Subproject C5.4

Protein assisted assembly of superlattices

Principle Investigator: Holger Puchta

CFN-Financed Scientists:
Further Scientists: Manfred Focke, Daniela Kobbe

Botanisches Institut II
KIT, Campus Süd

Protein assisted assembly of superlattices

Introduction and Summary

The genetic information of all organisms is encoded by DNA. Beyond this biological function, DNA has become an important material for nanotechnology [1]. Due to the fact that complementary DNA sequences can specifically anneal and will form stable double helices, specifically designed oligonucleotides can self-assemble and will form highly ordered, larger structures (bottom-up approach). The process of self-assembly and disassembly can either be influenced by physico-chemical parameters like pH, temperature, salt [2] or by the use of DNA processing enzymes.

Especially the use of proteins acting on DNA offers additional values for nanotechnology in terms of specificity and possibilities to modify either the DNA molecule or the resulting nanostructure. The deoxyribose-phosphate backbone of DNA can be cleaved specifically by endonucleases. The base pairing of the DNA can either be reversed by melting the hydrogen bonds between the bases by helicases or accelerated – surprisingly some helicases can catalyse this latter reaction as well [C5.4:1, C5.4:2]. Imperfect areas in the assembled nanostructure which may occur especially in an annealing procedure which only relies on physico-chemical processes might be repaired by strand exchange proteins like RAD51 or RecA. Deliberate base pair mismatches introduced into the DNA nanostructure for technical reasons, might be repaired by mismatch repair proteins. Furthermore the increasing knowledge of structural and enzymatic proteins involved in the parallel orientation of homologous DNA helices, a structure which is essential for the biological process of meiosis, might deliver new tools for DNA nanotechnology [C5.4:3].

The construction and assembly of three dimensional (3D) DNA crystals is a special challenge. Although one recent publication reported a success [3], the effort (estimated 20 years of a scientist working capacity) to obtain such a nanostructure demonstrates that further approaches are needed urgently. The general problem is that a 3D-structure needs branching in all three axes to guarantee an equal growth of the crystal in all three dimensions. However, the normal DNA is an unbranched molecule. Although nature offers some exceptions from this general rule (replication forks, Holliday junctions) a further problem still exists: natural branched DNA structures exhibit an unwanted flexibility such as conformeric transitions from unstagged to different stagged structures in the case of Holliday junctions. Although those structures can be influenced by the magnesium ion concentration, a regular crystal should be favored by fixed angles of the branches. One approach to facilitate defined branching is the use of hybrid molecules of DNA oligonucleotides and an organic or anorganic center. This allows to combine the advantages of DNA (addressable self-assembly by complementary sequences) with the extended possibilities for the geometry of the center part of the hybrid molecule. Nano-gold particles as anorganic centers resulted only in very small particles with a quasi-crystalline structure [4, 5]. Another approach is to use tetrameric streptavidin with biotinylated oligonucleotides to assemble high molecular weight DNA nanostructures. However, as reported recently, the preparation of larger constructs failed until a flexible linker was incorporated into the experimental design [6] such that the question arises of whether clearly defined crystals will be obtained with this approach.

The approach of the CFN C5 team was to use a symmetrical organic core molecule with covalently attached oligonucleotide arms. The other team members are the working groups of Wolfgang Wenzel (C5.1 theoretical calculation and modeling of the expected supra-molecular structures), Stefan Bräse (C5.2, synthesis of the organic cores) and Clemens Richert (C5.3, attachment of the oligonucleotide arms to the organic cores and crystallization trials).

The C5 team is pursuing two different sub-strategies for the assembly of ordered 3D DNA structures. On the one hand the team tries to focus on physico-chemical control of the assembly process using short (self-complementary) oligonucleotide arms (C5.3 Richert group). On the other hand the Puchta group wants to exploit the properties of proteins acting on and with DNA at different stages of the process:
First of all we want to be able to obtain different building blocks with the same basic geometrical structure determined by the organic core with the help of commercially available proteins and commercially available oligonucleotides. The ability to access the resulting diversity opens up different assembly strategies as detailed below.
Secondly, we want to use commercially available proteins with known specificities to challenge the DNA structures. This allows to conclude e.g. about formed double stranded regions.
Additionally, as mentioned above, proteins have the potential to channel and control reactions in a specific manner. We propose to control the assembly of 3D DNA structures by a network of DNA acting proteins in the future. Unfortunately most of the well characterized DNA helicases, DNA binding or DNA strand exchange proteins are either not commercially available, not enzymatically pure or can not be obtained in large quantities needed to set up complicated systems. Therefore we started to increase the pool of potentially useful DNA binding and processing enzymes within the framework of this project focusing on Arabidopsis proteins. The model plant Arabidopsis is known for its large protein families. The proteins within a family are thought to catalyse different reactions, which could already be shown [C5.4:2]. For a precise control of the assembly process it might therefore be useful to have similar but different proteins available.

Starting point for the generation of our different building blocks was a tetraphenylmethane-core (TPM) to which four oligonucleotide arms were attached in a manner that 3'-ends of the oligonucleotides were linked to phenyl-moiety of the TPM, whereas the 5'-end is free and extruding (figure 1). The DNA sequence of all four arms is identical, due to the synthesis strategy. Unfortunately the number of elongation steps is limited due to decreasing yield of completely elongated product in each step. We could show that we need a minimum length of six nucleotides to elongate the oligonucleotide arms enzymatically. To prevent an uncontrolled, unwanted early-annealing of the six-nucleotide long TPM-arms with each other the sequence of the arms was not self-complementary. The resulting synthesis product was $[5'\text{-GTCCAG-}3']_4$ –TPM and is called TPM-hybrid. This special TPM-hybrid was used for all experiments described in the following chapters.

