGG CTC TGG AGG CTC TGG AGT TGA TTT CCA CGG CTA TGC ACG TTC C and GTG CGG CCG CAA GCT TTA CCA CCA GAT TTC CAT CTG GGC A and cloned into plasmid pY27, using an In-Fusion Cloning Kit (Takara Bio USA,

Mountain View, USA). Plasmid pY27 contains the ompG gene from *E. coli* strain DH10B, with the sequence CAT CAT CAC CAT CAC CAT CCA TGG GGA GGC TCT GGA GGC TCT GGA (amino acid sequence HHHHHHPWGGSGGSG) inserted behind the sequence coding for the periplasmic export signal, in a pET21b(+) backbone between the NdeI and HindIII restriction sites. The plasmid was linearized using NcoI and HindIII restriction endonucleases. The deletion of loop 6 of LamB was introduced using the QuikChange PCR method (Cormack and Castaño, 2002) with primers CTC GAT GAC CAA CAA CAA CGG TCA CAT GC and CGT TGT TGT TGG TCA TCG AGT CAG TAG C.

Expression, Purification and Reconstitution of LamB. LamB samples were prepared essentially as described for the outer membrane protein FhuA (Thoma et al., 2012) with minor modifications (Figure S1B). LamB was expressed under leaky expression conditions in *E. coli* Bl21(DE3)omp8 (Prilipov et al., 1998) for 24 h at 23ÅãC in 4 l LB broth supplemented with Ampicillin (100 µg/ml). Cells were harvested (5000 xg for 12 min, 4ÅãC), resuspended in 100 ml buffer (20 mM Tris-HCl, 100 mM NaCl, pH 8), and broken by sonication. Cell envelopes were collected by centrifugation (100'000 xg for 1h, 4ÅãC) and the inner membrane was solubilized in 100 ml buffer (20 mM Tris-HCl, 2% (w/v) Nlauroylsarcosine, pH 8) for 2h at room temperature. The outer membrane was collected by centrifugation (100'000 xg for 1h, 4ÅãC) and solubilized in 100 ml buffer (20 mM Tris- HCl, 150 mM NaCl, 3% (w/v) n-octyl-β-D-glucopyranoside (OG), pH 8). LamB was bound to 1 ml Protino Ni-NTA agarose resin for 1 h at 4ÅãC (Macherey-Nagel, Düren, Germany), 73 collected on a 30x1 cm collection column (FLEX-COLUMNS, Kimble Chase, Rockwood, USA), washed with 100 ml buffer (20 mM Tris-HCl, 150 mM NaCl, 1.5% (w/v) OG, pH 8) and 100 ml buffer (20 mM Tris-HCl, 150 mM NaCl, 1% (w/v) OG, 10 mM imidazole, pH 8), and eluted with 2.5 ml buffer (20 mM Tris-HCl, 150 mM NaCl, 1% (w/v) OG, 500 mM Imidazole, pH 8). Elution buffer was exchanged to buffer (20 mM Tris-HCl, 150 mM NaCl, 1% (w/v) OG, pH 8) using PD-10 desalting columns (GE Healthcare, Chicago, USA) and the protein concentration was adjusted to 1 mg/ml. LamB was reconstituted into 1,2- dimyristoyl-sn-glycero-3-phosphocholine (DMPC (Avanti polar Lipids, Alabaster, USA), 1 mg/ml in 20 mM Tris-HCl, 150 mM NaCl, 1% (w/v) OG, pH 8) at a lipid to protein ratio of 0.2 (w/w) by dialysis-driven detergent removal. To this end LamB and DMPC were mixed in a total volume of 60 µl, equilibrated for 1 h at 30ÅãC, and dialyzed against 1 l buffer (20 mM Tris-HCl, 150 mM NaCl, 0.01% (w/v) NaN_3, pH8) using Spectra/Por 2 dialysis membranes (12 - 14 kDa MWCO, Spectrum Laboratories, Rancho Dominguez, USA) for 5 days at 30ÅãC with daily buffer exchange.

High-Resolution AFM Imaging. Proteoliposomes containing LamB were adsorbed to freshly cleaved mica for 30 min in buffer solution (20 mM Hepes, 300 mM NaCl, 20 mM $MgCl_2$, pH 7.8) at room temperature. The sample was gently rinsed with buffer (20 mM Hepes, 150 mM NaCl, 5 mM $MgCl_2$ pH 7.8) to remove non-adsorbed membranes (Müller et al., 1997). Then the sample was imaged using force-distance curve-based AFM (FD-based AFM, Nanoscope Multimode 8, Bruker, Santa Barbara, USA) in buffer solution at room temperature in the PeakForce Tapping mode as described (Pfreundschuh et al., 2014). For imaging, we applied a maximum force of 70 pN, an oscillation frequency of 2 kHz and oscillation amplitudes of 30–40 nm. The AFM was equipped with a 120 µm piezoelectric scanner and fluid cell. AFM cantilevers used PEAKFORCE-HIRS-F-A (Bruker, Santa Barbara, USA) had a nominal spring constant of 0.35 N m_{-1}, a resonance frequency of 165 kHz in liquid and sharpened silicon tip with a nominal radius of ≈ 1 nm. Image analysis was performed using the Nanoscope analysis software (version 1.5).

Single-Molecule Force Spectroscopy (SMFS). Proteoliposomes were adsorbed to freshly cleaved mica as described above for AFM imaging. Adsorbed membrane patches

were localized by AFM imaging (Nanowizard II, JPK Instruments, Berlin, Germany) in buffer solution. For AFM imaging and AFM-based SMFS we used OMCL-RC800PSA cantilevers (Olympus, Tokyo, Japan), which were calibrated using the thermal noise method (H -J Butt and M Jaschke, 1995). For SMFS the AFM stylus was pushed onto membrane patches showing densely packed LamB assemblies for 2 s applying a force of 1 nN. The AFM stylus was then retracted at a constant velocity of 500 nm s^{-1}. In about 0.02% of all cases (n = 591'380) the N-terminus of LamB attached non-specifically to the AFM stylus and an FD curve showing the complete unfolding of LamB could be recorded.

Data Analysis. FD curves were pre-processed as described (Bosshart et al., 2008). FD curves were corrected for deflection sensitivity of the cantilever and coarse filtered for force peak patterns exceeding a length of 90 nm, which corresponds to ≈ 2/3 of the contour length of the fully unfolded and extended LamB polypeptide (436 aa including the N-terminal extension). After filtering 99 FD curves remained, which were analyzed using an automated approach (Thoma et al., 2017). For alignment FD curves were transformed to force vs contour length space using the worm-like chain (WLC) model with a fixed persistence length of 0.4 nm (Bosshart et al., 2012b) and binned with a contour length bin-size of 1 nm. The FD curve with the highest similarity to all other FD curves was identified based on the minimal Euclidean distance against all curves, and used as a template for alignment. After alignment all force peaks in every FD curve were identified by discrete-wavelet-transform-based noise reduction (Daubechies, 1990; Torrence and Compo, 1998) and fitted using the WLC model (Siggia et al., 1994). The resulting force/contour length value pairs were pooled and clusters of high point density were identified using the DBSCAN algorithm (Ester et al., 1996) with the elliptical distance condition $\Delta f^2/r_f^2 + \Delta l^2/r_l^2 \leq 1$ (where Δf is the force distance between two points and Δl is the contour length distance between two points), force radius r_f of 50 pN, contour length radius r_l of 1 nm, and a core point threshold for clustering of 10 points. For every cluster the mean contour length and mean force were calculated. For all FD curves the cluster-allocation of all force peaks was determined. Data points not allocated to any cluster were neglected. The transition-frequency between all cluster pairs was calculated by counting the number of transitions between the respective clusters in all FD curves.

Acknowledgments We thank Stephanie Aoki for providing strain MG1655 and Jonne Helenius for fruitful discussions. This work was supported by the Swiss National Science Foundation (SNF; Grant 205320_160199 to D.J.M.), the Swiss Nanoscience Institute (SNI) of the University of Basel, and the ACRITAS Initial Training Network (FP7-PEOPLE-2012-ITN, Project 317348).

3.4 References

Bell, G.I. (1978). Models for the specific adhesion of cells to cells. Science 200, 618–627.

Bippes, C.A., and Muller, D.J. (2011). High-resolution atomic force microscopy and spectroscopy of native membrane proteins. Reports Prog. Phys. 74, 86601.

Bonhivers, M., Desmadril, M., and Moeck, G. (2001). Stability studies of FhuA, a two-domain outer membrane protein from Escherichia coli. Biochemistry 40, 2606–2613.

Borgia, A., Williams, P.M., and Clarke, J. (2008). Single-Molecule Studies of Protein Folding. Annu. Rev. Biochem. 77, 101–125.

Bosshart, P.D., Casagrande, F., Frederix, P.L.T.M., Ratera, M., Bippes, C. a, Müller, D.J., Palacin, M., Engel, A., and Fotiadis, D. (2008). High-throughput single-molecule force spectroscopy for membrane proteins. Nanotechnology 19, 384014.

Bosshart, P.D., Iordanov, I., Garzon-Coral, C., Demange, P., Engel, A., Milon, A., and Müller, D.J. (2012a). The transmembrane protein KpOmpA anchoring the outer membrane of Klebsiella pneumoniae unfolds and refolds in response to tensile load. Structure 20, 121–127.

Bosshart, P.D., Frederix, P.L.T.M., and Engel, A. (2012b). Reference-free alignment and sorting of single-molecule force spectroscopy data. Biophys. J. 102, 2202–2211.

Cormack, B., and Castaño, I. (2002). Introduction of point mutations into cloned genes. Methods Enzymol. 350, 199–218.

Damaghi, M., K.ster, S., Bippes, C. a, Yildiz, O., and Müller, D.J. (2011). One β hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane β-barrel protein OmpG. Angew. Chem. Int. Ed. Engl. 50, 7422–7424.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. Inf. Theory, IEEE Trans. 36, 961–1005.

Engel, A., and Gaub, H.E. (2008). Structure and mechanics of membrane proteins. Annu. Rev. Biochem. 77, 127–148.

Ester, M., Kriegel, H.P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Proc. 2nd Int. Conf. Knowl. Discov. Data Min. 96, 226–231.

Evans, E., and Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. Biophys. J. 72, 1541–1555.

H -J Butt and M Jaschke (1995). Calculation of thermal noise in atomic force microscopy. Nanotechnology 6, 1–7.

Hensen, U., and Mueller, D.J. (2013). Mechanistic explanation of different unfolding behaviors observed for transmembrane and soluble β-barrel proteins. Structure 21, 1317–1324.

Horne, J.E., and Radford, S.E. (2016). A growing toolbox of techniques for studying β-barrel outer membrane protein folding and biogenesis. Biochem. Soc. Trans. 44, 802–809.

Huysmans, G.H.M., Baldwin, S. a, Brockwell, D.J., and Radford, S.E. (2010). The transition state for folding of an outer membrane protein. Proc. Natl. Acad. Sci. U. S. A. 107, 4099–4104.

Janovjak, H., Kessler, M., Oesterhelt, D., Gaub, H., and Müller, D.J. (2003). Unfolding pathways of native bacteriorhodopsin depend on temperature. EMBO J. 22, 5220–5229.

Kedrov, A., Janovjak, H., Sapra, K.T., and Müller, D.J. (2007). Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu. Rev. Biophys. Biomol. Struct. 36, 233–260.

Koebnik, R., Locher, K., and Gelder, P. Van (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol. Microbiol. 37.

Luckey, M., and Nikaido, H. (1980). Specificity of diffusion channels produced by lambda phage receptor protein of Escherichia coli. Proc. Natl. Acad. Sci. 77, 167–171.

Möller, C., Fotiadis, D., Suda, K., Engel, A., Kessler, M., and Müller, D.J. (2003). Determining molecular forces that stabilize human aquaporin-1. J. Struct. Biol. 142, 369–378.
Müller, D.J., and Engel, A. (2007). Atomic force microscopy and spectroscopy of native membrane proteins. Nat. Protoc. 2, 2191–2197.
Müller, D., Amrein, M., and Engel, A. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. J. Struct. Biol. 188, 172–188.
Oesterhelt, F., Oesterhelt, D., Pfeiffer, M., Engel, A., Gaub, H.E., and D.J. Müller (2000). Unfolding Pathways of Individual Bacteriorhodopsins. Science. 288, 143–146.
Otto, K., Norbeck, J., Larsson, T., Karlsson, K.A., and Hermansson, M. (2001). Adhesion of type 1-fimbriated escherichia coli to abiotic surfaces leads to altered composition of outer membrane proteins. J. Bacteriol. 183, 2445–2453.
Otzen, D.E., and Andersen, K.K. (2013). Folding of outer membrane proteins. Arch. Biochem. Biophys. 531, 34–43.
Pfreundschuh, M., Martinez-Martin, D., Mulvihill, E., Wegmann, S., and Muller, D.J. (2014). Multiparametric high-resolution imaging of native proteins by force-distance curve-based AFM. Nat. Protoc. 9, 1113–1130.
Phale, P.S., Philippsen, A., Kiefhaber, T., Koebnik, R., Phale, V.P., Schirmer, T., and Rosenbusch, J.P. (1998). Stability of trimeric OmpF porin: The contributions of the latching loop L2. Biochemistry 37, 15663–15670.
Prilipov, A., Phale, P.S., Gelder, P., Rosenbusch, J.P., and Koebnik, R. (1998). Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from E. coli. FEMS Microbiol. Lett. 163, 65–72.
Renault, M., Saurel, O., Czaplicki, J., Demange, P., Gervais, V., Löhr, F., Réat, V., Piotto, M., and Milon, A. (2009). Solution state NMR structure and dynamics of KpOmpA, a 210 residue transmembrane domain possessing a high potential for immunological applications. J. Mol. Biol. 385, 117–130.
Sapra, K.T., Damaghi, M., Köster, S., Yildiz, O., Kühlbrandt, W., and Muller, D.J. (2009). One beta hairpin after the other: exploring mechanical unfolding pathways of the transmembrane beta-barrel protein OmpG. Angew. Chemie - Int. Ed. 48, 8306–8308.
Schirmer, T., Keller, T.A., Wang, Y., and Rosenbusch, J.P. (1995). Structural Basis for Sugar Translocation Through Maltoporin Channels at 3.1 Å Resolution. Science. 267, 512–514.
Schulz, G.E. (2002). The structure of bacterial outer membrane proteins. Biochim. Biophys. Acta - Biomembr. 1565, 308–317.
Serdiuk, T., Madej, M.G., Sugihara, J., Kawamura, S., Mari, S.A., Kaback, H.R., and Müller, D.J. (2014). Substrate-induced changes in the structural properties of LacY. Proc. Natl. Acad. Sci. U. S. A. 111, E1571-80.
Siggia, E., Bustamante, C., Marko, J., and Smith, S. (1994). Entropic elasticity of λ-phage DNA. Science. 265, 5–6.
Tamm, L.K., Hong, H., and Liang, B. (2004). Folding and assembly of beta-barrel membrane proteins. Biochim. Biophys. Acta 1666, 250–263.
Thoma, J., Bosshart, P., Pfreundschuh, M., and Müller, D.J. (2012). Out but not in: the large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. Structure 20, 2185–2190.
Thoma, J., Burmann, B.M., Hiller, S., and Müller, D.J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat. Struct. Mol. Biol. advance on, 795–802.

Thoma, J., Wolf, D., Blatman, D., and Hensen, U. (2017). Automated analysis of single molecule force-spectroscopy data. ETH Zurich.
Torrence, C., and Compo, G.P. (1998). A Practical Guide to Wavelet Analysis. Bull. Am. Meteor. Soc. 79, 61–78.
Wimley, W.C. (2003). The versatile β-barrel membrane protein. Curr. Opin. Struct. Biol. 13, 404–411.
Žoldák, G., and Rief, M. (2013). Force as a single molecule probe of multidimensional protein energy landscapes. Curr. Opin. Struct. Biol. 23, 48–57.

3.5 Supplementary Information

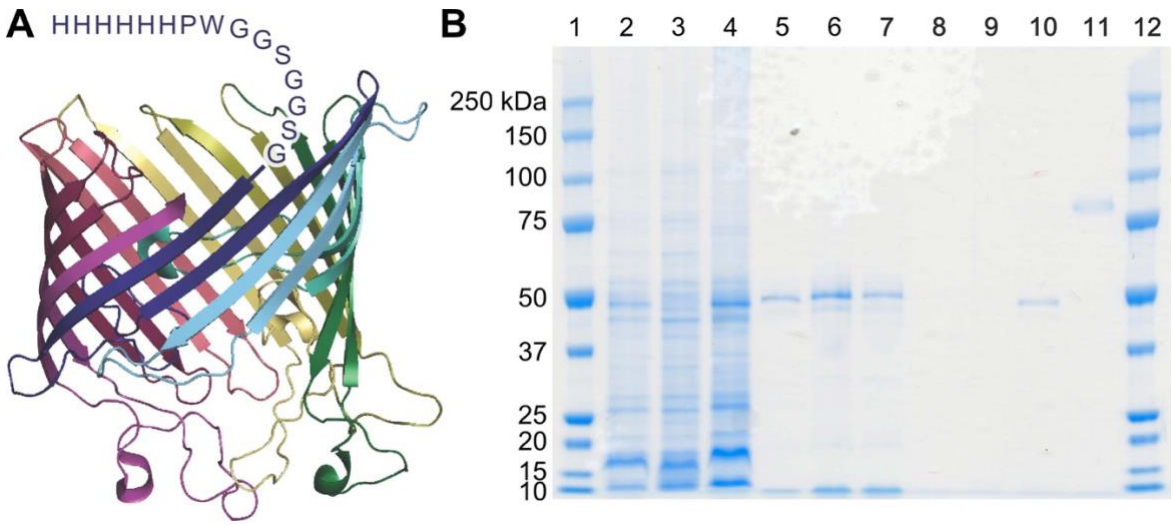


Figure S1, relates to Figure 1. Design and Purification of N-terminally Elongated LamB. (A) Illustration of N-terminally elongated LamB. The N-terminal end was elongated with the amino acid sequence HHHHHHPWGGSGGSG. (B) SDS-PAGE gel of the purification of N-terminally elongated LamB. Lanes contain the following: (1) molecular weight marker, (2) cell lysate, (3) cytosolic fraction, (4) solubilized inner membrane, (5) solubilized outer membrane, (6) cleared solubilized outer membrane, (7-9) flow-through, 1st wash, and 2nd wash of immobilized metal ion affinity chromatography, (10, 11) purified LamB, and (12) molecular weight marker. Samples in lanes 2–10 were heat-denatured prior to SDS-PAGE. The sample in lane 11 was not heat-treated, therefore LamB migrated as trimers at higher molecular weight.

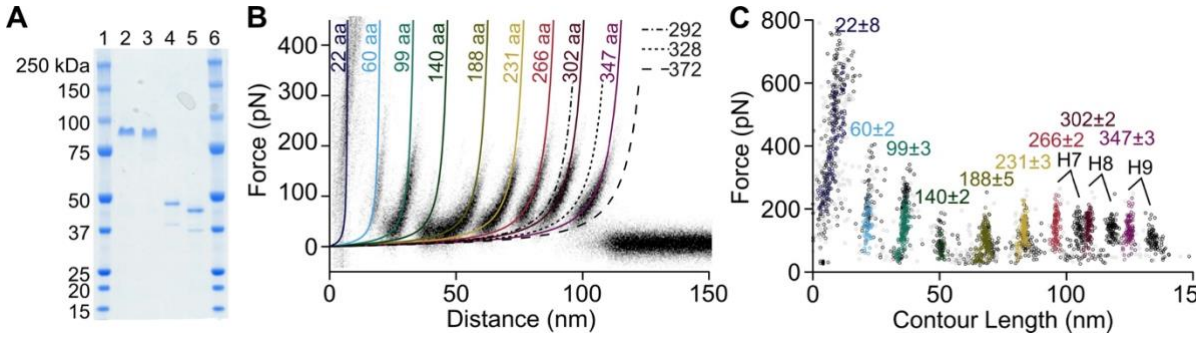
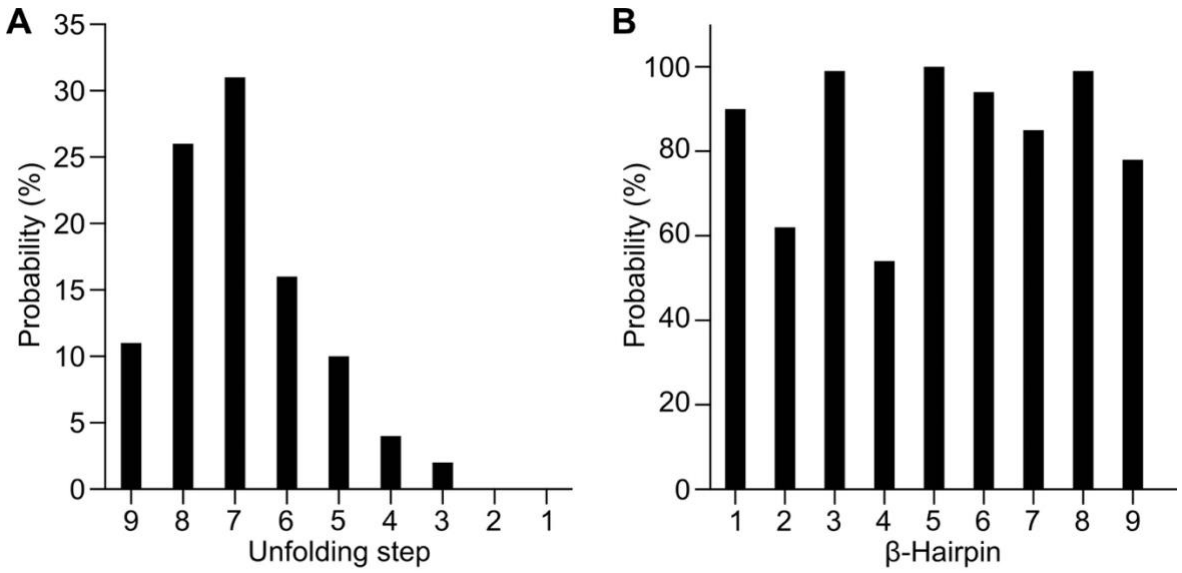
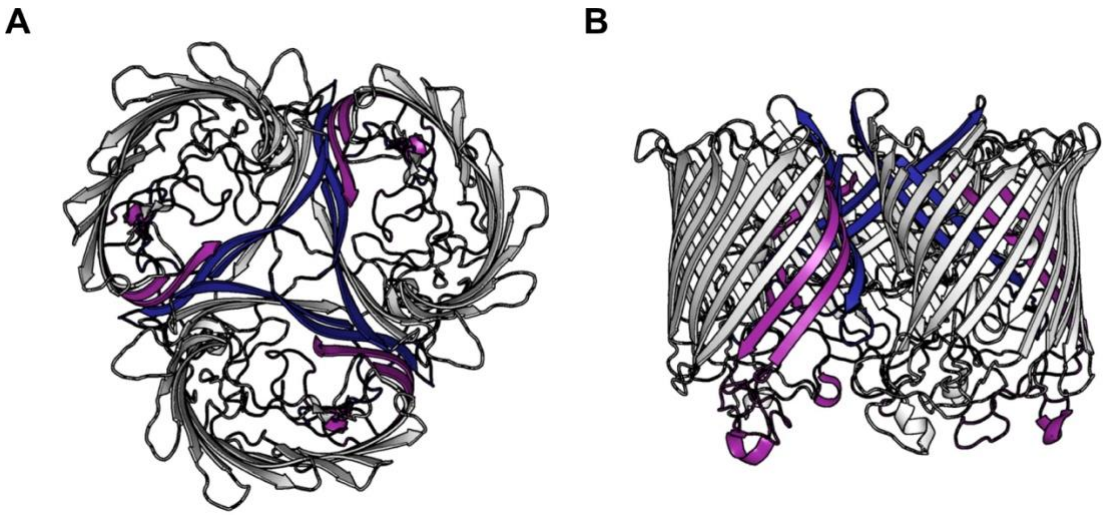


Figure S2, relates to Figure 2. Determination from which Terminal End LamB is Attached and Mechanically Unfolded by SMFS. To determine from which terminus the N-terminally elongated LamB (full-length LamB) was unfolded, 23 amino acids were deleted from loop 6 of LamB (LamBΔL6). LamBΔL6 with the shortened loop was expressed, purified and reconstituted as described for full-length LamB (Experimental Procedures). (A) SDS-PAGE gel showing purified full-length LamB (lane 2 and 4) and LamBΔL6 (lane 3 and 5). Molecular weight marker is in lanes 1 and 6. Samples in lanes 4 and 5 were heat-denatured prior to SDS-PAGE and therefore migrated at the weight of the LamB monomer. Samples in lanes 2 and 3 were not heat-treated and therefore migrated as trimers at higher molecular weight. (B) Density plot of 123 superimposed FD curves, each recorded upon mechanically unfolding of a single LamBΔL6. The common unfolding pattern of LamBΔL6 is shorter compared to full-length LamB (Figure 3A). Note that compared to full-length LamB the contour length positions of the last three unfolding force peaks are shifted by ≈ 26 aa towards shorter contour lengths. This shift is in agreement with the deletion of 23 aa of loop 6 and thus confirms that LamB was unfolded from the N-terminal end. Colored lines are worm-like chain (WLC) curves indicating the position of every force peak with contour lengths given at their ends in amino acids (aa). Black dashed lines are WLC curves indicating the positions of the last three force peaks detected for full-length LamB. (C) Unfolding force vs contour lengths of all force peaks detected in all FD curves recorded of LamBΔL6. Points of the same color belong to the same force peak class. Unclassified points are grey. Black points are data points detected for full-length LamB. Values above each force peak class provide the average contour length in amino acids ± SD. Force peak classes, which are shifted to shorter contour lengths are indicated (H7 – H9).



Supplementary Figure S3, relates to Figure 4. Unfolding Probabilities of Beta- Hairpins of LamB.
(A) Probabilities to detect 9, 8, etc. unfolding steps in an FD curve.
(B) Probability to detect beta-hairpin 1, 2, etc. in an FD curve.



Supplementary Figure S4, relates to Figure 4. Cartoon Illustrating the Position of Beta-Hairpin 1 in the Trimeric Arrangement of LamB. (A) Top view of the periplasmic side of LamB. Beta-hairpin 1 is shown in blue and beta-hairpin 9 in purple. Beta-hairpins 2 to 8 are shown in grey. (B) Side view with periplasmic surface of LamB is facing upwards. The structure of the Lamb trimer has been taken from PDB-code: 1MAL (Schirmer et al., 1995).

Unfolding step	**Contour length (aa)**		**Force (pN)**		**Number of events**
	Mean	± SD	Mean	± SD	
1	25	2	570	73	89
2	60	2	201	45	56
3	101	2	221	39	86
4	140	2	90	32	46
5	192	3	97	35	90
6	233	3	111	35	84
7	292	3	129	40	75
8	328	3	135	25	91
9	372	3	105	26	71

Supplementary Table S1, relates to Figure 3. Mean Contour length and mean force of every unfolding force peak class detected upon mechanically unfolding LamB by SMFS. Each unfolding force peak corresponds to one unfolding step of LamB.

4 POTRA Domains, Extracellular Lid, and Membrane Composition Modulate the Conformational Stability of the β Barrel Assembly Factor BamA

Johannes Thoma,[1] Yang Sun,[1,2] Noah Ritzmann,[1] and Daniel J. Müller[1,3,*]

[1]Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH) Zürich, Mattenstrasse 26, Basel 4058, Switzerland
[2]Collaborative Innovation Center of Advanced Microstructures, National Laboratory of Solid State Microstructure, Department of Physics, Nanjing University, 22 Hankou Road, Nanjing, Jiang Su 210093, China PR
[3]Lead Contact

***Correspondence:** daniel.mueller@bsse.ethz.ch

Author Contributions
J.T. cloned, purified and prepared all protein samples. J.T., Y.S., and N.R. performed SMFS of full-length BamA. J.T. performed SMFS of all other protein constructs and analyzed all data. J.T. and D.J.M. wrote the paper.

Published in: *Structure*, 26, 987-996, July3, 2018
Doi: 10.1016/j.str.2018.04.017

4.1 Summary

The core component BamA of the β barrel assembly machinery (BAM) adopts several conformations, which are thought to facilitate the insertion and folding of β barrel proteins into the bacterial outer membrane. Which factors alter the stability of these conformations remains to be quantified. Here, we apply single-molecule force spectroscopy to characterize the mechanical properties of BamA from *Escherichia coli*. In contrast to the N-terminal periplasmic polypeptide-transport-associated (POTRA) domains, the C-terminal transmembrane β barrel domain of BamA is mechanically much more stable. Exposed to mechanical stress this β barrel stepwise unfolds β hairpins until unfolding has been completed. Thereby, the mechanical stabilities of β barrel and β hairpins are modulated by the POTRA domains, the membrane composition and the extracellular lid closing the β barrel. We anticipate that these differences in stability, which are caused by factors contributing to BAM function, promote conformations of the BamA β barrel required to insert and fold outer membrane proteins.

4.2 Introduction

Transmembrane β barrel proteins equip outer membranes of Gram-negative bacteria, mitochondria, and chloroplasts with functional versatility, including cell adhesion, molecular transport, and signal transduction. Many structures of β barrel proteins have been solved thus far providing detailed insight into their unique structure-function relationship (Buchanan, 1999; Koebnik et al., 2000). However, how β barrel proteins insert and fold into cellular membranes remains largely unknown. In vivo, the folding and membrane insertion of these proteins is facilitated by members of the Omp85 protein family, including the mitochondrial Sam50, chloroplast Toc75, and bacterial FhaC, TamA, and BamA (Gruss et al., 2013; Hagan et al., 2011; Höhr et al., 2015; Jacob-Dubuisson et al., 2009; Plummer and Fleming, 2016; Rollauer et al., 2015; Selkrig et al., 2012). Most Omp85 family members form complexes with one or more accessory lipoproteins, the functions of which are not entirely clear yet. The structurally and functionally well-studied β barrel assembly machinery (BAM) from *Escherichia coli* comprises the 88 kDa transmembrane BamA and four lipoproteins BamB-E, of which only BamA and BamD are essential for outer membrane protein (Omp) insertion and folding (Ruiz et al., 2006; Wu et al., 2005). Like most Omp85 family members, BamA from *E. coli* comprises an N-terminal periplasmic domain and a C-terminal transmembrane β barrel domain (Figure S1). Whereas the N-terminal domain comprises five functionally important polypeptide-transport-associated (POTRA) domains, the β barrel domain is assembled by 16 β strands. Among Omp85 homologs the number of POTRA domains varies from one to seven, however, in vivo a single POTRA domain is sufficient for viability (Bos et al., 2007). The BamA β barrel domain accommodates a water-filled cavity, which is accessible from the periplasmic side, but sealed toward the extracellular space by a lid-like structure, formed by the extended extracellular loops 4 and 6 (Leonard-Rivera and Misra, 2012; Maier et al., 2015). The lid is held in place by the so-called lid-lock, a conserved VRGF motif in loop 6, of which arginine R661 forms a strong interaction with the inner wall of the β barrel (Maier et al., 2015). Replacement of R661 by an alanine (R661A) disrupts this locking mechanism and considerably reduces the in vivo insertion levels of Omps, whereas deletion of the entire VRGF motif is lethal (Leonard-Rivera and Misra, 2012).

The molecular mechanism by which BAM, and in particular BamA, facilitates the insertion of Omps remains poorly understood (Konovalova et al., 2017). Based on existing evidence from structural, biochemical, and genetic studies, the current discussion revolves largely around two models describing the insertion and folding of Omps in vivo. The budding model describes the transient formation of a hybrid β barrel between BamA and the incoming Omp polypeptide by beta-augmentation and subsequent release of the Omp by budding (Gruss et al., 2013; Noinaj et al., 2013, 2014). This model is supported by experimental results suggesting that the β strands 1 and 16 of the β barrel form a mechanically flexible lateral gate to release the Omp polypeptide toward the hydrophobic core of the membrane (Höhr et al., 2018). The BamA-assisted model suggests that the unfolded Omp polypeptide inserts by a self-guided thermodynamic mechanism into a locally thinned and disturbed membrane adjacent to the β barrel of BamA (Fleming, 2015; Gessmann et al., 2014; Patel and Kleinschmidt, 2013). Both models appear to be experimentally supported since in vivo studies show that BamA is functionally impaired by introducing disulfide bonds linking the two β strands 1 and 16 forming the lateral gate of the β barrel (Noinaj et al., 2014), and in vitro studies show that BamA having the lateral gate locked can still facilitate the insertion and folding of Omps (Doerner and Sousa, 2017). Whereas the majority of β barrel-forming Omps prevail in a closed conformation, with their first and last β strand forming a tight connection, β strands 1 and 16 of the β

barrel domain of BamA interact loosely (Noinaj et al., 2013). This loose seam of the β barrel is thought to form a flexible lateral gate and to permit register-sliding between β strands 1 and 16 (Doerner and Sousa, 2017; Noinaj et al., 2014). Recently solved structures of the entire BAM complex in the absence of a membrane revealed unusual arrangements of the β barrel seam, in which the β barrel is partially twisted resulting in β strands 1 and 16 facing each other at an angle of ≈45° (Gu et al., 2016; Han et al., 2016; Iadanza et al., 2016). BamA is, furthermore, an unusual Omp, as the hydrophobic belt surrounding the transmembrane β barrel is thinned at the seam forming the lateral gate, which is thought to interact with and locally destabilize the membrane (Fleming, 2015; Noinaj et al., 2013). However, whether the membrane can stabilize certain BamA conformations has not been studied. In summary, to which extent these three factors, the POTRA domains, extracellular lid-lock, and membrane, can alter the mechanical stability of the BamA structure toward supporting conformational variability remains to be characterized.

Atomic force microscopy (AFM)-based single-molecule force spectroscopy (SMFS) has been widely applied to study the mechanical properties of single membrane proteins as well as their unfolding and folding pathways (Engel and Gaub, 2008). In SMFS a mechanical pulling force applied to one terminal end of a membrane protein results in the sequential unfolding of the protein (Oesterhelt et al., 2000). This mechanical unfolding of membrane proteins typically occurs via a series of unfolding steps, which describe the distinctive unfolding pathways taken by the protein (Kedrov et al., 2007; Oesterhelt et al., 2000; Sapra et al., 2009). Recently, Omps received particular attention by SMFS, as their mechanical unfolding pathways are tightly linked to their secondary structure (Bosshart et al., 2012a; Sapra et al., 2009; Thoma et al., 2012, 2017). Commonly, the unfolding steps of transmembrane β barrel proteins are shaped by individual β hairpins, which has been observed for the 14-stranded β barrel of OmpG (Sapra et al., 2009), the 18-stranded β barrel of LamB (Thoma et al., 2017), the 19-stranded voltage-gated anion channel (VDAC) (Ge et al., 2016), as well as the 22-stranded β barrel of FhuA (Thoma et al., 2012). The unfolding patterns of transmembrane β barrel proteins recorded by SMFS are highly reproducible and sensitive to the fold and to the functional state. For example, SMFS can directly detect the misfolding of individual β strands of a β barrel (Thoma et al., 2012, 2015), the conformational change or the extension of a polypeptide loop connecting two β hairpins (Damaghi et al., 2010a; Sapra et al., 2009), or the conformational flexibility of transmembrane β strands and β barrels (Damaghi et al., 2010b; Ge et al., 2016).

Here, we apply SMFS to characterize the mechanical properties of BamA dependent on three factors having functional implications, the POTRA domains, the membrane environment, and the extracellular lid-lock. Our experiments show that, exposed to mechanical stress, the transmembrane β barrel domain of BamA unfolds stepwise via individual β hairpins until the entire β barrel unfolded. We find that the β barrel domain in the absence of the POTRA domains shows lower mechanical stability, suggesting a stabilizing effect of the POTRA domains. Furthermore, we observe that the composition of the membrane affects the mechanical stability of the β barrel. Finally, our experiments suggest a coupling between the extracellular lid sealing the β barrel and the stability of the seam of the BamA β barrel.

4.3 Results

Exposed to Mechanical Stress BamA Unfolds Stepwise

To characterize the stability of BamA in the native membrane, we prepared outer membrane vesicles (OMVs) from *E. coli* strain BL21(DE3)omp8, enriched in full-length BamA (Figures S2 and S3A) (Thoma et al., 2018). For SMFS, OMVs enriched with BamA were adsorbed to freshly cleaved mica, upon which the vesicles opened and exposed the inner, periplasmic surface (Figure S3B). After locating a membrane patch by AFM imaging, the AFM stylus was pushed onto the membrane protein assemblies, applying a non-destructive force of ≈1 nN to unspecifically attach a single BamA to the AFM stylus (Figure 1A) (Müller and Engel, 2007). Subsequent retraction of the AFM stylus applied a pulling force to the attached polypeptide, which induced the mechanical unfolding of the membrane protein. During retraction, the force deflecting the AFM cantilever and the distance traveled by the cantilever were recorded in a force-distance (FD) curve (Figure 1B). Individual force peaks in an FD curve marked the mechanically induced unfolding steps of BamA. To ensure that only fully unfolded BamA was analyzed, we selected FD curves exceeding a length of 95 nm, which corresponds to at least two-thirds of the contour length of the fully unfolded and stretched β barrel domain of BamA.

FD curves recorded upon mechanically unfolding single BamA were variable in length (Figure 1B). However, all FD curves showed a characteristic pattern of up to nine force peaks toward the end. This regular pattern was preceded by a region that varied in length and contained one or more force peaks. All 151 FD curves of full-length BamA were aligned in distance by superimposing their last force peaks (Figure 2A and STAR Methods). The superimposition highlighted the distinct unfolding pattern of nine unfolding force peaks at distances >100 nm. In contrast, the beginning of the superimposition (≤100 nm) showed a variable region, which did not show a reproducible pattern. It has been previously shown that membrane proteins exposing soluble domains preferentially adhere to an AFM stylus with these domains (Bosshart et al., 2012a). We thus assumed that, in our experiment, the AFM stylus pushed onto BamA preferentially adhered to the POTRA domains (Figure 1A). As the AFM stylus could adhere to any of the five POTRA domains this would explain the FD curves having different lengths and supports our assumption that BamA was unfolded from the N-terminal end.

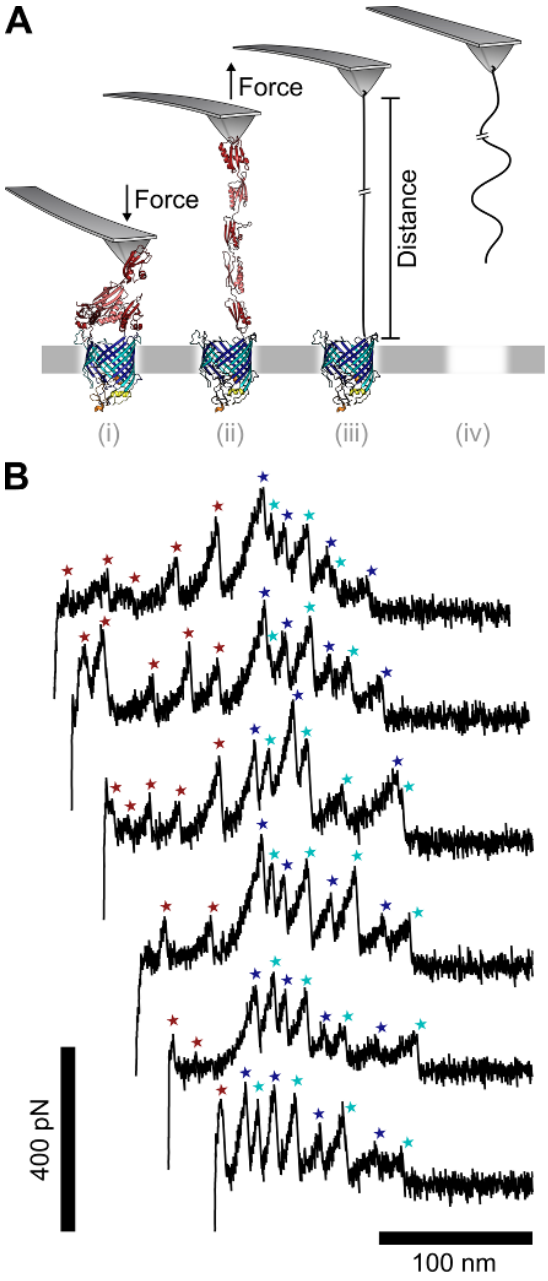


Figure 1. Mechanical Unfolding of Full-Length BamA by SMFS (A) Schematic of the mechanical unfolding experiment of BamA by AFM-based SMFS: (i) pushing the AFM stylus onto a single BamA unspecifically attaches the N-terminal polypeptide. The POTRA domain of BamA is colored red, the b barrel domain blue, the extracellular loop 4 yellow and the extracellular loop 6 orange (for structural details see Figure S1). (ii) Subsequent retraction of the AFM stylus applies a pulling force, which stretches BamA and (iii–iv) induces the stepwise mechanical unfolding of the protein. During retraction, the force required to unfold BamA is recorded in force-distance (FD) curves. (B) Exemplary FD curves recorded upon mechanically unfolding single full-length BamA. Individual force peaks are marked by colored asterisks and represent the unfolding steps of BamA. See also Figures S1 and S2.

To analyze the unfolding steps and the unfolding pathway of BamA, every unfolding force peak in every FD curve was fitted with the worm-like chain (WLC) model (Marko and Siggia, 1995). For every force peak, the rupture force was plotted against the contour length, and classes of recurring force peaks were identified using a density-based clustering approach (Figure 2B) (Ester et al., 1996; Thoma, 2017). For every force peak class, the mean rupture force and mean contour length were calculated. In addition, the frequency of occurrence of individual force peak classes and the transition frequencies between the classes were analyzed (Thoma, 2017; Yu et al., 2017). The unfolding force peaks of each force peak class represent one unfolding step, with all unfolding steps of all force peak classes describing the unfolding pathway of BamA (Figure 2C). The mechanical unfolding pathway of BamA was characterized by a heterogeneous unfolding pattern in the region preceding a contour length of ≈403 aa (≈145 nm) followed by a uniform unfolding pattern with nine discrete unfolding steps. On average, the unclassified force peaks contributing to the heterogeneous unfolding pattern were <100 pN (Figure 2B). In

contrast, the mean forces required to initiate unfolding of the first four unfolding steps of the uniform unfolding pattern ranged from ≈180 to 210 pN. The mean forces initiating the following five unfolding steps ranged from ≈80 to 120 pN (Table S1). The probability to detect any of the unfolding steps of the uniform unfolding pattern was almost 1, except for the seventh and ninth unfolding step, which were detected with probabilities of 0.5 and 0.4.

Based on our assumption that BamA was unfolded starting from the POTRA domains, we assigned the structural segments unfolded in each unfolding step to the structure of BamA. To this end we subtracted the mean contour length of every force peak class from the N-terminal end of BamA (Figure 2D). The assignment showed that each force peak class marked the unfolding step of a single β hairpin, except for the seventh force peak class, which marked an additional unfolding step in the extracellular loop of β hairpin 6. The unfolding pathway of BamA is therefore characterized by the stepwise unfolding of individual β hairpins, with the exception of β hairpin 6, which unfolded in two steps with a probability of 0.5, and β hairpin 8, which unfolded together with β hairpin 7 with a probability of 0.6. It has been previously shown that the unfolding pathways of Omps are largely shaped by β hairpins (Bosshart et al., 2012a; Ge et al., 2016; Sapra et al., 2009; Thoma et al., 2012, 2017). The observed unfolding pathway of BamA is consistent with these findings.

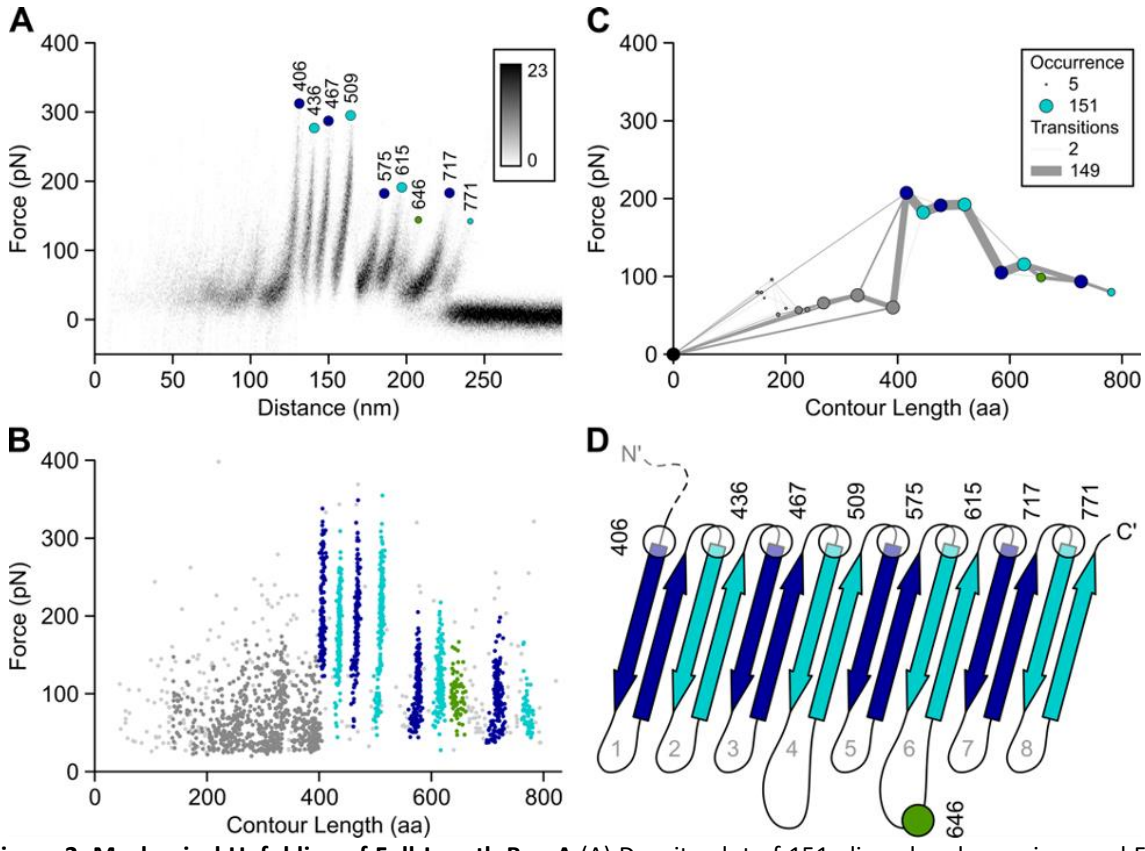


Figure 2. Mechanical Unfolding of Full-Length BamA (A) Density plot of 151 aligned and superimposed FD curves highlighting recurring force peaks. The color scale indicates the number of data points per bin. Colored circles mark the common unfolding force peaks of BamA, with their mean contour lengths given in amino acids (aa). (B) Unfolding force versus contour length of all force peaks detected in all FD curves. Data points of the same color indicate single force peaks belonging to the same force peak class. Unclassified points are gray. Each force peak class corresponds to one unfolding step. (C) Mechanical unfolding pathway of BamA. Colored circles represent mean forces and mean contour lengths of the force peak classes found in (B). The size of each circle indicates the probability of occurrence of the force peak class each representing an

unfolding step of BamA. Gray lines indicate transitions between unfolding steps, with the line thickness indicating the probability of transition. (D) Secondary structure cartoon of BamA showing the locations of predominant mechanical unfolding steps (circles). The mean contour length of each unfolding step is given in amino acids. β Hairpins are numbered 1–8 starting from the N-terminal end. Gray circles indicate the five POTRA domains. See also Figures S2 and S3 and Table S1.

Without POTRA Domains the BamA β Barrel Is Less Stable

To test our assignment of the unfolding steps of full-length BamA we engineered a BamA variant, BamAΔP, lacking the N-terminal POTRA domains. We then prepared OMVs enriched in BamAΔP, which were adsorbed to mica, imaged by AFM, and mechanically unfolded, as described for full-length BamA (Figure S4A). In contrast to the above experiments using full-length BamA, the FD curves recorded upon unfolding of only the β barrel domain were shorter in distance and showed a homogeneous distribution of force peaks. Superimposing 64 aligned FD curves revealed a distinct unfolding pattern of nine unfolding force peaks, which was similar to the pattern observed for the β barrel domain of full-length BamA (Figure 3A). Every unfolding force peak in every FD curve was fitted using the WLC model, and the resulting rupture force versus contour length value pairs were plotted and clustered (Figure 3B). For every force peak class, we determined the mean force, mean contour length, probability of occurrence, and the probability to transition to another force peak class (Figure 3C).

Compared to full-length BamA, the mean contour lengths of all force peak classes detected for BamAΔP were shifted to shorter lengths by ≈140 nm (≈389 aa). This shift corresponds to the length difference of 379 aa between full-length BamA and BamAΔP, and hence confirms our assumption that full-length BamA was unfolded by mechanically pulling the N-terminal end. We thus assigned the unfolding steps to the structure of the β barrel domain of BamA by subtracting the mean contour length of every force peak class from the N-terminal end of BamAΔP (Figure 3D). In agreement with our previous experiment, the assignment showed that each force peak class marked the unfolding step of a single β hairpin, with an additional unfolding step located in the extracellular loop of β hairpin 6. In agreement with our finding for full-length BamA, the mechanical unfolding pathway of the β barrel domain was characterized by seven prominent unfolding steps occurring at high probabilities (≈1) and marking the unfolding of β hairpins 1–7 (Table S1). Again, the seventh and ninth unfolding step of BamAΔP occurred at lower probabilities of 0.4 and 0.3, respectively, showing that β hairpin 6 could unfold in two unfolding steps, and β hairpin 8 preferentially unfolded together with β hairpin 7. The mean forces required to unfold the first four N-terminal β hairpins 1 to 4 of BamAΔP ranged from ≈135 to 180 pN, and thus lay considerably below the unfolding forces observed for full-length BamA (Table S1). The mean forces required to unfold the remaining five unfolding steps of BamAΔP remained at similar range (≈80 to 130 pN) as determined for full-length BamA. Taken together, our experiments show that deletion of the POTRA domains does not change the unfolding steps formed by the individual β hairpins of BamA. However, the forces required to unfold the first four β hairpins were lower, suggesting that the POTRA domains modulate the mechanical stability of the β barrel domain.

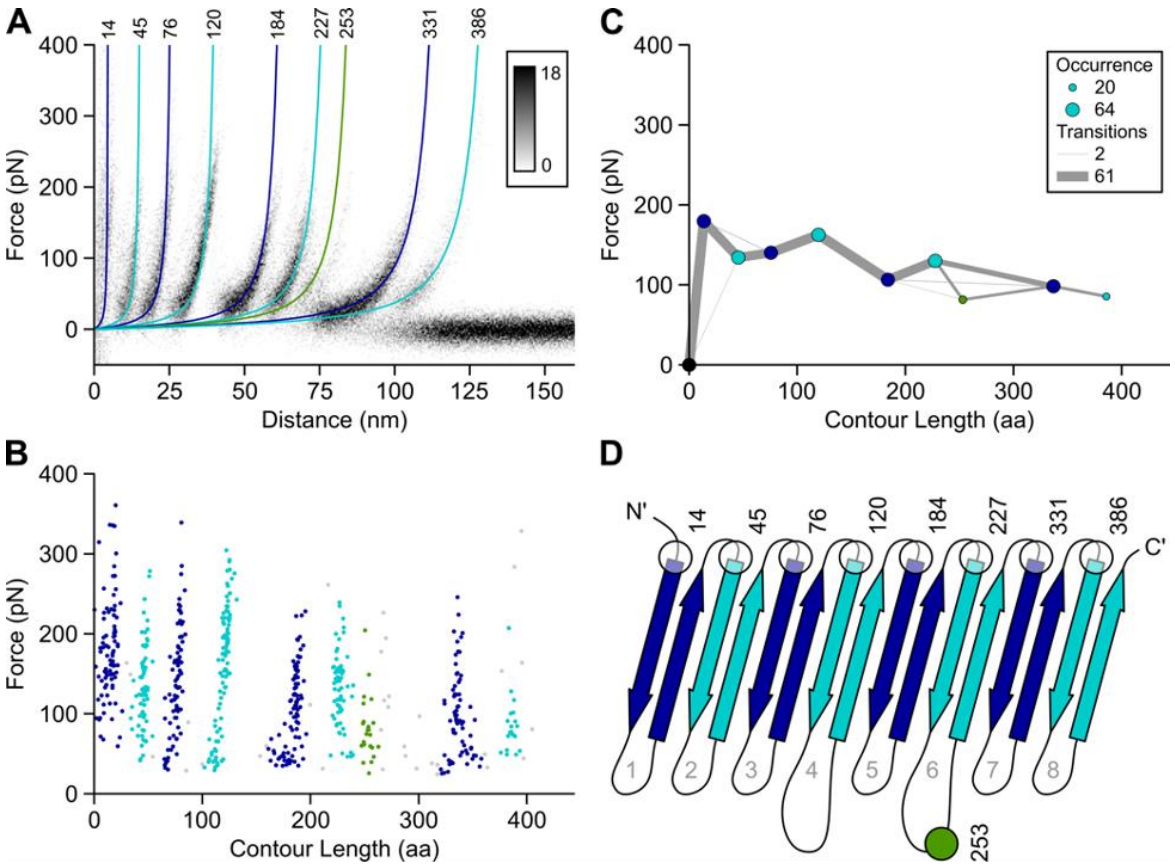


Figure 3. Mechanical Unfolding of the β Barrel Domain of BamA in the Absence of POTRA Domains (BamAΔP)
(A) Density plot of 64 aligned and superimposed FD curves highlighting recurring force peaks. The color scale indicates the number of data points per bin. Colored lines are worm-like chain (WLC) curves indicating the mean contour length of common unfolding force peaks, with their mean contour lengths given in amino acids (aa). (B) Unfolding force versus contour length of all force peaks detected in all FD curves. Data points of the same color belong to the same force peak class. Unclassified points are gray. Each force peak class corresponds to one unfolding step. (C) Mechanical unfolding pathway of BamA. Colored circles represent mean forces and mean contour lengths of the force peak classes found in (B), the size of each circle indicates the probability of occurrence of the force peak class. Gray lines indicate transitions between unfolding steps and the line thickness corresponds to the probability of transiting between unfolding steps. (D) Secondary structure cartoon of BamA showing the locations of predominant mechanical unfolding steps (circles). The mean contour length of each unfolding step is given in amino acids. β Hairpins are numbered 1–8 starting from the N-terminal end. See also Figure S4A and Table S1.

BamA β Barrel Stability Depends on Membrane Composition

Next, we wondered whether the membrane composition could influence the mechanical stability of the β barrel of BamA. To this end we purified and reconstituted BamA in *E. coli* polar lipid (EPL) membranes. However, we could not reconstitute full-length BamA but only BamAΔP in EPL liposomes at sufficiently high protein densities required for SMFS (Figure S2). Proteoliposomes containing BamAΔP were adsorbed to mica, imaged by AFM, and mechanically unfolded as described for BamA constructs in OMVs. FD curves recorded upon unfolding single BamAΔP from EPL membranes showed a homogeneous distribution of force peaks, similar to the ones recorded upon unfolding BamAΔP from the native membranes of OMVs (Figure S4B). Superimposing 93 aligned FD curves revealed a distinct pattern of ten unfolding force peaks (Figure 4A). As described above, every unfolding force peak in every FD curve was fitted using the WLC model and the resulting force versus contour length value pairs were plotted and clustered (Figure 4B). For every force peak class, the mean force, mean contour length, probability of occurrence, and the

probability to transition to another force peak class were determined (Figure 4C). To assign the unfolding steps to the BamA structure, the mean contour lengths of the force peak classes were subtracted from the N-terminal end of BamAΔP (Figure 4D).
Similar to the unfolding experiments of BamAΔP in OMVs, the unfolding pathway of BamAΔP in EPL membranes was characterized by seven predominant unfolding steps, each of which occurred at high probability (≈1) (Table S1). Each of these unfolding steps described the unfolding of one β hairpin 1–7. Three additional minor unfolding steps occurred at lower probabilities of 0.2, 0.5, and 0.2. The first less-probable unfolding event located in the extracellular loop 4 was not observed for full-length BamA or for BamAΔP in OMVs. It suggested that β hairpin 4 of BamAΔP could unfold in two steps, when unfolded from EPL membranes. The second minor unfolding step indicated that β hairpin 6 could unfold in two unfolding steps at a probability of 0.5, and the third minor unfolding step indicated that β hairpin 8 preferentially unfolded together with β hairpin 7. The second and third minor unfolding steps were similar to observations for the unfolding of full-length BamA or BamAΔP in OMVs. Except for the newly detected unfolding step of loop 4, the mean contour lengths of all force peak classes were at the same positions as determined (experimental accuracy ≈± 1.4 nm [4 aa]) upon unfolding BamAΔP from the native membrane of OMVs. The mean forces required to unfold individual β hairpins ranged from ≈170 to 240 pN for the first four unfolding steps, and from ≈80 to 150 pN for the remaining unfolding steps (Table S1). These forces were in a similar range as the ones determined for full-length BamA, but higher than the ones determined for BamAΔP in the native outer membrane. In summary, our experiments show that the membrane environment did not have a strong influence on the structural elements forming the unfolding steps of BamA, with the exception of the newly detected unfolding step in loop 4. However, the unfolding forces required to unfold individual β hairpins were higher than the ones determined for BamAΔP in the native membrane, suggesting that the β barrel of BamA is more stable in EPL bilayers, than in the native outer membrane.

