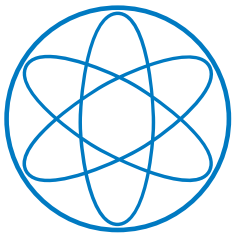

Physics Department

Probing RNA Aptamers, Guide RNAs and CRISPR-Cas9/12 Complexes with DNA Nanopores and AFM

Dissertation by

Daniela Ziegler

Technische Universität München

Technische Universität München
Fakultät für Physik
Physik synthetischer Biosysteme

Probing RNA Aptamers, Guide RNAs and CRISPR-Cas9/12 Complexes with DNA Nanopores and AFM

Daniela Ziegler

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Ulrich Gerland

Prüfer der Dissertation:

1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Hendrik Dietz

Die Dissertation wurde am 14.10.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 07.12.2020 angenommen.

Contents

Abstract

Nucleic acid structures are versatile in function. DNA and RNA convey information, serve as regulatory elements, and are the building material for the fabrication of functional DNA and RNA structures. Single-molecule studies assist in illuminating the function of DNA and RNA secondary structures.

A method for single-molecule studies is applying a voltage across a droplet interface bilayer (DIB) and recording the resulting current. The electrical recordings allow for the characterization of DNA pores and for single-molecule studies of translocating molecules. DNA pores characterized prior to this work were 'single-stranded tile' (SST) pores and a DNA origami pore. In this work, the first-time recording of a large-diameter DNA nanopore that incorporates into lipid bilayers was performed.

Additionally, the incorporation of two further DNA origami nanopores with a smaller diameter was recorded. The experimentally obtained ionic conductance values were shown to agree with the dimensions of the stem of the pore.

Further, proof of principle experiments were carried out, wherein the hydrophobic anchoring of the pores was replaced by elastin-like polypeptide (ELP) anchoring or by biotin-streptavidin bridging. Both replacements broaden the spectrum of incorporation strategies.

The large-diameter DNA nanopore was used as a single molecule tool for sensing the electrically driven translocation of nucleic acids. Prior to this work, the translocation of single-stranded DNA (ssDNA) through a small diameter DNA pore and the translocation of double-stranded DNA (dsDNA) through a DNA plate attached on top of a solid state pore had been demonstrated. Subsequent to our work, the presence of proteins in solution was sensed with a large-diameter DNA pore.

Here, the first-time recordings of the translocation of dsDNA, RNA, and rNTPs through a DNA nanopore that incorporates into lipid bilayers were performed.

The translocation data imply that the translocating molecules might interact with

the pore.

Another single-molecule technique to characterize functional DNA and RNA structures is atomic force microscopy (AFM). It was used to examine whether functional RNA structures can enable DNA loop formation. Since the advent of the clustered, regularly interspaced palindromic repeat (CRISPR) technique, it has been demonstrated that a DNA loop can be designed with two cross-linked CRISPR proteins and their respective guide RNAs. In this thesis, a different approach was explored. Here the focus was shifted towards the rational design of RNA structures.

Therefore, the DNase inactive CRISPR protein D908A Cas12a was mutated into a DNase and RNase inactive Cas12a protein. It was proven that the RNA processing of H800A D908A AsCas12a is deactivated and that the DNA binding capability of ddCas12a remains active. The ddCas12a protein devoid of DNase and RNase activity was subsequently used in experiments aiming towards DNA loop formation.

Zusammenfassung

Nukleinsäurestrukturen sind in ihrer Funktion vielseitig. DNS und RNS übermitteln Informationen, dienen als regulatorische Elemente und sind das Material für die Herstellung von funktionellen DNS- und RNS-Strukturen. Einzelmolekülstudien helfen bei der Aufklärung der Funktion von DNS- und RNS-Strukturen.

Eine Methode zur Studie einzelner Moleküle ist das Anlegen einer elektischen Spannung über eine Lipiddoppelschicht, welche über Tröpfchengrenzflächen keeirt wird, und die Aufzeichung des daraus resultierenden Stroms. Diese Methode ermöglicht Einzelmolekülmessungen, wie die Charaktariesierung von sich in die Lipidschicht einbauender DNS-Poren oder die Untersuchung translokierender Moleküle.

Bei den DNS Poren, die vor dieser Arbeit charakterisiert wurden, handelt es sich um "Single-Strand-Tile"-Poren und um eine DNS-Origami-Pore. Im Zuge dieser Arbeit wurde zum ersten Mal eine sich in Lipidschichten einbauende DNS-Nanopore mit großem Durchmesser untersucht und dokumentiert.

Zusätzlich wurde die Inkorporation von zwei weiteren DNS-Origami-Nanoporen mit kleinerem Durchmesser aufgezeichnet. Es konnte gezeigt werden, dass die experimentell ermittelten Werte der Ionenleitfähigkeit mit den Abmessungen des Porenstammes übereinstimmen.

Weiterhin wurde in Grundsatzbeweisen gezeigt, dass die hydrophoben Ankerelemente der Poren durch Elastin-ähnliche Polypeptidverankerung oder durch eine Biotin-Streptavidin Verankerung ersetzt werden können. Beide Ersetzungen erweitern das Spektrum der möglichen Inkorporationsstrategien.

Die DNS-Nanopore mit großem Durchmesser wurde als Einzelmolekül-Werkzeug zur Aufzeichnung der elektrisch getriebenen Translokation von Nukleinsäuren verwendet. Bereits vor dieser Arbeit wurde einerseits die Translokation von einzelsträngiger DNS durch eine DNS-Pore mit kleinem Durchmesser und andererseits die Translokation von doppelsträngiger DNS durch eine DNS-Platte, die auf einer Festkörperpore befestigt war, nachgewiesen. Im Anschluss an den bereits veröf-

fentlichten Teil der vorliegenden Arbeit, wurde mit einer DNS-Pore mit großem Durchmesser das Vorhandensein von Proteinen in Lösung nachgewiesen.

Im Zuge der vorliegenden Arbeit wurden zum ersten Mal Translokationen von dsDNS, RNS und rNTPs durch eine in Lipiddoppelschichten einbauende DNS-Nanopore durchgeführt. Die Translokationsdaten deuten darauf hin, dass die translokierenden Moleküle mit der Pore interagieren könnten.

Eine weitere Einzelmolekültechnik zur Charakterisierung funktioneller DNS- und RNS-Strukturen ist die Rasterkraftmikroskopie. Mit ihr wurde untersucht, ob funktionelle RNS-Strukturen die Bildung von DNS-Loops ermöglichen können. Der Ausgangspunkt hierfür ist, dass seit dem Aufkommen der 'Clustered, Regular Interspaced Palindromic Repeat' (CRISPR)-Technik bereits gezeigt werden konnte, dass ein DNS-Loop mittels zwei verlinkten CRISPR-Proteinen und ihren jeweiligen guide-RNSs kreeirt werden kann. In dieser Arbeit wurde ein alternativer Ansatz erforscht. Hier wurde der Schwerpunkt auf das rationale Design von RNS-Strukturen verlagert.

Dazu wurde das DNase-inaktive CRISPR-Protein D908A Cas12a zum DNase- und RNase-inaktiven Protein H800A D908A AsCas12a mutiert. Die Deaktivierung der RNS-Prozessierung sowie das Bestehenbleiben der DNS-Bindungsfähigkeit von H800A D908A AsCas12a konnten nachgewiesen werden. Abschließend wurde das ddCas12a-Protein ohne DNase- und RNase-Aktivität in Experimenten verwendet, die auf die Bildung von DNS-Loops abzielten.

1. Fundamentals

Life on earth is determined by a complex interplay of macromolecules. Key molecules among those are nucleic acids, proteins, and lipids, which are highlighted in Ch. 1.1 and 1.2 of the introductory chapter.

Chapter 1.1 focuses on nucleic acids and the central dogma of molecular biology. DNA, the blueprint of organisms, is a polymer chain that encodes information in the sequential arrangement of its elements, the nucleotides. The sugar-phosphate compound of each nucleotide allows for the polymerization of nucleotides into extended stretches of DNA and confers the strand directionality. Two strands can hybridize since each nucleotide contains one of four bases with predictable base-pairing properties. Hybridized strands form a double helix, which buries the bases inside and protects them from degradation.

Around the DNA molecule, a readout machinery evolved. RNA polymerase transcribes DNA into RNA. This single-stranded polymer of nucleotides is more susceptible to hydrolysis and hence to degradation. The RNA's temporary nature designates it as a versatile regulatory element.

Proteins are a sequence of amino acids condensated into polymer chains. Proteins are encoded by the messenger RNA, which the ribosome translates into the codon-assigned amino acid sequence. Driven by hydrophobic interactions and at times chaperon-assisted, the protein folds co-translationally in its three-dimensional structure. Proteins and other molecules often are located in defined volumes, which obviate the dilution of the molecules and boost reaction rates.

Nature's abundant macromolecule for encapsulation are lipids (Ch. 1.2). The amphiphilic molecules, inter alia, assemble into lipid sheets. Double layered lipid sheets, which are functionalized by embedded proteins, encapsulate substrates and can form cellular or subcellular compartments. Compartments are pivotal for the spatial division of tasks. Moreover, division is the foundation for molecular gradients across the bilayer. One of those, the protein gradient, fuels ATPases, which

generate energy. The molecular exchange between the in- and outside of the compartment can be regulated by carriers or nanopores.

Lipids are essential for building the insulating barrier in which nanopores can incorporate. Chapter 2 and Chapter 3 focus on those nanopores and the translocation of molecules through them. DNA is the building material for the nanopores, which in return implies that the DNA is functional.

Chapter 4 focuses on a functional RNA. There the functional RNA recruits two CRISPR-associated (Cas) proteins to enable the formation of DNA loops.

1.1. The Central Dogma of Molecular Biology

The classical view on DNA, RNA, and proteins includes the transcription of DNA into RNA and the subsequent translation of RNA into proteins, which both in turn are degraded [1]. The broader, modern approach, summarized in Figure 1, offers a more nuanced depiction. In there, DNA and RNA execute a range of regulatory functions on top of providing the code for protein translation [2].

Certain regulatory *in vivo* functions of RNA are its self-cleavage and its conformational changes upon external triggers, as observed in ribozymes and riboswitches [3,4]. RNAs trigger immune responses and can assist in editing the genome [5,6].

Two prominent *in vitro* uses of RNA are the systematic enrichment by exponential evolution (SELEX) and the design of RNA nanostructures [7,8].

1.1.1. Nucleic Acids

Nucleic acids are polymers, comprising repetitive units of ribonucleotides or deoxyribonucleotides, respectively forming ribonucleic acid (RNA) or deoxyribonucleic acid (DNA). Nucleotides are composed of a central sugar ring, either ribose or deoxyribose, a triphosphate at the 5' carbon atom of the (deoxy)ribose, and one of four nucleobases, connected to the 1' carbon atom of the (deoxy)ribose via a carbon-nitrogen bond. The four nucleobases subdivide into pyrimidines and purines, which either contain one or two carbon-nitrogen rings, respectively. The pyrimidines thymine (T), uracil (U), and cytosine (C) preferentially form hydrogen bonds with their complementary Watson Crick bases, respectively adenine (A) and guanine (G). Their pairing allows for molecular recognition of the bases. Besides hydrogen bonds, van der Waals forces play a decisive role in structure formation, specifically in the molding of the DNA double helix. The consecutive aromatic rings cause stacking interactions of adjacent bases, thereby stabilizing the helical conformation [9].

DNA Double-stranded DNA coils up into one of three forms, the A-, B- or Z-form, with B- and A-form being the most abundant in nature. The A-DNA forms within stretches of purines, the Z-DNA emerges in sequences in which purines and pyrimidines alter consecutively. Mixed sequences favor the B-form of the DNA. The disparities of the three types of DNA in the direction of winding and the number of bases per helical pitch result in varying groove formation propensity and rigidity

of the polymer. These features influence the interaction with proteins and vice versa. Certain DNA binding proteins force the DNA to change its conformation and certain conformations of DNA enhance the binding of proteins to DNA [10].

In nature, DNA is the primary hereditary material. Compared to RNA, DNA is less prone to self-cleavage, as the hydrogen atom at the 2' of the deoxyribose in DNA is less reactive than the hydroxyl group at the 2' of the ribose of RNA [11]. Additionally, the uracil in RNA is receptive to uracil to cysteine conversions. These properties render RNA an unfavorable material for long term information storage, but attractive for gene regulation processes [11].

DNA is used as a building material in DNA nanotechnology. The most frequently implemented subunit in DNA nanotechnology is the immobile Holliday junction. The characteristic of the branched nucleic acid structure is the double crossover, which, due to an asymmetric sequence, is incapable of branch migration. In the absence of cations, the DNA strands of the junction repel each other electrostatically and are a planar structure. Upon addition of multivalent cations, the backbones are electrostatically counterbalanced by a Stern layer and undergo a conformational change into a stacked conformation [12].

RNA Transcription Machinery The detailed description of "The molecular basis of eukaryotic transcription" won Roger Kronberg the Nobel Prize in Chemistry in 2006 [13]. RNA Polymerase (RNAP) binds to the promoter sequence upstream of the coding sequence. Transcription factors and helicases, which unwind the DNA, support the process of RNA synthesis. Transcription statistically results in one error per 10.000 bases. Correction processes secure the correction or the degradation in case of incorrectly transcribed RNA. Short poly(A)-tails can flag erroneous RNA for degradation.

Post-Transcriptional Poly(A) Tailing Poly(A) tailing is a post-transcriptional modification. The length and purpose of the tail varies depending on the organism. In eukaryotes, a long poly(A) tail, to which poly(A)-binding proteins bind, protects RNA from degradation. The proteins impose a steric hindrance for exonucleases. Furthermore the poly(A) tail is elementary for the initiation of translation [14]. In organelles and prokaryotes a short poly(A) tail of 20 to 50 nucleotides, added by poly(A) tailing polymerase, marks RNA for degradation [14,15]. However, in both, prokaryotes and eukaryotes, the removal of the poly(A) tail results in the initiation of RNA degradation [14].

RNA RNA serves a broad spectrum of functions, as indicated in Figure 1. Natural RNAs divide into messenger RNAs (mRNA) and non-coding RNAs (ncRNA), like the subunit of the ribosome (rRNA) and the adaptor RNA for amino acid recruitment (tRNA). *In vivo*, ncRNAs direct various regulatory functions ranging from sensing of environmental stimuli via ribozymes and riboswitches to DNA editing. Riboswitches are a combination of an expression platform and an aptamer, which senses changes in the concentration of an effector molecule [4]. The aptamer is a defined RNA architecture with specific tertiary structure, which binds a ligand molecules with high affinity and specificity [7].

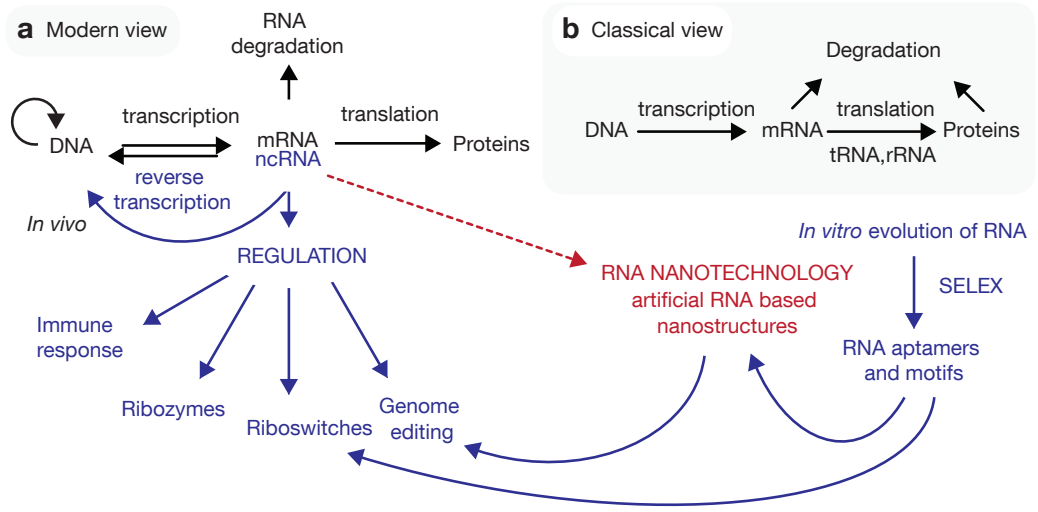

Figure 1. a, Modern and b, classical view of RNA. Image adapted, image by courtesy of A. Chopra [16]

RNA nanotechnology is an area of both *in vivo* and *in vitro* applications. RNA nanostructures are artificially designed RNAs which can include naturally occurring RNA aptamers or RNA aptamers evolved *in vitro* by SELEX.

RNA Aptamers Aptamers were discovered as parts of riboswitches and are conserved across all three domains of life. As the sensory part of the riboswitch, the aptamer reacts to altering effector concentration. Those induce a conformational change in the aptamer, which mostly resides at the 5' untranslated region upstream of the gene. The binding of the effector to the aptamer represses or activates the gene. A single aptamer can serve as a sensory element in a multitude of riboswitches and can control the translation of a variety of expression platforms. The aptamer region itself is a highly conserved domain.

Around 20 classes of natural aptamers with known ligands have been classified and several "orphan" classes have been identified by bioinformatic methods. These potentially function as aptamers of riboswitches, but for the "orphan" classes the ligand has not yet been determined [17].

In vitro aptamers were first evolved in an iterative directed evolution process, SELEX, in the early 1990s at the Gold and Szostak lab [18]. Similar to antibodies, aptamers bind to ligands with high affinity and specificity. In contrast to antibodies, aptamers offer the advantage of being non-immunogenetic and withstanding heat [19].

Ribonucleoproteins Ribonucleoprotein (RNPs) are RNA-protein complexes, like the ribosome. The ribosomal RNAs provide the enzymatic function and the ribosomal proteins are the catalyst for RNA function. The proteins serve as RNA chaperones, as they assemble the RNAs into functional shape [20]. Another set of RNPs, which are of specific relevance to this thesis, is the complex formed by guide RNAs and Cas proteins. The guide RNAs contain handles to which its according Cas protein can bind. The concept will be introduced in more detail in Chapter 1.4.

1.1.2. Proteins

The cell machinery translates mRNA into proteins, a biopolymer made up of amino acid chains. The ribosome mediates the condensation of single amino acid subunits into a chain. The folding process of the polymer is a complex pathway, and its proper realization is essential, as the correctly folded three-dimensional structure is decisive for the activity of the protein. Proteins of specific relevance to this thesis are membrane pores and the CRISPR-associated (Cas) proteins.

1.2. Lipid Membranes

Three of the most prominent functions of lipid molecules are the compact storage of energy, serving as a messenger molecule in signaling processes, and constituting barriers between compartments. The lipids that build up biological membranes require amphiphilicity: a polar and a non-polar moiety, the 'head' and the 'tail' group of a lipid [21]. Exposition of amphiphilic lipids to an aqueous environment results in entropically driven, spontaneous self-assembly into lipid substructures [22]. The compartments that are formed by lipids are the prerequisite for the subdivision of essential cellular processes.

1.2.1. Compartmentalization

Sealed compartments separate their internal constituents from the external environment, which is basal for the specific tuning of substance concentration. Encapsulation enables the decoupling of specific biochemical reactions and increases the efficiency of processes. It allows for the digestion of biopolymers by spatially delimited lysosomes, for the generation of energy in mitochondria and the confinement of DNA to the nucleus [23]. These features are essential for creating and sustaining large, cross-linked systems.

1.2.2. Determinants of Lipid Membrane Properties

The most abundant lipids in cellular membranes are glycerophospholipids [21]. A highly polar or charged group and two fatty acids are linked together by a glycerol moiety via a phosphodiester bond and ester bonds, respectively. The lipids subdivide into their polar or charged compound plus the glycerol, the 'head', and the two non-polar fatty acids, the 'tail'. The kind of lipid and their mixing ratio determine the properties of a membrane. Head and tail geometry define the packaging behavior and thereby the larger scale shape of the lipid assembly. The length of the tail and its degree of saturation influence the fluidity of the membrane. Based on lipid composition the functionality of the membrane alternates [21,24].

Packaging Parameter and Lipid Structures When exposed to water or oil, lipids tend to form structures that depend on the packaging parameter [25]. The packaging parameter is defined by the ratio of the tail volume to the product of tail length

and cross-sectional area of the head. For packing parameters below 0.3, lipids arrange in micelles when exposed to water. Packing parameters in the regime of 0.5 to 1 result in bilayer structures, whereas for a value larger than 1 inverse micelles are favored [26]. A bilayer unites the hydrophobic parts of two monolayers, sandwiching the hydrophobic layer between two hydrophilic ones and thereby shielding it from the aqueous phase. Structure formation in general is energetically favorable as it increases the number of microstates that can be occupied [9].

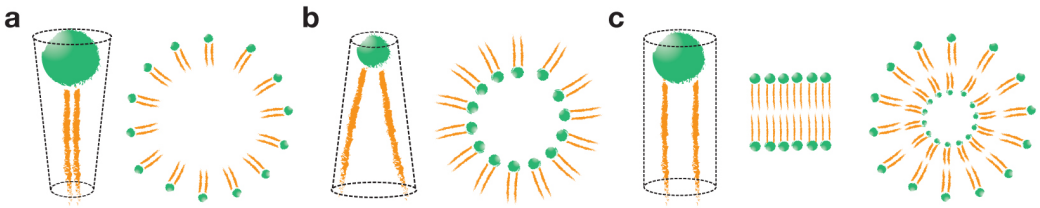

Figure 2. Amphiphilic lipids can form a, micelles for the packaging parameter $P = 1$, b, inverse micelles for $P < 1$, c, lipid bilayers and vesicles for $P < 1$. Image adapted, image by courtesy of S. Krishnan [27]

In living organisms, the cylindrical, conical, or inverted-conical lipids can support membrane modeling processes such as membrane fusion and fission [28].

Lipid Saturation and Membrane Fluidity Amphiphilic lipids subdivide into saturated lipids, where the non-polar carbon chain only contains single bonds, and unsaturated lipids, where the tail includes double bonds. A double bond, usually present in *cis*-form, introduces a kink in the hydrophobic tail, shortening the tail's effective length. Saturation influences the transition temperature. Double bonds and also short carbon tails decrease the temperature at which the lipids transition from the liquid crystalline phase into the rigid, ordered gel phase [21].

When required by environmental factors, bacteria and yeast produce a higher percentage of unsaturated lipids to counteract the decreasing fluidity of the membrane [21]. Furthermore membrane fluidity influences the exchange of molecules between the interior and the exterior and the diffusion coefficient of lipids and integral membrane proteins in it.

Lipid Composition and Functionalization of Membranes Another molecule that regulates membrane fluidity and permeability is cholesterol [21]. The cholesterol, predominately non-polar due to its steroid rings, is embedded in the hydrophobic

interior of lipid bilayers. Its polar hydroxyl group orients towards the polar head group of the lipids. The rigid, plate-like steroid rings partially immobilize the part of the hydrocarbon chain close to the lipid head. The presence of cholesterol stabilizes the membrane. At low temperature, the molecule serves as effective spacer between the lipids and thereby prevents phase transition. For the deformation of the membrane at high temperature in presence of cholesterol additional energy is required. Additionally, cholesterol-rich membranes are less permeable to water-soluble small molecules [29].

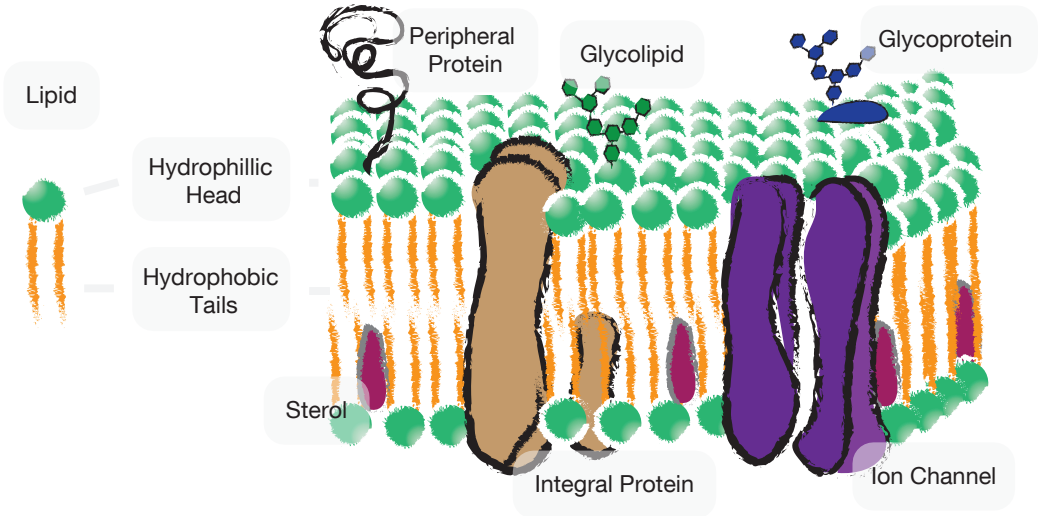

Figure 3. Membrane functionalized by integral and peripheral membrane proteins. Image modified, image by courtesy of S. Krishnan [27]

An assortment of functionalizing elements is depicted in Figure 3. Integral and peripheral proteins, glycolipids, glycoproteins and ion channels reside in or on the membrane. The membrane serves as the two-dimensional solvent in which lipids and proteins diffuse quickly in lateral direction [30,31].

1.3. Nanopores

Nanopores cause transmembrane conduits and subdivide into the subcategories natural protein pores, solid-state nanopores, and synthetic DNA membrane pores. The utilization of synthetic DNA membrane pores is described in detail in Chapter 2 and Chapter 3 and natural protein pores are used in Appendix A.

Solid-state nanopores are manufactured by drilling nanometer-scale holes with an ion or electron beam into a silicon compound membrane [32,33] or monolayer substrates [34,35]. In nature, protein nanopores evolved. Their *in situ* development *in situ* remains challenging [36]. Synthetic DNA membrane pores made up of DNA offer an additional design space, since the prediction of DNA folding into a three-dimensional shape is less computationally intensive than the modeling of the complex folding pathways of amino acid chains.

1.3.1. Protein Nanopores

Protein nanopores cover a broad spectrum of functions. The pores are selective or non-selective, laterally or unidirectionally permeable, and sensitive or insensitive to chemical signals or voltage drops across them. Nanopores are vital communication channels between the interior of a cell and its exterior environment and between the in- and outside of the nucleus.

Structure Nanopores self assemble from subunits into a final shape, topologically equivalent to a torus. Each of the subunits contains a hydrophobic moiety, which renders them capable of incorporating into bilayers. The assembly of the subunits is catalyzed by the presence of a lipid membrane. Four sample pores, alpha-hemolysin (aHl), *Mycobacterium smegmatis* porin A (MspA), the outer membrane lipoprotein CsgG and the ϕ29 DNA packaging motor are depicted in Figure 4. The pores consist of varying numbers of subunits, which form a vestibule and a stem. The vestibule introduces an asymmetry to the structure and a directionality to the pore. The narrowest confinement of the assembly is referred to as the constriction zone.

Exemplary Nanopores The aHl, a mostly heptameric pore and one of the model pores for nanopore experiments, belongs to the hemolysins, a toxic subgroup of nanopores. Secreted by *Staphylococcus aureus*, it causes detrimental damage to

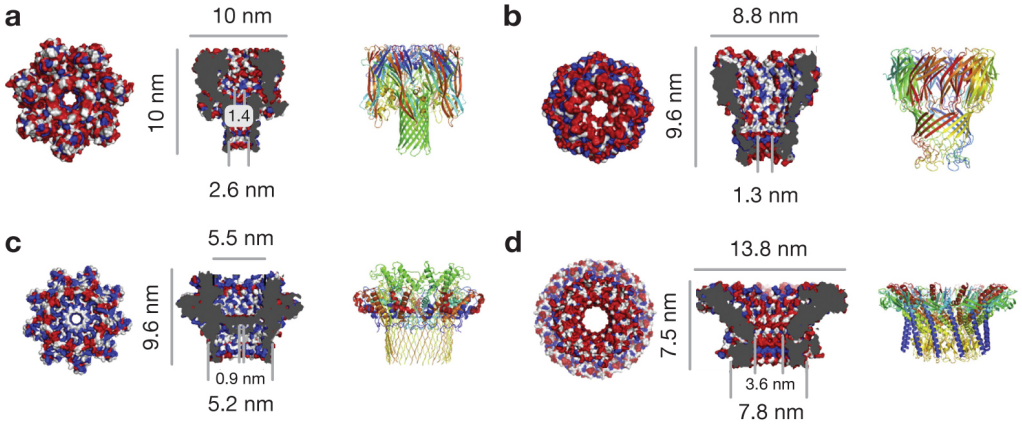

Figure 4. Top and side view of a, aHl b, MspA c, CsgG and d, ϕ29. Image adapted, image by courtesy of Wang *et al.* [37]

cells. Its incorporation into the cell membrane elicits an influx of calcium ions, which results in an imbalance of osmotic pressure and causes cell lysis [38].

The MspA, with octameric chalice-like structure, forms in the outer membrane of the bacterium to allow for the uptake of nutrients. It is a water-filled biological nanopore with a narrow constriction zone, a feature that brought it into the limelight of nanopore sequencing [39].

The general principle of resistive pulse sensing and the principle of nanopore sequencing will be introduced in Appendix A.

1.3.2. Synthetic DNA Membrane Pores

Synthetic DNA pore geometries draw inspiration from natural nanopores and use DNA as a building material. The DNA origami technique and the single-stranded tile (SST) approach allow for the design and folding of two and three-dimensional structures of almost arbitrary shape.

Folding DNA into Shapes The DNA origami technique evolved in 2006 and exploits the specificity and predictability which the hybridization of ssDNA offers [40].

The technique relies on a circular ssDNA which can originate from a M13 phage, termed 'scaffold', and ssDNA oligonucleotides, termed 'staples', which stack the scaffold together. The double crossover, two close and immobile Holliday junctions,

serves as a main building block. Appropriate routing of the scaffold and the choice of staples allow for the design of 2D and 3D structures.

The second approach for the design of DNA nanostructures is based on 'single-stranded tiles' (SSTs). Each single stranded DNA of 42 nt length forms a tile of 3 nm x 7 nm and hybridizes to four locally neighboring strands. A sheet of tiles can be shaped into a hollow SST tube by concatenating the boundary half-tiles into full tiles [41, 42].

Synthetic DNA Membrane Pores Both of the above-mentioned techniques were applied to design synthetic DNA membrane pores. The charged and hydrophilic DNA pores require hydrophobic modifications to render their incorporation into lipid bilayers energetically favorable.

For SST pores either the DNA backbone was chemically modified with 72 ethyl-phosphorothioate (PPT) moieties [43, 44], with two porphyrin moieties [45], or cholesterol staples were hybridized to the pore design [46, 47].

Most SST pores are based on the six-helix bundle barrel with a cross-sectional area of 4.9 nm^2. The pores vary in length from 9, over 13, 20 to 21 nm [43, 44, 46, 48] and exhibit conductance values of 0.25, 0.4 and 1.3 nS [43, 44, 46]. The six-helix bundle SST pore was administered to HeLa cells to inquire specificity of binding and its toxicity [48, 49]. An open and a closed version of the SST six-helix bundle pore were characterized and contrasted against each other [46].

A SST pore of similar length that is based on the square lattice has a cross-sectional area of 0.8 nm^2 and yields conductance values of 0.2 to 1.8 nS [47].

Published DNA origami pores show a higher degree of diversity in their shapes. The stem of the first generation pore and two pores evolving from the original design is, similar to most SST pores, based on the six-helix bundle [50, 51]. The translocation of ssDNA through DNA pores was first sensed with the first generation pore [50].

The design space was explored further and pores with large stem were created [51–54]. The larger stem allowed for the translocation of dsDNA [51] and of proteins [53]. Dye influx studies proved the size-selective translocation of molecules [51]. All-atom molecular dynamics simulations offered further insight into the ionic conductance of a large pore [52].

Functional versatility and a specifically large cross-sectional area were provided by the large size-selective DNA nanopore by Thomsen *et al.*. Three flaps, which

bury lipid moieties inside the structure, open upon actuation by external triggers. Further, the pores were hybridized together, forming a double pore. An additional blockage inside the pore lumen induces the size-selective blockage of molecules and is reversed through the addition of a key DNA [54].

Table 1 summarizes information about DNA pores which incorporate into bilayers. It lists the folding approach, the kind and number of hydrophobic modifications, their theoretical cross-sectional area, the length of the stem and their conductance values. The theoretical cross-sectional areas here were calculated based on plain geometrical assumptions, whereas the listed stem lengths are values that are stated in the papers cited.

	Technique	Anchor molecule	Anchor: mod. element	Number of anchors	Lattice	A / nm^2	l_{stem} / nm	G / nS
Langecker *et al.* [50]	DNA Origami	Cholesterol	Staples	26	hc	4.9	42	0.89
Burns/ Seifert *et al.* [43–46, 48]	SST	Ethane-PPT/ Porphyrin/ Cholesterol	DNA backbone / Staples	72 / 2	hc	4.9	21/ 20/ 13/ /-/ 9	0.4/ 0.25/ -/ /-/ 1.3
Göpfrich *et al.* [47]	SST	Cholesterol	Staples	2	s	0.9	11	0.2 - 1.8
Krishnan *et al.* [51]	DNA Origami	Tocopherol/ Biotin-Streptavidin	Staples	26/ 57/ 57	hc/ hc/ s	4.5/4.5/15.5	42/ 28/ 32	1.6/ 1.5/ 3.1
Göpfrich *et al.* [55]	dsDNA	Porphyrin	DNA backbone	6	-	-	5	0.08
Göpfrich *et al.* [52]	DNA Origami	Cholesterol	Staples	19	s	36	54	30
Diederichs *et al.* [53]	DNA Origami	Cholesterol	Staples	24	s	56	46	2.37
Thomsen *et al.* [54]	DNA Origami	Cholesterol & Palmitoyl	Staples	18 & 28	hc	87	32	-

Table 1. Synthetic DNA membrane pores which incorporate into lipid bilayers. Listed are the folding technique, the element that carries the modification, the number of hydrophobic elements, the theoretical cross-sectional area of the pores, the stem length and the conductance values.

1.4. Clustered Regularly Interspaced Palindromic Repeats

Bacteriophages are viruses that infect bacteria and archaea. Receptors mediate the docking of the phage and DNA is injected subsequently into the host. The phage can directly replicate in the cell, which leads to the lysis of the bacterium. An alternative pathway for the phage DNA is to integrate into the DNA of the host for later reactivation, which results in the lysogenic cycle.