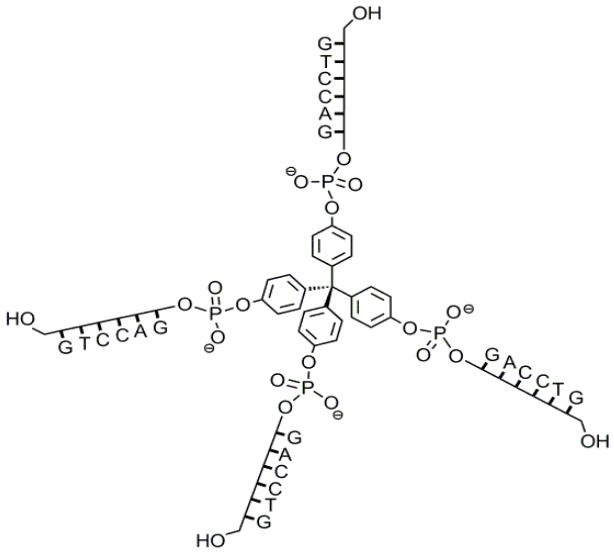

Fig. 1: Structure of the TPM-hybrid used in this project

1. Elongation of TPM hybrids with linear oligonucleotides

As our 3D structure is proposed to be based on DNA base pairing, complementary sequences which can anneal to each other are required. As described we opted for a non-selfcomplementary TPM-hybrid as a starting point. Beside the problem, that self-complementary sequences can form secondary structures (which might be a minor problem if the sequences are very short such as the six nucleotide long arms of the TPM-hybrid), the kinetic control of the annealing process is difficult in a solution of self-complementary molecules: if the annealing process is running to fast it might happen that not all four arms of the TPM hybrid find an annealing partner resulting in "cavities" in the crystal and possibly in an irregular shape of the crystal (asymmetric growth). This problem can be overcome with different species of molecules that are not self-complementary but at least partially complementary to each other. Those different species of molecules were generated by elongating the TPM-hybrid.
This elongation leads to a second advantage: this way it is possible to vary the distance homogenously between the TPM-cores by changing the duplex length of the DNA. Such flexibility in the distance would broaden the application options for the three-dimensional DNA crystals e.g. as scaffold for DNA binding proteins, sieving material for macromolecules or as a carrier for nanogold particles in nano-electronical techniques. Additionally for the aspect of controlled assembly by DNA processing enzymes a minimal length of the DNA is a requisite.
To circumvent the limitations of the chemical elongation steps described above we chose an enzymatic approach to elongate the TPM-hybrids with specifically designed, commercially available, oligonucleotides. In a multi-step process the 5' end of TPM-hybrid is phosphorylated by polynucleotide kinase (PNK) and after purification the phosphorylated TPM-hybrid is annealed with a splint-oligonucleotide (splint-oligo) spanning the sequence of the TPM-hybrid arm (6 nucleotides) and a part (8 nucleotides at the 3' end) of the oligonucleotide used for elongation (= linker oligo). The splint-oligo positions the 5' phosphate end of the TPM-hybrid oligonucleotide and to 3' end of the elongation/linker oligonucleotide optimally, enabling the enzyme T4 DNA ligase to covalently close the nick between the two oligonucleotides. Each step was optimized to obtain the highest possible yield. For example an excess of PNK will stick to TPM-hybrid and will inhibit the downstream steps. The ratio of splint- to linker oligonucleotide was changed with special emphasis on the identification ligation products. With increasing amount of the splint-oligo, the signals in the high molecular area of the gel increased while the signals in the low molecular area of the gel decrease. Figure 2 demonstrates, how the latter in combination with phenol-chloroform extraction and denaturing gel electrophoreses led to the assignment of completely (4 arms) and partially (1, 2 or 3 arms) elongated TPM-hybrids.

Cutting the native double stranded region of the elongated TPM/splint-oligo with the restriction endonuclease *Pst I* resulted in TPM molecules with shortened arms indicating that a duplex region inside the complex exists that can be recognized specifically by enzymes (data not shown).

The purification of the completely elongated TPM-hybrid was successfully optimized (see e.g. figure 4). Separation from by-products and splint-oligo is achieved by preparative denaturing gel electrophoreses followed by elution by diffusion.

