

CONTROLLING THE LOADING OF PROTEIN NANOCAGES



W.F. Rurup

CONTROLLING THE LOADING OF PROTEIN NANOCAGES

Willem Frederik Rurup

Promotiecommissie leden:

Prof. dr. Gerard van der Steenhoven	Universiteit Twente (Voorzitter)
Prof. dr. Jeroen J. L. M. Cornelissen	Universiteit Twente (Promotor)
Dr. Melissa S. T. Koay	Universiteit Twente (Assistent promotor)
Prof. dr. Albert J. R. Heck	Universiteit Utrecht
Dr. Maarten Merx	Technische Universiteit Eindhoven
Prof. dr. Johan F. J. Engbersen	Universiteit Twente
Prof. dr. ir. Pascal Jonkheijm	Universiteit Twente
Dr. Christian Blum	Universiteit Twente

Het onderzoek beschreven in dit proefschrift is uitgevoerd met de financiële ondersteuning van de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) en de European Science Foundation (ESF) in het kader van het EURYI Award programma.

Omslagontwerp: Thera Hijkoop
Drukker: Ipskamp Drukkers
ISBN: 978-90-365-0124-8
DOI: 10.3990/1.9789036501248

Copyright © 2013, W. F. Rurup, Enschede, Nederland. All rights reserved.

CONTROLLING THE LOADING OF PROTEIN NANOCAGES

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 20 september 2013 om 16.45 uur

door

Willem Frederik Rurup

geboren op 4 juni 1982
te Voorburg

Dit proefschrift is goedgekeurd door:

Promotor:

Prof. dr. Jeroen J. L. M. Cornelissen

Assistent promotor:

Dr. Melissa S. T. Koay

Table of Contents

CHAPTER 1	1
1.1 GENERAL INTRODUCTION	2
1.2 REFERENCES	4
CHAPTER 2: NANO-COMPARTMENTS AND THEIR LOADING	7
2.1 COMPARTMENTALIZATION	8
2.2 COMPARTMENT TYPES	9
2.2.1 LIPOSOMES AND POLYMERSOMES	9
2.2.2 PROTEIN CAGES	10
2.2.2.1 BACTERIAL MICROCOMPARTMENTS	10
2.2.2.2 VIRAL CAPSIDS	12
2.2.2.3 <i>DE NOVO</i> CAPSIDS	15
2.3 LOADING OF VIRAL CAPSIDS	15
2.3.1 <i>IN VITRO</i> LOADING	16
2.3.1.1 CHARGE INDUCED	16
2.3.1.2 COILED COIL PEPTIDES	16
2.3.2 <i>IN VIVO</i> LOADING	16
2.3.2.1 DNA APTAMERS	16
2.3.2.2 SIGNALLING PROTEINS	17
2.3.3 ESSENTIAL DIFFERENCES IN LOADING STRATEGIES	17
2.4 ENCAPSULATION EFFECTS	18
2.4.1 ENCAPSULATION OF FLUORESCENT PROTEINS	18
2.4.1.1 CCMV	19
2.4.1.2 BACTERIOPHAGE Q β	19
2.4.1.3 BACTERIOPHAGE P22	19
2.4.1.4 SV40	20
2.4.1.5 FLUORESCENT PROTEIN LOADING EFFECTS	20
2.4.2 ENCAPSULATION OF ENZYMES	20
2.4.2.1 CCMV	20
2.4.2.2 BACTERIOPHAGES MS2 AND Q β	20
2.4.2.3 BACTERIOPHAGE P22	21
2.4.2.4 SV40	21

Table of contents

2.4.2.5 ENZYME LOADING EFFECTS	22
2.5 FUTURE OUTLOOK	23
2.6 REFERENCES	25

CHAPTER 3: LOADING METHODS FOR CCMV VIRUS-LIKE PARTICLES **31**

3.1 INTRODUCTION	32
3.1.1 VIRUSES	32
3.1.2 COWPEA CHLOROTIC MOTTLE VIRUS (CCMV)	33
3.1.3 LOADING OF CCMV	34
3.2 RESULTS AND DISCUSSION	36
3.2.1 IMPROVING THE FLUORESCENT PROTEIN CHARACTERISTICS	36
3.2.2 IMPROVING TFP LOADING	38
3.2.3 ENCAPSULATION PROPERTIES	43
3.2.4 POSSIBLE FÖRSTER RESONANCE ENERGY TRANSFER (FRET) ACCEPTORS	44
3.3 CONCLUSION	49
3.4 ACKNOWLEDGEMENTS	49
3.5 EXPERIMENTAL	50
3.5.1 GENETIC ENGINEERING OF PET15B–MTEFP E–COIL	50
3.5.2 DESIGN AND MOLECULAR CLONING OF MTEFP – CCMV CAPSID FUSION PROTEIN	50
3.5.3 DESIGN AND MOLECULAR CLONING OF DSRED E COIL	51
3.5.4 RECOMBINANT PROTEIN EXPRESSION AND PURIFICATION	51
3.5.5 CAPSID ASSEMBLY	52
3.5.6 POLYACRYLAMIDE GEL ELECTROPHORESIS	52
3.5.7 UV–VISIBLE SPECTROSCOPIC ANALYSIS	52
3.5.8 STEADY–STATE FLUORESCENCE SPECTROSCOPY	54
3.6 REFERENCES	55

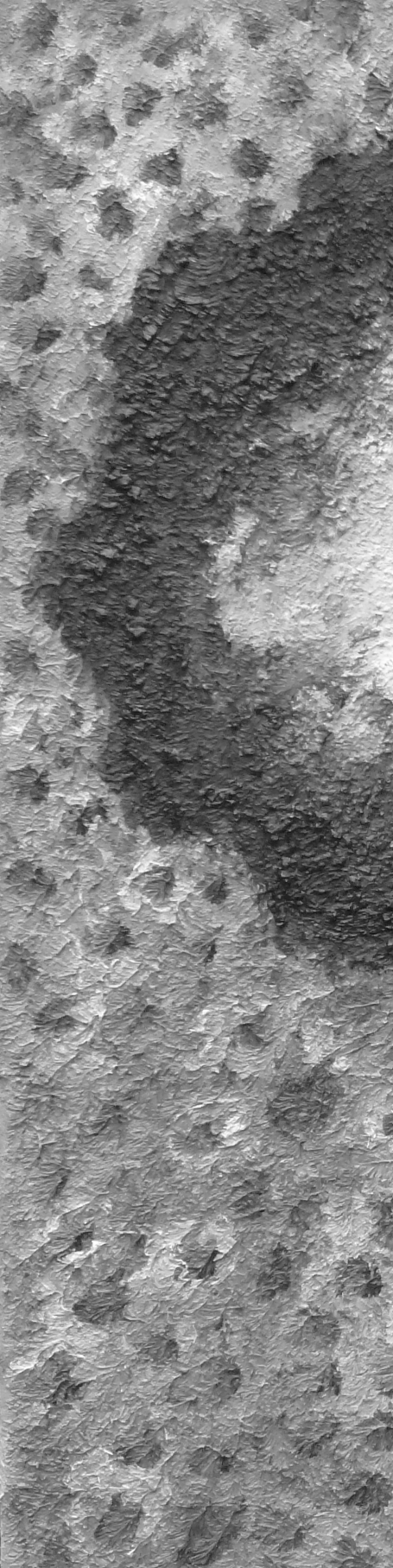
CHAPTER 4: CONTROLLED LOADING OF TFP FUSION PROTEIN IN CCMV VIRUS-LIKE PARTICLES **59**

4.1 INTRODUCTION	60
4.2 RESULTS AND DISCUSSION	60
4.2.1 VLP LOADING THROUGH CAPSID–TFP FUSION PROTEIN	60
4.2.2 DIRECTED LOADING	61

4.2.3	FLUORESCENCE ANISOTROPY	62
4.2.4	MAXIMUM THEORETICAL FLUORESCENT PROTEIN LOADING	67
4.3	CONCLUSION	69
4.4	ACKNOWLEDGEMENTS	70
4.5	EXPERIMENTAL	71
4.5.1	RECOMBINANT PROTEIN EXPRESSION AND PURIFICATION	71
4.5.2	CAPSID ASSEMBLY	71
4.5.3	UV–VIS ANALYSIS	71
4.5.4	FLUORESCENCE ANISOTROPY	71
4.5.5	TRANSMISSION ELECTRON MICROSCOPY (TEM)	71
4.6	REFERENCES	73
CHAPTER 5: ENCAPSULINS, BACTERIAL NANO CAGES		75
<hr/>		
5.1	INTRODUCTION	76
5.2	EXPRESSION AND SEQUENCES	80
5.3	EXPRESSION AND PURIFICATION PROTOCOL	82
5.3.1	ENCAPSULIN EXPRESSION	82
5.3.2	ISOLATION AND SEPARATION OF BACTERIAL ENCAPSULINS	83
5.3.3	ENCAPSULIN PURIFICATION BY SUCROSE DENSITY GRADIENT	84
5.3.4	ENCAPSULIN ANALYSIS	85
5.3.4.1	TRANSMISSION ELECTRON MICROSCOPY (TEM)	87
5.4	CONCLUSIONS	87
5.5	EXPERIENCED USER PROTOCOL	89
5.6	REFERENCES	91
CHAPTER 6: CONTROLLING ENCAPSULINS LOADING		93
<hr/>		
6.1	INTRODUCTION	94
6.1.1	ENCAPSULIN	94
6.1.2	TEAL FLUORESCENT PROTEIN CARGO	94
6.2	RESULTS AND DISCUSSION	95
6.2.1	UV–VISIBLE SPECTROSCOPY	95
6.2.2	SDS–PAGE DENSITOMETRY	97
6.2.3	MASS SPECTROMETRY	98
6.2.4	COMPARISON OF TFP LOADING ANALYSIS	99

Table of contents

6.2.5	DIFFERENT CARGO LOADING IN ENCAPSULIN	101
6.3	CONCLUSION	101
6.4	ACKNOWLEDGEMENTS	102
6.5	EXPERIMENTAL	103
6.5.1	ENCAPSULIN CO-EXPRESSION AND PURIFICATION	103
6.5.2	INSERTION OF CARGO IN BMT PLASMID	103
6.5.3	UV-VISIBLE SPECTROMETRY	103
6.5.4	SDS-PAGE DENSITOMETRY	104
6.5.5	MASS SPECTROMETRY	104
6.5.6	TRANSMISSION ELECTRON MICROSCOPY (TEM)	104
6.6	REFERENCES	105
<u>CHAPTER 7: SUMMARY</u>		<u>107</u>
7.1	SUMMARY	108
<u>CHAPTER 8: SAMENVATTING</u>		<u>113</u>
8.1	SAMENVATTING	114
<u>LIST OF ABBREVIATIONS</u>		<u>119</u>
9.1	ABBREVIATIONS IN THE TEXT	120
9.2	AMINO ACIDS	122
<u>DANKWOORD / ACKNOWLEDGEMENTS</u>		<u>125</u>
<u>LIST OF PUBLICATIONS</u>		<u>131</u>
11.1	PUBLICATIONS	132
11.2	CONFERENCE ABSTRACTS AND PRESENTATIONS	132
<u>CURRICULUM VITAE</u>		<u>135</u>



Chapter 1

General Introduction

1.1 General introduction

Throughout nature the biodiversity is tremendous, yet, all life forms are based on the same principle: 4 nucleotides that encode the total genome of every species. These nucleotides are translated into the limited amount of amino acids that form all the different proteins. These, in turn, are essential in all the different organisms on earth. How these ‘simple’ chains of amino acids are assembled with other biomolecules into something alive, is still one of the biggest mysteries of modern science.

To understand how ‘simple’ chemicals build up life, scientists try to mimic living systems. In that research field, one of the more specific areas focuses on how all the different, yet on a molecular level very similar, organisms are able to function so highly efficient in all their biochemical processes. Current opinions suggest that *compartmentalization* is a key factor in this biochemical efficiency.[1, 2]

The principle of natural compartmentalization was first observed when the first optical microscope was made (mid-16th century): Antoni van Leeuwenhoek saw that each organism was composed of cells. In later years, van Leeuwenhoek’s invention was used to identify individual organelles: microscopic membrane-based compartments inside eukaryotic cells. More recently, the existence of protein-based compartments was discovered in bacterial cells and has since changed the way scientists view such primitive organisms.[2, 3] Composed of small protein building blocks, detailed structural studies show that these protein building blocks closely resemble viral capsid proteins and that the nanocompartments they form closely resemble virus capsid morphologies. Recent evolutionary biology studies point out that protein-based nanocompartments exist throughout all domains of life, and they imply that molecular compartmentalization is much more important than once believed.[4, 5] The discovery of such assemblies with enzymes therein prompts the question of “What influence does nanoscale compartmentalization have on the biochemical efficiency of organisms?”

This thesis aims to address this fundamental question, by exploring the use of protein cages for the compartmentalization of molecular cargo, and studying the effects of compartmentalization on the properties of the cargo itself. These studies bridge the gap between understanding the role of compartmentalization in nature (*in vivo*) and how such compartments can be exploited for *in vitro* applications in nanotechnology. In this work, three different protein nanocages were investigated: the capsid of the Cowpea Chlorotic Mottle Virus (CCMV) and the encapsulins of

the bacteria *Brevibacterium linens* and *Thermotoga maritima*. Each protein cage has its own unique properties and challenges for directed loading.

Chapter 2 provides a brief introduction to the concept of compartmentalization, followed by an overview of the different compartments found in nature. The majority of the discussion focuses specifically on protein cages that are currently used for applications in nanotechnology, providing examples and strategies of protein cage loading. Since both fluorescent proteins and enzymes are of particular relevance to this thesis, these specific examples will be highlighted and the effects that confinement in nanocages has on such cargo will be discussed.

In Chapter 3, the protein cage of CCMV is used for the loading of fluorescent proteins. Different strategies for loading fluorescent proteins in CCMV, in particular, the use of either non-covalent or covalent (fusion) proteins to drive cargo loading, will be presented and critically compared.

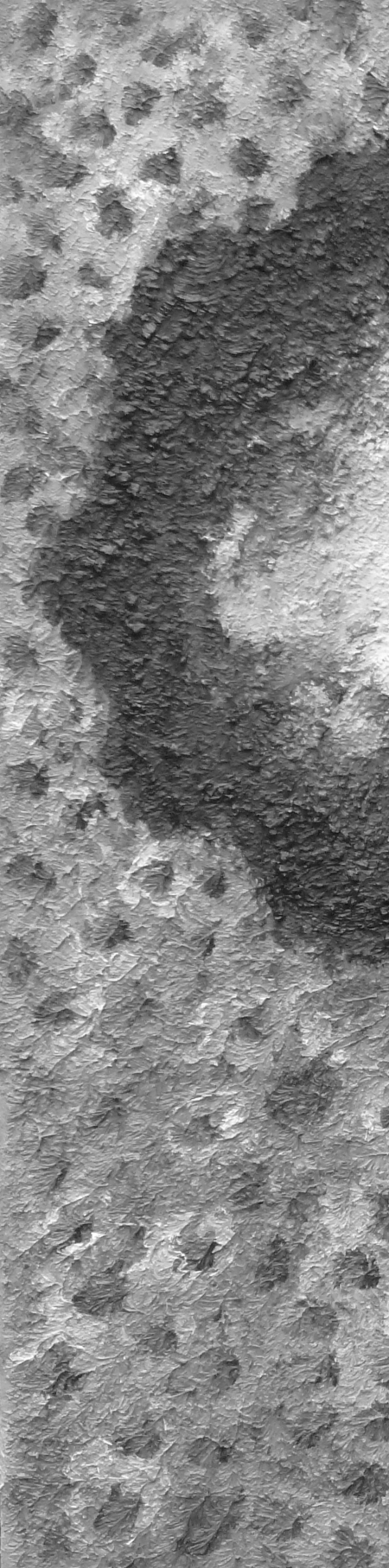
Based on the findings from the different loading strategies described in Chapter 3, in Chapter 4 the controlled (and predictive) loading of CCMV, based on the covalent (fusion) loading strategy is detailed. Via the controlled loading of cargo proteins, a systematic investigation is presented on the effect of confinement and crowding of fluorescent proteins in such CCMV based nanocages.