In bacteria, a defense mechanism against phages evolved. This adaptive bacterial immune response termed CRISPR bases on two main mechanisms. The system stores information regarding invasive phages in the DNA of the bacteria and it eliminates phages based on the recorded information. Both processes rely on CRISPR associated proteins, the Cas proteins [56].

Bacteria that survive an attack of a phage unbeknown to their defense mechanism integrate specific, approximately 30 base pair long snippets of the phage DNA in their genome. The Cas proteins Cas1 and Cas2 enable the integration and thereby build up a library of DNA segments. Additional to the phage DNA, each of the segments includes a conserved DNA sequence that transcribes into a specific handle structure. A Cas protein (eg Cas9 or Cas12a) binds to the handle, forming a functional RNP [57].

The RNP screens DNA for short stretches of conserved DNA motives, called protospacer adjacent motives (PAM). PAM recognition induces severe DNA bending, which leads to the unwinding of DNA and the formation of a seed bubble. In case the protospacer, the part of the RNA which was transcribed phage DNA, is complementary to the DNA, it hybridizes to the DNA strand and displaces the other one. Full R-loop formation activates the DNase function of the Cas protein. That way phage DNA is cleaved and rendered dysfunctional by the CRISPR system.

Discovered by Ishino *et al.* in 1987 and extensively studied by Mojica in the 1990s [58], CRISPR gained popularity in the field of gene editing when the Doudna lab simplified its usage. The CRISPR system became an analytic tool and a tool for gene-regulating and editing. In mammalian cells, gene editing was successfully demonstrated by the Zhang lab, which paved the way for further advances [59].

1.4.1. DNase Deactivated CRISPR Associated Proteins

Mutant versions of the Cas protein, devoid of the DNase activity (dCas9/dCas12a), can be used to interfere with gene activation and repression. The RNP there binds the target DNA without cutting it. DNase deactivated dCas9 or dCas12a can serve as a roadblock for polymerases, thereby suppressing the expression of the targeted gene. The deactivated protein fused to a transcription factor can activate or enhance transcription [60].

1.4.2. Cas proteins Cas9 and Cas12a

The Cas9 and Cas12a function similarly. Both proteins form RNPs with a guide R NA which contains the protospacer and the handle for the binding of the protein. They require a PAM for initial R-loop formation and a matching protospacer to initiate specific binding.

Their differences are illuminated in Figure 5. Cas9 belongs to the CRISPR II system. There CRISPR RNA (crRNA) and the transactivating RNA (tracrRNA) hybridize to form the full guide RNA. The handle of the sgRNA for Cas9 is located at the 3' end of the protospacer and includes a hairpin, the nexus, and the *S. pyogenes* terminator, which in total adds up to a structure of approximately 80 nt length (Fig. 5 b). Alterations in the sequence of the handle change the efficiency of protein binding. In the Doudna lab, a guide RNA was designed in which the cr- and traRNA are connected via a hairpin loop. Only a single guide RNA (sgRNA) hence is required for binding of the RNP to DNA. The sgRNA-Cas9 complex (Fig. 5 a) binds to the protospacer adjacent motif NGG which is found at the 3' end of the protospacer and cuts a blunt end between base three and four, counting from the nucleation site [61] (Fig. 5 c). The process is magnesium-dependent. In the absence of guide RNA and the presence of manganese ions, Cas9 degrades ssDNA [62].

The Zhang lab identified Cas12a as a viable alternative to Cas9. The protein is slightly smaller in size, 156 kDa for Cas12a of *Acidaminococcus* compared to the 170 kDa for Cas9 of *Streptococcus pyogenes* and the guide RNA is shorter. It has been reported that Cas12a is more susceptible to mismatches. The sensitivity for mismatches for Cas9 includes the eight bases in proximity to the PAM region, whereas Cas12 remains sensitive to the first 18 bases proximal to the PAM, a feature that has been attributed to features of the Cas12a [63].

The structure and guide RNA of Cas12a are shown in Figure 5a and b, respectively.

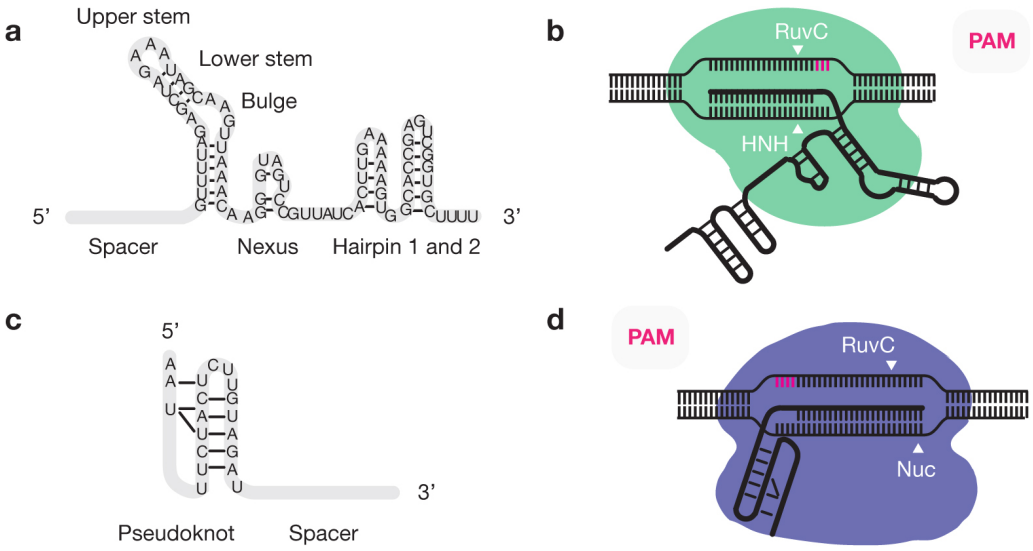

Figure 5. a, Secondary structure of Spy Cas9 sgRNA. b, Schema of sgRNA-Cas9 ribonucleoprotein binding to its target DNA. Arrows indicate the cutting sides. c, Secondary structure of As Cas12 c12aRNA. b, Schema of c12aRNA-Cas12 ribonucleoprotein binding to its target DNA. Arrows indicate the cutting sides. For b and d, image curtsey of Pickard *et al.* [64]; a and c based on Chen *et al.* [65]

The handle of the RNA is at the 5' end of the protospacer, the PAM (5' TTTN 3') resides at the 5' end of the protospacer. The 18 nt long handle which recruits Cas12a forms a hairpin and a pseudoknot. Cutting occurs in a staggered manner, between base 16 and 17 at the DNA strand equivalent to the protospacer and 7 to 9 bases further upward on the complementary strand (Fig. 5 f) [63]. Cas12a requires magnesium ions to initiate substrate binding [66]. In the absence of guide RNA and the presence of manganese ions, it is capable of nicking double-stranded plasmids and degrading ssDNA [62]. Additionally, Cas12a possesses RNase activity. It processes its guide RNAs and cuts them at the 5' end.

2. Characterization of DNA Pores

Four synthetic DNA membrane pores - termed the large connector, the pin pore, the wheel pore, and the T pore - were incorporated into droplet interface bilayers. Their conductances were recorded and compared with the theoretically expected values, which arise out of the pore dimensions. The pin, the wheel, and the T pore were examined regarding their tendency to incorporate. The wheel pore and the T pore were shown to fold in the condensing agent spermidine. Further, the insertion of the T pore into the bilayer was performed with two novel kinds of anchor molecules.

The characterization of the pores is based on the paper 'Molecular Transport through DNA Origami Pores' [51].

2.1. Theory

Chapter 2.1 will elaborate on introductory information about transmission electron microscopy [67], gel electrophoresis, the droplet interface bilayer (DIB) technique, and the conductance of nanopores.

2.1.1. Transmission Electron Microscopy

The direct observation of specimen is limited by the Rayleigh criterion

$$d = \frac{0.61\lambda}{n \sin\theta} = \frac{\lambda}{2\,NA}. \tag{2.1}$$

It is described by the refraction index n, the semi-angle of opening θ, the wavelength λ, and the numerical aperture NA of the lens which collects the light. The Rayleigh criterion constitutes that the distance between the maxima of the point spread functions of two resolvable points has to be equal or larger than the distance between maximum and minimum of the individual point spread functions. Light microscopy limits the resolution of the specimen to approximately 200 nm. Images of higher resolution are obtained using accelerated electrons. The wavelength of electrons depends on their energy and was described by de Broglie via the impulse of the electron p_e and therefore the mass m_e and charge e of an electron:

$$\lambda = \frac{h}{p_e} = \frac{h}{m_e v_e} = \frac{h}{m_e\sqrt{2\frac{e}{m_e}V}} \tag{2.2}$$

with h being the Planck constant. The kinetic energy of an electron here is defined by its acceleration energy and depends on the accelerating voltage V. A TEM with an accelerating voltage of a 100 kV delivers an electron wavelength of 4 pm, and with a thin opening angle of an electron beam of 1 to 2°, a maximum theoretical resolution of 0.1 to 0.2 nm can be reached. The inelastic impact of the electrons causes radiation damage and is one of the factors which decrease resolution [67].

Electrons are scattered with a probability proportional to the square of the atomic number. The difference in the interaction of electrons with the background and the interaction with the specimen creates contrast in TEM. DNA samples, which mainly consist of low atomic number atoms like carbon, hydrogen, and phosphate, are imaged on a carbon-coated polymer grid. Either the DNA structures are made to absorb a substance of high atomic number or the grid is coated with one [67].

Three-dimensional DNA origami structures are primarily assessed by the second method, the negative staining method. There, the bright specimen is accentuated on an uranyl formeat coated, dark background. At high resolution, the honeycomb or square lattice routing of the origami and the inter-helical areas of DNA origami can be identified. The non-native environment can distort of the structures. Furthermore, structures might be exposed to surface interaction forces, thereby flattening the structures.

Cryo-EM overcomes these obstacles. Here, the randomly orientated, unstained specimen is imaged in amorphous water. Its random orientation and the smaller degree of radiation damage allow for the three-dimensional reconstruction of the specimen at high resolution.

2.1.2. Agarose Gel Electrophoresis

Agarose gel electrophoresis separates molecules by an electrical force. Agarose, which is a natural linear polymer originating from seaweed, is known to form a porous matrix when melted and cooled. The mesh size of these pores depends on the polymer concentration. When an electric field is applied across an agarose gel submerged in ionic buffer, the gel matrix allows for the separation of molecules. The molecules move with velocity v which depends on the electric field E, their net charge q and the frictional coefficient f [68]:

$$v = \frac{q\,E}{f} \tag{2.3}$$

The f depends on factors like the molecular shape and the charge to mass ratio and accounts for the frictional resistance of the molecule [69, 70]. Agarose gel electrophoresis as a method serves analytical and preparative purposes.

2.1.3. Droplet Interface Bilayer

In nanopore force spectroscopy experiments a voltage gradient is established across a nanopore and the response current is measured. The experiments require an insulating membrane into which the nanopore can incorporate and which separates two aqueous solutions. Figure 6 illustrates the formation of a membrane with the DIB setup.

Two aqueous droplets are immersed in a lipid oil mix, whereby a lipid monolayer

forms around each droplet (Fig. 6 a). When those are brought in close proximity (Fig. 6 b), the two monolayers assemble into a lipid bilayer (Fig. 6 c). The bilayer separates the solution of the two droplets and when a voltage is applied across the bilayer, it insulates against current flow. The bilayer offers a platform for pore incorporation (Fig. 6 d). Across pores that are incorporated into the bilayer, a voltage can be applied. The ions translocating through the pores leave a detectable signature current. At standard conditions, electrophoretic translocation of analyte molecules through the pores leads to a steric exclusion of buffer ions and hence to a reduction of current. The reduction depends on the charge and size of the analyte, hence pores can be used as stochastic single-molecule sensors [71, 72].

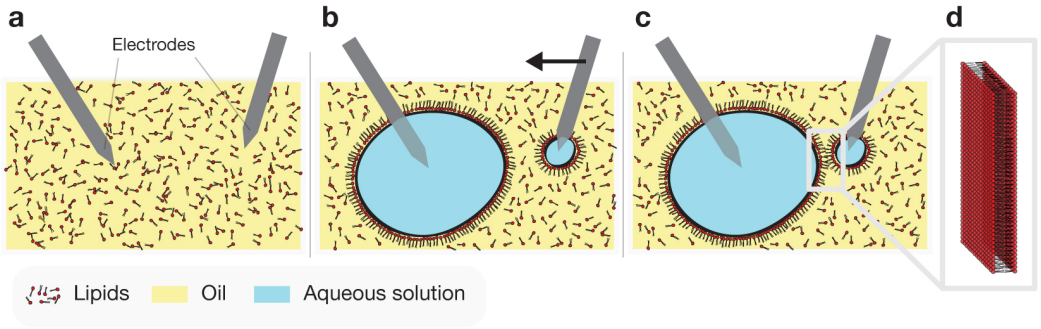

Figure 6. Generation of a droplet interface bilayer. a, Placement of electrodes in lipid containing oil, b, adding aqueous solution onto the electrodes and c, bringing droplets in close proximity leads to d, bilayer formation. Image adapted from D. Ziegler [73].

2.1.4. Conductance of Nanopores

The smallest constriction of the pore predominantly defines the current through a pore. The geometry of the pore determines the conductance of the pore.

Conductance - Theory The conductance is a defining pore characteristic. The total resistance of a nanopore is defined by the Hille equation [74]:

$$R_{total} = R_{pore} + 2R_{access} = \frac{l}{\sigma A} + \left(\frac{\pi r}{2\sigma \cdot A}\right) \tag{2.4}$$

where A denotes the cross-section of the stem and r its respective radius. The conductance expected for a simple pore with length l, diameter d, and minimum

cross-sectional area A equals [75]

$$G \equiv \frac{1}{R} = \frac{\sigma\, A}{\frac{\pi d}{4} + l} \approx \frac{\sigma\, A}{0.8d + l} \tag{2.5}$$

In this thesis conductance values were recorded at 1 M KCl. At 1 M KCl electrolyte, the conductivity σ equals $9.9\,\frac{\mathrm{S}}{\mathrm{m}}$ [76].

The measured conductance values allow an approximation of A:

$$A_{exp} \approx \frac{G\,(0.8d + l)}{\sigma}. \tag{2.6}$$

The value of A_{exp} can be discussed in regard to the theoretical cross-sectional area A_{th}. A A_{th} is defined by the DNA helices, which confine the cross-sectional area. For the calculation of the A_{th} the DNA double helices are considered as solid rods, with a diameter of 2 nm. The calculation is based on the design of the DNA structure.

2.2. Materials and Methods

Experiments with natural and synthetic DNA membrane pores require a lipid bilayer for incorporation. The synthetic DNA membrane pores were first designed *in silico*. The folding is examined by transmission electron microscopy and agarose gel electrophoresis. The droplet interface bilayer setup serves as a platform for pore incorporation.

2.2.1. DNA Origami Pores

Utilized in this thesis were the pin, the wheel, and the T pore and the large connector, designed by T. Martin and S. Krishnan, respectively.

The pores were designed with ssDNA extensions, which are complementary to adaptor strands at the bilayer facing side of the pore. The adaptor strands are ssDNAs which are modified at the 5' end with either tocopherol, biotin or elastin-like polypeptides (ELPs). The modifications are located in proximity to the vestibule of the pore.

Pin Pore The pin pore bases on the first version of synthetic DNA membrane pores [50]. In the original design, a free dangling scaffold loop resided at the top of the vestibule. For the pin pore, the scaffold loop was re-routed and transformed into a part of the well-defined DNA structure. The routing resulted in a design in which the vestibule is an additional 22 nm higher than in the first design. The 7249 nt long M13mp8 scaffold is folded into 54 double helices with 26 positions for anchoring modifications. The pore dimensions and TEM images of the pin pore are shown in Figure 7 a, b, c, and d.

Wheel Pore The wheel pore is a pore in which the vestibule offers more positions for anchoring elements. It was designed based on the 8064 nt DNA scaffold and accommodates 57 positions for anchor strands. Three large cavities were integrated into the large vestibule to obviate the burial of double helices in the structure. The cavities were introduced to obtain a high yield of nanostructures. Dimensions and TEM images are displayed in Figure 7 d, e, f, and h.

T Pore The T pore is the design which is predominantly used in this thesis. It consists of a square-shaped plate with a hole in the center and a stem that is

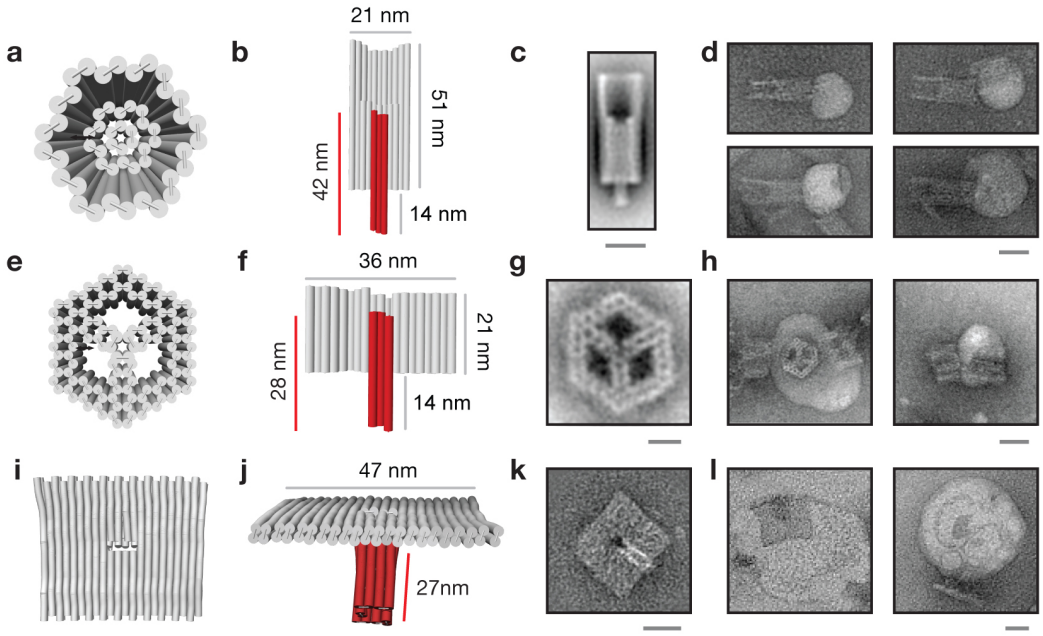

Figure 7. Three pore designs a, Schematic top and b, side view of the pin pore. The dimensions of the vestibule are shown and the transmembrane stem is indicated in red. c, TEM class averaged image of the side view of the pin pore. d, TEM images of tocopherol-modified pin pores interacting with small unilamellar vesicles (SUVs). e, Schematic top and f, side view of the pin pore. g, TEM class averaged image of the top view of the wheel pore. h, TEM images of tocopherol-modified wheel pores interacting with SUVs. i, Schematic top and j, side view of the pin pore. k, Class average of the plate and the stem of the T pore. k, TEM images of tocopherol-modified T pores interacting with SUVs. Scale bars: 25 nm. Image adapted from Krishnan *et al.* [51].

attached beneath the hole. Figure 7 i and j show its dimensions and TEM images of the T pore. The TEM images of T-pores that are incorporated into vesicles indicate bending of the plate along with the curvature of the vesicle (Fig 7 l). The images indicate a tight binding of the plate to the lipid bilayer. The structure is based on the 7560 nt long scaffold. The plate, which contains two layers, offers space to accommodate 57 anchoring modifications.

Large Connector The large connector is a pore of comparatively large diameter, with the potential to connect two compartments. It imitates the capability of natural occurring connexin to bind to and incorporate into membranes on two sides of the channel. Connexins enable the exchange of molecules between two vesicles. The large diameter is supposed to allow for increased translocation rates.

The structure is based on the 8064 nt scaffold. It accommodates 50 helper helices for modifications on both sides, which enable incorporation and additionally allow for specific binding. Folding the structures with different types of modification strands on each of the sides offers design space for side specific binding.

Folding and Purification DNA nanopores were assembled by mixing 50 nM scaffold strands with a three-fold excess of staple strands in a folding buffer containing 1× Tris-acetate-EDTA buffer (TAE, Carl Roth GmbH) and 20 mM $MgCl_2$ and applying a temperature ramp from 60 °C to 30 °C unless stated otherwise. Excess staples were removed by several rounds of filtering with a 100 kDa Amicon filter (Millipore). The scaffolds were modified versions of the single-stranded M13 phage genome, which were expressed and purified by the F. Praetorius [77]. Staple strands (Eurofins) were designed with the caDNAno software [78].

The structures were incubated with the adaptor strands, which were modified at the 5' end with either tocopherol (biomers.net GmbH), biotin (biomers.net GmbH) or ELPs (see below) for 45 min at room temperature. The tocopherol-modified strands were heated for 45 min at 60 °C before incubating the strands with the pores.

The sample was filtered with a 100 kDa Amicon filter. For experiments requiring biotin-streptavidin bridges, the DNA nanopores were incubated with biotinylated staples and purified with a 100 kDa Amicon filter. The pores were incubated with excess streptavidin directly in the filter. The structures were filter purified several rounds. The final concentration of DNA nanopores was determined by absorbance measurements with a nanophotometer (Implen GmbH).

2.2.2. Agarose Gel Electrophoresis

The DNA nanopores were analyzed with agarose gel electrophoresis with 0.8 % to 2 % w/v agarose gels in 1 x TAE, 21 mM at 100 V unless stated otherwise. The gel was subsequently stained with SYBR gold (Invitrogen, Thermo Fisher Scientific) for 15 min and scanned with a Typhoon FLA 9500 laser scanner (GE Healthcare). The images were analyzed with ImageJ (Image Processing and Analysis with Java).

2.2.3. Transmission Electron Microscopy

Collodion-supported carbon-coated Cu400 TEM grids were glow-discharged, incubated with 5 µl of 5 nM DNA nanostructures for 30 s, washed with 5 µl of 2 % aqueous uranyl formate solution containing 25 mM sodium hydroxide and incubated with 20 µl of it for 30 s. The grids were air-dried for 20 min. The negative stained nanopores were imaged with a CM100 electron microscope (Philips) operated at 100 kV at 28 500 fold magnification using an AMT 4 Megapixel CCD camera.

2.2.4. Droplet Interface Bilayer Setup

For the DIB experiments in this thesis, the synthetic lipid 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC, Avanti Polar Lipids) was used. This lipid possesses a high mechanical and chemical stability, a high electrical resistance [79], is stable against oxidation [80], and shows no phase transition between −120 °C to 120 °C [81]. It is a favorable choice for DIB measurements as the application of voltage strains the bilayer and demands for lipids which possess high stability [82, 83]. Current traces were recorded with an EPC 9 double amplifier (HEKA Elektronik GmbH).

2.2.5. Click Chemistry

ELPs are biopolymers that consist of the amino acid chain (glycine-X-glycine-valine-proline)$_n$, (G-X-G-V-P)$_n$ [84]. X denotes a variable amino acid and n the number of repeats of the five amino acid sequence. The transition temperature T_t of ELPs is one of their defining properties. It denotes their transition from a non-collapsed hydrophilic state to a collapsed hydrophobic state. The properties of the ELP are defined by the amino acid X and the surrounding salt conditions. It this thesis an ELP, provided by M. Götzfried and K. Vogele, was used, with leucine (L) as X. The MGH-(G-L-G-V-P)$_{40}$-GWP is an ELP that is hydrophobic at standard salt conditions at room temperature. The experimental conditions of 1 M KCl lower T_t further.

ELPs (10 µM final concentration) were incubated with 10 fold excess of N- Hydroxysuccinimide (NHS) in 100 µl PBS for 3 h at 4 °C. Samples were filtered twice with a 30 kDa filter (Amicon) with 1 fold PBS at 10 000 rms. The protein concentration was determined with a nanophotometer.

The NHS-modifed ELPs (approximately 10 µM final concentration) were incubated with alkyne modified DNA strands (modification C8-Alkyne-dU-CEP, 10 µM final concentration), Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, Sigma-Aldrich, 100 µM final concentration), Tris-(2-Carboxyethyl)phosphine (TCEP, Thermo Fisher Scientific, 1 mM final concentration), and $CuSO_4$ (1 mM final concentration) in 1 fold PBS overnight at 4 °C [85].

2.3. Results and Discussion

The conductance values of the pin pore, the wheel pore, the T pore, and the large connector were determined by electrical measurements with the DIB setup. In cooperation with N. Matejka [86] spermidine mediated folding as demonstrated in Chopra *et al.* [87] was performed for the wheel and the T pore.

2.3.1. Magnesium and Spermidine mediated Folding

The folding process was screened for the optimal $MgCl_2$ concentration and for the temperature interval which is most decisive for structure assembly. The standard folding condition is 50 nM scaffold, which is folded at 20 mM from 64 to 44 °C [88]. The reference structures (lane 3 in Fig. 8) assemble at the standard folding condition.

For the wheel pore, the comparison to the reference structure shows an optimal magnesium concentration of approximately 10 mM $MgCl_2$ (Fig. 8a). The temperature screen reveals that the majority of folding occurs between 60 °C and 51 °C (Fig. 8b).

As the scaffold and the folded structure run with similar velocity through the gel matrix, the folding is harder to assess for the T pore. Any concentration of $MgCl_2$ of 10 mM and more appears suitable (Fig. 8c). The main folding process proceeds between 60 °C and 53 °C (Fig. 8d).

Figure 8f shows that the large connector folds best at 10 to 15 mM $MgCl_2$. The majority of folding occurs between 60 °C and 53 °C (Fig. 8e).

Spermidine Folded Wheel Pore and T Pore Spermidine is a trivalent, naturally occurring polyamine, which at neural pH is a positively charged carbon chain with amino groups. The positive charges allow for spermidine mediated folding of DNA nanostructures. It offers an alternative to a folding process mediated by magnesium ions. Spermidine naturally serves as a condensing agent for DNA, as it is a means to overcome interhelical repulsion. Creating bridges between DNA helices, it interacts differently with DNA than magnesium cations [89–91].

The spermidine concentrations required for folding DNA structures have proven to in the range of spermidine concentration at certain physiological conditions [87]. Depending on the cell cycle state and the kind of cell, spermidine concentration in

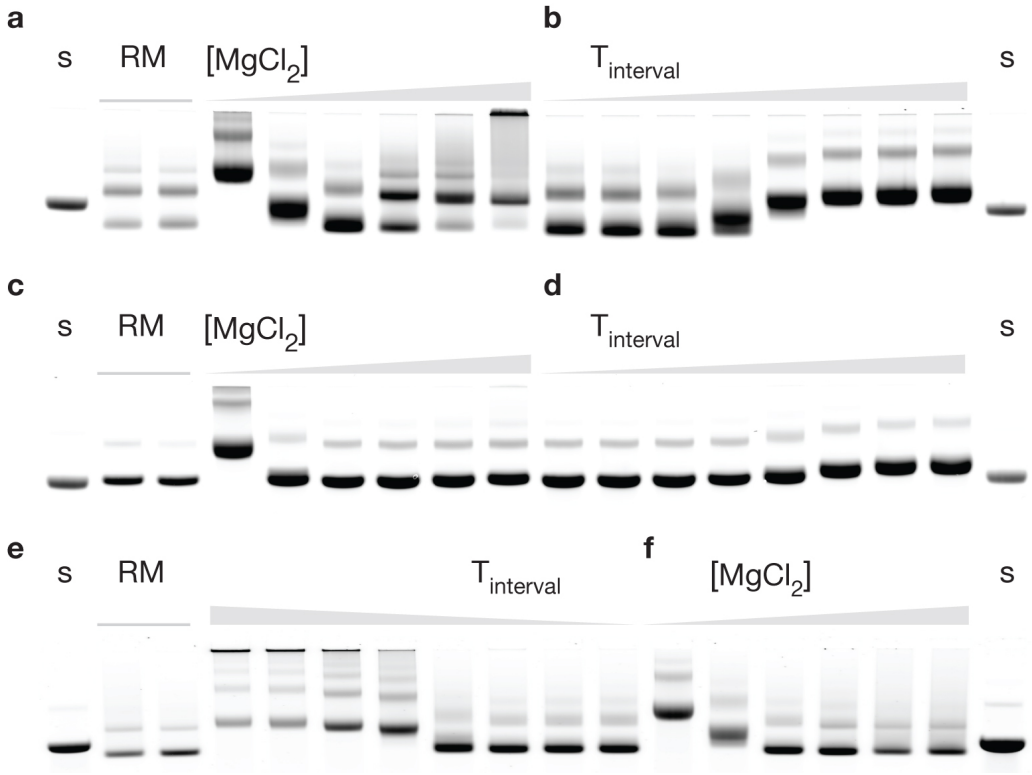

Figure 8. Folding screen of the wheel pore (a,b), the T pore (c,d) and the large connector (e,f). Samples from left to right: scaffold (s); the reaction mix (RM) at 20 and 50 nM scaffold concentration folding at 20 mM from 64 to 44 °C. In a, c, and f structure were folded at various $MgCl_2$ concentrations (0, 5, 10, 15, 20, 25, 30 mM) from 64 to 44 °C. In b, d, and e, the structure were folded in 20 mM at a range of temperature intervals ([50 °C to 47 °C], [52 °C to 49 °C], [54 °C to 51 °C], [56 °C to 53 °C], [58 °C to 55 °C], [60 °C to 57 °C], [62 °C to 59 °C], [64 °C to 61 °C]). Images a, b, c, d, by courtesy of Fabian Schneider. In e, temperature intervals decrease from the highest to the lowest interval.

the cell is at the order of 2 mM [92]. Free spermidine concentration ranges from around 15 to 200 µM [93,94]. DNA nanostructures that were folded in spermidine can be electroporated into bacteria. Exposing DNA origami folded with $MgCl_2$ to the voltages required for electroporation is detrimental to the nanostructures, whereas structures folded with spermidine survive the exposition to those voltages [87].

It here was screened for the optimal range of spermidine concentration required for folding the wheel pore and the T pore. TEM images indicate that the wheel pore folds around 1.4 mM to 1.6 mM spermidine (Fig. 10a). Structures imaged at 1.0 mM show improper folding. The batch folded for the gel folding screen appears

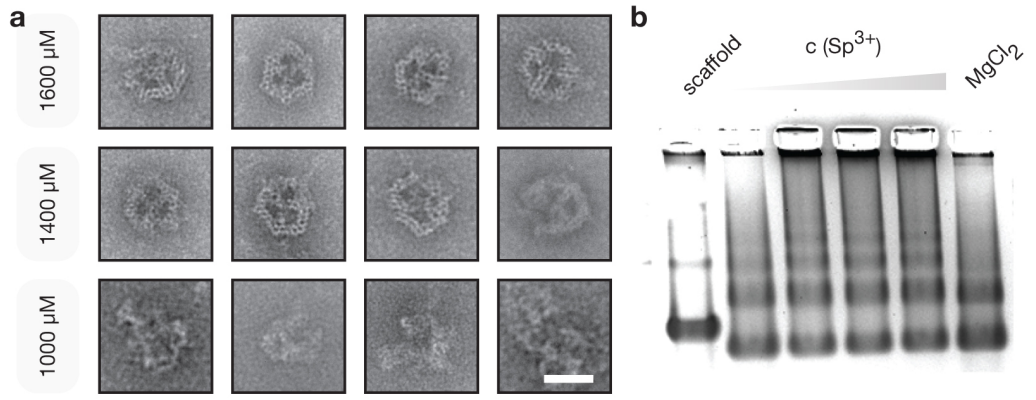

Figure 9. Folding of wheel pore with spermidine a, TEM images of wheel pores folded at a spermidine concentration of 1.0 mM, 1.4 mM and 1.6 mM. Structure folding was initiated at 1.4 mM, and 1.6 mM. Structures at 1.0 mM do not exhibit structural elements of the wheel pore. Scale bar: 25 nm b, Agarose gel electrophoresis of wheel pore structures folded at various spermidine concentrations, from left to right: scaffold, 1.0 mM, 1.2 mM, 1.4 mM, 1.6 mM, positive control was folded at 20 mM $MgCl_2$

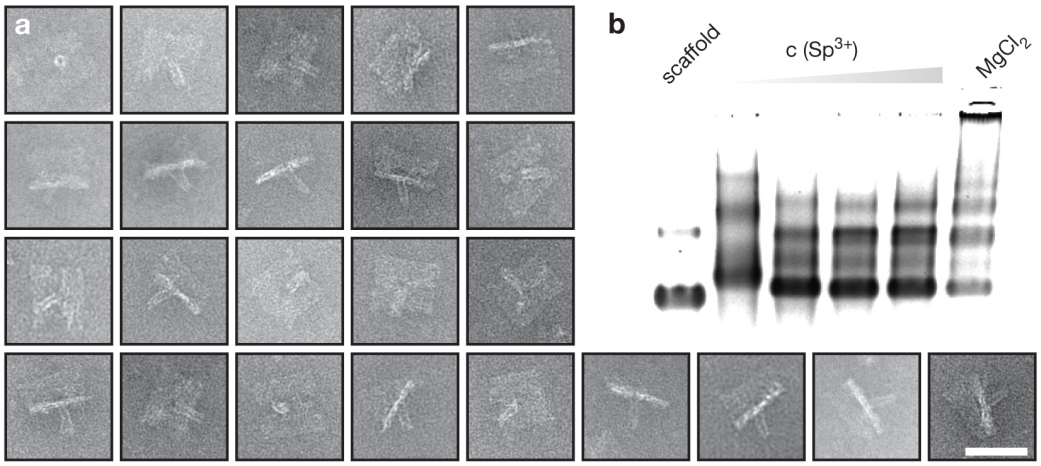

Figure 10. Folding of T pore with spermidine a, TEM images of T pores folded at 1.0 mM spermidine. The folding products reflect the structural details of the T pore. Scale bar: 50 nm b, Agarose gel electrophoresis of T pore structures folded at various spermidine concentrations, from left to right: scaffold, 1.0 mM, 1.2 mM, 1.4 mM, 1.6 mM, the positive control was folded at 20 mM $MgCl_2$

less sensitive regarding the spermidine concentration (Fig. 9 b). Here the DNA structures fold at concentrations ranging from 1.0 mM to 1.6 mM.

The gel folding screen of the T pore which was performed with spermidine con-

centrations ranging from 1.0 mM to 1.6 mM, indicates proper folding of the T pore between 1.2 mM to 1.6 mM spermidine (Fig. 10b). Folding of T pores at 1.0 mM spermidine results in structures which appear well defined on TEM images (Fig. 10a).