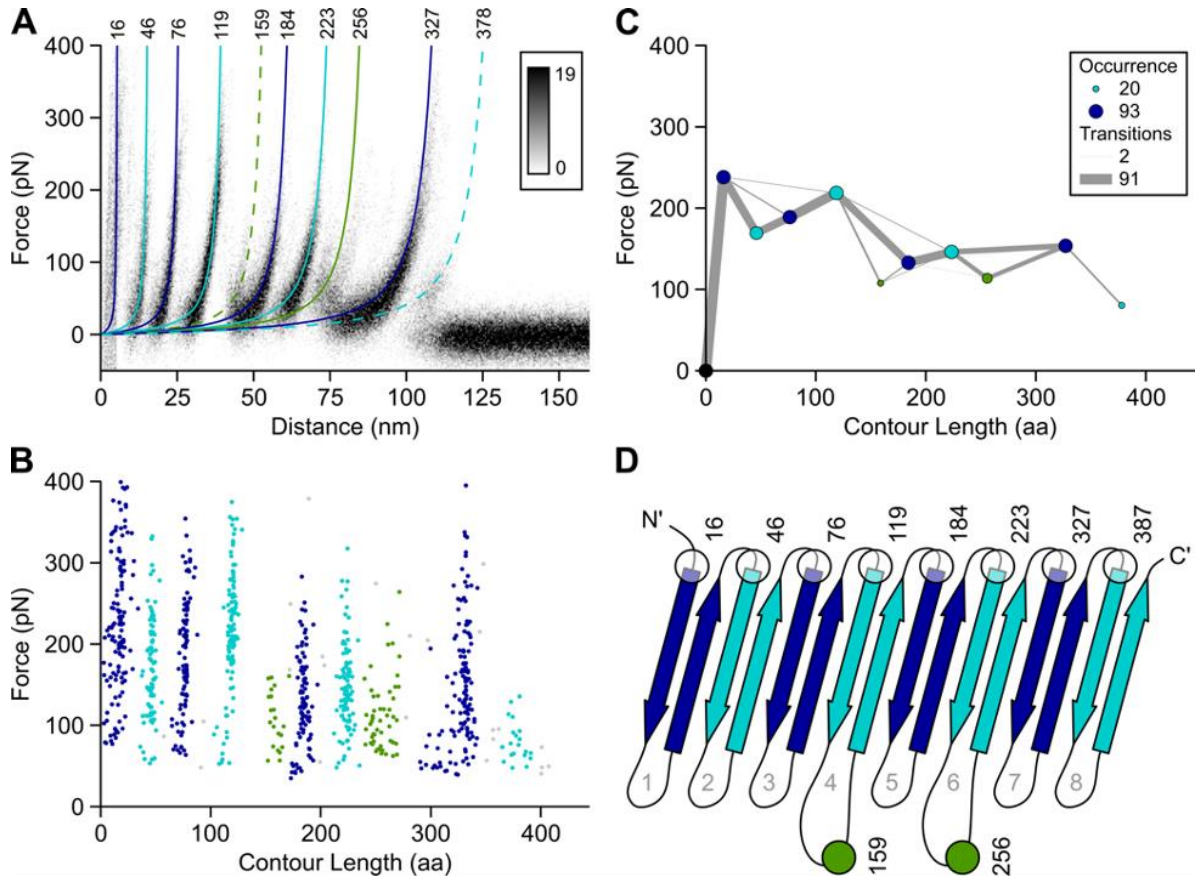


Figure 4. Mechanical Unfolding of the β Barrel Domain of BamA (BamAΔP) Reconstituted into *E. coli* Polar Lipid Bilayers (A) Density plot of 93 aligned and superimposed FD curves highlighting recurring force peaks. The color scale indicates the number of data points per bin. Colored lines are WLC curves indicating the mean contour length of common unfolding force peaks, with their mean contour lengths given in amino acids (aa). (B) Unfolding force versus contour length of all force peaks detected in all FD curves. Data points of the same color belong to the same force peak class. Unclassified points are gray. Each force peak class corresponds to one unfolding step. (C) Mechanical unfolding pathway of BamAΔP. Colored circles represent mean forces and mean contour lengths of the force peak classes found in (B), the size of each circle indicates the probability of occurrence of the force peak class. Gray lines indicate transitions between unfolding steps and the line thickness corresponds to the probability of transiting between unfolding steps. (D) Secondary structure cartoon of BamAΔP showing the locations of predominant mechanical unfolding steps (circles). The mean contour length of each unfolding step is given in amino acids. β Hairpins are numbered 1–8 starting from the N-terminal end. See also Figure S4B and Table S1.

Lid-Lock Modulates the Stability of the BamA β Barrel

In the above unfolding experiments, we localized one unfolding step in extracellular loop 6, which coincides with the position of the lid-lock structure of BamA. We thus wondered, whether the unfolding step of loop 6 could have been caused by interactions of the lid-lock (Figure S1D). To test this hypothesis, we introduced mutation R661A into the β barrel domain of BamAΔP, which had been shown to disrupt the locking mechanism and to considerably reduce the insertion levels of Omps in vivo (Leonard- Rivera and Misra, 2012). After expression of BamAΔP/R661A in *E. coli*, the protein was purified and reconstituted in EPL liposomes (Figure S2). Proteoliposomes containing BamAΔP/R661A were adsorbed to mica, imaged by AFM, and mechanically unfolded as described for all other BamA constructs.

Individual FD curves recorded upon unfolding single BamAΔP/R661A showed highly reproducible force peak patterns (Figure S4C). The superimposition of 35 aligned FD curves revealed a distinct unfolding pattern of nine unfolding force peaks, which was

similar to that observed for BamAΔP in EPL membranes (Figure 5A). Again, every unfolding force peak in every FD curve was fitted using the WLC model, and the resulting force versus contour length value pairs were plotted and clustered (Figure 5B). For every force peak class the mean force, mean contour length, probability of occurrence, and the probability to transition to another force peak class were determined (Figure 5C). The unfolding steps were then assigned to the structure of BamA by subtracting the mean contour length of every force peak class from the N-terminal end of BamAΔP/R661A (Figure 5D). The mechanical unfolding pathway of BamAΔP/R661A was highly similar to the one previously determined for BamAΔP, and the mean contour lengths of their force peak classes were located at the same distances (≈± 2 nm [6 aa]) (Table S1). Most of the unfolding steps of BamAΔP/R661A were shaped by individual β hairpins and occurred at probabilities approaching 1. One exception was observed for β hairpin 8, which unfolded in one step at a probability of 0.2, or together with β hairpin 7 at a probability of 0.8. The unfolding step in extracellular loop 6 occurred at a probability of 0.2, which was lower than detected for full-length BamA and BamAΔP (0.4 and 0.5, respectively). The mean forces required to unfold the first four β hairpins ranged from ≈140 to 190 pN and from ≈70 to 150 pN for the remaining β hairpins (Table S1). Our experiment thus showed that mutating the lid-lock did not have a strong influence on whether β hairpins formed unfolding steps of BamA. However, compared with unfolding BamAΔP from the same lipid membrane, the probability of β hairpin 6 to unfold in two unfolding steps reduced, and the additional unfolding step formed by extracellular loop 4 disappeared. Generally, the β barrel unfolded at lower forces, suggesting that introducing mutation R661A mechanically destabilized the BamA β barrel.

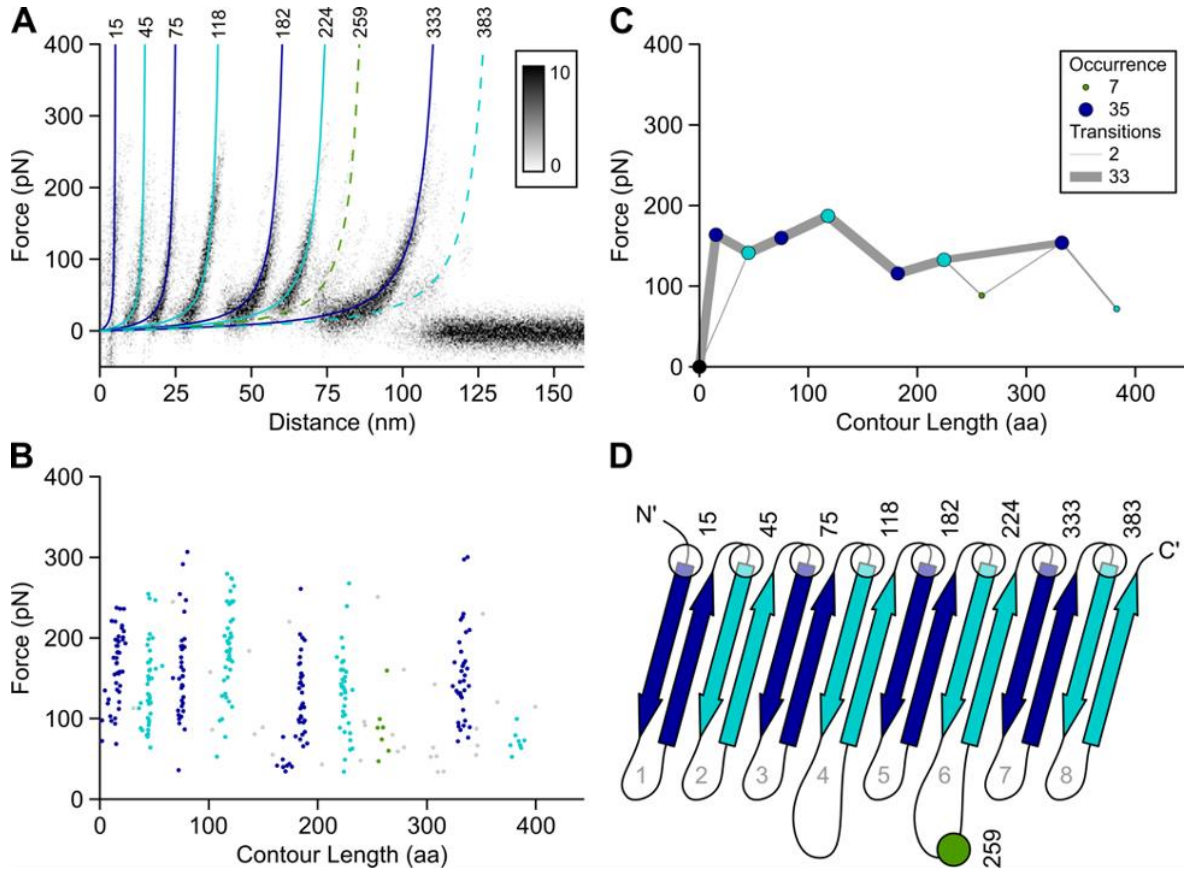


Figure 5. Mechanical Unfolding of the BamA β Barrel Domain Having Loop 6 Mutated by R661A (BamAΔP/R661A) (A) Density plot of 35 aligned and superimposed FD curves highlighting recurring force peaks. The color scale indicates the number of data points per bin. Colored lines are WLC curves indicating the mean

contour length of common unfolding force peaks, with their mean contour lengths given in amino acids (aa). (B) Unfolding force versus contour length of all force peaks detected in all FD curves. Data points of the same color belong to the same force peak class. Unclassified points are gray. Each force peak class corresponds to one unfolding step. (C)Mechanical unfolding pathway of BamA$_{\Delta P/R661A}$. Colored circles represent mean forces and mean contour lengths of the force peak classes found in (B), the size of each circle indicates the probability of occurrence of the force peak class. Gray lines indicate transitions between unfolding steps, and the line thickness corresponds to the probability of transiting between unfolding steps. (D) Secondary structure cartoon of BamA$_{\Delta P/R661A}$ showing the locations of predominant mechanical unfolding steps (circles). The mean contour length of each unfolding step is given in amino acids. β Hairpins are numbered 1–8 starting from the N-terminal end. The gray square indicates the conserved VRGF motif in loop 6. See also Figure S4C and Table S1.

4.4 Discussion

β Hairpins of the BamA β Barrel Are Mechanically Stable Units

Here we characterized the mechanical unfolding pathways of three BamA constructs (full-length BamA, BamAΔP, and BamAΔP/R661A) in two membrane environments (native OMV and EPL). The different conditions did not change the prominent unfolding steps along the unfolding pathway of BamA. The prominent unfolding steps were formed by the eight β hairpins assembling the transmembrane β barrel of BamA. Our experiments thus show that the β hairpins forming the transmembrane β barrel of BamA form mechanically stable units. Similarly, discrete unfolding steps formed by individual β hairpins of transmembrane β barrels of OmpG, LamB, FhuA, VDAC, and Tsx have been described previously during mechanical unfolding (Ge et al., 2016; Sapra et al., 2009; Thoma et al., 2012, 2017, 2018). However, compared with other β barrel proteins the mean forces required to mechanically unfold β hairpins of BamA, which ranged from ≈100 to 200 pN, were relatively low. In other β barrel-forming Omps these forces could reach 800 pN, in particular when mechanically unfolding the first β hairpin (Bosshart et al., 2012a; Sapra et al., 2009; Thoma et al., 2012, 2017). It has been reported that the forces required to unfold the first β hairpin of a β barrel initiate the mechanical unfolding process of stably folded Omps (Hensen and Müller, 2013). This force initiating the stepwise unfolding is comparably low in BamA, which, together with the generally low unfolding forces observed for the other β hairpins 2–8, indicates an overall reduced mechanical stability of the β barrel of BamA, especially in the region of the β barrel seam. Such indication is further supported by the reduced probability to detect β hairpin 8 in a separate unfolding step, which predominantly unfolded together with β hairpin 7. This finding, which suggests that β hairpin 8 interacts only weakly with the membrane, is supported by BamA structures showing β hairpin 8 being shorter than the other β hairpins and kinked toward the lumen of the β barrel (Albrecht et al., 2014; Noinaj et al., 2013).

Factors Modulating the Stability of the BamA β Barrel

Whereas the unfolding steps of BamA were the same in all experiments, the forces required to mechanically unfold individual β hairpins differed depending on the BamA construct and membrane environment. To identify these differences, we related the mean unfolding forces detected for each experimental condition to the BamA β barrel (Figures 6 and S5). Since quantification of the mechanical stability would require SMFS at multiple unfolding speeds in dynamic force spectroscopy mode (Evans and Ritchie, 1997; Heymann and Grubmüller, 2000), we here use the mean unfolding force to indicate differences in the mechanical stability of individual structural segments. Our analysis reveals that deletion of the POTRA domains mainly affected the mechanical stability of β

hairpins 1–4 (Figures 6A and 6B). The forces required to unfold these four hairpins from BamA$_{\Delta P}$ were ≈30 to 50 pN lower compared with full-length BamA. In contrast, the forces required to unfold β hairpins 5–8 remained unaffected by the deletion of the POTRA domains. Recent structural studies of BamA indicated a possible inward motion of the first three β hairpins of BamA, in particular in the absence of BamB (Gu et al., 2016; Han et al., 2016; Iadanza et al., 2016). The reduced mechanical stability of β hairpins 1–4 indicates a conformational instability of the BamA β barrel in the same region in absence of the POTRA domains. It may thus be speculated that the mechanical stability of the hinge-like region of BamA modulated by the POTRA domains supports the conformational change of the hinge-like motion induced by BamB.

The mean forces required to unfold the β barrel domain of BamA from the native membrane of OMVs were lower than the mean forces required to unfold the same construct from EPL membranes (Figures 6C and 6D). For all β hairpins the force differences ranged from ≈20 to 60 pN, with the largest differences observed for β hairpins 1, 4, and 7. Thus, the β barrel of BamA adopts a mechanically more stable conformation if reconstituted into EPL membranes. One of the main differences between the native OMV and EPL membranes is the presence of lipopolysaccharides (LPS) in the outer leaflet of the native outer membrane (Dowhan and Bogdanov, 2002; Nikaido, 2003). The lower stability of BamA in native OMVs might thus be attributed to the lipid composition and related to the presence of LPS. It may be speculated that LPS plays a major role in contributing to the BamA stability since previous studies indicated the involvement of LPS in Omp biogenesis (Arunmanee et al., 2016; Bulieris et al., 2003; Laird et al., 1994). Previous studies have found that the folding of Omps is influenced by the thickness of the lipid bilayer (Burgess et al., 2008; Schiffrin et al., 2017). Moreover, the folding of Omps by BamA varies depending on the membrane composition (Danoff and Fleming, 2015; Iadanza et al., 2016; Patel and Kleinschmidt, 2013; Plummer and Fleming, 2015). These dependencies of the BamA-mediated folding of Omps on the membrane environment, which have been observed in *in vitro* and *in vivo* studies (Doerner and Sousa, 2017; Noinaj et al., 2014), indicate the importance of the native membrane for BamA function. Our experiments show that BamA is mechanically less stable in the native outer membrane. It is thus possible that the intrinsic instability of the β barrel domain is required for BamA in order to adopt functionally important conformations (Noinaj et al., 2017). Hence, the membrane-dependent modulation of the mechanical stability of BamA could add another component toward understanding the BamA conformations required to facilitate Omp insertion and folding.

Besides reducing the probability to detect the unfolding step in extracellular loop 6, also unlocking the lid-lock of BamA by introducing the mutation R661A, did not change the unfolding steps. Instead, the mutation reduced the forces required to unfold the first four β hairpins 1–4 by ≈30 to 70 pN, and thus mechanically destabilized the BamA β barrel (Figures 6E and 6F). Whereas, the largest mechanical destabilization of 70 pN was observed for β hairpin 1, this destabilization decreased and disappeared toward the C-terminal end. While the interaction of R661 with the inner wall of the β barrel occurs largely within β hairpin 7 (Maier et al., 2015), the forces required to unfold β hairpin 7 were the same in the presence or absence of the mutation. Our experiments thus suggest that the lid-lock of BamA does not change the stability of the β barrel at the site of the interaction, but delocalized. Moreover, the large change observed in the mechanical stability of β hairpin 1 indicates a possible coupling between the extracellular lid of BamA and the seam of the β barrel.

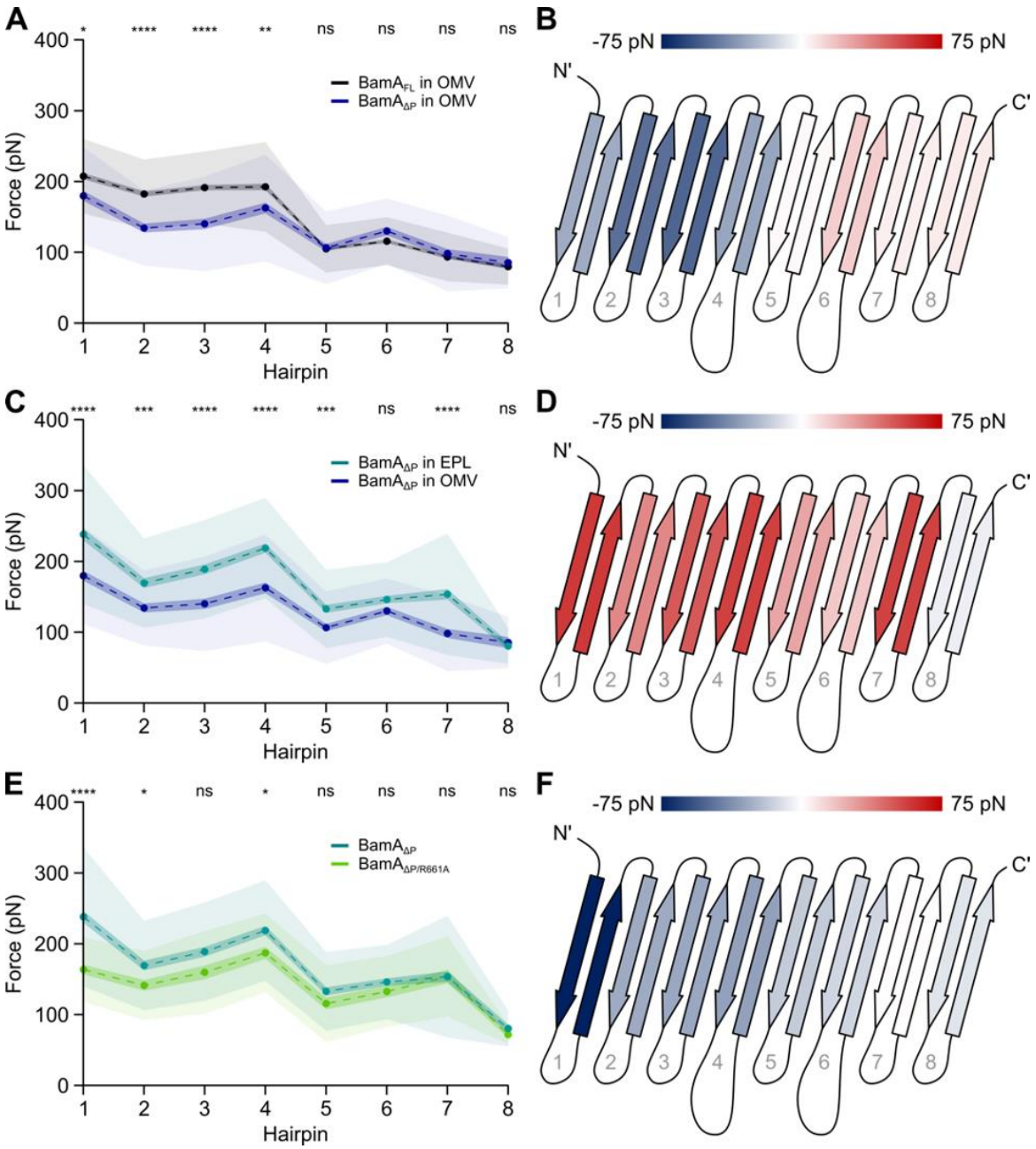


Figure 6. Mechanical Stabilities of the BamA β Barrel Dependent on the POTRA Domains, Membrane Environment, and Lid-Lock (A and B) Analyzing the mechanical stability of the BamA β barrel caused by removal of the POTRA domains. (A) Mean unfolding forces (circles) of β hairpins 1–8 of full-length BamA and BamADP. (B) Secondary structure cartoon of BamA. The color of β hairpins represents the difference in force required to mechanically unfold BamA$_{\Delta P}$ compared with full-length BamA. (C and D) Analyzing the mechanical stability of the BamA β barrel caused by changing the membrane environment. (C) Mean unfolding forces (circles) of β hairpins 1–8 of BamA$_{\Delta P}$ from reconstituted EPL bilayers and native membrane of OMVs. (D) Secondary structure cartoon of BamA. The color of β hairpins represents the difference in force required to mechanically unfold BamA$_{\Delta P}$ from EPL bilayers compared with native membrane of OMVs. (E) Mean unfolding forces (circles) of β hairpins 1–8 of BamA$_{\Delta P}$ and BamA$_{\Delta P/R661A}$. (F) Secondary structure cartoon of BamA. The color of β hairpins represents the difference in force required to mechanically unfold BamA$_{\Delta P/R661A}$ compared with BamA$_{\Delta P}$. Dark shaded areas in (A, C, and E) represent SE. Light shaded areas represent SD. Significance levels are *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; ns, not significant. β hairpins in (B, D, and F) are numbered 1–8 starting from the N-terminal end. See also Figure S5 and Tables S1 and S2.

Functional Implications

Taken together our experiments draw a new perspective of BamA existing as a metastable β barrel in the native outer membrane, the stability of which is influenced by several factors such as the POTRA domains, the membrane composition, and the extracellular lid. It is conceivable that such varying mechanical stabilities go hand in hand with different BamA conformations and local distortions of the membrane required to assist the process of insertion and folding of Omps into the membrane. Our finding complies with recent structural and functional studies revealing unusually high conformational dynamics of the β barrel of BamA (Doerner and Sousa, 2017; Gu et al., 2016; Han et al., 2016; Iadanza et al., 2016). However, to which particular conformations the altered mechanical stabilities of the BamA β barrel are correlated needs to be elucidated in future studies.

4.5 Method Details

Cloning

Plasmid pY219 used to express BamAΔP in the outer membrane, the sequence coding for the β-barrel region of BamA was amplified from plasmid pY161 (Thoma et al., 2018) using primers 5'-CGG CGA AAA CCT GTA CTT CCA GCA TAT GCG CAA CAC CGG TAG CTT C-3' and 5'-GTG CGG CCG CAA GCT TGT CGA CTT ACC AGG TTT TAC CGA TGT TAA AC-3' and cloned into plasmid pYOX (Thoma et al., 2018) using the RF-cloning method (Bond and Naus, 2012; van den Ent and Löwe, 2006). Plasmid pY224 used to purify BamAΔP from inclusion bodies was prepared by removing the periplasmic export sequence from plasmid pY219 using the QuikChange PCR method (Cormack and Castaño, 2002) with primers 5'-ATA CAT ATG GGC AGC AGC CAT CAT C-3' and 5'-GCT GCT GCC CAT ATG TAT ATC TCC TTC TTA AAG-3'. Plasmid pY225 used to express BamAΔP with mutation R661A was created with the QuikChange PCR method with primers 5'-GCA CCG TGG CTG GCT TCC AGT CCA ATA CC-3' and 5'- GGA AGC CAC GCA CGG TGC TGG AAC CAC-3' using the plasmid pY224 as template.

Preparation of BamA in Outer Membrane Vesicles

Full-length BamA was expressed into outer membrane vesicles (OMVs) in *E. coli* BL21(DE3)omp8 (Prilipov et al., 1998) from plasmid pY161 as described (Thoma et al., 2018). Briefly, cells were grown in 250 ml LB medium (Difco, Becton Dickinson) supplemented with Ampicillin (100 mg ml^{-1}, Sigma) at 37°C to OD_{600} z0.5, when expression was induced by addition of 0.5 mM isopropyl-β-D-thiogalactoside (IPTG, Sigma). Upon reaching OD_{600} z 0.95 cells were removed by centrifugation at 10'000 xg for 10 min and the supernatant was sterile-filtered using 0.45 mm filter units (Merk Millipore). All following steps were performed at 4°C. OMVs were collected from the supernatant by ultra-centrifugation at 38'400 xg for 2 h. The outer membrane pellet was washed by resuspension in 20 ml Dulbecco's phosphate buffered saline with added magnesium and calcium (DPBSS, Sigma) and OMVs were collected again by ultracentrifugation at 100'000 xg for 1 h. Outer membrane vesicles were resuspended in 250 ml DPBSS and stored at –80°C. OMVs from *E. coli* BL21(DE3)omp8 enriched in BamAΔP were prepared using plasmid pY219 as described for full-length BamA.

Preparation of Reconstituted BamA

The BamA β-barrel domain, carrying an N-terminal His_6-tag and was expressed in inclusion bodies from plasmid pY224 in *E. coli* Lemo21(DE3) cells (New England Biolabs). Cells were grown in 500 ml LB medium (Difco, Becton Dickinson) supplemented with ampicillin (100 mg ml^{-1}, Sigma) and chloramphenicol (36 mg ml^{-1}, Sigma) at 37°C to

OD_{600} ≈0.8, when expression was induced by addition of 1 mM IPTG (Sigma). After 5 h cells were harvested by centrifugation at 5'000 xg. All following steps were performed at 4°C. The cell pellet was resuspended in 50 ml buffer (50 mM Tris, 300 mM NaCl, pH 8) supplemented with Lysozyme (100 mg ml^{-1}, Sigma), and DNAseI (100 mg ml^{-1}, Roche) and cells were broken by sonication. Inclusion bodies were collected by centrifugation at 16'000 xg for 30 min and dissolved in 25 ml buffer (50 mM Tris, 300 mM NaCl, 6 M guanidine hydrochloride, pH 8) by homogenization. Insoluble material was removed by centrifugation at 16'000 xg for 30 min. The dissolved protein was precipitated by dialysis overnight against 2 l water. The precipitate was dissolved in buffer (50mM Tris, 300 mM NaCl, 6M guanidine hydrochloride, pH 8) and the protein concentration was adjusted to 5 mg ml^{-1}. To refold BamA, 1 ml of protein solution was added dropwise to 50 ml buffer (50 mM Tris, 300 mM NaCl, 0.5% Lauryldimethylamine-N-Oxid (LDAO; Anatrace), 500 mM L-Arginine, pH 8) under stirring and left stirring for 24 h. The solution was dialyzed twice against 2 l buffer (20 mM Tris, pH 8) for 8 h. Precipitate was removed by centrifugation at 10'000 xg for 30 min. The supernatant was loaded on a HiTrap Q HP 5 ml anion exchange column (GE Healthcare) and eluted with a NaCl gradient (20 mM Tris, 0–500 mM NaCl, 0.1% LDAO, pH 8). Refolded BamA eluted at ≈250 mM NaCl. BamA was concentrated to final protein concentration of 1 mg ml^{-1}. For reconstitution BamA was mixed with *E. coli* polar lipid extract (1 mg ml^{-1} in 20 mM Tris, 250 mM NaCl, 1% LDAO, pH 8; Avanti Polar Lipids) at a lipid-to-protein ratio of 0.5 (w/w) in a total volume of 250 ml. The lipid/protein mixture was incubated shaking at 37°C for 1 h, ≈35 mg of BioBeads (SM-2, BioRad) were added and the mixture was incubated shaking at 37°C for 2 h. BioBeads were removed and reconstituted BamA was stored at –80°C. The BamA β-barrel domain containing the additional mutation R661A was expressed, purified and reconstituted as described above for the BamA β-barrel domain.

Single-Molecule Force Spectroscopy

All SMFS measurements were performed at room temperature in buffer solution using a commercial AFM (Nanowizard II Ultra, JPK Instruments) and OMCL-RC800PSA cantilevers (Olympus), which were calibrated using the thermal noise method (Butt and Jaschke, 1995). For measurements of full-length BamA and BamA β-barrel domains 1 ml of BamA containing outer membrane vesicles or reconstituted BamA β-barrel domain containing proteoliposomes, respectively, were adsorbed in 50 ml DPBSS to freshly cleaved mica. After 15 min the sample was washed several times with DPBSS to remove non-adsorbed material (Müller et al., 1997). Adsorbed membrane patches were located by AFM-imaging in contact mode. For SMFS the AFM stylus was pushed onto membrane patches containing densely packed membrane protein assemblies for 500 ms applying a force of 1 nN. The AFM stylus was then retracted at a constant velocity of 1 mm s^{-1}.

SMFS Data Analysis

FD curves were pre-processed as described (Bosshart et al., 2008). Briefly, the FD curves were corrected for deflection sensitivity of the cantilever and coarse filtered for force peak patterns exceeding a length of 95 nm, which corresponds to at least 2/3 of the contour length of the fully unfolded and extended β-barrel domain of BamA (260 of 390 aa with 1 aa ≈0.36 nm). FD curves which remained after filtering were analyzed using an automated approach (Thoma, 2017). For alignment FD curves were transformed to force vs contour length space using the worm-like chain (WLC) model with a fixed persistence length of 0.4 nm (Bosshart et al., 2012b) and binned with a contour length bin-size of 1 nm. For each dataset, the FD curve with the highest similarity to all other FD curves was

identified based on the minimal Euclidean distance against all curves, and used as a template for alignment. Due to differences in length, FD curves recorded of full length BamA were additionally aligned in distance (omitting the transformation to force vs contour length space) to create the representative graph in Figure 2A. However, the successive analysis was based on the alignment in force vs contour length space. After alignment all force peaks in every FD curve were identified by discrete-wavelet-transform-based noise reduction (Daubechies, 1990; Torrence and Compo, 1998) and fitted using the WLC model (Siggia et al., 1994). The resulting force/contour length value pairs were pooled and clusters of high point density were identified using the DBSCAN algorithm (Ester et al., 1996) with the elliptical distance condition $\Delta f^2 \, r_f^{-2} + \Delta l^2 r_l^{-2} \leq 1$ (where Δf is the force distance between two points and Δl is the contour length distance between two points), force radius r_f of 100 pN, contour length radius r_l of 2 nm, and a core point threshold for clustering of 6 points. For every cluster, the mean contour length and mean force were calculated (Table S1). For all FD curves the cluster-allocation of all force peaks was determined. Data points not allocated to any cluster were neglected. The transition-frequency between all cluster pairs was calculated by counting the number of transitions between the respective clusters in all FD curves.

Statistical Analysis

For every force peak class the mean contour length and mean force, including S.D. and S.E.M. were determined (Table S1). To test the statistical significance of differences in the mean force of force peak classes, distributions of the respective force peak classes were compared using Tukey's multiple comparisons test (Figure S5) and Kolmogorov-Smirnov goodness-of-fit test (Table S2).

Acknowledgments

We thank T. Sapra and T. Serdiuk for critically reading and discussing the manuscript. We thank D. Kalbermatter and D. Fotiadis for providing the cryo-TEM image (Figure S3A), which had been taken from a larger previously published dataset. This work was supported by the Swiss National Science Foundation (SNF; grant 205320_160199 to D.J.M.), the NCCR Molecular Systems Engineering and the Swiss Nanoscience Institute (SNI) of the University of Basel. Y.S. was supported by the China Scholarship Council.

4.6 References

Albrecht, R., Schütz, M., Oberhettinger, P., Faulstich, M., Bermejo, I., Rudel, T., Diederichs, K., and Zeth, K. (2014). Structure of BamA, an essential factor in outer membrane protein biogenesis. Acta Crystallogr. D Biol. Crystallogr. 70, 1779–1789.

Arunmanee, W., Pathania, M., Solovyova, A.S., Le Brun, A.P., Ridley, H., Baslé , A., van den Berg, B., and Lakey, J.H. (2016). Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. Proc. Natl. Acad. Sci. USA 113, E5034–E5043.

Bond, S.R., and Naus, C.C. (2012). RF-Cloning.org: an online tool for the design of restriction-free cloning projects. Nucleic Acids Res. 40, 209–213.

Bos, M.P., Robert, V., and Tommassen, J. (2007). Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. EMBO Rep. 8, 1149–1154.

Bosshart, P.D., Casagrande, F., Frederix, P.L.T.M., Ratera, M., Bippes, C.A., Müller, D.J., Palacin, M., Engel, A., and Fotiadis, D. (2008). High-throughput single-molecule force spectroscopy for membrane proteins. Nanotechnology 19, 384014.

Bosshart, P.D., Iordanov, I., Garzon-Coral, C., Demange, P., Engel, A., Milon, A., and Müller, D.J. (2012a). The transmembrane protein KpOmpA anchoring the outer membrane of Klebsiella pneumoniae unfolds and refolds in response to tensile load. Structure 20, 121–127.

Bosshart, P.D., Frederix, P.L.T.M., and Engel, A. (2012b). Reference-free alignment and sorting of single-molecule force spectroscopy data. Biophys. J. 102, 2202–2211.

Buchanan, S.K. (1999). Beta-Barrel proteins from bacterial outer membranes: structure, function and refolding. Curr. Opin. Struct. Biol. 9, 455–461.

Bulieris, P.V., Behrens, S., Holst, O., and Kleinschmidt, J.H. (2003). Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. J. Biol. Chem. 278, 9092–9099.

Burgess, N.K., Dao, T.P., Stanley, A.M., and Fleming, K.G. (2008). β-Barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. J. Biol. Chem. 283, 26748–26758.

Butt, H.-J., and Jaschke, M. (1995). Calculation of thermal noise in atomic force microscopy. Nanotechnology 6, 1–7.

Cormack, B., and Castaño, I. (2002). Introduction of point mutations into cloned genes. Methods Enzymol. 350, 199–218.

Damaghi, M., Bippes, C., Köster, S., Yildiz,Ö., Mari, S.A., Kühlbrandt, W., and Muller, D.J. (2010a). pH-Dependent interactions guide the folding and gate the transmembrane pore of the β-barrel membrane protein OmpG.

J. Mol. Biol. 397, 878–882. Damaghi, M., Sapra, K.T., Köster, S., Yildiz, Ö., Kühlbrandt, W., and Muller, D.J. (2010b). Dual energy landscape: the functional state of the β-barrel outer membrane protein G molds its unfolding energy landscape. Proteomics 10, 4151–4162.

Danoff, E.J., and Fleming, K.G. (2015). Membrane defects accelerate outer membrane β-barrel protein folding. Biochemistry 54, 97–99.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Inf. Theory 36, 961–1005.

Doerner, P.A., and Sousa, M.C. (2017). Extreme dynamics in the BamA β-barrel seam. Biochemistry 56, 3142–3149.

Dowhan, W., and Bogdanov, M. (2002). Functional roles of lipids in membranes. In Biochemistry of Lipids, Lipoproteins and Membranes, Fourth Edition, J.E. Vance and D. Vance, eds. (Elsevier), pp. 1–35.

Engel, A., and Gaub, H.E. (2008). Structure and mechanics of membrane proteins. Annu. Rev. Biochem. 77, 127–148.

van den Ent, F., and Löwe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J. Biochem. Biophys. Methods 67, 67–74.

Ester, M., Kriegel, H.P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Proc. 2nd Int. Conf. Knowl. Discov. Data Min. 96, 226–231.

Evans, E., and Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. Biophys.J. 72, 1541–1555.

Fleming, K.G. (2015). A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, https://doi.org/10.1098/rstb. 2015.0026.

Ge, L., Villinger, S., Mari, S.A., Giller, K., Griesinger, C., Becker, S.,Müller, D.J., and Zweckstetter, M. (2016). Molecular plasticity of the human voltage-dependent anion channel embedded into a membrane. Structure 24, 585–594.

Gessmann, D., Chung, Y.H., Danoff, E.J., Plummer, A.M., Sandlin, C.W., Zaccai, N.R., and Fleming, K.G. (2014). Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. Proc. Natl. Acad. Sci. USA 111, 5878–5883.

Gruss, F., Zähringer, F., Jakob, R.P., Burmann, B.M., Hiller, S., and Maier, T. (2013). The structural basis of autotransporter translocation by TamA. Nat. Struct. Mol. Biol. 20, 1318–1320.

Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N.G., Stansfeld, P.J., Wang, Z., Zhang, Y., Wang, W., et al. (2016). Structural basis of outer membrane protein insertion by the BAM complex. Nature 531, 64–69.

Hagan, C.L., Silhavy, T.J., and Kahne, D. (2011). β-Barrel membrane protein assembly by the Bam complex. Annu. Rev. Biochem 80, 189–210.

Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., et al. (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. Nat. Struct. Mol. Biol. 23, 192–196.

Hensen, U., and Müller, D.J. (2013). Mechanistic explanation of different unfolding behaviors observed for transmembrane and soluble β-barrel proteins. Structure 21, 1317–1324.

Heymann, B., and Grubmüller, H. (2000). Dynamic force spectroscopy of molecular adhesion bonds. Phys. Rev. Lett. 84, 6126–6129.

Höhr, A.I.C., Straub, S.P., Warscheid, B., Becker, T., and Wiedemann, N. (2015). Assembly of β-barrel proteins in the mitochondrial outer membrane. Biochim. Biophys. Acta 1853, 74–88.

Höhr, A.I.C., Lindau, C., Wirth, C., Qiu, J., Stroud, D.A., Kutik, S., Guiard, B., Hunte, C., Becker, T., Pfanner, N., et al. (2018). Membrane protein insertion through a mitochondrial β-barrel gate. Science 359, eaah6834.

Iadanza, M.G., Higgins, A.J., Schiffrin, B., Calabrese, A.N., Brockwell, D.J., Ashcroft, A.E., Radford, S.E., and Ranson, N.A. (2016). Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. Nat. Commun 7, 12865.

Jacob-Dubuisson, F., Villeret, V., Clantin, B., Delattre, A.-S., and Saint, N. (2009). First structural insights into the TpsB/Omp85 superfamily. Biol. Chem. 390, 675–684.

Kedrov, A., Janovjak, H., Sapra, K.T., and Müller, D.J. (2007). Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu. Rev. Biophys. Biomol. Struct. 36, 233–260.

Koebnik, R., Locher, K., and Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol. Microbiol. 37, 239–253.

Konovalova, A., Kahne, D.E., and Silhavy, T.J. (2017). Outer membrane biogenesis. Annu. Rev. Microbiol. 71, 539–556.

Laird, M.W., Kloser, A.W., and Misra, R. (1994). Assembly of LamB and OmpF in deep rough lipopolysaccharide mutants of Escherichia coli K-12. J. Bacteriol. 176, 2259–2264.

Leonard-Rivera, M., and Misra, R. (2012). Conserved residues of the putative L6 loop of Escherichia coli BamA play a critical role in the assembly of β-barrel outer membrane proteins, including that of BamA itself. J. Bacteriol. 194, 4662–4668.

Maier, T., Clantin, B., Gruss, F., Dewitte, F., Delattre, A.-S., Jacob-Dubuisson, F., Hiller, S., and Villeret, V. (2015). Conserved Omp85 lid-lock structure and substrate recognition in FhaC. Nat. Commun. 6, 7452.

Marko, J.F., and Siggia, E.D. (1995). Stretching DNA. Macromolecules 28, 8759–8770.

Müller, D.J., and Engel, A. (2007). Atomic force microscopy and spectroscopy of native membrane proteins. Nat. Protoc. 2, 2191–2197.

Müller, D., Amrein, M., and Engel, A. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. J. Struct. Biol. 188, 172–188.

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. Microbiol. Mol. Biol. Rev. 67, 593–656.

Noinaj, N., Kuszak, A.J., Gumbart, J.C., Lukacik, P., Chang, H., Easley, N.C., Lithgow, T., and Buchanan, S.K. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. Nature 501, 385–390.

Noinaj, N., Kuszak, A.J., Balusek, C., Gumbart, J.C., and Buchanan, S.K. (2014). Lateral opening and exit pore formation are required for BamA function. Structure 22, 1055–1062.

Noinaj, N., Gumbart, J.C., and Buchanan, S.K. (2017). The β-barrel assembly machinery in motion. Nat. Rev. Microbiol. 15, 197–204.

Oesterhelt, F., Oesterhelt, D., Pfeiffer, M., Engel, A., Gaub, H.E., and Müller, D.J. (2000). Unfolding pathways of individual bacteriorhodopsins. Science 288, 143–146.

Patel, G.J., and Kleinschmidt, J.H. (2013). The lipid bilayer-inserted membrane protein BamA of Escherichia coli facilitates insertion and folding of outer membrane protein A from its complex with Skp. Biochemistry 52, 3974–3986.

Plummer, A.M., and Fleming, K.G. (2015). BamA alone accelerates outer membrane protein folding in vitro through a catalytic mechanism. Biochemistry 54, 6009–6011.

Plummer, A.M., and Fleming, K.G. (2016). From chaperones to the membrane with a BAM! Trends Biochem. Sci. 41, 872–882.

Prilipov, A., Phale, P.S., Gelder, P., Rosenbusch, J.P., and Koebnik, R. (1998). Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from E. coli. FEMS Microbiol. Lett. 163, 65–72.

Rollauer, S.E., Sooreshjani, M.A., Noinaj, N., and Buchanan, S.K. (2015). Outer membrane protein biogenesis in Gram-negative bacteria. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 1–10.

Ruiz, N., Kahne, D., and Silhavy, T.J. (2006). Advances in understanding bacterial outer-membrane biogenesis. Nat. Rev. Microbiol. 4, 57–66.

Sapra, K.T., Damaghi, M., Köster, S., Yildiz, Ö., Kühlbrandt, W., and Müller, D.J. (2009). One β hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG. Angew. Chem. Int. Ed. 48, 8306–8308.

Schiffrin, B., Calabrese, A.N., Higgins, A.J., Humes, J.R., Ashcroft, A.E., Kalli, A.C., Brockwell, D.J., and Radford, S.E. (2017). Effects of periplasmic chaperones and membrane thickness on BamA-catalyzed outer-membrane protein folding. J. Mol. Biol. 429, 3776–3792.

Selkrig, J., Mosbahi, K., Webb, C.T., Belousoff, M.J., Perry, A.J., Wells, T.J., Morris, F., Leyton, D.L., Totsika, M., Phan, M.-D., et al. (2012). Discovery of an archetypal protein transport system in bacterial outer membranes. Nat. Struct. Mol. Biol. 19, 506–510.

Siggia, E., Bustamante, C., Marko, J., and Smith, S. (1994). Entropic elasticity of l-phage DNA. Science 265, 5–6.

Thoma, J. (2017). Unfolding and folding pathways of Escherichia coli outer membrane proteins. PhD thesis (ETH Zürich).

Thoma, J., Bosshart, P., Pfreundschuh, M., and Müller, D.J. (2012). Out but not in: the large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. Structure 20, 2185–2190.

Thoma, J., Burmann, B.M., Hiller, S., and Müller, D.J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat. Struct. Mol. Biol. 22, 795–802.

Thoma, J., Ritzmann, N., Wolf, D., Mulvihill, E., Hiller, S., and Müller, D.J. (2017). Maltoporin LamB unfolds β hairpins along mechanical stress-dependent unfolding pathways. Structure 25, 1139–1144.e2.

Thoma, J., Manioglu, S., Kalbermatter, D., Bosshart, P.D., Fotiadis, D., and Müller, D.J. (2018). Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. Commun. Biol. 1, 23.

Torrence, C., and Compo, G.P. (1998). A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79, 61–78.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J., and Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell 121, 235–245.

Yu, H., Siewny, M.G.W., Edwards, D.T., Sanders, A.W., and Perkins, T.T. (2017). Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. Science 355, 945–950.

4.7 Supplemental Information

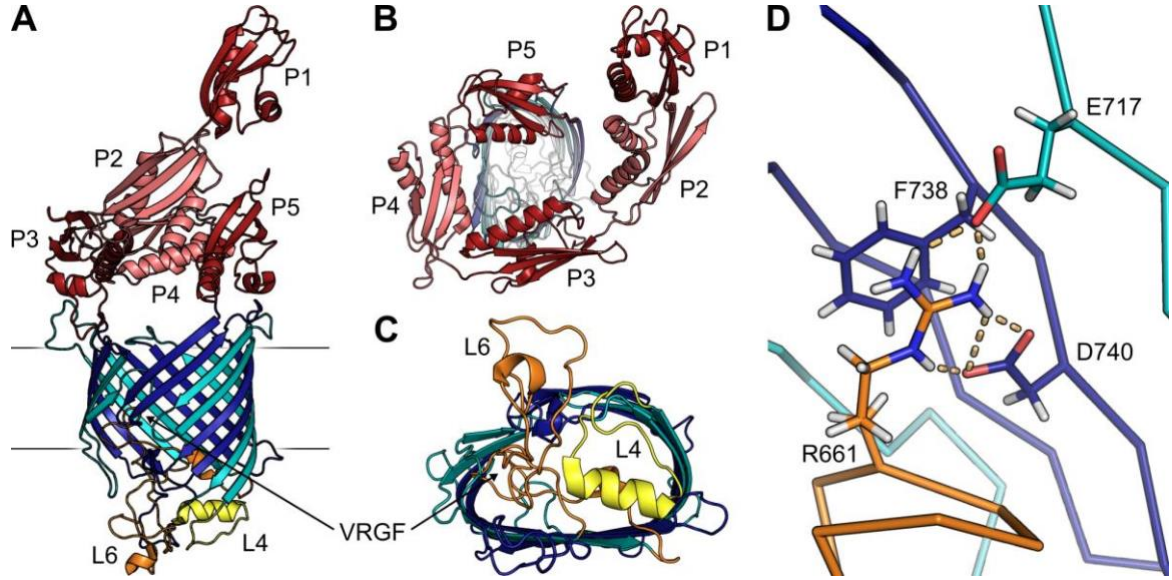


Figure S1. Related to Figure 1. Structure of the β-barrel assembly factor BamA. (A) Side view of BamA. The POTRA domains (P1–P5) are colored red, the β-barrel domain blue, the extracellular loop 4 (L4) yellow and the extracellular loop 6 (L6) orange. The membrane is indicated by black horizontal lines. The position of the conserved VRGC motif in L6 is indicated. (B) Top view of the periplasmic side of BamA. (C) Top view of the extracellular side of BamA. (D) Lid-lock structure of BamA. Involved amino acid residues are shown in stick representation. Structures were taken from PDB 5D0O.

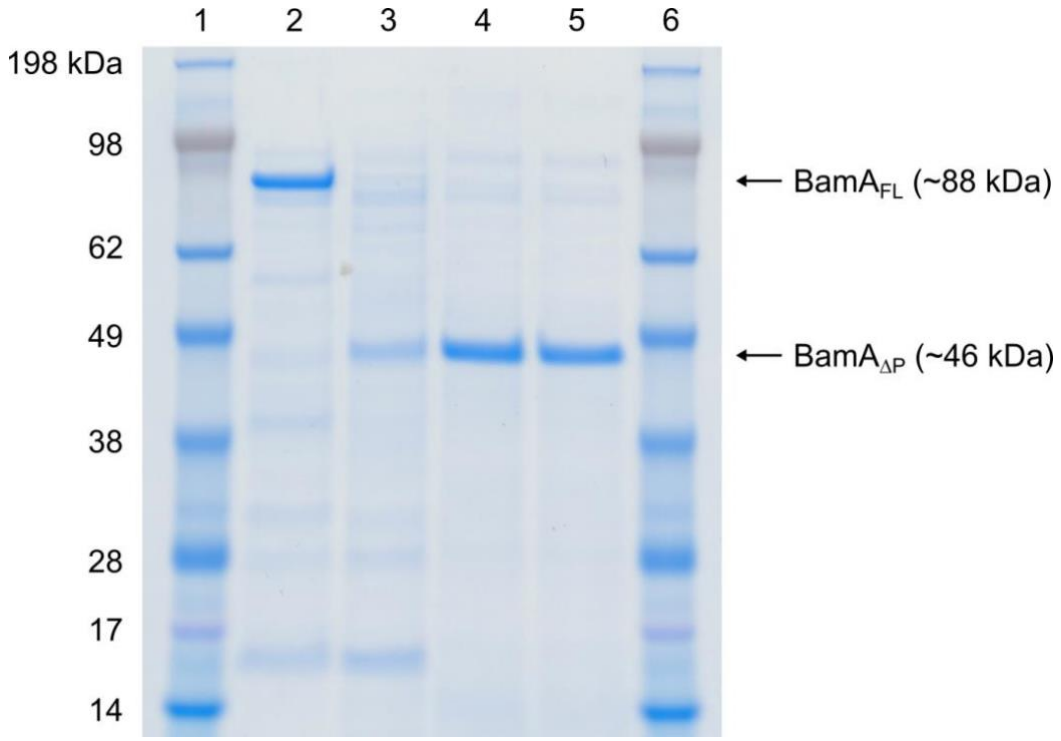


Figure S2. Related to Figures 2–5. SDS Page of BamA Constructs. Lanes 1 and 6) molecular weight marker, lane 2) OMVs enriched in full-length BamA (BamA$_{FL}$), lane 3) OMVs enriched in BamA$_{\Delta P}$, lane 4) purified BamA$_{\Delta P}$ reconstituted in E. coli polar lipid membranes, lane 5) purified BamA$_{\Delta P/R661A}$ reconstituted in *E. coli* polar lipid (EPL) membranes.

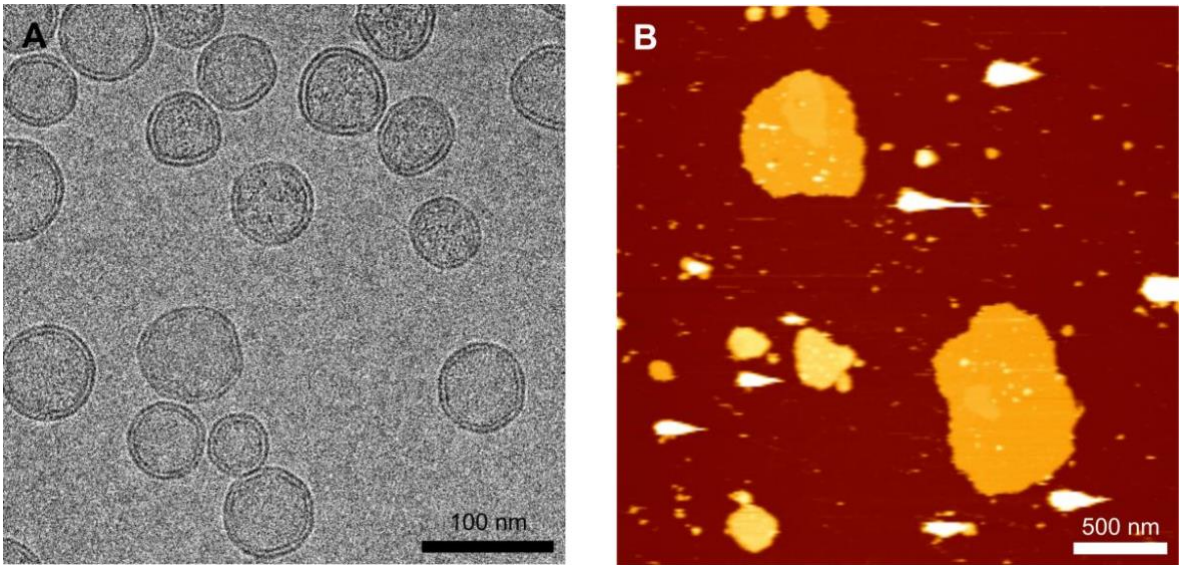


Figure S3. Related to Figure 2. Morphology of OMVs Enriched in BamA. (A) Cryo-transmission electron microscopy image of OMVs enriched in full-length BamA. Membrane bilayers of the OMVs are clearly visible. The image is representative for a larger data set collected at liquid nitrogen temperature at a magnification of 50,000 x (step size of 2.08 Å per pixel at the specimen level) in reference (Thoma et al., 2018). (B) AFM topograph of OMVs enriched in full-length BamA adsorbed to mica. Upon adsorption OMVs break open and expose their periplasmic surface. The AFM topograph was recorded in contact mode in buffer solution at room temperature as described (Thoma et al., 2018). The full color range of the topograph corresponds to a vertical range of 12 nm.

Thoma, J., Manioglu, S., Kalbermatter, D., Bosshart, P.D., Fotiadis, D., and Müller, D.J. (2018). Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. Commun. Biol. 1, 23.

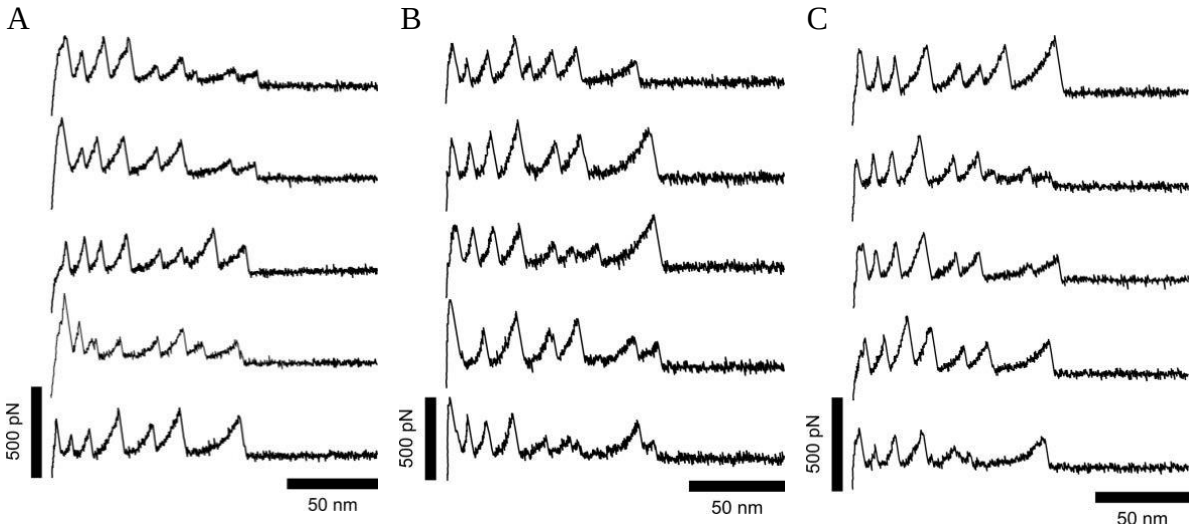


Figure S4. Related to Figures 3-5. Mechanical Unfolding of Single BamA$_{\Delta P}$ from OMVs, BamA$_{\Delta P}$ from EPL membranes, and BamA$_{\Delta P/R661A}$ from EPL membranes. (A) Exemplary FD curves recorded upon mechanically unfolding single BamA$_{\Delta P}$ from OMVs by SMFS in buffer solution at room temperature as described (Methods). (B) Exemplary FD curves were recorded upon mechanically unfolding single BamA$_{\Delta P}$ from EPL membranes by SMFS in buffer solution at room temperature as described (Methods). (C) Exemplary FD curves were recorded upon mechanically unfolding single BamA$_{\Delta P/R661A}$ from EPL membranes by SMFS in buffer solution at room temperature as described (Methods).