In Chapter 5, a new class of protein nanocages is introduced: the bacterial encapsulin. A brief introduction into the unique properties of two different encapsulins, from *Brevibacterium linens* and *Thermotoga maritima*, is presented. This chapter describes in detail the protocols for the (future) design and engineering of these protein cages and therefore provides a detailed description of their recombinant expression, purification and characterization.

Finally, Chapter 6 explores the use of encapsulins for applications in nanotechnology. The natural loading pathway of encapsulins is utilized to direct fluorescent protein cargo into the encapsulins as a first attempt to control the loading of these recently discovered protein nanocages.

1.2 References

1. Oehler, S. and B. Muller-Hill, *High local concentration: a fundamental strategy of life*. J Mol Biol, 2010. **395**(2): p. 242-53.
2. Kang, S. and T. Douglas, *Some Enzymes Just Need a Space of Their Own*. Science, 2010. **327**(5961): p. 42-43.
3. Tanaka, S., M.R. Sawaya, and T.O. Yeates, *Structure and mechanisms of a protein-based organelle in Escherichia coli*. Science, 2010. **327**(5961): p. 81-4.
4. Forterre, P. and D. Prangishvili, *The origin of viruses*. Res Microbiol, 2009. **160**(7): p. 466-72.
5. Weiss, R.A. and J.P. Stoye, *Virology. Our viral inheritance*. Science, 2013. **340**(6134): p. 820-1.



Chapter 2

Nanocompartments and their loading

A brief introduction to the concept of compartmentalization is provided and an overview of the different compartments found in nature, with a specific focus on protein cages that are currently used for applications in nanotechnology. State of the art examples of loading strategies for protein cages are discussed and the (confinement) effect on their cargo is evaluated. Since both fluorescent proteins and enzymes are of particular relevance, these two specific types of cargo are the main focus.

Part of this chapter has been accepted in the book titled “Enzyme Nanocarriers”: W. F. Rurup, M. S. T. Koay, J. J. L. M. Cornelissen, “Viruses as enzyme nanocarriers”

2.1 Compartmentalization

Compartmentalization as a principle is nothing new to an engineer and is used for many purposes in engineering, to a greater or lesser degree of success (*e.g.* submarines, passive fire protection in buildings, the HMS Titanic in 1914). However, long before mankind used this concept in design, nature already favored and implemented compartmentalization.[1] In the primordial soup, molecular compartmentalization was adopted by eukaryotes to improve cellular survival, biological function and efficiency. The evolutionary process led to the build-up of compartments within compartments, and ultimately the evolution of higher organisms on Earth. In all cases, these compartments function on their own length scale: from meter (humans) to centimeter (organs) to micrometer (cells) to nanometer (cellular organelles). In general, the enhanced cellular efficiency has proven to be highly beneficial, such that compartmentalization is one of the key requirements for the survival of higher organisms.[2, 3] However, there is still very little that is understood about how molecular reactions and processes in compartmentalization occur.

In nature, several benefits have been attributed to the compartmentalization of enzyme-catalyzed reactions within cells, and particularly, in organelles.[2-5] First, compartmentalization is proposed to shield enzymes from the exterior environment or other competing pathways. Second, the local confinement also creates a higher effective substrate and enzyme concentration, which is thought to improve the overall reaction kinetics. Since the volume inside compartments is much smaller, even one substrate molecule makes for a much higher effective concentration compared to bulk solution. Also, the compartment creates a barrier which can allow for increased control over the molecular flux of substrates and products entering and leaving the confined space, either by diffusion-based or via active transport. This control over molecular flux can be regulated by a selective pore size, by substrate hydrophobicity or hydrophilicity, charge-charge “gated” interactions or controlled by membrane pumps and proteins. Finally, the boundaries of the organelle help to retain intermediates on the inside, preventing their release into the cytosol. These intermediates (which, in some cases, could be either toxic or volatile) can then be rapidly converted to the desired product, therefore enhancing the overall substrate turnover and reaction efficiency.

To scientists the complexity of cells and organelles in nature, their assembly, their precise control over biochemical reactions, and their efficiency are of great inspiration for applications in nanotechnology.[6] To date, much research has been devoted to the design and assembly of artificial cells and organelles, which are

capable of replication and repair. However, despite impressive advancements in recent years [7], many designs do not shed light on the internal workings of the cell or the importance of compartmentalization in biology. Therefore, it still remains to be discovered why certain biological processes are sequestered in compartments, how the molecular flux of substrates and products in and out of organelles are controlled, and what the advantages are, on a chemical level, of performing reactions in such nanoscale compartments.

Many approaches are used to investigate the biochemical relevance of compartmentalization. These include various fundamentally different types of nanocompartments. This chapter provides a broad overview of the different types of synthetic and biological compartments, detailing their properties, benefits and limitations for applications in nanotechnology.

2.2 Compartment types

Several types of compartments have been used for the encapsulation (and release) of cargo for applications in drug release, for molecular storage, and as nanoreactors. A variety of building blocks are currently available, from lipids, (block co) polymers to proteins. An overview of lipid-based vesicles, such as liposomes and polymersomes that have been studied as compartments, and nanoscale protein assemblies, such as virus capsids and other self-assembling protein cages, will be discussed. [8-10]

2.2.1 *Liposomes and polymersomes*

Composed of lipid-based building blocks or amphiphilic block copolymers, which have a hydrophilic head and a hydrophobic tail, liposomes and polymersomes are nanometer- to micrometer-sized compartments that form double-layer (hollow) spheres. Since their composition also mimics the lipid membrane of organelles, with the hydrophilic groups presented on the outside and inside and the hydrophobic groups within the core layer, it has been suggested that polymersomes and liposomes could be used as artificial organelles.[11] To further mimic the properties of cellular organelles, liposomes and polymersomes can be functionalized with selective membrane proteins for controlled substrate uptake.[9] However, these types of compartments are also used in other fields. Their composition allows them to act as cargo carriers, of both hydrophilic and hydrophobic cargo, on the inside and outside. This in combination with their highly concentration dependent assembly and disassembly makes liposomes and polymersomes ideal for drug delivery and medical purposes, for which they are

widely explored.[12, 13] By selecting the building blocks, these compartments can be assembled to degrade over time, slowly releasing their content and they can even be targeted to specific parts of the body.

In a different application, polymersomes, which have the advantage over liposomes of having an increased stability, have been designed as local pH sensors[14] and as enzyme nanoreactors.[15, 16] These nanoreactors have been used for both one-step [15] and cascade enzyme reactions. In this cascade reaction, the first enzyme is localized in the membrane bilayer, the second in the interior of the polymersome and the third tethered to the exterior [16]. However despite the similarities of polymersomes and liposomes to that of membranes and cellular organelles, the polydispersity of such assemblies is large and therefore prevents accurate quantification of the confinement effects, such as the increased local concentration and changes in kinetics.

2.2.2 Protein cages

One of nature's own compartments that has found little use so far as model are protein cages.[17] Although protein cages are not commonly considered as subcellular compartments, it surprisingly became more apparent that different organisms use protein cages as primitive organelles,[18, 19] nanoreactors,[20] or storage compartments [8]. Furthermore, their combined roles in nature suggest that these compartments are permeable to substrate (and product) diffusion for their cargo. For applications in nanotechnology, protein cages also offer the advantage that they can be genetically modified, engineered and functionalized on the inside and outside. The protein cages can be classified in different categories: bacterial microcompartments, viral capsids and *de novo* protein cages. Each will be addressed below.

2.2.2.1 Bacterial microcompartments

Bacterial microcompartments are recently discovered protein cages that are involved in metabolic pathways. Of the three types of bacterial microcompartments currently identified, the first one that was discovered is the carboxysome (1970s), which is proposed to improve the Calvin cycle (a part of the photosynthesis where CO₂ is coupled to larger hydrocarbons). The carboxysome shell is composed entirely of proteins and glycoproteins, which are assembled into a cage-like structure of 100–150 nm in diameter. So far, the carboxysomes are divided in two groups, based on the components of the shell. However, both groups consist of at least eight different genes, necessary for correct polyhedral assembly.[18, 19] Although much research has focused on the role and the evolution of

carboxysomes, there is currently no known mechanism for their *in vitro* assembly, making them quite complex as compartments for nanotechnology.

The second type of bacterial microcompartment that is characterized, is the 1,2-propanediol utilization (Pdu) metabolosome. The Pdu metabolosome is assembled *in vivo* in the presence of 1,2-propanediol as a nutrient source. Although the exact role of the Pdu metabolosome remains unclear, it is thought that these compartments serve to protect the cytosol by confining the reaction inside. During the catabolism of 1,2-propanediol, the toxic intermediate, propionaldehyde, is produced. By confining the reaction inside, the catalytic conversion of the intermediate is thought to be more efficient. The locus of *pdu* consists of 23 different genes, of which seven are closely related to the carboxysome, including some of the shell proteins. Therefore it is not surprising that carboxysomes and Pdu compartments are of comparable morphology and size.[18, 21]

The third compartment is the ethanolamine utilization (Eut) metabolosome, of which the locus has 17 different genes. The Eut metabolosome is very similar to the Pdu metabolosome, and the shell proteins have a high sequence homology [4, 18], but the Eut metabolosome is involved in the catabolism of ethanolamine. Similar to 1,2-propanediol, the ethanolamine pathway produces a toxic intermediate, *i.e.* acetaldehyde, and the Eut compartment is thought to act as a barrier against the leakage of this toxic and volatile intermediate.

More gene clusters have been found with homologous shell protein sequences in other bacteria, but these are not yet fully characterized.[18] The bacterial microcompartments are interesting candidates to study the effects of confinement on enzyme activity and substrate turnover and to investigate the confinement of toxic or volatile intermediates. However, the number of genes that need to be expressed and the complex interplay between all the different proteins makes it difficult to create these protein compartments *in vitro*.[22] Similar to carboxysomes, there is currently no known *in vitro* assembly/disassembly pathway for metabolosomes.

The metabolosome shell has sub-nanometer pores and although many theories have been suggested, the mechanism for transport of substrates and products over the shell of the metabolosome is still unclear. In the case of carboxysomes, a transport mechanism has been postulated so far, which involves a combination of electrostatically charged pores and a primitive CO₂ translocator enzyme, in order to regulate substrate entry.

The difficulties of *in vitro* assembly, combined with the unknown transport mechanism over the compartment shell makes the study of the effects of

compartmentalization on biochemical processes using these bacterial microcompartments very complex.

2.2.2.2 Viral capsids

Viruses are intrinsically attractive compartments due to their ability to self-assemble from a single kind of protein into highly stable symmetrical assemblies.[23] Composed of multiple copies of protein building blocks, virus capsids can be readily modified (either chemically or genetically) and functionalized, making them highly versatile and tuneable compartments.

Although viruses can be categorized into different morphological classes, the most commonly used protein assemblies in nanotechnology are those of icosahedral and rod-shaped viruses. In particular, the plant-based icosahedral Cowpea Mosaic Virus (CPMV) and Cowpea Chlorotic Mottle Virus (CCMV) and the rod-shaped Tobacco Mosaic Virus (TMV) are amongst the most widely used and most well characterized viruses in nanotechnology.[24-37]

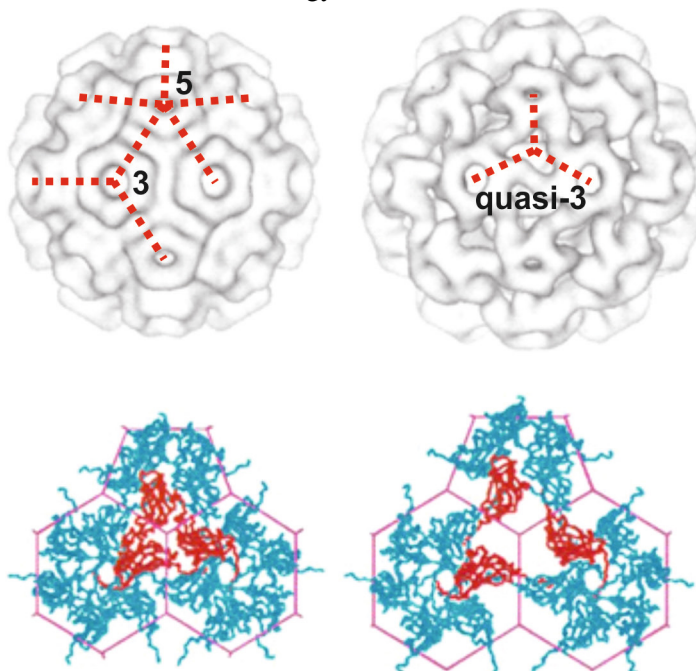


Figure 2.1: The CCMV capsid at pH 5 (left) and its swollen form at pH 7 (right), as edited from Liu *et al.*[38]

For the use of viruses as nanocompartments, the porosity of the capsid shell plays an important role in the diffusion of cargo and substrates. In particular, one of the most attractive features would be a stable protein assembly that contains pores that

can be reversibly controlled or triggered (for example, by altering the pH and ionic strength). By tuning and controlling the pH, the gradual increase in shell porosity could be extremely useful for the selective uptake and release of substrates and products. This is the case for cowpea chlorotic mottle virus (CCMV), which is composed of 180 identical copies of capsid protein that self-assemble at pH 5.0. CCMV forms a T = 3 capsid with an external diameter of 28 nm. CCMV capsids at pH 5.0 contain small < 1 nm pores on the threefold axis and < 0.5 nm on the fivefold axis (Figure 2.1).

At near physiological pH (pH ~ 7), a 10% swelling of the shell occurs, while the protein assembly itself remains intact. During this process, the porosity increases due to ~2 nm pores opening on the *quasi*-threefold axis, which allows diffusion of larger molecules.[35] At pH > 7.5, the CCMV capsid disassembles completely yielding free capsid proteins dimers. While the pH driven swelling of CCMV is an attractive approach to reversibly trigger changes in the porosity, gaining precise control over the pore size is extremely difficult.

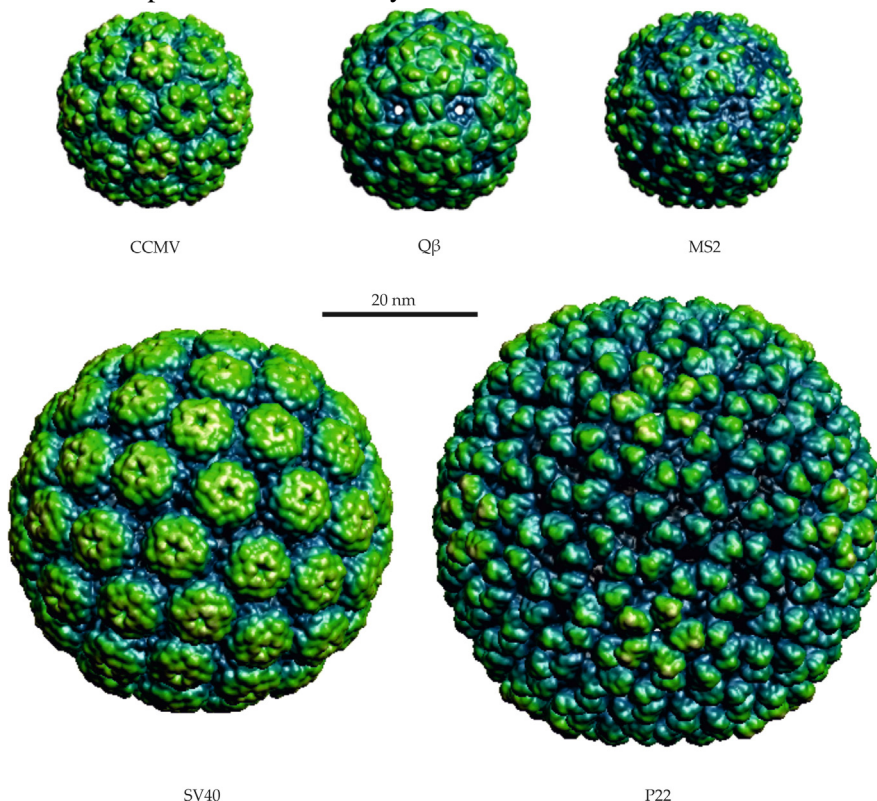


Figure 2.2: The size and morphology of the capsids of CCMV, Q β , MS2, SV40 and P22. The images are edited after rendering by ViperDB.

