A semisynthetic peptide-metalloporphyrin responsive matrix for artificial photosynthesis

Sun, Z.

A Semisynthetic Peptide-Metalloporphyrin Responsive Matrix for Artificial Photosynthesis

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op 20 November 2019
klokke 11:15 uur

door

Zhongwu Sun

geboren te WeiHai

in China

Promotiecommissie:

Promotor:	Prof.dr. H.J.M. de Groot
Referent:	Prof.dr. W.J. de Grip
Overige Leiden:	Prof.dr. A. Kros
	Dr. A. Jakobi
	Mw. Dr. A. Alia
	Prof. dr. Th. Vossegaard

ISBN 978-94-92597-32-8

Propositions

Belonging to the thesis entitled

"A Semisynthetic Peptide-Metalloporphyrin Responsive Matrix for Artificial Photosynthesis"

1. Engineering of responsive matrices is likely of vital importance for improving the efficiency of artificial photosynthesis devices. **(Chapter 1)**
2. Following the rules of natural systems, the derived semisynthetic peptide-metalloporphyrin responsive matrix can be developed for efficient light energy harvesting in artificial photosynthesis. **(Chapter 2)**
3. The self-assembly of CsmA homologues with ZnPP is a multiple functional problem involving the peptide sequence and its length, and environmental effects. **(Chapter 3)**
4. The biological wisdom "structure = function" is also reflected in the semisynthetic systems like the tCsmA-ZnPP complex. **(Chapter 4)**
5. Biology methods for protein harvesting can be widely used, but unpredictable at the same time. **(Chapter 5)**
6. Nature is complicated, but when we explore it, we can always find its rules.
7. Chinese wisdom and western European wisdom are less different than the way seems at first sight, because people are the same.
8. By giving up you can achieve more than with perseverance.
9. The expression *Keep it simple stupid* (KISS) may have its roots in the history of Leiden University around 1700.
10. We may be not the best, but we should do our best.

Index

Contents

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Abbreviations

3DEM	3D Electron Microscopy
BCA	Bicinchoninic acid
BChl	Bacteriochlorophyll
CD	Circular dichroism
CMC	Critical micelle concentration
Cryo-EM	Cryogenic electron microscopy
CsmA	Chlorosome protein A
CTF	Contrast transfer function
DMSO	Dimethyl sulfoxide
E.coli	Escherichia coli
ET	Electron tomography
FMO	Fenna matthews olson
GuHCl	Guanidine hydrochloride
HG	Hexyl β-D-glucopyranoside
ICP-OES	Inductive coupled plasma optical emission spectrometry
IPTG	Isopropyl β- d-1-thiogalactopyranoside
KPi	Potassium phosphate
LD	Linear dichroism
MBP	Maltose-binding protein
MV	Methyl viologen
NCAP	Non-adiabatic conversion by adiabatic passage
NMR	Nuclear magnetic resonance
OG	Octyl β-D-glucopyranoside
RP-HPLC	Reverse phase high-performance liquid chromatography
SEC	Size exclusion chromatography

.

Chapter 1

A general introduction

In recent years, the transition to sustainable energy has become one of the most pressing issues in the world. The search for efficient, clean and renewable energy has become a common goal for scientists all over the world, and solar energy is favored for its renewability, it is environmentally friendly, and has low cost. In nature, light-harvesting complexes absorb sunlight to transfer energy and maintain the activity of life through chemical reactions. After billions of years of evolution, protein matrices have been optimized to cooperate with light-harvesting molecules to efficiently capture and transport solar energy until it can be stored in a chemical bond.[1-3] The natural light-harvesting antenna systems are thought to carry on lossless Non-adiabatic Conversion by Adiabatic Passage (NCAP) processes for high yield.[4-7] Scientists try to create artificial light-harvesting antenna systems to create solar fuel. However, the uptake rate of solar energy by artificial photosynthesis is still very low. At present, most chemists' approach for solar fuel production is focused on changing the properties of dyes and catalysts, thus ignoring the matrix. Engineering of a responsive matrix is likely of vital importance to improve the efficiency of artificial photosynthesis devices. This thesis makes a first step in this direction.

1.1 Moving from natural photosynthesis to artificial photosynthesis

In nature, plants, algae and some bacteria are able to perform photosynthetic processes, converting light energy into chemical energy. In most cases, photosynthesis includes two processes called the 'light' and 'dark' reactions. Both ATP and reduced NADP are produced in the light-reaction after absorption of a photon. And these two products are consumed by the Calvin–Benson–Bassham cycle to reduce carbon dioxide into carbohydrates in the dark-reaction (Figure 1.1).

Photosynthesis is a complicated process and a diversity of strategies is used by different organisms. Commonly, the photosynthetic reactions can be divided into four steps.[8] The light is absorbed by light-harvesting antenna systems in the first step. The harvested light is transferred to photosynthetic reaction centers where charge separation takes place. The positive charges extract electrons from water in the third step, and the negative charges reduce carbon dioxide to carbohydrates in the last step.

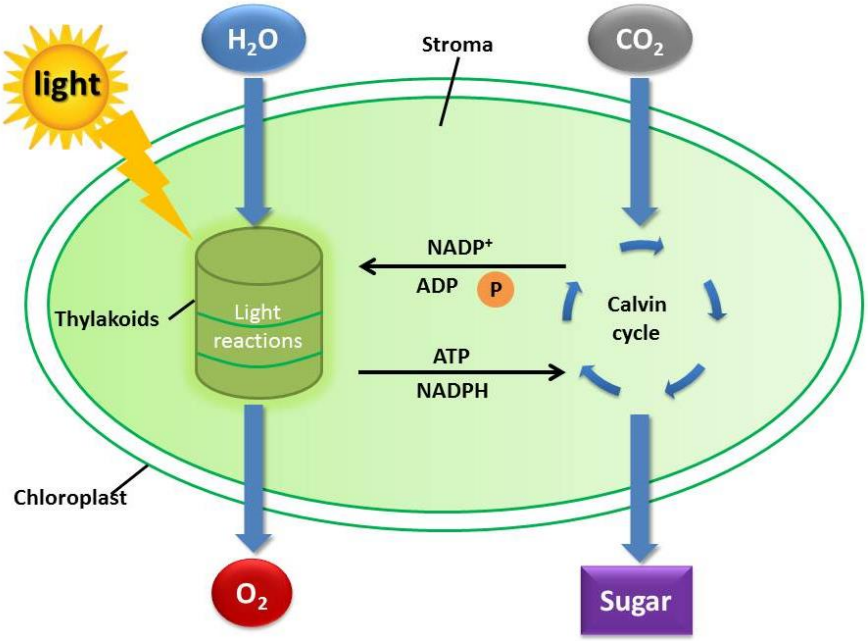


Figure 1.1 Diagram of the process of photosynthesis, showing the light reactions and the Calvin cycle.

During photosynthesis, light energy is converted into chemical energy. However, not all absorbed light energy finally converts into chemical energy, the overall photosynthetic efficiency normally reaches to 3% to 6% of the total incoming light energy.[9] The excess light energy dissipates as heat in a process called non-photochemical quenching, or is emitted as fluorescence to avoid damaging the photosynthetic apparatus. Although the final efficiency of photosynthesis is not high, the light energy can be transferred to reaction centers with nearly 100% efficiency.[10] Natural photosynthesis is highly conserved in evolution, and the biological success of natural photosynthesis inspires scientists to replicate nature and build an artificial photosynthesis system to convert sunlight energy into chemical energy.

To perform good artificial photosynthesis, two crucial factors, light harvesting and charge separation, should be optimized. The harvested light energy should be able to drive charge separation to an electron acceptor to promote chemical reactions. While many different artificial photosynthesis approaches are currently under development, genuinely efficient and robust components for capturing and storing solar energy in chemical bonds were not yet obtained. Research is mainly focused on designing and synthesizing electrocatalysts that can be linked to a light-driven charge separation system.[11-12] In this manner solar cells have been made that can achieve photo-electrochemical hydrogen production with photon to product yield, or external quantum efficiencies, as high as 60%.[13] A challenge is to build antenna complexes with highly ordered arrays of pigments which can transfer nearly 100% of the harvested energy to reaction center. This thesis focuses on necessary steps to be made towards the construction of an efficient and robust semisynthetic artificial light-harvesting antenna system.

1.2 The natural and artificial light-harvesting antenna

Light-harvesting antenna systems play a vital role in solar energy uptake by collecting photons and transferring excited state energy to the reaction center.[14] A wide variety of light harvesting complexes is found among the different photosynthetic species, *e.g.* green sulfur bacterial chlorosomes, cyanobacterial phycobilisomes, dinoflagellate peridinin-chlorophyll-protein, the Pcb protein, higher plant light-harvesting complex II (LHCII), and the purple bacterial light-harvesting complexes 1, 2 and 3 (LH1-LH3).[15-20] They all show similar properties for light harvesting. They all contain high-density arrays of pigments for exciton energy transfer via the Förster or Dexter exciton transfer mechanisms.[21-22] In these light-harvesting complexes, pigments are thought to be tuned by the matrix to perform non-adiabatic conversions that transiently violate detailed balance for high yield.[1, 3, 23] Among the high-density arrays inspired by nature that have been investigated for artificial photosynthesis, however, systems with high energy uptake efficiency have not yet been found.[24-26]

The smallest antenna motif among all natural light-harvesting complexes is the CsmA protein in the baseplate from *Chlorobaculum tepidum*. CsmA proteins (6.2 KDa) in complex with BChl *a* form a crystalline-like baseplate

for energy transfer, in between the surfaces of the chlorosome and the FMO protein complex (Figure 1.2). [27-29] Each CsmA only binds with one BChl *a* and self-assembles to form dimers and strands in the baseplate.[30]

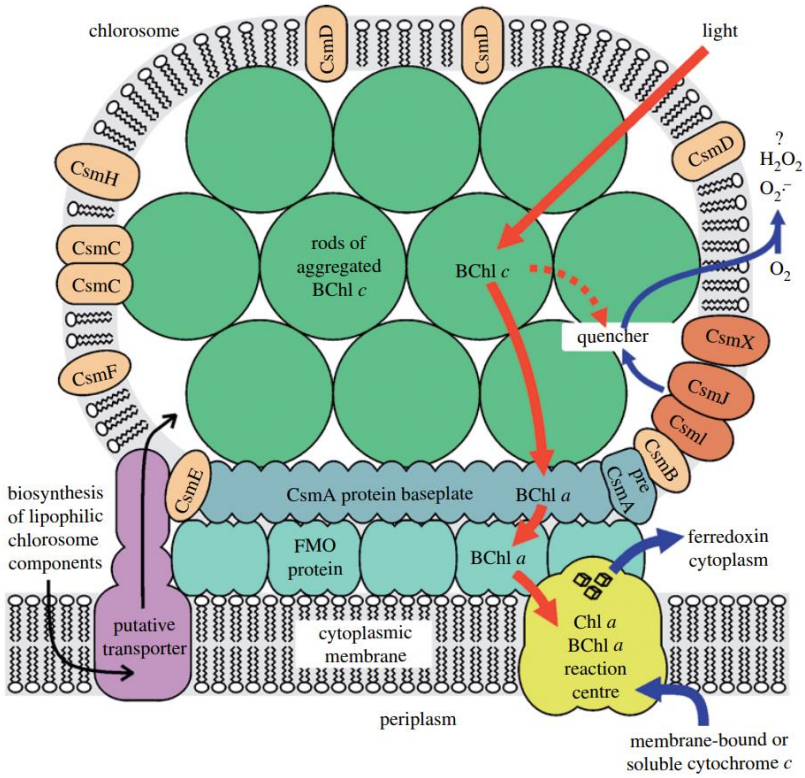


Figure 1.2 Schematic of the photosynthetic apparatus in *Chlorobium tepidum*.[31-32]

The liquid state NMR structure of the 59-residue CsmA protein shows a high α-helical content in the structure; the first and largest helical segment extends from residues 6–36, and contains a binding domain for BChl *a*. The other helical segment is shorter and extends over residues 41–49 towards the C-terminal part.[33] Amphipathic CsmA contains a hydrophobic and a hydrophilic side. With the exception of the only coordinating histidine, all potentially charged residues, e.g. lysine and arginine, are located at one side of the helix (Figure 1.3). The resulting strong hydrophobicity on the opposite side of the helix allows the protein to aggregate. The binding of the BChl *a* cofactor and the assistance of salt bridges formed by the charged residues on the hydrophilic side cause CsmA to self-assemble into long arrays resulting in the para-crystalline superstructure of the baseplate as observed by Cryogenic electron microscopy (cryo-EM).[30, 34-35]

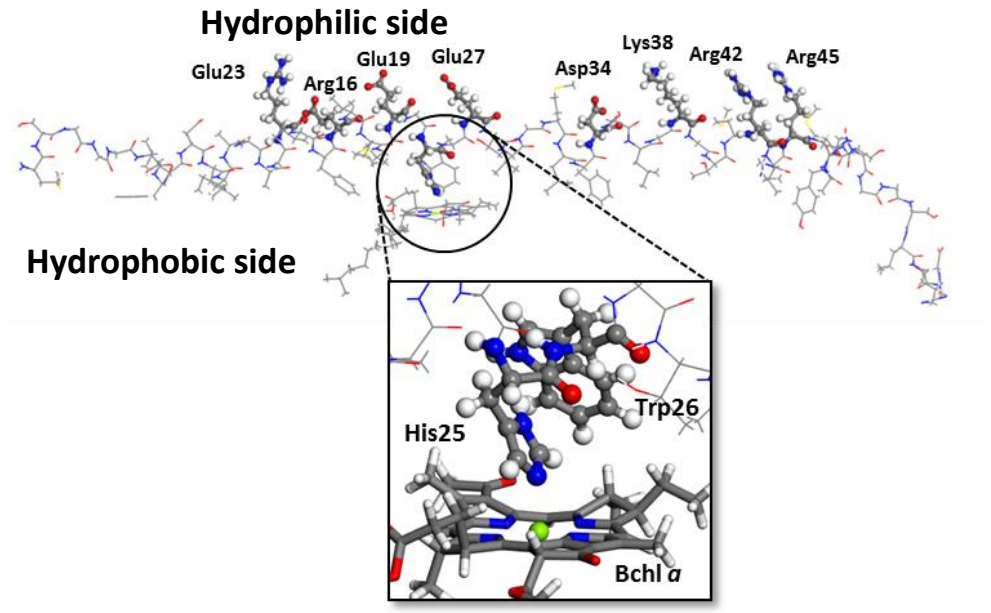


Figure 1.3 The CsmA–BChl *a* baseplate complex

Here we investigate a semisynthetic CsmA (tCsmA) construct, containing 43 amino acids, that was obtained by removing the flexible segments from the N and C termini and leaving the central α-helical part with the two functional helical segments. tCsmA can be synthesized using standard solid-phase Fmoc based chemistry and purified by Reverse phase High-performance liquid chromatography (RP-HPLC). For pigment, we use the chemically robust zinc-protoporphyrin IX (ZnPP) , a compound found in red blood cells that is often used in artificial light harvesting studies, to replace BChl *a*.[35-37]

Most research aiming for artificial antennae focuses on building entirely synthetic peptide scaffolds, such as four helix bundles, instead of modifying a natural complex.[37-41] Although bottom up design is not always straightforward, it allows for flexibility in the incorporation of chromophores. Often the stability is worse than for natural systems and it is difficult to match the ability of light harvesting compared to the performance of natural systems. In contrast, the semisynthetic tCsmA-ZnPP complex builds a bridge between the natural system and fully artificial systems. Using the outcome of evolution in biology as a starting point, with subsequent modification afterwards, we can possibly build a simple, stable and efficient semi-artificial light-harvesting antenna system.

1.3 Non-adiabatic Conversion by Adiabatic Passage (NCAP) processes

Traditionally, transmission of exciton energy is thought to proceed via the Förster or Dexter exciton transfer mechanisms.[21-22, 42] Most photosynthetic light harvesting complexes are chiral "responsive" matrix assemblies. It is believed that they can undergo a light-driven conformational change and self-select vibrations to develop vibronic states for driving semi-classical coherent transfer with nearly 100% yield in lossless Non-adiabatic Conversion by Adiabatic Passage (NCAP) processes.[4-7, 43]

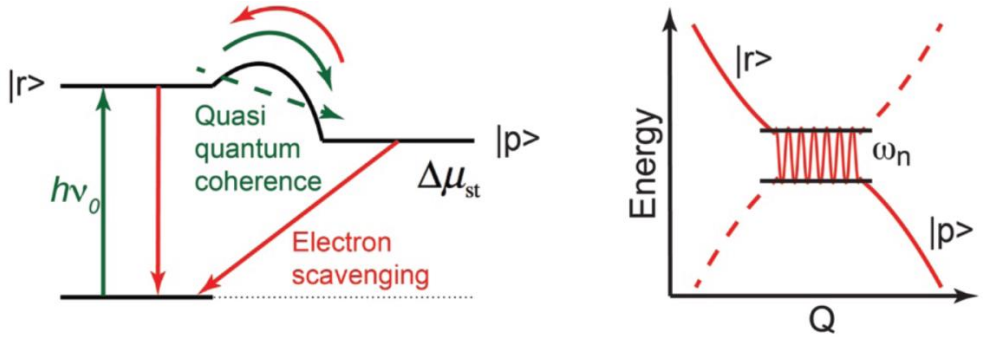


Figure 1.4 Conversion from a reactant state |r> to a product state |p>. Forward processes are shown in green and reverse processes in red. [43]

The Pauli Exclusion Principle states that two or more identical fermions cannot occupy the same quantum state within a given quantum system simultaneously. Thereby crossing of the electronic states is forbidden. The adiabatic conversion from a pure reactant state |r> into a product state |p> normally has a barrier to overcome, which is unfavorable with respect to energy efficiency. However, in a chiral "responsive matrix" the transfer from |r> to |p> proceeds non-adiabatically. A transient semi-classical channel through the barrier opens up when anticorrelated |r> and |p> states approach each other and resonant coupling to a vibrational mode producess off-diagonal elements in the Hamiltonian. This can occur most effectively in a chiral natural or semisynthetic matrix that is subject to a conformational change upon excitation. (Figure 1.4). The NCAP process differs from the adiabatic (Born-Oppenheimer type) process which is shown as a trajectory over the barrier (Figure 1.4, green curved arrow, right). The adiabatic process may be subject to back transfer (Figure 1.4, red curved arrow) and

subsequent relaxation (Figure 1.4, red downwards arrow) or direct recombination as a result of electron scavenging on the nanoscale (Figure 1.4, red diagonal arrow). However, due to dynamic symmetry breaking in chiral light-harvesting complexes and resonant coupling to a vibration, the NCAP process may proceed in a unidirectional lossless manner, driven by a ω_n twisting mode strengthened by the specific ordering of the system matrix (Figure 1.4, right).[43] The resonant coupling to a vibrational mode is a Goldilocks-type self-selection process, which is implicitly verified in this thesis that focuses on a chiral matrix to produce rapid energy transfer between cofactors.[3]

1.4 Constitution of an artificial light-harvesting antenna complex

1.4.1 The synthesis of tCsmA

There are two major paths to obtain protein products, these are chemical synthesis and protein biosynthesis, and both have their own advantages and disadvantages.

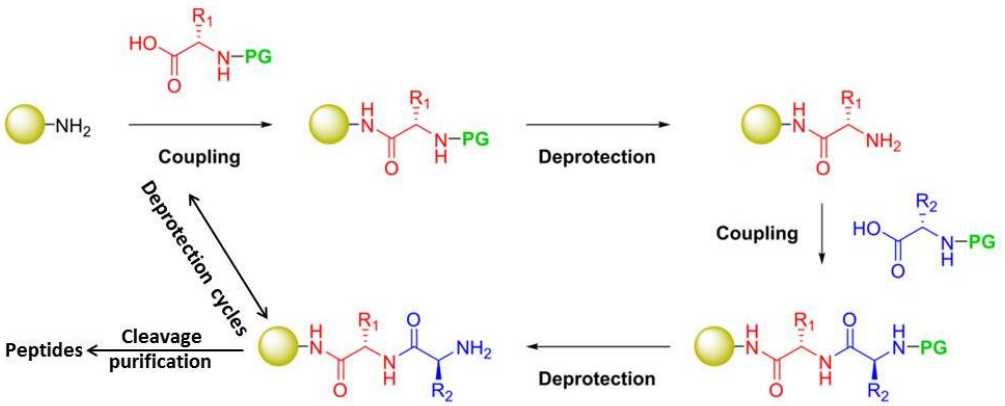


Figure 1.5 Solid-phase synthesis of peptide.

The most common and efficient chemical synthesis is known as solid-phase peptide synthesis (SPPS).[44-45] The SPPS procedure consists of repeated cycles of alternate N-terminal deprotection and coupling reactions (Figure 1.5). Amino acids are coupled to a resin which is washed before each step.[46] Subsequently, the deprotected amine can couple with the free acid of the

second amino acid. After a number of cycle repeats the desired peptide is stripped of all its protecting groups using a strong acid reagent, *e.g.* trifluoroacetic acid or a nucleophile and the resulting deprotected peptide strand is purified by RP-HPLC. However, the production of long protein sequences or isotopically labelled protein via SPPS is challenging due to the high costs of isotope labelled chemicals as well as due to technical limitations.

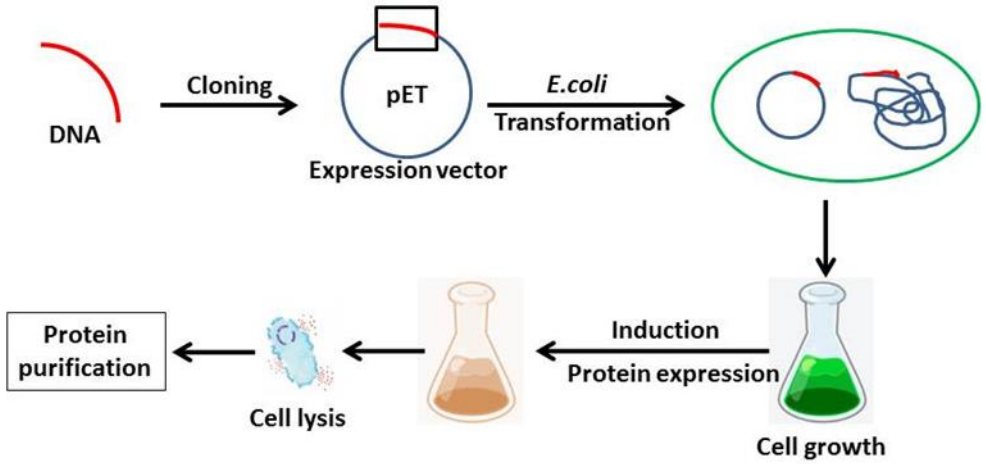


Figure 1.6 Steps for synthesis of the recombinant proteins in *E. coli*.

On the other hand, via protein biosynthesis large or isotope labelled proteins are more easily obtained. Generally, the gene of interest is cloned into the expression vector followed by transformation into the desired *Escherichia coli* (*E.coli*) strain. Upon induction, the cell will begin to overexpress protein. Other cell components are to be removed by purification after cell lysis (Figure 1.6). Protein expression in *E.coli* has been the most popular means of producing recombinant proteins for over the past two decades. *E.coli* is a well-established host that offers short culturing time, easy genetic manipulation and low cost media. Membrane proteins have been synthesized in the past using *E.coli* strains such as BL21.[47] Using these strains has the advantages of rapid, inexpensive and simple protein expression with a possibility of achieving high yields, even with proteins which are not native to *E.coli*. However, not all protein can be produced or easily purified. In addition, the extra residues can stay in the sequence if using fusion protein.

1.4.2 The constitution of the tCsmA-ZnPP complex

Amphipathic tCsmA and ZnPP have very poor solubility in water. To constitute the tCsmA and ZnPP, they should be co-dissolved with assistance of organic additives. However, if too much of these organic additives is present, the binding cannot easily take place. For instance, a detergent can be used for mediating constitution.[48] The amphiphilic nature of a detergent allows it to interact with hydrophobic membrane proteins and aqueous solvent to keep the protein in the dissolved state. Detergents are widely used for extraction, purification, and manipulation of membrane proteins.

Generally, detergents contain segregated polar and apolar domains (Figure 1.7 left).[49] In aqueous solution, a detergent is able to form micelles once a concentration above the critical micelle concentration (CMC) is reached. The shape and size of a micelle are a function of the molecular geometry of the detergent and solution conditions such as concentration, temperature, pH, and ionic strength. In addition, the concentration is important to the formation of the micelle. At below a certain temperature, the so called Krafft temperature, a saturated detergent solution will be in equilibrium with solid surfactant. In this case the temperature is too low for micelles to form. Above this temperature it is possible to form micelles and the detergent becomes much more soluble (Figure 1.7 right).[49]

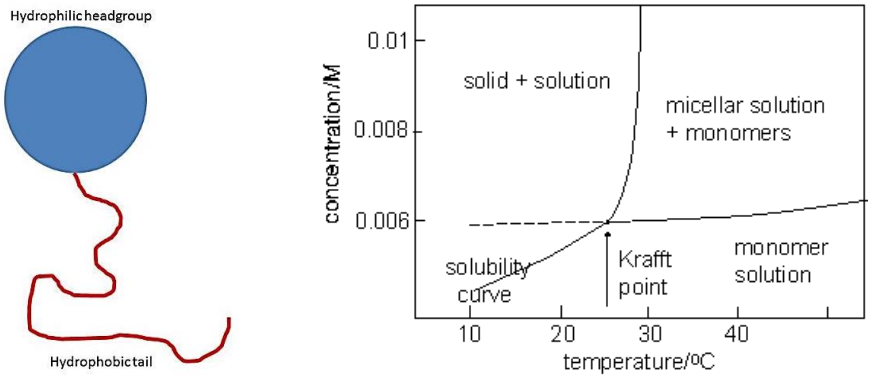


Figure 1.7 left : A typical illustration of a surfactant molecule. right : Temperature dependence of detergent solubility.[50]

Using these known properties we can control the solution conditions to solubilize the tCsmA and ZnPP above the CMC and while binding is allowed to take place below the CMC by dilution at low temperature. Here, detergents with high CMC such as Octyl β-D-glucopyranoside (OG) and Hexyl

β-D-glucopyranoside (HG) are a better choice for the constitution than detergents with low CMC, since a high enough final concentration of complex can be maintained in this manner. The solubility of tCsmA is proportional to the concentration of detergent (Figure 1.8). However, the detergent is difficult to remove while concentrating the protein due to the micelle formation.

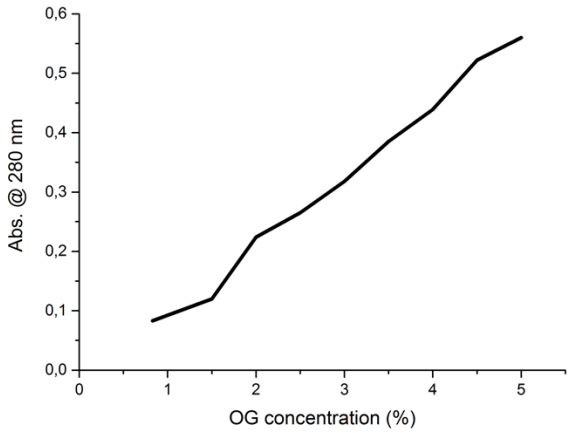


Figure 1.8 The OG-concentration dependent solubility of tCsmA.

In addition to dissolving the components with detergents, dialysis can also be used to control the solution conditions via buffer exchange. Compared to detergents, the complex can be obtained in pure buffer without dilution. Moreover, detergents are not necessary during dialysis for solubilization of protein, other compounds such as 1-propanol can also solubilize tCsmA well. In addition, compounds such as 1-propanol cannot form micelles which results in a faster dialysis speed compared to the dialysis of detergents.

1.5 Structure determination

X-ray crystallography, NMR spectroscopy, and electron microscopy are mainly used to determine the structures of proteins. X-ray crystallography can show atomic details in crystallized molecules. For NMR spectroscopy, the signal from the elements with nuclear spin ½ are able to offer information on the local conformation and distance between atoms that are close to one another to solve the structure. With 3D Electron Microscopy (3DEM) the 3D structure can be obtained from 2D projection images produced by transmission electron microscopy (TEM).

Table 1.1 Comparison of different techniques[51-54]

Methods	Advantages	Disadvantages
X-ray	1. High resolution can be obtained. 2. It costs less than other techniques. 3. The method for analysis is fully completed.	1. Crystallization is necessary. 2. We cannot examine the behavior of the molecules in solution. 3. Studying of motions is not possible. 4. Not all proteins crystallize well, the quality of the crystals influences the accuracy. 5. Cannot observe the hydrogen atoms. 6. The native state may be not observed.
NMR	1. The motion of protein can be examined. 2. The activation-thermodynamics and certain kinetic data can be resolved. 3. The influence of the dielectric constant, the polarity and any other property of the solvent or added material can be tested. 4. The sample can be in the solution phase or in the solid state so that the native state can be observed for both membrane and soluble proteins. 5. NMR can be used to obtain abundant information, *e.g.* angles, distances, coupling constants, chemical shifts, rate constants.	1. NMR is limited to relatively small proteins or parts of proteins 2. Costly. 3. Normally isotope labelling is usually necessary for the structure determination of protein. 4. High purity of sample is necessary. 5. Large amounts of sample are necessary.
3DEM	1. It is possible to observe large molecules or self-assemblies. 2. The native state of molecules can be observed. 3. It becomes possible to view cells, cell organelles as well as macromolecular complexes. 4. Only small amounts of sample are needed for 3DEM.	1. A large number of images is needed for 3D structure analysis. 2. Not easy to obtain images from a tilted specimen. 3. The electron-beam is easy to destroy the specimen. 4. Costly.

Each method has advantages and disadvantages (Table 1.1). Dependent on the type of sample and the experimental conditions, one or more methods can be chosen for structure determination. In this thesis, 3DEM is chosen for structure determination.

1.6 Single particle analysis for helical structure

TEM uses an electron beam to examine the structures of molecules and materials at the atomic scale. However, traditional TEM is not good enough for biomolecules due to the evaporation of surrounding water, and the risk of damage from the high energy electrons burning the sample. Hence, cryo-EM is used for structure determination of biomolecules to overcome these problems by using frozen samples, weaker electron beams with a low dose of electrons, and advanced image processing. Single-particle analysis (SPA) and electron tomography (ET) are two common techniques for cryo-EM to solve the structure puzzle of biomolecules. The technique used in this thesis is single particle analysis.