Discussion Spermidine concentrations required for the folding of the Rothemund rectangle, the hollow origami tube, and the 42 helix bundle range between 800 µM to 1 mM [87]. The wheel pore and the T pore are DNA structures that contain more DNA than those previously examined structures. This aligns with the wheel pore and the T pore requiring spermidine concentrations in the range of 1.0 mM to 1.6 mM, since spermidine needs to balance the phosphate charges of these larger structures.

For both pores, the samples imaged with TEM and those which were analyzed with gel did not originate from the same batch. The deterioration of spermidine over time [95] could offer an explanation for the different spermidine concentration ranges required for the TEM imaging sample and required for the samples of the gel folding screen.

The spermidine-mediate folding offers a strategy to electroporate DNA pores into living cells, where they could interact with phospholipid membranes.

2.3.2. Conductance of Pin, Wheel and T Pore

The probability to record incorporation for the pin pore, the wheel pore or the T pore differed.

The pin pore showed a very low incorporation probability, and thus only one I-V curve was obtained (Fig. 11 a). Higher incorporation frequency and better statistics were obtained for the wheel pore (Fig. 11 b, $N = 9$) and the T-pore (Fig. 11 c, $N = 45$).

Conductance of Pin, Wheel and T-Pore The conductance that was measured for the pin pore is 1.6 nS. For the pin pore only one incorporation into the membrane was recorded (Fig 11 d). The comparatively low number of hydrophobic modifications could be the reason behind the low insertion frequency.

The 57 hydrophobic anchors of the wheel pore resulted in superior membrane interaction. However, the wheel pore exhibited the tendency to aggregate in solution.

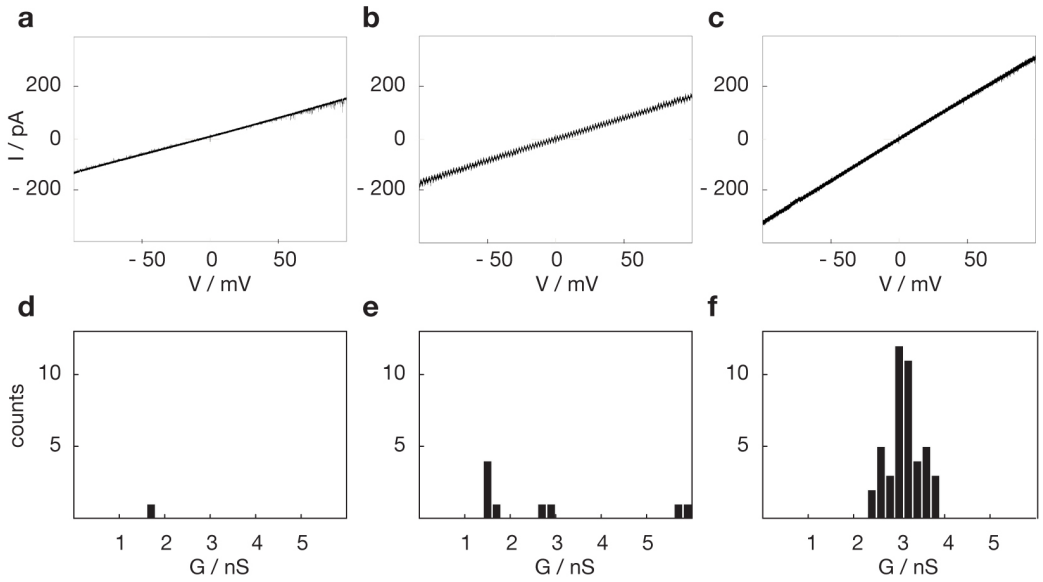

Figure 11. I-V responses for a, the pin, b, the wheel and c, the T pore. Histogram of the measured conductance values for d, the pin, e, the wheel and f, the T-pore. Adapted image by courtesy of Krishnan *et al.* [51].

TEM images frequently display groupings of two, three, four or more wheel pores. Correspondingly, higher conductance values (Fig 11 e) were recorded, which can be attributed to the incorporation of several pores. Based on the observation of several wheel pores stacking together and based on the conductance values of 1.5 nS, 2.8 nS, and 5.8 nS, we conclude that 1.5 nS are equivalent to the conductance of one pore. The values around 2.8 nS and 5.8 nS correspond to two and four pores, respectively.

The comparatively large cross-sectional area of the stem of the T pore resulted in a higher conductance value. The conductance of 3.1 ± 0.3 nS reflects the geometrical properties of the pore. Moreover, the T pore displayed the highest membrane insertion frequency and mostly incorporated as a single pore. Figure 12 depicts a rare example of nine T pores that are simultaneously present in the bilayer. The total conductance equals 27 nS.

Ahead of the recording of Figure 12 the size of the bilayer was steadily reduced by increasing the distance between the central points of the two droplets of the DIB setup. The size of the bilayer decreased steadily, as it was split into its two monolayers. The pores remain inside the bilayer for the entire separation process ahead of the recording of Figure 12. Assuming the pores were distributed across

the bilayer before starting to split the droplets, the observation implies that the pores were laterally mobile.

Figure 12 records the 30 ms before the complete separation of the droplets. There a voltage ramp of $1\,\frac{\mathrm{V}}{\mathrm{s}}$ was applied starting at $+100\,\mathrm{mV}$. Throughout 10 ms the nine pores disappear in serially from being recorded. Each step indicates the removal of one pore, which contributed to the total current before removal. Each of the steps allows for the derivation of a single pore conductance. The according conductance of every single pore is indicated next to each step, ranging from 2.9 nS to 3.2 nS.

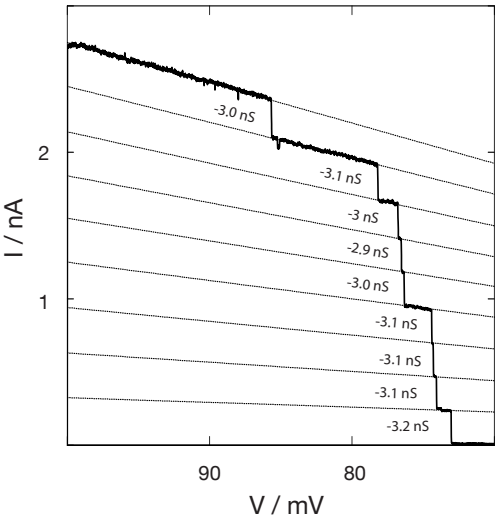

Figure 12. A conductance of G =27 nS was caused by the incorporation of multiple T pores into a DIB membrane. During the separation of the two bilayer forming droplets a voltage ramp of $1\,\frac{\mathrm{V}}{\mathrm{s}}$, starting at $+100\,\mathrm{mV}$ was applied. The separation of the droplets results in a loss of the connecting bilayer. During the separation process nine single and distinct current steps occurred. Dashed lines are a guide to the eye, indicating the contribution of the single pores. Image by courtesy of Krishnan *et al.* [51].

Negative controls with DNA nanopores devoid of the stem did not result in any detectable structure mediated transmembrane currents. The recordings of the pin pore, the wheel pore, and the T pore with the DIB barely exhibited any spontaneous channel closure, called gating, which is in contrast to previous recordings with planar suspended lipid bilayer set-ups [50]. The differences in how the membranes are built up might explain the diverging behavior. Suspended lipid bilayers appear to uphold a higher degree of lateral tension than free ones which hence might offer an explanation for the different experimental results.

Discussion For all pores the currents were predominately free of gating and the experimentally obtained conductance values showed ohmic behavior. Here, the conductance values are compared to the conductance values that are expected theoretically.

Figure 13a shows the theoretical cross-sectional area of pin and wheel pore, A_1, which equals $4.5\,\mathrm{nm}^2$.

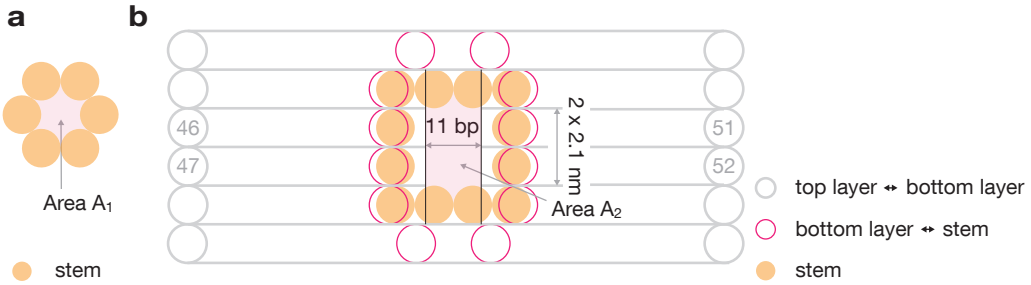

Figure 13. Schematic representation of the cross-sectional channel area for a, the pin and wheel pore and b, the T pore. Adapted image by courtesy of Krishnan *et al.* [51].

Figure 13b depicts the theoretical cross-sectional area of the T pore. Masked by the top layer, the theoretical cross-sectional area of the stem equals A_2 =15.5 nm^2. The narrowest constriction area A_2 is the intersection of the area of the stem and the hole in the top layer.

The conductivity σ of the electrolyte, which is predominately used in this thesis, a 1 M KCl solution, is 9.9 $\frac{\mathrm{S}}{\mathrm{m}}$.

For the pin and the wheel pore, the cross-sectional areas which we extract from the experimental conductance value are 4.4 nm^2 and 4.2 nm^2, respectively (Eq. 2.6).

For the T pore, with a stem of 32 nm length, the cross-sectional area extracted from experiment equals 10.9 nm^2 (Eq. 2.6). The discrepancy of 4.6 nm^2 between theoretically expected and experimentally calculated area could be attributed to the lipid bilayer compressing helices 46, 47, 51, and 52 approximately 0.6 nm inside the stem region ($2 \cdot 0.63\,\mathrm{nm} \cdot 3.7\,\mathrm{nm} = 4.7\,\mathrm{nm}^2$). The respective numbers of the helices are indicated in Figure 13b.

Also the area of 10.9 nm^2 corresponds to a pore radius of 1.9 nm, which has been determined to be the radius at which the total energy of the pore edge for a membrane made of a reference model lipid is at a local minimum [97]. Around the pore radius of 1.9 nm the membrane made of a reference model lipid is in a metastable state. For lipids 1,2-dioleoyl-sn-glycero-3- phosphocholine (DOPC), 1-palmitoyl-2-oleoyl-sn-glycero-3- phosphocholine (POPC) and 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) the radius of the metastable state has been modeled to be between 1 and 2 nm. It might be assumed that the radius of 1.9 nm is in the range where a hydrophilic pore in the DPhPC membrane reaches a metastable state.

	Pin Pore	Wheel Pore	T Pore
Stem length l / nm	42	28	32
Cross-sectional area, theoretical / nm^2	4.5	4.5	15.5
Conductance, experimental / nS	1.6	1.5	3.1 ± 0.3
Cross-sectional area, experimental / nm^2	4.4	4.2	10.9 ± 1.1

Table 2. Geometrical dimensions and conductance of the pin, the wheel, and the T pore. Adapted image by courtesy of Krishnan *et al.* [51].

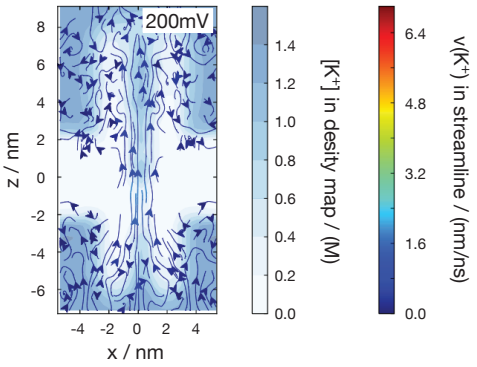

Figure 14. Density-flow maps of Cl^- ions at 200 mV through a six-helix-bundle pore. In this simulation the hydrophobic ethyl groups of the pore were replaced through hydrogen atoms. Gray-scale map indicates the average local density of Cl^-. Color-coded streamline plots indicate the average local velocity of Cl^-. Adapted with permission from J. Phys. Chem. Lett. 2015, 6, 23, 4680–4687. Copyright 2015 American Chemical Society [96].

The theoretical geometrical dimensions of the pore as well as the experimentally determined conductance of each pore are displayed in Table 2.

It cannot be excluded that the matching of the experimental and theoretical values emerges from a superposition of several contributions which compensate each other: The potential leakage through the sidewall, as indicated in Figure 14, and the compression of the pore structure are two examples for effects which were omitted here. The leakage would lead to an increase of effective diameter, the compression to the reduction of the effective diameter. Moreover, the electroosmotic effect and the fact that the current mainly runs through the bilayer and probably not along the whole length of the stem, are neglected. Molecular dynamics simulations, like those performed previously for other pores [96, 98, 99], might offer the means to address and clarify these open questions.

Lipid Anchor, Biotin-Streptavidin Anchor, Elastin-Like Polypeptide Anchor The lipid anchor was replaced with a biotin-streptavidin bridge and with hydrophobic ELPs.

Biotin-Streptavidin Anchor Here 5'-biotin DNA was used as an anchoring strand and the pores were incubated with streptavidin. For electrical measurements with

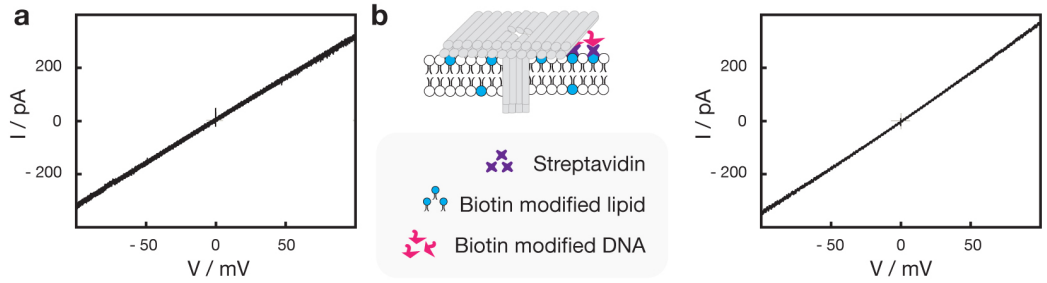

Figure 15. I-V curve of the T-pore when a, tocopherol or b, biotin-streptavidin are used for anchoring. Image of pore adapted from Krishnan *et al.* [51]

.

biotin-streptavidin anchoring 2 % of the DPhPC content was substituted with biotinylated DOPE. The biotinylated lipids were the bridging elements for potential pore incorporation. The I-V curve recorded at 1 M KCl is depicted in Figure 15 b.

The conductance of 3.5 nS is higher than the previously recorded conductances of 3.1±0.1 nS for the T pores with tocopherol anchors. However, since only a proof of principle experiment was performed, it is hard to assess whether this might be a general trend or a singular pore of higher conductance.

ELP Anchor We explored whether ELPs can serve as alternative anchor molecules. The ELP modified anchoring strands were fabricated with click chemistry as described in Methods 2.2.5 and were otherwise implemented according to the protocols for lipid modified strands. The ELP in this experiment is hydrophobic at room temperature, since the transaction temperature of the$(G\text{-}L\text{-}G\text{-}V\text{-}P)_{40}$ ELP in PBS was determined to be 5 °C [84].

It could be shown that the $(G\text{-}L\text{-}G\text{-}V\text{-}P)_{40}$ ELP anchored DNA pores incorporate in DPhPC bilayers. The measured conductance was lower than the statistical average of 3.1±0.3 nS of tocopherol anchored T pores. The according I-V curve is depicted in Figure 16 b. The conductance value of 2.2 nS, which is lower than the conductance of the T pore with tocopherol anchors either indicates a certain lateral pressure by the surrounding lipid molecules or is a statistical outlier.

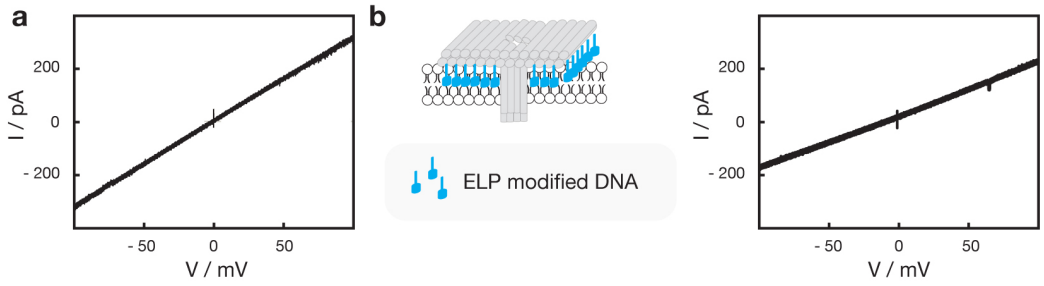

Figure 16. I-V curve of a, the T pore using tocopherol anchors and b, the T pore utilizing ELPs for anchoring. Image of pore modified from Krishnan *et al.* [51]

Discussion The conductance of 3.5 nS which was measured with biotin - streptavidin anchoring is higher than the previously recorded conductances of 3.1±0.1 nS. Reduced squeezing of the stem, due to a different kind of restructuring of the bilayer upon pore insertion and less lateral pressure could result in a larger cross-sectional area of the hole, which is formed.

The lower conductance value of 2.2 nS, which was measured with the ELP anchors, could indicate an increased lateral pressure by the surrounding lipid molecules. The $(\text{V-P-G-L-G})_{40}$ ELPs might induce a general perturbation of the membrane.

Both incorporation methods are proof of principle experiments devoid of the statistics which are required for drawing further conclusions from the conductance values. However, they prove that different incorporation strategies can be applied.

Biotin-streptavidin bridges offer an incorporation technique mediated by protein-ligand interaction which relies on non-hydrophobic interaction. It provides a possibility to reduce aggregation.

ELPs offer a potential means to switch the incorporation capability of the DNA nanopores, since ELPs can transition from hydrophilic to hydrophobic upon external stimuli. The external parameters include changes in temperature, salt concentration, and pH.

2.3.3. Conductance of Large Connector

The connector structure designed by S. Krishnan draws its inspiration from the complete intercellular gap junction channel, which is composed of two connexon structures [27]. The capability of the connector to incorporate into membranes

at two sides of the stem allows for a connection between two compartments, for instance between vesicle A and vesicle B.

Two distinct versions of connector structures were generated. A first version of the DNA based connector focused on devising a structure that exhibits a stem which can be anchored into a membrane on both its ends. For the first version the same stem geometry was employed as for the pin pore and the wheel pore. Is has a theoretical cross-sectional area of 4.5 nm^2 and accommodates 45 hydrophobic modifications.

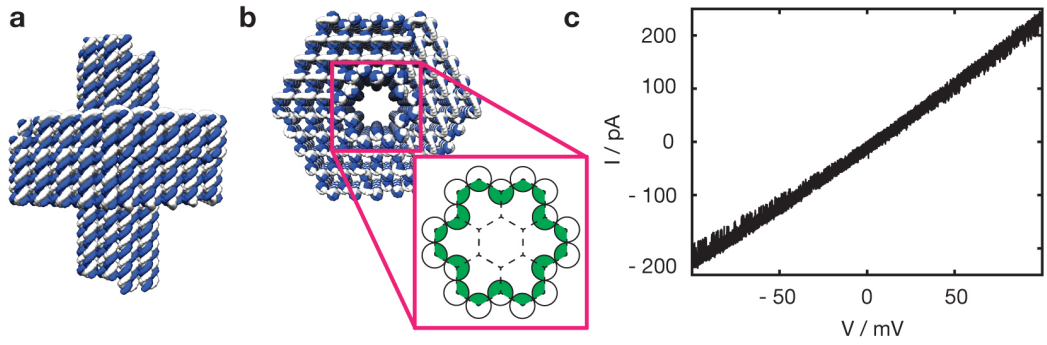

Figure 17. Conductance large connector a, Side and b, top view of the large connector. For the top view, a zoom in for the channel area is shown. Image adapted, image by courtesy of S. Krishnan [27]. c, I-V curve of the large connector at 500 mM KCl, 5 mM $MgCl_2$, utilizing tocopherol for anchoring

The large connector (Fig. 17 a) was endowed with a large diameter stem and a particular routing that allows for the positioning of 50 modifications. The theoretical cross-sectional area of the stem is approximately 48 nm^2 (Fig. 17 b). The incorporation of the large connector into a DPhPC bilayer yielded a conductance of approximately 4 nS (Fig. 17 c).

Discussion The conductance of 4 nS approximately equals half the value that is expected for the conductance of a pore with a theoretically calculated cross-sectional area of 48 nm^2 and a stem length of 47 nm. Assuming that the length of the stem is 47 nm, the conductance of 4 nS is equivalent to a cross-sectional area of 21 nm^2 or a radius of 2.6 nm.

The lateral pressure [100], which the lipid bilayer exerts on the stem, could be the cause for the discrepancy between theoretical and experimentally determined conductance. Potentially, the pressure squeezes the stem into a smaller than expected geometry, thereby reducing the cross-sectional area. Other effects in the counter

direction, which should not be neglected, are the leakage through and around the incorporated DNA strands, as well as a current that might run only through the bilayer area and not through the entire stem. Both would increase conductance values.

The degree of the influence of those two effects might be reduced in the connector, as the connector opposes its relatively massive structure against potential currents for a part of the channel (Fig. 17 a and b). In case the structure does not oppose those effects, it could be assumed that the lateral pressure reduces the area even further.

3. Translocation of Molecules through a DNA Pore

The passage of small molecules through the T pore was shown. The molecules which were proven to translocate through the pore are single and double-stranded DNA (dsDNA), RNA, and ribonucleic acid triphosphates (rNTPs). The translocation of the protein ubiquitin was not detected.

The herein described findings on the passage of DNA through the T pore are based on the paper 'Molecular Transport through DNA Origami Pores' [51].

3.1. Theory

The 1-D drift diffusion model [35, 101–109] was used to describe the translocation dynamics of molecules.

Translocation of Small Molecules

The 1-D drift diffusion model describes the probability $P(t)$ that a molecule translocates that the translocation of a molecule requires the time t for the translocation process.

The 1-D drift diffusion model assists in describing the translocation of a molecule through a nanopore.

$$P(t) = \frac{b}{\sqrt{4\pi D t^3}} \cdot exp\left(-\frac{(b-vt)^2}{4Dt}\right). \tag{3.1}$$

The v and D denote the drift velocity and the diffusion coefficient for a one-dimensional diffusion of the pore along the molecule. Parameter b is the sum of pore length and the contour length of the translocating molecule [102]. Equation 3.1, Schrödinger's first-passage-time distribution formula, offers values for the average drift velocities and diffusion constants of the translocating molecules [110].

The initial condition for the position-time probability densitiy function $p(x,t)$ is set by the capture of DNA into the pore

$$p(x,0) = \delta(x). \tag{3.2}$$

A second boundary condition is given by the translocation process being unidirectional and irreversible:

$$p(L,t) = 0. \tag{3.3}$$

When a molecule translocated through the molecules remains on the *trans* side.

The probability function is exposed to an electrophoretic force, which causes a drifting speed, and to the diffusion processes. The process can be described with the Fokker-Planck equation

$$\frac{\partial\, p(x,t)}{\partial_t} = D\frac{\partial^2\, p(x,t)}{\partial_x^2} - v\frac{\partial\, p(x,t)}{\partial_x}. \tag{3.4}$$

The first passage probability density function per unit time

$$P(t) = -\frac{d}{dt}\int_{-\infty}^{L} p(x,t)dx \tag{3.5}$$

describes the probability $P(t)$ for the molecule to translocate in the timespan t [102].

It has been shown that Equation 3.1 is the probability density function that can describe the probability $P(t)$ for the molecule to translocate in the timespan t [111,112].

3.2. Materials and Methods

The recording conditions, data evaluation, and sample conditions are listed.

3.2.1. Transcription

Template DNA (80 nM) was incubated with T7 RNA polymerase (100 nM, purified by S. Sagredo), 1 mM of each rNTP in 40 mM Tris-HCl, 20 mM $MgCl_2$, 1 mM DTT, 2 mM spermidine at 37 °C for 4 h or overnight. RNA was incubated with DNase I (New England Biolabs Inc.) for 30 min. RNA was run on an 8 % polyacrymaide gel in 1x TBE for 90 min at 90 V for the subsequent polyacrylamide gel electrophoresis (PAGE) purification. The gel was stained with Sybr Green II (Invitrogen, Thermo Fischer Scientific Inc), cut out under UV light, and extracted from the gel matrix with the ZR small-RNA PAGE recovery kit (Zymo Research). It was eluted in RNase-free water.

3.2.2. Investigated Molecules

The oligonucleotide sequences of translocated DNA and RNA strands and the salt concentrations in which the recordings were performed are summarized below.

Double-stranded DNA dsDNA of a length of 527 bp (sequence see Appendix B.1.1) was added to the large aqueous droplet which contained 1 M KCl and 5 mM $MgCl_2$.

Single and ds DNA First, ssDNA (5′ - (T)68-CCGG-(T)68-3′) was added to the large aqueous droplet, which contained 2 M KCl and 5 mM $MgCl_2$. Recordings were performed. Secondly, 115 bp long dsDNA (sequence see Appendix B.1.1) was added to the large aqueous droplet. Recordings were performed in presence of ssDNA and dsDNA.

sgRNA Either sgRNA or poly(A) tailed sgRNA (sequence see Appendix B.1.1) were added to the large aqueous droplet, which contained 1 M KCl and 5 mM $MgCl_2$).

rNTPs CTP or GTP were added to the large aqueous droplet, which contained 1 M KCl and 5 mM $MgCl_2$.

Ubiquitin Native or denatured ubiquitin (denaturing agent: 8 M urea) was added to the large aqueous droplet, which contained 1 M KCl.

3.2.3. Electrical Recordings

Current traces were recorded with an EPC 9 double amplifier (HEKA Elektronik GmbH) at 200 kHz which resulted in a temporal resolution of 5 µs per data point.

The inverse proportionality between event length and applied voltage proves the translocation of molecules through nanopores. Other current fluctuation and gating, the spontaneous closure of the nanopore, do not show inverse proportionality between event length and applied voltage. With translocation time being the crucial parameter for proving the translocation of molecules, it has to be ensured that the application of a filter does not skew the translocation times.

A median-filter was applied to improve the signal to noise ratio. Each value represents the median over 20 points of the original trace. The filter thereby introduces a cut off at 50 µs. Shorter events are not detected in the filtered trace. The median filter can affect the height of a peak, however, the event duration remains conserved.

For the detection of events a Matlab routine, which was customized by K. Kapsner, was used. The routine here averages the current values 1700 to 700 points ahead of the examined value. This average sets the upper baseline. The routine classifies a point as part of an event, if its value deviates from the upper baseline by a set threshold (here 10 times t, with t depending on the voltage). The routine classifies a point as the end of the event, when its value deviates less than a set threshold from the upper baseline (here 2 times t).

The set threshold t was increased with voltage. It here was increased linearly with voltage, as the background signal indicates an approximately linear increase of noise with increasing voltage.

The current reduction ΔI is defined by the difference between upper and lower baseline. The lower baseline is defined as the median of the current values of a single event.

3.3. Results and Discussion

The translocation of small molecules through DNA pores was investigated. The translocation behavior of DNA, RNA, rNTPs, and ubiquitin were examined.

3.3.1. Translocation of dsDNA

In contrast to previous DNA nanopores, the increased stem size of the T pore permits the passage of biomolecules with a larger diameter.

Translocation of dsDNA through the T pore The translocation of 527 base pair (bp) long DNA molecules through the T pore was examined. The translocation experiments were performed in an aqueous solution containing 1 M KCl and 5 mM $MgCl_2$. Sample traces in the absence and presence of dsDNA are shown in Fig. 18a and 18b, respectively. The zoom-ins show examples of current reductions that were detected as translocation events.

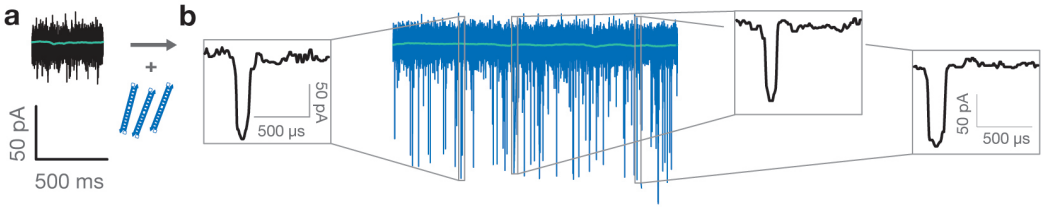

Figure 18. Filtered sample traces of a current measurement through a single DNA channel incorporated into a lipid bilayer membrane in a, the absence of analyte molecules, recorded at a transmembrane voltage of 125 mV and b, the presence of 527 bp long dsDNA. Three examples of translocation events are displayed as zoom-in. The green lines indicate the median of the current over 100 ms. Adapted image by courtesy of Krishnan *et al.* [51].

All events in the presence of dsDNA at 75 mV, 100 mV, and 125 mV are juxtaposed against those in the absence of dsDNA in scatter plots (Fig. 19 a).

The scatter plots of current reduction ΔI versus the event duration τ show that the presence of DNA leads to a distinct event population. The current reductions ΔI in the presence of dsDNA, depicted in Fig. 19b, are normal distributions around a ΔI. The ΔI is voltage specific, here 74 pA for 75 mV, 91 pA for 100 mV and 116 pA for 125 mV. As expected, the ΔI increases with voltage. The relative current reductions $\frac{\Delta I}{I}$ are of approximately 30 %. The relative current reduction

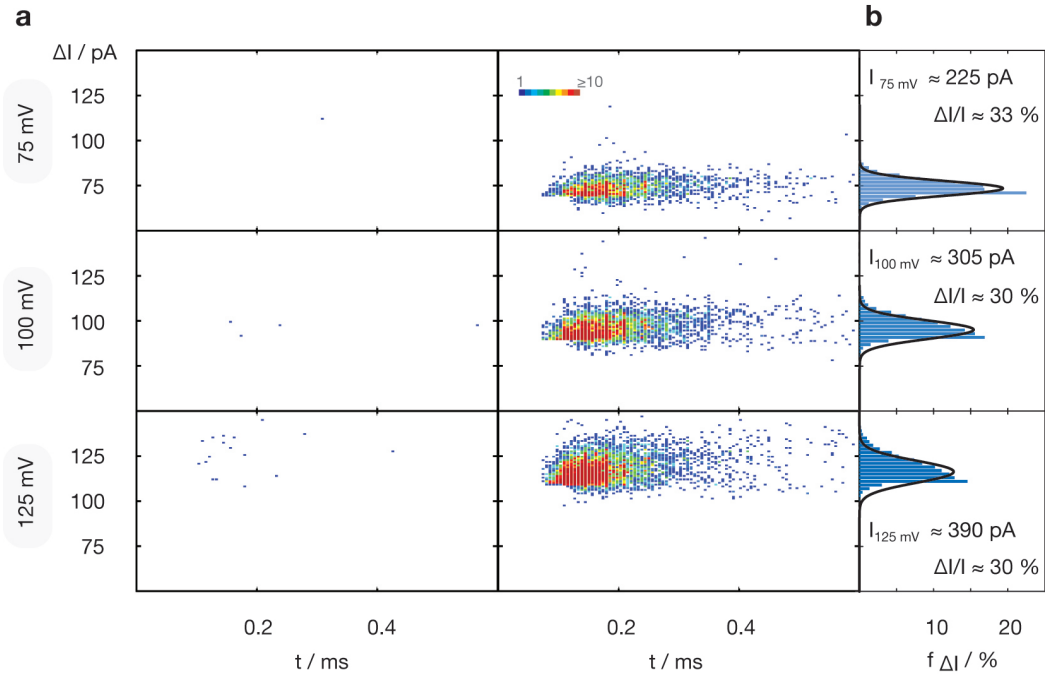

Figure 19. a, Translocation experiments of 527 bp long dsDNA analyte molecule through the T pore recorded with the droplet interface bilayer setup. Recordings were performed in absence and in presence of dsDNA at 75 mV, 100 mV, and 125 mV for 300 and 900 seconds, respectively. The number of translocation events is represented in the heat plot. b, The frequency count of ΔI is fit with a Gaussian distribution. The Gaussian distribution determines the most probable ΔI and the relative current reduction $\frac{\Delta I}{I}$. Adapted image by courtesy of Krishnan *et al.* [51].

allows us to reassess the cross-sectional area of the constriction zone of the T pore. In Ch. 2.3.2 a cross-sectional area that area equals $10.9 \pm 1.1\,\mathrm{nm}^2$ was derived from the conductance values.

The relative current reductions $\frac{\Delta I}{I}$ here are assumed to be equivalent to the relative reduction of cross-sectional area caused by the passage of the dsDNA. The DNA which threads through the hole excludes a certain percentage of the cross-sectional area from contributing to the flow of electrolyte.

Since the cross-sectional area of B-DNA with a diameter of 2 nm is $3.3\,\mathrm{nm}^2$ and $\frac{\Delta I}{I}$ equals approximately 30 %, the cross-sectional area is approximately equal to $11.5\,\mathrm{nm}^2$.

The approximation of the area of the hole of the T pore with approximately $11.5\,\mathrm{nm}^2$ coincides with the cross-sectional area determined from the conductance value of the T-pore ($10.9 \pm 1.1\,\mathrm{nm}^2$).

The relative current reduction $\frac{\Delta G}{G}$ in the presence of dsDNA corresponds well to the relative reduction in cross-sectional area of the pore, which might also indicate a relatively low level of leakage in our measurements.