Recently, MS2, Q β and P22 bacteriophages are more often used as an alternative scaffold for virus-based assemblies. Offering intrinsically unique properties, these bacteriophages are more stable than plant-based viruses, yet they can also form empty capsid shells, making them an attractive series of nanocompartments (Figure 2.2). The native MS2 virion is composed of three components: the virus capsid protein (13.7 kDa), the A protein (44.0 kDa) and a single-stranded RNA molecule. Similar to the plant-based viruses, MS2 can be reassembled from acid-driven denatured monomers into a stable virus-like particle (VLP). Moreover, the MS2 shell possesses 32 pores per capsid, each approximately 1.6 nm in diameter, which allow the transportation of potential substrates and cargos.[36] Interestingly Tullman-Ercek and co-workers showed that by using osmolytes such as trimethyl amine *N*-oxide (TMAO), the encapsulation efficiency of enzymes was enhanced.[39] Although osmolytes are used to regulate osmotic pressure inside cells, they have also been shown to be highly effective in stabilizing proteins.[40, 41] The use of osmolytes is an interesting approach to alter the capsid porosity, which could be used to confer substrate selectivity or regulate the diffusion of inhibitory molecules.

The Q β bacteriophage belongs to the same family as the MS2 virus. Nevertheless it is quite different, *i.e.* no assembly/disassembly pathway is known and the pores of the capsid are significantly smaller.[42]

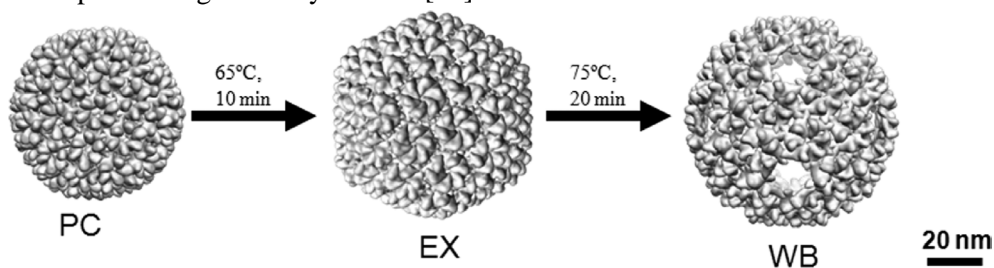


Figure 2.3: The different P22 morphologies (Adopted from Patterson *et al.*[43]).

P22 is a rather unique bacteriophage that is related to bacteriophage λ . Composed of 420 capsid proteins, P22 recruits 100–330 scaffold proteins in order to assemble into the icosahedral T = 7 procapsid (PC) of 58 nm diameter. Upon heating to 65°C for 10 minutes, the P22 procapsid transforms to the mature capsid, whereby the capsid swells to 64 nm and the scaffold proteins are released from the capsid. Interestingly, although the interior volume of the capsid is doubled in this expanded form (EX), the pore size of 2 nm remains unaffected. Upon raising the temperature further (75°C for 20 minutes), 12 pentamers are released, irreversibly creating a wiffle ball (WB) capsid with large 10 nm sized holes (Figure 2.3). In this

form, free diffusion of substrates and products in and out of the capsid is reported.[43]

The simian virus 40 (SV40) of the polyomavirus family is the only virus used for nanocompartments which has eukaryotic hosts. Therefore the particles are expressed through a different expression system, rendering non-infective SV40 capsids. The viral capsid is built up of 72 pentamers of the major capsid protein, which form the capsid, assisted by minor capsid proteins 2 and 3. The recombinant system expresses only the major capsid protein, which forms native capsids shaped VLPs. These VLPs are ~45 nm in diameter and assembles to a T=7 symmetry.[44]

2.2.2.3 *De novo capsids*

Hilvert and co-workers used a different approach to nanocompartmentalization. They used the icosahedral structure of the enzyme lumazine synthase from *Aquifex aeolicus* (AaLS) in a directed evolution experiment to improve its encapsulation abilities. By expressing a cell toxic enzyme with a 10-fold C-terminal positive tag together with random mutated AaLS, the bacteria with the improved encapsulation in cages survived better. These were randomly mutated again, constantly increasing the production of the toxic enzyme, until a AaLS species was reached that encapsulated the enzyme very strongly, due to many mutations on the AaLS interior.[45] This approach yielded a protein container that encapsulates anything with positive charges.[46]

2.3 Loading of viral capsids

Since the work in this thesis will focus on controlling the loading of small nanometer sized protein cages, this section will focus on loading strategies for those cages. Still, there is a multitude of examples of natural nanocompartments and their loading. Therefore the different methods for loading these < 100 nm cages will be considered below.

While virus-like particles form well-defined assemblies, the *controlled* loading of these capsids is a significant challenge that cannot be easily overcome. In the past, statistical encapsulation was used to load horseradish peroxidase into *in vitro* reassembling CCMV.[35] While convenient, statistical encapsulation is a highly concentration-dependent process and hence is largely uncontrollable and it suffers from wide loading distributions. Since the internal walls of many viruses are naturally positively charged to load the negatively charged genetic materials, negatively charged linkers have been explored to promote charge-charge

interactions for the controlled *in vitro* and *in vivo* loading of cargo inside virus-like particles.

2.3.1 *In vitro* loading

2.3.1.1 Charge induced

Inspired by the natural loading of DNA/RNA, the use of nucleic acid based linkers and their mimics have become increasingly popular to gain control over VLP loading. For example, MS2 bacteriophages rely on a targeting RNA-hairpin for the internalization of their genomic RNA. Tullman-Ercek and co-workers recently mimicked this approach to target a non-native enzyme inside MS2 capsids.[39] In this case, a poly(16) anionic peptide linker was used as an RNA mimic, which was fused to the C-terminus of *E. coli* alkaline phosphatase (PhoA) to facilitate loading inside MS2 (Figure 2.4b).

2.3.1.2 Coiled coil peptides

In CCMV, heterodimerizing ‘coiled-coil’ peptides were designed by fusing a positively-charged α -helix to the N-terminus of the CCMV capsid protein (Figure 2.4a). Fusion of a complementary negatively charged α -helix to the C-terminus of EGFP and lipase B from *Pseudozyma antarctica* (PalB) promoted the formation of a non-covalent cargo-capsid protein complex. By varying the mixing ratios of cargo-capsid protein complex to wild type capsid protein the amount of loaded enzyme could be influenced.[47] This approach will be explained in more detail in Chapter 3.

2.3.2 *In vivo* loading

2.3.2.1 DNA aptamers

In some cases, the use of specific RNA/DNA recognition loops allow further control and selectivity over cargo loading. An elegant example is that of the recombinant Q β bacteriophage, which could be loaded and assembled recombinantly under *in vivo* conditions. This design is comprised of two plasmids of which the first encodes expression of the virus capsid, the second the target enzyme. The mRNA used to transcribe the capsid protein was subsequently used as an anchor. The native RNA docking sequence was engineered on the 3'-end and on the 5'-end, an α -Rev aptamer was used. By fusing the Rev-tag to the N-terminal side of the target enzyme, the two mRNA ends form “baits” for Q β (3'-end) and the target enzyme (5'-end). This DNA-aptamer system has already been successfully used to encapsulate peptidase E, luciferase and a thermo stable

luciferase variant[48] and ‘superfolder’ GFP, BFP and CFP and mCherry[49], making it a versatile and effective approach for the *in vivo* loading of enzymes inside virus-based assemblies (Figure 2.4c). Furthermore, by tuning the recombinant expression conditions, the amount of encapsulated enzyme can be influenced, which allows further control over enzyme loading inside bacteriophage Q β .

2.3.2.2 Signalling proteins

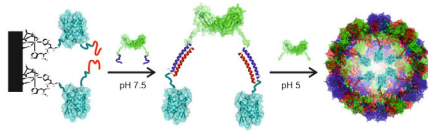
Of the viruses discussed thus far, P22 and SV40 are the only examples that rely on a scaffold protein during assembly (Figure 2.4d). For P22, a truncated version of the scaffold protein (amino acids 141–303; the yellow peptide in Figure 2.4d) fused to a target protein was shown to be sufficient to encapsulate fluorescent proteins and enzyme cargo during *in vivo* assembly.[50, 51] After the capsid is changed into the extended form (EX), the scaffold proteins are released and because they are tethered to the cargo they protrude out of the virus assembly. The scaffold proteins can subsequently be removed by enzymatic digestion of the linker between the scaffold and the cargo. Similarly, the capsid of the SV40 major capsid protein has a strong interaction with amino acids 273–308 of minor capsid protein 2. Co-infection of baculoviruses expressing both the SV40 capsid protein and the cargo protein from different viruses results in stable, EGFP or enzyme loaded capsids *in vivo*. By using their respective signalling proteins, alcohol dehydrogenase D (AdhD) from *Pyrococcus furiosus* and the homo-tetrameric β -glycosidase ‘CelB’ were both successfully encapsulated in P22 and yeast cytosine deaminase (yCD) was encapsulated into SV40.[43, 44] However, fully *controlled* enzyme loading has not been reported for either of these systems so far.

2.3.3 Essential differences in loading strategies

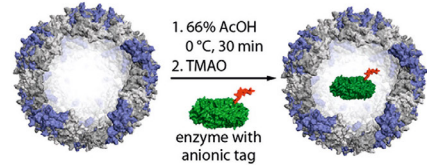
Most encapsulation strategies are based on native virus loading mechanisms. To date, the different loading strategies can be classified in two broad groups: *in vitro* and *in vivo* loading. For a virus with a known assembly/disassembly pathway, *in vitro* loading mechanisms offer a versatile approach for the controlled encapsulation of enzymes. The main benefit of *in vitro* loading is the possibility to purify the separate components before encapsulation. A secondary benefit is that the cargo loading can be more tightly controlled (and in some cases loading numbers can be predicted prior to the assembly). In contrast, viruses and virus-like assemblies that can be employed for *in vivo* loading are advantageous since many of these assemblies are intrinsically more stable. However, one major disadvantage

is that there is little control over the number of cargo proteins encapsulated, which is likely to lead to a disperse loading of virus-like particles.

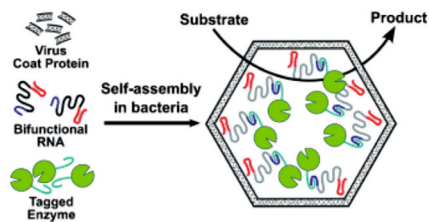
Coiled-coil loading



Charge induced loading



DNA aptamer



Signaling protein

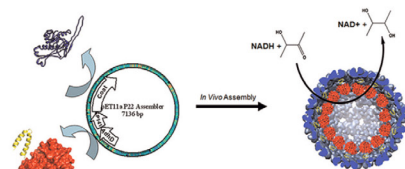


Figure 2.4: Different loading schemes used for loading of viral capsids as described in section 2.3.

2.4 Encapsulation effects

In nature, the cell membrane provides a physical boundary to confine enzymes and substrates within a local environment. This confinement imposes a large molecular crowding effect, whereby the reduced volume 1) increases the effective concentration of the cargo and 2) is proposed to significantly improve the kinetic efficiency of enzymes.[37, 52, 53] Similarly, the virus-like particles provide a confined environment, however, the effects of encapsulating fluorescent proteins and enzymes inside these VLPs thus far, are very diverse.

2.4.1 Encapsulation of fluorescent proteins

Early observations of fluorophores encapsulated in viral capsids were made with the T4 bacteriophage. The T4 head is about 50–110 nm in diameter and has an elongated icosahedral symmetry $T = 13$, $Q = 21$. [54] Black and co-workers calculated an interior of 4×10^{-19} L, or 4×10^5 nm³. In this volume they encapsulated ~ 100–200 GFP molecules through a signalling protein approach. They observed no change in fluorescence for these green fluorescent proteins if they were cleaved from the scaffold proteins after encapsulation. However, when

the fluorophores remained attached to the signalling protein, the fluorescence anisotropy drops significantly (from ~0.3 for GFP, to 0.226 for linked GFP). Since the signalling protein used normally binds DNA, this restricts the movement of the GFP and brings them close together, inducing a phenomenon known as homo-Förster resonance energy transfer (homo-FRET) [55], which will be further explained in Chapter 4.

2.4.1.1 CCMV

In CCMV the coiled-coil system was used to encapsulate EGFP. Here it showed that although loading was not as straightforward, up to 15 EGFP molecules could be encapsulated in the CCMV VLP. Transmission electron microscopy (TEM) analysis showed that this amount of cargo proteins does not interfere with the normal T = 3 capsid formation.[47, 56] The effect of this encapsulation on the fluorescent proteins was not discussed.

2.4.1.2 Bacteriophage Q β

The DNA-aptamer system described above was used to attach superfolder GFP (sfGFP), a blue and a cyan mutant of sfGFP and the red mCherry fluorescent protein in the Q β . It was reported that 10–11 proteins were encapsulated per Q β . The assembled filled particles were indistinguishable from native Q β capsids. The fluorescent properties of the cargo were virtually unchanged with respect to the free proteins, but it was noticed that the fluorescent proteins were ‘somewhat more resistant’ to thermal denaturation and fully protected to protease digestion.[49]

2.4.1.3 Bacteriophage P22

The signalling protein approach was used to load either EGFP or mCherry fluorescent proteins into the P22 capsid. It was found that when P22 is expanded to the EX form, the scaffold proteins protrude from the capsid and can be cleaved of, rendering the fluorescent proteins free in the capsid. When the capsid is expanded to the WB form, the pores become too large to contain the fluorescent proteins inside. O’Neil *et al.* reported ~ 250 fluorescent proteins on the inside of the P22, although they observed a discrepancy between multi-angle light scattering (MALS) (281 EGFP and 233 mCherry) and UV-visible (UV-vis) spectrometry (345 EGFP and 150 mCherry).[51] They continued this research by creating a fusion protein of GFP and mCherry, of which MALS measurements placed 111 copies inside the P22 capsid. They reported that FRET occurs for the encapsulated proteins as well as for the free fusion proteins, however due to a higher local concentration the encapsulated fusion proteins show 3–5 fold enhanced FRET.[50]

2.4.1.4 SV40

The approach of using the signalling protein was also used by Inoue *et al.* [44] to encapsulate EGFP in SV40 particles. However, they used it mainly to verify that encapsulation of foreign proteins by fusion to the signalling protein was possible. Therefore hardly any characterization was reported on the fluorescent protein filled SV40 particles.

2.4.1.5 Fluorescent protein loading effects

So far the most significant effect reported for encapsulation of fluorescent proteins in virus-like particles is the protection which is given by the capsid. Be it protection to antibodies, proteases or thermal denaturation, the capsid acts as a shield for the fluorescent proteins. However, the more fundamental nanotechnology effects remain neglected. Black and co-workers did see that crowding leads to an enhanced energy transfer between the fluorescent proteins, as did Douglas and co-workers. Nevertheless, this effect has not been thoroughly quantified.