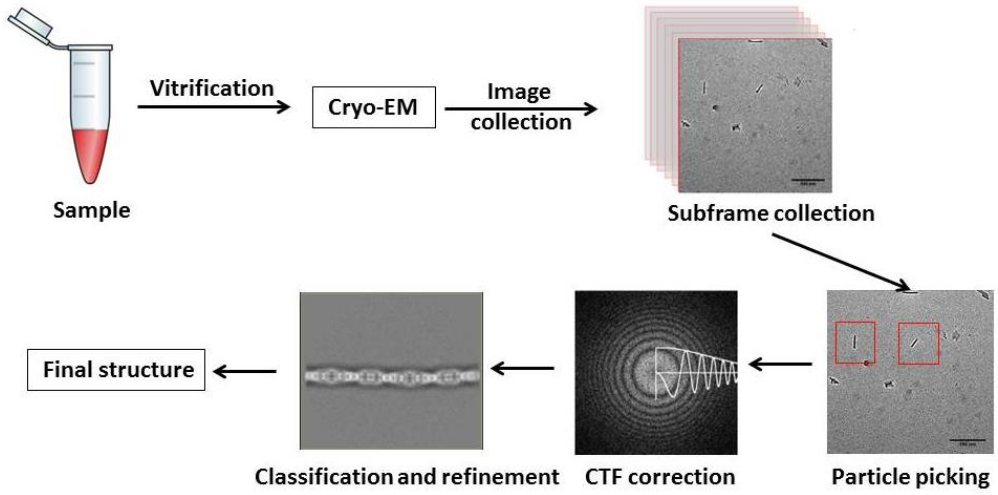


Figure 1.9 The process of single-particle cryo-EM.

The principle of SPA is based on averaging multiple views of a single molecule. A large number of images of the molecule is necessary for 3D reconstruction. The collected images from cryo-EM reflect the electron density of sample. To get less noise and higher resolution, the blotting and

vitrification process should be rapid to ensure a thin ice layer. Cryo-EM images are usually collected in defocused condition to increase phase contrast. The collected data is a series of images that need to be aligned to compensate for sample shift and beam-induced movement to reduce image blurring. These images are modulated in frequency space by the phase Contrast Transfer Function (CTF). By aligning several similar images to each other, an image with a higher signal to noise ratio can be obtained. There are several software packages such as SPIDER, EMAN2, RELION, SPARX and SPRING to process the collected images.[55-59] Generally, the particles should be picked and extracted for further classification and refinement (Figure 1.9).

Except for the single molecule or complex, SPA can be used for helical filaments if the accurate helical symmetry is applied to the 3D construction process by using RELION.[60] SPRING is a single-particle based helical reconstruction package for electron cryo-micrographs and has been used to determine the precise helical symmetry by using a class-based helical reconstruction scheme that explores the helical symmetry by reconstructing the 3D structures from a large number of symmetry combinations using a single class average. Afterwards, according to the cross-correlation of the power spectra of the class average and projections of the resulting 3D models, the helical symmetry can be extracted from a large number of possible symmetry combinations.

1.7 Scope of this thesis

This thesis is concerned with the constitution of chiral responsive matrix assemblies that may undergo light-driven conformational changes and self-select vibrations to develop vibronic states for driving semi-classical coherent transfer with nearly 100% yield in a lossless NCAP process. In this work, a truncated CsmA derived from the CsmA protein from the baseplate of the *Chlorobaculum tepidum* has been used to engineer a chiral matrix around a zinc protoporphyrin IX (ZnPP) chromophore (Chapter 2). After the study of the specific structure of tCsmA-ZnPP complex, the mechanism of self-assembly was resolved (Chapter 3 and Chapter 4). Expression and purification of tCsmA in *E.coli* was explored to pave the way for other methods, such as NMR (Chapter 5).

1.8 References

1. Chin, A. W.; Datta, A.; Caruso, F.; Huelga, S. F.; Plenio, M. B., *New Journal of Physics* **2010,** *12* (6), 065002.
2. Chin, A.; Prior, J.; Rosenbach, R.; Caycedo-Soler, F.; Huelga, S.; Plenio, M., *Nature Physics* **2013,** *9* (2), 113.
3. Huelga, S. F.; Plenio, M. B., *Contemporary Physics* **2013,** *54* (4), 181-207.
4. Mathies, R. A., *Nature chemistry* **2015,** *7* (12), 945.
5. Purchase, R.; De Groot, H., *Interface Focus* **2015,** *5* (3), 20150014.
6. Yin, X.; Struik, P. C., *Journal of experimental botany* **2015,** *66* (21), 6535-6549.
7. Hoffman, D. P.; Mathies, R. A., *Accounts of chemical research* **2016,** *49* (4), 616-625.
8. Cogdell, R. J.; Brotosudarmo, T. H.; Gardiner, A. T.; Sanchez, P. M.; Cronin, L., *Biofuels* **2010,** *1* (6), 861-876.
9. Hall, D.; House, J., *Biomass and Bioenergy* **1994,** *6* (1-2), 11-30.
10. Millan-Almaraz, J. R.; Guevara-Gonzalez, R. G.; Romero-Troncoso, R.; Osornio-Rios, R. A.; Torres-Pacheco, I., *African Journal of Biotechnology* **2009,** *8* (25).
11. Tran, P. D.; Wong, L. H.; Barber, J.; Loo, J. S., *Energ Environ Sci* **2012,** *5* (3), 5902-5918.
12. Barber, J.; Tran, P. D., *Journal of The Royal Society Interface* **2013,** *10* (81), 20120984.
13. Nann, T.; Ibrahim, S. K.; Woi, P. M.; Xu, S.; Ziegler, J.; Pickett, C. J., *Angewandte Chemie International Edition* **2010,** *49* (9), 1574-1577.
14. Blankenship, R. E., *Molecular mechanisms of photosynthesis*. John Wiley & Sons: 2014.
15. Adir, N., *Photosynthesis Research* **2005,** *85* (1), 15-32.
16. Larkum, T., *Trends in Plant Science* **1996,** *1* (8), 252.
17. Liu, Z.; Yan, H.; Wang, K.; Kuang, T.; Zhang, J.; Gui, L.; An, X.; Chang, W., *Nature* **2004,** *428* (6980), 287.
18. Pšenčík, J.; Ikonen, T.; Laurinmäki, P.; Merckel, M.; Butcher, S.; Serimaa, R.; Tuma, R., *Biophysical journal* **2004,** *87* (2), 1165-1172.
19. Ting, C. S.; Rocap, G.; King, J.; Chisholm, S. W., *Trends in microbiology* **2002,** *10* (3), 134-142.
20. van Grondelle, R.; Novoderezhkin, V., *Biochemistry* **2001,** *40* (50), 15057-15068.
21. Van Grondelle, R., *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics* **1985,** *811* (2), 147-195.
22. van Grondelle, R.; Dekker, J. P.; Gillbro, T.; Sundstrom, V., *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1994,** *1187* (1), 1-65.
23. Chin, A.; Prior, J.; Rosenbach, R.; Caycedo-Soler, F.; Huelga, S.; Plenio, M. B., *Nature Physics* **2013,** *9* (2), 113.

24. Hu, S.; Chi, C.-Y.; Fountaine, K. T.; Yao, M.; Atwater, H. A.; Dapkus, P. D.; Lewis, N. S.; Zhou, C., *Energ Environ Sci* **2013,** *6* (6), 1879-1890.
25. Yalamanchili, S.; Emmer, H. S.; Fountaine, K. T.; Chen, C. T.; Lewis, N. S.; Atwater, H. A., *Acs Photonics* **2016,** *3* (10), 1854-1861.
26. Gao, L.; Cui, Y.; Wang, J.; Cavalli, A.; Standing, A.; Vu, T. T.; Verheijen, M. A.; Haverkort, J. E.; Bakkers, E. P.; Notten, P. H., *Nano Lett* **2014,** *14* (7), 3715-3719.
27. Montano, G. A.; Wu, H.-M.; Lin, S.; Brune, D. C.; Blankenship, R. E., *Biochemistry* **2003,** *42* (34), 10246-10251.
28. Staehelin, L. A.; Golecki, J. R.; Drews, G., *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1980,** *589* (1), 30-45.
29. Sprague, S.; Staehelin, L.; DiBartolomeis, M.; Fuller, R., *Journal of bacteriology* **1981,** *147* (3), 1021-1031.
30. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.
31. Cogdell, R. J.; Gardiner, A. T.; Cronin, L., *Phil. Trans. R. Soc. A* **2012,** *370* (1972), 3819-3826.
32. Blankenship, R. E.; Olson, J. M.; Miller, M., Antenna complexes from green photosynthetic bacteria. In *Anoxygenic photosynthetic bacteria*, Springer: 1995; pp 399-435.
33. Pedersen, M. Ø.; Pham, L.; Steensgaard, D. B.; Miller, M., *Biochemistry* **2008,** *47* (5), 1435-1441.
34. Pedersen, M. Ø.; Linnanto, J.; Frigaard, N.-U.; Nielsen, N. C.; Miller, M., *Photosynthesis research* **2010,** *104* (2-3), 233-243.
35. Frigaard, N.-U.; Li, H.; Martinsson, P.; Das, S. K.; Frank, H. A.; Aartsma, T. J.; Bryant, D. A., *Photosynthesis research* **2005,** *86* (1-2), 101-111.
36. Luo, L.; Chang, C. H.; Chen, Y. C.; Wu, T. K.; Diau, E. W., *J Phys Chem B* **2007,** *111* (26), 7656-64.
37. Fry, H. C.; Garcia, J. M.; Medina, M. J.; Ricoy, U. M.; Gosztola, D. J.; Nikiforov, M. P.; Palmer, L. C.; Stupp, S. I., *J Am Chem Soc* **2012,** *134* (36), 14646-14649.
38. Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L., *Biochemistry* **1998,** *37* (13), 4635-4643.
39. Sharp, R. E.; Moser, C. C.; Rabanal, F.; Dutton, P. L., *Proceedings of the National Academy of Sciences* **1998,** *95* (18), 10465-10470.
40. Grosset, A. M.; Gibney, B. R.; Rabanal, F.; Moser, C. C.; Dutton, P. L., *Biochemistry* **2001,** *40* (18), 5474-5487.
41. Koder, R. L.; Valentine, K. G.; Cerda, J.; Noy, D.; Smith, K. M.; Wand, A. J.; Dutton, P. L., *J Am Chem Soc* **2006,** *128* (45), 14450-14451.
42. McConnell, I.; Li, G.; Brudvig, G. W., *Chemistry & biology* **2010,** *17* (5), 434-447.
43. R. Purchase, R. C., F. Breitling,V. Stadler,; N. van Hulst, G.-J. K., A. Ramirez,; R. Zwijnenberg, A. K., J. B. de Baan,; Groot, I. R. a. H. J. M. d., Semi-Synthetic Responsive Matrices for Artificial Photosynthesis. In *In Bioinspired Chemistry: From Enzymes to Synthetic Applications*, Reglier, M., Ed.

44. Synthesis—A, F. S. P. P., Practical Approach. Chan, WC: 2000.
45. Fields, G. B., *Peptide characterization and application protocols*. Humana Press: 2007; Vol. 386.
46. Chan, W. C.; White, P. D., *Fmoc solid phase peptide synthesis*. Oxford University Press: 2000.
47. Miroux, B.; Walker, J. E., *Journal of molecular biology* **1996,** *260* (3), 289-298.
48. Pedersen, M. O.; Pham, L.; Steensgaard, D. B.; Miller, M., *Biochemistry* **2008,** *47* (5), 1435-41.
49. Kronberg, B.; Lindman, B., *Surfactants and polymers in aqueous solution*. John Wiley & Sons Ltd., Chichester: 2003.
50. https://nova.disfarm.unimi.it/probuilder.htm.
51. Milne, J. L.; Borgnia, M. J.; Bartesaghi, A.; Tran, E. E.; Earl, L. A.; Schauder, D. M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S., *The FEBS journal* **2013,** *280* (1), 28-45.
52. Cheng, Y.; Grigorieff, N.; Penczek, P. A.; Walz, T., *Cell* **2015,** *161* (3), 438-449.
53. Warren, B. E., *X-ray Diffraction*. Courier Corporation: 1990.
54. Coates, G. R.; Xiao, L.; Prammer, M. G., *NMR logging: principles and applications*. Haliburton Energy Services Houston: 1999; Vol. 344.
55. Frank, J.; Shimkin, B.; Dowse, H., *Ultramicroscopy* **1981,** *6* (4), 343-357.
56. Hohn, M.; Tang, G.; Goodyear, G.; Baldwin, P. R.; Huang, Z.; Penczek, P. A.; Yang, C.; Glaeser, R. M.; Adams, P. D.; Ludtke, S. J., *Journal of structural biology* **2007,** *157* (1), 47-55.
57. Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J., *Journal of structural biology* **2007,** *157* (1), 38-46.
58. Scheres, S. H., *Journal of structural biology* **2012,** *180* (3), 519-530.
59. Desfosses, A.; Ciuffa, R.; Gutsche, I.; Sachse, C., *Journal of structural biology* **2014,** *185* (1), 15-26.
60. He, S.; Scheres, S. H., *Journal of structural biology* **2017,** *198* (3), 163-176.

Chapter 2

Self-assembly of a semisynthetic zinc-protoporphyrin IX peptide complex derived from the natural CsmA light harvesting antenna

2.0 Abstract

The CsmA protein from the baseplate of the *Chlorobaculum tepidum* is proposed as an attractive motif for the engineering of semisynthetic materials for artificial photosynthesis. We perform self-assembly of zinc protoporphyrin IX (ZnPP) with a truncated CsmA (tCsmA) amphipathic peptide scaffold for the bottom-up construction of a semisynthetic matrix for energy transfer. The resulting self-assembly shows an extended periodic structure with 7_2 helical symmetry. The complex is able to maintain a good photostability for at least 30 min when the concentration is more than 20 µM, and almost 100% harvested energy can be transferred to the acceptor cation radical methyl viologen (MV^+). The excitation-energy transfer rate of the complex was increased from the ns^{-1} scale to the ps^{-1} scale that is comparable to natural light-harvesting antenna systems. It contributes to converging evidence that chiral responsive matrix assemblies can be developed for efficient light energy harvesting in artificial photosynthesis.

This chapter has been published as: Zhongwu Sun, Christoph A. Diebolder, Ludovic Renault, Huub de Groot. A semisynthetic peptide-metalloporphyrin responsive matrix for artificial photosynthesis, 2019, *ChemPhotoChem*

2.1 Introduction

Solar energy is a major source of renewable energy with the potential to address major sustainability challenges the world is facing. A finite supply of low carbon electricity from photovoltaics and wind can be followed by electrolysis to produce fuels, but such adiabatic conversion and concentration chains are not optimally efficient. In contrast, nature makes abundant use of non-adiabatic conversions that transiently violate detailed balance for high yield.[1-3] Photosynthesis uses ordered arrays of pigments for energy or charge transfer.[4-5] Traditionally, transmission of exciton energy is thought to proceed via the Förster or Dexter exciton transfer mechanisms.[6-8] However, photosynthetic proteins are chiral matrix assemblies which are thought to undergo a light-driven conformational change and self-select vibrations to develop vibronic states for driving semi-classical coherent transfer with nearly 100% yield in lossless processes, a concept known as the responsive matrix.[9-13]

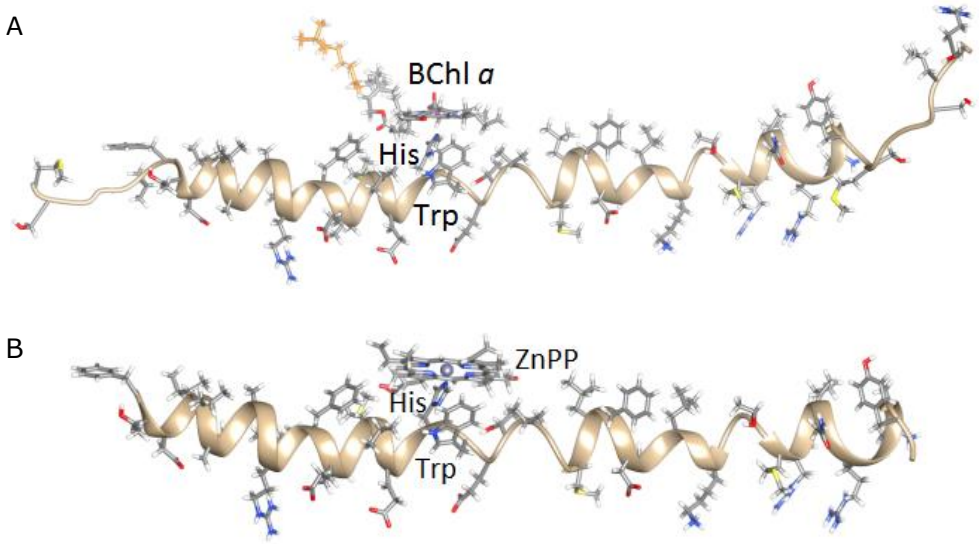


Figure 2.1 (A) The structure of the CsmA-BChl *a* complex from the baseplate (PDB 5LCB). (B) Putative model structure of the tCsmA in complex with ZnPP, as derived from the CsmA-BChl *a* complex.

The critical step in constructing a responsive matrix is the chemical engineering of a chiral module. Here we start from a zinc protoporphyrin IX (ZnPP) chromophore and provide a new paradigm for energy transfer in a semi-synthetic matrix for artificial photosynthesis.[14-16] We use tCsmA

construct, containing 43 amino acids, which is derived from the helical CsmA peptide from the baseplate of *Chlorobaculum tepidum*. This is the smallest light-harvesting protein known in nature, and coordinates a single BChl *a* (Figure 2.1B). The tCsmA was modified by removing the flexible segments from the N and C-terminal, which attach to the lipid monolayer confining the chromosome and the FMO complex respectively, leaving the functional central α-helix.[17] In nature, dimers of CsmA-BChl *a* self-assemble to form extended para-crystalline layers with BChl *a* in CsmA strands.[17] The system assists in energy transfer and storage from the chlorosome to the FMO antenna and the photosynthetic reaction center in *Chlorobaculum tepidum* (Figure 1.2).[18-19] NMR studies of monomeric CsmA in solution have shown that it is a α-helical amphipathic peptide.[20] With the exception of the single coordinating histidine in the sequence, all potentially charged residues, like glutamine, lysine and arginine, are located on one side of the helix (Figure 1.3 & 2.1A).

With the assistance of tCsmA scaffolds, ZnPP (SFigure 2.1) is bound to form a molecular wire with complexation via slow dialysis.[21] We perform initial assays of the light harvesting, indicating that excitons formed in the array upon photoexcitation have lifetimes on the ns scale, comparable to ZnPP in solution.[22] When energy transfer is significantly faster than this, a high efficiency of energy transfer can be obtained.[23] When a quencher methyl viologen (MV) is used in excess to accept the harvested light energy from ZnPP, nearly 100% of the fluorescence is rapidly quenched, consistent with a high efficiency of excitation-energy transfer from the ZnPP wires to the MV.[24-28]

2.2 Experimental section

2.2.1 Materials

Zn(II) Protoporphyrin IX (ZnPP) was purchased from Frontier Scientific. The tCsmA, with N-Terminal Acetylation and C-Terminal Amidation, was purchased from GL Biochem Ltd (98% purity). Dimethyl sulfoxide (DMSO, Sigma Aldrich), monobasic dihydrogen phosphate (Sigma Aldrich), dibasic monohydrogen phosphate (Sigma Aldrich), 1-propanol (Sigma Aldrich) and the microdialysis vial (Pierce™ 96-well Microdialysis Plate, 3.5K MWCO, Thermo Fisher Scientific) were used for constitution of tCsmA-ZnPP

complexes. MV and sodium dithionite (85% purity) were obtained from Sigma-Aldrich.

2.2.2 Equipments

The concentration of protein was measured with a Microvolume DS-11 Spectrophotometer (DeNovix). The steady state Fluorescence spectra were measured using a FLS900 Fluorescence Spectrometer (Edinburgh Instruments) and Aqualog® Water Treatment Plant Analyzer (HORIBA Scientific). The UV-Vis spectra were measured using a UV-1700 Spectrophotometer (Shimadzu). The molar extinction coefficient of the complex was determined by Vista MPX Simultaneous Inductive Coupled Plasma-Optical Emission Spectrometry (ICP-OES, Varian Inc.). The complex was loaded with a 10 µL injection loop on a SuperdexTM 200 Increase 5/150 GL column (GE Healthcare) to perform Size-Exclusion Chromatography (SEC). Negative stain images were obtained with a JEM-1010 transmission electron microscope with a maximum output voltage up to 70 KV equipped with an Olympus Megaview camera (JEOL Ltd.). The circular Dichroism (CD) spectra were measured using a J-815 CD spectrometer (JASCO). The Linear Dichroism (LD) spectra were measured using a Chirascan CD Spectrophotometer (Applied Photophysics). The time-resolved fluorescence measurements were performed with a FluoTime 300 instrument (PicoQuant).

2.2.3 Methods

The peptide was synthesized using standard solid-phase Fmoc based chemistry and purified by RP-HPLC from GL Biochem Ltd (98% purity). To ensure the similar biological activity with natural CsmA, the modification of N-Terminal Acetylation and C-Terminal Amidation were included. The sequence of the peptide corresponds to residues 7-49 of CsmA:
Ac-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG-NH_2

Constitution of the tCsmA-ZnPP complex was carried out in a molar ratio 5:1 mixture of ZnPP and tCsmA dissolved in pH 7.9 50 mM Potassium Phosphate buffer (KPi buffer) containing 6M 1-propanol. The ZnPP was added from a 200 mM DMSO solution. The reaction proceeded in the microdialysis vial against 4 times 1.4 mL KPi dialysis buffer at 4 °C in the dark. The dialysis

times were 2 hours except for the third one, where the sample was incubated overnight. After the last dialysis step, the sample volume was increased from 100 µL to 130 µL and the sample was centrifuged for 15 minutes at 4 °C with 30,000 × *g* to remove unbound CsmA and ZnPP precipitate.

To determine the extinction coefficient, different concentrations of tCsmA-ZnPP complex solutions were prepared. For each aliquot the absorption at 430 nm was measured and the Zn concentration was determined by ICP-OES to calculate the extinction coefficient of the complex. By combining a Bicinchoninic Acid (BCA) assay with the molar extinction coefficient of the complex, the stoichiometry was determined.

To observe the particles by TEM, the tCsmA-ZnPP complex was loaded onto an EM grid with a layer of amorphous carbon film that was rendered hydrophilic by glow discharge (EMS K100X, Agar Scientific) for 1 min at 30 mA. After quick washing with 10 µL distilled water 2 times, 5 µL sample was applied to the grid and left for 30 seconds. Subsequently the sample was stained quickly with 10 µL 2% (w/v) uranyl acetate solution, followed by blotting of excess solution and air-drying.

For Cryo-EM, grids were prepared by placing 4 µL of 200 µM tCsmA-ZnPP complex in KPi buffer onto 200 mesh grids with 2 µm holes (Quantifoil R2/2, Quantifoil Micro Tools, GmbH). Grids were glow discharged for 20 seconds prior to plunge freezing in liquid ethane cooled by liquid nitrogen, using a FEI Vitrobot IV at 100% relative humidity, chamber temperature of 4 °C. Data were collected at 300 kV with a Titan Krios electron microscope, using an electron dose of 51 $e^{-}/Å^{2}$ and a magnification of 75,000×. The final object sampling was therefore 0.88 Å/pixel. Exposures were recorded using the EPU automated acquisition software on a FEI Falcon II direct electron detector. The single particle analysis was performed with Relion 2.0, while Fiji was used for Fourier transforms of 2D class images.

For the time-resolved fluorescence measurements, the samples were prepared with 500 µM MV, 1.25 mM sodium dithionite and tCsmA-ZnPP complex in KPi buffer or the monomeric ZnPP in 6 M propanol buffer. The excitation beam was 440 nm.

For LD experiments, a squeezing gel of 5% acrylamide with 50% glycerol was used (STable 2.2). To avoid fast degradation of the sample by components of the squeezing gel, the procedure was optimized. All materials were cooled on ice before use. To prepare the gel, all components except the sample were mixed, and put at room temperature for 2.5 minutes. The complex was added and the solution was transferred to the gel squeezer. The system was kept on ice for 5 minutes and another 5 minutes at room temperature and immediately the LD was measured.

For steady state fluorescence measurements, the samples were prepared with 1.25 mM sodium dithionite and tCsmA-ZnPP complex in KPi buffer or the monomeric ZnPP in 6 M propanol buffer, and with different concentrations of MV. The excitation wavelength was 430 nm for the tCsmA-ZnPP complex and 416 nm for ZnPP monomer. In addition, steady state fluorescence was measured for both samples with 440 nm excitation light, corresponding with the wavelength used in time-resolved fluorescence experiments.

2.3 Results and discussion

2.3.1 Constitution and characterization

An important finding of our studies is that slow dialysis is key to form a tCsmA-ZnPP complex, since the complex is moderately soluble while the components precipitate upon dialysis. The 1-propanol, which is thought to cover the hydrophobic side of randomly aggregated tCsmA, is slowly removed by dialysis against KPi buffer. This exposes the histidine, which is the only charged residue at the hydrophobic side and is the binding site for coordination to ZnPP. UV-vis spectroscopy indicates histidyl axial binding of the chromophore by characteristic bathochromic shifts around 15 nm relative to the monomeric ZnPP in solution (Figure 2.2A).[15, 29] As a control experiment, we performed dialysis of either the ZnPP or the tCsmA in 50 mM KPi buffer containing 6 M 1-propanol and it was found that both precipitate upon dialysis against KPi buffer (blue trace in Figure 2.2A). This suggests that dialysis should be slow compared to the binding to prevent precipitation of separate components. The B-band shifts by 14 nm, from 416 nm to 430 nm, while the Q(1,0) and Q(0,0) bands shift by 8 and 7 nm to 554 nm and 590 nm, respectively, after constitution (Figure 2.2A). These shifts

are well in line with literature data collected from semi-artificial systems, which consistently show shifts in the range of 10-15 nm for the B-band.[14, 30] After binding, the absorption in the Q(1,0) band of the tCsmA-ZnPP complex is stronger than for the Q(0,0) band, while for the ZnPP monomer the absorption for the Q(1,0) band is less than for the Q(0,0) band. The relative intensities of these bands correlate with the stability of the complex, and the decrease of the Q(0,0) band relative to the Q(1,0) band upon formation of the complex indicates that binding enhances the stability of the complex with a square-planar Zn^{2+} in the porphyrin.[31-32]

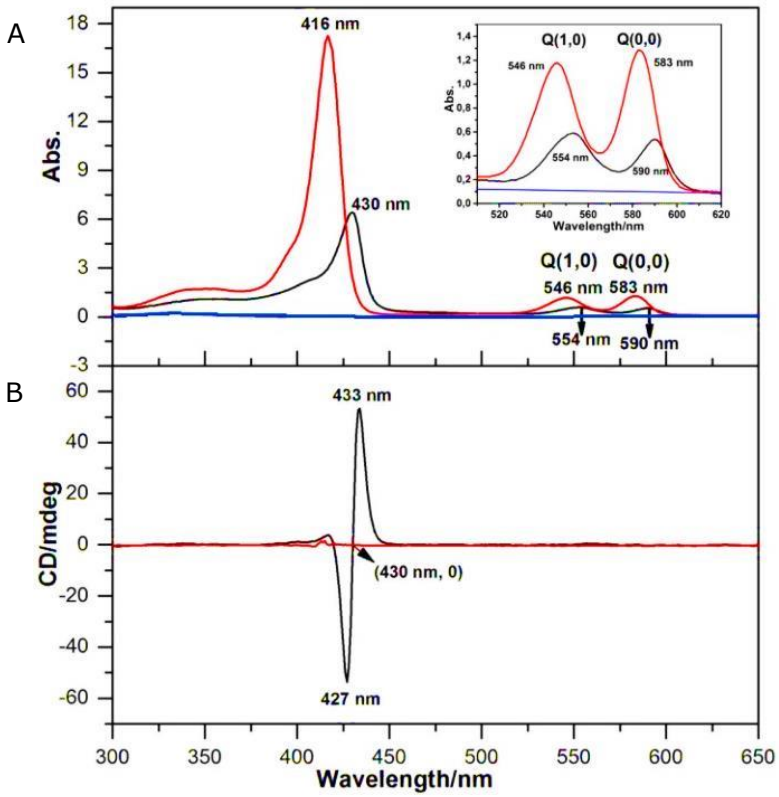


Figure 2.2 (A) UV-Vis and (B) corresponding CD spectra, collected from 50 µM tCsmA-ZnPP complex in 50 mM KPi buffer (black) and 50 µM ZnPP in 50 mM KPi buffer containing 6 M 1-propanol (red), respectively. The blue line shown in panel A is the residual UV response from KPi buffer after dialysis, *i.e.* without tCsmA, confirming that ZnPP has precipitated. The window of the enlarged Q bands region is shown in panel A. The concentration of peptide at the start of the dialysis is 50 mM in KPi buffer with 6 M 1-propanol and there is no CD response (B, red trace). Upon complex formation a CD response is detected (B, black trace).

