Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
- Dr. rer. nat. -

vorgelegt von

Andreas Sprengel, StEx.

geboren in Essen

Zentrum für Medizinische Biotechnologie
und Fakultät der Biologie
der
Universität Duisburg-Essen

2017

"In Hinsicht auf die Schätzung der Größe eines Menschen gilt für die geistige das umgekehrte Gesetz der physischen: Diese wird durch die Ferne verkleinert, jene vergrößert."

Parerga und Paralipomena; Arthur Schopenhauer

Dedicated to my family

Die vorliegende Arbeit wurde im Zeitraum von August 2012 bis Oktober 2016 am Zentrum für Medizinische Biotechnologie im Arbeitskreis von Dr. Barbara Saccà an der Universität Duisburg-Essen durchgeführt.

Vorsitzender:	Prof. Dr. Stefan Westermann
Gutachter:	1: Prof. Dr. Markus Kaiser
	2: Prof. Dr. Michael Ehrmann

Tag der mündlichen Prüfung: 8.11.2017

Während dieser Arbeit sind folgende Publikationen erschienen, die teilweise Inhalte dieser Dissertation enthalten:

Sprengel, A., Lill, P., Stegemann, P., Bravo-Rodriguez, K., Schöneweiß, E.-C., Merdanovic, M., Gudnason, D., Aznauryan, M., Gamrad, L., Barcikowski, S., Sanchez-Gracia, E., Birkedal, V., Gatsogiannis, C., Ehrmann, M., Saccà, B., Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions. *Nat Commun*, 2017.

Saccà, B., Sprengel, A. and Feldkamp, U., De Novo Design of Nucleic Acid Structures, *De novo Molecular Design* (ed G. Schneider). 2013, Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Index

1. Introduction

1.1 Nanotechnology

In 1959, during his lecture at the American Physical Society meeting in Caltech, Richard Feynman postulated the challenge of writing the *Encyclopaedia Britannica* on the head of a pin. In his visionary speech: "*There´s Plenty of Room at the Bottom*"[1] he paved the way to a new field of science, nowadays known as Nanotechnology. Inventions like the atomic force microscopy (AFM)[2] and the scanning tunneling microscopy (STM)[3, 4] made the nanoworld accessible through manipulation of single atoms.

In the field of nanotechnology, functional systems at the molecular scale from 1 nm to 100 nm are engineered either by top-down or the bottom-up approaches. The top-down approach relies on the traditional microfabrication methods, for example the use of focused ion beams to remove material, while the bottom-up approach is based on the hierarchical self-assembly to build up large structures from small units.

In 1985 the gradient student Tom Newman, from Stanford's University, succeeded in the challenge and reduced the first paragraph of *A Tale of Two Cities* by 25000-fold. He wrote this page on the head of a pin, using a beam of electrons[5].

1.1.1 DNA Nanotechnology

One famous example of the bottom-up approach is DNA nanotechnology which Ned Seeman developed in 1982 using the self-recognition properties of the DNA as an engine for the self-assembly of small DNA tiles[6].The DNA molecule, whose structure was first described by Watson and Crick in 1953[7], with its predictable base-pairing rules, is an ideal material to build nanostructures. Thus, DNA nanotechnology approaches, together with the opportunity to synthesize DNA molecules with a well-defined sequence, allow nowadays creating DNA "objects" of almost any desired shape and size, reaching, "*the finest possible level of control over the spatial and temporal structure of matter: Putting what you want where you want it in three dimensions, when you want it there.*"[8]

The DNA double helix is built up from two antiparallel right-handed strands, twisting around a central axis. (*Fig. 1*) An alternating deoxyribose-phosphate chain, connected through covalent bonds, forms the negative charged backbone of the DNA, here shown schematically

as blue stripes. The four bases, adenine (A), cytosine (C), guanine (G) and thymine (T) are connected via the 1′ carbon of the deoxyribose to the backbone and are perpendicularly oriented to the central axis of the DNA. Adenine with thymine and cytosine with guanine form a pair with different number of hydrogen bonds resulting in different bonding strengths. An increase of the stability of a certain double stranded DNA (dsDNA) segment can be achieved by raising, the number of cytosine-guanine pairs. The hydrogen bond and stacking interactions of the nitrogen-containing bases keep the double helix in place. The glycosidic linkages of the base pairs are not diametrically at opposite positions, causing the formation of the minor and major groove. The grooves are often binding sites for regulative enzymes.

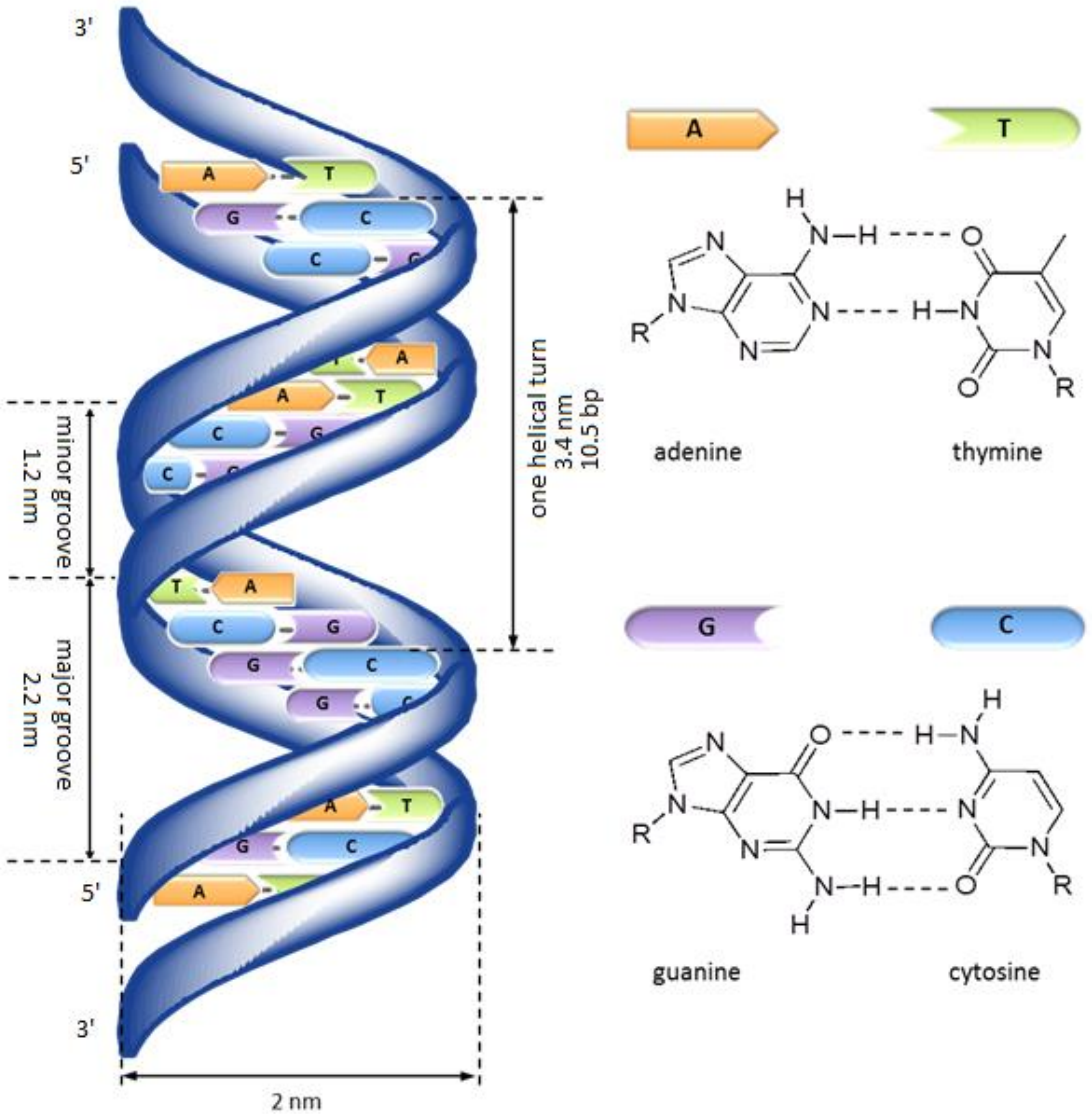


***Figure 1*:** The scheme of a DNA double helix in its common right-handed B form with its minor (1.2 nm) and major (2.2 nm) groove. Each 10.5 bases the helix turns 360° around its central axis. The DNA double helix has a diameter of 2 nm. The four bases are shown on the right, each with its complementary one. Adenine (A) with thymine (T) and guanine (G) with cytosine (C) form a pair. (adapted and modified from Ref.[9])

Using DNA as a construction material has several advantages: (i) DNA is easily available as it can be purchased or synthesized by solid phase synthesis, (ii) DNA is a nanometer-sized object, (iii) DNA has good chemical and physical stability, (iv) DNA is accessible in different chemically modified forms and (v) DNA can be processed by a large amount of enzymes, like DNA ligases, kinases and exonucleases.

1.1.2 The DNA crossover or Holliday junction (the multi-stranded approach)

The DNA crossover or Holliday junction is the easiest motif to connect four single stranded DNA (ssDNA) or two double stranded DNA strands. In nature, these junctions appear during meiosis in the recombination process and are unstable because of their sequence symmetry: they are indeed constituted by two identical pairs. If the molecule is small, it can totally resolve in two separated double stranded helices[10]. To avoid this "transformation" of the Holliday junction, which is called branch migration, the sequence symmetry at the branch point has to be reduced[11]. Consequentially the term immobile Holliday junction refers to a fixed branch point that lost the ability to migrate. To create sequences with minimal symmetry, Ned Seeman developed the software SEQUIN to assign oligonucleotides SEQUences INteractively. This software is still one of the most valid tools for designing complex DNA motifs based on minimal symmetry of the oligonucleotides.

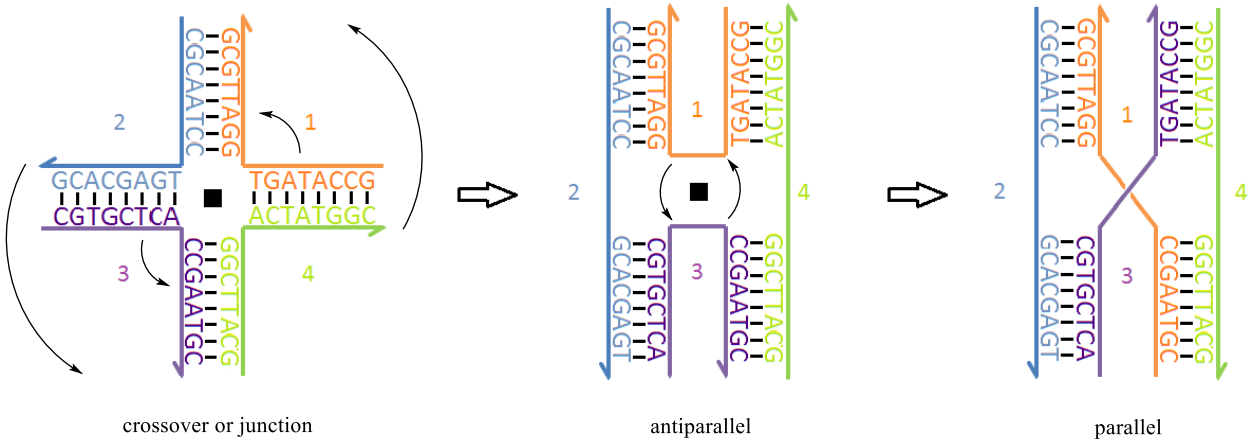


***Figure 2*:** The crossover or Holliday junction is the basic motif of DNA nanotechnology. Four single stranded DNA strands, each 16 bases long are divided into 2 segments consisting of 8 bases each complementary to two segments of the other DNA strands, e.g. the first 8 bases of strand 1 are complementary to the last 8 bases of strand 4 and the last 8 bases of strand 1 are complementary to the first bases of strand 2. This results in a central branch which can be immobilized by minimizing the sequence symmetry[11]. Depending on the way of rotation the central DNA strands proceed in an antiparallel or parallel way (taken and adopted from[8]).

Figure 2 schematically illustrates the design of the first immobile four-arm junction designed with SEQUIN (*Fig. 2*). The half-arrow heads represent the 3′-terminus, going clockwise

around the branch point and the four oligonucleotides are numbered according to the coordinate system. Depending on the rotation around the branch point the motif adopts an antiparallel or parallel configuration meaning that the two helices are orientated in the same (parallel) or opposite (antiparallel) directions. To build up super-structures, each motif has to be connected to another one in a certain and controlled way. Using the Watson-Crick base-pairing rule allows adding specific "sticky ends" which can be described as an oligonucleotide elongation of at least three bases. If the sticky ends contain fewer bases, room temperature would be enough to cause melting of the newly formed double stranded part. More stable binding between the sticky ends is achieved by increasing the amount of guanine and cytosine pairs. Such single stranded regions of the oligonucleotide of one motif are complementary to the single stranded region of a different motif. An example for a super-structure is the combination of two crossovers into a so-called double crossover motif (DX) or of two DX motifs spliced in the middle and arranged perpendicular to each other forming a 4x4 motif (*Fig. 3*)[12].

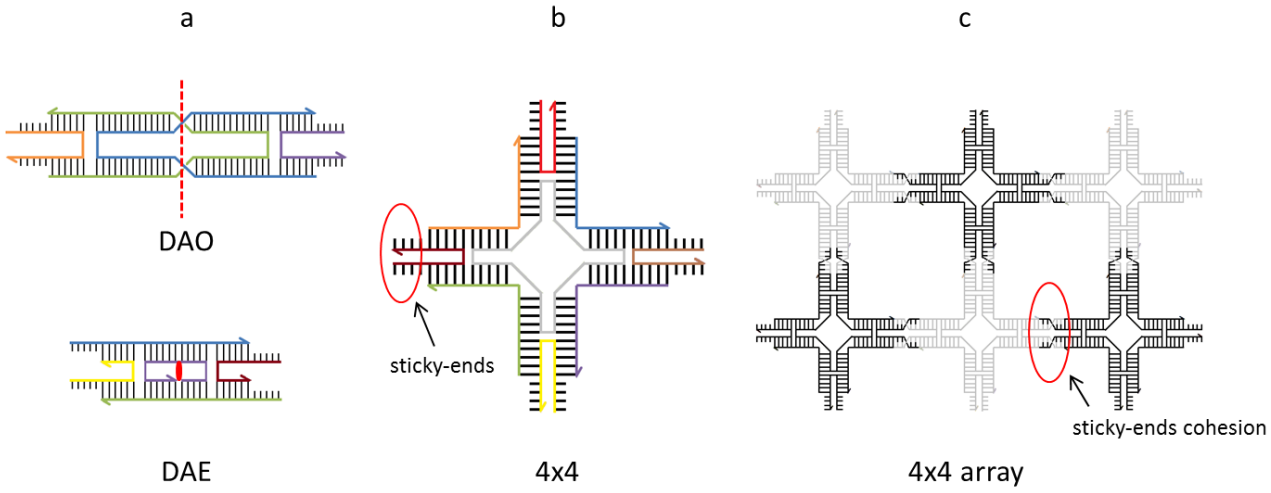


***Figure 3*:** a) Double crossover molecules (DX) motifs are formed by antiparallel DNA helices with an odd (DAO; 2n+1) or even (DAE; 2n) number of half helical turns (n = 5 bp). The red line in the DAO and the red dot in the DAE motif illustrate the symmetry of these structures. b) The 4x4 motif is built up from four Holliday junctions connected in a common junction and skewed perpendicular to each other. By mutually complementary sticky-ends, the 4x4 motifs can thus build large planar square lattices. c) Six 4x4 motifs forming an array by sticky-end cohesion. Atomic force microscopy allows to characterize these structures which in some cases may reach a size of several microns[13].

There are five possible DX motifs: DAO, DAE, DPE, DPOW and DPON. In this nomenclature, D stays for double-crossover, A and P stay, respectively, for antiparallel and parallel helices, E and O mean, respectively, even and odd number of half-helical turns, while W and N stay for wide and narrow minor-grooves. The DA motifs have an antiparallel relative orientation of the double-helical domains either with an odd (DAO) or even (DAE) number of half helical turns between the two branch points. Again the 3′ is represented by the

half-arrow heads. Both structures have a symmetry (shown by the red dashed line in the DAO and the red lens in the DAE motif, *Fig. 3a*). The same applies to the DP motifs with parallel relative orientation of the double-helical domains (DPE or DPO). In particular, the DPO motifs may differ in the number of minor and major grooves included between the two crossovers, depending on the number of base pairs (*Fig. 4a*). In the case of 14 bases between the branch points, there is an excess of narrow minor-grooves (DPON), while using 16 bases results in an excess of wide major-grooves (DPOW)[10] (*Fig. 4b*). All three molecules share their symmetry axis with the axis of the double helix. DPON and DPOW molecules also differ in the 5′ → 3′ direction of each crossover strand that points towards the center in DPOW and points away in DPON.

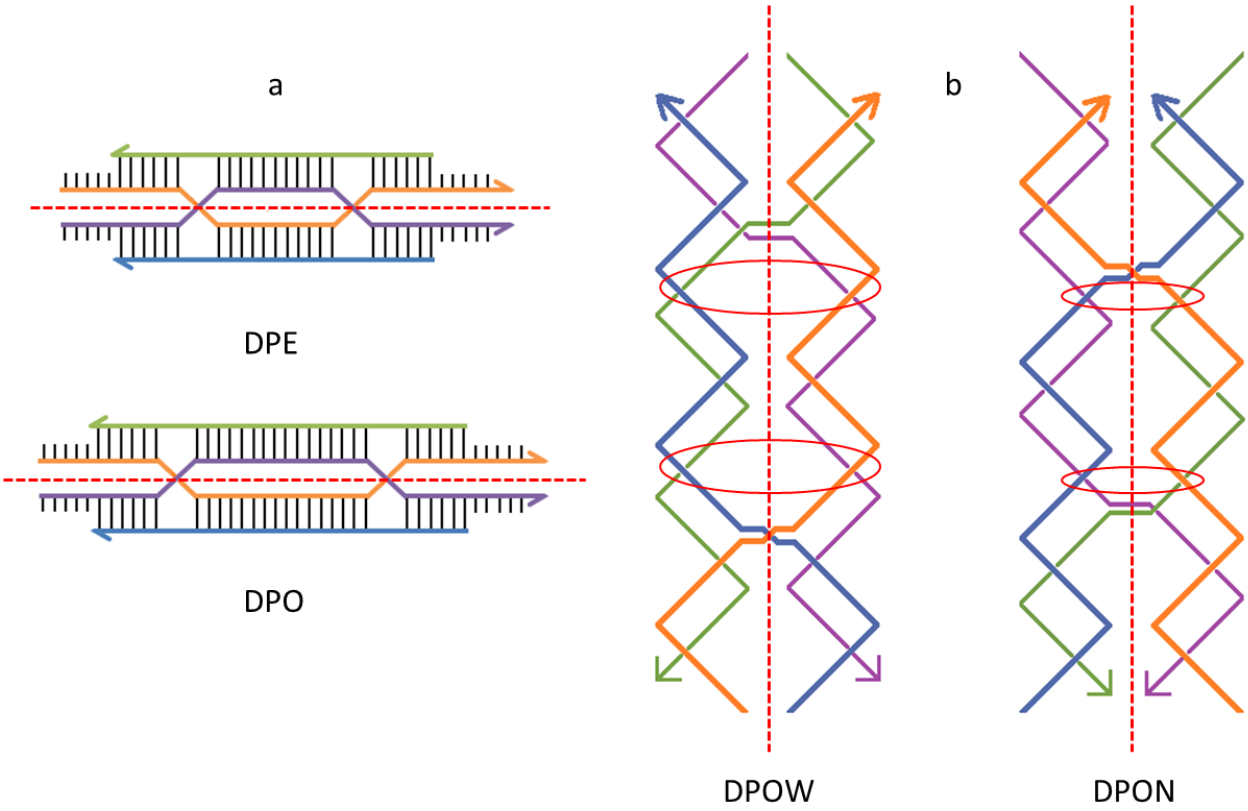


***Figure 4*:** The three possible parallel double crossover motifs with their symmetry axis, which is parallel to the central axis of the DNA molecule itself, represented by the interrupted red lines. (**a**) The two possible double-crossover motifs with a parallel orientation of the double-helical domain with an odd (2n+1) or even (2n) number of half helical turns. (**b**) To show the difference between the excess of minor and major-grooves in the DPO structures, the molecules are illustrated in a more spatial way. The red circles show either the excess of wide major-grooves (DPOW) or the excess of narrow minor-grooves (DPON).

Other important and more complex motifs than the DX molecules are illustrated below (*Fig. 5*). The DX structure on the left which is similar to the DAE molecule serves as comparison. By elongation of the central oligonucleotide perpendicular to the plane of the two double helices, another DNA domain is added to the DX motif. This additional domain may be used as a topographical marker or for functionalization of the molecule. Both, the PX and the JX_2

structures consist of two pairs of identical oligonucleotides and differ in the number of crossovers in the central region. While the PX motif contains crossovers at each possible position, the JX_2 motif lacks two of them. The TX structure is formed by three double helical domains and shows the versatility of possible DNA building blocks.

However, creating nanostructures out of a small number of short oligonucleotides, results in a system that is sensitive to the stoichiometry and purity of the oligonucleotides. Furthermore, the yields of these structures are often low.

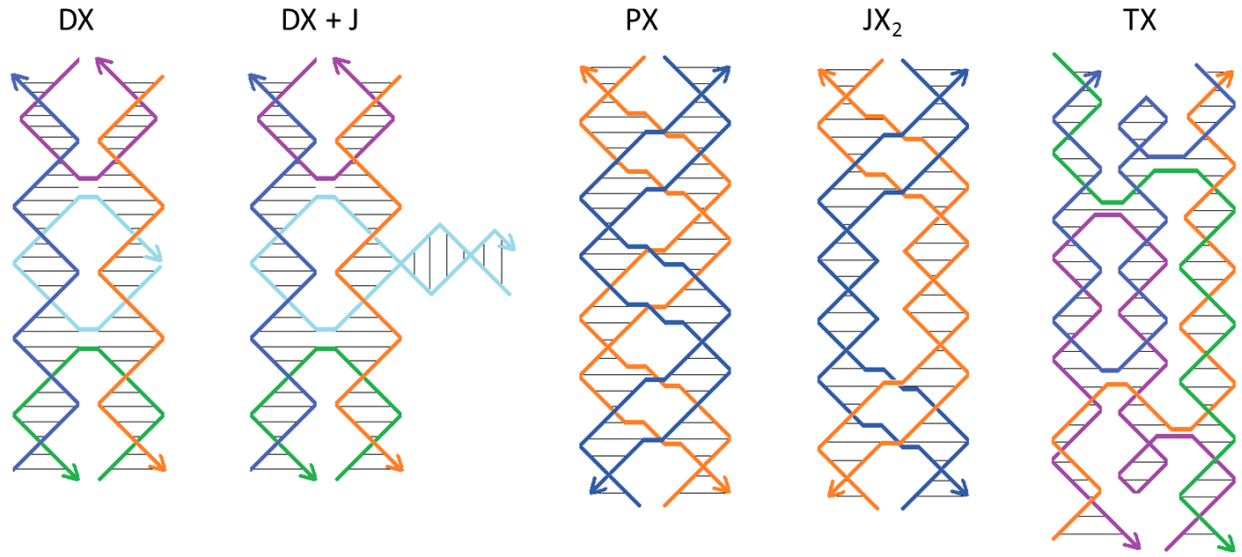


***Figure 5*:** Scheme of four furthers multiple crossover motifs; the DX motif is similar to the DAE molecule. By elongation of the central oligonucleotide, another perpendicular DNA domain is added to the DX + J molecule. The topoisomers PX and JX_2 are formed by identical juxtaposed helices and differ in the number of crossovers in the central region. While in the PX motif each possible crossover is formed, the JX_2 lacks in two central exchanges. Both motifs are central building blocks of robust nanostructures[8]. The TX structure is formed by three double helical domains which are formed by antiparallel DNA helices.

1.2 DNA origami

In 2006 Paul Rothemund developed a scaffold based approach to construct large nanometer sized objects via a circular viral single-stranded DNA strand (scaffold) folded into a desired shape by the help of hundreds of short helper-strands called staples[14] (*Fig. 6*). Referring to the Japanese art of paper folding, this approach was given the name “DNA origami”. In this design strategy, the 7249 bases-long viral genome of the M13mp18 phage was used as a scaffold and folded into six different shapes like a star or a triangle. Each of these structures can be described as an array of antiparallel helices formed by a repetition of periodic crossovers. Incorrect or partial hybridization of the staples to the scaffold are automatically displaced through strand invasion and exchange mechanisms because of the favored

completely hybridized sequences. Consequently, experimental mistakes are reduced and yields are higher. Each staple is added in a one-pot reaction, usually in large molar excess without the need to be previously purified.

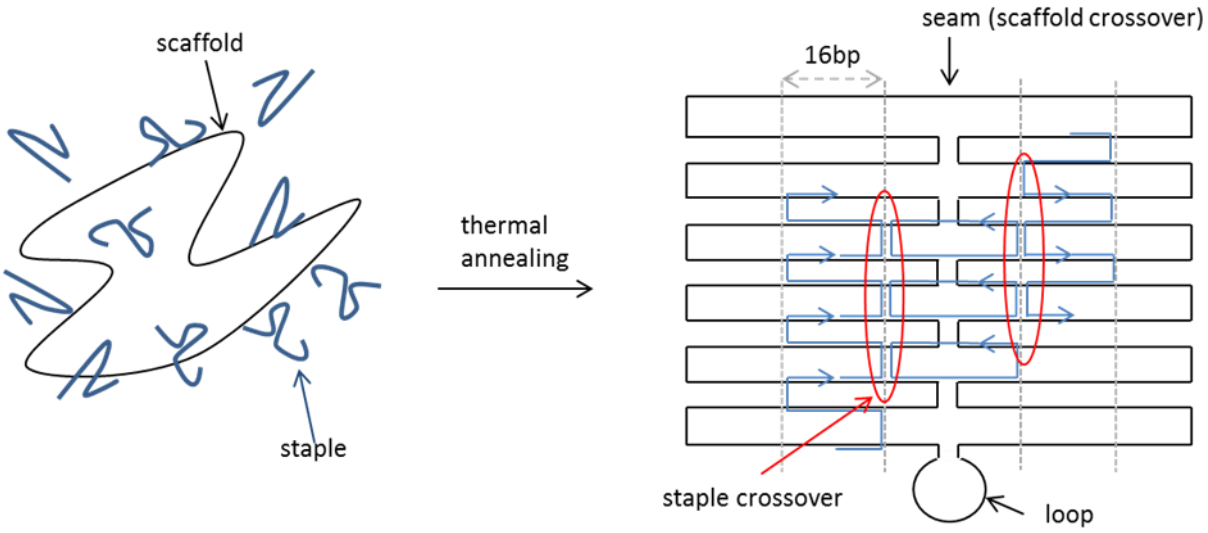


***Figure 6*:** Short overview of the DNA origami method. In a one pot reaction the scaffold is folded into its desired shape by the help of about 100-200 short staples hybridizing with its complementary sequence, just by thermal annealing. The sequences of the staples can easily be generated with the AutoCAD software CaDNAno[15]. Connecting the helices every 16 bases results in a planar structure, here e.g. a rectangle, with well-defined dimensions.

Furthermore, the time for the annealing process takes one to two hours (heating up to 90 °C, cooling down -1 °C per minute) in comparison to the usual 20 hours for the superstructures derived by the multi-stranded approach. Due to the already mentioned advantages and the astonishing high yields, the DNA origami method had a strong impact in the field of nanotechnology as a powerful tool for construction of DNA objects with desired shape and dimensions.

1.2.1 Single-layer planar DNA origami

Depending on the spatial connection between DNA helices, two or three dimensional architectures can be constructed. For simplification, the complex DNA molecule is idealized as a cylinder (*Fig. 7*). To arrange the helices in a planar sheet, the connection angle between adjacent helices has to be 180°. Each helix is connected every 16 bases with an adjacent helix, thus ca. every 1.5 helical turns. This leads to a periodic 180° connection of the same helix with the one directly above and below, such that two adjacent helices are mutually linked

every 32 bases (i.e. ca. every 3 helical turns). Computer-aided design (CAD) tools like caDNAno (http://cadnano.org), NanoEngineering[16] and SARSE are available to assist designing and visualizing of DNA origami structures. The size of a structure is limited by the number of available bases which is defined by the length of the scaffold. Commonly, the 7249 bases-long M13mp18 viral single stranded DNA genome is used as a scaffold. To get larger structures, two or more DNA origami architectures can be connected by sticky-end cohesion or stacking interactions[14, 17, 18]. In general, the double helical domains of a DNA origami should be long enough to ensure thermal stability at ambient temperature conditions but not too long to avoid exceeding structural tension and folding problems. As the scaffold is circular, a central seam is necessary to allow its back and forth folding along a closed pathway. The seam itself is stabilized by seam-crossing staples. Despite these limitations, almost each imaginable shape can be designed following the above-mentioned rules. To avoid unwanted stacking interactions between DNA origami structures, resulting from the π- π interactions between parallel bases, the staples at the edges are normally elongated with at least two thymine residues.

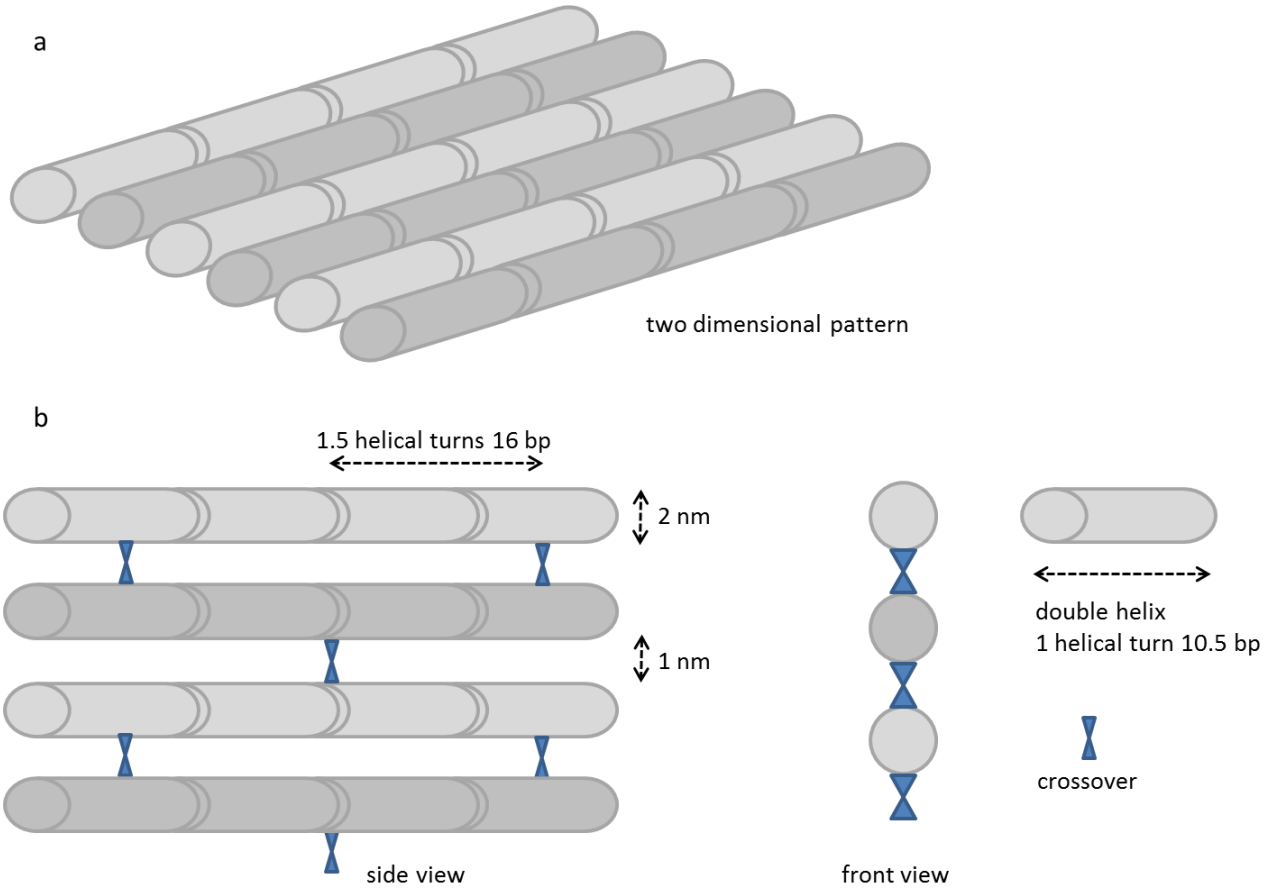


***Figure 7*:** Basic design principles of single-layer DNA origami. a) Scheme of a two dimensional pattern. Each cylinder represents a double helical domain, highlighted by light and dark grey shades. b) The connections between the helices in a side and front view. In two dimensional origami structures, each helix is connected every 16 base-pairs with an adjacent helix, either above or below the central axis leading to a register of crossovers every 180°. The size of a DNA single-layer

origami structure can easily be calculated using (3n-1; n = amount of helices) for the width and (n x 3.4/10.5; n = number of bases) for the length.

1.2.2 Single-layer three-dimensional DNA origami

To create a three-dimensional origami structure out of a planar sheet, as in this thesis, two planar sheets have to be connected at an angle that differs from 180 degrees. The rotation of one base pair around the central axis of a double helix results in a 34.29° angle. Using the predictable spatial orientation of a base pair along the helix, single strands, so called spatial or 3D staples, can be designed that create a "crossover" not lying in the plane of the sheets. A possible way for connecting two planar sheets one another relies in the insertion of three thymine bases at the virtual crossover between two adjacent sheets. In this way, a flexible hinge will be created to reduce the mechanical stress between the connected sheets (*Fig. 8a*).

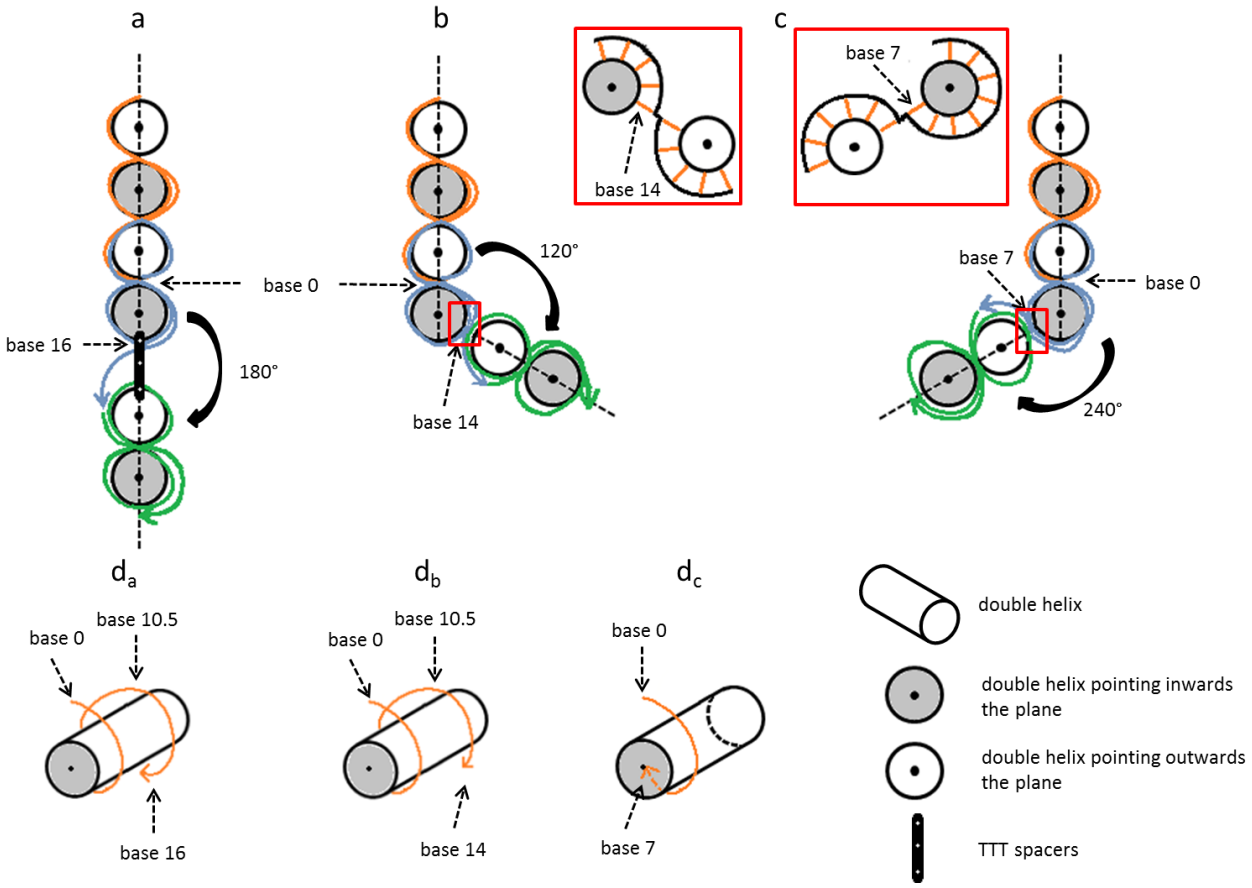


***Figure 8*:** Scheme of spatial staple connections of two shapes for a three dimensional structure. The upper four helices belong to one, the lower two to another shape. The staples are shown as colored lines, the scaffold as a helix to simplify visualization. (**a**) To reduce mechanical stress the space between the shapes is increased by adding three thymines to the spatial staples to allow correct arrangement of the shapes. Here, two shapes with an undefined spatial arrangement are shown, due to the flexible 180° angle. The helices are connected in the usual way; the three thymines, indicated by the black bar with the three white dots increase the distance between both shapes. (**b**) Calculating a rotation of 34.29° per base, the staple needs 3.5 more bases after the first helical turn to get in an angle of 120° related to the upper shape. The staple forms a crossover with the scaffold between base 14 of the upper and base 0 of the second shape to regain the common way of planar binding. Experiments show that the mechanical stress, due to this way of direct connection does not influence the yield and stability of the structures. (**c**) Connecting the helices after the rotation of 7 bases around the upper helix to base 0 of the second helix

creates an angle of about 240°. (**d_{a-c}**) Idealized rotation of the 3D staple around the central axis of the helix for the defined angles 180°, 120° and 240°.

Using thymine spacers to connect base 16 of the last helix of the upper sheet and base 0 of the first helix of the lower sheet will result in a flexible angle between those sheets. Defining a front and back for each planar sheet, it is not predictable whether they will be arranged with their front faces pointing inwards or outwards. Actually, both ways of folding are possible (chapter 3.2.1-3.2.2). Breaking the 16 base-pair rule by creating a 3D crossover after 14 or 7 base-pairs results, respectively, in an angle of 120° or 240°. In this way, the inner and outer surfaces can be distinguished.

Due to the missing thymine spacers, the mechanical stress is increased; however, this has surprisingly no effect on the yield and stability of the structures (chapter 3.2.2). Using this technique to connect helices, several three dimensional origami structures have been designed like barrels[19], cubes[20, 21] and prisms[22]. The size of the structures and therefore the size of the inner cavity are again limited by the length of the scaffold.

Ignoring the 10.5 base-pair rule to get planar structures, Yan and coworkers designed bended DNA origami structures by a network of latitudinal and longitudinal crossovers[23]. Connecting the helices in a range of 9 to 11 base pairs per helical turn, they succeeded in the construction of shapes like planar concentric rings, spheres and hemispheres.

1.2.3 Multi-layer DNA origami

One drawback for the complexity of three dimensional single-layer DNA origami structures is the low resistance against mechanical stress, resulting in deformation or even damage of the structure. Using a different design strategy, space-filled three dimensional origami structures can be constructed by connecting adjacent helices every 7 base pairs (*Fig. 9*). Assuming a central helix, which is surrounded by three other helices, the crossover from the central helix to one of the three surrounding helices occurs every 21 base pairs. Connecting several helices in such a way results in a honeycomb-like architecture wherein each helix is connected to the adjacent ones at an angle of 120°. Shih and co-workers published the first examples of multi-layer DNA origami structures in 2009[24].

The complexity of DNA origami structures rapidly increased after getting able to engineer shapes in a twisted and curved state[25]. By deleting or inserting a selected number of base pairs between adjacent crossovers, twisted DNA bundles of either handedness were created. Whereas base deletion results in left-handed twists, base insertion leads to right-

handed twists. Combining deletion and insertion of bases leads to untwisted but bended DNA bundles tunable in 5° steps. Using this technique, a beach ball, a concave and convex triangle was engineered[25].

Another way to overcome the low resistance against mechanical stress is a structural combination of tension and integrity according to the tensegrity concept of Kenneth Snelson.

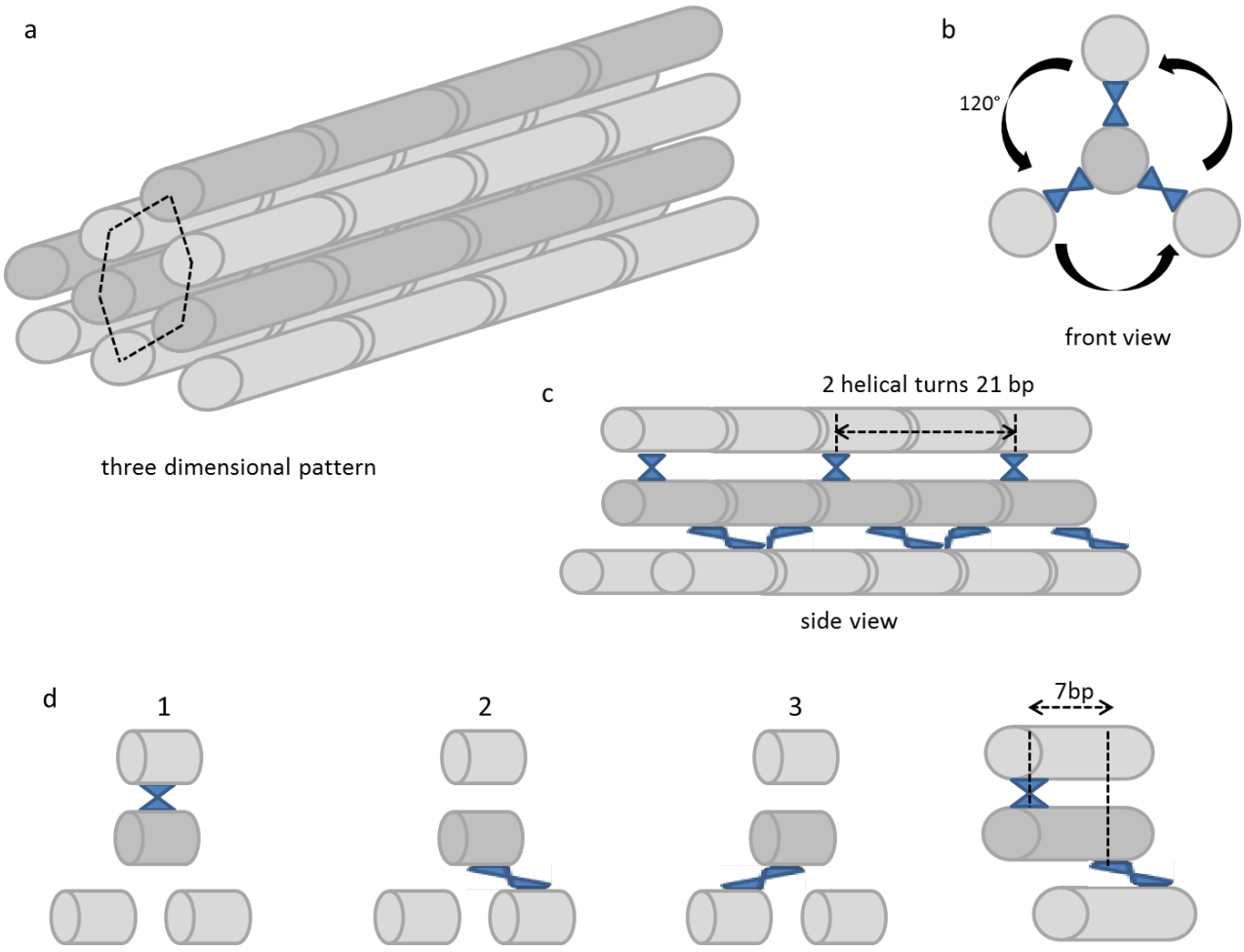


***Figure 9*:** Basic design principles of multi-layer DNA origami. (**a**) Scheme of a three dimensional pattern wherein the helices are connected every 120° resulting in some kind of a honeycomb structure. Therefore look at the marked honeycomb through the central axis of the double helices. (**b**) Front view to visualize the 120° crossover angles. (**c**) In common three dimensional patterns each helix is connected every 7 base-pairs with an adjacent helix, all 21 bases with the same one leading to a register of crossovers every 120°. (**d**) Every 7 base-pairs "cut" side view of the origami for a better visualization of the clockwise connection to the adjacent helices.

In this concept, pre-stressed origami structures with rigid bundles resisting against mechanical stress are supported and hold in position by tension-bearing single stranded DNA[26, 27]. Adjusting the right balance between forces pulling out- and inwards, stable and stress resistance structures can be designed.

1.2.4 Alternative design methods

Three-dimensional DNA origami structures are typically formed by tightly packed parallel DNA helices, which are therefore not fully addressable for functionalization or modification. In addition, the density of the DNA helices prevents DNA origami structures from being degraded in cells[28]. Due to the finite length of the chosen scaffold, either the size of the packed origami is limited, or the structure lacks in mechanical stability if hollows are integrated into it. To build up larger superstructures, longer scaffolds, interconnecting staples between two or more DNA origami structures or different design strategies have already been developed[17, 18].

1.2.4.1 Wireframe DNA origami

There are two different approaches to the wireframe DNA design[29, 30], which share a common idea but differ in the detailed design principles. Both concepts are based on single DNA helices or antiparallel DX motifs connected at common vertices of a wireframe structure.

B. Högberg *et al.* developed a highly automated design strategy that creates complex three-dimensional structures with triangular-mesh architecture. The method is based on the "Chinese postman tour" problem in graph theory and enables to fold arbitrary polygonal meshes, which are otherwise difficult to get using previous approaches. The scaffold is not allowed to cross itself at the vertices and should only pass once at each edge of the mesh using as less scaffold as possible. To reduce physical stress on the vertex junctions, the software modifies the length of the double helices, which occupy the edges of the mesh. Unpaired bases in the staple design could be added for fine-tuning smaller gaps within the relaxed model, thus providing flexibility and correct staple alignment.

In another approach Yan *et al*. designed various planar and three dimensional complex DNA structures, like a flower-and-bird pattern or a snub cuboctahedron, out of meshes of stiff rods represented by antiparallel DX motifs and vertices formed by the intersection of multiple DX motifs. By insertion of unpaired staple- and skipped scaffold-bases the junctions could be arranged in every desired angle. The length of the rods is determined by a multiple number of full helical turns. To overcome structural size limitation of the scaffold, two scaffolds were used and connected at each possible position by staple strands.

1.2.4.2 Single-stranded DNA tiles

Similarly to the multi-stranded approach of Ned Seeman used to create a big lattice through sticky-ends connected tiles, the single-stranded DNA tiles form a canvas of about 310 pixels with a total size of 24 helices in height and 28 helical turns in width. Each pixel is a unique tile designed according to the sequence minimization principle and is binding to four neighboring tiles. All staples are about 42 bases-long and are each divided into 4 domains being able to bind segments of four different staples. The canvas can be seen as parallel helices formed by short staples. Using this strategy, each shape was created just by leaving out the sequences that are not included in the desired structure. About 107 different shapes were assembled, including the 10 Arabic numbers, the 26 capital letters of the Latin alphabet and much more[31]. Almost all designed structures are planar except a long tube being formed by connecting both edges of the canvas.

1.2.4.3 DNA bricks

To build up three dimensional architectures P. Yin *et al.* published in 2012 a design strategy which is similar to the single-stranded tiles and reminds of LEGO bricks[32]. In a one-pot mixture of hundreds of staples, each with a unique sequence of 32 bases defined by the sequence minimization principle, every desired 3D shape was assembled. Again each staple is divided into four 8 base-pair domains, being able to bind segments of four different staples. One plane of the canvas is formed by bricks of the same spatial orientation, e.g. vertical. The adjacent plane is therefore built from bricks orientated perpendicular to the previous plane; i.e. horizontal. Creating a canvas consisting of 10 planes, each consisting of 10 bricks in width and height, increases the number of pixels or voxels to 1000. Using a three dimensional canvas, about 102 solid and hollow shapes were created, even without stoichiometric control.

1.2.5 Functionalization of DNA origami

Inserting bulky dumbbell hairpins at selected positions, Rothemund introduced a method to functionalize DNA origami structures which could easily be visualized by AFM[14]. These Dumbbell-coded DNA origami structures found widely use as Nano-chips for detection in single-molecule experiments[33].

Due to the solid phase synthesis of oligonucleotides, functional groups can be easily introduced into DNA origamis, commonly at the terminal phosphate by a carbon linker. Several functional groups like thiol-, amine-, biotin-, cyanine- and fluorophores can be added either to the 3' or 5' terminus of the DNA strand. Modifying oligonucleotides with a biotin linker enables streptavidin (STV) binding at predefined positions of the origami. This protein binds to biotin within some minutes and is, due to its size, detectable by AFM. Biotin and streptavidin conjugates are used in single-molecule experiments to immobilize DNA origami structures on a solid surface.

The principle of attaching ligands to a DNA strand to bind proteins at well-defined positions is not limited to biotin streptavidin conjugates. Other examples include the benzyl guanine (BG) and chlorohexane (CH) moieties as suicide ligands to bind fusion proteins containing, respectively, the "Snap" or "Halo" tag[34]. In addition, gold- or silver-nanoparticles of various sizes were bound to DNA origami structures by thiol-modified ssDNA[35, 36]. A general method to connect proteins to DNA structures is to crosslink e.g. a lysine residue of the protein with a thiol-modified oligonucleotide[37-39].

All these approaches share the same disadvantage: the irreversibility of the covalent bond between the oligonucleotide and the protein. This leads to permanent modification of the protein surface which is not easily controllable neither desired. In addition, as several reactive amino acid residues are exposed on the protein surfaces, the crosslinking reaction normally lacks regioselectivity and stoichiometric control, generating an "undefined" layer of DNA molecules surrounding the protein of interest.

To bypass the use of covalent chemistry on proteins, a supramolecular approach is shown in Chapter 3.4.3.

1.2.6 DNA origami for protein caging

Immobilizing and attaching proteins to artificial DNA structures to get protein-crystals for X-ray studies was the initial motivation that led Ned Seeman to the establishment of DNA nanotechnology. Since then, artificial DNA nanochambers have been used, (i) as carriers for transcription factors[40]; (ii) as logic gated nanorobots for molecular payloads[19]; (iii) to create bio-hybrid materials[41]; and (iv) to study the reconfigurability of a DNA box[42]. Despite the enormous potential of such structures, real applications are emerging only recently. A first study concerning the enhanced activity and increased stability of DNA encaged proteins has been published in 2016 by Yan *et al.*[43]. In this work, a pair of DNA-modified enzymes, i.e.

the glucose oxidase and the horseradish peroxidase, were attached to the inner cavity of a three dimensional nanocage consisting of two half-blocks hybridized by short staple strands. The enhanced activity of the system was attributed to the negatively charged backbone of the DNA cage surrounding the proteins. However, not all proteins studies showed the same effect: the largest protein, being closest to the negative DNA layer, showed indeed no activity increase.

1.3 Serine Proteases

Serine proteases or serine endopeptidases belong to a class of enzymes which cleave peptides at special positions within their active site by a catalytic triad. One third of all proteases belong to the class of serine proteases. Responsible for their activity (and also eponymous) is an extraordinary reactive serine which is preserved in all proteases of this family. Originally, this catalytic serine was only found in a catalytically relevant amino acid motif known as the catalytic triad that is formed by three amino acids, histidine (His), serine (Ser) and aspartate (Asp) and is localized in the active site of the protein[44]. Diverse serine proteases that differ in their catalytic triad or dyads have been later discovered, including Ser-His-Glu, Ser-Lys/His and His-Ser-His [45]. However, despite different structural solutions, all serine proteases differ from each other only in nuances concerning the spatial arrangement of their active site, due to their task to cleave peptides[46].

1.3.1 Chaperones

Chaperones are proteins with several functions, like assisting folding and assembly of macromolecular structures. Primarily, this type of proteins is involved in the control of protein folding. During translation of non-membrane associated proteins, the chaperones bind to the nascent polypeptide and assist in the folding process, thus avoiding misfolding or aggregation. For this reason, many chaperones belong to the class of heat shock proteins, which are expressed as a response to elevated temperatures to avoid increased aggregation of misfolded proteins. Additionally, some chaperones convey structural information onto proteins that cannot be folded spontaneously. However, there are several functions of chaperones, which still have to be discovered, e.g. their role in diseases caused by protein aggregation[47]. Such a folding disease is for example the poorly understood Alzheimer disease.

1.3.2 DegP

DegP is an extra-cytoplasmic heat-shock factor and belongs to the family of high-temperature requirements A (HtrA) proteins being essential for biological systems[48]. Working as an ATP independent protease chaperone complex, under protein folding stress conditions, DegP channels proteins into repair, assembly or degradation pathways. Interestingly, the protein is able to adopt at least three potentially coexisting modes: (i) in the presence of damaged and non-native cell-envelope proteins, DegP may act as an efficient endoprotease exclusively for unfolded substrates[49, 50]; (ii) respective to its chaperone function, DegP can aid other proteins being folded correctly[50] and (iii) in a third mode, DegP provides protection from uncontrolled proteolysis by encapsulation[51]. DegP itself is built up from monomers (48.6 kDa) including a serine protease domain and two C-terminal PDZ domains. The PDZ-1 domain is responsible for substrate binding and allosteric regulation; the PDZ-2 domain overtakes a different function and, stabilizes higher-order quaternary structures. All oligomeric states are cage-forming multiples of trimers with the proteolytic site inside the cavity. By adding substrates or allosteric effectors, the equilibrium between the different oligomeric states is shifted. In its hexameric form, DegP is in its resting state with a low proteolytic activity, whereas the 12- and 24-mers are the active conformations. Such a conformational change is for example induced by addition of the activating peptide DPMFKLV which binds the PDZ1 domain of the DegP protein with high affinity (K_d of 5 µM).

2. Purpose and working plan

This PhD thesis included aspects of biology, chemistry, biophysics and bio-informatics and therefore required also the contribution from other research groups. This will be cited and clearly outlined in each part of the thesis. The present chapter is divided into a purpose (chapter 2.1) and a working plan (chapter 2.2), this latter will be described in more detail in the experimental part (chapter 6).

2.1 Purpose

The aim of this thesis was to encapsulate different oligomers of the DegP protein inside a huge DNA origami structure by non-covalent supramolecular forces. Therefore, a suitable hexagonal hollow DNA origami nanochamber had to be designed, being accessible for proteins by diffusion. Using spatial out-of-plane crossovers, the angles between the planar faces of the shape were rationally designed, thus enabling to control the orientation of the catching arms. Three different structures were designed, connecting the faces one another at 120°, 180° and 240°. In the 120° design, the catching arms were oriented towards the center of the cavity: this allowed interaction between the arms and the protein diffusing through the origami channel. As a catching arm, a DNA-heptapeptide (DPMFKLV) conjugate was synthesized, which mimics a substrate with high affinity to the PDZ1 domain. After encapsulation and purification, several structural studies were performed to verify binding specificity and possible selectivity.
A general synthesis route for peptide-DNA conjugates without the use of standard in-solution conjugation chemistry was also established. Along these lines, the first supramolecular DNA origami-protein complex was constructed, purified and characterized.

2.2 Working plan

Mimicking natural host-guest systems that are usually stabilized by supramolecular forces leads to the creation of artificial compartmentalization systems for selective encapsulation of target proteins in their native form. In this PhD thesis, the protein DegP was chosen as an exemplary guest: due to its high symmetry, the binding sites exposed over the whole protein surface are accessible in any spatial orientation, thus favoring encapsulation. Each DegP

oligomer is build up from monomers, consisting of three domains (*Fig. 10*): (i) a serine protease domain with the active center (red), (ii) a PDZ1 domain for substrate recognition (green) and (iii) the PDZ2 domain that mediates stabilization of the quaternary structure (blue). The size of the oligomers ranges from the 6-mer (inactive; $DegP_6$; ca. 250 kDa) to the 12-mer (active; $DegP_{12}$: 500 kDa) till the highest 24-mer (active; $DegP_{24}$; 1 MDa). The oligomers therefore differ in their diffusion-rate, which is inversely proportional to their size, so that one would expect a difference in the binding efficiency of the host-system. The mutation of the serine residue in the active site of the catalytic triad with an alanine (DegP SA; chapter 3.1) avoided substrate digestion and auto-proteolysis. Once captured in a 1:1 ratio inside the cavity, the protein is surrounded by a negatively charged polyphosphate layer that is supposed to act as a primordial "chaperone"[52], thus opening the way to a series of future investigations.