2.4.2 Encapsulation of enzymes

2.4.2.1 CCMV

The encapsulation of *Pseudozyma antarctica* lipase B (PalB) inside CCMV was shown to increase the initial reaction velocity of the conversion of 6,8-difluoro-4-methylumbelliferyl octanoate (DiFMU-octanoate) to the fluorescent DiFMU with respect to the free enzyme in bulk solution. This suggests that the efficiency or turnover rate of substrate is enhanced upon encapsulation. However interestingly, the same study also showed that encapsulation of more than one enzyme within the same capsid decreased the relative reaction rate, suggesting that, for PalB in CCMV, molecular crowding plays a significant and almost immediate role on the encapsulated enzyme. In another example, horseradish peroxidase (HRP) was encapsulated inside CCMV. However, while no kinetic analyses were performed, it was shown that increasing the pH to induce capsid swelling also led to increased product diffusion.[57]

2.4.2.2 Bacteriophages MS2 and Q β

Although the MS2 capsid is similar in size to CCMV (27 and 28 nm respectively), it is unique as it contains 32 pores (1.6 nm in diameter) which allow diffusion of large substrates into the capsid without disassembling the virus.[58] The encapsulation of (on average) 1–2 alkaline phosphatase (PhoA) inside the MS2 bacteriophage resulted in a marginal increase in K_m and a decrease in k_{cat} . [39] Due

to the size of PhoA (94.9 kDa), Glasgow *et al.* attributed this decrease in enzyme efficiency to spatial crowding inside the MS2 capsid, making this a striking example of the delicate interplay between enhancement and inhibition of enzymes in confined environments.

In another example, Fiedler *et al.* reported the encapsulation of an aspartate dipeptidase, peptidase E, (PepE) in bacteriophage Q β . [48] In this example, the catalytic efficiency (k_{cat}/K_m) for the encapsulated PepE (2–18 enzymes per virus-like particle) increased 3-fold compared to free PepE. Importantly, this work showed that encapsulated PepE exhibits significantly improved thermostabilization. After heating to 50°C for 30 minutes, no loss in activity was observed for the encapsulated PepE, whereas under the same conditions, free PepE retained only 20% activity. Similarly, encapsulation was also shown to effectively protect PepE from enzymatic digestion by proteinase K (retaining 80% activity). In the same work, luciferase (Luc) and a thermo stable luciferase (tsLuc) encapsulated in Q β displayed an increase in K_m (*i.e.* increased apparent substrate binding affinity), suggesting that indeed the effective local substrate concentration increases inside the capsid. Interestingly however, no enhancement of k_{cat} was observed. The authors suggested that perhaps the enzyme's flexibility (or a restriction thereof) might cause this (negative) influence on kinetic performance.

2.4.2.3 Bacteriophage P22

The dense packing of approximately 250 alcohol dehydrogenase D (AdhD) enzymes per P22 bacteriophage led to an effective confinement molarity (M_{conf}) of 7.16 mM [43]. Intriguingly, the apparent k_{cat} was reported to be lower compared to the free enzyme and the overall enzyme efficiency (k_{cat}/K_m) also decreased. Similar observations were reported for the encapsulation of the homotetrameric β -glycosidase (CelB). [59] Encapsulation of 84–87 CelB monomers (21 homotetramers) per P22 showed no significant change in efficiency (k_{cat}/K_m).

Remarkably, Douglas and co-workers calculated that 250 AdhD led to ~27% occupancy of the total inner volume of the P22 capsid. This is a striking example since it closely resembles the conditions of the cell, whereby all enzymes and macromolecules combined are also estimated to occupy up to 30% of the physically available space. [37, 52, 60, 61] This is an interesting resemblance which might highlight the suitability of protein cages as cell or organelle mimics.

2.4.2.4 SV40

In the SV40 capsid *yeast cytosine deaminase* (yCD) was encapsulated. No kinetic studies of these constructs are reported, other than the affirmation of enzyme

function and diffusion of substrate and product over the capsid. Nevertheless, protection of yCD from antibodies and proteases is proven.[44]

Table 2.1. Overview of the confined kinetic effects observed in some model nanoreactors

Capsid	Pore size	Enzyme (E)	Substrate (S)	#E	Effect	Ref.
CCMV pH 5.0	<1 nm	PalB	DiFMU	1.3	~5-fold increase to non-encapsulated enzyme	[57]
		PalB	DiFMU	4	No significant change to non-encapsulated enzyme	[57]
		HRP	ABTS & dihydro-rhodamine 6G	1	Diffusion rates for encapsulated rhodamine 6G are 3ordersof magnitude lower than for free rhodamine 6G	[35]
CCMV pH 7.5	~2 nm	HRP	ABTS & dihydro-rhodamine 6G	1	Diffusion rates for encapsulated rhodamine 6G are 2-3-fold higher than for rhodamine 6G in CCMV at pH 5	[35]
MS2	1.6–1.8 nm	PhoA	p-nitrophenyl phosphate	~ 1.6	Marginal K_m decrease and k_{cat} decrease	[39]
Q β	0.7–1.4 nm	PepE	Asp-AMC	9	k_{cat}/K_m shows no significant difference due to encapsulation.	[48]
		Luc	Luciferin & ATP	4	Lower k_{cat} (factor 2), higher K_m (factor 17) for Luc than free Luc	[48]
		tsLuc	Luciferin & ATP	2	Equal k_{cat} , higher K_m (factor 10) for tsLuc than free tsLuc	[48]
		tsLuc	Luciferin & ATP	9	Lower k_{cat} (factor 2), higher K_m (factor 21) than free tsLuc	[48]
P22 PC (58 nm)	2 nm	AdhD	Acetoin	~ 250	k_{cat} is reduced 7-fold, K_m for acetoin is reduced 6 fold to free AdhD	[43]
P22 EX (64 nm)	2 nm	AdhD	Acetoin	~ 250	k_{cat} is equal to P22 PC, K_m for acetoin is doubled vs. P22 PC	[43]
P22 WB (64 nm)	10 nm	AdhD	Acetoin	~ 250	k_{cat} is equal to P22 EX, K_m for acetoin is equal to P22 EX	[43]
P22 PC (58 nm)	2 nm	CelB (tetramer)	PNPG	~ 85	No significant change to k_{cat} or K_m with respect to free CelB	[59]
P22 EX (64 nm)	2 nm	CelB (tetramer)	PNPG	~ 85	No significant change to k_{cat} or K_m with respect to free CelB	[59]
P22 WB (64 nm)	10 nm	CelB (tetramer)	PNPG	~ 85	No significant change to k_{cat} or K_m with respect to free CelB	[59]

2.4.2.5 Enzyme loading effects

In general, there appears to be a trend between the virus size and the enzyme loading number (#E) with its overall efficiency as an enzyme nanoreactor (Table 2.1). Smaller viruses such as CCMV, Q β and MS2, which are in the order of 24–28

nm in diameter, experience enhanced efficiency when fewer enzymes are encapsulated. Minten *et al.* showed that encapsulation of more than one PalB in the same CCMV capsid decreased the overall reaction rate. This suggests that molecular crowding plays a significant and direct role on the encapsulated enzymes in such small virus assemblies. This confinement effect shows that one enzyme is enough and that loading more just leads to steric crowding in the interior of the smaller viruses.

In contrast, the number of enzymes encapsulated in larger virus assemblies such as P22 is typically high (85–250 enzymes per capsid). However, despite the absolute number of encapsulated enzymes, these conditions closely resemble true molecular crowding and volume occupancy found in cells. Surprisingly, under these conditions, the overall efficiency and the apparent affinity were in both cases reported to be lower compared to the free enzyme. Nevertheless, it should be noted that, unlike the cell, whereby the actual concentration of a single enzyme in a cell is typically low, the encapsulation of a higher number of the same enzyme might lead to self-inhibition.

The virus capsid not only provides a confined and protected environment but also creates a physical barrier, which may influence the diffusion of substrates across the capsid shell and will behave differently compared to traditional membrane-based organelles. There are currently few examples reporting the substrate diffusion rates between different viruses. Comellas–Aragones *et al.* showed that the diffusion of rhodamine 6G in and out of CCMV is hindered for smaller pore sizes, whereas for the larger pore sizes (2 nm) free diffusion was observed.[35] Recently, Douglas and co-workers reported the same observation for bacteriophage P22.[43] In this work, it was shown that increasing the pore size from 2 to 10 nm did not result in an increase in substrate and cofactor diffusion, suggesting that a pore size of 2 nm is sufficient to allow free diffusion across the capsid.

2.5 Future outlook

All studies on both fluorescent proteins and enzymes show the protective character of the protein nanocage, be it for thermal, protease or other protection. This shielding feature of the protein cage is probably the strongest appeal to one's imagination. However, the nanoscopic effects of confinement for both fluorescent proteins and enzymes are what is most interesting to nanotechnology.

The ability to use fluorescent proteins in the interior of the virus-like particles is very interesting for virology. The knowledge that the fluorophore absorption and emission are virtually unchanged makes those assemblies very interesting for (fluorescent) microscopic probing. But the real challenge comes from the

observation that molecular crowding inside the nanometer sized cages can be studied by the fluorescent properties.

For the encapsulation of enzymes it has been shown that substrate diffusion, enzyme folding (and hence the apparent substrate binding affinity, K_m) and enzyme immobilization play important roles in the overall enzyme efficiency (k_{cat}/K_m rates). Similarly, the current approaches to encapsulate enzymes inside viruses typically involve chemical or genetic modification or tethering of the enzymes to the VLP. This may not only restrict the enzyme activity but also the mobility and conformation of the enzyme itself. Currently, for all models thus far, normal Michaelis–Menten enzyme kinetics has been assumed. However, it is becoming increasingly apparent that standard Michaelis–Menten kinetics, and hence the standard assumptions, may not be valid for enzymes that are within a confined environment. In particular, two parameters, namely the “free diffusion” assumption and the apparent affinity rates (K_{on}/K_{off}) need to be re-evaluated for confined systems. While fractal kinetic theory has been used to describe enzyme kinetics within the confined environment of cells and organelles, the exact parameters that define such models and their relevance to biochemical pathways are still currently under debate [62]. In the meantime, alternative enzyme loading strategies that do not alter or modify the enzyme itself, or binding through reversible linkers (*e.g.* photocleavable linkers) could be further explored.

Nevertheless, in recent years, more virus-like assemblies such as encapsulins (Chapter 5 and 6), carboxysomes and *de novo* nanocontainers (such as the AaLS) have emerged as highly promising examples for the future generation of nanocages and enzyme nanoreactors. Nonetheless, the examples that currently exist for virus cages have already contributed greatly to the understanding of the confined environment of organelles and cells and will probably do so quite some more in the future. The probing abilities of fluorescent proteins, the enhanced efficiency of enzymes and the possibility of confined cascade reactions, combined with the protective properties seen in the different systems described above, still only hint at the possibilities for future loaded nanocages.

2.6 References

1. Dawkins, R., *The selfish gene*. 1976, New York: Oxford University Press. xi, 224 p.
2. Oehler, S. and B. Muller-Hill, *High local concentration: a fundamental strategy of life*. *J Mol Biol*, 2010. **395**(2): p. 242-53.
3. Roodbeen, R. and J.C.M.v. Hest, *Synthetic cells and organelles: compartmentalization strategies*. *BioEssays*, 2009. **31**(12): p. 1299-1308.
4. Tanaka, S., M.R. Sawaya, and T.O. Yeates, *Structure and mechanisms of a protein-based organelle in Escherichia coli*. *Science*, 2010. **327**(5961): p. 81-4.
5. Kang, S. and T. Douglas, *Some Enzymes Just Need a Space of Their Own*. *Science*, 2010. **327**(5961): p. 42-43.
6. Service, R.F., *How far can we push chemical self-assembly?* *Science*, 2005. **309**(5731): p. 95.
7. Gibson, D.G., et al., *Creation of a bacterial cell controlled by a chemically synthesized genome*. *Science*, 2010. **329**(5987): p. 52-6.
8. Vriezema, D.M., et al., *Self-assembled nanoreactors*. *Chem Rev*, 2005. **105**(4): p. 1445-89.
9. Renggli, K., et al., *Selective and Responsive Nanoreactors*. *Advanced Functional Materials*, 2011. **21**(7): p. 1241-1259.
10. van Hest, J.C., *Micro- and nanocapsules: a lively field of research*. *Macromolecular Bioscience*, 2010. **10**(5): p. 463-4.
11. Tanner, P., et al., *Can polymeric vesicles that confine enzymatic reactions act as simplified organelles?* *FEBS Lett*, 2011. **585**(11): p. 1699-706.
12. Tanner, P., et al., *Polymeric vesicles: from drug carriers to nanoreactors and artificial organelles*. *Acc Chem Res*, 2011. **44**(10): p. 1039-49.
13. Algar, W.R., et al., *The controlled display of biomolecules on nanoparticles: a challenge suited to bioorthogonal chemistry*. *Bioconjug Chem*, 2011. **22**(5): p. 825-58.
14. Kreft, O., et al., *Polymer microcapsules as mobile local pH-sensors*. *Journal of Materials Chemistry*, 2007. **17**(42): p. 4471-4476.

15. Kim, K.T., et al., *A Polymersome Nanoreactor with Controllable Permeability Induced by Stimuli-Responsive Block Copolymers*. *Advanced Materials*, 2009. **21**(27): p. 2787-2791.
16. van Dongen, S.F.M., et al., *A Three-Enzyme Cascade Reaction through Positional Assembly of Enzymes in a Polymersome Nanoreactor*. *Chemistry-a European Journal*, 2009. **15**(5): p. 1107-1114.
17. Lalli, M., S.J. Facey, and B. Hauer, *Protein containers--promising tools for the future*. *Chembiochem*, 2011. **12**(10): p. 1519-21.
18. Bobik, T.A., *Polyhedral organelles compartmenting bacterial metabolic processes*. *Appl Microbiol Biotechnol*, 2006. **70**(5): p. 517-25.
19. Yeates, T.O., C.S. Crowley, and S. Tanaka, *Bacterial microcompartment organelles: protein shell structure and evolution*. *Annu Rev Biophys*, 2010. **39**: p. 185-205.
20. Uchida, M., et al., *Biological containers: Protein cages as multifunctional nanoplatforms*. *Advanced Materials*, 2007. **19**(8): p. 1025-1042.
21. Fan, C., et al., *Short N-terminal sequences package proteins into bacterial microcompartments*. *Proc Natl Acad Sci U S A*, 2010. **107**(16): p. 7509-14.
22. Kim, E.Y. and D. Tullman-Ercek, *Engineering nanoscale protein compartments for synthetic organelles*. *Curr Opin Biotechnol*, 2012. **24**: p. 1-6.
23. Arora, P.S. and K. Kirshenbaum, *Nano-tailoring: Stitching alterations on viral coats*. *Chemistry & Biology*, 2004. **11**(4): p. 418-420.
24. Aljabali, A.A., et al., *Controlled immobilisation of active enzymes on the cowpea mosaic virus capsid*. *Nanoscale*, 2012. **4**(18): p. 5640-5.
25. Destito, G., A. Schneemann, and M. Manchester, *Biomedical Nanotechnology Using Virus-Based Nanoparticles*, in *Viruses and Nanotechnology*, M. Manchester and N.F. Steinmetz, Editors. 2009. p. 95-122.
26. Manchester, M. and P. Singh, *Virus-based nanoparticles (VNPs): Platform technologies for diagnostic imaging*. *Advanced Drug Delivery Reviews*, 2006. **58**(14): p. 1505-1522.
27. Steinmetz, N.F., et al., *Structure-Based Engineering of an Icosahedral Virus for Nanomedicine and Nanotechnology*, in *Viruses and Nanotechnology*, M. Manchester and N.F. Steinmetz, Editors. 2009. p. 23-58.
28. Bruckman, M.A., et al., *Role of Hexahistidine in Directed Nanoassemblies of Tobacco Mosaic Virus Coat Protein*. *ACS Nano*, 2011. **5**(3): p. 1606-1616.