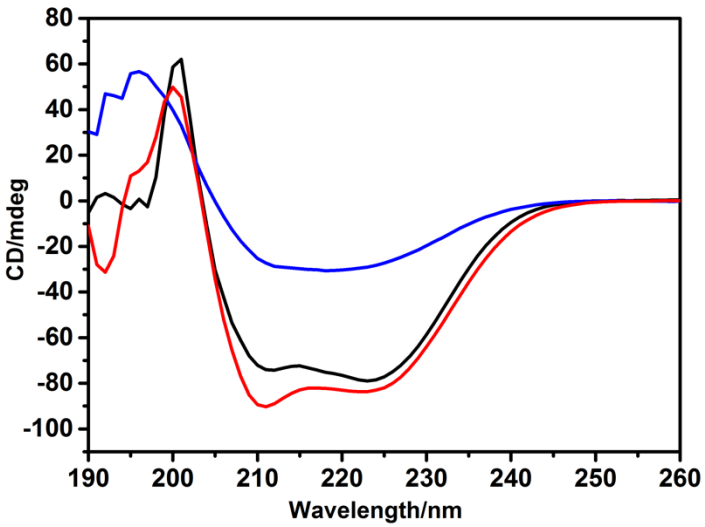


Figure 2.3 CD spectra. black: 50 µM tCsmA-ZnPP complex in pH 7.9 50 mM KPi buffer; red: 50 µM tCsmA in pH 7.9 6 M 1-propanol 50 mM KPi buffer. Both CD signals are characteristic for a helical secondary structure. In contrast, while the tCsmA possesses a maximum 50 µM solubility in MQ water, it shows a denatured β-sheet structure, and the binding cannot proceed in this condition (blue line). Samples were measured in quartz cuvettes with a 2 mm path length. Spectra were recorded from 260 to 190 nm at 1 nm intervals with a 1 nm bandwidth at 25 °C.

Moreover, the CD spectrum in Figure 2.3 for the secondary structure reveals the maintenance of a helical structure only when the ZnPP is bound. Therefore the slow dialysis is vitally important for the self-assembly of an intact helix structure with tCsmA bound to ZnPP.

Table 2.1 The concentration of zinc measured by ICP-OES with corresponding UV absorption at 430 nm

Zn (C,µM) Measured by ICP- OES	Complex Abs. at 430 nm
0.06295	0.070
0.82441	0.103
1.08106	0.147
1.30101	0.166
1.66412	0.210
2.02203	0.248
2.29428	0.266
2.59254	0.295
2.85408	0.322
3.12787	0.362

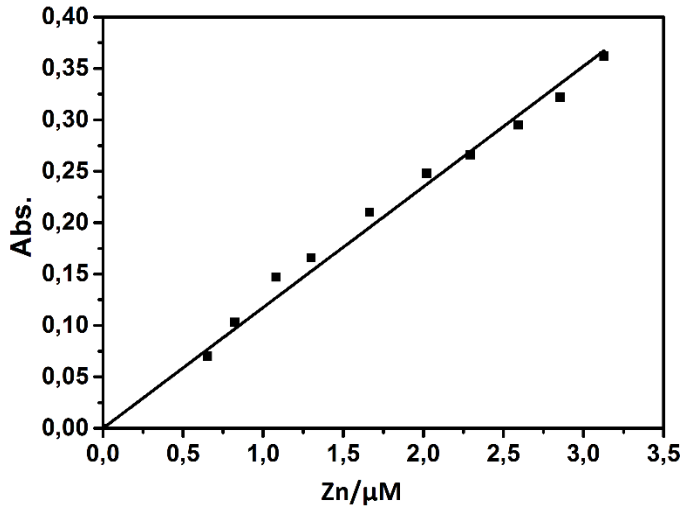


Figure 2.4 Standard curve represents the relationship between the Zn concentration in CsmA-ZnPP determined by ICP-OES and its absorbance maximum (λ = 430 nm). The resulting functional equation is Y = 0.1174X; R^2 = 0.9975.

The molar extinction coefficient ε can be determined using equation: ε = A / (c * l), which was converted based on the Beer-Lambert law, A = ε *c * l, where A represents an actual absorption at the given wavelength, path length l is 1 cm in our study, and the molar concentration c was obtained from the Zn concentration (C_{Zn}, µM) determined by ICP-OES. The molar extinction coefficient of the ZnPP in complex with tCsmA is 117400 M^{-1} cm^{-1} and was determined by the combined use of UV-Vis and ICP-OES (Table 2.1 & Figure 2.4).[34] With a subsequent BCA assay, a stoichiometry of tCsmA: ZnPP ~ 1 : 1 was determined (Table 2.2).

Table 2.2 The amount of protein in the complex was determined using the colorimetric BCA protein assay. The stoichiometry of complex can be calculated from the ratio of the protein concentration to the concentration of Zn.

ZnPP (C, µM) Measured from A430 nm	**tCsmA (C, µM) Measured by BCA assay**	**Ratio C_{ZnPP} / C_{tCsmA}**
13.8	13.6	1.01 ± 0.01
8.5	8.0	1.06 ± 0.02
3.3	3.5	0.94 ± 0.04

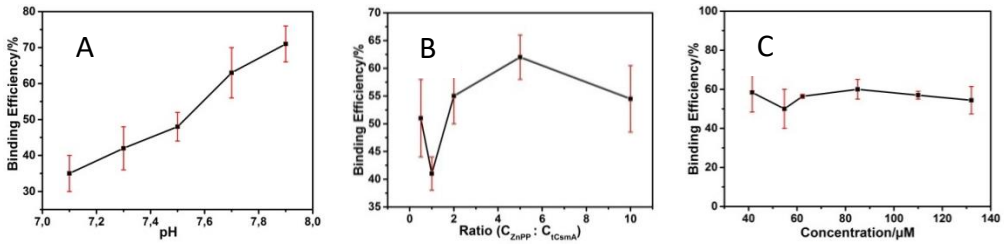


Figure 2.5 Optimization of the binding between CsmA and ZnPP during dialysis. A: The influence of the pH of KPi buffer (C_{tCsmA}=70 µM, C_{ZnPP}=350 µM). B: The influence of the ratio of CsmA and ZnPP in KPi pH 7.9 buffer (C_{tCsmA}=50 µM). C: The influence of the tCsmA concentration in KPi pH 7.9 buffer (C_{tCsmA} : C_{ZnPP}=1 : 5).

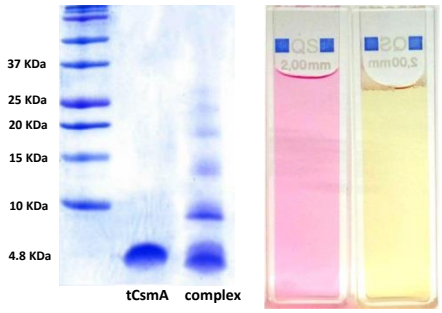


Figure 2.6 left: Tricine-SDS gel analysis of the complex and tCsmA. Prior to loading the samples were boiled for 4 min. Lane 1: Marker. Lane 2: tCsmA. Lane 3: the ZnPP-CsmA complex. right: cuvettes containing 50 µM ZnPP without tCsmA mixture in 6 M 1-propanol KPi buffer (left cuvette) and 50 µM tCsmA-ZnPP complex in KPi buffer (right cuvette).

At pH=7.9 and with five times molar excess of ZnPP relative to the concentration of peptide, the binding optimally proceeds, with an efficiency of ~ 60-70% over a peptide concentration range of 40-130 µM (Figure 2.5). UV-Vis spectra showed a narrow singlet peak at 430 nm, which indicates high purity. The concentration of protein has little influence on the binding efficiency. Similar to the natural CsmA-BChl *a* complex, the tCsmA-ZnPP complex shows oligomer bands on the SDS gel (Figure 2.6).[33] The complex is stable during three months storage in the dark at 4°C.

2.3.2 Structural information

Under cryo-conditions of TEM the images show long filaments (Figure 3A). We performed maximum likelihood 2D classification of 99142 particles using

a mask size of 535 Å, corresponding with a box size of 610 pixels. All 50 2D class averages show the same chain-like structure (SFigure 2.2).[17]

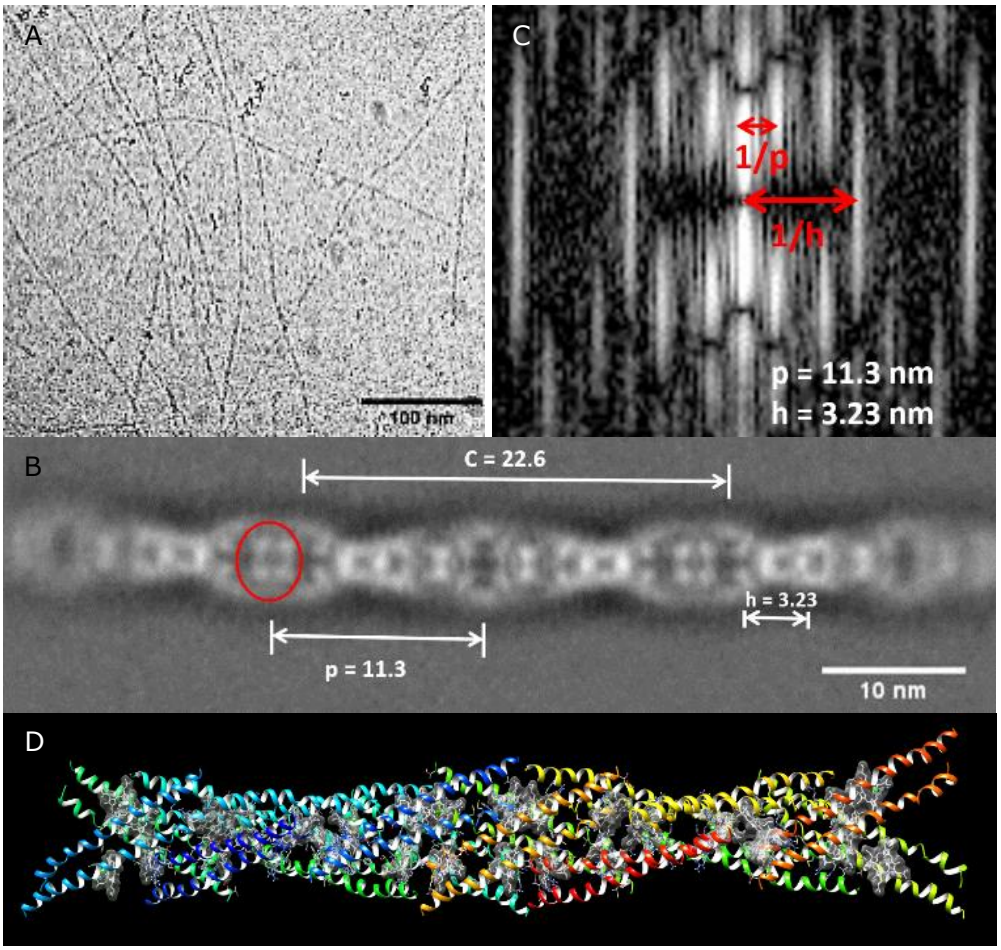


Figure 2.7 (A) cryo-EM image showing individual filaments in a sample of 200 µM tCsmA-ZnPP complex in KPi buffer. (B) 2D class average of particles from cryo-EM images. (C) Fast Fourier Transform of a 2D class average (Figure 3B). (D) Proposed assembly model according to the parameters from the 2D class average. Each subunit has a different colour and the ZnPP are depicted in light grey. Due to the image processing, there is a contrast inversion from A to B.

The diameter of a filament is about 5 nm (Figure 2.7B). As the Fast Fourier Transform (FFT) Figure 2.7C shows, multiple layer lines separated by 1/22.6 nm^{-1} provide converging and convincing evidence for a helical structure with a C= 22.6 ± 0.02 nm periodic repeat. The strongest peaks close to the meridian are on the second layer line, corresponding with a helix pitch p=11.3 ± 0.02 nm, *i.e.* half of the repeat (Figure 2.7C). The seventh layer line

with intensity on the meridian reveals a helical rise h=3.23 ± 0.02 nm, and according to the FFT the complex forms a 7_2 helix. A proposed assembly is shown in Figure 2.7D. The high density parts inside the filaments can be attributed to tetrameric units of tightly packed ZnPP (Figure 2.7B, labelled with the red circle). The size of the ZnPP clusters is ~ 1.2 nm, and they are separated by ~2 nm. In addition, most likely the system self-assembles as a symmetric double helix.

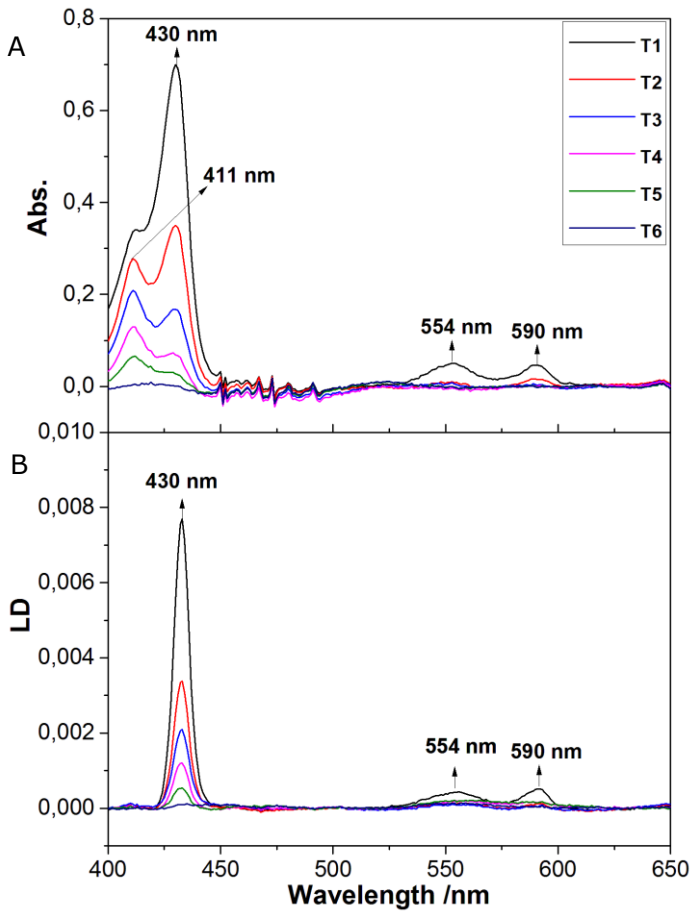


Figure 2.8 (A) UV-Vis and (B) corresponding LD spectra. The spectrum T1 corresponds to the first measurement, T2 was collected after 7 min, T3 after 13 min, T4 after 19 min, T5 after 30 min and T6 after 1 h 14 min.

Linear Dichroism (LD) allows to determine the macroscopic orientation from the absorption of parallel and perpendicular linearly polarized light, and to study the orientation of the ZnPP in the filaments, we collected LD data along with simultaneous UV-Vis absorbance spectra in a squeezing gel.

According to the UV-Vis results (Figure 2.8A), an additional 411 nm ZnPP peak appears, in contrast with the UV-Vis data collected from the complex in solution (black trace in Figure 2.2A). The shift to 411 nm is attributed to loss of binding and decomposition by squeezing gel components (Table 2.1, SFigure 2.4). The LD spectra measured over time show positive peaks at 430 nm with the two Q band signals at 554 nm and 590 nm, due to vertically oriented ZnPP-tCsmA filaments, along the squeezing direction of the gel (Figure 2.8B), which indicates a predominantly parallel orientation of ZnPP planes to tCsmA scaffolds as well.[36-38] This is in line with the results of the 2D class average of filaments (Figure 2.8B). The lack of orientation for the 411 nm component is consistent with dissociation of the complex, which leads to a randomly oriented ZnPP fraction. The LD confirms that the binding of ZnPP to the tCsmA is necessary to build the supramolecular photosynthetic system.

Compared to the CD response of the monomeric ZnPP in propanol buffer, the B band of the complex at 430 nm shows a high intensity and nearly symmetric Cotton effect (Figure 2.2B). The strong CD signal observed is from ZnPP motifs interacting with each other and forming a chiral exciton coupled unit, with dipole-dipole interactions between monomers, which is well in line with Figure 2.7B, where the ZnPP appears to form tetrameric repeat units. In contrast, while the optical response for the ZnPP in propanol buffer in Figure 2.2A is stronger than for the complex, the CD is virtually absent, which confirms that the ZnPP in solution is monomeric. On the other hand, the CD response for the Q bands is very weak in the CD data collected from the complex. Since the LD data in Figure 2.7 are in line with a specific orientation angle of the ZnPP relative to the tCsmA scaffolds, and the particle shows a helical structure, the vanished dipole-dipole interaction of the Q bands indicates that the ZnPP possibly form oligomers with exciton coupling at a specific rotation angle in the molecular system to offset the dipole-dipole interaction of the Q bands.[39-41] Based on LD and TEM data, the ZnPP-CsmA complex probably forms a helical structure that is composed of twisted peptide assemblies built from well-ordered chiral oligomeric units that give rise to a single CD response (Figure 2.2B).[17, 41-44]

2.3.3 Photostability

The tCsmA-ZnPP complex forms large soluble particles that show a single distribution peak around 1.07 mL in the SEC data, with the retention

indicating a large molecular weight (Figure 2.9). This is consistent with the formation of extended filaments shown in the TEM data in Figure 2.7A. In addition, SEC also shows that the complex is quite sensitive to light. The peak shifts to ~ 1.09 mL after 30 minutes of light exposure, which indicates that the particle size became smaller upon exposure to light. Moreover, steady state fluorescence measurements indicate slow photo-dissociation of the complex upon illumination (SFigure 2.3). The histidyl axial binding of the ZnPP appears unstable under light exposure.[28] While free ZnPP precipitates in buffer, in the gel the ZnPP can be dissolved and in the LD data in Figure 4A a pronounced free ZnPP signal is observed at 411 nm wavelength, indicative of partial dissociation in the gel. Both the 430 nm and the 411 nm response decrease upon illumination, indicating photo-degradation for prolonged exposure, but the relative intensity changes in favor of the free ZnPP. With more than 2 h exposure with sunlight or 430 nm UV light, the complex was totally dissociated. The reason for the photo-dissociation might be that the dissipation of heat is slow.[15]

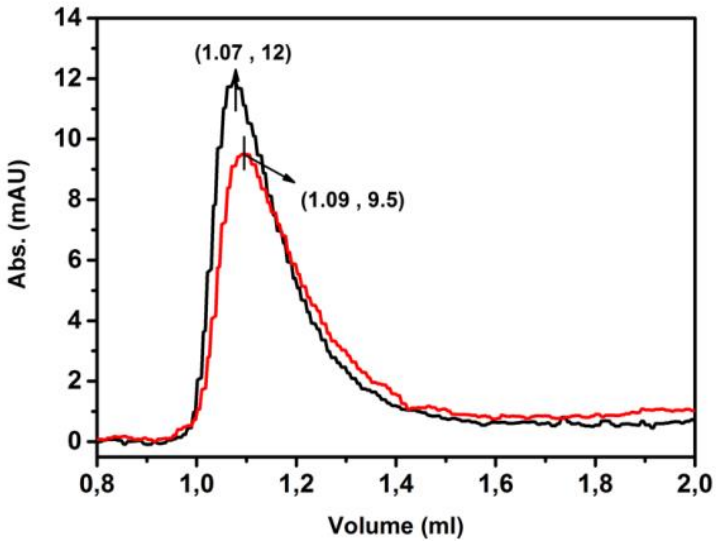


Figure 2.9 SEC elution spectra collected from the 50 µM tCsmA-ZnPP complex in pH 7.9, 50 mM KPi buffer (black trace) and the same sample after 30 minutes light exposure (red trace).

Despite photo-dissociation on longer time scales, the tCsmA-ZnPP complex is able to maintain a short time photostability. As Figure 2.10 shows, the stability is related to the concentration. At a concentration of about 20 µM the complex is stable for at least 30 minutes (Figure 2.10A). It has recently

been proposed that light-driven conformational changes in chiral protein scaffolds can develop vibronic states within or between linear arrays.[13] In such systems, excitons spread rapidly over the system, which may help to decrease the risk of damage and dissociation rate. However, during long time exposure the light still damages the complex which results in size decrease of the filament. Apparently a high concentration improves the stability, which suggests that distribution of energy helps to dissipate the excess energy and may be of vital importance for further improvement of the stability.

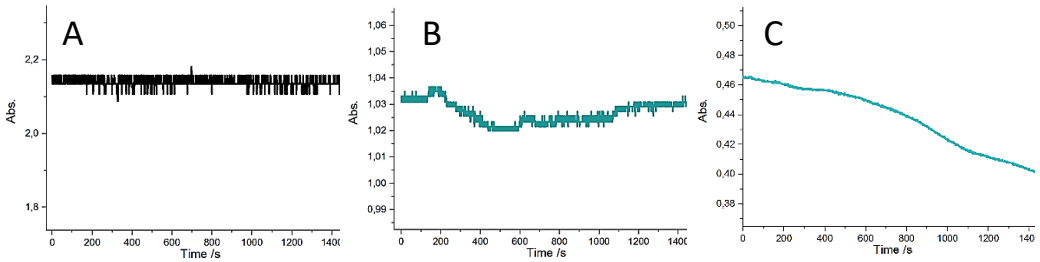


Figure 2.10 The photostability test by UV-Vis spectroscopy of the ZnPP-tCsmA complex with different concentration. A : 18 µM; B : 9 µM; C : 4.5 µM.

2.3.4 Exciton-energy transfer

To address further the energy transfer potential of the tCsmA-ZnPP self-assembly, an acceptor MV (SFigure 2.1) was added.[27-28, 45] Generally, MV exists in the dicationic form MV^{2+} (SFigure 2.1) in aqueous solution. However, MV has highly reversible redox reactions, at the reversible potential of E_0 = -0.4 ~ -0.55 V and -0.75 ~ -0.9 V vs NHE in aqueous solution.[26, 46-47] A solution of MV with sodium dithionite contains a mixture of MV^{+} and MV^{2+} (SFigure 2.5). When MV^{2+} is added to a suspension of tCsmA-ZnPP filaments in KPi buffer or the monomeric ZnPP in propanol buffer, fluorescence quenching is hardly observed (SFigure 2.6). This changes when the acceptor is converted to the cation radical form MV^{+} with excess sodium dithionite (SFigure 2.8).[48-51] As the data show, the tCsmA-ZnPP complex contains 594 nm and 650 nm emission peaks and the monomeric ZnPP contains 590 nm and 645 nm emission peaks after excitation (SFigure 2.6). On the other hand, the MV^{+} shows a strong broad absorption at 606 nm (ε = 13700 M^{-1} cm^{-1}), which overlaps with the tCsmA-ZnPP emission (SFigure 2.7A).[26] This leads us to conclude that intermolecular FRET or Dexter energy transfer between ZnPP and MV+ may cause fluorescence quenching by energy transfer. The

quenching is dependent on the concentration of MV^+. When the concentration was increased to about 300 µM MV^+ in the tCsmA-ZnPP solution, nearly all fluorescence of ZnPP was quenched (Figure 2.11).

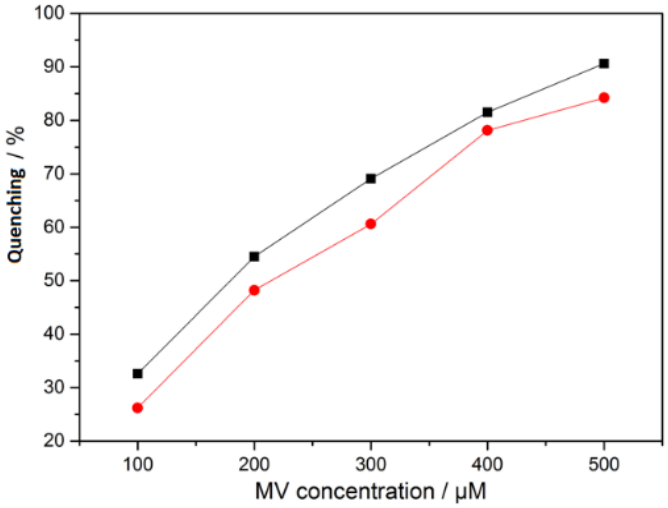


Figure 2.11 The fluorescence quenching diagram for 3.5 µM tCsmA-ZnPP complex at 594 nm (Black trace) and 650 nm, (Red trace) with different concentration of MV. The solution contains 1.25 mM sodium dithionite. The excitation wavelength is 440 nm.

Table 2.3 Fitted Time constants of Fluorescence Transients in tCsmA-ZnPP complex solution with different additives under a 440 nm excitation beam.

Additives	**594 nm Lifetime, ns**			
Lifetime	τ_1	τ_2	τ_3	$\tau_{average}$
No MV and $Na_2S_2O_4$	1.63 (9%)	0.5 (32%)	0.11 (59%)	0.37
MV^{2+} (no $Na_2S_2O_4$)	1.62 (8%)	0.63 (44%)	0.15 (48%)	0.48
MV^+ (excess $Na_2S_2O_4$)	1.1 (0.03%)	0.28 (0.12%)	0.001 (99.85%)	0.0017

Additives	**650 nm Lifetime, ns**			
Lifetime	τ_1	τ_2	τ_3	$\tau_{average}$
No MV and $Na_2S_2O_4$	3.13 (5%)	0.83 (29%)	0.14 (66%)	0.49
MV^{2+} (no $Na_2S_2O_4$)	2.82 (5%)	0.88 (44%)	0.21 (51%)	0.64
MV^+ (excess $Na_2S_2O_4$)	1.3 (0.01%)	0.4 (0.03%)	0.0015 (99.96%)	0.0017

As Table 2.3 shows, the exposure to the MV^+ acceptor reveals excellent energy transfer efficiency from the complex to the MV^+. The fluorescence decay responses at 594 nm and 650 nm can be described with 3 different lifetime components. All three lifetimes of the complex with the MV^+ additive are shortened, especially τ_3 was dramatically shortened to ~1 ps and the energy is almost 100% quenched by the acceptor. Although the time-resolved fluorescence spectroscopy has a limited resolution of 25 ps, τ_3 is about a few picoseconds or even shorter, at the femtosecond time scale. When sodium dithionite was depleted by letting it react with oxygen, the MV^+ was converted to MV^{2+} and there is no fluorescence quenching. The quenched complex recovered the same fluorescence as without acceptor, indicating that it maintains its integrity upon photoexcitation (SFigure 2.7). This highlights the importance of the MV^+ as the quenching species.

2.4 Conclusions

In conclusion, we have employed the dialysis method to gain control over metalloporphyrin forming an organized array with truncated CsmA peptides to develop helical scaffolds. The tCsmA-ZnPP complex forms soluble and stable long-aspect-ratio filaments to encapsulate strongly interacting oligomeric metalloporphyrins for photoexcitation. The solubility of the complex is good compared to the tCsmA starting material, which suggests that the amphipathic tCsmA has its hydrophobic residues on the inside and the hydrophilic residues on the outside. The tCsmA-ZnPP complex obtained by dialysis is well defined, and shows high purity. The single histidine available for binding with metalloporphyrin allows the material to get a stable 1:1 chromophore/peptide binding ratio. The self-assembly of the tCsmA-ZnPP complex shows certain photostability. The photoinduced energy of ZnPP-tCsmA complex can be transferred within few picoseconds with a nearly 100% quenching to MV^+. This artificial antenna system paves the way for the design of semisynthetic building blocks for light energy harvesting in artificial photosynthesis.

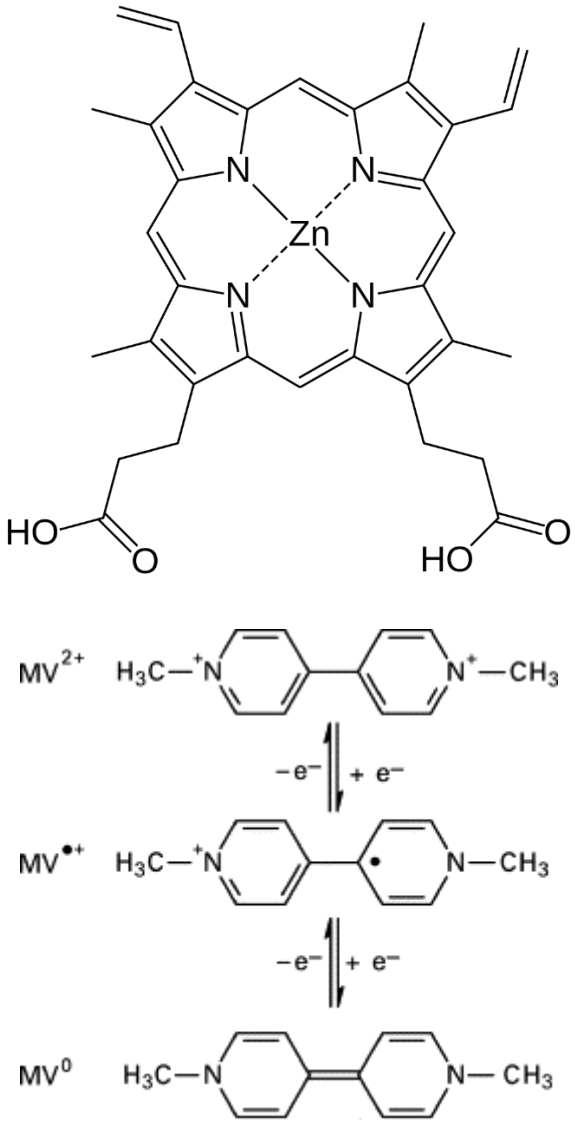
2.5 References

1. Chin, A. W.; Datta, A.; Caruso, F.; Huelga, S. F.; Plenio, M. B., *New Journal of Physics* **2010,** *12* (6), 065002.
2. Chin, A.; Prior, J.; Rosenbach, R.; Caycedo-Soler, F.; Huelga, S.; Plenio, M. B., *Nature Physics* **2013,** *9* (2), 113.
3. Huelga, S. F.; Plenio, M. B., *Contemporary Physics* **2013,** *54* (4), 181-207.
4. Pšenčík, J.; Ikonen, T.; Laurinmäki, P.; Merckel, M.; Butcher, S.; Serimaa, R.; Tuma, R., *Biophysical journal* **2004,** *87* (2), 1165-1172.
5. Adir, N., *Photosynthesis Research* **2005,** *85* (1), 15-32.
6. Van Grondelle, R., *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics* **1985,** *811* (2), 147-195.
7. van Grondelle, R.; Dekker, J. P.; Gillbro, T.; Sundstrom, V., *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1994,** *1187* (1), 1-65.
8. McConnell, I.; Li, G.; Brudvig, G. W., *Chemistry & biology* **2010,** *17* (5), 434-447.
9. Purchase, R.; de Groot, H., *Interface Focus* **2015,** *5* (3), 20150014.
10. Mathies, R. A., *Nature chemistry* **2015,** *7* (12), 945.
11. Yin, X.; Struik, P. C., *Journal of experimental botany* **2015,** *66* (21), 6535-6549.
12. Hoffman, D. P.; Mathies, R. A., *Accounts of chemical research* **2016,** *49* (4), 616-625.
13. R. Purchase, R. Cogdell, F. Breitling, V. Stadler, N. van Hulst, G.-J. Kramer, A. Ramirez, R. Zwijnenber, A. Kallergi (), J. B. de Baan, I. Rudra, and H. J. M de Groot *Bioinspired Chemistry* **2019**, 47-69.
14. Fry, H. C.; Garcia, J. M.; Medina, M. J.; Ricoy, U. M.; Gosztola, D. J.; Nikiforov, M. P.; Palmer, L. C.; Stupp, S. I., *J Am Chem Soc* **2012,** *134* (36), 14646-14649.
15. Luo, L.; Chang, C. H.; Chen, Y. C.; Wu, T. K.; Diau, E. W., *J Phys Chem B* **2007,** *111* (26), 7656-64.
16. Chromiński, M.; ó Proinsias, K.; Martin, E.; Gryko, D., *European journal of organic chemistry* **2013,** *2013* (8), 1530-1537.
17. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.
18. Milks, K. J.; Danielsen, M.; Persson, S.; Jensen, O. N.; Cox, R. P.; Miller, M., *Photosynthesis research* **2005,** *86* (1-2), 113-121.
19. Zobova, A.; Taisova, A.; Lukashev, E.; Fedorova, N.; Baratova, L.; Fetisova, Z., *Journal of Biophysics* **2011,** *2011*.
20. Kulminskaya, N. V.; Pedersen, M. O.; Bjerring, M.; Underhaug, J.; Miller, M.; Frigaard, N. U.; Nielsen, J. T.; Nielsen, N. C., *Angew Chem Int Ed Engl* **2012,** *51* (28), 6891-5.
21. Zhou, Q.; Swager, T. M., *J Am Chem Soc* **1995,** *117* (50), 12593-12602.
22. Leonard, J. J.; Yonetani, T.; Callis, J. B., *Biochemistry* **1974,** *13* (7), 1460-1464.