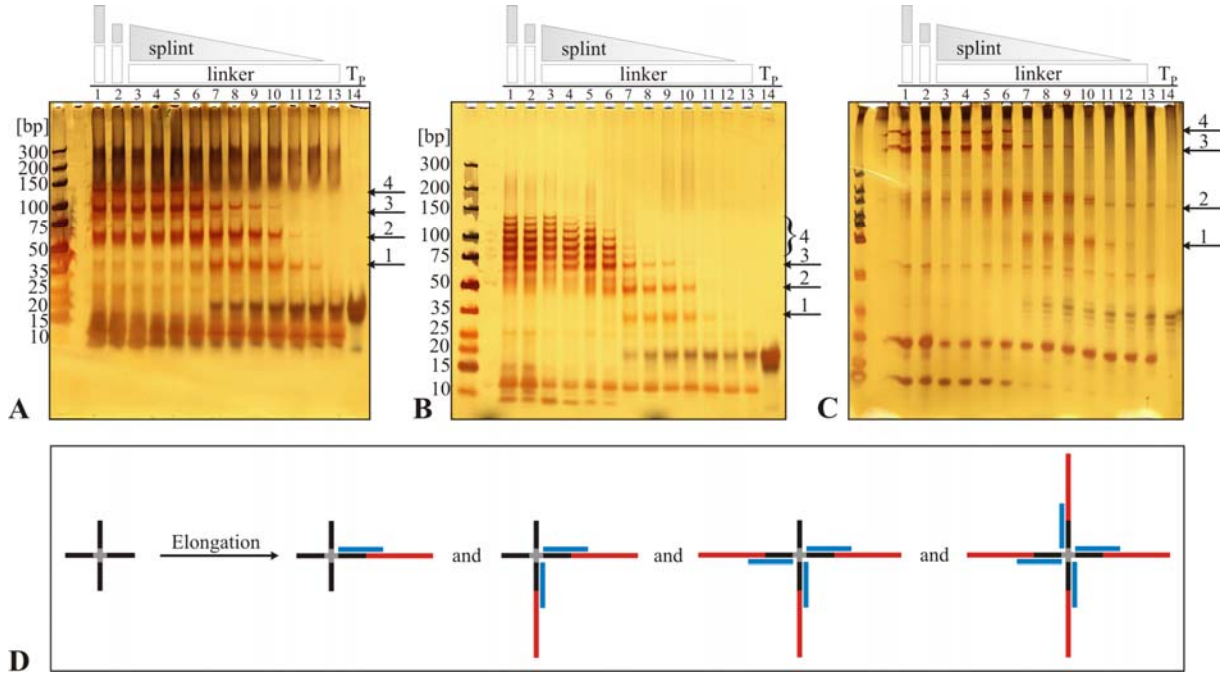

Fig. 2: Gel analysis to identify the by-products after elongation of the TPM-hybrid.
In each reaction 10 pmol phosphorylated $(GTCCAG)_4$-TPM were used for the elongation reaction. Lane 1) reaction with 80 pmol linker and 80 pmol splint, respectively, lane 2) with 80 pmol linker and 40 pmol splint, lane 3 – 13) 40 pmol linker, respectively, and varying amounts of splint: lane 3) 80 pmol, 4) 40 pmol, 5) 30 pmol, 6) 20 pmol, 7) 10 pmol, 8) 8 pmol, 9) 6 pmol, 10) 4 pmol, 11) 1 pmol, 12) 0,5 pmol und 13) 0 pmol splint. Lane 14) 5 pmol phosphorylated TPM-hybrid (T_P). **(A)** Elongated TPM-hybrid, native 5-12% TBE-PAGE. **(B)** Elongated and protein free TPM-hybrid, native 5-12% TBE-PAGE. **(C)** Elongated and protein free TPM-hybrid, denaturing 20% TBE-urea-PAGE. Because of the denaturing conditions, the marker is not labelled. The numbers on the arrows indicate the number of elongated ssDNA molecules of the TPM-hybrid. Silver staining. **(D)** Simplified schematic illustration of possible by-products after elongation. The splints are removed under denaturing conditions (panel C), whereas secondary structures as well as hybridized and replaced splints are visible under native conditions (panel A, B).

2. Annealing of elongated TPM-hybrids

As discussed above it is preferable to have the complementary DNA sequences on different TPM molecules. We decided therefore to elongate the TPM hybrids with two different oligonucleotides (30 nucleotides length) in separate reactions. The resulting elongated TPM hybrids were separated and purified as described in chapter 1 and were named TPM A and TPM B. Because the two oligonucleotides used are complementary to each other, TPM A will anneal with TPM B and *vice versa*. The strategy is to add an excess of the annealing partner that the TPM which is in shortfall should be annealed at all four arms. Therefore a perfect shell would be obtained avoiding the problem of cavities and irregular shape as discussed in chapter one. A schematic representation of this strategy is shown in figure 3.

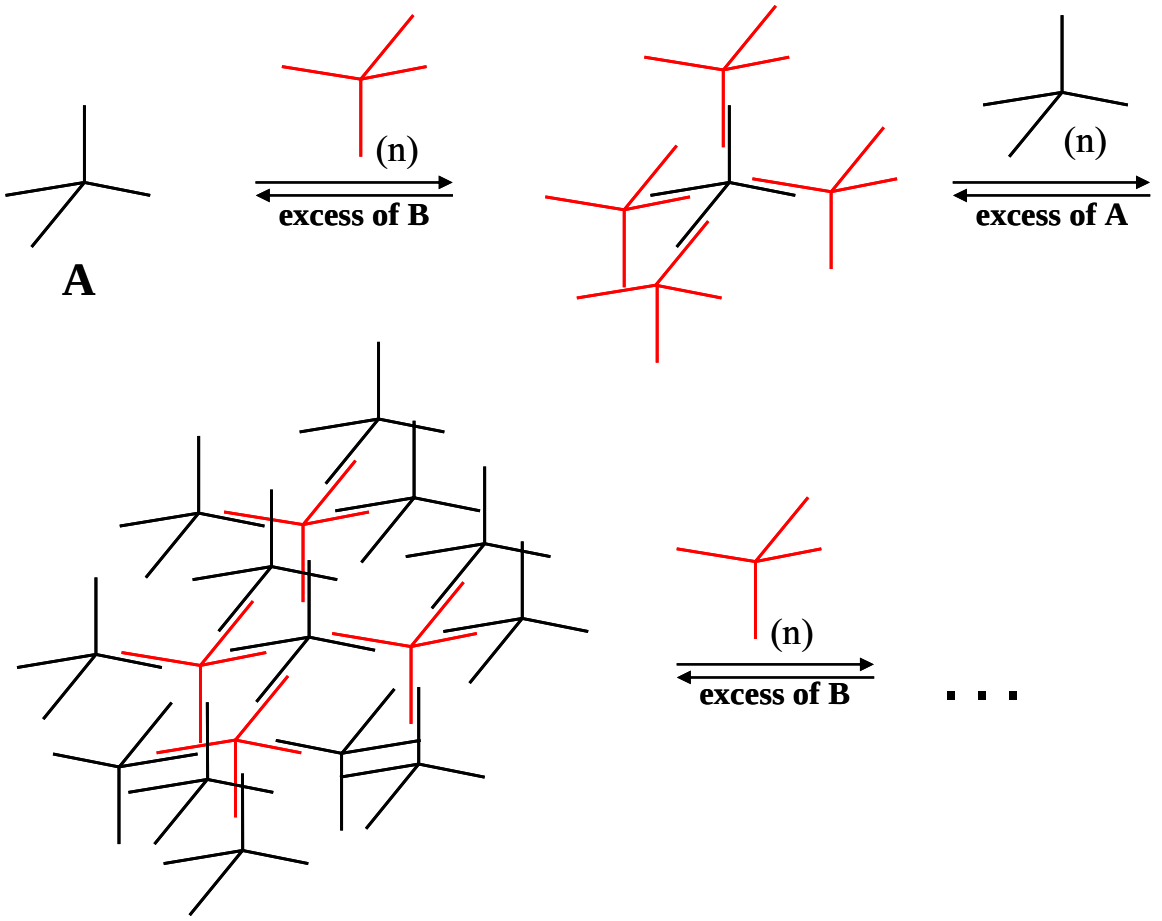

Fig. 3: Scheme of the strategy to obtain perfect shells. After each annealing step the perfect shell has to be separated from structures in which sub-stoichiometric annealing occurred (see chapter 3)