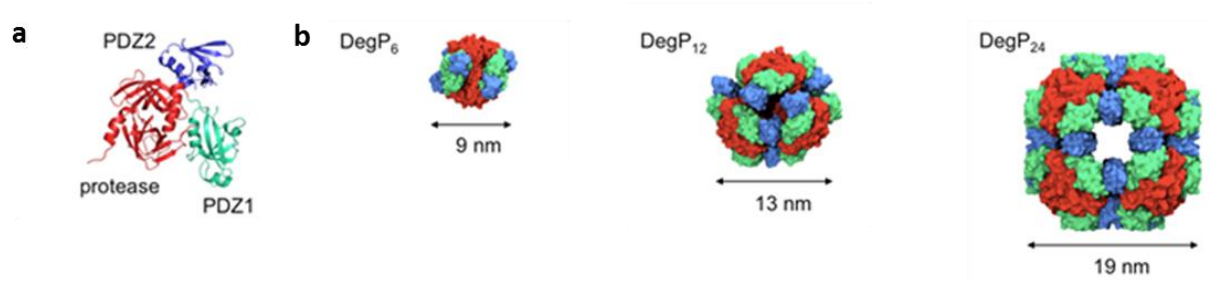


***Figure 10*: (a)** The monomeric unit with its protease domain (red), the substrate binding PDZ1 domain in (green) and the PDZ2 domain that mediates stabilization of the quaternary structure (blue). (**b**) The different guest proteins in their 6-, 12- and 24-mer oligomerization state shown with their relative size and different symmetry.

This PhD thesis can be divided into three milestones:

(i) Designing a suitable three-dimensional hollow DNA origami host that allows applying multivalent binding at short intermolecular distances by placing a layer of ligands in close proximity of the binding sites on the protein surface. The size of the inner cavity of the structure and consequently the size of the whole structure is determined by the largest oligomer of the DegP protein. With a protein radius of 9.5 nm and a host-to-guest bridge of 10 nm in length, providing also some degree of orientational freedom for an easier positioning of the ligand in the PDZ1 domain of the protein, the cavity of the DNA host should have a radius of ca. 20 nm. To create a chamber with predictable orientation of the ligands, several designs were developed that differ in their spatial out-of-plane crossovers between the helices of adjacent faces (Fig. 11). With a dihedral angle of 120° ($6p^{120}$) all protruding arms (PAs) are oriented inside the cavity, whereas all PAs are orientated outwards applying an angle of 240° ($6p^{240}$). Connecting the planes at their edges by flexible T-hinges (180°) results in a structure with an undefined direction of the PAs ($6p^{180}$). The structure was modified with a diverse

number of PAs for each face, namely zero, one or three arms indicated as 0cA1, 6cA1 or 18cA1. Applying all these criteria results in a prism with a vertex-to-center distance of about 23 nm (R_{out}), an inner radius of 20 nm (R_{in}), an accessible room of 10 nm (R_{free}) and a length of 49 nm, except for two opposite faces with slightly longer edges (56 nm), which offered a useful topographical feature for AFM characterization.

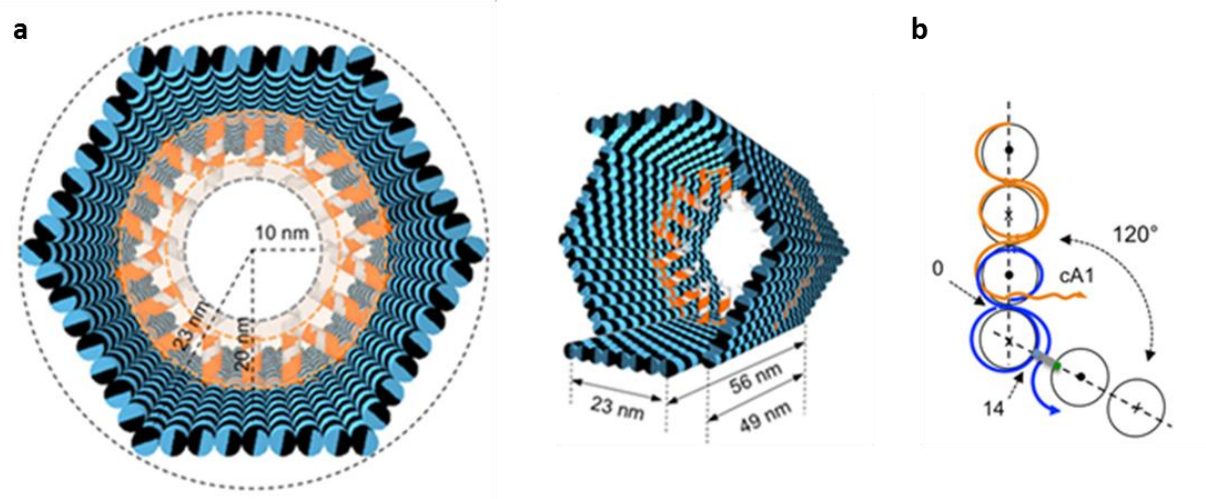


***Figure 11*: (a)** The DNA origami host with its internally orientated PAs (orange) and the complementary catching devices (grey). All PAs are in the center of the hollow structure to guarantee encapsulation inside of the structure. The host is made of six planar faces connected into a hexagonal prism with an edge and outer radius of 23 nm and a free inner room of ca. 10 nm in radius. Two opposite faces are slightly longer (56 nm) than the other four (49 nm). (**b**) Each face is 23 nm wide, connected via spatial out-of-plane crossovers, fixing the faces in an angle of 120°

Biotin-streptavidin experiments at the AFM were also performed to confirm the correct folding of the designed structure. Furthermore, despite the deformation of the structures at the AFM, the dimensions of the DNA nanocontainer were proven by transmission electron microscope (TEM), gel electrophoresis migration and dynamic light scattering (DLS).

(ii) According to the "key-and-lock" principle of enzymes-ligand recognition, the spatial arrangement of the supramolecular ligands of the host has to fit perfectly in the PDZ1 domain of DegP protein. Therefore, a supramolecular ligand that attaches to the PDZ1 domain of the DegP protein with high affinity had to be synthesized. The peptide of sequence aspartic acid (D), proline (P), methionine (M), phenylalanine (F), lysine (K), leucine (L) and valine (V) (DPMFKLV) was chosen because of its affinity with a K_d of 5 µM [48]. An efficient and general synthesis route had to be established to connect the N-terminal amino group with the thiol group of the single stranded DNA (*Fig. 12*). Using solid phase peptide synthesis (SPPS), introduction of a maleimide-function at the N-terminus of the peptide was performed directly on the solid phase. The oligonucleotides, modified with a thiol function at the 5'-terminal phosphate and, when necessary, with a fluorophore at the 3'-terminus, were linked to the

peptide via Michael addition. These DNA-peptide conjugates were characterized by matrix-assisted laser Desorption/Ionization (MALDI) and denaturing acrylamide gel-electrophoresis.

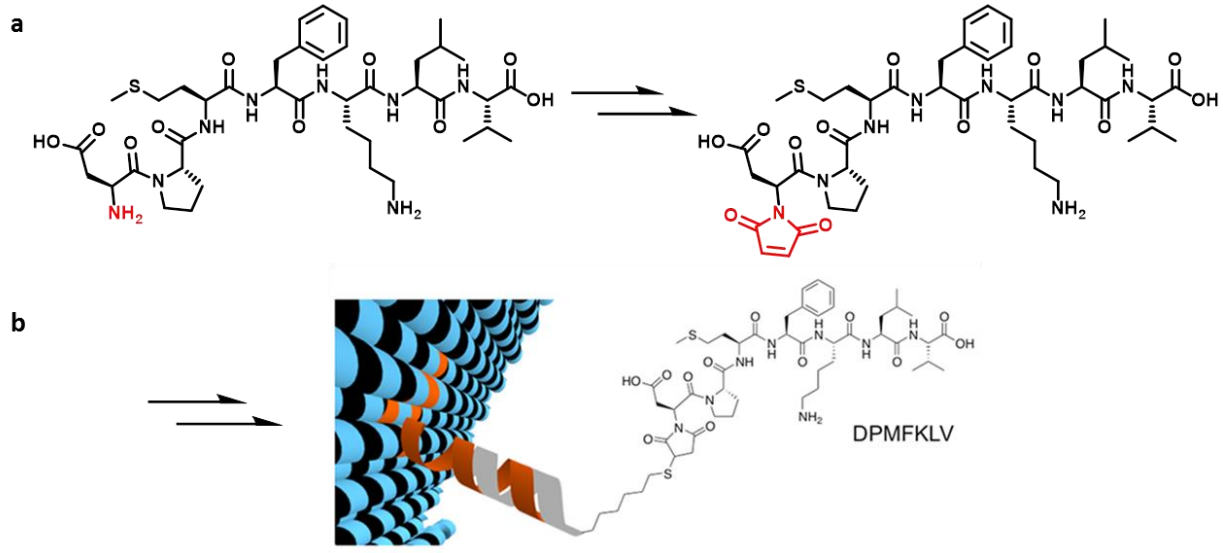


***Figure 12*: (a)** Chosen linker for the connection between the protein and the DNA nanostructure. The peptide sequence DPMFKLV with its C-terminus pointing inwards the protein. Several steps are needed to introduce the maleimide group (red) at the N-terminus of the peptide sequence. This functional group allows the nucleophilic attack of the thiol group of an oligonucleotide. (**b**) In the schematically zoom in. the 16 base-pair long oligonucleotide with its 3′ attached peptide deals as catching device[53].

(iii) Performing loading experiments with streptavidin and DegP should prove the general programmability of the DNA nanocage for protein loading, finally observed by different single-particle microscopy techniques and ensemble gel electrophoresis analysis. Using fluorescently-labeled peptide-DNA conjugates and proteins, the successful loading was proven by the colocalization of the fluorescent signals. To explore the effect of ligands multiplicity, several structures with a different number and orientation of ligands were designed. The effects were shown by comparing the illumination intensity of different bands by electrophoresis. Molecular modelling was performed to partially explain the binding mode between the DegP protein and the nanocages. All structures were characterized via AFM and TEM showing a preference for encapsulation of $DegP_{12}$.

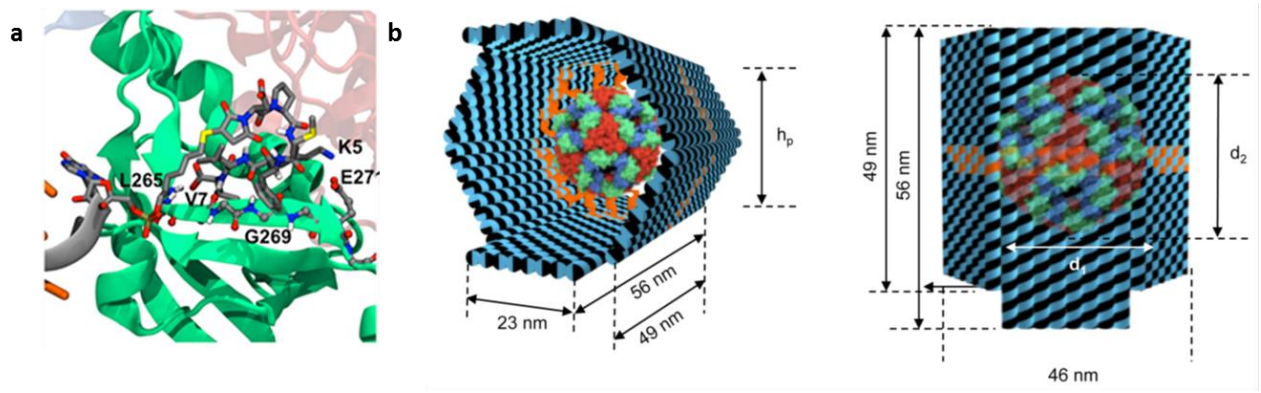


***Figure 13*: (a)** Detailed view of the molecular model showing the interaction between the DPMFKLV ligand and the PDZ1 binding site[53]. (**b**) Schematically representation of the encapsulated protein in its front and top view.

TEM characterization was performed at the MPI in Dortmund by Pascal Lill, currently a member of the newly established group of Dr. C. Gatsogiannis (AG Prof. S. Raunser, MPI Dortmund). Dr. Kenny Bravo-Rodriguez of the AG Prof. Sanchez-Garcia (Univ. Duisburg-Essen) performed the molecular dynamic simulations and geometric modeling studies. Pierre Stegemann and Dr. Melisa Merdanovic (AG Prof. M. Ehrmann, University Duisburg-Essen), provided the proteins. Daniel Gudnason and Dr. Mikayel Aznauryan, members of the AG Prof. V. Birkedal (Aarhus University, Denmark) performed the single molecule fluorescence experiments.

3. Results and Discussion

In the following chapter, the most significant results of the PhD thesis are shown. Additional results can be found in chapter 8.

3.1 DegP

The work, shown in this chapter (3.1) and subsections, was done in cooperation with Pierre Stegemann and Dr. Melisa Merdanovic (AG Ehrmann). Experimental details on DegP mutagenesis and expression can be found in the bachelor thesis of Pierre Stegemann[54].
Three different types of proteins were used (*Table 1*): (i) first experiments were performed with several inactive oligomers (DegP-SA) labeled at exposed lysines with Alexa488 fluorophores, (ii) unlabeled, inactive proteins and (iii) fluorescently-labeled, inactive and Cys-mutated DegP protein in the 6-mer form (chapter 3.1.1 and 3.1.2).

***Table 1*:** Detailed overview on the used proteins.

Oligomer DegP	Labeling (Alexa)	Location of label	Wild Type	SA inactive mutant	Shortcut
6-mer	-	-	-	+	$DegP_6SA$
12-mer	-	-	-	+	$DegP_{12}SA$
12/24-mer	-	-	-	+	$DegP_{12/24}SA$
24-mer	-	-	-	+	$DegP_{24}SA$
6-mer	-	-	+	-	$DegP_6WT$
6-mer	488	Lys at surface	+	-	$DegP_6^{A488}WT$
12/24-mer	488	Lys at surface	-	+	$DegP_{12/24}^{A488}SA$
6-mer	633	Cysteine (*Fig. 14*)	-	+	$DegP_6^{A633}SA$
6-mer	647	Cysteine (*Fig. 14*)	-	+	$DegP_6^{A647}SA$
12-mer	647	Cysteine (*Fig. 14*)	-	+	$DegP_{12}^{A647}SA$

3.1.1 Mutagenesis and expression of DegP

All plasmids were derivatives of pCS20 expressing wild-type *DegP* with a C-terminal His tag, verified by DNA sequencing. Point mutants were constructed by oligonucleotide-directed mutagenesis according to standard methods. The mutagenesis of C57A, C69A and S210A

was performed by single site mutagenesis. Substitution of the serine at position 210 by an alanine generated the catalytic inactive mutant Δss_3xCys_SA. Native cysteine residues in the LA-Loop (C57 and C69) were additionally mutated to avoid mislabeling. A schematic representation of all mutations is given in Figure 14.

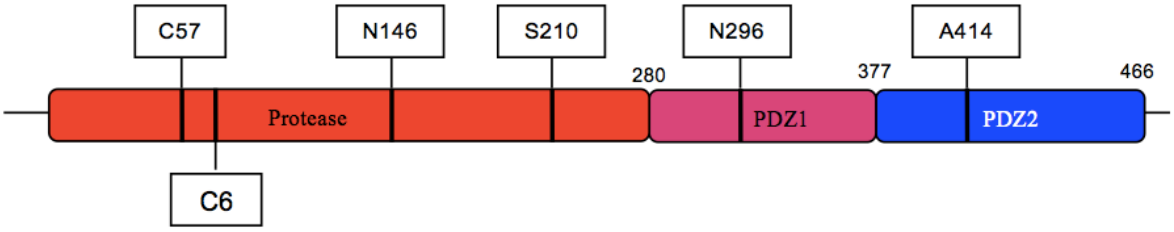


***Figure 14*:** Scheme of all mutations in DegP_Δss_3xCys_SA[53].

The weakly conserved amino acids N146, N296 and A414 were selected for mutation to cysteines to enable labeling with maleimide-activated fluorophores; due to their location on the surface of the 6-mer and 24-mer. The selected amino acids are located in distinct domains (*Fig. 15*) and are about 10 nm apart thus preventing formation of a disulfide bond. Using a Change-IT multiple mutation sites directed mutagenesis kit, the mutations N146C, N296C and A414C were introduced. MA001 cells were used to express the mutated protein.

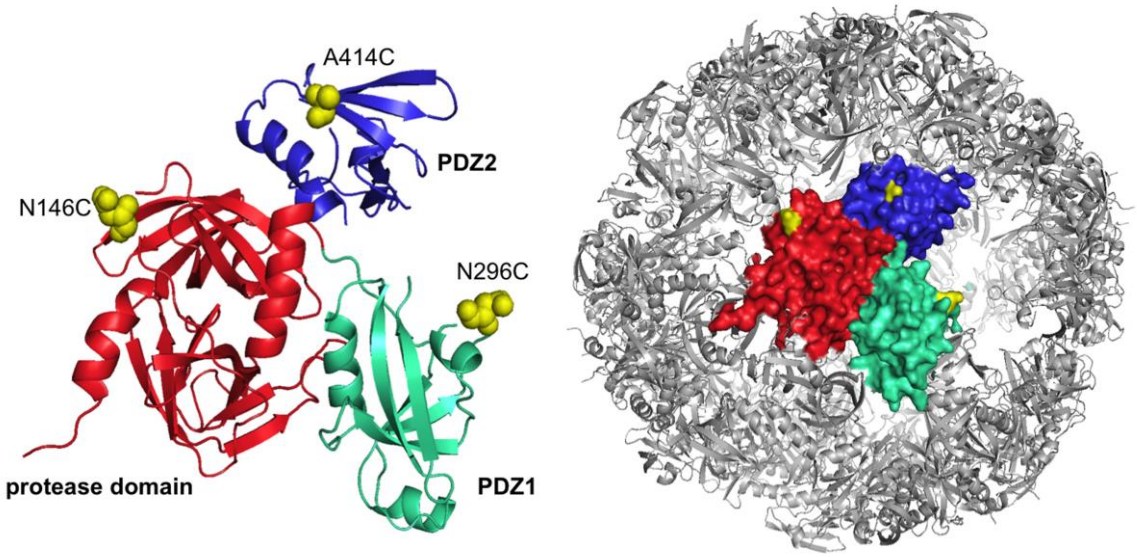


***Figure 15*:** Representation of the genetically modified DegP. For further functionalization with a maleimide-activated fluorophore, one residue of each domain was mutated into a cysteine (yellow). To avoid interaction with the PDZ-1 domain, the selected residues are located on the protein surface, being exposed to the solvent and far away from the peptide-binding domain[53].

3.1.2 Protein labelling

Purified samples of DegP Δss_3xCys_SA were labeled with DyLight™ 488 Maleimide, DyLight™ 633 Maleimide or DyLight™ 647 Maleimide (ThermoFisher). The DegP-SA (i.e. mutated only at the catalytic triad) was instead modified at the lysine residues with Alexa 488 N-hydroxysuccinimide ester. Excess labelling reagent was removed using dye removal columns[54]. Although the chemical formulas of Dylight are not available, their spectral properties are similar to the corresponding Alexa-type dyes. For this reason, all fluorophores will be indicated in the present thesis by an "A" followed by the associated excitation wavelength.
Various combination of unlabeled or fluorescently labeled DegP proteins in different oligomerization states and genetic forms were screened for their binding to the DNA cage, in order to estimate the role of different factors on the encapsulation properties of the guest (a more detailed discussion on this issue and corresponding gel results are provided in chapter 3.3). All protein forms used in this work are summarized in table 1 (chapter 3.1).

3.2 Host

Before the first experimental results are shown, the design and the basic ideas leading to the DNA origami construct will be presented.

3.2.1 Design strategies

To encapsulate a highly symmetric protein, which has in its 24-mer form a 4-3-2 symmetry and in all oligomeric states repetitive accessible PDZ1 domains for ligand binding, a large hexagonal shape with a quasi-radial distribution of peptide ligands within the cavity being compatible with the symmetry of the PDZ1 domains was designed (*Fig 16*). The size of the inner cavity of the DNA host is determined by the largest oligomeric form, with a radius of 9.5 nm. Modifying the inner surface with accessible attachment points for ligands, bearing the opportunity of single strand displacement, results in a double-stranded DNA corona of about 7.4 nm, due to the 22 bases long protruding arms leaving the plane of the faces orthogonally (*Fig. 16* ochre area). The DNA-peptide conjugate (detailed description in chapter 3.3 and chapter 6.5.2.8) is in total 8.4 nm long, including 5.4 nm for the 16 bases being

complementary to the last 16 bases of the protruding arms. The full hepta-peptide and the C-6 spacer connecting the DNA to the peptide are 3 nm long. Molecular modelling (see Chapter 3.4.1) shows that the last three amino acid residues interact with the PDZ1 domains of DegP (*Fig. 16* green area). In total, the host-to-guest bridge is about 10-11 nm long, leaving a free inner room of ca. 10 nm in radius. Respecting those requirements, a hexagonal shape with an edge length of 23 nm was designed. In this thesis, the M13mp18 scaffold was selected, resulting in a length of the channel of 49 nm, except for two opposite faces with slightly longer edges (56 nm).

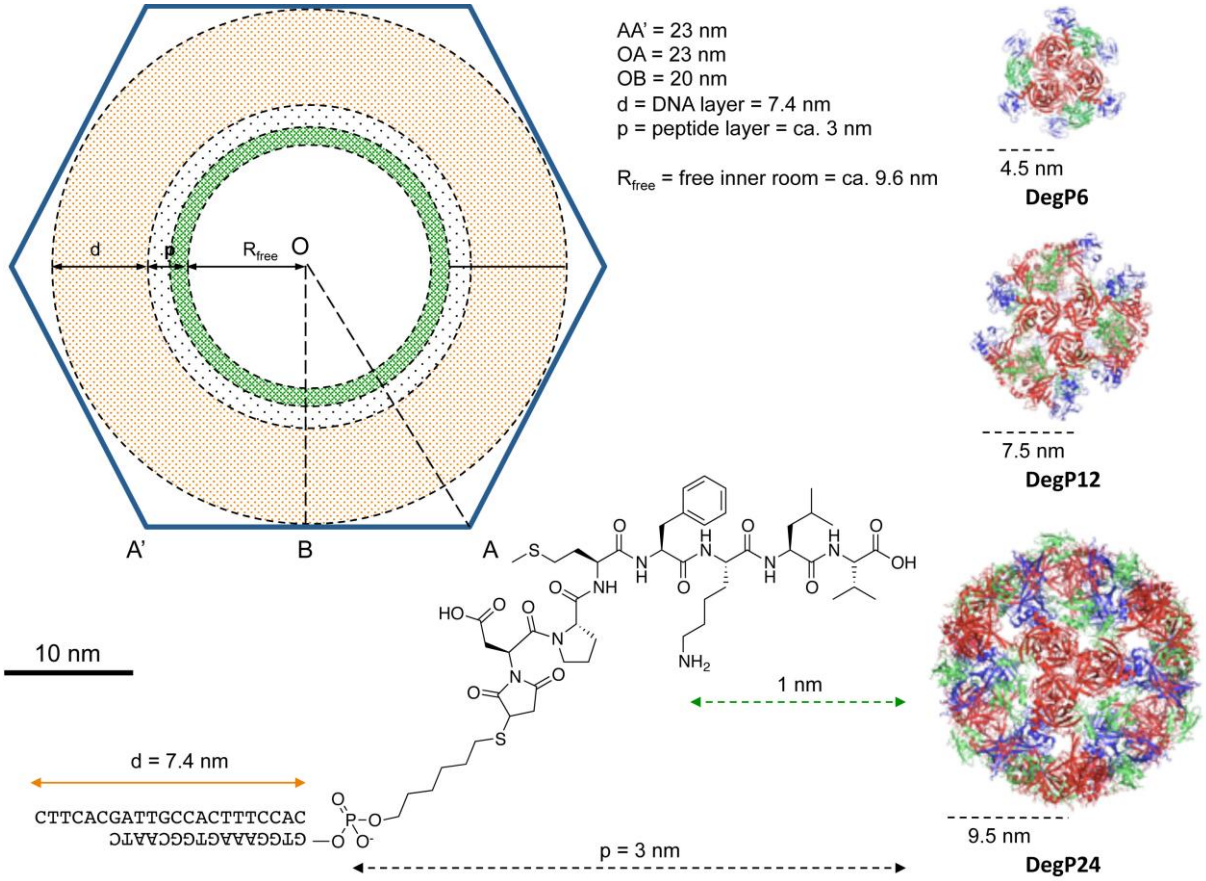


***Figure 16*:** Schematic geometric model of the idealized DNA host. The 22 bases long PAs are orientated inside the cavity for further hybridization with the DNA-peptide conjugates. Only the last three amino acid residues (Lys-Leu-Val) interact with the PDZ1 domains of DegP illustrated as green area with a length of 1 nm. The whole peptide and the C-6 spacer are around 3 nm long and are connected to the double-stranded DNA corona that has a length of 7.4 nm. In total, the linker has a length of 10-11 nm, leaving an inner room of around 10 nm in radius.

All designed DNA constructs that were used in this thesis are shown in Figure 17. The 6p-nc design lacks out-of-plane connecting staples and thus represents a convenient reference to control the correct design of the faces of the structure. To get a regularly hexagonal shape out of six faces, the faces have to be connected; (i) at 120°, leading to the $6p^{120}$ with possible protruding arms being orientated inside the structure, (ii) at 240°, leading to the $6p^{240}$ with

possible protruding arms being orientated outside the structure and (iii) at 180°, leading to the $6p^{180}$ design allowing the faces to orientate themselves such to reduce mechanical stress of the structure. Each design was modified with 0, 6 or 18 protruding arms indicated as 0cA1, 6cA1 and 18cA1. The increase in the number of the protruding arms leads to an increase of the local concentration of the ligand inside the cavity. The designs without protruding arms were used as control.

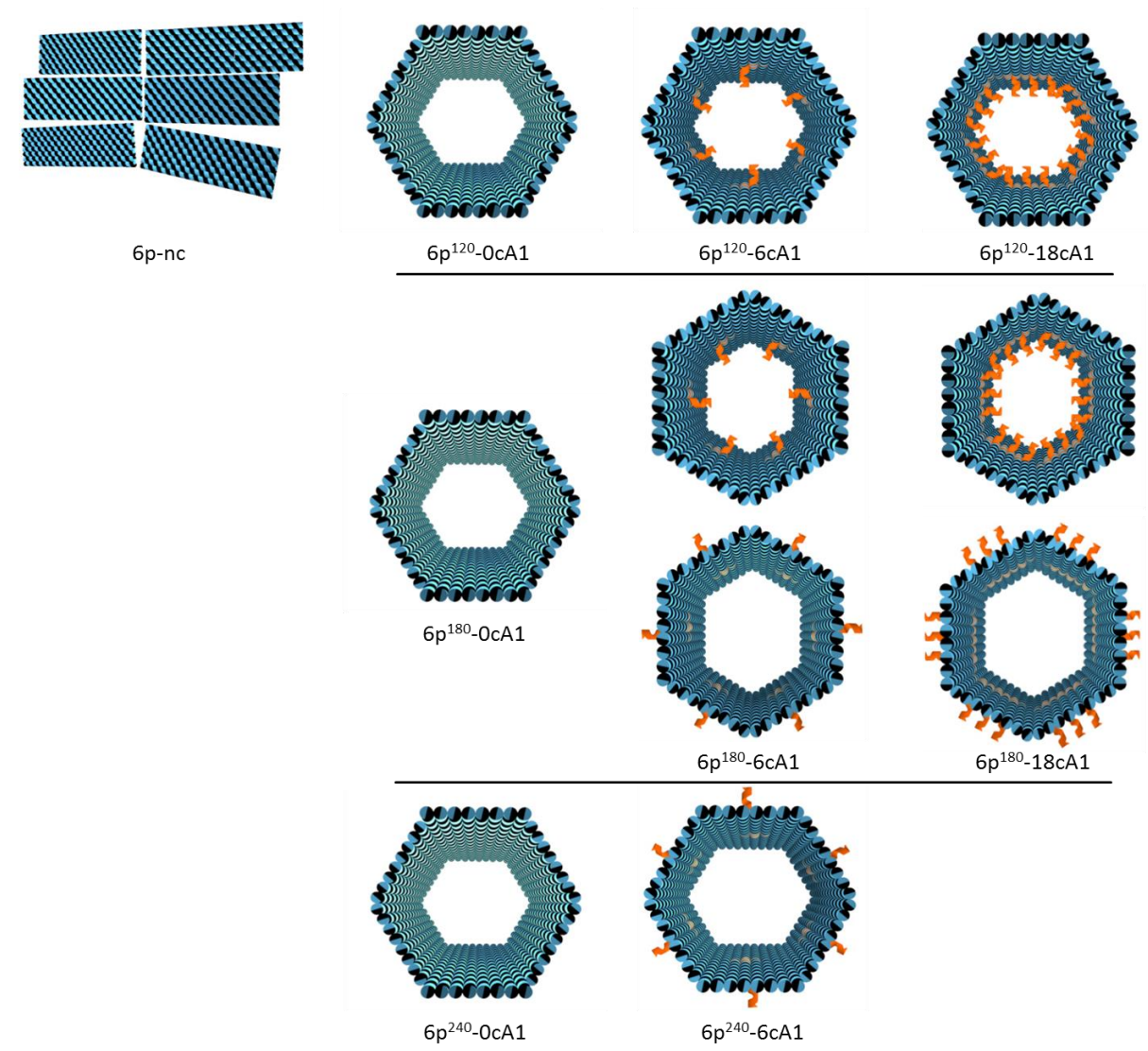


***Figure 17*:** All DNA constructs used in this PhD thesis. The basic design in its open form without PAs and spatial out-of-plane face-to-face connecting crossovers (6p-nc). A 6p host lacking the PAs for DNA-peptide conjugate anchoring will be indicated as 6p-0cA1. The other designs differ in their orientation of the PAs, resulting from special designed face-to-face connection staples. A more detailed illustration is shown in *Figure 18*. The hexagonal prism (6p) has been designed with PAs orientated towards the inner cavity ($6p^{120}$), or outwards ($6p^{240}$). Connecting the faces within the plane of the faces, by flexible T-hinges, ($6p^{180}$) leads to an arbitrary orientation of the PAs illustrated by the subordination of both possible orientations to the $6p^{180}$ nomenclature. The number of PAs is indicated as 6cA1 or 18cA1.

With the aim to create DNA hosts whose inner cavity is functionalized for encapsulation purposes, complete control of the face orientation is necessary (*Fig. 18*). To achieve this, each design ($6p^{120}$, $6p^{180}$ and $6p^{240}$) required a special set of spatial out-of-plane face-to-face

crossovers. Referring to the elongated staples cA1 (orange), which are used as protruding arms and defining this side as the inside of the structure; the oligo of the last helix of the upper domain is pointing into the screen.

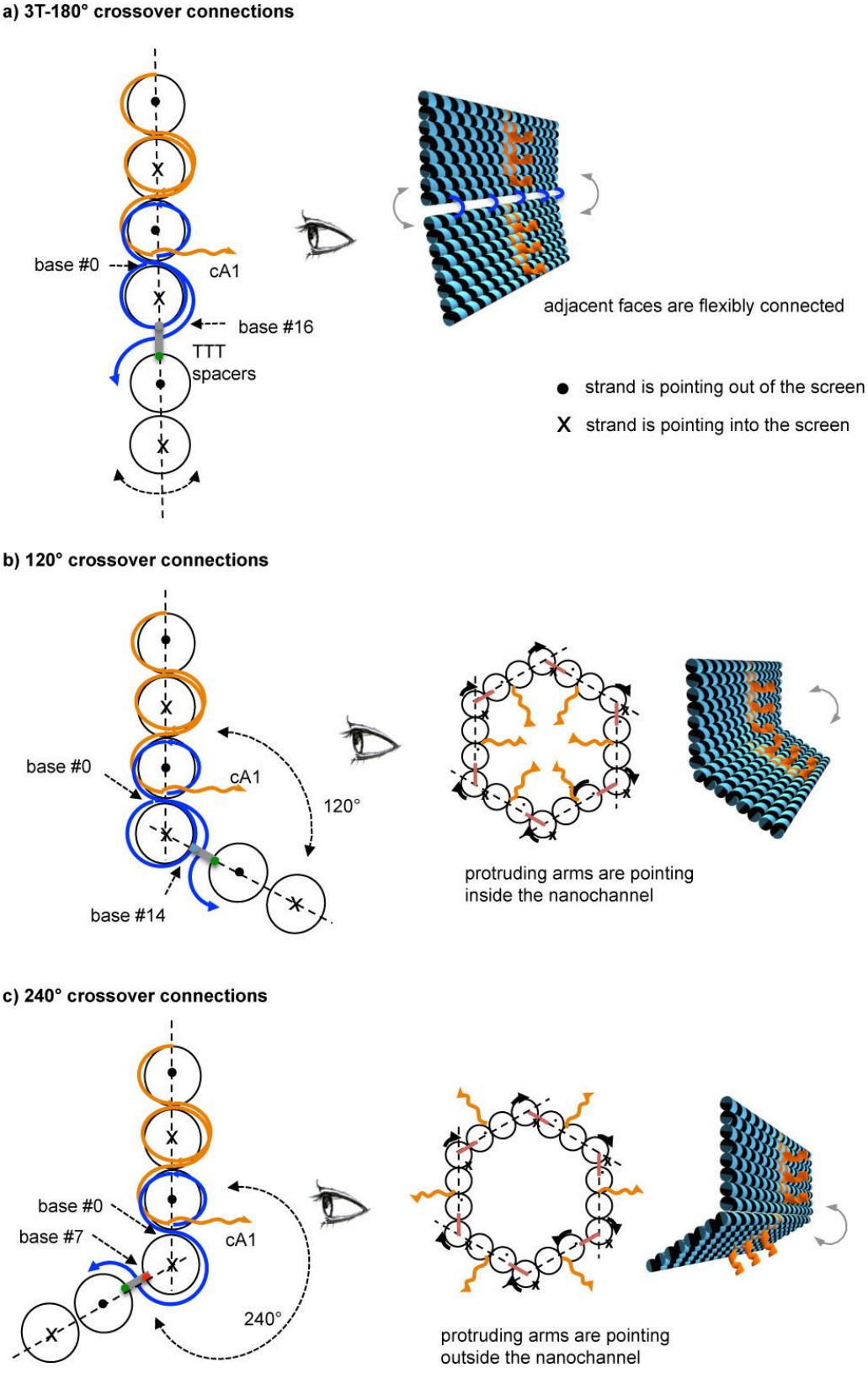


***Figure 18*:** Schematical representation of the different design strategies of the connection of the adjacent faces of the DNA nanocage. (**a**) As in common two dimensional designs, the crossovers between the adjacent faces are placed every 16 base-pairs. Including 3T spacers as "virtual" crossovers allow sufficient orientational freedom of two adjacent faces. (**b**) Placing

the crossover every 14 base-pairs, results in a relative orientation of 120° between the adjacent faces; and (**c**) placing the crossovers every 7 base-pairs, leads to a 240° angle between the faces. Elongating the staples (cA1) at the 3′- or 5′-terminus will direct the staple inwards or outwards of the cavity, allowing modifications at the inner or outer surface of the DNA nanocage.

The orientation of the adjacent helix on the next DNA domain is reverse. Illustrating the three dimensional out-of-plane oligonucleotides in blue, the first set of staples (*Fig. 18b*) connect the adjacent helices after 14 bases, start counting at base 0 of the lowest helix of the upper domain. Pointing into the screen, the three dimensional staple turns right-handed around the helix, being at the position of base 0 again after 10.5 bases, requiring 3.5 more bases to reach the crossover to the adjacent helix. This results in a 120° connection between the two adjacent faces. Applying the same design strategy, but connecting the helices directly after 7 bases (*Fig. 18c*) results in a 240° angle between the two adjacent faces. Contrarily, building a crossover after 16 bases and adding three thymines as a spacer between the two connected helices, results in a crossover lying theoretically within the plane of the two faces. During assembly, the orientation of the faces is stochastically distributed due to the flexible connection.

3.2.2 Experimental results

Figure 19 shows the agarose gel electrophoresis of the 6p complex without added magnesium ions (*Fig. 19*; lane 1) and with increasing magnesium concentration (*Fig. 19*; lane 2-11).

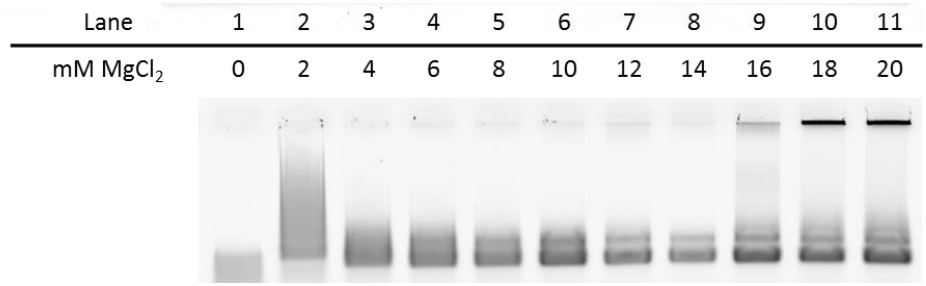


***Figure 19*:** Magnesium screening of the 6p construct of its closed and open form in its (**a**) original form and in its fitted form (b). In the following PhD thesis all gels will be shown in its reduced and fitted form like above. (**b**) The increase of the Magnesium concentration does not show a large effect on the correct assembly of the tube.

Unless specified differently, only most relevant gel bands will be shown in this work, omitting uninformative or repetitive bands, which do not bring additional information. This magnesium screening indicates just small differences between the different concentrations in a range of 8 to 14 mM. Considering the DNA origami paper of Paul Rothemund[14] the magnesium concentration of 12.5 mM was chosen. To prove the general loading capability and the correct folding of the different designs, several ligands were used to target different proteins (*table 2*).

***Table 2*:** List of DNA-ligands and corresponding targets used to proof the loading ability (A1-Flc) and the correct folding of the 6p constructs.

Ligand	A1-biotin	A1-Flc	A1-Flc
Target	STV	anti-Flc^{CF594}	anti-Flc^{CF5640R}

As a first binding study, the 6p^{120}-18cA1 cage functionalized with A1-Flc ligands complementary to the PAs was loaded with anti-fluorescein antibodies to show the general binding capability of the DNA-origami cage. Figure 20 shows successful binding by a significant gel shift of the DNA-origami-protein complex.

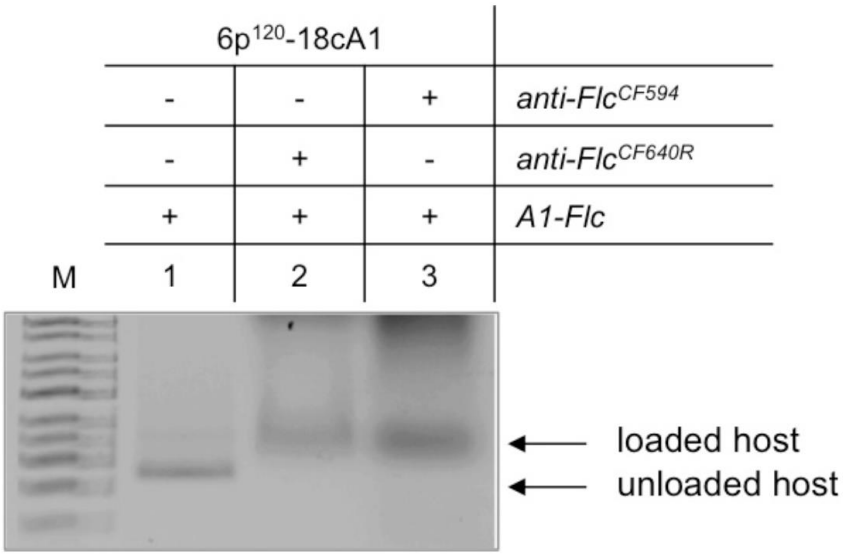


***Figure 20*:** Binding of the anti-fluorescein antibodies to the 6prism construct functionalized with 18-Flc-labeled A1-strands in its inner cavity shows a significant retardation in the electrophoretic mobility of the complex.

To verify the correct folding of the different designs, an agarose gel electrophoresis of streptavidin binding to biotinylated 6p nanocages with different connecting angles was performed (*Fig. 21*).

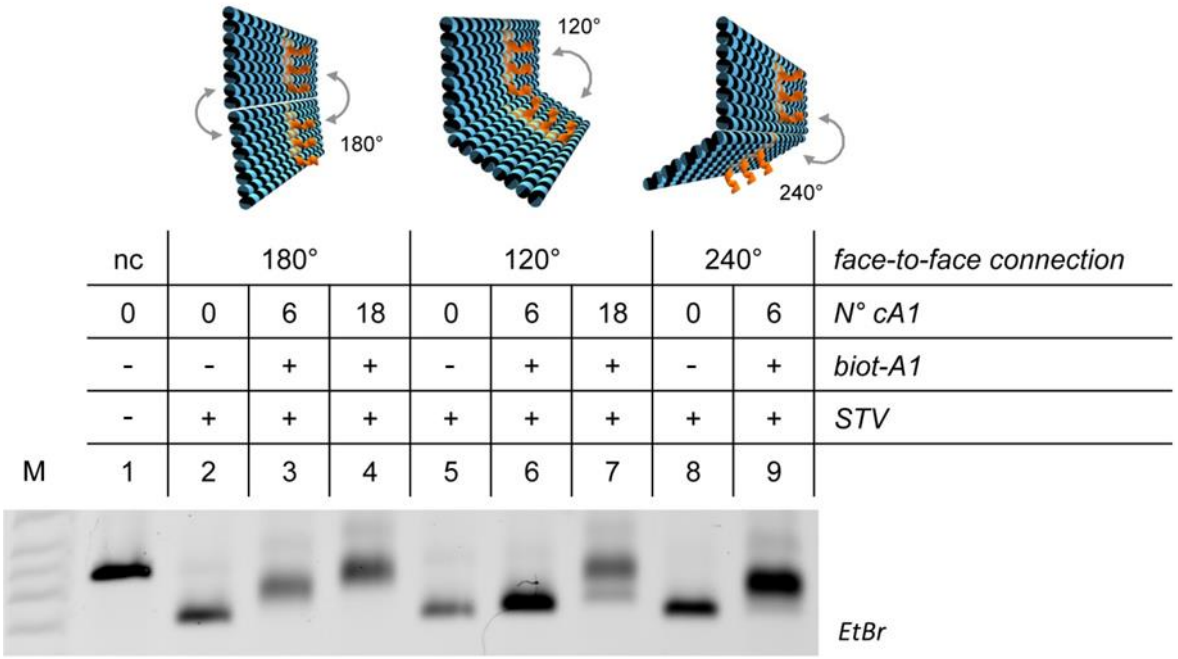


***Figure 21*:** Agarose gel electrophoresis characterization of streptavidin binding to biotinylated 6p nanocages with different connection angles of the faces. All 6p nanocages with protruding arms (lanes 3 and 4 for the $6p^{180}$ design; lanes 6 and 7 for the $6p^{120}$; and lane 9 for the $6p^{240}$ design) were modified with biotinylated DNA strands (biot-A1) complementary to the PAs (cA1). The corresponding DNA cages without PAs (Lane 2 for the $6p^{180}$-0cA1; Lane 5 for the $6p^{120}$-0cA1; and Lane 8 for the $6p^{240}$-0cA1) are used for comparison of the migration rates. Lane 1 shows the open structure with lacking face-to-face connections (nc: non-connected). Except to the open structure, streptavidin was added.

All 6p nanocages with protruding arms (lanes 3 and 4 for the $6p^{180}$ design; lanes 6 and 7 for the $6p^{120}$; and lane 9 for the $6p^{240}$ design) were modified with biotinylated DNA strands (biot-A1) complementary to the PAs (cA1). The corresponding DNA cages without PAs (lane 2 for the $6p^{180}$-0cA1; lane 5 for the $6p^{120}$-0cA1; and lane 8 for the $6p^{240}$-0cA1) are used for comparison of the migration rates. Due to the lower compactness of the open structure, caused by the lacking face-to-face connections (nc: non-connected) this structure migrated with a slower rate (lane 1) in comparison to the analog closed forms. Increasing the number of PAs within each design leads to a decrease in the migration rate of the DNA-streptavidin complexes (lane 2-4 for the $6p^{180}$ design; lane 5-7 for the $6p^{120}$ design and lanes 8 and 9 for the $6p^{240}$ design). Interesting conclusions can be drawn by comparing the DNA cages with identical number of PAs (lanes 3, 6 and 9, for the $6p^{180}$-, $6p^{120}$- and $6p^{240}$-design). The $6p^{120}$-6cA1-STV complex (lane 6) migrates faster than the corresponding $6p^{180}$-6cA1-STV and $6p^{240}$-6cA1-STV complexes (lanes 3 and 9), indicating that a convergent design of ligands effectively allows internalization of streptavidin with minimal effect on the surface charge of the host, thus resulting in a compact structure with migration properties similar to the unloaded nanocages. In the $6p^{240}$-design, the streptavidin molecules are placed outwards, covering the surface of the host and resulting in a larger complex with a slower migration rate. The stochastic orientation of the PAs in the $6p^{180}$-design yields instead a less defined band of lower intensity. To demonstrate the correct folding of the design, the seven comparable designs 6p-nc, $6p^{180}$-0cA1, $6p^{180}$-6cA1, $6p^{120}$-0cA1, $6p^{120}$-6cA1, 6p240-0cA1 and $6p^{240}$-6cA1 were assembled and analyzed at the AFM upon addition of ten-fold excess streptavidin. In presence of biotinylated protruding arms, streptavidin should bind almost quantitatively, resulting in increase of the height profile of the structure.

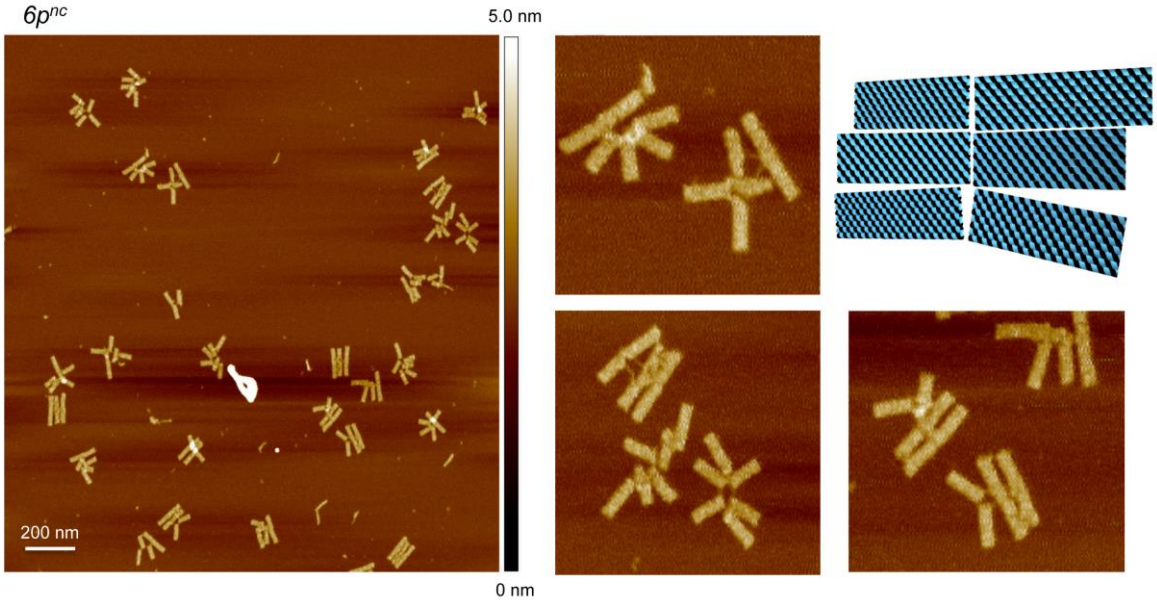


***Figure 22*:** AFM image of the 6p-nc design lacking of face-to-face connections.

First, the 6p-nc design was investigated proving the correct assembly of the six faces with their different dimensions. As shown in the enlarged panels, the correct folding of the six faces was successful; also the two longer faces were detectable. Due to the absence of protruding arms, no streptavidin could bind to the structure. The height of about 2 nm matched with a monolayer planar DNA structure. Correct assembly of the closed $6p^{180}$-0cA1 design is shown in Figure 23. Due to the high attractive forces during the AFM measurements, the structures collapsed onto the surface. The height of 4 nm fits with two monolayer DNA origami shapes, lying one on top of the other. This design also lacked biotinylated staples, showing no interaction with streptavidin.

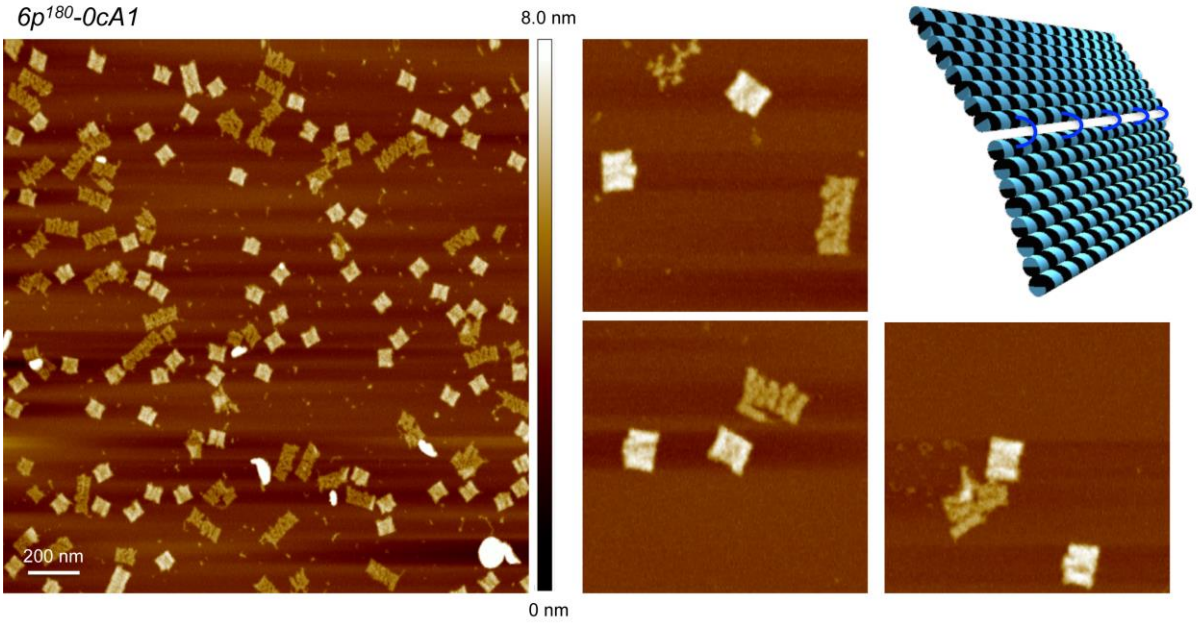


***Figure 23*:** AFM image of the 6p^{180}-0cA1design. Due to the lacking PAs, no interaction of the structure with STV is possible.

Successful binding of streptavidin to the 6p^{180}-6cA1 design is shown in Figure 24, where an increase in the height profile of the structure indicated the addition of streptavidin. The two different orientations of the inner surface, pointing to the center of the structure or from the center of the structure can be verified by the different resolutions of the bound protein. In the case of the biotinylated protruding arms pointing outwards, the bound streptavidin is easily detectable by AFM, resulting in sharp and bright spots forming a line in the center of the structure. Due to the ca. ten-fold excess of STV, all biotinylated protruding arms are saturated. This leads to three detectable bright and sharp spots (STV), at a well-defined distance one in respect to the other, on the surface of the DNA origami (*Fig. 24*; upper zoom-in). Encapsulation of streptavidin would therefore lead to a less evident increase of the height profile, lacking in bright and sharp spots. The deformation of the structure would exert a force on the STV-PA-complex, resulting in a bent protruding arm and in an undefined distance between the protein and the DNA cage.

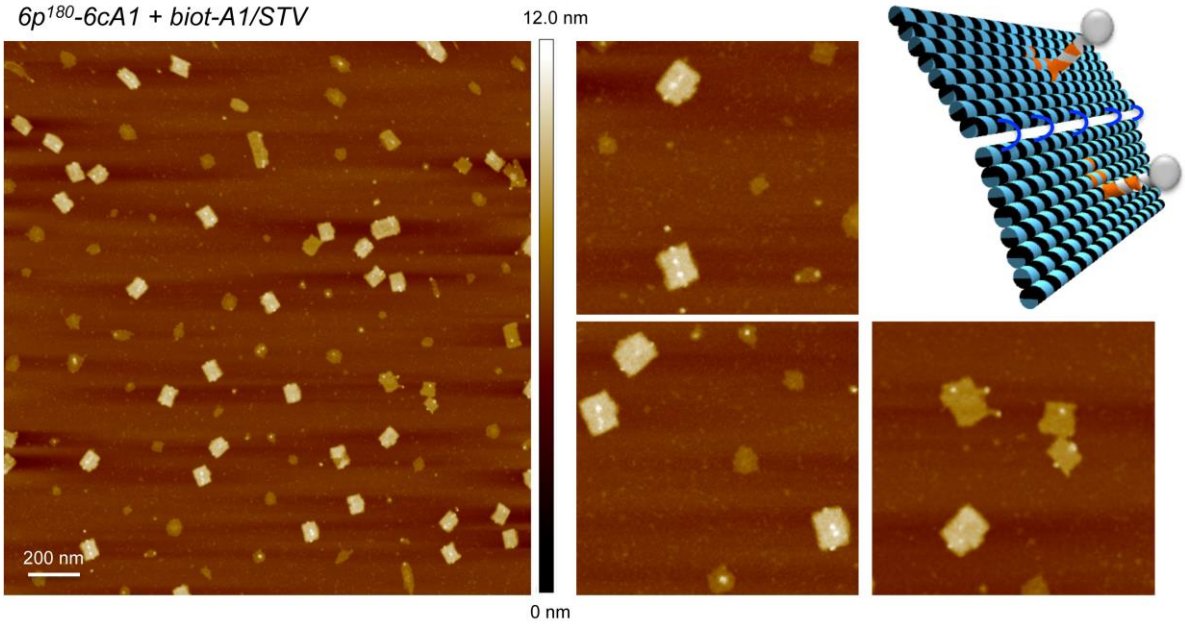


***Figure 24*:** AFM image of the 6p[180]-6cA1 design with biotinylated arms after addition of STV in solution. The two different orientations of the PAs within the two possible formations of the 6p[180]-6cA1 design can be shown by STV molecules bound to the structure. If the PAs point outwards, the STV molecules that bound to the structure can be detected as bright and sharp spots on top of them forming a line in the middle center of the structure (zoom in left from the model). Encapsulation (PAs pointing inwards) of the streptavidin and collapsing of the structure at the mica lead to deformation of the structure in its central region with a less evident increase of the height profile (zoom in in the lower line).

Due to the lack of biotinylated protruding arms, no interaction between the 6p[120]-0cA1 and streptavidin added in solution could be observed (*Figure 25*).

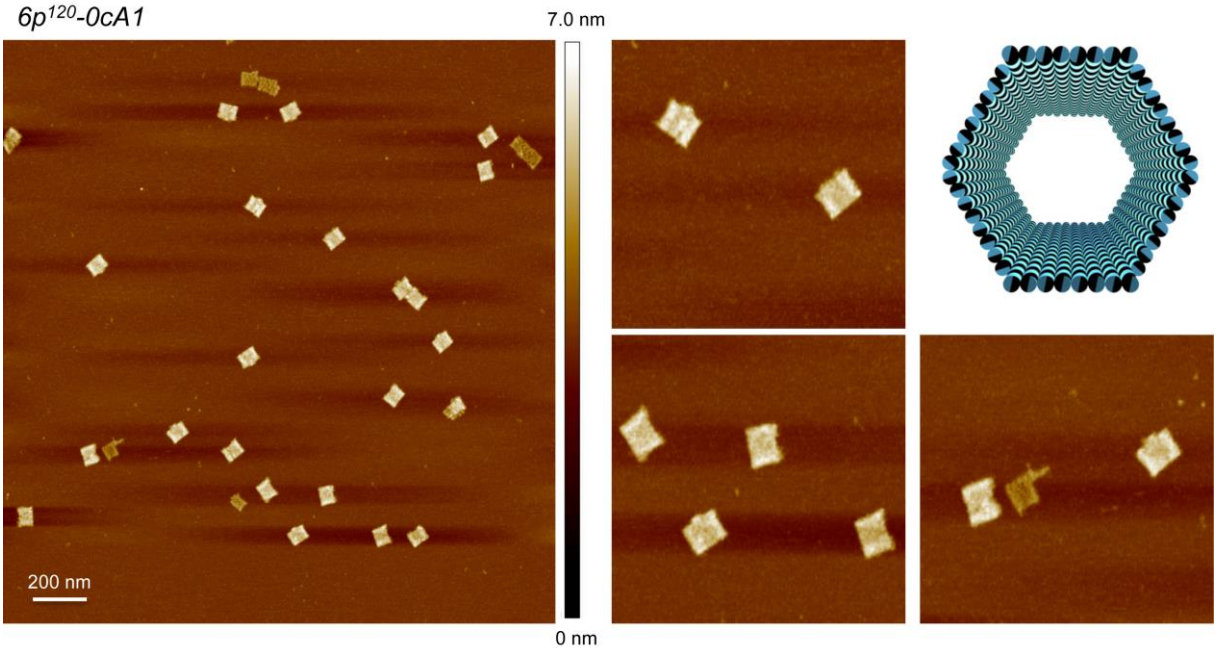


***Figure 25*:** AFM image of the 6p[120]-0cA1design. Due to the lacking PAs, no interaction of the structure with STV is possible.

AFM imaging of the 6p^{120}-6cA1 proved correct formation of the structure, encapsulating streptavidin exclusively inside the cavity (*Fig. 26*). Again the increase of the height profile in the central region without giving sharp bright spots suggests successful internalization of streptavidin.