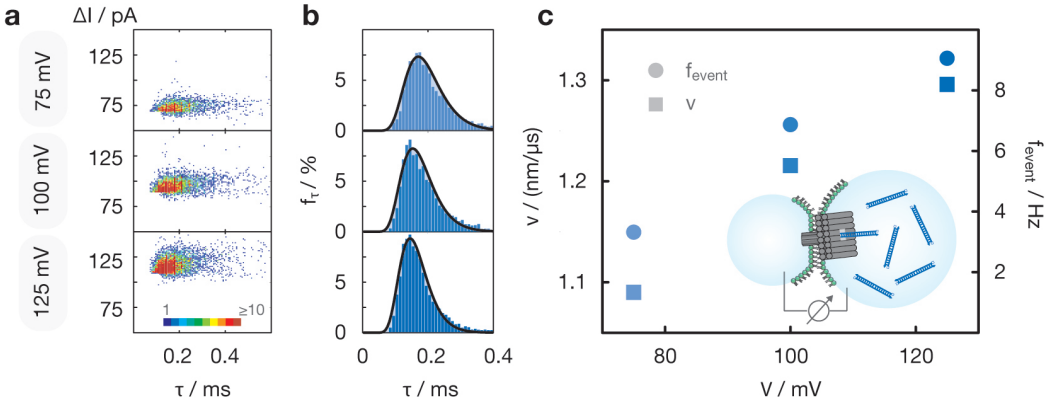

Figure 20. a, Scatter plots of events with a current reduction ΔI and an event duration τ extracted from current traces recorded at 75 mV, 100 mV, and 125 mV. b, Distribution of event durations at 75 mV, 100 mV, and 125 mV. The histograms were fit with a one-dimensional drift-diffusion model for nanopore translocation with the two fit parameters v and D (velocity and diffusion coefficient). The maxima of the fits (170 µs, 154 µs, and 149 µs at 75 mV, 100 mV, 125 mV, respectively) are defined as DNA translocation times at the corresponding voltages. c, Translocation velocities and event frequencies extracted from 1D drift diffusion model. Both parameters increase with increasing voltage, therefor proving translocation. Adapted image by courtesy of Krishnan *et al.* [51].

Besides the approximation of the cross-sectional area, the current reductions - or more specifically, their durations - allow for the confirmation of the translocation of dsDNA through the T pore. The mean event duration reduces from 207 µs over 190 µs to 167 µs for 75 mV, 100 mV, and 125 mV. The distribution of the event durations are modeled with the 1D drift diffusion model, which is determined by the molecule velocities v and diffusion constant D. The translocation time, the maximum of the 1D drift diffusion regression analysis, decreases from 170 µs at 75 mV over 154 µs at 100 mV to 149 µs at 125 mV, shown in Fig. 20 b. The frequency of translocation events and the velocity of the molecules increases with transmembrane voltage. It could be observed that both parameters increase with higher voltage, providing proof for the translocation of dsDNA through the T pore.

The calculated diffusion constant of 11 to 12 $\frac{\text{nm}^2}{\mu\text{s}}$ for the translocation of 500 bp dsDNA through the T pore is a relatively constant value. In solid-state pore experiments with pores of approximately 3 nm diameter, the diffusion constant was in the range of 1 $\frac{\text{nm}^2}{\mu\text{s}}$ to 20 $\frac{\text{nm}^2}{\mu\text{s}}$ for DNA of the length 35 bp to 1 kbp [104]. Further-

more, the diffusion constant was found to depend on the pore diameter. For 500 bp DNA Carson *et al.* obtained diffusion constants in the range of 1 to 100 $\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$, which increase with increasing pore diameter. It is not straightforward to assess which diameter of a solid-state pore a T pore compares to since the diameter of an approximately cylindrical-shaped DNA nanopore does not directly translate into the diameter of an hourglass-shaped solid-state pore [113].

However the value range of Carson *et al.* shows that the diffusion constants which were determined within the scope of this work are in the range of expected values [104].

3.3.2. Translocation of ssDNA and dsDNA through a modified T Pore

Experiments with a T pore, where 4 nt of ssDNA protrude inside the stem of the T pore, show that the speed of translocation is affected by the obstruction. Moreover, it was observed that the addition of ssDNA (140 nt) resulted in less pronounced current reductions than those for dsDNA analytes (115 bp), which is consistent with the differences in diameter for ssDNA and dsDNA.

In the middle of the central channel of the T pore, a 6 nt long ssDNA, with the sequence 5′ TTGGCC 3′, poses a potential obstacle for trespassing molecules. First, ssDNA analytes, then dsDNA analytes were added to one of the aqueous droplets (2 M KCl, 5 mM $MgCl_2$).

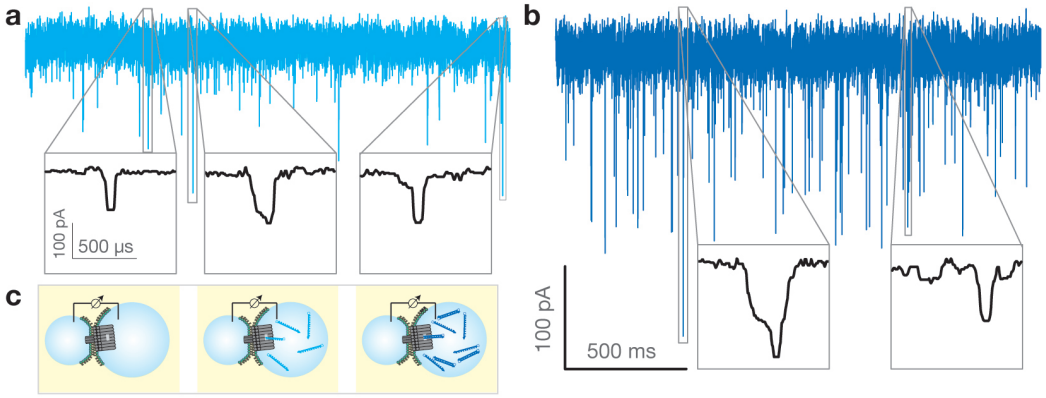

Figure 21. Current trace of 2 seconds a, in the presence of ssDNA and b, in the presence of ss and dsDNA at 100 mV. c, Setup in absence and presence of ssDNA or single and double stranded DNA. Images adapted from Krishnan *et al.* [51].

Sample current traces at 100 mV are shown in Fig. 21. Traces were recorded at100 mV, 125 mV, and 150 mV. While in the presence of ssDNA analyte (5′ -(T)68-CCGG-(T)68-3′) events were indeed detected (Fig. 21 a), the lengths and depths of the events were similar to those of the events present in the background signal. However, the signal indicates translocation, as event frequency and translocation velocity increase with increasing voltage, while translocation time decreases (Fig. 22 a). Event frequency increases significantly compared to the background. The diffusion coefficients of the various applied voltages here are independent of the applied voltage (2.4 $\frac{\text{nm}^2}{\mu\text{s}}$ at 50 mV, 2.4 $\frac{\text{nm}^2}{\mu\text{s}}$ at 75 mV and 2.8 $\frac{\text{nm}^2}{\mu\text{s}}$ at 100 mV).

Diffusion constants for the translocation of ssDNA of 27 nt length, previously has been discussed comparing it to the diffusion constant of a freely diffusing sphere with 1.5 nm diameter. There the diffusion constant was determined for the escape of the ssDNA from MspA. The escape times lead to diffusion constants which were between 13 % to 25 % of the expected diffusion constant of 330 $\frac{\text{nm}^2}{\mu\text{s}}$. The diffusion constant recorded here is 10 times lower than the diffusion constant extracted for the translocation of ssDNA through MspA, which might indicate an interaction of the ssDNA with the nanopore [114].

Upon addition of dsDNA analyte to the ssDNA events that are distinguishable from the background signal (Fig. 21 b) emerged. The overall translocation times decrease with increasing voltage from 184 µs at 50 mV over 170 µs at 75 mV to 150 µs at 100 mV. Moreover, the overall velocity increased with voltage from 0.26 $\frac{\text{nm}}{\mu\text{s}}$ at 50 mV, to 0.32 $\frac{\text{nm}}{\mu\text{s}}$ at 75 mV and to 0.42 $\frac{\text{nm}}{\mu\text{s}}$ at 100 mV. Compared to the prior recording of ssDNA (Fig. 23a), the overall event frequency increased significantly in the presence of dsDNA (Fig. 23b). These results prove the translocation of DNA through the pore.

The addition of dsDNA resulted in the emergence of a distinct, additional population of events in the scatter plot (Fig. 22d, non-faded). The population was further analyzed to obtain information about the translocation process of dsDNA molecules in the presence of an obstacle. The events were split into two populations for further analysis. There, the top population was attributed to the presence of dsDNA (Fig. 22d, non-faded). The bottom population was not taken into account (shown in faded colors in Fig. 22d). The translocation times extracted from the restricted population are depicted in Fig. 22c and the velocities are shown in Fig. 23b. With increasing voltage the translocation times decrease, while the velocities increase. The diffusion coefficients of 2.7 $\frac{\text{nm}^2}{\mu\text{s}}$ at 50 mV, 2.5 $\frac{\text{nm}^2}{\mu\text{s}}$ at 75 mV and 2.3 $\frac{\text{nm}^2}{\mu\text{s}}$ at 100 mV were obtained by modeling the time distribution event length

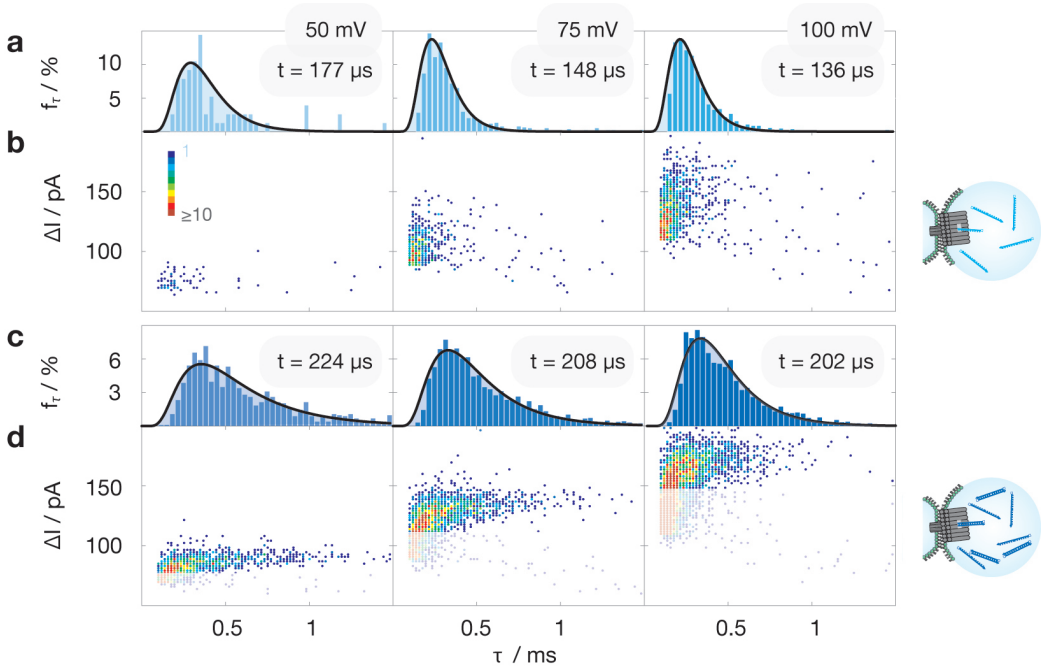

Figure 22. a, Distribution of the length τ of events that were caused in presence ssDNA at 100 mV, 125 mV, and 150 mV. b, Scatter plot ΔI vs. τ for all detected events. c, Distribution of the event length τ of the non-faded population at 100 mV, 125 mV, and 150 mV and d, all detected events at 100 mV, 125 mV, and 150 mV. All histograms were fitted with 1D drift diffusion regression analysis. The time where the fit of the translocation time distribution reaches its maximum is defined as the translocation time t, here stated for each voltage as an inset. Adapted image by courtesy of Krishnan *et al.* [51].

with the 1D drift-diffusion regression analysis. The obtained diffusion constants are approximately independent of the applied voltage.

The velocities of dsDNA compared to the velocities of ssDNA, in this particular setup, vary approximately by a factor of two. In accordance with the smaller persistence length of ssDNA, its smaller diameter, and its lower charge density the motion of ssDNA through the pore could be faster than the motion of dsDNA. The properties of ssDNA might result in an easier threading of ssDNA through the modified DNA channel.

The translocation of the longer 527 bp dsDNA without an obstacle inside the channel (Fig. 20 b and Fig. 23 b) is by a factor of approximately 5 faster than translocation of 115 bp dsDNA in the presence of an obstacle. In presence of the obstacle the diffusion constant is reduced by a factor of 5. This is an indication that the ssDNA obstruction has a significant effect on translocation dynamics.

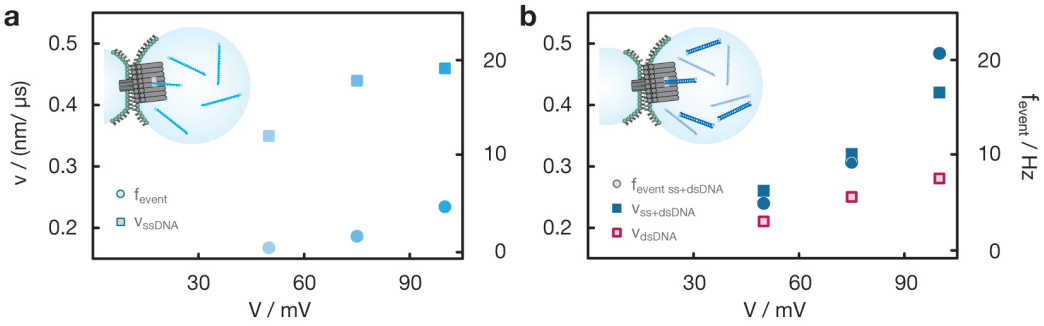

Figure 23. a, Translocation velocities and event frequencies for ssDNA at 2 M KCl and 5 mM $MgCl_2$, b, Translocation velocities and event frequencies for a mixture of 140 nt long ssDNA and 115 bp dsDNA and the translocation velocities and event frequencies that were extracted for dsDNA only. Adapted image by courtesy of Krishnan *et al.* [51].

3.3.3. Translocation of RNA

Furthermore, single guide RNA (sgRNA), poly(A)-tailed sgRNA and an RNA aptamer were translocated through the T pore. Data was acquired in cooperation with M. Fuchs [115].

Translocation of sgRNA The presence of RNA resulted in clear current reductions, which were analyzed as mentioned above. The extracted distribution of event length was modeled with the 1D drift diffusion model.

The modeling resulted in translocation velocities of 0.10 $\frac{\text{nm}}{\mu\text{s}}$ at 50 mV, 0.13 $\frac{\text{nm}}{\mu\text{s}}$ at 75 mV, 0.15 $\frac{\text{nm}}{\mu\text{s}}$ at 100 mV and 0.18 $\frac{\text{nm}}{\mu\text{s}}$ at 125 mV (Fig. 24a). The velocities and their uncertainties were obtained by regression analysis of the distribution of the event length. In Fig.24b The distributions are shown for all voltages. The translocation time, which is the maximum of the 1D drift diffusion model, is stated. It decreases with increasing voltage, whereas the velocity increases, proving the translocation of the sgRNA. Figure 24c depicts scatter plots of all detected events of event length τ against current reduction ΔI. The ΔI increase for increasing applied voltage.

Translocation of poly(A)-tailed sgRNA Furthermore, the presence of poly(A)-tailed sgRNA resulted in distinct current reductions. As shown in Fig. 25a the mean event duration of poly(A)-tailed sgRNA is longer than the one of the sgRNA. It can be concluded that the longer RNA molecule requires more time for the

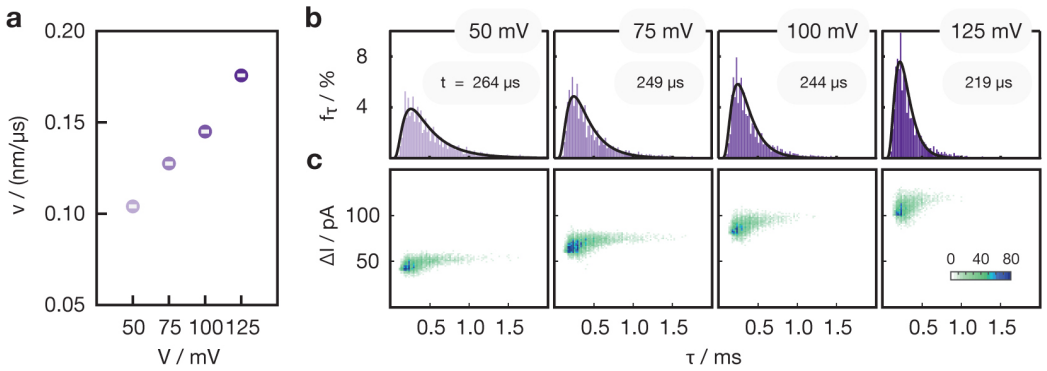

Figure 24. Evaluation of events in the presence of sgRNA a, Translocation velocities of the sgRNA increase with increasing voltage. Velocities and uncertainties were obtained from the 1D drift diffusion model. b, Distribution of event length at $50\,\mathrm{mV}$, $75\,\mathrm{mV}$, $100\,\mathrm{mV}$, and $125\,\mathrm{mV}$ The event lengths were modeled with the 1D drift diffusion model. Translocation times t are stated in the inset. c, Scatter plots show duration τ and current reduction ΔI of the events and display the distribution and amount of events for each voltage.

passage through the pore. Figure 25a also shows the decrease of the mean event duration with increasing voltage, which is an indication of the translocation of the poly(A)-tailed sgRNA.

For the 1D drift diffusion model the molecule length was set to the length of the sgRNA, since no precise information is available about the length of the poly(A)-tail. The definition determines the parameter b of the 1D drift diffusion model. Consequently, both, the event duration of the sgRNA and the event duration of the poly(A)-tailed sgRNA were fit with the same parameter b.

The results, therefore, have to be interpreted with caution, since the definition of the molecule length influences the translocation velocity and the diffusion constant resulting from the 1D drift diffusion model.

The determined translocation times and velocities of the poly(A)-tailed sgRNA are shown in Fig. 25b and c, respectively.

The higher velocity of poly(A)-tailed RNA (Fig. 25 c) should not *per se* be accounted for a definitive difference in translocation velocity. The obtained velocity values are influenced by the length parameter b. For the poly(A)-tailed sgRNA, which by default is longer than the sgRNA, b underestimates the length of the RNA, therefore the fit does not return the definitive translocation velocity for the poly(A) tailed RNA.

However, the velocities of the poly(A)-tailed RNA depict the trend that the translocation speed increases with increasing voltage (Fig. 25c). In combination with translocation time decreasing from 75mV up to 125mV (Fig. 25b) and the decrease of the mean event duration prove is provided that the poly(A)-tailed RNA translocates through the pore.

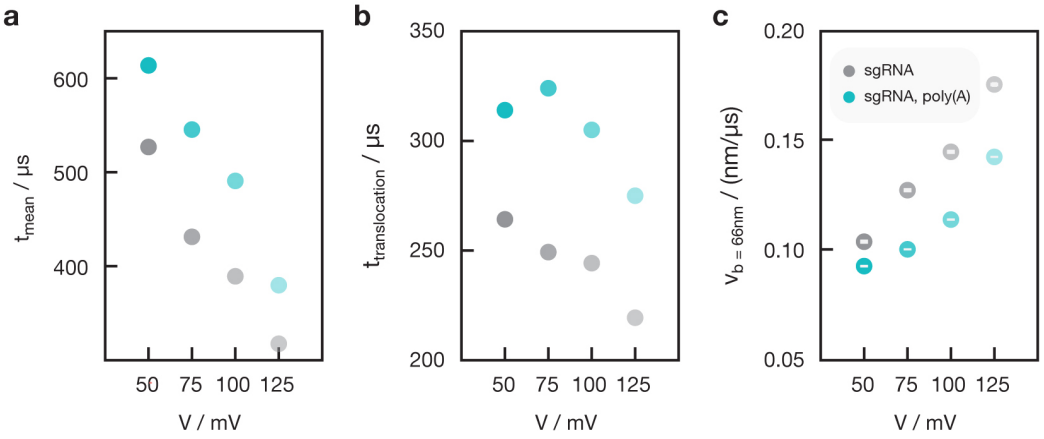

Figure 25. Translocation of poly(A)-tailed sgRNA a, Mean translocation times for poly(A)-tailed sgRNA and sgRNA decrease for both with increasing transmembrane voltage. b, Translocation time decreases for poly(A)-tailed sgRNA from 75 mV onward. c, Velocities extracted by fitting the distributions of event length with the 1D drift diffusion model. Here, in the parameter 'b = pore length + molecule length' the molecule length was set to the length of the sgRNA ($\approx 20\,\mathrm{nm}$), also for the poly(A)-tailed sgRNA. Error bars show a given error from 1D drift diffusion.

In Fig. 26a the duration τ and the current reduction ΔI of the events caused by poly(A)-tailed sgRNA at 50 mV, 75 mV, 100 mV, and 125 mV are shown in scatter plots. Figure 26b shows the duration τ and the current reduction ΔI of the events caused by the sgRNA. The plots visualize that τ reduces for increasing transmembrane voltage. Additionally, the plots illustrate that ΔI increases with increasing voltage.

Figure 26c depicts that diffusion constants decrease slightly with increasing transmembrane voltage, from $1.2\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ to $0.9\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ for the sgRNA and from $1.0\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ to $0.7\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ for the poly(A)-tailed sgRNA. The determined diffusion constant remains relatively constant. Since the length of poly(A)-tailed RNA was set to the length of the sgRNA, the velocity and the diffusion constant obtained from fitting the distribution have to be interpreted carefully.

Both recordings, those of sgRNA and the poly(A)-tailed sgRNA, result in the diffusion constants at a similar order of magnitude. In accordance with the Stokes-

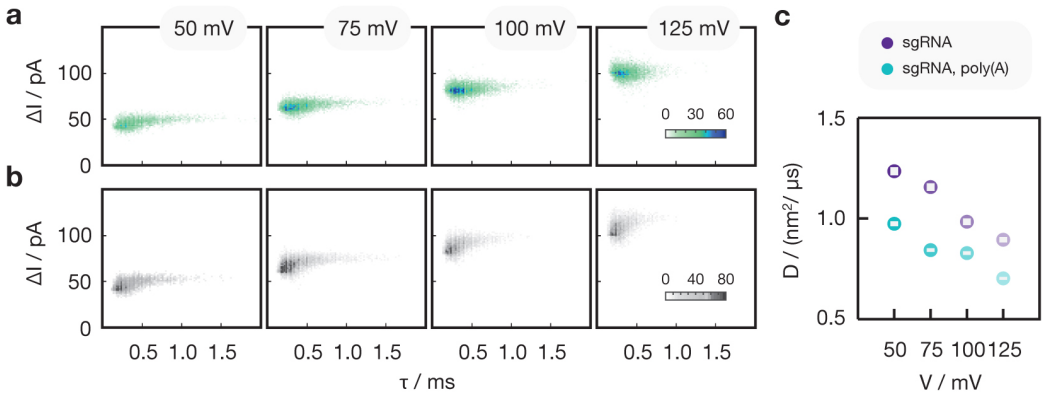

Figure 26. Juxtaposition of scatter plots of sgRNA and poly(A)-tailed sgRNA. a, Scatter plots of current reduction ΔI vs. the event duration τ of poly(A)-tailed sgRNA and b, sgRNA at $50\,\mathrm{mV}$, $75\,\mathrm{mV}$, $100\,\mathrm{mV}$, and $125\,\mathrm{mV}$. c, Diffusion coefficients extracted by applying the 1 D drift diffusion model to the distribution of event length of both the sgRNA and the poly(A)-tailed sgRNA at $50\,\mathrm{mV}$, $75\,\mathrm{mV}$, $100\,\mathrm{mV}$, and $125\,\mathrm{mV}$

Einstein relation, the lower diffusion constant was recorded for the larger molecule.

The diffusion coefficients, which were here obtained for RNA, compare to the order of the diffusion constant that was obtained for ssDNA, of $2.4\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ to $2.8\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$. The diffusion coefficients are two orders of magnitude smaller compared to those obtained in previously performed experiments in bulk or compared to simulation [116, 117].

RNA Aptamer The presence of RNA aptamer caused current reductions. Figure 27a shows the distribution of event length at $100\,\mathrm{mV}$, $125\,\mathrm{mV}$, and $150\,\mathrm{mV}$. The distribution was modeled with the 1D drift diffusion model.

The obtained velocities (Fig. 27b) are $0.11\,\frac{\mathrm{nm}}{\mu\mathrm{s}}$ at $100\,\mathrm{mV}$, $0.13\,\frac{\mathrm{nm}}{\mu\mathrm{s}}$ at $125\,\mathrm{mV}$ and $0.15\,\frac{\mathrm{nm}}{\mu\mathrm{s}}$ at $150\,\mathrm{mV}$. The translocation times of $235\,\mu\mathrm{s}$ at $100\,\mathrm{mV}$, $221\,\mu\mathrm{s}$ at $125\,\mathrm{mV}$ and $200\,\mu\mathrm{s}$ at $150\,\mathrm{mV}$ decrease with increasing voltage. The diffusion constant is $0.6\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ for all voltages. It hence is smaller than the one obtained for the sgRNA, however it is at the same order of magnitude.

The raw data of the results presented above can be filtered differently. For the data presented above, each value is the median over 20 points of the original trace. The data was additionally analyzed so that each value is the median over 2 points. The median over 2 points preserves shorter events therefore the values of the velocity and the diffusion constant are affected by changing the filtering, they increase to

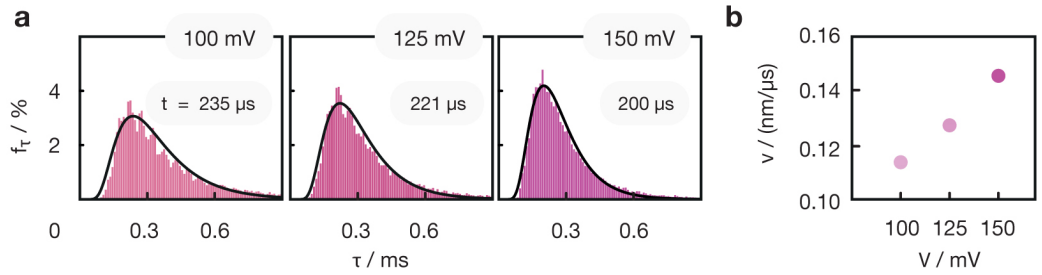

Figure 27. a, Duration of event length in the presence of the RNA aptamer at $100\,\mathrm{mV}$, $125\,\mathrm{mV}$, and $150\,\mathrm{mV}$. Translocation times t are stated in the inset. b, Translocation velocities of the RNA aptamer are obtained from the 1D drift diffusion model and increase with increasing voltage.

higher values. The obtained velocities are in the range of $0.3\,\frac{\mathrm{nm}}{\mu\mathrm{s}}$ and the obtained diffusion constants are in the range of 2 to $3\,\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$. The velocity increases with increasing voltage and the RNA aptamer translocates through the T pore.

3.3.4. Translocation of rNTPs

Additionally, with the DIB, the current through the T pore in presence of the rNTPs cytidine triphosphate (CTP) and guanosine triphosphate (GTP) was recorded. The data was acquired in cooperation with N. Matejka [86] and M. Fuchs [115].

CTP The presence of 1 to $10\,\mu\mathrm{l}$ of $10\,\mathrm{mM}$ CTP in the aqueous droplet does not result in a significant amount of signal when aHl is used for recording. Figure 28 reflects how the event numbers remain small in the presence of CTP at an estimated concentration of $1\,\mathrm{mM}$ CTP. The current through aHl was not significantly affected by the presence of rNTPs.

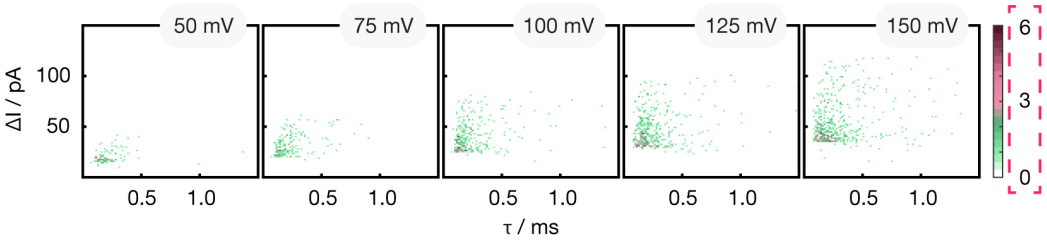

Figure 28. Events detected with aHl in the presence of CTP. Scatter plots of current reduction ΔI vs. the event duration τ at $50\,\mathrm{mV}$, $75\,\mathrm{mV}$, $100\,\mathrm{mV}$, and $125\,\mathrm{mV}$

When recording with the T pore, distinct peaks emerge when 1 to 5 µl of 10 mM CTP are added to the aqueous 1 M KCl, 5 mM $MgCl_2$ solution. CTP introduces an alteration of the current as can be seen in Fig. 29a and b. The zoom-ins demonstrate the well-defined nature of the peaks caused by the presence of CTP.

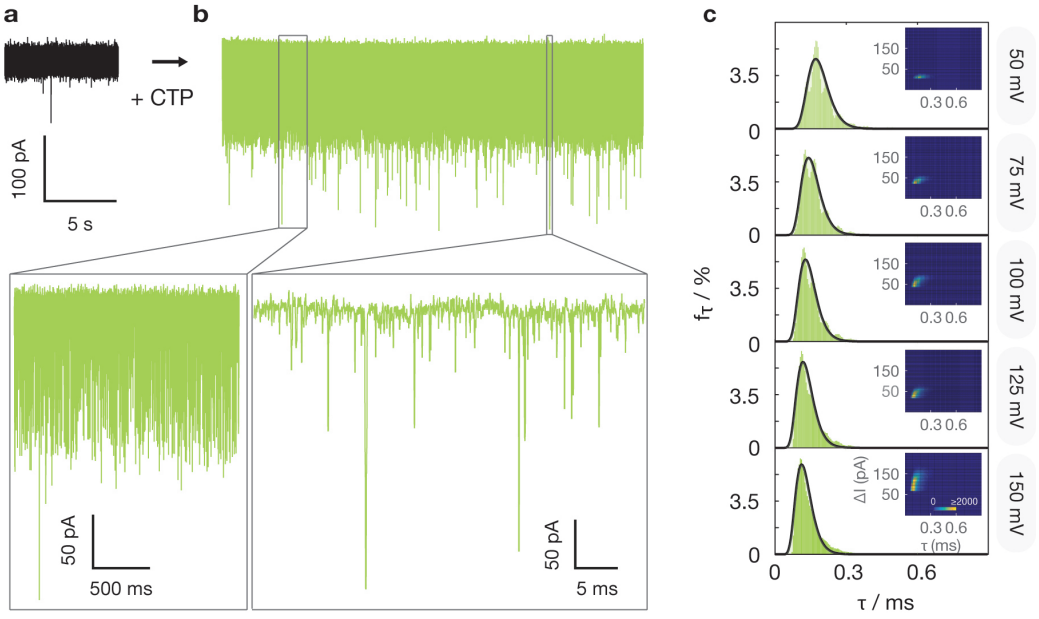

Figure 29. Translocation of CTP through large-diameter T pore. a, Sample trace of a current measurement through a single DNA channel incorporated into a lipid bilayer membrane in the absence of analyte molecules at a transmembrane voltage of 150 mV. b, Current trace recorded in the presence of CTP. Two examples show different zoom-in of the current trace. c, Distribution of event length extracted from current traces recorded at voltages between 50 mV and 150 mV. The histograms were fit with the 1D-drift-diffusion model for nanopore translocation. In the inset, a scatter plot of the current reduction ΔI vs. the event duration τ including all events, is shown.

Peak analysis, as described in section 3.2.3, reveals decreasing translocation times with increasing voltages. As depicted in Fig. 29c the translocation times decrease from 172 µs at 50 mV, 137 µs at 75 mV, 129 µs at 100 mV over 118 µs at 125 mV down to 113 µs at 150 mV. Translocation velocities increase from 0.20 $\frac{\text{nm}}{\mu\text{s}}$ at 50 mV, 0.24 $\frac{\text{nm}}{\mu\text{s}}$ at 75 mV, 0.26 $\frac{\text{nm}}{\mu\text{s}}$ at 100 mV over 0.28 $\frac{\text{nm}}{\mu\text{s}}$ at 125 mV up to 0.29 $\frac{\text{nm}}{\mu\text{s}}$ at 150 mV.

The diffusion constant increases with increasing voltage from 0.23 $\frac{\text{nm}^2}{\mu\text{s}}$ at 50 mV, 0.33 $\frac{\text{nm}^2}{\mu\text{s}}$ at 75 mV, 0.39 $\frac{\text{nm}^2}{\mu\text{s}}$ at 100 mV over 0.45 $\frac{\text{nm}^2}{\mu\text{s}}$ at 125 mV up to 0.47 $\frac{\text{nm}^2}{\mu\text{s}}$ at 150 mV. The scatter plots, shown as insets in Fig. 29c, with relative current

reduction $\frac{\Delta I}{I}$ and event duration τ, visualize the three observations. The shortening of event duration with increasing voltage, the increasing event numbers with increasing voltage, and the steadiness of the relative current reductions across all voltages are observed. These observations indicate the translocation of CTP.

GTP Similar observations were obtained in presence of GTP. Peaks occur when adding GTP to a T pore incorporated in a bilayer within a surrounding solution of 1 M KCl, 5 mM $MgCl_2$ solution (Fig. 30a, b). As shown in Fig. 30c translocation

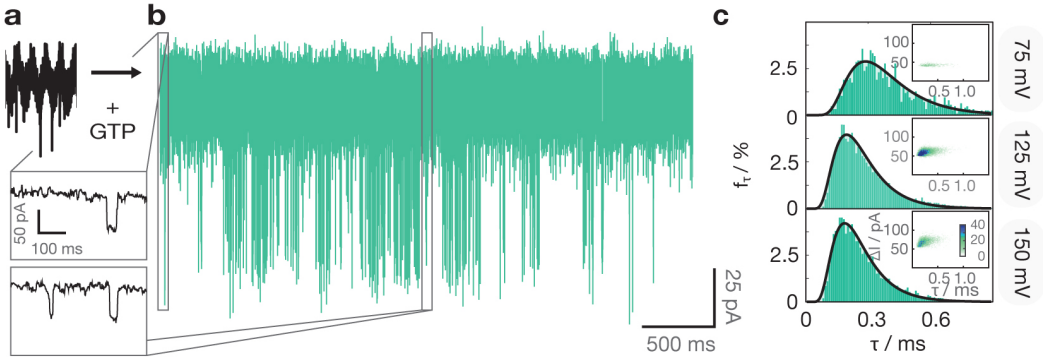

Figure 30. Translocation of GTP through large-diameter T pore. Samples of current trace in a, absence and b, presence of GTP. c, Event durations and the results of regression analysis with the 1D drift diffusion model at 75 mV, 125 mV, and 150 mV. Insets show the scatter plots of the events, with event length τ and current reduction ΔI

times reduce with increasing voltage, from 289 µs at 75 mV over 198 µs at 125 mV down to 188 µs at 150 mV, while event frequency increases. Velocities increase from 0.10 $\frac{\text{nm}}{\mu\text{s}}$ at 75 mV over 0.14 $\frac{\text{nm}}{\mu\text{s}}$ at 125 mV up to 0.15 $\frac{\text{nm}}{\mu\text{s}}$ at 150 mV. Diffusion constants are 0.37 $\frac{\text{nm}^2}{\mu\text{s}}$ at 75 mV, 0.58 $\frac{\text{nm}^2}{\mu\text{s}}$ at 125 mV, and 0.60 $\frac{\text{nm}^2}{\mu\text{s}}$ at 150 mV. The scatter plots of event length τ and the current reduction ΔI visualize that with increasing voltage event length decreases while event frequency increases. In sum the results indicate the translocation of GTP.