29. Gerasopoulos, K., et al., *Biofabrication methods for the patterned assembly and synthesis of viral nanotemplates*. Nanotechnology, 2010. **21**(5): p. 055304.
30. Liu, N., et al., *Au nanocrystals grown on a better-defined one-dimensional tobacco mosaic virus coated protein template genetically modified by a hexahistidine tag*. Nanotechnology, 2012. **23**(33): p. 335602.
31. Wen, A.M., et al., *Viral nanoparticles for in vivo tumor imaging*. Journal of visualized experiments : JoVE, 2012(69): p. e4352.
32. Soto, C.M. and B.R. Ratna, *Virus hybrids as nanomaterials for biotechnology*. Curr Opin Biotechnol, 2010. **21**(4): p. 426-38.
33. Cardinale, D., N. Carette, and T. Michon, *Virus scaffolds as enzyme nano-carriers*. Trends in Biotechnology, 2012. **30**(7): p. 369-376.
34. Carette, N., et al., *A virus-based biocatalyst*. Nature Nanotechnology, 2007. **2**(4): p. 226-229.
35. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor*. Nat Nanotechnol, 2007. **2**(10): p. 635-9.
36. de la Escosura, A., R.J.M. Nolte, and J.J.L.M. Cornelissen, *Viruses and protein cages as nanocontainers and nanoreactors*. Journal of Materials Chemistry, 2009. **19**(16): p. 2274-2278.
37. Dunker, A.K. and A. Fernandez, *Engineering productive enzyme confinement*. Trends in Biotechnology, 2007. **25**(5): p. 189-190.
38. Liu, H., et al., *Pseudo-atomic models of swollen CCMV from cryo-electron microscopy data*. J Struct Biol, 2003. **142**(3): p. 356-63.
39. Glasgow, J.E., et al., *Osmolyte-Mediated Encapsulation of Proteins inside MS2 Viral Capsids*. ACS Nano, 2012. **6**(10): p. 8658-64.
40. Auton, M., et al., *Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains*. Biophysical Chemistry, 2011. **159**(1): p. 90-99.
41. Khan, S.H., et al., *Naturally Occurring Organic Osmolytes: From Cell Physiology to Disease Prevention*. Iubmb Life, 2010. **62**(12): p. 891-895.
42. Golmohammadi, R., et al., *The crystal structure of bacteriophage Q β at 3.5 Å resolution*. Structure, 1996. **4**(5): p. 543-554.

43. Patterson, D.P., P.E. Prevelige, and T. Douglas, *Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage P22*. ACS Nano, 2012. **6**(6): p. 5000-9.
44. Inoue, T., et al., *Engineering of SV40-based nano-capsules for delivery of heterologous proteins as fusions with the minor capsid proteins VP2/3*. J Biotechnol, 2008. **134**(1-2): p. 181-92.
45. Worsdorfer, B., K.J. Woycechowsky, and D. Hilvert, *Directed evolution of a protein container*. Science, 2011. **331**(6017): p. 589-92.
46. Worsdorfer, B., Z. Pianowski, and D. Hilvert, *Efficient in vitro encapsulation of protein cargo by an engineered protein container*. J Am Chem Soc, 2012. **134**(2): p. 909-11.
47. Minten, I.J., R.J.M. Nolte, and J.J.L.M. Cornelissen, *Complex assembly behavior during the encapsulation of green fluorescent protein analogs in virus derived protein capsules*. Macromol Biosci, 2010. **10**(5): p. 539-45.
48. Fiedler, J.D., et al., *RNA-directed packaging of enzymes within virus-like particles*. Angew Chem Int Ed Engl, 2010. **49**(50): p. 9648-51.
49. Rhee, J.K., et al., *Colorful virus-like particles: fluorescent protein packaging by the Qbeta capsid*. Biomacromolecules, 2011. **12**(11): p. 3977-81.
50. O'Neil, A., et al., *Coconfinement of Fluorescent Proteins: Spatially Enforced Communication of GFP and mCherry Encapsulated within the P22 Capsid*. Biomacromolecules, 2012. **13**(12): p. 3902-3907.
51. O'Neil, A., et al., *Genetically programmed in vivo packaging of protein cargo and its controlled release from bacteriophage P22*. Angew Chem Int Ed Engl, 2011. **50**(32): p. 7425-8.
52. Ellis, R.J., *Macromolecular crowding: obvious but underappreciated*. Trends in Biochemical Sciences, 2001. **26**(10): p. 597-604.
53. Zhou, H.X., G.N. Rivas, and A.P. Minton, *Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences*, in *Annual Review of Biophysics*. 2008, Annual Reviews: Palo Alto. p. 375-397.
54. *Viralzone*. Available from: www.expasy.ch/viralzone
55. Mullaney, J.M., et al., *Green fluorescent protein as a probe of rotational mobility within bacteriophage T4*. J Virol Methods, 2000. **88**(1): p. 35-40.

-
56. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids*. Journal of the American Chemical Society, 2009. **131**(49): p. 17771-17773.
 57. Minten, I.J., et al., *Catalytic capsids: the art of confinement*. Chemical Science, 2011. **2**(2): p. 358-362.
 58. Witus, L.S. and M.B. Francis, *Using Synthetically Modified Proteins to Make New Materials*. Accounts of Chemical Research, 2011. **44**(9): p. 774-783.
 59. Patterson, D.P., et al., *Virus-like particle nanoreactors: programmed encapsulation of the thermostable CelB glycosidase inside the P22 capsid*. Soft Matter, 2012. **8**(39): p. 10158-10166.
 60. Ellis, R.J., *Macromolecular crowding: an important but neglected aspect of the intracellular environment*. Current Opinion in Structural Biology, 2001. **11**(1): p. 114-119.
 61. Minton, A.P., *The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media*. Journal of Biological Chemistry, 2001. **276**(14): p. 10577-10580.
 62. Xu, F. and H.S. Ding, *A new kinetic model for heterogeneous (or spatially confined) enzymatic catalysis: Contributions from the fractal and jamming (overcrowding) effects*. Applied Catalysis a-General, 2007. **317**(1): p. 70-81.



Chapter 3

Loading methods for CCMV virus-like particles

This chapter describes different strategies for loading (protein-based) molecular cargo inside the virus capsid of Cowpea Chlorotic Mottle Virus (CCMV). Viral capsids and virus like particles (VLPs) have been of increasing interest from both a biological and nanotechnology perspective in recent years. However, in many cases, the loading of molecular cargo in CCMV had a rather limited efficiency. Here, two different rational design approaches are studied, based on covalent and non-covalent coupling, for loading teal fluorescent proteins (TFP) into CCMV.

Part of this chapter is based on W. Frederik Rurup, Fabian C. Verbij, Melissa S. T. Koay, Christian Blum, Vinod Subramaniam, Jeroen J. L. M. Cornelissen “Controlled loading of virus-like particles” manuscript in preparation

3.1 Introduction

3.1.1 *Viruses*

Viruses are small infectious agents that generally consist of the combination of genetic information (DNA or RNA), an outer protein capsid shell and in some cases, a lipid envelope. According to the biological definition, viruses are not alive as they are dependent on the metabolic mechanisms of the host cell in order to survive. However, despite its simplicity, their ability to target all types of organisms (*i.e.* humans, plants and bacteria) and efficiently deliver genetic cargo serves as a constant source of inspiration to understand (complex) biological systems.

In nanotechnology, viruses have been explored extensively for applications in bionanomedicine, particularly in gene therapy, biomaterials and as drug delivery carrier systems. However, in addition to biomedical applications, viruses (in particular, the virus capsid of non-enveloped plant viruses) have proven to be highly attractive scaffolds for the formation of nanoassemblies, nanocompartments and for the fabrication of optic and electronic devices.[1, 2] In all cases, controlled hierarchical assembly of the macromolecular building blocks into the nanometer scaled architectures plays an important role. Although various approaches to design nanoscaled assemblies have been demonstrated [3-5], there are few examples that are as well-defined and monodisperse as virus-like particles (VLPs). One of the most well-studied viruses is that of Cowpea Chlorotic Mottle Virus (CCMV) [6, 7].

Upto now, the loading of CCMV was either done statistically or through a non-covalent loading mechanism, which will be discussed in detail shortly. It was found that the statistical loading was very inefficient and that the non-covalent system had possible drawbacks, inhibiting the VLP loading (Chapter 2). Other loading techniques and VLPs are also discussed in Chapter 2, yet significantly, it was found that in the large P22 cage ~250 fluorescent proteins could be encapsulated. The same P22 study also showed, that a fusion protein of two different fluorescent proteins shows slightly more energy transfer when encapsulated [8]. Interestingly, Black and co-workers [9] found energy transfer between fluorescent proteins of the same type in T4 capsids, upon the confinement of multiple copies of the protein. However, these effects only occur when >100 fluorescent proteins are encapsulated in these fairly large virus-like particles (Chapter 2). The smaller CCMV capsid likely allows for a more detailed analysis of the loading and the

confinement effect on the cargo, however, therefore *control* is needed over the CCMV loading.

3.1.2 *Cowpea Chlorotic Mottle Virus (CCMV)*

The Cowpea Chlorotic Mottle Virus (CCMV) is a member of the Bromovirus genus of the Bromoviridae family [10]. CCMV assembles into an icosahedral capsid composed of 180 chemically identical protein subunits in which each subunit is composed of 190 amino acids with a mass of 20.3 kDa.[11] The icosahedral capsid is described by a Caspar–Klug triangulation $T = 3$ symmetry, *i.e.* a composition of 12 pentameric and 20 hexameric proteins subunits, with an outer diameter of 28 nm and an inner diameter of 18 nm.[10-13]

The wild type viral capsids filled with RNA are known to be stable between pH 3 and pH 6 at low ionic strength ($I \sim 0.1$ M). When the pH is increased to $\text{pH} > 7.5$ and a high ionic strength ($I \sim 1$ M), the virus disassembles into capsid protein homodimers and RNA, from which the capsid protein can be separated (Figure 3.1).[14] The homo-dimeric state is stable whereby the C-termini of two neighboring capsid proteins interlock to form a stable hydrophilic dimer.[11] Due to its reversible pH driven assembly/disassembly,[6] we use the CCMV virus capsid to gain control over loading of cargo on a molecular level to study the VLP loading and effects of molecular confinement.[7]

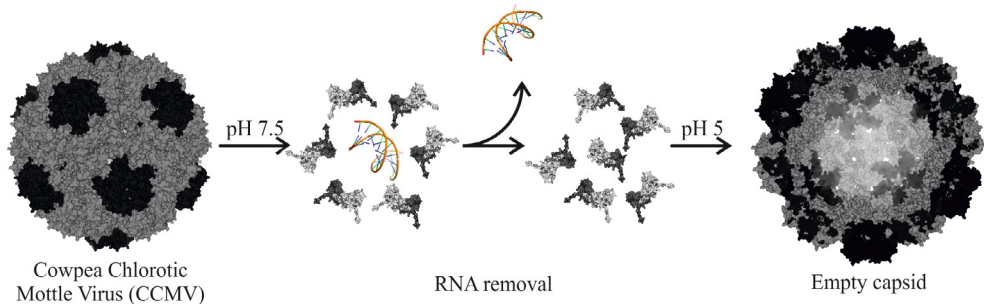


Figure 3.1: The disassembly/assembly cycle of CCMV and its empty VLP. (Based on PDB entry 1za7, rendered with PyMol.)

Isolated wild type capsid proteins (wtCP) form empty capsids under conditions of high ionic strength ($I \sim 0.2 - 1$ M) at $\text{pH} \leq 5$ (Figure 3.1), which can be further stabilized in the presence of Mg^{2+} -ions. At pH 5.0, CCMV capsids contain small < 1 nm pores on the threefold axis and < 0.5 nm on the fivefold axis, respectively. At near physiological conditions ($\text{pH} \sim 7.0$), the capsid is swollen $\sim 10\%$ resulting in

the opening of extra pores on the *quasi*-threefold axis (up to ~ 2 nm) while the virus assembly itself remains stable (Figure 3.2).[15, 16]

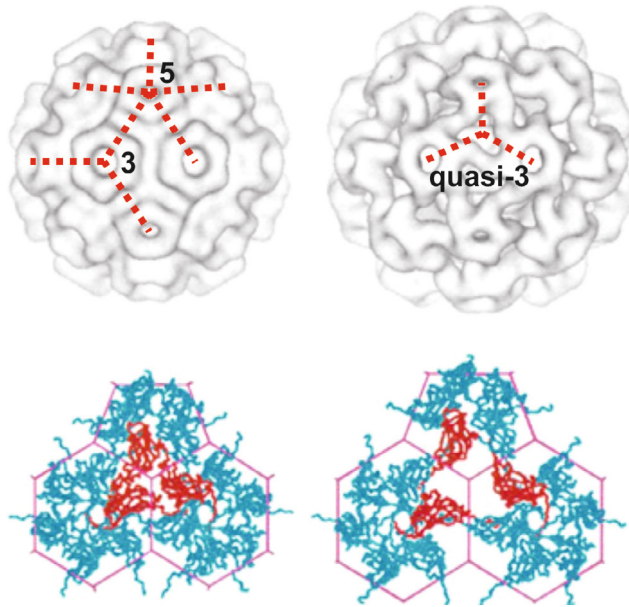


Figure 3.2: The capsid in normal condition (left) and swollen (right). The upper images are cryo-EM reconstructions and the lower images are schematic representations of the proteins. Adopted from Lui *et al.*[15]

3.1.3 Loading of CCMV

The native encapsulation of RNA in CCMV relies on electrostatic interactions between the positively charged N-terminus of the CCMV capsid protein and the negatively charged RNA phosphate backbone. Upon removal of the native RNA, this affinity of the wtCP for negative charges can be, and has been, used for the *in vitro* encapsulation of various synthetic materials, such as anionic polymers and nanoparticles.[17, 18]

However, although CCMV virus-like particles form well-defined assemblies in the presence of negatively charged templates, the *controlled* loading of proteins is more challenging. While statistical encapsulation has been shown to be possible, it is a highly concentration dependent process that is very inefficient, causing a large amount of cargo material to be lost upon purification of the loaded CCMV particles.

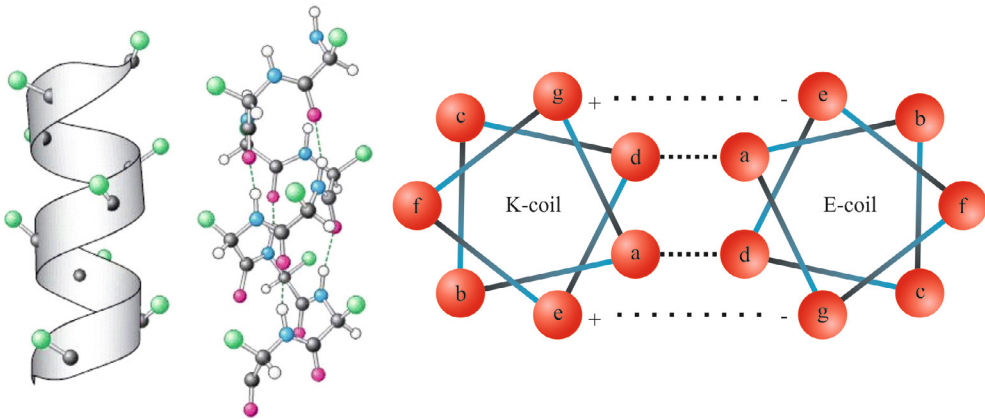


Figure 3.3: Leucine zipper build-up. Left shows the repeating nature of amino acid α -helices, right shows the amino acid position in top-down view of the two coils interacting.

To increase the encapsulation efficiency of neutral or positively charged proteins, non-covalent affinity tags have been used to promote and direct the loading of proteins inside CCMV.[4, 19] Introduction of a positively charged (K-coil) α -helix to the N-terminus of the CCMV capsid protein and a complementary negatively charged (E-coil) α -helix to the C terminus of the cargo protein induces the formation of a non-covalent capsid protein-cargo complex. The resulting heterodimerization of the leucine zipper E-K coiled-coil peptides has a very low dissociation constant ($K_d = 7 \times 10^{-8}$ M at pH 7), indicating a very stable non-covalent complex, and can effectively guide cargo proteins inside the assembling CCMV capsid. In this example, the coiled-coil peptides are composed of $(abcdefg)_n$ repeating units, in which the K-coil contains an $(IAALKEK)_3$ repeat and the E-coil an $(IAALEKE)_3$ repeat (Figure 3.3). The isoleucine and leucine residues (a & d position) form strong hydrophobic interactions whereas the alanines (b & c position) are good α -helix builders and stabilize the helix. Finally, the e , f and g positions are charged at neutral pH and promote the heterodimerization, *i.e.* K to E binding.