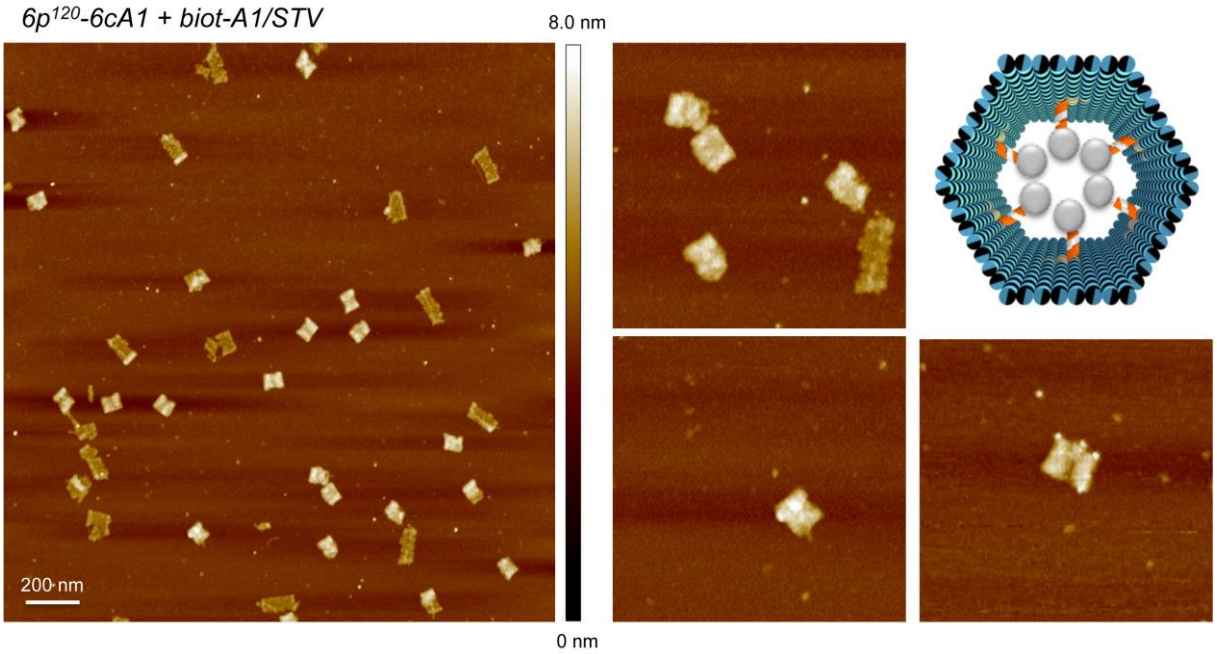


***Figure 26*:** AFM image of the 6p^{120}-6cA1 design with biotinylated arms after addition of STV in solution. Encapsulation of the streptavidin can be shown by the increase of the height profile in the central region, but lacking in a resolution.

No interaction took place between the structures of the 6p^{240}-0cA1 design (*Fig. 27*) and the added streptavidin.

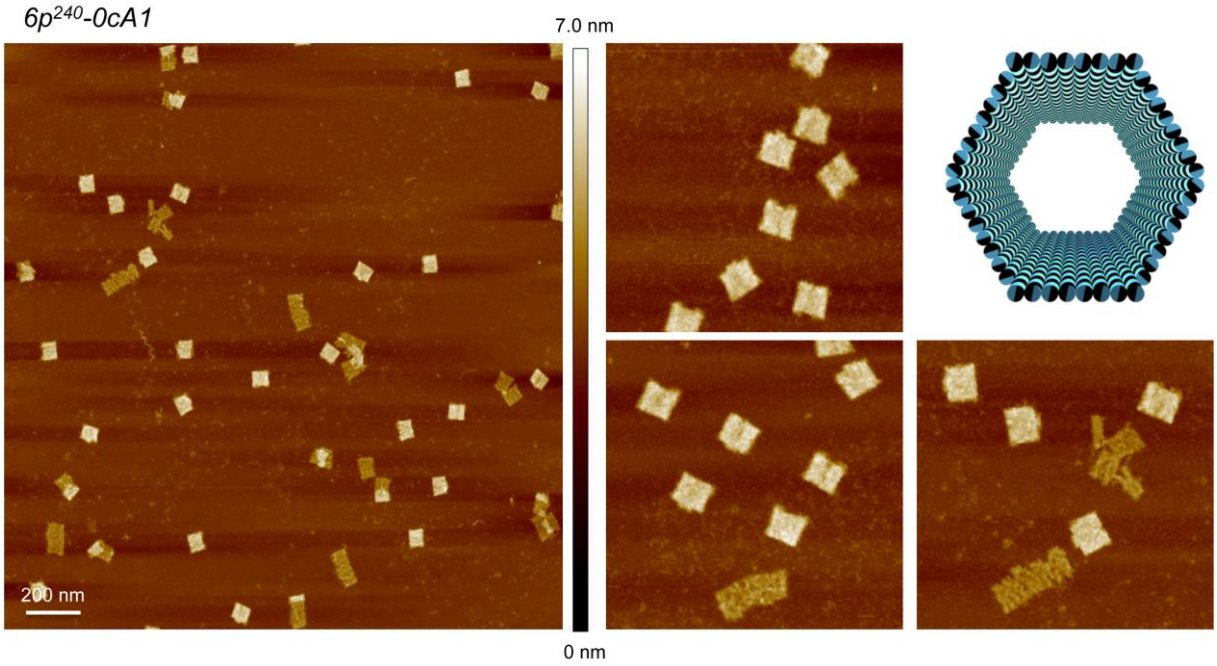


***Figure 27*:** AFM image of the 6p^{240}-0cA1design. Due to the lacking PAs, no interaction of the structure with STV is possible.

In most cases, three aligned bright and sharp spots in the central region of the 6p^{240}-6cA1 could be visualized by AFM (*Fig. 28*), indicating the successful binding of streptavidin to the external surface of the host structure.

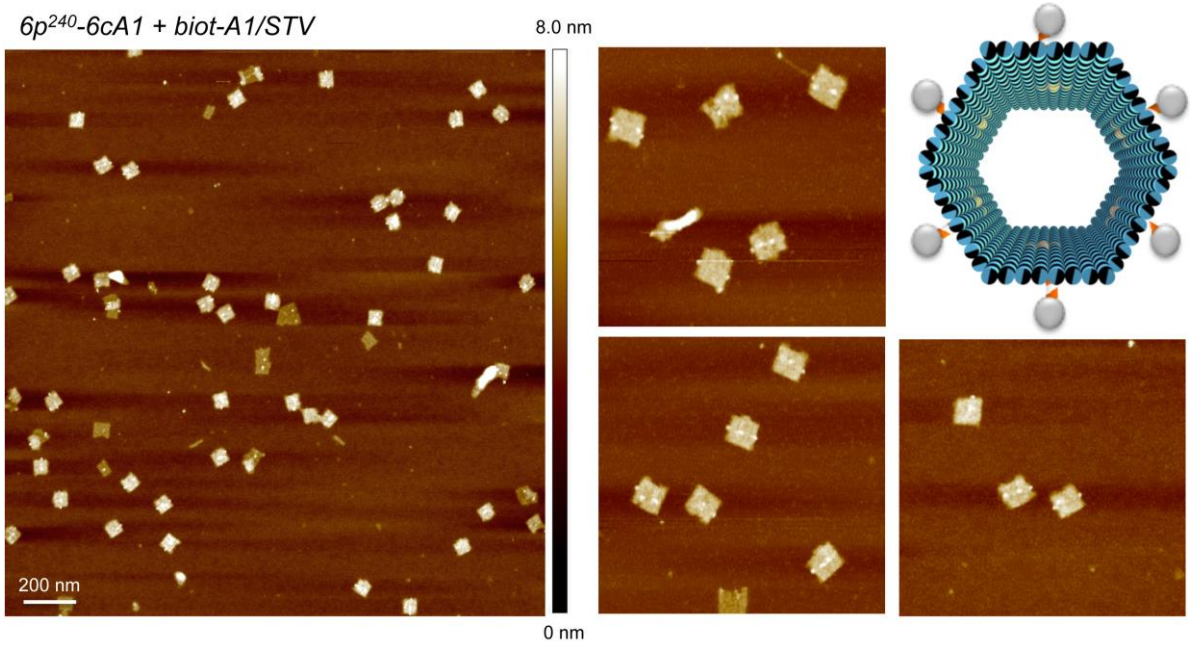
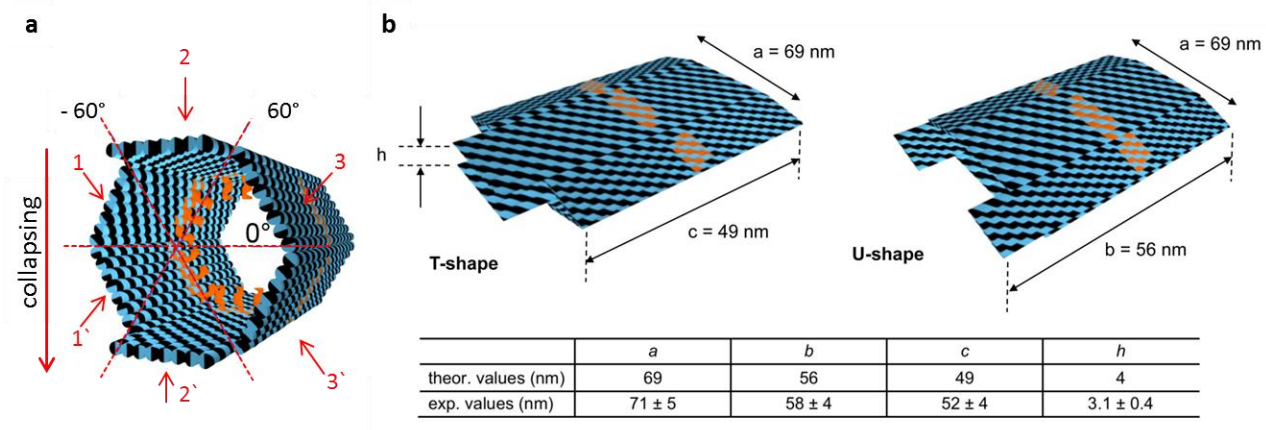


***Figure 28*:** AFM image of the 6p^{240}-6cA1 design with biotinylated arms after addition of STV in solution. The three bright and sharp spots in the central of the structure forming a line show the addition of streptavidin to the structure.

Due to the high mechanical forces exerted during the AFM measurements and the drying procedure used for imaging, the DNA nanocages collapsed onto the surface. The two longer opposite faces of the DNA origami were used as topographical markers to verify the intactness of the structure (*Fig. 29*).



	a	*b*	*c*	*h*
theor. values (nm)	69	56	49	4
exp. values (nm)	71 ± 5	58 ± 4	52 ± 4	3.1 ± 0.4

***Figure 29*:** (**a**) Schematically illustration of the orientation of the longer (56 nm) face of the hexagonal tube. The longer face can be at position 1 or 3, resulting in a collapsed U-shape. If the longer shape is at position 2, the collapsed structure will form a T-shape. (**b**) Analysis of the dimensions of the unloaded 6p construct by AFM imaging in air. Collapsing of the structure under AFM imaging leads to two possible flat shapes, depending on the axis along which compression is acting: a T-shape (35%) and a U-shape (65%). This corresponds to the theoretical distribution of 1/3 T-shape and 2/3 U-shape due to 2 out of 3 axes may lead to this configuration. In total 703 structures were analyzed.

Three different collapsing scenarios are possible: (i) bending of the structure along the horizontal symmetry axis indicated as 0° forming a double layer T shape. The two longer shapes are therefore situated at position 2 and 2´ (*Fig. 29b*). (ii) Rotating the structure clockwise and collapsing along the -60° axis results in a U shape formed by the upper longer shape being situated at position 3 and the lower one at position 1´. (iii) Rotating the structure anticlockwise and collapsing along the +60° axis results in a U shape formed by the upper longer shape being situated at position 1 and the lower one at position 3´. The theoretical distribution of 1/3 T-shape and 2/3 U-shape, due to the fact that 2 out of 3 axes lead to the same configuration, corresponds to the experimental AFM distribution of the measured collapsed structures in air: a T-shape (35%) and a U-shape (65%). Although all observed structures at the AFM were collapsed, the total yield of intact structures for the three different 6p designs was at least 82.7 % (*table 3*). To evaluate whether the AFM is responsible for the damage of the DNA nanocages, TEM characterization was performed (*Fig. 30).* The yield of the intact structures at the TEM is quantitative, supporting the hypothesis of an AFM-induced deformation of the structures. A representative class average is shown in Figure 29a.

***Table 3*:** Yields of the correct folded different 6p designs, by AFM imaging.

	defect	intact	total	yield
$6p^{120}$-0cA1	249	3079	3328	92.5 %
$6p^{180}$-0cA1	631	3016	3647	82.7 %
$6p^{240}$-0cA1	235	4412	4647	94.9 %

The fitting dimensions of the simulated 3D models with their corresponding TEM images (*Fig. 30b*) denoted that the DNA nanocage does not collapse at the TEM grid. This shows that the damage of the structures takes place, probably completely, during the AFM measurement.
To prove that the structures are well-formed in solution, dynamic light scattering was performed resulting in a good agreement between the theoretical dimensions and the experimental values (*Fig. 31*).

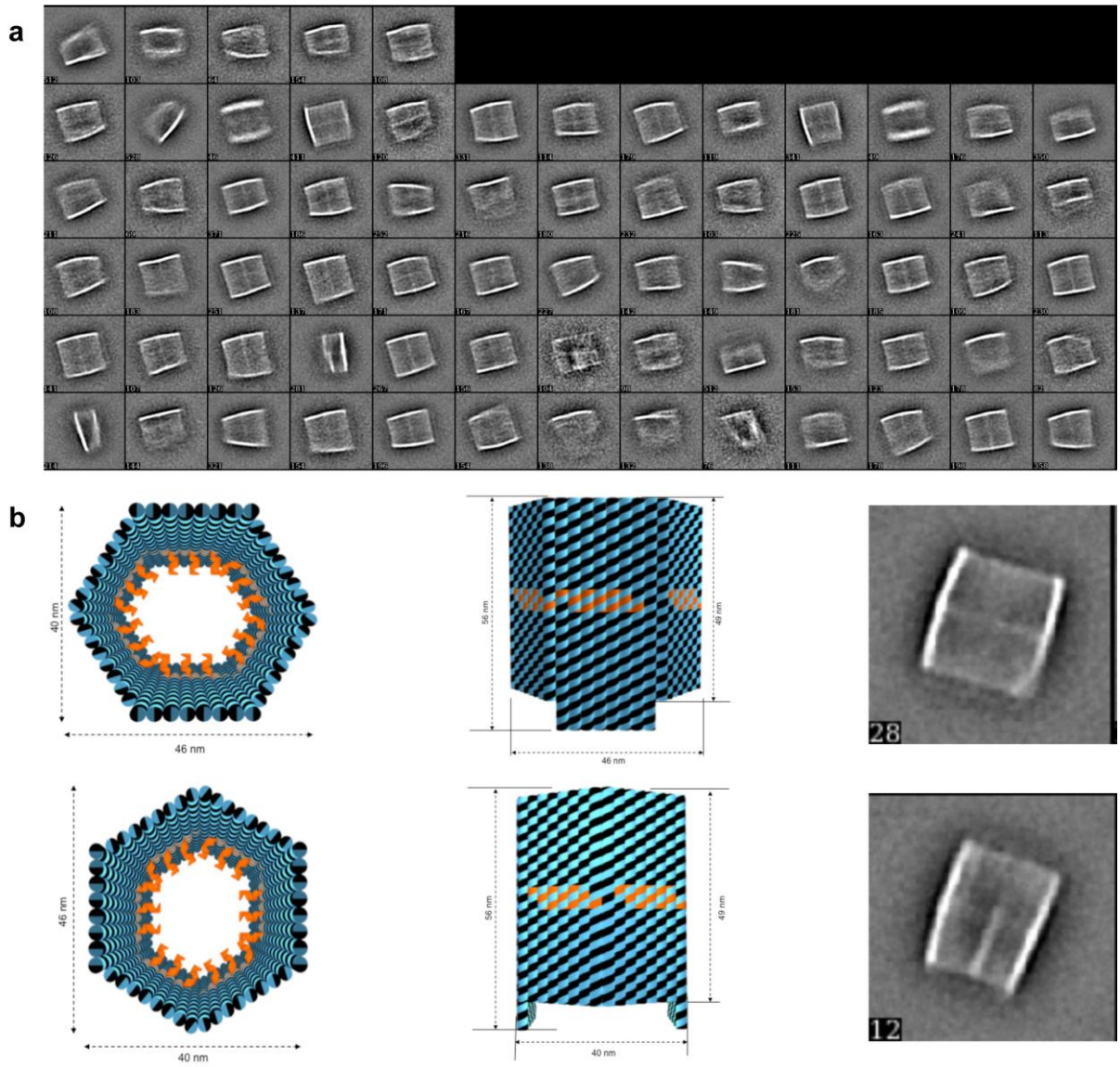


***Figure 30*:** Two dimensional TEM analysis of the unloaded 6p construct. (**a**) Representative class averages of the empty channels and (**b**) simulated 3D models showing the construct from its front or top view in two possible perspectives with their corresponding TEM images on the right.

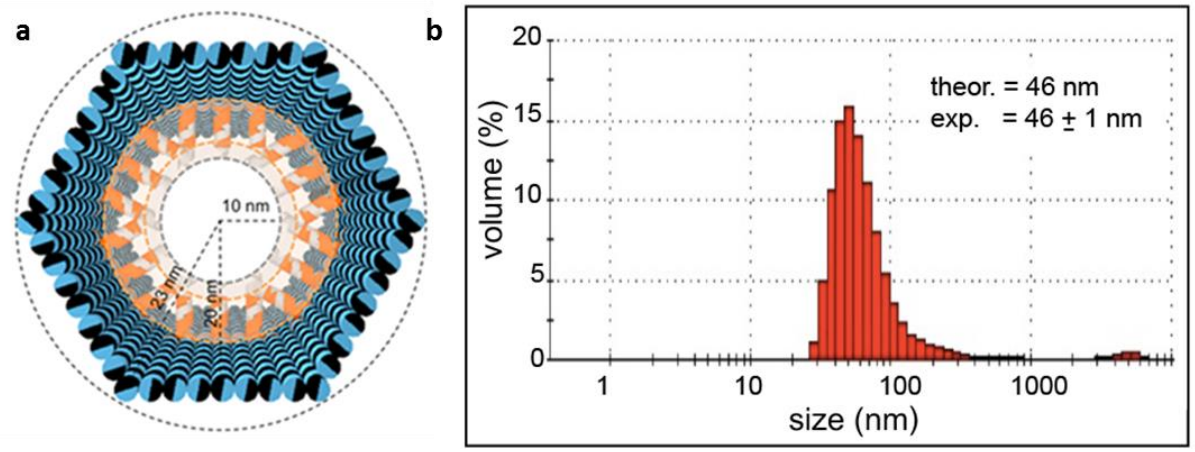


Figure 31: (**a**) Schematically illustration of the 6p design with its inner (20 nm) and outer radius (23 nm). (**b**) The theoretical values match with the experimental ones, measured by dynamic light scattering (DLS).

3.3 Linker

The use of a supramolecular linker is based on the choice of a suitable peptide sequence, which binds with high affinity to the protein of interest. In this case the peptide sequence aspartic acid (D), proline (P), methionine (M), phenylalanine (F), lysine (K), leucine (L) and valine (V) (DPMFKLV) was chosen because of its affinity to bind to the PDZ1 domain of the DegP protein in all oligomeric states (K_d of 5 µM)[48]. A more detailed description of the synthesis can be found in chapter 6.5.2.8.

The length of 16 bases of the DNA sequence (*Table 4*) guarantees stable binding to the protruding arms at room temperature and during gel electrophoresis, but also allows removing them by single strand displacement using a sequence of 22 bases fully complementary to the PAs (chapter 3.4.3.8). Using labeled oligonucleotides and proteins offered the opportunity to perform co-localization gel-electrophoresis studies and single molecule FRET experiments (chapter 3.4.2.6).

***Table 4*:** List of available supramolecular DNA-peptide conjugates and. The amino terminus of the peptide is linked to the 5`-terminus of the DNA sequence.

Peptide 5′	Sequence	Label 3′	Shortcut
DPMFKLV-OH	GTGGAAAGTGGCAATC	-	DPMFKLV-A1
DPMFKLV-OH	GTGGAAAGTGGCAATC	Flc	DPMFKLV-A1-Flc
DPMFKLV-OH	GTGGAAAGTGGCAATC	TAMRA	DPMFKLV-A1-TAMRA

3.3.1 Synthesis and characterization of the DNA-peptide conjugate

Solid phase peptide synthesis was carried out on Wang resin applying common Fmoc/tBu chemistry. The peptide Asp-Pro-Met-Phe-Lys-Leu-Val was synthesized and modified at the N-terminus using maleic anhydride. Within the use of standard coupling reagents, the coupled maleic anhydride was transformed into a maleimide function. Modifying the peptide at the resin avoided side-reactions at other functional groups, due to the orthogonal protecting groups. The modified peptide sequence was removed from the resin and purified by HPLC. After removing the solvent, the product was lyophilized and immediately used for covalent binding to the thiol-modified oligonucleotide A1. For this, the oligonucleotide was firstly reduced to its thiol form using TCEP. The oligonucleotide was further purified using NAP^{TM}-5 and NAP^{TM}-10 columns and concentrated in 3000 Da MWCO ultra centrifugal filter units. The peptide was added in 25-fold excess to the oligonucleotide for 48 h and purified by

denaturing PAGE (chapter 6.5.4). The purified and concentrated DNA-peptide conjugate was characterized by PAGE (*Fig. 32a*) and MALDI TOF TOF (*Fig. 32b*). The final yield was up to 56% (based on the amount of maleimide-DPMFKLV).

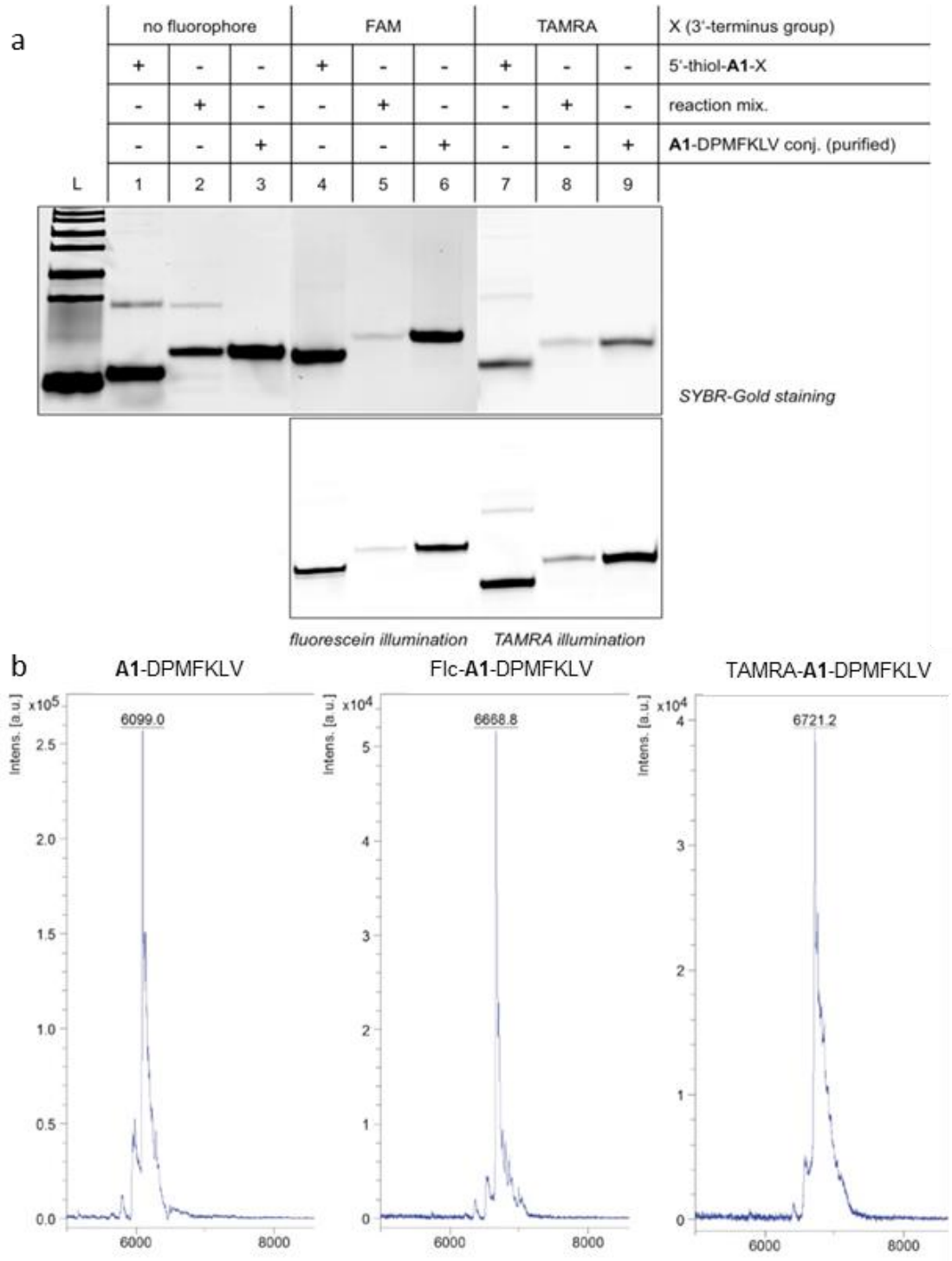


***Figure 32*: (a)** Denaturing PAGE characterization of the available DNA-peptide conjugates. The DNA strand (A1) used for peptide conjugation was either unlabeled at the 3'-terminus (lanes 1 to 3), or modified with a Flc (lanes 4 to 6) or a TAMRA fluorophore (lanes 7 to 9). The adduct was loaded before (lanes 1, 4 and 7) and after reaction with the maleimide-activated peptide (lanes 2, 5 and 7) and finally purified by denaturing PAGE (lanes 3, 6 and 9). (**b**) MALDI characterization of the available DNA-peptide conjugates. Theoretical masses are: 6099 g/mol (A1-DPMFKLV), 6668 g/mol (Flc-A1-DPMFKLV) and 6722 g/mol (TAMRA-A1-DPMFKLV). Deviations from the experimental values are less than 0.05%.

3.4 Loading

Loading experiments were performed, if not mentioned differently, with 25-fold excess of the respective protein. The mixture was shaken overnight (200 u/min) at room temperature under light exclusion.

3.4.1 Molecular Modeling

In all three oligomeric states of DegP, the monomers are arranged around a central cavity of increasing size. This arrangement of the monomers leads to orientation of the PDZ1 domains in $DegP_6$ towards the solvent. In $DegP_{12}$ and $DegP_{24}$ instead the domains are located at the entrance of a central cavity. In $DegP_{12}$ the PDZ domains are grouped in units of three domains arranged in a tetrahedral orientation. The resulting space between those PDZ domains just allows the entrance of one DNA helix. Conversely, the $DegP_{24}$ features a cube with six entrances to the inner cavity, each formed by four PDZ1 domains. Molecular Dynamics (MD) simulations suggest that these entrances are big enough to host up to three DNA helices without compromising the binding of the incoming peptide DPMFKLV (*Fig. 33a*). Maximal probability of binding should be thus provided in the presence of three ligands per DNA origami face; 18 ligands in total.

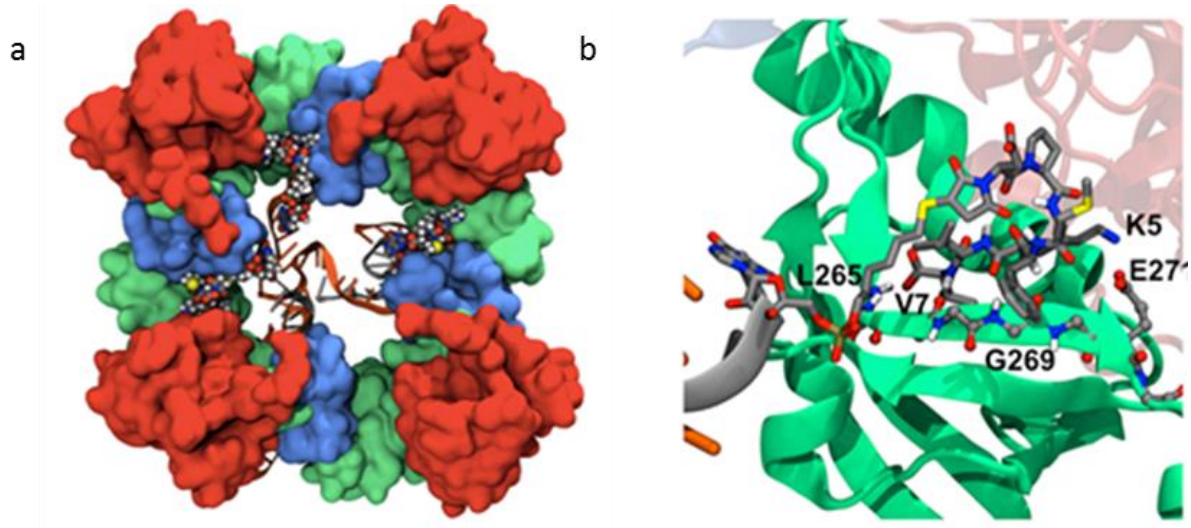


***Figure 33*:** Results of the molecular modelling. (**a**) Illustration of the molecular dynamics simulation of the binding of the DNA-peptide ligand to the PDZ1 domains of $DegP_{24}$. The PDZ1 domains are orientated into the cavity of the protein. (**b**) Detailed view of the molecular model showing the interaction between the DPMFKLV ligand and the PDZ1 binding site[53].

The placement of the side chain V7 in the hydrophobic pocket of the PDZ1 domain and the hydrogen bonds established between the backbone of the peptide (residues 5 to 7) and the

corresponding residues of the PDZ1 domain (residues 265 to 269, *Fig. 33b*) are the driving forces for the interaction between the peptide and the PDZ1 domain. All DegP forms, although with distinct space-filling capabilities, can be hosted inside the DNA-origami cage and furthermore, more than one protein molecule could bind inside the cavity (AFM image; see chapter 3.4.4.1).

3.4.2 Theoretical consideration on molecular diffusion

The size of the three different oligomeric states of the protein ($DegP_6$ with a diameter of 9 nm; $DegP_{12}$ with a diameter of 13 nm and $DegP_{24}$ with a diameter of 19 nm) has an effect concerning the diffusion properties of the proteins through the DNA cage. Assuming that the radius of the encapsulated protein is r_{prot} and the radius of the available space within the cavity is R_{free}, the flux (I) of a protein traversing through the DNA host in a typical experiment (30 µL of a 20 nM origami solution, at room temperature in an aqueous buffer) is given by $I = 166\ R_{free}/r_{prot}$ (particles s^{-1}).
This results in an inverse proportionality between the size and the flux of molecules per unit of time according to a 4 : 2 : 1 ratio, for the 6-, 12- and 24-mer, respectively.

3.4.3 Loading of the protein

In this chapter, the results of the loading experiments are shown. Due to some remarkable information concerning the migration of the protein and DNA, the gels are shown in their full size and explained in detail.
In general, proteins and DNA migrate in opposite directions due to their different net charge (in this work all gels are depicted with the anode at the bottom and the cathode at the top of the image). All protein/DNA containing bands, (*Fig 34*; i.e. lane 7; Alexa 488 illumination) show a specific interaction between the two species. Binding of the protein to the DNA is associated to a decrease of the migration rate (compare lanes 3, 4, 5 and 7 of *Figure 34*). Migration of the protein upwards is maximal in the absence of any DNA. It has to be mentioned that the excess of staples as well as the unbound peptide-linker were removed prior to gel loading by 100 kDa MWCO ultra centrifugal filter units. Successful and specific binding of the $DegP_{12/24}{}^{A488}$SA protein (*Fig. 34*; lane 7, merge of Alexa 488 illumination and

TAMRA illumination) to the host 6p^{120}-18cA1 functionalized with TAMRA labeled ligands in its inner cavity could be demonstrated.

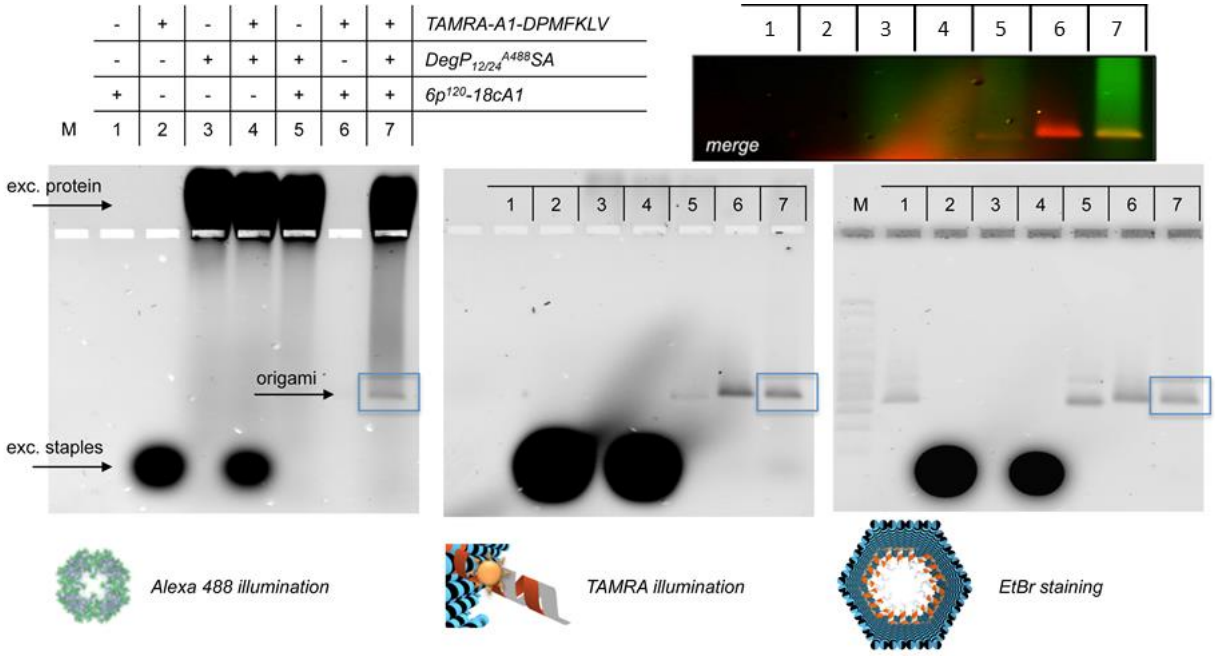


***Figure 34*:** Binding of the DegP$_{12/24}^{A488}$SA to the 6p^{120}-18cA1 construct functionalized with TAMRA labeled ligands in its inner cavity. Successful binding only takes place in the presence of ligands.

The binding of the protein to the peptide-modified origami did not result in any remarkable gel mobility shift, (*Fig. 34*; comparing lane 6 and 7). This supports the fact that the protein is encapsulated inside the cage. Binding to the exterior surface would decrease the mobility in agarose gel electrophoresis, due to the coverage of the surface charge of the DNA origami cage (see *Fig. 21*).

Specific and successful loading of the DegP$_6^{A633}$SA to the host 6p^{120}-18cA1 functionalized with fluorescein labeled ligands in its inner cavity was demonstrated by comparing lanes 5, 6 and 7 of Figure 35 (*Fig. 35*). In the absence of the peptide-ligand (*Fig. 35*; lane 5) no protein signal occurs; only in the presence of the ligands (*Fig. 35*; lane 7) an Alexa 633 illumination (protein) and a fluorescein (linker) illumination signal could be measured. Again, no migration difference between the loaded and unloaded DNA origami cages (*Fig. 35*; lane 6 and 7) was detected, supporting the already mentioned idea of correct internalization, due to lack of surface charge covering.

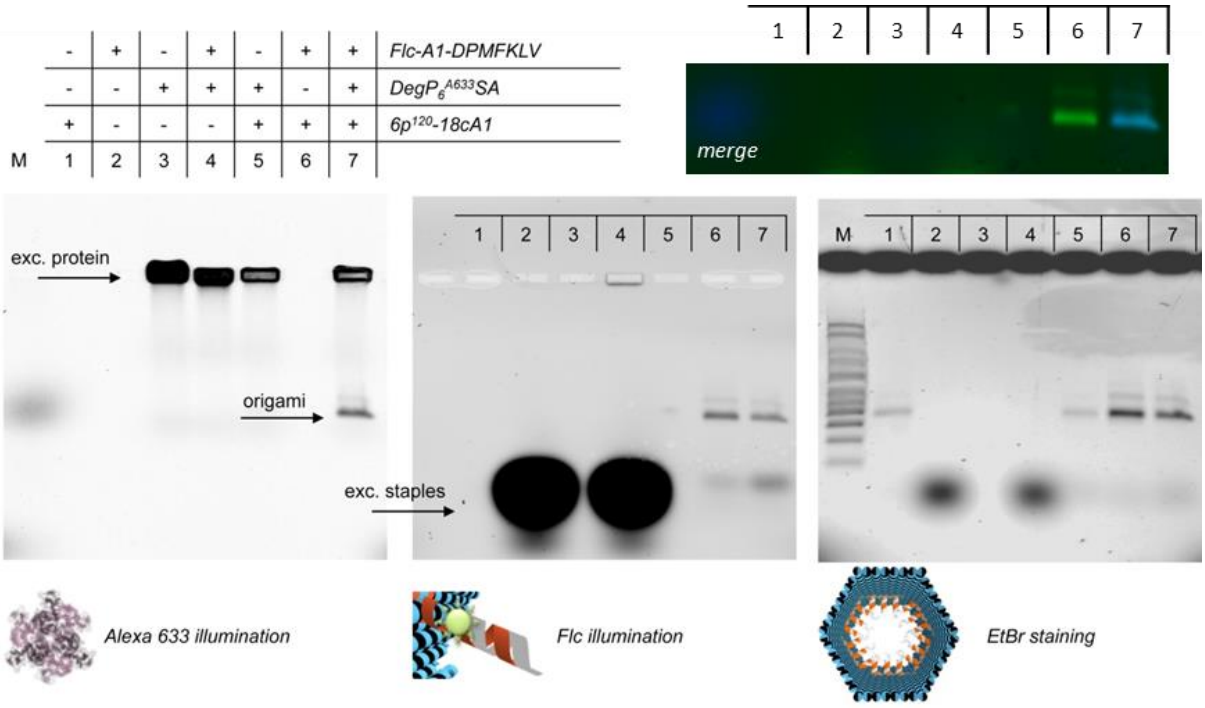


***Figure 35*:** Binding of the $DegP_6^{A633}SA$ to the $6p^{120}$-18cA1 construct functionalized with Fluorescein (Flc) labeled ligands in its inner cavity. Successful binding only takes place in the presence of ligands.

Information about the time needed to form the DNA-origami protein complex can be gained from Figure 36. After 3 hours, the maximal binding efficiency was reached.

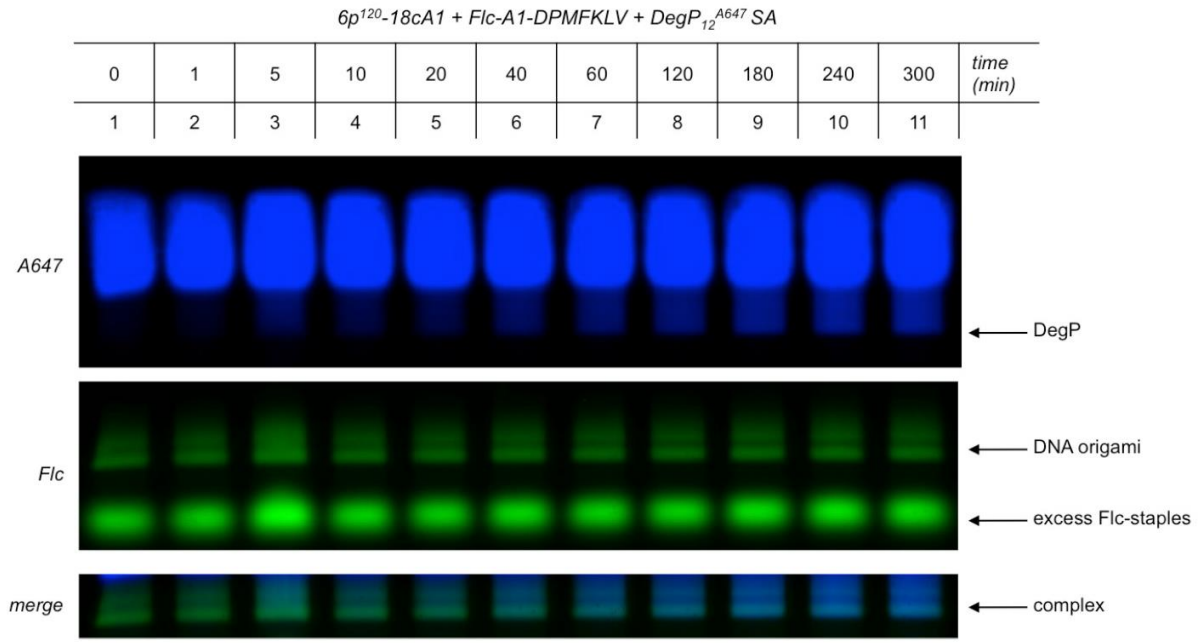


***Figure 36*:** Time dependent binding of the DegP12A647SA protein to the 6p host functionalized with 18 Flc-labeled A1-peptide ligands in its inner cavity. After 180 minutes the maximal binding efficiency is reached.

3.4.3.1 Titration studies

In order to establish the minimal amount of DegP protein that is necessary to visualize successful binding, either in the presence of 6 or 18 ligands, titrations experiments were performed (*Fig. 37*). These experiments also served to show whether the use of too much protein would cause aggregation and unspecific binding. As expected, the efficiency of protein binding increased with an increasing amount of molar equivalents. In the presence of an equal excess of protein, the higher number of ligands in the $6p^{120}$-18cA1 leads to a higher binding efficiency, with ca. 36 % yield in the presence of 25 molar equivalents of $DegP_{12/24}^{A488}SA$ and almost saturated (82 %) at 50-fold excess of protein (a detailed description of the calculation of the yields of gel electrophoresis can be found in chapter 3.4.3.4). On the contrary, using the $6p^{120}$-6cA1 construct, saturation was reached at 100 molar equivalents of the protein. Both constructs appear to be fully bound with 100-fold excess of protein. These results indicate that the binding efficiency of DegP increases with the number of ligands, which are accessible in the vicinity of the protein surface.

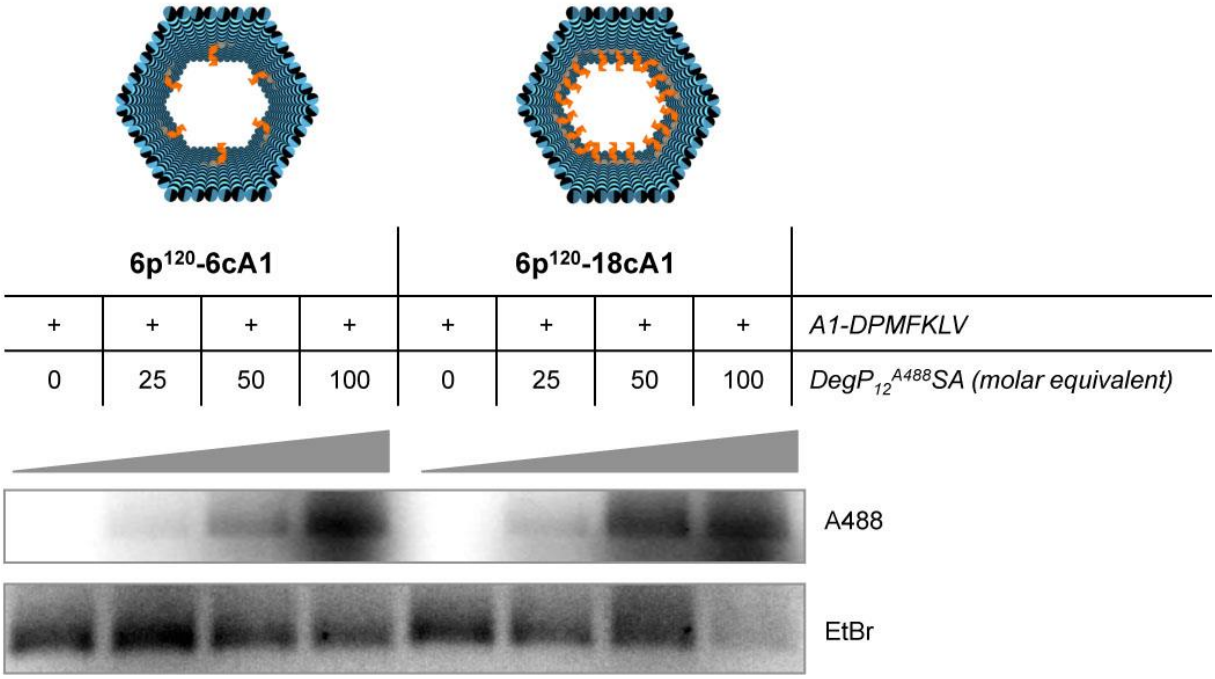


***Figure 37*:** Titration studies of the binding of $DegP_{12/24}^{A488}SA$ to the 6prism modified with either 6 or 18 A1-DEPMFKLV ligands in its inner cavity.

3.4.3.2 Effect of pH value on DegP binding to the DNA host

In all performed agarose gels, the lanes containing the labeled DegP protein showed a large band directly above and on the loading pocket migrating towards the cathode (*Fig. 34*).

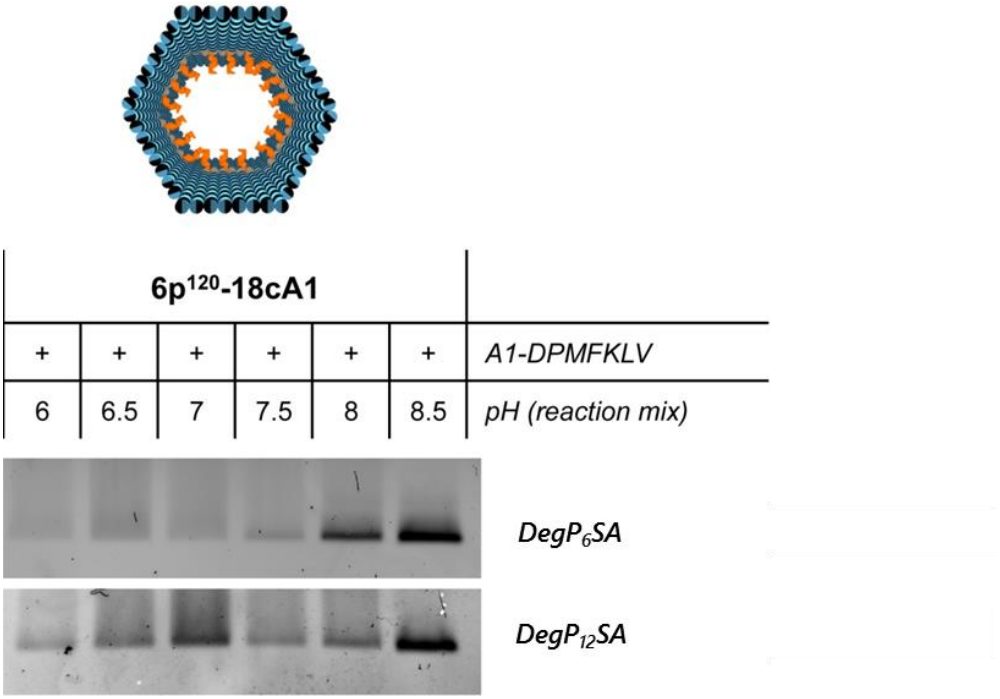


***Figure 38*:** EtBr stained gel of the binding of $DegP_6SA$ and $DegP_{12}SA$ to the $6p^{120}$-18cA1 host functionalized with A1-DPMFKLV ligands in its inner cavity. The different pH values show effects concerning loading efficiency and host stability.

This band clearly belongs to the protein; it is only visible at the excitation wavelength of the fluorophores that were used for DegP labeling (*Fig. 34 and 35*). The low migration properties of the protein are not treated in this work, but reveal an interesting relation between the surface charge and the loading efficiency, suggesting possible electrostatic forces between host and guest.

It has to be noted that the presence of fluorophores has an effect on the surface charge and therefore, on the electrostatic forces, due to the charge of the fluorophores themselves or to the partial screening of the fluorophore modified residues. To understand whether the charge of the protein has any effect on specific or unspecific binding to the host, pH dependent experiments were performed (*Fig. 38*). The unlabeled proteins and single mutated $DegP_6SA$ and $DegP_{12}SA$ were used (the labeled proteins were not used for this experiment, due to the effect of the fluorophore charges on the protein). For both proteins, increasing the pH from 6 to 8.5 in 0.5 steps, results in improved resolution and intensity of the ethidium bromide bands. This indicates a slightly lower (less positive) net charge of the protein and consequently a better and more specific interaction with the DNA host. Increasing the pH causes a

deprotonation of the lysine residues on the surface, reducing the unspecific electrostatic interactions and allowing the specificity of the peptide-PDZ1 domain interaction to emerge. The strength of the effect of the unspecific binding can be seen in lane 1 to 4 of Figure 38, indicated by a less compact band and a gel smear.

3.4.3.3 Binding of the wild-type of $DegP_6$ to the DNA origami host

To prove whether the mutation, which stops the auto digestion of the DegP protein (DegP-SA) has any effect on the binding capability to the DNA origami host, loading experiments of the wild type were performed. Therefore, the labeled $DegP_6^{A488}WT$ was used, due to its slightly remaining auto digesting activity. In higher oligomeric forms ($DegP_{12}WT$ would be preferential for loading) DegP has an increased protease activity and would digest itself too rapidly for routine handling and analysis. Successful binding of the $DegP_6^{A488}WT$ to the $6p^{120}$-18cA1 host was shown by matching Alexa 488 illumination and ethidium bromide signals, only in the presence of A1-DPMFKLV ligands (*Fig. 39*; lane 7). The oligomerization state of the loaded protein (either 6-mer or 12-mer) does not appear to affect the binding specificity to the ligand modified host, thus confirming the general applicability of the system.

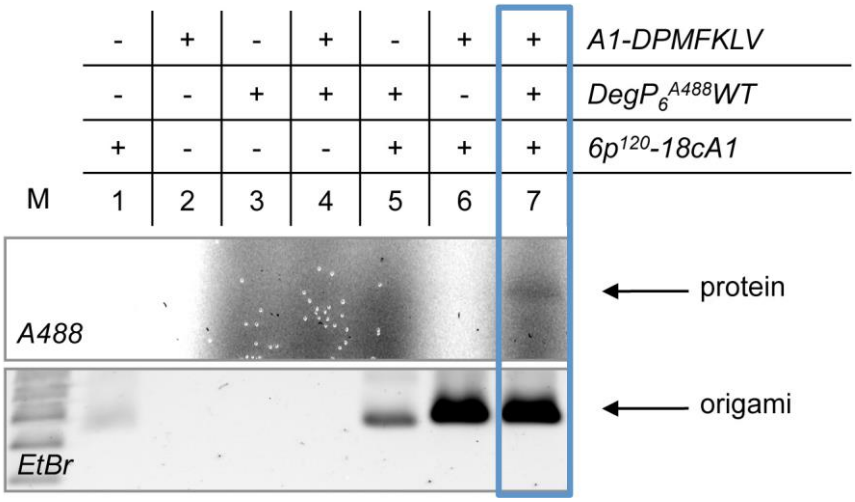


***Figure 39*:** Binding of the $DegP_6^{A488}WT$ to the 6prism functionalized with 18 A1-DPMFKLV ligands in its inner cavity.

3.4.3.4 Comparison of the loading ability and efficiency of different designs

Successful and specific binding to the host was shown for $6p^{120}$ constructs functionalized with ligands bound to the PAs of the inner cavity in well-defined orientations. To investigate the effect of distinct face-to-face connections, resulting in DNA origami hosts with randomly outwards and inwards oriented PAs, an experiment was performed (compare with *Fig. 21*; chapter 3.2.2) to test whether, $DegP_{12/24}{}^{A488}SA$ can also be attached to the outer surface of the prism, both in the absence and presence of PAs. Again, protein binding was observed only in presence of PAs and their complementary A1-DPMFKLV ligands, confirming specificity of the binding interactions (*Fig. 40*; lanes 3, 5 and 7).

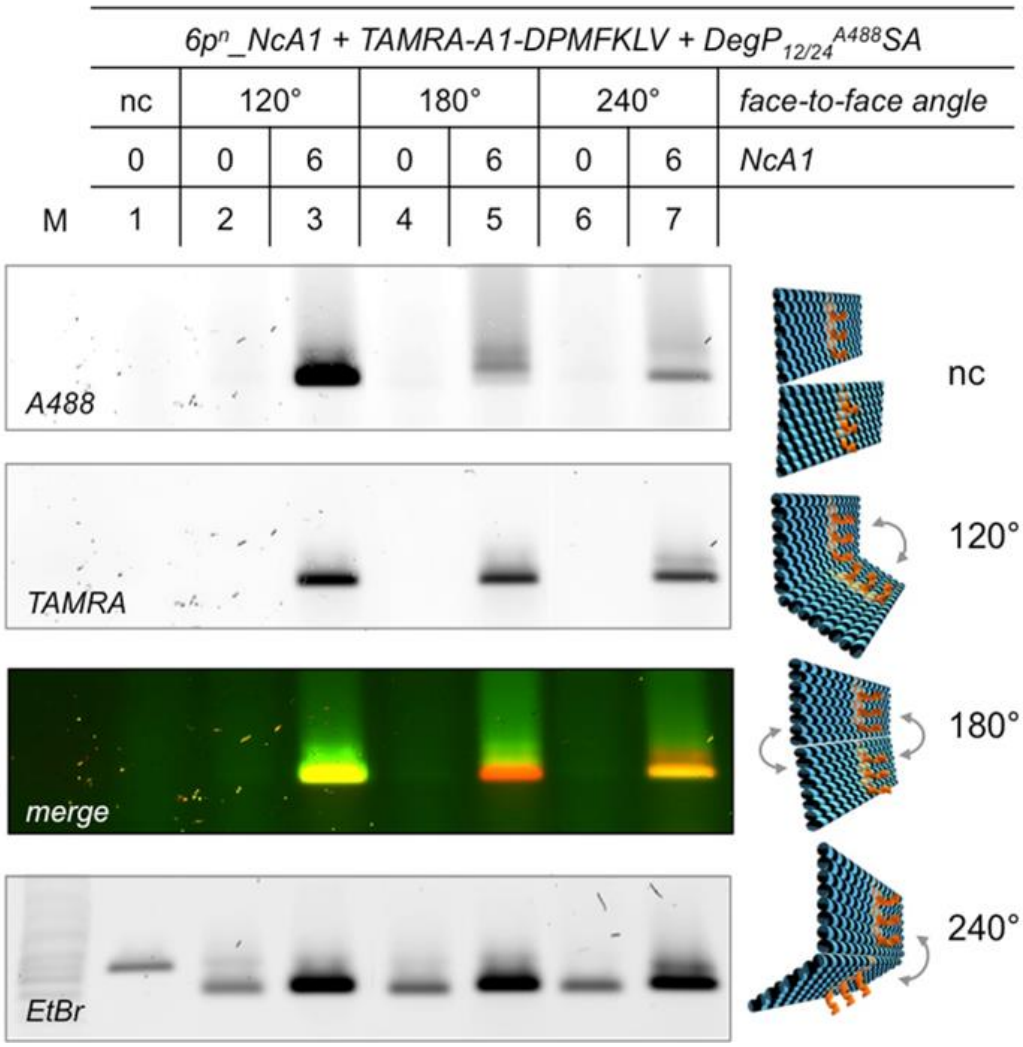


***Figure 40*:** $DegP_{12/24}{}^{A488}SA$ binds to the 6p constructs depending on the presence of PAs (lane 3, 5 and 7) and on their convergent (120°), randomly oriented (180°) or divergent (240°) arrangement. The strength of Ethidium Bromide staining is almost similar, indicating a comparable amount of the 6prisms; in comparison to the different strength of the Alexa 488 illumination. Full gel is shown in chapter 10.3.3..

A comparison of these bands, which are associated to the DNA-protein complex in the different face-to-face connection designs, shows that although there is a similar TAMRA

(ligand) and ethidium bromide (cage) signal, there is a huge difference in the intensity of the protein signal (Alexa 488 illumination). Gel analysis using ImageJ (the method will be exemplarily explained in the following chapter) revealed yields in a 8 : 1.4 : 1 ratio, for the $6p^{120}$, $6p^{180}$ and $6p^{240}$ designs. These results indicate a higher loading efficiency for host structures whose PAs are oriented inwards ($6p^{120}$), suggesting that convergent ligands within a restricted environment promote protein caging, presumably as a consequence of their high local concentration and spatially oriented interactions towards a common target. The $6p^{180}$ and $6p^{240}$ designs do not offer these conditions, due to their either stochastically (180°) or outwards (240°) oriented PAs.

3.4.3.4.1 Calculation of the yield by gel electrophoresis

Using ImageJ, the yield of the protein binding to the host could be calculated, marking the selected bands with an identical rectangular shape and measuring the area below the intensity curve. Relative intensities corresponding to the products analyzed under similar conditions were compared. All DNA constructs have been prepared and loaded in the same nominal amount, resulting in similar intensities of the bands after ethidium bromide staining, at least within the experimental errors. For variations of the signal intensity inferior to 10 %, it was reasonable to quantify the signals visible under different wavelengths and the relative yield of protein loading as P/D = intensity protein band / intensity DNA band was calculated. The yield of DNA labeling was calculated similarly as L/D = intensity labeled-peptide band / intensity DNA band (*Fig. 41*).