Alternatively, the raw data can be analyzed by filtering differently. For the data presented above, each value represents the median over 20 points of the original trace. Furthermore, the presented data was analyzed with a filter, where each value represents the median over 2 points of the original trace. In contrast to the median over 20 points, the median over 2 points does not erase short events. Therefore the values of the velocity and the diffusion constant are affected by changing the filtering. They both increase to higher values. For CTP the velocities range from

$0.4\,\frac{\text{nm}}{\mu\text{s}}$ to $0.5\,\frac{\text{nm}}{\mu\text{s}}$ and the diffusion constants are in the range of $2\,\frac{\text{nm}^2}{\mu\text{s}}$. For GTP the velocities are in the range of $0.1\,\frac{\text{nm}}{\mu\text{s}}$ and the diffusion constants are in the range of 1 to $2\,\frac{\text{nm}^2}{\mu\text{s}}$. However, for both, CTP and GTP the velocity increase with increasing voltage. It can be concluded that with both, a median over 2 and over 20 points, GTP and CTP translocate through the T pore.

3.3.5. Experiments with Ubiquitin

Recordings of ubiquitin through solid-state nanopore have been reported [118]. The ionic current was recorded in the absence and presence of ubiquitin. Ubiquitin was either added in its native form or was denatured with 8 M urea before addition. A ssDNA tag was attached to the protein to increase its capture probability.

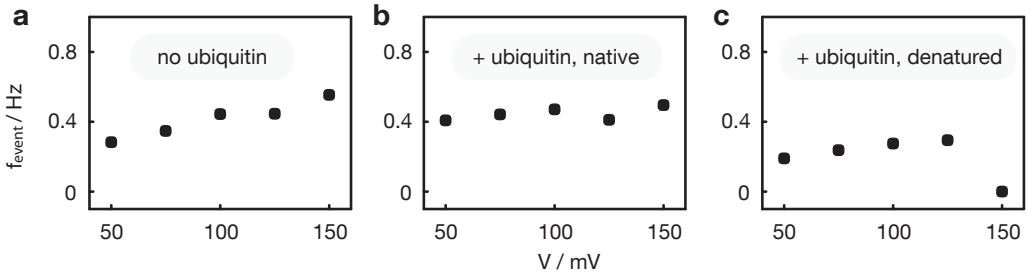

Figure 31. Event frequencies in a, absence and presence of b, native, or c, denatured ubiquitin

The modified ubiquitin was kindly provided by M. Grison, Rief AG [119]. The current signal was not altered by the presence of ubiquitin. The event frequency remained steady. No indication for the presence of molecules was found.

3.3.6. Discussion

The sensitivity of the pore for the detection of translocating molecules is determined by two major factors: the geometry of the sensing zone and the chemical components present in the sensing zone. The geometrically relevant factors are the length of the zone and the number of sensing zones of the pore. Two relevant attributes of the chemical components that are present in the sensing zone are their charges and aromatic rings. All factors can be decisive for the passage of molecules through the pore.

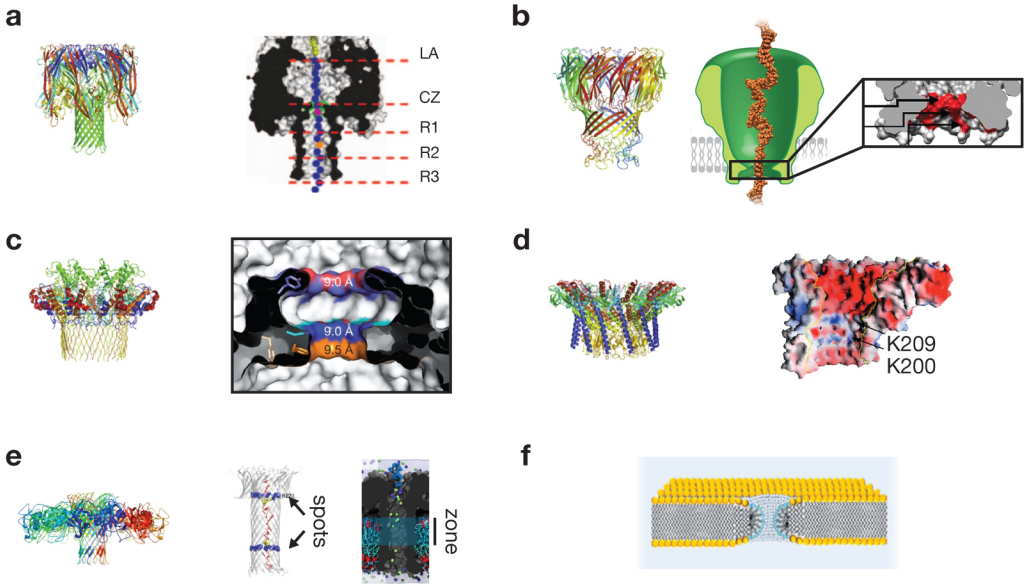

Figure 32. Sensing zones of aHl, MspA, CsgG, ϕ29, aerolysin, and α hederin. Image by courtesy of structures to Wang *et al.* [37] and Cao *et al.* [120]. a, The main sensing zone of aHl resides at the constriction indicated here with CZ. Further sensing zones have been identified at R1, R2, and R3 within the β-barrel domain of the aHL nanopore. A further sensing zone in the latch can be found around LA [121]. Image by courtesy of Stoddart *et al.* [122] b, The constriction zone of the wild type MspA pore contains the negatively charged residues D90, D91, and D93. The D90N, D91N, and D93N mutations to neutrally charged residues N allow for DNA translocation. Image by courtesy of Deamer *et al.* [123] and Bulter *et al.*, Copyright 2008 National Academy of Sciences [124] c, Cross-section of CsgG channel constriction, which includes several sensing zones. Image by courtesy of Goyal *et al.* [125] d, Half a ϕ29, where Lys209 and Lys200 are indicated. Image by courtesy of Guasch *et al.* [126] e, Sensing spots vs sensing zone. MD snapshot of dA14 fully inserted in the pore, showing the two main constriction points (R220 and K238, in blue) Image by courtesy of Cao *et al.* [127] Initial state of an SMD simulation where the sensing region of the aerolysin pore is highlighted in cyan. Image by courtesy of Ouldali *et al.* [128] f, An α hederin pore with short sensing zone. Reprinted with permission from Jeong, KB *et al.*, Alpha-Hederin nanopore for single nucleotide discrimination, ACS nano 13.2, 1719-1727(2019). Copyright 2019 American Chemical Society [129].

The first aspect to be elucidated is how the dimensions of the constriction zone and the number of sensing zones affect the detection of trespassing molecules.

The narrowest constriction of the nanopore defines the region of the highest ion density and its dimensions are crucial for sensing [130]. Certain regions along the stem can evoke additional signal, which then interferes with the main signal originating from the constriction zone [122].

For aHl multiple regions generate a signal when a molecule translocates [122]. One is its constriction zone, a second one resides in the latch of the vestibule [121] and more are distributed across the β barrel (Fig. 32a) [122, 123]. If signals are generated simultaneously, the signal is convoluted and the extraction of information becomes challenging. For more precise sensing, which is required for DNA sequencing, research shifted its focus on pores with narrow and single sensing zones, like the MspA porin. The constriction zone of MspA porin, with a diameter of 1.2 nm and a short longitudinal direction of 0.6 nm offers enhanced read-out capabilities (Fig. 32b) [130]. Mutated versions of the pore are capable to detect signal that originates from DNA strands which are as short as four nucleotides [130].

Other sensitive protein pores are aerolysin, which either has two sensing spots [127] or a whole sensing zone (Fig. 32e), and CsgG, a nanopore with two narrow constrictions that reside in close proximity, each of a diameter of 0.9 Å [125] (Fig. 32c). A mutant of CsgG was commercially applied in the sequencing apparatus MinION for sequencing purposes [131].

The detection of dNTPs or rNTPs which translocate through protein pores has not been reported to date [132]. However, non-protein nanopores showed sensitivity for trespassing dNTPs. To the best of my knowledge, the sensing of dNTPs has been demonstrated with three pores to date [129, 133, 134].

One of them is the α hederin pore, an amphipathic glycoside which forms in cholesterol-rich membranes with a diameter of 0.8 nm [129]. The height of the pore is equivalent to the thickness of the membrane (Fig. 32f). It hence is a short and narrow pore. The second type of pore with which dNTPs were detected is an atomically thin monolayer of molybdenum disulfide immersed in room-temperature ionic liquids with a hole of approximately 2 nm diameter [133]. For both pores, it was assumed that the short constriction zone was decisive for the detection of dNTPs.

The detection of dNTPs with a silicon nitride nanopore of 1.8 nm diameter and 10 nm height was attributed to the electroosmotic flow [134]. The dNTPs passed the pore in the direction opposite to the electrophoretic force.

The second aspect besides geometrically relevant factors concerns how charges and aromatic rings inside the constriction zone reportedly determine the translocation dynamics of the trespassing molecule [135].

For protein pores, the translocation dynamics through the constriction zone depend on the amino acids present in the constriction zone. The positively charged R220

and K238 in the stem of aerolysin (Fig. 32e) are associated with the formation of salt bridges with the phosphate of the translocating DNA. MD simulations support the experimental observation [127]. Moreover, the simulations show that the nucleobases form π-cation interactions with the positive side chains of the amino acids. Furthermore, the simulations indicate that hydrogen bonds and van der Waals interactions are established when negative amino acid side chains are present in the stem. Similar observations were obtained for the phi29 nanopore. There, 24 lysines, which are charged positively, are relevant for the translocation of DNA (Fig. 32d). The 24 lysins, K209 and K200, form two rings which both have negative counterparts, D202 and D208, in close proximity [126].

In aHl the E111N and the K147N mutations in the constriction zone result in reduced interaction with DNA [122]. Instead of positively (K) and negatively (E) charged amino acids there the polar but uncharged asparagine (N) dominates the translocation behavior, which results in a reduced DNA-pore interaction. Interestingly, an additional M113Y or M113W mutation of the third residue in the constriction zone turns the weakly interacting NN site into a strong NNY [122] or NNW interaction site [37]. In contrast to methionine (M) both tyrosine (Y) and tryptophan (W) have aromatic rings, which both could stack with aromatic rings or form hydrogen bond interactions with passing molecules.

In CsgG the constriction zone contains three rings, one formed by phenylalanines (F), one by Ns and one by Ys (Fig. 32c). The rings formed by the two amino acids with aromatic rings, F and Y, are regarded as crucial for molecule sensing [125].

For the T pore, we hypothesize that the translocating molecules interact with the various charges and the aromatic rings which all are present in the T pore.

The main voltage drop in the T pore potentially occurs in proximity to the DNA plate and the lipid bilayer or around one of the two components. The height of the sensing zone hence might be equal to the double layer plate of the pore plus the height of the bilayer or to only one of the two. The hole in the double layered plate of the T pore exhibits four blunt ends in each layer (Fig. 13b) which offer additional interaction sites for the trespassing molecule. The specifics of the T pore, like the absence of its lumen might, be an advantage for molecule sensing [136]. Other pores taper towards their constriction zone more gradually, which can deteriorate the electrical signal [136].

With the T pore, it was possible to observe the translocation of nucleic acids and nucleotides, the presence of ubiquitin however did not cause any change in cur-

rent.

The pI of 6.9 of ubiquitin [137, 138] and the pH of approximately 8 of the aqueous droplet solution offer an explanation for the absence of signal in the presence of ubiquitin. In pH 8 solution the ubiquitin is negatively charged. It previously has been observed that negatively charged proteins are not detected with DNA pores [53]. It was proposed that negatively charged proteins, in contrast to positively charged ones, do not interact with the pore walls. Our results are in line with the assumption.

In contrast, oligonucleotides and nucleotides seem to interact with the T pore. Both differ concerning their translocation velocities and their diffusion constants. For dsDNA a diffusion constant of approximately 10 $\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$ and velocities of approximately 1 $\frac{\mathrm{nm}}{\mu\mathrm{s}}$ were determined. The value for the diffusion constant is in the range of previously determined diffusion constants [104].

The sgRNA translocations with velocities at the order of 0.2 $\frac{\mathrm{nm}}{\mu\mathrm{s}}$ and with diffusion constants of approximately 2 $\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$. The diffusion constant is one order of magnitude lower compared to previous models [116]. The diffusion constant of ssDNA in the presence of an additional obstacle is 2.5 $\frac{\mathrm{nm}^2}{\mu\mathrm{s}}$, which implies that it is also an order of magnitude lower compared to previous models [116]. Its velocity of translocation is of approximately 0.4 $\frac{\mathrm{nm}}{\mu\mathrm{s}}$.

The recordings demonstrate that nucleic acids can be sensed with the T pore. The diffusion constants and translocation velocities of dsDNA are higher than those of RNA and those of ssDNA in the presence of an obstacle.

The results indicate faster translocation of dsDNA. The dsDNA offers less moieties for hydrogen bond formation and the interactions of the T pore with the π-orbitals of the nucleotides than ssDNA and RNA. Nucleobases of the ssDNA and the RNA can interact more with the T pore. Hence, dsDNA would interact less with the DNA of the T pore or with blunt ends at the hole of the T pore. Relative to the electric forces acting on the dsDNA, these interactions might be less pronounced for dsDNA than for oligomers. This offers a potential explanation of why translocations of dsDNA are faster than those of oligomers.

RNA translocates with similar diffusion constants as the ssDNA which trespasses the T pore in the presence of an obstacle out of ssDNA. Here, for both molecules, the diffusion constants are two orders of magnitude lower than the free diffusion of ssDNA or RNA predicted by models [116]. The lower values might be a reflection of π-π, π-cation, hydrogen bond or van der Waals interactions between the oligomers

and the pore. Previously, a diffusion constant, which was one order of magnitude lower than freely diffusing ssDNA has been reported for the translocation of ssDNA through the MspA nanopore [114]. Consequently, the oligonucleotides might interact more with the T pore than they interact with MspA. Both the low diffusion constant and the comparatively low translocation velocity indicate interactions of the single-stranded nucleic acids with the sensing zone of the T pore.

Similar to RNA and ssDNA, CTP and GTP can form hydrogen bonds, van der Waals, π-π and π-cation interactions. DNA pores in general and the T pore specifically might enhance all of the interactions. The pore offers positive Mg^{2+} charges, negative charges that are locally close to the positive charges, and an abundance of aromatic rings. As the constriction zone is rather large in comparison to α hederin and the MoS_2 nanopore, the very specific kind of charge distribution in a DNA pore and the simultaneous presence of aromatic rings might be the decisive factor for molecule detection here.

For both MoS_2 and α hederin, dCTP translocates faster than dGTP [129, 133]. These observations are in accordance with our recordings. These are the third documented recordings of NTPs with nanopore sensing and to our knowledge the first experiments that prove the translocation process of GTP and CTP.

4. CRISPR - Double Deactivated Cas12a

Chromatin looping holds a key position in the regulation of gene expression [139]. The existence of loops and their dynamic changes contribute to enhancer-promoter contacts or serve as an insulator, thereupon affecting cellular mechanisms [140, 141]. Diverse highly specific long-range looping interactions have been identified by genome-wide chromatin capture studies [142–144]. The exact mechanisms, however, remain undetermined and studying them remains tedious, as analytical tools are still under development.

Before the advent of CRISPR the available methods either required knowledge of the looping factors and were technically demanding or the approach involved major alteration of the DNA sequence [145–149].

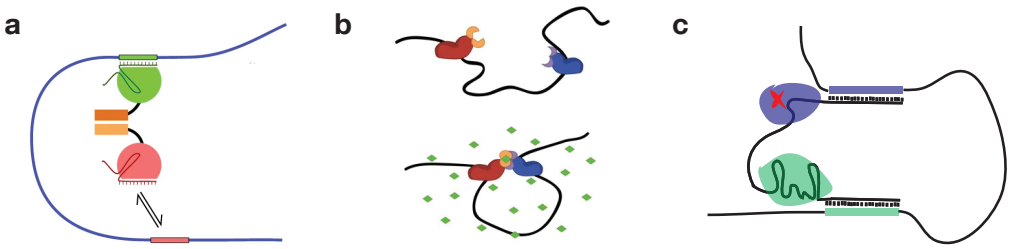

Figure 33. CRISPR enabled DNA looping with a, engineered bivalent dCas9 complexes, image by courtesy of Hao *et al.*, [150] or with the b, the CLOuD9 technique. Image by courtesy of Morgan *et al.* [145]. c, Schematic illustration of the DNA looping approach taken in this thesis. Here the focus is directed towards the rational design of RNA.

In bacteria, DNA loops generated with CRISPR have previously been proven to reinforce enhancer-promoter communication (Fig. 33a) [150]. It was shown that the presence of loops elevates gene expression levels. Additionally, it was shown that in bacteria CRISPR DNA loops can assist the natural formation of loops that repress genes [150]. In mammalian cells, reversible chromatin loops were concep-

tualized similarly and gene expression was modulated with the 'chromatin loop reorganization using CRISPR-dCas9' (CLOuD9) technique (Fig. 33b) [145].

The approaches rely on the use of two Cas9 proteins that originate from different species, which require different RNA handles. The two orthogonal proteins either are fused into one protein [150] or are crosslinked post-transcriptionally [145, 150]. The crosslinking can be permanent [150] or reversible [145]. Reversibility, as it was demonstrated in CLOuD9, is provided by the dimerization of proteins [145]. Both techniques are described in more detail in Ch. 4.1.4.

In our experiments, the looping of DNA evolves around a single RNA that guides both proteins (Fig. 33c). It can bind two Cas proteins, each capable of binding to DNA. The approach focuses on a loop formation that is based on RNA design.

For both dCas9 and dCas12a, the crRNA-Cas complex alters in binding efficiency when the distal end of the spacer is extended by additional bases [151]. When the 5' end of the Cas9 crRNA and the 3' end of the Cas12a crRNA are extended, on-target activity is reduced. Therefore our single RNA that guides dCas9 and ddCas12a to their targets simultaneously is composed of, from 5' to 3': the dCas9 sgRNA, a linker element and the dCas12a crRNA (c9c12aRNA) (Fig. 34).

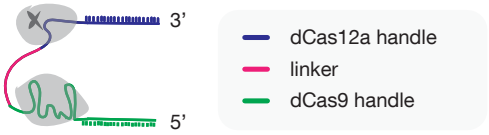

Figure 34. Single RNA that guides dCas9 and ddCas12a to their targets simultaneously (c9c12aRNA). The structure contains the RNA handles that complex with dCas9 and ddCas12a and a linker element.

One obstacle for complexing dCas9, c9c12aRNA, and dCas12a into one functional unit, is the RNA processing activity of dCas12a. Both Cas12a and dCas12a cleave their crRNA at the 5' end of their handle. In the scope of this thesis a Cas12a variant devoid of DNAse and RNAse activity (ddCas12a) was cloned and expressed (Ch. 4.3.1 and 4.3.2).

It is based on the non-DNA cutting D908A AsCas12a, which we obtained as a kind gift by the Finkelstein lab [63]. The second mutation, which eliminates the RNA processing activity, was chosen based on results acquired by Fonfora *et al.* [152] and by Zetsche *et al.*. The H800A mutation which previously displayed the strongest reduction of RNA cutting activity [153] here was chosen to be combined with the D908A mutation. The deactivation of the RNA processing activity of ddCas12a was proven (Ch. 4.3.3) and it was shown that the ddCas12a protein maintains

its DNA binding capability (Ch. 4.3.4). This work examines whether the devised double mutant H800A D908A Cas12a in combination with dCas9 and c9c12aRNA loops DNA.

4.1. Theory

Electrophoretic mobility shift assays, fluorescence anisotropy, size exclusion chromatography, and atomic force microscopy were employed to assess the functionality of ddCas12a.

4.1.1. Electrophoretic Mobility Shift Assay

Electrophoretic mobility shift assays (EMSA) measure the binding of a protein to fluorescently labeled DNA or RNA. The change of electrophoretic mobility of oligonucleotides upon binding to protein results in a band shift in the gel assay. Upon entering the gel matrix the "caging" of the protein-ribonucleotide complex in the gel matrix supports the revoking of dissociation. The proximity in which the two components remain upon dissociation results in a high probability of re-association. Osmolytes like glycerol further support re-association. The repulsion between the osmolyte and the protein backbone creates an environment energetically favorable for the protein in its native bound state [154, 155, 155, 156].

4.1.2. Size Exclusion Chromatography

Size exclusion chromatography (SEC) was used to assess protein-oligonucleotide interactions. The samples dissolved in aqueous solution cross a column that contains the stationary phase and are retained in there for a period which depends on their size. The free energy change of the process can be described by

$$\Delta G = \Delta H - T\Delta S = RT \ln k \tag{4.1}$$

with ΔG, ΔH and ΔS being the free energy, enthalpy and entropy differences, respectively, R is the gas constant, T the absolute temperature, and k the partition coefficient. The majority of chromatographic methods include particle adsorption. For these methods, enthalpy is the main contributor to ΔG. In contrast to other chromatic processes, SEC is entirely driven by entropy, hence $\Delta H = 0$. SEC is described via the thermodynamic retention factor K_D

$$\ln K_D = -\frac{\Delta S}{R} \tag{4.2}$$

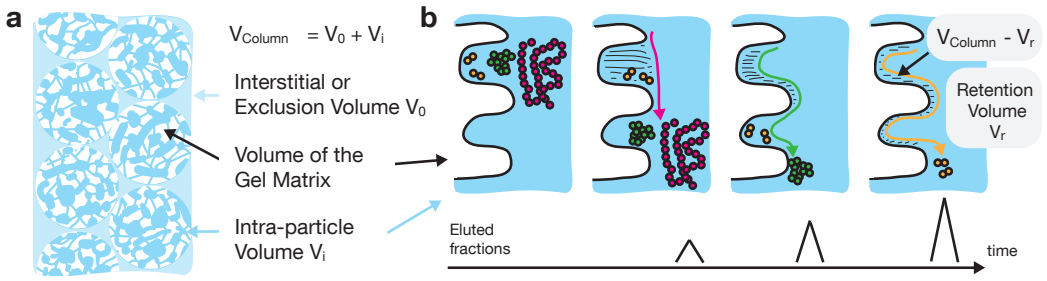

Figure 35. a, Scheme of the inside of a size exclusion column. It is filled with porous beats (white), where the porous inside of the beads constitutes the intra-particle volume V_i. V_i plus interstitial volume V_0 make up for the total column volume V_C. Image adapted from [157] b, Size-exclusion principle. In white the porous material, in yellow, green and pink, particles of increasing size. The largest particle comes along with the smallest retention volume and hence is eluted first. Image adapted from [158]

The retention factor K_D describes the fraction of intraparticle pore volume accessible to the analyte:

$$K_D = \frac{V_r - V_0}{V_i} \tag{4.3}$$

with V_r, V_0, and V_i being the accessible retention volume of the analyte, the interstitial or exclusion volume, and the intra-particle or pore volume ($V_{column} = V_0 + V_i$). Molecules larger than the pore size will not enter the pores, hence the retention volume equals the interstitial volume and $K_D = 0$. Very small molecules, in contrast, fill nearly all the intra-particle volume, hence their retention volume equals ($V_{column} - V_0$) consequently implying a K_D of ≈ 1. Molecule sizes in between the extrema are associated with a K_D between 0 and 1 [159].

4.1.3. Atomic Force Microscopy

In atomic force microscopy (AFM) an image of the specimen is obtained by scanning the surface with a sharp tip that is mounted on a cantilever. Interaction forces between tip and surface, mainly Van-der-Waals and dipole-dipole forces, offer information about the shape of the surface. The interactions cause deflections of the tip that are transmitted to the cantilever. The degree of cantilever bending is recorded by an optical laser-detector system. The small force in the pN to nN regime requires a small mass cantilever with high rigidity, proper vibrational damping, and a piezoelectric stage that moves the sample with a few Angstrom precision relative to the tip.

Due to their weak interaction forces, biological samples require a cantilever with low spring constant or a mode in which the tip oscillates near the resonance frequency. In oscillation mode, the interference of the tip with the sample is reduced.

4.1.4. DNA Looping

CRISPR and its deactivated proteins offer a novel tool kit that can aid in the comprehensive understanding of looping. Any sequence exhibiting a PAM can be connected to another one by fusing two dCas9 which bind to each of the sequences. One study characterized loop formation with either orthogonal dCas9 that were dimerized or orthogonal dCas9 that were fused during translation. There, the loops enhanced the translation of lacZ as well as the translation of an endogenous gene, while establishing loop lengths of up to 11 kbp [150]. Another study explored loop formation in the presence of external stimuli [145]. There, the three-dimensional organization of the chromatin was modified with two orthogonal dCas9, each fused to a reversible, chemically inducible proximity system. Upon addition of the inducer abscisic acid (ABA) two dCas9 in proximity dimerize, thereby remodeling the chromatin structure. Upon washout of ABA the two dCas9 detach which renders the chromatin loop reorganization using CRISPR-dCas9 (CLOuD9) method a reversible process.

The aim here was to expand the CRISPR tool kit by adding another CRISPR mediated loop formation tool. In our case, looping is based on a single RNA that contains handles for both dCas9 and dCas12a (c9c12aRNA). With each handle recruiting one of the proteins the whole complex can bind to two defined positions, as shown in Figure 36.

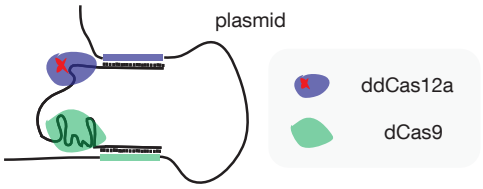

Figure 36. Schematic illustration of DNA looping mediated by an RNA which dCas9 and ddCas12a bind to (c9c12aRNA). The plasmid is indicated in grey, the connecting element consists of dCas9, ddCas12a and c9c12aRNA.

The c9c12aRNA contains three parts: the sgRNA for Cas9, which consists of the spacer and the handle for recruiting dCas9, and the cas12aRNA made up of the handle recruiting dCas12a and the spacer. The two are connected via the third region, from here on referred to as the linker. Hence the final structure corresponds to 5'- sgRNA(cas9)-linker-cas12aRNA -3' (c9c12aRNA).

In this RNA design, no 5' extension resides ahead of the Cas9 spacer, as such an extension could inhibit binding efficiency [151]. Similarly the RNA design enforces that there is no 3' extension following the spacer of Cas12a, which is located at the 3' end of the c9c12aRNA. The extension-free spacers ensure that for neither of the Cas proteins the binding is reduced due to extensions at the spacer, which binds to the target.

The scheme in Figure 37 depicts the desired looping of a plasmid: a dCas9-c9c12aRNA-ddCas12 complex is supposed to form a connection between positions 1 and 6, position 2 and 5 and position 3 and 4. The dCas9 side of the complex bind to positions 1, 2, and 3 respectively, the ddCas12a side of the complex to 4, 5, and 6.

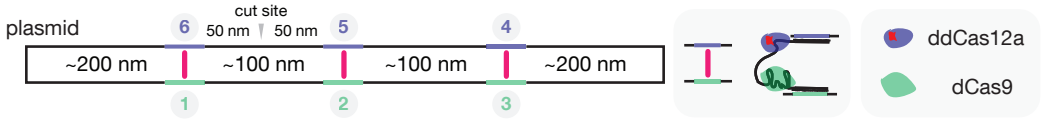

Figure 37. Schema of a sample plasmid, with positions 1 to 6. Cas9 can bind to positions 1, 2, and 3 and ddCas12a can bind to positions 4, 5, and 6. The c9c12aRNA-dCas9-ddCas12a complex can connect 1-6, 2-5, and 3-4.

The plasmid was linearised to enable the identification of the protein positions: positions 5 and 6 both are located at a distance of approximately 50 nm to the cutting site, 4 resides at a distance of 150 nm or 650 nm, 2 and 3 at either 350 nm or 450 nm.

An energy penalty for bending DNA into a defined loop size occurs due to the stiffness of DNA. The looping of DNA inflicts an energy penalty, since DNA is a stiff polymer with a persistence length of l_p of approximately 50 nm [160]. Looping behavior is often described by the worm-like chain model [161, 162]. It offers an approximation of the energy required for bending DNA:

$$E_{bend} = \frac{1}{\beta}\frac{l_p}{2}\int_0^L \kappa(s)^2 ds. \tag{4.4}$$

Here β denotes the thermodynamic energy and κ the curvature of the loop. A loop with a circumference of 200 nm requires a bending energy of approximately 4.9 k_BT. A loop of 600 nm circumference demands for 1.6 k_BT bending energy [162, 163].

DNA looping is described in more detail by the Shimada-Yamakawa equation for DNA shorter than 2000 bp [164] and with the freely jointed chain model by Rippe *et al.* for larger loops [165]. The models describe the Jacobson–Stockmayer (J)

factor, which is the ring closure probability. It is described as the effective relative concentration of the two ends of the loop in respect to each other.

In the freely jointed chain model, the loop J factor as a function of loop length is [165]

$$J_{FJC}(bp) = (1M)\left(\frac{4 \cdot 10^5\, l_p}{bp}\right)^{\frac{3}{2}}.$$

Here *bp* denotes the length of the loop in base pairs.

The freely jointed chain model was unified with the Shimada-Yamakawa equation by Ringrose *et al.*. The analytical description assumes l_p of 50 nm [161]:

$$J(bp) = (1M)\left(\frac{2}{10^2 bp}\right)^{\frac{3}{2}} e^{-\frac{184 \cdot 10^3}{bp^2}}$$

All three of the models are depicted in Figure 38.

The J factor reaches its maximum at a loop length of approximately 500 bp. The length of approximately 500 bp offers the optimum balance between bending energy and the probability that the ends are in close proximity.

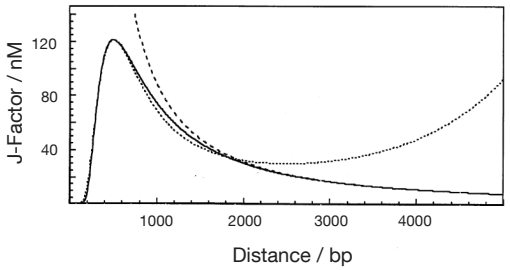

Figure 38. Model for the average J-factor as a function of the distance between two interaction sites [161]. The model unifies the Shimada-Yamakawa equation for DNA shorter than 2000 bp (dotted line) [164] and the freely jointed chain model by Rippe *et al.* for larger loops (dashed line) [165]. Adapted with permission from Ringrose *et al.* [161].

Here, in the case of dCas9-ddCas12 mediated looping, ends of the loop which is formed have a certain end-to-end distance. The end-to-end distance of the starting and end segment of the loop can be approximated with 10 nm, which is approximately equivalent to the sum of the radii of dCas9 and ddCas12a. The analytical description of the J-factor *with an end to end distance* of 10 nm is given by [161]

$$J_{prot}(bp) = 1M\left(\frac{2}{10^2 bp}\right)^{\frac{3}{2}} e^{-\frac{1275 \cdot 10^3}{6.25\, bp^2 + 50^3}}$$

The loops with distances of 200 nm and 600 nm and end-to-end distances of 10 nm leads to J-factors of 33 nM and 113 nM, respectively.

The J factor relates to the free energy ΔF for the deformation of DNA [162]:

$$J = (1M)e^{-\beta\Delta F}.$$

Looping the distances of 200 nm and 600 nm are associated with an energy penalty ΔF of 18.4 k_BT and 16.0 k_BT, respectively.

4.2. Materials and Methods

4.2.1. RNA Synthesis

DNA template sequences were obtained by PCR amplification of gBlocks (Integrated DNA Technologies). The obtained PCR product was purified with the Monarch PCR and DNA Cleanup Kit (New England BioLabs Inc) and quantified by its absorption on a NanoPhotometer (Implen).

4.2.2. H800A D908A Cas12a

Mutation Site-directed mutagenesis was used to introduce H800A to the plasmid pdAsCas12a, kindly provided by the Finkelstein lab. Site-directed mutagenesis by substitution includes a linearization step with Phusion High-Fidelity PCR DNA Master Mix (New England Biolabs), which introduces the point mutation [166]. The mutated linear DNA was phosphorylated and ligated with a T4 kinase-T4 ligase mix and a DpnI digest was performed.

The plasmid was chemically transformed into the *E. coli* strain DH5α [167], the integrity of the protein sequence (Fig. 40 b) and the presence of the mutation (Fig. 40 c) were confirmed by DNA sequencing (GATC) of the region of interest (B.2.1). Protein was expressed in *E. coli* BL21$*$(DE3) which were grown to an OD of 0.6, induced with 1 mM IPTG for 24 hours at 18 °C, pelleted, flash-frozen, and stored at −80 °C.

Purification Cells were lysed with a French press (Fig. 41 a) or via sonication (50 % amplitude, 10 s sonication and 10 s break for 10 min) in lysis buffer (1 M NaCl, 20 mM HEPES, 5 % glycerol, 1 mM EDTA, 0.1 % Tween20). The lysate was spun for 1 h at 4 °C, the supernatant was separated from debris with 22 µm filters. With the protein containing both a His and a Strep Tag, either a HisTrap HP column (GE Healthcare) or StrepTrap HP column (GE Healthcare) was equilibrated with wash buffer (1 M NaCl, 20 mM Hepes, 5 % glycerol, 5 mM $MgCl_2$) and the bacterial lysate was applied on the column. HisTrap columns were incubated with 5 ml HisTrap elution buffer (wash buffer plus imidazole, Carl Roth) for 5 min and another 40 ml of HisTrap elution buffer were utilized for protein elution (Fig. 41 b). For StrepTrap columns, 1 ml of StrepTrap elution buffer (wash buffer and 10 mM desbiotin, Sigma-Aldrich) was applied onto the column and incubated for 5 min and 6 ml of StrepTrap

elution buffer were used for protein elution.