We tested several annealing conditions to optimize the yield of perfect shells. The molar ratio of TPM A to TPM B was typically 1:5. Beside temperature programs with a thermocycler, a combination of volume exclusion by freezing and crowding with polyethylene glycol [7] and different chemicals like magnesium chloride, lithium chloride or tetraethylammonium chloride (TEAC) were tested on their influence on the yield of annealing products. The chemicals mentioned above are used for the acceleration of the hybridization [8], because these chemicals reduce the electric repulsion of DNA strands. It turned out that an incubation in 2,4 M TEAC at 42 °C favors the occurrence of higher molecular annealing products at the expense of lower molecular weight signals. It was estimated that compared to a single elongated TPM molecule a perfect AB_4-shell (one elongated TPM A annealed with four elongated TPM B molecules, see figure 3) should exhibit a twofold larger diameter. As can be seen in figure 4 beside a putative AB_4-shell higher molecular weight signals can be detected in the gel. This implicates that the putative AB_4-shell has to be purified as will be reported in chapter 3.

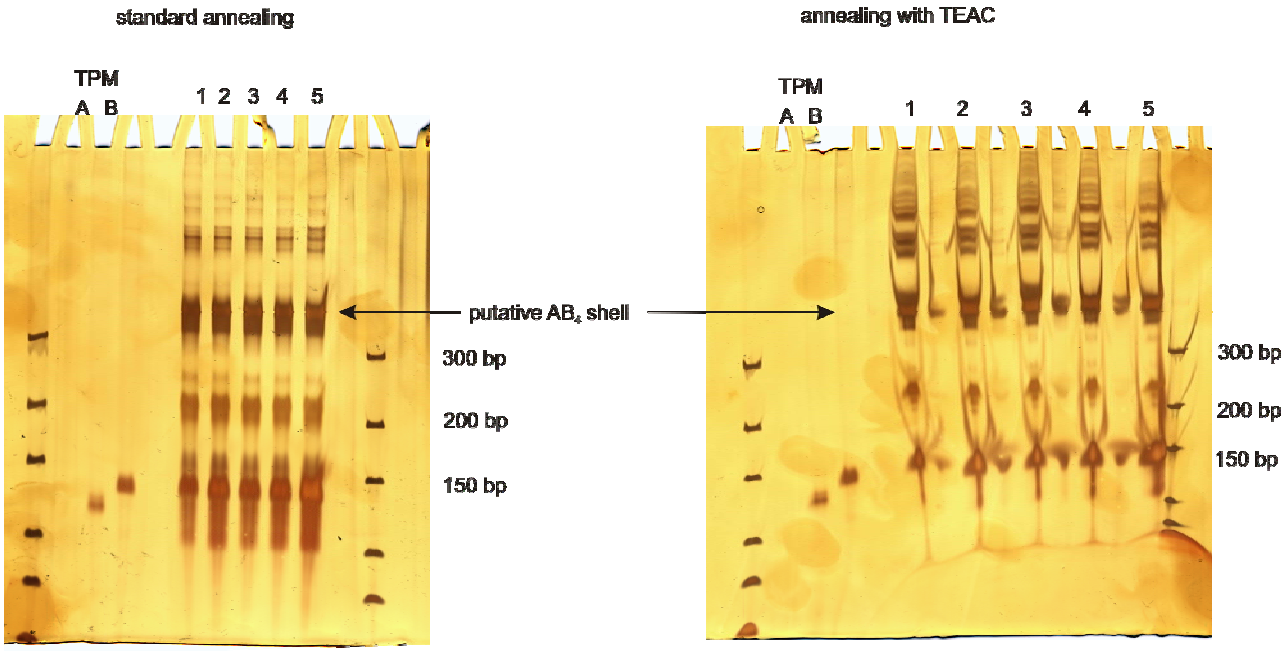

Fig. 4: Gel analysis to compare different annealing conditions
5% TBE silver stained polyacrylamide gels. TPM A/B: elongated hybrid TPM A or B loaded as a reference, due to secondary structures their migration pattern are slightly different. Standard annealing in the presence of 1 mM magnesium chloride, 95 °C 10 min followed by 95 °C → 5 °C in dT= - 0,1 °C/min, TEAC annealing 1h at 42 °C with 2,4 M TEAC (the TEAC annealing was used later on as routine procedure) 1: 10 pmol TPM A + 10 pmol TPM B, 2: 10 pmol + 20 pmol TPM B, 3: 10 pmol TPM A + 30 pmol TPM b, 4: 10 pmol TPM a + 40 pmol TPM B = stoichiometric ratio for complete annealing of all TPM A arms, 5: 10 pmol TPM A + 100 pmol TPM B, it should be noticed that a quantitative analysis of silver stained gels are difficult due to limited linear dynamic range of silver staining. In each of the marker lanes on the left sides of the gels the marker (GeneRuler ™ ultra low range DNA ladder) was loaded with exact the same buffer as the applied samples

In the future we will test, if the annealing process can be influenced enzymatically. For this purpose different helicases (e.g. SRS2, RECQ 3) have been cloned, expressed, purified and characterized. As these enzymes are able to both melt DNA helices and anneal single stranded complementary DNA strands [C5.4:1, C5.4:2], but to different extents, and possibly via different mechanisms, it will be interesting to compare their behavior and those of further helicases and choose the optimal enzyme for this application.