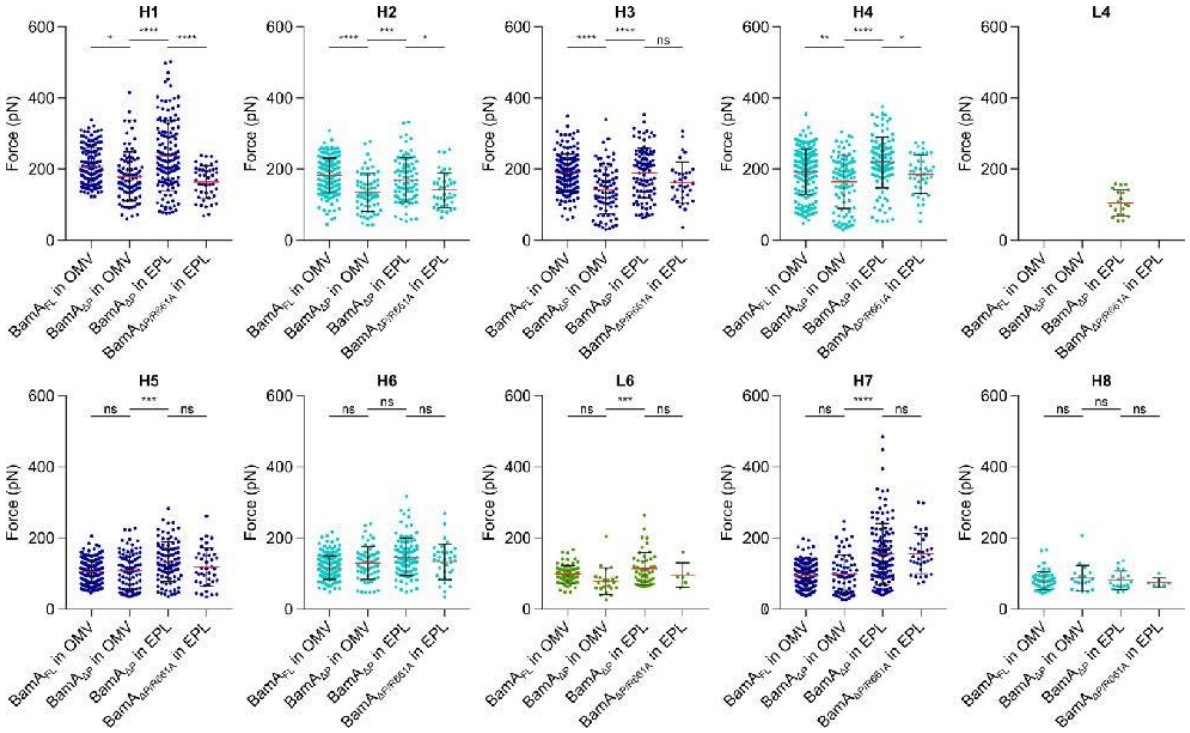


Figure S5. Related to Figure 6. Statistical analysis of mechanical forces required to unfold individual β-hairpins (H1-H8) from different BamA constructs by SMFS. Colored dots are rupture forces of a single force peak of every single FD curve clustered in a force peak class. Red lines are mean forces, error bars give S.D.. Significance levels of pair-wise comparison with Tukey's multiple comparisons test are given above, significance levels are * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$. See also Table S2.

	Segment	Number of Events	Probability	Contour Length (aa)			Force (pN)		
				Mean	S.D.	S.E.M.	Mean	S.D.	S.E.M.
$BamA_{FL}$ in OMV	H1	149	0.99	405.84	2.40	0.19	207.55	52.69	4.26
	H2	151	1	435.55	2.38	0.19	182.35	48.80	3.81
	H3	149	0.99	466.70	3.01	0.22	191.47	50.97	3.77
	H4	151	1	509.41	3.67	0.26	192.41	63.40	4.44
	H5	151	1	574.92	4.52	0.33	104.97	33.48	2.46
	H6	151	1	615.33	3.76	0.29	115.79	33.98	2.58
	L6	73	0.48	645.53	7.09	0.83	98.72	25.67	3.00
	H7	151	1	717.01	6.29	0.47	93.50	34.48	2.58
	H8	55	0.36	770.97	5.60	0.76	79.82	25.31	3.41
$BamA_{\Delta P}$ in OMV	H1	64	1	13.52	6.51	0.68	179.77	68.52	7.11
	H2	63	0.99	45.39	3.84	0.47	134.26	52.91	6.46
	H3	64	1	75.68	4.82	0.54	140.15	66.61	7.40
	H4	64	1	119.58	5.58	0.56	162.66	75.33	7.61
	H5	64	1	183.78	7.35	0.77	106.48	51.12	5.36
	H6	63	0.98	227.49	4.58	0.58	130.03	46.00	5.80
	L6	23	0.37	253.03	3.98	0.83	81.69	39.61	8.26
	H7	64	1	331.29	7.98	0.96	98.33	53.38	6.43
	H8	20	0.31	385.82	3.96	0.89	85.51	36.07	8.07
$BamA_{\Delta P}$ in EPL	H1	93	1	16.02	6.04	0.52	238.12	98.13	8.41
	H2	83	0.9	45.88	3.21	0.35	169.46	63.14	6.93
	H3	93	1	76.32	3.67	0.37	188.96	69.43	6.94
	H4	93	1	118.55	3.63	0.34	218.80	70.89	6.64
	L4	21	0.22	158.69	5.18	1.13	107.86	34.89	7.61
	H5	93	1	184.06	4.35	0.45	132.88	55.46	5.75
	H6	93	1	223.21	3.74	0.37	146.23	52.45	5.24
	L6	51	0.54	255.74	9.44	1.32	113.86	44.66	6.25
	H7	93	1	327.03	10.61	0.93	153.66	85.79	7.52
	H8	20	0.21	377.97	7.39	1.65	80.45	25.25	5.65
$BamA_{\Delta P/R661A}$ in EPL	H1	35	1	15.24	4.87	0.73	163.77	45.49	6.78
	H2	35	0.99	44.92	3.54	0.55	141.42	48.76	7.62
	H3	35	1	75.12	2.39	0.40	159.77	58.31	9.86
	H4	35	1	117.85	3.33	0.51	187.29	55.49	8.46
	H5	34	0.98	181.97	6.59	1.03	115.81	54.19	8.46
	H6	35	1	224.44	3.21	0.53	132.84	49.93	8.32
	L6	7	0.2	259.05	3.81	1.44	88.36	36.24	13.70
	H7	35	1	332.60	3.74	0.63	154.02	56.58	9.56
	H8	8	0.22	382.93	4.30	1.52	71.83	13.64	4.82

Table S1. Related to Figures 2–6. Probability, contour length, and unfolding force determined of individual β-hairpins (H1-H8) unfolded from different BamA constructs by SMFS. Contour lengths are given in numbers of amino acids (aa) with 0.36 nm corresponding to the length of one aa.

	$BamA_{FL}$ in OMV / $BamA_{\Delta P}$ in OMV	$BamA_{\Delta P}$ in OMV / $BamA_{\Delta P}$ in EPL	$BamA_{\Delta P}$ in EPL / $BamA_{\Delta P/R661A}$ in EPL
H1	$5.48\ 10^{-4}$	$1.93\ 10^{-6}$	$5.33\ 10^{-7}$
H2	$1.66\ 10^{-8}$	$6.86\ 10^{-3}$	$4.54\ 10^{-2}$
H3	$2.81\ 10^{-8}$	$8.40\ 10^{-5}$	$2.10\ 10^{-3}$
H4	$3.03\ 10^{-3}$	$7.80\ 10^{-7}$	$4.97\ 10^{-3}$
H5	$5.36\ 10^{-3}$	$3.32\ 10^{-2}$	$1.62\ 10^{-1}$
H6	$1.07\ 10^{-2}$	$8.55\ 10^{-2}$	$5.89\ 10^{-1}$
L6	$4.85\ 10^{-3}$	$5.30\ 10^{-3}$	$1.94\ 10^{-1}$
H7	$1.29\ 10^{-1}$	$1.99\ 10^{-4}$	$1.70\ 10^{-1}$
H8	$6.48\ 10^{-1}$	$7.94\ 10^{-1}$	$5.19\ 10^{-1}$

Table S2. Related to Figure S5. Statistical analysis of mechanical forces required to unfold individual β-hairpins (H1-H8) from different BamA constructs by SMFS. *P* values derived from pair-wise comparison of force distributions (Figure S5) with Kolmogorov-Smirnov goodness-of-fit test.

5 Structural Properties of the β-Barrel Assembly Factor BamA Modulated by the Antibiotic Darobactin

Noah Ritzmann[1], Selen Manioglu[1], Sebastian Hiller[2], Daniel J. Müller[1,*]

[1] Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH) Zürich, Mattenstrasse 26, 4058 Basel, Switzerland
[2] Biozentrum, University of Basel, Spitalstrasse 41, 4056 Basel, Switzerland

Author Information
Corresponding Author:
*Email: (D.J.M) daniel.mueller@bsse.ethz.ch

Author Contributions. N.R. performed cloning, prepared outer membrane vesicles, conducted SMFS experiments and analyzed all data. S.M. recorded high-resolution AFM images. N.R., S.H. and D.J.M. designed the study, discussed and wrote the manuscript.

Submitted Manuscript

5.1 Abstract

The β-barrel assembly machinery (BAM) complex is an essential component of *Escherichia coli* that inserts and folds outer membrane proteins. The natural antibiotic compound darobactin inhibits BamA, the central unit of BAM. Here, we employ single-molecule force spectroscopy (SMFS) to better understand the structure-function relationship of BamA and its inhibition by darobactin. The five POTRA domains show low mechanical, kinetic and energetic stabilities. In contrast, the structural region linking the POTRA domains to the transmembrane β-barrel exposes the highest mechanical stiffness and lowest kinetic stability within BamA, thus indicating a mechano-functional role. Within the β-barrel, the four N-terminal β-hairpins H1–H4 expose the highest mechanical stabilities and stiffnesses, while the four C-terminal β-hairpins H5–H6 show lower stabilities and higher flexibilities. This asymmetry within the β-barrel suggests that substrates funneling into the lateral gate formed by β-hairpins H1 and H8 can force the flexible C-terminal β-hairpins to change conformations. Darobactin-binding considerably changes the mechanical, kinetic and energetic stability of the BamA β-barrel, increasing the mechanical stability by around 20%. The local stability parameters of BamA provide insight into the structure-function relationship of BamA and its inhibition by darobactin.

5.2 Introduction

Integral outer membrane proteins (OMPs) of Gram-negative bacteria participate in a variety of crucial cellular functions, including nutrient transport[1], protein secretion[2] and adhesion[3]. Most integral OMPs adopt a β-barrel conformation, which spans the outer membrane. In *Escherichia coli* the insertion and folding of β-barrel proteins into the outer membrane is facilitated by the β-barrel assembly machinery (BAM) complex. The BAM complex consists of five protein components. These are BamA, an integral OMP and a member of the Omp85 superfamily[4,5], and four lipoproteins, BamB, BamC, BamD and BamE, which are anchored to the inner leaflet of the outer membrane[6]. BamA, the central component of the BAM complex consists of a C-terminal 16-stranded transmembrane β-barrel domain and five periplasmic N-terminal polypeptide transport (POTRA) domains[7,8]. Although all components of the BAM complex are necessary to efficiently insert proteins into the outer membrane[9], only BamA and BamD are essential[4,10].

The exact mechanism by which β-barrel OMPs are rapidly folded and integrated into the outer membrane by the BAM complex in the absence of an energy-source such as ATP presumably occurs via the "budding model" but remains debated[11]. The budding model proposes the formation of a hybrid β-barrel, formed by BamA and the OMP substrate, followed by the lateral release of the OMP into the membrane[12,13]. Recent structures of multiple folding intermediates support the budding model and extend the proposed process of BAM-mediated OMP folding and membrane insertion[14–16]. The first step of the insertion process is initiated by hydrogen bonding of a C-terminal signal sequence of the incoming substrate to the N-terminal β-strand of BamA[14,17,18]. β-sheets of the substrate are then thought to fold starting from the C-terminus into the interior of the β-barrel. Upon reaching sufficient folding the substrate is released into the outer membrane through a lateral gate formed by the first N-terminal and the last C-terminal β-strand of the transmembrane BamA β-barrel[19]. Compared to most other β-barrel-forming OMPs, the first and the last β-strand of the BamA β-barrel establish loose interactions. This results in a structurally dynamic region, referred to as gate region, that takes a central role in inserting, folding and releasing OMPs into the outer membrane[13]. However, which mechanical and kinetic properties of the BamA β-barrel retain the unique dynamic state of the gate region remain to be characterized.

Because the BAM complex is essential for Gram-negative bacteria, it renders a promising antibiotic target. Compared to other essential components of Gram-negatives, the BAM complex is located at the cellular periphery and thus directly exposed to the external environment. Recently, the natural compound darobactin was identified to inhibit BamA. Darobactin is a bicyclic heptapeptide produced by *Photorhabdus*, found in the microbiome of entomopathogenic nematodes[20]. The recent structure of the BAM complex bound to darobactin revealed darobactin binding to the functionally relevant gate region of the BamA β-barrel[21]. In order to inhibit BamA, darobactin mimics the C-terminal β-signal of native BamA substrates and binds to the first and last β-strand of the BamA β-barrel, which seals the lateral gate. Although the structure of the BamA-darobactin complex describes in great detail how darobactin inhibits BamA, it does not describe how darobactin affects the mechanical, kinetic, and energetic properties of BamA. Thus, quantifying how darobactin modulates these properties at a resolution of individual secondary structural elements would complement the structure of the BAM complex and allow to draw a more complete picture of the darobactin mediated BamA-inhibition. Additionally, such insight is relevant for the targeted engineering of novel antibiotics, which could inhibit BamA similarly to darobactin.

Here, we characterize quantitatively and structurally how darobactin modulates the mechanical, kinetic, and energetic properties of BamA in the native outer membrane. To do so, we apply atomic force microscopy (AFM), which is a versatile tool for the study of single membrane proteins under physiologically relevant conditions[22,23]. AFM has been applied to membrane proteins in a wide range of applications, including high resolution imaging[24], the imaging and mapping of their mechanical and chemical properties[25], or the quantification of their interactions with ligands[26,27]. Importantly for this work, AFM-based single-molecule force spectroscopy (SMFS) allows to quantify the mechanical properties of individual membrane proteins and to map them to their structural regions such as α-helices, β-hairpins or polypeptide loops[28–30]. We thus apply SMFS to mechanically unfold BamA from the native outer membrane in the presence and absence of darobactin. The resolution of our experiments allows to measure the mechanical stability of individual POTRA domains and β-hairpins forming the transmembrane β-barrel of BamA. The experiments show that darobactin considerably increases the mechanical stability of BamA, especially of the transmembrane β-barrel. We further quantify how darobactin modulates the free-energy wells and barriers stabilizing distinct structural regions of BamA as well as their kinetic stabilities. The results, which provide important and unique insight into the mechanisms by which darobactin inhibits BamA, guide towards the mechanistic understanding of how to functionally inhibit BamA.

5.3 Results and Discussion

BamA Stabilizes Distinct Structural Segments

To characterize BamA in the native membrane, we prepared outer membrane vesicles (OMVs) enriched with BamA (Supporting Figure S1A). For AFM imaging and SMFS, the OMVs were adsorbed to mica where they opened up and formed planar membrane patches with diameters ranging from 100 nm to 1 μm (Supporting Figure S1B). Membrane patches of densely packed BamA showed corrugated topographs and elevated 10 – 15 nm in height from the mica (Supporting Figure S1C). For SMFS we pushed the tip of the AFM cantilever onto BamA enriched regions with a force of 1 nN for 0.5 s to promote the non-specific attachment of the tip to the N-terminal region[31] of individual BamA proteins (Figure 1A). Upon the subsequent retraction of the cantilever from the membrane we recorded force-distance (FD) curves. Occasionally, the FD curves recorded a sawtooth-like pattern of force peaks (Figure 1B), which was recently correlated to describe the stepwise unfolding of BamA[31]. Such FD curves showed the mechanical extension and stretching of the N-terminal region, which was followed by the stepwise unfolding of structural segments until a single BamA was fully unfolded and extracted from the outer membrane. We repeated this procedure thousands of times to identify the reoccurring mechanical unfolding pattern specific to BamA in the absence of the antibiotic darobactin.

To compare the FD curves recorded upon mechanically unfolding single BamA we aligned them in the force contour length space, where the prominent force peak at a contour length of 406 aa was used as a reference (Figure 1C, Supporting Figure S2A). The alignment revealed 15 force peak classes (Figure 1D, Supporting Figure S2B). Each of the force peak classes detected the unfolding of a structural segment of BamA in response to mechanical stress[31]. Counting from the N-terminal region, which was picked up by the AFM tip, we used the contour length of each force peak class to assign the structural segment BamA unfolded in response to mechanical stress (Figure 1E). The first five force peak classes describe the unfolding of the five N-terminal POTRA domains of BamA (P1–P5), followed by force peak class describing the short linker region (LR) between the

POTRA domain P5 and the transmembrane β-barrel. The next force peak classes are located in the β-barrel region of BamA and describe the unfolding of one β-hairpin after the other until the eight β-hairpins (H1–H8) are completely unfolded. One additional force peak class located between β-hairpins H6 and H7 corresponds to the unfolding of extracellular loop 6, which is locked in a lid-like structure (LL) to seal the β-barrel lumen from the extracellular space[32].

In summary, the FD curves recorded upon mechanically unfolding single BamA from the native outer membrane reveal a sequence of force peaks, each of which assigning the unfolding of a structural segment of the membrane protein. The forces at which each segment unfold are a direct measure of their mechanical stability.

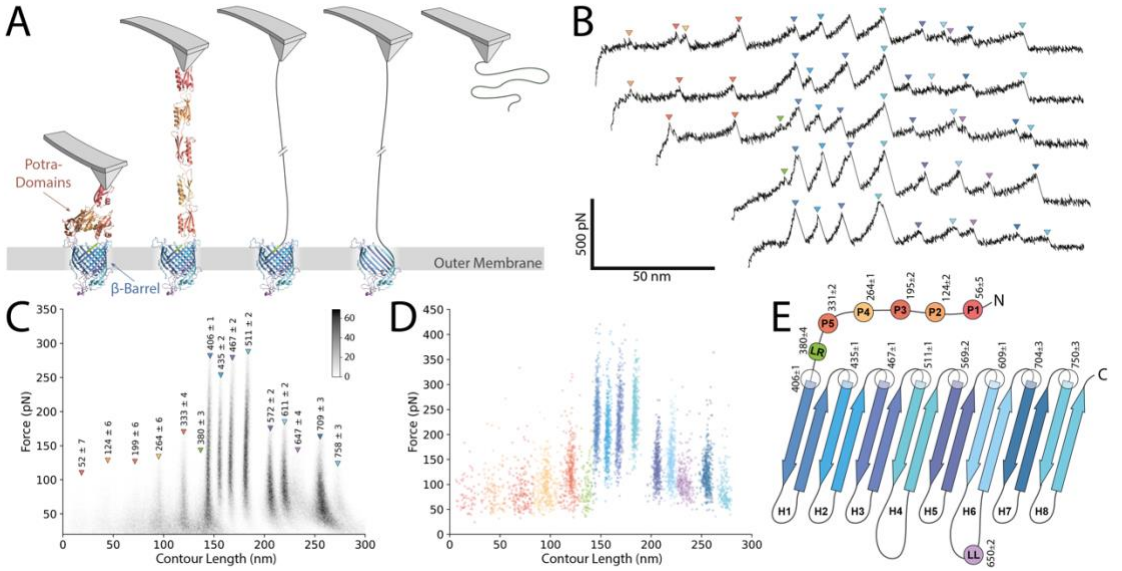


Figure 1. Single-Molecule Force Spectroscopy (SMFS) of BamA in Native Outer Membrane. (A) Schematic representation of the mechanical unfolding of a single BamA from the native outer membrane by AFM-based SMFS. From left to right: the tip of the AFM-cantilever is pushed onto the sample to establish unspecific interactions between the tip of the AFM-cantilever and the N-terminal region of BamA. Retraction of the cantilever mechanically stretches the N-terminal region until it is fully extended. Further retraction of the cantilever results in the stepwise mechanical unfolding of individual structural segments of BamA, until it is completely extracted from the membrane. (B) Exemplary force-distance (FD) curves recorded during the mechanical unfolding of single BamA by SMFS. Colored triangles indicate force peaks that denote unfolding events of BamA. This stepwise unfolding of structural segments continues until the entire BamA has been unfolded and extracted from the outer membrane. (C) Density plot of 418 aligned and superimposed FD curves plotted in the force *versus* contour length space. The colored triangles indicate the force peak classes of BamA. Numbers above each triangle represent mean contour lengths and standard deviations of the respective force peak class in amino acids. The grey scale bar indicates the number of data points per bin. (D) Force peak positions of all 418 FD curves shown in (C), plotted in the force *versus* contour length space. Each point indicates the force and contour length of a single rupture event, as indicated by the colored triangles in (B). Points with identical coloration belong to the same force peak class. The mean contour lengths and standard deviations are indicated in (C) for all force peak classes. (E) Assignment of structural segment unfolded by BamA. Contour lengths of force peak classes (C, D) were used to assign the structural segments of BamA. Mean contour lengths and standard deviations in amino acids are indicated next to each structural segment. The secondary structure of BamA is based on (PDB: 5D0O)[55].

Darobactin Increases the Mechanical Stability of BamA

In the next step we unfolded BamA in the presence of darobactin in order to determine how darobactin-binding changes the mechanical stability of BamA. At first sight, the force peak classes and thus the structural segments stabilizing BamA in the absence and

presence of darobactin were very similar (Supporting Figure S2, S3). Thus, darobactin did not change the structural segments BamA stabilizes against mechanical unfolding. Also, the mechanical forces required to unfold the individual POTRA domains were mostly unaffected by darobactin, except for POTRA domain P5, which increased from 136 pN to 149 pN, i.e. by ≈ 10% (Figure 2A,B). Similarly, the mechanical stability of the linker-region increased only weakly in the presence of darobactin. In large contrast, however, the forces required to mechanically unfold structural segments of the β-barrel region of BamA increased considerably in the presence of darobactin. In principle, every structural segment of the β-barrel, most of which representing a single b-hairpin, required higher forces to unfold than in the absence of darobactin. Upon adding up these forces we found that 1.45 ± 0.15 nN (mean ± sd) was required to unfold all structural segments of the BamA β-barrel in the absence and that 1.67 ± 0.08 nN was required to completely unfold the β-barrel in the presence of darobactin (Figure 2B).

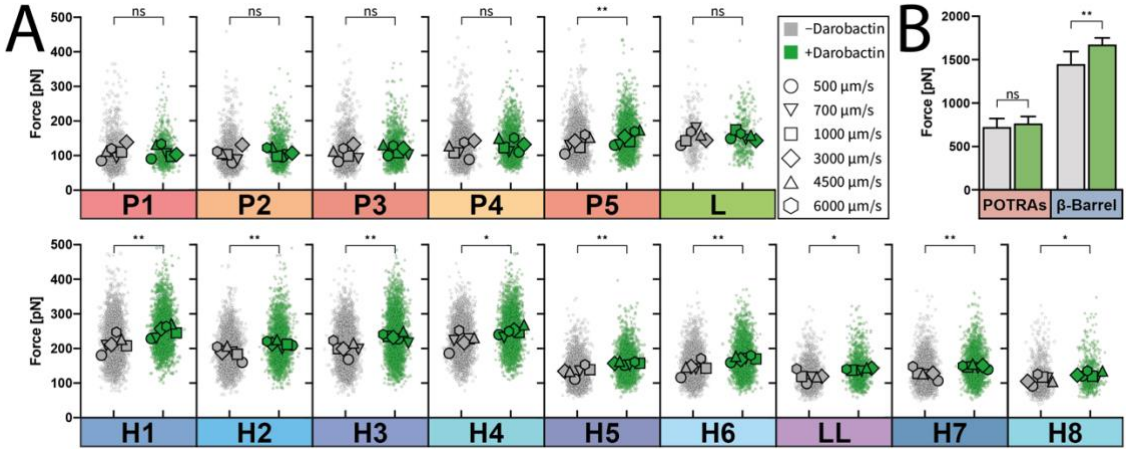


Figure 2. Unfolding Forces of BamA in the Absence and the Presence of Darobactin. (A) Forces required to mechanically unfold BamA in absence (grey) and in presence (green) of darobactin. The structural segments unfolded by BamA in response to mechanical force correspond to the POTRA domains P1 – P5, the linker region LR, the b-hairpins H1 – H8 of the transmembrane b-barrel and the lid-like structure LL in the extracellular loop 6. The datapoints within each plot show the forces of all unfolding force peaks belonging to the force peak class (Figure 1). The mean unfolding forces at different unfolding speeds are indicated by different markers. Statistical significances were calculated using two-tailed paired t-tests comparing each pair of mean forces originating from the same pulling speed. $^{**}P<0.01$; $^{*}P<0.05$; ns, non-significant, $P\geq0.05$. (B) Mechanical stability of the POTRA domains and the β-barrel domain of BamA in absence (grey) and presence (green) of darobactin. Bars represent the sum of the mean unfolding forces of all POTRA domains, including the gate class G, and of the β-barrel domain. Error bars indicate the standard deviation of the mean. Statistical significances were calculated using two-tailed paired t-tests comparing the summed up forces in absence and in presence of darobactin. $^{**}P<0.01$; $^{*}P<0.05$; ns, non-significant, $P\geq0.05$.

Close inspection of the unfolding forces showed that the first four N-terminal β-hairpins H1–H4 were the most stable structural segments of the β-barrel, which further increased stability upon darobactin-binding. Among the four N-terminal β-hairpins, the first β-hairpin H1 showed the highest mechanical stability, both in the absence and in the presence of darobactin. That H1 increases stability upon darobactin-binding can be explained by the structural model, showing darobactin to bind to the gate region of BamA, located between the first β-strand and the last β-strand of the β-barrel domain[21]. However, the presented SMFS data also shows that darobactin affected the mechanical stability of all other β-hairpins shaping the transmembrane β-barrel. Upon darobactin-binding the mechanical stability of the β-barrel domain increased by ≈ 16%, while the

remaining structural segments increased stability by ≈ 6%. Similar effects have been observed in previous SMFS studies, where localized interactions established by the binding of a ligand or a molecular compound to a membrane protein could stabilize or destabilize structural regions, not directly involved in the binding process[33,34].

In summary, the forces required to mechanically unfold structural segments of BamA increase upon darobactin binding. This indicates that BamA increases mechanical stability, with the biggest increase observed for the first four N-terminal β-hairpins of the β-barrel domain. However, the observations also indicate that the localized darobactin-binding to BamA modulates the stability of structural regions distant from the binding site.

Darobactin Modulates the Free-Energy Landscape of BamA

The mechanical forces at which a protein unfolds depends on the pulling speed (*i.e.*, the loading rate describing the force applied over time)[35–37]. In response to mechanically applied force membrane proteins, stepwise unfold structural segments[37], such as observed here for BamA. The mechanical properties as well as the kinetic or energetic stability of the structural segments at equilibrium (*i.e.,* no force applied) can be approached by measuring their unfolding forces over a broad range of pulling speeds. We hence decided to apply SMFS in the dynamic force spectroscopy (DFS) mode and mechanically unfolded BamA at retraction speeds of 500, 700, 1000, 3'000, 4'500 and 6'000 nm s^{-1} in the absence and presence of darobactin (Supporting Figure S2, S3).

The DFS data show the mean forces required to unfold individual structural segments of BamA to increase linearly with the logarithm of the loading rate (Figure 3, raw data shown in Supporting Figure S4). Fitting the Bell-Evans model[35,36,38] to the DFS plots allows to extrapolate the mechanical, kinetic and energetic properties of each structural segment at equilibrium[37]. These fits approximate the distance x_u every structural segment has to be stretched along the pulling trajectory to reach the transition state towards unfolding and the unfolding rate k_0. The x_u values can be used to describe the width of the free-energy valley hosting the folded state of a structural segment[34]. Accordingly, a larger x_u value describes a wider free-energy valley that can host a larger number of conformational substates of the structural segment. The reciprocal of k_0 represents the lifetime of a structural segment. Using x_u and k_0, the height of the free-energy barrier $\Delta G_u^{\ddagger}$ stabilizing a structural segment against unfolding and the mechanical stiffness κ (*i.e.*, spring constant) of the segment can be calculated. The free-energy landscape parameters x_u, k_0, $\Delta G_u^{\ddagger}$, and κ, which were approximated for each structural segment of BamA in the presence and in the absence of darobactin (Table 1), are described in the following.

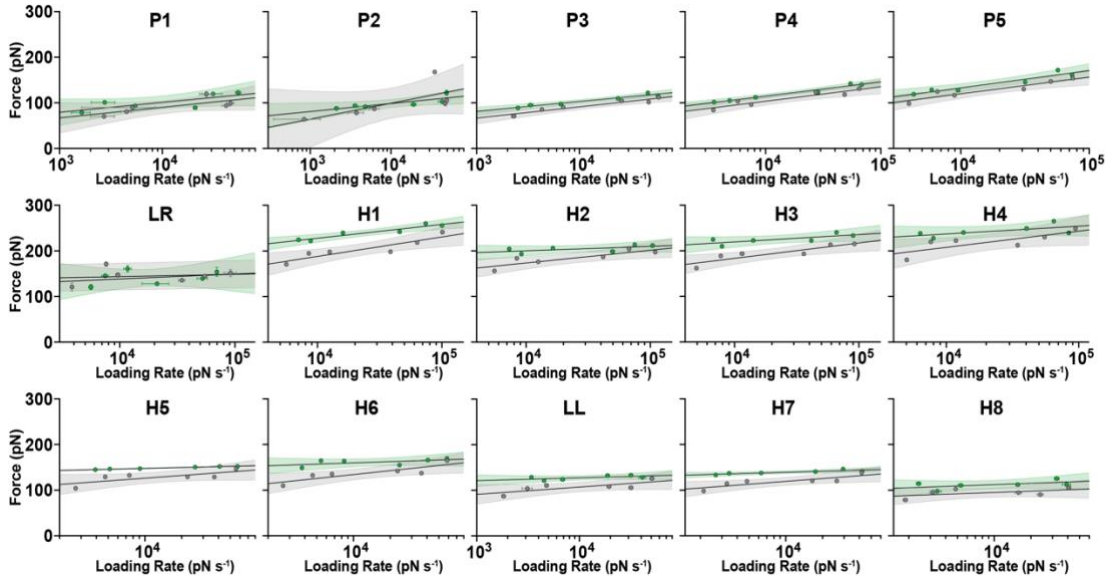


Figure 3. Estimating the Free-Energy Landscape Parameters of the Sturctural Segments of BamA in the Absence and the Presence of Darobactin by Dynamic Force Spectroscopy (DFS). The DFS-plots shows the mean unfolding forces *versus* the mean loading rates (data points) of each structural segment in the absence (grey) and the presence (green) of darobactin. Black lines result from fitting the Bell-Evans model[35,36,38] to the data with the shaded areas indicating the 95% confidence bands. Error bars represent the standard error of each data point. Free-energy parameters x_u and k_0 obtained from the fits are given in Table 1. Raw data underlying the DFS plots is shown in Supporting Figure S4.

Darobactin Widens Free-Energy Wells

The x_u values of the five POTRA domains increased from 0.32 nm to 0.39 nm (≈ 122%) in the presence of darobactin (Table 1). The only POTRA domain, which slightly decreased x_u upon darobactin-binding was P5. Among all structural segments of BamA the linker region LR showed the smallest x_u values in the absence (0.17 nm) and presence (0.20 nm) of darobactin. In contrast, the structural segments (*e.g.*, β-hairpins) of the transmembrane β-barrel region widened their free-energy wells from 0.33 nm to 0.76 nm (≈ 235%) in the presence of darobactin. The strongest increase in x_u was observed for β-hairpin H6, which moved its distance to the transition state from the initial 0.33 nm to 1.05 nm and thus increased the width of the free-energy well to 315%.

In summary, the results show that darobactin-binding to BamA only slightly widens the x_u values of the POTRA domains and of the linker region, but considerably increases the x_u values of the b-barrel domain. Thus, darobactin-binding increases the x_u value of almost every structural segment of BamA. Consequently, the free-energy valleys accommodating the structural segments widen so that they could host more and/or different conformational states. Amongst all structural segments of BamA the region linking the POTRA domains to the transmembrane b-barrel show the smallest x_u values, thus indicating to display the smallest conformational variability amongst all segments in the absence and in the presence of darobactin.

Darobactin Increases Kinetic Stability

On average the POTRA domains decreased their k_0 values to 24% in the presence of darobactin and thus considerably increased their lifetimes (Table 1). Amongst all structural segments of BamA the linker region LR showed the largest k_0 value of 1.05 s^{-1} in the absence of darobactin, which upon darobactin-binding reduced to 0.53 s^{-1}. In stark contrast the structural segments (*e.g.,* β-hairpins) of the β-barrel of BamA decreased on average their k_0 values to 6% in the presence of darobactin.

In summary, every structural segment of BamA lowers the unfolding rate k_0 und thus increases lifetime upon darobactin-binding. Naturally, by their insertion into the outer membrane, the β-hairpins of the transmembrane β-barrel expose high lifetimes, which, in the presence of darobactin is further increased by orders of magnitude. In contrast, the shortest lifetime is observed for the polypeptide stretch linking the POTRA domains to the transmembrane β-barrel.

Darobactin Increases Energetic Stability

Darobactin-binding of BamA increased the free-energy barriers stabilizing the structural segments against mechanical unfolding (Table 1). On average, the β-hairpins of the BamA β-barrel increased their $\Delta G_u^{\ddagger}$ values drastically from 25.8 k_BT in the absence to 45.0 k_BT in the presence of darobactin. Whereas the relative increase of the free-energy barriers of individual β-hairpins ranged from 103% (H8) up to 229% (H5), the average increase of all β-hairpins forming the β-barrel approached 171%. At the same time, the average $\Delta G_u^{\ddagger}$ values of the POTRA domains and the linker region LR increased slightly from 21.7 k_BT to 23.7 k_BT.

Taken together, whereas the free-energy barriers of all β-hairpins and the lid-like structure of the β-barrel increase on average to 171% upon darobactin-binding, the POTRA domains and the gate region increase the free-energy barriers on average to only 109%. We thus conclude that the binding of darobactin to BamA considerably increases the energetic stability of the b-barrel but only slightly that of the soluble POTRA domains and linker region.

Darobactin Decreases Structural Stiffness

In the absence of darobactin, the spring constants κ of the five POTRA domains ranged from 1.12 N m^{-1} (P1) to 2.67 N m^{-1} (P5) (Table 1). Among all POTRA domains, P1–P4 considerably decreased stiffness (on average to ≈77%) upon darobactin-binding. One exception was POTRA domain P5, which slightly increased stiffness from 2.67 to 2.95 N m^{-1} (to ≈111%). Among all structural segments of BamA the linker region LR showed the highest stiffness both in the unbound (5.69 N m^{-1}) and in the darobactin-bound (4.40 N m^{-1}) state. In the absence of darobactin, the structural segments representing individual β-hairpins of the transmembrane β-barrel showed spring constants ranging from 1.02 N m^{-1} (H5) to 4.09 N m^{-1} (H1). Upon darobactin-binding the β-hairpins considerably decreased spring constants now ranging from 0.25 N m^{-1} (H5) to 2.68 N m^{-1} (H8). The largest relative decrease in mechanical stiffness was measured for β-hairpin H6, which lowered from 1.85 N m^{-1} (unbound state) to 0.40 N m^{-1} (to ≈22%). In the presence of darobactin, structural segments of the β-barrel on average reduced their spring constants to ≈42% of their initial values.

In summary, upon darobactin-binding BamA reduces the structural stiffness of almost every structural segment. In the unbound state the mechanically stiffest segments are the linker region and the N-terminal β-hairpin H1. Also, the C-terminal β-hairpin H8, which together with H1 forms the gate region of BamA, shows a relatively high spring constant (mechanical stiffness) in the unbound state. Upon darobactin-binding we find that these and the other β-hairpins of the transmembrane β-barrel region considerably reduce mechanical stiffness.

		x_u (nm)		k_0 (s^{-1})		$\Delta G_u^{\ddagger}$ (k_BT)		κ (N m^{-1})	
Segment	CL (aa)	–Darobactin	+Darobactin	–Darobactin	+Darobactin	–Darobactin	+Darobactin	–Darobactin	+Darobactin
P1	56±5	0.41±0.18	0.44±0.18	0.13±0.45	0.02±0.10	22.80±3.58	24.47±4.10	1.12±0.98	1.06±0.90
P2	124±2	0.27±0.14	0.54±0.18	0.94±2.85	0.00±0.01	20.79±3.03	26.45±4.02	2.30±2.41	0.75±0.51
P3	195±2	0.38±0.06	0.44±0.05	0.18±0.22	0.02±0.02	22.42±1.19	24.75±1.28	1.24±0.38	1.05±0.27
P4	264±1	0.29±0.06	0.30±0.03	0.46±0.68	0.15±0.12	21.49±1.46	22.61±0.76	2.11±0.91	2.00±0.39
P5	331±2	0.26±0.05	0.25±0.04	0.33±0.43	0.19±0.24	21.84±1.32	22.37±1.25	2.67±0.99	2.95±0.97
LR	380±4	0.17±0.03	0.20±0.07	1.05±1.04	0.53±1.21	20.68±0.99	21.36±2.30	5.69±1.95	4.40±3.11
H1	406±0	0.22±0.05	0.31±0.05	0.02±0.05	$2.3 \times 10^{-5} \pm 6.7 \times 10^{-5}$	24.60±2.34	31.43±2.96	4.09±1.94	2.64±0.94
H2	435±1	0.34±0.10	0.96±0.60	$5.1 \times 10^{-4} \pm 2.1 \times 10^{-3}$	$1.4 \times 10^{-17} \pm 4.1 \times 10^{-16}$	28.30±4.03	59.54±29.30	2.03±1.19	0.53±0.72
H3	467±1	0.28±0.07	0.59±0.26	$2.4 \times 10^{-3} \pm 7.5 \times 10^{-3}$	$3.4 \times 10^{-11} \pm 4.7 \times 10^{-10}$	26.74±3.07	44.84±13.82	2.78±1.42	1.07±1.00
H4	511±1	0.27±0.10	0.55±0.32	$9.1 \times 10^{-4} \pm 4.6 \times 10^{-3}$	$1.8 \times 10^{-11} \pm 3.3 \times 10^{-10}$	27.73±5.09	45.48±18.57	3.20±2.53	1.22±1.50
H5	569±2	0.48±0.21	1.46±0.00	$4.7 \times 10^{-4} \pm 2.9 \times 10^{-3}$	$5.7 \times 10^{-20} \pm 3.7 \times 10^{-21}$	28.38±6.04	65.04±0.06	1.02±0.90	0.25±0.00
H6	609±1	0.33±0.10	1.05±0.75	0.02±0.05	$5.6 \times 10^{-15} \pm 1.6 \times 10^{-13}$	24.88±3.00	53.55±28.55	1.85±1.12	0.40±0.61
LL	650±2	0.32±0.14	0.47±0.28	0.05±0.20	$3.5 \times 10^{-4} \pm 2.9 \times 10^{-3}$	23.75±4.04	28.69±8.34	1.90±1.73	1.07±1.29
H7	704±3	0.45±0.14	1.23±0.36	$2.2 \times 10^{-3} \pm 8.1 \times 10^{-3}$	$2.4 \times 10^{-15} \pm 2.9 \times 10^{-14}$	26.83±3.64	54.38±11.74	1.08±0.67	0.29±0.18
H8	750±3	0.23±0.09	0.26±0.15	0.53±1.31	0.30±1.20	21.35±2.45	21.93±4.03	3.18±2.50	2.68±3.17

Table 1. Free-Energy Landscape Parameters of the Structural Segments of BamA in the Absence and in the Presence of Darobactin. For each structural segment of BamA the force peak positions are indicated by the contour lengths (CL) in amino acids. The parameters x_u, k_0, $\Delta G_u^{\ddagger}$ and κ are obtained from the DFS plots, which were recorded in the absence (-) and the presence (+) of darobactin (Figure 3). All values give mean ± sd.

5.4 Conclusions

SMFS experiments on integral β-barrel OMPs have thus far observed common mechanical unfolding pathways. By applying a mechanical pulling force to one terminal end, individual β-hairpins forming the transmembrane β-barrel unfold and extract sequentially from the membrane. This sequential unfolding of β-hairpins was observed for several OMPs, including OmpA[39], OmpG[40], FhuA[41], LamB[29], and BamA[31]. The findings presented here confirm the previous studies. The extensive data recorded in our study allows us to identify an additional force peak class of BamA (Supporting Figure S5), which locates between the POTRA domains and the β-barrel, and is referred to as linker region LR. Thus, upon mechanically pulling its N-terminal end, BamA displays 15 distinct force peak classes, each originating from the mechanical unfolding of a structural segment. The soluble POTRA domains and the linker region, which are located at the N-terminus of BamA, unfold first at comparably low forces. Afterwards the region linking POTRA domains and transmembrane b-barrel unfolds. Eight of the nine following force peak classes correspond to the mechanical unfolding of individual β-hairpins and display relatively high mechanical stability, such as previously measured for other OMPs[29,41].

Upon mechanically unfolding BamA in the absence and in the presence of darobactin, we observe the same 15 force peak classes, describing the same sequential unfolding pathways (Supporting Figure S6). This observation indicates that a characteristic network of interactions stabilizes the same structural segments of BamA in the absence and presence of darobactin. The finding is in full agreement with the structural model of BamA that shows no large structural rearrangement upon darobactin-binding[21]. However, the force peak classes recorded in the presence of darobactin display considerably higher unfolding forces (≈116%) as compared to those recorded in the absence of darobactin. Thus, darobactin considerably increases the mechanical stability of BamA. This stabilizing effect is most prominent for the β-barrel region of BamA (Figure 4). Thereby, the first four N-terminal β-hairpins H1 – H4 were the most stable β-hairpins with the first N-terminal β-hairpin H1 showing the highest mechanical stability. This observation indicates that the first β-hairpin H1 forms a mechanically stable gate together with the last C-terminal β-hairpin H8 of the BamA β-barrel. Upon binding to this gate region darobactin further stabilizes β-hairpins H1 and H8 as well as the other β-hairpins of the β-barrel. However, darobactin also increases the stability of the only essential POTRA domain P5[42]. Although the structural model does not provide conclusive evidences for a specific interaction of POTRA domain P5 and darobactin, our data shows

that the properties of POTRA domain P5 are modulated *via* the darobactin-mediated inhibition of BamA. Interestingly, it has been previously observed that darobactin-resistant strains show mutations of residue F394 in the POTRA domain P5[20]. The darobactin-dependent modulation of the mechanical stability of P5 thus adds new insight into the manifold interactions of darobactin with BamA.

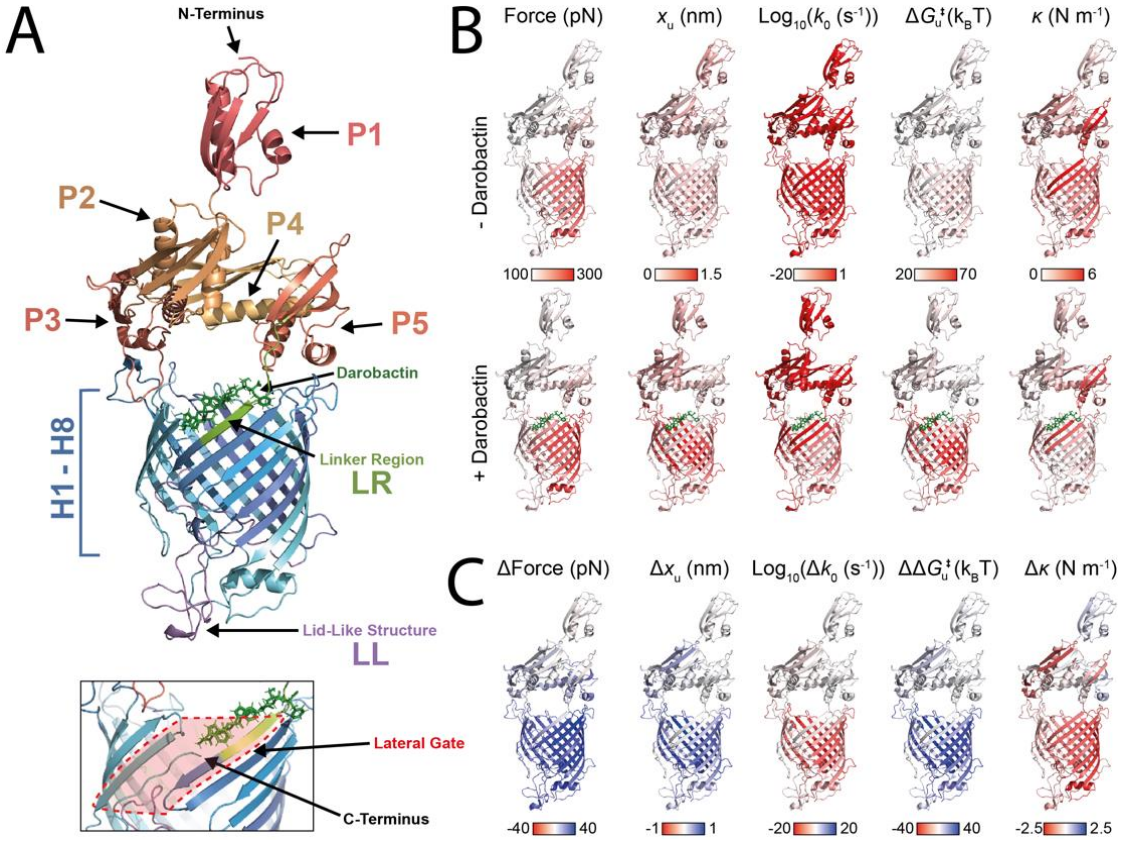


Figure 4. Effect of Darobactin on the Mechanical and Energetic Properties of BamA. (A) Structure of BamA bound to darobactin (PDB: 7NRI)[21]. Darobactin is shown in green. Colored segments of BamA represent force peak classes, which were assigned to structural segments, as shown in Figure 1. Both termini are indicated in black. Inset at the bottom left shows the lateral gate of BamA, indicated by the red area. (B) Mapping the forces required to mechanically unfold BamA and the parameters describing the free-energy landscape to the unliganded (PDB: 5D0O)[55] and to the darobactin-bound BamA structure. (C) Mapping the relative changes in the forces required to mechanically unfold BamA and in the parameters describing the free-energy landscape induced by darobactin to the BamA structure.

To further investigate the complex effects of darobactin on the structural properties of BamA, we employed SMFS in the dynamic mode. For every structural segment of BamA, we find a linear relationship between the mean unfolding force and the logarithm of the mean loading rate. This suggests, that the folded state of every segment is separated from the unfolded state by a single free-energy barrier[35]. We observe the general trend that darobactin increases the distance to the transition state (x_u values), increases the lifetime, enhances the free-energy barrier, but lowers the mechanical stiffness (or rigidity) of almost every structural segment of BamA (Figure 4). As these changes apply at different magnitudes in the structural segments, we will discuss them in the following in detail.
One prominent effect is observed for the structural region linking the POTRA domains to the transmembrane β-barrel. This linker region shows the lowest kinetic stability amongst all structural segments of BamA in the absence and presence of darobactin.

Interestingly the linker region also shows the smallest transition state distances in the absence and presence of darobactin, thus indicating to adopt the smallest conformational variability amongst all structural segments of BamA. Moreover, the linker region exposes the highest spring constant amongst all structural segments of BamA, thus showing the highest mechanical stiffness. Together the low kinetic stability and conformational variability, and the high mechanical stiffness of the linker region provide new functional insight of how the POTRA domains connect to the β-barrel domain.

Another remarkable insight is provided for the transmembrane β-barrel. Although darobactin increases the height of every free-energy barrier stabilizing a structural segment of BamA (Figure 4), which means that all segments stabilize energetically, the highest increase (≈ 171%) was observed for the structural segments (e.g., β-hairpins) of the transmembrane β-barrel. In the unbound state, the mechanically stiffest segments of BamA are the linker region and the N-terminal β-hairpin H1. Also, the N-terminal β-hairpins H2, H3, H4 and the C-terminal β-hairpin H8 show a relatively high mechanical stiffness in the unbound state. Comparatively, β-hairpins H5 – H7 are among the mechanically most flexible regions of BamA. This suggests, that if a substrate is funneled into the membrane by the two relatively stiff and stable β-hairpins H1 and H8, which form the lateral gate of BamA, the b-hairpins H5 – H7 provide a mechanically flexible region such as needed to enlarge the transmembrane β-barrel and to insert and fold the substrate. As darobactin considerably changes the mechanical stiffnesses of almost every structural region of BamA as well as their mechanical, kinetic and energetic stabilities, these fine-tuned properties come out of balance. Particularly, the stiffnesses of all structural segments of the β-barrel decrease. In summary, the data thus shows that in the presence of darobactin the β-barrel domain of BamA increases its relatively high energetic stability and decreases its mechanical stiffness to structurally soften. Most importantly, however, darobactin binds between b-hairpins H1 and H8 thus sealing the lateral gate to prevent substrate inclusion[21].

Taken together, we observe darobactin to mechanically stabilize BamA considerably. The stabilizing effects are most prominent for the entire β-barrel domain of BamA, although darobactin only binds to the lateral gate region[21]. The free-energy landscape parameters characterizing the properties of every structural segment of BamA describe how darobactin modulates the mechanical, kinetic and energetic properties of BamA at great detail (Figure 4). Together with the structural models of BamA in the unbound state and in the darobactin-bound state, the parameters provide mechanistic insight into how the rather complex BamA machinery is inhibited. Further SMFS studies may be conducted in presence of β-signals of varying OMPs to investigate how transient interactions modulate the mechanical, kinetic and energetic properties of BamA in contrast to darobactin. Similarly, the effects of different antibiotic compounds targeting the BAM complex could be studied by SMFS to elaborate on current findings. Gaining a deeper understanding of different factors shaping BamA-mediated OMP insertion and folding will be beneficial for the rational design of novel antibiotic compounds against the BAM complex and will hopefully advance our efforts to resolve the current antimicrobial resistance crisis.

5.5 Materials and Methods

Cloning. Plasmid pNR48, which was used to overexpress BamA into the *E. coli* outer membrane was cloned in two steps, in which an N-terminal flexible linker was attached to facilitate unfolding experiments. First the sequence coding for FFSARGGSGSGS was inserted into plasmid pY161[31] after the N-terminal periplasmic export sequence using primers 5'-GCGTGGCGGCTCTGGTTCCGGTTCTGCTGAAGGGTTCGTAGTGA-3' and 5'-CCAGAGCCGCCACGCGCAGAAAAGAAACCGTATACGGTGGCGC-3' with the QuikChange PCR method[43]. In a second step the sequence coding for NEEG was inserted in front of the newly inserted sequence with primers5'-GGTAACGAAGAAGGCTTCTTTTCTGCGCGTGG-3'and 5'-GAAGCCTTCTTCGTTACCGTATACGGTGGCG-3' again using the QuikChange PCR method.

Preparation of Outer Membrane Vesicles (OMVs) Enriched with BamA. The plasmid pNR48 was transformed into *E. coli* BL21(DE3)omp8[44]. OMVs were prepared as described[54]. Briefly, the *E. coli* were grown on a Luria-Bertani (LB) Agar plate (Lennox L Agar, Invitrogen) supplemented with 100 μg ml^{-1} ampicillin (Ampicillin sodium salt, Sigma) at 37°C. A single colony was picked and grown over night in LB medium (Difco, Becton Dickinson) supplemented with 100 μg ml^{-1} ampicillin at 37°C under constant shaking at 220 rpm. 300 ml LB medium in a baffled Erlenmeyer flask containing 100 μg ml^{-1} ampicillin were inoculated 1:100 with the overnight culture and grown at 37°C, 220 rpm while the optical density at 600 nm (OD_{600}) was closely monitored. Overexpression of BamA was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma) when $OD_{600} \approx 0.4$ was reached. After induction the cells were incubated until the onset of saturation in growth was reached. The cells were removed by centrifugation at 10'000 xg for 10 min. The supernatant was sterile-filtered using a 450 nm filter unit (Merk Millipore) and stored overnight at 4°C. On the following day, OMVs were collected by centrifugation at 38'400 xg for 1.5 h and resuspended in 12 ml Dulbecco's phosphate buffered saline with added magnesium and calcium (DPBSS, Sigma). The OMVs were washed in an additional centrifugation step at 108'000 xg for 1 h and resuspended in 1 ml of DPBSS. The resulting solution was stored in aliquots at –80°C.

Single-Molecule Force Spectroscopy (SMFS) and Dynamic Force Spectroscopy (DFS). For each SMFS experiment, an aliquot containing 5 μl OMVs in DPBSS was thawed at 4°C and filled up to 50 μl with DPBSS. The sample was centrifuged for 15 min at 16'100 xg at 4°C and the pellet was resuspended in 40 μl DPBSS. 5 μl of the resulting sample were diluted in 45 μl DPBSS and adsorbed to freshly cleaved mica. After 15 min the sample was rinsed several times with DPBSS to remove non-adsorbed material. The sample was filled up to 1.5 ml and covered with a silicon skirt to prevent evaporation. SMFS was performed at ≈ 25°C using a commercial AFM (Nanowizard II Ultra, JPK Instruments) and OMCL-RC800PSA cantilevers (Olympus), which were calibrated with the thermal noise method[45]. Membrane patches were located by contact mode AFM imaging. The tip of the AFM cantilever was pushed onto densely packed regions of the membrane for 500 ms with a force of 1 nN to non-specifically attach BamA. DFS was conducted at six different retraction speeds (500, 700, 1'000, 3'000, 4'500 and 6'000 nm s^{-1}). At least five different cantilevers were used for each experimental condition to minimize errors due to uncertainties in cantilever spring constant calibration. SMFS experiments conducted at different retraction speeds were recorded in mixed and random order to reduce the influence of external factors on the experimental outcome. For SFMS in presence of darobactin, the sample preparation, AFM imaging and SMFS was conducted in DPBSS

containing 20 μM darobactin. An external 16-bit data acquisition card (NI PCI-6221, National Instruments) was used to record SMFS data at high-frequencies.