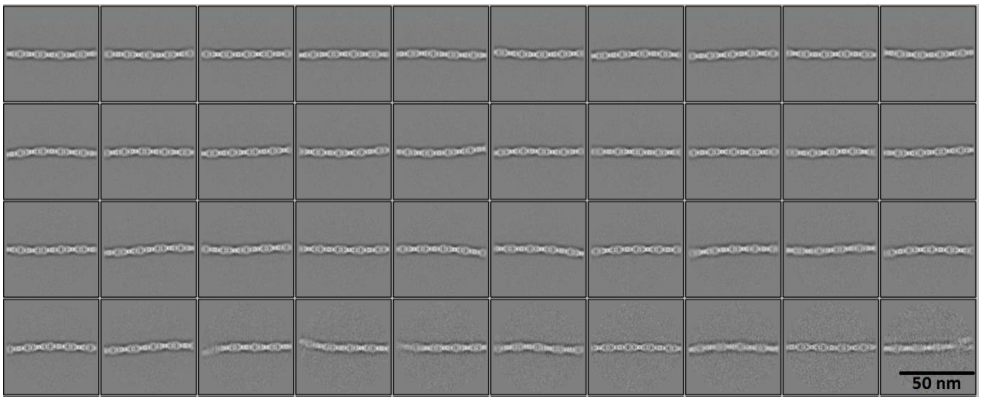
23. Cogdell, R. J.; Gardiner, A. T.; Cronin, L., *Phil. Trans. R. Soc. A* **2012,** *370* (1972), 3819-3826.
24. Summers, L. A., *The bipyridinium herbicides*. Academic Press Inc.: 1980.
25. Bus, J. S.; Gibson, J. E., *Environmental health perspectives* **1984,** *55*, 37.
26. Watanabe, T.; Honda, K., *The Journal of Physical Chemistry* **1982,** *86* (14), 2617-2619.
27. Proux-Delrouyre, V.; Demaille, C.; Leibl, W.; Sétif, P.; Bottin, H.; Bourdillon, C., *J Am Chem Soc* **2003,** *125* (45), 13686-13692.
28. Komatsu, T.; Wang, R.-M.; Zunszain, P. A.; Curry, S.; Tsuchida, E., *J Am Chem Soc* **2006,** *128* (50), 16297-16301.
29. Kamlet, M. J.; Taft, R., *J Am Chem Soc* **1976,** *98* (2), 377-383.
30. Kodali, G.; Mancini, J. A.; Solomon, L. A.; Episova, T. V.; Roach, N.; Hobbs, C. J.; Wagner, P.; Mass, O. A.; Aravindu, K.; Barnsley, J. E., *Chemical science* **2017,** *8* (1), 316-324.
31. Milgrom, L. R.; Warren, M. J., **1997**.
32. Giovannetti, R., The use of Spectrophotometry UV-Vis for the Study of Porphyrins. In *Macro To Nano Spectroscopy*, InTech: 2012.
33. Montano, G. A.; Wu, H.-M.; Lin, S.; Brune, D. C.; Blankenship, R. E., *Biochemistry* **2003,** *42* (34), 10246-10251.
34. Li, Y.; Scales, N.; Blankenship, R. E.; Willows, R. D.; Chen, M., *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2012,** *1817* (8), 1292-1298.
35. He, S.; Scheres, S. H., *Journal of structural biology* **2017,** *198* (3), 163-176.
36. Furumaki, S.; Vacha, F.; Habuchi, S.; Tsukatani, Y.; Bryant, D. A.; Vacha, M., *J Am Chem Soc* **2011,** *133* (17), 6703-6710.
37. Abdourakhmanov, I. A.; Ganago, A. O.; Erokhin, Y. E.; Solov'ev, A. A.; Chugunov, V. A., *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1979,** *546* (1), 183-186.
38. Van Amerongen, H.; Vasmel, H.; Van Grondelle, R., *Biophysical journal* **1988,** *54* (1), 65-76.
39. Agranovich, V.; Galanin, M., *Electronic Excitation Energy Transfer in Condensed Matter* **1982**, 81-170.
40. Linnanto, J. M.; Korppi-Tommola, J. E., *The Journal of Physical Chemistry B* **2013,** *117* (38), 11144-11161.
41. Farid, T. A.; Kodali, G.; Solomon, L. A.; Lichtenstein, B. R.; Sheehan, M. M.; Fry, B. A.; Bialas, C.; Ennist, N. M.; Siedlecki, J. A.; Zhao, Z., *Nature chemical biology* **2013,** *9* (12), 826.
42. Pescitelli, G.; Gabriel, S.; Wang, Y.; Fleischhauer, J.; Woody, R. W.; Berova, N., *J Am Chem Soc* **2003,** *125* (25), 7613-7628.
43. Schellman, J. A., *Accounts of chemical research* **1968,** *1* (5), 144-151.
44. Harada, N.; Nakanishi, K., *Accounts of Chemical Research* **1972,** *5* (8), 257-263.
45. Tyson, D. S.; Luman, C. R.; Castellano, F. N., *Inorganic chemistry* **2002,** *41* (13), 3578-3586.
46. Steckhan, E.; Kuwana, T., *Berichte der Bunsengesellschaft für physikalische Chemie* **1974,** *78* (3), 253-259.

47. Bird, C.; Kuhn, A., *Chemical Society Reviews* **1981,** *10* (1), 49-82.
48. Kosower, E. M.; Cotter, J. L., *J Am Chem Soc* **1964,** *86* (24), 5524-5527.
49. Senn, D. R.; Carr, P. W.; Klatt, L. N., *Analytical biochemistry* **1976,** *75* (2), 464-471.
50. Bockman, T.; Kochi, J., *The Journal of Organic Chemistry* **1990,** *55* (13), 4127-4135.
51. MAYHEW, S. G., *European journal of biochemistry* **1978,** *85* (2), 535-547.

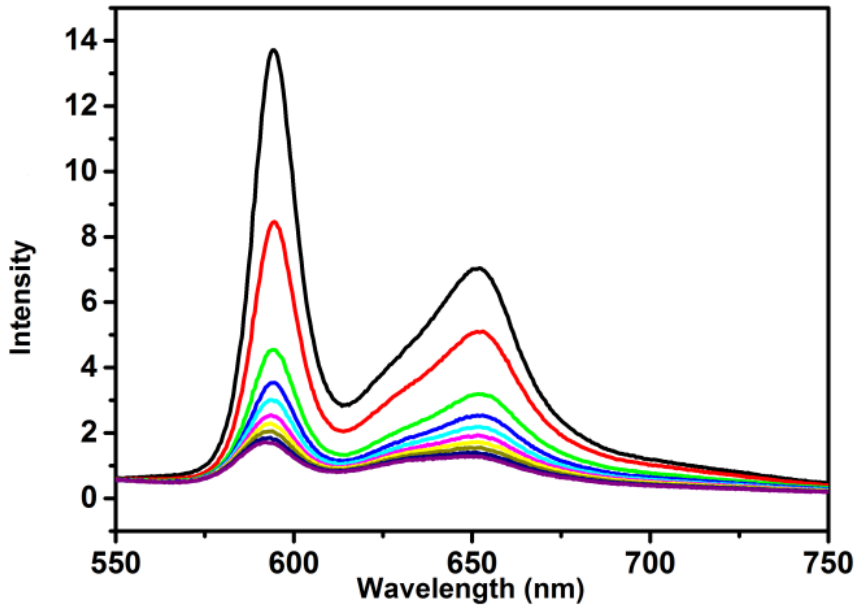
2.6 Appendix



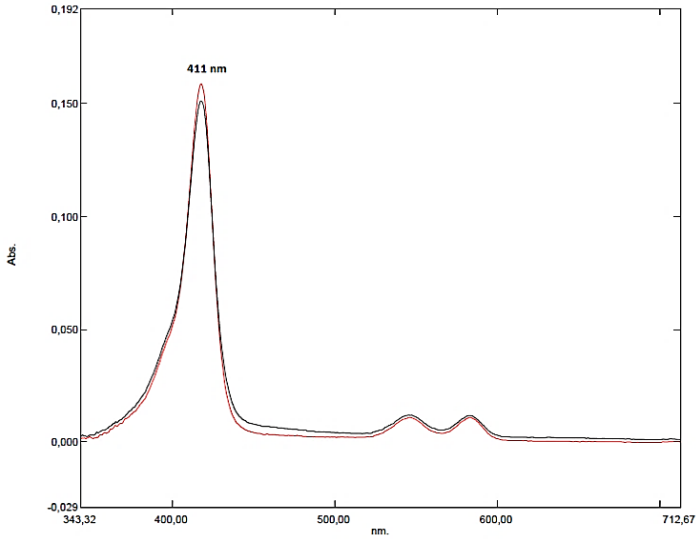
SFigure 2.1 The skeletal formula of zinc protoporphyrin and methyl viologen with different charge.



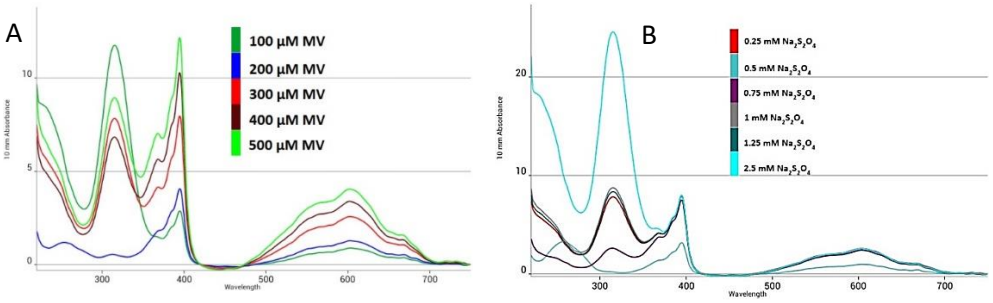
SFigure 2.2 50 classes 2D classification from 994142 particles by RELION2.0.



SFigure 2.3 Fluorescence spectroscopy of the tCsmA-ZnPP complex. The complex was excited at 430 nm for 5 minutes and followed 15 minutes incubation time. The concentration of the complex is 1.6 µM in pH 7.9 50 mM KPi buffer.



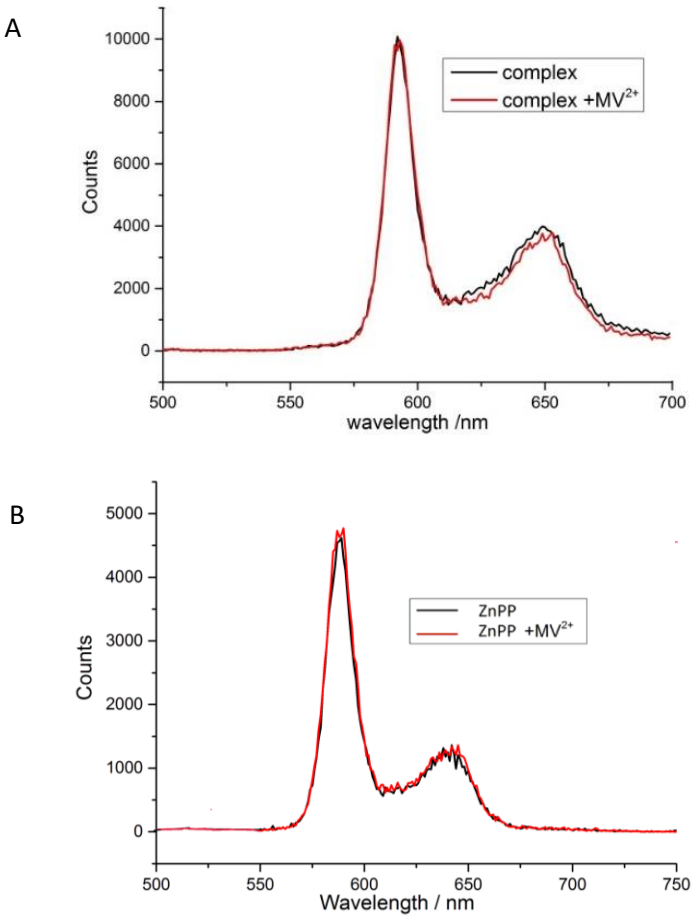
SFigure 2.4 UV-Vis spectra of the ZnPP-tCsmA complex after two hours in squeezing gel (black trace) and pure ZnPP (red trace) in squeezing gel. For both samples the major response is at 411 nm.



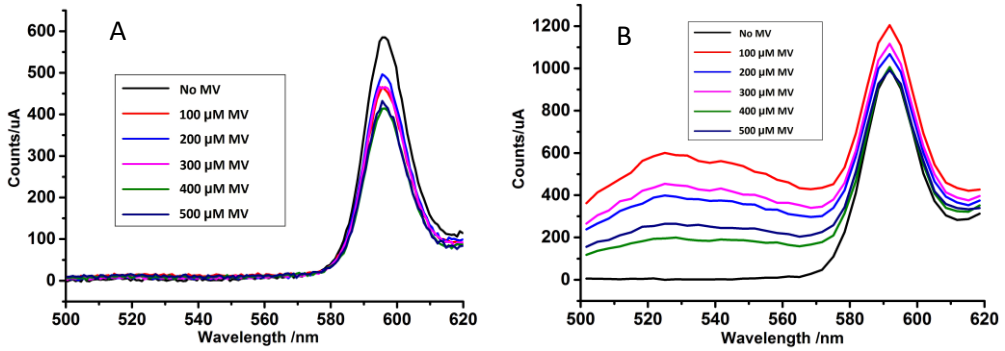
SFigure 2.5 A: the UV spectra of MV with different concentration with 1.25 mM sodium dithionite in the solution. The ratio of A396 nm/A606 nm is about 3 which is same as the ratio in spectra of MV^+. The peak at 606 nm is attributed to MV^+, whereas the peak 396 nm can be from MV^+ and MV^0. That means the sodium dithionite only converts MV to MV^+. If there is MV^0, the MV^0 will contribute to the UV absorption at 396 nm. In this case, the ratio A396 nm/A606 nm should be larger than 3. This reaction was also reported in the literature.[45] B: the UV spectra of 300 µM MV with different amounts of sodium dithionite. Adding more excess sodium dithionite does not improve the conversion rate. The MV can be saturated at about 60% conversion rate.

STable 2.1 The conversion rate of MV^+ (The concentration of sodium dithionite is 1.25 mM, ε606 nm of MV^+ is 13700 M^{-1} cm^{-1}), the data is from the results of SFigure 2.5B.

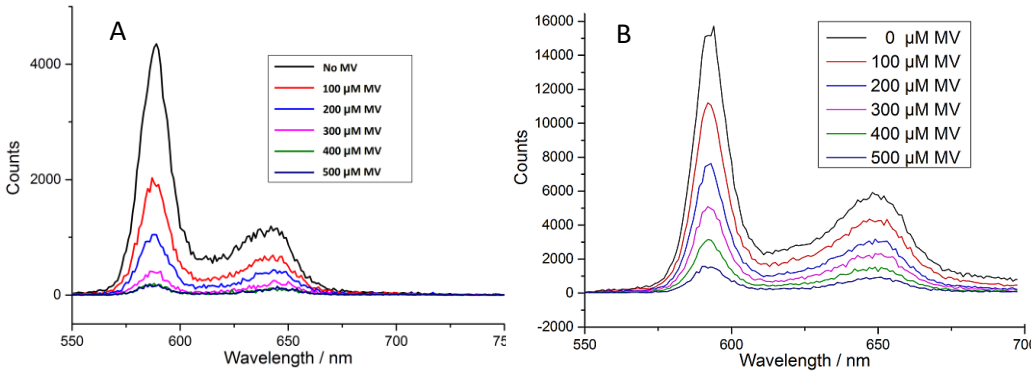
MV (µM)	Abs. 606 nm	MV^+ (µM)	MV^+/MV (%)	A396 nm/A 606
100	0.88	64	64	3.1
200	1.27	93	47	3.09
300	2.56	187	62	3.07
400	3.35	245	61	3.03
500	4.02	293	59	3.02



SFigure 2.6 A: Steady state fluorescence spectra of 4.5 µM tCsmA-ZnPP complex and 4.5 µM tCsmA-ZnPP complex with 500 µM MV in KPi buffer with 430 nm excitation light. B: Steady state fluorescence spectra of 10 µM ZnPP and 10 µM ZnPP with 500 µM MV in 6M 1-propanol KPi buffer with 416 nm excitation light.



SFigure 2.7 A: Steady state fluorescence spectra of 3.5 µM tCsmA-ZnPP complex with different concentration of MV and the depleted 1.25 mM sodium dithionite in KPi buffer with 440 nm excitation light. B: Steady state fluorescence spectra of 3.5 µM ZnPP with different concentration of MV and depleted 1.25 mM sodium dithionite in 6M 1-propanol KPi buffer with 440 nm excitation light.



SFigure 2.8 A: Steady state fluorescence spectra of 3.5 µM tCsmA-ZnPP complex with different concentration of MV and 1.25 mM sodium dithionite in KPi buffer with 430 nm excitation light. B: Steady state fluorescence spectra of 10 µM ZnPP with different concentration of MV and 1.25 mM sodium dithionite in 6M 1-propanol KPi buffer with 416 nm excitation light.

STable 2.2 Preparation of the squeezing gel

Components	Volume (µL)
Glycerol (100%)	1500
37% (29:1) acrylamide	405
APS (10%)	30
TEMED	4.5
Complex (50 µM)	500
KPi buffer (3M, pH 7.9)	41.5
MQ water	519
Total	3000

STable 2.3 Fitted Time constants of Fluorescence Transients in monomer ZnPP solution with different additives under a 440 nm excitation beam.

Additives	590 nm Lifetime, ns			
Lifetime	τ_1	τ_2	τ_3	$\tau_{average}$
No MV and $Na_2S_2O_4$	2.767 (1.2%)	2.11 (98.8%)	-	2.12
MV^{2+} (no $Na_2S_2O_4$)	13.4 (0.1%)	2.12 (99.9%)	-	2.13
MV^{+} (excess $Na_2S_2O_4$)	22.65 (0.02%)	1.98 (99.98%)	-	1.98

Additives	645 nm Lifetime, ns			
Lifetime	τ_1	τ_2	τ_3	$\tau_{average}$
No MV and $Na_2S_2O_4$	2.94 (0.2%)	2.12 (99.8%)	-	2.12
MV^{2+} (no $Na_2S_2O_4$)	17.03 (0.02%)	2.12 (99.98%)	-	2.12
MV^{+} (excess $Na_2S_2O_4$)	10.94 (0.01%)	1.95 (99.99%)	-	1.95

Chapter 3

Contrasting modes of modulated self-assembly for CsmA homologues coordinated to ZnPP

3.0 Abstract

The self-assembly of different homologues of tCsmA with ZnPP is studied. The results indicate that the high pH of dialysis buffer, the net charge, and α-helical structure of the peptide and the binding of ligand are important for the self-assembly. The tCsmA and ZnPP form two different modes of modulated self-assembly. The major type shows a 7_2 helical structure. When dialyzing the 7_2 helical assembly from the KPi buffer to the Tris buffer, a para-crystalline sheet similar to the natural baseplate can be observed by cryo-EM, showing that changing buffer can lead to a different structure. The para-crystalline sheets can be established by the untwisting of the helical structure and the formation of salt bridges between antiparallel peptide strands with arrays of dimeric ZnPP coordinated to the histidyl between the strands. In contrast, in the helical filaments the polar residues are solvent exposed on the outside of the 7_2 helix, and the self-assembly is stabilized predominantly by dipole-dipole interactions for chromophores and helices in the Tris buffer.

Zhongwu Sun, Christoph A. Diebolder, Ludovic Renault, Huub de Groot. Contrasting modes of modulated self-assembly for truncated CsmA and ZnPP, *To be submitted*

3.1 Introduction

As the smallest photosynthetic antenna known in nature, the CsmA-BChl *a* complex from the baseplate of *Chlorobaculum tepidum* is of interest for the design of semisynthetic building blocks for light harvesting in artificial photosynthesis.[1] The system assists in energy transfer and storage from the chlorosome to the FMO antenna and the photosynthetic reaction center.[2] The 59-residue CsmA protein is predominantly α-helical. It includes a long helical domain extending over residues 6–36, and a short helix in the C-terminal part formed by residues 41–49. Two flexible segments in both terminals attach to the chlorosome and the FMO antenna, respectively.[3] The BChl *a* coordinates to the His 25. CsmA is an amphipathic peptide. Except for the single coordinating histidine on the hydrophobic side, all potentially charged residues, like lysine and arginine, are located at the hydrophilic side of the helix.[4-6] Antiparallel dimers of CsmA-BChl *a* self-assemble to form extended para-crystalline layers with BChl *a* in between two CsmA strands running in opposite directions, bound to the next pair of strands by salt bridges.[7-8]

α-Helical peptides are often used as building blocks for self-assembly. In this chapter, I prepare different homologues of the natural CsmA by removing flexible segments from the N and C-terminals (Table 3.1). This leaves the functional central α-helix as the peptide motif to bind zinc-protoporphyrin (ZnPP) and form a semisynthetic scaffold. The semisynthetic tCsmA construct that is also used in the previous chapter contains 43 amino acids and is capped at the C-terminus with acetate and at the N-terminus with amine. In the F-CsmA, both caps are removed, while VF-CsmA and GF-CsmA are forms where a V or G residue was added at the C terminal. The rationale behind comparing these residues is that although tCsmA works, the additional groups may affect the self-assembly process, and by adding another residue instead of the caps it is possible to make a comparison between capped and bare peptides without changing the molecular weight and length significantly.

For the tCsmA, two different modes of self-assembly are found. It forms filaments that are predominantly helical, and the formation of a para-crystalline sheet similar to the natural baseplate is a rare event. It can be observed after dialysis to the Tris buffer, which appears to provoke untwisting of the helical structure and formation of salt bridges between

antiparallel peptide strands. Interestingly, the capping appears to be essential for the self-assembly, since for the uncapped versions modulated self-assembly by dialysis is difficult.

3.2 Experimental section

3.2.1 Materials

Zn (II) Protoporphyrin IX (ZnPP) was purchased from Frontier Scientific. Variants of truncated CsmA were purchased from GL Biochem Ltd (98% purity). DMSO (Sigma Aldrich), monobasic dihydrogen phosphate (Sigma Aldrich), dibasic monohydrogen phosphate (Sigma Aldrich), HG, OG and 1-propanol (Sigma Aldrich), Amicon® Ultra Centrifugal Filters (Sigma Aldrich) and microdialysis vials (Pierce™ 96-well Microdialysis Plate, 3.5K MWCO, Thermo Fisher Scientific) were used for the constitution.

3.2.2 Equipment

The concentration of protein was measured with a Microvolume DS-11 Spectrophotometer (DeNovix). The UV-Vis spectra were measured using a UV-1700 Spectrophotometer (Shimadzu). Negative stain images were obtained with a JEM-1010 transmission electron microscope with a maximum output voltage up to 70 kV, equipped with an Olympus Megaview camera (JEOL Ltd.). The CD spectra were measured using a J-815 CD spectrometer (JASCO). The time-resolved fluorescence measurements were performed with a FluoTime 300 spectrometer (PicoQuant). Cryo-EM data were collected at 300 kV with a Titan Krios electron microscope (Thermo Fisher Scientific).

3.2.3 Methods

Constitution of the tCsmA-ZnPP complex was carried out in a 5:1 mixture of ZnPP (Frontier Scientific) and tCsmA dissolved in a pH 7.9 50 mM KPi buffer containing either 6M 1-propanol or 8% HG detergent (Sigma Aldrich). The ZnPP was added from a 200 mM DMSO (Sigma Aldrich) stock solution. The coordination and self-assembly proceeded in the microdialysis vial (Pierce™ 96-well Microdialysis Plate, 3.5K MWCO, Thermo Fisher Scientific) against a

4 times 1.4 mL KPi dialysis buffer at 4 °C in the dark. The first two and the last dialysis steps were for 2 hours, while for the third dialysis step the sample was incubated overnight. For the OG solution, dialysis was performed twice overnight. All other conditions except for the dialysis times were kept the same as for the procedure with 1-propanol or 8% HG detergent solution. After the last dialysis step, the samples were centrifuged for 15 minutes at 4 °C with 4,000 $\times$ *g* to remove remaining free CsmA and ZnPP precipitate. The dialysis conditions were modified with respect to pH, temperature and dialysis time tor the other truncated variants. Dialysis times of 1h, 4 h or overnight were used, and. the pH of the dialysis buffer was varied over a range of 6.1-7.9. Finally, different peptide to ZnPP molar ratios 2:1, 1:1, 1:2,..., 1:10 were used in addition to 1:5 for optimizing the protocol.

For Cryo-EM and negative staining TEM, the method was described in Chapter 2.

For the time-resolved fluorescence measurements, samples were prepared with 3.5 µM tCsmA-ZnPP complex in the KPi buffer or the Tris buffer. The excitation beam was 440 nm.

The binding of tCsmA and ZnPP was determined from the optical shift of the ZnPP using a UV-1700 Spectrophotometer (Shimadzu) as described in Chapter 2. The binding condition was optimized by varying the pH over a range of 6.1-7.9.

To assess the secondary structure of variants of truncated CsmA and the tCsmA-ZnPP complex, CD spectra were measured using a J-815 CD spectrometer (JASCO).

3.3 Results and discussion

3.3.1 Conditions for self-assembly of the tCsmA-ZnPP complex

As an amphipathic α-helical peptide, tCsmA contains a hydrophobic side that can be stabilized by the hydrophobic sides of neighboring peptides in a

packed assembly, and a hydrophilic side to associate with water to increase the solubility.[9] Apart from the amphipathic property, intermolecular interactions are also necessary to drive the peptides to form an ordered supramolecular structure.[10] The driving force can be hydrogen bonds, salt bridges, the effect of ligands and specific functional groups to form *e.g.* disulfide bridges or π-π stacking interactions.[11]

Table 3.1 Information of different truncated CsmA

Name	Sequence	Calculated pI	MW
F-CsmA	FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG	5.65	4718
VF-CsmA	VFTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG	5.65	4817
GF-CsmA	GFTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG	5.65	4775
tCsmA	Ac-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG-NH_2	6.52	4759

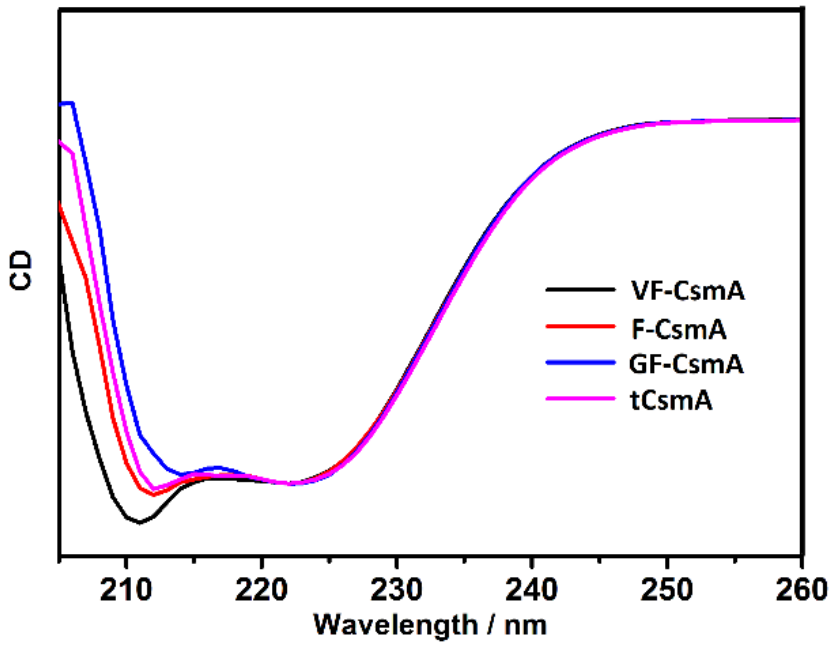


Figure 3.1 CD spectra for various truncated CsmA peptides dissolved in a propanol buffer.