3. Purification of annealing products

As described in chapter 2 after the annealing reaction itself, the perfect shells will be contaminated with non-perfect structures. For our approach it is important only to use perfect shells, because otherwise the DNA crystal growth might be disturbed and become irregular. We tested several purification methods, which in contrast to the purification of the elongated TPM hybrids has to be under non-denaturing conditions. Two main parameters had to be optimized. First, the in-gel detection of the annealing products has to be sensitive and at the same moment gentle to the analyte. For ethidium bromide detection UV illumination of the gel is needed and this might induce DNA damage such as thymine dimers. Additionally the ethidium bromide method is not as sensitive as

other methods. Silver staining is much more sensitive but the applied formaldehyde might induce crosslinks. We developed a home-made technique in which the migration pattern of the annealing products can be seen on-line due to excitation of a dye (GelStar^{TM}) with blue light. The agarose gels were run at 4 °C in the dark (to observe the fluorescence of the dye) in sodium borate (SB) buffer [9]. The use of SB buffer is important because in other buffers the heat development even with cooling during electrophoresis is melting the annealed TPM shells. The second parameter is the recovery of the perfect shell from the gel. Electro-elution or diffusion-based elution of the TPM-shells out of gel slices gave unsatisfactory results in terms of by-products, yield and time consumption. The best methods in our hands to extract the desired DNA was not to cut a gel slice, but to run the DNA into a second well several centimeters distant from the loading well and to pipett it out there. This necessitates the above mentioned on-line detection during the electrophoretic run. If the desired fraction is entering the well the voltage is shut down and the sample is taken up (in running buffer) with a micropipette. The different fractions can be re-run on a gel and appear to be stable (figure 5).
Additionally, via this method unreacted molecules can be recycled.

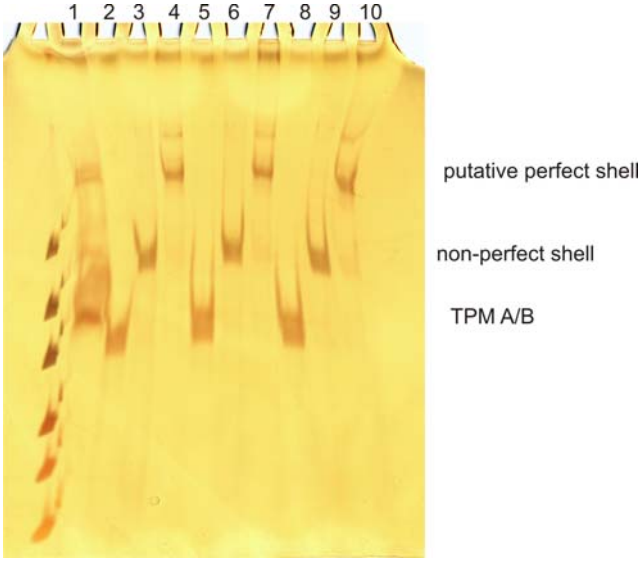

Fig. 5: Gel analysis of annealed and purified TPM A/B shells.
SB PAGE and silver staining. TPMA/B were annealed according to the TEAC protocol. The annealing products were separated and recovered from a SB agarose gel and re-run on the here shown PAGE gel. 1: non-separated annealing aliquot as reference, 2+5+8: non-annealed TPMA/B recovered from the SB gel, 3+6+9: non-perfect shell recovered from the SB gel, 4+7+10: putative perfect shell (AB_4) recovered from the SB gel, 2+3+4 were stored for 2 days at 4 °C, 5+6+7 were stored slowly frozen and stored at -20 °C, 8+9+10 were quickly frozen in liquid nitrogen and stored at -80 °C

4. Analysis of annealing products

To verify that the AB_4-shell has indeed the expected confirmation we started cooperations with partners using electron microscopy and atomic force microscopy. In the following first preliminary results of these analyses are presented.

Electron microscopy analysis (TEM) of the interesting fractions were done in a co-operation with the group of Holger Jeske (Universität Stuttgart, Molekularbiologie und Virologie der Pflanzen). The samples were spread on a drop of buffer containing cytochrome c. The samples were transferred onto a copper grid covered with collodium. The grids were first sputtered with platinum and afterwards with carbon. The grids were loaded in a transmission electron microscope (TEM) and searched for particles. Some representative images are shown in figure 6.