SMFS Data Analysis. For each experimental condition, the FD curves were transformed to force *versus* contour length space using the worm-like chain (WLC) model[46],

$$F(x) = \frac{k_B T}{P} = \left[\frac{1}{4}\left(1 - \frac{x}{L}\right)^{-2} - \frac{1}{4} + \frac{x}{L}\right] \qquad \text{Equation 1}$$

where k_B is the Boltzmann constant, T the absolute temperature in Kelvin, P the average polypeptide persistence length (assuming a length of 0.36 nm per amino acid (aa)), x the extension in nm and L the polypeptide contour length in aa. A fixed persistence length of 0.4 nm was used[47]. The BamA unfolding curves were aligned to each other in force versus contour length space using a self-written software. In the next step the position of every force peak in every FD curve was detected. To assign individual force peaks to force peak classes, we used the OPTICS clustering from scikit learn[48,49] to automatically cluster the force peaks one SMFS data set recorded at 1000 nm s^{-1} retraction speed in presence of darobactin. After the force peak positions were clustered and the mean contour lengths determined, we used the resulting force peak clusters as a template to classify the FD curves recorded under other experimental conditions.
The aligned and classified FD curves were analysed as described[29,33]. Briefly, the loading rate of every force peak was determined by the slope of a linear fit to the last data points before of the rupture event described by the respective force peak in the force-time curve[26]. For each experimental condition the rupture forces and loading rates of all force peaks belonging to a force peak class were binned and the resulting histograms were fitted with Gaussians to determine the most probable rupture force and loading rate (Supporting Figure S4). The resulting means and standard errors of each force peak class were fitted with the Bell-Evans model[35,36,38],

$$F^* = \frac{k_B T}{x_u} ln\left(\frac{x_u r}{k_B T k_0}\right) \qquad \text{Equation 2}$$

where F^* is the most probable rupture force in pN and r is the loading rate in pN s^{-1} (Figure 3). Both values allowed to approximate the distance x_u, separating the free-energy valley of the folded structural segment from the unfolded state, and the unfolding rate k_0 of the structural segment. The parameter $\Delta G_u^{\ddagger}$ describing the height of the free-energy barrier stabilizing a structural segment against unfolding was calculated using

$$\Delta G_u^{\ddagger} = -k_B T ln(\tau_A k_0) \qquad \text{Equation 3}$$

where τ_A is the Arrhenius frequency[50], for which we chose 10^{-8} s^{-1} (ref. [51]). The parameter κ, which describes the stiffness of a structural segment, was calculated by:

$$\kappa = \frac{2\Delta G_u^{\ddagger}}{x_u^2} \qquad \text{Equation 4}$$

Errors were propagated using the python uncertainties package[52].

Force–Distance Curve-Based AFM (FD-Based AFM) Imaging of OMVs. OMVs enriched with BamA were diluted (1:50) in DPBSS and adsorbed onto freshly cleaved mica for

15 min at room temperature. After the adsorption, the sample was gently washed with DPBSS to remove non-adsorbed membranes. FD-based AFM imaging of OMVs was performed with a Nanoscope Multimode 8 (Bruker, USA) operated in PeakForce Tapping mode in DPBSS at room temperature as described[53,54]. The AFM was placed in a temperature-controlled acoustic isolation box and equipped with a 120-µm piezoelectric J scanner and fluid cell. AFM topographs were recorded using PEAKFORCE-HiRs-F-A (Bruker Nano Inc., USA) cantilevers having a nominal spring constant of 0.4 N m^{-1}, a resonance frequency of ≈ 165 kHz in liquid, and a sharpened silicon tip with a nominal radius of ≈ 1 nm. Before imaging, cantilevers were calibrated by ramping on the mica surface and the thermal noise method[45]. AFM topographs were recorded by applying an imaging force of 100–120 pN at 2 kHz oscillation frequency, with a vertical oscillation amplitude of 30 nm and a resolution of 512 × 512 pixels. Image post-processing and analysis were performed using the Nanoscope Analysis software v.1.8.

Acknowledgements
This work was supported by the Swiss National Science Foundation, the NCCR Molecular Systems Engineering and the Swiss Nanoscience Institute (SNI, Basel, Switzerland).

Abbreviations
aa, amino acid; AFM, atomic force microscopy; BAM, β-barrel assembly machinery; BamA, outer membrane protein assembly factor; BamA-E, subunits of the BAM complex; DFS, dynamic force spectroscopy; DPBSS, Dulbecco's phosphate buffered saline with added magnesium and calcium; FD curve, force-distance curve; FhuA, LamB, OmpA, OmpG, outer membrane proteins; IPTG, isopropyl-β-D-thiogalactopyranoside; LB, Luria-Bertani; OD_{600}, optical density at a wavelength of 600 nm; OMP, outer membrane protein; OMV, outer membrane vesicle; P1-P5, LR, H1-H8, LL, structural segments of BamA; POTRA, polypeptide transport; sd, standard deviation; SMFS, single-molecule force spectroscopy; WLC, worm-like chain;

5.6 References

(1) Klebba, P. E. and Newton, S. M. C. Mechanisms of solute transport through outer membrane porins: burning down the house. *Curr. Opin. Microbiol.* 1998, 1, 238–248.
(2) Lauber, F.; Deme, J. C.; Lea, S. M. and Berks, B. C. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* 2018, 564, 77–82.
(3) Heras, B.; Totsika, M.; Peters, K. M.; Paxman, J. J.; Gee, C. L.; Jarrott, R. J.; Perugini, M. A.; Whitten, A. E. and Schembri, M. A. The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proc. Natl. Acad. Sci. U. S. A.* 2014, 111, 457–462.
(4) Voulhoux, R.; Bos, M. P.; Geurtsen, J.; Mols, M. and Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 2003, 299, 262–265.
(5) Gentle, I.; Gabriel, K.; Beech, P.; Waller, R. and Lithgow, T. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol.* 2004, 164, 19–24.
(6) Wu, T.; Malinverni, J.; Ruiz, N.; Kim, S.; Silhavy, T. J. and Kahne, D. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* 2005, 121, 235–245.
(7) Sánchez-Pulido, L.; Devos, D.; Genevrois, S.; Vicente, M. and Valencia, A. POTRA: a conserved domain in the FtsQ family and a class of β-barrel outer membrane proteins. *Trends Biochem. Sci.* 2003, 28, 523–526.
(8) Gentle, I. E.; Burri, L. and Lithgow, T. Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* 2005, 58, 1216–1225.
(9) Hagan, C. L.; Kim, S. and Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* 2010, 328, 890–892.
(10) Malinverni, J. C.; Werner, J.; Kim, S.; Sklar, J. G.; Kahne, D.; Misra, R. and Silhavy, T. J. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. *Mol. Microbiol.* 2006, 61, 151–164.
(11) Horne, J. E.; Brockwell, D. J. and Radford, S. E. Role of the lipid bilayer in outer membrane protein folding in Gram-negative bacteria. *J. Biol. Chem.* 2020, 295, 10340–10367.
(12) Gruss, F.; Zähringer, F.; Jakob, R. P.; Burmann, B. M.; Hiller, S. and Maier, T. The structural basis of autotransporter translocation by TamA. *Nat. Struct. Mol. Biol.* 2013, 20, 1318–1320.
(13) Noinaj, N.; Kuszak, A. J.; Gumbart, J. C.; Lukacik, P.; Chang, H.; Easley, N. C.; Lithgow, T. and Buchanan, S. K. Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* 2013, 501, 385–390.
(14) Lee, J.; Tomasek, D.; Santos, T. M.; May, M. D.; Meuskens, I. and Kahne, D. Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *Elife* 2019, 8, 1–20.
(15) Höhr, A. I. C.; Lindau, C.; Wirth, C.; Qiu, J.; Stroud, D. A.; Kutik, S.; Guiard, B.; Hunte, C.; Becker, T.; Pfanner, N. and Wiedemann, N. Membrane protein insertion through a mitochondrial β-barrel gate. *Science* 2018, 359, 1–12.
(16) Tomasek, D.; Rawson, S.; Lee, J.; Wzorek, J. S.; Harrison, S. C.; Li, Z. and Kahne, D. Structure of a nascent membrane protein as it folds on the BAM complex. *Nature* 2020, 583, 473–478.
(17) Xiao, L.; Han, L.; Li, B.; Zhang, M.; Zhou, H.; Luo, Q.; Zhang, X. and Huang, Y. Structures of the β-barrel assembly machine recognizing outer membrane protein substrates. *FASEB J.* 2021, 35, 1–13.

(18) Robert, V.; Volokhina, E. B.; Senf, F.; Bos, M. P.; Van Gelder, P. and Tommassen, J. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* 2006, 4, 1984–1995.
(19) Tomasek, D. and Kahne, D. The assembly of β-barrel outer membrane proteins. *Curr. Opin. Microbiol.* 2021, 60, 16–23.
(20) Imai, Y.; Meyer, K. J.; Iinishi, A.; Favre-Godal, Q.; Green, R.; Manuse, S.; Caboni, M.; Mori, M.; Niles, S.; Ghiglieri, M.; Honrao, C.; Ma, X.; Guo, J. J.; Makriyannis, A.; Linares-Otoya, L.; Böhringer, N.; Wuisan, Z. G.; Kaur, H.; Wu, R.; Mateus, A; Typas, A; Savitski, M.M.; Espinoza, J.L.; O'Rourke, A; Nelson, K.E.; Hiller, S.; Noinaj, N; Schäberle, T.F; D'Onofrio, A and Lewis, K. A new antibiotic selectively kills Gram-negative pathogens. *Nature* 2019, 576, 459–464.
(21) Kaur, H.; Jakob, R. P.; Marzinek, J. K.; Green, R.; Imai, Y.; Bolla, J. R.; Agustoni, E.; Robinson, C. V.; Bond, P. J.; Lewis, K.; Maier, T. and Hiller, S. The antibiotic darobactin mimics a β-strand to inhibit outer membrane insertase. *Nature* 2021, 593, 125–129.
(22) Dufrêne, Y. F.; Ando, T.; Garcia, R.; Alsteens, D.; Martinez-Martin, D.; Engel, A.; Gerber, C. and Müller, D. J. Imaging modes of atomic force microscopy for application in molecular and cell biology. *Nat. Nanotechnol.* 2017, 12, 295–307.
(23) Fotiadis, D.; Scheuring, S.; Müller, S. A.; Engel, A. and Müller, D. J. Imaging and manipulation of biological structures with the AFM. *Micron* 2002, 33, 385–397.
(24) Müller, D. J.; Schabert, F. A.; Büldt, G. and Engel, A. Imaging purple membranes in aqueous solutions at sub-nanometer resolution by atomic force microscopy. *Biophys. J.* 1995, 68, 1681–1686.
(25) Mulvihill, E.; Pfreundschuh, M.; Thoma, J.; Ritzmann, N. and Müller, D. J. High-resolution imaging of maltoporin LamB while quantifying the free-energy landscape and asymmetry of sugar binding. *Nano Lett.* 2019, 19, 6442–6453.
(26) Alsteens, D.; Pfreundschuh, M.; Zhang, C.; Spoerri, P. M.; Coughlin, S. R.; Kobilka, B. K. and Müller, D. J. Imaging G protein-coupled receptors while quantifying their ligand-binding free-energy landscape. *Nat. Methods* 2015, 12, 845–851.
(27) Lo Giudice, C.; Dumitru, A. C. and Alsteens, D. Probing ligand-receptor bonds in physiologically relevant conditions using AFM. *Anal. Bioanal. Chem.* 2019, 411, 6549–6559.
(28) Engel, A. and Gaub, H. E. Structure and mechanics of membrane proteins. *Annu. Rev. Biochem.* 2008, 77, 127–148.
(29) Thoma, J.; Ritzmann, N.; Wolf, D.; Mulvihill, E.; Hiller, S. and Müller, D. J. Maltoporin LamB unfolds β hairpins along mechanical stress-dependent unfolding pathways. *Structure* 2017, 25, 1139–1144.
(30) Oesterhelt, F.; Oesterhelt, D.; Pfeiffer, M.; Engel, A.; Gaub, H. E. and Müller, D. J. Unfolding pathways of individual bacteriorhodopsins. *Science* 2000, 288, 143–146.
(31) Thoma, J.; Sun, Y.; Ritzmann, N. and Müller, D. J. POTRA domains, extracellular lid, and membrane composition modulate the conformational stability of the β barrel assembly factor BamA. *Structure* 2018, 26, 987–996.
(32) Maier, T.; Clantin, B.; Gruss, F.; Dewitte, F.; Delattre, A. S.; Jacob-Dubuisson, F.; Hiller, S. and Villeret, V. Conserved Omp85 lid-lock structure and substrate recognition in FhaC. *Nat. Commun.* 2015, 6, 1–9.
(33) Spoerri, P. M.; Sapra, K. T.; Zhang, C.; Mari, S. A.; Kato, H. E.; Kobilka, B. K. and Müller, D. J. Conformational plasticity of human protease-activated receptor 1 upon antagonist- and agonist-binding. *Structure* 2019, 27, 1517–1526.

(34) Serdiuk, T.; Madej, M. G.; Sugihara, J.; Kawamura, S.; Mari, S. A.; Kaback, H. R. and Müller, D. J. Substrate-induced changes in the structural properties of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 2014, 111, 1571–1580.
(35) Evans, E. and Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 1997, 72, 1541–1555.
(36) Evans, E. Probing the relation between force - lifetime - and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* 2001, 30, 105–128.
(37) Bippes, C. A. and Muller, D. J. High-resolution atomic force microscopy and spectroscopy of native membrane proteins. *Reports Prog. Phys.* 2011, 74, 1–43.
(38) Bell, G. I. Models for the specific adhesion of cells to cells. *Science* 1978, 200, 618–627.
(39) Bosshart, P. D.; Iordanov, I.; Garzon-Coral, C.; Demange, P.; Engel, A.; Milon, A. and Müller, D. J. The transmembrane protein KpOmpA anchoring the outer membrane of Klebsiella pneumoniae unfolds and refolds in response to tensile load. *Structure* 2012, 20, 121–127.
(40) Sapra, K. T.; Damaghi, M.; Köster, S.; Yildiz, Ö.; Kühlbrandt, W. and Müller, D. J. One β hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG. *Angew. Chemie* 2009, 121, 8456–8458.
(41) Thoma, J.; Bosshart, P.; Pfreundschuh, M. and Müller, D. J. Out but not in: The large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. *Structure* 2012, 20, 2185–2190.
(42) Bos, M. P.; Robert, V. and Tommassen, J. Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep.* 2007, 8, 1149–1154.
(43) Cormack, B. and Castaño, I. Introduction of point mutations into cloned genes. *Methods Enzymol.* 2002, 350, 199–218.
(44) Prilipov, A.; Phale, P. S.; Van Gelder, P.; Rosenbusch, J. P. and Koebnik, R. Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from E. coli. *FEMS Microbiol. Lett.* 1998, 163, 65–72.
(45) Butt, H. J. and Jaschke, M. Calculation of thermal noise in atomic force microscopy. *Nanotechnology* 1995, 6, 1–7.
(46) Bustamante, C.; Marko, J. F.; Siggia, E. D. and Smith, S. Entropic elasticity of lambda-phage DNA. *Science* 1994, 265, 1599–1600.
(47) Bosshart, P. D.; Frederix, P. L. T. M. and Engel, A. Reference-free alignment and sorting of single-molecule force spectroscopy data. *Biophys. J.* 2012, 102, 2202–2211.
(48) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M. and Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 2011, 12, 2825–2830.
(49) Ankerst, M.; Breunig, M. M.; Kriegel, H. P. and Sander, J. OPTICS: ordering points to identify the clustering structure. *ACM SIGMOD Rec.* 1999, 28, 49–60.
(50) Gräter, F.; Shen, J.; Jiang, H.; Gautel, M. and Grubmüller, H. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. *Biophys. J.* 2005, 88, 790–804.
(51) Zocher, M.; Fung, J. J.; Kobilka, B. K. and Müller, D. J. Ligand-specific interactions modulate kinetic, energetic, and mechanical properties of the human β2 adrenergic receptor. *Structure* 2012, 20, 1391–1402.
(52) Lebigot, E. O. Uncertainties: a Python package for calculations with uncertainties, http://pythonhosted.org/uncertainties/. 2018.

(53) Pfreundschuh, M.; Martinez-Martin, D.; Mulvihill, E.; Wegmann, S. and Muller, D. J. Multiparametric high-resolution imaging of native proteins by force-distance curve-based AFM. *Nat. Protoc.* 2014, 9, 1113–1130.
(54) Thoma, J.; Manioglu, S.; Kalbermatter, D.; Bosshart, P. D.; Fotiadis, D. and Müller, D. J. Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. *Commun. Biol.* 2018, 1, 1–9.
(55) Gu, Y.; Li, H.; Dong, H.; Zeng, Y.; Zhang, Z.; Paterson, N. G.; Stansfeld, P. J.; Wang, Z.; Zhang, Y.; Wang, W. and Dong, C. Structural basis of outer membrane protein insertion by the BAM complex. *Nature* 2016, 531, 64–69.

5.7 Supporting Information

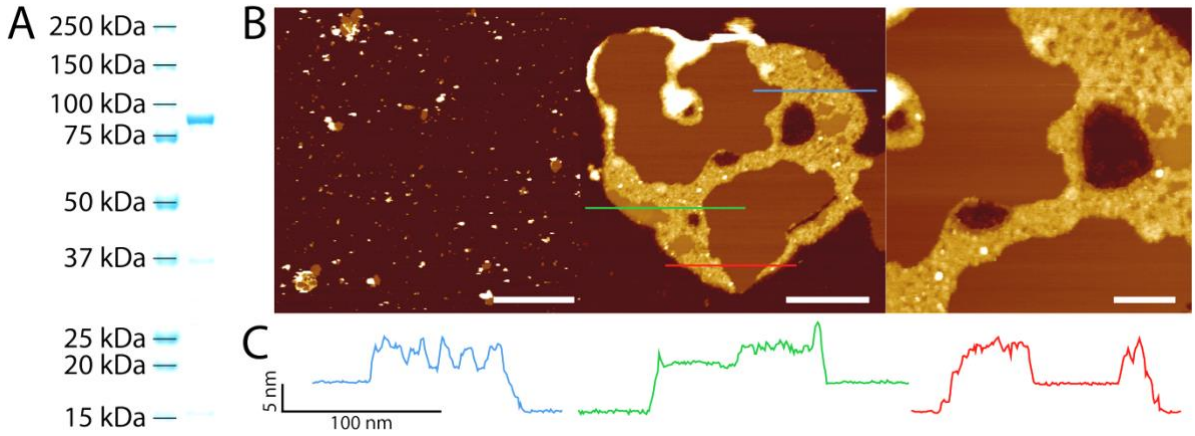


Figure S1. Biochemical Analysis and Atomic Force Microscopy (AFM) Imaging of Outer Membrane Vesicles (OMVs) enriched with BamA. (**A**) SDS gel of OMVs enriched with BamA as characterized in this study (**Materials and Methods**). (**B**) AFM topographs of OMVs enriched with BamA. Upon adsorption to freshly cleaved mica the OMVs open up and form planar membrane patches (left). Middle and right topographs, the densely packed areas of BamA appear higher (yellow, heights of ≈ 10 – 15 nm) than the surrounding membrane (heights of 5 – 8 nm). Blue, green, and red lines indicate height profiles taken for analysis (**C**). Scale bars from left to right represent 5 μm, 160 nm and 130 nm, respectively. The full color range of the topographs corresponds to a vertical scale of 30 nm. (**C**) Height profiles taken along the colored lines as indicated in the middle AFM topography.

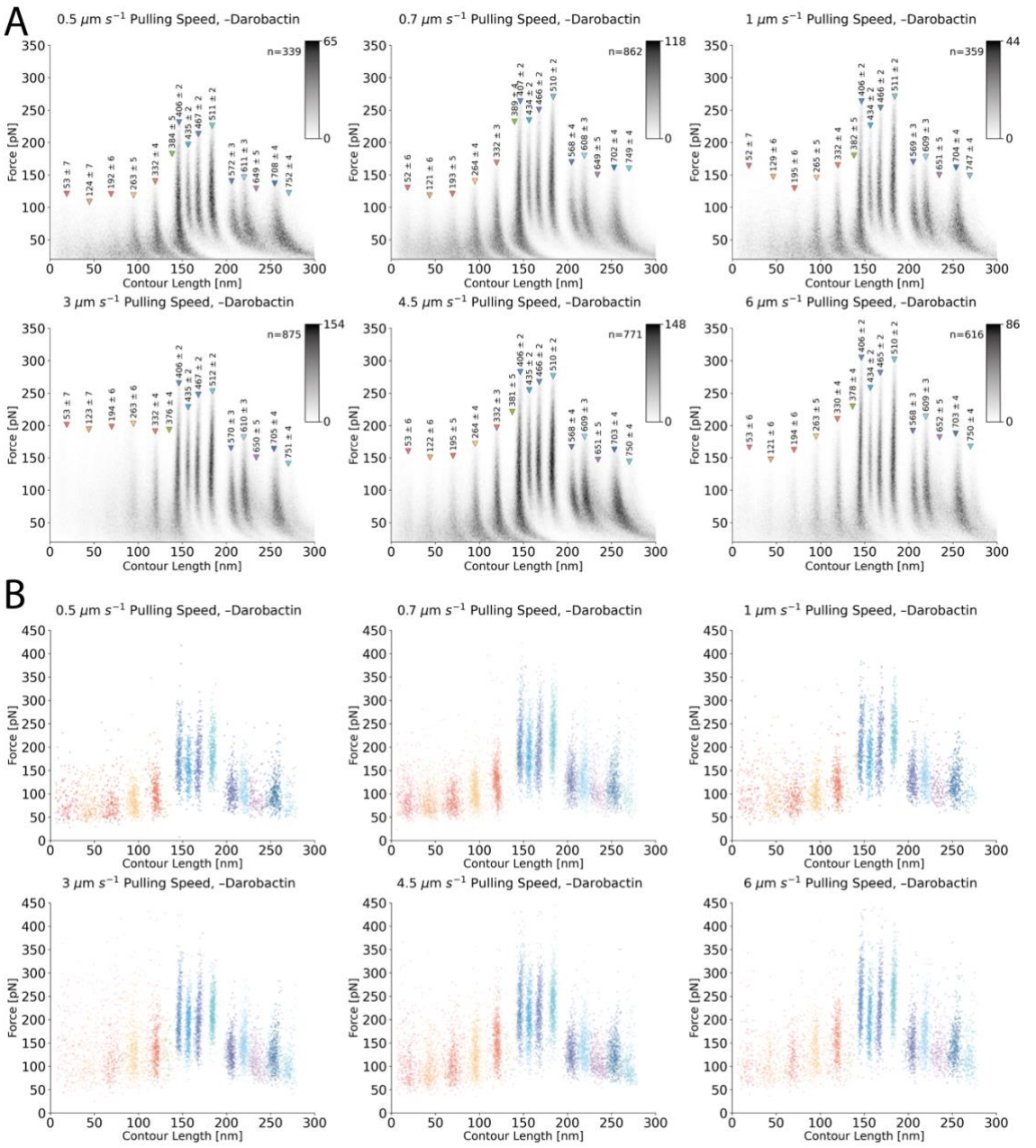


Figure S2. Superimposition of FD Curves and Classification of Force-Peaks recorded upon the mechanical unfolding of BamA in the Absence of Darobactin (**A**) Superimposition of force curves each recording the unfolding of a single BamA from the native outer membrane. The force curves were recorded at different pulling speeds in Dulbecco's phosphate buffered saline with added magnesium and calcium (DPBSS). The unfolding force curves were aligned and superimposed in force *versus* contour length space. The force-contour length curves are displayed as density plots and grey-scale bars indicate the number of data points superimposed per bin. Coloured triangles indicate the mean contour lengths of unfolding force peak classes of BamA. Each force peak class denotes the unfolding of a structural segment of BamA (**Figure 1**). Numbers above the triangles represent the means and standard deviations of the force peak class in amino acids. The number *n* in the top-right of each plot indicates the number of superimposed force curves. (**B**) Classification of force-peaks. Each plot displays all force-peaks of all force curves recording the mechanical unfolding of individual BamA proteins. The curves were recorded at the indicated unfolding speed. Each dot represents the force and contour length of one unfolding force peak. The coloration indicates the class to which a specific force-peak belongs to.

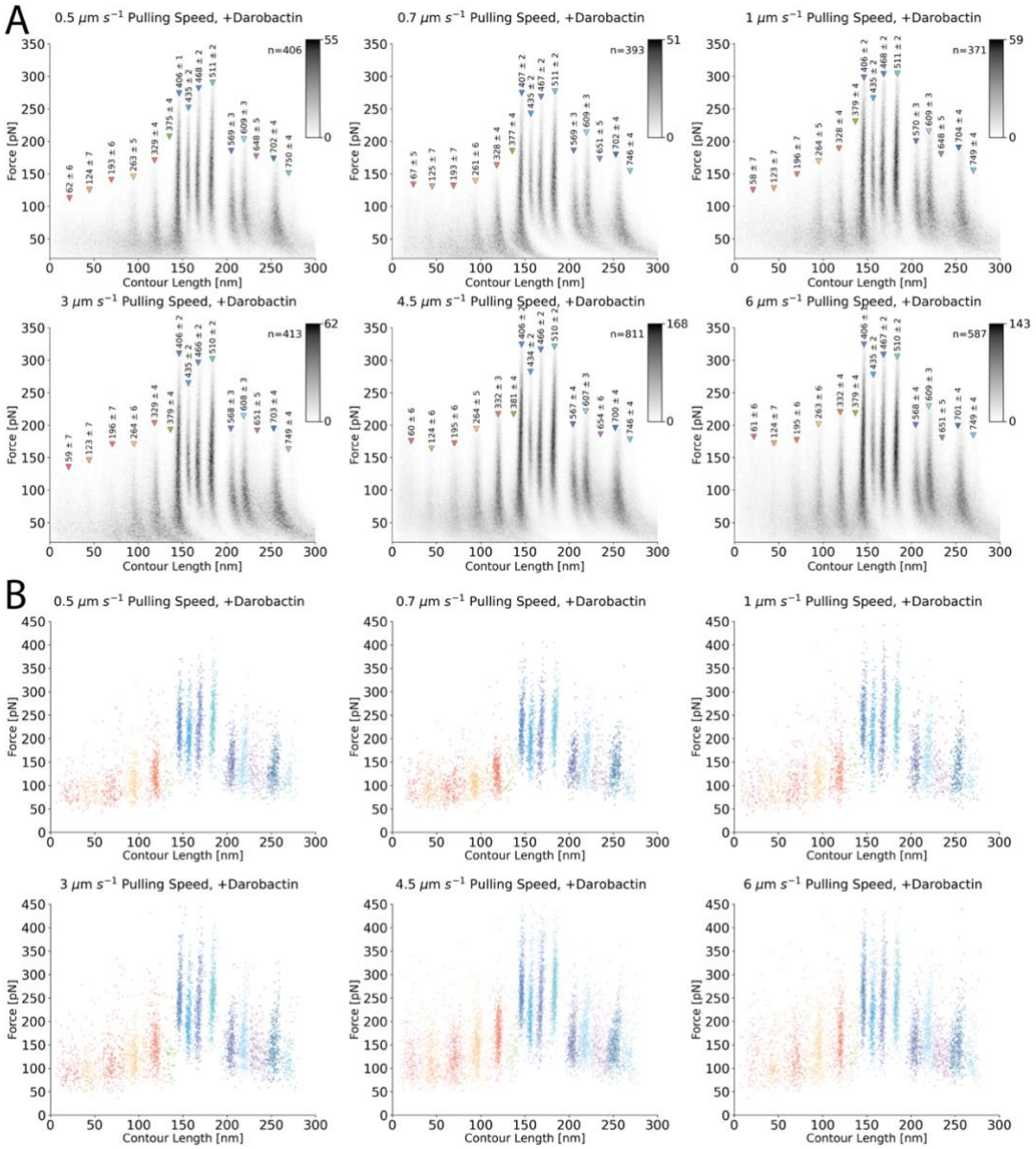
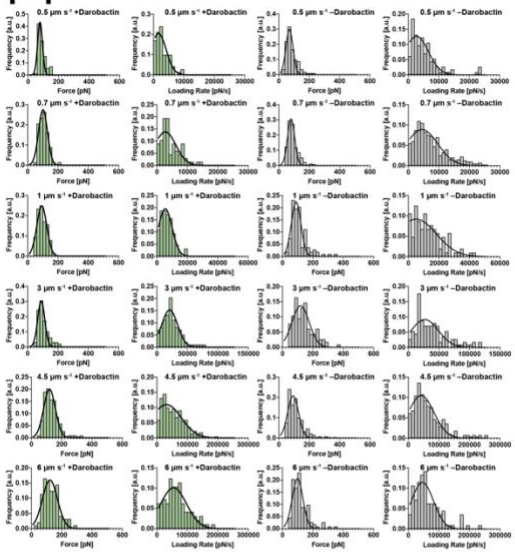
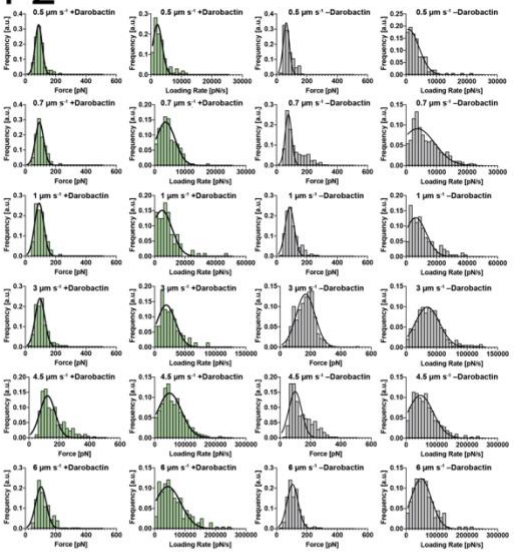


Figure S3. Superimposition of FD Curves and Classification of Force-Peaks recorded upon the mechanical unfolding of BamA in the Presence of Darobactin (**A**) Superimposition of force curves each recording the unfolding of a single BamA from the native outer membrane. The force curves were recorded at different pulling speeds in the presence of 20 μM darobactin, measured in Dulbecco's phosphate buffered saline with added magnesium and calcium (DPBSS). The unfolding force curves were aligned and superimposed in force *versus* contour length space. The force-contour length curves are displayed as density plots and grey-scale bars indicate the number of data points superimposed per bin. Coloured triangles indicate the mean contour lengths of unfolding force peak classes of BamA. Each force peak class denotes the unfolding of a structural segment of BamA (**Figure 1**). Numbers above the triangles represent the means and standard deviations of the force peak class in amino acids. The number *n* in the top-right of each plot indicates the number of superimposed force curves. (**B**) Classification of force-peaks. Each plot displays all force-peaks of all force curves recording the mechanical unfolding of individual BamA proteins. The curves were recorded at the indicated unfolding speed. Each dot represents the force and contour length of one unfolding force peak. The coloration indicates the class to which a specific force-peak belongs to.

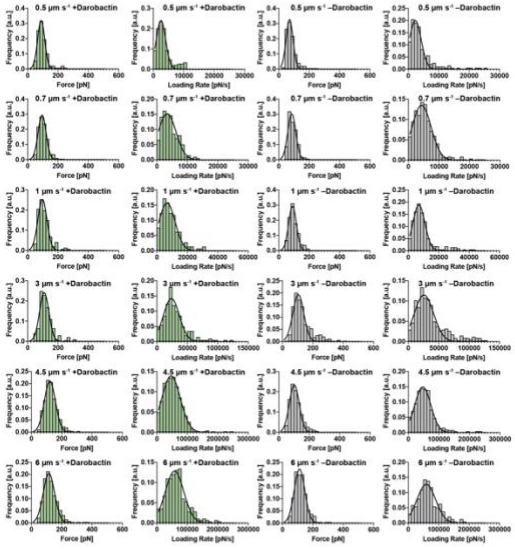
P1



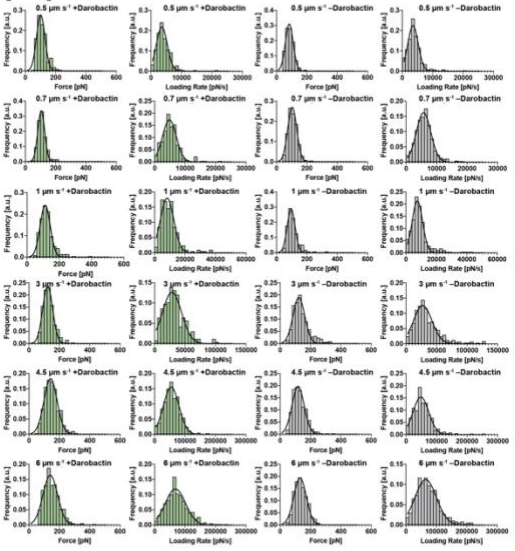
P2



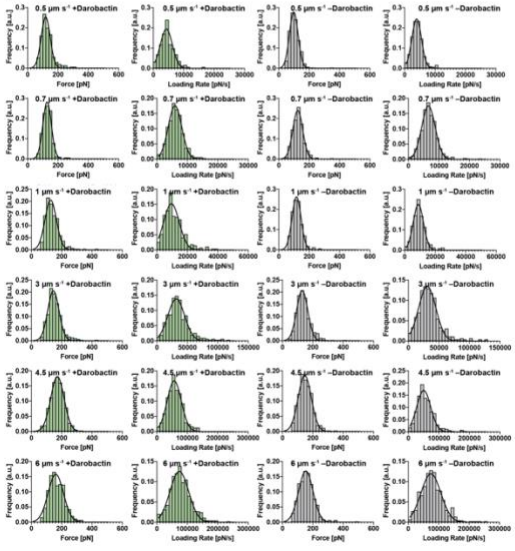
P3



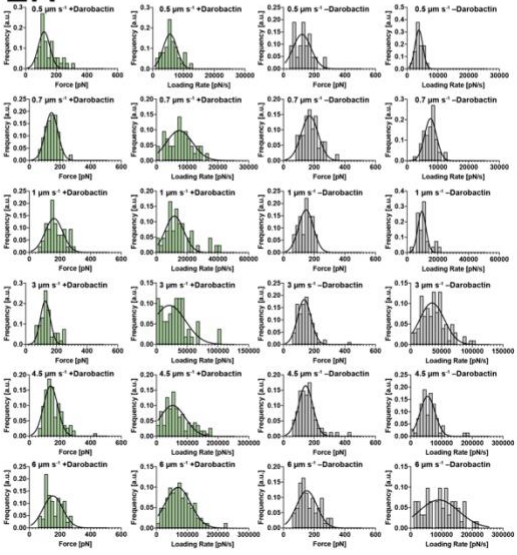
P4



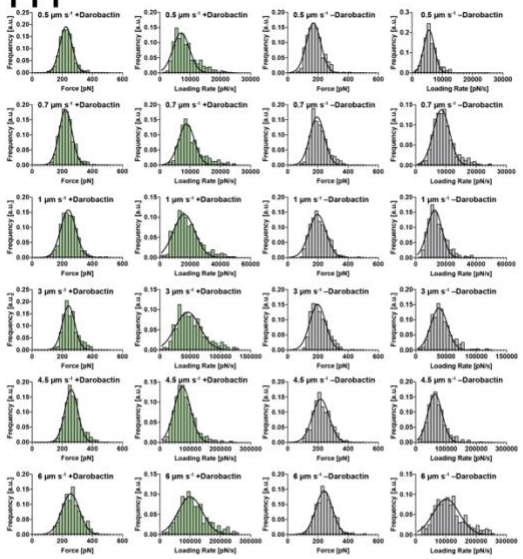
P5



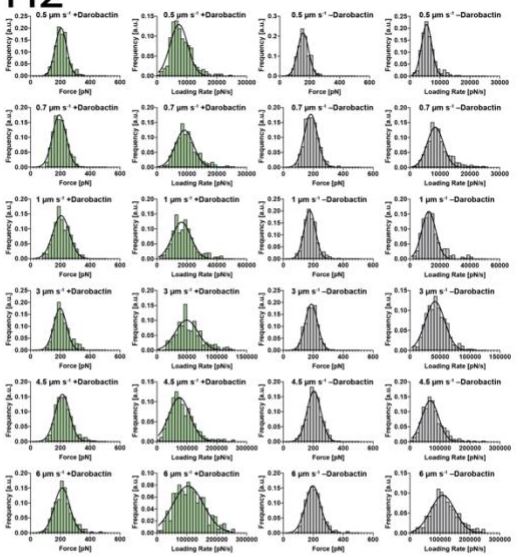
LR



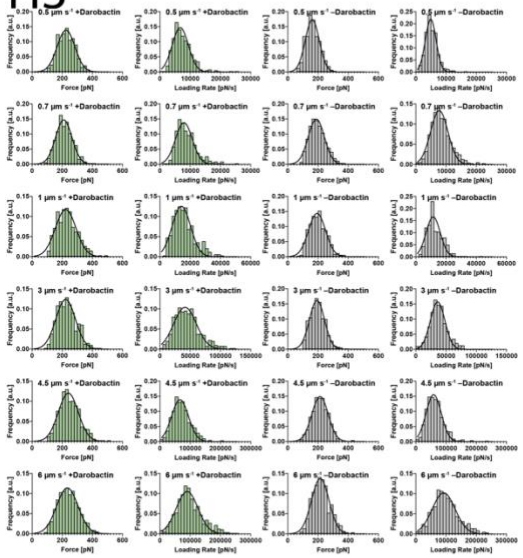
H1



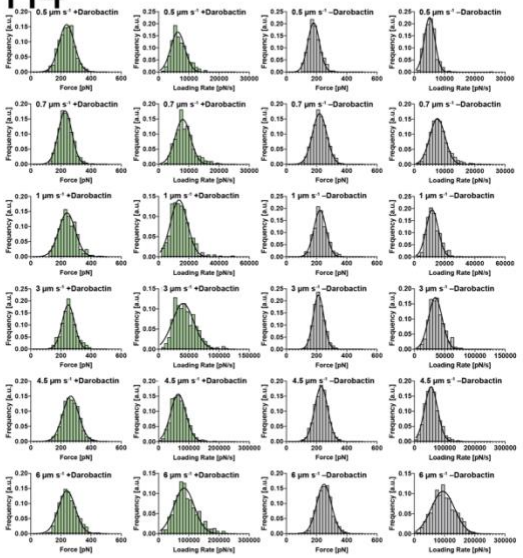
H2



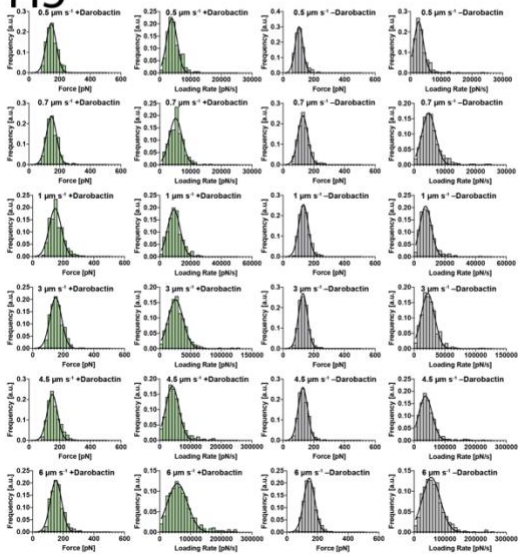
H3



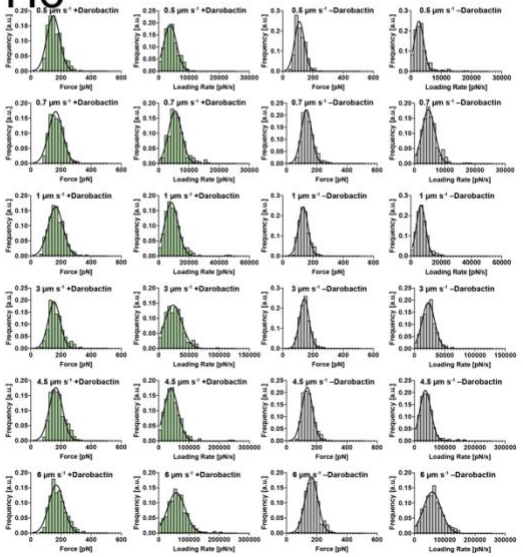
H4



H5



H6



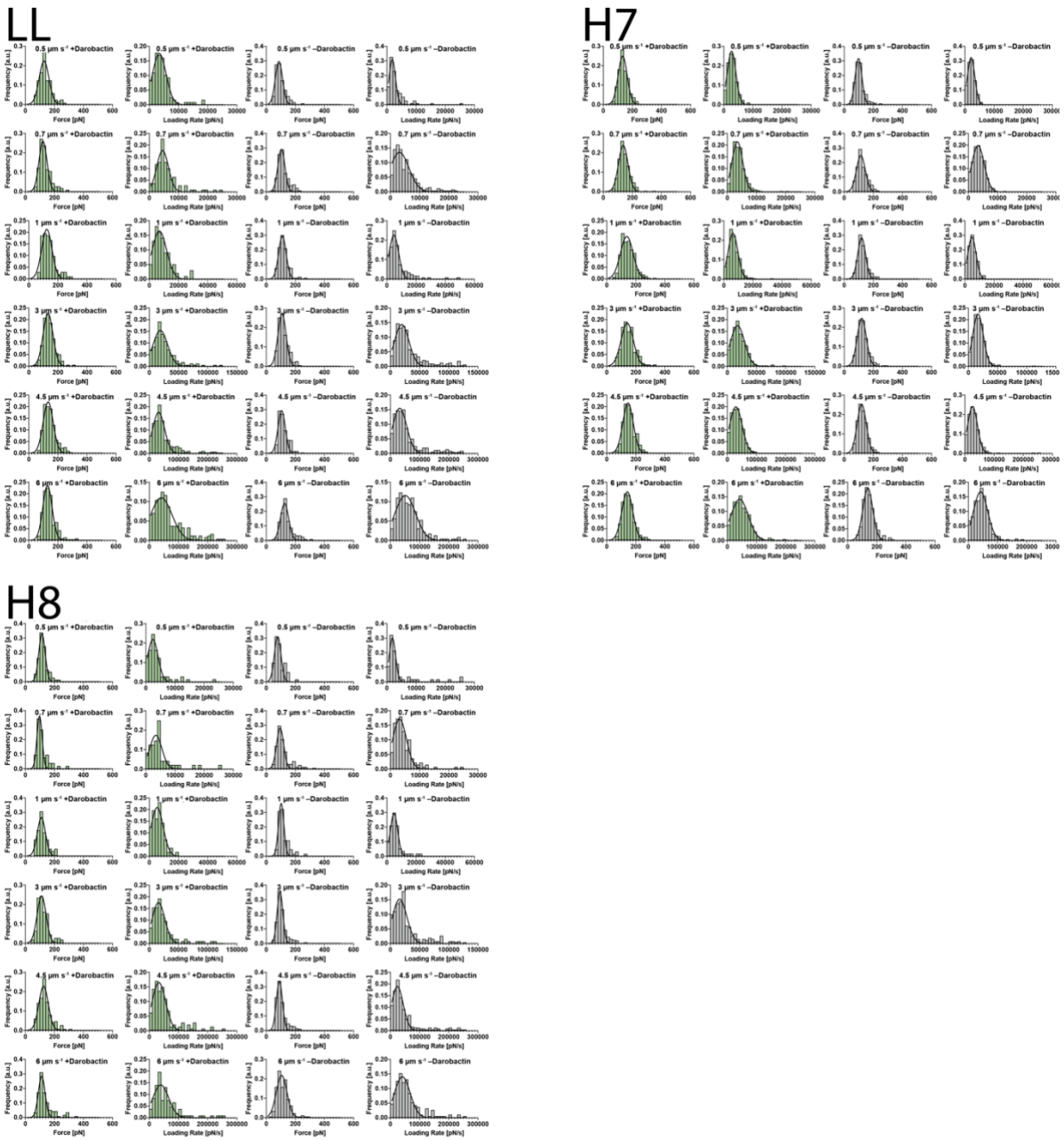


Figure S4. Unfolding Force and Loading Rate Distributions. The plots display the unfolding force and loading rate distributions of all force peak classes recorded upon mechanically unfolding BamA at different pulling speeds (500, 700, 1’000, 3’000, 4’500 and 6’000 nm s^{-1}) in the presence (green) and the absence (grey) of darobactin. The structural segments of BamA unfolded in each force peaks class are indicated. Black lines indicate Gaussian fits to the force and loading rate distributions. The titles of all plots indicate the respective unfolding speed in nm s^{-1}.

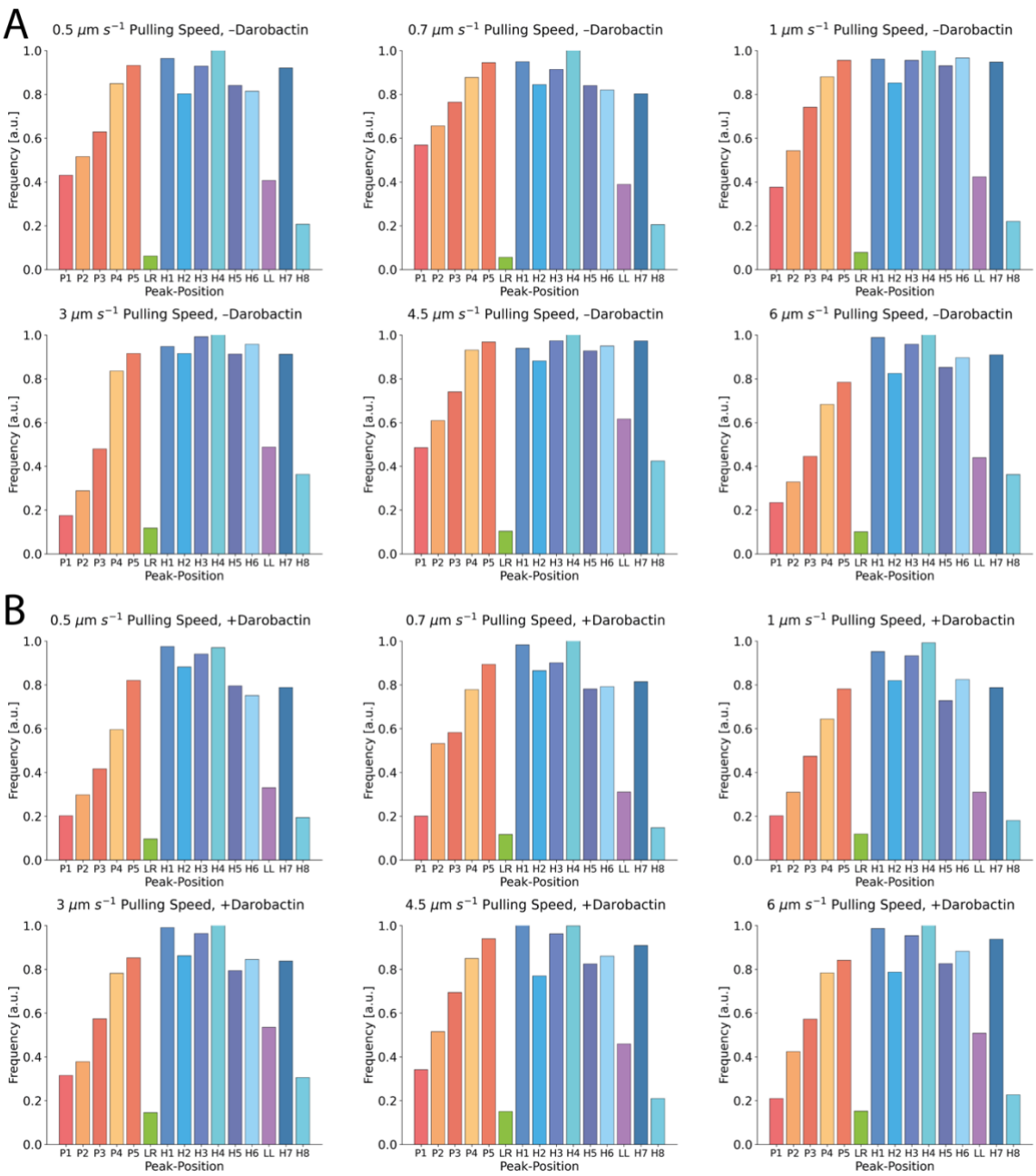


Figure S5. Unfolding Probabilities of Structural Segments of BamA. Each plot displays the unfolding probabilities of the structural segments detected upon mechanically unfolding BamA in the absence (**A**) or the presence (**B**) of darobactin. Histograms are given for different unfolding speeds as indicated.

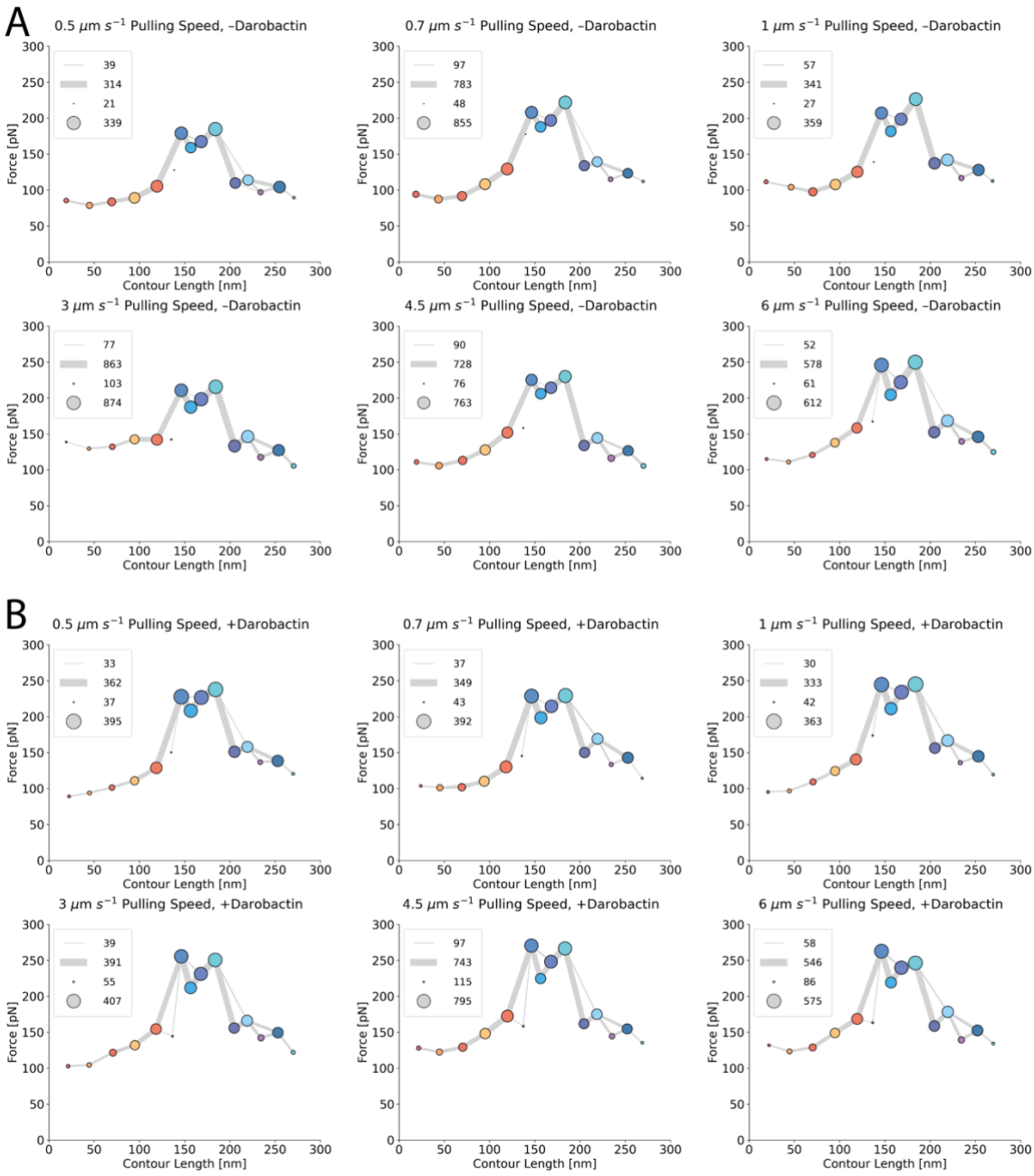


Figure S6. Mechanical Unfolding Pathways of BamA in the Absence and Presence of Darobactin. Each plot shows the mechanical unfolding pathways of BamA in absence (**A**) or presence (**B**) of darobactin. The marker-positions indicate the mean forces and contour lengths of force peak classes and the colors indicate the corresponding structural segment. The marker diameter indicates the relative number of events observed for each unfolding step and the grey lines between markers indicate transitions from one unfolding intermediate to another one. The thickness of the grey lines indicates how many BamA unfolding curves followed the respective transitions. The inset boxes display the marker sizes and line widths represent the minimal and maximal number of events and transitions. In all SMFS experiments BamA has been mechanically unfolded from the N-terminal end.

6 Fusion Domains Guide the Oriented Insertion of Light-Driven Proton Pumps into Liposomes

Noah Ritzmann[1]*, Johannes Thoma[1]*, Stephan Hirschi[2], David Kalbermatter[2], Dimitrios Fotiadis[2], and Daniel J. Müller[1†]

[1] Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland
[2] Institute of Biochemistry and Molecular Medicine, University of Bern, 3012 Bern, Switzerland.

*Contributed equally to this work

† **Correspondence:**
Daniel J. Müller, Dr.

Author Contributions: All authors designed and discussed the experimental approach. N.R. and J.T. cloned, expressed and purified the PR constructs. N.R., J.T., S.H., and D.F. reconstituted the PR constructs. N.R. and S.H. characterized the light-driven proton pumping activity. D.K., S.H. and D.F. recorded electron microscopy images. All authors discussed and wrote the manuscript.

Published in: *Biophysical Journal*, 113, 1-6, September 19, 2017
Doi: 10.1016/j.bpj.2017.06.022

6.1 Abstract

One major objective of synthetic biology is the bottom-up assembly of minimalistic nanocells consisting of lipid or polymer vesicles as architectural scaffolds and of membrane and soluble proteins as functional elements. However, there is no reliable method to orient membrane proteins reconstituted into vesicles. Here, we introduce a simple approach to orient the insertion of the light-driven proton pump proteorhodopsin (PR) into liposomes. To this end, we engineered red or green fluorescent proteins to the N- or C-terminus of PR, respectively. The fluorescent proteins optically identified the PR-constructs and guided the insertion of PR into liposomes with the unoccupied terminal end facing inwards. Using the PR-constructs, we generated proton gradients across the vesicle membrane along predefined directions such as required to power (bio)chemical processes in nanocells. Our approach may be adapted to direct the insertion of other membrane proteins into vesicles.

6.2 Introduction

The assembly of molecular systems is a key element in synthetic biology and aims for the engineering of novel devices with functionalities not found in nature. Hereto, minimalistic nanocells built from lipid or polymer vesicles with integrated membrane proteins experience particular interest since they provide a broad spectrum of potential applications in biotechnology, cell biology and medicine (1,2). Despite varying functionalities of nanocells the underlying engineering principles remain the same. A membrane defines the boundary between interior and exterior and forms an impermeable barrier for hydrophilic molecules. This separation of intra- and extravesicular space creates a nano-environment, which can be utilized for internal reactions (3,4). Regulated transport of molecules to supply and energize these reactions can be achieved by embedding specific transporting and energy converting proteins in the membrane (5-7). Utilizing membrane proteins as functional building blocks for such synthetic systems requires the ability to control their orientation in the membrane. In the living cell the orientation of membrane proteins is determined during insertion and folding, which is assisted by chaperones, insertases and translocases (8,9). This orientation of membrane proteins is not preserved throughout purification and reconstitution into synthetic membranes (10,11), wherein membrane proteins adopt either inward or/and outward facing orientations. However, to our knowledge no suitable and easily applicable method to control this task exists. Here we introduce an approach for the directed reconstitution of membrane proteins into synthetic vesicles.

In analogy to cellular membranes a proton-motive force can provide synthetic systems with energy to power (bio)chemical processes. In this context, light-driven proton pumps are of particular interest (6,7,12). By generating a proton gradient upon illumination such membrane proteins add an element of external control, thus allowing downstream energy-dependent processes such as uptake and release of solutes to be triggered (13). Proteorhodopsin (PR), a light-driven proton pump from proteobacteria, can be employed as an energy-converting module (14,15). PR translocates protons from the C-terminal cytoplasmic towards the N-terminal extracellular space upon illumination (16). Thus, in bacteria expressing PR illumination induces proton outflux and generates a proton gradient across the cellular membrane (17). Here we wanted to engineer PR as energy-supplying module to convert light energy into proton gradients of predefined directionalities across vesicular membranes. Thus far, the production of PR-containing proteoliposomes relies on protocols in which PR is solubilized, purified and reconstituted. Thereby, the two possible orientations of PR cannot be controlled to direct the light-driven proton gradient across vesicular membranes (Fig. 1 *A*). We hence thought to establish a rationale to direct the PR insertion into vesicles.

Insertion of membrane proteins with a preferred directionality due to steric effects has previously been reported for proteins that natively comprise bulky domains (18). It has been also observed that membrane proteins tend to insert with their most hydrophobic domain facing the liposomal lumen (19). To control the proton pumping of PR, we wanted to exploit these effects and expected that the hydrophilic properties of soluble protein domains fused to PR would only allow the unoccupied terminal end to traverse the hydrophobic core of the membrane of preformed liposomes (Fig. 1, *B* and *C*). Based on this consideration we engineered two different PR-constructs to guide their oriented insertion: PR with a C-terminal green fluorescent protein (PR-GFP) and PR with an N-terminal red fluorescent protein (mCherry-PR). The use of green and red fluorescent proteins as soluble fusion domains allowed the optical identification of the constructs. Furthermore, we designed the PR-constructs so that they would not interfere with the

native orientation of PR in the inner membrane of *Escherichia coli*, which we used for overexpression. PR-GFP was designed to carry GFP in the cytoplasm by adding GFP to the C-terminal end of PR connected by a flexible polypeptide linker (Sequence S1 in the Supporting Material). To locate mCherry in the periplasm of *E. coli* mCherry-PR was designed by replacing the native N-terminal signal sequence of PR by a signal sequence of the periplasmic Skp protein, the mCherry sequence, and a short polypeptide linker (Sequence S2).