To study the self-assembly behaviour, different truncated CsmA homologues were prepared (Table 3.1). tCsmA was prepared by N-terminal acetylation and C-terminal amidation to replace the flexible segments in the natural peptide. As was shown in Chapter 2, this leaves essential biological properties such as self-assembly and cofactor binding intact. As the CD spectra show, an α-helical structure was preserved for all variants studied, and in particular the CD spectra of tCsmA and F-CsmA virtually overlap

(Figure 3.1). This validates our hypothesis that the different truncations have little influence on the secondary structure. In Fig 3.2 UV-Vis spectra are shown, which shows whether the tCsmA and the ZnPP ligand coordinate to form an ordered assembly at different 6.1<pH<7.9. In contrast with the tCsmA, although these other peptides maintain the helical secondary structure of the natural system, little evidence for binding and complex formation is obtained, also when the pH, or the dialysis time and temperature are changed. The pI changes from pI=6.52 for the tCsmA to pI=5.65 for the other peptides.[12-13] When the pH > pI, a protein has a net negative charge and when the pH < pI, a protein has a net positive charge.[14] From pH 7.1 to pH 7.9 the other variants carry more significant negative charge than tCsmA. This may prohibit binding in the alkaline environment required for histidine coordination to the ZnPP to proceed. As the net charge of a protein directly influences many of its properties, such as solubility, assembly and crystallization, these other factors can be important as well.[15-16]

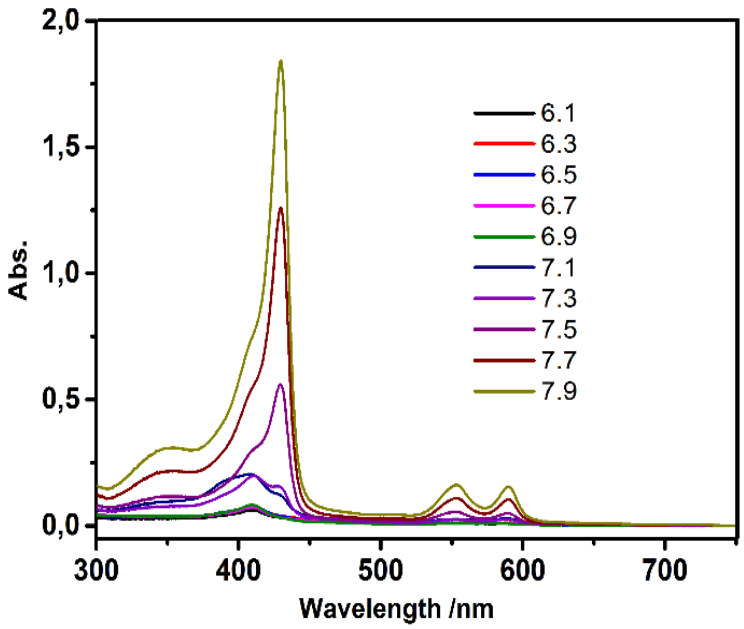


Figure 3.2 UV-vis spectra for the pH dependent constitution of tCsmA-ZnPP.

On the other hand, the pH of the dialysis buffer is important for the binding of ligand. As Figure 3.2 shows, when the buffer became basic pH>7 the binding could occur, and the binding became more efficient with the higher pH of the dialysis buffer, up to pH 7.9. This confirms that the basic

nitrogen atom in the imidazole of histidine with a lone electron pair can coordinate to metallic cations like the Zn^{2+} of ZnPP when the pH becomes basic.[17]

Because tCsmA and ZnPP both are insoluble in aqueous solution before self-assembly, the choice of solvent also influences the constitution. Detergents are used to solubilize tCsmA and ZnPP and keep the original secondary structure.[18] To avoid the interruption of binding by the micelles of detergents, the solution has to be diluted until the concentration of detergents is below the CMC.[19] This eliminates the micelles to promote the binding of tCsmA and ZnPP.[19] When the complex is prepared in this way, the complex cannot be concentrated since micelles will re-form during the concentration when using centrifugal filters. However, the detergents can be removed by dialysis. In addition, the unbound ZnPP and tCsmA precipitate upon removal of the detergent. As a result, according the UV data we are able to obtain a very pure complex when using the dialysis method for the constitution (Figure. 3.2).

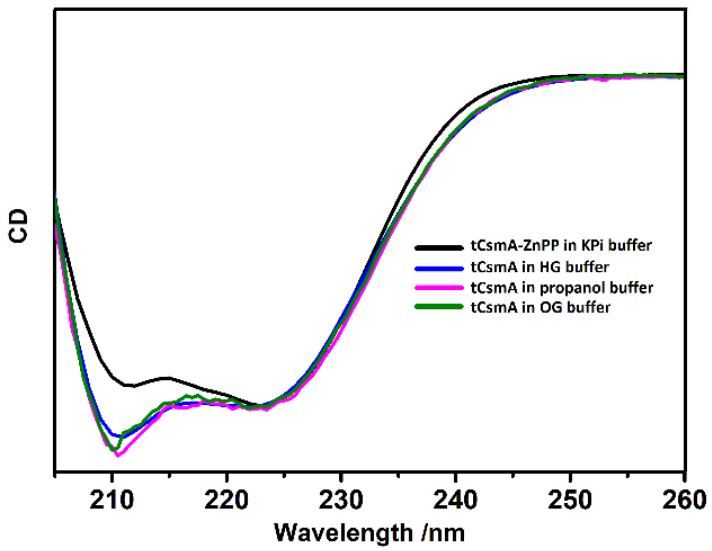


Figure 3.3 CD spectra for the tCsmA in different solvents and for the tCsmA-ZnPP construct in the KPi buffer.

The detergents with low CMC value are difficult to remove by dialysis, since the micelles are too large to pass through the membrane. Therefore, detergents with high CMC such as HG and OG are recommended. 1-propanol also increases the solubility of tCsmA, but with no risk of micelles.

The CD spectra of tCsmA in the propanol buffer and the detergent buffer overlap, and both show the characteristic shape for an α-helical structure (Figure 3.3). Thus, the 1-propanol and detergents keep the natural state of the tCsmA for ligand coordination.[20] However, although the detergents and 1-propanol cover the hydrophobic side to increase the solubility of tCsmA, the tCsmA peptides still randomly aggregate together. The solubility is limited to about 130 μM. The histidine coordination position is located at the hydrophobic side, and the detergents and 1-propanol have to be dialyzed out to expose the histidine for the coordination of ZnPP. For the HG and 1-propanol buffer, the complex is formed during the first 2 h of dialysis. However, the OG detergent with lower CMC is much slower to be dialyzed out of the membrane, and the formation of complex occurs after a 2 times dialysis overnight. The experiments indicate that the tCsmA peptides aggregate easily. Only after binding with ZnPP the peptides start to form large homogeneous filaments. Therefore the ZnPP also plays an important role in the self-assembly. The CD spectra in Figure 3.3 also show the change of tCsmA after the binding. The α-helical structure was still maintained after the binding, but the intensity at 211 nm and 216 nm increased. This is probably caused by the change in the loop between two helical regions of tCsmA. After self-assembly, the loop of peptides was aligned with two helical regions to form a straight α-helical peptide.

3.3.2 The helical structure of the self-assembly

The tCsmA-ZnPP complex is able to self-assemble to form 7_2 helical filaments after dialysis in the KPi buffer (cf. Chapter 2). The length of most filaments is a few micrometers (Figure 3.4A). Under negative stain TEM conditions, the images show strands or bundles (Figure 3.4B). The diameter of most bundles is around 10-20 nm, and the diameter of a single filament is about 5 nm. Self-assembly over extended time periods of 1 month leads to a very limited solubility, around 250 μM, which is attributed to the formation of larger aggregates over time. When the solution of the complex was centrifuged at 50,000 rpm, nearly all filaments precipitate in 30 minutes. On the other hand, the complex can be precipitated without centrifugation when the concentration of the KPi buffer is increased to 1 M. Both precipitates still maintain the same optical absorption characteristics when the sample is re-dissolved. The details of the helical structure are presented in Chapter 4.

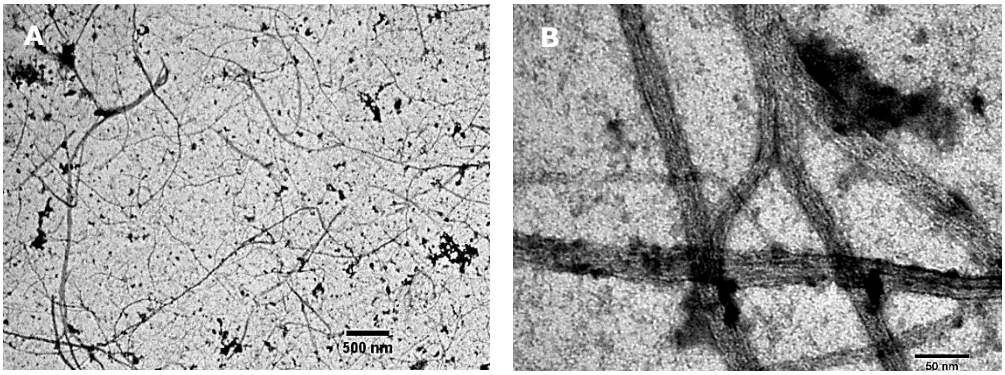


Figure 3.4 (A), (B) TEM image of the tCsmA-ZnPP complex in the KPi buffer.

3.3.3 The para-crystalline structure of the self-assembly

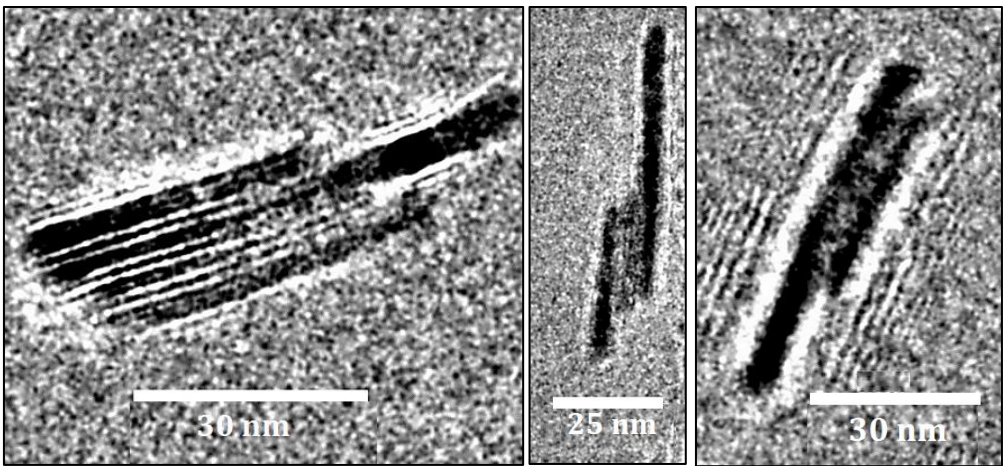


Figure 3.5 The cryo-EM images of transitional particles in the Tris buffer.

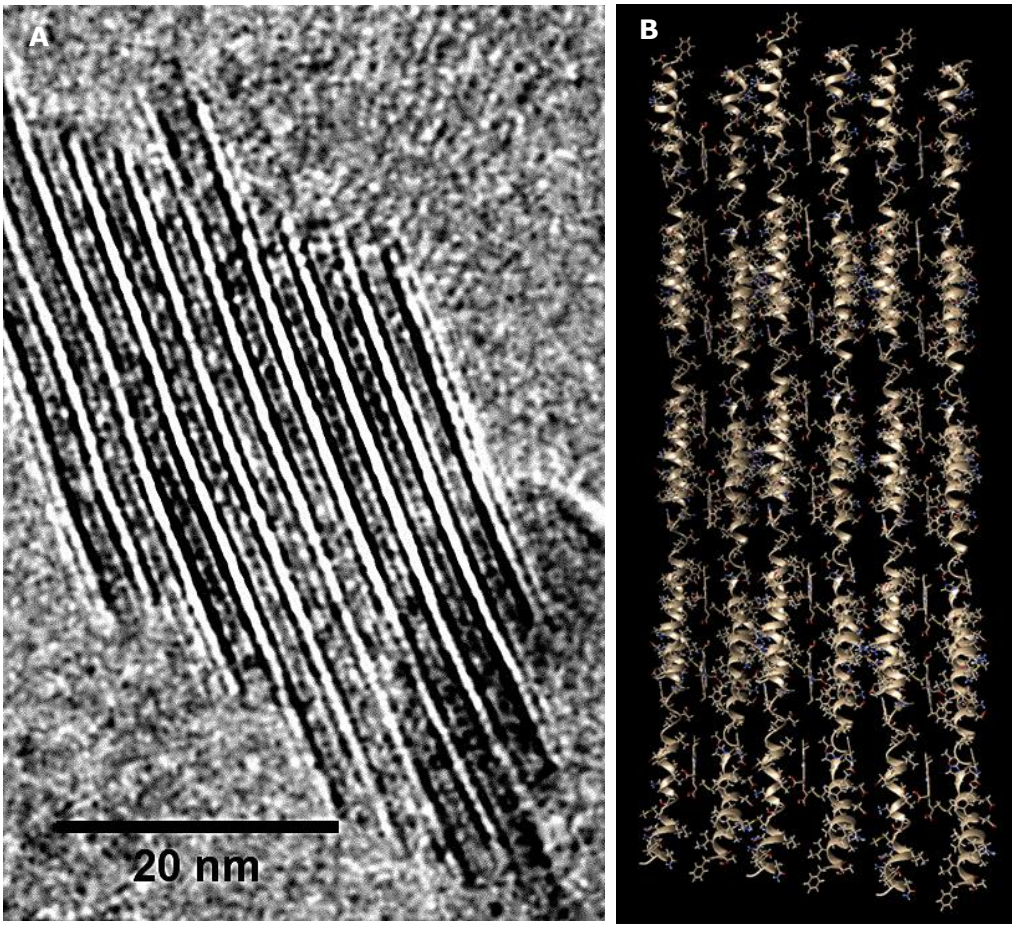


Figure 3.6 (A) Cryo-EM image of a para-crystalline self-assembly of the tCsmA-ZnPP complex in the Tris buffer. (B) The simulated para-crystalline self-assembly of the tCsmA-ZnPP complex in the Tris buffer.

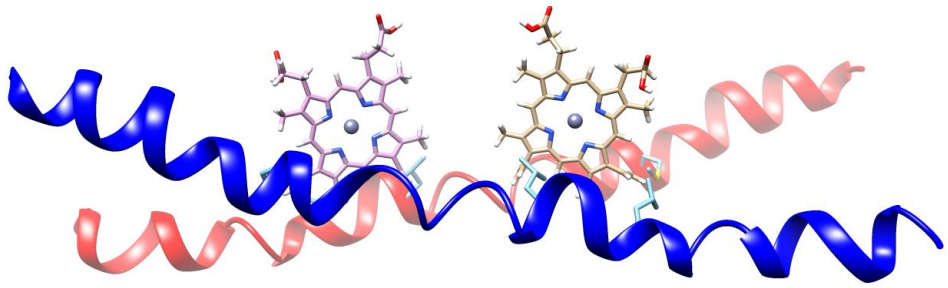


Figure 3.7 The dimeric unit of the para-crystalline baseplate

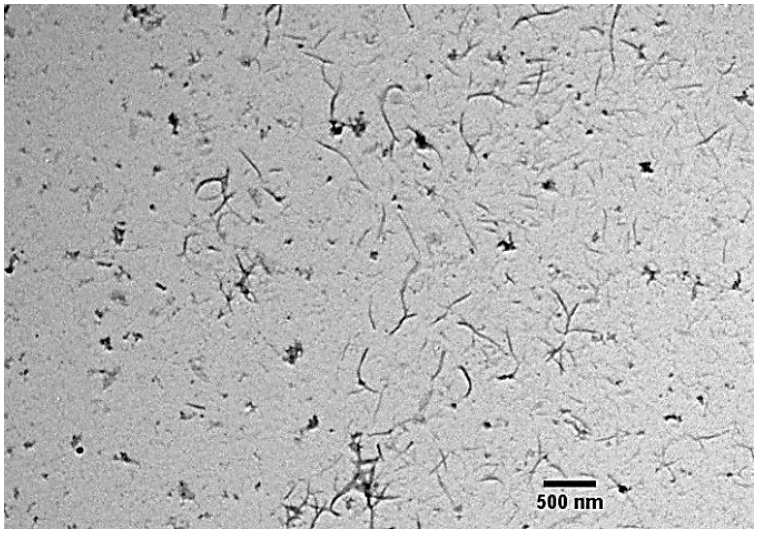


Figure 3.8 Negative staining TEM image of the tCsmA-ZnPP complex in the Tris buffer.

Tris buffer cannot be used as the dialysis buffer for the complex constitution. However, when the complex in the KPi buffer is dialyzed to the Tris buffer, the complex remains intact. Most filaments break up into smaller pieces during the dialysis, but without breaking down the complex (Figure 3.8).

In selected images, planar structures are observed that resemble the natural CsmA baseplate antenna. We attribute these planar structures to the untwisting of helical filaments upon changing the buffer (Figure 3.5). In the helix, the ZnPP interact to form a tetramer (Chapter 2), while in the planar form the ZnPP forms extended linear arrays of dimeric ZnPP in between two tCsmA strands (Figure 3.6). It is tempting to conclude that, similar to the natural baseplate, salt bridges can stabilize the planar form of the tCsmA-ZnPP complex. The baseplate structure is provoked by the untwisting of the helical structure and the formation of salt bridges between antiparallel peptide strands with arrays of dimeric ZnPP coordinated to the histidyl between the strands. The ZnPP shows a vertical orientation between tCsmA strands (Figure 3.7).

As Figure 3.6A shows, the distance between strands is about 31.3 Å, which is the same as for the natural baseplate. According to the EM data, the ZnPP is embedded in the middle of the strands and is oriented perpendicular to the plane of the assembly. Probably the hydrophobic side of the self-assembly is turned nearly 90° relative to the helical form to interact with the adjacent

complex in the other strand directly. As a result, the dimeric ZnPP cannot interact with the ZnPP motifs in neighboring strands. Thus, in the crystalline structure, the ZnPP forms a dimeric unit rather than a tetrameric unit as Figure 3.7 shows.

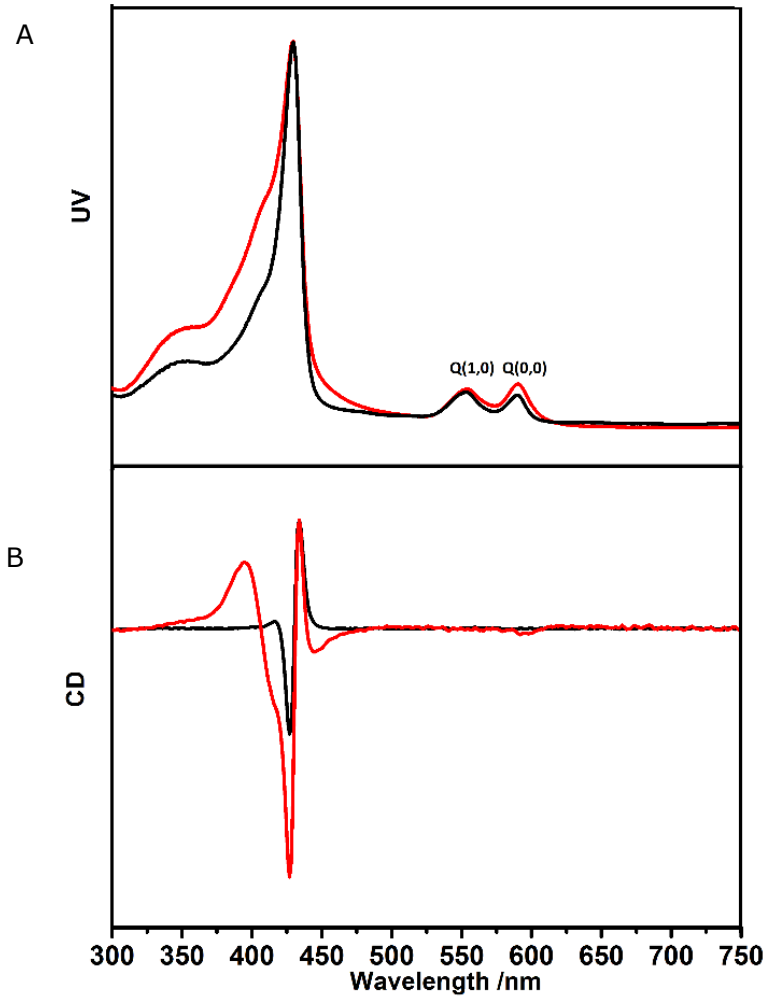


Figure 3.9 UV-vis and corresponding CD spectra of the tCsmA-ZnPP complex in the KPi buffer (black trace) and the Tris buffer (red trace) respectively.

Figure 3.9 indicates the conformational changes that are observed in optical spectroscopy for the two sample species. The UV-Vis spectrum of the complex in the Tris buffer shows a broad peak at 430 nm which indicates less homogeneity in the Tris buffer (Figure 3.9A). Moreover, the UV-Vis absorption of Q(0,0) becomes higher than for the Q(1,0). This indicates that the ZnPP may be less planar in the Tris buffer relative to the KPi buffer.[21-22]

This is corroborated by the CD response of the complex. The CD response of the complex in Tris buffer is asymmetric, which contrasts with the symmetric spectrum in the KPi buffer due to the tetrameric ZnPP in the 7_2 helix (Figure 3.9B and Chapter 2). However, the CD response of the Q bands in the Tris buffer is quenched, similar to the helical form. This suggests reorientation of the molecule going from the tetramers in the helical form to the linear arrays in the planar form into a configuration to offset any dipole-dipole interactions that could give rise to a CD response.[23-24]

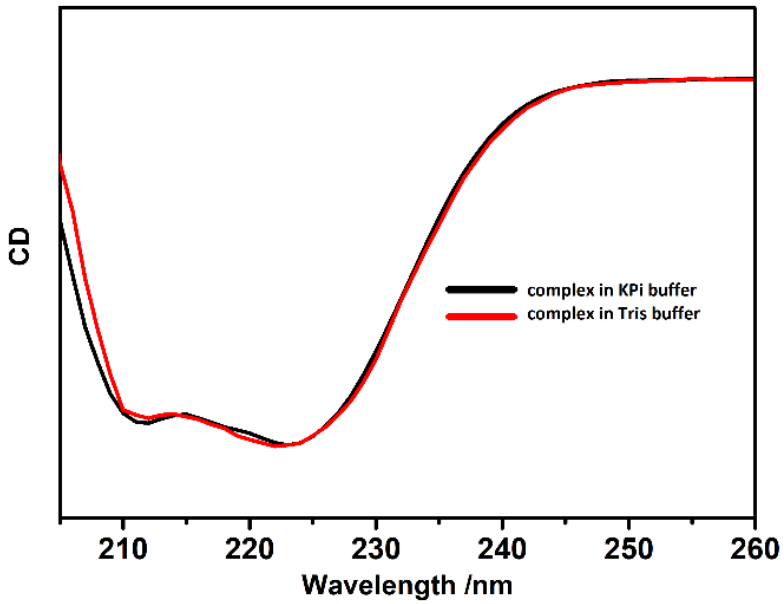


Figure 3.10 CD spectra of the tCsmA-ZnPP complex in the KPi buffer (black trace) and the Tris buffer (red trace) respectively.

As Figure 3.9B shows, the intensity ratio of the negative and positive CD responses for the ZnPP is more than 2, which indicates structural rearrangement at the level of the complex when it is transferred to another buffer. However, in the CD data in Figure 3.10, there is considerable spectral overlap, which indicates that overall the tCsmA helical peptide has little conformational change of secondary structure on the peptide level upon buffer change.

Table 3.2 Fitted Time constants of Fluorescence Transients of the tCsmA-ZnPP complex in the Tris buffer under a 440 nm excitation.

Buffer	594 nm Lifetime, ns				650 nm Lifetime, ns			
Lifetime	τ_1	τ_2	τ_3	$\tau_{average}$	τ_1	τ_2	τ_3	$\tau_{average}$
Tris buffer	2.64 (3%)	1.21 (45%)	0.4 (52%)	0.84	5.15 (2.4%)	1.36 (44%)	0.45 (53.6%)	0.97

Finally, for the ZnPP in the Tris buffer the photochemical properties are very similar to the system in the KPi buffer (Table 3.2 and Chapter 2 Table 2.3). The complexes in the Tris and the KPi buffer both have 3 similar lifetimes at the same emission peaks at 594 nm and 650 nm. Apparently the loss of symmetry in the Tris buffer has little effect on the decay characteristics.

3.4 Conclusion

At near neutral pH 7.1-7.9, binding of tCsmA to ZnPP is observed with dialysis in the KPi buffer. In addition, a high pH of the dialysis buffer is important for the self-assembly of tCsmA and ZnPP. The net charge of the protein was implied to be important for the self-assembly. Hence N-terminal acetylation and C-terminal amidation promote self-assembly. While the principle structure of self-assembly is a 7_2 helix, the tCsmA-ZnPP complex is able to form another type of ordered chiral matrix in the Tris buffer that mimics the baseplate array in the natural system and appears stabilised by hydrophobic effects, salt bridges, and dipole-dipole interactions of chromophores. It is tempting to conclude that the para-crystalline baseplate structure is formed by the untwisting of the 7_2 helix and the reorientation of the peptide strands to accommodate salt bridges between strands.

3.5 References

1. Frigaard, N.-U.; Bryant, D. A., Chlorosomes: Antenna Organelles in Photosynthetic Green Bacteria. In *Complex Intracellular Structures in Prokaryotes*, Shively, J. M., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; pp 79-114.
2. Li, H.; Frigaard, N.-U.; Bryant, D. A., *Biochemistry* **2006,** *45* (30), 9095-9103.
3. Pedersen, M. Ø.; Underhaug, J.; Dittmer, J.; Miller, M.; Nielsen, N. C., *FEBS Letters* **2008,** *582* (19), 2869-2874.
4. Bryant, D. A.; Vassilieva, E. V.; Frigaard, N.-U.; Li, H., *Biochemistry* **2002,** *41* (48), 14403-14411.
5. Zobova, A.; Taisova, A.; Lukashev, E.; Fedorova, N.; Baratova, L.; Fetisova, Z., *Journal of Biophysics* **2011,** *2011*.
6. Kulminskaya, N. V.; Pedersen, M. Ø.; Bjerring, M.; Underhaug, J.; Miller, M.; Frigaard, N. U.; Nielsen, J. T.; Nielsen, N. C., *Angewandte Chemie International Edition* **2012,** *51* (28), 6891-6895.
7. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.
8. Andrew Staehelin, L.; Golecki, J. R.; Drews, G., *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1980,** *589* (1), 30-45.
9. Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J., *Nature Chemistry* **2013,** *5*, 161.
10. Hu, K.; Jiang, Y.; Xiong, W.; Li, H.; Zhang, P.-Y.; Yin, F.; Zhang, Q.; Geng, H.; Jiang, F.; Li, Z.; Wang, X.; Li, Z., *Science Advances* **2018,** *4* (5), eaar5907.
11. Bai, Y.; Luo, Q.; Liu, J., *Chemical Society Reviews* **2016,** *45* (10), 2756-2767.
12. Bjellqvist, B.; Hughes, G. J.; Pasquali, C.; Paquet, N.; Ravier, F.; Sanchez, J.-C.; Frutiger, S.; Hochstrasser, D., *ELECTROPHORESIS* **1993,** *14* (1), 1023-1031.
13. Gauci, S.; van Breukelen, B.; Lemeer, S. M.; Krijgsveld, J.; Heck, A. J. R., *PROTEOMICS* **2008,** *8* (23‐24), 4898-4906.
14. Kozlowski, L. P., *Nucleic acids research* **2017,** *45* (D1), D1112-D1116.
15. Gitlin, I.; Carbeck, J. D.; Whitesides, G. M., *Angewandte Chemie International Edition* **2006,** *45* (19), 3022-3060.
16. Shaw, K. L.; Grimsley, G. R.; Yakovlev, G. I.; Makarov, A. A.; Pace, C. N., *Protein science : a publication of the Protein Society* **2001,** *10* (6), 1206-1215.
17. Liao, S.-M.; Du, Q.-S.; Meng, J.-Z.; Pang, Z.-W.; Huang, R.-B., *Chemistry Central journal* **2013,** *7* (1), 44-44.
18. Nagarajan, R.; Ganesh, K., *Macromolecules* **1989,** *22* (11), 4312-4325.
19. Pedersen, M. Ø.; Pham, L.; Steensgaard, D. B.; Miller, M., *Biochemistry* **2008,** *47* (5), 1435-1441.
20. Singh, S. M.; Sharma, A.; Upadhyay, A. K.; Singh, A.; Garg, L. C.; Panda, A. K., *Protein expression and purification* **2012,** *81* (1), 75-82.
21. Milgrom, L. R.; Warren, M. J., **1997**.

22. Giovannetti, R., The use of spectrophotometry UV-Vis for the study of porphyrins. In *Macro to nano spectroscopy*, IntechOpen: 2012.
23. Harada, N.; Nakanishi, K., *Accounts of Chemical Research* **1972,** *5* (8), 257-263.
24. Pescitelli, G.; Gabriel, S.; Wang, Y.; Fleischhauer, J.; Woody, R. W.; Berova, N., *Journal of the American Chemical Society* **2003,** *125* (25), 7613-7628.

Chapter 4

Cryo-EM structure of the self-assembly of tCsmA and ZnPP

4.0 Abstract

The 7_2 helical structure of the self-assembled tCsmA-ZnPP complex is resolved at about 7 Å resolution by the cryo-EM technique. Each strand of the 2-start helix contains two tCsmA peptide strands running antiparallel. The threads of tCsmA have the monomers organized in a head to tail manner. They twist and form a hydrophobic core, to embed ZnPP tetramers coordinated to the imidazole groups of the histidine. The study of the structure at moderate resolution is helpful for understanding the mechanism of the self-assembly and the interactions between chromophores, and chromophore-peptide interactions. This provides a structural basis for the rational design of semisynthetic building blocks for artificial photosynthesis.

Zhongwu Sun, Stefan Huber, Ludovic Renault, Christoph A. Diebolder, Arjen Jakobi, Huub de Groot. Cryo-EM structure of the self-assembly of tCsmA and ZnPP, *To be submitted*

4.1 Introduction

As described in chapter 2, the truncated CsmA (tCsmA) and zinc protoporphyrin IX (ZnPP) chromophore are able to form 7_2 helical filaments in KPi buffer. The specific aim of this chapter is to provide structural information at moderate solution to understand the self-assembly and the interactions between chromophores. Structures of macromolecular assemblies have been determined by X-ray crystallography, NMR spectroscopy and single particle cryo-electron microscopy (cryo-EM).[1-3] Cryo-EM is the most suitable method for helical reconstruction of the tCsmA-ZnPP self-assembly. Cryo-EM is able to determine structures of assemblies with molecular weights larger than 150 kDa with very small sample amounts.[4] It does not require crystals and it is possible to work from paracrystalline materials.