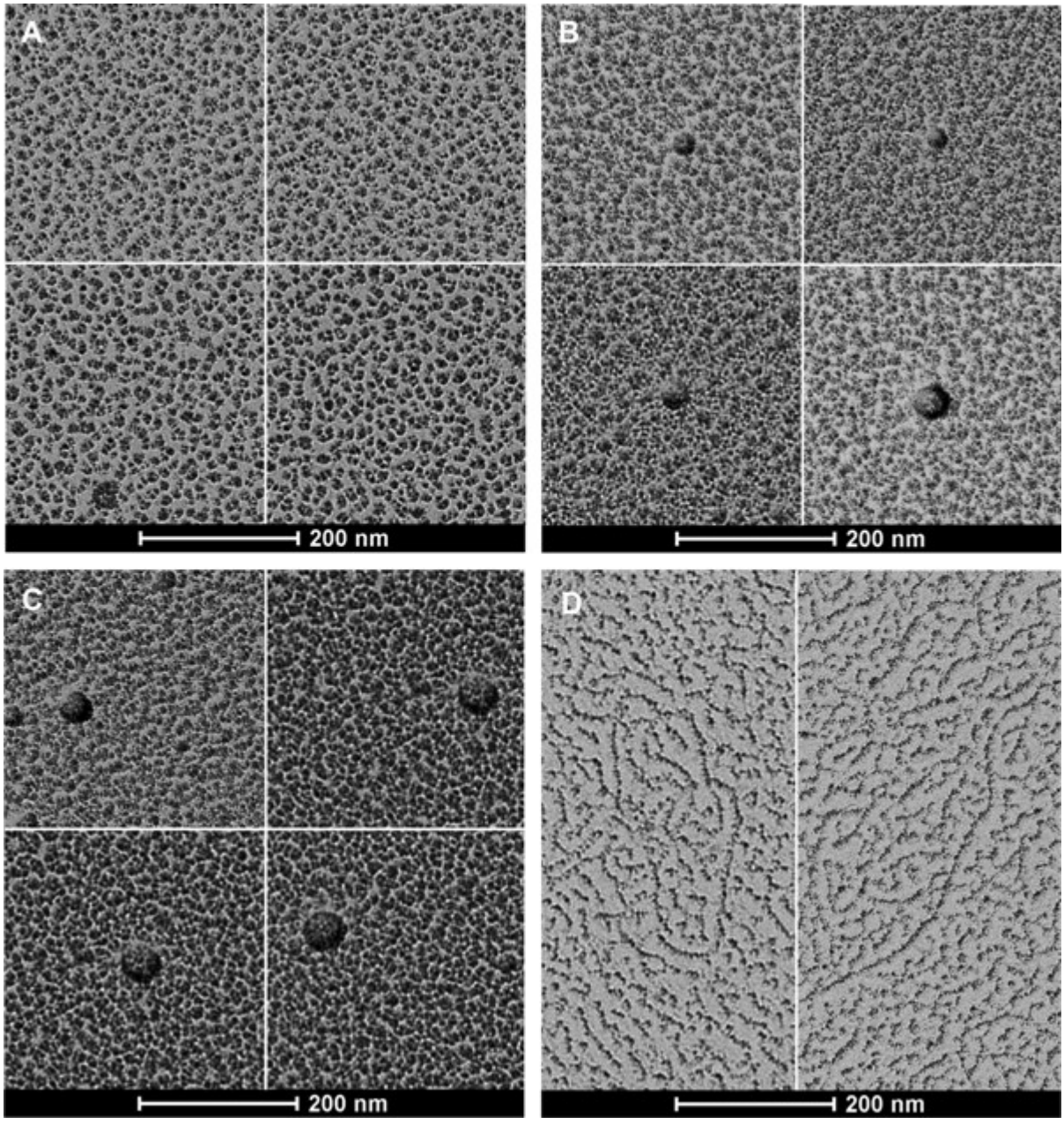

Fig. 6: Analysis of annealing products with TEM. A: background without DNA samples (water control), B: TPM A/B, C: AB_4-shell, D: 6000 nucleotide long single-stranded RNA (tobacco mosaic virus) for comparison (courtesy of Fabian Eber).

For a more quantitative analysis we measured the diameter of the particles. The diameter of the particles in the water control was 26 nm (n=13), for TPM A/B 47,2 nm (n= 20) and for AB_4 80,4 nm. This indicates in relative scales there is a increase of about two-fold in diameter for AB_4 as

compared to TPM A/B alone (the annealing double stranded regions in the AB_4-shell does not contribute to the diameter increase. However at absolute scale the TPM A/B appears too large, because from the dimension of the elongated tetraeder (personal communication Martin Meng) projected on a flat surface (as done in TEM) a diameter of approximately 26 nm would be expected. However due to the adsorption of cytochrom c (a basic protein) to the DNA and the sputtering with platinium afterwards the apparent size of the particles increases. Furthermore it is interesting that some substructures can be seen in the particles, especially in the AB_4-fraction. However, the resolution of this technique is at the limit.
We therefore started a co-operation with Clemens Franz (CFN) to analyze our fractions by atomic force microscopy (AFM). First preliminary results support our hypothesis, because the particles might have an appropriate size and are branched (figure 7).

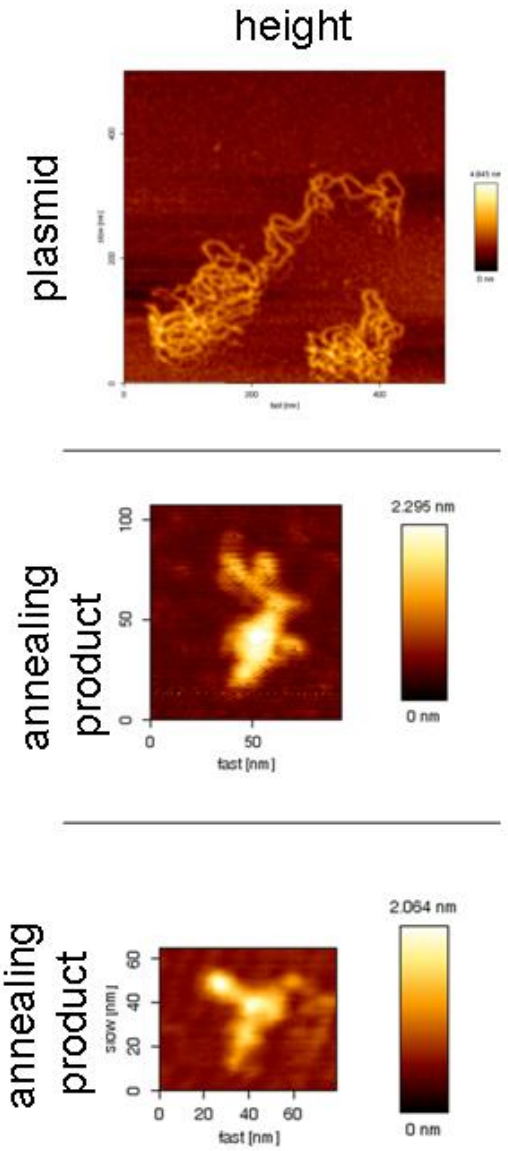

Fig. 7: AFM image of a plasmid (reference) and two structures of our samples (courtesy of Clemens Franz)

However, at the moment only a limited number of particles can be detected on the mica surface. Therefore we are optimizing the binding and washing protocols to increase the number of detectable particles. This work is in progress. In the near future we will also analyze the AB_4A_{12}-shell, which

should be easily visible using both microscopy techniques. In the long run we aim to construct and analyze higher order shells.

5. Synthesis and annealing of stem-loop-TPM hybrids

In many cases when two single stranded nucleic acids anneal, at least one strand has a freely diffusible end. However it is also possible that a single stranded nucleic acid loop may hybridize to another nucleic acid strand in a region where no free nucleic acid end is present. This sort of hybridization is realized e.g. in the codon/anticodon recognition during translation in which the anticodon of the t-RNA is located in a loop. It is also possible that two loops anneal to each other. This so called “kissing loop” and is realized in retroviruses such as the human immunodeficiency virus (HIV) in which the annealing partners are RNA [10]. A homologous structure on a DNA basis could be described, too [11]. We wanted to test whether kissing loop annealing could be an alternative to free-end annealing in our nano-technology project. As a proof-of-concept we designed first a stem-loop structure that possess a self-complementary 5’-overhang. By phosphorylating, annealing and ligating two stem loop structures a new dumb-bell like structure appeared with one stem twice as long as in the ancestor molecules and two loops (see scheme figure 8). Two different oligonucleotides were designed that form such structures (A and B). Because the loops A and B have complementary sequences it was possible to anneal these dumb-bell structures to higher-ordered structures (figure 9).