The tags are fused to ddCas12a via a small ubiquitin-like modifier (SUMO)-tag which can be cleaved off by SUMO protease. The SUMO tag allows for the removal of the His, Strep and Flag tags which are fused to the N-terminus, and creates a native N-terminus of ddCas12a. SUMO protease was expressed in *E. coli* BL21∗(D3), lysis was performed as described above and the filtered lysate was purified with Ni-NTA resin columns (ThermoFisher Scientific).

The His- or StrepTrap purified ddCas12a was incubated and dialyzed with an excess of SUMO in a Slide-A-Lyzer dialysis cassette kit overnight (10 kDa, 3 ml, Thermo Fischer Scientific) in S200 Proteolysis Buffer (500 mM NaCl, 50 mM HEPES, 5 mM $MgCl_2$, 1 mM DTT). A Superdex 200 size exclusion column (GE Healthcare) was equilibrated with S200 buffer (500 mM NaCl, 20 mM HEPES, 5 mM $MgCl_2$, 2 mM DTT) in an ÄKTA Pure (GE Healthcare). The content of the dialysis cassette was applied to the S200 column, the fractions (Fig. 41 d) were analyzed on SDS PAGE (Fig. 41 c). Protein was stored in 50 % glycerol at −20 °C for further usage.

4.2.3. Target DNA

The target plasmid was extracted with a plasmid purification kit (QIAGEN, Maxiprep), separated on a 1 % agarose gel, and purified with the Monarch DNA Gel Extraction Kit. For linearization, plasmids were incubated with a restriction enzyme (eg PvuII/ScaI-HF, New England BioLabs Inc) at 37 °C for several hours, gel purified and extracted with the Monarch DNA Gel Extraction Kit (sequences see Appendix B.2.1).

The 50bp long, linear target DNA was annealed contain one ssDNA fluorescently labeled oligonucleotide (Eurofins) either tagged with Cy3 or Cy5 (sequences see Appendix B.2.1).

4.2.4. RNA Transcription and Design

RNA was transcribed as stated in 3.2.1.

Linking elements The linking element can be chosen freely and five types of linking element were examined here. Two straight connections of 20 and 30bp length (A20 and N30), two elements which include a hairpin (2hp and 2hplong), and one

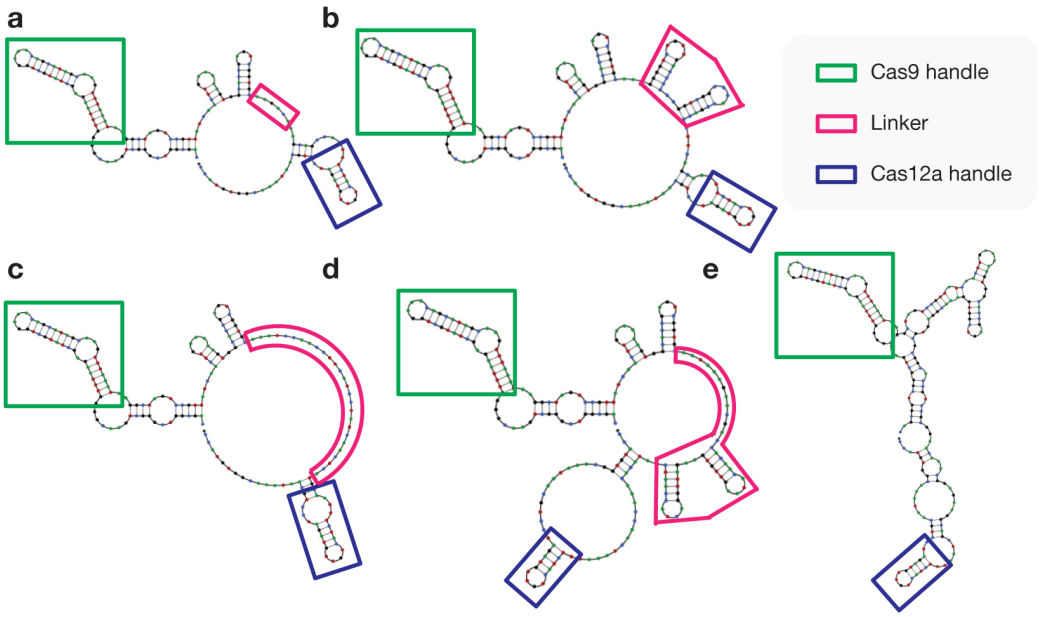

Figure 39. The c9c12RNA with five kinds of linker elements. All NUPACK designs show the handle for Cas9 (green) and Cas12a (red). The structures vary in the linker (blue): a, the natural separator b, the natural separator plus a hairpin c, a straight 30 nucleotide sequence d, two hairpins and a straight sequence of 20 nucleotides e, the beet western yellow virus pseudoknot

element which forms a pseudoknot (PK). The PK is derived from the beet western yellow virus (BWYV). Sequences are listed in the Appendix B.2.1.

4.2.5. Electrophoretic Mobility Shift Assay

EMSAs were performed either at room temperature or at 4 °C to minimize dissociation. The 1 x TBE running buffer was supplemented with magnesium acetate. The running buffer was adjusted to pH 8. A 4 % stacking gel was placed on top of the 8 % to 10 % running gel. Alternatively, precast 4 % to 20 % gradient gels (Biorad) were used. PAGE was run at 90 V for 20 min to equilibrate the gel. Gels were stained with Sybr Green II RNA Gel Stain (Invitrogen, Thermo Fischer Scientific), the readout was performed with a Typhoon FLA 9500 laser scanner (GE Healthcare).

4.2.6. Fluorescence Anisotropy

Fluorescence anisotropy measurements were performed as an alternative to EMSA for assessing the binding of protein to labeled DNA or RNA or DNA with a fluo-

rescence spectrometer (Fluorolog 3, HORIBA). A starting volume of 80 µl contains either 20 nM to 50 nM fluorescently labeled DNA or 1 µl of gel purified, Atto 488 labeled RNA in the range of 1 µM in 1 x NEB 3.1 buffer. Upon titration of protein or ribonucleoprotein the intensities perpendicular and parallel to the polarizer were measured (Atto 488 excitation at 501 nm, emission at 523 nm; Cy3 excitation at 548 nm emission at 561 nm; Cy5 excitation at 647 nm emission at 665 nm)

4.2.7. Atomic Force Microscopy

For imaging Cas proteins and plasmid, a concentration of 0.5 to 1 $\frac{\text{ng}}{\mu\text{l}}$ of the approximately 2.8 kbp plasmid (Appendix B.2.1, pSB3A1-T7wt-mVenus, kindly provided by E. Falgenhauer) was incubated on mica for 2 min and imaged in air. For pinning the complex onto the mica 10 to 20 mM NaCl and 2 mM $MgCl_2$ were used.

4.3. Results and Discussion

Within the scope of this thesis, a dCas12a was mutated into a version devoid of its RNA processing capability. Its functionality was proven and it was examined whether ddCas12a can be used for RNA-guided DNA looping.

4.3.1. H800A D908A Mutated Cas12a

The PCR mutated, linearized, recircularized pdAsCas12a was sequenced, the results are shown in Figure 40. Sequencing confirms the successful mutation of his-

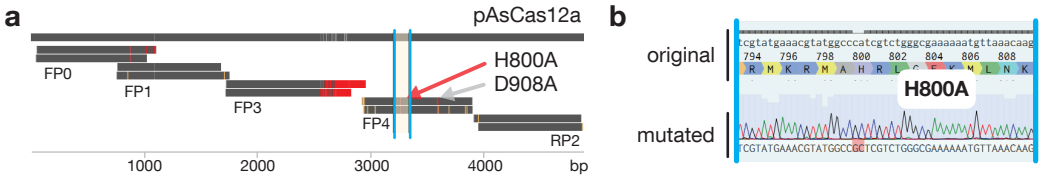

Figure 40. Site directed mutagenisis of pdAsCas12a. a, Sequencing data of the entire sequence that encodes ddAsCas12a with FP and RP indicating the forward and reverse primer, respectively. b, The sequencing data of pddAsCas12a proves the successful H800A mutation.

tidin to alanine at amino acid 800, which here is equivalent to a change of CA to GC (Fig 40b). The plasmid of H800A D908A AsCas12a from here on is referred to as pddAsCas12a. Besides the two intentional mutations, the protein-coding region of pddCas12a is mutation free, as confirmed by sequencing (Fig 40a). The primer indicated in Figure 40a are listed in the Appendix B.2.1.

4.3.2. Expression and Purification of ddCas12a

Figure 41 shows successful expression and purification of ddCas12a. The lysate of the BL21∗(D3) cells indicates the expression of the 171 kDa His6-TS-SUMO-ddAsCpf1-3xFLAG protein, as the SDS PAGE in Figure 41a shows.

HisTrap purification removed the majority of unwanted cell material and concentrates the tagged protein. The result is shown in the SDS PAGE in Figure 41b. The tags are cleaved off at the SUMO recognition site by the SUMO protease. The successful removal reflect in the absorbance values of the size exclusion chromatography. The absorbance values depicted in Figure 41c exhibit a peak at fraction 6 and 7, one between fractions 9 to 11 and a third one at fraction 16 and 17. The first

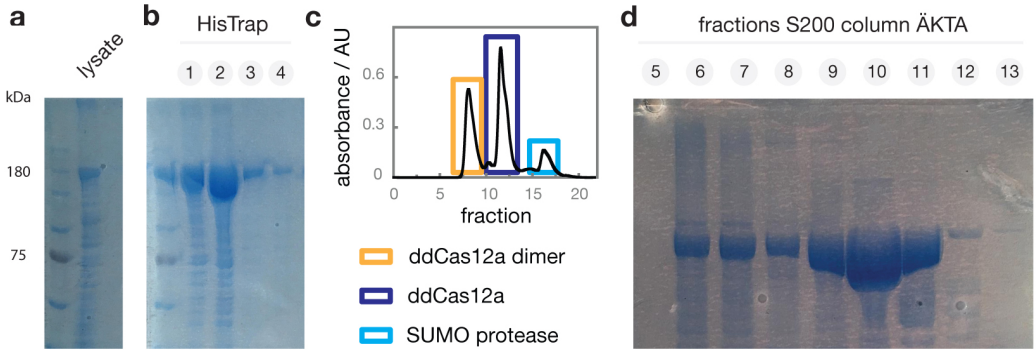

Figure 41. Purification of ddCas12a. a, 8 % SDS-PAGE of cell lysate. BL21∗(D3)-pddCas12a were induced with IPTG at 16 °C for 12 h and lysed with a French press. b, 8 % SDS-PAGE of the fractions obtained by His-tag purification. c, Chromatogram of the fractions obtained by size exclusion chromatography d, 8 % SDS-PAGE of the corresponding fractions

peak, equaling the largest molecule, is caused by a ddCas12a dimers or aggregates, which is confirmed by the SDS PAGE (Fig. 41d). The second peak originates from 154 kDa ddAsCas12a. The 21 kDa SUMO protease causes the third peak. The results are confirmed by 10 % SDS PAGE (Fig. 41d).

4.3.3. Deactivated RNA Processing of ddCas12a

The ddCas12a was examined upon its RNase activity. A concentration of 70 nM of

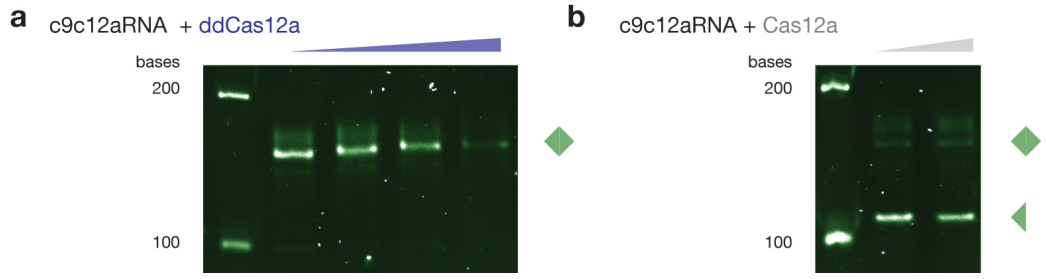

Figure 42. Absence of RNA processing in H800A D908A Cas12a. 10 % denaturing PAGE of 70 nM c9c12RNA and increasing concentration of a, ddCas12a (0.4 µM, 0.7 µM, 1.4 µM and 2.8 µM) or b, Cas12a (0.7 µM and 2.8 µM). The diamond shape indicates the c9c12RNA, the triangle indicates one part of the processed RNA.

c9c12RNA (7-8c9c12RNAstr) were exposed to 0.4 µM, 0.7 µM, 1.4 µM and 2.8 µM ddCas12a and 0.7 µM and 2.8 µM Cas12a for one hour. Figure 42b shows that both 0.7 µM and 2.8 µM Cas12a processed the RNA close to entirely and to the same

degree. In the presence of ddCas12a the percentage of cut RNA remains zero for all ddCas12a concentrations (Fig. 42b). For 2.8 µM of ddCas12a, a high percentage of c9c12RNA disappears from the gel. Its disappearance can be attributed to the formation of aggregates of the highly concentrated ddCas12a and RNA. However, even at the high concentration of 2.8 µM c9c12aRNA is not processed.

4.3.4. DNA Binding Capability of ddCas12a

With EMSA it was examined, whether the ddCas12a-RNA ribonucleoprotein is capable of binding to DNA. Fig. 43 shows 50 nM target DNA, here DNA08, in

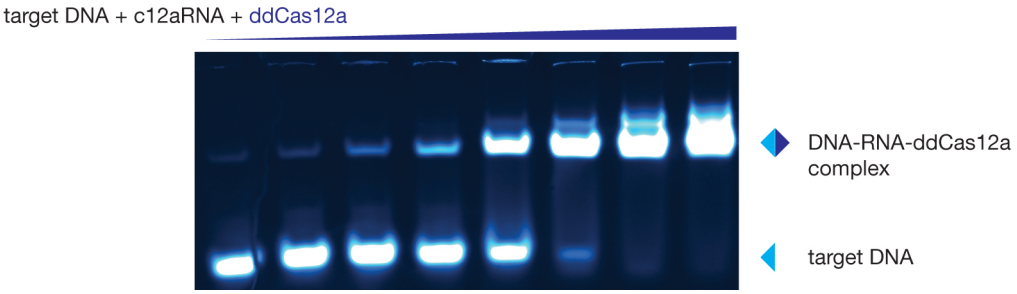

Figure 43. EMSA of the ddCas12a-c12RNA RNP binding to its target DNA. Protein concentration from left to right are 0.4 µM, 0.5 µM, 0.7 µM, 0.9 µM, 1.1 µM, 1.3 µM, 1.4 µM and 1.6 µM, the concentration of target DNA is 50 nM.

presence of the ribonucleoprotein. The c12aRNA-ddCas12a ribonucleoprotein was pre-assembled. The ddCas12a with final concentrations of 0.4 µM, 0.5 µM, 0.7 µM, 0.9 µM, 1.1 µM, 1.3 µM, 1.4 µM and 1.6 µM was added to a constant concentration of c12aRNA of approximately 500 nM. The EMSA shows an increase of the DNA-RNA-ddCas12a complex with increasing ddCas12a concentration. For 1.1 µM around 50 % of target DNA is bound, for 1.4 µM all DNA is bound. The high excess of RNA might account for the fact that 1.1 µM of ddCas12a are required for binding 50 % of the DNA.

The binding of the c12aRNA-ddCas12a ribonucleoprotein to DNA was imaged with AFM (Fig. 44). Rough approximations of the distances were performed contouring the DNA strand and performing a rule of three. Fig. 44 shows example images of dCas9 and ddCas12a binding to the linearized 2.8 kbp plasmid. The design of the 1-6 c9c12aRNA_N14, 2-5 c9c12aRNA_N14 and 3-4 c9c12aRNA_N14 were chosen, such that dCas9 binds in positions 1, 2 and 3 and ddCas12a binds in positions 4, 5 and 6.

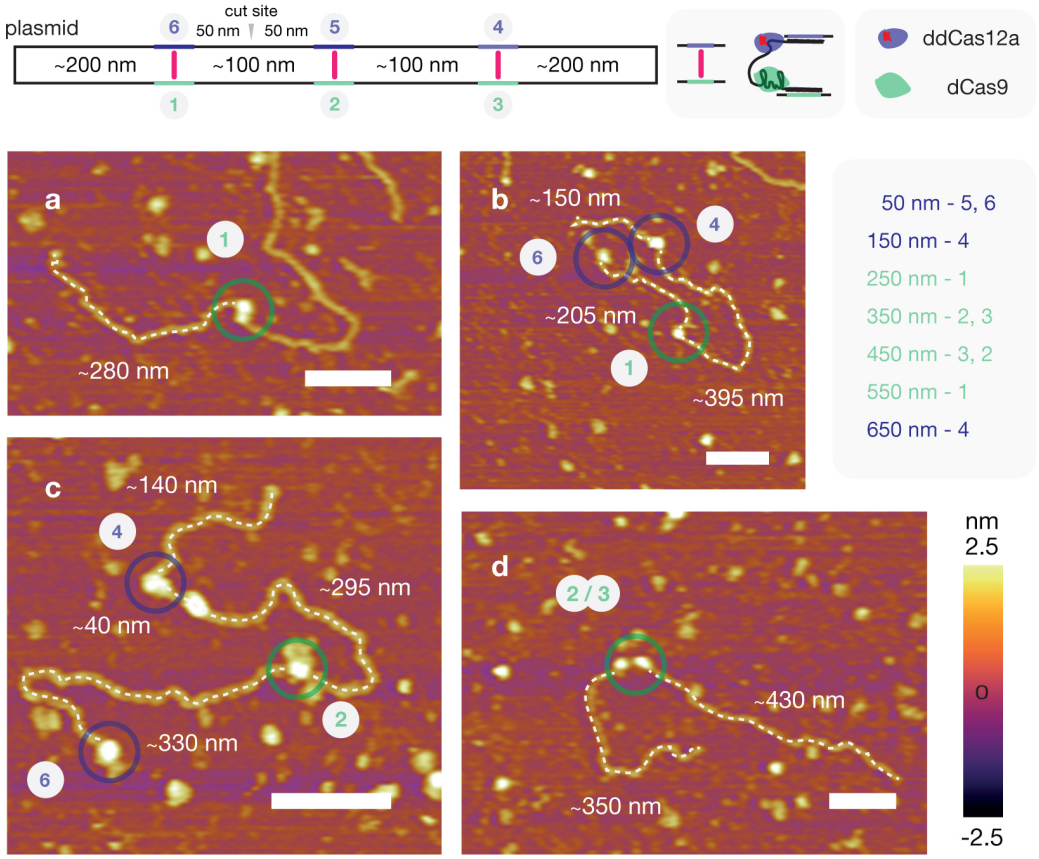

Figure 44. AFM images of dCas9 and dCas12a binding to dsDNA. The dCas9 is supposed to bind to positions 1, 2, and 3, the ddCas12a to positions 4, 5, and 6 at the linearized 2.8 kbp plasmid. The positions can be identified with regard to the distance to the blunt ends. The distances of position 1 to the blunt ends are 250 and 550 nm. Those of positions 2 and 3 are 350 and 450 nm, those of positions 4 are 150 and 650 nm, those of positions 5 and 6 are 50 nm. The assigned binding positions and according proteins are a, position 1 and dCas9 b, positions 4, 1, and 6 and ddCas12a, dCas9 and ddCas12a, respectively, c, positions 4, 2 and 6 and ddCas12a, dCas9 and ddCas12a, respectively, d, an example of ambiguous position identification, with dCas9 either binding to position 2 or 3. Scale bars: 100 nm.

The proteins were located at the positions, where dCas9 and ddCas12a are supposed to bind, as indicated with numbering in Fig. 44. The AFM images show distinct binding positions that were identified as positions 1, 2, 3, 4 and 6. Positions 2 and 3 cannot always be identified unambiguously, as can be seen in image Fig. 44 d.

A schematic illustration of two linearized pddCas12 plasmids and the binding positions 7 for dCas9 and 8 for ddCas12a are shown in Figure 45 a. As indicated, a dCas9-c9c12RNA-ddCas12a complex can link two linearized plasmids. Fig. 45b depicts a pristine plasmid without CRISPR proteins for reference. In Fig. 45c, two

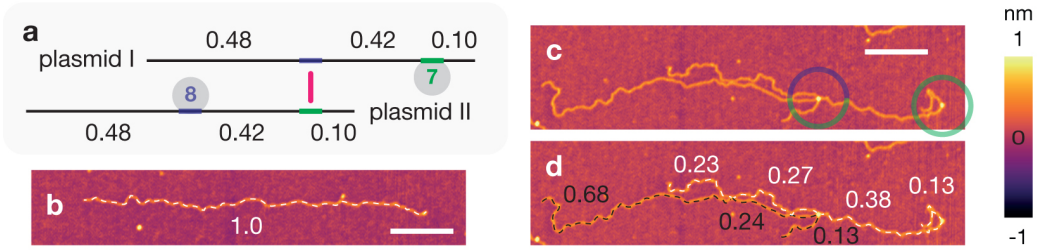

Figure 45. AFM images of Cas proteins binding to the 10 kbp plasmid pddCas12a. a, Schematic illustration of dCas9 and ddCas12a binding sites 7 and 8 on the 10kbp plasmid pddCas12a, indicated in green and purple, respectively. The numbers indicate the length as a fraction of the total plasmid length. b, A pristine pddCas12a plasmid for reference. c, Two pddCas12a plasmids with two possible binding sites each. The green circle indicates binding of dCas9 and the green/purple circle indicates binding of either a dCas9, a ddCas12a, or a dCas9-ddCas12a complex. These conclusions were drawn from the information presented in d. d, The numbers indicate distances as a fraction of the total plasmid length. The two plasmids are highlighted in white and black. The distances of the proteins to the end of the plasmids imply that the protein binding on the right end is a dCas9 and that the protein in the center-right is either a dCas9, a ddCas12a, or a dCas9-ddCas12a complex. Scale bars: 500 nm.

potential interaction locations of Cas proteins with DNA are pointed out by circles. The lengths of plasmid sections as the fraction of the total length are shown in Fig. 45d.

The protein located at the right side resides at a distance of approximately 1300 bp, or 13 % of the total plasmid length, from the end of the plasmid. The protein at the medium right position is in the middle of the plasmid that is highlighted in white. It simultaneously is located at a distance of 13 % of total plasmid length from one blunt end of the plasmid that is contoured in black.

According to the positioning, the protein located at the right can be identifed as dCas9. The protein at the medium right position could either be identified as a ddCas12a residing on the plasmid that is contoured in white, a dCas9 residing at the plasmid that is contoured in black, or as a c9c12aRNA-dCas9-ddCas12a complex, which connects both proteins.

4.3.5. Discussion

The protein D908A AsCas12a was successfully mutated into H800A D908A AsCas12a. It demonstrated that the newly obtained ddCas12a is devoid of RNase activity. With EMSA and AFM studies it was proven that the DNA binding capability of ddCas12a was not inhibited by the introduction of the H800A mutation.

AFM clarified that the rationally designed c9c12aRNA is functionally available for each of the proteins, dCas9 and ddCas12a, individually. AFM showed that both ribonucleoproteins, c9c12aRNA-dCas9 and c9c12aRNA-ddCas12a, bind to defined positions on a plasmid. Further, a potential c9c12aRNA-dCas9-ddCas12a-mediated connection of two plasmids was identified.

These steps lay the foundation for c9c12aRNA-dCas9-ddCas12a-mediated DNA-loop formation.

5. Outlook

In summary, several functional DNA and RNA secondary structures were investigated with single molecule methods. All the structures rely on the natural versatility of nucleic acids.

The large-diameter, membrane-spanning DNA nanopore showed ohmic conductance and a stable transmembrane current. This current was mostly devoid of gating events, which made the DNA nanopore an appropriate instrument for sensing the electrically driven translocation of single molecules. The use of the DNA nanopore allowed to detect the translocation of dsDNA, sgRNA and rNTPs. The findings suggests to elucidate whether the T pore or differently designed DNA nanopores hold the potential to discriminate different kinds of rNTPs, NMPs or dNMPs. Further investigation could unravel whether the use of ionic liquids or a change in stem design can improve both detection and differentiation of these molecules.

ELP-mediated and streptavidin-mediated anchoring strategies were implemented for DNA pores. They expand the set of available anchoring options and both approaches entail future advantages.

The streptavidin-mediated membrane incorporation of DNA nanopores allows for pores that lack hydrophobic modifications and can contribute to the reduction of pore aggregation. It also demonstrates that the incorporation of a DNA nanopore into a membrane can be mediated by protein-ligand interaction, an approach that could be expanded to other protein-ligand interactions.

The introduction of ELPs offers a potential means to use a DNA pore of which the incorporation status can be switched, since the hydrophobicity of ELPs can depend on external stimuli. ELPs as anchoring elements could open a pathway towards externally triggered DNA nanopore incorporation. Timed alterations in pH or salt concentration could enable time specific incorporation. Spatial differences in pH or

salt concentration could define localized spaces for pore incorporation.

A new variant of Cas12a, ddCas12a, was designed, expressed, purified and its activity was demonstrated. The newly designed ddCas12a can be further exploited. It can therefore be used in place of dCas9 or as an orthogonal additional to dCas9.

Finally, the functional secondary RNA structure c9c12aRNA merits further investigation. The rational design of RNA nanostructures could lead to novel insights on how CRISPR-RNA mediated DNA loops can be generated. Synthetically engineered DNA loops remain a largely unexplored domain, which could be investigated with this method.

Bibliography

[1] Watson, J. *et al.* Molecular Biology of the Gene (1965).

[2] Villa, J. K., Su, Y., Contreras, L. M. & Hammond, M. C. Synthetic biology of small RNAs and riboswitches. *Regulating with RNA in Bacteria and Archaea* **6**, 527–545 (2018).

[3] Kruger, K. *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**, 147–157 (1982).

[4] Nudler, E. & Mironov, A. S. The riboswitch control of bacterial metabolism. *Trends in Biochemical Sciences* **29**, 11–17 (2004).

[5] Hadjicharalambous, M. R. & Lindsay, M. A. Long non-coding RNAs and the innate immune response. *Non-Coding RNA* **5**, 34 (2019).

[6] Jinek, M. *et al.* A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

[7] Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).

[8] Geary, C., Rothemund, P. W. & Andersen, E. S. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science* **345**, 799–804 (2014).

[9] Berg, J. M., Stryer, L. & Tymoczko, J. L. *Stryer Biochemie* (Springer-Verlag, 2015).

[10] Ussery, D. W. DNA structure: A-, B-and Z-DNA helix families. *e LS* (2001).

[11] Forterre, P. & Grosjean, H. The interplay between RNA and DNA modifications: back to the rna world. In *Madame Curie Bioscience Database [Internet]* (Landes Bioscience, 2013).

[12] Mao, C., Sun, W. & Seeman, N. C. Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy. *Journal of the American Chemical Society* **121**, 5437–5443 (1999).

[13] Thelander, L. Advanced information on the nobel prize in chemistry 2006 (2006).

[14] Régnier, P. & Marujo, P. E. Polyadenylation and degradation of RNA in prokaryotes. In *Madame Curie Bioscience Database [Internet]* (Landes Bioscience, 2013).

[15] Condon, C. *Molecular biology of RNA processing and decay in prokaryotes* (Academic Press, 2009).

[16] Chopra, A. *Nucleic Acid Nanostructures: From Biomimetic Assembly to in vivo Applications.* Dissertation, TU München (2018).

[17] Weinberg, Z. *et al.* Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome biology* **11**, R31–R31 (2009).

[18] Burgstaller, P. & Famulok, M. Isolation of RNA aptamers for biological cofactors by in vitro selection. *Angewandte Chemie International Edition in English* **33**, 1084–1087 (1994).

[19] Sun, H. *et al.* Oligonucleotide aptamers: new tools for targeted cancer therapy. *Molecular Therapy* **3**, e182–e182 (2013).

[20] Beckmann, B. M., Castello, A. & Medenbach, J. The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. *Pflügers Archiv-European Journal of Physiology* **468**, 1029–1040 (2016).

[21] Alberts, B. *et al.* The lipid bilayer. In *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).

[22] Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology* **9**, 112–124 (2008).

[23] Alberts, B. *et al.* The compartmentalization of cells. In *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).

[24] Sitte, P., Ziegler, H., Ehrendorfer, F. & Bresinsky, A. *Strasburger - Lehrbuch der Botanik für Hochschulen* (G. Fischer, 1998).

[25] Cullis, P. R., Hope, M. J. & Tilcock, C. P. Lipid polymorphism and the roles of lipids in membranes. *Chemistry and Physics of Lipids* **40**, 127–144 (1986).

[26] Ashrafuzzaman, M. & Tuszynski, J. A. *Membrane Biophysics* (Springer Science & Business Media, 2012).

[27] Krishnan, S. *Building Biomimetic Systems using Vesicles and DNA Channels.* Dissertation, TU München (2018).

[28] Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nature Ecology & Evolution* 1–18 (2018).

[29] Levitan, I. & Barrantes, F. *Cholesterol Regulation of Ion Channels and Receptors* (John Wiley & Sons, 2012).

[30] Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720–731 (1972).

[31] Almeida, P. & Vaz, W. Lateral diffusion in membranes. *Handbook of Biological Physics* **1**, 305–357 (1995).

[32] Wanunu, M. *et al.* Rapid electronic detection of probe-specific micrornas using thin nanopore sensors. *Nature Nanotechnology* **5**, 807 (2010).

[33] Venta, K. *et al.* Differentiation of short, single-stranded DNA homopolymers in solid-state nanopores. *ACS Nano* **7**, 4629–4636 (2013).

[34] Liu, K., Feng, J., Kis, A. & Radenovic, A. Atomically thin molybdenum disulfide nanopores with high sensitivity for DNA translocation. *ACS Nano* **8**, 2504–2511 (2014).

[35] Zhou, Z. *et al.* Dna translocation through hydrophilic nanopore in hexagonal boron nitride. *Scientific Reports* **3**, 3287 (2013).

[36] Marx, S. K. *et al.* De novo design of ion conducting transmembrane protein nanopores. *Biophysical Journal* **118**, 483a–484a (2020).

[37] Wang, S., Zhao, Z., Haque, F. & Guo, P. Engineering of protein nanopores for sequencing, chemical or protein sensing and disease diagnosis. *Current Opinion in Biotechnology* **51**, 80–89 (2018).

[38] Bhakdi, S. & Tranum-Jensen, J. Alpha-toxin of staphylococcus aureus. *Microbiological Reviews* **55**, 733–751 (1991).

[39] Manrao, E. A., Derrington, I. M., Pavlenok, M., Niederweis, M. & Gundlach, J. H. Nucleotide discrimination with DNA immobilized in the MspA nanopore. *PloS one* **6**, e25723 (2011).

[40] Rothemund, P. W. Design of DNA origami. In *ICCAD-2005. IEEE/ACM International Conference on Computer-Aided Design, 2005.*, 471–478 (IEEE, 2005).

[41] Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485**, 623–626 (2012).

[42] Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338**, 1177–1183 (2012).

[43] Burns, J. R., Stulz, E. & Howorka, S. Self-assembled DNA nanopores that span lipid bilayers. *Nano Letters* **13**, 2351–2356 (2013).

[44] Burns, J. R. *et al.* Lipid-bilayer-spanning DNA nanopores with a bifunctional porphyrin anchor. *Angewandte Chemie International Edition* **52**, 12069–12072 (2013).

[45] Seifert, A. *et al.* Bilayer-spanning DNA nanopores with voltage-switching between open and closed state. *ACS Nano* **9**, 1117–1126 (2015).

[46] Burns, J. R., Seifert, A., Fertig, N. & Howorka, S. A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nature Nanotechnology* **11**, 152 (2016).

[47] Göpfrich, K. *et al.* DNA-tile structures induce ionic currents through lipid membranes. *Nano Letters* **15**, 3134–3138 (2015).

[48] Burns, J. R., Al-Juffali, N., Janes, S. M. & Howorka, S. Membrane-spanning DNA nanopores with cytotoxic effect. *Angewandte Chemie International Edition* **53**, 12466–12470 (2014).

[49] Krishnan, S. & Simmel, F. C. Nanotechnology: deadly DNA. *Nature Chemistry* **7**, 17 (2015).

[50] Langecker, M. *et al.* Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338**, 932–936 (2012).

[51] Krishnan, S. *et al.* Molecular transport through large-diameter DNA nanopores. *Nature Communications* **7**, 1–7 (2016).

[52] Göpfrich, K. *et al.* Large-conductance transmembrane porin made from DNA origami. *ACS Nano* **10**, 8207–8214 (2016).

[53] Diederichs, T. *et al.* Synthetic protein-conductive membrane nanopores built with DNA. *Nature Communications* **10**, 1–11 (2019).

[54] Thomsen, R. P. *et al.* A large size-selective DNA nanopore with sensing applications. *Nature Communications* **10**, 1–10 (2019).

[55] Göpfrich, K. *et al.* Ion channels made from a single membrane-spanning DNA duplex. *Nano Letters* **16**, 4665–4669 (2016).

[56] Xiao, Y., Ng, S., Nam, K. H. & Ke, A. How type ii crispr–cas establish immunity through cas1–cas2-mediated spacer integration. *Nature* **550**, 137–141 (2017).

[57] Fagerlund, R. D. *et al.* Spacer capture and integration by a type if cas1–cas2-3 crispr adaptation complex. *Proceedings of the National Academy of Sciences* **114**, E5122–E5128 (2017).

[58] Lander, E. S. The heroes of crispr. *Cell* **164**, 18–28 (2016).

[59] Pyzocha, N. K., Ran, F. A., Hsu, P. D. & Zhang, F. Rna-guided genome editing of mammalian cells. In *Gene Correction*, 269–277 (Springer, 2014).

[60] Xu, X. & Qi, L. S. A crispr–dcas toolbox for genetic engineering and synthetic biology. *Journal of Molecular Biology* **431**, 34–47 (2019).

[61] Ran, F. A. *et al.* Genome engineering using the crispr-cas9 system. *Nature Protocols* **8**, 2281 (2013).

[62] Sundaresan, R., Parameshwaran, H. P., Yogesha, S., Keilbarth, M. W. & Rajan, R. Rna-independent DNA cleavage activities of cas9 and cas12a. *Cell Reports* **21**, 3728–3739 (2017).

[63] Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J. & Russell, R. Kinetic basis for DNA target specificity of crispr-cas12a. *Molecular Cell* **71**, 816–824 (2018).

[64] Pickar-Oliver, A. & Gersbach, C. A. The next generation of crispr–cas technologies and applications. *Nature Reviews Molecular Cell Biology* **20**, 490–507 (2019).