To perform helical reconstruction in 3D, a single particle like analysis was applied to cryo-EM data collected from large helical filaments of tCsmA-ZnPP using the RELION empirical Bayesian approach.[5-9] Helical structures provide a view from all sides in every single image of a filament. Although in single-particle analysis all relative orientations of the particle projection images are necessary, the repeating, asymmetrical units of helical structures have fixed relative orientations. Therefore, helical symmetry and the orientation of the image of an entire filament can also reduce the experimental noise efficiently by averaging over a large number of asymmetrical units. To determine the precise symmetry during the 3D reconstruction process, the RELION software was used to obtain class averages, perform 3D reconstruction and determine the helical symmetry. Moreover, the additional point group symmetry that appears to be present in the tCsmA-ZnPP self-assembly, was also determined with RELION. However, although RELION has the tools for helix reconstruction with little external guidance, it can give rise to artifacts.[10] The helical symmetry was therefore validated with SPRING, which is an excellent single-particle procedure to determine precise helical symmetry in Fourier space.[11] As a helical reconstruction package for electron cryo-micrographs, it is based on a class-based helical reconstruction scheme to explore the helical symmetry by reconstructing the 3D structure from a large number of symmetry combinations using a single class average. According to the cross-correlation of the power spectra of the class average and projections of the

reconstructed 3D model, the accurate helical symmetry can be determined from a large number of possible symmetry combinations.[10]

Using this combination approach, we present here the cryo-EM structure at moderate resolution of the tCsmA in complex with ZnPP and the resulting helical structure of the self-assembly. The structure reveals the overall organization of tCsmA scaffolds, the interaction between ZnPP and tCsmA, and how tetramer formation affects the ZnPP and its dipole-dipole interactions.

4.2 Experimental section

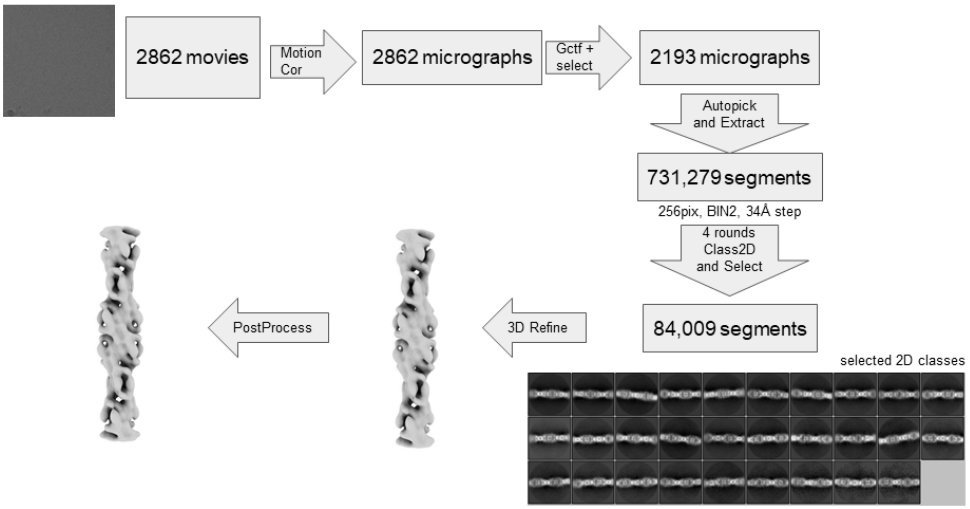


Figure 4.1 Processing scheme of cryo-EM data

Cryo-EM data were collected using a Titan Krios microscope at 300 kV. The experimental procedures are described in detail in Chapter 2. The data were transferred to a computer containing 4 NVIDIA GPU image processing units for pre-processing and helical symmetry determination with RELION 3.0. The helical symmetry determination with SPRING was performed with a CPU-based cluster. The final 3D reconstruction was performed in RELION 3.0. The processing scheme of cryo-EM data is shown in Figure 4.1. A total of 2193 micrographs were selected from 2862 raw movies. After 4 rounds of 2D classification, 84009 particles were selected for the 3D refinement. In addition, a cylinder mask with 6 nm diameter was applied during the initial 3D refinement. The helical rise was set to 34 Å in RELION. To make sure that a full helical unit is included in the analysis, this parameter should marginally

exceed the experimentally measured helical rise, which is about 32.9 Å. To fully observe the helix, a final box size of 256 Å was used to cover the 230 Å repeat for the 7_2 helical assembly. The details of the image processing are discussed in Results and Discussion.

4.3 Results and discussion

4.3.1 Cryo-EM images processing of the helical self-assembly

$$\text{Helical rise}_{1\text{-start}} = \text{rise}_{n\text{-start}}/n$$
$$\text{Helical twist}_{1\text{-start}} = (\text{twist}_{n\text{-start}} - 360°)/n \text{ for twist}_{n\text{-start}} < 0$$
$$\text{Helical twist}_{1\text{-start}} = (\text{twist}_{n\text{-start}} + 360°)/n \text{ for twist}_{n\text{-start}} > 0$$

Equation 4.1

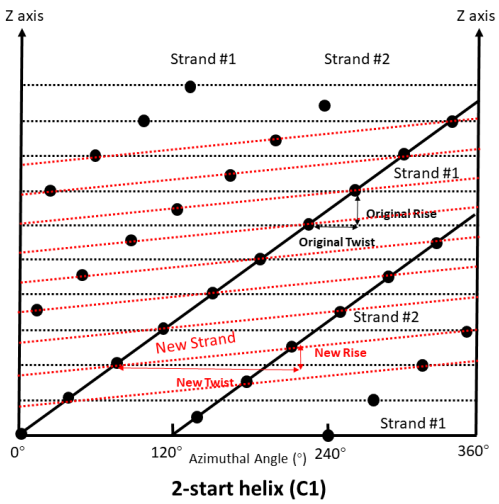


Figure 4.2 Two-start helix in a C1 point group symmetry representation, where instead of using the original twist and rise of the individual two strands (in black), RELION expresses the helical symmetry using a new twist and rise (in red).[10]

From the class averages the diameter of the filaments was determined as about 5 nm (Chapter 2, Figure 2.7B). The first step in resolving the structure was to apply a cylinder mask of 6 nm diameter to provide an initial electron density model, shown in Figure 4.3. In RELION only 1-start helices can be handled for 3D reconstruction. The procedure to convert the 2-start helix into a 1-start helix is schematically depicted in Figure 4.2. [10] This can be visualized by opening up the helix to form a 2D surface. n-start helices

always lead to a new helical twist and rise from the original twist and rise to represent the 1-start helix, where the new rise and twist are easily calculated using Equation 4.1.

The C1 point group symmetry analysis provides an unbiased (*i.e.* without a priori symmetry assumption) initial estimate of the helical reconstruction. In particular it indicates a 2-start helix (Figure 4.3).

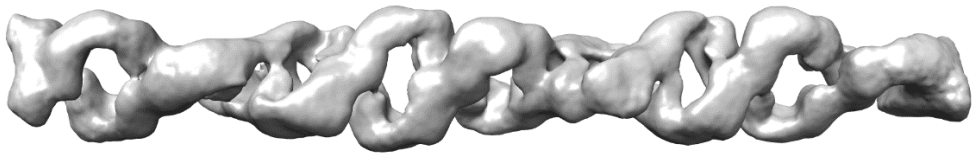


Figure 4.3 The initial electron density model without a priori symmetry assumption.

Except for helical symmetry, the self-assembly may also contain additional point group symmetry. The 2D class averages indicate rotational symmetry along the helix and perpendicular to the helix, pointing to D2 symmetry (SFigure 4.1). In addition, according to the optical data in Chapter 2, the tetrameric ZnPP clusters are highly symmetric with only one unique site, which indicates D2 point group symmetry (with three twofold axes through the tetrameric cluster) as well. For helices with D2 symmetry, the two separate strands of molecules are stacked against each other and all start from the same level along the Z axis. With point symmetry taken into account and validated, RELION can perform its empirical Bayesian single particle like analysis on the actual symmetry model, instead of the C1 reconstructed model according to Figure 4.2. To check the point symmetry guess, D2 symmetry without helical symmetry was applied for 3D reconstruction (Figure 4.4). Compared to the initial map (Figure 4.3), the 3D map with D2 symmetry shows more details (Figure 4.4B). In addition, the projection of the resulting map (Figure 4.4A) from D2 symmetry compares well with the 2D class average data (Figure 2.7B), with almost perfect overlap. This confirms that the helix contains additional D2 symmetry.

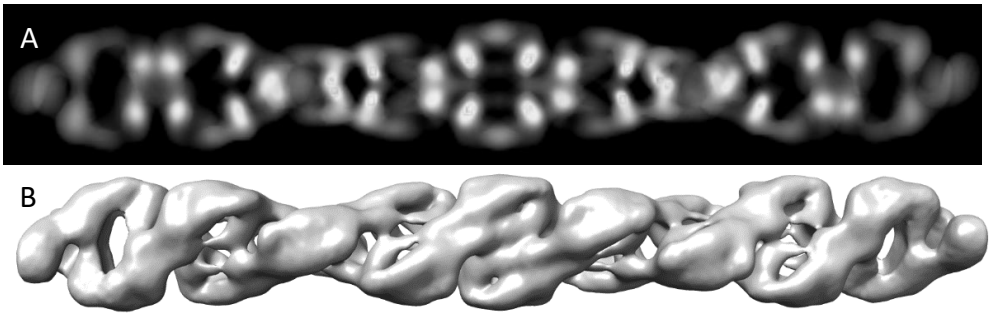


Figure 4.4 The projection image of resulting electron density map with D2 symmetry (A) and the corresponding electron density map (B).

To proceed with resolving the helical symmetry with D2 point symmetry, we determined from the 2D class averages initial estimates for the helical rise and twist (Figure 4.5). Here the initial guess is determined in Fourier space by taking the Fourier transform of a class average and determine the position of the layer lines. In Figure 4.5 a clear layer line at 1/32.9 Å is observed with intensity on the meridian, which serves as the initial guess for the helical rise in the reconstruction procedure. Since the image repeats itself after 113 Å, corresponding with a turn over 180°, a full turn corresponds to 226 Å, leading to 6.87 units in one helical turn with a 52.4° helical twist. This initial analysis is in line with the direct observations from the 2D class images in chapter 2.

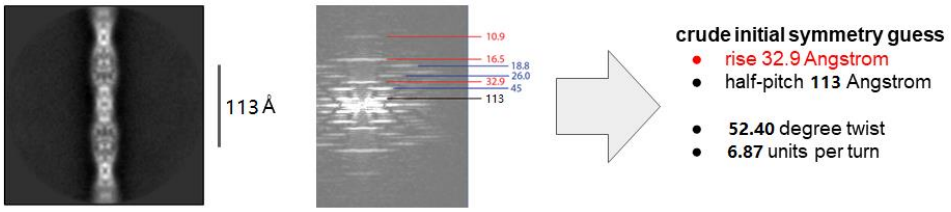


Figure 4.5 The crude initial helical symmetry guess based on the Fourier transform of a class average.

For refinement both D2 symmetry and helical symmetry with 52.4° for the helical twist and 32.9 Å for the helical rise were applied. However, comparing SFigure 4.1 to Figure 4.4B, the resolution is not improved after applying the helical symmetry. To validate the initial guess and get the precise helical symmetry, the helical parameters were confirmed using the SPRING helical analysis software (SFigure 4.2). The simulated power spectra fits the experimental data well. It confirms that the initial guess should be close to the real helical symmetry.

4.3.2 Cryo-EM structure of the helical self-assembly

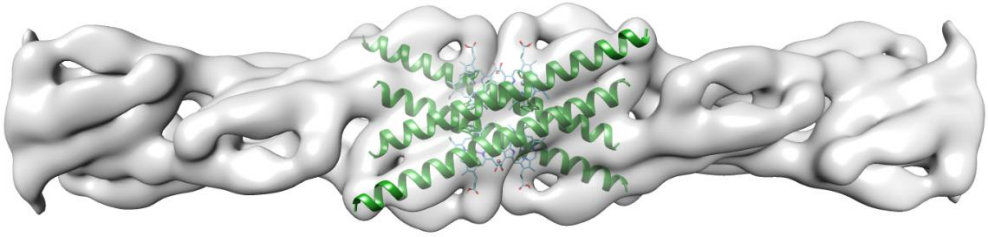


Figure 4.7 The electron density map fitted by the tetramer unit.

Under cryo-conditions, only single filaments stayed on the grid. They represent the same type of helical self-assembly which is helpful for the image processing. The 3D reconstruction map was obtained at about 7 Å resolution (Figure 4.7), which is sufficient to identify the helices and the ZnPP motifs. Interestingly, the homology model derived from the CsmA from the natural baseplate determined by NMR was found too long for matching to the electron density map. This can be attributed to the loop fragment from residue 30 to residue 35 in the homology structure (blue model in Figure 4.8). When this loop fragment is made helical, a fully extended rigid helix is obtained that is well in line with the EM structural data (Figure 4.7). Hence we propose there is no loop in the helix for the tCsmA-ZnPP complex. A fully helical tCsmA fragment is corroborated by the continuous electron density in the map, as a loop fragment should give rise to weaker electron density than for helical motifs. The straight alpha-helical tCsmA model simulated with the protein builder was used to fit the electron density map (Figure 4.7).[12] The stretched peptide helix model matches the NMR structure for residues 1-31 (red model in Figure 4.8) and fits well into the density map (Figure 4.7). Although a stretched helix derived from the structure proposed for the CsmA in the natural baseplate was also used to analyse the electron density (green model in Figure 4.8), this may be overstretched in the region of residues 1-31 and 35-43 as the peptide helix is too long and protrudes on both sides from the electron density map, giving a very poor fit. The filament helix is built from two parallel strands of tCsmA-ZnPP. Each strand is composed of antiparallel dimers of tCsmA-ZnPP with a P2 symmetry. Between the two strands, pairs of dimeric ZnPP are close to one another, and as a result, tetrameric units of ZnPP are formed. The complex forms a hydrophobic core, while the hydrophilic residues are

exposed to outside (Figure 4.10). This validates the interpretation of the optical data presented in Chapter 3.

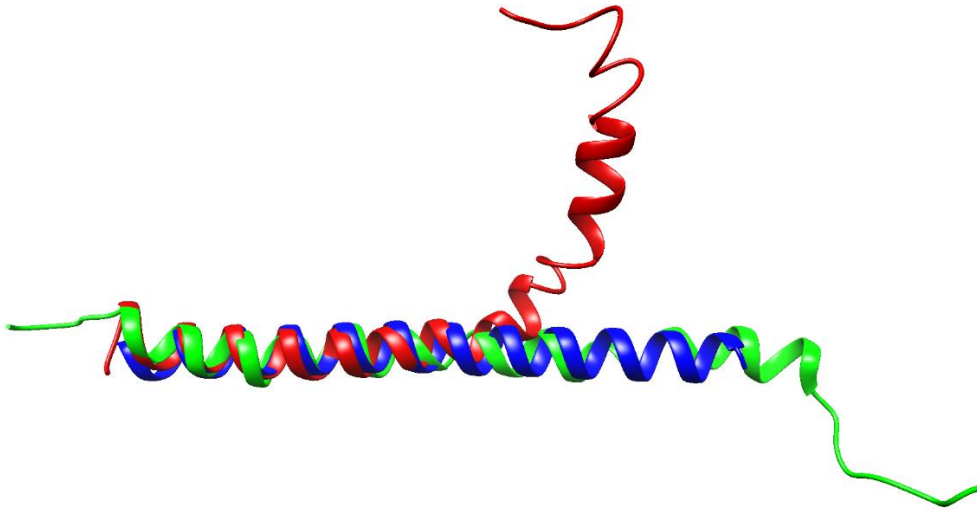


Figure 4.8 The simulated model of tCsmA (blue), the model from the natural baseplate (green) and the model of CsmA from NMR (red).

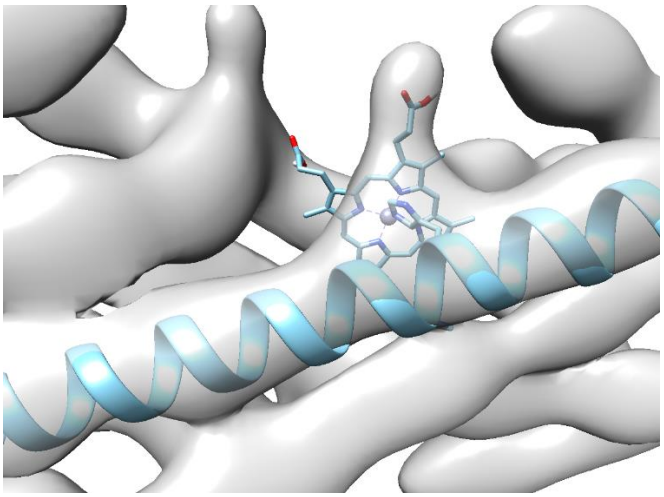


Figure 4.9 The possible fitting part of the propionic acid groups of ZnPP.

The electron density map is in line with coordination of the ZnPP at the H25 binding site (Figure 4.9). It also shows a small bump around the ZnPP, which can be attributed to one of the propionic acid groups -CH_2-CH_2-COOH (Figure 4.9). The rigid density indicates that the propionic acid group may interact with the peptide, for instance by forming a salt bridge, to make itself fixed. The dimeric unit in the tCsmA-ZnPP self-assembly shows some

similarity with the natural CsmA-BChl *a* complex.[13] The distance between the two ZnPP within an antiparallel tCsmA-ZnPP dimer is 21.1 Å, which is close to the distance between BChl *a* in the natural baseplate (Figure 4.10).[13] In addition, the two complexes cross with a similar angle as in the natural system. Having access to a model structure like tCsmA-ZnPP will be helpful to understand the mechanism of self-assembly and provide further refinement of the baseplate structure in the natural system.

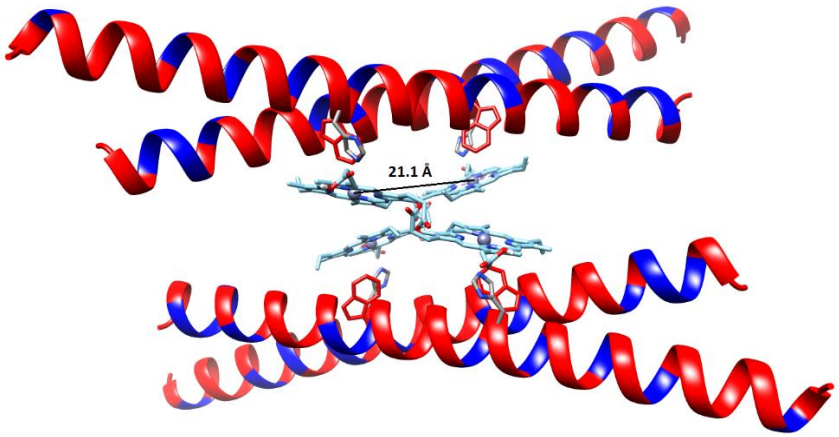


Figure 4.10 The tetrameric unit of the helical filament. The red colour presents hydrophobic parts and the blue colour is the hydrophilic part.

For the tCsmA-ZnPP complex in KPi buffer, each strand prefers to twist to stabilize a helical structure. As a result, the orientation of the complexes is nearly 90° rotated away from the orientation of the dimers in the planar arrangement observed in the natural system. On the one hand this allows to form a well-protected hydrophobic core with the ZnPP located inside, shielded from the environment by the outer peptide scaffold. In this arrangement, the salt bridges between peptide strands cannot contribute to stabilizing the self-assembly. In this way the structure indicates that the driving forces are mainly the hydrophobic effect and dipole-dipole interactions between the ZnPP. Since the charged residues are exposed towards the outside of the helix they can form salt bridges with other helixes to form bundles in KPi buffer when not on a grid (Chapter 3, Figure 3.4).

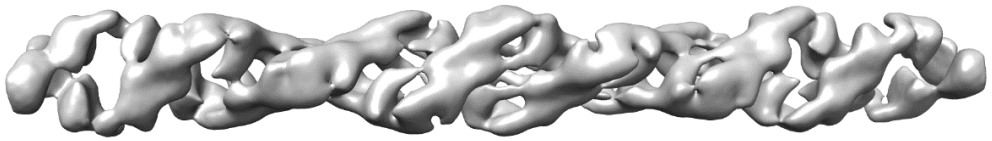
4.4 Conclusion

In conclusion, a refined 3D electron density map is obtained at a resolution of about 7 Å. The electron density map indicates a full alpha-helical structure and a fully helical simulated tCsmA-ZnPP model fits the map quite well. The complex forms a 7_2 helix in KPi buffer. Each strand is composed by dimeric tCsmA-ZnPP complexes which contain two antiparallel complexes with P2 symmetry. Dimeric ZnPP from the two strands are quite close to one another, and as a result, tetrameric units are formed.

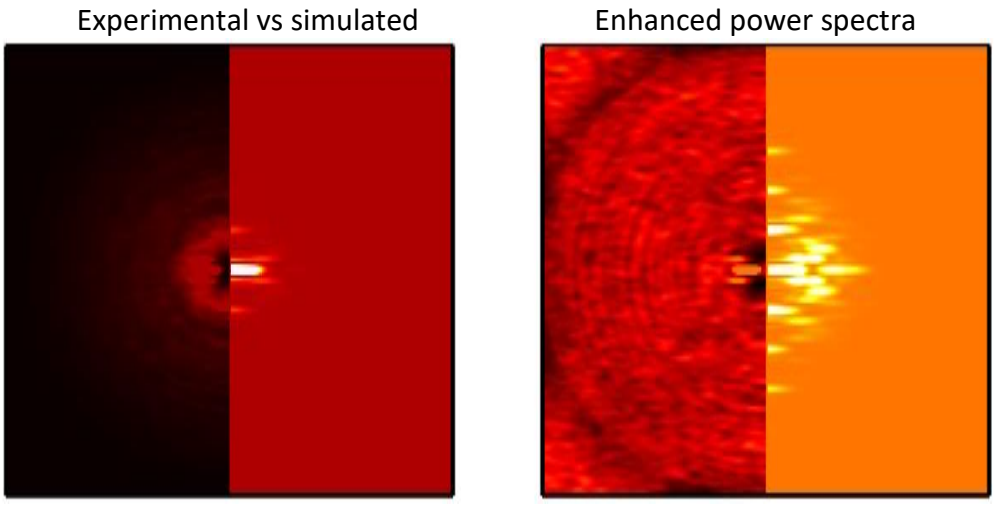
4.5 References

1. Sali, A.; Glaeser, R.; Earnest, T.; Baumeister, W., *Nature* **2003,** *422* (6928), 216-225.
2. Jiang, W.; Ludtke, S. J., *Current Opinion in Structural Biology* **2005,** *15* (5), 571-577.
3. Chiu, W.; Baker, M. L.; Jiang, W.; Dougherty, M.; Schmid, M. F., *Structure* **2005,** *13* (3), 363-372.
4. Topf, M.; Lasker, K.; Webb, B.; Wolfson, H.; Chiu, W.; Sali, A., *Structure* **2008,** *16* (2), 295-307.
5. Scheres, S. H. W., *Journal of Molecular Biology* **2012,** *415* (2), 406-418.
6. Bricogne, G., *Acta Crystallographica Section A* **1984,** *40* (4), 410-445.
7. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., *Acta Crystallographica Section D* **1997,** *53* (3), 240-255.
8. Rieping, W.; Habeck, M.; Nilges, M., *Science* **2005,** *309* (5732), 303-306.
9. Frank, J., *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state*. Oxford University Press: 2006.
10. He, S.; Scheres, S. H., *Journal of structural biology* **2017,** *198* (3), 163-176.
11. Desfosses, A.; Ciuffa, R.; Gutsche, I.; Sachse, C., *Journal of Structural Biology* **2014,** *185* (1), 15-26.
12. https://nova.disfarm.unimi.it/probuilder.htm.
13. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.

4.6 Appendix



SFigure 4.1 The electron density map with D2 symmetry and helical symmetry (52.4° for the helical twist and 32.9 Å for the helical rise).



SFigure 4.2 The comparison chart of experimental and simulated power spectra based on the 2D class average images and the initial guess of the helical symmetry respectively (52.4° for the helical twist and 32.9 Å for the helical rise).

Chapter 5

Purification of recombinant variants of tCsmA

5.0 Abstract

The tCsmA was derived from the natural system by removing the flexible segments at the N and C-terminal, leaving the two functional central α-helix segments. Without a large fusion tag, tCsmA rapidly degrades. Two tags were studied to prevent the degradation of tCsmA, TRX and MBP. With TRX the stability can be increased, but the sample precipitates making further treatment difficult. With the MBP tag a strongly associated soluble tCsmA-MBP construct is obtained. While SEC can separate the tCsmA and MBP with the help of 6 M GuHCl, and the tCsmA denatures, a small fraction of tCsmA was recovered during the dialysis. This paves the way for further improvement towards obtaining recombinant variants of tCsmA.

5.1 Introduction

The harvesting of sunlight using artificial photosynthesis has been widely studied for decades. In nature, materials for light harvesting, charge separation, and catalysis are integrated into high-density matrix arrays with proteins.[1] High-density arrays of pigments inspired by nature have been explored for the design of artificial photosynthetic units.[2-5] In particular, the chemically robust ZnPP, a compound found in red blood cells, is often used for studies into artificial light harvesting.[6-8] To build a semi-artificial matrix around the ZnPP chromophore, we employ a truncated version of the natural 59-residue CsmA as the peptide motif to bind ZnPP and form a semisynthetic 2D scaffold. The semisynthetic tCsmA construct, containing 43 amino acids, was derived from the natural system by removing the flexible segments from the N and C-terminal which attach to the membrane of the chlorosome and the FMO complex, leaving the functional central α-helix.[9]

NMR, X-ray crystallography and EM are major techniques to study the structure of proteins.[10] As is shown in chapter 4, it is possible to obtain at moderate solution for the tCsmA-ZnPP complex with cryo-EM, however, it is not possible to obtain dynamic information with this method.[11-12] On the other hand, NMR can be used to obtain abundant microscopic information, *e.g.* angles, distances, coupling constants, chemical shifts, relaxation rates, to study the activation thermodynamics.[13] Generally, the isotope labelled of proteins is necessary for NMR. SPPS is too costly to obtain labelling protein.[14-15] However, via protein biosynthesis large or isotope labelled proteins are more easily obtained. Protein expression in *E.coli* has been one of the most popular means of producing recombinant proteins for over the past two decades.[16] *E.coli* is a well-established host that offers short culturing time, easy genetic manipulation and low cost media. In this chapter, our specific aim is using *E.coli* to express the tCsmA peptide.

5.2 Experimental section

5.2.1 Materials

Bacterial strains and plasmids

All cloning experiments were performed in the Leiden Cell observatory facility by its support staff. For the expression experiments two bacterial strains BL21(DE3) and BL21(DE3)/pLys (CamR) were used. Plasmid vectors pETMBP1a, pETMBP1b, pETTRX1a, pETTRX1b (EMBL, Hamburg) and pET16SK-SUMO2 are used for expression to get variants of tCsmA to ensure that at least one of them works as well as tCsmA from SPPS. All these plasmids carry the kanamycine resistance gene.[17] For production of tCsmA with a His tag plasmid pET3His was used (AmpR).[18] The tCsmA without a tag was also tested. The sequences are listed in Table 5.1 and include leftovers of the various tags that were used. Purification was performed with the buffers, proteins and protease inhibitors shown in Table 5.2.

Table 5.1 The sequence of recombinant variants of tCsmA

Name	Sequence
GAMG-tCsmA	GAMG-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG
M-tCsmA	M-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG
tCsmA	FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG
G-tCsmA	G-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG
MG-tCsmA	MG-FTDILAAAGRIFEVMVEGHWETVGMLFDSLGKGTMRINRNAYG

Table 5.2 Buffers, proteins and protease inhibitors

Buffer	Components
NPI-10	50 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazole, pH 8.0
NPI-250	50 mM NaH_2PO_4, 300 mM NaCl, 250 mM Imidazole, pH 8.0
Sol Buffer	50 mM Tris-HCl, 6 M 1-propanol, 2 M Urea, 0.5 mM EDTA, pH 8.5[19]
MBP-Lysis Buffer	20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol, pH 7.5
MBP-Column Buffer	20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol, pH 7.5
MBP-Column Elution Buffer	MBP-Column Buffer with 10 mM Maltose
Phosphate Buffer (Kpi)	50 mM KPi buffer pH 7.9 or pH 7.5
Propanol-Buffer	6M 1-propanol, 50 mM KPi pH 7.9. Phosphate Buffer: 50 mM KPi pH 7.9 or pH 7.5
Phosphate Buffer	50 mM KPi pH7.9 or pH7.5
TEV Buffer	50 mM Tris-HCl, 0.5 mM EDTA, 14 mM 2-mercaptoethanol, pH 8.0
2x Osmosis buffer	40 mM Tris-HCl, 5 mM EDTA, pH 8.0
Sucrose buffer	40 % m/v sucrose in MQ water and 2x Osmosis buffer

For bacterial protein extraction, B-PER reagent in the KPi buffer and Inclusion Body Solubilisation Reagent (IBS) were purchased from Thermo Scientific.

Pefabloc SC (Sigma-Aldrich): a stock of 200 mM in MQ was stored at 4°C and used at 1 mM final concentration, which was added to prevent proteolytic degradation of peptides. Complete (Roche) 25× stock was made by dissolving 1 tablet in 2 mL MQ and stored at -20 °C. One mini Complete Ultra EDTA-free tablet (Roche) was dissolved in 10 mL Lysis Buffer and used immediately.

AcTEV protease 10 u/µL was purchased from In Vitrogen. For cleavage 50 units/mg MBP-CsmA fusion protein was used overnight at 22 °C in the TEV Buffer. Lysozyme stock was freshly prepared 50 mg/mL in the MBP-Lysis buffer.

ZnPP was bought from Frontier Scientific. Commercial tCsmA was synthesized by GL Biochem Ltd. (Shanghai).