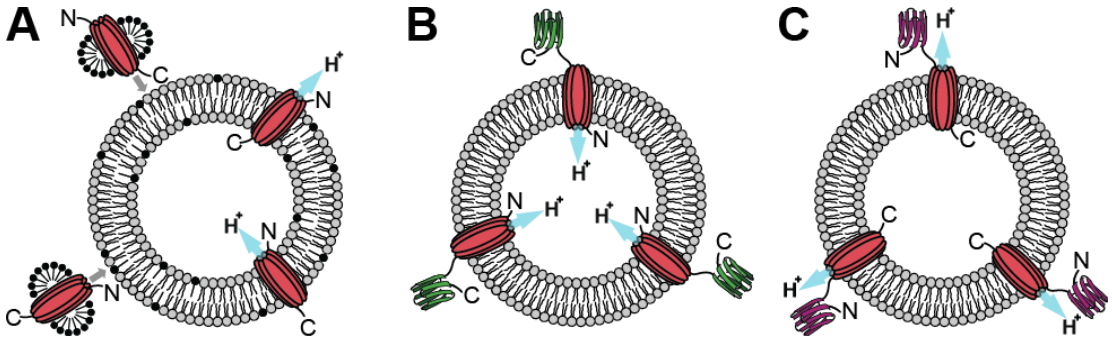


FIGURE 1 Orientations and proton pumping directions of proteorhodopsin (PR) reconstituted into liposomes. (*A*) During detergent-mediated reconstitution solubilized wild-type PR (red) can insert into liposomes in two opposing directions. (*B*) Fusion of the C-terminus of PR to a soluble GFP-domain prevents the terminus to insert. PR thus inserting from the N-terminal side transports protons into the vesicle. (*C*) Fusion of the N-terminus of PR to a soluble mCherry-domain prevents the terminus to insert. PR inserting from the C-terminal side transports protons out of the vesicle.

6.3 Results and Discussion

Variations of both PR-constructs were expressed in *E. coli* and colonies were selected for high-level overexpression based on their red color and fluorescence intensity (Fig. S1). To assess the proton-pumping activity of the PR-constructs we evaluated their activity in *E. coli* using a photoactivity assay (Fig. 2 *A*) (15). Illumination of bacteria overexpressing the PR-constructs caused a pH-drop in the unbuffered solution, thus indicating that both PR-constructs translocated protons from the cytosol to the extracellular solution (Fig. 2 *B*). Control *E. coli*, not expressing PR did not show such an effect. These results showed that PRs from both constructs inserted in the same direction in the bacterial membrane (Fig. 2 *C*).

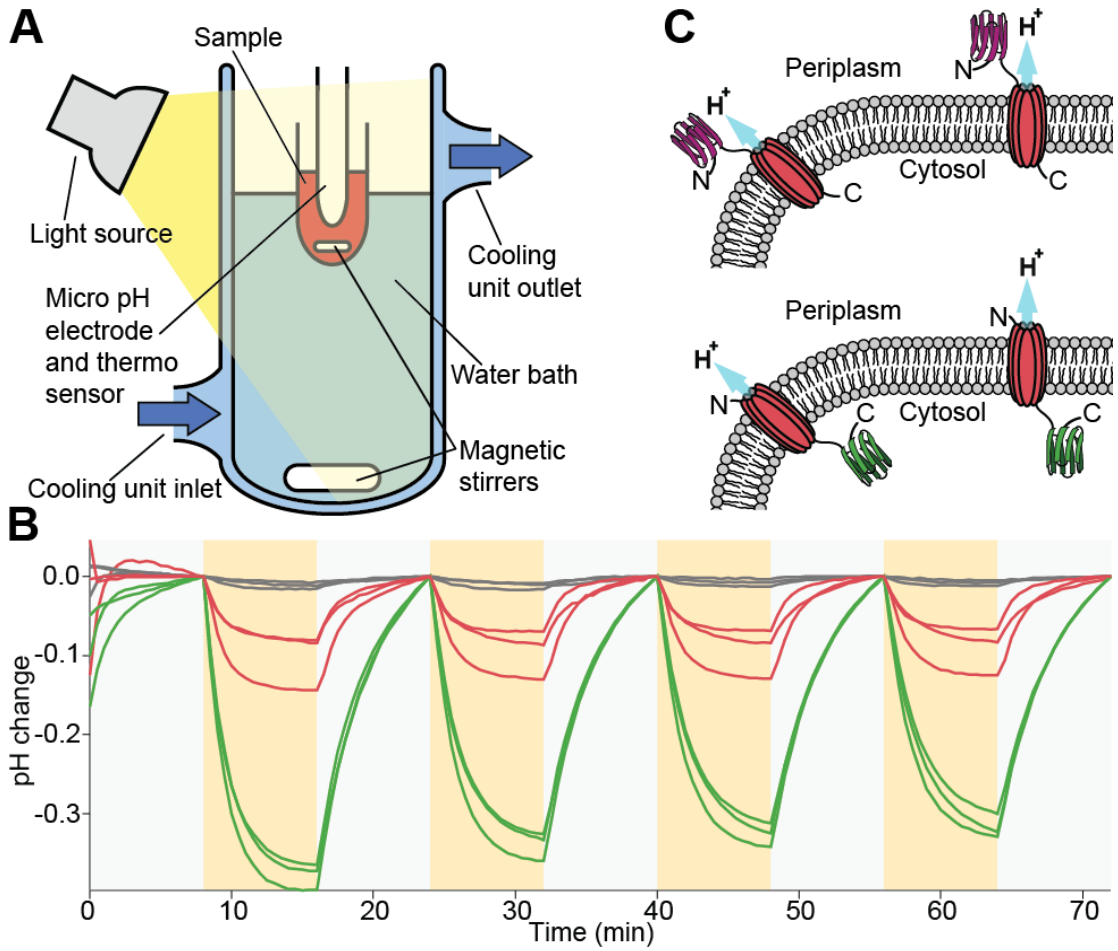


FIGURE 2 Photoactivity measurements of *E. coli*. (*A*) Experimental setup: the pH of the sample is recorded with a micro pH electrode, while it is stirred, temperature controlled and protected from external light. An internal light-source activates proton-pumping. (*B*) Measurements of *E. coli* either overexpressing mCherry-PR (red lines) or PR-GFP (green lines). Grey lines are control measurements performed with *E. coli* containing an empty pET21a(+) plasmid. Yellow areas indicate periods of illumination, grey areas dark periods. (*C*) Orientation of the two PR-constructs in the *E. coli* inner membrane. Both constructs translocate protons to the periplasm.

Once the expression was optimized for both PR-constructs they were purified (Materials and Methods). The yield of purified PR-GFP reached ≈ 2.5 mg protein per g cell pellet (wet weight) whereas that of purified mCherry-PR reached ≈ 0.5 mg per g cell pellet. This variation in expression level explained the differing magnitudes of the pH change detected in the photoactivity measurements of *E. coli* (Fig. 2). The purity of the overexpressed PR-constructs was evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 3 *A*). Wild-type (wt)-PR migrated at ≈ 22 kDa, PR-GFP at ≈ 39 kDa and mCherry-PR at ≈ 41 kDa. Absorption spectra of both purified PR-constructs were recorded (Fig. 3, *B* and *C*) and compared to that of purified wt-PR, GFP, and mCherry. The spectra of both PR-constructs appeared as a convolution of the absorption spectra of the individual proteins from which they were engineered, with coinciding peak positions. The presence of the GFP- and mCherry-absorption peaks of the PR-constructs demonstrated that detergent treatment during purification did not affect the soluble proteins.

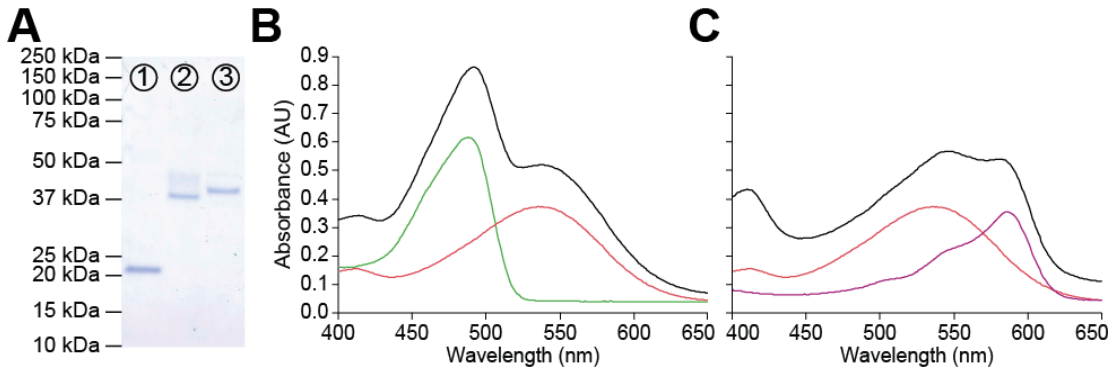


FIGURE 3 Analytical description of PR-constructs. (*A*) 4–12% SDS-PAGE gel of wild-type (wt)-PR in lane 1, PR-GFP in lane 2 and mCherry-PR in lane 3. (*B*) Absorption spectra of wt-PR (red), GFP (green) and PR-GFP (black). (*C*) Absorption spectra of wt-PR (red), mCherry (purple) and mCherry-PR (black).

Next, we reconstituted both PR-constructs into preformed liposomes under identical conditions (Materials and Methods). Cryo-electron microscopy (Cryo-EM) showed mostly unilamellar proteoliposomes having average diameters of ≈ 100 nm (Fig. 4 *A*). To assess the orientation of insertion and the functionality of the reconstituted PR-constructs we evaluated their activity in proteoliposomes using the photoactivity assay (Fig. 4 *B*). The assay monitored the extravesicular change in pH over time, similar to above measurements performed with *E. coli* overexpressing the constructs. Each sample underwent multiple light-dark cycles to periodically activate and inactivate proton-pumping of the reconstituted PR. Upon illumination proteoliposomes containing PR-GFP increased the extravesicular pH, which recovered in the dark phase. In contrast, illumination of proteoliposomes containing mCherry-PR decreased the extravesicular pH, which recovered in the dark phase. Proteoliposomes reconstituted with either of both PR-constructs showed reproducible behaviour over several light-dark cycles. Control measurements performed with proteoliposomes prepared under identical conditions with wt-PR did not show significant pH changes. The photoactivity assay thus demonstrated that proteoliposomes containing the respective PR-constructs translocated protons across the vesicular membrane in opposing directions. To evaluate the ratio of PRs in the predefined orientation, we followed an established proteolytic digestion assay with the nonspecific serine protease proteinase K (Fig. 4 *C*) (20,21). The fluorescent moieties of the PR-constructs are only accessible to proteinase K in the extravesicular space, while the lipid bilayer acts as a diffusion barrier and prevents inwards facing moieties from digestion by proteinase K (22,23). The analysis of the fluorescence before and after digestion on an SDS-PAGE gel revealed that a vast majority of proteins were inserted in the predefined orientation. We can thus conclude that the PR-constructs directed the insertion of PR into liposomes.

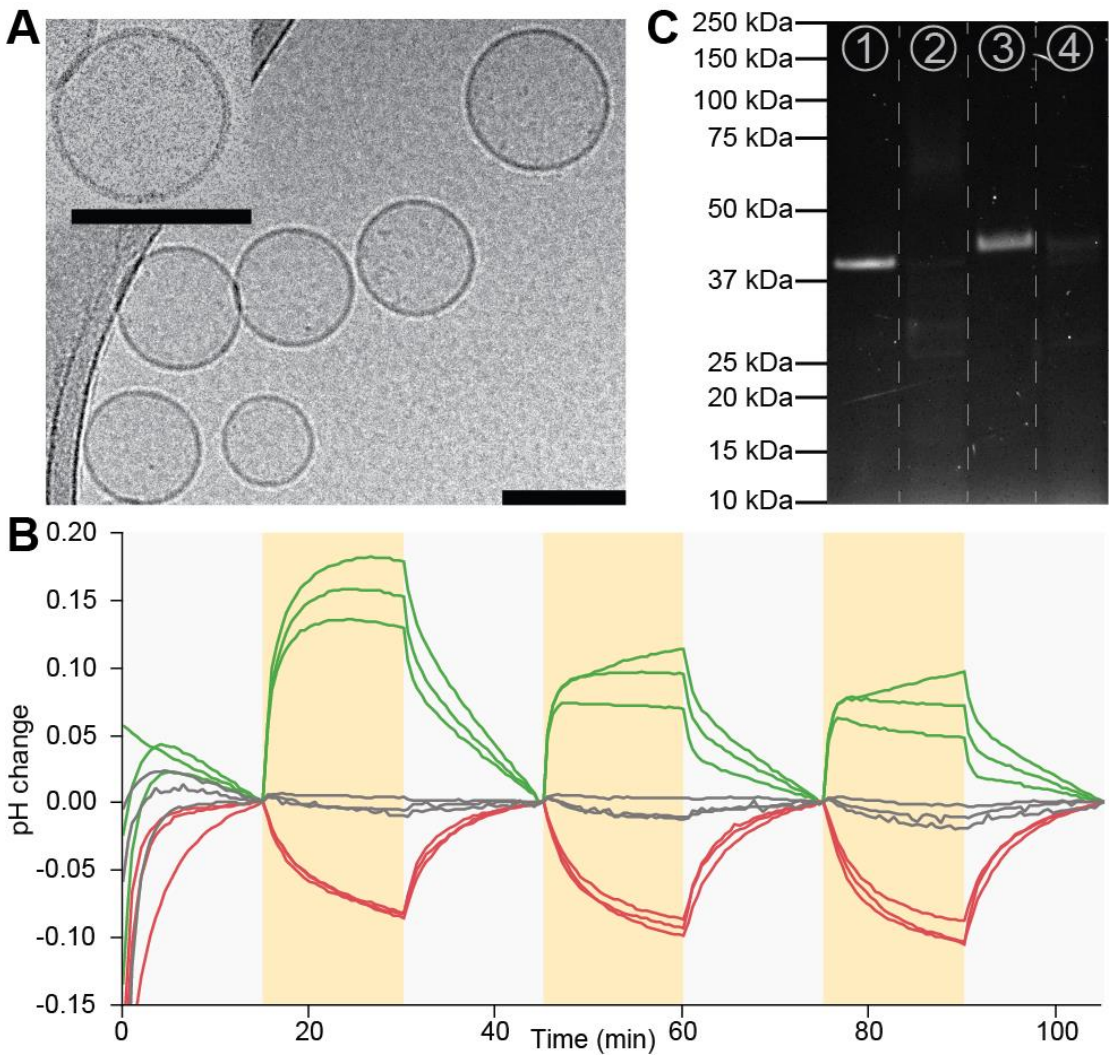


FIGURE 4 Analysis of proteoliposomes. (*A*) Cryo-EM of PR-GFP. Scale bars, 100 nm. (*B*) Photoactivity measurements of proteoliposomes reconstituted with either mCherry-PR or PR-GFP. Yellow areas indicate periods of illumination, grey areas dark periods. Green lines show the proton-translocation activity of proteoliposomes containing PR-GFP and red lines show the activity of proteoliposomes containing mCherry-PR. Grey lines show controls of wt-PR proteoliposomes. Each line represents an individually prepared sample. (*C*) Fluorescence analysis of an SDS-PAGE gel of PR-GFP proteoliposomes before (lane 1) and after proteinase treatment (lane 2) and mCherry-PR proteoliposomes before (lane 3) and after proteinase treatment (lane 4).

6.4 Conclusions

Previous studies controlled the directionality of proton-pumping either by altering the lipid composition of the vesicles (21) or by chemically deactivating one orientation of PR reconstituted into vesicles (15). However, such efforts are limited to the usage of specific lipids or, in the latter example, to one particular pumping direction. Here, we introduced an approach to control the directionality of membrane protein insertion by exploiting the repulsive properties of membranes against hydrophilic proteins. Thereby, the position of a soluble protein fused to a membrane protein determined how it inserted into preformed vesicles. This control of directed membrane protein insertion represents a crucial step towards engineering nanocellular systems of higher complexity. We believe that our approach of fusion-protein controlled directed insertion of membrane proteins can be readily adapted for many other membrane proteins, such as needed to functionally equip nanocells and to drive their (bio-) chemical reactions. En route towards

engineering of multi-modular systems it is conceivable to expand our method to the fusion of two or more membrane proteins for their directed insertion into synthetic vesicles and nanocells. The use of such constructs would not only provide control over the absolute and relative orientation of the membrane proteins but also over their stoichiometry.

Materials and Methods

Cloning. Proteorhodopsin (PR) variants were engineered based on green-light absorbing PR (GenBank: AY601905.1) from plasmid pZUDF-rbs-PR-GGS-3C10H (15). Plasmid pNR03, harboring PR fused with GFP (PR-GFP) was assembled by adding the gene coding for sf-GFP to the C-terminal end of PR, separated by a flexible linker sequence. The resulting sequence (Sequence S1) was subcloned into a pET-21a(+) plasmid between the *NdeI* and *HindIII* restriction sites. Plasmid pNR09, harboring PR with mCherry linked to its N-terminal end (PR-mCherry) was assembled by replacing the coding sequence for PRs native signal sequence with the signal sequence of the periplasmic protein Skp from *E. coli* MG1655 followed by the gene coding for mCherry and a short linker. The resulting sequence (Sequence S2) was subcloned into a pET-21a(+) plasmid between the *NdeI* and *NotI* restriction sites. Both PR constructs contain a His_6-tag from the pET21a(+) backbone on the C-terminal end.

Overexpression of PR-constructs. pNR03 and pNR09 were transformed into *E. coli* Lemo21(DE3) cells. Colonies were selected for high-level expression from small-scale test expressions based on their red color intensity and the fluorescence intensity of GFP or mCherry, respectively. 4 L Luria-Bertani (LB) liquid cultures (100 µg mL^{-1} ampicillin, 36 µg mL^{-1} chloramphenicol) were inoculated 1:100 from overnight cultures. For PR-GFP expression cells were grown under vigorous shaking at 30°C to OD_{600} ≈ 0.5, 5 µM all-*trans*-retinal was added, and expression was induced with addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). For mCherry-PR expression media were supplemented with 150 µM L-rhamnose. Cells were grown under vigorous shaking at 30°C to OD_{600} ≈ 0.5, 5 µM all-*trans*-retinal was added, and expression was induced with addition of 0.4 mM IPTG. Following induction, cells were incubated for 4 h at 30°C. The cells were harvested by centrifugation (5,000 ×g, 12 min, 4°C), resuspended in lysis buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM TCEP) and stored at –20°C until further use.

Photoactivity measurements of *E. coli* overexpressing PR-constructs. Photoactivity measurements were performed as described (15). 100 mL LB liquid cultures containing overexpressing *E. coli* were prepared as described above. Cells were washed twice with 10 mL 150 mM NaCl pH 7.4 followed by centrifugation (3200 ×g, 10 min, 4°C). Immediately before the photoactivity-measurements another washing step was performed and the concentration of the cells was adjusted to OD_{600} = 40. 800 µL of the sample were used to measure the light-driven proton translocation activity of the PR-constructs (Fig. 1 *B*). The activity was monitored by recording the pH in the unbuffered extracellular solution using a micro pH-electrode with integrated temperature sensor (InLab Micro Pro, Mettler Toledo). During the measurement, the sample was stirred and the temperature was kept constant at 18°C using a cooling water bath (setup shown in Fig. 1 *A*). The sample was illuminated by a 2 W, warm white (3,000 K) LED lamp (JANSÖ, IKEA) for 8 min during four consecutive light-dark cycles. After each period of illumination, the sample was kept in the dark for 8 min to recover. To prevent background illumination, the whole setup was guarded from light. The pH and the temperature were recorded in intervals of 30 s. The pH-drift was corrected by subtracting a piecewise linear function from the raw data (Fig. S2). The slope of the function was defined for each peak by two sequential starting points of illumination cycles as described before (15).

Purification of PR-constructs. Cell pellets were thawed, DNAseI from bovine pancreas (Roche Diagnostics), Lysozyme from hen egg white (Fluka Analytical), and cOmplete EDTA-free protease inhibitor (Roche Diagnostics) were added. Cells were lysed by sonication with a Branson digital sonifier with a total pulse time of 25 min (25% amplitude). Unbroken cells were removed by centrifugation (3,200 ×g, 10 min, 4°C). The

membrane fraction was collected by centrifugation (75,000 ×g, 1 h, 4°C), resuspended in lysis buffer, and homogenized. The centrifugation step was repeated and the pellet was resuspended and homogenized in 6 mL membrane storage buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10% (v/v) glycerol). The sample was stored at – 80°C in aliquots of 1 mL until further use. For purification, one aliquot was thawed and solubilized in 10 mL solubilization buffer (20 mM Tris-HCl, pH 7.4, 300 mM NaCl, 10% glycerol, 20 mM imidazole, 1 mM TCEP, 3% (w/v) n-Octyl-β-D-Glucoside (OG, Anatrace), pH 7.4) over night on a roller shaker at 4°C. Unsolubilized material was removed by centrifugation (75000 ×g, 20 min, 4°C). 1 mL Ni-NTA agarose resin (Protino, Macherey-Nagel) was added to the solubilized membrane together with 5 mL binding buffer (20 mM Tris-HCl, pH 7.4, 300 mM NaCl, 10% glycerol, 30 mM imidazole, 1 mM TCEP, 3% (w/v) OG, pH 7.4) and incubated rolling for 3 h at 4°C. The sample was washed twice with 10 mL wash buffer (20 mM Tris-HCl, pH 7.4, 300 mM NaCl, 10% glycerol, 10 mM imidazole, 1 mM TCEP, 1% (w/v) OG, pH 7.4). The protein was eluted by stepwise addition of elution buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10% glycerol, 400 mM imidazole, 1 mM TCEP, 1% (w/v) OG, pH 7.4). Elution fractions with strong red color intensity were pooled. Protein concentrations were determined by measuring their absorption at 280 nm (NanoDrop 2000c, Thermo Scientific). Molecular weight and extinction coefficients were predicted for each protein, based on the amino acid sequence, using the Expasy ProtParam tool (http://www.expasy.ch/tools/protparam.html). For PR-GFP a molecular weight of 58.8 kDa and extinction coefficient of 96510 $M^{-1}cm^{-1}$ were predicted and for mCherry-PR 57.4 kDa and 115,865 $M^{-1}cm^{-1}$. Purified protein was stored at 4°C in the dark.

Reconstitution of PR-constructs into liposomes. 10 mg 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC, Avanti Polar Lipids) in 1 mL chloroform was dried in a 25 mL round bottom Duran glass flask (Lenz) under a nitrogen stream. Residual chloroform was removed by applying vacuum over night at room temperature. The lipid film was hydrated with 2 mL hydration buffer (20 mM KP_i, pH 7.2, 100 mM KCl, 1 mM TCEP) and liposomes were generated by shaking at 700 rpm for 1 h. Liposomes were destabilized by addition of 0.75% (w/v) OG under shaking for 1 h. The liposomes were extruded (LipX Liposome Extruder) through a pore diameter of 200 nm by 19 passes. 0.5 mM $CaCl_2$ was added to the liposomes, followed by shaking for 30 min at 700 rpm. The protein was added to the preformed liposomes and the concentration was adjusted to a final lipid-to-protein ratio (LPR) of 11.5 (w/w), which results in ≈ 80 proteins per proteoliposome, assuming a liposome diameter of 100 nm and a molar LPR of ≈ 1,000. The amount of detergent added with the protein increased the total OG-concentration to 0.8%. The sample was shaken for 3 h at 700 rpm. The sample was transferred to a dialysis tube (14 kDa MWCO, Visking dialysis tubing, Medicell Membranes) and dialyzed against 2 L of dialysis buffer (20 mM Tris-HCl, pH 7.5, 0.5 mM $CaCl_2$) overnight at room temperature.

Cryo-EM. A 3 µl drop of PR-GFP proteoliposomes was deposited on a Lacey carbon grid (Cu 300 mesh, Agar Scientific). The grid was blotted on both sides for 4 s in a Vitrobot (FEI) at 100% humidity and 4°C, and frozen rapidly by plunging into liquid ethane. Images were collected at liquid nitrogen temperature on a Tecnai F20 electron microscope (FEI) operated at 200 kV equipped with a FEI Falcon 3 direct electron detector at a magnification of 50,000x (step size of 2.08 Å/pixel at the specimen level) and at a defocus value of – 5.0 µm (Fig. 4 *A*, overview) and – 2.5 µm (Fig. 4 *A, inset*). The exposure time was 2 s resulting in a total electron dose of ≈ 29 $e^-/Å^2$.

Photoactivity measurements of proteoliposomes. Proteoliposomes were washed twice with 800 µL 150 mM NaCl pH 7.4 followed by centrifugation (200,000 ×g, 20 min,

4°C) as described (15). Another washing step was performed immediately before measurement. The photoactivity of the proteoliposomes was monitored in a total volume of 800 µL. Measurements were performed as described for the bacteria, except for longer illumination-/dark-periods of 15 min each.

Limited Proteolysis. Proteinase K (800 U/mL, New England Biolabs) was used to digest exposed moieties of reconstituted PR-constructs. Limited proteolysis was performed as described (21). Proteinase K was added to proteoliposomes to a final concentration of 2.5 mg/mL. Samples were incubated at 37°C for 2 h. Subsequently, phenylmethanesulfonyl fluoride (PMSF, AppliChem, 0.2 M in ethanol) was added to a final concentration of 10 mM and the samples were cooled on ice for 30 min to inhibit protease-activity. The samples were loaded on a 4–12% SDS-PAGE gel and analyzed with a Gel Doc XR+ Imager and the Image Lab 4.1 software (Bio-Rad) using UV trans illumination and an exposure time of 20 s.

Acknowledgments

The Swiss Nanoscience Institute (SNI, Basel), the Swiss National Science Foundation (SNF; Grant 205320_160199) and the NCCR Molecular Systems Engineering supported this work.

6.5 References

1. Tanner, P., P. Baumann, R. Enea, O. Onaca, C. Palivan, and W. Meier. 2011. Polymeric vesicles: from drug carriers to nanoreactors and artificial organelles. Acc. Chem. Res. 44:1039-1049.
2. Samad, A., Y. Sultana, and M. Aqil. 2007. Liposomal drug delivery systems: an update review. Curr. Drug Del. 4:297-305.
3. Adamala, K. P., D. A. Martin-Alarcon, K. R. Guthrie-Honea, and E. S. Boyden. 2016. Engineering genetic circuit interactions within and between synthetic minimal cells. Nat. Chem.:doi:10.1038/nchem.2644.
4. Nardin, C., J. Widmer, M. Winterhalter, and W. Meier. 2001. Amphiphilic block copolymer nanocontainers as bioreactors. Eur. Phys. J. E 4:403-410.
5. Abramson, J., S. Iwata, and H. R. Kaback. 2004. Lactose permease as a paradigm for membrane transport proteins. Mol. Membr. Biol. 21:227-236.
6. Claassens, N. J., M. Volpers, V. A. dos Santos, J. van der Oost, and W. M. de Vos. 2013. Potential of proton-pumping rhodopsins: engineering photosystems into microorganisms. Trends Biotechnol. 31:633-642.
7. Walter, J. M., D. Greenfield, and J. Liphardt. 2010. Potential of light-harvesting proton pumps for bioenergy applications. Curr. Opin. Biotechnol. 21:265-270.
8. Dalbey, R. E., P. Wang, and A. Kuhn. 2011. Assembly of bacterial inner membrane proteins. Annu. Rev. Biochem. 80:161-187.
9. Heijne, G. v. 1997. Getting greasy: how transmembrane polypeptide segments integrate into the lipid bilayer. Mol. Microbiol. 24:249-253.
10. Rigaud, J. L., M. T. Paternostre, and A. Bluzat. 1988. Mechanisms of membrane protein insertion into liposomes during reconstitution procedures involving the use of detergents. 2. Incorporation of the light-driven proton pump bacteriorhodopsin. Biochemistry 27:2677-2688.
11. Rémigy, H. W., D. Caujolle-Bert, K. Suda, A. Schenk, M. Chami, and A. Engel. 2003. Membrane protein reconstitution and crystallization by controlled dilution. FEBS Lett. 555:160-169.
12. Hirschi, S., M. Stauffer, D. Harder, D. J. Müller, W. Meier, and D. Fotiadis. 2016. Engineering and assembly of protein modules into functional molecular systems. CHIMIA 70:398-401.
13. LaVan, D. A., and J. N. Cha. 2006. Approaches for biological and biomimetic energy conversion. PNAS 103:5251-5255.
14. Bamann, C., E. Bamberg, J. Wachtveitl, and C. Glaubitz. 2014. Proteorhodopsin. Biochim. Biophys. Acta 1837:614-625.
15. Harder, D., S. Hirschi, Z. Ucurum, R. Goers, W. Meier, D. J. Müller, and D. Fotiadis. 2016. Engineering a chemical switch into the light-driven proton pump proteorhodopsin by cysteine mutagenesis and thiol modification. Angew. Chem. Int. Ed. Engl. 55:8846-8849.
16. Friedrich, T., S. Geibel, R. Kalmbach, I. Chizhov, K. Ataka, J. Heberle, M. Engelhard, and E. Bamberg. 2002. Proteorhodopsin is a light-driven proton pump with variable vectoriality. J. Mol. Biol. 321:821-838.
17. Béjà, O., L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, and E. F. DeLong. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science 289:1902-1906.

18. Romsicki, Y., and F. J. Sharom. 2001. Phospholipid flippase activity of the reconstituted P-glycoprotein multidrug transporter. Biochemistry 40:6937-6947.
19. Rigaud, J.-L., B. Pitard, and D. Levy. 1995. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. Biochim. Biophys. Acta. 1231:223-246.
20. Kalmbach, R., I. Chizhov, M. C. Schumacher, T. Friedrich, E. Bamberg, and M. Engelhard. 2007. Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. J. Mol. Biol. 371:639-648.
21. Tunuguntla, R., M. Bangar, K. Kim, P. Stroeve, C. M. Ajo-Franklin, and A. Noy. 2013. Lipid bilayer composition can influence the orientation of proteorhodopsin in artificial membranes. Biophys. J. 105:1388-1396.
22. Nomura, S. M., K. Tsumoto, T. Hamada, K. Akiyoshi, Y. Nakatani, and K. Yoshikawa. 2003. Gene expression within cell-sized lipid vesicles. Chembiochem 4:1172-1175.
23. Yoshimoto, M., S. Wang, K. Fukunaga, M. Treyer, P. Walde, R. Kuboi, and K. Nakao. 2004. Enhancement of apparent substrate selectivity of proteinase K encapsulated in liposomes through a cholate-induced alteration of the bilayer permeability. Biotechnol. Bioeng. 85:222-233.

6.6 Supplemental Information

Supporting sequences of proteorhodopsin (PR)-constructs

Sequence S1

PR-GFP DNA Sequence

```
   1 CATATGGGTA AATTATTACT GATATTAGGT AGTGTTATTG CACTTCCTAC ATTTGCTGCA
  61 GGTGGTGGTG ACCTTGATGC TAGTGATTAC ACTGGTGTTT CTTTTTGGTT AGTTACTGCT
 121 GCTTTATTAG CATCTACTGT ATTTTTCTTT GTTGAAAGAG ATAGAGTTTC TGCAAAATGG
 181 AAAACATCAT TAACTGTATC TGGTCTTGTT ACTGGTATTG CTTTCTGGCA TTACATGTAC
 241 ATGAGAGGGG TATGGATTGA AACTGGTGAT TCGCCAACTG TATTTAGATA CATTGATTGG
 301 TTACTAACAG TTCCTCTATT AATATGTGAA TTCTACTTAA TTCTTGCTGC TGCAACTAAT
 361 GTTGCTGGAT CATTATTTAA GAAATTACTA GTTGGTTCTC TTGTTATGCT TGTGTTTGGT
 421 TACATGGGTG AAGCAGGAAT CATGGCTGCA TGGCCTGCAT TCATTATTGG GTGTTTAGCT
 481 TGGGTATACA TGATTTATGA ATTATGGGCT GGAGAAGGAA AATCTGCATG TAATACTGCA
 541 AGTCCTGCTG TGCAATCAGC TTACAACACA ATGATGTATA TTATCATCTT TGGTTGGGCG
 601 ATTTATCCTG TAGGTTATTT CACAGGTTAC CTGATGGGTG ACGGTGGATC AGCTCTTAAC
 661 TTAAACCTTA TCTATAACCT TGCTGACTTT GTTAACAAGA TTCTATTTGG TTTAATTATA
 721 TGGAATGTTG CTGTTAAAGA ATCTTCTAAT GCTCTCGAGG GAGGAAGTCT GGAAGTTCTG
 781 TTCCAGGGGC CCGTCGACGG CGGCTCCGGA TCCGAAAACT TGTATTTCCA GGGCATGAGT
 841 AAAGGAGAAG AACTTTTCAC TGGAGTTGTC CCAATTCTTG TTGAATTAGA TGGTGATGTT
 901 AATGGGCACA AATTTTCTGT CCGTGGAGAG GGTGAAGGTG ATGCTACAAA CGGAAAACTC
 961 ACCCTTAAAT TTATTTGCAC TACTGGAAAA CTACCTGTTC CGTGGCCAAC ACTTGTCACT
1021 ACTCTGACCT ATGGTGTTCA ATGCTTTTCC CGTTATCCGG ATCACATGAA ACGGCATGAC
1081 TTTTTCAAGA GTGCCATGCC CGAAGGTTAT GTACAGGAAC GCACTATATC TTTCAAAGAT
1141 GACGGGACCT ACAAGACGCG TGCTGAAGTC AAGTTTGAAG GTGATACCCT TGTTAATCGT
1201 ATCGAGTTAA AGGGTATTGA TTTTAAAGAA GATGGAAACA TTCTTGGACA CAAACTCGAG
1261 TACAACTTTA ACTCACACAA TGTATACATC ACGGCAGACA AACAAAAGAA TGGAATCAAA
1321 GCTAACTTCA AAATTCGCCA CAACGTTGAA GATGGTTCCG TTCAACTAGC AGACCATTAT
1381 CAACAAAATA CTCCAATTGG CGATGGCCCT GTCCTTTTAC CAGACAACCA TTACCTGTCG
1441 ACACAATCTG TCCTTTCGAA AGATCCCAAC GAAAAGCGTG ACCACATGGT CCTTCTTGAG
1501 TTTGTAACTG CTGCTGGGAT TACACATGGC ATGGATGAGC TCTACAAAGG AGGATCTGGT
1561 GGTTCTGGGA AGCTT
```

PR-GFP Amino-Acid Sequence

```
MGKLLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFFVERDRVSAKWKTSLTVSGLVT
GIAFWHYMYMRGVWIETGDSPTVFRYIDWLLTVPLLICEFYLILAAATNVAGSLFKKLLVGSLVMLVFGY
MGEAGIMAAWPAFIIGCLAWVYMIYELWAGEGKSACNTASPAVQSAYNTMMYIIIFGWAIYPVGYFTGYL
MGDGGSALNLNLIYNLADFVNKILFGLIIWNVAVKESSNALEGGSLEVLFQGPVDGGSGSENLYFQGMSK
GEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSR
YPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEY
NFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNE
KRDHMVLLEFVTAAGITHGMDELYKGGSGGSGKLAAALEHHHHHH*
```

- Native PR Signal Sequence
- Proteorhodopsin
- Sf-GFP
- His_6-Tag

Sequence S2

mCherry-PR DNA Sequence

```
   1 CATATGAAAA AGTGGTTATT AGCTGCAGGT CTCGGTTTAG CACTGGCAAC TTCTGCTCAG
  61 GCGGGTGGTG GTGGTGGTGC TAGCGGAGGC TCTGGAGGCT CTGGAATGCA TAGCAAGGGC
 121 GAGGAGGATA ACATGGCCAT CATCAAGGAG TTCATGCGCT TCAAGGTGCA CATGGAGGGC
 181 TCCGTGAACG GCCACGAGTT CGAGATCGAG GGCGAGGGCG AGGGCCGCCC CTACGAGGGC
 241 ACCCAGACCG CCAAGCTGAA GGTGACCAAG GGTGGCCCCC TGCCCTTCGC CTGGGACATC
 301 CTGTCCCCTC AGTTCATGTA CGGCTCCAAG GCCTACGTGA AGCACCCCGC CGACATCCCC
 361 GACTACTTGA AGCTGTCCTT CCCCGAGGGC TTCAAGTGGG AGCGCGTGAT GAACTTCGAG
 421 GACGGCGGCG TGGTGACCGT GACCCAGGAC TCCTCCTTGC AGGACGGCGA GTTCATCTAC
 481 AAGGTGAAGC TGCGCGGCAC CAACTTCCCC TCCGACGGCC CCGTAATGCA GAAGAAGACC
 541 ATGGGCTGGG AGGCCTCCTC CGAGCGGATG TACCCCGAGG ACGGCGCCCT GAAGGGCGAG
 601 ATCAAGCAGA GGCTGAAGCT GAAGGACGGC GGCCACTACG ACGCTGAGGT CAAGACCACC
 661 TACAAGGCCA AGAAGCCCGT GCAGCTGCCC GGCGCCTACA ACGTCAACAT CAAGTTGGAC
 721 ATCACCTCCC ACAACGAGGA CTACACCATC GTGGAACAGT ACGAACGCGC CGAGGGCCGC
 781 CACTCCACCG GCGGCATGGA CGAGCTGTAC AAGAAGCTTG GCGGCTCCGG CTCCGCAGGT
 841 GGTGGTGACC TTGATGCTAG TGATTACACT GGTGTTTCTT TTTGGTTAGT TACTGCTGCT
 901 TTATTAGCAT CTACTGTATT TTTCTTTGTT GAAAGAGATA GAGTTTCTGC AAAATGGAAA
 961 ACATCATTAA CTGTATCTGG TCTTGTTACT GGTATTGCTT TCTGGCATTA CATGTACATG
1021 AGAGGGGTAT GGATTGAAAC TGGTGATTCG CCAACTGTAT TTAGATACAT TGATTGGTTA
1081 CTAACAGTTC CTCTATTAAT ATGTGAATTC TACTTAATTC TTGCTGCTGC AACTAATGTT
1141 GCTGGATCAT TATTTAAGAA ATTACTAGTT GGTTCTCTTG TTATGCTTGT GTTTGGTTAC
1201 ATGGGTGAAG CAGGAATCAT GGCTGCATGG CCTGCATTCA TTATTGGGTG TTTAGCTTGG
1261 GTATACATGA TTTATGAATT ATGGGCTGGA GAAGGAAAAT CTGCATGTAA TACTGCAAGT
1321 CCTGCTGTGC AATCAGCTTA CAACACAATG ATGTATATTA TCATCTTTGG TTGGGCGATT
1381 TATCCTGTAG GTTATTTCAC AGGTTACCTG ATGGGTGACG GTGGATCAGC TCTTAACTTA
1441 AACCTTATCT ATAACCTTGC TGACTTTGTT AACAAGATTC TATTTGGTTT AATTATATGG
1501 AATGTTGCTG TTAAAGAATC TTCTAATGCT CTCGAGGGAG GAATAGCGGC CGC
```

mCherry-PR Amino-Acid Sequence

```
MKKWLLAAGLGLALATSAQAGGGGGASGGSGGSGMHSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEG
EGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFED
GGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGG
HYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKLGGSGSAGG
GDLDASDYTGVSFWLVTAALLASTVFFFVERDRVSAKWKTSLTVSGLVTGIAFWHYMYMRGVWIETGDSP
TVFRYIDWLLTVPLLICEFYLILAAATNVAGSLFKKLLVGSLVMLVFGYMGEAGIMAAWPAFIIGCLAWV
YMIYELWAGEGKSACNTASPAVQSAYNTMMYIIIFGWAIYPVGYFTGYLMGDGGSALNLNLIYNLADFVN
KILFGLIIWNVAVKESSNALEGGIAAALEHHHHHH*
```

- Skp Signal Sequence
- mCherry
- Proteorhodopsin
- His_6-Tag

Supporting Figures

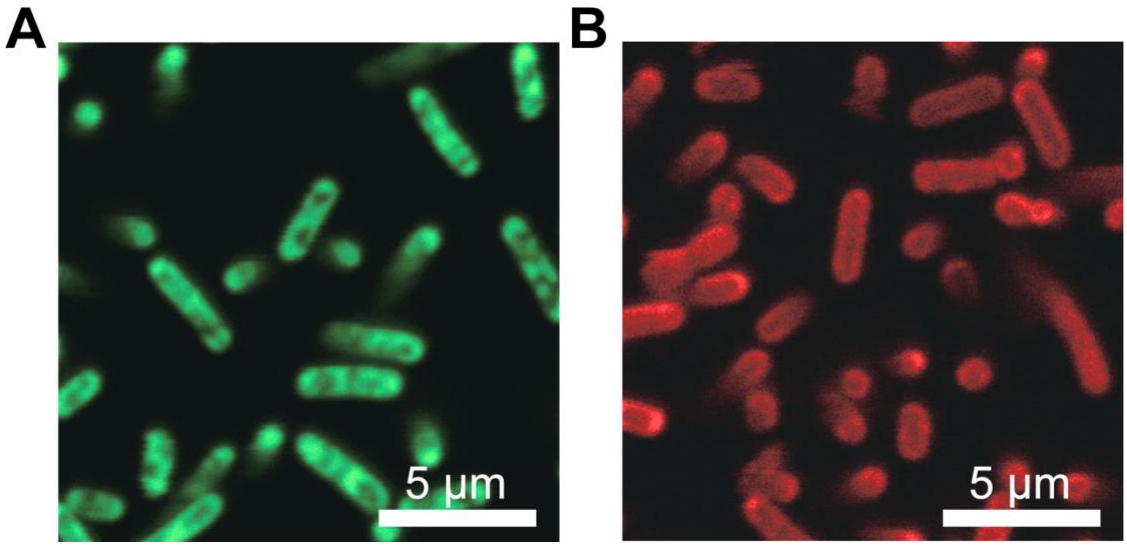


FIGURE S1 Images of *E. coli* overexpressing the PR-constructs. (*A*) *E. coli* overexpressing PR-GFP. (*B*) *E. coli* overexpressing mCherry-PR. Fluorescence images were recorded with an inverted confocal microscope (LSM 700, Zeiss AG) using an α Plan-FLUAR 100x objective.

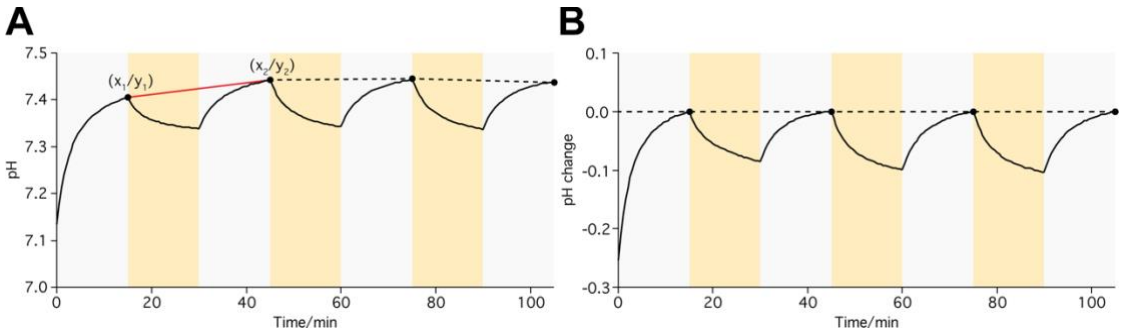


FIGURE S2 pH drift correction of the photoactivity measurements. (*A*) Raw data (black line) of the pH change induced by the PR-constructs was recorded with the pH-meter. A linear function (red line), defined by two sequential starting points of illumination, is calculated and subtracted from the measured data. (*B*) Data after drift correction using the piecewise linear function determined in (*A*).

7 Proton gradients from light-harvesting E. coli control DNA assemblies for synthetic cells

Kevin Jahnke[1,2], Noah Ritzmann[3], Julius Fichtler[1,2], Anna Nitschke[1,2], Yannik Dreher[1,2], Tobias Abele[1,2], Götz Hofhaus[4], Ilia Platzman[5,6], Rasmus Schröder[4], Daniel J. Müller[3], Joachim P. Spatz[5,6,7], and Kerstin Göpfrich[1,2†]

[1]Biophysical Engineering Group, Max Planck Institute for Medical Research, Heidelberg, Germany.
[2]Department of Physics and Astronomy, Heidelberg University, Heidelberg, Germany.
[3]Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH) Zurich, Basel, Switzerland.
[4]Centre for Advanced Materials, Heidelberg, Germany.
[5]Max Planck Institute for Medical Research, Department of Cellular Biophysics, Heidelberg, Germany.
[6]Institute for Molecular Systems Engineering (IMSE), Heidelberg University, Heidelberg, Germany.
[7]Max Planck School Matter to Life, Heidelberg, Germany.

[†]Correspondence:
Kerstin Göpfrich

Author contributions
K.J. performed most experiments and analysis. N.R. designed and prepared genetically engineered E. coli and performed pH electrode measurements. J.F. and K.G. designed the DNA origami. J.F. and K.J. carried out pH-sensitive deformation experiments. A.N. and K.J. established the use of pyranine as pH-sensor within droplets or GUVs. Y.D. and K.J. performed and analyzed pH-sensitive DNA attachment to GUVs. T.A. helped in analyzing the fluorescence intensity ratios. G.H. and R.R.S. designed and carried out cryo-EM experiments. K.J., K.G., I.P., D.J.M. and J.P.S. designed the study. K.J. and K.G. wrote the manuscript with help from all authors.

Published in: *Nature Communications*, 12, 1-9, June 25, 2021
Doi: 10.1038/s41467-021-24103-x

7.1 Abstract

Bottom-up and top-down approaches to synthetic biology each employ distinct methodologies with the common aim to harness living systems. Here, we realize a strategic merger of both approaches to convert light into proton gradients for the actuation of synthetic cellular systems. We genetically engineer *E. coli* to overexpress the light-driven inward-directed proton pump xenorhodopsin and encapsulate them in artificial cell-sized compartments. Exposing the compartments to light-dark cycles, we reversibly switch the pH by almost one pH unit and employ these pH gradients to trigger the attachment of DNA structures to the compartment periphery. For this purpose, a DNA triplex motif serves as a nanomechanical switch responding to the pH-trigger of the *E. coli*. When DNA origami plates are modified with the pH-sensitive triplex motif, the proton-pumping *E. coli* can trigger their attachment to giant unilamellar lipid vesicles (GUVs) upon illumination. A DNA cortex is formed upon DNA origami polymerization, which sculpts and deforms the GUVs. We foresee that the combination of bottom-up and top down approaches is an efficient way to engineer synthetic cells.

7.2 Introduction

Synthetic biology cultivates an engineering approach to biology with the aim to create or to re-purpose biological parts for specific tasks. The field is commonly divided into two branches with distinct tools and methodologies, but also distinct challenges—top-down and bottom-up synthetic biology[1,2]. The top-down approach uses genetic engineering techniques to manipulate natural cells, reprogramming their behaviour and equipping them with unique and exciting functions[3]. *Escherichia coli* (*E. coli*) bacteria, for instance, have been engineered for a variety of tasks, including biofuel production[4], cancer cell targeting[5] or light harvesting[6,7]. Yet living cells remain too complex to achieve full control and not all added functions are compatible with the host[8].

The bottom-up approach, on the other hand, has been successful at reconstituting natural biomolecules, or artificial components in cell-sized confinement like microfluidic droplets or lipid vesicles[9–11]. Noteworthy modules have been implemented so far, each mimicking a specific function of a living cell, including energy generation[12,13], metabolism[14], motility[15,16], cytoskeletal contraction[17] or division[18]. Yet the combination of these modules towards complex signalling pathways for dynamic systems remains challenging. Merging the capacities of top-down and bottom-up approaches to synthetic biology can be a leap forward towards complex bottom-up assemblies but also more versatile and well-defined top-down systems. Leading to this direction, communication between natural and synthetic cells has been implemented[19–21] and bottom-up assembled vesicles were used as organelle mimics in living cells[22]. Furthermore, engineered prokaryotes have recently been used as artificial organelles in living cells[23,24], yet this has never been translated into synthetic cells.

Here, we use top-down genetic engineering to equip *E. coli* with light-harvesting capabilities. We employ them as synthetic organelle mimics inside bottom-up assembled synthetic cellular compartments. Thereby, we can reversibly switch the pH upon illumination to trigger an optical or a mechanical response. The latter is based on the pH-sensitive membrane attachment of a triplex-forming DNA motif triggered by proton gradients from light-harvesting *E. coli*. Furthermore, we employ the pH-gradients to sculpt synthetic cellular compartments by attaching a DNA origami plate to the pH-sensitive DNA strand.

7.3 Results

Top-down engineering of *E. coli* for light-harvesting. To equip synthetic cells with the capability to generate proton gradients, we set out to assemble an energy module. We genetically engineered *E. coli* to overexpress the light-driven proton pump xenorhodopsin, a transmembrane protein from nanohalosarchaeon *Nanosalina*[25]. It contains a retinal which, upon illumination, undergoes a trans-cis conformational change and shuttles a proton across the lipid membrane. We chose xenorhodopsin because it shows unique features compared to other proton pumps, such as bacteriorhodopsin or proteorhodopsin: First of all, xenorhodopsin exhibits a substantially faster photocycle, which can result in larger proton gradients[25]. Second, as an inward-directed pump[26], xenorhodopsin increases the pH (instead of decreasing it) in the extracellular space upon illumination (Fig. 1a). As an additional feature, we introduced a C-terminal fluorescent GFP or mCherry tag to xenorhodopsin for visualization of the *E. coli*. The choice of two dyes allows us to work with different combinations of fluorophores as required. To assess and quantify the proton pumping capabilities of the genetically engineered *E. coli*, we performed a photoactivity assay, where we inserted a micro pH electrode into the *E. coli* suspension and exposed it to multiple light-dark cycles. Since the absorption

spectrum of xenorhodopsin covers a broad range, the use of a white light lamp is more effective than excitation witha specific wavelength (Supplementary Fig. 1). Illumination increased the pH in the extracellular space by almost one pH unit within five minutes (Fig. 1b), because protons are translocated from the extracellular solution to the cytosol. Longer illumination times resulted in saturation of the pH change (Supplementary Fig. 2). In the range from OD_{600} = 8 to OD_{600} = 40, the *E. coli* concentration did not significantly change the obtained pH gradients and we observed only a very minor increase in the kinetics at higher concentrations (Supplementary Fig. 3). The pH quickly returned to its initial value after the light was turned off due to the dissipation of protons. Even after three complete light-dark cycles, we observed only little decrease in the pH gradient. Compared to previous reports where proton pumps were reconstituted in lipid vesicles[7,27], we could achieve faster and higher pH gradients using genetically engineered *E. coli*. Moreover, the use of *E. coli* circumvented the need for cumbersome protein purification and reconstitution to prepare proteoliposomes[28], which highlights a key advantage of merging top-down and bottom-up synthetic biology.

As a next step, we aimed to encapsulate the *E. coli* as a pH switch in synthetic cells, which makes pH monitoring with an electrode impractical. We thus supplement the *E. coli* suspension with the ratiometric pH-sensitive fluorescent dye pyranine. The fluorescence properties of pyranine depend on its protonation state (Fig. 1c, Supplementary Fig. 4). After suitable calibration measurements (Supplementary Fig. 5), we could hence monitor the pH optically[29]. Figure 1d plots the fluorescence intensity ratio over time while the system was exposed to light-dark cycles (Supplementary Movie 1, Supplementary Fig. 6). Notably, we obtained the same results as previously with the pH electrode.

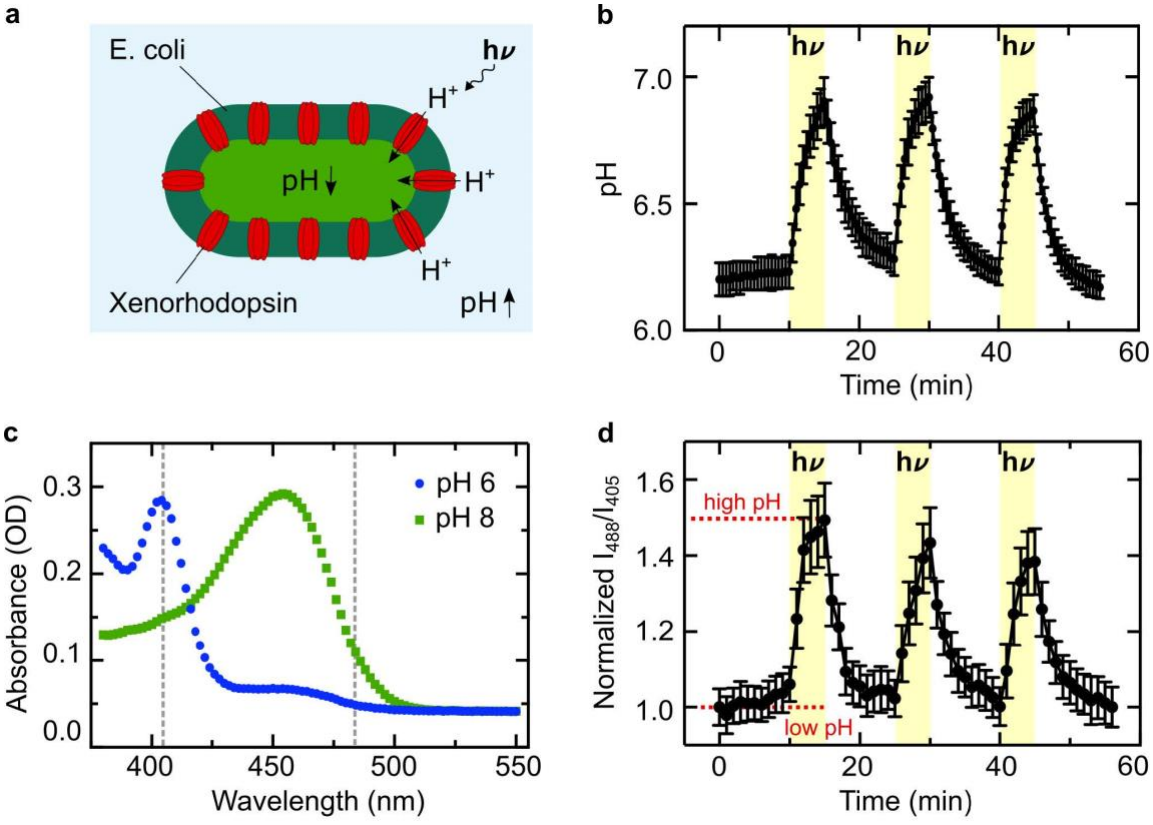


Fig. 1 Genetically engineered xenorhodopsin-expressing *E. coli* generate a pH gradient upon illumination with white light. a Schematic illustration of an *E. coli* expressing xenorhodopsin, a light-driven proton pump

(red), allowing for the reversible generation of a directional pH gradient during illumination with white light. The inward pump increases the pH of the external solution. **b** Photoactivity generated by the *E. coli* (OD_{600} =20, in 150 mM NaCl) measured with an external pH electrode. The pH is plotted over time during three light-dark cycles (periods of illumination are indicated in yellow). The pH increases by almost one pH unit within 5 min of illumination and nearly returns to its original value after 10 min in the dark (mean ± s.d., n = 3). **c** Absorbance measurements of the pH-sensitive ratiometric fluorophore pyranine at pH 6 (blue) and pH 8 (green). The pH can be quantified as the fluorescence intensity ratio at the excitation wavelengths 488 nm and 405 nm (gray dashed lines). **d** Normalized fluorescence intensity ratio I_{488}/I_{405} of pyranine (50 µM) over time in a solution containing *E. coli* and lipid vesicles as determined with confocal fluorescence microscopy (mean ± s.d., n = 4). Periods of illumination are indicated in yellow. Source data is available for Fig. 1b–d.