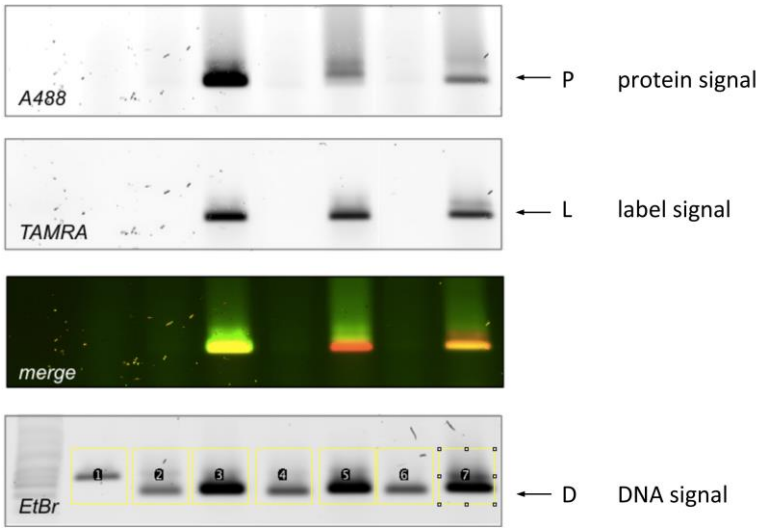


***Figure 41*:** Exemplary calculation of the yields by gel electrophoresis.

3.4.3.5 The effect of ligands multiplicity

The results shown in chapter 3.4.3.4 indicate that a higher local concentration of PAs and consequently of the peptide-modified ligands leads to a higher loading efficiency. To explore this effect, which can be described as an effect of ligands multiplicity, several distinct host constructs of the $6p^{120}$ design, differing in the number and spatial arrangement of the cavity, were prepared (*Fig. 42*).

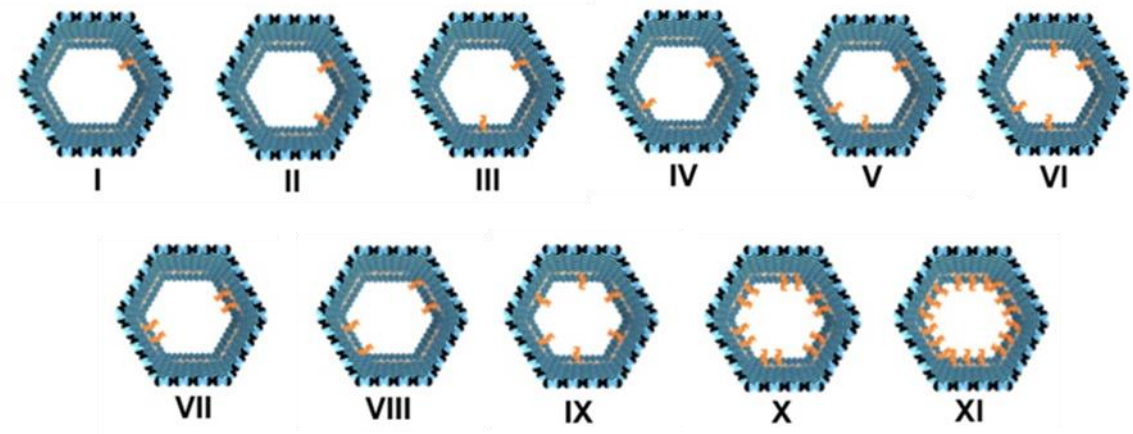


***Figure 42*:** Schematic illustration of several $6p^{120}$-designs differing in the number and spatial arrangement of the ligands. The different spatial orientation of the same number of ligands are used as a tool to prove whether the proximity or distance of ligands have any effect on the protein binding.

In this experiment, the almost completely purified $DegP_6{}^{A633}SA$ (purity of 97 %, calculated by TEM measurements) was used, because of the identical illumination strength of the fluorophores, due to the similar number of fluorophores in each 6-mer. Using a mixture of e.g. $DegP_{12/24}{}^{A488}SA$ resulted in the encapsulation of a 12-mer and in another case a 24-mer with a twice stronger illumination leading to invalid data. One should note that the DegP samples, which were used in this thesis, are difficult to isolate if the oligomeric state is above the 6-mer state. Next to the various numbers of PAs, the structures also differ in the spatial arrangement of the PAs for structures with an identical amount of arms.

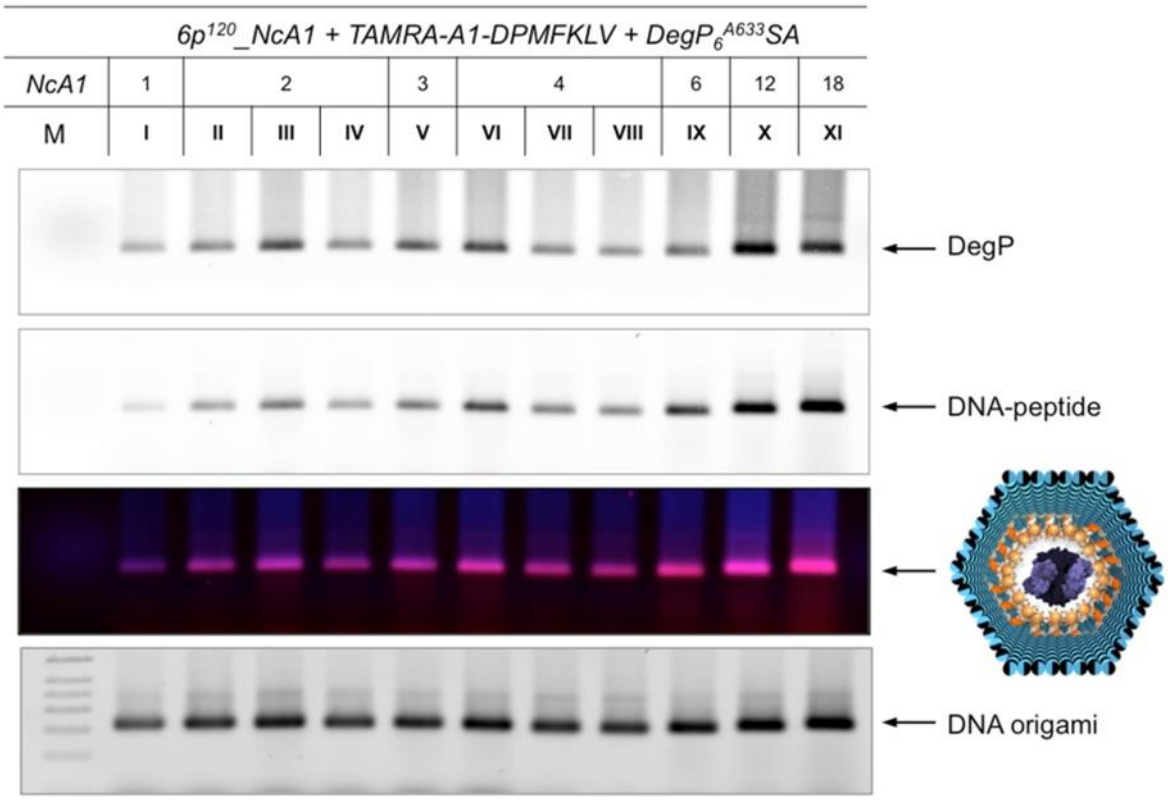


***Figure 43*:** Agarose gel electrophoresis characterization of the binding of $DegP_6^{A633}SA$ to diverse $6p^{120}$ constructs (see *Fig.* 42). The different yields were calculated with the same method as in chapter 3.4.3.4.1. Binding took place for all constructs with increasing efficiency for a radial distribution of ligands.

Two facts were noticed from this experiment: (i) the binding of the protein inside the cavity of the DNA origami cage occurs even in the presence of one single ligand; although the yield is low (*Fig. 43*) and (ii) the binding efficiency is proportional to the number of internalized peptide ligands and assumed a maximal value for a distribution of 12 or 18 ligands.

3.4.3.6 Single molecule co-localization of the DegP and the host

Additional single-molecule characterization of gel-purified compounds using total internal reflection fluorescence (TIRF) microscopy was performed (*Fig. 44*). Therefore, a $6p^{120}$ DNA host functionalized in its cavity with 18 TAMRA-A1-DPMFKLV ligands was further modified with biotin handles appended on the outer surface of one of its six faces for immobilization on a streptavidin coated glass support (*Fig. 44*; right panel).

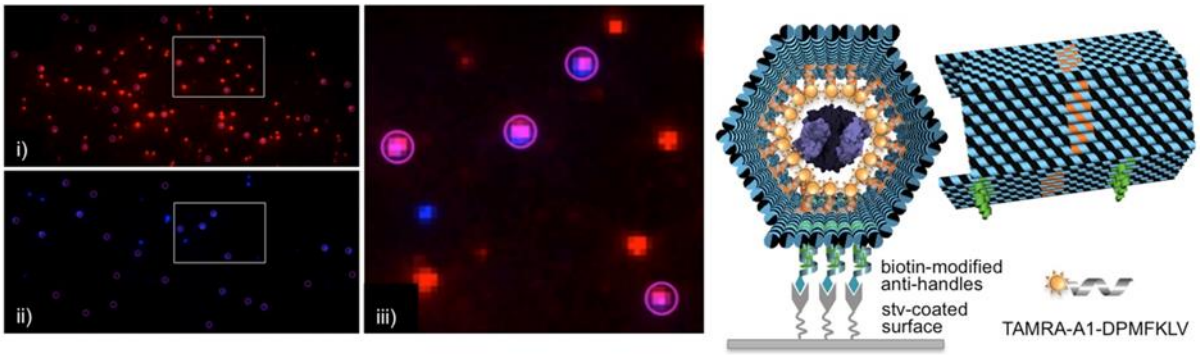


***Figure 44*:** Single-molecule fluorescence characterization of the 6p constructs bearing 18 convergent PAs hybridized with TAMRA labeled A1-DPMFKLV ligands and loaded with $DegP_6^{A647}SA$ proteins. DNA cages were immobilized on a coverslip surface and were measured by TIRF microscopy. TAMRA (red spots in i) and Alexa 647 (blue spots in ii) detection channels have been overlapped, indicating co-localization of the two species (FRET spots in purple, iii, which shows a zoom-in view of the highlighted region in panels i and ii).

Red spots in (i) (*Fig. 44*; left upper panel) show the TAMRA labeled DNA-origami hosts and in the Alexa647 signals in the panel below (ii). Overlapping these detection channels (iii) indicates a co-localization of both, the TAMRA labeled DNA host and the Alexa647 labeled DegP guest following donor excitation (*Fig. 44*; panel iii zoom-in). Binding of the protein resulted in an observable energy transfer from the donor to the acceptor fluorophore, thus proving that the binding partner were in close proximity.

3.4.3.7 Effect of lysine-selective molecular tweezers on DegP binding to the host

High attractive electrostatic forces between the unlabeled DegP proteins, e.g. $DegP_{12}SA$, whose surface is covered with positive lysine residues, and the DNA origami host are probably responsible for unspecific binding. Therefore, it was investigated whether, next to the pH value, reducing the net charge of the protein via lysine selective molecular tweezers has any effect on the specificity of binding of the protein to the DNA host and consequently reducing unspecific binding between them. Two different pH values were chosen (referring to chapter 3.4.3.2): (i) pH 6 was chosen due to high unspecific interactions, resulting in a barely visible migration band with a lot of smear and (ii) a pH value of 7.6, thus resembling standard assembly and loading conditions for the DNA origami. The effect of the peptide ligand TAMRA-A1-DPMFKLV and the different pH values (*Fig. 45*; lanes 1 to 5) are explained separately form the effect of the lysine selective molecular tweezer (*Fig. 45*; lanes 6 to 9).

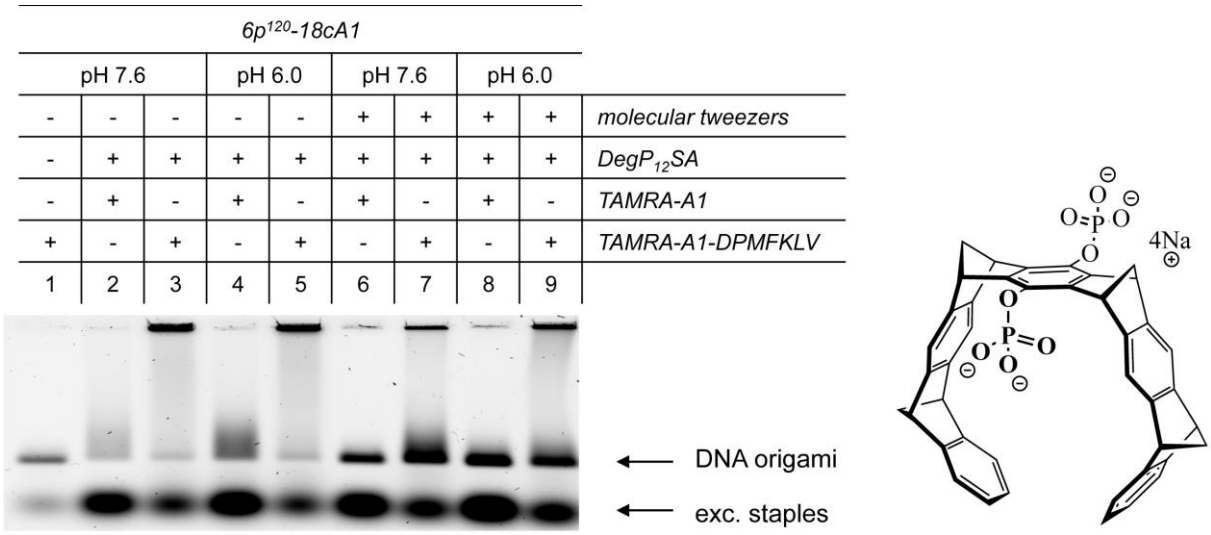


***Figure 45*:** Binding of unlabeled $DegP_{12}SA$ to the $6p^{120}$-18cA1 construct functionalized with TAMRA-A1 or TAMRA-A1-DPMFKLV ligands orientated inside the cavity in absence or presence of lysine selective molecular tweezers using different pH-values. Due to the almost identical TAMRA illumination and Ethidium Bromide staining, only the TAMRA illumination is shown.

Effect of peptide ligands and pH: $6p^{120}$-18cA1 functionalized with TAMRA-A1-DPMFKLV ligands was used as a control (lane 1), giving a well-defined band without any remaining signal inside the loading pocket. Addition of the $DegP_{12}SA$ in the absence of the peptide ligands (lanes 2 and 4) leads to an undefined mixture of products of similar migration properties, evidenced by comparable gel smear. This indicates that in the absence of the peptide modified ligands the unspecific electrostatic interaction between the positive charged DegP protein and the negative charged DNA origami dominate, with a stronger effect at lower pH. Adding protein in the presence of peptide ligands (lane 3 and 5) leads to the formation of a well-defined band of low intensity and identical migration rate.

Effects of lysine selective molecular tweezers (lane 6 to 9): Addition of the tweezers in 25 fold excess related to the protein leads to the formation of well-defined bands of higher intensity and almost complete disappearance of gel smears. This suggests an enormous reduction of unspecific electrostatic interactions between the negatively charged DNA host and the positively charged protein, due to the selective interaction of the tweezer with the lysine residues on the surface. There is no remarkable difference between the two different pH values.

All samples with the TAMRA modified peptide ligands and protein show huge aggregates in the gel pockets, just differing in the illumination strength in absence or presence of tweezers. Interestingly, the strength of the illumination signal in the pockets corresponds with a decrease of the illumination signal strength in the excess staples, which indicates the formation of aggregates due to a combination of the DNA-origami cage, the excess staples

and the protein (compare lanes 3 with 7 and 5 with 9). However, lanes 6 and 7 were analyzed by AFM proving specific and successful binding of the $DegP_{12}SA$ in presence of the tweezer with even higher loading yields (AFM images see chapter 3.4.4.1.1).

3.4.3.8 Release or exchange studies

After successful encapsulation of the protein in the cavity of the $6p^{120}$ constructs, experiments for further applications were performed. One possible function would be to release the encapsulated protein from the cage via single strand displacement (ssdp) of the A1-DPMFKLV ligand (*Fig. 46*). Lanes 1-9 show the corresponding controls without the addition of $DegP_6^{A647}SA$ for the $6p^{120}$-1cA1 (lanes 1 to 3), $6p^{120}$-6cA1 (lanes 4 to 6) and $6p^{120}$-18cA1 (lanes 7 to 9) designs. To lanes 2, 5 and 8, 10-fold and to lanes 3, 6 and 9, 100-fold excess of the full complementary sequence of the protruding arms was added, proving the almost completely removal of the Flc-A1-DPMFKLV ligands (the darker fluorescein illumination panel shows that there is an almost negligible excess of Flc-A1-DPMFKLV ligands left).

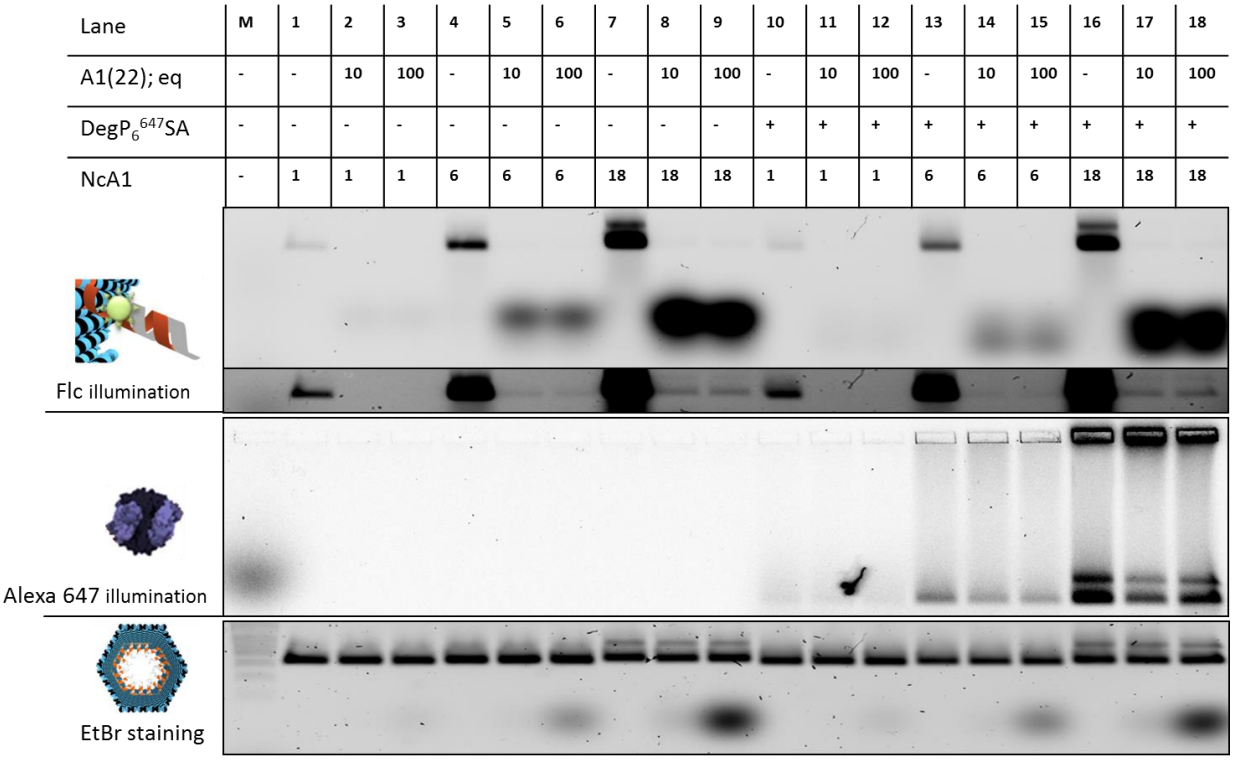


***Figure 46*:** Agarose gel electrophoresis of the first release studies of the $DegP_6^{A647}SA$ protein out of the various $6p^{120}$ constructs with 1, 6 and 18 PAs. The Fluorescein labeled A1-DPMFKLV ligands are removed by adding the full complementary sequence (A1(22)) to the PAs in excess (no A1(22) was added to lane 1, 4, 7, 10, 13 and 16).

Lanes 10-18 correspond to lanes 1-9 and only differ in the addition of $DegP_6^{A647}SA$ in 25-fold excess to the DNA origami cage. Adding 10-fold or 100-fold excess of the A1(22) sequence results in a comparable removal of the Flc-A1-DPMFKLV ligands to lanes 2, 3, 5, 6, 8 and 9, but lacks in the complete release of the protein. Comparing the Alexa647 illumination signals of lanes 16 and 17 shows a decrease in intensity of 49 % and between lanes 16 and 18 a decrease of only 19 %. Due to the present fluorescein signal after single strand displacement with almost identical illumination strength in absence or presence of the protein (compare lanes 8 and 9 with 17 and 18), the only explanation for the remaining DegP protein is an unspecific electrostatic interaction between the protein and the DNA cage.
Consequently, an experiment to reduce the electrostatic interaction during single strand displacement, via adapting the pH value, was performed (*Fig. 47*).

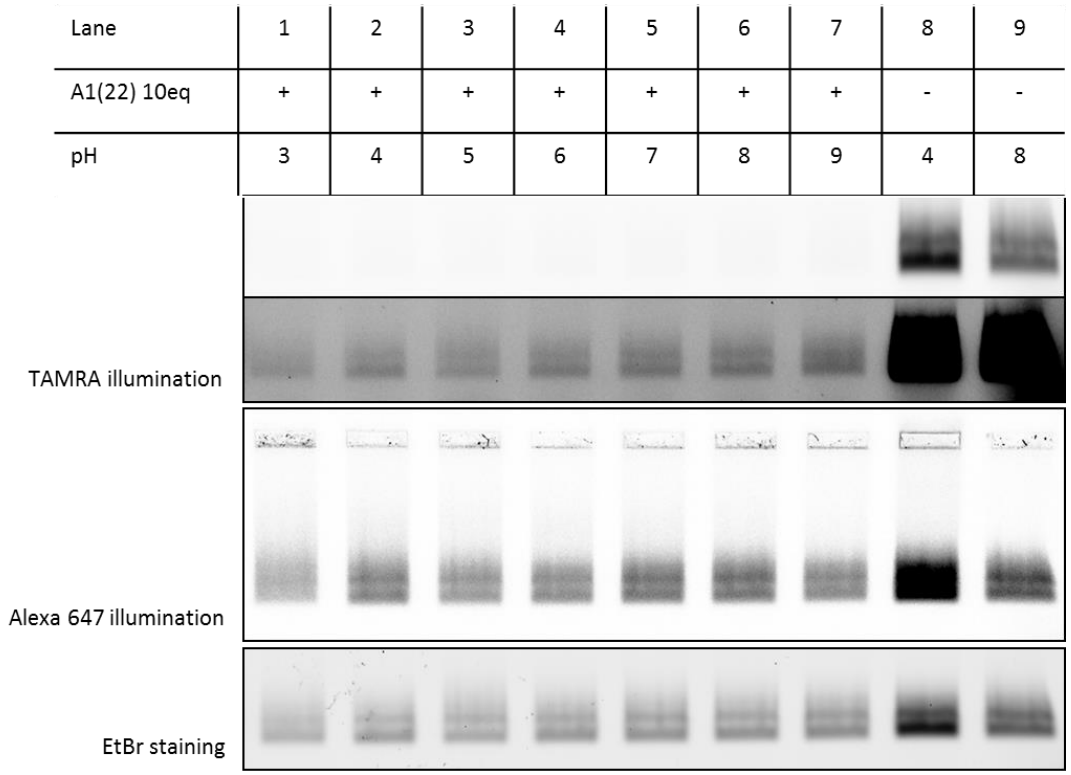


***Figure 47*:** Agarose gel electrophoresis of pH dependent release studies of the $DegP_6^{A647}SA$ protein in TEMg buffer with different pH values. The TAMRA labeled A1-DPMFKLV ligands were removed by adding 10-fold excess of the full complementary sequence (A1(22)) to the PAs (no A1(22) was added to lane 8 and 9).

The pH values were chosen based on the stability of the protein at different pH values. Using PEG purification[55] the $6p^{120}$-18cA1 constructs were dissolved in the TEMg buffer with adjusted pH and the A1(22) sequence for single strand displacement was added in 10-fold excess (*Fig. 47*; lanes 1-7). Agarose gel electrophoresis analysis revealed that the protein was

not released from the cage, suggesting that, whereas the pH has an effect for loading (*Fig. 45* page 52), it has instead no relevant role for the release of the protein.

According to release studies performed by Zhao and coauthors[43], increasing salt concentration avoids interaction between protein and DNA-origami. Adapting this experiment, single-strand displacement experiments were performed for testing the effect of an increasing sodium chloride concentration to the escape capabilities of the protein (*Fig. 48*). The increasing salt concentration has no effect on the ligands displacement, (*Fig. 48*; TAMRA panel), but showed also no significant effect on protein release. Due to the lower protein and DNA origami band intensity observed at higher NaCl concentrations and the shift of these bands, it can just be concluded that even high ionic strength conditions have no detectable effect on the protein release.

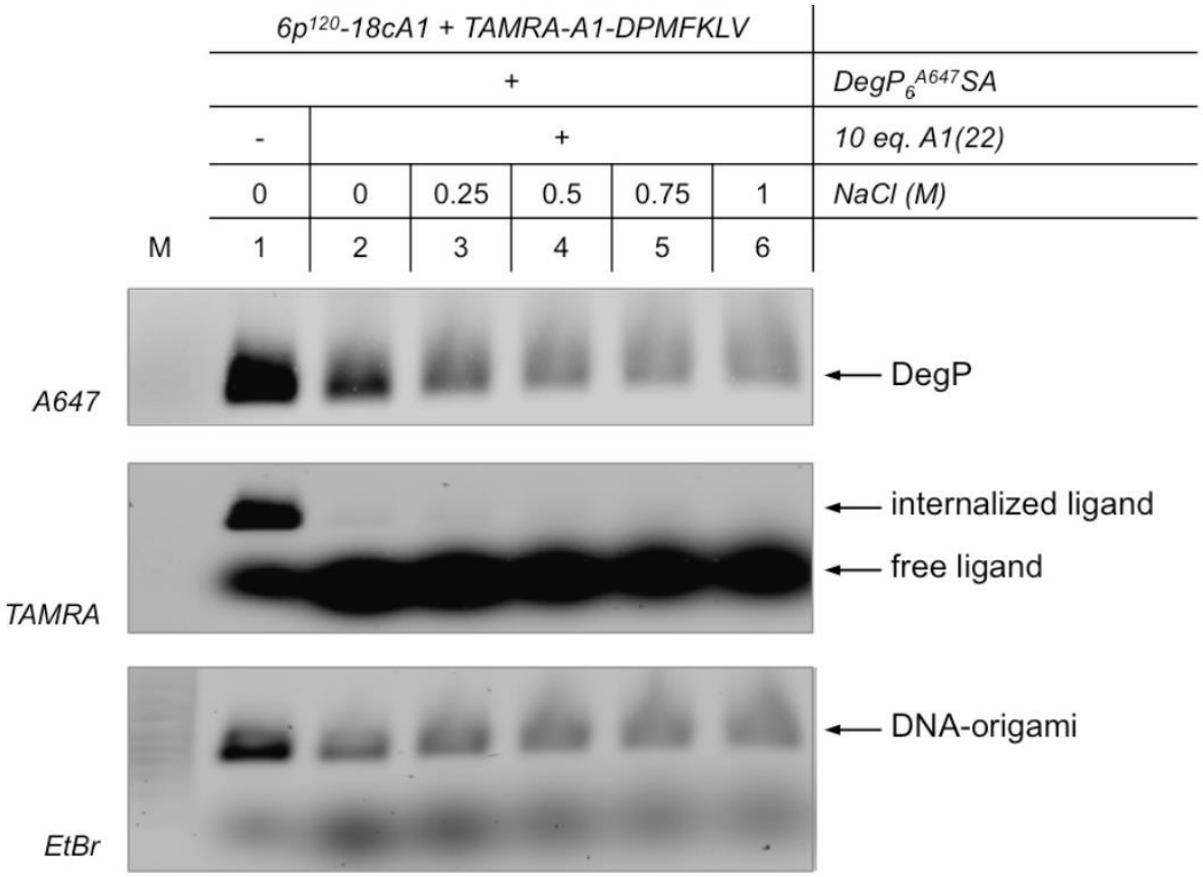


***Figure 48*:** Agarose gel characterization of increasing NaCl concentration dependent single-strand displacement studies actuated on DegP6A647SA bound to 6p120-18cA1. Addition of 10 equimolar amounts of A1(22) in absence or presence of sodium ions resulted in a complete release of the TAMRA-labeled ligands (*Fig. 48*; TAMRA panel). Nevertheless, the DegP protein signal was still present, indicating incomplete protein escape (*Fig. 48*; A647 panel). The presence of high salt concentrations in the reaction mixture affects the migration properties of the DNA/protein samples, leading to broader and less-defined gel bands.

To investigate whether the binding of the ligand to the PDZ1 domain of the protein prevents the single-strand displacement and subsequently hinder the delivery of the protein, a

competitive binding motive was added. A 50-fold excess of the peptide DPMFKLV was added to release the peptide-oligonucleotide ligand, in absence or presence of the tweezer. This experiment (*Fig. 49*) demonstrates that even in presence of Lys-specific molecular tweezers and the competitive peptide DPMFKLV to the PDZ1 linkers, the delivery of the protein from the DNA host is still not complete. This suggests that other forces keep the protein firmly anchored to the inner surface of the DNA cage and that these forces are not depleted in the conditions used until now.

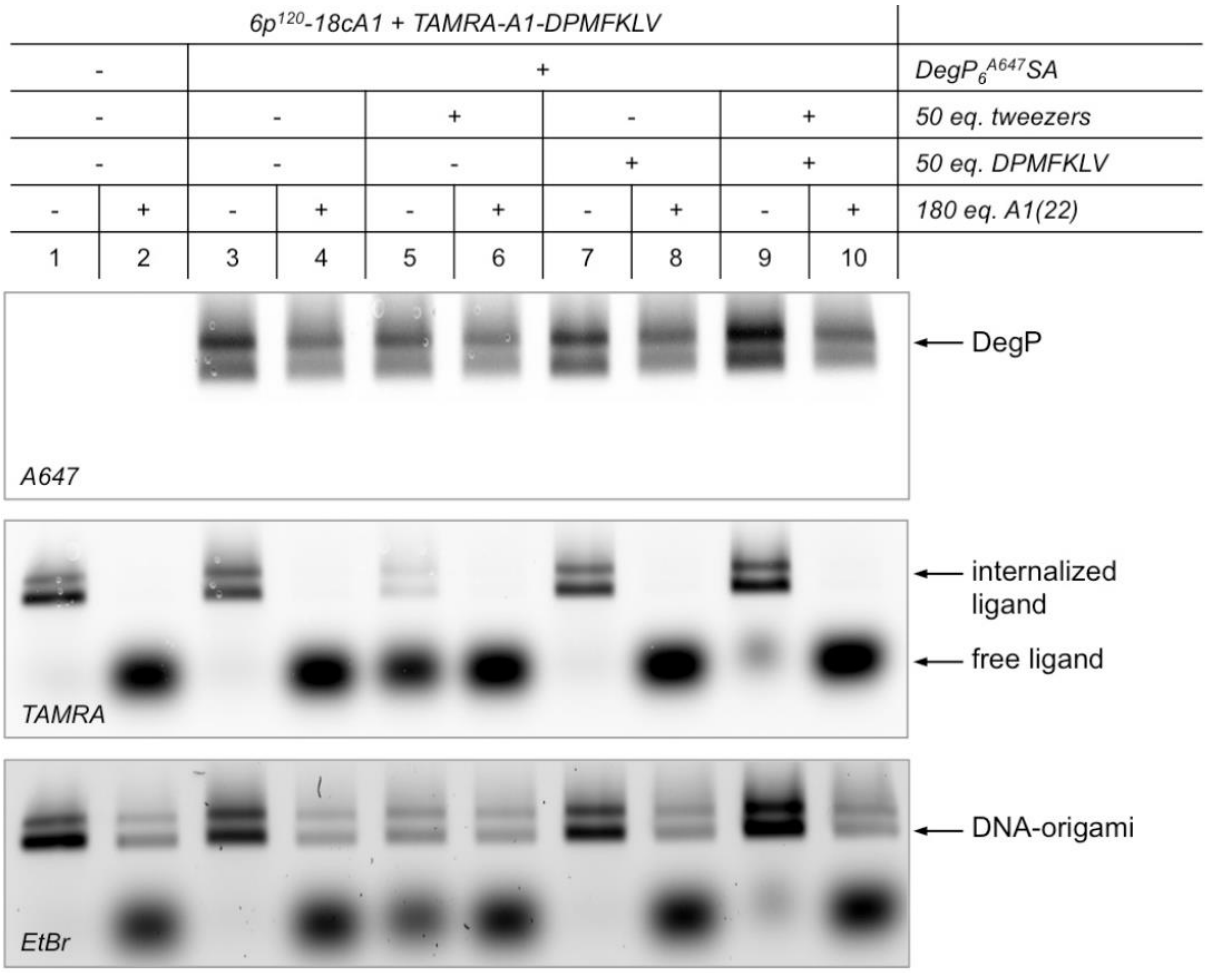
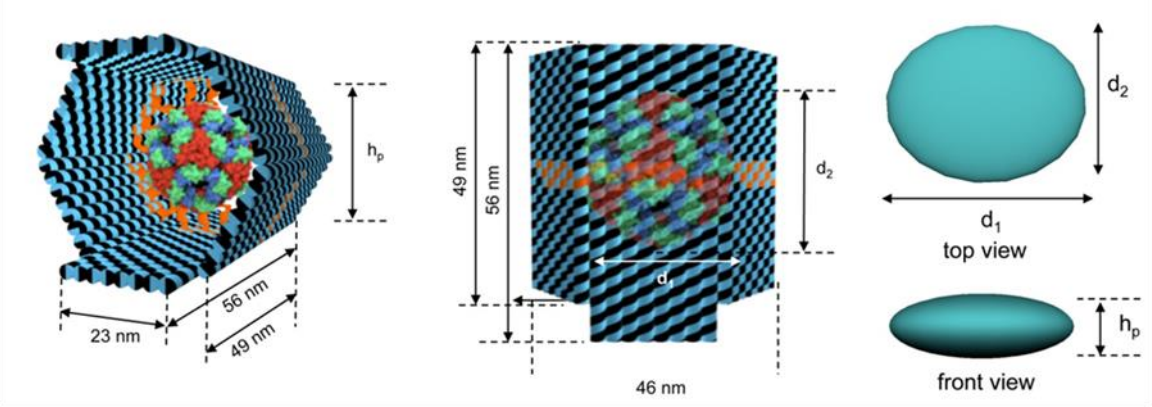


***Figure 49*:** Agarose gel characterization of a single-strand displacement experiment actuated on $DegP_6^{A647}SA$ bound to 6p120-18cA1 in presence of added peptide substrate or tweezers. Upon DNA origami assembly and loading with DegP as usual, the samples were purified by agarose gel extraction (chapter 6.5.6) and treated with 180 equimolar amounts of A1(22) effector strands. The displacement of the inner bound ligands was performed either in absence (lane 4) or presence (lane 6) of lysine-specific molecular tweezers (50 eq.); alternatively 50 eq. of the DPMFKLV peptide were also added to the solution mixture either in absence (lane 8) or presence (lane 10) of tweezers. As shown by the A647 signal (top panel) the protein is still trapped inside the cage, although the decreased band intensity suggests its partial escape. Remarkably, ligand displacement successfully occurs in all cases, as demonstrated by the disappearance of the TAMRA signal from the DNA origami band (middle panel, lanes 4, 6, 8 and 10) and simultaneous appearance of a new TAMRA-visible band with higher electrophilic mobility. The loss of the DNA-peptide conjugates from the inner cavity of the host is also visible from the decreased intensity of the DNA origami bands upon ethidium bromide staining and appearance of low molecular weight band associated to the free ligands in solution (bottom panel, lanes 2, 4, 6, 8 and 10).

3.4.4 Analysis of the loading affinities of the different DegP oligomers

Compression of the DNA origami constructs during AFM imaging leads to deformation of the different type of proteins, resulting in a decrease of the height and expansion of width and length. To get valid data, theoretical considerations concerning volume in collapsed and not collapsed state of an idealized sphere into an ellipsoid (*Fig. 50*) were carried out.



	theoretical values			6p^{120}-18cA1 + A1-peptide		
	6-mer	***12-mer***	***24-mer***	***6-mer***	***12-mer***	***24-mer***
d_1 (nm)	9	14	19	**21 ± 5**	**27 ± 6**	**37 ± 7**
d_2 (nm)				**18 ± 2**	**22 ± 5**	**32 ± 5**
h_p (nm)				**3 ± 1.5**	**5 ± 3**	**7.5 ± 1**
vol (nm^3)	4/3 π 91	4/3 π 341	4/3 π 857	4/3 π 92	4/3 π 355	4/3 π 862

***Figure 50*:** Detailed description of the 6p constructs loaded with various DegP forms under AFM imaging in air. Loading the 6prism with DegP leads to a partial increase in the height profile of the structure. The contribution given by encapsulation of the protein is h_p; d_1 and d_2 describe the dimensions of the protein along its two axes. An ellipsoidal shape that may be attributed to the high compression forces imaging conditions was observed. Analysis of the following AFM images allowed distinguishing three distributions, which seem to be reasonably attributed to DegP in its three oligomerization states.

Providing the volume is constant, deformation of the protein would cause a decrease in the height, but contrary an increase in width and length. The calculated data (*Fig. 50*; left side) match very well with the measured ones, showing an enormous decrease in the height and consequently an increase in width and length. Several AFM images were analyzed and the averages of the measured dimensions are shown in the right side of the table of figure 50; showing matching volumes. The slightly increased experimental values can be explained by the required space of the single layer DNA envelope. Due to the mixture of 12-mer and 24-mer within the DegP$_{12/24}^{A488}$SA samples, both types of proteins can be measured at the AFM. Comparing the loading efficiency of different designs, the 6p^{120}-18cA1 constructs show the highest yields. No loading experiments were performed with the 6p^{180}-18cA1 design, due to the lower loading efficiency of the corresponding designs with 6cA1. Furthermore, the

orientation of the PAs is stochastically distributed, which leads to invalid data concerning the loading efficiency. Table 5 summarizes the loading yields as obtained from AFM-based measurements.

***Table 5*:** AFM yields of the loaded 6p constructs with different number of PAs.

		120°			180°			240°	
ncA1		0	6	18	0	6	18	0	6
AFM yield	$DegP_6$	-	5.8	6.9	-	> 0	-	-	> 0
	$DegP_{12/24}$	-	12	20	-	2.9	-	-	1

3.4.4.1 AFM imaging and data analysis

The corresponding representative AFM images measured in air of the $6p^{120}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity, purified upon agarose gel electrophoresis extraction of the corresponding band (chapter 6.5.6) are depicted in Fig. 51.

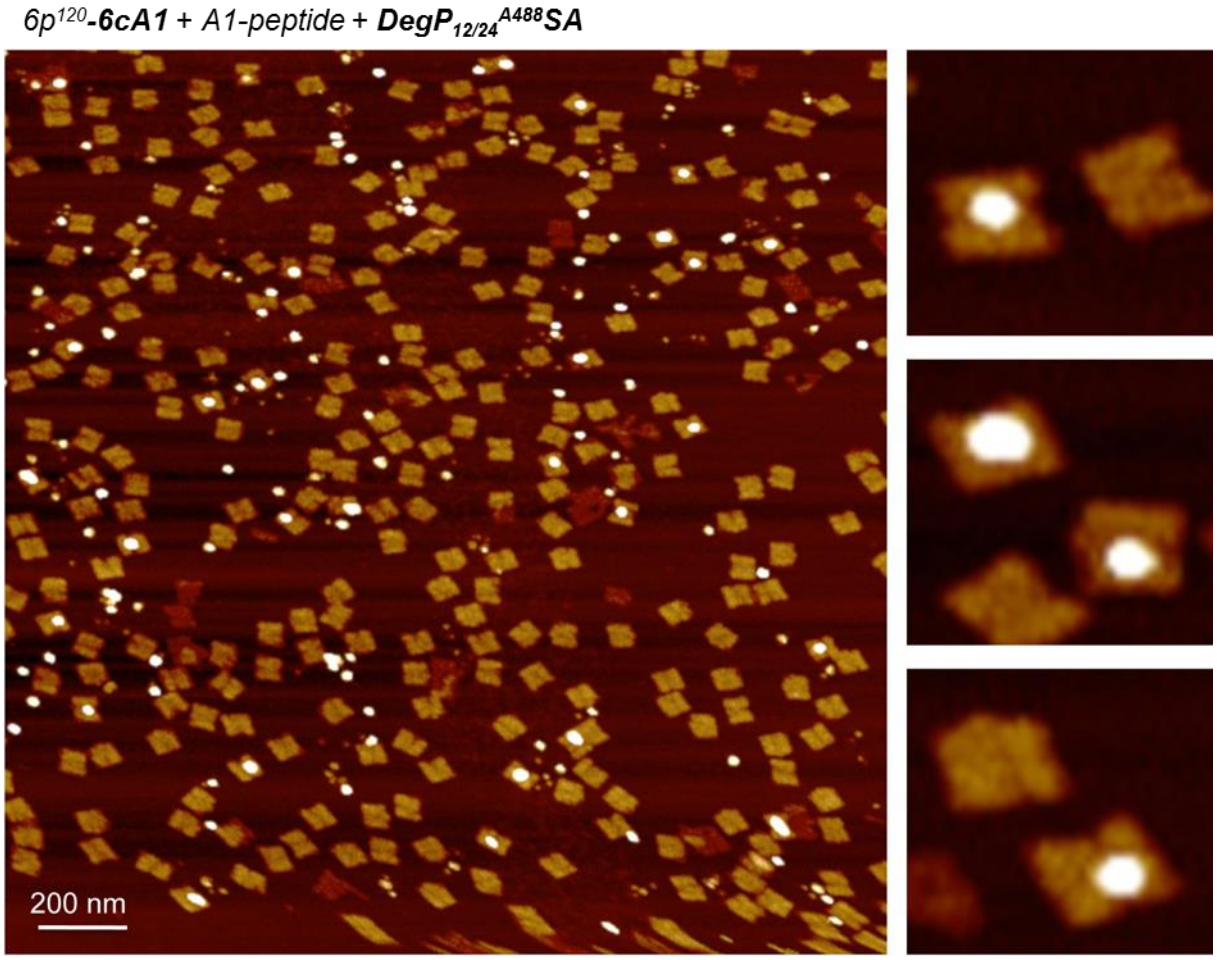


***Figure 51*:** AFM imaging of $DegP_{12/24}{}^{A488}SA$ bound to $6p^{120}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of successful loaded $DegP_{12/24}{}^{A488}SA$ proteins.

The zoom-in images on the right show successful loaded $DegP_{12}^{A488}SA$ (upper and lower panels) and $DegP_{24}^{A488}SA$ (middle panel; left structure), which can be differentiated by their size. In the overview, a lot of unloaded protein can be seen, probably due to damage of the structure during AFM imaging or sample manipulation. Increasing the number of PAs and consequently the number of A1-peptide ligands targeting the PDZ1 domains, leads to a higher probability of encapsulating two DegP proteins inside the cavity (*Fig. 52*; middle zoom in). Figure 48 shows a representative AFM image measured in air of the $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity, purified by gel electrophoresis extraction.

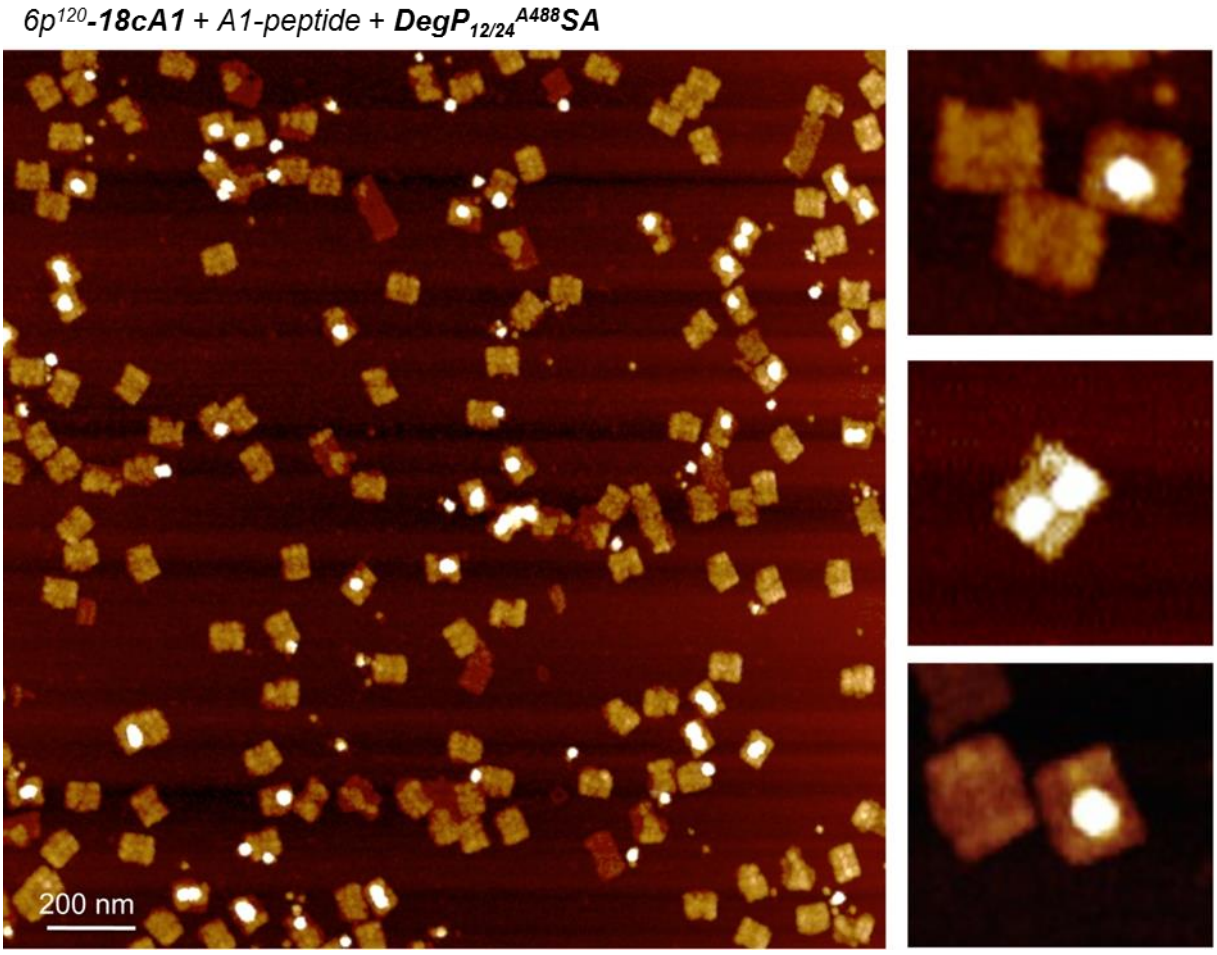


***Figure 52*:** AFM imaging of $DegP_{12/24}^{A488}SA$ bound to $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of successful loaded $DegP_{12/24}^{A488}SA$ proteins.

The enlarged images on the right show successful loaded proteins. Such a binding event can take place for the 12-mer or 24-mer of DegP. Another effect of the increased number of PAs is the higher yield (*Table 5*) of successful loaded structures.

Successful binding, although with lower yields (*Table 5*), could also be shown for the $6p^{120}$-6cA1 construct functionalized with 6 A1-DPMFKLV ligands in its inner cavity and $DegP_6^{A633}SA$ protein. Measuring the dimensions of the encapsulated proteins revealed that also larger oligomers of DegP, in this case DegP 12-mers, were caught inside the DNA host (*Fig. 53*; two upper zoom in images on the right). Probably, the higher affinity of the DNA cage towards the $DegP_{12}$ protein, which can be seen as impurities within the $DegP_6$ sample; or remaining A1-DPMFKLV ligands in solution, which could initiate a change in the oligomeric state of the $DegP_6$ proteins to 12-mers, could explain these results. The lowest panel shows a successful encapsulated $DegP_6$ protein.

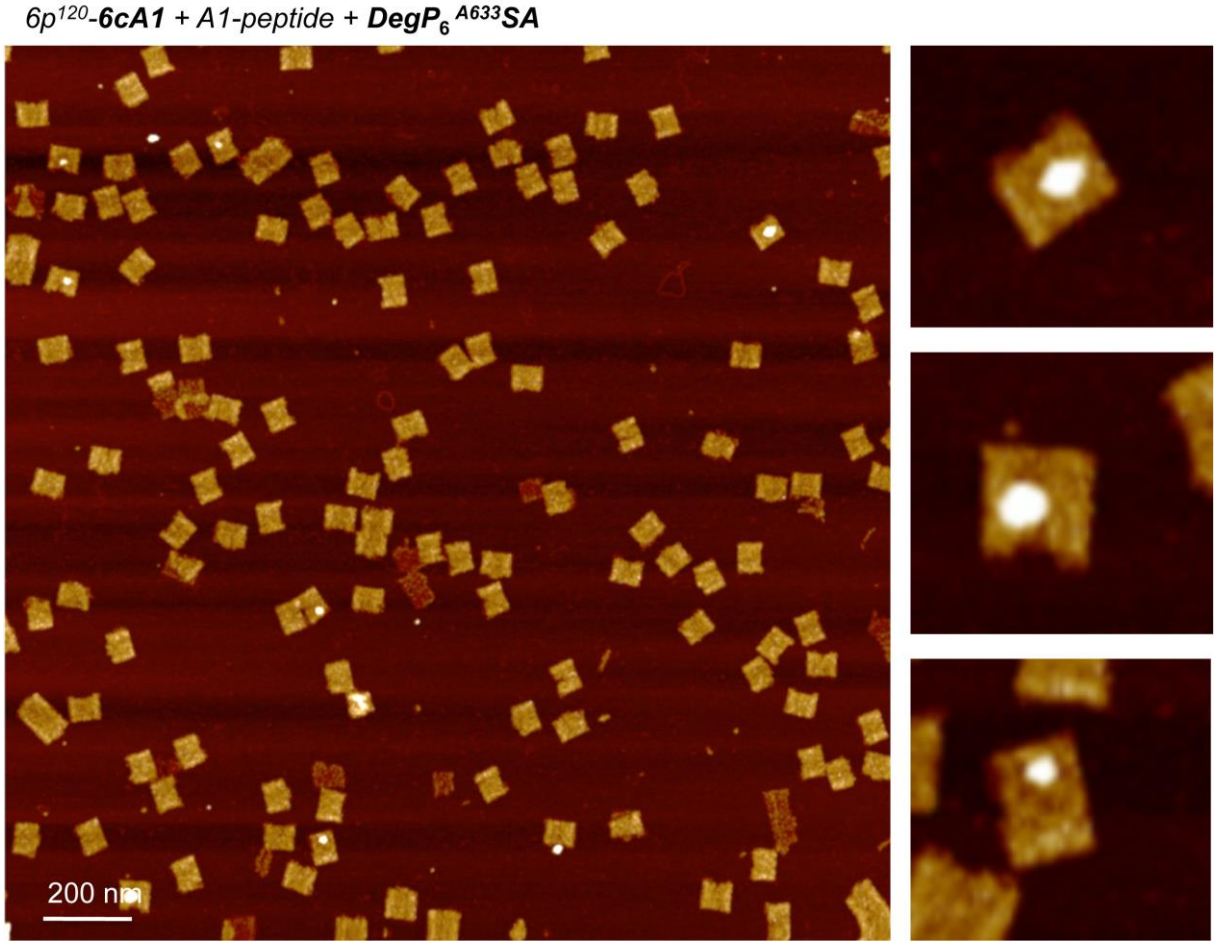


***Figure 53*:** AFM imaging of $DegP_6^{A647}SA$ bound to $6p^{120}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of successful loaded $DegP_6$ proteins.

Again, the increase in the number of PAs and consequently the number of A1-peptide ligands binding to the PDZ1 domains, leads to a slightly higher encapsulation yield (*Table 5*). Similar to the encapsulation experiment of the $6p^{120}$-6cA1 design with $DegP_6^{A633}SA$, a mixture of 6-

mer (*Fig. 54*; right panel, middle zoom in) and 12-mer (right panel, upper zoom in) was encapsulated inside the cavity; probably for identical reasons. As observed in the loading experiment of the $6p^{120}$-18cA1 with $DegP_{12/24}{}^{A488}SA$, two $DegP_6{}^{A633}SA$ proteins could bind to the DNA host simultaneously (right panel, lower zoom in).

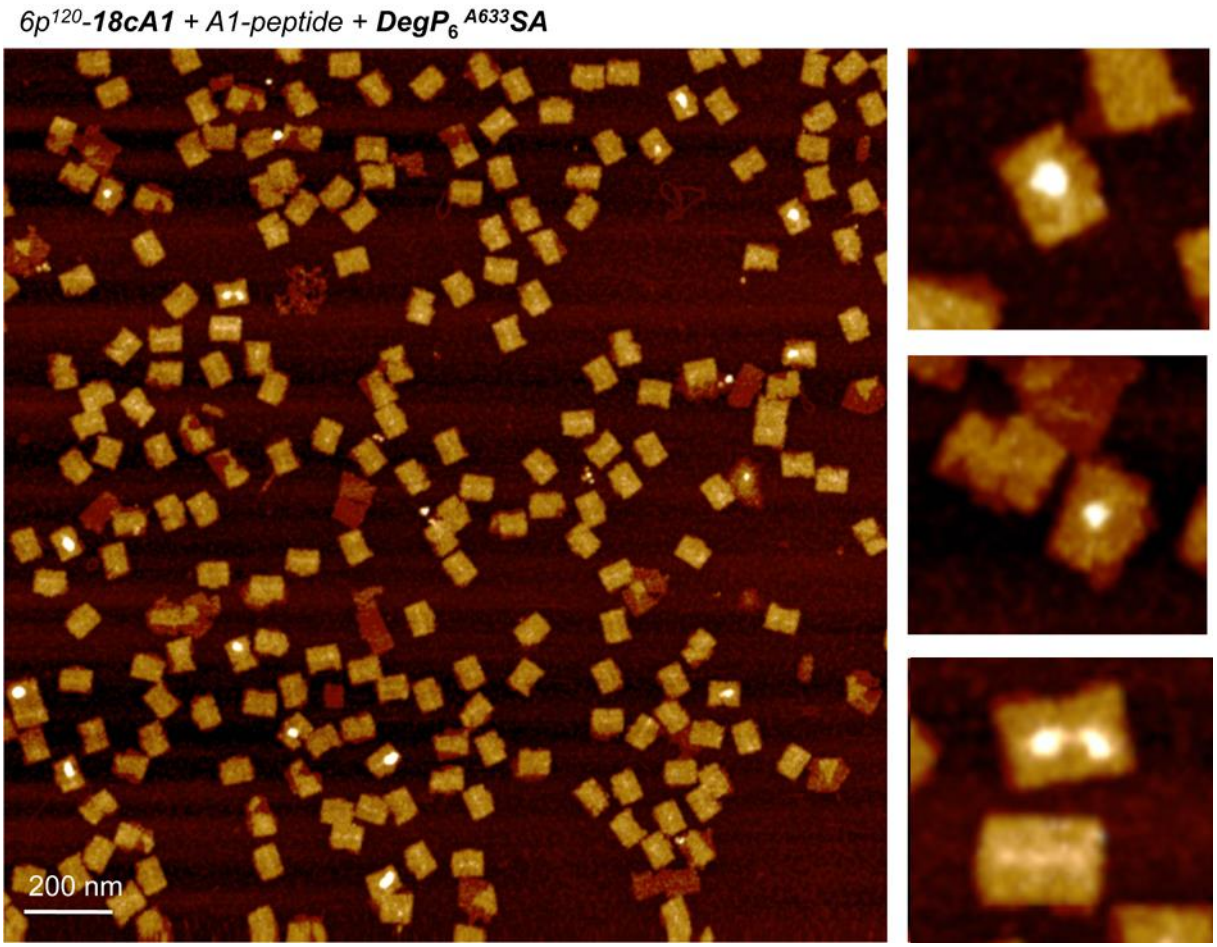


***Figure 54*:** AFM imaging of $DegP_6{}^{A647}SA$ bound to $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of successful loaded $DegP_6$ proteins.

Binding of the $DegP_{12/24}{}^{A488}SA$ to the $6p^{180}$-6cA1 host takes place with a 4-fold lower yield in comparison to the $6p^{120}$-6cA1 construct (*Table 5*). This lower yield fits with the lower yield determined by gel electrophoresis analysis shown in chapter 3.4.3.4.1 and results from the stochastic orientation of the PAs. The upper zoom in (*Fig. 55*; right vertical panel) shows a possible successful encapsulated $DegP_{24}{}^{A488}SA$. Due to the high compressing force during AFM measurements, some proteins seem to be pressed out of the cavity (*Fig. 55*; lower zoom in). In these cases the proteins are directly located at the entrance of the cavity, which is identifiable as the longer side of the DNA cages in their collapsed form. In contrast to the encapsulated protein, two $DegP_{12}{}^{A488}SA$ proteins have bound at the outer face (*Fig. 55*;

middle zoom in), distinguishable through their position half on the surface of the shorter side of the origami and half on the mica. All protruding arms are located in the middle of the structure, forming a line that is parallel to the entrance of the cavity. The fact that the proteins are situated in the middle of the shorter and closed origami side enhances the impression of the proteins having bound to the PAs on the outer face.