Equipment

Cell-free *in vitro* protein synthesis kit (PurExpress) and Gibson Assembly Master Mix (Bioké); Slide-A-Lyzer Dialysis cassettes G2 (3 mL and 0.5 mL MWCO 3.5 kDa) and Pierce 96-well Micro-dialysis Plate (3.5K MWCO) (Thermo Scientific); Amicon Ultra-0.5ml and Ultra-2 mL centrifugal filters Ultracel-3K (Sigma-Aldrich); 1 mL Ni-NTA columns (Qiagen); Superdex 75 Increase 10/300 GL and 5 mL MBPTrap HP columns (GE Healthcare); DeNovix DS-11 nanodrop spectrophotometer (DeNovix); Size Exclusion Chromatography (SEC) was done using an Äkta Explorer 10S and all other columns were used on an Äkta Start purification system (GE Healthcare).

5.2.2 Methods

The cells were grow in LB medium with 50 µg/mL ampicillin at 37 or 18 °C in an orbital shaker at 150 rpm. The frozen glycerol stocks of transformed cells were diluted 1 : 100 to grow with different culture time. When OD600 reached about 1.0, expression of the opsin was induced by the addition of IPTG to a final concentration of 1 mM. The detail of culturing and further processes of different tCsmA fusion proteins are shown as below.

Untagged-tCsmA and His-tagged tCsmA
His-tagged tCsmA producing plasmid pET28a-tCsmA was transformed into *E.coli* strains BL21 (DE3) and BL21 (DE3)/pLys. Both strains were induced and cultured at 37 °C for 2h and 18 °C overnight.

SUMO-tagged tCsmA
pET16SK-SUMO2-tCsmA was derived from MBP1b-tCsmA by PCR using primers sumoFW (CAGGGTCTCGCATGGGCTTTACCGATATTCTGGCAGCAGC) and sumoREV (GAGGGATCCTTTCAGCAAAAAACCCCTCAAGACCC). The resulting sequence is MG-CsmA.

GAMG-tCsmA, tCsmA and G-tCsmA from MBP tags
pETMBP1a-tCsmA (Fusion 6xHis-MBP-TEV-tCsmA) and pETMBP1b-tCsmA (fusion MBP-6xHis-TEV-tCsmA) produce the GAMG-tCsmA protein after TEV cleavage. pETMBP1b-ΔGAM-tCsmA and pETMBP1b-ΔGAMG-tCsmA were derived from pETMBP1b-tCsmA using Gibson assembly following the manufacturers protocol resulting in G-tCsmA and tCsmA respectively.

The pellet from 1 L cells was lysed in 22 mL MBP-Lysis Buffer containing 1 mg/mL lysozyme, 1 mM Pefabloc and 2 mM DTT. After 10 minutes incubation and 4 minutes sonication the lysate was centrifuged for 30 min at 100,000 × *g* using a 70 Ti rotor (Beckman Coulter). The supernatant was loaded and washed on a 5 mL MBP-Trap column equilibrated with the MBP-Column Buffer at 4 °C. The protein was collected by isocratic elution using the MBP-Column Buffer containing 10 mM Maltose. The fractions collected were dialyzed against 1 L TEV Buffer containing 14 mM 2-mercaptoethanol at 4 °C for 2 hours and one more time for the overnight. The purified proteins were flash frozen with nitrogen and stored at -80 °C.

After TEV cleavage GuHCl was added to get a final concentration of 6 M. Then the sample was put on a rotary shaker for 90 minutes. The 480 µL samples were injected using a 0.5 mL loop on a Superdex 75 SEC column equilibrated in the TEV buffer containing 6 M GuHCl and eluted at 0.6 mL/min to collect 0.2 mL fractions. The collection was concentrated to a final volume of 0.5 mL by Amicon-2ml centrifugal filter to dialyze and obtain the precipitated tCsmA. The sample was dialyzed first using 0.5 mL Slide-A-Lyzer in 2 L 10 mM KPi buffer pH 5.0 at room temperature for 2 hours and one more time overnight. The final dialysis proceeded for 6 hours into the next day. The sample was centrifuged 30 minutes at 30,000 × *g*,

washed with 500 µL pH 5.0 KPi buffer one time, and 500 µL pH 7.9 KPi buffer for two more times. The final pellet of protein was freeze-dried overnight.

GAMG-tCsmA from TRX tag

pETTRX1a-tCsmA (fusion 6xHis-TRX-TEV-tCsmA) and pETTRX1b-tCsmA (fusion TRX-6xHis-TEV-tCsmA) were derived from pETMBP1a-CsmA using NcoI and KpnI. The resulting protein is GAMG-tCsmA. The inclusion bodies were purified from 500 mL IPTG induced cell culture and divided into 40 mL batches. The pellets were stored at -80 °C. The Osmotic lysis and B-PER lysis were both used for the purification from inclusion bodies.

For Osmotic lysis, the cell pellet was washed gently by cold 10 mL 10 mM pH 7.0 Tris buffer containing 30 mM NaCl. The supernatant was removed after 10 minutes centrifuge at 4000 rpm. The pellet was resuspended with 2mL sucrose buffer, and the supernatant was removed by 4 minutes centrifugation at 15,000 × *g* after 10 minutes incubation on ice. The pellet was resuspended by 2 mL Osmosis buffer. After 10 minutes incubation on ice, 80 µL complete stock was added to the solution. In the final step, after 4 minutes centrifugation at 15,000 × *g* the supernatant and pellet were separated.

For B-PER lysis, the pellet was dissolved in cold 5 mL B-PER containing Complete and was set on a rotary shaker for 10 minutes incubation on ice. The pellet was resuspended in 5 mL B-PER with 0.2 mg/mL lysozyme solution after 15 minutes centrifugation at 15,000 × *g* at 4 °C. The suspension was further incubated 5 min at room temperature after shaking vigourously for 1 min using a vortexer. 15 mL 1:10 diluted B-PER in MQ water was added and the suspension was centrifuged for 15 min at 15,000 × *g* at 4 °C after 1 minute. The pellet was resuspended in 20 mL 1:10 B-PER and centrifuged three times. In the final step, the inclusion body pellet (26 mg) was resuspended for 30 min with a rotary shaker in either 210 µL IBS with EDTA-free Complete or Sol-Buffer.

SDS-PAGE

Two types of SDS-PAGE gel were used in this chapter. Tris-glycine gel is used for separating protiens >30 KDa, whereas Tris-tricine gel is used for proteins <30 KDa. Aliquots of samples were diluted with SDS sample buffer and incubated for 30 min at 37 °C, then run on polyacrylamide gel at 20 mA for 2 hours. The gel was stained using Coomassie brilliant-blue G for 2 hours and

destained overnight in a solution which contains methanol, glacial acetic acid and water (volume ratio = 1 : 1 : 8).

5.3 Results and discussion

5.3.1 Protein expression

In Figure 5.1 the gels for expression of His-tagged tCsmA, and the untagged construct are shown. His tag is the smallest tag and although it may not be the most suitable for expression of tCsmA, it was tried to verify that expression cannot be realized by standard methods and more complicated approaches are warranted. The gels in Figure 5.1A and C are essentially identical, and in particular there is no evidence for the presence of the His-tCsmA. Likewise, SUMO-tag expression turned out to be unsuccessful (data not shown). In contrast, when the same plasmid for untagged-tCsmA was used in the *in vitro* protein synthesis system, protein expression could be observed (Figure 5.1B, Lane 2). This indicates that *in vivo* the small tCsmA peptide is probably digested by proteases. To overcome this problem, larger protective tags such as TRX and MBP tag were considered to obtain tCsmA *in vivo*.

An example of a tag that has been used successfully in the past for protein expression is TRX.[20] The most important advantage of this tag is enhanced solubility. In adddition this 12 kDa tag is much larger than His (1 kDa), which implies that it can protect the tCsmA better against digestion. As Figure 5.2A shows, pETTRX1b-tCsmA (tCsmA gene is behind TRX gene) and pETTRX1a-tCsmA (tCsmA gene is in front of TRX gene) both are able to produce the tCsmA. However, the protein obtained from the pETTRX1b-tCsmA plasmid was partially degraded, there are two protein protection bands which present TRX-tCsmA and TRX protein respectively. The degraded TRX protein fraction band is labelled with a star in Figure 5.2A Lane 2. The pETTRX1a-tCsmA on the other hand remains intact and was used for further steps.

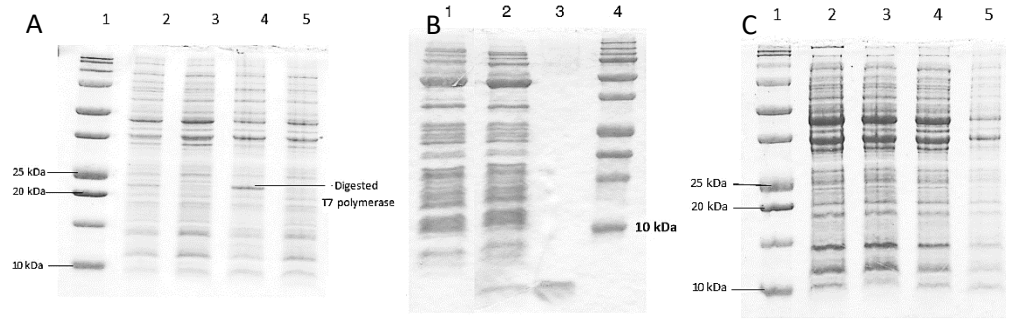


Figure 5.1 A : Protein gel showing the expression of His-tCsmA from BL21 His-tCsmA at 37 °C. Lane 1: Tricine gel containing MW marker; Lane 2: cell pellets from BL21 His-tCsmA 37 °C induced 0 h; Lane 3: cell pellets from BL21 His-tCsmA pLys 37 °C induced 0 h; Lane 4: cell pellets from BL21 His-tCsmA 37 °C induced 2 h; Lane 5: cell pellets from BL21 His-tCsmA pLys 37 °C induced 2 h.
B : Protein gel showing the *in vitro* translation tCsmA proteins. Lane 1: without plasmid; Lane 2: pET3-tCsmA; Lane 3: synthetic tCsmA peptide; Lane 4: MW marker.
C : Protein gel showing the expression of His-tCsmA from BL21 His-tCsmA at 18 °C. Lane 1: Tricine gel containing MW marker; Lane 2: cell pellets from BL21 His-tCsmA 18 °C induced 0 h; Lane 3: cell pellets from BL21 His-tCsmA pLys 18 °C induced 0 h; Lane 4: cell pellets from BL21 His-tCsmA 18 °C induced 2 h; Lane 5: cell pellets from BL21 His-tCsmA pLys 18 °C induced 2 h.

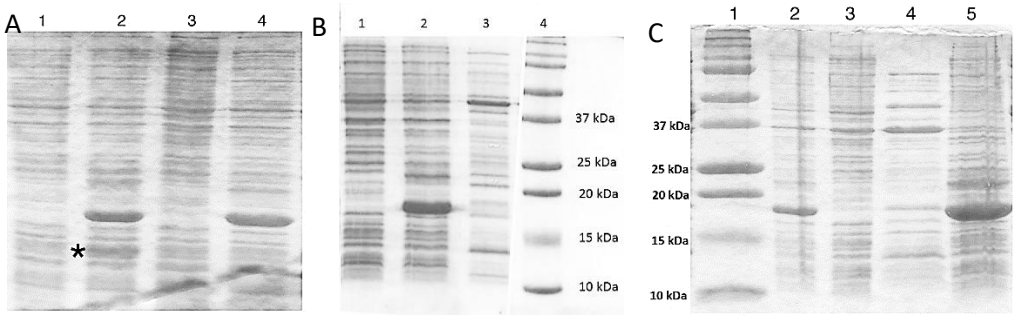


Figure 5.2 A : Protein gel showing the expression of TRX-tCsmA fusion protein. Lane 1: pETTRX1b-tCsmA -IPTG; Lane 2: pETTRX1b-tCsmA +IPTG; Lane 3: pETTRX1a-tCsmA -IPTG; Lane 4: pETTRX1a-tCsmA +IPTG.
B : Protein gel showing the effect of osmotic lysis. Lane 1: BL21(DE3)/pETTRX1a-tCsmA -IPTG; Lane 2: BL21(DE3)/pETTRX1a-tCsmA +IPTG; Lane 3: supernatant; Lane 4: MW marker.
C : Protein gel showing the effect of B-PER lysis. Lane 1: MW marker; Lane 2: BL21(DE3)/pETTRX1a-tCsmA +IPTG; Lane 3: BL21(DE3)/pETTRX1a-tCsmA -IPTG; Lane 4: supernatant; Lane 5: pellet.

The Figure 5.2B and C show the results of lysis. The fusion protein from the pETTRX1a-tCsmA plasmid only appears in the pellets obtained by centrifugation after cell lysis using B-PER and not in the supernatant (Figure 5.2C Lanes 4 and 5). Osmotic lysis gives similar results (Figure 5.2B, pellet was not shown). Thus, the TRX tag protects the tCsmA from degradation, however, it is not able to solubilize the peptide in the buffer.

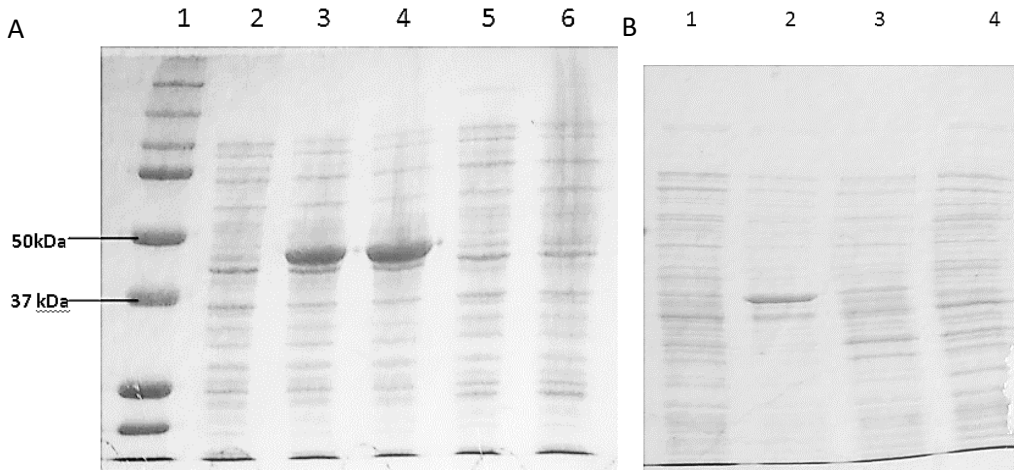


Figure 5.3 A : Protein gel showing the expression of MBP-tCsmA fusion protein in different strains. Lane 1 : MW marker; Lane 2, 3 and 4 : cell pellets from BL21 pET-MBP1a-tCsmA induced for 0 h, 1 h and 2 h; Lanes 5 and 6 : BL21 pET-MBP1a-tCsmA pLys induced for 0 h and 1 h respectively.
B : Protein gel showing the expression of MBP-tCsmA fusion protein in BL21 pET-MBP1a-tCsmA at different temperature. Lane 1 and 2 : cell pellets from BL21 pET-MBP1a-tCsmA induced at 37 °C for 0 h and 2h; Lane 3 and 4 : cell pellets from BL21 pET-MBP1-CsmA induced at 18 °C 0 h and the overnight respectively.

In Figure 5.3A, an overproduction band that is clearly present in Lane 3 and 4, corresponds with the molecular weight of MBP-tCsmA of 45 kDa. This indicates a successful expression of induced BL21(DE3) pET-MBP1a-tCsmA (tCsmA gene is behind MBP gene) in *E.coli* BL21 (DE3) strain at 37 °C. In contrast, there is no evidence for overproduction in the culture of BL21 (DE3)/pLys strain at 37 °C (Figure 5.3A Lane 5 and 6).

Two overproduction tests were performed for the BL21 (DE3) pET-MBP1a-CsmA strain at both 37 °C for 2h (Figure 5.3B Lane 1 and 2) and at 18 °C overnight (Figure 5.3B Lane 3 and 4). There appears to be no

overproduction band at 18 °C, and we take 37 °C as a suitable temperature for the culture of BL21 (DE3) pET-MBP1a-tCsmA.

5.3.2 Protein purification

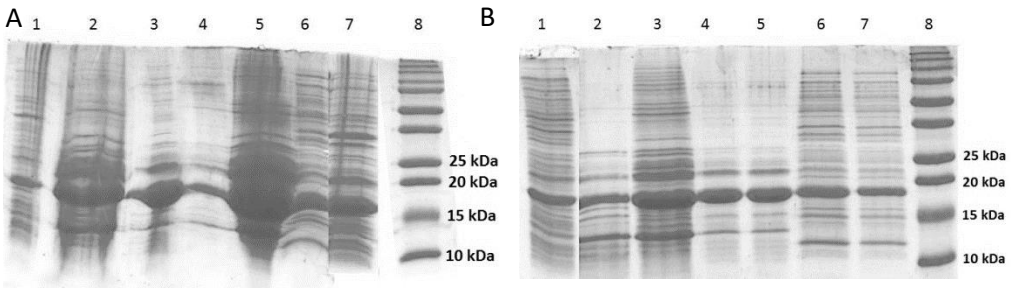


Figure 5.4 A : Protein gel showing the extraction of TRX-tCsmA fusion protein by B-PER lysis and osmotic lysis. Lane 1: BL21(DE3)/pETTRX1a-tCsmA cells +IPTG; Lane 2: supernatant after treatment of the inclusion body pellet with IBS buffer from B-PER lysis; Lane 3: supernatant after treatment the inclusion body pellet with Sol-Buffer from B-PER lysis; Lane 4: pellet after treatment of the inclusion body pellet with IBS buffer from B-PER lysis; Lane 5: pellet after treatment of the inclusion body pellet with Sol-buffer from B-PER lysis; Lane 6: supernatant after treatment of the inclusion body pellet with Sol-buffer from Osmotic lysis; Lane 7: pellet after treatment of the inclusion body pellet with Sol-buffer from Osmotic lysis; 8 : MW marker.
B : Protein gel showing the solublization of TRX-tCsmA fusion protein after extraction. Lane 1: BL21(DE3)/pETTRX1a-tCsmA cells +IPTG; Lane 2: supernatant of TRX-tCsmA after dialysis in the propanol buffer (solubilized by the IBS buffer from B-PER lysis); Lane 3: supernatant of TRX-tCsmA after dialysis in 2 M Urea, 50 mM KPi pH 7.5 (solubilized by IBS buffer from B-PER lysis); Lane 4: supernatant of TRX-tCsmA after dialysis in 6 M Urea (solubilized by the Sol buffer from the B-PER lysis); Lane 5: TRX-tCsmA from Lane 6 after dialysis in 2 M Urea, 50 mM KPi pH 7.5; Lane 6: supernatant of TRX-tCsmA after dialysis in the propanol buffer (solubilized by Sol buffer from Osmotic lysis); Lane 7: supernatant of TRX-tCsmA after dialysis in 2 M Urea, 50 mM KPi pH 7.5 (solubilized by the Sol buffer from Osmotic lysis); Lane 8 : MW marker.

To purify the tCsmA from TRX fusion protein, the first step is to extract the fusion protein in a soluble state to continue the cleavage. In Figure 5.4, most fusion protein is present in the supernatant after the treatment with IBS buffer (Figure 5.4A Lane 2) and Sol-Buffer (Figure 5.4A Lane 3). According to the width of protein bands, the IBS buffer dissolved more fusion protein

than the Sol-buffer, but also contained more contaminants (Figure 5.4A Lane 2 and 3).

Propanol buffer and urea were also used to solubilize the fusion proteins. According to the Figure 5.4B, the TRX fusion protein from the IBS buffer treatment remained mostly in the supernatant after dialysis in 2 M Urea, 50 mM KPi pH 7.5 (Figure 5.4B Lane 6). The fusion proteins from Sol buffer contain much fewer contaminants, but more protein precipitated during the dialysis (Figure 5.4A Lane 3 and Lane 6). After B-PER lysis more TRX fusion protein was recovered than for Osmotic lysis (Figure 5.4A Lane 2 and Figure 5.4B Lane 2). Although the dialysis extraction from the IBS and Sol buffer with urea both yield encouraging results, for TEV cleavage another step, involving dialysis against the KPi buffer, is needed. Unfortunately, here all fusion protein precipitated.

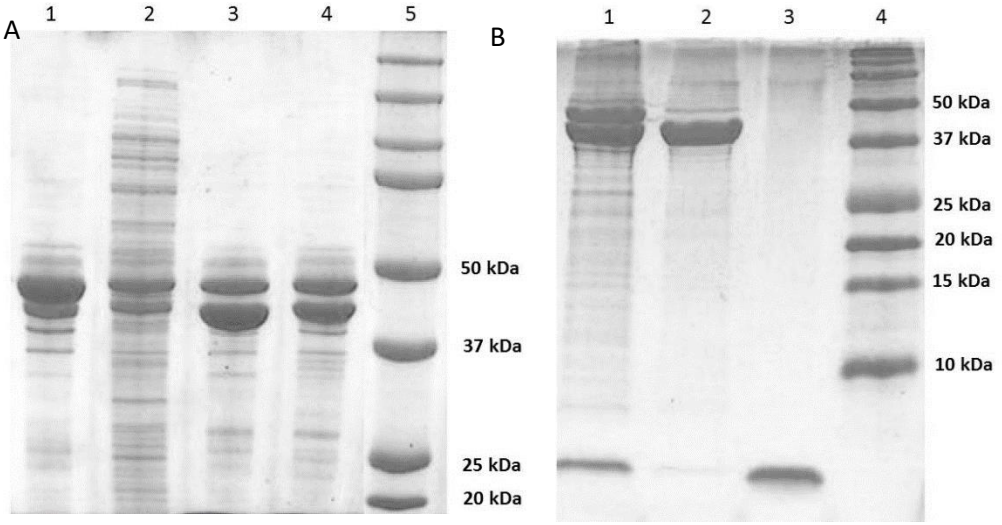


Figure 5.5 A : Protein gel showing the cleavage effect of MBP-tCsmA fusion protein. Lane 1: MBP1a-tCsmA before cleavage; Lane 2: MBP1b-tCsmA before cleavage; Lane 3: MBP1a-tCsmA after cleavage; Lane 4: MBP1b-tCsmA after cleavage; Lane 5: MW marker.
B : Protein gel showing the cleaved MBP1b-tCsmA fusion protein in Tricine gel to show the amount of tCsmA. Lane 1: pellet from MBP1b-tCsmA after cleavage; Lane 2: supernatant from MBP1b-tCsmA after cleavage; Lane 3: 1 µg commercial tCsmA; Lane 4: MW marker.

In contrast with the TRX tag, the MBP tag is able to solubilize tCsmA in the lysis buffer. MBP1a-tCsmA and MBP1b-tCsmA both show broad cleaved

protein bands which indicates a good cleavage efficiency (Figure 5.5A). Although a minor fraction of the tCsmA precipitated during cleavage, the cleaved tCsmA hardly shows bands below 10 KDa, probably most of the cleaved tCsmA still remains associated with the MBP protein in solution (Figure 5.5B Lane 1 and Lane 2).

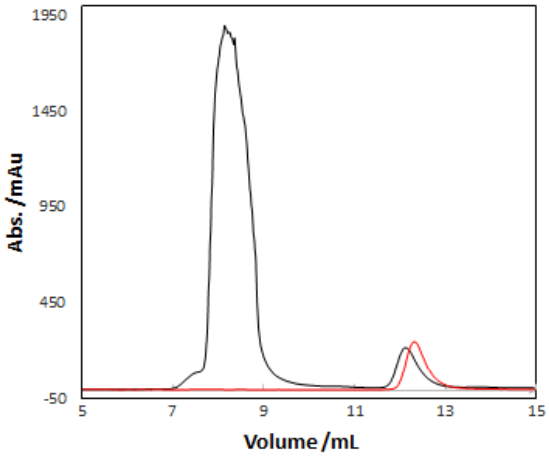


Figure 5.6 SEC elution profile (Superdex 75 10/300) of recombinant MBP-GAMG-tCsmA after TEV cleavage and commercial tCsmA peptide in the TEV Buffer containing 6 M GuHCl. red: 480 µg commercial tCsmA peptide (4.8 kDa), Ve=12.3 mL; Black: 4.8 mg MBP-GAMG-tCsmA treated with TEV protease. At Ve=8.5 mL the MBP tag (48 kDa) and non-cleaved MBP-GAMG-tCsmA (53.1 kDa) are eluting. At Ve=12.1 mL GAMG-tCsmA (5.1 kDa).

Using a Superdex-75 column with Size Exclusion Chromatography (SEC) the two proteins could be separated in the presence of GuHCl. In Figure 5.6 the SEC trace for the cleavage product is compared with benchmark data collected from a sample of synthetic tCsmA. This provides convincing evidence that the tCsmA can be separated from the MBP by SEC. However, since the running buffer of SEC contains 6 M GuHCl, another dialysis step is needed to collect the tCsmA fraction from the SEC buffer. This last step is very inefficient, with very low yield. When 300 µg of synthetic tCsmA is dissolved in the GuHCl buffer and dialyzed against the KPi buffer, ~ 160 µg is recovered after freeze drying the precipitated peptide and redissolving in the propanol buffer, while only 20 µg is recovered when a SEC step is performed in between. Since most protein was lost during the dialysis after SEC purification, most probably the SEC protocol, with purification in GuHCl

after SEC, leads to denaturation and unfolded protein, which passes through the membrane of the dialysis cassette.

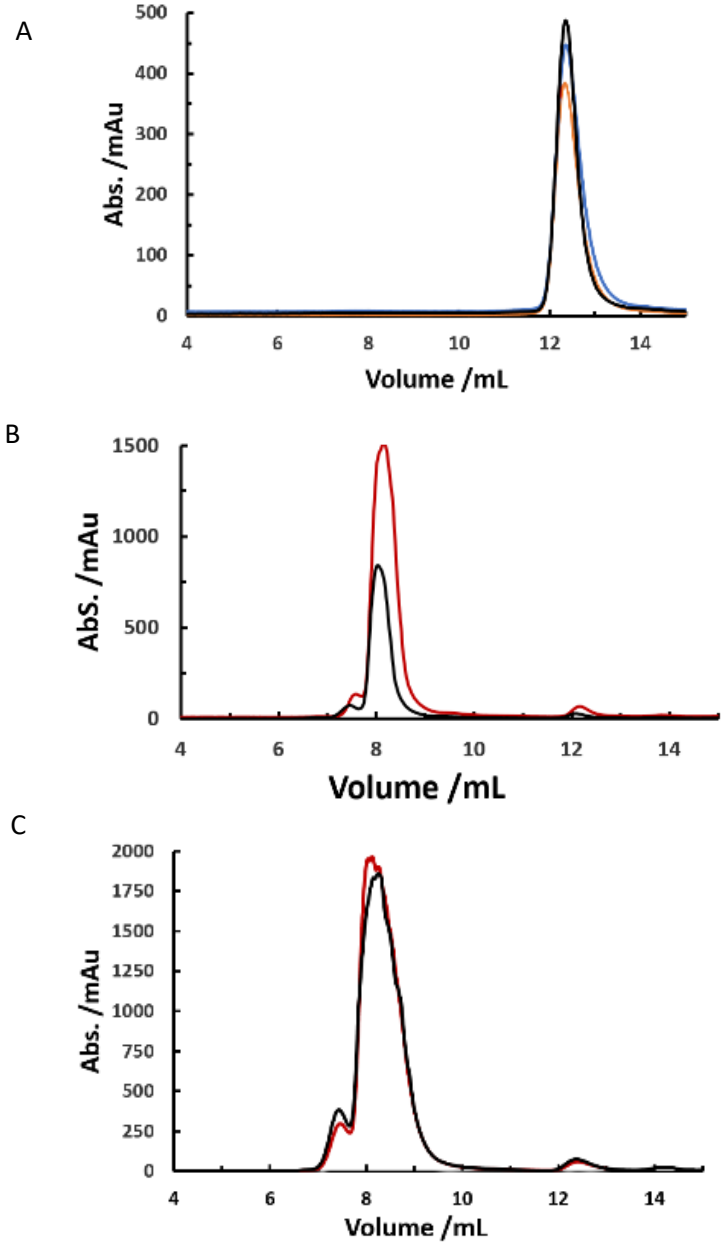
5.4 Conclusions

The tCsmA degrades during expression without the assistance of big fusion tags. TRX and MBP tags both prevent degradation of tCsmA. However, for the TRX tag, the inclusion bodies containing the fusion protein precipitate during expression. Although it is possible to extract the fusion protein in the SOL buffer, this is not a good environment for cleavage. Although the MBP tag is able to solubilize the tCsmA and the SEC can separate the tCsmA and MBP with the help of 6 M GuHCl, during the purification, the tCsmA was denatured too much, and the stretched tCsmA was lost in dialysis. According to UV analysis, a small fraction of purified tCsmA peptides can be recovered, around 10% of the total amount of protein after cleavage.