[65] Chen, J. S. & Doudna, J. A. The chemistry of cas9 and its crispr colleagues. *Nature Reviews Chemistry* **1**, 0078 (2017).

[66] Swarts, D. C., van der Oost, J. & Jinek, M. Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-cas12a. *Molecular Cell* **66**, 221–233 (2017).

[67] Bozzola, J. J. & Russell, L. D. *Electron microscopy: Principles and Techniques for Biologists* (Jones & Bartlett Learning, 1999).

[68] Adamson, N. & Reynolds, E. Rules relating electrophoretic mobility, charge and molecular size of peptides and proteins. *Journal of Chromatography B: Biomedical Sciences and Applications* **699**, 133–147 (1997).

[69] Magdeldin, S. *Gel Electrophoresis-Principles and Basics* (2012).

[70] Sheehan, D. *Physical Biochemistry: Principles and Applications* (John Wiley & Sons, 2013).

[71] Iwamoto, M. & Oiki, S. Lipid bilayer experiments with contact bubble bilayers for patch-clampers. *JoVE (Journal of Visualized Experiments)* e58840 (2019).

[72] Howorka, S. & Siwy, Z. Nanopore analytics: sensing of single molecules. *Chemical Society Reviews* **38**, 2360–2384 (2009).

[73] Ziegler, D. *Nanopore Force Spectroscopy using the Droplet Bilayer Interface Setup - Characterization of poly(A) tailed RNA Aptamers and Synthetic DNA Membrane Pores.* Master thesis", school = TU München, year = 2015,.

[74] Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, 2001).

[75] Wen, C., Zhang, Z. & Zhang, S.-L. Physical model for rapid and accurate determination of nanopore size via conductance measurement. *ACS Sensors* **2**, 1523–1530 (2017).

[76] Feng, D. & Holland, L. A dc method for the absolute determination of conductivities of the primary standard kcl solutions from 0c to 50c. *Journal of Research of the National Institute of Standards and Technology* **99** (1994).

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

[78] Douglas, S. M. *et al.* Rapid prototyping of 3d DNA-origami shapes with caDNAno. *Nucleic Acid Research* **37**, 5001–5006 (2009).

[79] Andersson, M. *et al.* Vesicle and bilayer formation of diphytanoylphosphatidylcholine (DPhPC) and diphytanoylphosphatidylethanolamine (DPhPE) mixtures and their bilayers' electrical stability. *Colloids and Surfaces B: Biointerfaces* **82**, 550–561 (2011).

[80] Maglia, G., Heron, A. J., Stoddart, D., Japrung, D. & Bayley, H. Analysis of Single Nucleic Acid Molecules with Protein Nanopores. In *Single Molecule Tools, Part B:Super-Resolution, Particle Tracking, Multiparameter, and Force Based Methods*, 591–623 (Elsevier, 2010).

[81] Lindsey, H., Petersen, N. O. & Chan, S. I. Physicochemical characterization of 1,2-diphytanoyl-sn-glycero-3-phosphocholine in model membrane systems. *Biochimica et Biophysica Acta* **555**, 147–167 (1979).

[82] Yasmann, A. & Sukharev, S. Properties of diphytanoyl phospholipids at the air-water interface. *Langmuir* **31**, 350–357 (2015).

[83] Barriga, H. M. *et al.* Droplet interface bilayer reconstitution and activity measurement of the mechanosensitive channel of large conductance from escherichia coli. *Journal of the Royal Society Interface* **11**, 20140404 (2014).

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

[85] Rudolf, G. C. & Sieber, S. A. Copper-assisted click reactions for activity-based proteomics: fine-tuned ligands and refined conditions extend the scope of application. *ChemBioChem* **14**, 2447–2455 (2013).

[86] Matejka, N. Bachelor Thesis - Characterization and Translocation Experiments with a DNA-based Nanopore (2016).

[87] Chopra, A., Krishnan, S. & Simmel, F. C. Electrotransfection of polyamine folded DNA origami structures. *Nano Letters* **16**, 6683–6690 (2016).

[88] Wagenbauer, K. F. *et al.* How we make DNA origami. *ChemBioChem* **18**, 1873–1885 (2017).

[89] Krishnan, Y. & Simmel, F. C. Nucleic acid based molecular devices. *Angewandte Chemie International Edition* **50**, 3124–3156 (2011).

[90] Gelbart, W. M., Bruinsma, R. F., Pincus, P. A. & Parsegian, V. A. Dna-inspired electrostatics. *Physics Today* **53**, 38–45 (2000).

[91] Bloomfield, V. A. DNA condensation by multivalent cations. *Biopolymers: Original Research on Biomolecules* **44**, 269–282 (1997).

[92] Davis, R. H., Morris, D. & Coffino, P. Sequestered end products and enzyme regulation: the case of ornithine decarboxylase. *Microbiology and Molecular Biology Reviews* **56**, 280–290 (1992).

[93] Yan, D.-H. *et al.* Different intracellular polyamine concentrations underlie the difference in the inward rectifier k+ currents in atria and ventricles of the guinea-pig heart. *The Journal of Physiology* **563**, 713–724 (2005).

[94] Watanabe, S.-i., Kusama-Eguchi, K., Kobayashi, H. & Igarashi, K. Estimation of polyamine binding to macromolecules and atp in bovine lymphocytes and rat liver. *Journal of Biological Chemistry* **266**, 20803–20809 (1991).

[95] Sigma-Aldrich, I. Productinformation spermidine (2015-04-23T11:04:01). URL `https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Product_Information_Sheet/1/s4139pis.pdf`.

[96] Yoo, J. & Aksimentiev, A. Molecular dynamics of membrane-spanning DNA channels: conductance mechanism, electro-osmotic transport, and mechanical gating. *The Journal of Physical Chemistry Letters* **6**, 4680–4687 (2015).

[97] Akimov, S. A. *et al.* Pore formation in lipid membrane i: Continuous reversible trajectory from intact bilayer through hydrophobic defect to transversal pore. *Scientific Reports* **7**, 1–20 (2017).

[98] Gopfrich, K. *et al.* Ion channels made from a single membrane-spanning DNA duplex. *Nano Letters* **16**, 4665–4669 (2016).

[99] Li, C.-Y. *et al.* Ionic conductivity, structural deformation, and programmable anisotropy of DNA origami in electric field. *ACS Nano* **9**, 1420–1433 (2015).

[100] Akimov, S. A. *et al.* Pore formation in lipid membrane ii: Energy landscape under external stress. *Scientific Reports* **7**, 1–20 (2017).

[101] Li, J. & Talaga, D. S. The distribution of DNA translocation times in solid-state nanopores. *Journal of Physics: Condensed Matter* **22**, 454129 (2010).

[102] Ling, D. Y. & Ling, X. S. On the distribution of DNA translocation times in solid-state nanopores: an analysis using schrödinger's first-passage-time theory. *Journal of Physics: Condensed Matter* **25**, 375102 (2013).

[103] Wu, H. *et al.* The estimation of field-dependent conductance change of nanopore by field-induced charge in the translocations of AuNPs-DNA conjugates. *The Journal of Physical Chemistry C* **118**, 26825–26835 (2014).

[104] Carson, S., Wilson, J., Aksimentiev, A. & Wanunu, M. Smooth DNA transport through a narrowed pore geometry. *Bioohysical Journal* **107**, 2381–2393 (2014).

[105] Carson, S., Wick, S. T., Carr, P. A., Wanunu, M. & Aguilar, C. A. Direct analysis of gene synthesis reactions using solid-state nanopores. *ACS Nano* **9**, 12417–12424 (2015).

[106] Pitchford, W. H. *et al.* Synchronized optical and electronic detection of biomolecules using a low noise nanopore platform. *ACS Nano* **9**, 1740–1748 (2015).

[107] Sze, J. Y., Kumar, S., Ivanov, A. P., Oh, S.-H. & Edel, J. B. Fine tuning of nanopipettes using atomic layer deposition for single molecule sensing. *Analyst* **140**, 4828–4834 (2015).

[108] Angeli, E. *et al.* Simultaneous electro-optical tracking for nanoparticle recognition and counting. *Nano Letters* **15**, 5696–5701 (2015).

[109] Carson, S., Wilson, J., Aksimentiev, A., Weigele, P. R. & Wanunu, M. Hydroxymethyluracil modifications enhance the flexibility and hydrophilicity of double-stranded DNA. *Nucleic Acid Research* **44**, 2085–2092 (2016).

[110] Hall, A. R. *et al.* Hybrid pore formation by directed insertion of α-haemolysin into solid-state nanopores. *Nature Nanotechnology* **5**, 874 (2010).

[111] Cox, D. & Miller, H. *The Theory of Stochastic Processes.* Wiley publications in statistics (Wiley, 1965).

[112] Ammenti, A. & Cecconi, F. Marconi umb and vulpiani a. *J. Phys. Chem. B* **2009**, 113 (2009).

[113] Kowalczyk, S. W., Grosberg, A. Y., Rabin, Y. & Dekker, C. Modeling the conductance and DNA blockade of solid-state nanopores. *Nanotechnology* **22**, 315101 (2011).

[114] Fleming, S. J., Lu, B. & Golovchenko, J. A. Charge, diffusion, and current fluctuations of single-stranded DNA trapped in an MspA nanopore. *Biophysical Journal* **112**, 368–375 (2017).

[115] Fuchs, M. *Charakterisierung und Translokation von Makromolekülen mittels künstlichen Nanoporen.* Master thesis, TU München (2016).

[116] Yeh, I.-C. & Hummer, G. Diffusion and electrophoretic mobility of single-stranded rna from molecular dynamics simulations. *Biophysical Journal* **86**, 681–689 (2004).

[117] Lukacs, G. L. *et al.* Size-dependent DNA mobility in cytoplasm and nucleus. *Journal of Biological Chemistry* **275**, 1625–1629 (2000).

[118] Nir, I., Huttner, D. & Meller, A. Direct sensing and discrimination among ubiquitin and ubiquitin chains using solid-state nanopores. *Biophysical Journal* **108**, 2340–2349 (2015).

[119] Grison, M. *Single-Molecule Cohesion and Adhesion in Muscle Cells.* Dissertation, TU München (2017).

[120] Cao, C., Liao, D.-F., Yu, J., Tian, H. & Long, Y.-T. Construction of an aerolysin nanopore in a lipid bilayer for single-oligonucleotide analysis. *Nature Protocols* **12**, 1901 (2017).

[121] Johnson, R. P., Fleming, A. M., Jin, Q., Burrows, C. J. & White, H. S. Temperature and electrolyte optimization of the α-hemolysin latch sensing zone for detection of base modification in double-stranded DNA. *Biophysical Journal* **107**, 924–931 (2014).

[122] Stoddart, D., Maglia, G., Mikhailova, E., Heron, A. J. & Bayley, H. Multiple base-recognition sites in a biological nanopore: Two heads are better than one. *Angewandte Chemie International Edition* **49**, 556–559 (2010).

[123] Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nature Biotechnology* **34**, 518 (2016).

[124] Butler, T. Z., Pavlenok, M., Derrington, I. M., Niederweis, M. & Gundlach, J. H. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proceedings of the National Academy of Sciences* **105**, 20647–20652 (2008).

[125] Goyal, P. *et al.* Structural and mechanistic insights into the bacterial amyloid secretion channel csgg. *Nature* **516**, 250–253 (2014).

[126] Guasch, A. *et al.* Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage φ29 connector particle. *Journal of Molecular Biology* **315**, 663–676 (2002).

[127] Cao, C. *et al.* Mapping the sensing spots of aerolysin for single oligonucleotides analysis. *Nature Communications* **9**, 1–9 (2018).

[128] Ouldali, H. *et al.* Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore. *Nature Biotechnology* **38**, 176–181 (2020).

[129] Jeong, K.-B. *et al.* Alpha-hederin nanopore for single nucleotide discrimination. *ACS Nano* **13**, 1719–1727 (2019).

[130] Laszlo, A. H., Derrington, I. M. & Gundlach, J. H. MspA nanopore as a single-molecule tool: From sequencing to sprnt. *Methods* **105**, 75–89 (2016).

[131] Brown, C. G. & Clarke, J. Nanopore development at oxford nanopore. *Nature Biotechnology* **34**, 810–811 (2016).

[132] McRuer, R. N. & Kokoris, M. S. Translocation control for sensing by a nanopore (2019). US Patent 10,457,979.

[133] Feng, J. *et al.* Identification of single nucleotides in mos 2 nanopores. *Nature Nanotechnology* **10**, 1070 (2015).

[134] Yang, H. *et al.* Identification of single nucleotides by a tiny charged solid-state nanopore. *The Journal of Physical Chemistry B* **122**, 7929–7935 (2018).

[135] Zhou, W., Qiu, H., Guo, Y. & Guo, W. Molecular insights into distinct detection properties of α-hemolysin, MspA, csgg and aerolysin nanopore sensors. *The Journal of Physical Chemistry B* (2020).

[136] Servick, K. Amid patent lawsuit, genetic sequencing upstart unveils new technology. *Science* (2016).

[137] Low, T. & Goldstein, A. L. The chemistry and biology of thymosin. ii. amino acid sequence analysis of thymosin alpha1 and polypeptide beta1. *Journal of Biological Chemistry* **254**, 987–995 (1979).

[138] Wilkinson, K., Urban, M. & Haas, A. Ubiquitin is the atp-dependent proteolysis factor i of rabbit reticulocytes. *Journal of Biological Chemistry* **255**, 7529–7532 (1980).

[139] Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics* **14**, 390 (2013).

[140] Chetverina, D., Aoki, T., Erokhin, M., Georgiev, P. & Schedl, P. Making connections: Insulators organize eukaryotic chromosomes into independent cis-regulatory networks. *Bioessays* **36**, 163–172 (2014).

[141] Kwon, D. *et al.* Enhancer-promoter communication at the drosophila engrailed locus. *Development* **136**, 3067–3075 (2009).

[142] Denker, A. & De Laat, W. The second decade of 3c technologies: detailed insights into nuclear organization. *Genes & Development* **30**, 1357–1382 (2016).

[143] Stadhouders, R. *et al.* Transcription regulation by distal enhancers: who's in the loop? *Transcription* **3**, 181–186 (2012).

[144] Van Steensel, B. & Dekker, J. Genomics tools for unraveling chromosome architecture. *Nature Biotechnology* **28**, 1089 (2010).

[145] Morgan, S. L. *et al.* Manipulation of nuclear architecture through crispr-mediated chromosomal looping. *Nature Communications* **8**, 15993 (2017).

[146] Petrascheck, M. *et al.* Dna looping induced by a transcriptional enhancer in vivo. *Nucleic Acid Research* **33**, 3743–3750 (2005).

[147] Ameres, S. L. *et al.* Inducible DNA-loop formation blocks transcriptional activation by an sv40 enhancer. *The EMBO journal* **24**, 358–367 (2005).

[148] Deng, W. *et al.* Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).

[149] Deng, W. *et al.* Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).

[150] Hao, N., Shearwin, K. E. & Dodd, I. B. Programmable DNA looping using engineered bivalent dcas9 complexes. *Nature Communications* **8**, 1628 (2017).

[151] Cho, S. W. *et al.* Analysis of off-target effects of crispr/cas-derived rna-guided endonucleases and nickases. *Genome Research* **24**, 132–141 (2014).

[152] Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A. & Charpentier, E. The crispr-associated DNA-cleaving enzyme cpf1 also processes precursor crispr rna. *Nature* **532**, 517–521 (2016).

[153] Zetsche, B. *et al.* Multiplex gene editing by crispr–cpf1 using a single crrna array. *Nature Biotechnology* **35**, 31 (2017).

[154] Linhananta, A., Hadizadeh, S. & Plotkin, S. S. An effective solvent theory connecting the underlying mechanisms of osmolytes and denaturants for protein stability. *Biophysical Journal* **100**, 459–468 (2011).

[155] Auton, M. & Bolen, D. W. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences* **102**, 15065–15068 (2005).

[156] Bolen, D. & Baskakov, I. V. The osmophobic effect: natural selection of a thermodynamic force in protein folding. *Journal of Molecular Biology* **310**, 955–963 (2001).

[157] Shandilya, M. Size exclusion chromatography (2013). URL https://www.slideshare.net/mshandilya1/size-exclusion-chromatography-28288632.

[158] YMC. Analysis and purification of biomolecules, such as nucleic acids, peptides, and proteins (2013). URL http://www.ymc.co.jp/en/bio/.

[159] Hong, P., Koza, S. & Bouvier, E. S. A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *Journal of Liquid Chromatography & Related Technologies* **35**, 2923–2950 (2012).

[160] Hagerman, P. J. Flexibility of DNA. *Annual Review of Biophysics and Biophysical Chemistry* **17**, 265–286 (1988).

[161] Ringrose, L., Chabanis, S., Angrand, P.-O., Woodroofe, C. & Stewart, A. F. Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *The EMBO Journal* **18**, 6630–6641 (1999).

[162] Chen, Y.-J., Johnson, S., Mulligan, P., Spakowitz, A. J. & Phillips, R. Modulation of DNA loop lifetimes by the free energy of loop formation. *Proceedings of the National Academy of Sciences* **111**, 17396–17401 (2014).

[163] Vologodskii, A. & D. Frank-Kamenetskii, M. Strong bending of the DNA double helix. *Nucleic Acid Research* **41**, 6785–6792 (2013).

[164] Shimada, J. & Yamakawa, H. Ring-closure probabilities for twisted wormlike chains. application to DNA. *Macromolecules* **17**, 689–698 (1984).

[165] Rippe, K., von Hippel, P. H. & Langowski, J. Action at a distance: DNA-looping and initiation of transcription. *Trends in Biochemical Sciences* **20**, 500–506 (1995).

[166] Site directed mutagenesis (2020). URL https://international.neb.com/applications/cloning-and-synthetic-biology/site-directed-mutagenesis.

[167] Chemical transformation tips (2020). URL https://international.neb.com/tools-and-resources/usage-guidelines/chemical-transformation-tips.

[168] Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E. & Deamer, D. W. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic

acid, and polyuridylic acid as homopolymers or as segments within single rna molecules. *Biophysical Journal* **77**, 3227–3233 (1999).

[169] Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences* **93**, 13770–13773 (1996).

[170] Kasianowicz, J., Kellermayer, M. & Deamer, D. *Structure and Dynamics of Confined Polymers: Proceedings of the NATO Advanced Research Workshop on Biological, Biophysical & Theoretical Aspects of Polymer Structure and Transport Bikal, Hungary 20–25 June 1999.* Nato Science Partnership Subseries: 3 (Springer Netherlands, 2012).

[171] Meller, A., Nivon, L., Brandin, E., Golovchenko, J. & Branton, D. Rapid nanopore discrimination between single polynucleotide molecules. *Proceedings of the National Academy of Sciences* **97**, 1079–1084 (2000).

[172] Meller, A., Nivon, L. & Branton, D. Voltage-driven DNA translocations through a nanopore. *Physical Review Letters* **86**, 3435 (2001).

[173] Wanunu, M. & Meller, A. *Single-Molecule Analysis of Nucleic Acids and DNA-protein Interactions Using Nanopores* (Cold Spring Harbor Laboratory Press, 2008).

[174] Dudko, O. K., Mathé, J. & Meller, A. Chapter twenty-one-nanopore force spectroscopy tools for analyzing single biomolecular complexes. *Methods in Enzymology* **475**, 565–589 (2010).

[175] Risken, H. *Fokker-Planck Equation* (Springer, 1984).

[176] Dudko, O., Filippov, A., Klafter, J. & Urbakh, M. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. *Proceedings of the National Academy of Sciences* **100**, 11378–11381 (2003).

[177] Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Physical Review Letters* **96**, 108101–108101 (2006).

[178] Bullerjahn, J. T., Sturm, S. & Kroy, K. Theory of rapid force spectroscopy. *Nature Communications* **5** (2014).

[179] Dudko, O., Mathé, J. & Meller, A. Nanopore force spectroscopy tools for analyzing single biomolecular complexes. *Methods in Enzymology* **475**, 565–89 (2010).

[180] Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences* **105**, 15755–15760 (2008).

[181] Schulten, K. Theoretical physics, 4. smoluchowski diffusion equation (1999). URL `https://www.ks.uiuc.edu/Services/Class/PHYS498/LectureNotes/chp4.pdf`.

[182] Hummer, G. & Szabo, A. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophysical Journal* **85**, 5–15 (2003).

[183] Garg, A. Escape-field distribution for escape from a metastable potential well subject to a steadily increasing bias field. *Physical Review B* **51**, 15592 (1995).

[184] Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Physical Review Letters* **96**, 108101 (2006).

[185] Mathé, J., Visram, H., Viasnoff, V., Rabin, Y. & Meller, A. Nanopore unzipping of individual DNA hairpin molecules. *Biophysical Journal* **87**, 3205–3212 (2004).

[186] Keyser, U. F. *et al.* Direct force measurements on DNA in a solid-state nanopore. *Nature Physics* **2**, 473–477 (2006).

[187] Mathé, J. The richness of the eye of a needle: Comment on "nanopores: A journey towards DNA sequencing" by m. wanunu. *Physics of Life Reviews* **9**, 159–160 (2012).

[188] Lin, J., Fabian, M., Sonenberg, N. & Meller, A. Nanopore detachment kinetics of poly(A) binding proteins from RNA molecules reveals the critical role of C-terminus interactions. *Biophysical Journal* **102**, 1427–34 (2012).

[189] Mathé, J., Aksimentiev, A., Nelson, D. R., Schulten, K. & Meller, A. Orientation discrimination of single-stranded DNA inside the α-hemolysin membrane channel. *Proceedings of the National Academy of Sciences* **102**, 12377–12382 (2005).

[190] Renner, S. *Kraftspektroskopie mit Nanoporen.* Dissertation, Technische Universität München (2011).

[191] Poly(A) Tailing Kit (Accessed: 2015-06-14). URL `https://www.lifetechnologies.com/order/catalog/product/AM1350`.

[192] Chang, A. L., McKeague, M., Liang, J. C. & Smolke, C. D. Kinetic and equilibrium binding characterization of aptamers to small molecules using a label-free, sensitive, and scalable platform. *Analytical Chemistry* **86**, 3273–3278 (2014).

[193] Schink, S. *et al.* Quantitative analysis of the nanopore translocation dynamics of simple structured polynucleotides. *Biophysical Journal* **102**, 85–95 (2012).

A. Nanopore Force Spectroscopy of an RNA Aptamer

Nanopore force spectroscopy (NFS), a single-molecule force spectroscopy method, allows to obtain information about binding kinetics of oligonucleotide-ligand complexes. Similar to other single-molecule force spectroscopy methods, like atomic force microscopy (AFM) and optical or magnetic tweezers, a force is exerted to examine the connection strength of the oligonucleotide to its ligand. In force spectroscopy that is performed with the AFM and tweezers, a mechanical force is exerted on the complex, whereas in NFS the rupturing is linked to an electrical force. The technique allows for the recording of many subsequently translocating molecules, rendering it into a high throughput method. [168–172].

Principle NFS Upon the application of a voltage across a pore which is incorporated in a bilayer the main voltage drop occurs around the pore, as the resistance of the pore and the electrolyte differ orders of magnitude. The electric field around the pore lumen exerts a force on a diffusing charged molecule. Molecules smaller than the constriction zone of the pore, directly translocate through the pore, the passage of K^+ and Cl^- trough aHl occurs in the range of nanoseconds. A compound of a size comparable to the constriction zone which enters the channel at standard ionic conditions [1], causes a drop in current, as the exclusion volume created by the entering molecules reduces the ionic current. Molecules larger than the constriction zone either leave the vestibule by thermal motion or they rupture and pass through the pore. Rupturing implies that the pore exerts enough shear force on the molecule to destabilize the molecular bonds to rupture. The rupturing time τ is force-dependent. A simple description of the force is

$$F = \int \frac{Q_{eff} \cdot E(x)}{l} dx = \frac{Q_{eff} \cdot V}{l} \tag{A1}$$

[1] Standard ionic conditions in NFS include a concentration of monovalent ions in the molar range

where l denotes the length of the pore, $E(x)$ the electric field, Q_{eff} the effective charge residing in the stem region of the pore [173] and V the drop of electrical voltage across the pore [174]. Chapter A.1 follows the argumentation of my master thesis [73] which is mainly based on the reasoning of Meller *et al.* [172]. In Chapter A.2 materials and methods are listed that were required to obtain the experimental results of examining the FMN aptamer with NFS (Chapter A.3).

A.1. Theory

Here, we focus on the analysis of FMN aptamer complexing with its ligand. We utilize the aHl nanopore for NFS. The aHl is suitable for the NFS of the FMN aptamer, since the secondary structure of the aptamer is wider than the aHl constriction zone. Hence the folded aptamer outsizes the pore, whereas the ruptured, then ssRNA, with a diameter of 1.1 nm, fits through the narrowest constriction of aHl.

Constant and Dynamic Force Experiments Both, the application of constant and the application of dynamic electrical force, cause mechanical tension on the molecule in the pore. Constant force experiments directly probe the voltage-dependent lifetime $\tau(V)$ of a given complex. Dynamic force experiments generate a distribution of rupture voltages $p(V)$ which indirectly offer information about the voltage-dependent lifetime $\tau(V)$ and the voltage of rupture.

Summary of Theory of Force Driven Molecular Rupture Early work evaluated constant force experiments by using Bell's formula [175]. Dudko *et al.* established a model [176,177] derived from Kramer's high barrier theory which takes influences of force onto the energy landscape into account. Bullerjahn *et al.* [178] provided a generalized theory for constant force experiments that overcame the shortcomings of the previous models.

Dudko *et al.* explored dynamic force spectroscopy experiments and built a model for it [179].

The crucial component for all force driven molecular rupture is the underlying energy landscape ΔG on which the molecule moves. The transition from closed to ruptured state demands for the crossing of an energy barrier. The barrier can be lowered by the application of an external force that simultaneously subjects the whole landscape to tilting.

A.1.1. Constant Force Experiments

Bell's equation offers a rough estimation for the rupturing time $\tau(F)$ of the molecule and the rate of rupture $k(F)$.

Bell's Formula The earliest description of the rupturing process was postulated by Bell before the first single-molecule experiments. It was "intrinsically a phenomenological description of the forced unbinding of adhesive contacts" [178]. The argument follows Van't Hoff-Arrhenius law, wherein the rate coefficient or the rate of escape k under zero force equals

$$k = a\, e^{-\beta E_{unbinding}} \tag{A2}$$

where β denotes the thermodynamic beta and a a constant factor. Most simplified k equals the number of openings per second, a the number of potentially possible openings per second, and $e^{-\beta E_{unbinding}}$ the probability of an opening event. The rupturing time at zero force τ_0 is given by the inverse of the rate coefficient

$$\tau_0 = \frac{1}{k} = \frac{1}{a}\, e^{\beta E_{unbinding}} \tag{A3}$$

Additional force increases the rate coefficient, respectively decreasing the rupturing time $\tau(F)$. The force reduces the energy landscape linearly by $-\beta F \Delta x$. The rupturing time hence decreases exponentially with increasing force

$$\tau(F) = \tau_0\, e^{-\beta\, F\, \Delta x_{trans}} \tag{A4}$$

where x_{trans} describes the distance between the equilibrium well and the transition state. Force lowers the energy barrier between the secondary structure of the aptamer and the unwound aptamer conformation. Van't Hoff-Arrhenius equation serves as the most simplistic description of the process which still credits the stochastic character of the rupturing process and the rupturing time by including the exponential dependence on the unbinding energy in its description. The amount of applied energy depends on the force exerted by the drop of potential between *cis* and *trans* side of the pore. The force can be replaced by A

$$\tau(V) = \tau_0\, \exp\left(-\beta\, \frac{Q_{eff} V}{l}\, \Delta x_{trans}\right). \tag{A5}$$

Bell'f formula serves well in describing experiments in the regime of low forces, however, it fails to describe the non-linearity of $\ln \tau$ vs F for higher forces. In Bell's formula, the force dependence of Δx_{trans} is not taken into consideration. The non-linearity is well described by Dudko *et al.* [180] and recapitulated in the following section.

Kramer's High Barrier Theory The probability current $j(x,t)$ across a high barrier is equivalent to the product of the rupturing rate k and the probability that the particle is close to the well P_{well}

$$j \equiv k(F) \cdot P_{well} = \frac{1}{\tau(F)} P_{well} \tag{A6}$$

Hence the force-dependent rupturing time equals

$$\tau(F) = \frac{P_{well}}{j}. \tag{A7}$$

The probability of a particle close to the well, around point A, can be expressed as

$$P_{well,A} = \int_{-a}^{a} p(x,t)\, dx \tag{A8}$$

The assumption of a high barrier ΔG in Kramer's high barrier theory implies that the distribution function near $x = 0$ is approximately a stationary distribution

$$p(x,t) = p(0,t) e^{-\beta U(x)} \tag{A9}$$

hence

$$P_{well,A} = \int_{-a}^{a} p(0,t) e^{-\beta U(x)} \tag{A10}$$

For the description of the force-dependent rupture time, a description of the probability current $j(r,t)$ is required additionally. It is obtained through the diffusion equation

$$\partial_t\, p(r,t) = -\nabla\, j(r,t) \tag{A11}$$

and the Langevin equation. The overdamped Langevin equation

$$0 = m\ddot{x}(t) = F(x) + \gamma\dot{x} + \eta(t) \tag{A12}$$

describes the system. Here $F(x)$ denotes an external force, γ the drag or friction coefficient, $\eta(t)$ the fluctuating or random force [181], a stochastic variable, describing the effect of the random force of Brownian motion. The Fokker Planck equation offers a description equivalent to the Langevin equation. It describes the time evolution of the probability density function of the *velocity* and *location* of

the particle. In the overdamped system, the Fokker Planck equation directly reduces to the Smoluchowski equation: with D as isotropic diffusion coefficient. As velocity remains constant in the high friction limit, the Smoluchowski equation describes the probability density function of the *location* of a Brownian particle. Reducing the Smoluchowski equation to one dimension and substituting the force by $F(x) = -\partial_x U(x)$ results in a differential equation for $j(x,t)$

$$j(x,t) = -D\ (\partial_x + \partial_x \beta U(x))\ p(x,t) \tag{A13}$$

$$\Rightarrow \quad j(x,t) = -D\, e^{-\beta U(x(t))}\, \partial_x e^{\beta U(x(t))}\, p(x,t). \tag{A14}$$

The assumption of a high barrier separating the ruptured from the non-ruptured state implies that the flux over the potential barrier is small. Hence the temporal change of the probability density $p(x,t)$ is small, which results in a quasi-stationary state. There the probability current is nearly independent of x. Integrating equation A.1.1 from 0 to B results in

$$j \int_0^B e^{\beta U(x)} dx = -D \left(e^{\beta U(b)} p(b,t) - e^{\beta U(0)} p(0,t) \right) \tag{A15}$$

With $e^{\beta U(B)} \approx 0$ and $U(0) := 0$ the expression of j reduces to

$$j = D\, p(0,t) \left(\int_0^b e^{\beta U(x)} dx \right)^{-1} \tag{A16}$$

Substituting P_{well} and j in $\tau(F)$ the force-dependent first mean passage time under the assumption of a high barrier is

$$\tau(F) = \frac{P_{well,A}}{j} \tag{A17}$$

$$= \int_{-a}^{a} p(0,t) e^{-\beta U(x)}\, \frac{\int_0^b e^{\beta U(x)} dx}{D\, p(0,t)} \tag{A18}$$

$$= \frac{1}{D} \int_{barrier} e^{\beta U(x)} dx \int_{well} e^{-\beta U(x)} dx \tag{A19}$$

Based on Kramer's high barrier theory Dudko *et al.* could obtain analytic solutions for different energy landscapes for the force-dependent energy. A specific energy landscape has to be assumed to acquire an analytic term for $\tau(F)$. The energy landscape can be described in approximation either by a harmonic potential or a cusp-like potential. The harmonic potential contains linear and cubic terms, like $U_0(x) = \Delta G(\frac{x}{x_{trans}})^2(3 - \frac{2x}{x_{trans}})$ [182, 183]. The cusp-like potential is either a

potential with a harmonic well with a cusp-like barrier or a cusp-like well with a harmonic barrier as described in [182]. Here an example is $U_0(x) = \Delta G \cdot (\frac{x}{x_{trans}})^2$ for $x < x_{trans}$ and $G_0(x) = \infty$ for $x \geq x_{trans}$ [184]. Inserting the potentials into the lifetime $\tau(F)$ derived by Kramer's high energy barrier approximation led Dudko *et al.* [184] to

$$\tau(F) = \tau_0 \left(1 - \frac{\nu F \Delta x}{\Delta G}\right)^{1-1/\nu} \exp\left(-\beta \Delta G[1 - (1 - \frac{\nu F \Delta x}{\Delta G})^{1-1/\nu}]\right) \qquad \text{(A20)}$$

where $\nu = \frac{2}{3}$ applies for a linear cubic surface and $\nu = \frac{1}{2}$ for a linear harmonic-cusp surface. Limitations for the applicability of the equation occur as only a limited range of forces - $0 \leq F \leq \frac{\Delta G}{\nu x_{trans}}$ - can be inserted. Additionally, the expression for $k(F)$ gets invalid for forces approaching $F_C = \frac{\Delta G}{\nu x_{trans}}$. The rupturing rate $k(F)$ decreases down to a final value of $k(F_c) = 0$, however, it should increase by the application of force. These limitations can be traced back to Kramer's *high barrier* theory. With increasing force, the assumption of a high barrier loses its validity. Dudko *et al.* pointed out that these false predictions in the high energy regime could be rectified by the usage of the theory of mean first passage time. Nevertheless, the analytic expression for $\tau(F)$ by Dudko *et al.* is one of the most substantial tools for acquiring information about rupturing times.

The Langevin description allows a more accurate solution for $\tau(F)$ than Bell's formula, as the force dependence of Δx_{trans} is taken into account

$$\tau(F) = \tau_0 \exp\left(-\beta \int_0^F < x(F')_{trans} > dF'\right). \qquad \text{(A21)}$$

Mean First Passage Time The limited predictions in the high force regime can be avoided by numerically solving the mean first passage time (MFPT) equation

$$\tau(F) = \frac{1}{D} \int_0^{x_{trans}} e^{\beta U(x)} \int_{-\infty}^{x} e^{-\beta U(x)} dy dx. \qquad \text{(A22)}$$

It does not require the assumption of a high barrier, its validity is therefore not limited to high barriers or low forces.