Light-harvesting *E. coli* as internal pH actuators. Having demonstrated light-activated pH switching in bulk, we wanted to integrate the engineered *E. coli* as artificial mitochondria mimics in synthetic cell-sized confinements. Using a microfluidic droplet formation device (Fig. 2a), *E. coli* and pyranine were encapsulatedin surfactant-stabilized water-in-oil droplets (Fig. 2b; Supplementary Fig. 7). We obtained *E. coli*-containing compartments with a radius of 27 ± 5 µm (mean ± s.d., n = 53, Fig. 2c). Pyranine served as a fluorescent pH indicator inside the compartments (Fig. 2d; Supplementary Fig. 8). We exposed the system to three consecutive light-dark cycles. Illumination with white light triggered a pH increase inside the cell-sized compartments due to thelight-driven proton transport by the *E. coli*, resulting in an opticalresponse of the compartments themselves (Fig. 2e; Supplementary Movie 2). Taken together, we demonstrated that the genetically engineered *E. coli* can provide light-activated proton gradients in cell-sized compartments.

pH-sensitive attachment of DNA to the compartment periphery. Proton gradients in synthetic systems are especially exciting if they can be utilized to control and energize downstream processes.Instead of relying on purified proteins, an increasingly popular approach is to construct such pH-dependent machineries *de novo*from molecular building blocks. DNA nanotechnology, in particular, has been employed to build a variety of functional components for synthetic cells[17,30,31], including membrane-sculpting[32–35] and pH-responsive components such as filaments[36] or rotors[37,38].However, pH-responsive actuation is challenging after encapsulation into a compartment. With the *E. coli*, we can circumvent this by converting light into a proton gradient.

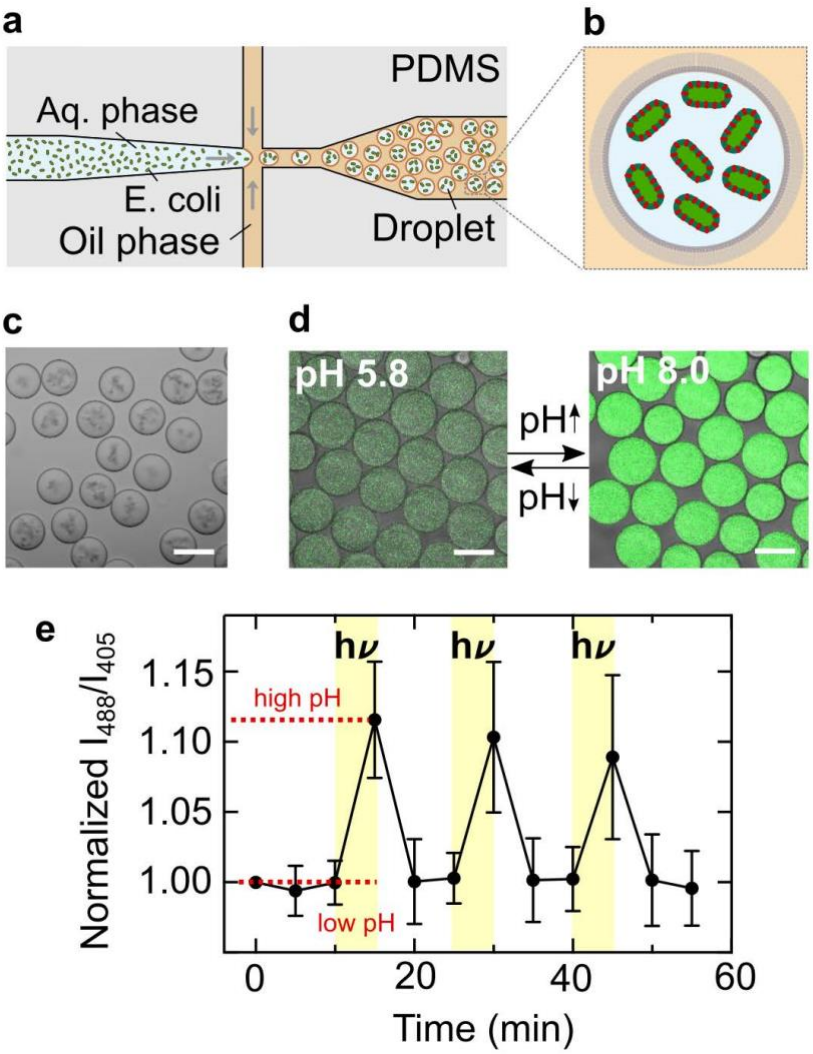


Fig. 2 E. coli as light-activated synthetic organelles that change the pH inside cell-sized confinement. a Schematic illustration of the microfluidic device used to encapsulate engineered *E. coli* and pyranine into cell-sized compartments. Water-in-oil droplets were generated at a flow-focusing T-junction of a PDMS-based device. **b** Schematic illustration of a surfactant-stabilized water-in-oil droplet containing engineered E. coli. **c** Brightfield image of monodisperse water-in-oil droplets with a radius of 27 ± 5 µm (mean ± s.d., n = 53) containing engineered *E. coli* (OD_{600} = 20). Scale bar: 50 µm. **d** Overlay of confocal fluorescence and brightfield images of pyranine (c= 50 µM, λex = 488 nm) inside droplet-based compartments at pH 5.8 and pH 8.0. Scale bar: 50 µm. **e** Normalized fluorescence intensity ratio I_{488}/I_{405} of *E. coli* and pyranine-containing droplets over time. The fluorescence intensity ratio (mean ± s.d., n = 11 droplets) of pyranine (and hence the pH) increases reversibly during periods of illumination with white light (30W halogen bulb, highlighted in yellow). Note that the number of recorded frames was reduced because the illumination light had to be turned off each time an image was acquired, which will bias the proton pumping activity. Source data is available for Fig. 2e.

Towards this goal, we want to implement pH-induced membrane modification and remodeling. For this purpose, we employ a single-stranded DNA sequence, which consists of specifically designed sections[36,39]: First, it contains a self-complementary section, which forms a DNA duplex following the Watson-Crick base-pairing rules. A single-stranded hairpin loop connects the duplex-forming sections. Another critical single-stranded region is located at the 3' end. At acidic pH it wraps around the DNA duplex to form a triplex, held together by Hoogsten interactions. At basic pH, the triplex becomes unstable.The remaining duplex can now also open up, if a second DNA strand

with higher affinity binds to the hairpin loop[36]. By functionalizing this second DNA strand with a terminal cholesterol tag, it self-assembles at the compartment periphery due to hydrophobic interactions[40]. Thereby, we can recruit the triplex-motif strand to the compartment-periphery in a pH-reversible manner (Fig. 3a). At basic pH, the triplex-motif strandis bound to the periphery (Fig. 3b, inset top right and Supplementary Fig. 9). At acidic pH, on the other hand, it remains homogeneously distributed inside the compartment (Fig. 3b, inset bottom left). Note that the periphery attachmentis due to specific interactions between the opened DNA triplex and the complementary cholesterol-tagged DNA. Unspecific absorption in the absence of the cholesterol-tagged DNA was not observed (Supplementary Fig. 10)[41]. To characterize the pH-sensitive membrane attachment, we assessed the fluorescence intensity inside the compartment as a function of pH. The fluorescence intensity decrease with increasing pH follows a sigmoidal fit with a pK_a of 6.05, which is compatible with the pHrange of the *E. coli* and previous works[39]. It is important to note that the choice of fluorophore can affect the pH switching point[41]. As a next step, we need to verify that membrane attachment of the DNA can also be triggered by the engineered *E. coli*. We hence co-encapsulated them with the cholesterol-tagged as well as the triplex-forming DNA strand using a microfluidic two-inlet device (Supplementary Figs. 7 and 11). A second inlet proved to be advantageous, because the cholesterol-tagged DNA could bind to the droplet periphery before encountering the *E. coli*, hence preventing unwanted attachment to the *E. coli* due to hydrophobic interactions[42].

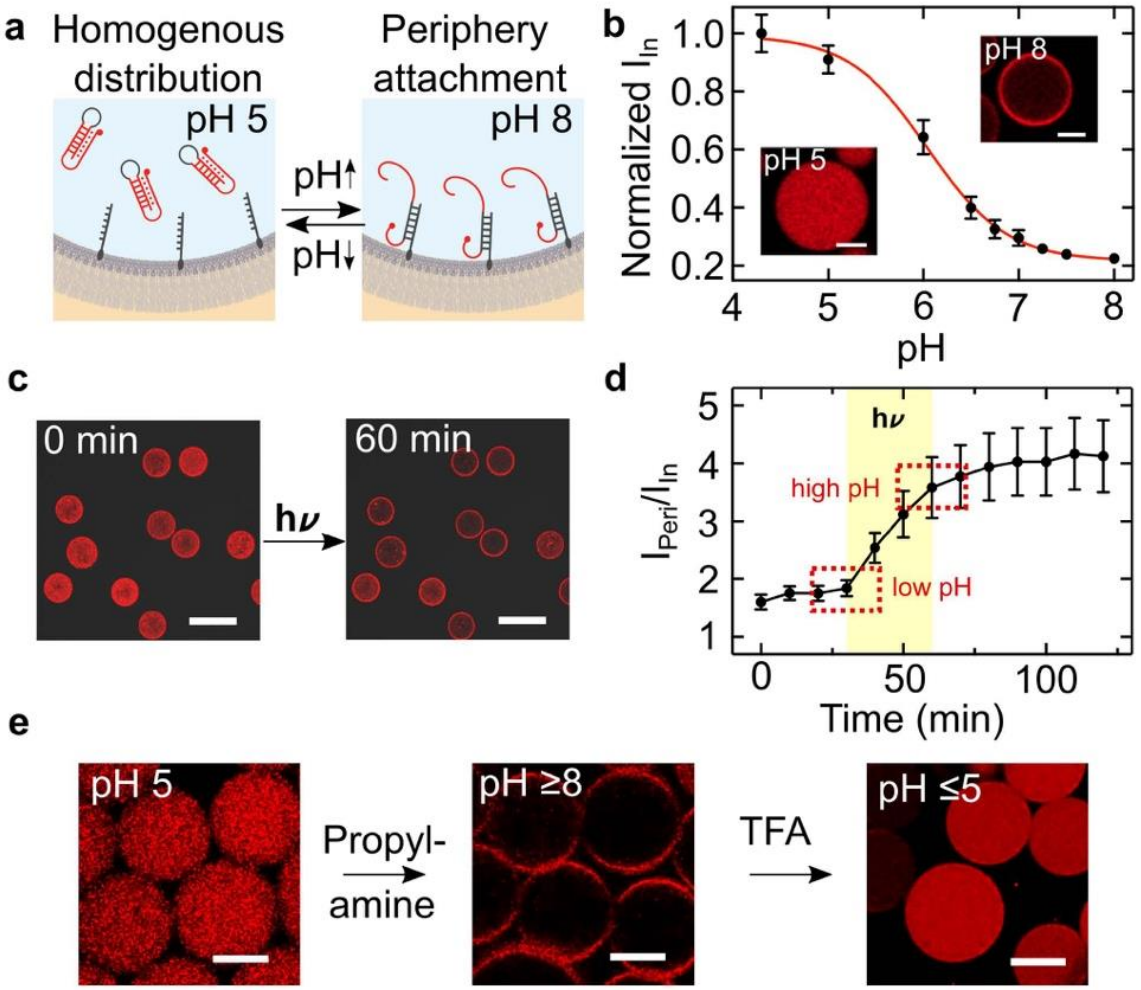


Fig. 3 pH-sensitive DNA attachment to the droplet periphery stimulated with engineered *E. coli*. a Schematic illustration of pH-sensitive duplex formation at the droplet periphery. In response to higher pH, the DNA triplex motif opens up and reversibly attaches to the cholesterol-tagged DNA handles at the compartment periphery. **b** Normalized fluorescence intensity of triplex-forming DNA inside the droplet (excluding the periphery) dependent on the pH (mean ± s.d., n = 20). The sigmoidal fit (red curve) has a turning point at pH 6.05. The insets depict confocal fluorescence images of Cy5-labeled triplex-forming DNA (λex = 633 nm, 1 µM) inside a water-in-oil droplet (containing 1.5 µM cholesterol-tagged DNA) at pH 5 (bottom left) and pH 8 (top right). At pH 8, the triplex-forming DNA is located at the droplet periphery, whereas it is homogeneously distributed at pH 5. Scale bars: 20 µm. **c** Confocal images of microfluidic water-in-oil droplets containing the triplex-forming DNA (λex = 633 nm), cholesterol-tagged DNA and engineered E. coli before (0 min) and after (60 min) illumination with white light. Scale bars: 100 µm. **d** Fluorescence intensity ratio I_{peri}/I_{in} (mean ± s.d., n = 20) of the triplex-forming DNA over time. The ratio increases during light illumination due to binding of the triplex-forming DNA to the droplet periphery. The time period of illumination is indicated in yellow. **e** Confocal images of microfluidic water-in-oil droplets containing the triplex-forming DNA (λex = 633 nm) and cholesterol-tagged DNA produced at pH 5 (left image). Flushing of the proton acceptor propylamine (1 vol% in HFE) led to a pH increase of the aqueous solution inside the droplets and hence attachment of the triplex-forming DNA (middle). Subsequent flushing of the proton donor trifluoroacetic acid (1 vol% in HFE) decreased the pH and hence causes DNA detachment (right). The attachment of triplex-forming DNA to the droplet periphery is reversible. Scale bars: 30 µm. Source data is available for Fig. 3b, d.

After microfluidic droplet formation in the dark, the triplex-forming DNA was homogeneously distributed inside the compartment with some attachment to the periphery (Fig. 3c). From the calibration curve, we could deduce a starting pH value of around 6.2 inside the droplets, consistent with previous experiments in Fig. 1. Upon illumination, the DNA attached to the compartment periphery over the course of 30 min (Fig. 3d, Supplementary Movie 3). We can deduce a pH increase of approximately one pH unit to about pH 7.25, consistent with the bulk experiments in Fig. 1 (Supplementary Note 1). The dynamic opening of the triplex and subsequent attachment to the periphery was considerably slower than the pyranine response[36]. We observed that the DNA remained attached to the compartment periphery after the light was turned off. We found that this is due to an interesting hysteresis effect: Once the DNA duplex at the droplet periphery was formed, the detachment of the triplex-forming DNA was shifted to substantially lower pH values (Supplementary Fig. 12). Therefore, the DNA did not detach when the pH returned to its original value after turning the light off. This effect can likely be attributed to an effective stabilization of the duplex conformation[41]. Detachment could, however, be achieved with larger pH gradients (Supplementary Movie 4): Fig. 3e shows the reversible attachment of the DNA triplex to the compartment periphery, triggered by the addition of a proton acceptor (1 vol% propylamine in HFE) and subsequent addition of a proton donor (1 vol% trifluoroacetic acid (TFA) in HFE). The increase in fluorescence after addition of TFA can be attributed to the pH-sensitive nature of the Cy5 dye[43].

We have thus realized a complex reaction pathway, where illumination activates the internal organelle mimics, causing a proton gradient which, in turn, leads to the stable modification of the compartment periphery. Moreover, the pH-sensitive membrane attachment and the discovered hysteresis effect extend the scope of the DNA triplex motif in DNA nanotechnology.

pH-induced morphology change. Next, we can exploit the pH-responsive modification of the compartment periphery to provide a meaningful function. Assuming that the DNA triplex motif could serve as a shuttle to bring components to the periphery, we set out to develop a cytoskeleton mimic, which could sculpt synthetic cellular compartments in a pH-responsive manner. For this purpose, we designed a DNA origami plate made of two layers of DNA helices (Fig. 4a, Supplementary Fig. 13). The two layers were twisted

at a 90° angle as visible in the cryo electron micrographs (Fig. 4b), providing interaction sites for blunt-end stacking[44] on all four sides of the DNA origami. This, in turn, leads to efficient polymerization of the DNA origami monomers into large flat sheets when the edge staples at the scaffold seam are included as verified via cryo-electron microscopy (Supplementary Fig. 14), atomic force microscopy (Supplementary Fig. 15) and agarose gel electrophoresis (Supplementary Fig. 16). The bottom-side of the DNA origami was functionalized with the DNA triplex motif at four positions. At basic pH, the DNA origami thus attached to the periphery of cell-sized droplets functionalized with the complementary cholesterol-tagged strand. However, the droplets remained spherical (Supplementary Fig. 17). This is not surprising given that droplets could also not be deformed with cytoskeletal proteins due to their interfacial properties[17,45]. We thus moved to a compartment system which better mimics the mechanical properties of cellular membranes. We produced giant unilamellar lipid vesicles (GUVs) and functionalized them externally with the cholesterol-tagged DNA. We then added the pH-sensitive DNA origami to the GUVs at pH 8.3. At this pH, the DNA origami binds to the GUV membrane. Upon addition of the staples at the scaffold seam, which enable blunt-end stacking, we observed considerable deviations from the initially spherical shape of the GUV (Fig. 4c, d). Large flat sections appeared on the GUV with kinks at the phase boundaries between the polymerized flat DNA sheets. In fluorescence recovery after photobleaching (FRAP) experiments, we find that the polymerized DNA origami layer is not diffusive, as expected for large interconnected sheets, in particular in the presence of Mg^{2+} [46]. For the underlying deformed lipid membrane, we obtain a diffusion coefficient of $1.23 \pm 0.14\ \mu m^2 s^{-1}$ which is comparable to the lipid diffusion in bare GUVs[47] (Supplementary Fig. 18). In addition to the morphological change, we observe a suppression of membrane fluctuations (Supplementary Fig. 19, Supplementary Movie 5), indicating a mechanical stabilization of the compartment[35,48] by the DNA-based exoskeleton mimic (Supplementary Note 2). The stabilization effect could potentially be exploited for drug delivery applications. Both the morphological and the mechanical alterations are reversible (Supplementary Fig. 20): Addition of an acid led to pH decrease and hence to the detachment of the DNA origami from the GUV membrane. Notably, the GUVs relax back into their initial spherical shape (Fig. 4d, more images in Supplementary Figs. 21 and 22). Note that a pH decrease to pH 5.6 is required to fully detach the DNA origami from the GUV membrane, which is below the pH decrease that can be provided by the *E. coli* (see Fig. 1). The larger pH gradients required for attachment and detachment of the DNA origami compared to the triplex strand alone can be explained with a cooperativity effect. Each DNA origami is modified with four triplexes. Therefore, complete detachment took several hours and hence the addition of an acid was necessary. The histograms in Fig. 4e quantify the pH-reversible morphology change of the GUVs, revealing lower and more broadly distributed circularities when the DNA origami was attached at high pH. Taken together, the self-assembly of nanoscopic pH-responsive building blocks could trigger the microscopic morphological remodeling of the shape of lipid membrane-based synthetic cellular compartments.

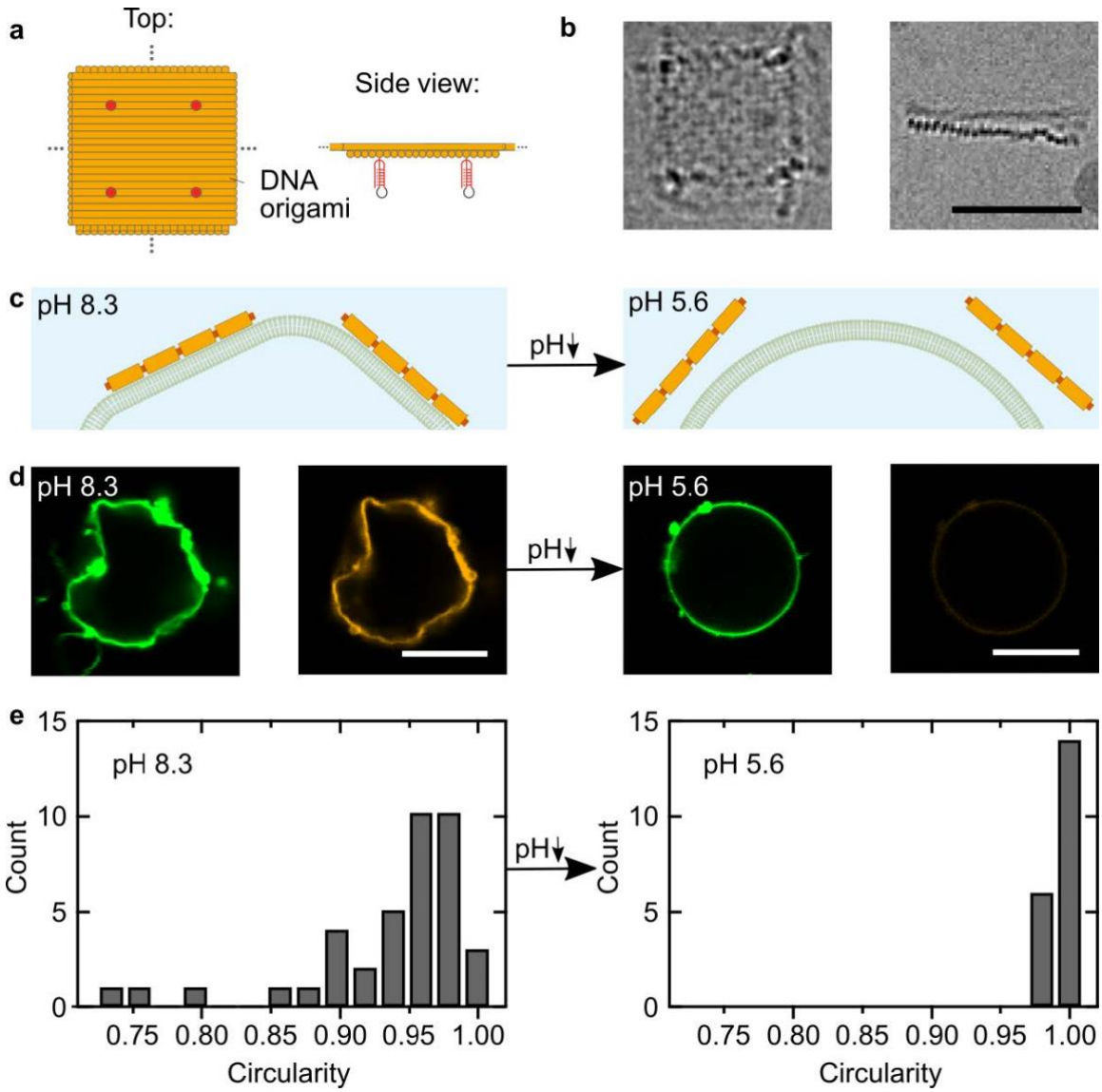


Fig. 4 Deformation of GUVs with pH-sensitive DNA origami. a Schematic illustration of the DNA origami, which can polymerize into flat DNA origami sheets due to blunt end stacking. The DNA origami was functionalized with four DNA triplex motifs (red, two are shown), such that its assembly on the GUV membrane is pH-dependent. **b** Cryo-EM micrographs of the DNA origami plates. The top view (left) and the side view (right) showing the two DNA layers connected at a 90° angle. Scale bar: 50 nm. **c** Schematic illustration of a section of a GUV membrane functionalized with cholesterol-tagged pH-sensitive polymerized DNA origami. At high pH the DNA origami sculpts the GUV membrane. At low pH, it detaches and the GUV relaxes into its spherical shape. (Continued on the following page) **d** Confocal images of GUVs before (left) and after (right) decreasing the pH from pH 8.3 to pH 5.6 by addition of iso-osmotic potassium dihydrogenphosphate buffer. The GUV (lipids labeled with Atto488, λex = 488 nm) is initially deformed due to the membrane-bound polymerized DNA origami (labeled with Cy3, λex = 561 nm). The DNA origami detaches upon lowering the pH (the fluorescence from the detached DNA origami in the background is too weak to be visible). Scale bars: 10 µm. **e** Histograms of GUV circularity before (left) and after (right) lowering the pH. At pH 8.3, the mean circularity is 0.94 ± 0.06 (n = 39) compared to 0.991 ± 0.004 (n=20) at pH 5.6, respectively. Source data is available for Fig. 4e.

Finally, we set out to combine the DNA origami-mediated pH-sensitive deformation of GUVs with the light-responsive proton pumping capabilities of the *E. coli*. First, we showed that the GUVs remained stable in the *E. coli* culture and that we can attach the plain triplex-forming DNA to the GUV membrane upon illumination (Supplementary Figs. 23 and 24; Supplementary Movie 6). Thus, the pH-signal-transduction between the top

down engineered *E. coli* and bottom-up assembled synthetic cellsis also successful when the *E. coli* are used as external actuators. Next, we immersed the GUVs in a solution of *E. coli* and pH-sensitive DNA origami. We observed the attachment of the pH-sensitive DNA origami to the GUV membrane upon illumination(Fig. 5a, Supplementary Movie 7). Fig. 5b quantifies the amount of DNA origami attachment, i.e., normalized fluorescence intensity at the GUV periphery I_{Peri}, over time during periodsof illumination and darkness. DNA origami attachment happens repeatedly and only during periods of illumination, untilsaturation is reached. The attachment occurs due to the pH increase triggered by the light-harvesting *E. coli*. After illumination, in periods of darkness, the pH gradient dissipates. Nevertheless, the amount of DNA origami attachment remains roughly constant (Fig. 5b, t = 50 min until t = 110 min). This can beattributed to the observed pH hysteresis effect (Supplementary Fig. 12): The pH would have to drop below the starting value for detachment, which cannot be achieved with the *E. coli* alone. Nevertheless, a second illumination cycle (from t = 110 min)showed that the *E. coli* remain active and that the DNA origami attachment continued until saturation was reached.
After attachment, we enabled the polymerization of individualDNA origami monomers by adding the staple strands at thescaffold seam, which induce blunt-end stacking. This leads to thedeformation of GUVs within two hours (Fig. 5c, Supplementary Movie 8), when the solution was illuminated with light, whereas GUVs remained spherical when they were left in the dark (Supplementary Figs. 25 and 26). Note that the deformation is weaker compared to the deformation achieved with conventionalpH switching due to the smaller pH gradient. We can thus exploit the light-harvesting *E. coli* to actuate a morphology change ofthe GUVs.

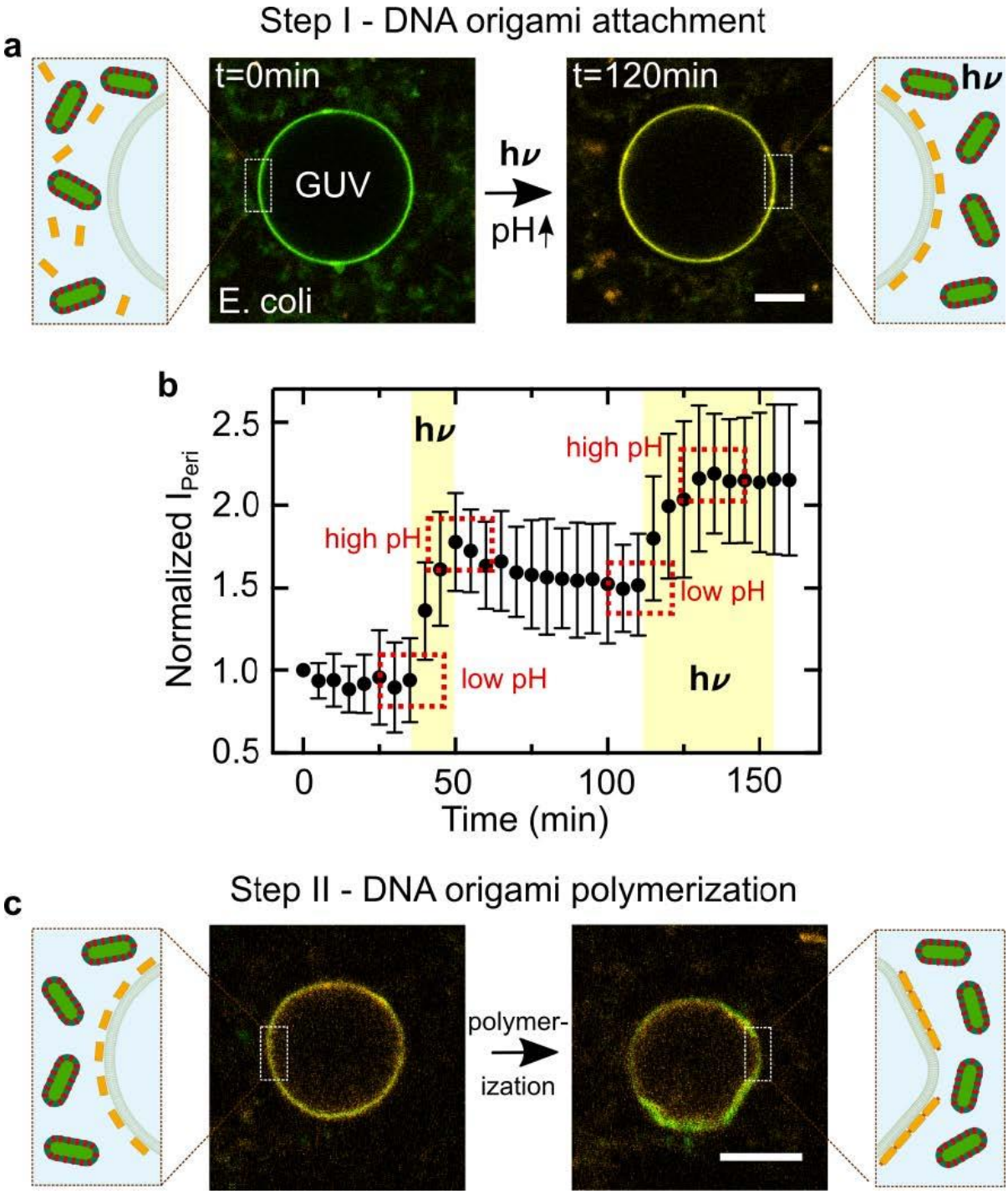


Fig. 5 Light-harvesting *E. coli* trigger DNA origami attachment and GUV deformation. a Confocal images of a GUV (λex = 488 nm, green) functionalized with cholesterol-tagged DNA immersed in an *E. coli* and DNA origami (λex= 561 nm, orange) containing solution before (left) and after (right) light illumination. Scale bar: 10 µm. After light illumination, the DNA origami (orange) attaches to the GUV due to the pH increase triggered by the *E. coli*. **b** Normalized fluorescence intensity Iperi (mean ± s.d., n = 11) of the triplex-forming DNA at the GUV periphery monitored over time. The time period of illumination is indicated in yellow, illumination leads to a pH increase and hence DNA origami attachment. **c** Confocal images of a GUV (λex = 488 nm, green) after light-mediated DNA origami (λex = 561 nm, orange) attachment to the membrane and addition of the DNA staple strands at the scaffold seam which enable blunt-end stacking. DNA origami polymerization leads to the deformation of the GUV membrane within 2 h. Scale bar: 10 µm. Source data is available for Fig. 5b.

7.4 Discussion

In summary, we have shown that the use of top-down engineered bacteria can enhance bottom-up assembled synthetic cells. The light-induced proton gradients we achieve with xenorhodopsin-overexpressing *E. coli* are not only larger than what was previously achieved with purified and reconstituted proteins – we also circumvent the laborious processes involved in their preparation. Especially membrane proteins, which can provide transient or chemically storable forms of energy as well as signal transduction and molecular transport in living cells, can be challenging to purify and reconstitute. Therefore, we can exploit the engineered *E. coli* to drive sophisticated downstream dynamics in synthetic cells. In particular, we demonstrate the pH-sensitive attachment of a triplex-motif-carrying DNA origami to the compartment periphery upon illumination. The polymerized DNA origami, in turn, leads to a shape change of the GUVs triggered by the proton-pumping activity of the *E. coli*. Thepossibility to manipulate lipid membranes and not just the DNA nanostructures themselves broadens the scope of the popularDNA triplex-motif. For biotechnological applications, compartments that modify themselves as a response to environmental factors are highly desirable. More general, the integration of top-down engineered cells into bottom-up synthetic biology, contributing to bridge a long-standing divide, will provide the potential to realize diverse functions beyond light-harvesting[49]. We envision that the integration of top-down engineered components in synthetic cells will be a leap forward in their complexity and functionality.

7.5 Methods

Cloning. The plasmid pNR31 harboring the xenorhodopsin gene from Nanosalina (NsXeR) fused to the gene coding for superfolder-GFP (sf-GFP) was assembled by replacing the gene coding for proteorhodopsin in plasmid pNR03[7] with the NsXeR gene (Supplementary Table 1). Therefore, a codon-optimized NsXeR gene based on the amino-acid sequence[25] with a 5' NdeI and a 3' BamHI restriction site was synthesized by GenScript (https://www.genscript.com) and cloned into the pUC57 plasmid. Using these two restriction enzymes (New England Biolabs, Ipswich, MA), the NsXeR gene was then subcloned into the pNR03 plasmid. The plasmid pNR33 harboring the NsXeR gene fused to mCherry (Supplementary Table 1) was assembled in multiple steps. First the sf-GFP gene in pNR03 was replaced by the gene coding for mCherry. To that end, the mCherry gene was amplified from the pNR09 plasmid using primers 5'-GGC GGA TCC ATG CAT AGC AAG GGC GAG-3' and 5'-GCC AAG CTT CTT GTA CAG C-3' (Microsynth AG) to introduce 5' BamHI and 3' HindIII restriction sites[7]. The resulting PCR-product was then cloned into plasmid pNR03 where it replaced the sf-GFP gene. Subsequently the same subcloning as for plasmid pNR31 was performed to replace the gene coding for proteorhodopsin with the NsXeR gene.

Overexpression of fusion-proteins in *E. coli*. *E. coli* C41 (DE3) cells (Sigma-Aldrich) were transformed with the plasmids pNR31 and pNR33. 100 mL Luria-Bertani (Fisher Scientific) liquid cultures (100 µg/mL ampicillin, Sigma-Aldrich) were inoculated 1:100 from overnight cultures. The *E. coli* cells were grown at 37°C while shaking at 220 rpm until an OD_{600} of 0.4 was reached. Then, all-trans-retinal (Sigma-Aldrich) was added to a concentration of 10 µM and the expression of the fusion-proteins was induced with the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich). The cells were incubated for another 4 h at 37°C while shaking at 220 rpm. Subsequently they were harvested by centrifugation (3200 × *g* for 10 min at 4°C) and resuspended in 150 mM NaCl. The cells were stored at 4°C and protected from light until further use.

Photoactivity measurements with a micro pH-electrode. *E. coli* cells over-expressing either XeR-GFP or XeR-mCherry were washed twice with 150 mM NaCl (3200 × g for 10

min at 4°C) prior to photoactivity measurements. Immediately before the measurement, another washing step was performed. The bacteria were concentrated to an OD_{600} of 20. Photoactivity measurements were conducted using a micro pH-electrode (InLab Micro Pro, Mettler Toledo, Columbus, OH) and a sample volume of 800 µL. The pH was recorded every 10 s. During the measurements the bacteria were protected from ambient light and continuously stirred to prevent sedimentation. The sample was illuminated with a KL 1500 LCD halogen lamp (Schott) for 5 min during each light-dark cycle. After each illumination-period the sample was kept in the dark for 10 min. All measurements were performed at room temperature.

Confocal fluorescence microscopy. A confocal laser scanning microscope LSM 880, LSM 800 or LSM 700 (Carl Zeiss AG) was used for confocal imaging. The pinhole aperture was set to one Airy Unit and experiments were performed at room temperature. The images were acquired using a 20x objective (Plan-Apochromat 20x/0.8 M27, Carl Zeiss AG). Images were analyzed and processed with ImageJ (NIH, brightness and contrast adjusted) and Zen imaging software (Version 2.3). Confocal images were analyzed using a custom-written ImageJ macro (available here: https://doi.org/10.5281/zenodo.4738934). Cell-sized compartments were identified, their radius calculated and the intensity within the compartment center defined as mean inner intensity I_{In}. The peripheral intensity was determined by quantifying the maximum intensity along a line orthogonal to the compartment periphery. This was repeated every 18° and the mean value taken as I_{Peri}. The resulting data was plotted with Prism 8 (Version 8.4.3) and figures were compiled with Inkscape (Version 1.0rc1).

Formation of surfactant-stabilized water-in-oil droplets. Microfluidic PDMS-based (Sylgard 184, Dow Corning) devices for the formation of water-in-oil droplets were produced and assembled[40]. The device layouts of the single and two-inlet devices are shown in the Supplementary Fig. 7. The oil-phase contained 1.4 wt% of Perflouro-polyether-polyethylene glycol (PFPE-PEG) block-copolymer fluorosurfactants (PEG-based fluorosurfactant, Ran Biotechnologies, Inc.) dissolved in HFE-7500 oil (DuPont). The aqueous phase contained the encapsulated content and was varied as described in the corresponding sections. The fluid pressures to induce droplet formation were controlled by an Elveflow microfluidic flow control system or syringe pumps (Harvard Apparatus). The fluids for the syringe pumps were injected into the channels with 1 ml syringes (Omnifix, B. Braun, Germany) connected by a cannula (Sterican®0.4 ×20 mm, BL/LB, B.Braun) as well as PTFE-tubing (0.4 × 0.9 mm, Bola). To observe the droplet production process, an Axio Vert.A1 (Carl Zeiss AG) inverse microscope was used. As an alternative to the microfluidic formation of droplets, the aqueous phase was layered on top of the oil phase within a microtube (Eppendorf) and droplet formation was induced by manual shaking[50].

Photoactivity measurements in droplets. Photoactivity measurements in droplets were performed by encapsulating *E. coli* (OD_{600} ≈ 20) with pyranine (50 µM) into surfactant-stabilized droplets using the microfluidic device described above. The droplets were stored at 4°C after formation to allow for equilibration of the pH inside the droplet. Subsequently, droplets were sealed in an observation chamber and observed with confocal fluorescence microscopy. After 10 min of imaging in the dark, the sample was illuminated for 5 min using a Photonic PL 1000 lamp (light intensity 8 Mlx using a 30 W halogen bulb). The lightguide was placed 5-10 cm above the sample. These cycles were repeated for 1 h.

pH-sensitive attachment of DNA to the droplet periphery. Cholesterol-tagged DNA (sequence: 5' (Cy3)-ACCAGACAATACCACACAATTTT-CholTEG 3', HPLC purified) and the

Cy-5 labeled triplex-forming DNA (sequence: 5' Cy5-TTCTCTTCTCGTTTGCTCTTCTCTTGTGTGGTATTGTCTAAGAGAAGAG 3', adapted from Green et al.[36], HPLC purified) were purchased from Biomers or Integrated DNA Technologies. Both DNA sequences were encapsulated in microfluidic droplets at a concentration of 1.5 µM and 1 µM, respectively. For the calibration measurement (Fig. 3b), the aqueous solution inside the droplets additionally contained 50 mM potassium phosphate buffer at the respective pH. Propylamine (from Sigma Aldrich) and Trifluoracetic Acid (TFA, from Sigma Aldrich) were flushed to dynamically change the pH of the droplets' aqueous phase. For the co-encapsulation of the DNA together with the *E. coli* (OD_{600} ≈ 20), a two-inlet droplet formation device was used (see Supplementary Fig. 7). Droplets were sealed in an observation chamber for confocal fluorescence imaging experiments.

GUVs electroformation and DNA attachment. GUVs consisting of 99 % DOPC (1,2-dioleoyl-sn-glycero-3-phosphocholine, from Avanti Polar Lipids) and 1 % Atto488-DOPE (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-Atto488, from AttoTEC) in 120 mM sucrose were produced via electroformation using a Vesicle Prep Pro (Nanion)[34]. An AC-current with an amplitude of 3 V and a frequency of 5 Hz was applied for 2 h at 37°C. The cholesterol-tagged DNA and the triplex-forming DNA were added to the GUVs at a concentration of 0.6 µM and 0.4 µM, respectively, before the addition of the *E. coli* (OD_{600} ≈ 20), in an unbuffered solution containing 150 mM NaCl and 5 mM $MgCl_2$.

DNA origami design and assembly. DNA origami structures were adapted from an earlier design by Kopperger et al.[51] using the open-access software *cadnano*[52]. Several changes were introduced, in particular: (1) Addition of nine DNA staple strand overhangs on the top layer, complementary to single stranded fluorescent Cy3-tagged DNA; (2) Addition of four single stranded overhangs on the bottom layer, complementary to the triplex-forming DNA; (3) Complete redesign of the edge staples resulting in a cross-shaped plate. The sticky cross DNA origami contained edge staples that finish the scaffold seam, enabling blunt-end stacking with neighboring origami. (4) Use of the longer single-stranded scaffold DNA, type p8064. A complete list of the DNA sequences is shown in Supplementary Data 1, the details of the design are shown in Supplementary Fig. 13. DNA origami was assembled by adding all unmodified staple strands (Integrated DNA Technologies, Inc., purification: standard desalting) in fivefold excess compared to the p8064 scaffold strand (tilibit nanosystems GmbH). The solution contained 1 × TAE (Tris-Acetate-EDTA, Sigma-Aldrich) and 20 mM $MgCl_2$ (Sigma-Aldrich) at pH 7.4. The DNA origami was annealed in a thermocycler (Bio-Rad T100) that controls a temperature ramp from 70°C to 20°C over 12 h and successively holds the temperature at 40°C for at least 3 h[51]. The unpurified samples were stored at 4°C until further use.

Purification of the DNA origami. Prior to purification from excess staples, the DNA origami was mixed with 1 µM Cy3-tagged single-stranded DNA (Integrated DNA Technologies, Inc., DNA sequence: 5' Cy3- AAAAAAAAAAAAAAAAAA 3', purification: HPLC) as well as a pH-sensitive triplex-forming DNA motif (Integrated DNA Technologies, Inc., DNA sequence: 5' TTCTCTTCTCGTTTGCTCTTCTCTTGTGTGGTATTGTCTAAGAGAAGAG TTTGATGCATAGAAGG 3'). The DNA origami was then suspendend in 500 µL of 1 × TAE, 5 mM $MgCl_2$ and purification was performed by spin filtration in a Biofuge Fresco microlitre centrifuge (Heraeus 75005521) using 100 kDa cutoff filters from Amicon (Amicon Ultra-15, PLHK Ultracel-PL Membran, UFC910008)[31]. After filtration, the $MgCl_2$ concentration was raised to 20 mM again. Before the experiment, microvolume

spectrophotometry (PEQLAB Biotechnologie GmbH) yielded a DNA origami concentration of 6.54 ± 0.42 nM.

Cryo electron microscopy. In total, 3 µL of the assembled DNA origami in 10 mM sodium phosphate pH 8.3 containing 20 mM $MgCl_2$ were blotted for 5-10 s in a (Vitrobot Mark IV, Thermo Fischer) on Quantifoli 2/1 grids with zero blot force at 100% humidity. Plunge frozen samples were imaged in a Krios equipped with a K3 camera behind an energy filter at a pixel size of 0.137 nm. Images were taken by single particle program (EPU, Thermo Fischer) with a a total dose of 20 e/A2. Movies of 20 frames were corrected[53] then cropped, normalized, low-pass filtered (0.0625) and 4x binned[54].

GUV deformation with pH-sensitive DNA origami. The DNA origami (in 1 x TAE, 20 mM $MgCl_2$) was incubated with cholesterol-tagged DNA at 50 nM for 25 minutes and immediately mixed with Atto488-labeled iso-osomotic (120 mOsmol) GUVs in a ratio of one to three. DNA origami-coated GUVs were imaged after 24 hours of incubation in the fridge. Subsequently, the GUVs were incubated for another 24 hours with 48 mM KH_2PO_4 buffer in order to detach the DNA origami from the GUV membrane.

GUV deformation with pH-sensitive DNA origami and light-responsive *E. coli*. For these experiments, the DNA origami was suspended in a solution of 75 mM NaCl and 10 mM $MgCl_2$ without addition of a pH-buffering agent. In order to prevent DNA origami polymerization prior to GUV-attachment, the staple strands at the scaffold seam were excluded. The single-stranded scaffold loops prevent base stacking. Subsequently, Atto488-labeled GUVs in sucrose (195 mOsmol) were diluted in 75 mM NaCl and 10 mM $MgCl_2$ and mixed with 2 µM cholesterol-tagged DNA. After 5 min incubation, *E. coli* resuspended at an OD_{600} = 60 in 75 mM NaCl and 10 mM $MgCl_2$ and DNA origami were mixed in a 1:1:1 ratio with GUVs. For the DNA origami attachment the solution was put into an observation chamber and illuminated with white light during confocal imaging. For subsequent GUV deformation, the staples at the scaffold seam were added to induce base stacking interactions between the membrane-bound DNA origami. The solution was illuminated for 30 min in bulk before the addition of 50 nM staple strands and imaged 14 h later.

Statistics and reproducibility. The experiments were performed independently at least two times. In particular, values in Fig. 1b correspond to three independent experiments and in Fig. 1d to four independent experiments. Experiments for Figs. 2e, 3b, d and 4e were performed two times or more. Fig. 5 was only replicated once, however due to the sequential attachment via light and appropriate controls, we believe that this is adequate. All representative confocal, atomic force and electron microscopy images are only a subset of at least 10 or more images showing similar results.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Source data for main figures 1b, c, d, 2e, 3b, d, 4e, 5b and supplementary figures 1a, c, d, 2, 3a, b, 5, 8, 12, 16, 18b, 19a, b and 24c are provided with this paper.

Code availability

Image J macro for the analysis of the intensity inside the compartment and at the compartment periphery is provided under the following link: https://doi.org/10.5281/zenodo.4738934.

Acknowledgements
K.J., R.R.S., J.P.S. and K.G. acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy via the Excellence Cluster 3D Matter Made to Order (EXC-2082/1 – 390761711). K.J. thanks the Carl Zeiss Foundation. J.P.S and I.P. acknowledge funding from the European Research Council, Grant Agreement no. 294852, SynAd, the MaxSynBio Consortium (jointly funded by the Federal Ministry of Education and Research of Germany and the Max Planck Society), from the SFB 1129 of the German Science Foundation and from the VolkswagenStiftung (priority call 'Life?'). J.P.S. is the Weston Visiting Professor at the Weizmann Institute of Science. G.H. and R.R.S. acknowledge the services SDShd and bwHPC supported by the Ministry of Science, Research and the Arts Baden-Württemberg. The Krios microscope is part of the Cluster of Excellence "Cell-Networks" (Exc 81) at the Universität Heidelberg. K.G. received funding from the European Union Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 792270. The authors acknowledge the Max Planck Society for its general support. N.R. and D.J.M. thank the Swiss Nanoscience Institute (SNI, Basel, Switzerland), the Swiss National Science Foundation and NCCR Molecular Systems Engineering for their support.

7.6 References

1. Benner, S. A. & Sismour, A. M. Synthetic biology. Nat. Rev. Genet. 6, 533–543 (2005).
2. Schwille, P. Jump-starting life? Fundamental aspects of synthetic biology. J. Cell Biol. 210, 687–690 (2015).
3. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. Nat. Rev. Microbiol. 12, 381–390 (2014).
4. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat. Biotechnol. 30, 354–359 (2012).
5. Anderson, J., Clarke, E., Arkin, A. & Voigt, C. Environmentally Controlled Invasion of Cancer Cells by Engineered Bacteria. J. Mol. Biol. 355, 619–27 (2006).
6. Schuergers, N., Werlang, C. & Boghossian, A. A. A synthetic biology approach to engineering living photovoltaics. Energy Environ. Sci. 10, 1102–1115 (2017).
7. Ritzmann, N. et al. Fusion domains guide the oriented insertion of light-driven proton pumps into liposomes. Biophys. J. 113, 1181–1186 (2017).
8. Ausländer, S., Ausländer, D. & Fussenegger, M. Synthetic biology-the synthesis of biology. Angewandte Chemie International Edition 56, 6396–6419 (2017).
9. Göpfrich, K., Platzman, I. & Spatz, J. P. Mastering complexity: towards bottom-up construction of multifunctional eukaryotic synthetic cells. Trends Biotechnol. 36, 938–951 (2018).
10. Schwille, P. et al. MaxSynBio: avenues towards creating cells from the bottom up. Angewandte Chemie Int. Ed. 57, 13382–13392 (2018).
11. Supramaniam, P., Ces, O. & Salehi-Reyhani, A. Microfluidics for artificial life: techniques for bottom-up synthetic biology. Micromachines 10, 299 (2019).
12. Otrin, L. et al. Toward artificial mitochondrion: mimicking oxidative phosphorylation in polymer and hybrid membranes. Nano Lett. 17, 6816–6821 (2017).
13. Berhanu, S., Ueda, T. & Kuruma, Y. Artificial photosynthetic cell producing energy for protein synthesis. Nat. Commun. 10, 1325 (2019).
14. Miller, T. E. et al. Light-powered CO2 fixation in a chloroplast mimic with natural and synthetic parts. Science 368, 649–654 (2020).
15. Keber, F. C. et al. Topology and dynamics of active nematic vesicles. Science 345, 1135–1139 (2014).
16. Bartelt, S. M., Steinkühler, J., Dimova, R. & Wegner, S. V. Light-Guided Motility of a Minimal Synthetic Cell. Nano Lett. 18, 7268–7274 (2018).
17. Jahnke, K. et al. Engineering light-responsive contractile actomyosin networks with DNA nanotechnology. Adv. Biosyst. 2020, 2000102.
18. Steinkühler, J. et al. Controlled division of cell-sized vesicles by low densities of membrane-bound proteins. Nat. Commun. 11, 1–11 (2020).
19. 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 Biol. 8, 564–570 (2016).
20. Lentini, R. et al. Two-way chemical communication between artificial and natural cells. ACS Central Sci. 3, 117–123 (2017).
21. Lentini, R. et al. Integrating artificial with natural cells to translate chemical messages that direct E. coli behaviour. Nat. Commun. 5, 1–6 (2014).
22. Staufer, O.. Schröter, M., Platzman, I. & Spatz, J. P. Bottom-up assembly of functional intracellular synthetic organelles by droplet-based microfluidics. Small 2020, 1906424.

23. Mehta, A. P. et al. Engineering yeast endosymbionts as a step toward the evolution of mitochondria. Proc. Natl. Acad. Sci. USA 115, 11796–11801 (2018).
24. Lau, Y. H., Giessen, T. W., Altenburg, W. J. & Silver, P. A. Prokaryotic nanocompartments form synthetic organelles in a eukaryote. Nat. Commun. 2018, 9.
25. Shevchenko, V. et al. Inward H+ pump xenorhodopsin: mechanism and alternative optogenetic approach. Sci. Adv. 3, e1603187 (2017).
26. Inoue, K. et al. A natural light-driven inward proton pump. Nat. Commun. 7, 13415 (2016).
27. Marušič, N. et al. Constructing artificial respiratory chain in polymer compartments: Insights into the interplay betweenbo3oxidase and the membrane. Proc. Natl. Acad. Sci. 117, 15006–15017 (2020).
28. Goers, R. et al. Optimized reconstitution of membrane proteins into synthetic membranes. Commun. Chem. 2018, 1.
29. Kano, K. & Fendler, J. H. Pyranine as a sensitive pH probe for liposome interiors and surfaces. pH gradients across phospholipid vesicles. Biochimica et Biophysica Acta (BBA) - Biomembranes 509, 289–299 (1978).
30. Langecker, M. et al. Synthetic lipid membrane channels formed by designed DNA nanostructures. Science 338, 932–6 (2012).
31. Göpfrich, K. et al. Large-conductance transmembrane porin made from DNA origami. ACS Nano 2016, 10.
32. Czogalla, A. et al. Amphipathic DNA origami nanoparticles to scaffold and deform lipid membrane vesicles. Angewandte Chemie International Edition 54, 6501–5 (2015).
33. Franquelim, H. G., Khmelinskaia, A., Sobczak, J.-P., Dietz, H. & Schwille, P. Membrane sculpting by curved DNA origami scaffolds. Nat. Commun. 9, 811 (2018).
34. Göpfrich, K. et al. DNA-tile structures lead to ionic currents through lipid membranes. Nano Lett. 15, 3134–3138 (2015).
35. Franquelim, H. G., Dietz, H. & Schwille, P. Reversible membrane deformations by straight DNA origami filaments. Soft Matter 17, 276 (2020).
36. Green, L., Amodio, A., Subramanian, H. K. K., Ricci, F. & Franco, E. pHdriven reversible self-assembly of micron-scale DNA scaffolds. Nano Lett. 17, 7283–7288 (2017).
37. Göpfrich, K., Urban, M. J., Frey, C., Platzman, I. & Spatz, J. P. Dynamic actuation of DNA-assembled plasmonic nanostructures in microfluidic cell-sized compartments. Nano Lett. 20, 1571–1577 (2020).
38. Kuzyk, A., Urban, M. J., Idili, A., Ricci, F. & Liu, N. Selective control of reconfigurable chiral plasmonic metamolecules. Sci. Adv. 3, e1602803 (2017).
39. Porchetta, A., Idili, A., Vallée-Bélisle, A. & Ricci, F. General strategy to introduce pH-induced allostery in DNA-based receptors to achieve controlled release of ligands. Nano Lett. 15, 4467–4471 (2015).
40. Jahnke, K. et al. Programmable functionalization of surfactant-stabilized microfluidic droplets via DNA-tags. Adv. Funct. Mater. 29, 1808647 (2019).
41. Jahnke, K., Grubmüller, H., Igaev, M. & Göpfrich, K. Choice of fluorophore affects dynamic DNA nanostructures. Nucleic Acids Res. 2021, gkab201.
42. Jia, H.-R., Zhu, Y.-X., Chen, Z. & Wu, F.-G. Cholesterol-assisted bacterial cell surface engineering for photodynamic inactivation of gram-positive and gram-negative bacteria. ACS Appl. Mater. Interfaces 9, 15943–15951 (2017).
43. Cooper, M. E., Gregory, S., Adie, E. & Kalinka, S. pH-Sensitive Cyanine Dyes for Biological Applications. J. Fluorescence 12, 425–429 (2002).

44. Kilchherr, F. et al. Single-molecule dissection of stacking forces in DNA. Science 353, aaf5508 (2016).
45. Vogel, S. K. et al. Symmetry breaking and emergence of directional flows in minimal actomyosin cortices. Cells 9, 1432 (2020).
46. Kocabey, S. et al. Membrane-assisted growth of DNA origami nanostructure arrays. ACS Nano 9, 3530–3539 (2015).
47. Schaich, M., Sobota, D., Sleath, H., Cama, J. & Keyser, U. F. Characterization of lipid composition and diffusivity in OLA generated vesicles. Biochimica et Biophysica Acta (BBA) - Biomembranes 1862, 183359 (2020).
48. Baumann, K. N. et al. Coating and stabilization of liposomes by clathrin-inspired DNA self-assembly. ACS Nano 14, 2316–2323 (2020).
49. Elani, Y. Interfacing living and synthetic cells as an emerging frontier in synthetic biology. Angewandte Chemie Int. Edi. 59, 2–12 (2020).
50. Göpfrich, K. et al. One-pot assembly of complex giant unilamellar vesicle-based synthetic cells. ACS Synthetic Biology 8, 937–947 (2019).
51. Kopperger, E. et al. A self-assembled nanoscale robotic arm controlled by electric fields. Science 359, 296–301 (2018).
52. Douglas, S. M. et al. Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic Acids Res. 37, 5001–5006 (2009).
53. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017).
54. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012).

7.7 Supplemental Information

Supplementary Figure S1: Activation of xenorhodopsin with different light sources.