5.5 References

1. Ganapathy, S.; Oostergetel, G. T.; Wawrzyniak, P. K.; Reus, M.; Gomez Maqueo Chew, A.; Buda, F.; Boekema, E. J.; Bryant, D. A.; Holzwarth, A. R.; de Groot, H. J., *Proc Natl Acad Sci U S A* **2009,** *106* (21), 8525-30.
2. Yalamanchili, S.; Emmer, H. S.; Fountaine, K. T.; Chen, C. T.; Lewis, N. S.; Atwater, H. A., *Acs Photonics* **2016,** *3* (10), 1854-1861.
3. Hu, S.; Chi, C. Y.; Fountaine, K. T.; Yao, M. Q.; Atwater, H. A.; Dapkus, P. D.; Lewis, N. S.; Zhou, C. W., *Energ Environ Sci* **2013,** *6* (6), 1879-1890.
4. Gao, L.; Cui, Y. C.; Wang, J.; Cavalli, A.; Standing, A.; Vu, T. T. T.; Verheijen, M. A.; Haverkort, J. E. M.; Bakkers, E. P. A. M.; Notten, P. H. L., *Nano Lett* **2014,** *14* (7), 3715-3719.
5. Katterle, M.; Prokhorenko, V. I.; Holzwarth, A. R.; Jesorka, A., *Chemical Physics Letters* **2007,** *447* (4-6), 284-288.
6. Fry, H. C.; Garcia, J. M.; Medina, M. J.; Ricoy, U. M.; Gosztola, D. J.; Nikiforov, M. P.; Palmer, L. C.; Stupp, S. I., *J Am Chem Soc* **2012,** *134* (36), 14646-14649.
7. Luo, L.; Chang, C. H.; Chen, Y. C.; Wu, T. K.; Diau, E. W., *J Phys Chem B* **2007,** *111* (26), 7656-64.
8. Chromiński, M.; ó Proinsias, K.; Martin, E.; Gryko, D., *European journal of organic chemistry* **2013,** *2013* (8), 1530-1537.
9. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.
10. Warren, B. E., *X-ray Diffraction*. Courier Corporation: 1990.
11. Milne, J. L.; Borgnia, M. J.; Bartesaghi, A.; Tran, E. E.; Earl, L. A.; Schauder, D. M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S., *The FEBS journal* **2013,** *280* (1), 28-45.
12. Cheng, Y.; Grigorieff, N.; Penczek, P. A.; Walz, T., *Cell* **2015,** *161* (3), 438-449.
13. Coates, G. R.; Xiao, L.; Prammer, M. G., *NMR logging: principles and applications*. Haliburton Energy Services Houston: 1999; Vol. 344.
14. Synthesis—A, F. S. P. P., Practical Approach. Chan, WC: 2000.
15. Fields, G. B., *Peptide characterization and application protocols*. Humana Press: 2007; Vol. 386.
16. Miroux, B.; Walker, J. E., *Journal of molecular biology* **1996,** *260* (3), 289-298.
17. Studier, F. W., *Methods Enzymol.* **1990,** *185*, 60-89.
18. Chen, B. P.; Hai, T., *Gene* **1994,** *139* (1), 73-75.
19. Singh, S. M.; Sharma, A.; Upadhyay, A. K.; Singh, A.; Garg, L. C.; Panda, A. K., *Protein expression and purification* **2012,** *81* (1), 75-82.
20. Hirota, K., Nakamura, H., Masutani, H., & Yodoi, J. *Annals of the New York Academy of Sciences.* **2002,** 957(1), 189-99.

5.6 Appendix



SFigure 5.1 Size exclusion chromatography elution profile (Superdex 75 10/300) of recombinant MBP-GAMG-tCsmA after TEV cleavage and synthetic CsmA peptide in TEV Buffer containing 6 M GuHCl. A : commercial tCsmA with different amounts (350, 400 and 450 µL respectively); B : MBP1a-tCsmA after TEV cleavage (1 mg and 2 mg respectively); C : 3 mg MBP1b-ΔGAM-tCsmA after TEV cleavage (two times

repeat).

The accuracy of the SEC was confirmed by different experiments (SFigure 5.1). As SFigure 5.1A shows, when we injected different volumes of commercial tCsmA solution, the intensity is proportional to the volume injected. For the MBP-tCsmA sample, it also shows same result (SFigure 5.1B). The results are repeatable, the duplo sample showed the overlap in SEC spectrum (SFigure 5.1C). It indicates that the SEC is a precise tool to compare the amounts of proteins.

Chapter 6

General discussion and outlook

6.1 Structure assessment

Advances in the constitution of the self-assembly of a semi-synthetic zinc-protoporphyrin IX (ZnPP) peptide (tCsmA) complex derived from the natural CsmA light harvesting antenna has led to more insight in the responsive chiral matrix in the photosystems. The tCsmA-ZnPP complex is able to self-assemble to form extended periodic 7_2 helical filaments after dialysis in KPi buffer (Chapter 2 and Chapter 4). The 7_2 helical structure of the self-assembled tCsmA-ZnPP complex is resolved at about 7 Å resolution by the cryo-EM technique. The tCsmA peptide, which contains 43 amino acids, is derived from the CsmA scaffolding peptide in the baseplate of *C. tepidum* by removing the flexible segments from the N and C-termini. The tCsmA successfully maintains the functional central α-helix (Chapter 3). The ZnPP coordinate to the only histidine of tCsmA. For the tCsmA-ZnPP complex in KPi buffer, each strand prefers to twist to form a stable helical structure. As a result, the hydrophobic side of the tCsma-ZnPP structural motif in the helical fibrils are rotated nearly 90° compared to the CsmA in the planar baseplate in the natural system. The tCsmA forms a hydrophobic core which leads the ZnPP to locate at the inside of the peptide scaffold. The charged residues are exposed towards the outside and can form salt bridges with other filaments to form bundles, which was shown in Chapter 3.

The electron density map for the monomeric building block shows continuous density of protein, which indicates an extended alpha helical structure without any loop regions. Therefore the loop from residue 30 to residue 35 that was observed in NMR data collected from CsmA in solution probably transforms to an extended helical structure for the tCsmA-ZnPP complex self-assembly. Each strand in the 7_2 helix is composed of dimeric tCsmA-ZnPP complexes in an antiparallel arrangement with P2 symmetry. The two ZnPP in each dimeric unit are close to the neighbouring ZnPP from a building block in the other strand and form a stable tetrameric unit.

Different from the baseplate structure, for the 7_2 helix salt bridges are of little importance for the self-assembly. The packing strongly suggests that dipole-dipole interactions involving ZnPP may play a role in the self-assembly.

The electron density map also shows a small bump around the ZnPP, which can be one of the propionic acid groups -CH_2-CH_2-COOH (Chapter 4). A rigid density indicates that the propionic acid group may interact with the peptide and is fixed by *e.g.* a hydrogen bond interaction. The tCsmA-ZnPP complex shows some similarity with the natural CsmA-BChl *a* complex. The distance between the ZnPP in the dimeric units also is about 21.1 Å. Moreover, two complexes cross with a similar angle as for the natural system.[1]

Although Tris buffer cannot be used as the dialysis buffer for the complex constitution, the complex obtained in KPi buffer remains intact when it is dialyzed against Tris buffer. During the buffer change, most filaments turn into smaller pieces without degradation. However, part of the structure of helix transforms to a planar structure, which resembles the natural CsmA baseplate antenna. The distance between strands is found to be same as for the natural baseplate, which is about 31.3 Å. The ZnPP are also embedded in the middle of the strands. Their macrocycles are oriented perpendicular to the plane of the assembly. We attribute these planar structures to untwisting of helical filaments upon changing the buffer. While in the helix the ZnPP interact to form a tetramer, in the planar form the ZnPP form extended linear arrays of dimeric ZnPP in between two tCsmA strands (Chapter 3). The optical data indicate that the ZnPP may be less planar in Tris buffer relative to KPi buffer. Although the packings of the two types of self-assembly are different, planar versus helical, their CD responses strongly overlap (Chapter 3). That indicates that overall the tCsmA helical peptide undergoes little conformational change of the secondary structure on the peptide level upon buffer change.

6.2 Constitution

We tested different truncated CsmA homologues for the constitution with ZnPP. However, only tCsmA prepared by N-terminal acetylation and C-terminal amidation to replace the flexible segments in the natural peptide is able to bind with ZnPP and form a large self-assembly. It is probably the

net charge of the peptides that influences the constitution (Chapter 3). Because the basic nitrogen atom in the imidazole of histidine with a lone electron pair can coordinate to metallic cations like the Zn^{2+} of ZnPP when pH>7. The binding is efficient at higher pH of the dialysis buffer, up to pH 7.9.

The molar extinction coefficient of the ZnPP in complex with tCsmA is 117400 M^{-1} cm^{-1} with a stoichiometry of tCsmA : ZnPP ~ 1 : 1 (Chapter 2). The constitution can be done in 1-propanol KPi buffer or in KPi buffer containing detergents with high CMC such as HG or OG detergents. Compared to the 1-propanol KPi buffer, the detergents buffers lead to a slow dialysis process but higher binding efficiency (Chapter 3).

6.3 Photochemical relevance

The photo-degradation occurs during a prolonged light exposure. The complex can be dissociated when it is exposed to sunlight for a few hours. The reason for the photo-dissociation might be that the dissipation of heat is slow. However, despite photo-dissociation on longer time scales, the tCsmA-ZnPP complex is able to maintain a short time photostability. It occurs when a high concentration more than 20 µM of the complex is excited by 430 nm UV light. Recently, it has been proposed that light-driven conformational changes in chiral protein scaffolds can develop vibronic states within or between linear arrays. It may suggest that distribution of energy helps to dissipate the excess energy and may be of vital importance for further improvement of the stability (Chapter 2).

In Chapter 2, the tCsmA-ZnPP complex shows 594 nm and 650 nm emission peaks after excitation. The MV^+ has a strong broad absorption at 606 nm (ε = 13700 M^{-1} cm^{-1}). When the acceptor cation radical methyl viologen (MV^+) was added in the complex solution, the harvested light from the tCsmA-ZnPP complex can be 100% quenched by MV^+. This leads us to conclude that intermolecular FRET or Dexter energy transfer between ZnPP and MV^+ may cause fluorescence quenching by energy transfer due to overlapping spectra. On the other hand, the quenching is dependent on the concentration of MV^+. When the concentration was increased to about 300 µM MV^+ in the tCsmA-ZnPP solution, nearly all fluorescence of ZnPP was quenched.

The fluorescence decay responses of the tCsmA-ZnPP complex at 594 nm and 650 nm can be described with 3 different lifetime components. It is found that all three lifetimes of the complex with the MV^+ additive are shortened, especially τ_3 was dramatically shortened to ~1 ps and the energy is almost 100% quenched by the acceptor. This is comparable to natural light-harvesting antenna systems (Chapter 2).

The MV^+ in solution was obtained with the help of sodium dithionite. The sodium dithionite is able to offer one electron to MV^{2+} to form MV^+. When sodium dithionite is depleted by letting it react with oxygen, the MV^+ is converted back to MV^{2+}. When MV^+ is oxidized to MV^{2+}, it shows no fluorescence quenching any more. Moreover, the quenched complex recovered the same fluorescence as without acceptor. It indicates that the complex still maintains its integrity upon photoexcitation (Chapter 2). This highlights the importance of the MV^+ as the quenching species.

6.4 Future experiments

6.4.1 Functional exploration

In Chapter 2, the tCsmA-ZnPP complex shows the ability of light-harvesting and energy transfer. Further experiments can be done to explore how to increase the stability of the complex, and to improve the efficiency and mechanism of energy transfer. This will be helpful for the design of semisynthetic systems for light harvesting and their application in solar fuel cells. The Cyclic Voltammetry (CV) measures the current that develops in an electrochemical cell under conditions where voltage is in excess of that predicted by the Nernst equation.[2-4] Therefore CV experiments are able to observe the electron transfer to check the functional relevance of the tCsmA-ZnPP complex for charge transfer.[5-6] For the further application, it is also interesting to attach the tCsmA-ZnPP complex to an electrode and build a prototypic solar cell device, as schematically depicted in Figure 6.1.

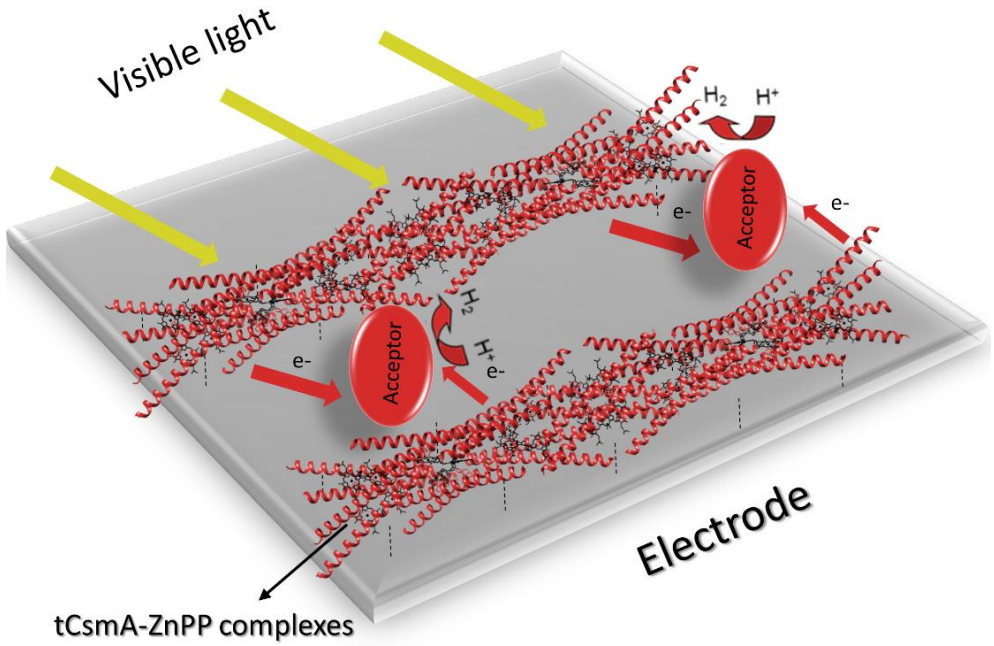


Figure 6.1 Schematic design of a solar cell device. The tCsmA-ZnPP complex could be modified to attach to the electrode surface.

6.4.2 Atomic structure determination

Despite good progress, atomic information was not yet obtained. There are two methods, NMR and cryo-EM, that can be used for resolving structure at high resolution. For NMR, some sample preparation work has been done in Chapter 5. It shows that without a large fusion tag, tCsmA rapidly degrades. In addition, TRX and MBP tags are used to get protein. However, with TRX the stability can be increased, but the sample precipitates making further treatment difficult. With the MBP tag a strongly associated soluble tCsmA-MBP construct is obtained, but only little protein can be purified. Thus, to continue the further experiments, more tags should be tried. Since preliminary data have shown that the peptides are expressed in vitro and can be purified without degradation, this may be the best route to follow for further optimization. Before the purification of recombinant variants of truncated CsmA, the sequence should be modified and optimized for constitution. As the data in Chapter 3 shows, the net charge of the protein influences the binding. The truncated CsmA are derived from the natural

CsmA by cutting off the flexible segments. The net charge is changed by the truncation. Therefore the sequence should be optimized first.

For cryo-EM, the low resolution is caused by a thick ice layer. This results in a high scattering mass and more noise in comparison to thin layer samples.[7] Hence the sample preparation process can be optimized for better resolution. In particular, the sample can be improved to obtain a thinner ice layer with a longer blotting time or by adjusting the humidity during the blotting. During the image collection, a higher dose of electrons can be tried to get more contrast.

6.5 References

1. Nielsen, J. T.; Kulminskaya, N. V.; Bjerring, M.; Linnanto, J. M.; Rätsep, M.; Pedersen, M. Ø.; Lambrev, P. H.; Dorogi, M.; Garab, G.; Thomsen, K., *Nature communications* **2016,** *7*, 12454.
2. Nicholson, R. S.; Shain, I., *Analytical Chemistry* **1964,** *36* (4), 706-723.
3. Heinze, J., *Angewandte Chemie International Edition in English* **1984,** *23* (11), 831-847.
4. Elgrishi, N.; Rountree, K. J.; McCarthy, B. D.; Rountree, E. S.; Eisenhart, T. T.; Dempsey, J. L., *Journal of Chemical Education* **2018,** *95* (2), 197-206.
5. Li, Q.; Luo, G.; Feng, J., *Electroanalysis* **2001,** *13* (5), 359-363.
6. Hirst, J.; Armstrong, F. A., *Analytical Chemistry* **1998,** *70* (23), 5062-5071.
7. Wu, S.; Armache, J.-P.; Cheng, Y., *Microscopy (Oxf)* **2016,** *65* (1), 35-41.

Summary

In nature, photosynthesis is a complicated process and a diversity of strategies is used by different organisms. Under low light conditions, light energy can be transferred from the antenna to a reaction center with nearly 100% efficiency. The high yield of natural photosynthesis inspires scientists to replicate the natural process of photosynthesis and build artificial photosynthesis systems to convert CO_2 and H_2O into fuel and chemicals with sunlight. In this thesis, I focus on building an artificial light-harvesting system to explore how a chiral responsive matrix can be formed and if the harvested light energy can be transferred with nearly 100% to an acceptor. **Chapter 1** presents a general introduction to the semi-artificial light-harvest system that is studied in this thesis and that is derived from the CsmA protein from the baseplate of the *Chlorobaculum tepidum*.

In **Chapter 2** the tCsmA, a truncated version of the CsmA, has been successfully applied to form a protein scaffold and bind with the chromophore ZnPP using a dialysis procedure specifically developed for this purpose. The photochemical properties such as the molar extinction coefficient, and the stoichiometry of the tCsmA-ZnPP complex, are determined. Structural information is obtained by cryo-EM, and validated with LD and CD optical spectroscopy, and SEC. The results show that the complex forms an extended periodic 7_2 helical structure composed of well-ordered tetramer units. The ZnPP are oriented perpendicular to the long axis of the peptide helix and subject to strong dipole-dipole couplings in the self-assembly. And almost 100% harvested energy can be transferred to the acceptor cation radical methyl viologen. The excitation-energy transfer rate of the complex was increased from the ns^{-1} scale monomeric ZnPP in solution compared to the ps^{-1} scale for the complex, which is comparable to natural light-harvesting antenna systems.

Chapter 3 presents a detailed study on the mechanism of the self-assembly of the tCsmA-ZnPP complex. The high pH of the dialysis buffer, the net charge of the complex, an α-helical peptide structure, and the binding of ligand are shown to be important for modulating the self-assembly of truncated CsmA and ZnPP. The hydrophobic effect, the salt bridges or the

dipole-dipole interactions of chromophores drive the complex to form an ordered chiral matrix. In addition, a minor fraction of para-crystalline sheets similar to the natural baseplate was observed by cryo-EM. The baseplate structure is provoked by untwisting of the helical structure and possible formation of salt bridges between antiparallel peptide strands with arrays of dimeric ZnPP coordinated to the histidyl between the strands.

The 3D structure of the helical tCsmA-ZnPP assembly was solved by single particle analysis and is discussed in **Chapter 4**. In the helical structure, two strands of tCsmA twist and form a hydrophobic core, to embed the ZnPP coordinated to the imidazole group of the histidine. The helix is built from antiparallel strands of tCsmA with the monomers organized in a head to tail manner within each strand. In this case, the ZnPP interact and form a tetrameric unit. The charged residues are on the outside of the helices and form salt bridges to stabilize helical filaments that form bundles. In the baseplate structure, the ZnPP shows a vertical orientation to the plate.

To complement the peptide synthesis, in **Chapter 5** protein expression in *E.coli* is explored. Without a large fusion tag, tCsmA rapidly degrades. With a TRX tag the stability can be improved, but the sample precipitates making further treatment difficult. With the MBP tag a strongly associated soluble tCsmA-MBP construct is obtained. While SEC can separate the tCsmA and MBP with the help of 6 M GuHCl, and the tCsmA denatures, a small fraction of tCsmA was recovered during the dialysis. This paves the way for further improvement towards obtaining recombinant variants of tCsmA for *e.g.* isotope labeling of samples for NMR studies.

Samenvatting

Natuurlijke fotosynthese is een gecompliceerd proces waarvoor door verschillende organismen een verscheidenheid aan strategieën wordt toegepast. Onder lichtarme condities kan lichtenergie via de antenne naar het reactiecentrum vervoerd worden met bijna 100 % efficiëntie. De hoge opbrengst van natuurlijke fotosynthese inspireert wetenschappers om het natuurlijke proces van fotosynthese te repliceren en om artificiële fotosynthese systemen te bouwen die CO_2 en H_2O in brandstof en chemicaliën om te zetten met gebruik van zonlicht. De focus in deze thesis zal op het bouwen van een artificieel licht verzamelend complex liggen, waarbij onderzocht wordt hoe een chirale responsieve matrix kan worden gevormd en of de geoogste lichtenergie met 100 % efficiëntie kan worden overgedragen aan een acceptor. **Hoofdstuk 1** is een algemene introductie van het semi-artificiële licht verzamelende systeem dat wordt bestudeerd in deze thesis, dit systeem is afgeleid van het CsmA eiwit van de baseplate van *Chlorobaculum tepidum.*

In **Hoofdstuk 2** wordt uiteengezet hoe tCsmA, de verkorte versie van CsmA, succesvol is toegepast om de eiwit structuur te vormen en om te binden met het chromofoor ZnPP, gebruik makende van een speciaal voor dit doel ontwikkelde dialyse procedure. De fotochemische eigenschappen zoals de molaire extinctie coëfficiënt en de stoichiometrie van het tCsmA-ZnPP complex zijn bepaald. Structurele informatie is verkregen met behulp van cryo-EM, en gevalideerd met LD en CD optische spectroscopie en SEC. Het resultaat laat zien dat het complex een verlengde periodieke 7_2 helische structuur, bestaande uit welgeordende tetramerische eenheden, vormt. De ZnPP zijn loodrecht ten opzichte van de lange as van de eiwit helix georiënteerd en zijn aan sterke dipool-dipool koppelingen onderhevig in de zelfassemblage. En bijna 100 % van de geoogste energie kan worden overgedragen naar het positieve acceptor ion radicaal, methyl viologen. De excitatie-energie overdrachtssnelheid in het complex is van de ns^{-1} schaal geobserveerd in monomerisch ZnPP in oplossing naar de ps^{-1} schaal, vergelijkbaar met natuurlijke licht verzamelende antenne systemen, overgegaan.

In **Hoofstuk 3** wordt een gedetailleerde studie van het mechanisme van de zelfassemblage van het tCsmA-ZnPP complex gepresenteerd. Er wordt aangetoond dat de hoge pH van de dialyse buffer, de netto lading van het complex, een α-helische eiwitstructuur, en het binden van ligand belangrijk zijn voor de modulatie van de zelfassemblage van verkort CsmA en ZnPP. De organisatie van het complex in een geordende chirale matrix zou gedreven kunnen worden door het hydrofobe effect, door de zoutbruggen of door de dipool-dipool interacties van de chromoforen. Daarnaast wordt er een kleinere fractie van para-kristallijne platen geobserveerd met cryo-EM, vergelijkbaar met de structuur die in de natuurlijke baseplate gevonden wordt. De baseplate-structuur komt tot stand door het loswinden van de helische structuur en mogelijk ook door de formatie van zoutbruggen tussen antiparallel gelegen eiwit strengen en matrices van dimerisch ZnPP gecoördineerd aan de histidyl tussen de strengen.

De 3D structuur van de helische tCsmA-ZnPP assemblage is opgelost met single particle analysis en wordt bediscussieerd in **Hoofdstuk 4**. In de helische structuur draaien twee tCsmA strengen om elkaar heen en vormen zo een hydrofobe kern waarin ZnPP, gecoördineerd aan de imidazool groep van de histidine, verankert wordt. De helix is opgebouwd uit antiparallel gelegen strengen van tCsmA waarvan de monomeren kop aan staart georganiseerd zijn binnen elke streng. In dit geval is er interactie tussen de ZnPP en vormen deze een tetramerische eenheid. De geladen residuen zitten aan de buitenkant van de helices en vormen zoutbruggen om de helische filamenten in bundels te stabiliseren. In de baseplate structuur is ZnPP op een verticale manier ten opzichte van de plaat georiënteerd.

Om de peptide synthese te complimenteren is in **Hoofdstuk 5** de expressie van tCsmA in *E-coli* onderzocht. Zonder een lange fusie tag valt tCsmA snel uiteen. Met een TRX tag wordt de stabiliteit significant verbeterd, maar slaat het sample ook sneller neer, wat verdere bewerkingen compliceert. Met de MBP tag wordt een sterk geassocieerd, oplosbaar tCsmA-MBP construct verkregen. Met behulp van denaturatie met 6 M guanidine hydrochloride kunnen tCsmA en MBP via SEC van elkaar gescheiden worden, een kleine hoeveelheid tCsmA kan vervolgens verkregen worden uit dialyse. Dit maakt de weg vrij voor verdere verbeteringen gericht op het verkrijgen van recombinante variaties van tCsmA voor onder andere het prepareren van isotoop gelabelde samples voor toekomstige NMR studies.

总结

自然界中光合作用拥有复杂的过程，它以不同形式存在于各种组织中。在低光条件下，光能几乎可以 100%通过叶绿素传递到光合反应中心。高产率的自然光合作用启发了许多科研工作者建立一个可以转化光能，CO_2 和 H_2O 为化学燃料的人工光合作用体系。 在本论文工作主要集中在构建人工捕光系统来探索手性响应矩阵的形成及作用机理，并验证捕获的光能是否可以几乎 100%地传递到电子受体。 第一章总体介绍了本论文的半人工光合捕光系统，这个人工光合捕光系统是衍生于 *Chlorobaculum tepidum* 中底盘结构中的 CsmA 蛋白质。

在第 2 章中，改性于 CsmA 的蛋白质 tCsmA 通过离析手段被成功地制备成支架来连接色素分子 ZnPP。光化学性质如摩尔消光系数和 tCsmA-ZnPP 络合物的化学计量也在本章测出。结构信息由 cryo-EM 获得，并通过 LD 和 CD 光谱以及 SEC 来验证。 结果表明，该配合物由一个有序四聚体单元组装成周期性延伸的 7_2螺旋体结构。ZnPP 受强偶极-偶极耦合作用的影响下垂直于肽螺旋的长轴取向。 并且几乎 100％捕获的光能可以转移到单价阳离子自由基甲基紫精中。与单体 ZnPP 相比，在有受体情况下复合物激发的能量转移速率从纳秒级别提升到皮秒级别，这与自然界中的光捕获系统相当。

第 3 章详细研究了 tCsmA-ZnPP 复合物的自组装机制。透析缓冲液的 pH，蛋白质的净电荷， α-螺旋肽结构和配体结合等对于 tCsmA 和 ZnPP 的自组装起着重要作用。蛋白质的疏水作用，盐桥或偶极-偶极相互作用驱使络合物形成有序的手性矩阵。此外，通过 cryo-EM 观察到少量与自然体系里底盘类似的结构。 这种底盘结构的形成可能是由于原来的旋螺结构扭转成反平行肽链，然后通过盐桥的形成， 这种结构下与组氨酸配位的 ZnPP 以二聚体阵列形式排列。

在第 4 章中讨论了通过单粒子分析解决了 tCsmA-ZnPP 复合物自组螺旋的三维结构。在螺旋结构中，两股 tCsmA 蛋白质支架扭曲并形成疏水核心，将与组氨酸的咪唑基团连接的 ZnPP 嵌入到中心。 螺旋由 tCsmA 的反平行链构成，单体在每条链内以头尾方式组织。 在这种情况下，ZnPP 相互作用并形成四聚体单元。 带电氨基酸残基位于螺旋的外侧并形成盐桥

以稳定形成纤维束。 而在底盘结构中，ZnPP 显示出与蛋白质支架垂直的取向。

在第 5 章中探讨了大肠杆菌中的蛋白质表达来制备多肽。 没有大的融合标签，tCsmA 会迅速降解。使用 TRX 标签可以提高稳定性，但样品会沉淀，难以进行下一步纯化。使用 MBP 标签，可以获得了可溶性的 tCsmA-MBP 复合体。虽然 SEC 可以在 6M 盐酸胍的帮助下分离 tCsmA 和 MBP，但是 tCsmA 已经变性，所以在透析过程中只能得到一小部分 tCsmA。不过这些工作为 tCsmA 的后续工作比如用于 NMR 研究的样品的同位素标记奠定了基础。

Acknowledgements

The research work presented in this thesis was sponsored by Chinese Scholarship Council (CSC) and Leiden University. They deserve the greatest acknowledgement and appreciation.

I want to give my first thank to Huub for your support, encouragement and trust. I am glad to have you as my supervisor. Every time when I found out your “secret” help, I always feel warm. I appreciate the wisdom you shared to me. I will keep them in my mind. Thanks for your kindness.

A lot of thanks to Liesbeth, I will never forget the first day we met. I followed you to finish all paperwork and living problems. I feel lucky to work with you. You are a so opened, kind and caring person. You always makes everyone’s work easier and always comforts everyone in the group. I do hope I can still knock your door and say “Can I make an appointment?” in the future. I will never forget your kindness.

I also feel lucky to have nice seniors. Thanks to Chen Song (宋辰) to teach me how to live in Leiden and take care of me like a brother. I always feel warm when you gave me some Chinese stuff to release my homesick. I also want to thank Yuliya. Thanks for your company and encouragement. I will never forget our happy memories. You are always positive and I hardly heard any complaining about work and life from you. Thanks to show me how to be a strong and positive person.

Many thanks to SSNMR group. Thanks to Anjali for giving me many suggestions on work, I admire your scientific spirit a lot. Thanks to Franco for your patience, you always explain all the questions I asked very clearly. Thanks to Wim to be the secretary for my defense, and generously offered your instruments to me. Thanks to Laura and Yang Shao (邵洋) for everything you did for me and always be “real” friends. I feel honored to work with so many friendly people here. I am also sorry that I had to spend a lot of time on work and hardly had many social talks with many of you. But when I recall all memory, I found I always got smiles and help. Agur, Alia, Fons, Brijth, Karthick, Robin, Fatemeh, Faezeh, Dieuwertje, Jessica, Remco,

Jan-Paul, Rubin, Xinmeng (李新蒙), Dario, Lijin (田利金), Vidya: I will never forget you all.

My work cannot be done without many people's support. I would like to thank Ludo, Chris and Bart from NeCEN, Arjen and Stefan from Delft University of Technology. Thank you for the help and teaching to make me a "half expert" in cryo-EM. Many thanks to Geri and Nora for the time on the protein expression and purification. I also want to acknowledge Sandrine from Vrije Universiteit Amsterdam, your passion for science will always inspire me. Thanks to Anneloes for starting working on my protein, the time following behind your fast heels will be always remembered by me. I also feel lucky to work in Leiden University, because I got a lot of help from technicians here. Jos, Hans, Viorica, Stefan: thank you for the help.

Most of my time is focus on work. But I am still lucky to make some heart friends here. Laugh, talk, support, trust and cares are all from them. Especially thank to my Chinese friend in Leiden, Kun Cao & Yanhong (曹坤&常艳红), Zhiwei & Ningning (翟志伟&许宁宁), Yao Xiao (肖遥), Han Hu(胡翰), and Jingli (辛景莉): we may separate in different places, but I am always with you all.

Many thanks to the Joostens', you show to me how lovely Dutch family can be, that always reminds me, my Chinese family. Thanks for all happy memories. For Peter, thank you for standing for me all the time. You show to me how kind Dutch can be. Of course, special thanks to my family, thank you for loving me so much.