A.1.2. Dynamic Force Experiments

In dynamic force experiments, a whole range of voltages that can rupture the molecule is screened. The approach allows for a more efficient data acquisition.

Theory described in the following section assists in extracting information of the measured voltage-dependent probability distributions $p(V)$. It additionally hands a mean to obtain information about the voltage-dependent $\tau(V)$.

Theory for the force-dependent rupturing probabilities and times, $p(F)$ and $\tau(F)$ help obtain models for $p(V)$ and $\tau(V)$. For a rupturing process, the number of ruptures $\dot{S}(t)$ at time t equals the rupturing rate $k(t)$ times the number of non-perturbed, intact molecules $S(t)$

$$\dot{S} = -S(t) \cdot k(F(t)) = -\frac{S(t)}{\tau(F(t))}. \tag{A23}$$

or the number of intact molecules over the rupturing time $\tau(F(t))$. Hence the probability $S(t)$ that the system is still intact at time t equals the total number of molecules breaking open $\dot{S}$ times the rupturing time $\tau(F(t))$:

$$S(t) = -\dot{S}\,\tau(F(t)) \tag{A24}$$

The number of intact molecules at force $F(t)$ is equivalent to the integral over the rupturing probability $p(f)$ from F to infinity

$$S(t) = 1 - \int_0^F p(f)\,df = \int_F^\infty p(f)\,df \tag{A25}$$

Hence the rupturing time $\tau(F)$ in the dynamic force model equals

$$\tau(F) = \frac{S(t)}{\dot{F}\,p(F)} = \frac{\int_F^\infty p(f)\,df}{\dot{F}\,p(F)} \tag{A26}$$

A description of $p(F)$ is offered by reformulating equation A.1.2

$$p(F)\,dF = \dot{S}dt \tag{A27}$$

$$\Rightarrow \quad p(F) = -\frac{\dot{S}}{\dot{F}} \overset{(A.1.2)}{=} \frac{S(t)}{\dot{F}\,\tau(F)} \tag{A28}$$

We replace $S(t)$ by solving the differential equation A.1.2

$$\int_1^S \frac{dS'}{S'} = \int_0^F -\frac{1}{\tau(f)}\frac{1}{\dot{F}}\,df \tag{A29}$$

$$\Rightarrow \quad S(t) = \exp\left(-\int_0^F \frac{1}{\tau(f)\,\dot{F}(f)} df\right) \tag{A30}$$

Substituting $S(t)$ in equation A.1.2 defines $p(F)$ as

$$p(F) = \frac{\exp\left(-\int_0^F \frac{1}{\tau(f)\,\dot{F}(f)} df\right)}{\dot{F}\tau(F)} \tag{A31}$$

In NFS experiments the applied voltage is measured, therefore force dependence is replaced by voltage-dependence using

$$\beta \cdot F \cdot x = a \cdot x_\beta \cdot V \tag{A32}$$

where x_β denotes the width of the potential well, a a coupling parameter, and V the unzipping voltage of the RNA aptamer. It results in a voltage probability distribution:

$$p(V) = \frac{\exp(-\int_0^V \frac{1}{\dot{V}(v)\tau(v)} dv)}{\dot{V}\tau(V)}. \tag{A33}$$

Equation A.1.1 can be inserted and for $\nu = 1$ the probability distribution is [185]

$$p(V) = \frac{1}{\tau_0 \dot{V}} \exp\left(\frac{V}{V_\beta} - \frac{V_\beta}{\tau_0 \dot{V}}\left(\exp\left(\frac{V}{V_\beta}\right) - 1\right)\right). \tag{A34}$$

where $V_\beta = \frac{1}{\beta \cdot a \cdot x_\beta}$.

NFS experiments yield $p(V)$ and regression analysis of $p(V)$ offers information about τ_0 and V_β. For $\nu = 1$, $a \cdot x_\beta$ equals the effective charge. The charge residing in the stem region is described by $Q_{eff} = \frac{k_B T}{V_\beta}$.

Knowledge about the effective charge is required to quantify the applied force. In solid-state pores the effective charge comprises approximately 25 % of the actual amount of charge residing in the constriction zone, independent of the surrounding salt conditions [186].
Combining knowledge about translocations through aHl and AFM measurements, the conclusion was reached that the effective charge is reduced down to approximately 20 % of the actual charge being present [187].

A.1.3. Translocation Dynamics and Geometrical Properties

The direction of the translocating DNA and its sequence composition can influence measured translocation times and current blockage levels [188]. Directional influ-

ences, like entering the pore with the 3' or 5' end or entering from the *cis* or the *trans* side, feed into the resulting current reductions. Unidirectional translocation experiments exhibit subpopulations that were attributed to the primary entry of the 3' and 5' end of DNA. Molecular dynamics studies provided further proof and more detailed insight. The entries with the 3' end interact more strongly with the channel and introduce a more enhanced and longer current blockage [189]. Differences in *cis* versus *trans* translocation are attributed to a stabilizing effect of the vestibule of the pore on the structure of the oligonucleotide [190].

Catching rates from the *trans* direction are low. A method to circumvent the stabilizing effect of the vestibule is the anchor technique. The anchor technique relies on an anchor at the 3' end of the oligonucleotide. The 5' end of the molecule is caught on the *cis* side and threaded through the pore to the *trans* side. The anchor at the 3' end prevents the escape of the molecule. Upon the application of a reverse voltage, the DNA molecule can be ruptured while entering from the *trans* side and the stabilizing effect of the vestibule can be circumvented.

Here, another approach for *trans* side rupture was implemented. Molecules here were directly caught in *trans* direction. To allow for translocation from the *trans* side an unstructured, loose poly(A) tail is added to the RNA. It is supposed to support the initial threading of the RNA structure into the pore. Poly(A) offers the advantage of exhibiting minimal interaction with the RNA tertiary structure. This approach offers a high throughput characterization method and circumvents the stabilizing effect of the vestibule.

The translocation of molecules from the vestibule facing side, the *cis* side, to the stem facing side, the *trans* side is defined as forward translocations. Backward translocations are defined as passages of molecules from the *trans* to the *cis* side.

A.2. Materials and Methods

The required materials comprise the DIB setup, the poly(A) tailed aptamer, and its ligand.

A.2.1. DIB Setup

The DIB was used as previously described in Chapter 2.2.4. Instead of synthetic DNA pores, here aHl pores were used at 0.5 mg/ml in 1 M KCl. A small amount of solution, in the range of several microliters, was added close to the bilayer. Upon incorporation of an aHl pore the bilayer was moved away from the locally high concentration of aHl. Thereby incorporation of a second pore was rendered improbable.

A.2.2. RNA Aptamer and Poly(A) Tailing

Transcription was performed as previously described in Chapter 3.2.1. Two units of DNase I were added for 30 min. The mixture was phenol chloroform purified. 100 µl of sample were mixed thoroughly with 100 µl Chloroform in a PhaseLock Gel (PLG Light, VWR). After centrifugation (15,000 rcf for 4 min) 100 µl of Chloroform were added, mixed thoroughly, and spun for 4 min at 15,000 rcf. The supernatant is transferred in a new PhaseLock Gel, 100 µl of phenol were added, mixed thoroughly and spun for 4 min at 15,000 rcf. The supernatant was mixed with 300 ul ethanol, 100 % and stored at −80 °C for 1 h. The precipitate was centrifuged at 20,000 rcf for 15 min at 4 °C, the pellet was washed with 70 % ethanol, dried, and re-suspended in nuclease-free water. *E. coli* poly(A) tailing polymerase post-transcriptionally added a poly(A) tail to the RNA.

EPAP The RNA was poly(A) tailed post-transcriptionally by *E-coli* poly(A) tailing polymerase (EPAP). The length of the poly(A) tail is varying, depending on enzyme concentration and incubation time [191]. Transcribed RNA was incubated with 1 mM ATP, 2.5 mM $MnCl_2$, 1x EPAP buffer and 0.08 u/µl EPAP (Thermo Fischer Scientific) at 37 °C.

A.2.3. Electrical Recordings

Electrical recordings were performed with a HEKA EPC 9 double amplifier. Pores were incorporated in the bilayer of the DIB, so that molecules translocate from trans to cis side. An applied constant voltage caused the capture of poly(A)-tailed RNA molecules at the trans side. Upon catching, a voltage ramp of 1 V/s was triggered and the current response was recorded. The voltage where a steep current increase occurs is identified as the rupturing voltage by a MatLab routine.

A.3. Results and Discussion

The FMN aptamer was translocated through aHl in the absence and presence of FMN.

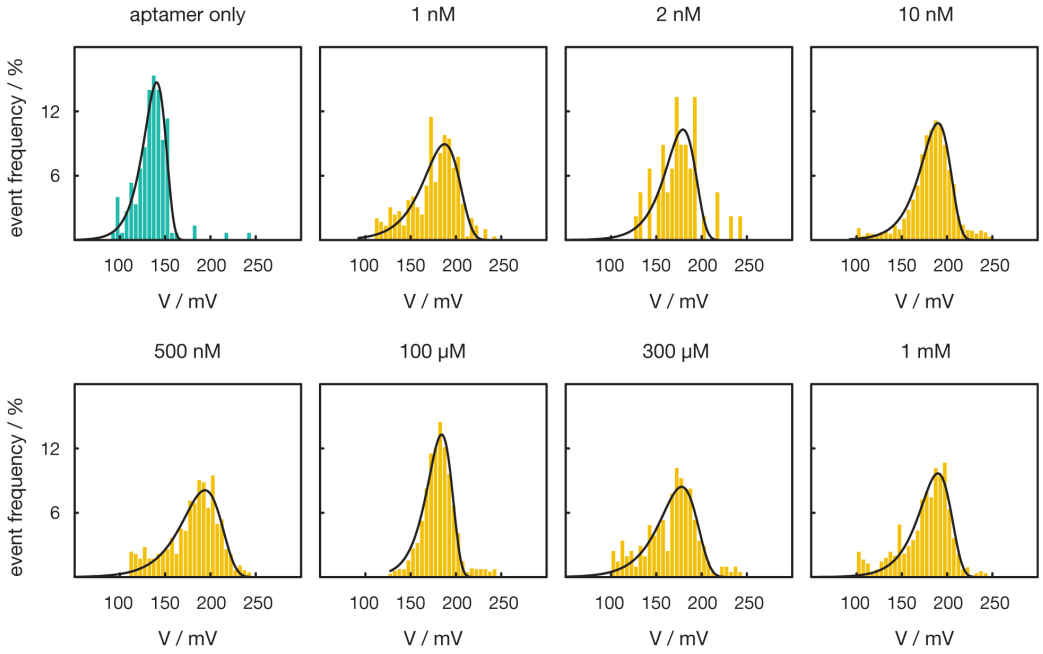

Figure A1. a, Probability distribution of rupturing voltages p(V) of the FMN aptamer in the absence and presence of FMN at various concentrations of FMN.

The probability distribution of rupturing voltages of the FMN aptamer in the absence and presence of FMN are displayed in Figure A1. Several FMN concentration values were examined. Regression analysis was performed with Equation A.1.2. The corresponding maxima of regression analysis are listed in Table A1. The shift of the maxima of the probability distributions equals a shift towards an increased rupturing voltage in presence of FMN. In Figure A1 the population which emerges around 90 mV was discarded, since it remained the same for all voltages.

c (FMN)	0 nM	1 nM	2 nM	10 nM	500 nM	100 µM	300 µM	1 mM
Max / mV	140, 145	187	179	190	194	183	178	190

Table A1. Table displays the maxima of the regression analysis for the distribution of rupturing voltages at several FMN concentration values

The probability distributions were converted into voltage-dependent lifetimes for

$\nu = 1$. The distributions are depicted in a logarithmic scale in Figure A2.

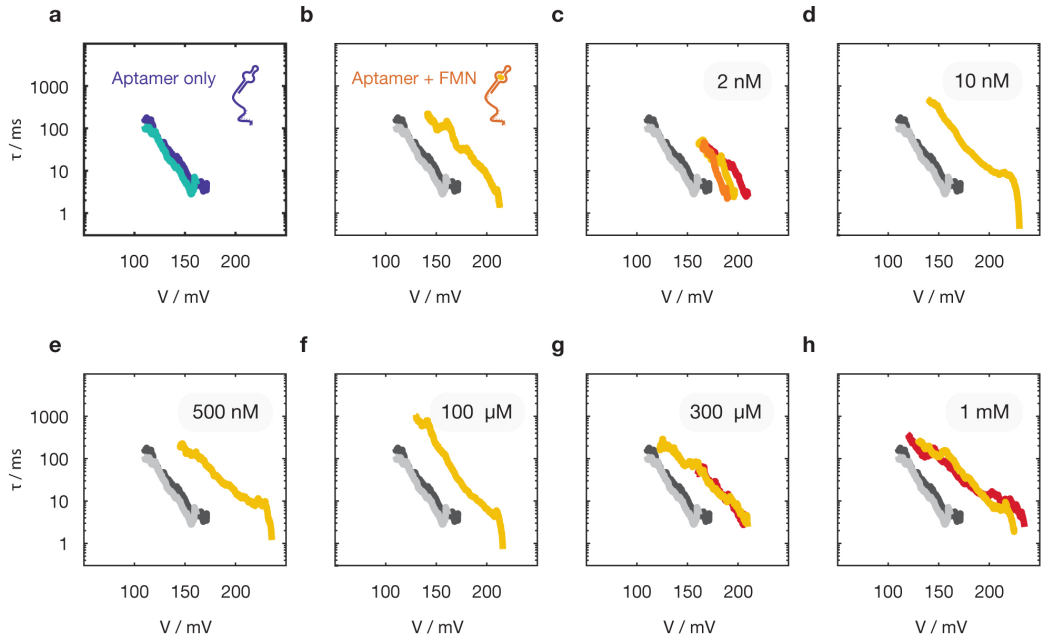

Figure A2. Voltage-dependent lifetimes $\tau(V)$ for a, FMN aptamer only and FMN aptamer in presence of a, 1 nM, b, 2 nM, c, 10 nM, d, 500 nM, e, 100 µM, f, 300 µM, g, 1 mM FMN

Figure A2 a shows two recordings of the voltage-dependent lifetime of the FMN aptamer in the absence of FMN. The two recordings of the aptamer in the absence of FMN overlap strongly. Figure A2b to h show the voltage-dependent lifetimes of the aptamer plus FMN (1 nM, 2 nM, 10 nM, 500 nM, 100 µM, 300 µM, 1 mM). In each graph, the data of aptamer without FMN is additionally depicted in greyscale. Compared to the aptamer in the absence of FMN, the rupturing time of the aptamer in presence of FMN is shifted towards higher voltages for all FMN concentration values.

Discussion The probability distributions of the rupturing voltages deviate from a classical two-population behavior. For increasing FMN concentration, a shift from an aptamer population towards an aptamer-FMN population would be the expected result. In this setup, however only singular populations were observed. All populations in presence of FMN are more stable than the populations in absence of FMN.

The implications of the voltage-dependent lifetimes are in accordance with the

implications of the probability distributions of the rupturing voltages. The shift towards longer rupturing times of aptamers in presence of FMN shows that the aptamers in presence of FMN are more stable than the aptamers in the absence of FMN. However, the stabilities of the aptamers in presence of FMN are similar for all concentrations of FMN. The voltage-dependent lifetimes do not exhibit a clear trend with increasing FMN concentration. The data suggests that from 1 nM on the entire population is shifted to a stable conformation. However, this observation does not align with literature values, where the dissociation constant is stated to be of around 500 to 700 nM [18, 192].

When Kramer's high barrier theory is applied for either $\nu = \frac{1}{2}, \frac{2}{3}$ or 1, all three energy landscapes yield infinite energy barriers. It indicates that those models might not be applicable to our data. A potential reason for this might be that the rupturing voltages we obtain here are relatively high in comparison to prior studies [185, 193].

B. DNA Sequences

B.1. DNA Pores

B.1.1. Examined Molecules and Conditions

Double stranded DNA sequence dsDNA, 527 bp: 5′ - ATAATACGACTCACT ATAGGGCGCCGCGAAAACGCGGCGCCCAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAGGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATC CAAAAA - 3 ′

Single and double stranded DNA sequence ssDNA, 140 bases: 5′ - (T)68-CCGG-(T)68-3′; sequence dsDNA, 115 bp: 5′ - ATAATACGACTCACTATAGGGCGC CGCGAAAACGCGGCGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA GGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATCCAAAAA - 3 ′

sgRNA sgRNA sequence: 5′ - CCAUCCAGUUCAACCAGGAUGUUUUAGAG CUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUUUUUUU - 3 ′

RNA aptamer FMN aptamer sequence: 5′ - GGCGUGUAGGAUAUGCUUCG-GCAGAAGGACACGCC - 3 ′

B.2. CRISPR - Cas12a Double Mutant

B.2.1. H800A D908A Cas12a

pddCas12a: 5′ -CCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCAT TAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCG CAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCA

CGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGA
AGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGC
CAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGG
CGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTAT
AGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTT
TAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATC
ATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCTGGA
GCCATCCGCAGTTTGAAAAAGGTGGTGGTAGCGGTGGTGGTTCAGGT
GGTAGTGCATGGTCACACCCTCAGTTTGAGAAAATGTCGGACTCAGAA
GTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAG
ACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTC
AAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCT
AAAAGACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGT
ATTAGAATCCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGAT
AACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGATCCATGACA
CAGTTTGAAGGCTTTACCAATCTGTATCAGGTGAGCAAAACGTTACGC
TTTGAACTGATTCCTCAGGGCAAAACCCTGAAACATATCCAGGAACAG
GGTTTTATCGAAGAAGATAAAGCCCGCAATGATCATTATAAGGAACTG
AAACCTATCATCGATCGCATCTATAAGACGTATGCGGATCAGTGTCTT
CAGCTGGTGCAGTTAGATTGGGAAAATCTGAGTGCGGCTATCGATAG
CTATCGCAAAGAAAAAACCGAAGAAACACGCAATGCCCTGATCGAAGA
ACAGGCCACCTATCGCAATGCTATTCATGATTATTTTATTGGTCGTAC
GGATAATCTCACGGATGCCATCAACAAACGTCATGCCGAAATCTACAA
AGGCCTGTTTAAAGCGGAACTGTTTAATGGTAAAGTGCTGAAACAGTT
AGGTACGGTGACGACGACCGAACATGAAAATGCCTTACTGCGTAGCTT
TGATAAGTTCACGACATATTTTAGCGGCTTTTATGAAAATCGCAAAAA
TGTGTTTAGTGCAGAAGATATTTCTACGGCCATTCCACATCGCATTGT
GCAGGATAATTTTCCTAAATTCAAAGAAAATTGTCATATCTTTACACG
CCTGATCACCGCCGTTCCTAGCTTACGCGAACATTTTGAAAATGTTAA
AAAAGCCATTGGCATCTTTGTGAGTACAAGCATTGAAGAAGTGTTTTC
TTTTCCGTTTTATAATCAGTTACTGACACAGACACAGATTGATCTGTA
TAATCAGCTGCTGGGTGGCATCTCACGCGAAGCCGGCACCGAAAAAAT
CAAAGGCCTGAATGAGGTGCTGAACCTTGCCATCCAGAAAAATGATGA
AACGGCCCATATTATTGCCAGCTTACCTCATCGCTTTATTCCTCTGTT
TAAACAGATCCTGAGCGATCGTAATACCCTGTCTTTTATTCTGGAAGA
GTTTAAATCAGATGAAGAAGTGATCCAGAGCTTTTGCAAATACAAGAC
CTTACTGCGCAATGAGAATGTGCTGGAAACCGCCGAAGCACTGTTTAA

TGAGCTGAATAGCATCGACCTCACCCATATCTTTATTAGTCATAAAAA
ACTGGAAACAATTAGTAGCGCCCTGTGCGATCATTGGGATACCTTACG
AAATGCACTGTATGAACGTCGCATTAGCGAACTGACCGGCAAAATTAC
CAAATCAGCCAAAGAAAAAGTGCAGCGCTCTCTGAAACATGAGGACAT
TAATCTGCAAGAAATTATTAGCGCCGCGGGCAAAGAACTGAGCGAAGC
CTTTAAACAGAAAACGAGCGAAATTCTGTCACATGCACATGCAGCACT
GGATCAGCCGTTACCTACAACGCTGAAAAAACAGGAAGAAAAAGAAAT
CCTGAAATCACAGTTAGATAGTCTGCTGGGCCTGTATCATCTGCTGGA
TTGGTTTGCGGTGGATGAATCTAACGAAGTAGATCCGGAGTTTAGTG
CCCGCCTGACGGGCATCAAACTGGAGATGGAACCTTCACTGAGCTTTT
ATAACAAGGCCCGTAATTATGCCACCAAAAAACCATATAGCGTGGAAA
AGTTCAAACTGAATTTTCAGATGCCGACGCTGGCGAGCGGTTGGGATG
TTAACAAAGAGAAGAATAACGGCGCCATTCTGTTTGTTAAAAATGGCC
TGTATTATCTCGGCATTATGCCGAAACAGAAAGGTCGCTACAAAGCGC
TGAGCTTTGAACCGACCGAAAAAACCTCAGAAGGCTTTGATAAAATGT
ATTATGATTATTTTCCGGATGCCGCCAAAATGATCCCGAAATGCTCAA
CCCAGCTGAAAGCCGTGACCGCCCATTTTCAGACACATACGACCCCGA
TCCTGCTGTCTAATAACTTTATCGAACCACTGGAAATTACCAAAGAAA
TCTATGATCTGAATAACCCGGAAAAAGAACCCAAAAAGTTTCAGACCG
CCTATGCCAAAAAGACGGGCGATCAGAAAGGCTATCGCGAAGCCCTGT
GCAAGTGGATCGATTTTACACGCGATTTTCTGTCTAAATATACCAAAA
CCACGAGCATCGATCTGTCAAGTCTGCGTCCGAGTAGCCAGTACAAGG
ATTTGGGCGAATATTATGCGGAACTGAATCCACTGCTGTATCATATTA
GCTTTCAGCGCATTGCCGAAAAAGAAATTATGGATGCGGTGGAAACCG
GTAAACTGTATCTGTTTCAAATATACAACAAAGATTTTGCTAAAGGTC
ATCATGGCAAACCTAACTTACATACCTTATATTGGACGGGCCTGTTTT
CTCCGGAAAACCTTGCTAAAACGAGCATCAAACTGAATGGTCAGGCCG
AACTGTTTTATCGTCCTAAATCTCGTATGAAACGTATGGCCCATCGTC
TGGGCGAAAAAATGTTAAACAAGAAACTGAAAGATCAGAAAACCCCTA
TCCCGGATACCCTGTATCAGGAACTGTATGATTATGTTAATCATCGCC
TGTCACATGACCTCAGCGATGAAGCTCGCGCCTTACTGCCTAATGTGA
TCACCAAAGAAGTGAGTCATGAAATCATCAAAGATCGTCGCTTTACGA
GCGATAAATTTTTCTTTCATGTTCCTATCACCCTGAATTATCAGGCCG
CTAATAGTCCTAGCAAGTTTAATCAGCGCGTTAATGCGTATCTGAAAG
AACATCCGGAAACCCCTATTATTGGCATCGATCGCGGCGAACGCAACC
TGATCTATATTACCGTGATCGATAGTACGGGAAAGATTCTGGAACAGC
GCTCACTGAATACCATCCAGCAGTTTGATTATCAGAAAAAACTGGATA

ATCGCGAAAAAGAGAGGGTTGCGGCACGCCAGGCGTGGAGTGTGGTT
GGTACAATCAAGGATCTGAAACAGGGCTATCTGTCACAAGTCATTCAT
GAAATTGTGGATCTGATGATTCATTATCAGGCGGTGGTTGTGTTAGAA
AATCTGAATTTTGGCTTTAAATCTAAACGTACGGGCATTGCGGAAAAA
GCCGTGTATCAGCAGTTTGAAAAAATGTTAATCGATAAACTGAATTGT
CTGGTTCTGAAAGATTATCCAGCGGAAAAAGTTGGTGGCGTGCTGAAT
CCGTATCAGTTAACGGATCAGTTTACGTCGTTTGCCAAAATGGGTACG
CAGAGTGGCTTTCTGTTTTATGTTCCGGCCCCGTATACAAGCAAAATC
GATCCGCTGACCGGCTTTGTAGACCCGTTTGTGTGGAAAACAATCAAA
AATCATGAATCTCGCAAACATTTTCTGGAAGGCTTTGATTTTCTGCAT
TATGATGTTAAAACGGGCGATTTTATCTTACATTTTAAGATGAATCGT
AATCTTAGTTTTCAGCGTGGCTTACCGGGTTTTATGCCGGCTTGGGAT
ATTGTGTTTGAAAAGAACGAAACCCAGTTTGATGCCAAAGGTACACCG
TTTATTGCAGGTAAACGCATTGTTCCAGTGATCGAAAATCATCGCTTT
ACGGGTCGTTATCGCGATCTGTATCCAGCCAATGAGTTAATTGCTCTG
TTAGAAGAAAAAGGTATTGTGTTTCGTGATGGTAGTAATATCTTACCT
AAACTGCTGGAAAATGATGATAGTCATGCCATCGATACAATGGTTGCC
CTGATTCGTAGCGTGCTCCAAATGCGCAATAGTAATGCGGCTACGGGC
GAAGATTATATCAATAGTCCGGTTCGTGATCTGAATGGTGTGTGCTTT
GATAGTCGCTTTCAGAATCCGGAATGGCCGATGGATGCCGATGCCAAT
GGCGCCTATCATATTGCGCTGAAAGGTCAGTTACTGCTGAATCATCTG
AAAGAATCTAAGGATCTGAAACTTCAAAATGGCATCTCTAATCAGGAT
TGGTTAGCATATATTCAGGAACTGCGCAATGATTATAAGGATCATGAT
GGCGATTACAAGGATCATGATATAGATTACAAAGATGATGATGATAAA
TAAGCGGCCGCACTCGAGGCCCGAAAGGAAGCTGAGTTGGCTGCTGC
CACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGT
CTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCCGC
AAGAGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCC
AGGGTGACGGTGCCGAGGATGACGATGAGCGCATTGTTAGATTTCAT
ACACGGTGCCTGACTGCGTTAGCAATTTAACTGTGATAAACTACCGCA
TTAAAGCTTATCGATGATAAGCTGTCAAACATGAGAAACACCATCGAA
TGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAG
TCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGC
AGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCA
GGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGA
TGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGG
GCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC

ACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAA
GCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGG
CTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAA
GCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAG
ACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTG
GGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTA
GCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGG
CATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAA
GGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTG
AATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATG
GCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGT
GCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGT
TATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGG
CAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTG
AAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACC
CTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCA
TTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTG
AGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACCGGGAT
CTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGT
GGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTA
TCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCG
GCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTG
CGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTC
CCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGG
CGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGCT
GGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGA
TGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATC
AGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGA
TCACTGGACCGCTGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCA
CATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACCTTGTCT
GCCTCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCT
GAATGGAAGCCGGCGGCACCTCGCTAACGGATTCACCACTCCAAGAAT
TGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCC
TTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCG
GCGCATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGATCGT
GCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGG

TTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTG
CTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTT
CCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCAT
TATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAAC
ACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTT
TCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACG
TTCCAGTAACCGGGCATGTTCATCATCAGTAACCCGTATCGTGAGCAT
CCTCTCTCGTTTCATCGGTATCATTACCCCCATGAACAGAAATCCCCC
TTACACGGAGGCATCAGTGACCAAACAGGAAAAAACCGCCCTTAACAT
GGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAACTCAA
CGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGA
CCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGA
CGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTG
TCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAG
CGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGC
GATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTG
TACTGAGAGTGCACCATATATGCGGTGTGAAATACCGCACAGATGCGT
AAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGA
CTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC
AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA
AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT
CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA
TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG
GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC
CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCA
GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAA
TTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTG

GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT
CTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCA
GCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC
GCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGT
GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA
ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGT
TAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGT
GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC
AACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCAT
CATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCT
GTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTC
AGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA
TACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTT
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCT
AAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCA
CGAGGCCCTTTCGTCTTCAAGAA - 3 ′

Primer sequences:

pAsCpf1 seq FP 1: 5′ - CGCTAAAAGACAGGGTAAG - 3 ′
pAsCpf1 seq FP 3: 5′ - GAGGTGCTGAACCTTGCC - 3 ′
pAsCpf1 seq FP 2: 5′ - GATCAGAAAGGCTATCGC - 3 ′
pAsCpf1 seq FP 4: 5′ - GAAGCCCTGTGCAAGTGG - 3 ′
pAsCpf1 seq RP 1: 5′ - GTTCCGCATAATATTCGCC - 3 ′
pAsCpf1 seq RP 2: 5′ - GTTATTGCTCAGCGGTGG - 3 ′

Target positions (AFM):

7 for dCas9: 5′ - CCCGTTCAGCCCGACCGCTG - 3 ′
8 for ddCas12a: 5′ - GGTACACCGTTTATTGCAGG - 3 ′

pSB3A1-T7wt-mVenus: 5′ -GCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCACA
GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGAT
ACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT
CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCG
GTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC
TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA
AAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGG
ATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATC
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA
TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCT
CCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA
AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGC
CGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTT
GTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG
GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC
CCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCA
CTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTG
ACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA
CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACAT
AGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATATAACCC
ACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTT

TCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC
CCGAAAAGTGCCACCTGACGTCTAAGAATAATACGACTCACTATAGGG
TCTATACTAGAGAAAGAGGAGAAATACTAGATGAGCAAAGGCGAAGA
ACTGTTCACGGGTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGT
GAACGGTCATAAATTTAGCGTGTCTGGTGAAGGCGAAGGTGATGCGA
CCTACGGCAAACTGACGCTGAAACTGATTTGCACCACGGGTAAACTGC
CGGTTCCGTGGCCGACCCTGGTGACCACGCTGGGTTATGGTCTGATGT
GTTTCGCACGTTACCCGGATCACATGAAACGCCATGATTTCTTTAAAT
CTGCGATGCCGGAAGGCTATGTGCAGGAACGTACCATCTTTTTCAAAG
ATGATGGTAACTACAAAACCCGCGCGGAAGTTAAATTTGAAGGCGATA
CGCTGGTGAACCGTATTGAACTGAAAGGTATCGATTTCAAAGAAGATG
GCAATATTCTGGGTCACAAACTGGAATACAACTACAACAGTCATAACG
TGTACATTACCGCCGATAAACAGAAAAACGGTATCAAAGCAAACTTCA
AAATCCGTCACAACATCGAAGATGGCGGTGTTCAGCTGGCCGATCATT
ACCAGCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCTGCCGGATA
ATCATTATCTGAGTTACCAGAGCAAACTGTCTAAAGATCCGAATGAAA
AACGCGATCACATGGTTCTGCTGGAATTTGTGACCGCGGCCGGCATTA
CGCATGGTATGGATGAACTGTATAAAAGGCCTGCAGCAAACGACGAAA
ACTACGCTTTAGTAGCTTTAACCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACG
CTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCG
TTTATAACTAGTGCG - 3 ′

Target positions (AFM):

1 for dCas9: 5′ - TTGAAGTGGTGGCCTAACTA - 3 ′
2 for dCas9: 5′ - AAAGTATATATGAGTAAACT - 3 ′
3 for dCas9: 5′ - GTCACGCTCGTCGTTTGGTA - 3 ′
4 for ddCas12a: 5′ - GAGGAGAAATACTAGATGAG - 3 ′
5 for ddCas12a: 5′ - CTGGTGAACCGTATTGAACT - 3 ′
6 for ddCas12a: 5′ - CATGGTATGGATGAACTGTA - 3 ′

Target DNA DNA 08: Cy5 5′ -ATGCCAAAGGTACACCGTTTATTGCAGGT
AAACGCATTGTTCCAGTGATC - 3 ′
5′ -GATCACTGGAACAATGCGTTTACCTGCAATAAACGGTGTACCTTTG
GCAT - 3 ′

Connection RNA - c9c12RNA c9c12a RNA for pSB3A1-T7wt-mVenus:
1-6 c9c12RNA:
5′ -UUGAAGUGGUGGCCUAACUAGUUUAAGAGCUAUGCUGGAAACAGC
AUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG
CACCGAGUCGGUGCAAAAAAAAAAAAAUUUCUACUGUUGUAGAUCAU
GGUAUGGAUGAACUGUA - 3 ′

2-5 c9c12RNA:
5′ -AAAGUAUAUAUGAGUAAACUGUUUAAGAGCUAUGCUGGAAACAGC
AUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG
CACCGAGUCGGUGCAAAAAAAAAAAAAUUUCUACUGUUGUAGAUCUG
GUGAACCGUAUUGAACU - 3 ′

3-4 c9c12RNA:
5′ -GUCACGCUCGUCGUUUGGUAGUUUAAGAGCUAUGCUGGAAACAGC
AUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG
CACCGAGUCGGUGCAAAAAAAAAAAAAUUUCUACUGUUGUAGAUGAGG
AGAAAUACUAGAUGAG - 3 ′

c9c12a RNA for pddCas12a:
7-8 c9c12RNA_N14:
5′ -CAGCGGUCGGGCUGAACGGGGUUUAAGAGCUAUGCUGGAAACAGC
AUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG
CACCGAGUCGGUGCAAGACAAAGCAGCCAAUUUCUACUGUUGUAGAU
CCUGCAAUAAACGGUGUACC - 3 ′

Acknowledgments

First and foremost, my gratitude is directed to Prof. Friedrich C. Simmel. Thank you Fritz, for offering me the opportunity to perform research at your lab and for all your guidance through this highly educational learning experience. Also I want to thank you for the truly international environment in your lab. For bringing me close to Indian, Persian, Moldavian and Ukrainian culture. It was a great honor to have been part of the lab and its international experience.

Lots of additional gratitude to:

To Susanne Kinzel and Helene Budjarek for reducing the general daily burdens and keeping the lab very well-organized.

To all of E14 and to E69. For the good times and for being surrounded by kind and clever people.

For guidance and generally everything to Aradhana, Swati, Aurore, Anna and Vera .

To Aradhana Chopra, Mario Teichmann and Martin Langecker for proofreading.

To all my office mates, Kilian, Katzi, Enzo, Lukas, Aradhana, Swati, Aurore, and Ali. What a great office!

To Lisa, Hannah, Katha, David, Mitti, Michi, Benni. The pizza ppl, the D&D crew, Trevor, the boulder group, the hikers, bikers and the climbing lot. To everyone else who supported me in one way or another throughout the journey.

To my **family**. Danke!