By genetically engineering the complementary coiled-coil peptides on CCMV and the desired fluorescent protein cargo, fluorescent proteins could be directed inside CCMV. However, despite the stoichiometric heterodimerization of E-K coiled-coils at pH 7, during CCMV assembly at pH 5.0, the K-coil is known to unfold and the E-coil homotrimerizes, thereby significantly reducing the loading efficiency.[20, 21]

In this chapter, we sought to optimize the design for the rational loading of CCMV by creating more stable ‘capsid protein-cargo’ complex variants. Two different

approaches were investigated: (1) increasing the intrinsic ratio between the cargo and CCMV capsid protein and (2) genetically designing a flexible peptide linker whereby the cargo is fused to the CCMV capsid protein.

3.2 Results and discussion

3.2.1 *Improving the fluorescent protein characteristics*

In the past, directed loading studies in CCMV were performed using the enhanced green fluorescent protein (EGFP). EGFP E-coil (EGE) was used with the CCMV capsid protein K-coil (CK) to form the EGECK complex. Despite its improved photophysical properties, EGFP still has a unstable fluorophore at lower pH (pKa = 6.0), suggesting it is not favorable for studying capsid formation and encapsulation at pH 5.0.[22, 23] Therefore, in this work, EGFP was replaced with the monomeric teal fluorescent protein (mTFP), which has a pKa of 4.3 (Table 3.1).[24]

Table 3.1: Overview of the different photophysical properties of mEGFP compared to mTFP.[22] * Brightness is defined as the product of the molar extinction coefficient and the quantum yield.

Protein	Excitation peak (nm)	Emission peak (nm)	Brightness ($\text{mM}^{-1}\cdot\text{cm}^{-1}$)*	Photostability (s)	pKa	Association state
mEGFP	488	507	34	174	6.0	monomer
mTFP	462	492	54	110	4.3	monomer

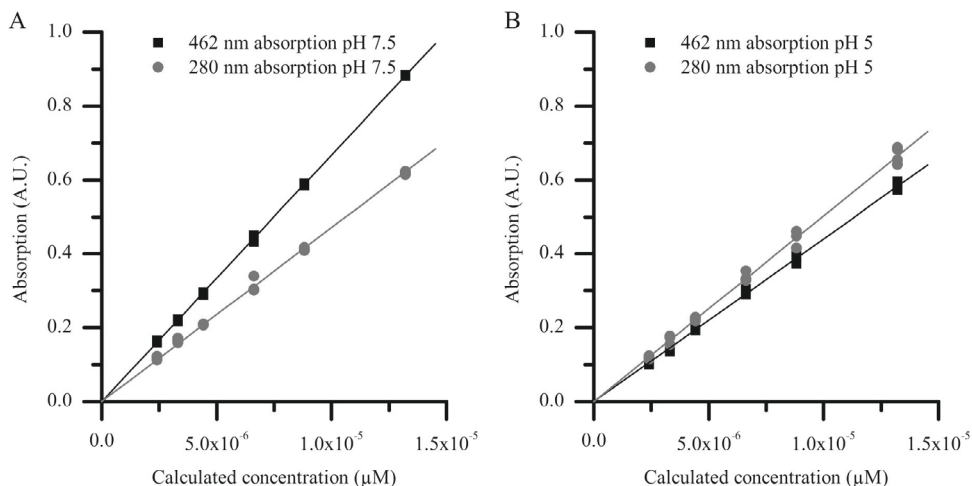


Figure 3.4: mTFP concentration dependent standard curve series at (A) pH 7.5 and (B) pH 5.

mTFP originates from a tetrameric fluorescent protein from *Clavularia sp.* coral that has been genetically engineered to be monomeric, bright and pH stable.[24] The free mTFP exhibits fluorescence excitation and emission maxima at $\lambda=462$ nm and $\lambda=492$ nm, respectively, the full spectra are part of Figure 3.12.[24] The reported extinction coefficient at $\lambda=462$ nm (ϵ_{462}) is $64\,000\text{ M}^{-1}\text{ cm}^{-1}$.[24] This was confirmed in these studies based on a concentration dependent standard curve series at pH 7.5 (Figure 3.4A), using mTFP as prepared in Section 3.5.

Based on the same set of experiments, the extinction coefficient at $\lambda=280$ nm at pH 7.5 was determined to be $45\times 10^3\text{ M}^{-1}\text{ cm}^{-1}$. Since the extinction coefficient at pH 5.0 is unknown, another series was performed at pH 5.0 (Figure 3.4B) that mimics the conditions during encapsulation in CCMV. Assuming the concentrations in each dilution series are consistent, the extinction coefficients for mTFP at pH 5.0 were $\epsilon_{462} = 42\,000\text{ M}^{-1}\text{ cm}^{-1}$ and $\epsilon_{280} = 48\,000\text{ M}^{-1}\text{ cm}^{-1}$.

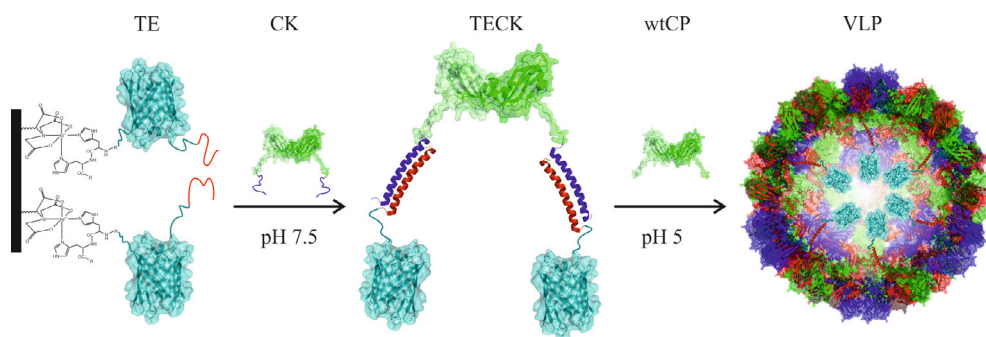


Figure 3.5: TFP with an E-coil (TE) bearing an N-terminal His-tag is immobilized on Ni-NTA resin and combined with CCMV capsid protein with a K-coil (CK) to form the TECK complex at pH 7.5. Upon mixing with wild type CCMV capsid proteins and lowering to pH 5.0, CCMV virus-like particles encapsulating mTFP are formed.

For encapsulation, mTFP was equipped with an E-coil (TE) and immobilized on a Ni-NTA affinity column via the N-terminal His₆-tag. The immobilized TE was allowed to complex with CK to form the non-covalent TFP capsid protein complex (TECK). TECK is subsequently mixed with wild type capsid proteins (wtCP), *i.e.* capsid protein from the native virus source, and upon lowering the pH to pH = 5.0 mTFP is encapsulated inside CCMV VLPs (Figure 3.5). The ratios of TECK and wtCP can be varied to change the amount of mTFP encapsulated by this procedure. Although mTFP has the same global shape as EGFP, the amino acid sequence differs. This may give rise to a different loading behavior since the residues on the exterior surface of the fluorescent protein are different.

Based on the absorbance intensities, the loading efficiency in CCMV could be determined. For detailed calculations, see section 3.5.7. Interestingly, TECK showed a significantly lower loading efficiency compared to EGECK (Figure 3.6). Instead, the loading efficiency for TECK appears to be in the same order as that of unbound EGECK. In the past, this ratio was attributed to leucine zipper dissociation during assembly at pH 5.0,[21] and may suggest that dissociation of the TECK complex is also playing a role. Nevertheless, for both EGE and TE, the maximum loading was observed at a mixing ratio of 30% complex (*e.g.* EGECK and TECK) with 70% wtCP.

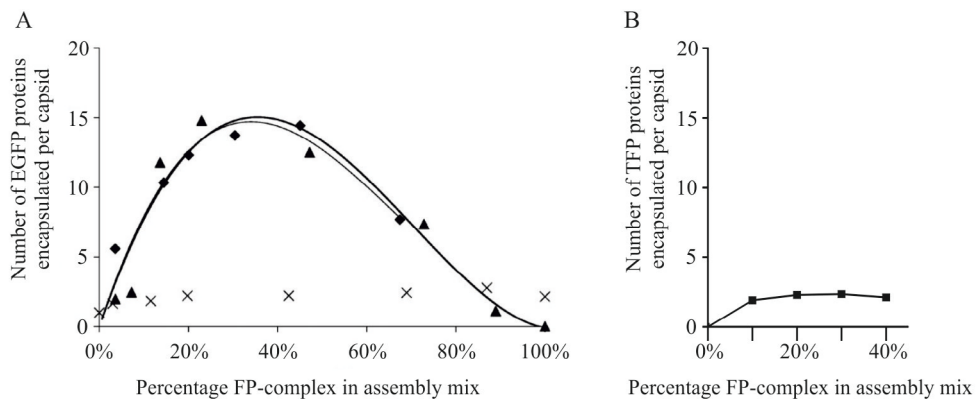


Figure 3.6: The loading efficiency for (A) EGFP as reported by Minten *et al.*[4] and (B) TFP (this work). Diamonds and triangles represent data points of duplicate experiments by Minten. Crosses represent negative control experiments in which unbound EGFP with the E-coil was used.

3.2.2 Improving TFP loading

To understand the mechanism and improve the loading of TFP into CCMV a more stable loading complex was designed. Since the CCMV capsid proteins form a strong dimer and hence, present two (K-coil) binding sites per capsid subunit cargo that also favors dimer formation was engineered.

To achieve this, mTFP was mutated to re-introduce one of its natural dimerization interfaces. Ai *et al.* reported two interfaces (A-B and A-C) of dimerization in the original *Clavularia sp.* TFP, which were genetically engineered in the monomeric form. The A-B interface is a parallel interface, where the A-C interface is under a strong angle (Figure 3.7).

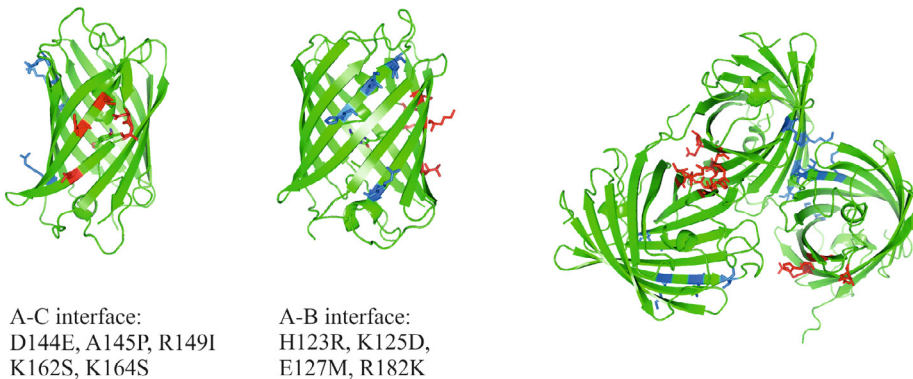


Figure 3.7: The locations of the mutations on mTFP (PDB: 2HQB) and their arrangement with respect to each other in the ‘tetrameric’ arrangement. The A–C interface is shown in red, the A–B interface in blue.

Since the A–C interface forms dimerization at an angle, it was thought that this would minimize any potential steric strain between CCMV dimerization and TFP dimerization. Hence, the A–C interface was initially restored by reintroducing the following mutations: D144E, A145P, R149I, K162S and K164S. Site-directed mutagenesis to re-convert the A–C interface resulted in the dimeric TFP E-coil (dTE) which was confirmed by DNA sequencing. dTFP protein dimerization was also verified by dynamic light scattering (DLS) and native polyacrylamide gel electrophoresis (PAGE) (Figure 3.8). To further investigate whether increasing the number of E-coils on the cargo can further improve the loading efficiency, the A–B interface was mutated to restore the native TFP tetramer by reintroducing the mutations: H123R, K125D, E127M and R182K. These mutations corresponding to the tetrameric TFP E-coil (tTE) construct were also confirmed by DNA sequencing. To verify that the coil dissociation is the underlying problem during TECK loading, a TFP capsid fusion protein was also designed, consisting of an N-terminal His₆-tag directly followed by the mTFP sequence, a flexible linker with a thrombin cleavage site and the capsid protein (see Section 3.5.2). The resultant HTC (His₆-TFP-capsid) sequence was also confirmed by DNA sequencing.

The individual proteins (TE, dTE, tTE, CK and HTC) were expressed recombinantly in *Escherichia coli* and purified by a combination of Immobilized Metal ion Affinity Chromatography (IMAC) and size exclusion fast protein liquid chromatography (FPLC) to obtain the purified TECK, dTECK and tTECK complexes and HTC. The individual TFP variants and their complexes with CCMV capsid proteins were analyzed by FPLC, dynamic light scattering and gel

electrophoresis at pH 7.5 (Figure 3.8). No changes were observed in the absorption spectra dTE and tTE with respect to TE.

As can be seen in Figure 3.8A, the presence of the E-coil linker to the TFP (TE) significantly decreases the elution time in the FPLC compared to TFP lacking the E coil (from 11.5 to 10.3 mL), consistent with a larger apparent size. Surprisingly, both the dimer and tetramer with E-coil (dTE and tTE) appear to have a slightly larger elution volume than the monomer with E-coil (10.4 and 10.5 mL respectively), where we expected it to be less. Upon complexation with the capsid K-coil, the elution volume of the fluorescent proteins are further decreased to 8.7–8.8 mL, consistent with complex formation (Figure 3.8B). The monomeric and dimeric TFP E-coil (TE and dTE) were also analyzed by DLS, which revealed an apparent size of ~5 nm for TE (including the flexible E-coil) and ~8 nm for dTE, which is significantly larger than the TE size, both at ~50 μ M (Figure 3.8C). These observations are consistent with reported X-ray crystallographic data which shows that TFP is β -barrel shaped, 3 nm in diameter and 4 nm long[24]. The apparent doubling of the size as observed by DLS strongly suggest that dimerization of dTE is restored. Similarly, the native PAGE shows a steady decrease in apparent molecular weight between TE, dTE and tTE variants (Figure 3.8D lanes 1–4, 20 μ L of 10–50 μ M). Although the increase in mass for dimeric and tetrameric TE (dTE and tTE) would be expected to perturb the protein migration rate, the increase in overall negative charge for the dimer and tetramer (*i.e.* single molecules of T = -4, TE = -8, dTE = -11 and tTE = -11, respectively) is much more pronounced and increases the migration rate. Nevertheless, the difference in migration between TE, dTE and tTE indicates a clear, step-wise change in TFP conformational states. Since the denaturing SDS-PAGE shows only one major protein band corresponding to the TFP E-coil (at 31.5 kDa), these changes in conformational states are attributed to non-covalent multimer formation.