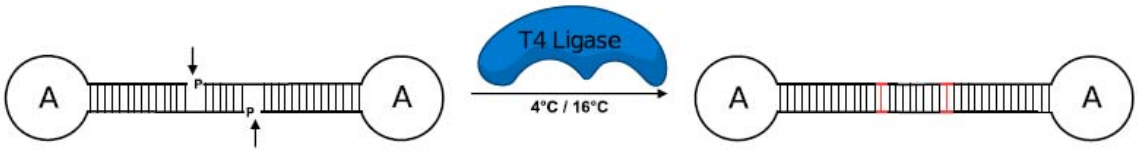

Fig. 8: Synthesis of a dumb-bell structure from two phosphorylated stem-loop structures

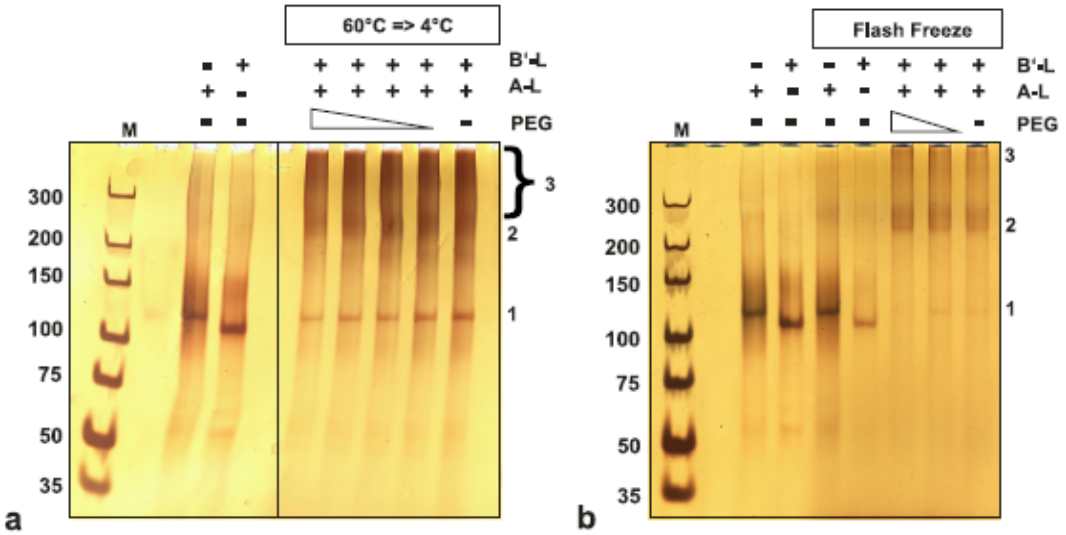

Fig 9: Gel analysis of the optimization of the annealing of the two dumb-bell structures A and B. PAGE in TB magnesium acetate buffer, a: a temperature program was used to optimize the annealing, b: a flesh freeze technique (see [7]) was applied. 1: monomeric dumb-bell structure, 2: dimer of dumb-bells A+B, 3: higher ordered dumb-bell structures. Notice that at b) both spots 1+3 disappear in favor to spot 2.

Due to the feasibility of our concept we decided to link the stem loop to our hybrid TPM. The new stem-loop molecules had now a 5'-overhang complementary to the oligonucleotide arms of the TPM. Analogous to the approach with linear oligonucleotides the TPM was phosphorylated, annealed to the stem-loop structures (with loop A or B respectively) and ligated. The ligation products were purified by PAGE and the spots were extracted. The results are shown in figure 10.