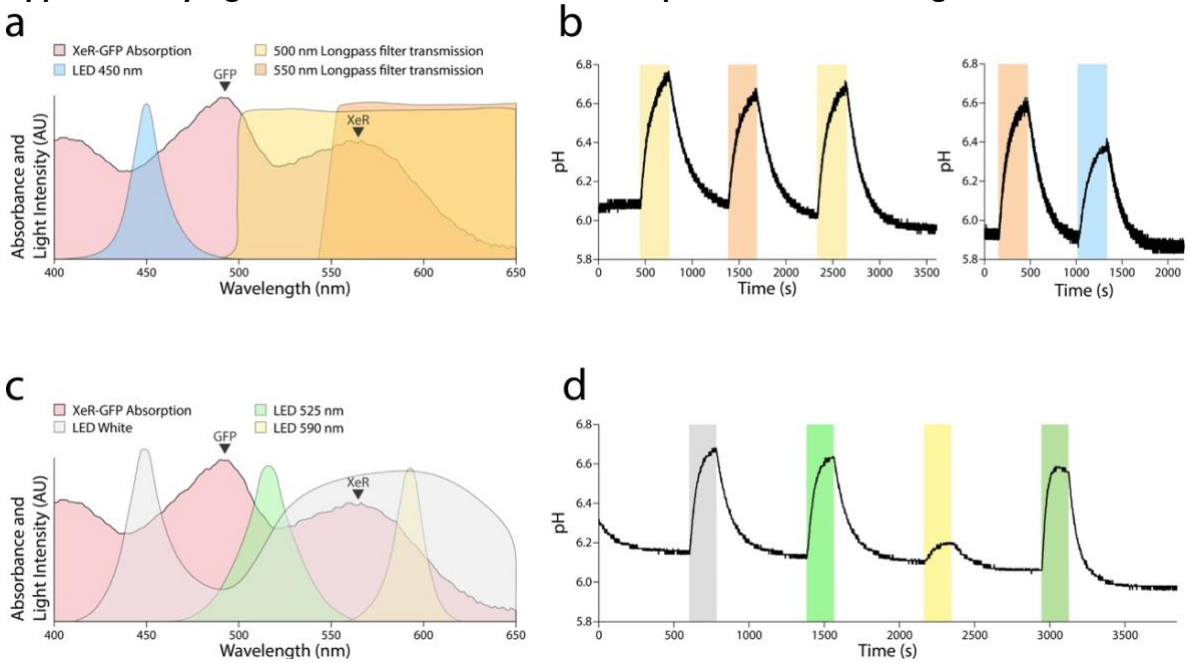


Figure 1: Activation of xenorhodopsin (XeR) with different light sources. a Absorption spectrum of the xenorhodopsin-GFP (XeR-GFP) fusion protein shown in light red. Respective absorption peak positions of xenorhodopsin and GFP are indicated. Relative light intensity and transmission profiles of filters and light sources are also shown. b Bulk photoactivity traces induced with the light sources presented in a. c Light intensity profiles of white, green and yellow LEDs plotted together with the XeR-GFP absorption spectrum. d Bulk photoactivity traces induced with the light sources shown in c. The dark green shaded region represents a combination of all three light sources shown in c.

Supplementary Figure S2: Saturation of pH-gradients upon white light illumination

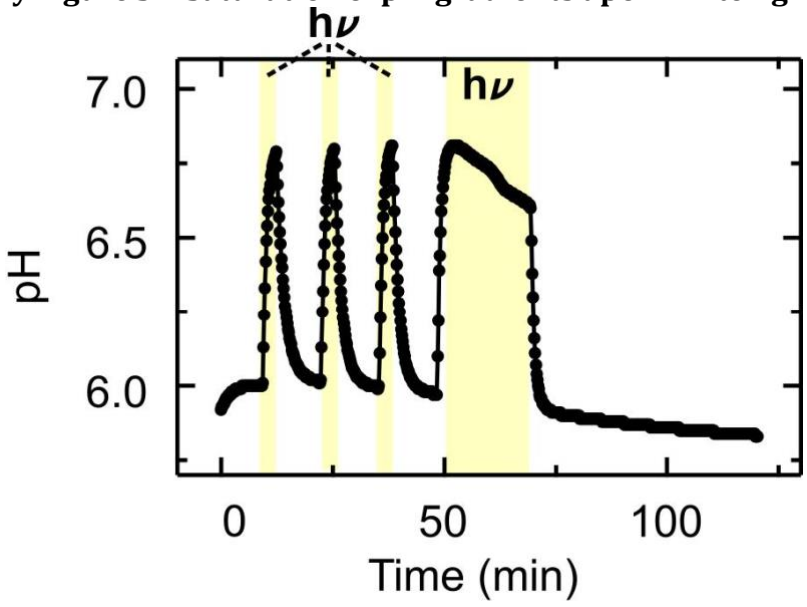


Figure 2: Engineered xenorhodopsin-expressing *E. coli* generate a pH-gradient upon illumination with white light, which saturates after ~5 min of illumination. Photoactivity measurements of engineered *E. coli* (at OD_{600}=20, in 150 mM NaCl) with a pH-electrode. The pH is plotted as a function of time during four light-dark cycles (periods of illumination are indicated in yellow). The pH increases by ~0.8 within 5 min of illumination and nearly returns to its original value after 10 min in the dark. A longer illumination time of 20 min (see final illumination period starting at t = 50 min) shows saturation of the pHgradient and a slow decrease of the pH during continuous illumination.

Supplementary Figure S3: Photoactivity measurements with different *E. coli* densities

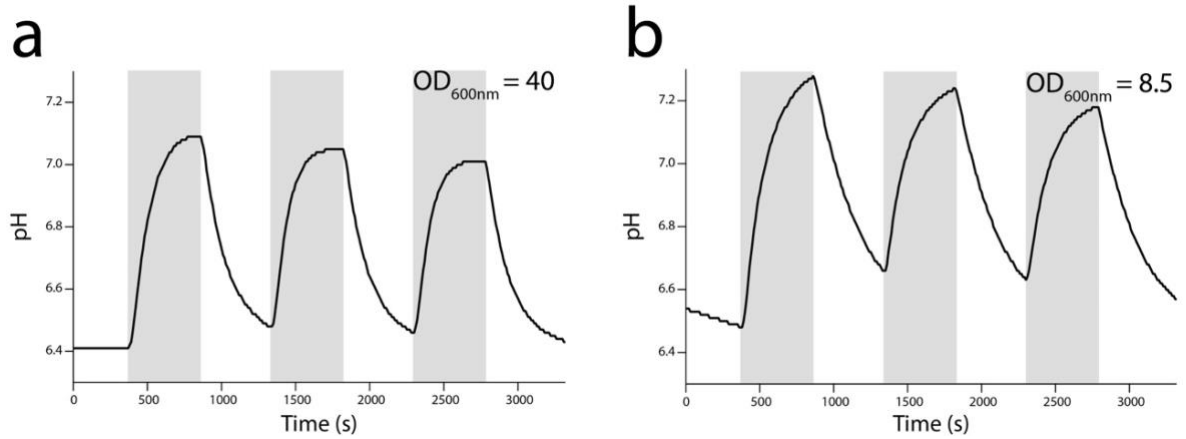


Figure 3: Photoactivity measurements with different *E. coli* densities. a Bulk photoactivity measurements with xenorhodopsin-GFP expressing *E. coli* at an OD_{600} of 40. b Bulk photoactivity measurements with the same XeR-GFP expressing *E. coli* as shown in a, diluted to an OD_{600} of 8.5. Illumination in both experiments was performed with a Schott, 1500 LCD lamp. The obtained pH gradients are very similar at both ODs, the kinetics are slightly increased at higher ODs. This could be explained by the hypothesis that the *E. coli* membrane itself cannot maintain pH gradients above one pH unit. Alternatively, it is conceivable that the light is partially absorbed by the denser *E. coli* solution.

Supplementary Figure S4: Chemical structure of pyranine

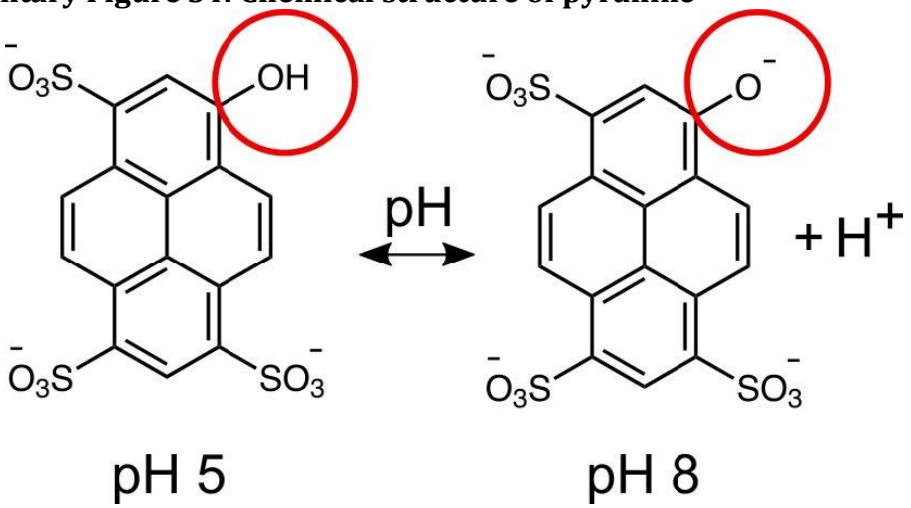


Figure 4: Chemical structure of pyranine at different pH values. The hydroxyl group of pyranine (indicated with a red circle) is deprotonated at high pH values leading to a change of the molecule's fluorescent properties (see Figure 1c). Therefore, pyranine can be used as pH-indicator by measuring the ratio of fluorescence emission upon excitation with the wavelengths 488nm and 405nm. A high ratio I_{488}/I_{405} indicates high pH values and low ratios a low pH value, respectively.

Supplementary Figure S5: Calibration curve of pyranine fluorescence in presence of GUVs

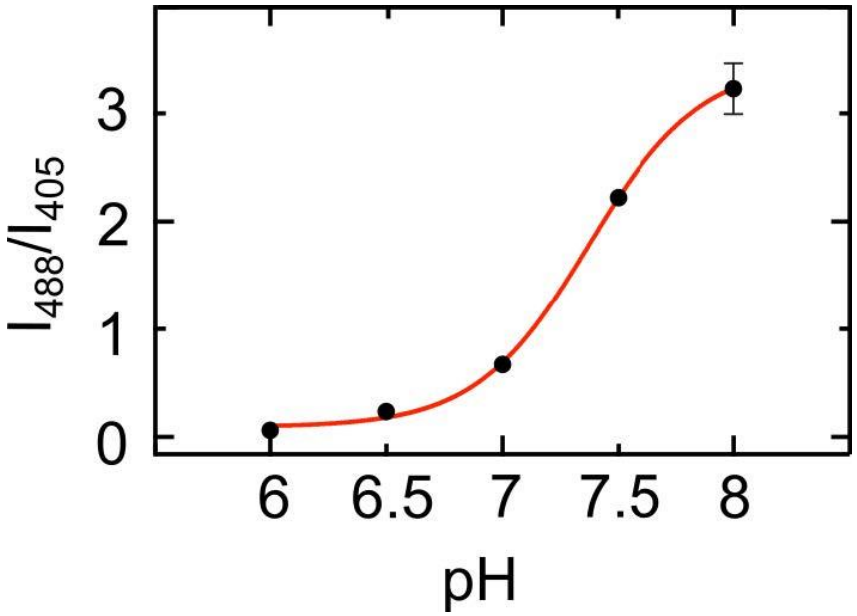


Figure 5: Calibration curve of pyranine fluorescence intensity (I_{488}/I_{405}) as a function of pH in a GUV-containing solution (150mM sucrose, 50 µM pyranine, 5mM $MgCl_2$, 100mM sodium phosphate buffer pre-adjusted to the respective pH value). A sigmoidal fit (red) shows a pK_a value of pH 7.37. This is in very good agreement with the calibration performed in droplet-based compartments (see Supplementary Figure S8) and literature values.[1] Note that the pyranine fluorescence is strongly dependent on the buffer conditions. Therefore, the absolute values cannot be compared to the measurements in an *E. coli*-containing solution.

Supplementary Figure S6: Monitoring pH changes with pyranine in the presence of GUVs

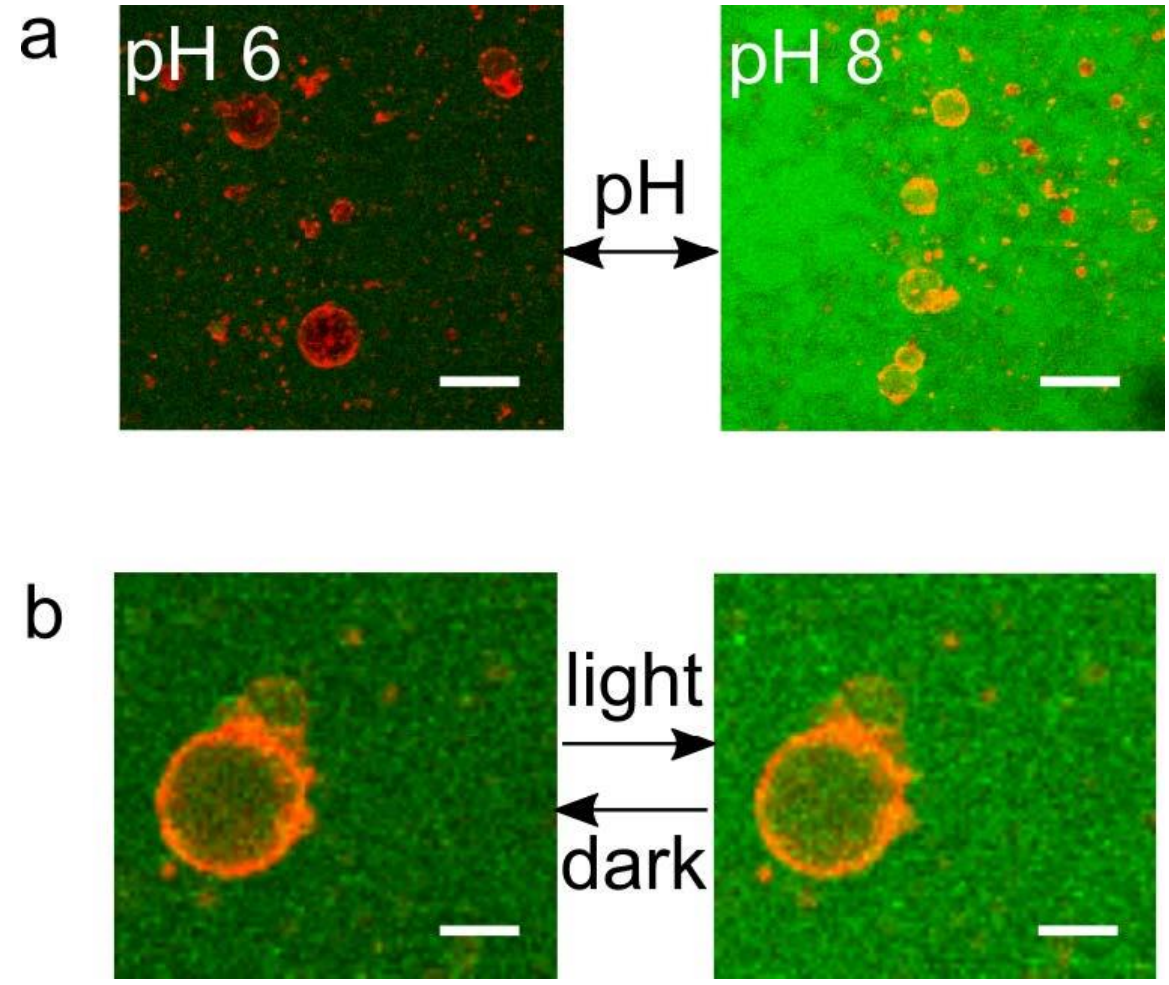


Figure 6: Monitoring pH changes with pyranine in the presence of GUVs. a Confocal fluorescence images of GUVs (red, λ_{ex}=647nm) and pyranine (green, λ_{ex}=488nm) at pH 6 and pH 8. The pyranine intensity upon 488nm excitation increases with increasing pH. Scale bar: 50 µm. b Confocal fluorescence images of GUVs (red, λ_{ex} =647nm) and pyranine (green, λ_{ex} =488nm) in presence of engineered *E. coli* before (left) and after (right) white light illumination. White light illumination leads to an increase of the pH due to the proton-pumping activity of the *E. coli*, which can be visualized by an increase in pyranine fluorescence emission. Scale bar: 20 µm.

Supplementary Figure S7: Layout of the microfluidic device for the formation of water-in-oil droplets

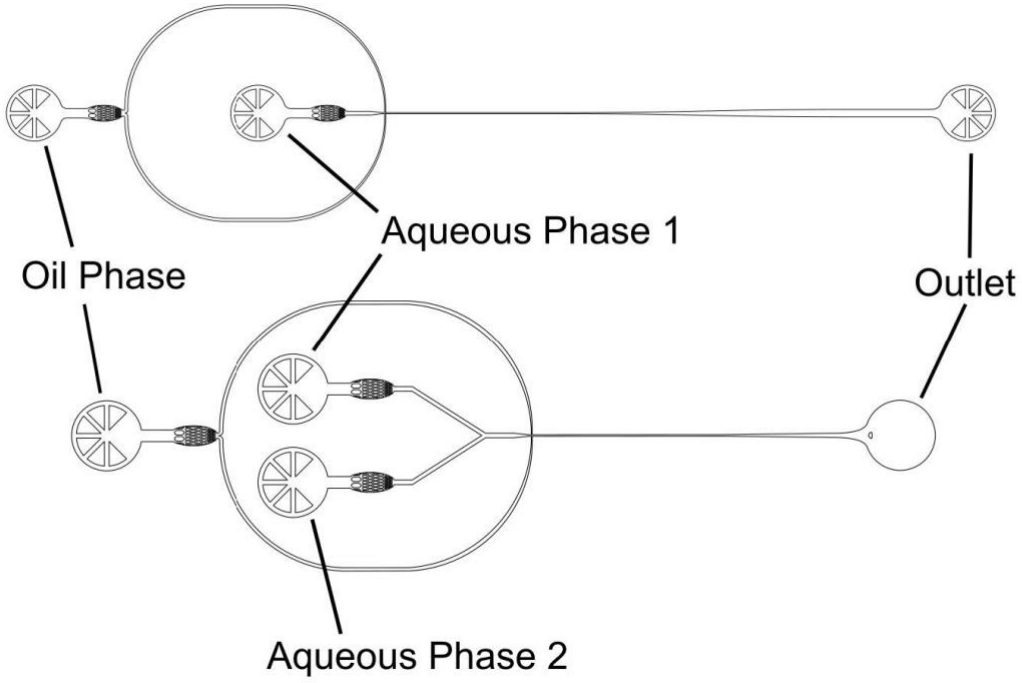


Figure 7: Layouts of single- (top) and two-inlet (bottom) microfluidic devices for the coencapsulation of DNA and *E. coli* into surfactant-stabilized water-in-oil droplets. The cholesterol-tagged and triplex-forming DNA were supplied via one inlet and the *E. coli* via the second one to avoid attachment of the cholesterol-tagged DNA to the *E. coli* prior to droplet formation. The microfluidic PDMS devices (Sylgard184, Dow Corning, USA) were fabricated according to a previously published protocol[2] (see Methods). For confocal fluorescence imaging, the droplets were collected from the outlet and sealed in a simple observation chamber as described previously.[3]

Supplementary Figure S8: Calibration curve of pyranine fluorescence in water-in-oil droplets

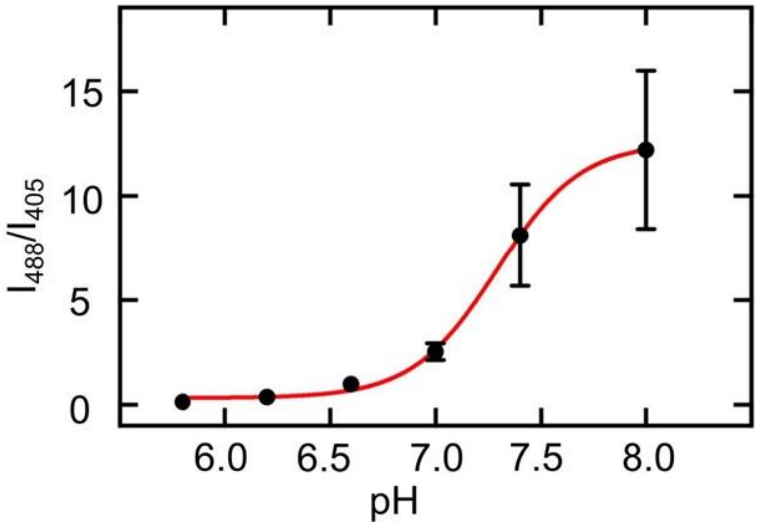


Figure 8: Calibration curve of pyranine fluorescence intensity I_{488}/I_{405} as a function of pH within water-in-oil droplets (50 µM pyranine, 5mM $MgCl_2$, 50mM sodium phosphate buffer pre-adjusted to the respective pH value). A sigmoidal fit (red) has a pK_a value of 7.29. This is in very good agreement with the calibration performed in presence of GUVs (see Supplementary Information Figure S5) and literature values.[1] Note that the pyranine fluorescence is strongly dependent on the buffer conditions. Therefore, the absolute values cannot be compared quantitatively to the measurements in *E. coli*-containing droplets. Error bars correspond to the standard deviation of n≥20 droplets.

Supplementary Figure S9: pH-sensitive attachment of triplex-forming DNA

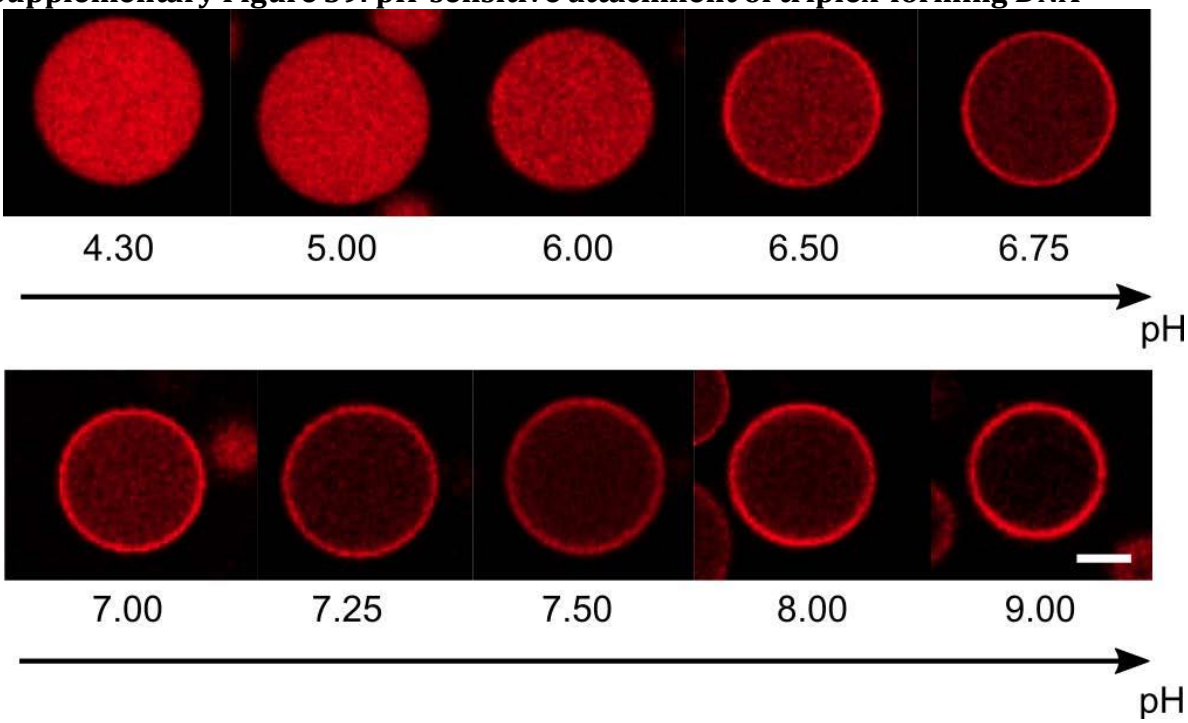


Figure 9: pH-sensitive attachment of triplex-forming DNA. Representative confocal fluorescence images of 1 µM Cy5-labeled triplex-forming DNA (red, λex =647nm) in droplet-based compartments at different pH values as indicated. The droplet periphery was functionalized with 1:5 µM cholesterol-tagged DNA (complementary to the hairpin region of the triplex, for DNA sequences see Materials and Methods). With increasing pH, Hoogsten interactions become weaker and an increasing amount of the triplex-forming DNA binds to the droplet periphery. For a quantitative plot of the fluorescence intensity inside the droplet at the different pH values, see Figure 3b (main text). Scale bar: 20 µm.

Supplementary Figure S10: Fluorophore-tagged single-stranded DNA does not interact with droplet-stabilizing surfactants

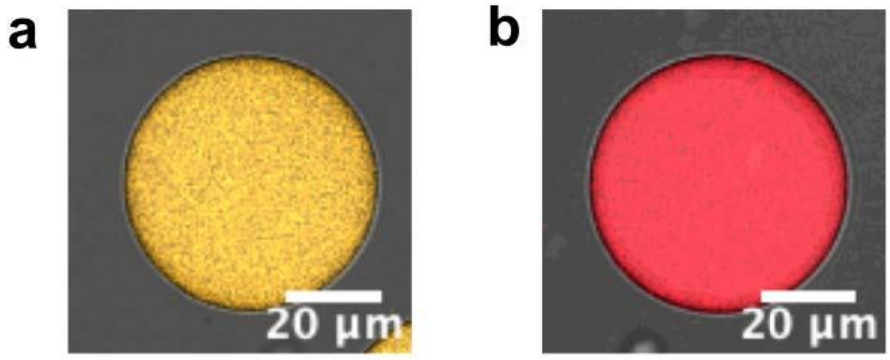


Figure 10: Fluorophore-tagged single-stranded DNA does not interact with droplet-stabilizing surfactants. Representative confocal images of water-in-oil droplets containing Cy3-(a, λ_{ex} = 561nm) and Cy5-labeled ssDNA (b, λ_{ex} = 647nm) without cholesterol-modification at pH8. The solution contained 20mM potassium phosphate buffer, 10mM $MgCl_2$ and 1:5 µM DNA. This confirms that there is no unspecific pH-dependent adsorption of the DNA to the droplet periphery at elevated pH. Scale bars: 20 µm.

Supplementary Figure S11: Brightfield and confocal images of microfluidic droplets containing engineered *E. coli*

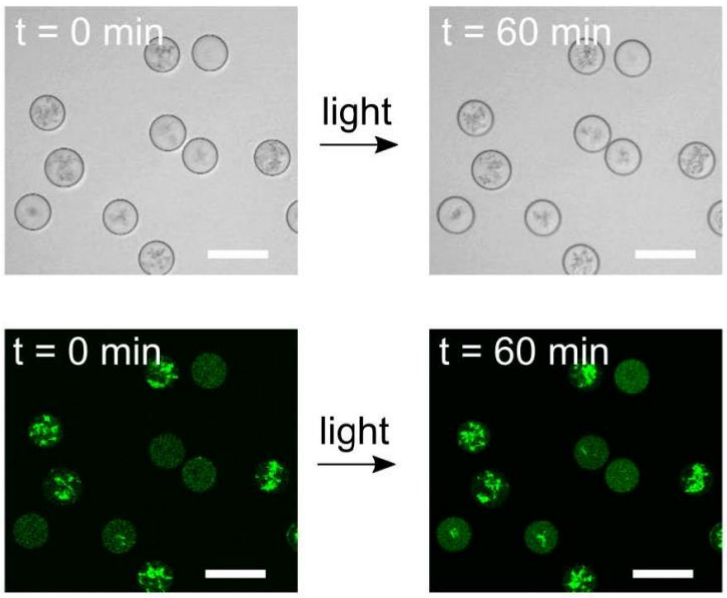


Figure 11: Brightfield (top) and confocal (bottom) images of microfluidic water-in-oil droplets containing the triplex-forming DNA, cholesterol-tagged DNA, pyranine and engineered *E. coli* before (0 min) and after (60 min) illumination with white light. The images confirm the presence of the *E. coli* inside the droplets and their stable confinement. Scale bars: 100 µm.

Supplementary Figure S12: Hysteresis of DNA triplex attachment and detachment

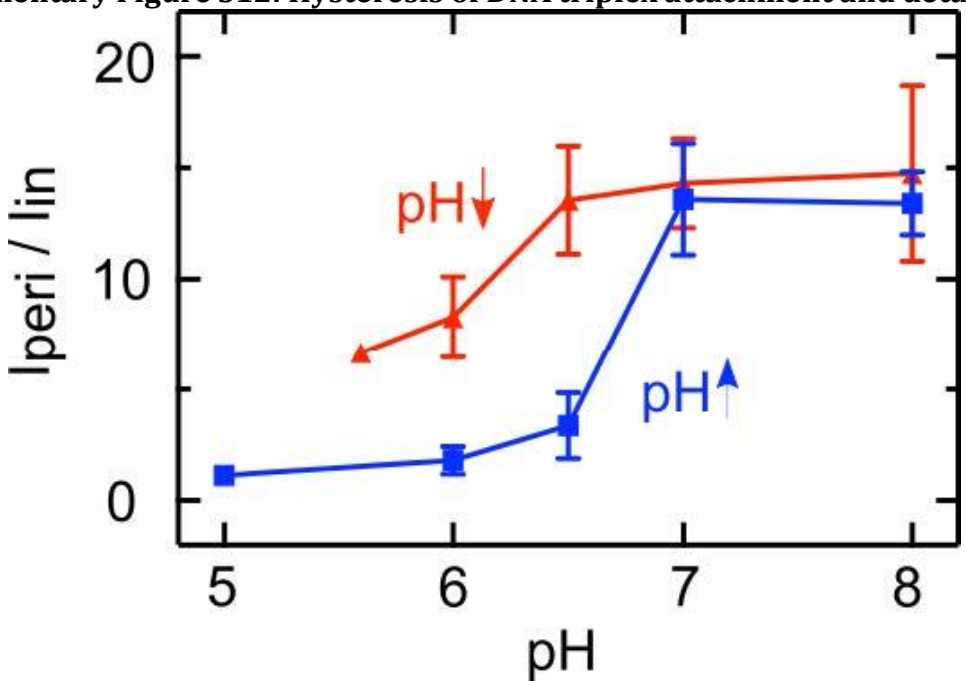


Figure 12: Hysteresis of DNA triplex attachment and detachment. Fluorescence intensity ratio I_{Peri}/I_{In} of the Cy5-labeled triplex-forming DNA strand at the droplet periphery over the droplet lumen at different pH values. The triplex-forming DNA was incubated with the complementary cholesterol-tagged strand for 10 min before encapsulation into droplets with 10mM sodium phosphate buffer and 20mM $MgCl_2$ at pH 5 (blue curve) or pH 8 (red curve). After incubation the solutions were mixed 1:1 with 200mM phosphate buffers ranging from pH 5 to 8. Droplets were then imaged with confocal fluorescence microscopy. The plot clearly indicates that the duplex dissociation happens at lower pH values compared to the duplex formation. This explains why the *E. coli* can induce attachment but not detachment of the triplex-forming DNA. Error bars correspond to the standard deviation of n≥15 droplets.

Supplementary Figure S13: Cadnano design of the membrane-sculpting DNA origami

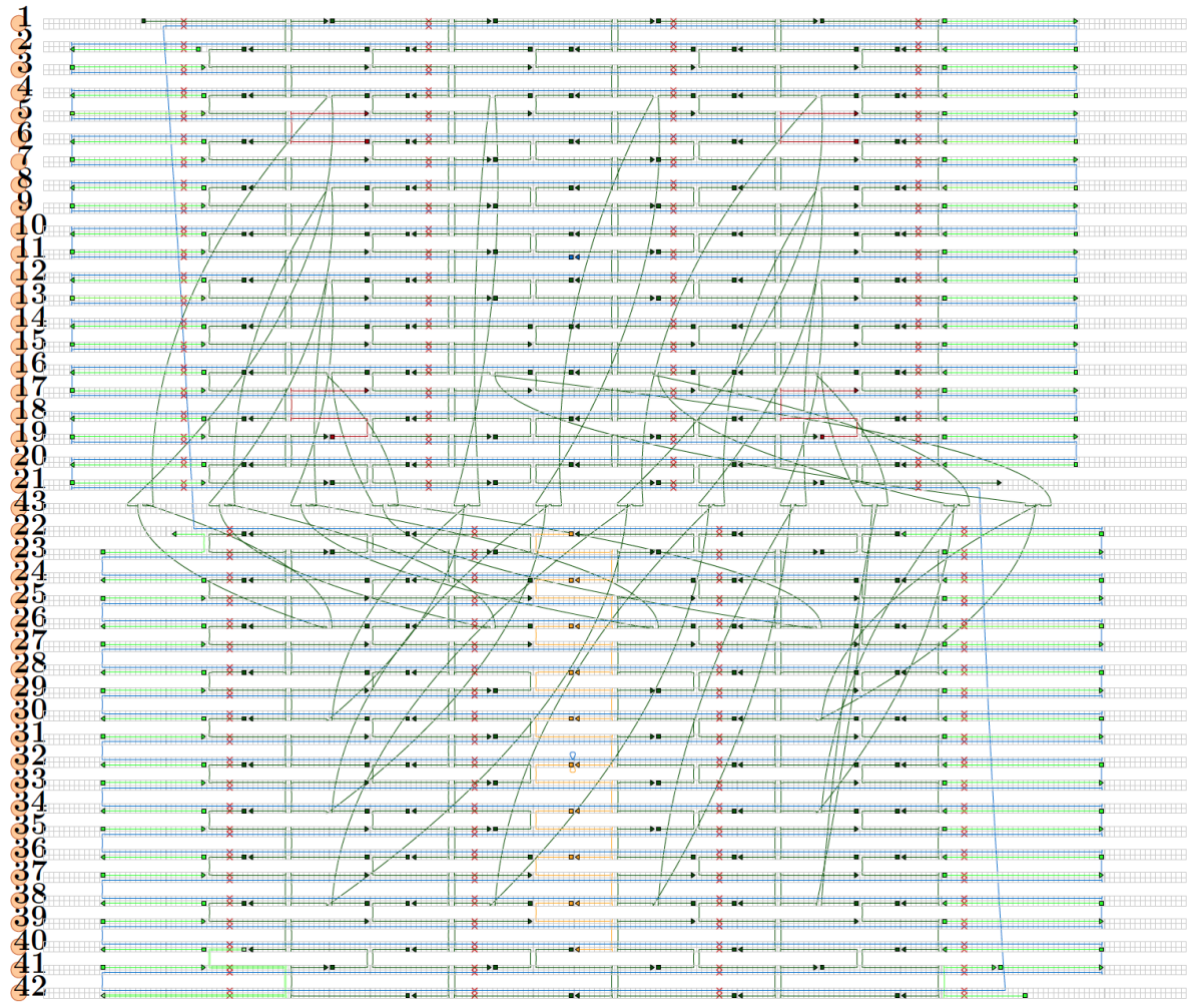


Figure 13: Cadnano design of the membrane-sculpting DNA origami. The scaffold (p8064) is shown in blue, bright green staples induce blunt-end stacking, red staples carry overhangs for the triplex-forming DNA on their 3' end, yellow staples carry overhangs for a complementary Cy3-tagged DNA strand on their 3' end.

Supplementary Figure S14: Blunt-end stacking induces polymerization of DNA origami plates

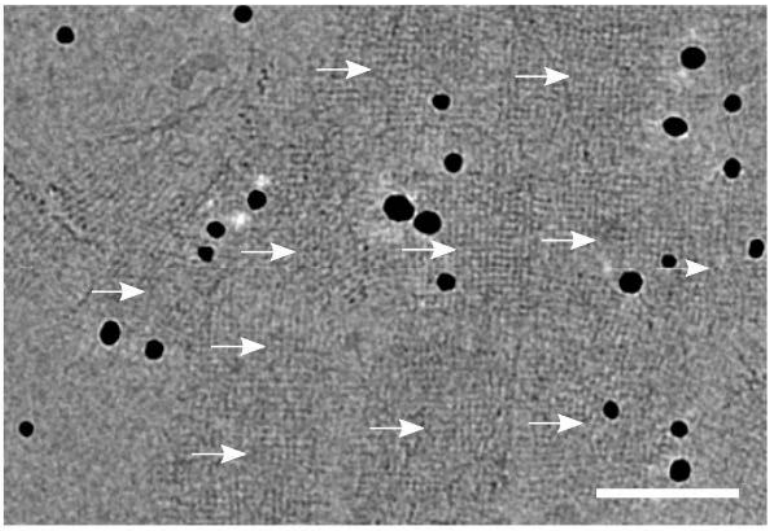


Figure 14: Cryo-EM micrograph of the polymerized DNA origami. The image depicts an arrangement of fused DNA origami squares (indicated by white arrows). Black spots correspond to gold fiducials. This arrangement was used to sculpt the membrane of GUVs. Scale bar: 50nm.

Supplementary Figure S15: Atomic force microscopy images of the DNA origami

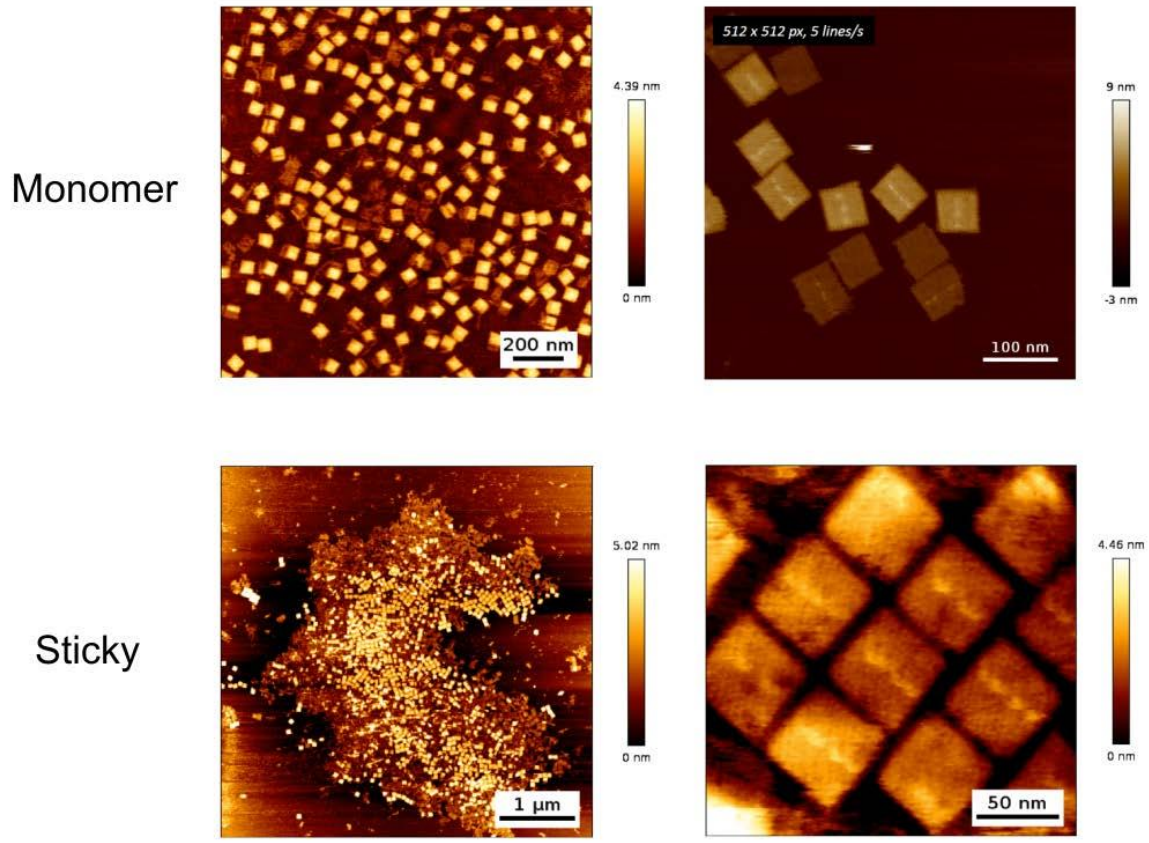


Figure 15: Atomic force microscopy images of the DNA origami without (monomer) and with the overhang strands (sticky). The sticky DNA origami shows a tight packing and cluster formation due to blunt-end stacking, whereas the DNA origami remain loosely distributed when the overhang strands were omitted (monomer). The line that appears across the DNA origami corresponds to the binding sites of the fluorophores. 1 nM of the DNA origami was added to a mica surface, incubated for 90 s and then washed with buffer. Atomic force microscopy was conducted in liquid using a Nanowizzard Ultra Speed 2 (Bruker).

Supplementary Figure S16: Agarose gel electrophoresis of the DNA origami

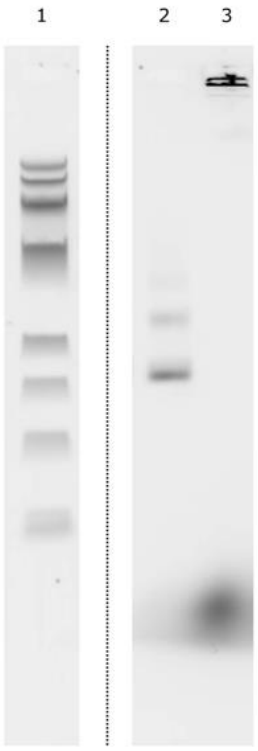


Figure 16: Agarose gel electrophoresis (0.7% agarose) of the DNA origami. Lane 1) 1 kbp DNA ladder; Lane 2) DNA origami without the staples at the scaffold seam (i.e. with single-stranded scaffold loops); Lane 3) DNA origami with the staples at the scaffold seam. Without the the staples at the scaffold seam (Lane 2), there is a clear band for the monomeric DNA origami and weaker bands from oligomers. The single-stranded scaffold loops prevent blunt-end stacking. In presence of the staples at the scaffold seam, blunt-end stacking occurs and the DNA origami does not leave the pocket due to its highly polymerized state. The gel was run at 60 V for 3.5 h at 4°C.

Supplementary Figure S17: Droplets are not deformed by attaching DNA origami to the droplet periphery

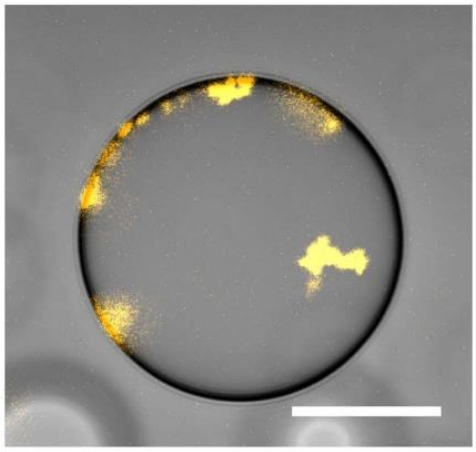


Figure 17: Droplets are not deformed by attaching DNA origami to the droplet periphery. Confocal image of a surfactant-stabilized water-in-oil droplet containing 10nM cholesterol-tagged Cy3-labeled DNA origami (λ_{ex}=561nm). The droplet remains spherical even though DNA origami clusters were successfully attached to the droplet periphery. Scale bar: 20 µm.

Supplementary Figure S18: FRAP experiments of GUVs with membrane-bound DNA origami

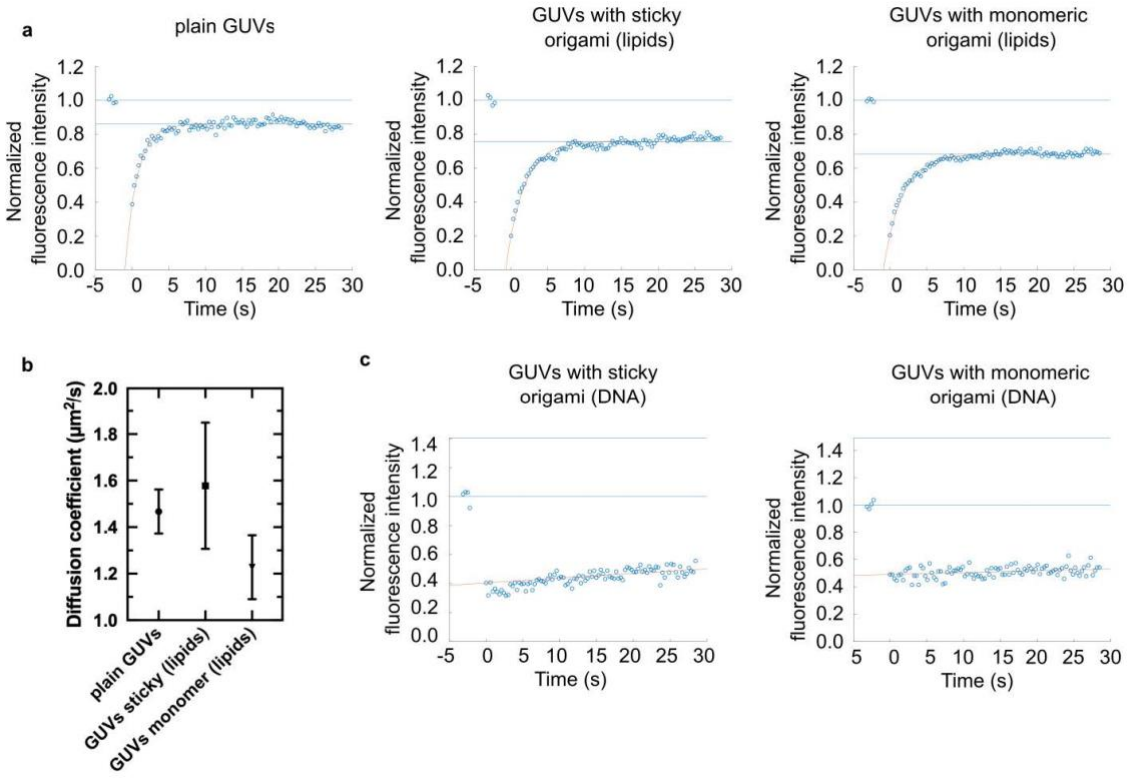


Figure 18: a FRAP of lipids (λ_{ex} =488nm). Exemplary normalized fluorescence recovery after photobleaching (FRAP) traces for plain GUVs, deformed GUVs with membrane-bound polymerized DNA origami (lipid recovery) and spherical GUVs with membrane-bound monomeric DNA origami (with single-stranded scaffold loops, lipid recovery). b Diffusion coefficients of lipids in absence and presence of DNA origami (mean ± SD; n=3 for each condition). The lipid diffusivity is not affected significantly by the presence of membrane-bound DNA origami. The diffusion coefficients were calculated according to previous works.[4] c FRAP of DNA origami (λ_{ex} =561nm). Exemplary normalized fluorescence recovery after photobleaching (FRAP) traces for deformed GUVs with membrane-bound polymerized DNA origami (DNA recovery) and spherical GUVs with membrane-bound monomeric DNA origami (with single-stranded scaffold loops, DNA recovery). The DNA origami do not recover after photobleaching as expected in the presence of divalent ions.[5]

Supplementary Figure S19: DNA origami cortex suppresses membrane Fluctuations

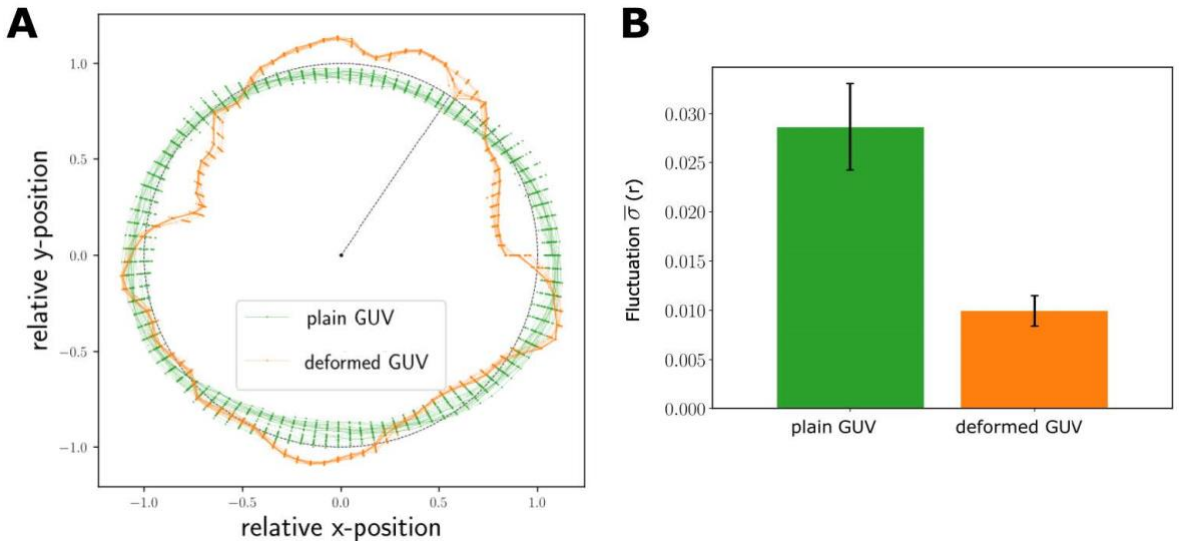


Figure 19: Membrane fluctuations of osmotically deflated GUVs (c/c_0 = 1:8) with and without membrane-bound DNA origami. A Outline of a plain GUV without membrane-bound DNA origami (green) and a deformed GUV with membrane-bound polymerized DNA origami (orange). The outline was traced from a confocal cross section over time (see also Supplementary Video S5). B Standard deviation of the radius r from the mean radius for plain (green) and deformed GUVs (n=5 individual GUVs tracked over time, mean ± std.). Higher deviations from the mean radius correspond to larger membrane fluctuations. For the deformed GUV, membrane fluctuations are approximately three times lower than for the plain GUV.

Supplementary Figure S20: Deformation process of GUVs with DNA origami

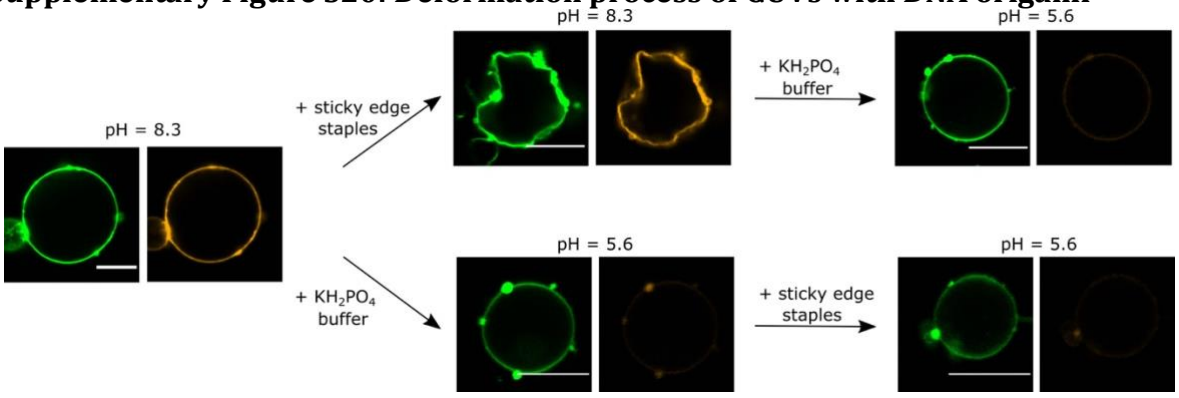


Figure 20: Flow diagram with confocal images of GUVs (lipids labelled with Atto488, λ_{ex}= 488nm) and DNA origami (labelled with Cy3, λ_{ex}= 561nm) depicting the deformation process. First of all, we attached the monomeric pH-sensitive DNA origami to the GUVs at pH 8.3 using cholesterol-tags which bind to the hairpin loop of the triplex motif at elevated pH. Subsequently, we added the sticky edge staples, which allow for blunt-end stacking of the DNA origami and thus induce polymerization of the DNA origami. Polymerization, in turn, deforms the GUVs. Following this step, the DNA origami were detached from the GUVs by lowering the pH to 5.6 (upper panels, the fluorescence from the detached DNA origami in the background is too weak to be visible). As a control, we also first lowered the pH to 5.6 and then added the overhang strands (lower panels). This did not affect the morphology of GUVs. After each mixing step, the GUVs were incubated for 24 h either with overhang strands or at a different pH value. Note that the deformation process takes about two hours. Scale bars: 15 µm.

Supplementary Figure S21: Confocal images of deformed GUVs

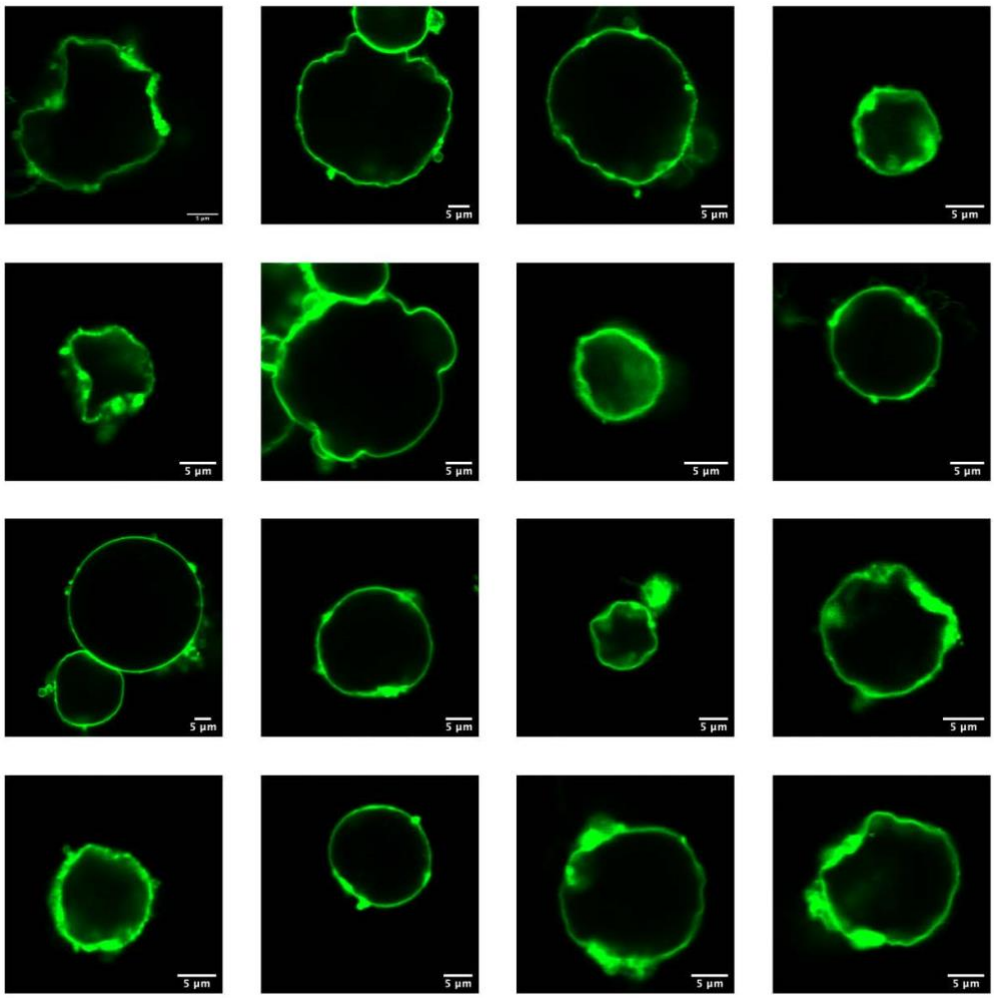


Figure 21: Confocal images of deformed GUVs (λ_{ex}= 561nm) in presence of membrane-bound polymerized DNA origami at pH 8.3. The corresponding quantification of the GUV circularity is shown in Figure 4e (main text). Scale bars: 5 µm.

Supplementary Figure S22: Confocal images of GUVs after detachment of membrane-bound DNA-origami

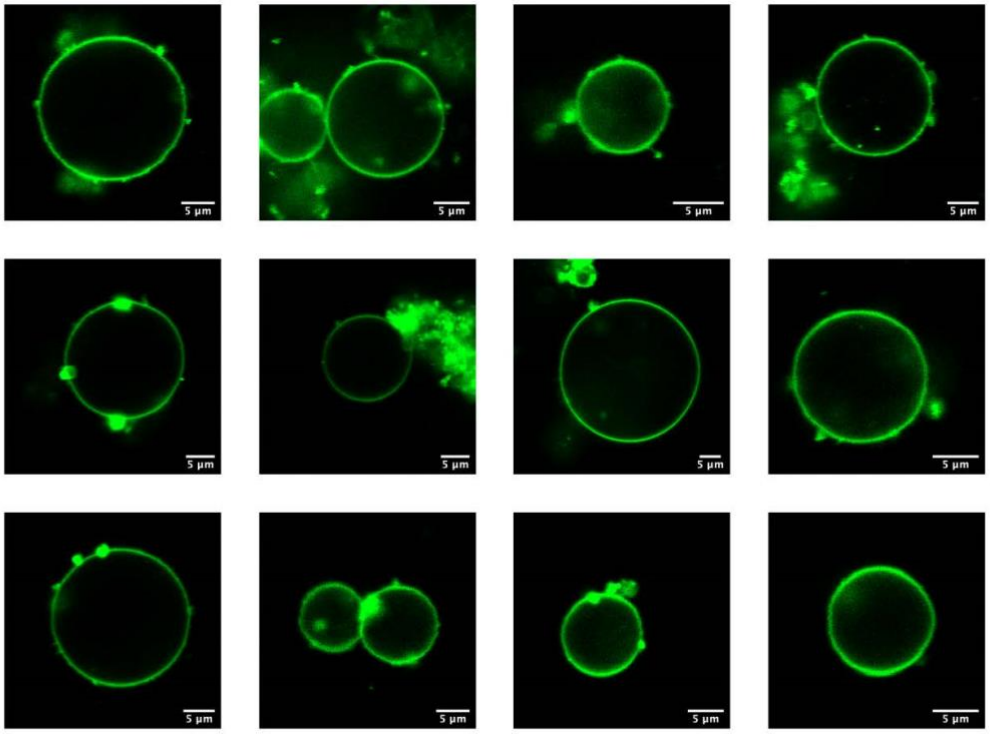


Figure 22: Confocal images of GUVs (λ_{ex}= 561nm) after decreasing the pH from pH 8.3 to pH 5.6 by addition of iso-osmotic potassium dihydrogenphosphate buffer. The DNA origami detaches from the GUV upon lowering the pH and the GUVs return to a spherical shape. The corresponding quantification of the GUV circularity is shown in Figure 4e (main text). Scale bars: 5 µm.