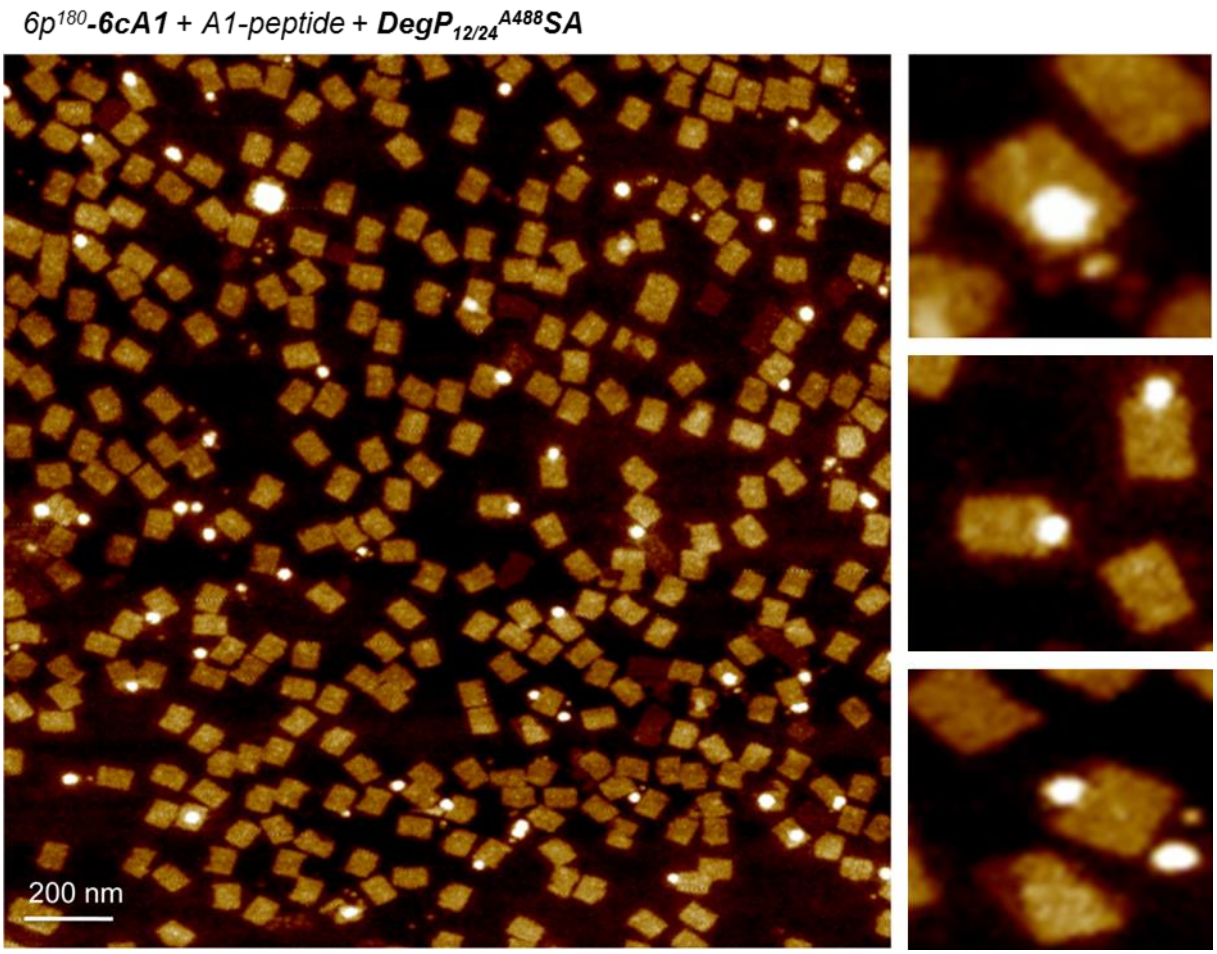


***Figure 55*:** AFM imaging of $DegP_{12/24}^{A488}SA$ bound to $6p^{180}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of successful loaded $DegP_{12/24}^{A488}SA$ proteins.

The importance of the stochastically distributed protruding arms becomes more apparent by the fact that almost no bound $DegP_6^{A633}SA$ to the $6p^{180}$-6cA1, functionalized with 6 A1-DPMFKLV ligands in its inner cavity, could be detected by AFM measurements (*Fig. 56*).
Only a single encapsulated 6-mer protein was detectable indicating that successful binding to the 6prism was possible (*Fig. 56*; right panel, upper zoom in). A possible explanation for the low binding efficiency of the $DegP_6$ proteins at the outer face of the 6prism (including the $6p^{180}$ and $6p^{240}$ design) may be that binding to the peptide DPMFKLV of the A1-peptide

ligand may induce a change of the oligomeric state of the DegP proteins, starting with a breakup into trimers[56]. In this case, a high local concentration of peptides attaching to the PDZ2 domain and a high local concentration of other trimers are necessary for folding back into the 6-mer or alternatively into a 12-mer or 24-mer. If the proteins are located inside the cavity ($6p^{120}$-designs), the local concentrations seem to be high enough for rearrangement of the protein. This does not happen if the PAs and consequently the proteins are located at the outer face.

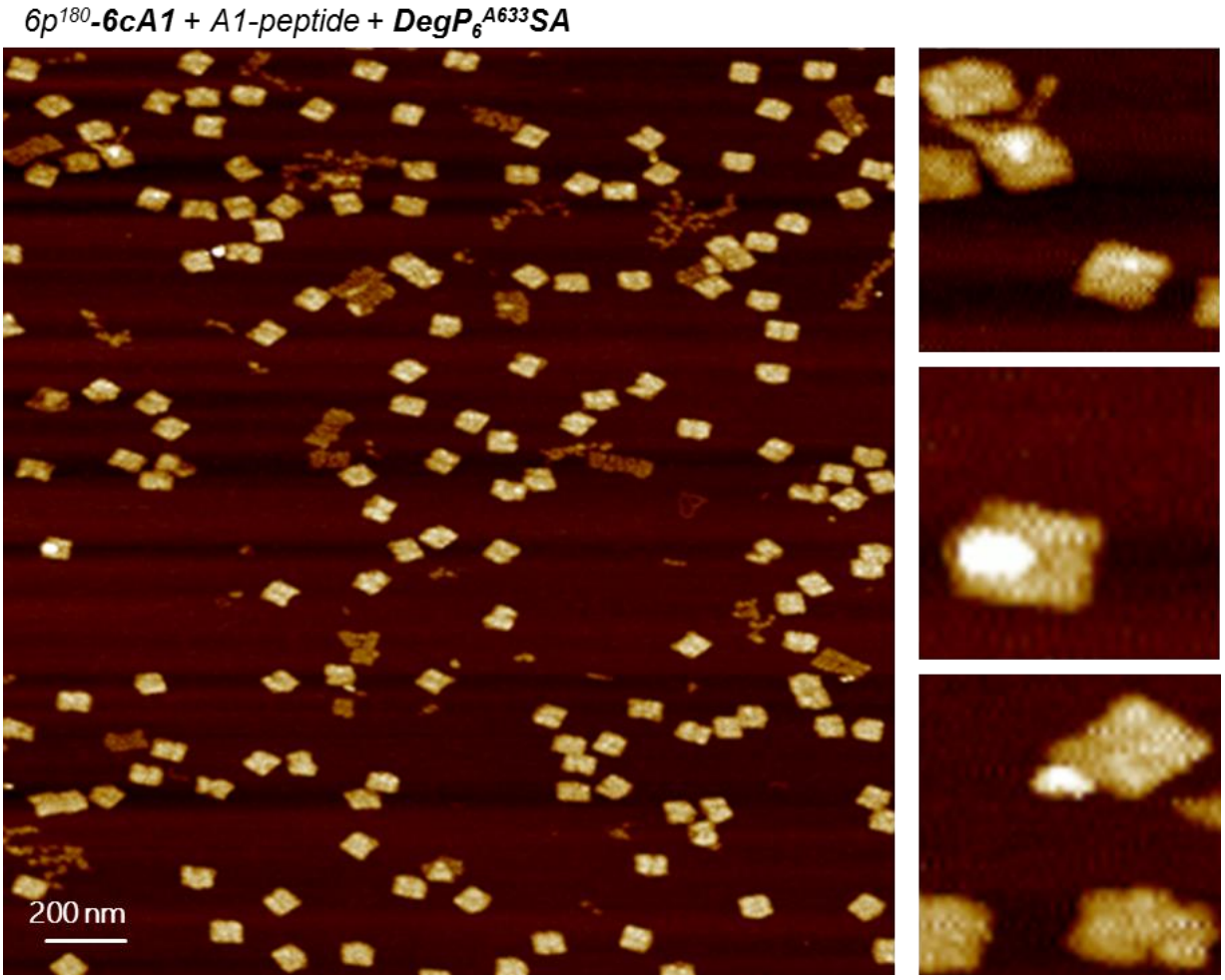


***Figure 56*:** AFM imaging of $DegP_6^{A633}SA$ bound to $6p^{180}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity. The right vertical panel shows zoom in images of the $6p^{180}$ construct.

Low concentrations and the opportunity of the trimers to diffuse unhindered into all directions, prohibit binding of the proteins to the cage, at least for 6-mers. Probably, $DegP_{12}$ and $DegP_{24}$ behave differently if they bind to the outer face of the DNA origami cages, which can be seen by the already mentioned proteins bound to the outer faces (*Fig. 55*; middle zoom in; *Fig. 56*; lower zoom in and *Fig. 57*; upper zoom in). These reasons would explain why

almost only higher oligomers of DegP were detected during AFM measurements of the $6p^{180}$-6cA1 samples with $DegP_6^{A633}SA$ (*Fig. 56*; upper zoom in).

Similar results are shown in Figure 53, which illustrates an AFM image in air of the $6p^{240}$-6cA1 design functionalized with 6 A1-DPMFKLV ligands in its inner cavity loaded with $DegP_6^{A633}SA$. Again, only larger oligomers of DegP can be detected (*Fig. 57*; zoom in images).

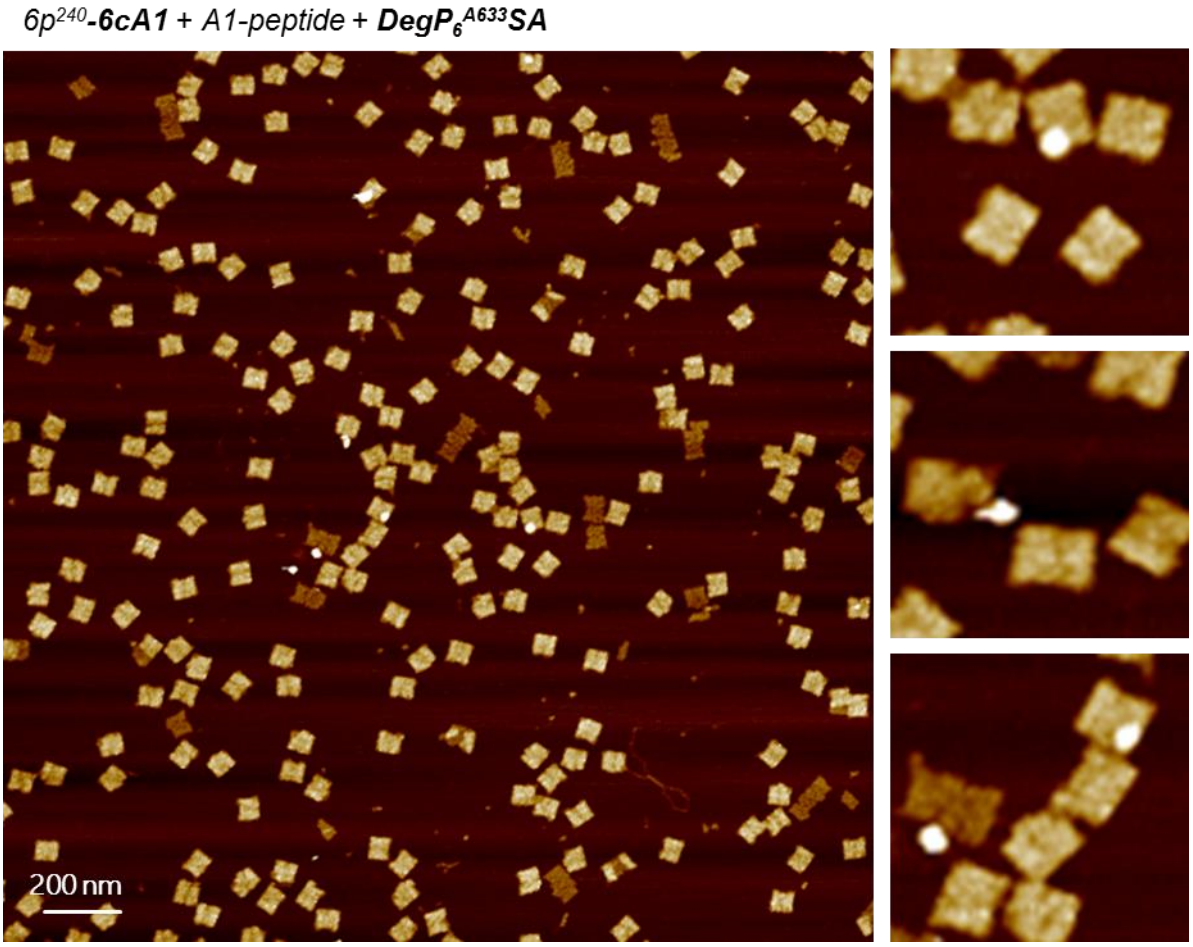


***Figure 57*:** AFM imaging of $DegP_6^{A633}SA$ bound to $6p^{240}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity.

It cannot be determined, whether the higher oligomeric states of DegP were induced by the peptide ligands, or have already been present as "impurities" in the $DegP_6^{A633}SA$ samples. Binding of the $DegP_{12/24}^{A488}SA$ to the $6p^{240}$-6cA1 constructs takes place with low yields (*Fig. 58* and *table 5*). The middle and the lower zoom in images indicate a slightly increase of the height profile in the center of the structure, which could be caused by bound trimers on the surface of the DNA origami cage. This would support the hypothesis of an induced breakup into trimers, if the protein attaches to a peptide ligand and low efficiency of rearrangement if

the PAs are orientated outwards. Alternatively, this result suggests that a high local concentration of ligands is necessary to induce protein binding.

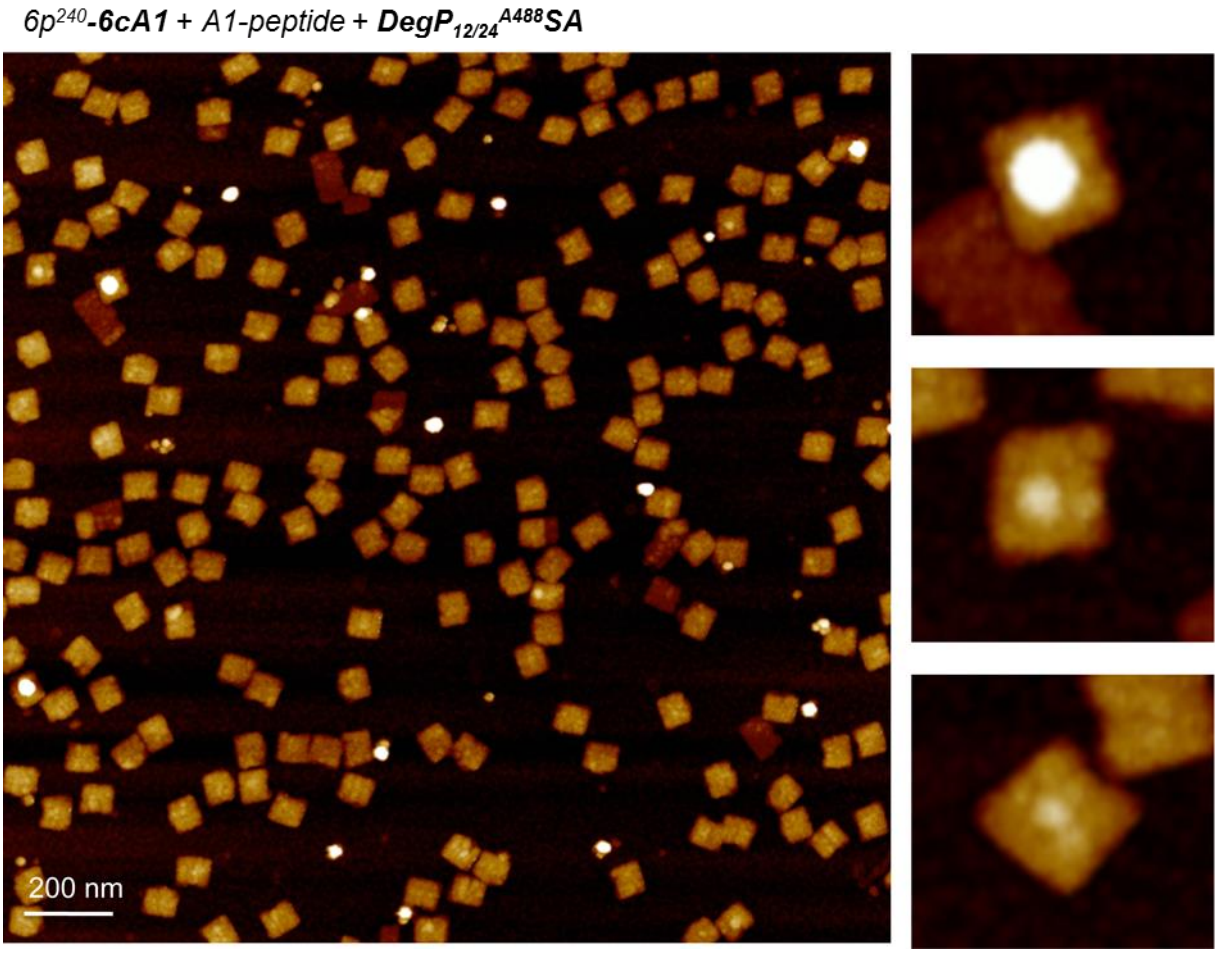


***Figure 58*:** AFM imaging of $DegP_{12/24}^{A488}SA$ bound to $6p^{240}$-6cA1 functionalized with 6 A1-DPMFKLV ligands in its inner cavity.

3.4.4.1.1 AFM imaging of tweezer modified DegP

In chapter 3.4.3.7, the positive effect of the tweezer, reducing the net charge of the protein and consequently an improved loading yield, could be shown (*Table 5*). Adapting these conditions, unlabeled DegP was preincubated with the molecular tweezer and added to the $6p^{120}$-18cA1 DNA origami cages. After gel electrophoresis, the DNA origami bands were cut out and analyzed by AFM measurements in air, showing low binding efficiency for $DegP_6SA$ (*Fig. 59*) and a high binding efficiency for $DegP_{12}SA$ (*Fig. 60*).

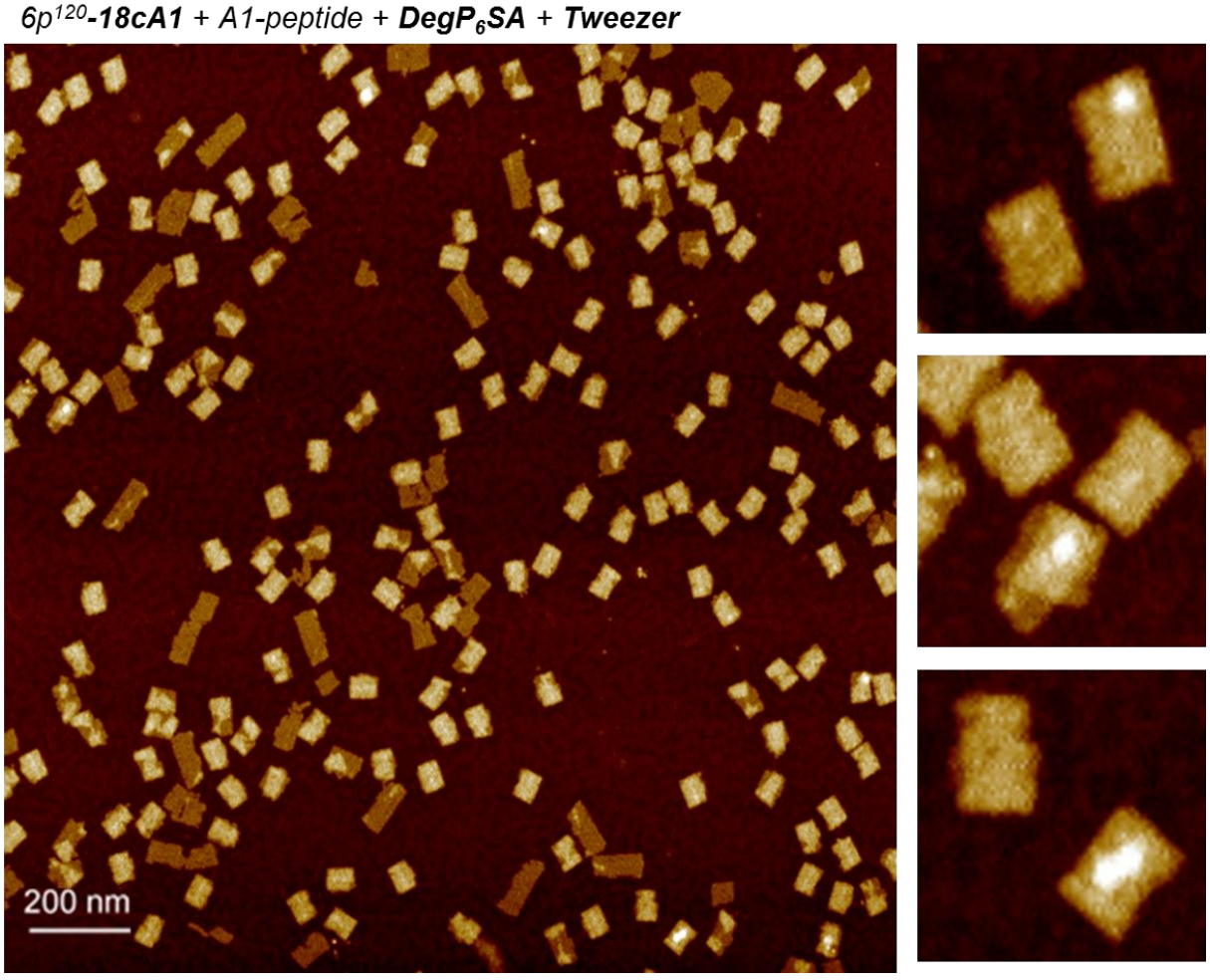


***Figure 59*:** AFM imaging of $DegP_6SA$ bound to $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity. Protein was preincubated with 25eq of molecular tweezers before adding to the DNA origami cage.

Probably, the low binding efficiency of the $DegP_6SA$ protein can be explained by the hindering of the PDZ1 domain through the molecular tweezers. The zoom in images of figure 55 show successfully bound $DegP_6SA$.

Adding the molecular tweezer to the $DegP_{12}SA$ protein increases the loading yield up to 60 %. This supports the idea that covering the net charge enhances the loading efficiency

through the increased specificity of peptide binding to the PDZ1 domain. Interestingly, the number of chambers, which carry two proteins, also increases.

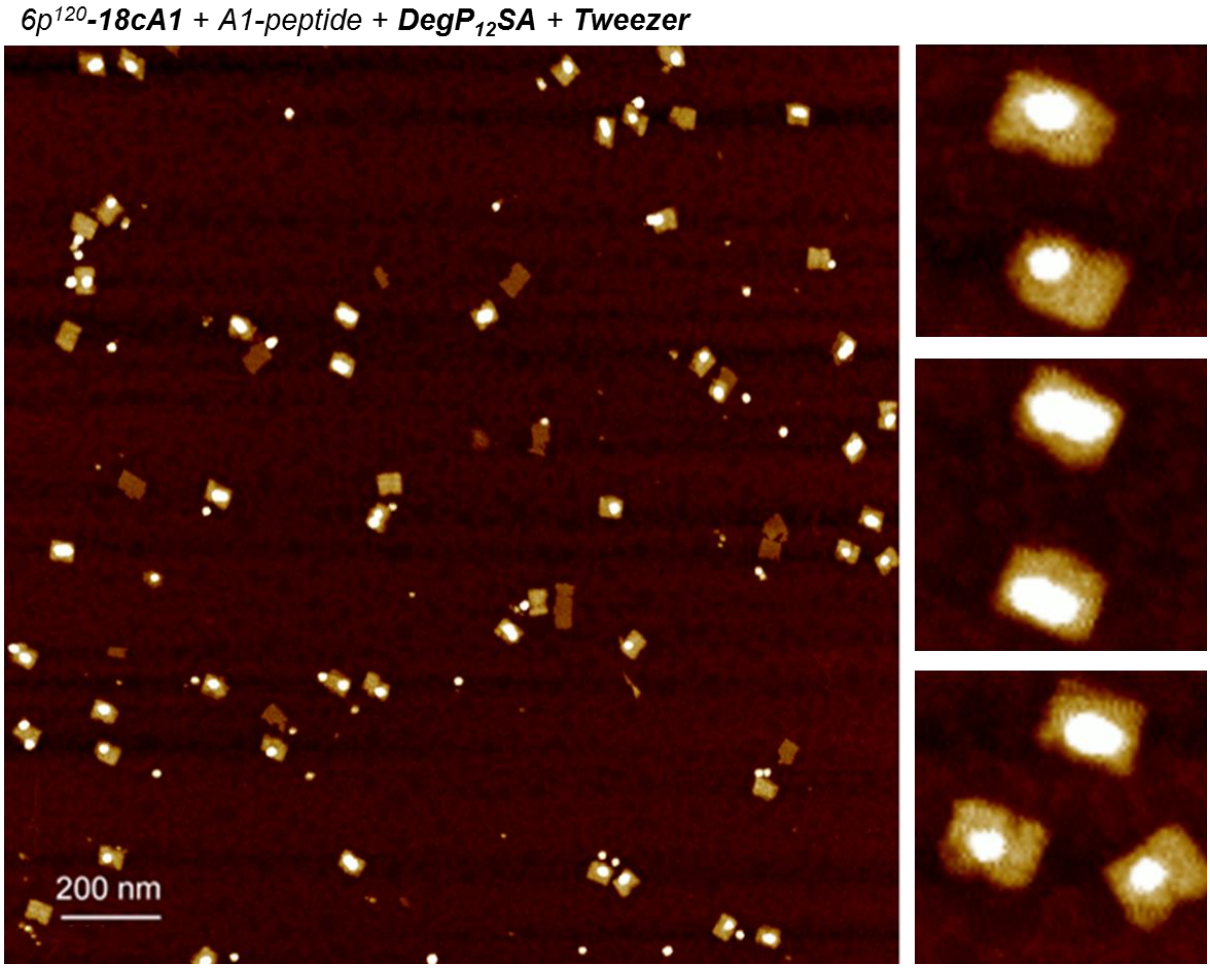


***Figure 60*:** AFM imaging of $DegP_6SA$ bound to $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity. Protein was preincubated with 25eq of molecular tweezers before adding to the DNA origami cage.

3.4.4.2 TEM imaging

TEM imaging was performed by Pascal Lill at the MPI in Dortmund[57]. Figure 61 shows representative digital micrographs of negatively stained $DegP_{12/24}$ on the left and $DegP_6$ on the right (in total 149). In the following chapter, class averages of about 10000 particles for $DegP_6$ and 9804 particles for $DegP_{12/24}$ are shown. TEM imaging of the proteins has been done for the following reasons: First, it has to be proven, whether the protein itself is intact and pure. Next, and most importantly, a statistical analysis of the percentage of 6-mer, 12-mer and 24-mer in the initial protein sample would allow to postulate some hypotheses on the change of the oligomerization state upon protein encapsulation (chapter 3.4.4.2.2)[57].

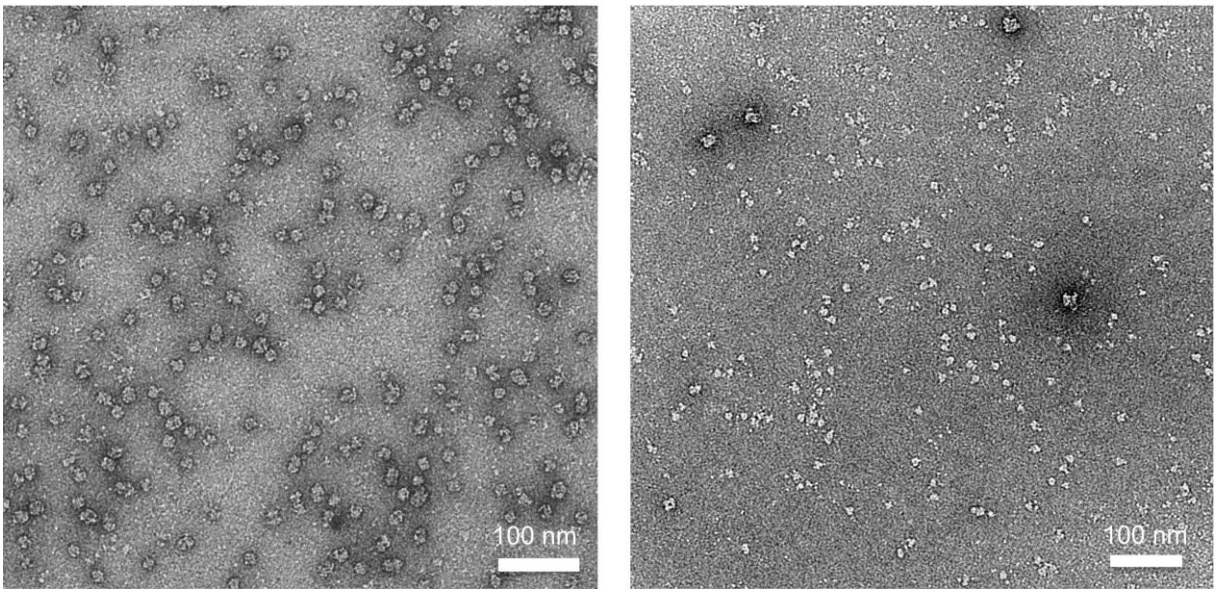


***Figure 61*:** Representative digital micrographs of negatively stained $DegP_{12/24}SA$ and $DegP_6SA$.

3.4.4.2.1 DegP oligomers

A representative class average of about 100 members of $DegP_6SA$ is shown in Figure 62 upper panel. The left part of the lower panel shows several orientations of the $DegP_6$ protein with its diameter of 10 nm. Simulated three-dimensional models from the respective crystal structures downfiltered to a resolution of 15 Å are shown on the right side of the lower panel. The two trimer subunits (upper part and lower part are both a trimer) can be seen at the left crystal structure, which shows the side view, being connected at each monomer; the right one shows the top view with its distinctive form.

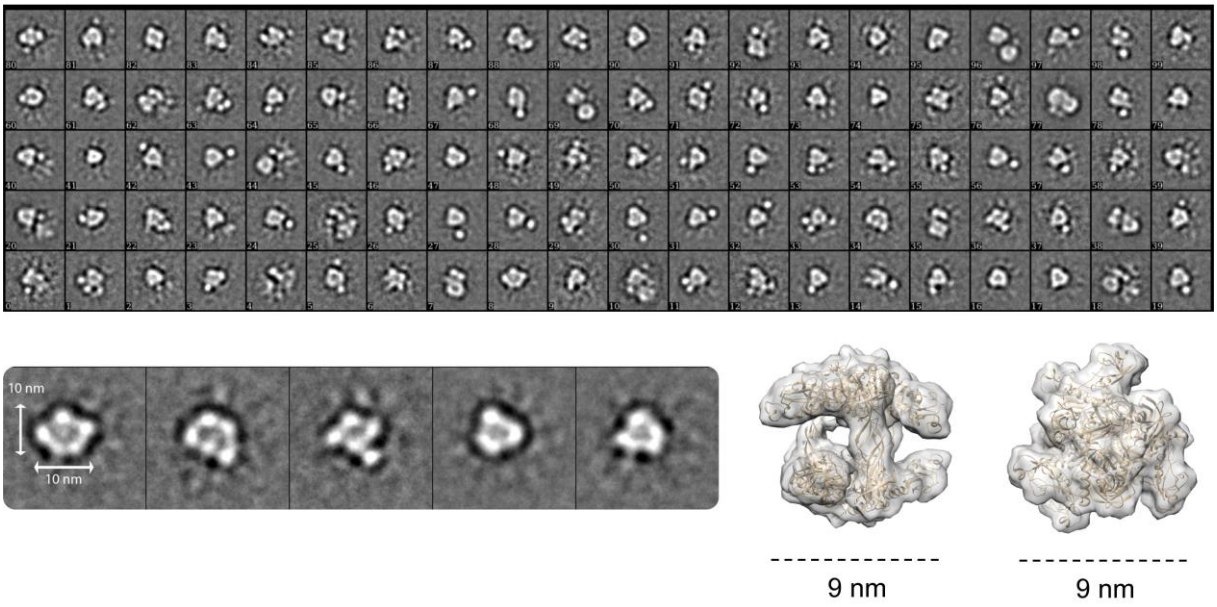


***Figure 62*:** Two dimensional analysis of $DegP_6SA$. Representative class average of $DegP_6SA$ (*upper panel*) and simulated 3-dimensional models from the respective crystal structures downfiltered to a resolution of 15 Å (*lower panel*; PDB-ID 1ky9).

Figure 63 shows a representative class average of 56 members of $DegP_{12}$ in the upper panel. About 28 images were taken from a mixture of $DegP_{12}$ and $DegP_{24}$ and analyzed separately. The lower panel compares two class averages of similar orientation to the downfiltered crystal structures of $DegP_{12}$, showing identical shapes with slightly larger dimensions. This increase may be caused by deformation of the protein along its height during grid preparation.

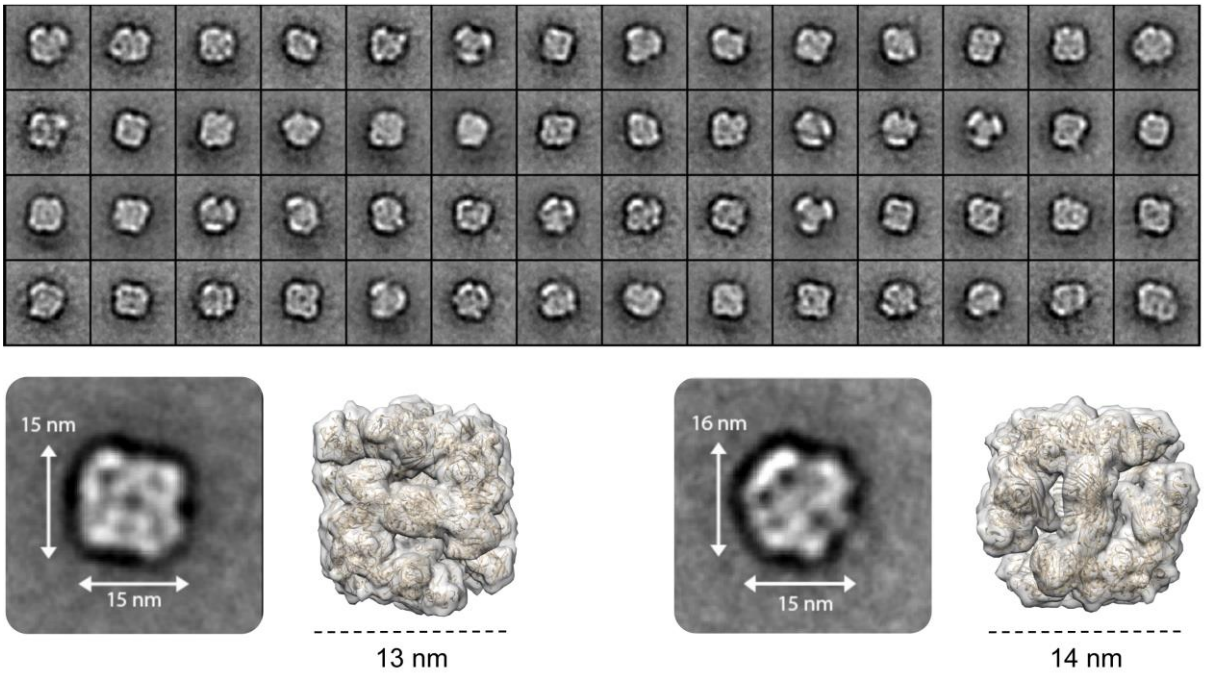


***Figure 63*:** Two dimensional analysis of $DegP_{12}SA$. Representative class average of $DegP_{12}SA$ (*upper panel*) and simulated 3-dimensional models from the respective crystal structures downfiltered to a resolution of 15 Å (*lower panel*; PDB-ID 2zle).

Similar deformations appeared for the $DegP_{24}$ protein, as shown in Figure 64 in the lower panel. Comparing the crystal structures with two averages reveals a slightly increase and decrease along width and length of about two nanometers. However, it is obvious that the measured protein is the $DegP_{24}$ protein, keeping almost its structure. The upper panel shows a representative class average of about 100 members of this protein.

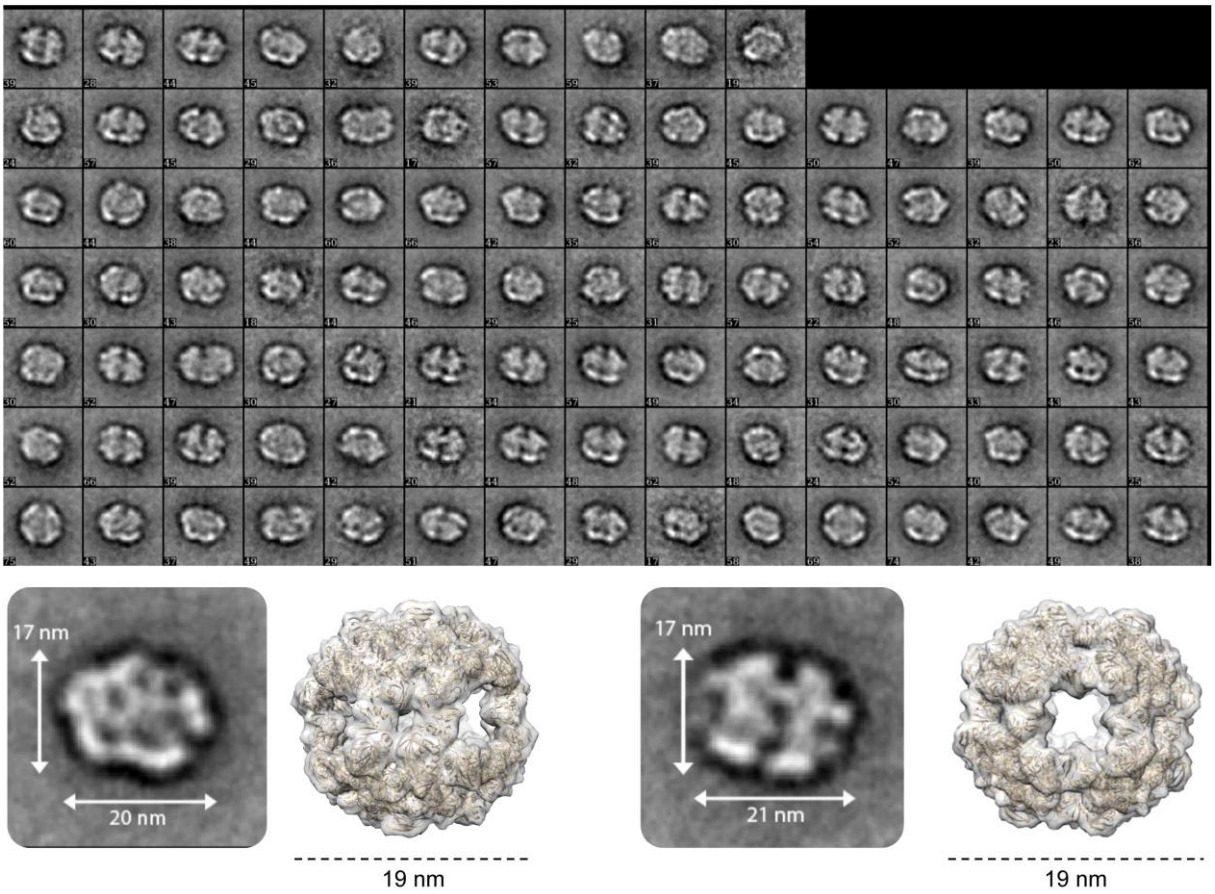


***Figure 64*:** Two dimensional analysis of $DegP_{24}SA$. Representative class average of $DegP_{24}SA$ (*upper panel*) and simulated 3dimensional models from the respective crystal structures downfiltered to a resolution of 15 Å (*lower panel*; PDB-ID 3cS0).

It is of note that the dimensions of the three measured proteins of the TEM images differ at most two nanometers from the respective crystal structures (*Fig. 61-65*), what allows clear differentiation from each other.

3.4.4.2.2 Loaded structures

About 18084 particles out of 1304 images of the $6p^{120}$-18cA1 functionalized with 18 A1-DPMFKLV ligands in its inner cavity with loaded $DegP_{12/24}SA$ were analyzed, boxed out and visually inspected manually. Figure 65 shows a representative class average of the loaded

DNA origami cage with DegP6 (lower panel) and a mixture of $DegP_{12}$ and $DegP_{24}$ (upper panel).

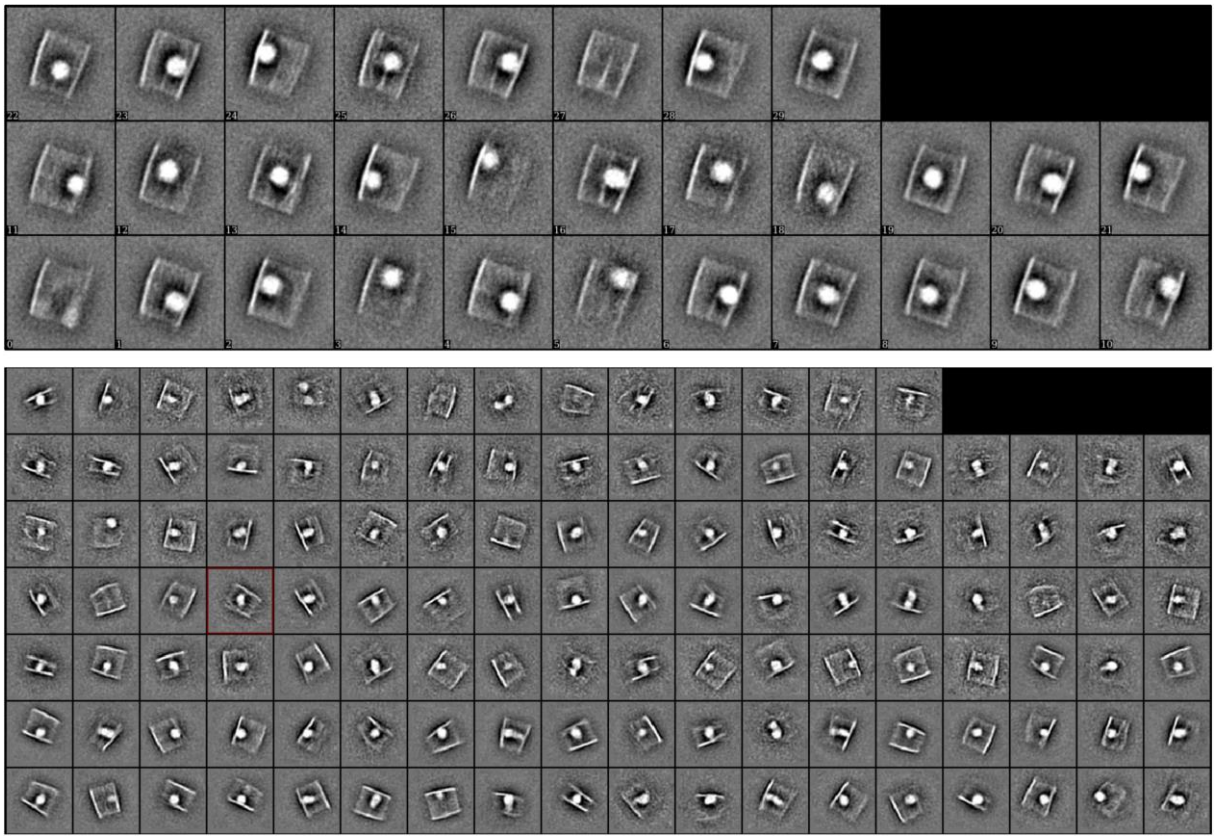


***Figure 65*:** Two-dimensional analysis of the DegP loaded DNA-origami host structures. Representative class average of the DNA cage loaded with $DegP_{12/24}SA$ (*upper panel*) and $DegP_6SA$ (*lower panel*).

Different from the AFM imaging, TEM characterization better preserves the 3-dimensional structure of either the protein or the DNA cage, loaded and empty. This allows a differentiation between the different species of the protein being encapsulated inside the DNA nanochamber. A representative large view micrograph of the DNA host bound to the $DegP_{12/24}SA$ is shown in Figure 66a. For comparison, the three protein species, which could be easily distinguished by the TEM images, are shown in figure 66b. In the first to fourth column (*Fig. 66c*) class averages, as well as representative raw TEM images and molecular models of the origami constructs, either with or without loaded protein in the three oligomerization states, are given. The DNA host appears as a rectangular shape of about 44 nm x 48 nm, confirming correct formation of the hexagonal tubular prism structure (*Fig. 66c*; first column). The TEM images of the loaded DNA host do not reveal any data concerning the position of the protein in the middle or at the edge of the structure, due to the not differentiatable orientation of the complex.

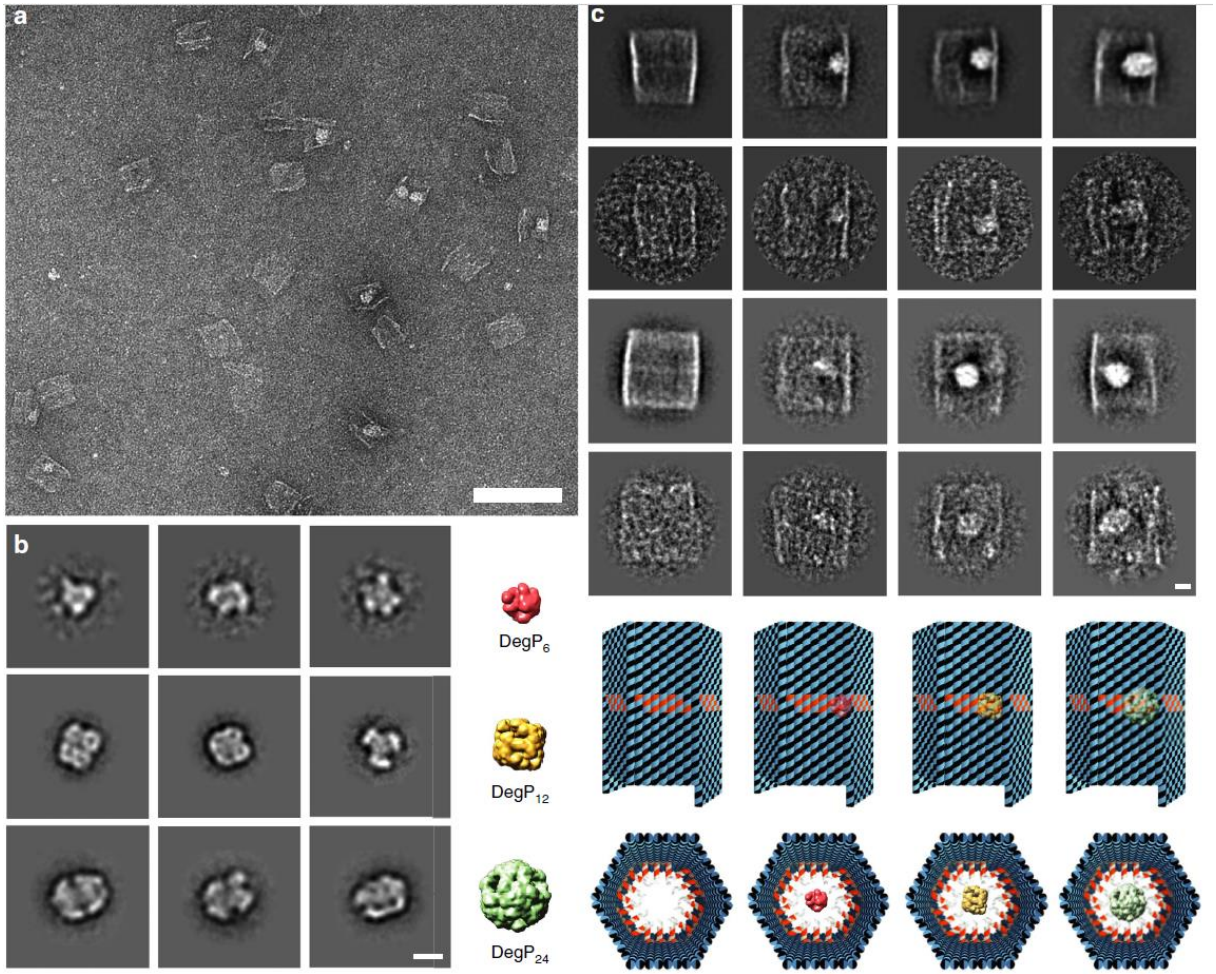


***Figure 66*:** Negative stain EM of DegP and DegP loaded DNA-origami cages. (a) Representative digital micrograph area of negatively stained $DegP_{12}SA$ and $DegP_{24}SA$ loaded DNA cages; Scale bar is 100 nm. (b) *Left panels*: Representative class averages, each containing approximately 50 to 100 particles, for $DegP_6SA$ (1st row), $DegP_{12}SA$ (2nd row) and $DegP_{24}SA$ (3rd row). Scale bar is 10 nm. *Right panels*: simulated 3D models from the respective crystal structures downfiltered to a resolution of 15 Å (PDB-ID $DegP_6SA$: lky9, PDB-ID $DegP_{12}SA$: 2zle and PDB-ID $DegP_{24}SA$: 3cs0). (c) *Upper panel from left to right*: 2D analysis of empty, $DegP_6SA$, $Degp_{12}SA$ and $DegP_{24}SA$ loaded DNA-origami $6p^{120}$-18cA1 hosts functionalized with 18 A1-DPMFKLV ligands in its cavity. Representative class averages, each containing approximately 25 to 100 particles (1st and 3rd row) and raw particle images of the corresponding classes (2nd and 4th row). Scale bar is 10 nm. *Lower panels, from left to right*: Schematic representation of the empty, $DegP_6SA$, $DegP_{12}SA$ and $DegP_{24}SA$ loaded DNA-origami hosts in top (1st row) and front (2nd row) view.

Most of the complexes analyzed by TEM showed encapsulation of one single protein with comparable yields of the $6p^{120}$-18cA1 structure (*Table 6*).

***Table 6*:** List of the respective loading yields compared with the different imaging methods.

imaging method	$DegP_6$	$DegP_{12/24}$
AFM	6.9	20
TEM	4	26

3.4.4.2.3 Loading efficiency

Calculations concerning loading efficiency were only based on AFM analysis, due to the problem that class averages (TEM imaging) are the result of image processing and averaging which leads to a loss of resolution. All dimensions of the loaded structures were manually measured and differentiated between the three oligomeric states. Before analyzing the different protein sizes, one has to ensure that the loaded DNA origami host systems are formed correctly in solution, excluding any other effects causing a change in the height profile. Dynamic light scattering experiments on the $6p^{120}$-18cA1 DNA cage were performed (*Fig. 67a*), both in absence and presence of the $DegP_{12/24}{}^{A488}SA$ protein, showing that the unloaded structure may be approximated to a spheroidal particle of about 40 nm in diameter (grey bars), corresponding to the expected theoretical values of 46 nm. Successful encapsulation of the protein leads to a slight decrease in size (yellow bars) and may be attributed to a partial squeezing of the complex when grasping the protein guest, thus corroborating proper internalization. Data from the statistical analysis showed that three different populations with a height profile centered at 7, 9, and 10.5 nm, which can be attributed to the encapsulated protein, could be distinguished (*Fig. 67b*). This distribution can be related to the 6-mer, 12-mer and 24-mer (red, yellow and green bars), demonstrating preferential binding of the host to the $DegP_{12}$ and almost a 2-fold lower selectivity for $DegP_6$ and $DegP_{24}$.

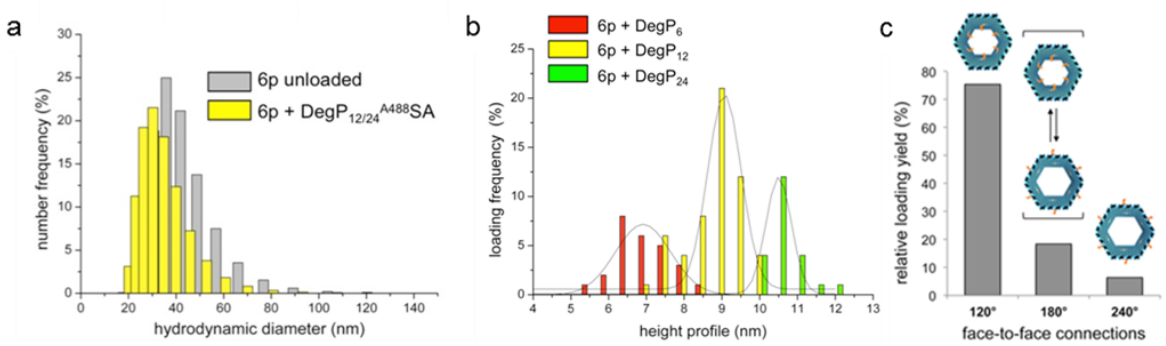


***Figure 67*:** Analysis of the data from atomic force microscopy characterization of the $6p^{120}$-18cA1. (**a**) The hydrodynamic size distribution of the by gel electrophoresis purified DNA-origami host either unloaded (grey bars) or loaded with $DegP_{12/24}{}^{A488}SA$ (yellow bars) was measured by dynamic light scattering, demonstrating correct formation of the hollow structure in solution. (**b**) Analysis of the height profile of the structures revealed three distributions centered at ca. 7, 9 and 10.5 nm, corresponding respectively to encapsulation of the 6-, 12- and 24-mer, with preferential selectivity of binding for the 12-mer. (**c**) Systematic analysis of the loading yield revealed most efficient protein encapsulation for a convergent design of multiple ligands.

Consequently, in the case of similar binding affinity, higher loading efficiency occurs for guest molecules, which are small enough to diffuse through the host channel but are

sufficiently large to allow short-range interactions between the exposed binding sites on the proteins surface and the ligands bound to the inner side of the host. Comparing the effect of inwards, outwards and stochastically orientated PAs on the loading efficiency of $DegP_6$ and $DegP_{12/24}$, confirms again the importance of multiple convergent ligands for successful protein encapsulation (*Fig. 67c*), showing an eight-fold higher loading efficiency of inward orientated PAs ($6p^{120}$) compared with outwards orientated PAs ($6p^{240}$).

4. Summary and outlook

During this PhD project, aspects of DNA nanotechnology, biology and supramolecular chemistry have been merged.

This work can be divided into three main parts: (i) the design of a suitable tubular DNA cage, which is able to encapsulate all different oligomers of DegP; (ii) the synthesis of a DNA-hepta-peptide conjugate, which binds non-covalently to the PDZ1 domain of the target protein and (iii) the loading of the protein.

The DNA origami cage was successfully designed and realized as a single-layer hexagonal DNA prism with an inner radius of 20 nm and an outer radius of 23 nm. Using special spatial staples strands for the face-to-face connections, the orientation of the faces towards the cavity could be programmed ($6p^{120}$ and $6p^{240}$). Alternatively, stochastically oriented faces ($6p^{180}$) were obtained using three thymines as flexible hinges (see *chapter 3.2.1 and 3.2.2*). Additionally, each face was equipped with zero, one, two or three orthogonal protruding arms (for a total of 0cA1, 6cA1 or 18cA1 arms, respectively). These arms were used for hybridization with DNA-peptide conjugates equipped with optional fluorophores (chapter 3.3). Correct formation of the different designs (6p120, 6p180 and 6p240) was proven by AFM and gel electrophoresis, after adding streptavidin to the biotinylated ligands. TEM characterization (performed by Pascal Lill at the MPI in Dortmund) and dynamic light scattering confirmed the correct dimensions of the DNA structures in solution.

The peptide sequence DPMFKLV was synthesized via solid phase peptide synthesis under standard coupling conditions. Transforming the N-terminal amino group of the peptide directly into a maleimide function allowed reaction with the thiol group of an oligonucleotide without the use of any crosslinking agents. This successful method represents a general method to link the terminal amino-group of a peptide to a thiol bearing oligonucleotide. Purification via HPLC allowed characterization of the DNA-DPMFKLV ligands per MALDI-TOF.

Loading of diverse DegP proteins (compare *chapter 3.1, table 1*) only took place in the presence of the A1-DPMFKLV ligands, demonstrating the validity of the encapsulation strategy (compare *chapter 3.4.3*). Successful and specific binding was shown at the single-molecule level using total internal reflection fluorescence (TIRF) microscopy, (performed by AG Birkedal, Aarhus University). Statistical evaluation of AFM images revealed preferential encapsulation of the $DegP_{12}$ protein, with a ratio of 1.3 : 2.2 : 1 for the $DegP_6$, $DegP_{12}$ and $DegP_{24}$, respectively. Loading experiments performed with $DegP_6^{A633}SA$ and cages with a

different number of PAs (see *chapter 3.4.3.5*) and a correspondingly different number of A1-peptide ligands, showed that one A1-DPMFKLV ligand is sufficient for encapsulation of the protein, with a loading efficiency proportional to the number of ligands. DNA origami cages with identical number of ligands but different orientations of the PAs ($6p^{120}$, $6p^{180}$ and $6p^{240}$, see *chapter 3.4.3.4*) were loaded with $DegP_{12/24}{}^{A488}SA$. Gel electrophoresis analysis showed a highest binding efficiency for the $6p^{120}$ design in a ratio of 8 : 1.4 : 1 for the $6p^{120}$-, $6p^{180}$- and $6p^{240}$-designs, respectively, thus indicating the importance of a high local concentration of peptide ligands. After successful encapsulation of the protein, experiments to release the protein from the cage were performed, using single strand displacement reactions. Disappearance of the fluorophore signals from the cage sample showed successful displacement of the ligands; however, without releasing the protein. Variation of the pH of the solution and its ionic strength did not result in any beneficial effects (see *chapter 3.4.3.8*).