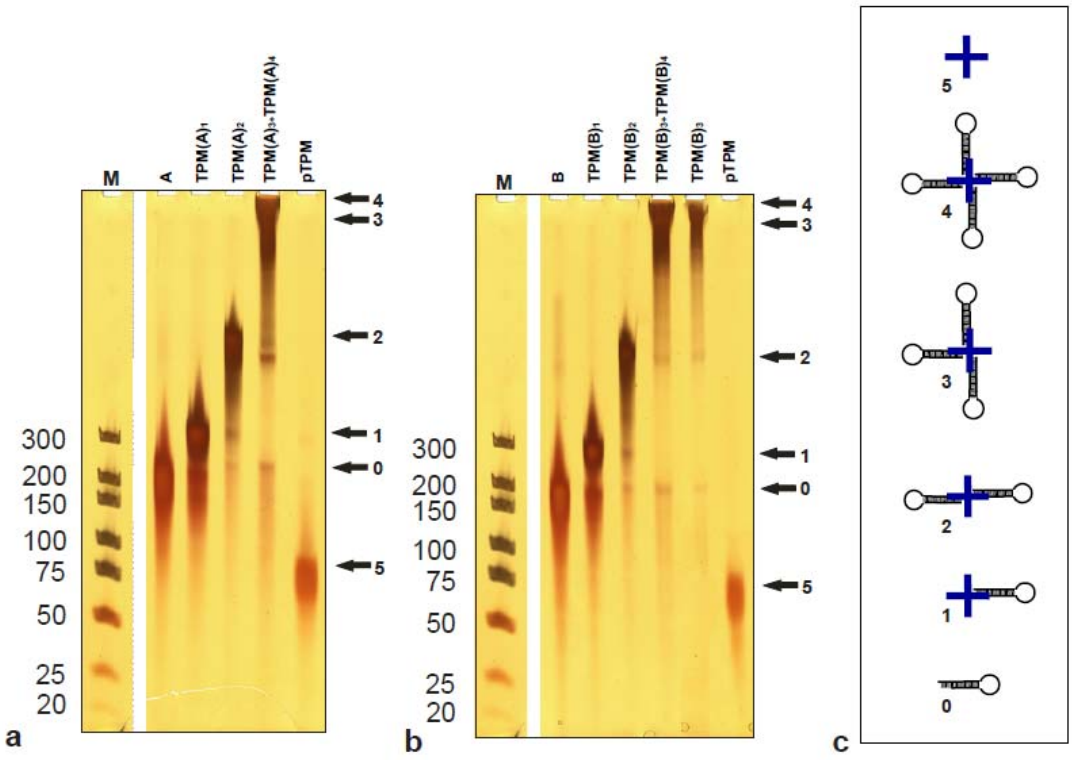

Fig. 10: Gel-electrophoretic characterization of the purified ligated TPM-stem loop structures, TBE gel with 9M urea. The structures 0 (non-ligated excess stem loop), 1 (TPM with one stem loop), 2 (TPM with two stem loops), and 5 (free TPM) can be well separated, whereas structures 3 (TPM with 3 stem loops) and 4 (TPM with 4 stem loops) are not well separated due to their large size.

To test whether the "TPM with stem loop" molecules can anneal via kissing loops we annealed the structures 1 and 2 and as a control the free stem loop (structure 0) with their corresponding complementary molecules. The electrophoretic characterization of these annealing samples indicated that indeed a kissing loop annealing with the TPM constructs is possible (figure 11). This indicates that TPM can not only assemble to higher ordered structures via linear sequences with free ends but also with the kissing loop approach. This result might be extended to other DNA nanostructures such as a recently described tetrahedral structure which can be replicated in biological system and exhibits its structure without an organic core [12]. If the vertices of such a structure could be extended via a stem loop structure such new structures could assemble to higher ordered structures via kissing loop annealing. This opens a new exciting class of nanostructures based on DNA with the advantage of relatively low costs due the synthesis in biological systems.

C5.4 Puchta

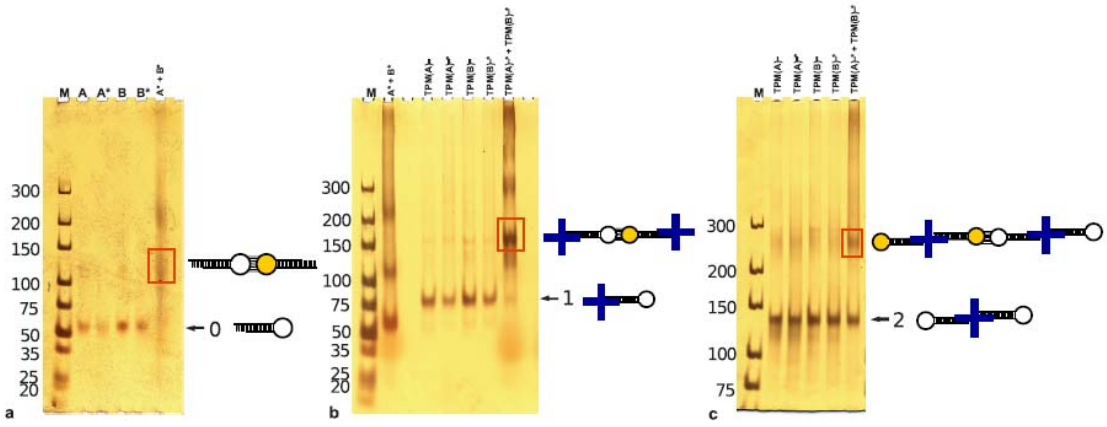

Fig. 11: Kissing loop annealing of different structures with and without a TPM core

References

[C5.4:1] S. Blanck, D. Kobbe, F. Hartung, K. Fengler, M. Focke and H. Puchta, A SRS2 homologue from Arabidopsis thaliana disrupts recombinogenic DNA intermediates and facilitates single strand annealing. Nucl. Acids Res. 37, 7163 (2009)

[C5.4:2] D. Kobbe, S. Blanck, M. Focke and H. Puchta, Biochemical characterization of AtRECQ3 reveals significant differences to other RecQ helicases. Plant Physiol. 151, 1658 (2009)

[C5.4:3] A. Knoll and H. Puchta, The role of DNA helicases and their interaction partners in genome stability and meitic recombination in plants. J. Exp. Bot. in press, Epub available (2011)

[1] N.C. Seeman, Ann. Rev. Biochem. **79**, 65 (2010)

[2] T. Lidl and F.C. Simmel, Nano Lett. **5**, 1894 (2005)

[3] J. Zheng, J.J. Birktoft, Y. Cheng, T. Wang, R. Sha, P.E. Constantinou, S.L. Ginell, C. Mao and N.C. Seeman, Nature **461**, 74 (2009)

[4] D. Nykypanchuk, M.M. Maye, D. van der Lelie and O. Gang, Nature **451**, 549 (2008)

[5] S.Y. Park, A.K.R. Lytton-Jean, B. Lee, S. Weigand, G.C. Schatz, C.A. Mirkin Nature **451**, 553 (2008)

[6] J. Kim, C.-A. Im, Y. Jung, A. Quazi and J.-S. Shin, Biomacromolecules **11**, 1705 (2010)

[7] R.K. Ranjan and K. Rajagopal, Anal. Biochem. **402**, 91 (2010)

[8] J.G. Wetmur, Crit. Rev. Biochem. Molec. Biol. **26**, 227 (1991)

[9] J.R. Brody and S.E. Kern, BioTechniques **36**, 214 (2004)

[10] K.Y. Chang and I. Tinoco, Proc. Nat. Acad. Sci. USA **91**, 8705 (1994)

[11] F. Barbault, T. Huynh-Dinh, J. Paoletti and G. Lancelotti, J. Biomol. Struc. Dyn. **19**, 649 (2002)

[12] Z. Li, B. Wei, J. Nangreave, C. Lin, Y. Liu,Y. Mi and H. Yan, J. Am. Chem. Soc. **131**, 13093 (2009)