Supplementary Figure S23: Exemplary confocal images of DNA-functionalized GUVs

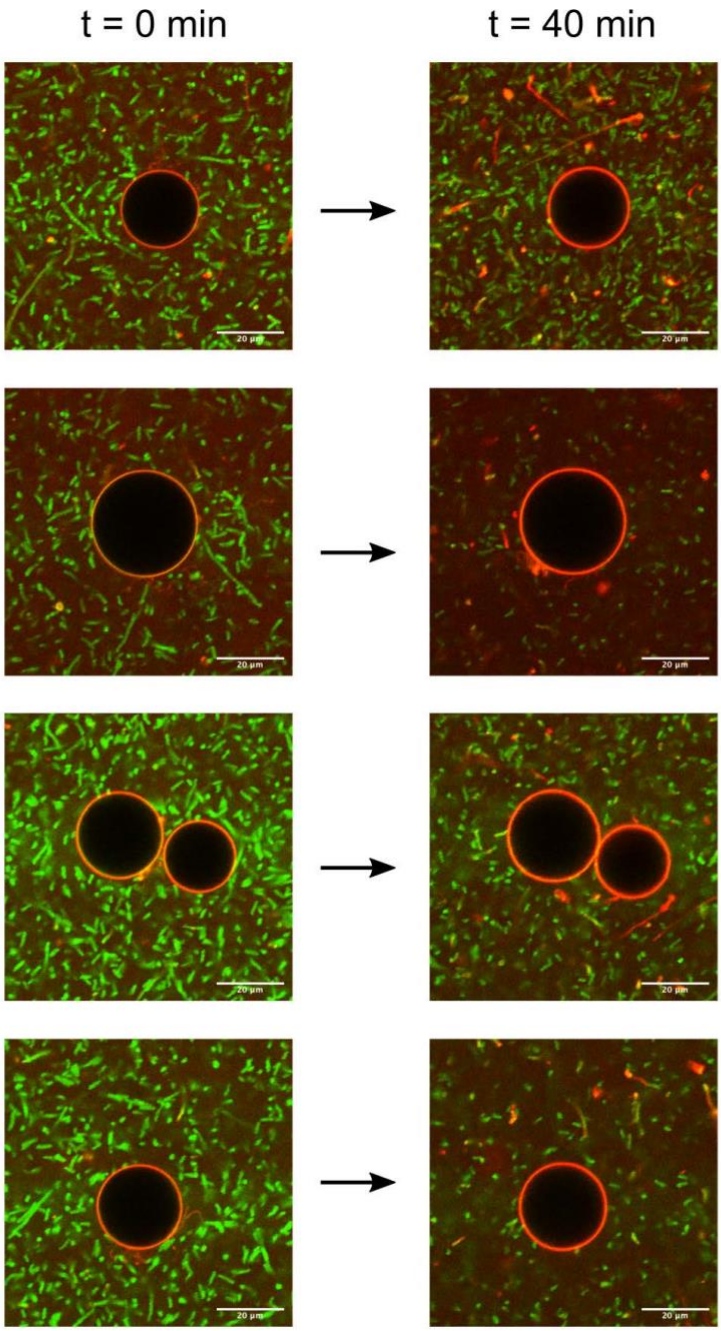


Figure 23: Exemplary confocal images of GUVs coated with cholesterol-tagged DNA (0.6 µM) and surrounded by *E. coli* and triplex-forming DNA (0.4 µM, λ_{ex}=561nm). The images show the attachment of the triplex-forming DNA after illumination (40 min time point) as well as the settling of *E. coli*. The droplet was illuminated for 15 min after 25 min in the dark. The corresponding quantification of the peripheral DNA intensity is shown in Figure S24. Scale bars: 20 µm.

Supplementary Figure S24: Attachment of the single-stranded DNA triplex to GUVs during light illumination

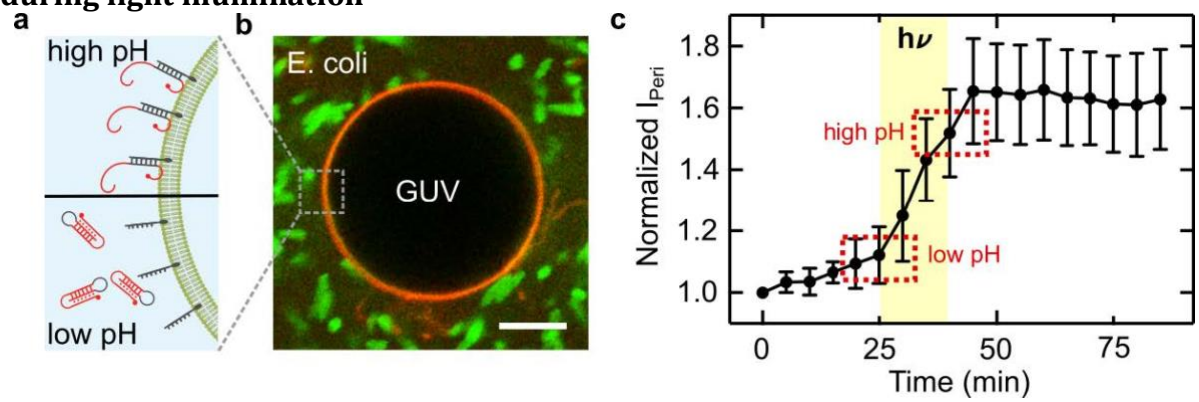


Figure 24: a Schematic illustration of a GUV membrane functionalized with cholesterol-tagged DNA in presence of triplex-forming DNA at high and low pH values. b Confocal image of a DNA-coated GUV surrounded by *E. coli* as described in a (0.4 µM triplex-forming DNA, λ_{ex}=561nm; 0.6 µM cholesterol-tagged DNA). Scale bar: 10 µm. c Normalized fluorescence intensity I_{peri}(mean ± s.d., n=15) of the triplex-forming DNA at the GUV periphery monitored over time. The time period of illumination is indicated in yellow, illumination leads to a pH increase and hence DNA attachment. The data is extracted from GUVs as shown in Figure S23.

Supplementary Figure S25: Confocal images of deformed GUVs after light-mediated attachment of DNA origami

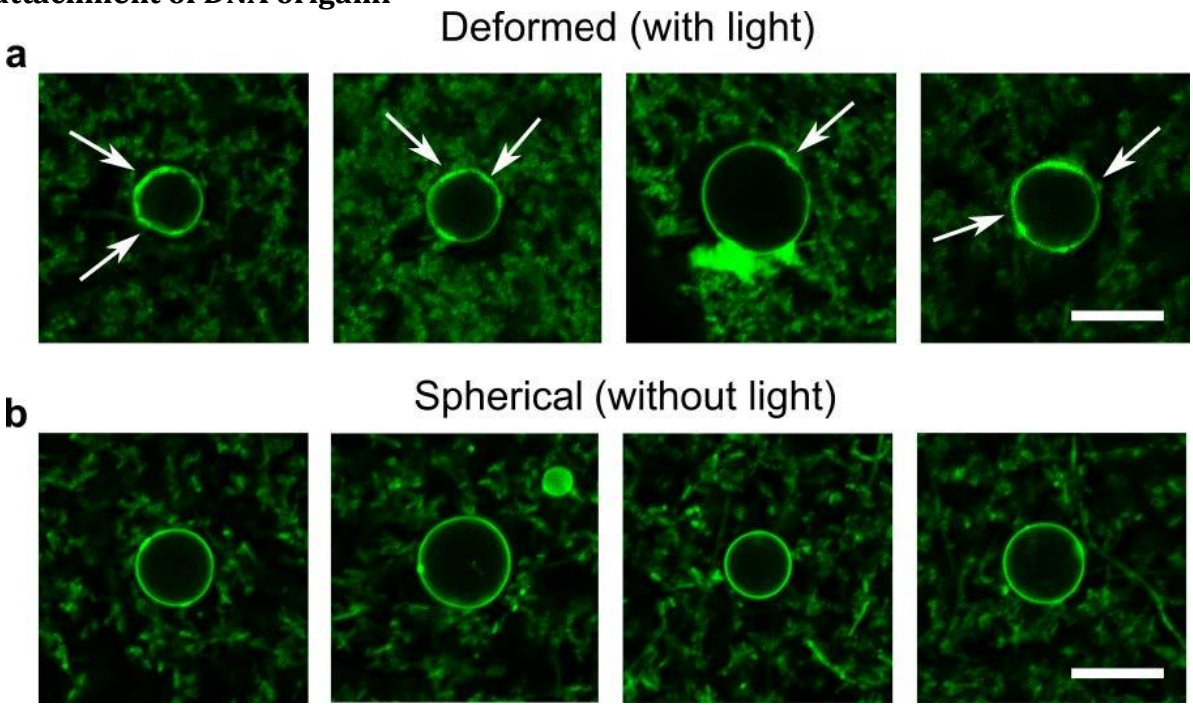


Figure 25: Confocal images of deformed GUVs (λ_{ex}= 488nm) in presence of membrane-bound polymerized DNA origami and *E. coli* after light-mediated attachment of DNA origami and spherical GUVs that without light illumination. Note that the deformation is weaker compared to the deformation achieved with conventional pH switching due to the smaller pH gradient. Scale bars: 10 µm.

Supplementary Table S1: DNA and amino-acid sequences of xenorhodopsin-constructs

XeR-GFP DNA sequence:

1 ATGGTGTATG AAGCAATTAC CGCAGGCGGT TTTGGTAGCC AGCCGTTTAT TCTGGCATAT
61 ATCATTACCG CAATGATTAG CGGTCTGCTG TTTCTGTATC TGCCTCGTAA ACTGGATGTT
121 CCGCAGAAAT TTGGCATCAT CCATTTTTTC ATTGTGGTTT GGAGCGGTCT GATGTATACC
181 AATTTTCTGA ATCAGAGCTT CCTGAGCGAT TATGCATGGT ATATGGATTG GATGGTTAGC
241 ACACCGCTGA TTCTGCTGGC ACTGGGTCTG ACCGCATTTC ATGGTGCAGA TACCAAACGT
301 TATGATCTGC TGGGTGCACT GCTGGGAGCA GAATTTACCC TGGTTATTAC AGGTCTGCTG
361 GCCCAGGCAC AGGGTAGCAT TACCCCGTAT TATGTTGGTG TTCTGCTGCT GCTGGGCGTT
421 GTTTATCTGC TGGCGAAACC GTTTCGTGAA ATTGCCGAAG AAAGCAGTGA TGGTCTGGCA
481 CGTGCGTATA AAATCCTGGC AGGTTATATT GGCATCTTTT TTCTGAGCTA TCCGACCGTG
541 TGGTATATTA GCGGTATTGA TGCACTGCCT GGTAGCCTGA ATATTCTGGA CCCGACCCAG
601 ACCAGCATTG CACTGGTTGT TCTGCCGTTT TTTTGCAAAC AGGTTTATGG CTTCCTGGAC
661 ATGTATCTGA TTCATAAAGC AGAAGCTCTC GAGGGAGGAA GTCTGGAAGT TCTGTTCCAG
721 GGGCCCGTCG ACGGATCCGA AAACTTGTAT TTCCAGGGCA TGAGTAAAGG AGAAGAACTT
781 TTCACTGGAG TTGTCCCAAT TCTTGTTGAA TTAGATGGTG ATGTTAATGG GCACAAATTT
841 TCTGTCCGTG GAGAGGGTGA AGGTGATGCT ACAAACGGAA AACTCACCCT TAAATTTATT
901 TGCACTACTG GAAAACTACC TGTTCCGTGG CCAACACTTG TCACTACTCT GACCTATGGT
961 GTTCAATGCT TTTCCCGTTA TCCGGATCAC ATGAAACGGC ATGACTTTTT CAAGAGTGCC
1021 ATGCCCGAAG GTTATGTACA GGAACGCACT ATATCTTTCA AAGATGACGG GACCTACAAG
1081 ACGCGTGCTG AAGTCAAGTT TGAAGGTGAT ACCCTTGTTA ATCGTATCGA GTTAAAGGGT
1141 ATTGATTTTA AAGAAGATGG AAACATTCTT GGACACAAAC TCGAGTACAA CTTTAACTCA
1201 CACAATGTAT ACATCACGGC AGACAAACAA AAGAATGGAA TCAAAGCTAA CTTCAAAATT
1261 CGCCACAACG TTGAAGATGG TTCCGTTCAA CTAGCAGACC ATTATCAACA AAATACTCCA
1321 ATTGGCGATG GCCCTGTCCT TTTACCAGAC AACCATTACC TGTCGACACA ATCTGTCCTT
1381 TCGAAAGATC CCAACGAAAA GCGTGACCAC ATGGTCCTTC TTGAGTTTGT AACTGCTGCT
1441 GGGATTACAC ATGGCATGGA TGAGCTCTAC AAAGGAGGAT CTGGTGGTTC TGGGAAGCTT
1501 GCGGCCGCAC TCGAGCACCA CCACCACCAC CACTGA

XeR-GFP amino-acid sequence:

MVYEAITAGGFGSQPFILAYIITAMISGLLFLYLPRKLDVPQKFGIIHFFIVVWSGLMYTNFLNQSFLSD
YAWYMDWMVSTPLILLALGLTAFHGADTKRYDLLGALLGAEFTLVITGLLAQAQGSITPYYVGVLLLLGV
VYLLAKPFREIAEESSDGLARAYKILAGYIGIFFLSYPTVWYISGIDALPGSLNILDPTQTSIALVVLPF
FCKQVYGFLDMYLIHKAEALEGGSLEVLFQGPVDGSENLYFQGMSKGEELFTGVVPILVELDGDVNGHKF
SVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERT
ISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKI
RHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMDELY
KGGSGGSGKLAAALEHHHHHH

NsXeR Sf-GFP His6-Tag

XeR-mCherry DNA sequence:

1 ATGGTGTATG AAGCAATTAC CGCAGGCGGT TTTGGTAGCC AGCCGTTTAT TCTGGCATAT
61 ATCATTACCG CAATGATTAG CGGTCTGCTG TTTCTGTATC TGCCTCGTAA ACTGGATGTT
121 CCGCAGAAAT TTGGCATCAT CCATTTTTTC ATTGTGGTTT GGAGCGGTCT GATGTATACC
181 AATTTTCTGA ATCAGAGCTT CCTGAGCGAT TATGCATGGT ATATGGATTG GATGGTTAGC
241 ACACCGCTGA TTCTGCTGGC ACTGGGTCTG ACCGCATTTC ATGGTGCAGA TACCAAACGT
301 TATGATCTGC TGGGTGCACT GCTGGGAGCA GAATTTACCC TGGTTATTAC AGGTCTGCTG
361 GCCCAGGCAC AGGGTAGCAT TACCCCGTAT TATGTTGGTG TTCTGCTGCT GCTGGGCGTT
421 GTTTATCTGC TGGCGAAACC GTTTCGTGAA ATTGCCGAAG AAAGCAGTGA TGGTCTGGCA
481 CGTGCGTATA AAATCCTGGC AGGTTATATT GGCATCTTTT TTCTGAGCTA TCCGACCGTG
541 TGGTATATTA GCGGTATTGA TGCACTGCCT GGTAGCCTGA ATATTCTGGA CCCGACCCAG
601 ACCAGCATTG CACTGGTTGT TCTGCCGTTT TTTTGCAAAC AGGTTTATGG CTTCCTGGAC
661 ATGTATCTGA TTCATAAAGC AGAAGCTCTC GAGGGAGGAA GTCTGGAAGT TCTGTTCCAG
721 GGGCCCGTCG ACGGATCCAT GCATAGCAAG GGCGAGGAGG ATAACATGGC CATCATCAAG
781 GAGTTCATGC GCTTCAAGGT GCACATGGAG GGCTCCGTGA ACGGCCACGA GTTCGAGATC
841 GAGGGCGAGG GCGAGGGCCG CCCCTACGAG GGCACCCAGA CCGCCAAGCT GAAGGTGACC
901 AAGGGTGGCC CCCTGCCCTT CGCCTGGGAC ATCCTGTCCC CTCAGTTCAT GTACGGCTCC
961 AAGGCCTACG TGAAGCACCC CGCCGACATC CCCGACTACT TGAAGCTGTC CTTCCCCGAG
1021 GGCTTCAAGT GGGAGCGCGT GATGAACTTC GAGGACGGCG GCGTGGTGAC CGTGACCCAG
1081 GACTCCTCCT TGCAGGACGG CGAGTTCATC TACAAGGTGA AGCTGCGCGG CACCAACTTC
1141 CCCTCCGACG GCCCCGTAAT GCAGAAGAAG ACCATGGGCT GGGAGGCCTC CTCCGAGCGG
1201 ATGTACCCCG AGGACGGCGC CCTGAAGGGC GAGATCAAGC AGAGGCTGAA GCTGAAGGAC
1261 GGCGGCCACT ACGACGCTGA GGTCAAGACC ACCTACAAGG CCAAGAAGCC CGTGCAGCTG
1321 CCCGGCGCCT ACAACGTCAA CATCAAGTTG GACATCACCT CCCACAACGA GGACTACACC
1381 ATCGTGGAAC AGTACGAACG CGCCGAGGGC CGCCACTCCA CCGGCGGCAT GGACGAGCTG
1441 TACAAGAAGC TTGCGGCCGC ACTCGAGCAC CACCACCACC ACCACTGA

XeR-mCherry amino-acid sequence:

MVYEAITAGGFGSQPFILAYIITAMISGLLFLYLPRKLDVPQKFGIIHFFIVVWSGLMYTNFLNQSFLSD
YAWYMDWMVSTPLILLALGLTAFHGADTKRYDLLGALLGAEFTLVITGLLAQAQGSITPYYVGVLLLLGV
VYLLAKPFREIAEESSDGLARAYKILAGYIGIFFLSYPTVWYISGIDALPGSLNILDPTQTSIALVVLPF
FCKQVYGFLDMYLIHKAEALEGGSLEVLFQGPVDGSMHSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEI
EGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNF
EDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKLAAALE
HHHHHH

NsXeR mCherry His6-Tag

Supplementary Table S3: List of DNA sequences for blunt-end stacking between the DNA origami plates

B_35[120]_33[119]	TGACAGGAGGGTCAGTGCCTTGAGGAGGTTTA
B_38[127]_11[159]	ATAATAACTTTGAAAGCGATTTTTTGGGGTCGAGG
B_16[71]_18[72]	AATAGATTCATTATCATTTTGCGGATTCGC
B_42[135]_41[119]	ACAAAAGGTAAAGTAATTCTGTCGCCGTTTT
B_42[167]_41[151]	TCGAGCCAGTAATAAGAGAATATACCGCGCCC
B_8[159]_37[95]	TGTGCACTTTTTCCTTATAAAAGGGCGACATTC
B_36[167]_38[168]	TGAAACCAGTCACCGACTTGAGCCAAGCAGAT
B_18[71]_20[72]	CTGATTGCCATTTGAATTACCTTAGACAAA
B_24[167]_26[168]	GCTGTAGCGTCAGGATTAGAGAGTAAACCAAA
B_36[159]_35[159]	TCGATAGCAGCACCGTGACGATTGGCCTTGAT
B_22[39]_24[40]	ATTTTTGAAAGGTGGCATCAATATTTCGCA
B_10[167]_12[168]	TTTTCTTTTGCCGTAAAGCACTAAACGTGG
B_18[135]_20[136]	AACGTCAGTGCTTCTGTAAATCGTTTAACCTC
B_20[135]_21[151]	CGGCTTAGAACGCTCAACAGTAGGGCTTAATT
B_24[95]_23[87]	GATTCCCAATTCTGCAATCATA
B_34[159]_15[127]	TACAGGAGTTTGGAATTGAACCACCAGCAGAAGA
B_4[63]_25[63]	ACAATCGGTTTCCTCAAATGAAGCAAAGCGGATTG
B_10[63]_9[63]	AAAATCCTGTTTGATGCCACACAACATACGAG
B_8[63]_25[95]	CGTGAGCCTTTGCGGAATAGGAAGCCCGAAAGA
B_10[159]_9[159]	TCACCAGTGAGACGGGCTTTCCAGTCGGGAAA
B_8[135]_10[136]	AGAATGCGTTGCGCTCACTGCCCGCAACAGCT
B_30[135]_32[136]	CGGCTACGATACCGATAGTTGCCAGACAGC
B_2[159]_1[151]	AACCAATAGGAACGCCCCCCAAAA
B_12[63]_25[127]	CGCTACAGTTTTAAAATGTTTAATTCGAGCTTCAA
B_40[135]_42[136]	AGCTACATCGTAGGAATCATTAAAGTACCG
B_12[39]_14[40]	GCTTTGACAGGCCGATTAAAGGGAAAAAGGGA
B_2[39]_4[40]	ATTGACCGCCAGTTTGAGGGGACGACAGCGCC
B_7[120]_5[127]	ATAAACACGGACTTGTAGAACGACATCGACATAAAAA
B_14[127]_13[119]	AATACCTACATTTTGGTAAAAGA
B_14[159]_13[159]	CCAGCCATTGCAACAGCGTTGTAGCAATACTT
B_26[127]_11[63]	GGTAATAGTTTGGCGCGTAGTCCACTATTAAAGAA
B_31[88]_29[87]	TTTATCAGGGAAGTTTCCATTAATCGCCTG
B_10[135]_12[136]	GATTGCCCCATCACCCAAATCAAGAAGGAGCG
B_6[135]_8[136]	GAACGTGCTCCCTTACACTGGTGTCTGCGGCC
B_28[39]_30[40]	TTCAACTAGCGATTATACCAAGACCAACCT
B_14[167]_16[168]	TATTACCGCGCCTGCAACAGTGCCAAATAT
B_26[63]_3[63]	TGAATCCCTTTCGAAACGTACGACAGTATCGGCCT

B_26[159]_15[63]	AGGCTTTTTTTCAATAGATGCTATTAGTCTTTAAT
B_23[152]_22[136]	TGTACCAAGCCTGAGTAATGTGTAGGTAAAGA
B_28[71]_30[72]	CGAGTAGTACGGAGATTTGTATCAACGGGTAA
B_16[135]_18[136]	AATCAACATTATCAGATGATGGCATCAGGTTT
B_12[71]_14[72]	TAATGCGCTACGCCAGAATCCTGTGGCAGA
B_33[88]_31[87]	GAATAGGTGTAAATGAATTTTCTGTATCGG
B_24[135]_26[136]	AAAGTACAGCGAACCAGACCGGAGTTTTGC
B_18[39]_20[40]	CGCAGAGGAAGAAAACAAAATTAAATATTTTA
B_2[167]_4[168]	CATTTTTTGGAAGGGCGATCGGTATTAAGT
B_38[159]_15[159]	AAAGTTACTTTATCTGGTCCGGTCAGTATTAACAC
B_23[88]_22[72]	CAGGCAAGATCAATATGATATTCAACCGTTC
B_42[103]_41[87]	GACAATAAACAACATGTTCAGCTACCAAGT
B_2[103]_4[104]	TTCATCAACACCGCTTCTGGTGCCAGCTTTCA
B_20[167]_21[187]	TCATAGGTCCATATTTAACAACGCCAACATGTTTT
B_26[135]_28[136]	CAGAGGGAGATTTAGGAATACCATTACCCA
B_20[103]_21[119]	AATGCTGAGCGTTATACAAATTCTTACCAGTA
B_29[120]_27[127]	TGTTACTTGTAACAAAGCTGCTCAAAAGATTCATCAGTTG
B_6[103]_8[104]	GTGCTGGTTGGGTAAAGGTTTCTTGGTCATAC
B_40[103]_42[104]	CTAACGAGCATCGAGAACAAGCAACAGACGAC
B_38[167]_40[168]	AGCCGAACTGAGTTAAGCCCAATATTTGAAGC
B_16[103]_18[104]	ATCTAAAAAGAAGGAGCGGAATTACAGTACCT
B_18[103]_20[104]	TTTACATCACATAAATCAATATATTATATGTA
B_24[71]_26[72]	AGATTTAGCATCAAAAAGATTAAGCGTCATAA
B_13[120]_11[119]	GTCTGTCGGCGCTGGCAAGTGTGCCCACTA
B_6[167]_8[168]	GCAGCCTCACGGCATCAGATGCCGTGTCAC
B_1[56]_2[72]	AGTCTGGAGCAAACAAGAGAATCGTAACAA
B_28[135]_30[136]	AATCAACAGCCGGAACGAGGCGGGTAGCAA
B_10[71]_12[72]	AGCAGGCGCGTGGACTCCAACGTCCGCGCT
B_38[135]_40[136]	AAACGCACGCTAATATCAGAGATTGCACCC
B_19[88]_17[95]	GAAACAGTGGGAGAAACAATAACGGAACAAAGAAACCACC
B_1[88]_2[104]	GTAATCGTAAAACTAGCATGTCAAAGCCAGCT
B_16[39]_18[40]	TGAGGATTCCGAACGTTATTAATTCAAAATCG
B_2[95]_1[87]	CATTAAATGTGAGCGAGATGAACG
B_21[56]_19[55]	GAATCATAGAAAACTTTTTCAAATTTACATTT
B_34[71]_36[72]	TTCGGAACGAGCCACCACCCTCAGTTTTCGGT
B_32[71]_34[72]	TTTTGCTAAGGGTTGATATAAGTACTGCCTAT
B_34[63]_3[127]	CTATTATTTTTCAGTGCCAGGAAACCAGGCAAAG
B_8[71]_10[72]	TTCGCGTCCCGGAAGCATAAAGTTTGCCCC
B_41[120]_39[127]	TATTTTCAATTTTATCCTGAATCTTCAGAGGGTAATTGAG
B_10[103]_12[104]	TTGCAGCAGTCTATCAGGGCGATGAGCGGTCA
B_1[120]_2[136]	ACCCCGGTTGATAATCAGAAAAGATCAAAAA

B_6[71]_8[72]	TGCCAACGAGCCGGGTCACTGTTCCTGTTC
B_40[39]_42[32]	CCAATCCATCAATAATCGGCTGTAAGTCCTGAACAAG
Ochol_6[159]_5[159]	CGGCCAGAGCACATCCCCGCACAGGCGGCCTTTTCCTTCTATGCATCA
Ochol_19[56]_17[63]	AACAATTTTTTGAATACCAAGTTATTAAAAGTTTGAGTTTCCTTCTATGCATCA
Ochol_19[152]_17[159]	ATTAATTTGAAATTGCGTAGATTTATTCATCAATATAATTCCTTCTATGCATCA
Ochol_6[63]_5[63]	GCAGCACCGTCGGTGGGCAGAAACAGCGGATTCCTTCTATGCATCA
Otop_30[103]_32[104]	TTTCATGACTTGCTTTCGAGGTGAAAAGTTTTTTTTTTTTTTTTTTTTTTT
Otop_22[103]_24[104]	AAATCACCGCAAAGAATTAGCAAAATTCCATTTTTTTTTTTTTTTTTTTTT
Otop_28[103]_30[104]	ATAAGGCTTGTCGAAATCCGCGACAAAGACTTTTTTTTTTTTTTTTTTTTT
Otop_32[103]_34[104]	GTCGTCTTTCCAGACGTTAGTATCACCGTACTCAGTAACAGTTTTTTTTTTTTTTTTTTTTT
Otop_36[103]_38[104]	ACTGTAGCAACCGATTGAGGGAGGCAAAAGATTTTTTTTTTTTTTTTTTTT
Otop_38[103]_40[104]	CTGGCATGAACACCCTGAACAAAGTACCAACTTTTTTTTTTTTTTTTTTTT
Otop_26[103]_28[104]	GATAGCGTACATTATTACAGGTAGTTCAGTGTTTTTTTTTTTTTTTTTTT
Otop_24[103]_26[104]	TAACAGTTCTTCAAATATCGCGTTTTAGACTTTTTTTTTTTTTTTTTTTT
Otop_34[103]_36[104]	CCCGTATACACCACCAGAGCCGCCAGCGTCATTTTTTTTTTTTTTTTTTTT
Fl. Strand	Cy3-AAAAAAAAAAAAAAAAAAA

Table 2: List of DNA sequences (from the 5’ to the 3’ end) forming the membrane-sculpting DNA origami plates.

Supplementary Table S3: List of DNA sequences for blunt-end stacking between the DNA origami plates

Estick_01[176]_01[202]	ATTTAAATTGTAAACGTTAATATTTTG
Estick_02[202]_02[176]	TTAAAATTCGCATTAAATTTTTGTTAA
Estick_03[176]_03[202]	GCGGGCCTCTTCGCTATTACGCCAGCT
Estick_04[202]_04[176]	GGCGAAAGGGGGATGTGCTGCAAGGCG
Estick_05[176]_05[202]	AAGTTAAACGATGCTGATTGCCGTTCC
Estick_06[202]_06[176]	GGCAAACGCGGTCCGTTTTTTCGTCTC
Estick_07[176]_07[202]	GGGTTACCTGCAGCCAGCGGTGCCGGT
Estick_08[202]_08[176]	GCCCCCTGCATCAGACGATCCAGCGCA
Estick_09[176]_09[202]	GCATTAATGAATCGGCCAACGCGCGGG
Estick_10[202]_10[176]	GAGAGGCGGTTTGCGTATTGGGCGCCA
Estick_11[176]_11[202]	AATCGGAACCCTAAAGGGAGCCCCCGA
Estick_12[202]_12[176]	TTTAGAGCTTGACGGGGAAAGCCGGCG
Estick_13[176]_13[202]	ACATCACTTGCCTGAGTAGAAGAACTC
Estick_14[202]_14[176]	AAACTATCGGCCTTGCTGGTAATATCC
Estick_15[176]_15[202]	CACGCTGAGAGCCAGCAGCAAATGAAA
Estick_16[202]_16[176]	AATCTAAAGCATCACCTTGCTGAACCT
Estick_17[176]_17[202]	ATACTTCTGAATAATGGAAGGGTTAGA
Estick_18[202]_18[176]	ACCTACCATATCAAAATTATTTGCACG
Estick_19[176]_19[202]	GAAAACATAGCGATAGCTTAGATTAAG
Estick_20[202]_20[176]	ACGCTGAGAAGAGTCAATAGTGAATTT
Estick_22[207]_22[168]	ACGCAAGGATAAAAATTTTTAGAACCCTCATATATTTTA
Estick_23[176]_23[207]	TAATACTTTTGCGGGAGAAGCCTTTATTTCA
Estick_24[207]_24[176]	TTTTTGCGGATGGCTTAGAGCTTAATTGCTG
Estick_25[176]_25[207]	ACCTTTAATTGCTCCTTTTGATAAGAGGTCA
Estick_26[207]_26[176]	ATCATAACCCTCGTTTACCAGACGACGATAA
Estick_27[176]_27[207]	AGGAATTACGAGGCATAGTAAGAGCAACACT
Estick_28[207]_28[176]	CAGACCAGGCGCATAGGCTGGCTGACCTTCA
Estick_29[176]_29[207]	CAACTTTGAAAGAGGACAGATGAACGGTGTA
Estick_30[207]_30[176]	GGAGTTAAAGGCCGCTTTTGCGGGATCGTCA
Estick_31[176]_31[207]	AACCGATATATTCGGTCGCTGAGGCTTGCAG
Estick_32[207]_32[176]	AACCCATGTACCGTAACACTGAGTTTCGTCA
Estick_33[176]_33[207]	CCTCATTTTCAGGGATAGCAAGCCCAATAGG
Estick_34[207]_34[176]	TCTCTGAATTTACCGTTCCAGTAAGCGTCAT
Estick_35[176]_35[207]	ATCCTCATTAAAGCCAGAATGGAAAGCGCAG
Estick_36[207]_36[176]	TAGCACCATTACCATTAGCAAGGCCGGAAAC
Estick_37[176]_37[207]	ATTTGGGAATTAGAGCCAGCAAAATCACCAG
Estick_38[207]_38[176]	CTATCTTACCGAAGCCCTTTTTAAGAAAAGT

Estick_39[176]_39[207]	ATAAGAGCAAGAAACAATGAAATAGCAATAG
Estick_40[207]_40[176]	GTTTTAGCGAACCTCCCGACTTGCGGGAGGT
Estick_42[192]_41[207]	TTTTTAATTTAGGCAGAGGCTTATCCGGTATTCTAAGAACGCGAGGC
Estick_02[31]_02[5]	TAATGGGATAGGTCACGTTGGTGTAG
Estick_03[5]_03[31]	ATGGGCGCATCGTAACCGTGCATCTG
Estick_04[31]_04[5]	CAGTCCCGGAATTTGTGAGAGATAGA
Estick_05[5]_05[31]	CTTTCTCCGTGGTGAAGGGATAGCTC
Estick_06[31]_06[5]	CAGCTTACGGCTGGAGGTGTCCAGCA
Estick_07[5]_07[31]	TCAGCGGGGTCATTGCAGGCGCTTTC
Estick_08[31]_08[5]	TACCGAGCTCGAATTCGTAATCATGG
Estick_09[5]_09[31]	TCATAGCTGTTTCCTGTGTGAAATTG
Estick_10[31]_10[5]	CAAAATCCCTTATAAATCAAAAGAAT
Estick_11[5]_11[31]	AGCCCGAGATAGGGTTGAGTGTTGTT
Estick_12[31]_12[5]	GAGCACGTATAACGTGCTTTCCTCGT
Estick_13[5]_13[31]	TAGAATCAGAGCGGGAGCTAAACAGG
Estick_14[31]_14[5]	CCAACAGAGATAGAACCCTTCTGACC
Estick_15[5]_15[31]	TGAAAGCGTAAGAATACGTGGCACAG
Estick_16[31]_16[5]	TAGAAGTATTAGACTTTACAAACAAT
Estick_17[5]_17[31]	TCGACAACTCGTATTAAATCCTTTGC
Estick_18[31]_18[5]	CGAATTATTCATTTCAATTACCTGAG
Estick_19[5]_19[31]	CAAAAGAAGATGATGAAACAAACATC
Estick_20[31]_20[5]	CATCTTCTGACCTAAATTTAATGGTT
Estick_21[5]_21[31]	TGAAATACCGACCGTGTGATAAATAA
Estick_23[11]_22[25]	TTCATTTGGGGCGCGAGCTGAAGATTTT
Estick_24[31]_24[11]	AATAACCTGTTTAGCTATATT
Estick_25[11]_25[31]	TCAAAAATCAGGTCTTTACCC
Estick_26[31]_26[11]	GAAAACGAGAATGACCATAAA
Estick_27[11]_27[31]	ATGCGATTTTAAGAACTGGCT
Estick_28[31]_28[11]	TTAATCATTGTGAATTACCTT
Estick_29[11]_29[31]	AAACACTCATCTTTGACCCCC
Estick_30[31]_30[11]	AAGAGGCAAAAGAATACACTA
Estick_31[11]_31[31]	AATTGCGAATAATAATTTTTT
Estick_32[31]_32[11]	ATAGAAAGGAACAACTAAAGG
Estick_33[11]_33[31]	GATTAGGATTAGCGGGGTTTT
Estick_34[31]_34[11]	GCTGAGACTCCTCAAGAGAAG
Estick_35[11]_35[31]	ACCACCGGAACCGCCTCCCTC
Estick_36[31]_36[11]	AAAATCACCGGAACCAGAGCC
Estick_37[11]_37[31]	ACACCACGGAATAAGTTTATT
Estick_38[31]_38[11]	ACATATAAAAGAAACGCAAAG
Estick_39[11]_39[31]	TTAACGTCAAAAATGAAAATA
Estick_40[31]_40[11]	AAATAAGAAACGATTTTTTGT
Estick_41[11]_41[11]	TTTACGAGCATGTAGAAACCA
Estick_42[33]_42[11]	AAAAATAATATCCCATCCTAA

Table 3: List of DNA sequences (from the 5' to the 3' end) that were added to the membrane-sculpting DNA origami to induce blunt-end stacking and hence polymerization of the DNA origami plates into larger sheets.

Supplementary Notes

Supplementary Note S1: Estimation of the pH change for DNA attachment

In principle, we can quantify the pH change from the quantification of the DNA attachment and the corresponding calibration curve (see Figure 3b). However, the problem in comparing the fluorescent ratios from the calibration measurement in Figure 3b and the DNA attachment mediated via *E. coli* is the fact that the *E. coli* containing droplets absorb some of the excitation as well as emission light from the pH-sensitive DNA which makes us cautious regarding fully quantitative statements on the pH. However, we can still approximate the rise in pH. During light illumination of the droplets with *E. coli* the ratio I_{peri}/I_{in} rises from 1.83 to 3.58 and thus by a factor of 1.96. From the confocal images in Figure S24 and the pH electrode measurements in Figure 1b, we can deduce the starting point of pH 6.2 before light illumination takes place. This is also in line with the calibration curve in Figure 3b and Supplementary Figure S9, where the DNA starts to attach in between pH 6-6.5. A pH of 6.2 equals a ratio in the calibration measurement of $I_{peri}/I_{in}=1.36$ and is thus smaller than the starting value of the measurement with *E. coli* due to absorption of light by *E. coli* on the inside of the droplet and hence a smaller I_{in}. Multiplying the ratio $I_{peri}/I_{in}=1.36$ at pH 6.2 by 1.96 yields 2.72, which corresponds to a pH of 7.25. This means that by this approximation the pH within droplets rises from pH 6.20 to pH 7.25 during light illumination. This is comparable to the pH range obtained from bulk measurements where the pH increases from pH 6.2 to 7.0 and further in line with a comparison of the images from the calibration measurement with the ones of the droplets after light illumination.

Supplementary Note S2: Estimation of DNA origami density per GUV

To obtain a lower bound estimate for the area coverage of the DNA origami on the GUVs we need to estimate the lipid concentration first. For the electroformation 40 µL of 1mM lipids in chloroform are spread on the whole ITO slide until all chloroform evaporated. Subsequently, 275 µL of aqueous buffer solution is added into the ring covering 1/5 of the area covered with
lipids leading to a final lipid concentration of 29 µM. Note that this is an upper bound estimate for the lipid concentration since most likely not all lipids will detach from the ITO slide. After mixing the GUV solution with the DNA origami solution the effective concentrations are $c_{lipid} \sim 20$ µM and $c_{DNA} = 2$nM. The relative area of the GUV covered with DNA origami is then given by:

$$\alpha = \frac{c_{DNA} * A_{DNA}}{2 * C_{lipid} * A_{lipid}} = 0.175$$

with $A_{DNA} = 250\text{nm}^2$ and $A_{lipid} = 0.7\text{nm}^2$. Thus, at least 17% of the membrane area of the GUVs should be covered with DNA origami.

Supplementary Videos

Supplementary Video S1: Monitoring of the pH change induced by light-harvesting *E. coli* with pyranine in bulk

Confocal fluorescence time series of xenorhodopsin-overexpressing *E. coli* mixed with 50 µM pyranine and GUVs. The sample was illuminated for 5 min after 10, 25 and 40 min. During illumination, the pH of the bulk solution becomes more basic, leading to an increase in pyranine emission upon excitation with 488nm, hence the apparent blinking in the video. Scale bar: 100 µm.

Supplementary Video S2: Monitoring of the pH change induced by light-harvesting *E. coli* with pyranine in water-in-oil droplets

Confocal fluorescence time series of xenorhodopsin-overexpressing *E. coli* mixed with 50 µM pyranine and encapsulated into surfactant-stabilized water-in-oil droplets. The sample was illuminated for 5 min after 10, 25 and 40 min. During illumination, the pH of the bulk solution becomes more basic leading to an increase in pyranine emission upon excitation with 488nm. Scale bar: 100 µm.

Supplementary Video S3: DNA attachment to the compartment periphery of water-in-oil droplets by light illumination

Confocal fluorescence time series of Cy5-labeled triplex-forming pH.sensitive DNA (λ_{ex} =647nm within microfluidic droplets containing xenorhodopsin-overexpressing *E. coli*. The light was turned on for 30 min after 30 min in the dark. During light illumination, the triplex-forming DNA attaches to the droplet periphery. Scale bar: 100 µm.

Supplementary Video S4: Reversible binding of triplex-forming DNA to the droplet periphery

Reversible binding of triplex-forming DNA to the droplet periphery. A change of the internal pH of the droplets was achieved by flushing a proton donor or a proton acceptor via the oil phase. Left: Confocal fluorescence time series of microfluidic droplets containing 1 µM triplex-forming DNA (green, λ_{ex} =488nm) and 1:5 µM cholesterol-tagged DNA. The proton acceptor propylamine (1 vol% in HFE) was flushed into the observation chamber. This causes the pH inside the droplets to increase, leading to the binding of triplex-forming DNA to the periphery. Right: Confocal fluorescence images of microfluidic droplets after flushing the propylamine-containing oil phase. To reversibly switch the pH inside droplet-based compartments and to cause unbinding of the triplex-forming DNA from the droplet periphery, the oil phase was modified with trifluoroacetic acid (1 vol% in HFE) and subsequently flushed into the observation chamber. This causes the pH inside the droplets to decrease leading to the unbinding of triplex-forming DNA. All in all, the video confirms the pH-reversible attachment of the triplex-forming DNA to the droplet periphery. Scale bar: 50 µm.

Supplementary Video S5: DNA origami cortex suppresses membrane fluctuations

Membrane fluctuations of osmotically deflated GUVs (c/c_0 = 1.8) with and without membrane-bound DNA origami. The plain GUV (green, lipids labelled with Atto488) shows significant membrane fluctuations, while for the deformed GUV with membrane-bound polymerized DNA origami (orange, DNA origami labelled with Cy3) the fluctuations are suppressed. For a quantitative analysis of the fluctuations, see Supplementary Figure S19.

Supplementary Video S6: Light-mediated attachment of triplexforming DNA to GUVs

Exemplary confocal time series of a DNA-coated GUV surrounded by *E. coli* (0.4 µM triplex-forming DNA, λ_{ex}=561nm; 0.6 µM cholesterol-tagged DNA). The time lapse shows the DNA attachment over time after light illumination as well as the settling of *E. coli*. The

droplet was illuminated for 15 min after 25 min in the dark. The corresponding quantification of the peripheral DNA intensity is shown in Figure 4e (main text). Scale bar: 20 μm.

Supplementary Video S7: DNA origami attaches to GUVs during light-illumination

Exemplary confocal time series of a GUV surrounded by *E. coli* (λ_{ex}=561nm; 0.6 μM cholesterol-tagged DNA). The time lapse shows the DNA origami attachment over time during light illumination. Note that the *E. coli* settle over time and therefore disappear from the confocal plane. The GUV was illuminated from 35-50 min and from 110-155 min. The corresponding quantification of the peripheral DNA intensity is shown in Figure 5b (main text). Scale bar: 20 μm.

Supplementary Video S8: DNA origami deform GUVs after attachment

Exemplary confocal time series of a GUV surrounded by *E. coli* (λ_{ex}=561nm; 0.6 μM cholesterol-tagged DNA). The time lapse shows the DNA origami-mediated deformation over time after light illumination and addition of the DNA staple strands at the scaffold seam which enable blunt-end stacking and thus polymerization of the DNA origami. Scale bar: 20 μm.

Supplemental References

(1) Paxton, W. F.; Price, D.; Richardson, N. J. Hydroxide ion flux and pH-gradient driven ester hydrolysis in polymer vesicle reactors. Soft Matter 2013, 9, 11295.
(2) Platzman, I.; Janiesch, J.-W.; Spatz, J. P. Synthesis of Nanostructured and Biofunctionalized Water-in-Oil Droplets as Tools for Homing T Cells. Journal of the American Chemical Society 2013, 135, 3339–3342.
(3) Weiss, M. et al. Sequential bottom-up assembly of mechanically stabilized synthetic cells by microfluidics. Nature Materials 2017, 17, 89–96.
(4) Jahnke, K.; Weiss, M.; Weber, C.; Platzman, I.; Göpfrich, K.; Spatz, J. P. Engineering Light-Responsive Contractile Actomyosin Networks with DNA Nanotechnology. Advanced Biosystems 2020, 2000102.
(5) Kocabey, S.; Kempter, S.; List, J.; Xing, Y.; Bae, W.; Schiffels, D.; Shih, W. M.; Simmel, F. C.; Liedl, T. Membrane-Assisted Growth of DNA Origami Nanostructure Arrays. ACS Nano 2015, 9, 3530–3539.

8 Conclusion and Outlook

The goal of this thesis was to broaden the methodology, required to assemble artificial cells. Thereby, the presented studies are mainly focused on membrane protein-based approaches. While the studies presented in chapters 3-5 are focused on the investigation of membrane proteins by employing atomic force microscopy (AFM)-based single-molecule force spectroscopy (SMFS), chapters 6 and 7 demonstrate applications of membrane proteins in artificial cells.

The studies presented in chapters 3-5 investigate the intricate network of forces, stabilizing structural segments of outer membrane proteins (OMPs). Hence, the two OMPs LamB and BamA from *Escherichia Coli* were mechanically unfolded with AFM-based SMFS. The two proteins were selected for investigation in order to identify common structural features, shaping the unfolding pathways of OMPs. Previous studies have identified, that by applying a mechanical pulling force at one terminus, β-barrel OMPs typically unfold stepwise, one β-hairpin after the other, until the whole protein is extracted from the membrane. Upon mechanically unfolding LamB, a similar unfolding pathway was observed (chapter 3). However, the study also revealed, that β-hairpins can unfold in groups, especially if they display high mechanical stabilities. This suggests, that the probability to observe collective unfolding of β-hairpins is increased in far-from-equilibrium processes. In order to be investigated with the AFM, LamB was purified, reconstituted and adsorbed to a mica surface, which resulted in two dimensional, densely packed and highly ordered structures. This rendered the sample suitable for high resolution AFM imaging, which was used to study the interaction of different maltosaccharide substrates with LamB, while simultaneously recording topographs of the sample with sub-nanometer resolution (E. Mulvihill et. Al., 2019, not included in this thesis). The results of this study suggest asymmetric sugar binding and uptake, which supports one of two contradicting models describing LamB-mediated sugar import.

While LamB was purified and reconstituted into a new membrane environment, SMFS studies conducted with BamA were performed in both, a reconstituted state, as well as in its natural outer membrane environment (chapter 4). Hence, conclusions could not only be drawn about the membrane protein under investigation, but also how it is affected by the membrane in which it is embedded. Furthermore, the influences of additional factors on the mechanical stability of BamA were investigated, such as the extracellular lid and the POTRA domains. To this end, BamA was mechanically unfolded by SMFS under those different conditions. In all conducted experiments, BamA unfolded hairpinwise, similar to the unfolding behavior of LamB and other OMPs. However, altering the protein or the membrane environment modulated the mechanical stability of BamA. Mutating BamA to unlock the extracellular lid, as well as deleting the POTRA domains, decreased the mechanical stabilities of the first four β-hairpins of the BamA β-barrel. This indicates that various factors stabilize BamA in a delocalized manner, since the mechanical stabilities of β-hairpins were influenced, which were not mutated in the experiments. In contrast to the introduced mutations, BamA reconstituted into a new membrane environment displays higher mechanical stability than in the native outer membrane. This indicates the importance of the native membrane for BamA functionality, which might be required to establish local distortions, necessary for the successful integration of incoming OMPs. The different factors modulating the mechanical stability of BamA further suggest, that it exists in a metastable state in the native outer membrane. Those findings are in line with structural studies, which have identified high conformational dynamics of the BamA β-barrel domain.

Chapter 5 presents an additional SFMS study conducted with BamA, which investigated the effects of the novel antibiotic compound darobactin on BamA. Darobactin was recently identified to bind to and inhibit BamA. In this study, dynamic force spectroscopy (DFS) was employed to investigate how the mechanical, kinetic and energetic properties of BamA are affected upon darobactin binding. The results show, that mechanically unfolding BamA in the absence and in the presence of darobactin identified the same structural segments and thus describes the same sequential unfolding pathways, which indicates a characteristic network of interactions stabilizing BamA. However, the results show, that the mechanical stability of BamA is substantially increased in the presence of darobactin, especially in the β-barrel domain. The increased mechanical stability in the inhibited form of BamA is in line with the results presented in chapter 4, as it substantiates the importance of a metastable state, required for BamA functionality. The DFS experiments further contoured the free-energy landscape of individual structural segments of BamA in presence and absence of darobactin. This revealed, that the first and the last β-hairpins of the BamA β-barrel form a mechanically stable gate, which is further stabilized by darobactin. However, the β-hairpins H5-H7 provide a mechanically flexible region, potentially required to insert and fold OMP-substrates. Furthermore, the structural region linking the POTRA domains to the transmembrane β-barrel displays low kinetic stability and conformational variability as well as high mechanical stiffness, which indicates a mechano-functional role. What this role might be, remains to be elucidated in further studies.

It should be mentioned here, that considerable time and effort was invested in studying how BamA folds its substrate into the membrane by conducting numerous SMFS refolding experiments with BamA. However, to this date, the results remain unconclusive and are therefore neither published nor included in this thesis. The main reason was that, due to experimental restrictions, the insertase activity of BamA was studied with BamA as a substrate. With the goal to measure single molecule refolding events and to answer, whether BamA can promote the folding and insertion of other BamA molecules, millions of FD-curves were recorded and analysed. However, multiple factors persistently prevented a meaningful interpretation of the results. Nonetheless, during the refolding experiments, BamA was mechanically unfolded in the presence of darobactin and what was initially intended as a control experiment ultimately led to the findings presented in Chapter 5.

When collecting data for the presented SMFS studies, the AFM has proven to be a valuable and well-established tool for studying membrane proteins under physiologically relevant conditions. However, initially the sample preparation, data acquisition as well as data analysis have emerged as bottlenecks in the process. Meanwhile, the sample preparation for the study of OMPs was substantially facilitated by employing outer membrane vesicles enriched with the protein of interest, instead of purified and reconstituted proteins. The pipeline for analysing the large amount of data was also significantly accelerated with the help of self-written analysis software. Hence, data acquisition remains the single biggest factor slowing down the process of AFM-based SMFS studies. This is mainly due to the low probability of establishing an unspecific interaction between the cantilever tip and the protein of interest, which renders successful unfolding events rare. If future studies can achieve increased unfolding probabilities, for example by establishing a specific interaction between the cantilever tip and the protein of interest or by designing improved linker moieties, the process of AFM-based SMFS would become a more efficient tool to investigate different membrane proteins under varying conditions. Additionally,

many of the obstacles encountered during the refolding experiments conducted with BamA could be overcome with the added efficiency and control.
In contrast to chapters 3-5, the studies presented in chapters 6 and 7 are less focused on the investigation of membrane proteins and more on their applications in artificial cells. Hence, in order to manipulate membrane proteins to improve their applicability in artificial cells, two proteorhodopsin (PR) fusion constructs were designed and engineered in chapter 6 to control their orientation in artificial membranes. The presented methodology provides a solution to the random orientation of membrane proteins upon reconstitution, a common obstacle in artificial cell engineering. PR translocates protons across the membrane upon illumination, which allows verification of the final protein orientation by measuring the extra-vesicular pH change upon illumination. Indeed, the two constructs were functionally integrated and translocated protons along the predicted directions upon reconstitution, which demonstrated that the approach was successful. Moreover, reconstituted PR provides an energizing building block for future artificial cell engineering efforts. The proton-pumping activity of PR can be utilized to energize downstream energy consuming components in more complex artificial cells. The presented fusion proteins were also successfully employed in an extensive study to optimize the detergent mediated reconstitution of PR into polymer vesicles (R. Goers et al., 2018, not included in this thesis). The fluorescence of the soluble fusion domains, as well as the properties to guide the directed insertion were beneficial for conducting this study.
The PR constructs presented in chapter 6 also provide a method to externally control pH dependent processes, which was utilized in a subsequent study, presented in chapter 7. However, PR was replaced with xenorhodopsin (XeR), another light driven proton pump, which was modified, similar to PR in chapter 6. In this study XeR was chosen over PR, due to a substantially faster photocycle, which results in a stronger proton-flux and thus in a more drastic change in pH. Also, compared to the previous study, XeR-overexpressing bacteria were used directly, instead of purified and reconstituted XeR. The bacteria were encapsulated in microfluidic water-in-oil droplets, where they were used to externally trigger pH dependent processes upon illumination. In the presented study the engineered bacteria were thus utilized to demonstrate externally controllable immobilization of pH dependent DNA constructs to the droplet periphery. The bacteria and the DNA-constructs were combined in different systems to mimic different natural processes, such as cell-cell communication, formation of sub-compartments in the form of organelles and the assembly of a simplified cytoskeleton. In a final experiment a reaction cascade was assembled, which ultimately resulted in the deformation of giant unilamellar vesicles upon illumination.
Although the presented artificial cells in chapters 6 and 7 are a proof of principle, they provide new building blocks and methodologies for further approaches in the future. The presented systems could be further improved by either spectrally tune individual light-driven proton pumps or by combining different pumps with different absorption properties. In combination with the presented strategy to control protein orientation in the membrane, different light-driven proton pumps could be combined in one artificial cell to increase or decrease the internal pH, depending on the wavelength of the incident light. Promising initial experiments pointing in this direction have been conducted (unpublished results).
Finally, the functional integration of a membrane protein insertase has the potential to act as a scaffold for the controlled insertion and combination of numerous membrane proteins, which could advance the field significantly.

9 Acknowledgment

During the time I spent working on this thesis, I received enormous support from many people, who made time fly by, which is an enormous achievement, considering that it took five years. Since this thesis now has over 200 pages, I cannot thank everyone to whom I owe this great time, in order not to fill another 200 pages. So, for everyone I forgot to mention here: Thank you very much, you are awesome!

First of all, I want to thank Daniel Müller. I am really grateful, that you gave me the opportunity to be part of your group. You somehow always managed to meet every single one of us on a regular basis, even if that meant, you had to turn your relaxing walks through Basel into "Dackel-Walk" meetings. You also gave me a lot of freedom to implement my own ideas, and develop little side-projects, which I really appreciate. Thank you for your support Daniel – Stay "Z'wääääg".

I also want to thank Sebastian Hiller. Working with you and the members of your group was highly interesting. The meetings with you were always helpful, and the feedback from your side drastically improved our projects.

In addition, I want to thank Dimitrios Fotiadis. You and your lab members were always very helpful and a good company to hang out with. Without your prompt and uncomplicated support, proteorhodopsin likely would not have entered any membrane in either direction.

I also want to thank Yosh (Johannes Thoma). You teached me so much during the time you have mentored me and my projects. Your "can't be that difficult, others can do that too" mentality was a real inspiration for me and helped me to tackle projects that I would not have dared to without you as a role model. I hope you are doing well in the north and that I manage to visit you soon.

I want to thank all the members of the Müller Lab. I really really enjoyed the time I spent with all of you. Whenever I needed some help, wanted to discuss anything, or just drink a "covfefe", I knew you would all be there for me, which is truly amazing. I am honored to be part of this group and hope we will all stay in touch. I don't want to name any people individually here because I couldn't do you all justice and would probably even forget somebody and then feel bad about it. So please all, feel hugged and may the force be with you.

I also want to thank all the great collaborators I worked with in the last years. Special thanks to Kerstin Göpfrich, Kevin Jahnke, Stephan Hirschi and Roland Goers. It was a great pleasure to work with you.

I thank Nico Strohmeyer and Pawel Laskowski for taking the time to carefully read my thesis and providing valuable feedback. Your input is much appreciated.

I also want to thank my family and all my friends. You all are just the best and probably the main reason I never made it out of Basel. Special thanks to my parents. Your unconditional support is fundamental for everything I do. Thank you so much for everything you did for me!

Finally, I want to thank Jolanda. At this point I wanted to thank you, for making the time since we first met the happiest time of my life and write that I truly love you. However, since you're not the romantic type and I am not skilled at writing flowery texts, I guess a simple "thank you" is fine too. (Seriously though, thank you for everything!)