To sum up, it could be shown that a protein can be encapsulated within a DNA origami cage by weak non-covalent supramolecular interactions without any previous chemical treatment of the protein. The arrangement of a distinct number of peptide ligands in the vicinity of the corresponding binding sites on the protein surface allowed modulation of local concentration effects and multivalent short-range interactions in a single system.

The importance of the net charge of the protein for encapsulation within the cavity of the DNA nanochamber and the different binding efficiencies observed for distinct fluorophores, require further investigation. Evidence for the important role of the net charge is the successful loading of the $DegP_{12}$ protein in presence of molecular tweezers, targeting the lysine residues on the surface of the protein.

Results of analogue experiments performed with a second DNA cage, (HoneyComb design, see chapter 1.2.3) are shown in chapter 8.

5. Zusammenfassung und Ausblick

Im Rahmen dieser Dissertation wurden verschiedene Aspekte der DNA-Nanotechnologie, Biologie und der supramolekularen Chemie miteinander verknüpft und zur Anwendung gebracht.

Die Arbeit kann in drei Teilbereiche unterteilt werden:

(i) Das Designen eines geeigneten hexagonalen turbularen DNA-Käfigs mittels der Software CaDNAno2, dessen räumliche Dimensionen zur Immobilisierung des Proteins DegP in seinen unterschiedlichen Oligomerisierungs-Zuständen entsprechen musste.

(ii) (ii) Die Synthese eines Oligonukleotid-Heptapeptid-Konjugates, welches supramolekular an die PDZ1 Domäne des Zielproteins bindet.

(iii) (iii) Das Laden des Proteins in den Käfig mit anschließendem Versuch der gerichteten Freisetzung des Proteins.

Der DNA-Origami-Käfig wurde mittels der Software CaDNAno2 als einlagiges, turbuläres und hexagonales DNA-Röhrchen mit einem inneren Radius von 20 nm und einem äußeren Radius von 23 nm entworfen und erfolgreich realisiert. Die einzelnen Seitenwände des Käfigs wurden durch jeweils speziell entworfene Oligonukleotide miteinander verknüpft, so dass zum einen die Orientierung der Seitenflächen zum Zentrum der Kavität ($6p^{120}$ und $6p^{240}$) determiniert werden konnte, (vgl. *Kapitel 3.2.1 und 3.2.2*) und zum anderen durch eine flexible Verbindung der Seitenflächen mittels drei Thyminen eine zufällige Orientierung ($6p^{180}$) der Seitenflächen zum Zentrum der Kavität vorlag. Zusätzlich wurde jede Seitenfläche mit der Option entworfen, mit null, ein, zwei oder drei (insgesamt 0cA1, 6cA1 oder 18cA1) orthogonal herausragenden Oligonukleotiden (Protruding-Arme), die als Verlängerung von zentral liegenden Oligonukleotiden zu sehen sind, ausgestattet zu werden. Die Protruding-Arme dienen als Träger für die mit Peptiden modifizierten komplementären Oligonukleotid-Liganden (Kapitel 3,3), welche wahlweise mit Fluorescein (Flc-A1-DPMFKLV) oder TAMRA (TAMRA-A1-DPMFKLV) ausgestattet waren. Mittels eines Biotin-modifizierten Liganden konnte durch Zugabe von Streptavidin die korrekte Formation der Konstrukte $6p^{120}$, $6p^{180}$ und $6p^{240}$ durch Gel-Elektrophorese und Raster-Kraft-Mikroskopie (AFM) bestätigt werden. Messungen durch das Transmissions-Elektronen-Mikroskop (durchgeführt von Pascal Lill am MPI in Dortmund im Rahmen seiner Masterarbeit) und der dynamischer Lichtstreuung bestätigten ebenfalls die korrekte Formation in Lösung. Unter Berücksichtigung des antiproportionalen Verhältnisses der Diffusionsrate durch den Käfig zur

Größe des Proteins wurde der DNA-Origami-Komplex so geplant, dass das 12-mer des Proteins DegP bevorzugt binden sollte.

Die Synthese des Peptidfragmentes der Liganden erfolgte mittels Festphasenpeptidsynthese nach Standardbedingungen. Die N-terminale Aminogruppe der Peptidsequenz konnte direkt in eine Maleimide Funktion überführt werden (Kapitel 6.5.2.8.2). Ohne die Verwendung eines üblichen Crosslinkers konnten Thiol-modifizierte Oligonukleotide über eine kovalente Bindung an das Maleimid gebunden werden. Dieser erfolgreich etablierte Weg stellt somit eine universelle Methode dar, N-terminale Peptide direkt an Thiol-modifizierte Oligonukleotide zu binden. Eine Charakterisierung der Konjugate erfolgte anschließend mittels MALDI-TOF.

Das Laden diverser DegP Protein-Varianten (vgl. *Kapitel 3.1*, *Tabelle 1*) erfolgte nur in Anwesenheit des Peptid-Liganden (vgl. *Kapitel 3.4.3*), was die Selektivität der Methode erfolgreich demonstrierte und unspezifische Wechselwirkungen ausschloß. Das spezifische und erfolgreiche Binden an die DNA-Nano-Käfige konnte mittels interner Totalreflexionsfluoreszenzmikroskopie durch Überlagerung der Fluoreszenzsignale in Einzelmolekül-Experimenten bestätigt werden (durchgeführt durch die AG Birkedal, Aarhus Universität). Nach statistischer Auswertung der AFM-Bilder konnte gezeigt werden, dass eine Präferenz zur Immobilisierung von 12-meren im Verhältnis 1 : 2.2 : 1.3 für $DegP_6$, $DegP_{12}$ und $DegP_{24}$ vorlag, was der Zielsetzung bezüglich der Selektivtät entsprach. Durch das Assemblieren eines $6p^{120}$ Käfigs mit einer unterschiedlichen Anzahl an Protruding-Armen (vgl. *Kapitel 3.4.3.5*) und dementsprechend mit einer unterschiedlichen Anzahl von Liganden konnte gezeigt werden, dass zum einen ein Ligand ausreicht, um ein Protein ($DegP_6$) erfolgreich zu immobilisieren und zum anderen ein proportionales Verhältnis zwischen Ladungseffiziens und der Anzahl der Protruding-Arme besteht. Um den positiven Effekt einer hohen lokalen Konzentration der Peptid-Liganden auf die Bindeeffiziens der Proteine zu zeigen, wurden die DNA-Käfige mit den jeweiligen Ausrichtungen der Protruding-Arme ($6p^{120}$, $6p^{180}$ und $6p^{240}$, vgl. *Kapitel 3.4.3.4*) und gleicher Anzahl der Liganden mit $DegP_{12/24}{}^{A488}SA$ versetzt und mittels eines Vergleiches der Intensität des Fluorescein-Signals die höchste Bindeeffiziens im Verhältnis 8 : 1.4 : 1 ($6p^{120}$: $6p^{180}$: $6p^{240}$) für die Käfige mit nach innen-orientierten Liganden bestimmt. Nach dem erfolgreichen Immobilisieren des Proteins innerhalb der Kavität des Käfigs wurde anschließend eine Freisetzung des Proteins durch einen Austausch des zum Protruding-Arm nicht vollständig komplementären, mit Fluorophoren markierten Peptid-Liganden durch ein vollständig komplementäres Oligonukleotid versucht. Der erfolgreiche Austausch der markierten Liganden konnte mittels

Gel-Elektrophorese gezeigt werden, jedoch nicht unter vollständiger Freisetzung des Proteins. Eine verbesserte Freisetzung konnte trotz Änderung der Nettoladung des Proteins ebenfalls nicht erreicht werden (vgl. *Kapitel 3.4.3.8*).

Zusammenfassend kann festgehalten werden, dass zum ersten Mal ein Protein durch schwache supramolekulare Interaktionen innerhalb einer DNA-Origamistruktur immobilisiert werden konnte. Diese Proteine konnten ohne chemische Veränderung nur aufgrund der räumlichen Nähe der Peptid-Liganden, die an die PDZ1-Domänen der Proteine bindeten, und der hohen lokalen Konzentration innerhalb der Kavität durch multivalente Wechselwirkungen mit geringer Reichweite im Käfig gehalten werden.

Die Bedeutung der Oberflächenladung der Proteine für die Immobilisierung innerhalb der Kavität bedarf noch weiterer Untersuchungen. Ein Indiz für diese Bedeutung liefern die Versuche mit dem molekularen Tweezer, welche an die Lysine an der Oberfläche der Proteine bindet und deren positive Ladung abschirmt. Folglich führt dies zur Abschwächung der unspezifischen Wechselwirkungen zwischen den Proteinen und dem Käfig, so dass diese erfolgreich geladen werden konnten.

Die analogen Ergebnisse der Experimente mit einer weiteren Struktur eines offenen Prismas im HoneyComb-Design (vgl. *Kapitel 1.2.3*) werden in Kapitel 8 gezeigt, da diese keine neuen, bzw. nur die bisherigen Ergebnisse bestätigende Resultate erbracht haben.

6. Experimental Part

6.1 Chemicals

All chemicals including anhydrous solventswere purchased from Bernd Kraft, ABCR, Biozym, Carl Roth, Merck, Novabiochem, Sigma Aldrich or VWR and were used without further purification.

Acetonitrile	VWR
MeOH	VWR
DCM	VWR
DMF	Bernd Kraft
Fmoc-Val-OH	Sigma Aldrich
Fmoc-Leu-OH	Sigma Aldrich
Fmoc-Phe-OH	Sigma Aldrich
Fmoc-Met-OH	Sigma Aldrich
Fmoc-Lys(Boc)-OH	Sigma Aldrich
Fmoc-Pro-OH	Sigma Aldrich
Fmoc-Asp(O*t*Bu)-OH	Sigma Aldrich
HOBt	ABCR
HBTU	ABCR
DIPEA	ABCR
$DIPEA_{(dry)}$	Sigma Aldrich
$DMF_{(dry)}$	Sigma Aldrich
$DCM_{(dry)}$	Sigma Aldrich
EDT	Sigma Aldrich
TIS	Sigma Aldrich
TFA	Sigma Aldrich
Et_2O	VWR
Wang-Resin	Novabiochem
Maleic anhydride	Merck
Triethylamine	Sigma Aldrich
Acetic acid	Sigma Aldrich
DMAP	Sigma Aldrich

3-HPA	Sigma Aldrich
Ammonium acetate dibasic	Sigma Aldrich
TEMED	Roth
Acrylamide/bis solution (37.5:1; 30 % w/v)	Sigma Aldrich
Urea	Sigma Aldrich
Bromophenol blue	Sigma Aldrich
APS	Sigma Aldrich
$MgCl_2$	Roth
$Mg(AcO)_2$	Merck
Na_2HPO_4	Sigma Aldrich
NaH_2PO_4	Sigma Aldrich
Agarose	Biozym
Boric acid	Roth
EDTA	Roth
TRIS base	Roth
NaCl	VWR
Ethidium bromide	Merck
SYBR-Gold	Merck
Ladder 10 bp	GeneON
Ladder 1 kbp	Roth
Scaffold	Affymetrix
Oligonucleotides	Sigma Aldrich (HPLC grade)
TCEP	Sigma Aldrich
Acetic anhydride	Sigma Aldrich
Formamide	Sigma Aldrich

6.2 Buffers

All listed buffer concentrations conform to one-fold buffer.

TBEMg	40 mM Tris base, 20 mM boric acid, 2 mM EDTA, 12.5 mM $Mg(AcO)_2$, pH 8.0
TEMg	20 mM Tris base, 2 mM EDTA, 12.5 mM $MgCl_2$, pH 7.6
TBE	89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.0
PBS	3.5 mM NaH_2PO_4, 8 mM Na_2HPO_4, 50 mM NaCl, pH 7.6

TAEMg	40 mM Tris base, 2 mM EDTA, 12.5 $Mg(AcO)_2$, 20 mM Acetic Acid, pH 8.0
TE	10 mM Tris base, 1mM EDTA, pH 8.0

Agarose loading buffer:		PAGE loading buffer:	
TBEMg 5x	70 %	Formamide	80 %
Glycerin	30 %	TBE 5x	20 %
		Bromophenol blue (optional)	~ 2 mg

6.3 Instruments

Instruments with various settings and exchangeable parts are described in detail. Further instruments are listed in chapter 6.1.3.8.

6.3.1 Preparative reversed-phase high performance liquid chromatography (HPLC)

Peptide compounds were purified on a Shimadzu HPLC system (Prominence UFLC) with a RP-C_{18}-column from Phenomenex (Phenomenex Luna® 5 µm C18(2), 100 x 21.20 mm) and peak detection at 210 nm and 254 nm. The used linear gradients were delinated from the measured retention time of the compounds in LC-MS measurements and were based on increasing solvent B (0.1 % TFA in ACN) levels in solvent A (0.1 % TFA in ddH_20) at a 25 ml/min flow rate.

6.3.2 Semi-Prep reversed-phase high performance liquid chromatography (HPLC)

DNA-peptide conjugates were purified on a Shimadzu HPLC system (Prominence UFLC) with a RP-C_{18}-column from Shodex (Shodex Asahipak® ODP-50 4E 5µm C18, 250 mm x 4.6 mm) and peak detection at 254 nm and 260 nm. Linear gradients of solvent B (MeOH) in solvent A (50 mM TEAA, 5% ACN in ddH_20) were applied at a 1 ml/min flow rate.

6.3.3 Freeze drying

Purification of substrates solved in pure water was carried out on a Freeze Dryer ALPHA 2-4 LD plus (CHRIST). Time required for lyophilisation depends on the thickness of the ice layer. To get a thin ice layer, the samples were frozen in liquid nitrogen under rotation inside a glass vessel or alternatively inside a micro tubes (2 ml). The samples were submitted to the freeze dryer until complete removal of the solvent.

6.3.4 Reversed-phase liquid chromatography – electrospray ionization mass spectrometry (LC-MS)

LC-MS analyses of substrates were performed on a LC-MS system from Thermo Scientific with an Eclipse XDB-C18 (5 µm) column from Agilent and a Thermo Scientific LCQ Fleet^{TM} ESI-Spectrometer with a peak detection at 210 nm, 254 nm and 280 nm. Flow rate was set to 1 ml/min for positive and negative mode measurements, applying for the positive mode a linear gradient of solvent B (0.1 % formic acid in ACN) in solvent A (0.1 % formic acid in ddH_20). For negative mode measurements, a linear gradient of D (5 mM NH_4OAc in ACN) in solvent C (5 mM NH_4OAc in ddH_20) was used.

Gradient for positive measurements: 0 min - 5 min (10 % B constant) → 5 min - 10 min (10 % B increasing to 100 % B) → 10 min - 12 min (100 % B constant) → 12 min - 15 min 10 % B constant)
Gradient for negative measurements: 0 min - 5 min (10 % D constant) → 5 min - 10 min (10 % D increasing to 100 % D) → 10 min - 12 min (100 % D constant) → 12 min - 15 min 10 % D constant)

6.3.5 Matrix-assisted laser desorption ionization – Time of Flight – Time of Flight (MALDI-TOF-TOF)

1 µL of the matrix solution was added to 50 pmol - 250 pmol of the DNA-peptide conjugates which were previously desalted and solved in ddH_2O (maximal 3 µL). After mixing, the solution was applied onto the MALDI plate (divided into at least two separate spots) and the solvent evaporated to dryness. To calibrate the MALDI-TOF, a mixture of 1 µL oligo

calibration standard and 1 µL matrix solution was also applied onto the MALDI plate. Setting the offset to 75 % and consequently the range to 25 %, the laser power was adjusted in a way that the standard could be measured as sharp peaks, without showing a large background. The mass deviation was corrected manually within the software and the DNA-peptide conjugates were analyzed without any change of the laser power

Target MALDI plate: MTP Small Anchor 384

Matrix: 3-HPA in ddH_20 (10 mg / mL) / (ddH_20 contains 1 mg / mL Ammonium acetate dibasic)

Oligo standard for calibration: Bruker oligo calibration standard (12-mer 3645.4; 20-mer 6117.0; 30-mer 9191.0)

Settings: Ion Source 1: 20 kV, Ion Source 2: 18.2 kV, Lens: 9 kV, Detector Gain: Linear 15 x

6.3.6 AFM imaging - MultiMode 8

Origami samples were deposited on a freshly cleaved mica surface (Plano GmbH) and adsorbed for three minutes at room temperature protected from air flow. After washing with ddH_2O, the samples were dried under argon flow and measured in ScanAsyst Mode using a MultiMode^{TM} 8 microscope from Bruker attached to a Nanoscope V controller. Scanning was performed with sharpened pyramidal tips (ScanAsyst-Air tips, Bruker) with a force constant of 0.4 N/m. Usually 3 µm large images were taken applying a peak force set point of typically 0.02 V and a scan rate of about 1 Hz. Several AFM images at different locations of the mica surface were taken to ensure reproducibility. The images were analyzed using Gwyddion (version 2.46) software.

6.3.7 Negative stain electron microscopy imaging

For each sample, 4 µL of either DNA-origami or DNA-origami-protein or protein samples were applied on glow discharger copper grids (Agar Scientific; G2400C), coated by an 8 nm thick continuous carbon film. Buffer excess was blotted from the side with a piece of filter paper (Whatman no. 4) after 60 seconds. This procedure was repeated twice for DNA origami samples, to increase the concentration of the particles on the grid, which was washed afterwards twice with two drops of ddH_2O and subsequently stained with two drops 1% Uranyl acetate. Images were taken with a FEI Tecnai G2 Spirit electron microscope equipped

with a Lab_6 cathode at an operation voltage of 120 kV. All digital micrographs were recorded with a 4k x 4k CMOS Camera F416 (TVIPS) using low-dose conditions.

6.3.8 Further instruments

All listed instruments below are mentioned without further description, due to their standard usage without modifications or specific handling.

Typhoon FLA 9000 - Laser Scanner for biomolecular imaging	GE Healthcare
Milli-Q® Integral Water Purification System cat. # Z00QSVC01	VWR
DLS Zetasizer Nano ZS	Malvern
UV/VIS Spectrophotometer Smart Spec™ Plus	BioRad
Thermocycler Mastercycler nexus gradient; nexus eco; nexus X2e	Eppendorf

6.4 Materials

All materials that are necessary for the reproducibility of the results are listed below. Not listed materials were ordered from various producers. Special materials for instruments are listed in chapter 6.1.3.

Ultrafree-DA Amicon 100,000 Da MWCO, cat. # UFC5100BK	Merck Millipore
Ultrafree-DA Amicon 50,000 Da MWCO, cat. # UFC5050BK	Merck Millipore
Ultrafree-DA Amicon 3,000 Da MWCO, cat. # UFC5003BK	Merck Millipore
Vivaspin centrifugal filters 100,000 Da MWCO, cat. # VS0142	Sartorius Biolab
Vivaspin centrifugal filters 3,000 Da MWCO, cat. # VS0192	Sartorius Biolab
NAP™-5 column, cat. # 17-0853-02	GE Healthcare
NAP™-10 column, cat. # 17-0854-02	GE Healthcare
0.22 µm membrane filters (cellulose acetate, sterile cat. # 28145-477)	VWR
Quantum Prep Freeze ´N´ Squeeze gel extraction kit cat. # 732-6165	BioRad
Anti-fluorescein IgG (H + L) CF™ 594 antibody cat. # SAB4600115	Sigma Aldrich
Anti-fluorescein IgG CF™ 640R antibody cat. # SAB4600169	Sigma Aldrich

6.5 Methods

6.5.1 Argon Balloon technique

To perform air or moisture sensitive reactions, all glassware was dried for at least 24 hours in an oven at 100°C. Before using the glassware, it was attached to a two-way vacuum line and evacuated and flushed repetitively for three times by argon gas. Next, the flasks were closed via a septum and connected to an argon filled balloon.

6.5.2 Methods used during SPPS and synthesis of the peptide-compound

In the following chapter, all methods that were used during synthesis of the peptide ligand are described.

6.5.2.1 Fmoc-determination (Method A)

To determine the loading efficiency of the first coupled Fmoc-protected amino acid to the resin, a UV/Vis spectrophotometer-based detection of the cleaved Fmoc residue was used. Therefore, a small amount of the dried resin (~ 5 mg) was transferred to a 25 ml conical flask. Five milliliters of a 20% (v/v) piperidine in DMF solution was added (in total 10 mL of this solution are prepared, 5 ml for cleaving the Fmoc-group and 5 ml for a serial dilution and as reference). The suspension was shaken for 30 minutes at room temperature on a shaker. After the resin settled down, the reference and the diluted samples (each 1 mL) were measured at λ = 301 nm. Each sample was measured for at least three times. Using the dilution factor and the data of the measurements, the loading of the resin was calculated according to formula [1].

$$c = \frac{A_{301} \cdot V \cdot F}{\varepsilon_{301} \cdot d \cdot m} \qquad [1]$$

c = resin load (mmol/g)

A_{301} = absorption value, λ = 301 nm

V = volume of cleavage solution

F = dilution factor

ε_{301} = extinction coefficient (7800 M^{-1} cm^{-1})

d = thickness of cuvette

m = resin mass

6.5.2.2 Cleavage of Fmoc protection group (Method B)

Base induced cleavage of the Fmoc group was achieved by adding 4 ml of a 20 % piperidine in DMF solution per 250 mg resin. The resulting suspension was shaken at rt on an orbital shaker for 30 min. After washing the resin with DMF → MeOH → DMF → DCM → DMF, further 4 ml of the cleavage solution were added to the resin and shaken again for 30 minutes. The following washing step was repeated twice (2x DMF → MeOH → DMF → DCM → DMF).

6.5.2.3 Amino acid coupling conditions (Method C)

Each coupling step of the corresponding Fmoc-amino acids was performed in a syringe reactor using identical conditions. If not mentioned explicitly in the experimental description, all reagents (HOBt, HBTU, DIPEA and the Fmoc-amino acid) were used in 4-fold excess (4 eq.) related to the amount of resin. The Fmoc-amino acid, HOBt, HBTU and DIPEA are dissolved in a minimal amount of DMF in a separate flask for preactivation of the amino acid for some minutes and were subsequently added to the syringe reactor containing the resin. After shaking for 2 hours, the resin was washed twice (2x DMF → MeOH → DMF → DCM → DMF).

6.5.2.4 Cleavage for LC/MS analysis (Method D)

Mircocleavage of the resin was performed after critical steps during synthesis to evaluate coupling efficiency. A small amount (1-2 mg resin) was transferred into a separate reactor and cleaved by a mixture of A (95 % TFA; 2.5 % H_20; 2.5 % TIS; 1000 µL) or B (92.5 % TFA; 2.5 % H_20; 2.5 % TIS; 2.5 % EDT; 1000 µL), if methionine was present in the synthesized peptide, for 30 minutes. The solution was transferred into a 5 mL flask and the solvent was removed under reduced pressure. The crude product was dissolved in 200 µL of a H_2O/ACN

mixture (1:1), filtered and transferred into a LC/MS vial. Further treatment of the sample is described in chapter 6.3.4.

6.5.2.5 Total cleavage of the peptide compound (Method E)

Total cleavage of the peptide compound, including the side-chain protecting groups, was performed by adding 1 mL cleavage mixture for 50 mg resin (92.5 % TFA; 2.5 % H_20; 2.5 % TIS; 2.5 % EDT) and the resulting suspension was shaken for 2 hours, filtered and transferred into a 50 mL falcon tube filled with ice-cold diethyl ether to precipitate. The falcon tube was put into a freezer (-20°C) for 30 min, centrifuged for 30 min at 4°C, decanted and the crude product was washed again with 45 mL ice-cold diethyl ether. After the second centrifugation step (30 min at 4°C), the pellet was dried under a continuous stream of argon.

6.5.2.6 HPLC purification of the peptide compound (Method F)

HPLC purification was performed using the HPLC system described in chapter 6.3.1. The crude product was dissolved in 5 mL ACN/H_2O (1:1) containing 0.1 % TFA, applying a gradient that fitted to the polarity of the product. All product containing fractions were collected, dried under reduced pressure, dissolved again in ddH_2O and lyophilized as described in chapter 6.3.3.

6.5.2.7 HPLC (semi-prep) purification of the DNA peptide compound

HPLC purification of the DNA peptide conjugate was performed using the HPLC system described in chapter 6.3.2. The reaction mixture was concentrated to an amount of at least 100 µL using a 3000 MWCO ultra centrifugal device. All fractions corresponding to the product peak were collected, dried under reduced pressure at 40°C. Characterization was performed either by denaturing gel electrophoresis (chapter 6.5.4) or MALDI TOF (Chapter 6.1.3.5).

6.5.2.8 Synthesis of peptide compound

All equivalents mentioned are relative to the amount of resin used for this reaction.

6.5.2.8.1 Attachment of the first amino acid to the Wang resin

Wang resin (250 mg, c = 1.1 mmol/g) was given into a round bottom flask and suspended in a CH_2Cl_2/DMF mixture (v/v 9:1). Separately, 4 equivalents of the Fmoc-amino acid, 4 equivalents of HOBt were dissolved in a minimal amount of DMF and subsequently added to the resin. Afterwards, 1.0 equivalents of DIC and 0.1 equivalents of DMAP were added and the flask was sealed with a drying tube. The mixture was shaken overnight (at least 12 hours) on an orbital shaker at room temperature. To cap the reaction 2 equivalents of acetic anhydride and DIPEA were added, mixing it for an additional hour. Afterwards, the resin was transferred into a reactor, washed twice (2x DMF → MeOH → DMF → DCM → DMF) and in a final step twice with Diethyl ether. It was dried under reduced pressure and the loading efficiency was calculated following method **A** (c = 0.7 mm/g).

6.5.2.8.2 Coupling of further amino acids and N-terminal modification

The Fmoc protecting group of the preloaded resin (chapter 6.5.2.8.1; 250 mg; c = 0.7 mmol/g) was removed according to method **B**. Coupling of Fmoc-Leu-OH was performed as described in method **C** using following quantities.

HBTU:	106.18 mg, 0.28 mmol
HOBt:	37.83 mg, 0.28 mmol
DIPEA:	47.62 µL, 0.28 mmol
DMF:	6 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Leu-Val loaded resin was removed according to method **B**. Coupling of Fmoc-Lys(Boc)-OH was performed as described in method **C** using following quantities.

HBTU:	106.18 mg, 0.28 mmol
HOBt:	37.83 mg, 0.28 mmol
DIPEA:	47.62 µL, 0.28 mmol

DMF: 6 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Lys(Boc)-Leu-Val loaded resin was removed according to method **B**. Coupling of Fmoc-Phe-OH was performed as described in method **C** using following quantities.

HBTU: 106.18 mg, 0.28 mmol
HOBt: 37.83 mg, 0.28 mmol
DIPEA: 47.62 µL, 0.28 mmol
DMF: 6 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Phe-Lys(Boc)-Leu-Val loaded resin was removed according to method **B**. Coupling of Fmoc-Met-OH was performed as described in method **C** using following quantities.

HBTU: 106.18 mg, 0.28 mmol
HOBt: 37.83 mg, 0.28 mmol
DIPEA: 47.62 µL, 0.28 mmol
DMF: 6 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Met-Phe-Lys(Boc)-Leu-Val loaded resin was removed according to method **B**. Coupling of Fmoc-Pro-OH was performed as described in method **C** using following quantities.

HBTU: 106.18 mg, 0.28 mmol

HOBt:	37.83 mg, 0.28 mmol
DIPEA:	47.62 µL, 0.28 mmol
DMF:	6 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Pro-Met-Phe-Lys(Boc)-Leu-Val loaded resin was removed according to method **B**. Coupling of Fmoc-Asp(O*t*Bu)-OH was performed as described in method **C** using following quantities, increasing the amount of equivalents to 6 and the time to 3 hours.

HBTU:	159.27 mg, 0.42 mmol
HOBt:	56.75 mg, 0.42 mmol
DIPEA:	71.43 µL, 0.42 mmol
DMF:	9 mL

After addition of the amino acid, a microcleavage according to method **D** was performed, showing a complete turnover.

The Fmoc protection group of the Fmoc-Asp(O*t*Bu)-Pro-Met-Phe-Lys(Boc)-Leu-Val loaded resin was removed according to method **B**. Coupling of maleic anhydride was performed by adding 6 equivalents (41.2 mg, 0.42 mmol) to the syringe reactor containing the resin. After shaking for 3 hours, the resin was washed twice (2x DMF → DCM → DMF → MeOH → DMF).

After addition of maleic anhydride, a microcleavage according to method **D** was performed, showing a complete turnover.

Cyclization was achieved by adding HOBt, HBTU and DIPEA to the resin as described in method **C** using following quantities, increasing the amount of equivalents to 6 and the time to 4 hours.

HBTU:	159.27 mg, 0.42 mmol

HOBt:	56.75 mg, 0.42 mmol
DIPEA:	71.43 µL, 0.42 mmol
DMF:	9 mL

After shaking for 3 hours, the resin was washed twice (2x DMF → DCM → DMF → MeOH → DMF). The complete cyclization was controlled by a microcleavage according to method **D**. Final cleavage of the peptide compound, including the side-chain protection groups was performed according to method **E** and purified by HPLC according to method **F**.

Yield: 37.44 mg (0.0403 mmol, 23 % overall yield) as white solid
LC-MS (ESI): t_R 5.70 min, calcd. for $C_{44}H_{64}N_8O_{12}S$ $[M+H]^+$: 929.44 , found 929.32
HPLC: 0 min 0 % → 5 min 10 % → 35 % 70 min → 100 % 80 min

6.5.2.8.3 Addition of the thiol modified oligonucleotide to the peptide compound and purification.

The 5`thiol-modified oligonucleotide (named as A1: GTGGAAAGTGGCAATC) was dissolved in PBS 1x treated with 50 equivalents of TCEP and further purified by gel filtration using NAP™-5 and NAP™-10 columns. The solution was then concentrated to ca. 20 µL using a 3000 Da MWCO ultra centrifugal filter unit. The lyophilized peptide was dissolved in in dd H_2O:ACN (1:10, v/v) to a concentration of 126 mM and mixed with 0.02 equimolar amounts of the thiol-modified oligonucleotide. 100 µL of dd H_2O:ACN (1:1, v/v) were added to a final concentration of 29 mM. The reaction mixture was gently shaken for 48 hrs at 30°C, then concentrated to a final volume of about 20 µL using a 3000 MWCO ultra centrifugal device.
The target product was purified by denaturing PAGE (25% acrylamide in TBE 1x buffer at 220 V for 45 minutes at room temperature).
The purified DNA-peptide conjugate was finally characterized by denaturing gel electrophoresis, using a Typhoon FLA 9000 from GE healthcare Life Sciences and by MALDI spectrometry.

6.5.3 Analytical denaturing Urea Polyacrylamide Gel Electrophoresis (denat. PAGE)

Analysis of ssDNA (0.1 nmol) was performed by denaturing PAGE 25% acrylamide in TBE 1x buffer at 220 V for 30 min at room temperature. If fluorescently labeled oligonucleotides were used, the ssDNA was first characterized by using a Typhoon FLA 9000 from GE healthcare Life Sciences, stained afterwards with SYBR-gold in TBE 1x for 15 minutes and characterized a second time by the Typhoon scanner. In the case of the characterization of an unlabeled ssDNA, it was stained first with SYBR-gold in TBE 1x for 15 minutes and scanned afterwards.

6.5.4 Preparative denaturing Urea Polyacrylamide Gel Electrophoresis (prep. PAGE)

To the reaction mixture, bromophenol blue colored loading dye was added in a one to one ratio. Purification of the oligonucleotides or the DNA-peptide conjugate were performed by denaturing PAGE 25% acrylamide in TBE 1x buffer at 220 V for 45 min at room temperature. The fluorescent bands were cut out under UV-light, transferred into 15 mL falcon tubes and squashed into small pieces. Five mL of TE 1x buffer were added and eluted overnight in a rotary shaker. Finally, the eluate was concentrated by using a 3000 MWCO ultra centrifugal device.
The purified DNA-oligonucleotides or the DNA-peptide conjugate was characterized by denaturing gel electrophoresis, using a Typhoon FLA 9000 from GE healthcare Life Sciences and by MALDI spectrometry.

6.5.4 Assembly protocol

Unless stated differently, the origami structures were assembled using a 1:10 molar ration between the M13mp18 viral DNA (20 nM) and each of the staple strands, in ddH_2O. Afterwards, 10x TEMg buffer was added to a final concentration of (20 mM Tris, 2 mM EDTA, 12.5 mM $MgCl_2$, pH 7.6).

6.5.4.1 Assembly of the DNA origami host (6prism)

Thermal annealing was performed by decreasing the temperature from 90°C to 20°C at -1°C/min on a Thermocycler Mastercycler nexus gradient.

6.5.4.2 Assembly of the DNA origami host (6prism)

Thermal annealing was performed by decreasing the temperature from 90°C to 20°C at -1°C/min followed by a slow cooling from 44°C to 20°C (-1°C/150 min) on a Thermocycler Mastercycler nexus gradient.

6.5.5 Agarose Gel Electrophoresis

Unless stated differently, all agarose gels (0.75 % agarose in TBEMg 1x) were performed at 80 V for 2.15 h at 4°C. The gels were scanned (in case of fluorescence DNA or proteins) with a Typhoon FLA 9000, stained in EtBr and scanned again.

6.5.6 Gel extraction of origami compounds (Agarose gels)

Preparative agarose gels were prepared with 0.5 % agarose in TBEMg 1x and performed at 80 V for 1.5 h at 4°C. The desired samples were excised with a clean scalpel directly under UV light, loaded into a Freeze 'N' Squeeze™ spin column and cooled for 5 minutes at -20°C. The target compounds were recovered after centrifugation at 7500 rcf for 2 min at 4°C and directly used for AFM, TEM or further characterizations.

7. Supplementary Materials

7.1 Peptide Synthesis

The value of bioactive peptides as drugs has been identified a long time ago. *In vivo* the biosynthesis of peptides lasts a few minutes or even seconds. There is no strict definition for the length of a peptide; a peptide that consists of up to ten amino acids is classified as an oligopeptide and from 10 to 100 amino acids it is called polypeptide. Today a polypeptide that consists of about 50 amino acids can be notated as a protein[58]. Solid supported synthesized peptides are used in all areas of biomedical research[59].

7.1.1 Basic principles of Peptide Synthesis

The formation of a dipeptide by coupling two single amino acids is a simple chemical condensation reaction (*Fig. 68*).

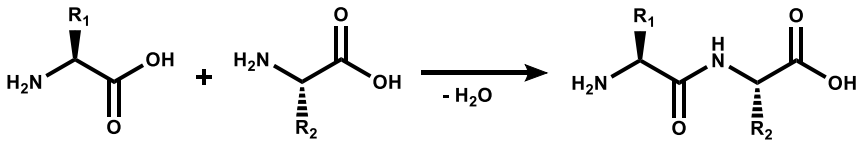


***Figure 68*:** Simplified scheme of the formation of a dipeptide.

Although the reaction is rather simple, some precautions against side reactions are necessary, considering additional functional groups at the side chain. To get complete control of the reaction, the amino acids have to be modified with protecting groups (PG) in a way, that just the desired reaction takes place. Under mild conditions the reaction needs to be activated at the carboxyl component of one amino acid (*Fig. 69*; peptide **A**), carrying a selectively blocked *N*-terminus. The functional groups of the second amino acid are protected in an opposite way, being protected at its carboxylic function and holding an unprotected amino group (*Fig. 69*; peptide **B**). Mixing amino acid A and B results in the formation of the dipeptide A-B, through a reaction between the nucleophilic amino group and the activated carboxylic function.

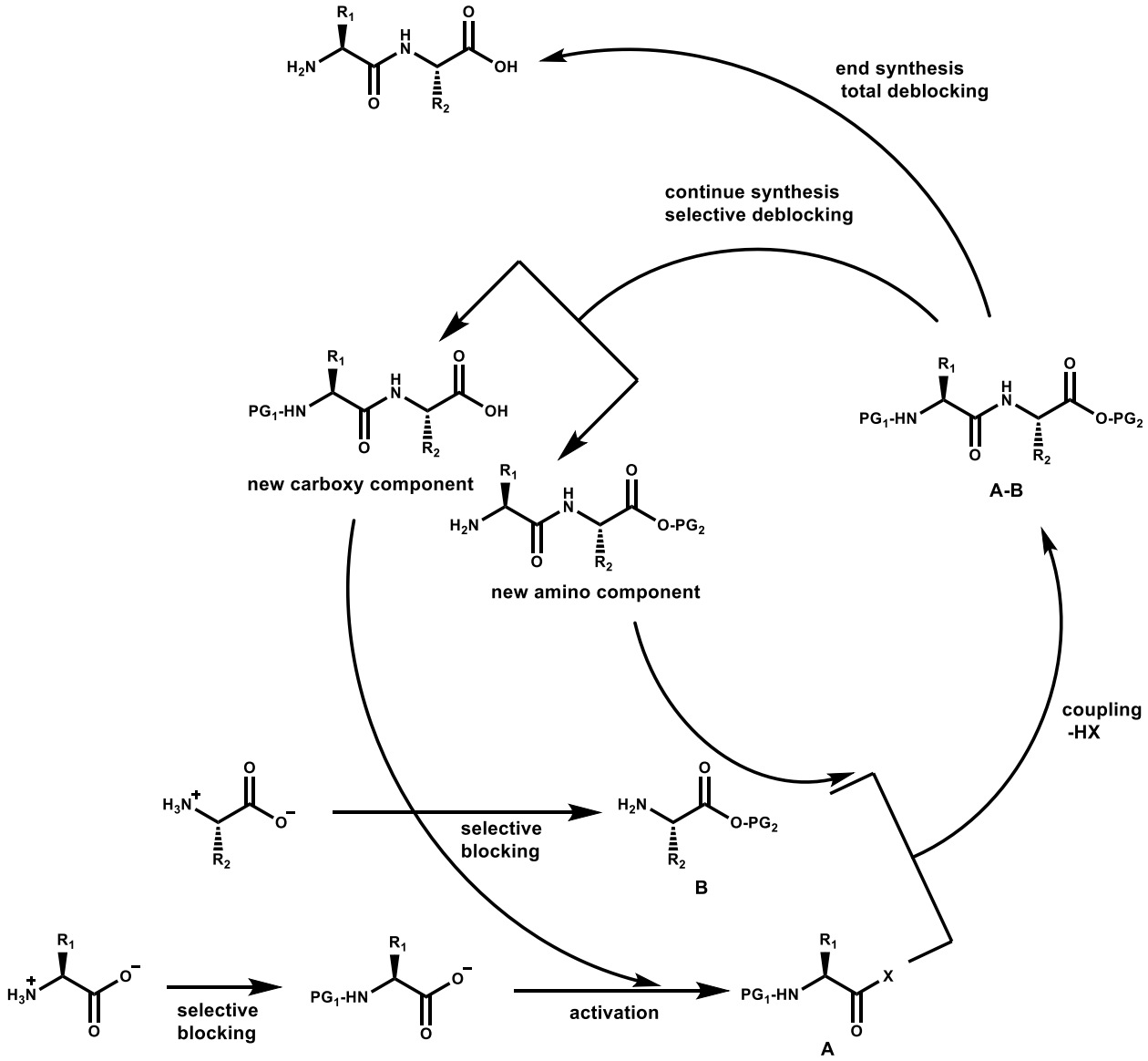


***Figure 69*:** Circular scheme of a multi-step peptide synthesis (adapted and modified from Ref.[58]).

After the successful synthesis of the dipeptide A-B, the PGs of the molecule can be completely removed to end synthesis or partially removed to continue synthesis. Hence, the multistep peptide synthesis can be described as a procedure following these three steps:

1. The amino acid has to be partially protected, resulting in a loss of the zwitterionic structure.
2. Within two steps, the *N*-terminal protected amino acid has to be activated at its carboxylic function forming a reactive intermediate. The formation of the peptide bond takes place immediately.
3. Either all PGs are removed to end synthesis, or selected PGs are cleaved to continue synthesis.

To activate the carboxylic acid a broad spectrum of reagents is available[60], like acyl azides[61], carbodiimides[62], mixed anhydrides[63] and active esters of HOBt or HOAt.

7.1.2 Protection group strategies

A challenging task of peptide chemistry is to avoid side reactions, especially in the case of additional functional groups in the side chains. These groups are blocked by so called protecting groups, which can be divided into two different classes of PGs. The transient protecting groups are used to protect the amino groups forming the peptide bond (*Fig. 69*) and should be easily removable under conditions which do not affect the other PGs or the stability of the peptide sequence. As a second class of PGs, the permanent protecting groups are used to avoid undesired side chain reactions. These permanent PGs have to be stable enough to sustain the cleavage of the transient PGs. Two orthogonal protecting group strategies are commonly used in peptide chemistry, the Boc[64] and the Fmoc[65] strategy. Some advantages and disadvantages are listed in table **7**.

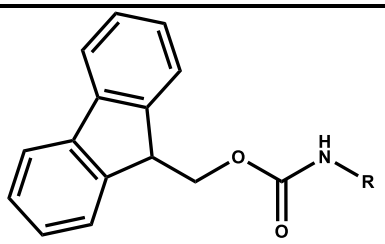
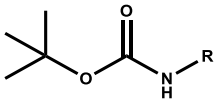
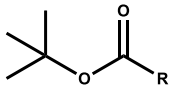
***Table 7*:** Advantages and Disadvantages of Boc and Fmoc strategy (adapted and modified from Ref.[66])

Strategy	Boc strategy	Fmoc strategy
Deprotection	mild acidic conditions with TFA: commonly TFA:DCM in 1:1 ratio	mild basic conditions with primary or secondary amines: commonly 20% (v/v) piperidine in DMF
Advantage	possible Z groups for side chain protection; stable towards catalytic hydrogenation	stable towards tertiary amines; deprotection of the transient PG does not affect amide or t-butyl protected side chain esters
Disadvantage	final cleavage with HF	strong secondary structures if the peptide chain is too long

The older Boc strategy has mostly been replaced through the Fmoc strategy, which is predominantly superior due its milder reaction conditions, a larger number of side chain PGs and the lack of irritant chemicals like hydrogen fluoride[67]. Hereafter, only the orthogonal Fmoc/tBu, Fmoc/Boc or SHEPARD-strategy[68] will be described. As the name denotes, the Fmoc (Fluorenylmethyloxycarbonyl) strategy is based on the base-sensitive *N*-amino transient protecting group, which is usually removed by piperidine or DBU in DMF. The orthogonal side chain PGs which are primarily used are *t*-butyl ethers (Ser, Thr, Tyr), esters (Glu, Asp) and Boc carbamates (Lys). All these acid-labile side chain PGs can be simultaneously

deprotected with the cleavage of an acid-labile linker, which connects the resin via the *N*-terminus of the peptide chain, to produce the free peptide in solution within a single step. In table 8, all protecting groups which were used in this thesis are shown.

***Table 8*:** Structure of the protecting groups which were used in this thesis.

Protection group	Fmoc	Boc	O*t*Bu
Structure			
Used for Amino acids	all amino acids	Lys	Asp

7.1.3 Steps in Solid-Phase Peptide Synthesis

In the solid phase synthesis (SPS) a chemical reaction takes place at the interface of the solid support and the solution. Such type of chemical reaction are inherently heterogenous. Target molecules are coupled via functional groups to the solid support. Commonly, functionalized polystyrene resin beads are used, which are part of the gelatinous solid supports[69]. This single- or multistep synthesis was introduced in the 1960´s by Merrifield and was honored with the Nobel Prize in 1984[70]. Nowadays, the solid phase synthesis (SPS) is a standard technique for the synthesis of oligopeptides and oligonucleotides.

Solid phase peptide synthesis (SPPS) can be performed automatically by synthesizers. Through the cleavage of the Fmoc-protection group the yield of each coupling step can be measured directly. Even on-bead spectroscopic and spectrometric methods are routinely used[67]. By-products and excess of coupling reagents can be easily removed by filtration. Due to all these advantages, the solid phase peptide synthesis revolutionized the peptide chemistry.

Solid phase peptide synthesis starts with the attachment of the first amino acid via its carboxylic function to the resin (*Fig. 70*; Step 1). In the following, the transient protection group (TPG) is removed (*Fig. 70*; Step 2) and the next amino acid, which is preactivated at its carboxylic function and protected at its amino function is coupled (*Fig. 70*; Step 3) to the deprotected amino acid. Steps 2 and 3 are repeated n-times until the desired sequence is

synthesized. In a final step, the peptide is completely deprotected and cleaved from the resin simultaneously (*Fig. 70*; Step 4).

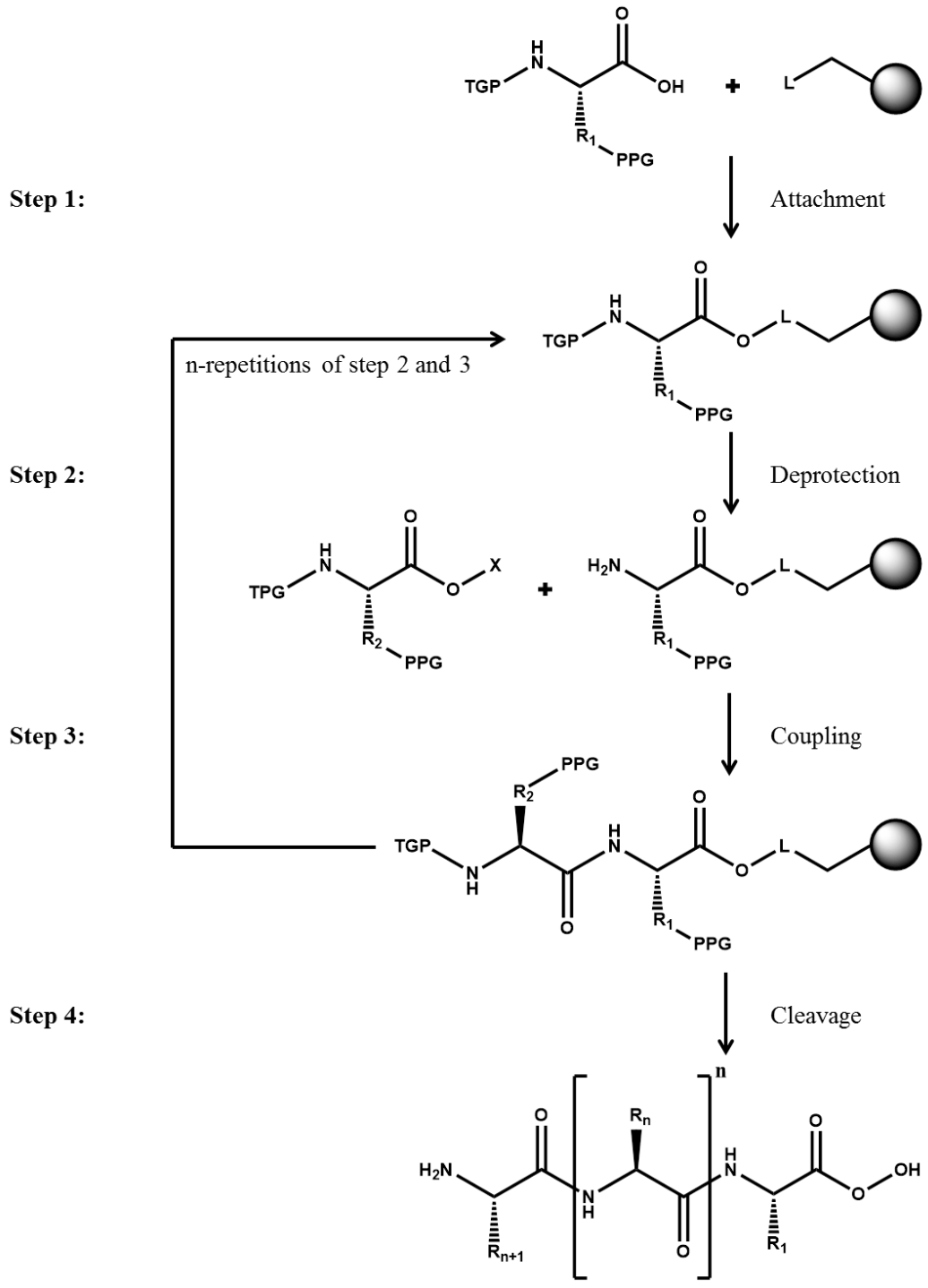


***Figure 70*:** Generalized steps in solid phase peptide synthesis. (TPG = transient protecting group; PPG = permanent protecting group; R = amino acid side chain; X = activating molecule; L = linker)

7.2 Atomic Force Microscopy

Scanning probe microscopy (SPMs) including scanning force (SFM) and atomic force microscopy (AFM) is based on the distance-dependent interaction between the sample and the probe (tip). Without the use of lenses or optical devices, the information is gathered by the existing forces between the sample and the probe. AFMs can operate in liquid, vacuum and air, which open a wide field of applications. Next to the topographical measurements, several mechanical forces like adhesion strength, stiffness and electrical properties like conductivity or surface potential can be measured and even the surface of a sample can be manipulated by the probe being used lithographically.

7.2.1 Basic principles of the AFM

There are several different setups of the atomic force microscope, but they all correspond in their main principles. A common AFM uses a small nanometer-sharp tip to “feel” the surface of the sample by moving the tip in x and y axis by oscillations of the piezo element and in z axis, either through movement of the probe or through movement of the piezo tube (either the sample could be fixed and the tip is moving). Moving above the sample changes the interaction between the sample and the tip due to the different material and distance, what causes a deflection of the laser beam on the photodiode. This deflection is transformed into a topographical image by converting the normal and lateral forces (F_N and F_L) as well as the total intensity of light ($\sum$) by transferring them via the digital signal processor (DSP) to the computer software[71]. Reacting to these signals, the computer orders the change of the XYZ voltage, sending the order to the high voltage electronics (HV) via the DSP element. The HV element finally changes the voltage of the piezo tube, resulting in a change of the x, y and z coordinates.

Using the laser beam deflection method, the deflection of the cantilever can be measured by the shift of the lasers reflection at the photodiode (*Fig. 71*). This semi conductive diode converts the intensity of light into electrical voltage and is divided into four parts to detect horizontal and vertical movements of the cantilever. Vertical movement of the cantilever originates from the normal forces (F_N) and horizontal movement from the lateral forces (F_L).

The force F acting at the end of a lever results in an angle:

$$\theta = \frac{Fl^2}{2EI} \qquad [2]$$

I is the area moment of inertia and *E* the Young´s Modulus. Regarding a rectangular cantilever, $EI = kl^3/3$, where *k* fulfills *F* = *kz* as the spring constant and *l* is the cantilever length. Applying these values results in:

$$\theta = \frac{3\,z}{2\,l} \qquad [3]$$

Assuming that for small angles tan $\theta \sim \theta$ and tan $\theta = D/S$:

$$D = \frac{3\,S}{2\,l} z \qquad [4]$$

For a cantilever with a length of $l \sim 200$ µm at a distance of $S \sim 5$ cm, the movement of the cantilever is amplified by a factor of 375, resulting in a less sensitive method in comparison to other electronic methods. Due to the simplification of this method, it is preferred to detect the cantilever deflection for measurements in air and in liquid.

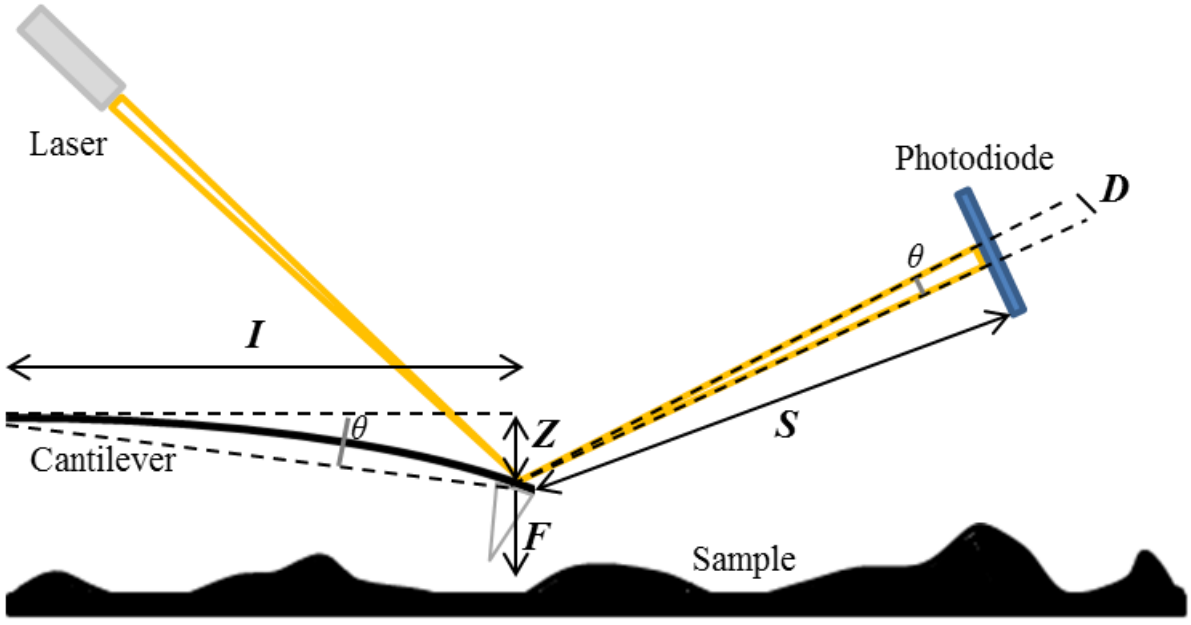


***Figure 71*:** Side view of the beam deflection mode setup with important dimensions (taken and adapted from Ref.[71]).

7.2.2 Tapping mode

Imaging modes of the AFM can be divided into two general classes according to the oscillation of the tip. During contact and jumping or pulsed force modes, the tip is static and it does not oscillate. In contrast to this setup, the tip oscillates in the dynamic mode including, among others, the amplitude-modulation atomic force microscopy (AM-AFM; tapping mode) and the frequency-modulation atomic force microscopy (FM-AFM). In the following section, only the tapping mode is described due to exclusively usage in this thesis.

To minimize the time of contact between the sample and the tip and avoiding damaging either, the cantilever is driven to oscillate at, or close to, its free resonant frequency f_0. The amplitude A_0 and frequency f_0 of the oscillation are kept at a constant level. Approaching to the surface, the tip is been attracted by Van der Waals forces, dipole-dipole interaction, electrostatic forces, etc. causing a change (usually decrease) of the amplitude. To keep the amplitude constant, the height of the cantilever above the sample is adapted by the piezo element. In tapping mode the force of the intermittent contacts between the tip and the surface of the sample produces the topographical image[72].

7.2.3 Artifacts in AFM imaging

Artifacts in AFM imaging are related to either the geometry of the tip or to the feedback cycle. The tip dilation or convolution artifact is based on the finite dimensions of the AFM tip and causes a broadening of the samples (*Fig. 72a*). Two models exist to explain the broadening, one defining the tip as conical with an apex angle of 2α and a square object with dimensions of h x D (*Fig. 72b*). Geometry allows deriving following equation:

$$FWHM = \sqrt{D + h\tan(\alpha)} \qquad [5]$$

The alternative model considers the tip as spherical with the radius R_{tip} and a cylindrical radius R_{sample}. By knowing the dimensions of a certain object, the radius R_{tip} can be easily calculated based on geometry yields.

$$FWHM = 2\sqrt{2\,R_{tip}R_{sample} + R_{sample}^2} \qquad [6]$$

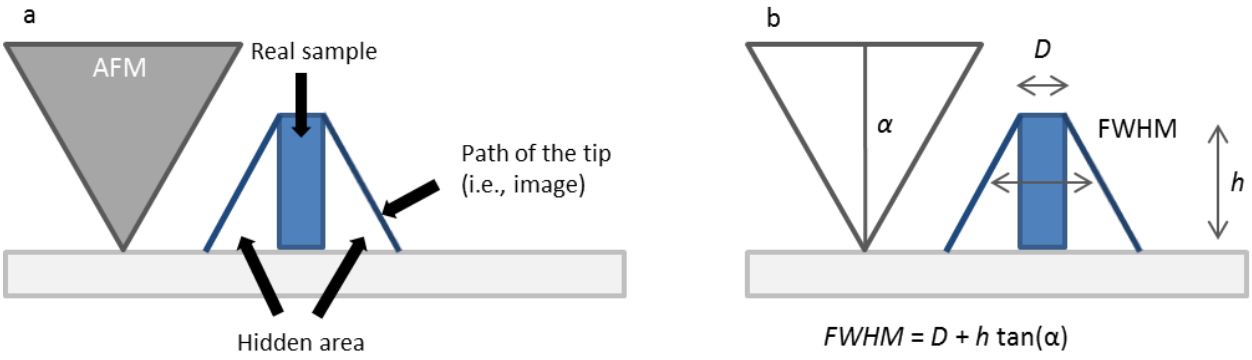


***Figure 72*:** Schemes of the tip dilation or convolution phenomenon (**a**) and (**b**) a model used to estimate tip dimensions from measurements with a full width and half maximum being defined by the angle of the conical tip.

Concerning the artifacts related to the feedback loop, it has to be said, that the number of possible artifacts is too large to be described here. If the response system reacts to slow, the surface protrusions are elongated along the scan direction. A conversely too high feedback results in an unstable system and a high-frequency oscillation can be observed in the image; resulting in a wrongly attribution to a real surface.

7.3 Transmission Electron Microscopy (TEM)

To increase the resolution of optical microscopy, which is limited due to the wavelength of light, the wavelike character of electrons is used. Accelerating electrons in vacuum with a potential of e.g. 50 kV results in an electron wavelength of λ = 5 pm. Focusing these electrons with lenses, comparable to optical microscopes, images of specimen with a higher spatial resolution can be produced.