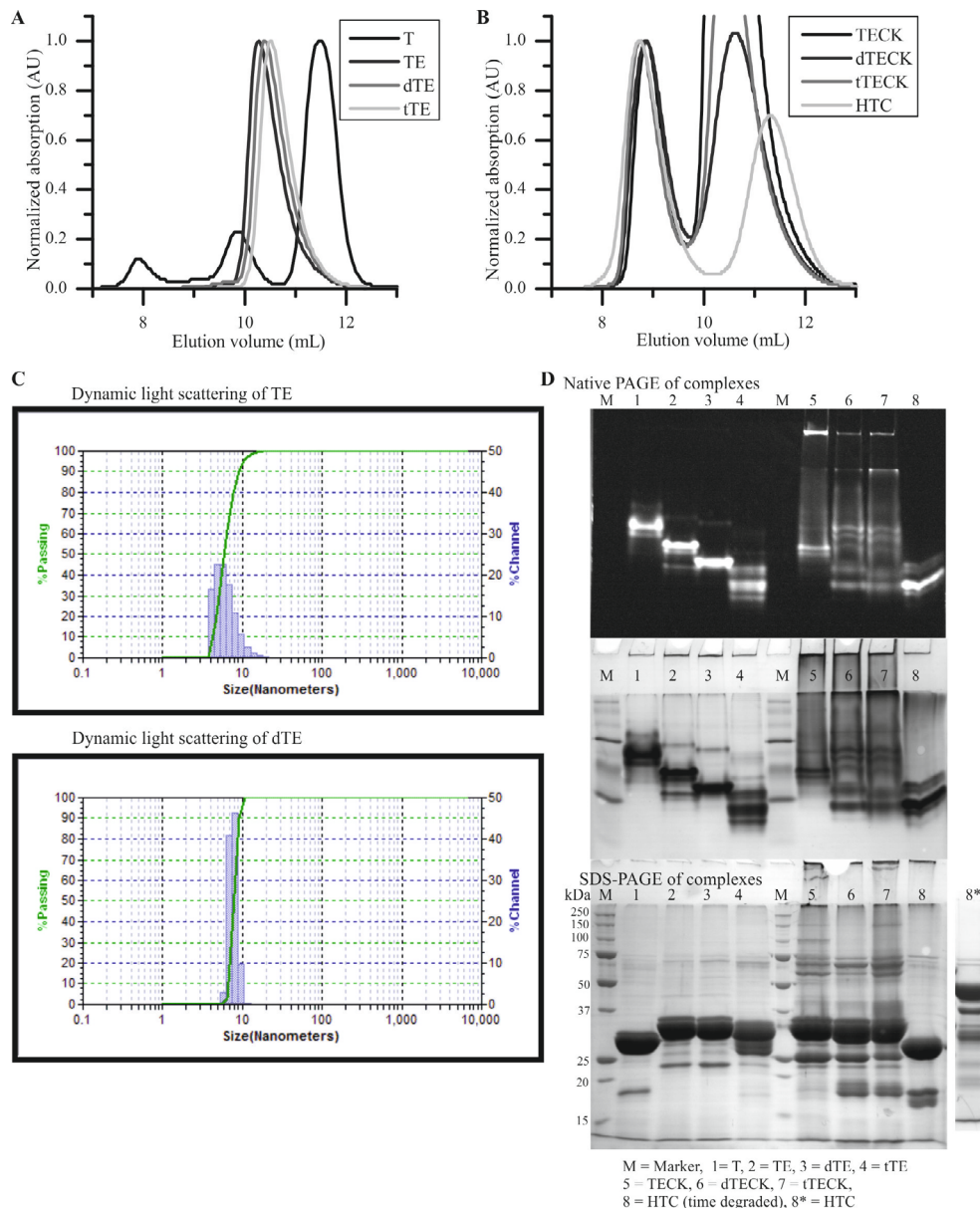


Figure 3.8: Characterization of the dimerization states of the TFP variants at pH 7.5. In all cases, T = TFP lacking E coil whereas TE, dTE, tTE = monomeric, dimeric and tetrameric TFP, respectively, with E coils. (A) normalized FPLC elution profiles of the TFP cargo proteins, (B) normalized FPLC elution profile of the TFP complexes, (C) DLS measurements of the TE and dTE, (D) native PAGE (top: fluorescence imaging, middle: white light imaging) and bottom: denaturing SDS-PAGE of the TFP complexes. Since HTC is susceptible to auto-degradation, a fresh sample is also show, scaled and aligned to the marker bands.

In contrast to the individual TE variants which showed a clear migration trend, the native PAGE of the TFP complexes (at pH 7.5) showed largely unresolved bands. Instead, multiple bands are visible suggesting that the electrostatic interactions that keep the E–K coils together are disrupted during electrophoresis (Figure 3.8D lanes 5–8). However, the formation of each TFP complex is evident from the denaturing SDS–PAGE which shows the presence of TE (31.5 kDa) and CK (25 kDa) as major materials and some minor bands of degradation products and aggregates, which are commonly found upon denaturing the proteins with an E–coil or K–coil addition (lanes 1–4). The presence of the two major bands in SDS–PAGE indicates that the multiple bands in native PAGE are likely due to different E–K complexation states. Importantly, the SDS–PAGE also confirms that these interactions are reversible and hence, non–covalent. It should be noted that although native PAGE and FPLC analyses are consistent with the formation of tTE, the results for the tTECK complex remain partly inconclusive since the native PAGE unexpectedly resembles the dTECK complex despite the overall increase in negative charge. In turn the FPLC elution profile for tTECK shows a slight shift towards higher apparent mass compared to dTECK, the elution volume (at 8 mL) is at the limit of resolution.

Since SDS–PAGE is a semi–quantitative technique, the relative intensities between TFP and the capsid protein bands can be used to estimate the TFP complex ratios. The relative intensities of CK and the cargo proteins were analyzed using ImageJ software, in which the background was subtracted and the intensities were corrected for the molecular weight of each protein. Based on these analyses, the TFP monomer per capsid monomer ratio for TE is ~ 1.8, for dTE ~ 2.1 and for tTE ~3.5. We know that all collected complexes have some additional free TFP co–purified in the FPLC elution, due to the overlap of the complex and unbound cargo peaks (Figure 3.8B), which explains why the TFP per capsid ratio still appears to be 1.8 for TE. The values for dTE and tTE are in good agreement with the expected ratios of 2 and 4 respectively and are clearly indicative of the multimer state.

For the HTC fusion protein, the SDS–PAGE shows the expected mass of 50.7 kDa in lane 8*. However, unlike the TE variants, HTC appears to be susceptible to degradation over prolonged storage. Lane 8 shows HTC after 3 months of storage, where the linker appears to be cleaved. Whether this is caused by the thrombin site or the natural degradation of the capsid N–terminus remains unclear.

3.2.3 *Encapsulation properties*

After purification and characterization, each TFP complex was mixed with wild type CCMV capsid proteins in different ratios and dialyzed overnight to pH 5.0, to induce the formation of CCMV VLPs. As shown in Figure 3.9, there is a significant improvement in the loading efficiency, particularly between the monomeric and dimeric forms of TFP, where the TECK maximum was determined to be on average 2.4 TFP per capsid whilst the dTECK contained on average 6.6 TFP per capsid. Interestingly, the tTECK led to only a slight increase in encapsulation efficiency compared to dTECK with 7.2 TFP per capsid.

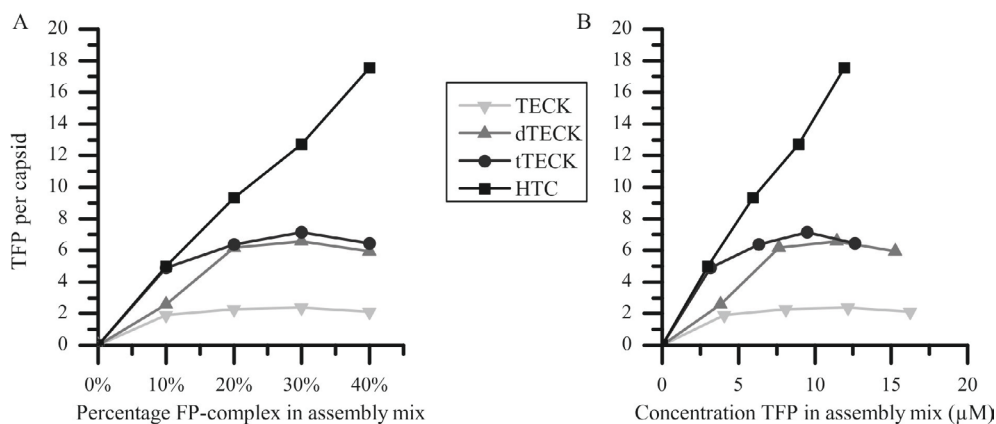


Figure 3.9: (A) Capsid TFP loading efficiency as a function of percentage of complex (TECK, dTECK, tTECK or HTC). (B) Capsid TFP loading efficiency as a function of TFP concentration incubated in the assembly mix. All measurements were performed at pH 5.0.

The loading behavior of the monomer appears to reach a maximum already at 10%, suggesting that E–K dissociation may occur at very low concentrations, consequently causing a low loading efficiency. In contrast, both the dimer and tetramer show more stable E–K binding affinity and hence an improved loading up to higher mixing percentages. Nevertheless, it is clearly shown that for TFP E–coil concentrations in the capsid there is a maximum around 10 μM in each multimer state (Figure 3.9B). In all cases, the loading efficiency reaches a maximum value, after which no more TFP appears to be encapsulated. This may be due to trimerization of the E–coils and dissociation from the capsid K–coils at high concentration, and consequently limiting encapsulation into the capsids. The FPLC chromatograms (Figure 3.10) support this, as seen by the increased 462 nm signal at the 2 mL elution peak (free proteins), with an almost constant 1.2 mL capsid elution peak. On the other hand, the HTC, which is not subject to the E–K coiled

coil hetero-trimerization, shows an almost linear loading behavior (Figure 3.9) and a constant ratio for encapsulated and non-encapsulated TFP (Figure 3.10).

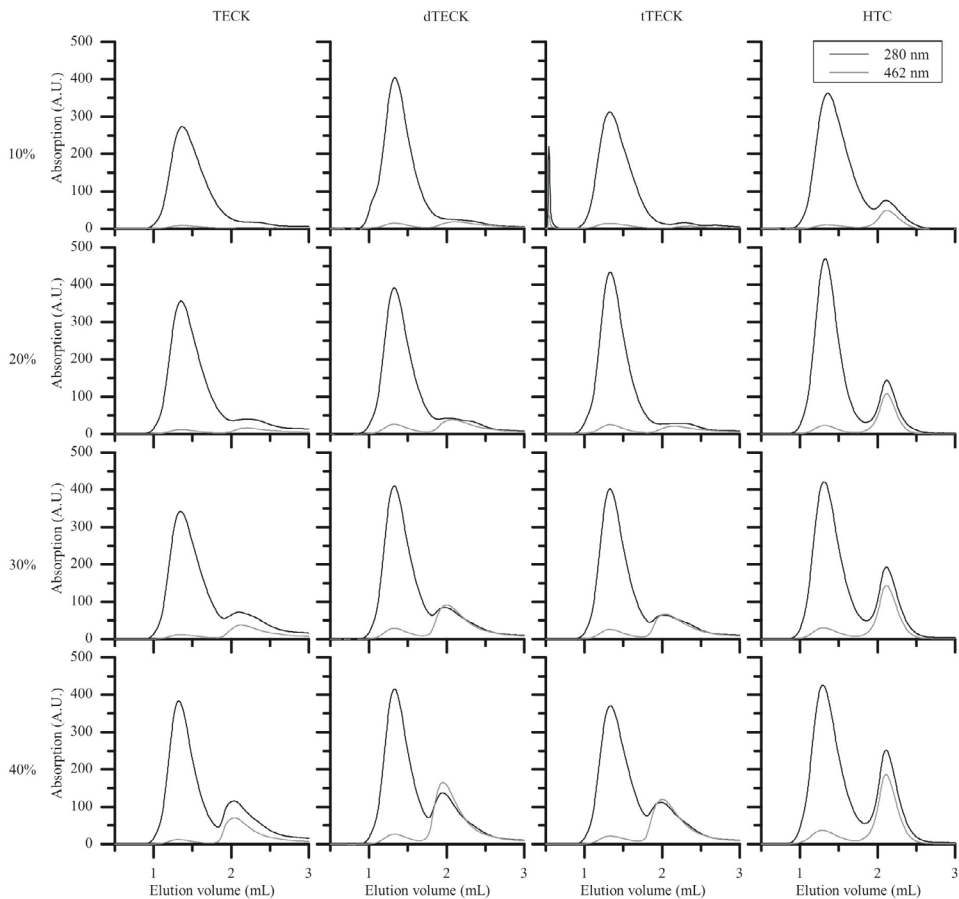


Figure 3.10: FPLC chromatograms of TECK, dTECK, tTECK and HTC for 10%, 20%, 30% and 40% complex of the total capsid protein concentration.

3.2.4 Possible Förster Resonance Energy Transfer (FRET) acceptors

Since the interactions of encapsulated and confined cargo are of interest, a technique was sought that gives information on confinement and interactions. Our cargo is the TFP fluorescent protein and therefore fluorescence spectroscopy to detect Förster Resonance Energy Transfer (FRET) can be highly informative on close range fluorophore interactions. FRET is a phenomenon used to study the molecular interactions between closely packed fluorophores. Upon fluorescence excitation, an excited donor fluorophore transfers its fluorescence energy via long range dipole-dipole interactions to an appropriate acceptor fluorophore without the

emission or loss of an intermediate photon. The efficiency of this energy transfer (E) is dependent of inter-fluorophore distance (r) and the Förster distance (R_0) (Equation 3-1).[25, 26]

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

Equation 3-1

The Förster distance is given by Equation 3-2:

$$R_0 = \left(\frac{9000(\ln 10)\kappa^2 Q_D J(\lambda)}{128\pi^5 N n^4} \right)^{1/6} \approx 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{1/6}$$

Equation 3-2

In Equation 3-2 there are four variables that influence the R_0 and thus the FRET efficiency. The factor κ^2 describes the orientation of the donor and acceptor dipoles with respect to each other, n is the refractive index of the medium in which the molecules are dissolved, Q_D is the quantum yield of the donor and $J(\lambda)$ is the spectral overlap integral of the donor emission spectrum and the acceptor absorption spectrum. The value for Q_D is fluorophore specific, n is assumed to be 1.4 for aqueous solutions of biological molecules (*e.g.* proteins). κ^2 is assumed to be 2/3 for random orientation of the dipole moments.[26] $J(\lambda)$ is the spectral overlap, *i.e.* a partial overlap in the emission spectrum of the donor and the absorption spectrum of the acceptor and thus depends on the fluorophore pair that is chosen. However, more often the R_0 is empirically determined and reported for specific fluorophore pairs.

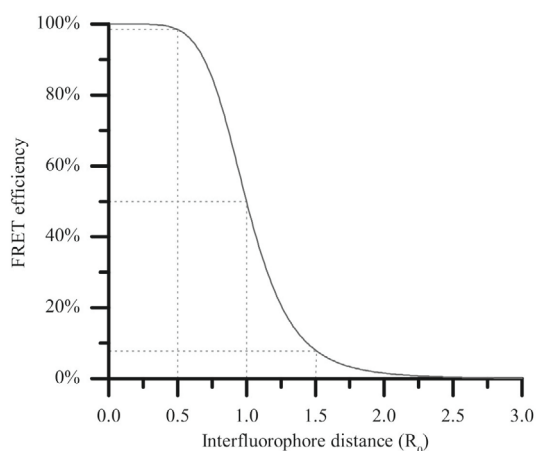


Figure 3.11: FRET efficiency E (Equation 3-1) as a function to multiples of the Förster distance (R_0)

In general, the FRET efficiency is mainly influenced by the inter-fluorophore distance (Equation 3-1). Three specific inter-fluorophore distances (r) are essential for understanding FRET efficiency. The first occurs when r is $1.5R_0$, giving an efficiency of $\sim 10\%$. Any further and the efficiency quickly drops to 0. Second, for $r = R_0$ the efficiency is exact 50% and finally, when r approaches $0.5 R_0$ the efficiency is almost 100 % (Figure 3.11). Thus, to optimize FRET efficiency, the distance should be $1.5 R_0$ or less. However, due to the shape and size of the fluorescent proteins, $0.5 R_0$ cannot be reached for fluorescent proteins because of steric hindrance.

Based on the fluorescent properties of TFP, the first choice of FRET partner was that of the red mStrawberry fluorescent protein.[27] In Figure 3.12 the fluorescence absorption and emission spectra of mTFP and mStrawberry are shown. The overlapping region of the emission of mTFP (donor) and absorption of mStrawberry (acceptor), which is essential for FRET, is shown in grey.