7.3.1 Basic principles of TEM

Transmission electron microscopes (TEM) are based on electrons penetrating a thin specimen and function exactly like their optical counterparts concerning yielding information as to its structure and composition (*Fig. 73*). As an illumination source a filament is used as the source of electrons, usually a hairpin shaped-tungsten wire (or LaB6). To get monochromatic emitted electrons from a small part of the filament (with similar energy), a negative potential is applied to its surrounding cap (Wehnelt Cap) creating a point source (*Fig. 73*; zoom in). Due to the positive potential of the anode plate, the electrons which are close to the opening of the anode plate are accelerated into the column for imaging. All lenses are used to focus the

electron beam and are not described in detail. It just has to be mentioned that focusing of the electron beam through the lenses occurs by applied magnetic fields.

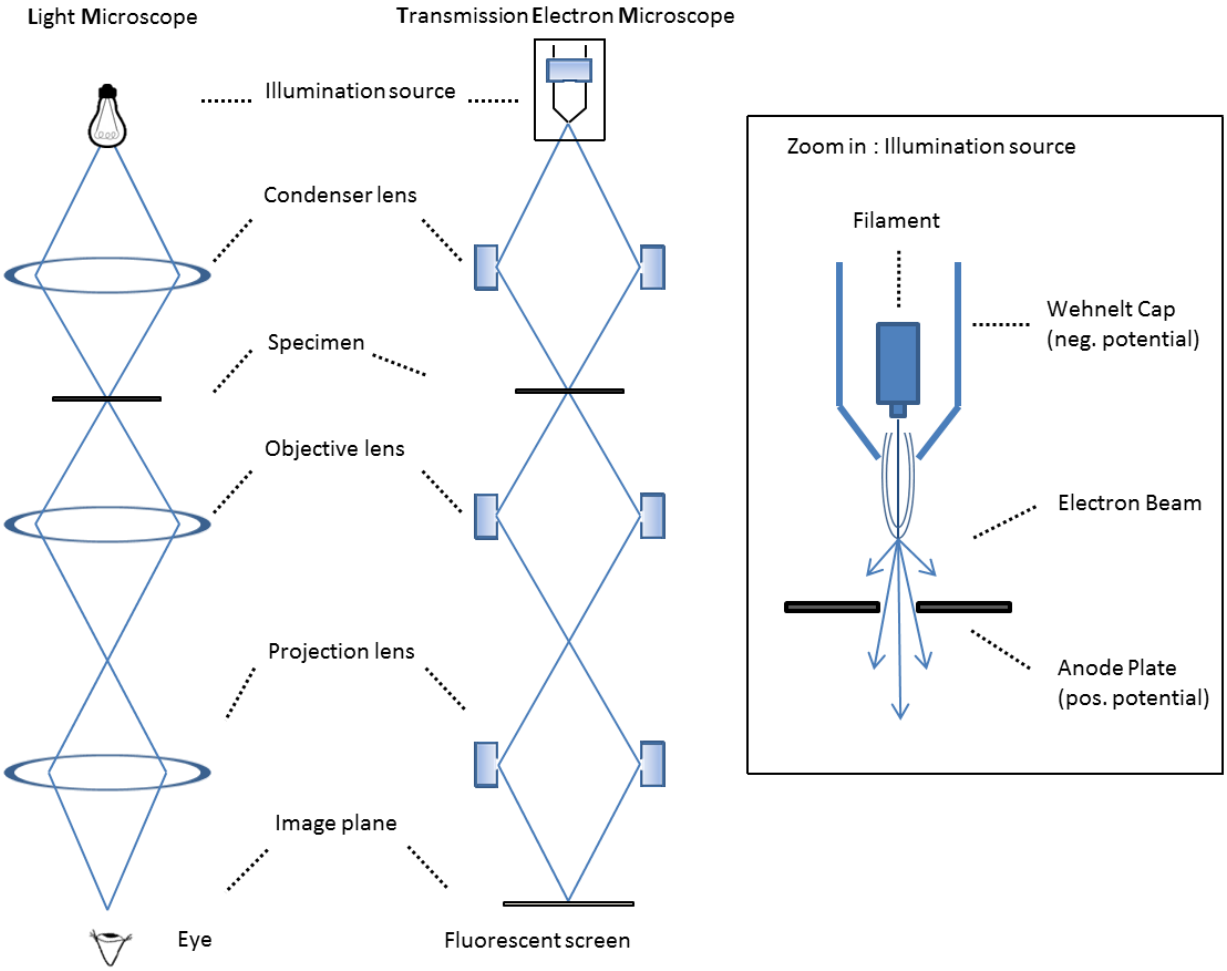


***Figure 73*:** Similar basic setup of the light microscope and the transmission electron microscope. Analogy is shown to simplify functionality.

Regarding the electron-specimen interaction, there are three different types of interactions in transmission electron microscopy; elastically scattered electrons (diffracted beam), inelastically scattered electrons and unscattered electrons (transmitted beam).

The beam of incident electrons passing the specimen without any interaction is called transmitted beam. Regarding on the thickness of the specimen the number of transmitted electrons is inversely proportional. Incident electrons, which are deflected from their original path without losing energy, are scattered according to Bragg´s Law.:

$$n * \lambda = 2d * \sin(\theta) \qquad [7]$$

Where λ is the electrons wavelength, θ is the complementary angle of the angle of entry and d is the distance between two aligned layers of atoms.
In the case of identical atomic spaces, all electrons are scattered in the same angle. Electrons which are deflected are collected using magnetic lenses to form a pattern of spots. Such a pattern may yield information concerning atomic arrangements, the orientations and phases being present in the examined area. Inelastically scattered electrons interact with the specimen while passing it resulting in a loss of energy. These electrons can be used in another type of experiments, e.g. the Electron Energy Loss Spectroscopy (EELS) to get information concerning the elemental composition and atomic bonding.
The electron beam passing through the specimen gets through a magnification system and forms an image on a fluorescent screen or is photographed by an extra camera.

7.3.2 Sample preparation and staining

Referring to the chapter above, the samples measured in transmission electron microscopy have to fulfill several requirements. On the one side the sample which has to be examined needs to scatter enough electrons from the beam to give a good contrast and on the other side, the opaqueness for electrons should not be too high.
If the contrast of the sample is too low, like for biological samples (like DNA origami) negative staining is necessary. Several different chemicals for staining are available, like ammonium molybdate, uranyl acetate, uranyl formate, phosphotungstic acid and osmium tetroxide. These molecules fill the space between the sample-molecules and enhance the contrast of the image by scattering electrons. To get good images, a special staining protocol for each type of sample is needed.

7.4 Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry – time of flight (TOF)

In the late 1980s the matrix-assisted laser desorption/ionization mass spectrometry was introduced and coined by *M. Karas* and *F. Hillenkamp*[73, 74]. Nowadays, this technology is used in a brought field, e.g. to assess compound identity of solid phase synthesized oligonucleotides requiring 1-3 pmol of the sample[75].

7.4.1 Basic principles of MALDI-TOF

This technique is based on a laser light energy absorbing matrixes, usually ultraviolet (UV)-absorbing weak organic acids being vaporizing by laser radiation carrying the intact analyte indirectly into a "cloud" of analyte and matrix. Therefore, the matrix is added in high molar access, causing co-crystallization of the analyte and serving also as a proton donor or acceptor for ionization. A high yield of the analyte is provided by the efficient and directed energy transfer, allowing measurements of compounds with high accuracy and subpicomole sensitivity[76, 77]. Desorption process of large molecules during MALDI-TOF mass spectrometry is not completely understood but several theories have been developed, the thermal-spike model[78] and the pressure pulse theory[79]. Generally, it is supposed that ionization occurs through proton transfer-depending on the matrix analyte combination. This interaction between the analyte and the matrix seems to be more than simple acid-base chemistry, due to the lack of correlation concerning ionization between the number of acidic or basic groups of the analyte and the matrix.

After ionization the matrix-analyte mixture is accelerated in the instrument by an electrostatic field to a common kinetic energy. Ions obey the rule,

$$E = \frac{1}{2}mv^2 \qquad [8]$$

where *m* is the mass of the ion, *E* is the kinetic energy and *v* is the velocity of the moving ion. Define *E* as equal to *zA* (under established condition), where *z* is the charge of the ion and *A* is the constant accelerating potential applied to the ions, formula [8] can be rearranged to:

$$v = \sqrt{2\frac{z}{m}} \qquad [9]$$

Under the aspect that velocity is a function of time and distance and that in the MALDI-TOF instrument velocity of ions are determined by their mass and charge, the time for a fixed distance can be described by:

$$t = \sqrt{\frac{m}{2z}} \qquad [10]$$

Lighter ions will arrive at the detector faster than heavier ions, if they are at the same charge state. The entire mass spectrum is recorded in a fraction of seconds as ion flux versus time. In the reflection mode the accuracy of the measurements can be improved by reflecting the ions to a second detector.

8. Additional Results

This chapter includes a summary of additional results and information concerning a second DNA-origami cage confirming the results of the first 6prism design.

8.1 Design details

A second DNA-origami cage was designed in the honeycomb lattice strategy (see chapter 1.2.3.) with interhelical connections relatively oriented at 120° and periodically repeated every 2 helical turns (i.e. every 21 bp). Almost the entire scaffold was used (9 bases left) with a six-helix bundle at each edge of the prism, which ensures a fixed 60° angle between two adjacent faces. Figure 74 shows the geometric model of the new DNA cage (3prsim) being highly symmetric, with a C3 symmetry which is compatible to the regular symmetry of the DegP protein. A nanocage with a 3.5-fold smaller volume, $3x10^3$ nm^3 for the 6prism and $9x10^3$ nm^3 for the 3prism was designed to investigate the role of the inner cavity size to test whether the proposed encapsulation strategy is of general applicability.

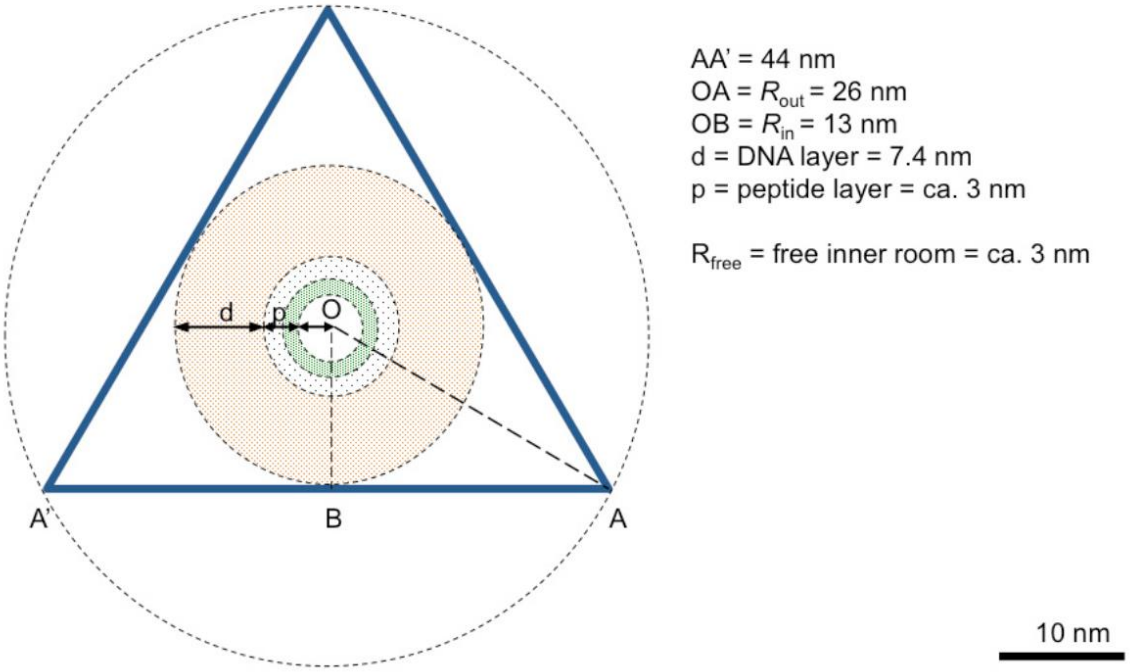


***Figure 74*:** Geometric model of the 3prism DNA host. The inner surface of the chamber is decorated with DNA strands for further hybridization with complementary DNA-peptide conjugates. Only the last 3 amino acids residues interact with the PDZ1 domains of DegP (green area about 1 nm long). The full peptide is about 3 nm long (including a C6 spacer) and is connected to a double-stranded DNA corona of about 7.4 nm. The total length of the host-to-guest bridge is about 10-11 nm, leaving a free inner room of ca 3 nm in radius.

Further design details are shown in Figure 75. With a width of 44 nm, a height of 38 nm and a length of 31 nm the 3prism has a larger outer radius (R_{out} = 25 nm) but a smaller inner radius (R_{in} = 13 nm) in comparison to the 6prism. Such difference in the inner room of the channels is expected to affect the loading capability of the hosts towards the same molecular guests.

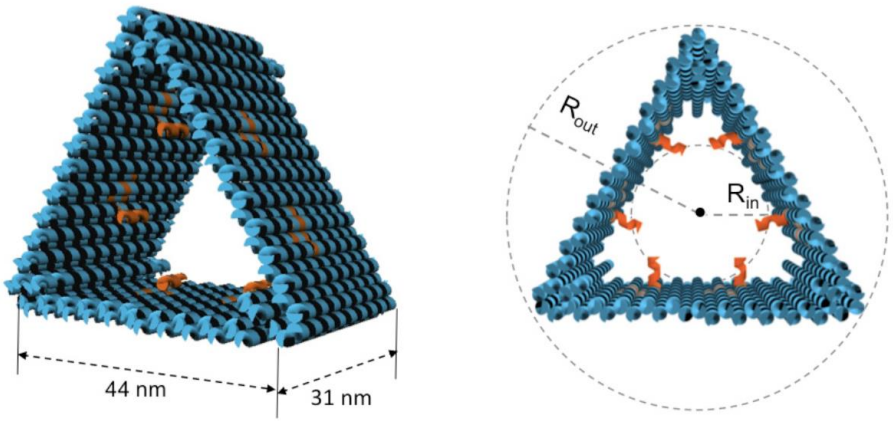


***Figure 75*:** Dimensions and values of the 3prism construct used in this work (44 nm in width x 31 nm in length and 38 nm in height). The prism with a triangular section was prepared from the same M13 scaffold as the 6prism, thus leading to a chamber with a hydrodynamic radius of 25 nm (R_{out} = 25 nm) and an inner radius of 13 nm (R_{in} = 13 nm).

Again the number of attaching points for the DNA-peptide linkers was increased to improve the loading yield. Both available 3prism designs are shown in Figure 76.

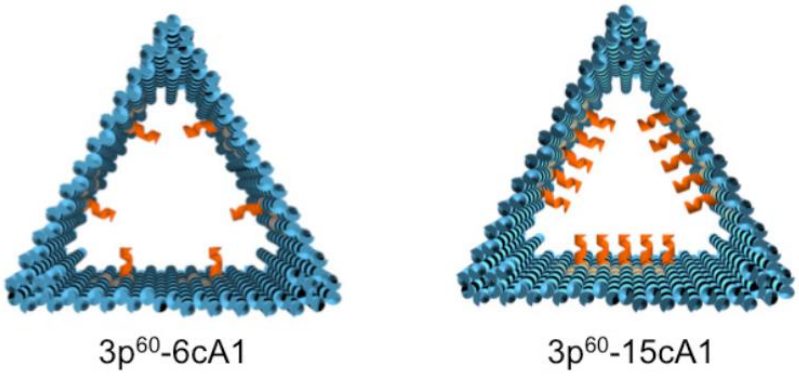


***Figure 76*:** Schematic representation of the available 3prism constructs. The triangular prism was prepared in two possible forms, bearing either 6 or 15 inner protruding arms. Due to the different design strategy used to construct this origami structure, the face-to-face angles are always fixed at 60°.

8.2 Experimental results

For the 3prism chambers, a screening of different magnesium buffer concentrations and annealing temperatures was performed in order to find the optimal conditions for folding (*Fig. 77*). Best yields were obtained at 12 mM magnesium concentration and a thermal annealing from 90°C to 45°C (-1°C/min) followed by a slow cooling from 44°C to 20°C (-1°C/150 min).

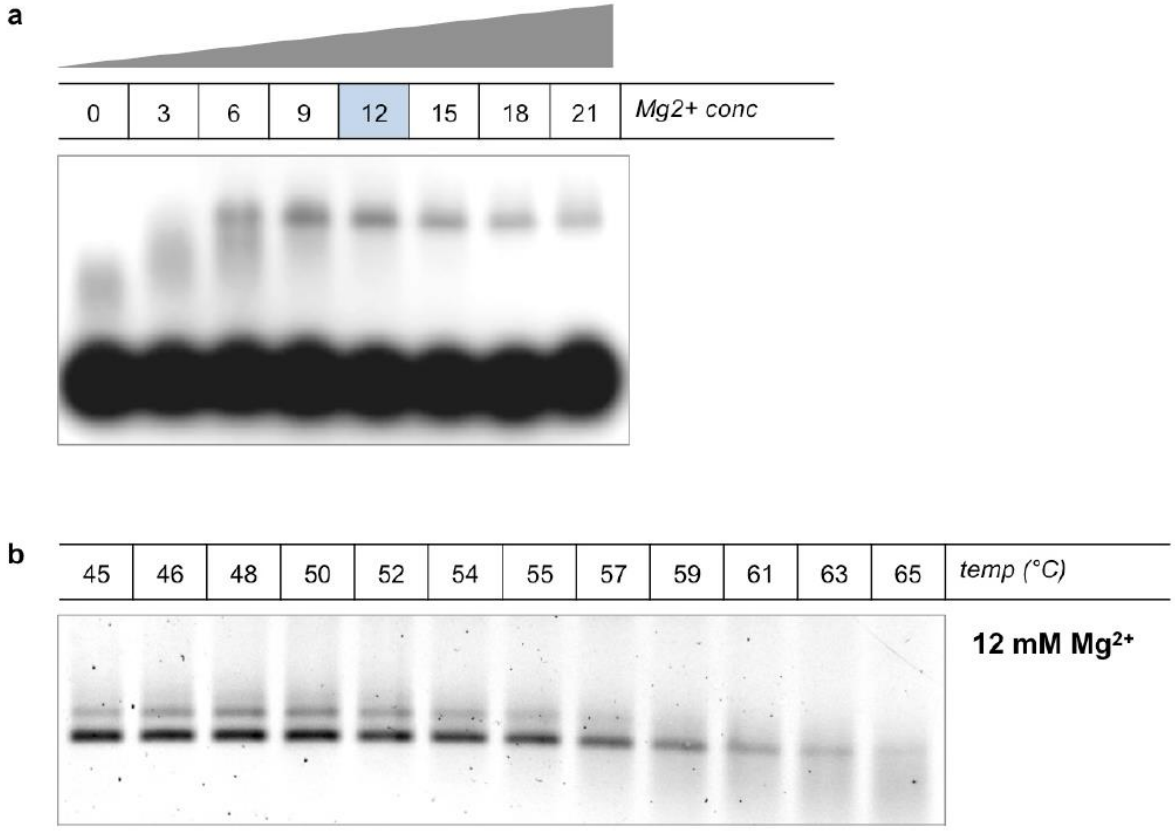


***Figure 77*:** Agarose gel characterization of magnesium and temperature screening for the 3prism. (**a**) Optimal conditions for folding were found at 12 mM magnesium concentration. (**b**) At this concentration, different isothermal annealing assemblies were performed in the range between 45 and 65°C. Higher yields were obtained for a long annealing at moderate temperature. These conditions were then used for assembly of the 3prism structure.

AFM pictures of the successful assembly of the 3prism are shown in chapter 8.4.

8.3 Loading experiments

To better investigate the role played by the diffusion of the protein inside the DNA chamber, a series of loading experiments using a DNA host of different geometry and inner size was performed (*Fig. 74-76*). At this purpose, the M13 scaffold was folded into a regular triangular prism, which theoretically leaves an inner room of only 3 nm. This would in principle exclude the binding in every oligomerization state. Successful binding, even of larger forms of the DegP protein, was observed (*Fig. 78*).

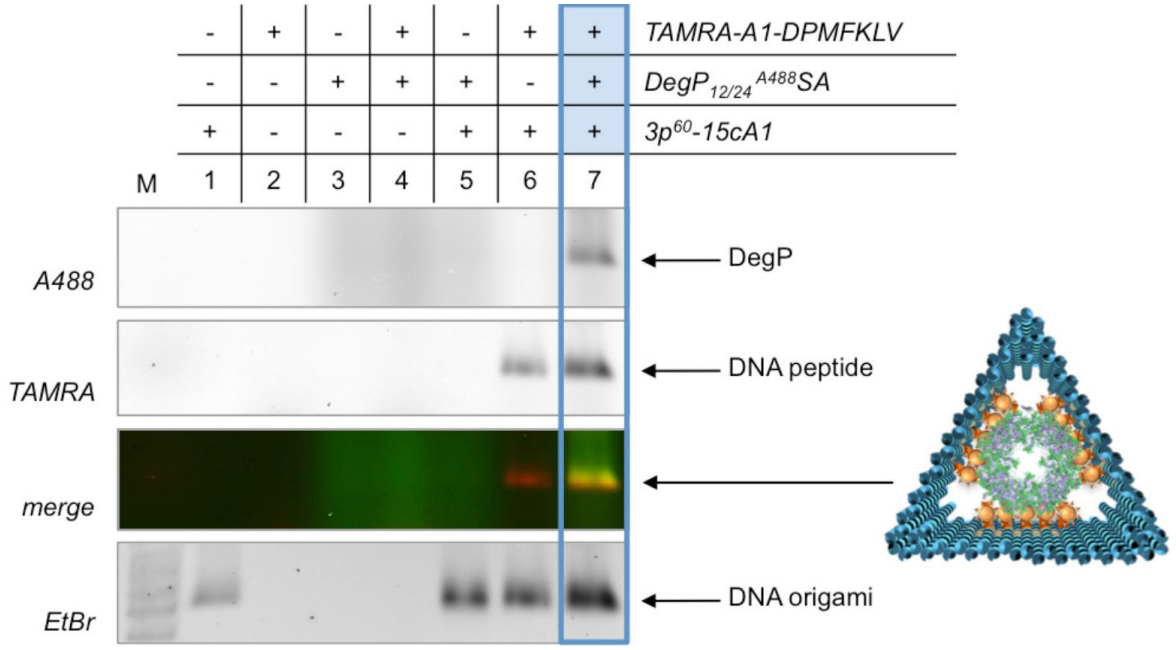


***Figure 78*:** Agarose gel characterization of binding studies of $DegP_{12/24}^{A488}SA$ to the 3prism construct functionalized with 18 TAMRA-labeled A1-peptide ligands in its inner cavity. Successful binding of the protein appears only in presence of the peptide-ligand (lane 7).

These results suggest that the DNA cages are not structurally rigid but may be partially deformed, thus allowing accommodation of large proteins inside the inner cavity. This was shown by AFM characterization of the empty cages (*Fig. 79* for the 3prism and chapter 3.2.2 for the 6prism). In addition, as confirmed by DLS, such a structural flexibility does not hamper full integrity and folding of the structure in solution (*Fig. 79c*).

8.4 AFM characterization of the loaded and unloaded 3p-construct

Figure 79 shows the AFM characterization of the 3p DNA host both in presence and absence of the DegP protein. As also shown for the 6prism construct, AFM imaging of the 3prism leads to its partial deformation, although the structure still keeps integer in solution.

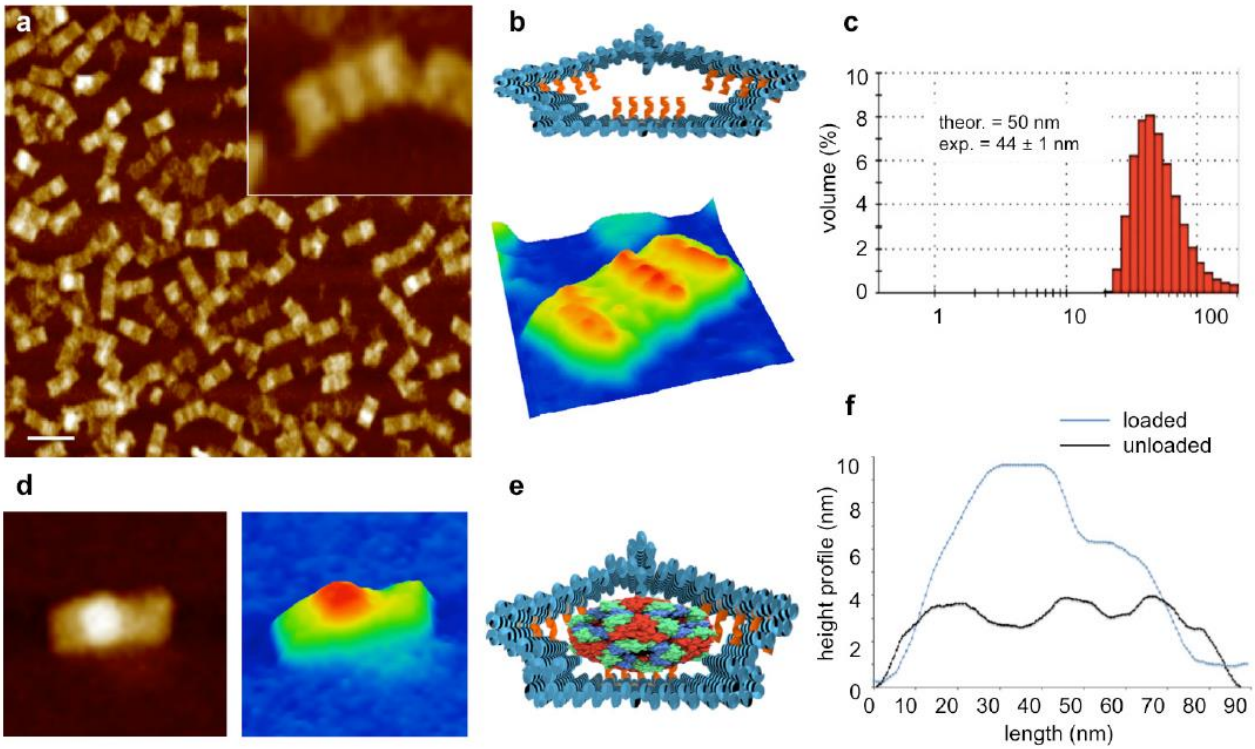


***Figure 79*:** Compilation of AFM and DLS characterization of the $3p^{60}$-15cA1 construct either unloaded or loaded with a $DegP_{12/24}{}^{A488}SA$. **(a)** AFM characterization of the empty DNA chamber results in compression of the structure, giving rise to a rectangular shape of variable width, same length as the 3prism and a height profile of ca. 3 nm displaying three features about 4 nm high (**b** and black curve in **f**). Nevertheless, the structure adopts the expected hollow shape in solution (**c**). When loaded with DegP, the structure displays a brighter spot in its center (**d**) associated to a higher height profile (blue curve in **f**), which we attribute to successful incorporation of the protein inside the chamber (**e**). This latter might be facilitated by the capability of the chamber to deform.

In particular, the DNA prism adopts the shape that would result by compression from the top, resulting into a rectangular shape of variable width with three high-profile features (*Fig. 79a and 79b*; upper panel). These latter correspond to the six-helix bundles originally positioned at the vertexes of the triangular prism; under AFM imaging, such bundles appear as brighter regions spanning the whole length of the rectangular shape (*Fig. 79b*; lower panel). Despite structure deformation in air conditions and under the effect of applied mechanical forces, the DNA host displays the expected hydrodynamic radius in solution (*Fig. 79c*). Loading of the 3p host with $DegP_{12/24}{}^{A488}SA$ leads to a notable increase of the height profile of the structure (ca. 6-7 nm) and appearance of a single brighter spot in its center (*Fig. 79d-f*). Despite the reduced inner size of the DNA host, binding the protein do occurs. This might be due to the

capability of the structure to deform and therefore to adapt to the size/shape of the protein guest (*Fig. 79c*).

9. References

1. Feynman, R.P., *There is plenty of room at the bottom.* Engineering and Science, 1960. **23**: p. 22-36.
2. Junno T., D.K., Montelius L., Samuelson L., *Controlled manipulation of nanoparticles with an atomic force microscope.* Appl. Phys. Lett., 1995. **66**: p. 3627–3629.
3. Eigler D. M., S., E.K., *Positioning single atoms with a scanning tunneling microscope.* Nature, 1990. **344**: p. 524-526.
4. Heinrich, A.J., Lutz, C.P., Gupta, J.A., Eigler, D.M., *Molecular cascades.* Science, 2002. **298**: p. 1381-1387.
5. Gribbin, J.a.G., M., *Richard Feynman: A Life in Science*. 1997, U.S.A., New York: Dutton 320.
6. Seeman, N.C., *Nucleic acid junctions and lattices.* J Theor Biol, 1982. **99**(2): p. 237-47.
7. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.
8. Seeman, N.C., *Nanomaterials based on DNA.* Annu Rev Biochem, 2010. **79**: p. 65-87.
9. Saccà, B., Sprengel, A. and Feldkamp, U., *De Novo Design of Nucleic Acid Structures, in De novo Molecular Design (ed G. Schneider)*. 2013, Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.
10. Fu, T.J. and N.C. Seeman, *DNA double-crossover molecules.* Biochemistry, 1993. **32**(13): p. 3211-20.
11. Seeman, N.C., *De novo design of sequences for nucleic acid structural engineering.* J Biomol Struct Dyn, 1990. **8**(3): p. 573-81.
12. Brun, Y., Gopalkrishnan, M., Reishus, D., Shaw, B., Chelyapov, N., Adleman, L., *Building Blocks fpr DNA Self-Assembly*, in *Foundations of Nanoscience: self-assembled architectures and devices (FNANO04)*, J.H. Reif, Editor. 2004: Snowbird Cliff Lodge, Snowbird, Utah. p. 2-15.
13. Yan, H., et al., *DNA-templated self-assembly of protein arrays and highly conductive nanowires.* Science, 2003. **301**(5641): p. 1882-4.
14. Rothemund, P.W., *Folding DNA to create nanoscale shapes and patterns.* Nature, 2006. **440**(7082): p. 297-302.
15. Douglas, S.M., et al., *Rapid prototyping of 3D DNA-origami shapes with caDNAno.* Nucleic Acids Res, 2009. **37**(15): p. 5001-6.
16. Aksimentiev, A., Brunner, R., Cohen, J., Comer, J., Cruz-Chu, E., Hardy, D., Rajan, A., Shih, A., Sigaliv, G., Yin, Y. and Schulten, K., *Computer modeling in biotechnology, a partner in development.* Protocols in Nanostructure Design, 2008: p. 181-234.
17. Iinuma, R., et al., *Polyhedra self-assembled from DNA tripods and characterized with 3D DNA-PAINT.* Science, 2014. **344**(6179): p. 65-9.
18. Sun, S., et al., *DNA polygonal cavities with tunable shapes and sizes.* Chem Commun (Camb), 2015. **51**(90): p. 16247-50.
19. Douglas, S.M., I. Bachelet, and G.M. Church, *A logic-gated nanorobot for targeted transport of molecular payloads.* Science, 2012. **335**(6070): p. 831-4.
20. Andersen, E.S., et al., *Self-assembly of a nanoscale DNA box with a controllable lid.* Nature, 2009. **459**(7243): p. 73-6.
21. Kuzuya, A. and M. Komiyama, *Design and construction of a box-shaped 3D-DNA origami.* Chem Commun (Camb), 2009(28): p. 4182-4.

22. Endo, M., et al., *DNA prism structures constructed by folding of multiple rectangular arms.* J Am Chem Soc, 2009. **131**(43): p. 15570-1.
23. Han, D., et al., *DNA origami with complex curvatures in three-dimensional space.* Science, 2011. **332**(6027): p. 342-6.
24. Douglas, S.M., et al., *Self-assembly of DNA into nanoscale three-dimensional shapes.* Nature, 2009. **459**(7245): p. 414-8.
25. Dietz, H., S.M. Douglas, and W.M. Shih, *Folding DNA into twisted and curved nanoscale shapes.* Science, 2009. **325**(5941): p. 725-30.
26. Goodman, R.P., et al., *Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication.* Science, 2005. **310**(5754): p. 1661-5.
27. von Kiedrowski, G., et al., *Toward replicatable, multifunctional, nanoscaffolded machines. A chemical manifesto.* Pure and Applied Chemistry 2003. **75**(5): p. 609-619.
28. Mei, Q., Wei, X., Su, F., Liu, Y., Youngbull, C., Johnson, R., Lindsay, S., Yan, H., Meldrum, D., *Stability of DNA Origami Nanoarrays in Cell Lysate.* Nano Lett., 2011. **11**: p. 1477-1482.
29. Zhang, F., et al., *Complex wireframe DNA origami nanostructures with multi-arm junction vertices.* Nat Nanotechnol, 2015. **10**(9): p. 779-84.
30. Benson, E., et al., *DNA rendering of polyhedral meshes at the nanoscale.* Nature, 2015. **523**(7561): p. 441-4.
31. Wei, B., M. Dai, and P. Yin, *Complex shapes self-assembled from single-stranded DNA tiles.* Nature, 2012. **485**(7400): p. 623-6.
32. Ke, Y., et al., *Three-dimensional structures self-assembled from DNA bricks.* Science, 2012. **338**(6111): p. 1177-83.
33. Voigt, N.V., et al., *Single-molecule chemical reactions on DNA origami.* Nat Nanotechnol, 2010. **5**(3): p. 200-3.
34. Sacca, B., et al., *Orthogonal Protein Decoration of DNA Origami.* Angew Chem Int Ed Engl, 2010.
35. Ding, B., et al., *Gold nanoparticle self-similar chain structure organized by DNA origami.* J Am Chem Soc, 2010. **132**(10): p. 3248-9.
36. Pal, S., et al., *DNA-origami-directed self-assembly of discrete silver-nanoparticle architectures.* Angew Chem Int Ed Engl, 2010. **49**(15): p. 2700-4.
37. Fu, J., et al., *Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures.* J Am Chem Soc, 2012. **134**(12): p. 5516-9.
38. Liu, M., et al., *A DNA tweezer-actuated enzyme nanoreactor.* Nat Commun, 2013. **4**: p. 2127.
39. Fu, J., et al., *Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm.* Nat Nanotechnol, 2014. **9**(7): p. 531-6.
40. Crawford, R., et al., *Non-covalent Single Transcription Factor Encapsulation Inside a DNA Cage.* Angew Chem Int Ed Engl, 2013. **52**(8): p. 2284-8.
41. Maune, H.T., et al., *Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates.* Nat Nanotechnol, 2010. **5**(1): p. 61-6.
42. Zadegan, R.M., et al., *Construction of a 4 zeptoliters switchable 3D DNA box origami.* ACS Nano, 2012. **6**(11): p. 10050-3.
43. Zhao, T., Fu, J., Dhakal, S., Johnson-Buck, A., Liu, M., Zhang, T., Woodbury, N. W., Liu, Y., Walter, N. G., Yan, H., *Nanocaged enzymes with enhanced catalytic activity and increased stability against protease disgestion.* Nat Commun, 2016.

44. Hedstrom, L., *Serine Protease Mechanism and Specificity.* Chem Rev, 2002. **102**: p. 4501-4523.
45. Dodson, G., Wlodawer, A., *Catalytic triads and their relatives.* Trends Biochem Sci, 1998. **23**: p. 347-352.
46. Mueller-Esterl, W., *Biochemie*. 2 ed. 2011, Heidelberg: Spektrum Akademischer Verlag.
47. Sadigh-Eteghad, S., Majdi, A., Talebi, M.m Mahmoudi, J., Babri, S., *Regulation of nicotinic acetylcholine receptors in Alzheimer´s disease: A possible role of chaperones.* European Journal of Pharmacology, 2015. **755**: p. 34-41.
48. Merdanovic, M., et al., *Determinants of structural and functional plasticity of a widely conserved protease chaperone complex.* Nat Struct Mol Biol, 2010. **17**(7): p. 837-43.
49. Kolmar, H., Waller, P., Sauer, R. , *The DegP and DegQ periplasmic endoproteases of Escherichia coli: specificity for cleavage sites and substrate conformation.* J. Bacteriol., 1996. **178**: p. 5925-5929.
50. Spiess, C., Beil, A., Ehrmann, M., *A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein.* Cell, 1999. **97**: p. 339-347.
51. Krojer, T., Sawa, J., Schaefer, E., Saibil, H. R., Ehrmann, M., Clausen, T., *Structural basis for the regulated protease and chaperone function of DegP.* Nature, 2008. **453**: p. 885-890.
52. Gray, M.J., Wholey, W.-Y., Wagner, N. O., Cremers, C. M., Mueller-Schickert, A., Hock, N. T., Krieger, A. G., Smith, E. M., Bedner, R. A., Bardwell, J. C. A., Jakob, U., *Polyphosphate is a Primordial Chaperone.* Mol Cell, 2014. **53**: p. 689-699.
53. Sprengel, A., Lill, P., Stegemann, P., Bravo-Rodriguez, K., Schöneweiß, E.-C., Merdanovic, M., Gudnason, D., Aznauryan, M., Gamrad, L., Barcikowski, S., Sanchez-Gracia, E., Birkedal, V., Gatsogiannis, C., Ehrmann, M., Saccà, B., *Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions.* Nat Commun, 2017.
54. Stegemann, P., *Etablierung von Labelingsystemen für den Nachweis der Interaktion von DegP mit DNA-Nanopartikeln*, in *ZMB*. 2016, Duisburg-Essen.
55. Stahl, E., Martin, T. G., Praetorius, F. and Dietz, H., *Facile and Scalable Preparation of Pure and Dense DNA Origami Solutions.* Angew Chem Int Ed Engl, 2014. **126**: p. 12949–12954.
56. Meltzer, M., et al., *Allosteric activation of HtrA protease DegP by stress signals during bacterial protein quality control.* Angew Chem Int Ed Engl, 2008. **47**(7): p. 1332-4.
57. Lill., P., *Electron microscopy of DNA origami structures*, in *Biology*. 2016, MPI Dortmund.
58. Sewald, N., Jakubke, H.-D., *Peptides: Chemistry and Biology*. second edition ed. 2009, Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA.
59. Stevenson, C.L., *Advances in Peptide Pharmaceuticals.* Curr. Pharm. Biotechnol., 2009. **10**: p. 122-137.
60. Montalbetti, C.A.G.N., Falque, V., *Amide bond formation and peptide coupling.* Tetrahedron, 2005. **61**: p. 10827-10852.
61. Klausner, Y.S., Bodanszk, M., *Azide Method in Peptide-Synthesis - Its Scope and Limitations.* Synthesis (Stuttg.), 1974: p. 549-559.
62. Izdebski, J., Kunce, D., *Evaluation of carbodiimides using a competition method.* J. Pept. Sci., 1997. **3**: p. 141-144.

63. Wittenberger, S.J., McLaughlin, M. A., *Preparation of endothelin antagonist ABT-627.* Tetrahedron Lett., 1999. **40**: p. 7175-7178.
64. Alewood, P., Alewood, D., Miranda, L., Love, S., Meutermans, W. and Wilson, D., *Rapid in situ neutralization protocols for Boc and Fmoc solid-phase chemistries.* Methods Enzymol, 1997. **289**: p. 14-29.
65. Wellings, D.A.a.A., E., *Standard Fmoc protocols.* Methods Enzymol, 1997. **289**: p. 44-67.
66. Corporation, T.E., *Peptide synthesis - chemistry and modifications.* p. 1-19.
67. Seneci, P., *Solid-Phase Synthesis: Basic Principles, in Solid-Phase Synthesis and Combinatorial Technologies*. 2000, New York, USA: John Wiley & Sons, Inc.
68. Chan, W.C.a.W., P.D., *Fmoc solid phase peptide synthesis: a practical approach*, ed. E.D. Hames. 1999, U.S.A. New York: Oxford University Press.
69. Hodge, P., *Polymer-supported organic reactions: what takes place in the beads?* Chem. Soc. Rev., 1997. **26**: p. 417-424.
70. Merrifield, R.B., *Solid Phase Peptide Synthesis. 1. Synthesis of a Tetrapeptide.* J. Am. Chem. Soc. , 1963. **85**: p. 2149-2154.
71. Baro, A.M.a.R.R.G., *Atomic Force Microscopy in Liquid: Biological Applications*. 2012: Wiley-VCH Verlag GmbH & Co.
72. Geisse, N.A., *AFM and Combined Optical Techniques.* Materials Today, 2009. **12**(7-8): p. 40-45.
73. Karas, M.a.H.F., *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Anal Chem, 1988. **60**(20): p. 2299-2301.
74. Karas, M., Bahr, U. and Gießmann, U., *Matrix-assisted laser desorption ionization mass spectrometry. .* Mass Spectrom. Rev, 1991. **10**: p. 335-357.
75. Sabel, J., *Mass Spectrometry Analysis of Oligonucleotides Syntheses*. 2011, Integrated DNA Technologies
76. Cotter, R.J., *Time-of-flight Mass Spectrometry for Structural Analysis of Biological Molecules* Anal. Chem. , 1992. **64**: p. A1027-A1039.
77. Chait, B.T.a.K., S. B. H. , *Weighing Naked Proteins - Practical, High-accuracy Mass Measurement of Peptides and Proteins.* Science, 1992. **257**: p. 1885-1894.
78. Vertes, A.a.L., R. D., *Sublimation Versus Fragmentation in Matrix-assisted Laser Desorption.* Chem. Phys. Lett., 1990. **171**: p. 284-290.
79. Johnson, R.E.a.S., B. U. R., *Laser-pulse Ejection of Organic Molecules from A Matrix - Lessons from Fast-ion-induced Ejection* Rapid. Commun. Mass Spectrom, 1991. **5**: p. 574-578.
80. (JCBN), J.C.o.B.N., *(1984) Nomenclature and symbolism for amino acids and peptides. .* European Journal of Biochemistry, 1983. **138**: p. 9-37.

10. Appendix

10.1 Abbreviations

According to the recommended IUPAC-IUB joint commission on biochemical nomenclature all used amino acids were abbreviated according to the 1-letter or 3-letter code[80].

aa	amino acid(s)
Å	Ångström
ACN	Acetonitrile
APS	Ammoniumperoxodisulfat
Boc	*tert*-butoxycarbonyl-
Da	Dalton
DCM	Dichlormethane
DIPEA	*N*,*N*`-diisopropylethylamine
DMAP	4-dimethylamino pyridine
DMF	dimethyl formamide
DNA	desoxyribonuclein acid
E. coli	*Escherichia coli*
EDT	1,2-ethanedithiol
EDTA	ethylenediaminetetraacetic acid
eq.	equivalent(s)
ESI	electrospray ionization
Et-	ethyl-
Et_2O	Diethylether
FA	formic acid
Fmoc	9-fluorenylmethoxy carbonyl-
h	hours
HATU	2-(1H-azabenzotriazole-1-yl)1,1,3,3-tetramethylaminium hexafluorophosphate
HBTU	2-(1H-benzotriazole-1-yl)1,1,3,3-tetramethylaminium hexafluorophosphate
HOBt	*N*-hydroxybenzotriazole
HPLC	high pressure liquid chromatography
LC-MS	liquid chromatography mass spectrometry
MeOH	Methanol

min	minute(s)
mL	milliliter(s)
o.n.	overnight
O*t*Bu	*tert*-butanol
PA	protruding arm
NH_4OAc	ammonium acetate
nm	nanometer
pH	pondus Hydrogenii
rt	room temperature
SPPS	solid phase peptide synthesis
TCEP	tris(2-carboxyethyl)phosphine
TEAA	triethylammonium acetate
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TIS	triisopropyl silane
TRIS	tris(hydroxymethyl)aminomethane
UV	ultraviolet
WT	wild type
°C	degree celsius
µ	micro
3-HPA	3-hydroxy picolinic acid

10.2 List of publications

1. Sprengel, A., Lill, P., Stegemann, P., Bravo-Rodriguez, K., Schöneweiß, E.-C., Merdanovic, M., Gudnason, D., Aznauryan, M., Gamrad, L., Barcikowski, S., Sanchez-Gracia, E., Birkedal, V., Gatsogiannis, C., Ehrmann, M., Saccà, B., Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions. *Nat Commun.*, **2017**.

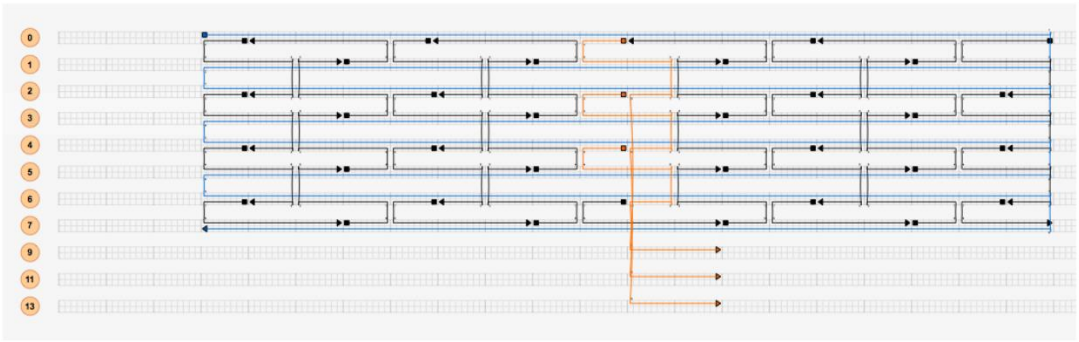
2. Poepsel S, Sprengel A, Sacca B, Kaschani F, Kaiser M, Gatsogiannis C, Raunser S, Clausen T, Ehrmann M. Determinants of amyloid fibril degradation by the PDZ protease HTRA1. *Nat. Chem. Biol*. **2015** Oct 5

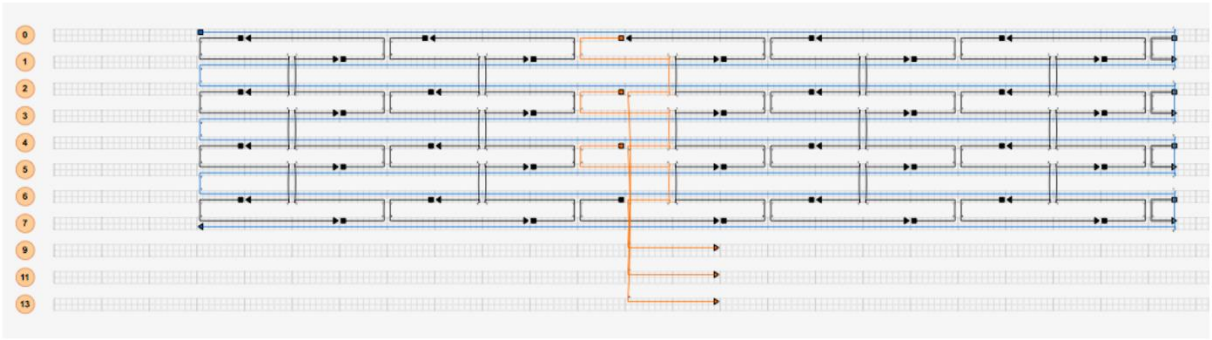
3. Saccà B, Ishitsuka Y, Meyer R, Sprengel A, Schöneweiß EC, Nienhaus GU, Niemeyer CM. Reversible Reconfiguration of DNA Origami Nanochambers Monitored by Single-Molecule FRET. *Angew Chem Int Ed Engl*. **2015** Jan 28.

4. Saccà, B., Sprengel, A. and Feldkamp, U., De Novo Design of Nucleic Acid Structures, *De novo Molecular Design* (ed G. Schneider). 2013, Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

10.3 Supplementary Information

10.3.1 6prism – design scheme and staples





Full list of sequences for the 6prism DNA host

I B-23,31	CCCTTATAAAGCCGGCGAACGTGGCGAGAAAG
I B-23,63	GGTTGAGTAAGGGAGCCCCCGATTTAGAGCTT
I B-23,95	CCACTATTGAGGTGCCGTAAAGCACTAAATCG
I B-23,127	AGGGCGAAGAACCATCACCCAAATCAAGTTTT
I B-43,31	CTGGGGTGGGTTTGCCCCAGCAGGGGCAAAAT
I B-43,63	TAATTGCGCCCTGAGAGAGTTGCACGAGATAG
I B-43,95	TCGGGAAACGGGCAACAGCTGATTACAAGAGT
I B-43,127	AATCGGCCCGCCAGGGTGGTTTTTAACGTCAA
I B-63,31	CGCAACTGTGTTATCCGCTCACAATGTAAAGC
I B-63,63	CTCTTCGCAATCATGGTCATAGCTACTCACAT
I B-63,95	GATGTGCTAGAGGATCCCCGGGTACTTTCCAG
I B-63,127	CCAGGGTTGTGCCAAGCTTGCATGCATTAATG
I B-83,31	ATAGGAACCCGGCACCGCTTCTGGTCAGGCTG
I B-83,63	GCCTTCCTGGCCTCAGGAAGATCGGTGCGGGC
I B-83,95	ATGTGAGCATCTGCCAGTTTGAGGGAAAGGGG

I B-83,127	GTGGGAACCGTTGGTGTAGATGGGGGGTAACG
I G-13,150	TTCGATGGCCCACTACGTAAACCGTCTATCAGGGTTTTTCGGTTTGC
I G-17,14	GAAGGGAATTTTTTGGTGGTTCCGAAATCCGAAAATC
I G-17,46	GACGGGGAAATCAAAAGAATAGCCGCAAGCGG
I G-17,78	GAACCCTAGTTGTTCCAGTTTGGAGCCCTTCA
I G-17,110	TTGGGGTCAAAGAACGTGGACTCCCTTTTCAC
I G-37,14	CTGTTTGATTTTTAGCCGGAAGCATAAAGTTCCACAC
I G-37,46	TCCACGCTCCTAATGAGTGAGCTAGTTTCCTG
I G-37,78	CCGCCTGGTTGCGCTCACTGCCCGCCGAGCTC
I G-37,110	CAGTGAGACCTGTCGTGCCAGCTGCCTGCAGG
I G-37,142	GTATTGGGAACGCGCGGGGAGAGGTTTTTTGTAAAAC
I G-57,14	AACATACGTTTTTAAGCGCCATTCGCCATTGCCGGAA
I G-57,46	TGTGAAATTTGGGAAGGGCGATCGCACTCCAG
I G-57,78	GAATTCGTTATTACGCCAGCTGGCGGACGACG
I G-57,110	TCGACTCTGCAAGGCGATTAAGTTCGCATCGT
I G-57,142	GACGGCCATTCCCAGTCACGACGTTTTTTGTAATGGG
I G-77,14	ACCAGGCATTTTTTTAAATCAGCTCATTTTTTAACCA
I G-77,46	CCAGCTTTGCCATCAAAAATAATTCGCGTCTG
I G-77,78	ACAGTATCGTAGCCAGCTTTCATCAACATTAA
I G-77,110	AACCGTGCGAGTAACAACCCGTCGGATTCTCC
I G-77,142	ATAGGTCAAAACGGCGGATTGACCTT
II B-23,31	ATCAATATAATATTTTGTTAAAATTCGCATTA
II B-23,63	AATTAATGTTGTATAAGCAAATATTTAAATTG
II B-23,95	AGATCTACTGATAATCAGAAAAGCCCCAAAAA
II B-23,127	GAGTCTGGAAAACTAGCATGTCAATCATATGT
II B-43,31	AGCTGAAATAATGTGTAGGTAAAGAAATCACC
II B-43,63	TAGTAGCAAACCCTCATATATTTTAGCTGATA
II B-43,95	CAAGGCAACTTTATTTCAACGCAATTTTTGAG
II B-43,127	AGCCTCAGTTATGACCCTGTAATATTGCCTGA
II B-63,31	AAAGCGAAGATACATTTCGCAAATGGGGCGCG
II B-63,63	GGTCAGGATCTGCGAACGAGTAGAACTAATAG
II B-63,95	TTTTGATAAGTTTCATTCCATATACATACAGG
II B-63,127	AGAGCTTATTTAAATATGCAACTAAGCAATAA
II B-83,31	CAGACGACAAAAGATTAAGAGGAACGAGCTTC
Ii B-83,63	CTTTTGCATATTATAGTCAGAAGCCTCCAACA
II B-83,95	ATAGTAAACCATAAATCAAAAATCATTGCTCC
II B-83,127	TACTGCGGAATGCTTTAAACAGTTGATGGCTT
II G-13,150	TTGATGAACGGTAATCGTAGCAAACAAGAGAATCTTTTTGGTTGTAC
II G-17,14	AATTTTTGTTTTTAAGGCCGGAGACAGTCATTCAAAA
II G-17,46	TAAACGTTGATATTCAACCGTTCTAAATGCAA
II G-17,78	CAGGAAGACCGGAGAGGGTAGCTAGGATAAAA
II G-17,110	ACCCCGGTAAAGGCTATCAGGTCACTTTTGCG
II G-37,14	GGGTGAGATTTTTAGCTATATTTTCATTTGGTCAATA
II G-37,46	TGCCTGAGAGGTGGCATCAATTCTTTTAGTTT
II G-37,78	ATTTTTAGTTAACATCCAATAAATACAGTTGA
II G-37,110	GGAGAAGCAGAATTAGCAAAATTAAAGTACGG

II G-37,142	CAAAAACAAGCATAAAGCTAAATCTTTTTCTGTAGCT
II G-57,14	ACCTGTTTTTTTTATATCGCGTTTTAATTGCCCGAAA
II G-57,46	GACCATTACCAGACCGGAAGCAAAAAAGCGGA
II G-57,78	TTCCCAATTTAGAGAGTACCTTTAAGGTCTTT
II G-57,110	TGTCTGGAAGAGGTCATTTTTGCGCAGAAAAC
II G-57,142	CAACATGTATTGCTGAATATAATGTTTTTCATTGAAT
II G-77,14	GACTTCAATTTTTAACACTATCATAACCCTCGTTTAC
II G-77,46	TTGCATCAGATAAAAACCAAAATAGCGAGAGG
II G-77,78	ACCCTGACAAAGAAGTTTTGCCAGAGGGGGTA
II G-77,110	GAGAATGAATGTTTAGACTGGATAGCGTCCAA
II G-77,142	CCCCCTCAAATCGTCATAAATATTTT
III B-23,31	CAACGTAACGCCAAAAGGAATTACGAGGCATA
III B-23,63	GGCTTGCCGAATACCACATTCAACTAATGCAG
III B-23,95	AGTAAATTACAGGTAGAAAGATTCATCAGTTG
III B-23,127	CTTTAATCTTAATAAAACGAACTAACGGAACA
III B-23,159	TAAGAACTACCAGTCAGGACGTTGGGAAGAAA
III B-43,31	TAAAATACTGGCTGACCTTCATCAACCCAAAT
III B-43,63	CCTAAAACAGATGAACGGTGTACAGTGAATAA
III B-43,95	AACACTCAGGGAACCGAACTGACCGAACGAGT
III B-43,127	AAGCGCGAAGCCGGAACGAGGCGCAATTTCAA
III B-43,159	TCATCGCCTGTCGAAATCCGCGACTGCGATTT
III B-63,31	CTAAAGGACTACAGAGGCTTTGAGTAAACGGG
III B-63,63	TTGAAAATGAAAGACAGCATCGGAGGCACCAA
III B-63,95	AGCCTTTAGCCGCTTTTGCGGGATTACACTAA
III B-63,127	TTCGAGGTATTCGGTCGCTGAGGCATTATACC
III B-63,159	GATAGTTGGACAACAACCATCGCCGATTTGTA
III B-83,31	TATAAGTAGGATTTTGCTAAACAAAGGAACAA
III B-83,63	TACTCAGGTTCCAGACGTTAGTAATTTTCACG
III B-83,95	ACCGCCACAGTTAGCGTAACGATCCCAAAAGG
III B-83,127	CACCACCCAACGCCTGTAGCATTCAGCTTGCT
III B-83,159	ATAGGAACTAACACTGAGTTTCGTTTGATACC
III G-17,14	GTAAGAGCTTTTTGAACCGGATATTCATTAGAGTAAT
III G-17,46	ATACATAACAAAGCTGCTCATTCAGACCAGGC
III G-17,78	AGATTTAGCTGACGAGAAACACCAAACTTTGA
III G-17,110	ACATTATTGGGCTTGAGATGGTTTAGACGGTC
III G-17,142	AATCTACGATTGTGAATTACCTTACTGCTCCA
III G-17,171	TTATTATGGCTCTT
III G-37,14	CTTGACAATTTTTATGAGGAAGTTTCCATGACTAAAG
III G-37,46	GCATAGGCGTAATGCCACTACGAAACGAGGGT
III G-37,78	AAGAGGACGAAAGAGGCAAAAGAACGTCACCC
III G-37,110	AATCATAATCTTTGACCCCCAGCGTTGCAGGG
III G-37,142	TGTTACTTAACAAAGTACAACGGACACGCATA
III G-37,171	TTAATTGTGATATT
III G-57,14	ACTTTTTCTTTTTCGGAGTGAGAATAGAACTTTCAAC
III G-57,46	AGCAACGGATTGCGAATAATAATTATGAATTT
III G-57,78	TCAGCAGCCTCCAAAAAAAAGGCTTAAAGTTT

III G-57,110	AGTTAAAGATTGTATCGGTTTATCCACAGACA
III G-57,142	ACCGATATGAATTTCTTAAACAGCCACCAGTA
III G-57,171	TTACAATCGCCGTT
III G-77,14	AGTTTCAGTTTTTGATAAGTGCCGTCGAGAGGGTTGA
III G-77,46	TCTGTATGTAGCCCGGAATAGGTGTATCACCG
III G-77,78	TGTCGTCTAGGTTTAGTACCGCCACCCTCAGA
III G-77,110	GCCCTCATCCTCAGAACCGCCACCCTCAGAGC
III G-77,142	CAAACTACTCATTTTCAGGGATAGCAAGCCCA
III G-77,171	TTTACCGCCATGTT
IV B-23,31	ATAAATCCAGGATTAGCGGGGTTTTGCTCAGT
IV B-23,63	CAGTCTCTTTAAGAGGCTGAGACTCCTCAAGA
IV B-23,95	ATACATGGTTCGGAACCTATTATTCTGAAACA
IV B-23,127	GGTAATAAGCCCGTATAAACAGTTAATGCCCC
IV B-43,31	AATCAGTAAGCATTGACAGGAGGTACAAACAA
IV B-43,63	GCGTCAGACCGCCACCAGAACCACGGAAAGCG
IV B-43,95	TTCGGTCAAACCGCCACCCTCAGATAAGCGTC
IV B-43,127	TCTTTTCAGGAACCGCCTCCCTCAAGTGTACT
IV B-63,31	TAAAGGTGATTACCATTAGCAAGGAGCACCGT
IV B-63,63	ACACCACGGAATTAGAGCCAGCAATGCCTTTA
IV B-63,95	AATAGAAAAATTATCACCGTCACCTCGGCATT
IV B-63,127	AGACAAAATAAATATTGACGGAAAGTTTGCCA
IV B-83,31	GACGGGAGGCATGATTAAGACTCCTACATACA
IV B-83,63	GTCAGAGGACGCAATAATAACGGAACGCAAAG
IV B-83,95	GATAACCCCCGAACAAAGTTACCATCACAATC
IV B-83,127	ATAAGAGCCGAAGCCCTTTTTAAGAGCGCCAA
IV G-13,150	TTTGCCTTGAGTAACAGTGTTTTAACGGGGTCAGTTTTTAACCAGAG
IV G-17,14	ACCAGGCGTTTTTATTGGCCTTGATATTCTGAGGCAG
IV G-17,46	GAAGGATTTCATTAAAGCCAGAATCACCAGAG
IV G-17,78	TGAAAGTAGAATTTACCGTTCCAGGCCACCAC
IV G-17,110	CTGCCTATCTTTTGATGATACAGGGAGCCGCC
IV G-37,14	GTCAGACGTTTTTTGAAACCATCGATAGCCCGGAAAC
IV G-37,46	CCGCCGCCGCGACAGAATCAAGTTAATCACCA
IV G-37,78	CCTCAGAGCTGTAGCGCGTTTTCAGACTTGAG
IV G-37,110	ACCCTCAGTAGCCCCCTTATTAGCTTATTCAT
IV G-37,142	CCACCACCTAATCAAAATCACCGGTTTTTATTGAGGG
IV G-57,14	GTCACCAATTTTTTTAGCAAACGTAGAAAATTATTACG
IV G-57,46	GTAGCACCGCAACATATAAAAGAAATACCCAA
IV G-57,78	CCATTTGGGAATAAGTTTATTTTGGAAGGAAA
IV G-57,110	TAAAGGTGATTCATATGGTTTACCAAAAGTAA
IV G-57,142	AGGGAAGGGGGCGACATTCAACCGTTTTTGCAATAGC
IV G-77,14	CAGTATGTTTTTTACATAAAAACAGGGAAGCGCATTA
IV G-77,46	AAGAACTGAATTAACTGAACACCCTGAACAAA
IV G-77,78	CCGAGGAAGTAATTGAGCGCTAATATCAGAGA
IV G-77,110	GCAGATAGACAAGAATTGAGTTAAGCCCAATA
IV G-77,142	TATCTTACAAGAAACAATGAAATATT
V B-23,31	GCTTATCCAAATGAAAATAGCAGCCTTTACAG

V B-23,63	TTAGCGAACAAATAAGAAACGATTTTTTGTTT
V B-23,95	AAGCCTTACAAAATAAACAGCCATATTATTTA
V B-23,127	ACCCAGCTGAGCGTCTTTCCAGAGCCTAATTT
V B-43,31	CGACAAAAATTTTCATCGTAGGAATATAGAAG
V B-43,63	CGACAATACCGCACTCATCGAGAAGAGGCGTT
V B-43,95	ACGCGCCTTATCATTCCAAGAACGAGGTTTTG
V B-43,127	GAACAAGATAGAAACCAATCAATATATTTTGC
V B-63,31	CGCGAGAAAACATGTAATTTAGGCTAAAGTAC
V B-63,63	AATTTCATCTTAATTGAGAATCGCCCAGACGA
V B-63,95	GAAATACCTACCAGTATAAAGCCAAATGCAGA
V B-63,127	AATAAGAACTGTTTAGTATCATATTAAGTCCT
V B-83,31	ATTAATTATTGGGTTATATAACTAACAAAGAA
V B-83,63	CCTTTTTTTGAGAGACTACCTTTTTTTTAGTT
V B-83,95	ATATGTGAAGTCAATAGTGAATTTAATGGTTT
V B-83,127	CGTCGCTAATAGCGATAGCTTAGAAGGCGTTA
V G-13,150	TTTCTTACCAACGCTAACACAATTTTATCCTGAATTTTTCTAATTTA
V G-17,14	AGAGAATATTTTTTAGCAAGCAAATCAGATCATTACC
V G-17,46	AACGTCAAGGTATTCTAAGAACGCCAAGCAAG
V G-17,78	TCCCAATCCCTCCCGACTTGCGGGGGTATTAA
V G-17,110	GCCAGTTAAATCAAGATTAGTTGCATCGGCTG
V G-37,14	GCGCCCAATTTTTCAGTAATAAGAGAATAAGAGGCAT
V G-37,46	CCGTTTTTGGTAAAGTAATTCTGTCATATTTA
V G-37,78	ACCAAGTAAACAACATGTTCAGCTACGCTCAA
V G-37,110	TCTTTCCTGTTTATCAACAATAGAGCGTTATA
V G-37,142	CGAGCATGAAAATAATATCCCATCTTTTTAATTACTA
V G-57,14	TTTCGAGCTTTTTCAAATCCAATCGCAAGTATGTAAA
V G-57,46	ACAACGCCAACTTTTTCAAATATATAACCTCC
V G-57,78	CAGTAGGGCTTCTGACCTAAATTTATCAAAAT
V G-57,110	CAAATTCTGACCGTGTGATAAATATTAAGACG
V G-57,142	GAAAAAGCTAAACACCGGAATCATTTTTTTAGAATCC
V G-77,14	TGCTGATGTTTTTTGAAACAAACATCAAGAAAACAAA
V G-77,46	GGCTTAGGCATTTAACAATTTCATTTGAATTA
V G-77,78	CATAGGTCAATGGAAACAGTACATAAATCAAT
V G-77,110	CTGAGAAGGTGAATAACCTTGCTTCTGTAAAT
V G-77,142	TTGAAAACTTAATTAATTTTCCCTTT
VI B-23,31	AGCGGAATTTCATTTCAATTACCTGAGCAAAA
VI B-23,63	ATGATGGCACCAAGTTACAAAATCGCGCAGAG
VI B-23,95	TTTGGATTAACAATAACGGATTCGCCTGATTG
VI B-23,127	AGAACCTATATACAGTAACAGTACCTTTTACA
VI B-23,159	AAACAGAATGCGTAGATTTTCAGGTTTAACGT
VI B-43,31	GCCTGCAAATTTTAAAAGTTTGAGCCAGAAGG
VI B-43,63	CAAATGAACTCGTATTAAATCCTTATTATCAG
VI B-43,95	ACCTCAAAAGAAGTATTAGACTTTCCTGATTG
VI B-43,127	GTCAGTTGGAGCCGTCAATAGATAGAAGGGTT
VI B-43,159	GAGGAAGGTATCTTTAGGAGCACTTGCACGTA
VI B-63,31	GGTAATATCCGAACGAACCACCAGTTAACACC

VI B-63,63	TTGCAACACTGATAGCCCTAAAACGCCAGCAG
VI B-63,95	ACATTTTGTTTTGAATGGCTATTACTTGCTGA
VI B-63,127	ATTTACATGAAAGCGTAAGAATACAATATCTG
VI B-63,159	CAGTAATATCTGGCCAACAGAGATAAGGAATT
VI B-83,31	GCGCTGGCTGATTAGTAATAACATGCCTTGCT
VI B-83,63	AACCACCACACGCAAATTAACCGTGCCAGCCA
VI B-83,95	TACAGGGCAGTGAGGCCACCGAGTAAATACCT
VI B-83,127	CACGTATAACGCCAGAATCCTGAGAATGGATT
VI B-83,159	AGCGGGAGGGCCGATTAAAGGGATCACACGAC
VI G-17,14	GAAGATGATTTTTCGGAACAAAGAAACCATAACATTA
VI G-17,46	GCGAATTATATCATCATATTCCTGTGCCCGAA
VI G-17,78	CTTTGAATAATTCATCAATATAATACAAACAA
VI G-17,110	TCGGGAGAATACTTCTGAATAATGATACATTT
VI G-17,142	CAGATGAACCATATCAAAATTATTAACAACTA
VI G-17,171	TTGAAATATAAATT
VI G-37,14	TCATTTTGTTTTTGGTGAGGCGGTCAGTACAGAAGAT
VI G-37,46	CGTTATTACAGTGCCACGCTGAGAATCGCCAT
VI G-37,78	TTCGACAAAAATCTAAAGCATCACGTCTTTAA
VI G-37,110	GAGGATTTTATCAAACCCTCAATCGTGGCACA
VI G-37,142	ATAGATTAGCAAATCAACAGTTGAAGAACCCT
VI G-37,171	TTTAAAATTATCTT
VI G-57,14	AAAACAGATTTTTAGAACTCAAACTATCGCACTTGCC
VI G-57,46	TAAAAATACCAGAACAATATTACCTGTAGCAA
VI G-57,78	TGCGCGAAGGAAAAACGCTCATGGAAAAGAGT
VI G-57,110	GACAATATACGCTCAATCGTCTGAAAGTGTTT
VI G-57,142	TCTGACCTTGGCAGATTCACCAGTTTTAGACA
VI G-57,171	TTGACATAAAGGTT
VI G-77,14	TGAGTAGATTTTTAAAGCGAAAGGAGCGGGCGCTAGG
VI G-77,46	TACTTCTTAAGTGTAGCGGTCACGCTGCGCGT
VI G-77,78	CTGTCCATCACCCGCCGCGCTTAATGCGCCGC
VI G-77,110	TTATAATCGCGTACTATGGTTGCTTTGACGAG
VI G-77,142	GGAACGGTACGTGCTTTCCTCGTTAGAATCAG
VI G-77,171	TTCAGGACTAAATT

3D staples for 2T-180° face-to-face connections

I/II G-77,14	ACCAGGCATTTTTTTAAATCAGCTCATTTTTTCGCATTA
I/II G-77,46	CCAGCTTTGCCATCAAAAATAATTTTTTAAATTG
I/II G-77,78	ACAGTATCGTAGCCAGCTTTCATCTTCCCAAAAA
I/II G-77,110	AACCGTGCGAGTAACAACCCGTCGTTTCATATGT
I/II G-77,142	ATAGGTCAAAACGGCGGATTGACCTTGATGAACG
II G-17,142	GTAATCGTAGCAAACAAGAGAATCTTTTTGGTTGTAC
II/I B-23,31	ATCAATATAATATTTTGTTAAAATTTTTTAACCA
II/I B-23,63	AATTAATGTTGTATAAGCAAATATTTCGCGTCTG

II/I B-23,95	AGATCTACTGATAATCAGAAAAGCTTAACATTAA
II/I B-23,127	GAGTCTGGAAAACTAGCATGTCAATTGATTCTCC
II/III G-77,14	GACTTCAATTTTTAACACTATCATAACCCTTGAGGCATA
II/III G-77,46	TTGCATCAGATAAAAACCAAAATATTTAATGCAG
II/III G-77,78	ACCCTGACAAAGAAGTTTTGCCAGTTATCAGTTG
II/III G-77,11	GAGAATGAATGTTTAGACTGGATATTACGGAACA
II/III G-77,142	CCCCCTCAAATCGTCATAAATATTTTGGAAGAAA
III/II B-23,31	CAACGTAACGCCAAAAGGAATTACTTTCGTTTAC
III/II B-23,63	GGCTTGCCGAATACCACATTCAACTTGCGAGAGG
III/II B-23,95	AGTAAATTACAGGTAGAAAGATTCTTAGGGGGTA
III/II B-23,127	CTTTAATCTTAATAAAACGAACTATTGCGTCCAA
III B-23,159s	TAAGAACTACCAGTCAGGACGTTG
III/IV G-77,14	AGTTTCAGTTTTTGATAAGTGCCGTCGAGTTTGCTCAGT
III/IV G-77,46	TCTGTATGTAGCCCGGAATAGGTGTTCCTCAAGA
III/IV G-77,78	TGTCGTCTAGGTTTAGTACCGCCATTCTGAAACA
III/IV G-77,110	GCCCTCATCCTCAGAACCGCCACCTTAATGCCCC
III/IV G-77,142	CAAACTACTCATTTTCAGGGATAGTTTGCCTTGA
III B-83,150	CAAGCCCAATAGGAACTAACACTGAGTTTCGTTTGATACC
IV/III B-23,31	ATAAATCCAGGATTAGCGGGGTTTTTAGGGTTGA
IV/III B-23,63	CAGTCTCTTTAAGAGGCTGAGACTTTTATCACCG
IV/III B-23,95	ATACATGGTTCGGAACCTATTATTTTCCCTCAGA
IV/III B-23,127	GGTAATAAGCCCGTATAAACAGTTTTCTCAGAGC
IV G-17,142	GTAACAGTGTTTTAACGGGGTCAGTTTTTAACCAGAG
IV/V G-77,14	CAGTATGTTTTTTACATAAAAACAGGGAATTCTTTACAG
IV/V G-77,46	AAGAACTGAATTAACTGAACACCCTTTTTTGTTT
IV/V G-77,78	CCGAGGAAGTAATTGAGCGCTAATTTATTATTTA
IV/V G-77,110	GCAGATAGACAAGAATTGAGTTAATTCCTAATTT
IV/V G-77,142	TATCTTACAAGAAACAATGAAATATTTCTTACCA
V/IV B-23,31	GCTTATCCAAATGAAAATAGCAGCTTGCGCATTA
V/IV B-23,63	TTAGCGAACAAATAAGAAACGATTTTTGAACAAA
V/IV B-23,95	AAGCCTTACAAAATAAACAGCCATTTATCAGAGA
V/IV B-23,127	ACCCAGCTGAGCGTCTTTCCAGAGTTGCCCAATA
V G-17,142	ACGCTAACACAATTTTATCCTGAATTTTTCTAATTTA
V/VI G-77,14	TGCTGATGTTTTTTGAAACAAACATCAAGTTGAGCAAAA
V/VI G-77,46	GGCTTAGGCATTTAACAATTTCATTTGCGCAGAG
V/VI G-77,78	CATAGGTCAATGGAAACAGTACATTTCCTGATTG
V/VI G-77,110	CTGAGAAGGTGAATAACCTTGCTTTTCTTTTACA
V/VI G-77,142	TTGAAAACTTAATTAATTTTCCCTTTTTTAACGT
VI/V B-23,31	AGCGGAATTTCATTTCAATTACCTTTAAAACAAA
VI/V B-23,63	ATGATGGCACCAAGTTACAAAATCTTTTGAATTA
VI/V B-23,95	TTTGGATTAACAATAACGGATTCGTTAAATCAAT
VI/V B-23,127	AGAACCTATATACAGTAACAGTACTTCTGTAAAT
VI B-23,159s	AAACAGAATGCGTAGATTTTCAGG
VI/I G-77,14	TGAGTAGATTTTTAAAGCGAAAGGAGCGGTTCGAGAAAG
VI/I G-77,46	TACTTCTTAAGTGTAGCGGTCACGTTTAGAGCTT
VI/I G-77,78	CTGTCCATCACCCGCCGCGCTTAATTCTAAATCG

VI/I G-77,110	TTATAATCGCGTACTATGGTTGCTTTCAAGTTTT
VI/I G-77,142	GGAACGGTACGTGCTTTCCTCGTTTTCGATGGCC
VI B-83,150	AGAATCAGAGCGGGAGGGCCGATTAAAGGGATCACACGAC
I/VI B-23,31	CCCTTATAAAGCCGGCGAACGTGGTTGCGCTAGG
I/VI B-23,63	GGTTGAGTAAGGGAGCCCCCGATTTTCTGCGCGT
I/VI B-23,95	CCACTATTGAGGTGCCGTAAAGCATTTGCGCCGC
I/VI B-23,127	AGGGCGAAGAACCATCACCCAAATTTTTGACGAG
I G-17,142	CACTACGTAAACCGTCTATCAGGGTTTTTCGGTTTGC

3D staples for 120° face-to-face connections

I/II_G-120(1)	ACCAGGCATTTTTTTAAATCAGCTCATTTTATTCGCATTA
I/II_G-120(2)	CCAGCTTTGCCATCAAAAATAATTCATTTAAATTG
I/II_G-120(3)	ACAGTATCGTAGCCAGCTTTCATCAGCCCCAAAAA
I/II_G-120(4)	AACCGTGCGAGTAACAACCCGTCGGAATCATATGT
I/II_G-120(5)	ATAGGTCAAAACGGCGGATTGACCTTTGATGAACG
II/I B-120(1)	ATCAATATAATATTTTGTTAAATTAACCA
II/I B-120(2)	AATTAATGTTGTATAAGCAAATGCGTCTG
II/I B-120(3)	AGATCTACTGATAATCAGAAAAACATTAA
II/I B-120(4)	GAGTCTGGAAAACTAGCATGTCATTCTCC
II/I B-120(5)	GTAATCGTAGCAAACAAGAGAATCTTTTTGGTTGTAC
II/III_G-120(1)	GACTTCAATTTTTAACACTATCATAACCCTACGAGGCATA
II/III_G-120(2)	TTGCATCAGATAAAAACCAAAATAGACTAATGCAG
II/III_G-120(3)	ACCCTGACAAAGAAGTTTTGCCAGATCATCAGTTG
II/III_G-120(4)	GAGAATGAATGTTTAGACTGGATAGTAACGGAACA
II/III_G-120(5)	CCCCCTCAAATCGTCATAAATATTTTGGGAAGAAA
III/II B-120(1)	CAACGTAACGCCAAAAGGAATTCGTTTAC
III/II B-120(2)	GGCTTGCCGAATACCACATTCACGAGAGG
III/II B-120(3)	AGTAAATTACAGGTAGAAAGATGGGGGTA
III/II B-120(4)	CTTTAATCTTAATAAAACGAACCGTCCAA
III/II B-120(5)	TAAGAACTACCAGTCAGGACGT
III/IV_G-120(1)	AGTTTCAGTTTTTGATAAGTGCCGTCGAGATTTGCTCAGT
III/IV_G-120(2)	TCTGTATGTAGCCCGGAATAGGTGTCTCCTCAAGA
III/IV_G-120(3)	TGTCGTCTAGGTTTAGTACCGCCACTTCTGAAACA
III/IV_G-120(4)	GCCCTCATCCTCAGAACCGCCACCCTTAATGCCCC
III/IV_G-120(5)	CAAACTACTCATTTTCAGGGATAGCTTTGCCTTGA
III/IV_G-120(6)	AAGCCCAATAGGAACTAACACTGAGTTTCGTTTGATACC
IV/III B-120(1)	ATAAATCCAGGATTAGCGGGGTGGGTTGA
IV/III B-120(2)	CAGTCTCTTTAAGAGGCTGAGAATCACCG
IV/III B-120(3)	ATACATGGTTCGGAACCTATTACCTCAGA
IV/III B-120(4)	GGTAATAAGCCCGTATAAACAGTCAGAGC
IV/III B-120(5)	GTAACAGTGTTTTAACGGGGTCAGTTTTTAACCAGAG
IV/V_G-120(1)	CAGTATGTTTTTTACATAAAAACAGGGAAGGCCTTTACAG
IV/V_G-120(2)	AAGAACTGAATTAACTGAACACCCTTTTTTTGTTT

IV/V_G-120(3)	CCGAGGAAGTAATTGAGCGCTAATAATATTATTTA
IV/V_G-120(4)	GCAGATAGACAAGAATTGAGTTAAGAGCCTAATTT
IV/V_G-120(5)	TATCTTACAAGAAACAATGAAATATTTTCTTACCA
V/IV B-120(1)	GCTTATCCAAATGAAAATAGCACGCATTA
V/IV B-120(2)	TTAGCGAACAAATAAGAAACGAGAACAAA
V/IV B-120(3)	AAGCCTTACAAAATAAACAGCCTCAGAGA
V/IV B-120(4)	ACCCAGCTGAGCGTCTTTCCAGCCCAATA
V/IV B-120(5)	ACGCTAACACAATTTTATCCTGAATTTTTCTAATTTA
V/VI_G-120(1)	TGCTGATGTTTTTTGAAACAAACATCAAGACTGAGCAAAA
V/VI_G-120(2)	GGCTTAGGCATTTAACAATTTCATTTCGCGCAGAG
V/VI_G-120(3)	CATAGGTCAATGGAAACAGTACATACGCCTGATTG
V/VI_G-120(4)	CTGAGAAGGTGAATAACCTTGCTTCACCTTTTACA
V/VI_G-120(5)	TTGAAAACTTAATTAATTTTCCCTTGGTTTAACGT
VI/V B-120(1)	AGCGGAATTTCATTTCAATTACAAACAAA
VI/V B-120(2)	ATGATGGCACCAAGTTACAAAATGAATTA
VI/V B-120(3)	TTTGGATTAACAATAACGGATTAATCAAT
VI/V B-120(4)	AGAACCTATATACAGTAACAGTTGTAAAT
VI/V B-120(5)	AAACAGAATGCGTAGATTTTCA
VI/I_G-120(1)	TGAGTAGATTTTTAAAGCGAAAGGAGCGGGGGCGAGAAAG
VI/I_G-120(2)	TACTTCTTAAGTGTAGCGGTCACGCTTTAGAGCTT
VI/I_G-120(3)	CTGTCCATCACCCGCCGCGCTTAATCACTAAATCG
VI/I_G-120(4)	TTATAATCGCGTACTATGGTTGCTTATCAAGTTTT
VI/I_G-120(5)	GGAACGGTACGTGCTTTCCTCGTTATTCGATGGCC
VI/I_G-120(6)	GAATCAGAGCGGGAGGGCCGATTAAAGGGATCACACGAC
VI/I B-120(1)	CCCTTATAAAGCCGGCGAACGTCGCTAGG
VI/I B-120(2)	GGTTGAGTAAGGGAGCCCCCGATGCGCGT
VI/I B-120(3)	CCACTATTGAGGTGCCGTAAAGGCGCCGC
VI/I B-120(4)	AGGGCGAAGAACCATCACCCAATGACGAG
VI/I B-120(5)	CACTACGTAAACCGTCTATCAGGGTTTTTCGGTTTGC

3D staples for 240° face-to-face connections

I/II G-240(1)	ACCAGGCATTTTTTTAAATCAGCAAATTTTTGTTTTTAAGGCCGGAGACAGTCATTCAAAA
I/II G-240(2)	CCAGCTTTGCCATCAAAAGTAAACGTTGATATTCAACCGTTCTAAATGCAA
I/II G-240(3)	ACAGTATCGTAGCCAGCTACAGGAAGACCGGAGAGGGTAGCTAGGATAAAA
I/II G-240(4)	AACCGTGCGAGTAACAACTACCCCGGTAAAGGCTATCAGGTCACTTTTGCG
I/II G-240(5)	ATAGGTCAAAACGGCGGAGGTAATCGTAGCAAACAAGAGAATCTTTTTGGTTGTAC
II/I B-240(1)	ATCAATATAATATTTTGTTAAAATTCGCATTTCATTTTTTAACCA
II/I B-240(2)	AATTAATGTTGTATAAGCAAATATTTAAATTATAATTCGCGTCTG
II/I B-240(3)	AGATCTACTGATAATCAGAAAAGCCCCAAAATTCATCAACATTAA
II/I B-240(4)	GAGTCTGGAAAACTAGCATGTCAATCATATGCCGTCGGATTCTCC
II/I B-240(5)	TTGATGAACTTGACCTT
II/III G-240(1)	GACTTCAATTTTTAACACTATCAAGTAAGAGCTTTTTGAACCGGATATTCATTAGAGTAAT
II/III G-240(2)	TTGCATCAGATAAAAACCGATACATAACAAAGCTGCTCATTCAGACCAGGC

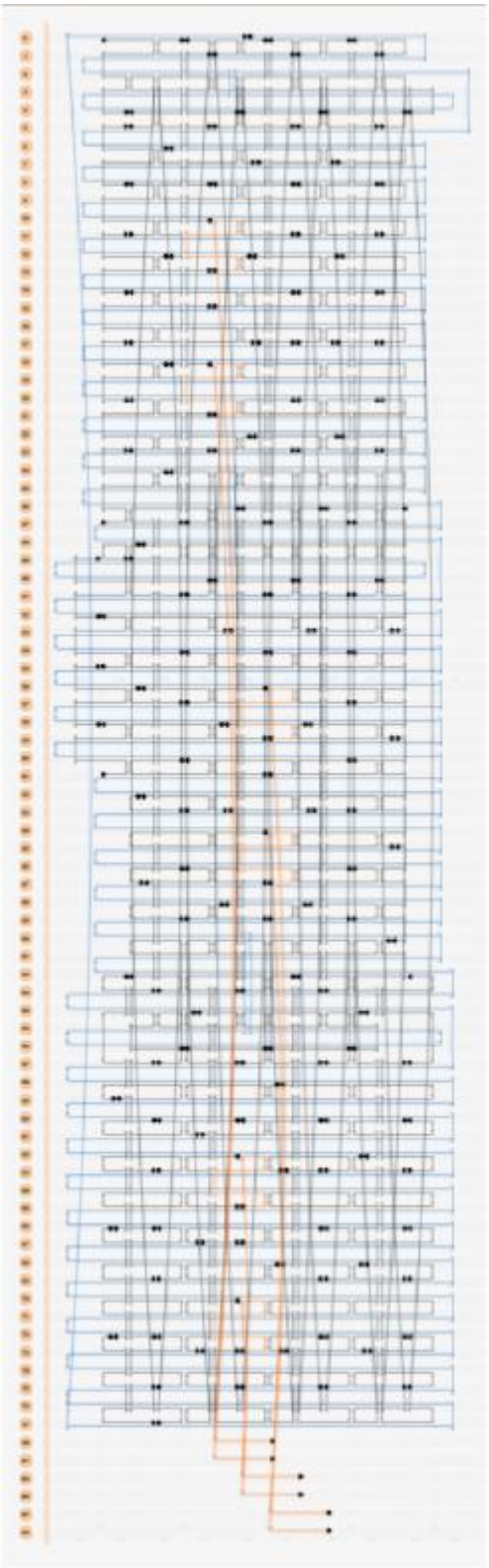
II/III G-240(3)	ACCCTGACAAAGAAGTTTGAGATTTAGCTGACGAGAAACACCAAACTTTGA
II/III G-240(4)	GAGAATGAATGTTTAGACAACATTATTGGGCTTGAGATGGTTTAGACGGTC
II/III G-240(5)	CCCCCTCAAATCGTCATAAAATCTACGATTGTGAATTACCTTACTGCTCCA
III/II B-240(1)	CAACGTAACGCCAAAAGGAATTACGAGGCATTAACCCTCGTTTAC
III/II B-240(2)	GGCTTGCCGAATACCACATTCAACTAATGCAAAAATAGCGAGAGG
III/II B-240(3)	AGTAAATTACAGGTAGAAAGATTCATCAGTTTGCCAGAGGGGGTA
III/II B-240(4)	CTTTAATCTTAATAAAACGAACTAACGGAACTGGATAGCGTCCAA
III/II B-240(5)	TAAGAACTACCAGTCAGGACGTTGGGAAGAAAATATTTT
III/IV G-240(1)	AGTTTCAGTTTTTGATAAGTGCCTACCAGGCGTTTTTATTGGCCTTGATATTCTGAGGCAG
III/IV G-240(2)	TCTGTATGTAGCCCGGAAAGAAGGATTTCATTAAAGCCAGAATCACCAGAG
III/IV G-240(3)	TGTCGTCTAGGTTTAGTAATGAAAGTAGAATTTACCGTTCCAGGCCACCAC
III/IV G-240(4)	GCCCTCATCCTCAGAACCCCTGCCTATCTTTTGATGATACAGGGAGCCGCC
III/IV G-240(5)	CAAACTACTCATTTTCAGAGTAACAGTGTTTTAACGGGGTCAGTTTTTAACCAGAG
IV/III B-240(1)	ATAAATCCAGGATTAGCGGGGTTTTGCTCAGGTCGAGAGGGTTGA
IV/III B-240(2)	CAGTCTCTTTAAGAGGCTGAGACTCCTCAAGTAGGTGTATCACCG
IV/III B-240(3)	ATACATGGTTCGGAACCTATTATTCTGAAACCCGCCACCCTCAGA
IV/III B-240(4)	GGTAATAAGCCCGTATAAACAGTTAATGCCCGCCACCCTCAGAGC
IV/III B-240(5)	TTTGCCTTGGGATAGCAAGCCCA
IV/V G-240(1)	CAGTATGTTTTTTACATAAAAACGAGAGAATATTTTTTAGCAAGCAAATCAGATCATTACC
IV/V G-240(2)	AAGAACTGAATTAACTGATAACGTCAAGGTATTCTAAGAACGCCAAGCAAG
IV/V G-240(3)	CCGAGGAAGTAATTGAGCATCCCAATCCCTCCCGACTTGCGGGGGTATTAA
IV/V G-240(4)	GCAGATAGACAAGAATTGTGCCAGTTAAATCAAGATTAGTTGCATCGGCTG
IV/V G-240(5)	TATCTTACAAGAAACAATAACGCTAACACAATTTTATCCTGAATTTTTCTAATTTA
V/IV B-240(1)	GCTTATCCAAATGAAAATAGCAGCCTTTACAAGGGAAGCGCATTA
V/IV B-240(2)	TTAGCGAACAAATAAGAAACGATTTTTTGTTACACCCTGAACAAA
V/IV B-240(3)	AAGCCTTACAAAATAAACAGCCATATTATTTGCTAATATCAGAGA
V/IV B-240(4)	ACCCAGCTGAGCGTCTTTCCAGAGCCTAATTAGTTAAGCCCAATA
V/IV B-240(5)	TTTCTTACCGAAATATT
V/VI G-240(1)	TGCTGATGTTTTTTGAAACAAACAGAAGATGATTTTTCGGAACAAAGAAACCATAACATTA
V/VI G-240(2)	GGCTTAGGCATTTAACAAGGCGAATTATATCATCATATTCCTGTGCCCGAA
V/VI G-240(3)	CATAGGTCAATGGAAACAGCTTTGAATAATTCATCAATATAATACAAACAA
V/VI G-240(4)	CTGAGAAGGTGAATAACCATCGGGAGAATACTTCTGAATAATGATACATTT
V/VI G-240(5)	TTGAAAACTTAATTAATTTCAGATGAACCATATCAAAATTATTAACAACTA
VI/V B-240(1)	AGCGGAATTTCATTTCAATTACCTGAGCAAAATCAAGAAAACAAA
VI/V B-240(2)	ATGATGGCACCAAGTTACAAAATCGCGCAGATTTCATTTGAATTA
VI/V B-240(3)	TTTGGATTAACAATAACGGATTCGCCTGATTGTACATAAATCAAT
I_G-17,78_cF9	GAACCCTAGTTGTTCCAGTTTGGAGCCCTTCACTTCACGATTGCCACTTTCCAC
VI/V B-240(4)	AGAACCTATATACAGTAACAGTACCTTTTACTTGCTTCTGTAAAT
VI/V B-240(5)	AAACAGAATGCGTAGATTTTCAGGTTTAACGTTCCCTTT
VI/I G-240(1)	TGAGTAGATTTTTAAAGCGAAAGGGAAGGGAATTTTTTGGTGGTTCCGAAATCCGAAAATC
VI/I G-240(2)	TACTTCTTAAGTGTAGCGTGACGGGGAAATCAAAAGAATAGCCGCAAGCGG
VI/I G-240(3)	CTGTCCATCACCCGCCGCGGAACCCTAGTTGTTCCAGTTTGGAGCCCTTCA
VI/I G-240(4)	TTATAATCGCGTACTATG TTTGGGGTCAAAGAACGTGGACTCCCTTTTCAC
VI/I G-240(5)	GGAACGGTACGTGCTTTCCCACTACGTAAACCGTCTATCAGGGTTTTTCGGTTTGC
VI/I B-240(1)	CCCTTATAAAGCCGGCGAACGTGGCGAGAAAGAGCGGGCGCTAGG
VI/I B-240(2)	GGTTGAGTAAGGGAGCCCCCGATTTAGAGCTGTCACGCTGCGCGT

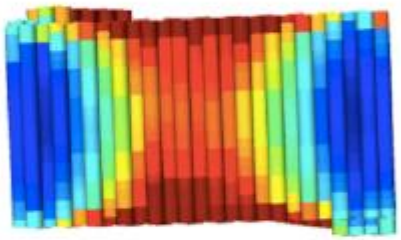
VI/I B-240(3)	CCACTATTGAGGTGCCGTAAAGCACTAAATCGCTTAATGCGCCGC
VI/I B-240(4)	AGGGCGAAGAACCATCACCCAAATCAAGTTTGTTGCTTTGACGAG
VI/I B-240(5)	TTCGATGGCCTCGTTAGAATCAG

Extended staples terminating with cA1 sequence

I_G-17,78_cF9	GAACCCTAGTTGTTCCAGTTTGGAGCCCTTCACTTCACGATTGCCACTTTCCAC
I_G-37,78_cF9	CCGCCTGGTTGCGCTCACTGCCCGCCGAGCTCCTTCACGATTGCCACTTTCCAC
I_G-57,78_cF9	GAATTCGTTATTACGCCAGCTGGCGGACGACGCTTCACGATTGCCACTTTCCAC
II_G-17,78_cF9	CAGGAAGACCGGAGAGGGTAGCTAGGATAAAACTTCACGATTGCCACTTTCCAC
II_G-37,78_cF9	ATTTTTAGTTAACATCCAATAAATACAGTTGACTTCACGATTGCCACTTTCCAC
II_G-57,78_cF9	TTCCCAATTTAGAGAGTACCTTTAAGGTCTTTCTTCACGATTGCCACTTTCCAC
III_G-17,78_cF9	AGATTTAGCTGACGAGAAACACCAAACTTTGACTTCACGATTGCCACTTTCCAC
III_G-37,78_cF9	AAGAGGACGAAAGAGGCAAAAGAACGTCACCCCTTCACGATTGCCACTTTCCAC
III_G-57,78_cF9	TCAGCAGCCTCCAAAAAAAAGGCTTAAAGTTTCTTCACGATTGCCACTTTCCAC
IV_G-17,78_cF9	TGAAAGTAGAATTTACCGTTCCAGGCCACCACCTTCACGATTGCCACTTTCCAC
IV_G-37,78_cF9	CCTCAGAGCTGTAGCGCGTTTTCAGACTTGAGCTTCACGATTGCCACTTTCCAC
IV_G-57,78_cF9	CCATTTGGGAATAAGTTTATTTTGGAAGGAAACTTCACGATTGCCACTTTCCAC
V_G-17,78_cF9	TCCCAATCCCTCCCGACTTGCGGGGGTATTAACTTCACGATTGCCACTTTCCAC
V_G-37,78_cF9	ACCAAGTAAACAACATGTTCAGCTACGCTCAACTTCACGATTGCCACTTTCCAC
V_G-57,78_cF9	CAGTAGGGCTTCTGACCTAAATTTATCAAAATCTTCACGATTGCCACTTTCCAC
VI_G-17,78_cF9	CTTTGAATAATTCATCAATATAATACAAACAACTTCACGATTGCCACTTTCCAC
VI_G-37,78_cF9	TTCGACAAAAATCTAAAGCATCACGTCTTTAACTTCACGATTGCCACTTTCCAC
VI_G-57,78_cF9	TGCGCGAAGGAAAAACGCTCATGGAAAAGAGTCTTCACGATTGCCACTTTCCAC

10.3.2 3prism – design scheme and staples





Full list of sequences for the 3prism DNA host

Start	End	Sequence
0[34]	6[31]	GCGCTGGTCAAGTTACCGTCTGCTTTCCTAAT
0[55]	7[52]	GGGAAAGTGCCGTACCCTTCATCGTGCCAAGTGTATCGA
0[76]	7[72]	GCGAACGAACCCTAGAGACGGATCGGCCAACATACTAG
1[42]	4[49]	GTCGAGGGAGCGGGAGGGCGCGTACTATGGTACGCCTGGCCCATCAAAA
1[63]	4[70]	AAATCGGTGGCGAGTTGACGAGCACGTAAAAGGGACAAGCGGAATCGGC
1[84]	4[91]	CCCCCGAGACGGGGCTTTCCTCGTTAGAGAGCTAACCCCAGCCCTGTTT
10[41]	8[42]	TCCGGCACAACTGTTGGGAAGGTCACGA
11[21]	8[21]	ACAAACGTGGTGCCGGAAACCCCATTCGCCATTCAGGCCAGT
11[63]	8[63]	TAGATGGCTCAGGAAGATCGCGGTGCGGGCCTCTTGGTAACG
11[84]	8[84]	TGCATCTGGGGACGACGACAGACGCCAGCTGGCGATGCTGCA
12[30]	18[31]	GAGGTAAACGAAAAGCCGTAGCTAATATGATGAAGCCTTCGG
12[51]	17[52]	GCTAAATTCGTGTACCCACAAAGGGAGACAGGGAT
12[73]	17[72]	GCGTTAAATCAAAAACTACCTGAGATCAAAAGTCA
14[20]	11[20]	CAGGAAGCAAATATTTAAATTCGAGTAACAACCCGCGTGGGA
14[62]	11[62]	AATCATACATTAAATTTTTGTCTGGCCTTCCTGTAGTTGGTG
14[83]	11[83]	TAATCGTGCTCATTTTTTAACGCCATCAAAAATAAGTAACCG
15[42]	13[41]	GAGATCTCGGTTGATAATCAGTTAATAT
17[21]	14[21]	TGCGGGAATTCAACCGTTCTATAATGCCGGAGAGGCCAAAAA
17[53]	22[52]	AAAGCCTCAGGTGGCATGTAGATTAGCTCAAGGAT
17[63]	14[63]	AGAACCCGGTGAGAAAGGCCGCTATCAGGTCATTGGCATGTC
17[73]	22[74]	TATATTAGCAAGTAGTAGGATTCCCGCAACTAGAA
17[84]	14[84]	TGCAATGGTGTAGGTAAAGATGTCTGGAGCAAACATGAACGG
18[30]	24[31]	TTGTTTGGGGGATACATAATTGCTTGCTCCTAAGCGGAAAAT
18[41]	15[41]	AGCTAAATTATTTCAACGCAATCAAATCACCATCATTTTTGA
20[20]	17[20]	ATGGTCATAGCTATATTTTCATACCAAAAACATTAATACTTT
20[62]	17[62]	CGAACGACAATTCTACTAATAAATTAAGCAATAAAAATTTTT
20[83]	17[83]	AACAGTTCATTAACATCCAATGGCAAGGCAAAGAATTTTAAA
22[51]	26[49]	TAGTTAAGAGGCATAGTTATTACAATGCGAT
22[73]	26[70]	GCAATTCAAATAACGCCATCAGTTGTTAATCA
23[21]	20[21]	AGAAGCATTTGATAAGAGGTCATGGCTTAGAGCTTTTCGCAA
23[42]	21[41]	AAAAAGAAGAGTACCTTTAATGAATATA
23[63]	20[63]	GAAAGACACTCCAACAGGTCACATGTTTTAAATATAATTCTG
23[84]	20[84]	TTTAATTAGCGAACCAGACCGAAGTACGGTGTCTGTCCATAT
24[30]	28[24]	CAAAACGAACCATTATACCAGTCAAAATCAAGACA
26[48]	33[45]	TTTAAGATCATTCAACCTTCAGTTACTTAGCGATTCATG
26[69]	33[66]	TTGTGAACCCTGACACCAGGCGCAGACGAAACACTCGGG
26[90]	33[87]	GTTTAATCGAGTAGGGACAGAAACCGAAAAGAGGCCTAC
27[14]	23[20]	ATTACCCGGACGTTATCTACGTTAATAAAAATCAGGTCTTTATATAGTC
27[35]	30[42]	AAGCTGCACTGGCTTAACGGAACAACATAAGAGAAAAAACGATAAAATG
27[56]	30[63]	AGGCTTGTTACCTTGGTAGAAAGATTCAAAAGGAAACCCTCGTTTTGCC
27[77]	30[84]	ACCAGAATTCAACTAGATTTAGGAATACTAATGCAATAAAAAGAGAGGC
28[23]	36[24]	AGAGAAATCCACAAAGTTTTGAGGAGCGAAAAGCTTGCCAAA

29[21]	29[13]	CAAATGCCGTCCAATACTGCGGAATCGTCATTGAA
30[41]	27[34]	TTTAGACTGGATAGTTTAAACTAATCTTCGTAACA
30[62]	27[55]	AGAGGGGGTAATAGGCAACACCTGGCTGGTGAATA
30[83]	27[76]	TTTTGCAAAAGAAGTTTACCATGTACAGGAGAAAC
31[35]	23[41]	GCTCCATTCAAGAGAGTTCAGTGACCATTTGCATC
31[56]	23[62]	ACGAGGCGCATAGGTATCATATTACGAGGAAGCCC
31[77]	23[83]	ATAAGGGTGAACGGGACGACGGATACATATCGCGT
32[13]	29[20]	AGATTTGTGATAAATTGTGTCACCGGATTCCCCCT
33[46]	38[45]	AGGTGCGGGATCTTAAATTTTTTCTCCAGACTTTC
33[67]	38[66]	TAACAGGGAGTAGTTGCTAAAGGATTTCTGTCCCA
33[88]	39[87]	GAAATATTCGCAACCATGTGAGAAAACAACTGGATAGCTCAG
34[34]	31[34]	CCTCAGCACTAAAGACTTTTTATACCAAGCGCGAAGCGACCT
34[55]	31[55]	CCGCTTTAAGTTTCCATTAAACATCTTTGACCCCCAGCCGGA
34[76]	31[76]	AGGCTTGAATACGTAATGCCAAAAAGAATACACTAGTCAATC
35[14]	32[14]	GTTTATCGACAGCATCGGAACACGGCTACAGAGGCACAACGG
36[23]	42[24]	AAAGATCTAACAACGCCTCACCGTGTGCCGTAGTGCCTACTG
36[55]	34[56]	ATAATAACAGCTTGATACCGATTAAAGG
37[35]	34[35]	TCGTCTTACGTTGAAAATCTCTTTCGAGGTGAATTTCGTCAC
37[77]	34[77]	TTTGCTATAGAAAGGAACAACGCCGACAATGACAAGTCGCTG
38[13]	35[13]	TTCCACAATAGTTAGCGTAACAAGGCTCCAAAAGGTGTATCG
38[44]	43[45]	GTCGTACCGCTGCTCAGCGTATAACTTTTGAGGCA
38[65]	43[66]	TGTCCACCCTGGATTAGCTGCCTATCCAGTATGAT
39[88]	45[87]	AGCGGCTGAGATTCTGACGCAGTCTCCTCATGCCTCCCCGGA
4[20]	0[14]	TTCCAGTAGTCCACCGATGGCAACCATCACCCAAACAAGTGTAGCGGTC
4[48]	1[41]	GAATAGCCCGAGATGACTCCAGCGAAAATTTTGGG
4[69]	1[62]	AAAATCCCTTATAATGAGAGACTGATTGAAGCACT
4[90]	1[83]	GATGGTGGTTCCGATCCACGCCACCAGTAAGGGAG
40[34]	37[34]	CGGATAAACTCAGGAGGTTTAACCAGTACAAACTAAGTTTTG
40[76]	37[76]	AAGAGAACAGAACCGCCACCCAAGCCCAATAGGAAATGGGAT
41[14]	38[14]	CGGGGTCCGAGAGGGTTGATACCGGAATAGGTGTATGTAGCA
41[56]	39[55]	ATGCCCCGATTAGCGGGGTTTCACCCTC
42[23]	47[24]	GTACCAGCATACCAGAACATCGGCTTAGCGTTGAA
43[35]	40[35]	AGGTTGATGATACAGGAGTGTTGAGTAACAGTGCCTACCAGG
43[46]	48[45]	GGTGCCACCACCCCCTTGCGACAGTTGAGCCATTG
43[67]	48[66]	ATTCTCAGAATCTTTTCATAGCAGGCCAGCACAAA
43[77]	40[77]	AAATAAATCTGAATTTACCGTTTTCGGAACCTATTACTCCTC
44[55]	41[55]	CCTCAGACAGACGATTGGCCTAGCGTCATACATGGACAGTTA
45[88]	50[87]	ACCACGTCACACCATTAATGGTTTTTGTCACGACG
46[34]	43[34]	TTTGCCTATTTTCGGTCATAGCCCTCAGAGCCGCCTGACAGG
46[76]	43[76]	ACCATCGATAATCAAAATCACTCAGAGCCGCCACCCACAAAC
47[25]	52[21]	TTAATTGACGAATACATACGCAGTCAAATAAAACAGCC
48[44]	54[38]	AGGACATATAGCATGATTTTAACGTAATTTGCCAGTTATTAGCGAACCG
48[65]	52[63]	AGGACACCACCCCTGAAAGCAGCCAACGAGC
49[35]	46[35]	GGTGGCAGAGGGAAGGTAAATTCACCGTCACCGACAATCAAG
49[56]	47[55]	CGCAAAGGCGACATTCAACCGATTTGGG
49[77]	46[77]	GTTTATTACCAGCGCCAAAGAAAATCACCAGTAGCCAATGAA

5[21]	75[27]	ACTGCCCATCAGGGTATTAAATATAATCCACGCAA
5[42]	75[48]	AAACCTGCCAAAGGACGTCGCCAGAATCCAATACT
5[63]	75[69]	TTAATGAGCAACAGGTTGCAGTTTTAGATAACATC
5[84]	75[90]	GGGGAGATTCTTTTTGGTTTGACAGGAGAAGAACT
50[86]	54[80]	GGAAACATAAATCCTGAATCTTACTCAAGATGAGA
52[20]	59[17]	ATATTATATTCTAAATCAGATAAAAATACAACAATTAGG
52[62]	58[59]	GTCTTTCGGGAGGTATTTTCAATCAATAAATA
53[28]	56[35]	AGGCGTTCAAAATAGAAACGATTTTTTGTAAGACTGAATACCTTAAGAA
53[49]	56[56]	GACTTGCCAGAGCCTCAAAAATGAAAATCAAAGTCTGAGCGCGCTATCT
53[70]	56[77]	CCTTAAACAACGCTTTTACAGAGAGAATGAATTAATAACCCAAGAAACA
54[37]	61[38]	CGCTTTACGAATGCAGACATATTTTATGCGTTATA
54[79]	62[80]	ACACTTATCATAATTCTTCGAGCCACACCGGTTGAAATCTGA
56[34]	53[27]	AAGTAAGCAGATAGACGCAATCAAGCAAGAACGCG
56[55]	53[48]	TACCGAAGCCCTTTCAAAAGAAATCATTACCTCCC
56[76]	53[69]	ATGAAATAGCAATATAATATCCGTTTTTTTTGAAG
57[28]	49[34]	ATCCTAACCAATAGAATAACGCCTTATTACATAAA
57[49]	49[55]	GAAACCATCGTAGGACGTAATAGAGGTGAAAGAAA
57[70]	49[76]	GTCTTTCAGCAAGCAGAGAGACTGAACAGGAATAA
57[91]	52[91]	AACGGGTGTACCGCACTCATCTAGTTGCTATTTTGCAATTTT
58[58]	63[59]	AACGTAATTTGAAAAAGTTCTGACCTCCGGCAACA
59[18]	66[17]	GCTCCAGTATACGCGAGTGCAAATTAAATCGATGTGAGGCCTGATCATC
6[30]	12[31]	GAGGAGGATCAAACGACGGCTGCGCCGCTTCGCGGATTATGT
60[27]	57[27]	ATTCTTATAATTGAGAATCGCACGCGCCTGTTTATATATCCC
60[48]	57[48]	AGTATCAAACAACGCCAACATAACATGTTCAGCTAGCATGTA
60[69]	57[69]	ATTACTAAGGCAGAGGCATTTGTCCAGACGACGACATCGGCT
60[90]	57[90]	AGAATAAAGTAATAAGAGAATGACAAAAGGTAAAGTTCCAAG
61[39]	67[38]	TTTATAACTATTTTCCCAAACAGTAGTTACAATACAGTTATA
62[48]	60[49]	GGGTTATTAGTTAATTTCATCCCTGTTT
62[79]	68[80]	GAGAAGACGCAATTACACTGAGCATAAAGAAGGTTAGATTAA
63[28]	60[28]	TAATTAATATGTAAATGCTGAAAAACTTTTTCAAATATACAA
63[60]	68[59]	TAGTTGAATTCGAATTACAGGTTTTATACTTACAT
63[70]	60[70]	TTAGATTACTACCTTTTTAACCTAAATTTAATGGTAATCATA
63[91]	60[91]	AGTCAATTCAAAATCATAGGTACCGACCGTGTGATTTAAATA
66[16]	72[17]	GGGCAGATGAGAATTATACAACTAATCTAAATAAAACATACC
66[27]	63[27]	CCTTTTATGCTTTGAATACCAACATAAATCAATATTCGCTAT
66[69]	63[69]	AGATTTTTTCATTTCAATTACTTTAACAATTTCATCGATAGC
66[90]	63[90]	ACAGAAAAAAGAAGATGATGAAAGAAAACAAAATTTGAGAAG
67[39]	73[38]	ATCAAGAAACTCAATAGTGAAAGGAGTATTACTAAAACCACG
67[49]	65[48]	TTTGGATAACGTCAGATGAATAAATCGC
68[58]	73[59]	TATGATTTAGGTCAGTTAGTGCCATTAATGCCATT
68[79]	73[80]	TTTACAAACACAAACCCAGCAAATTTTTTGAAACC
69[28]	66[28]	AGAGCCGCACCAGAAGGAGCGTGGCAATTCATCAAAACAGTA
69[70]	66[70]	AGACTTTTAAAAGTTTGAGTACTGAATAATGGAAGATTGCGT
69[91]	66[91]	AACTCGTTTGCCCGAACGTTAACCTACCATATCAAACGTAAA
7[53]	12[52]	ATTTTTCCCAGGCGATCACTCCAGAGGTCACGCCA
7[73]	12[74]	CTGTTAAGTTGCGCTATTTATCGGCGCGCATCTTC

70[48]	67[48]	CAACAGTATAATACATTTGAGCATTTTGCGGAACACTGATTG
72[16]	77[27]	GAACATTGGCTCAATCGTGTCCATAGTGAGGACACCCGCCGCGCT
72[27]	69[27]	TTAAAAAGAGGTGAGGCGGTCAATTGAGGAAGGTTATAGATT
72[69]	69[69]	TTAGTCTCGCTGAGAGCCAGCTCAATCAATATCTGAAGTATT
72[90]	69[90]	GACAATAGAAAAATCTAAAGCTGAACCTCAAATATATTCGAC
73[39]	0[35]	ACCGAAATACGTTGTAGCTGAGAACCGCTACCGCTAGG
73[60]	0[56]	CTGGCAACAGTTAGTAACAGGAACGGTTGCTAAAGGAA
73[81]	0[77]	CTTACAATATTGAGTAGGCCGATTTAACGTGAAAGCCG
75[28]	72[28]	ATTAACCCTACATTTTGACGCAGATTCACCAGTCAATCGCCA
75[49]	73[48]	TCTTTGAGAAAAACGCTCATGAGTAATA
75[70]	72[70]	ACTTGCCTACCGCCAGCCATTGCCAACAGAGATAGATGGCTA
75[91]	72[91]	CAAACTATGGTAATATCCAGACTGACCTGAAAGCGTGGCACA
77[28]	4[21]	TAATGCGGTGTTTTGAACGTGAGGGTTGAGTGTTG
8[20]	5[20]	GCCAAGCGCAGGTCGACTCTATGAGCTAACTCACATGCGCTC
8[41]	5[41]	CGTTGTACCCGGGTACCGAGCAAGCCTGGGGTGCCAGTCGGG
8[62]	5[62]	CCAGGGTCGTAATCATGGTCAGAGCCGGAAGCATAAGCTGCA
8[83]	5[83]	AGGCGATTTCCTGTGTGAAATTCACAATTCCACACAACGCGC

Extended staples terminating with cA1 sequence

Start	End	Sequence
13[42]	81[57]	TTTGTTATTCATCAACATTAAGACCGTAATGGGATCCAGCTTCTTCACGATTGCCACTTTCCAC
21[42]	79[57]	ATGCTGTTAGTTTGACCATTACGCGAGCTGAAAAGAGCATAACTTCACGATTGCCACTTTCCAC
39[56]	89[71]	AGAACCGACCGTAACACTGAGGTTAGTAAATGAATATTGCGACTTCACGATTGCCACTTTCCAC
47[56]	87[71]	AATTAGACACCGTAATCAGTAATTAGCGTTTGCCACCGCCACCTTCACGATTGCCACTTTCCAC
65[49]	85[64]	GCAGAGGACCTTTTTTAATGGTTAGAATCCTTGAATTAGGTTCTTCACGATTGCCACTTTCCAC
73[49]	83[64]	AAAGGGAGCGAACTGATAGCCACACCGCCTGCAACGGCAAATCTTCACGATTGCCACTTTCCAC

10.3.3 Complete gel from chapter 3.4.3.4

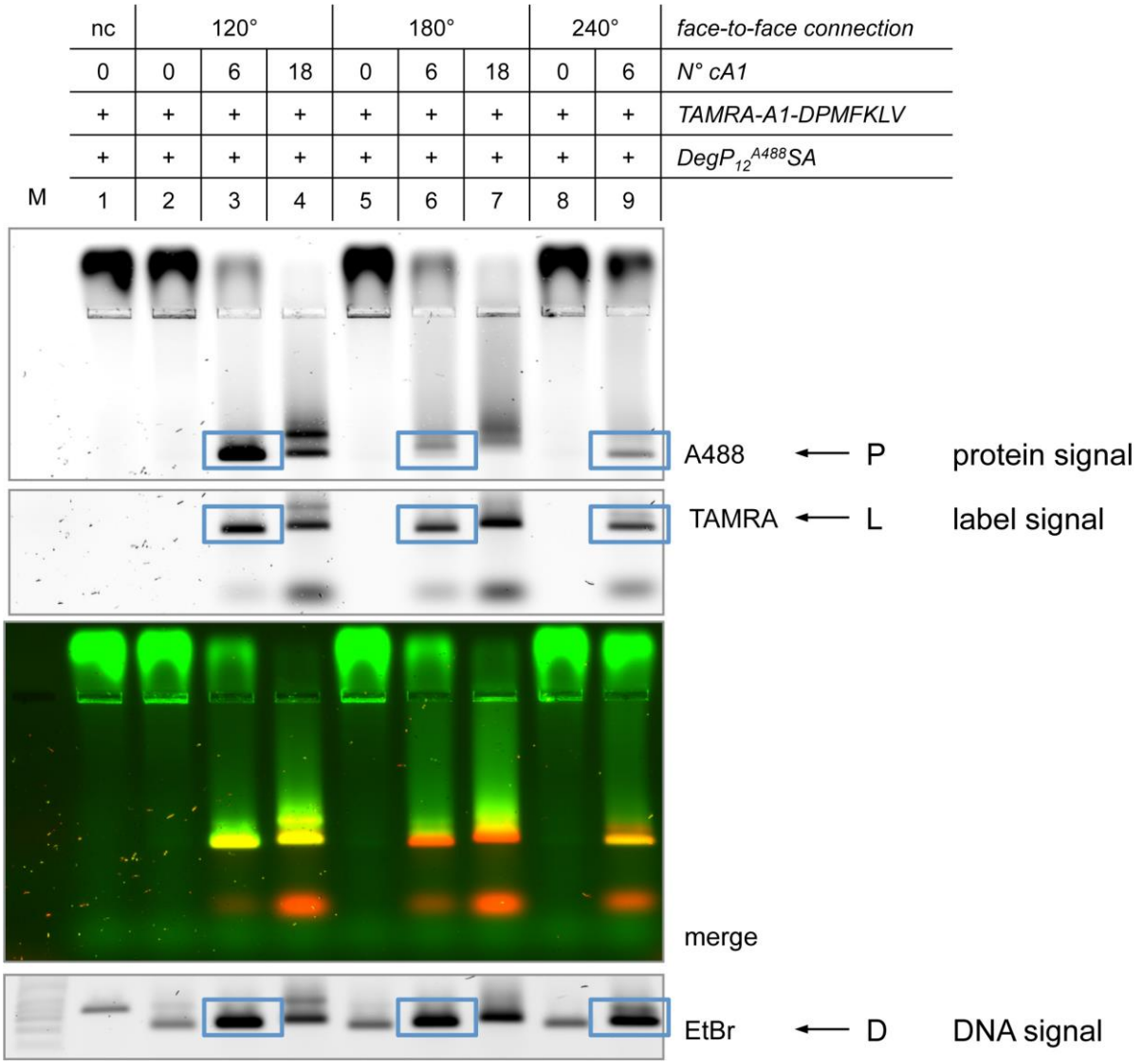


***Figure 80*:** Full view of the gel given in chapter 3.4.3.4. Lane M contained a 1 kbp DNA ladder. The origami (lane 1) migrates between the 1.5 and 2.0 kbp bands of the ladder.

10.4 Curriculum Vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

10.5 Danksagung

Die Danksagung ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

10.6 Erklärung

Hiermit erkläre ich, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient, bei der Abfassung der Dissertation nur die angegeben Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Essen, den _____ _____

Unterschrift des/r Doktoranden/in

Hiermit erkläre ich, gem. § 7 Abs. (2) e) + g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den _____ _____

Unterschrift des Doktoranden

Hiermit erkläre ich, gem. § 6 Abs. (2) g) der Promotionsordnung der Fakultät für Biologie zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions“ zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Andreas Sprengel befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Essen, den _____ _____

Unterschrift eines Mitglieds der Universität Duisburg-Essen