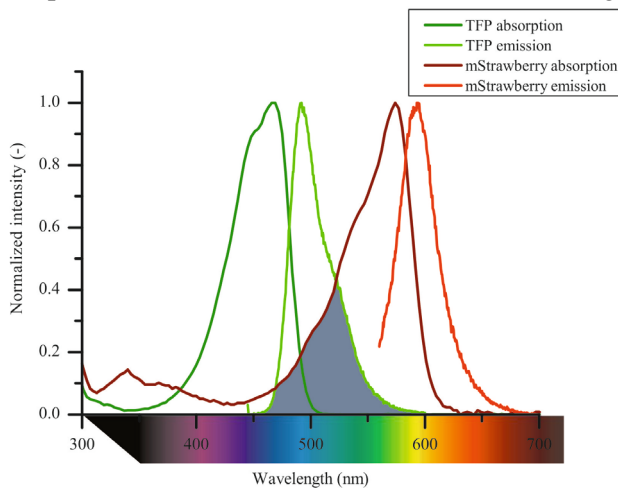


Figure 3.12: Illustration of spectral overlap (gray area) between the absorption of mStrawberry and emission of mTFP. All spectra are normalized to their respective maximum intensity.

The vector for mStrawberry[28] with E-coil (SE) was engineered and provided by Dr. Inge Minten. SE was expressed and purified as for the TE variants and was pre-assembled with the capsid K-coil forming the SECK complex (Figure 3.13). The UV-visible absorbance spectra of both mStrawberry and SECK are shown in Figure 3.14A. The SECK spectrum has a relatively larger 280 signal, due to coiled coils and the capsid proteins.

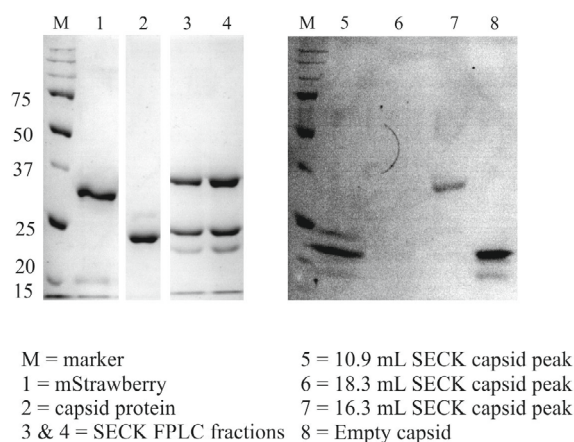


Figure 3.13: SDS–PAGE analysis of mStrawberry and SECK (lanes 1–4) and the FPLC fractions of the encapsulation of mStrawberry as shown in Figure 3.14B.

However, unlike the TECK complexes, dialysis of SECK to pH 5.0 in the presence of wild type CCMV capsid protein (ratio 35:65 SECK: wtCP) mainly yielded a red-colored precipitate, suggesting strong aggregation of SE occurs at low pH. After removal of the aggregated complexes, the remaining sample was purified by FPLC (Figure 3.14B). Although CCMV capsid formation was apparent due to the elution peak at 10.9 mL, the absorbance signal at 280 nm was significantly lower compared to the initial protein concentration (about 10% yield), indicating loss of more than 80% of capsid protein. Similarly, the lack of an absorbance band at 574 nm at 10.9 mL indicated that no mStrawberry was encapsulated in CCMV during assembly. Instead, free unassembled SE and CK eluted at ~ 16.3 mL and 18.3 mL respectively. SDS–PAGE analysis of the different elution peaks confirmed the absence of mStrawberry in the formed capsids (10.9 mL peak), but showed its presence in the 16.3 mL peak (Figure 3.13).

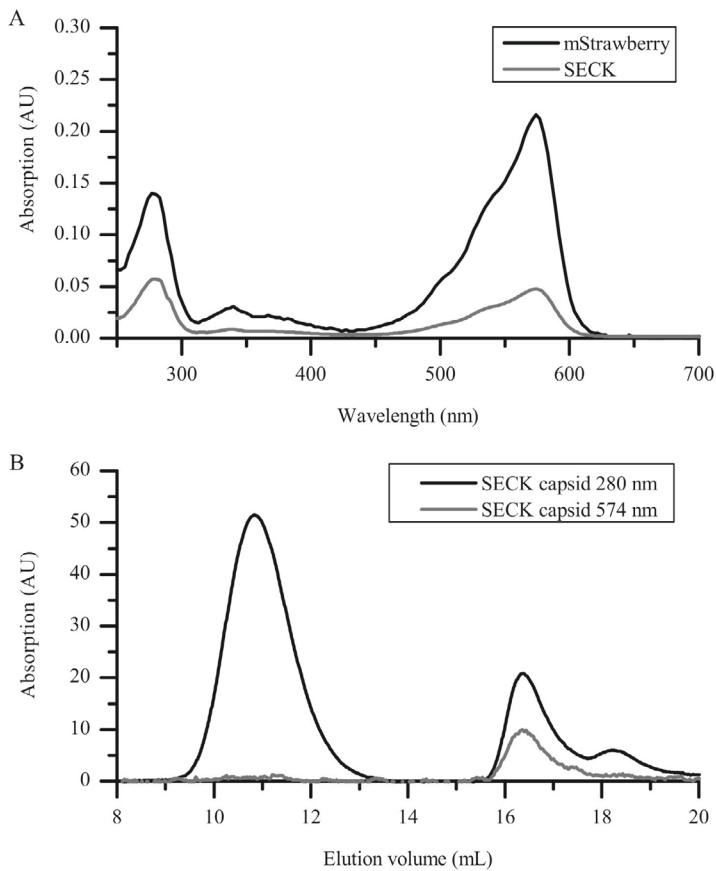


Figure 3.14: (A) UV-Vis spectrum of mStrawberry and the mStrawberry capsid complex (SECK). The intense bands at 280 nm are due to protein absorbance whereas the band at 574 nm is due to the mStrawberry chromophore. (B) The FPLC elution profile of CCMV VLPs encapsulating mStrawberry at pH 5.0 (upon mixing 35% SECK with 65% wild type capsid protein), monitoring at 280 nm for protein absorbance and 574 nm for mStrawberry absorbance.

It was shown that for TE, the dimerization and tetramerization improved the loading efficiency. Although the monomeric mStrawberry did not show effective encapsulation using the E-K coil strategy, attempts were made to use a tetrameric red fluorescent protein. Instead of genetically engineering mStrawberry to form a tetramer, the dsRed fluorescent protein, from *Discosoma* coral [29, 30] was engineered with an E-coil instead. dsRed is a naturally occurring homotetramer with fluorescent properties similar to mStrawberry (absorption 554 nm, emission 586 nm). The dsRed was engineered to form the dsRed E-coil (dsRE), which was confirmed by DNA sequencing. Recombinant expression of dsRE in *E. coli*

resulted in bright pink colored cell pellets, however the purification was quite troublesome due to poor solubility of dsRed E-coil in buffer. IMAC purification was possible, however, strong aggregation was evident which led to irreversible binding to the column. This may be due to the aggregation of E-coils at high concentrations combined with the strong dsRed tetramerization. No further attempts to obtain the above materials were undertaken.

3.3 Conclusion

In conclusion, the non-covalent E-K coil strategy for encapsulation was improved by (re)-introducing the homo-dimerization and tetramerization of the cargo. Based on monomeric TFP, it was shown that dimerization significantly improved the loading efficiency and tetramerization slightly improved the loading efficiency further. In all cases, it was observed that the number of proteins encapsulated reached a maximum value, suggesting that E-K coiled-coil dissociation was the main cause for the limited loading efficiency, which was also postulated for EGFP [21]. Another drawback of the E-K coil system is that it appears to be strongly cargo dependent. Overall, the EGFP (from *Aequorea victoria*) shown by Minten *et al.* showed much more efficient loading compared to the TFP (from *Clavularia sp.*) used in these studies. Furthermore, the mStrawberry (originating from *Discosoma sp.*) performed even less efficiently, which may suggest that even small differences in the cargo (i.e. amino acid composition, overall charge, pKa) greatly influence the loading of CCMV using the E-K coiled coil system. Although the loading of TFP was improved by using different multimers of the protein, the intrinsic tendency of the E-K coiled-coils to homotrimerize leads to a significant loss of cargo and hence, the E-K coiled coil approach does not appear to be suitable for *controlled* loading of cargo into CCMV VLPs.

3.4 Acknowledgements

F.C. Verbij and B. Grunewald are gratefully acknowledged for their help performing experiments in this chapter. Prof. Dr. V. Subramaniam and Dr. I Minten are gratefully acknowledged for providing crucial genetic material.

3.5 Experimental

3.5.1 Genetic engineering of pET15b-mTFP E-coil

To implement the mTFP in our protein production, it was cloned into the pET15b bacterial expression vectors described by Minten *et al.*[4, 19, 21]. Using site-directed mutagenesis (QuikChange mutagenesis, Stratagene), a unique restriction site for BamHI was introduced into the *pET15b EGFP-Ecoil* expression vector using the following primers: forward 5'-GGT GCC GCG CGG GAT CCA TAT GCT CGA GAA AAG AG-3' and its reverse complement (BamHI restriction site underlined).

The DNA fragment encoding mTFP was excised from the *pNCS-mTFP* DNA vector (Allele BioTech, USA) by double digestion with the BamHI and BsrGI restriction enzymes. *pET15b EGFP-Ecoil* was double-digested using the same enzymes, hence removing EGFP for subsequent replacement by TFP. The frame shift caused by the cloning was corrected by site-directed mutagenesis with primer: 5'-GGT GCC GCG CGG GAG CCA TAT GGT GAG CAA GGG CG-3' and its reverse complement. This resulted in a *pET15b TFP-Ecoil* (TE) vector. All sequences were confirmed by DNA sequencing (Eurofins MWG Operon, Germany).

The native dimerization of TFP was restored by re-introducing the mutations into the A-C dimerization site (D144E, A145P, R149I, K162S, K164SCK by site-directed mutagenesis, using the primers: 5'- CCA CCG GCT GGG AGC CCT CCA CCG AGA TCA TGT ACG TGC GCG-3' (D144E, A145P, R149I, mutations are underlined), 5'-GCT GAA GGG CGA CGT CAG CCA CAG CCT GCT GCT GGA GGG CG-3' (K162S, K164S, mutations are underlined) and their reverse complementary strands. The dimeric TFP-Ecoil (dTE) variant was confirmed by DNA sequencing (Eurofins MWG Operon, Germany).

3.5.2 Design and molecular cloning of mTFP – CCMV Capsid fusion protein

The XhoI site (underlined) was introduced into the *pET15b CCMV K-coil[4]* (CK) vector, between the K-coil and the capsid protein, by site-directed mutagenesis using the following primer: 5'-CGC CGC CCT GAA GGA GCT CGA GAT GTC TAC AGT CGG AAC AGG G-3'. The sequence for His₆-tagged TFP was amplified by PCR from the *pET15b TFP-Ecoil* vector using the following forward primer: 5'-GGA ATT GTG AGC GGA TAA-3' and reverse primer: 5'-GCG GCT CGA GGC CAC CAG AGC CGC TCC CGC GCG GCA CCA GAC CCT

GCC ACC *CGT GTA CAG CTC GTC CAT GCC GTC*–3' (XhoI is underlined, the TFP sequence italic). These primers also introduce the sequence for the flexible linker between TFP and the capsid protein and a unique XhoI site in the PCR product. Double digestion by XbaI and XhoI of both the PCR product and the mutated *pET15b CCMV K-coil* vector and subsequent ligation resulted in the *pET15b His-TFP-Capsid* (HTC) vector.

3.5.3 *Design and molecular cloning of dsRed E coil*

The vector for mStrawberry E-coil was provided by Dr. Inge Minten (Radboud University of Nijmegen, The Netherlands). The vector for dsRed was provided by Professor Vinod Subramaniam (University of Twente, The Netherlands). The sequence for dsRed to incorporate in the E-coil vector was amplified by PCR with the following primers: forward 5'– CCA CCG GGA TCC GGC ATG GTB CGC TCC TCC AAG AAC GTC ATC AAG G –3' and reverse 5'– GAG TCG CTA GCG CTA CAG GAA CAG GTG GTG GCG GCC C –3'. The restriction sites for BamHI and NheI are underlined and used to clone the sequence into the pET15 EGFP E-coil vector, replacing EGFP.

3.5.4 *Recombinant protein expression and purification*

The pET15b plasmids containing CK, TE, dTE, tTE and HTC were transformed into *Escherichia coli* BL21(DE3)pLysS (Novagen) for protein expression. Starting cultures were grown overnight at 37°C from glycerol stock cells in 7 mL LB medium (Sigma) containing 100 µg/mL ampicillin, 34 µg/mL chloramphenicol (Sigma). The overnight cultures were used to inoculate 0.5 L LB medium containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) and grown to an optical density of OD₆₀₀ = 0.6–0.8 A.U.. Protein expression was induced following addition of isopropyl β-D-thiogalactoside (IPTG) to a final concentration of 0.1 mM at 30 °C for 4–5 hours.

The cells were harvested by centrifuging (10,000 g, 15 min) and the cells were lysed using BugBuster according to the manufacturer's protocol (Novagen). The TE, dTE, tTE and HTC proteins (all bearing a His-tag) were purified using nickel-affinity column chromatography with a modified version of the suppliers protocol (Novagen). TE, dTE, tTE and HTC were bound and washed with 0.1 M phosphate buffer (also containing 0.3 M NaCl and 12.5 mM Imidazole at pH 8.0). Only for the non-covalent variants, the nickel immobilized TE, dTE and tTE were then mixed for one hour with the CK to form the TECK, dTECK and tTECK

complexes, respectively. TECK, dTECK, tTECK and HTC were then washed with 10 column volumes of 0.1 M phosphate buffer (also containing 1.5 M NaCl and 12.5 mM imidazole at pH 8.0) and eluted with 0.1 M phosphate buffer (also containing 1.5 M NaCl and 0.25 M imidazole at pH 8.0). TECK, dTECK, tTECK and HTC were stored in 50 mM Tris, 0.5 M NaCl, 10 mM MgCl₂, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.5.

All formed complexes, namely TECK, dTECK, tTECK and HTC, were further purified by fast protein liquid chromatography (FPLC) using a Superdex 75 column (GE healthcare) coupled to an ÄKTA purifier (1 mL injection volume and as eluent the same buffer was used as the complexes were stored in), to remove non-complexed fluorescent proteins.

3.5.5 *Capsid assembly*

The growth and purification of wild type CCMV virus (wtCP) from *Vigna unguiculata* leaves was performed as reported previously.[31] To reassemble the CCMV capsid, the isolated wtCPs were mixed with the recombinant protein complexes (TECK, dTECK, tTECK and HTC) in various ratios in a buffer containing 50 mM Tris-Cl, 500 mM NaCl and 1 mM EDTA (pH 7.5) For all capsid reassembly mixtures, a final concentration of 75 μM CCMV protein was maintained (combined wtCP and recombinant CK proteins). The mixtures were dialyzed overnight to a buffer of 50 mM sodium acetate, 1 M NaCl and, 1 mM NaN₃ (pH 5.0) to induce capsid formation. The assembled capsids were purified and analyzed by FPLC using a Superose 6 column.

3.5.6 *Polyacrylamide gel electrophoresis*

Poly Acrylamide Gel Electrophoresis (PAGE) was performed both in the presence of sodium dodecylsulfate (SDS) and under native conditions in 12% gels. For SDS-PAGE, samples were heated to 99°C for 5 min in the presence of 2-mercaptoethanol and 1% SDS.

3.5.7 *UV-visible spectroscopic analysis*

All UV-visible absorbance measurements were performed on a Perkin Elmer Lambda 850 spectrometer. Standard quartz cuvettes with a 1 cm path length were used.

Absorption curves were measured by creating two dilution series (one at pH 7.5 and one at pH 5) from a mTFP stock. The stock concentration was determined by

UV-vis, based the literature extinction coefficient. The dilution steps were equal in each series, to standardize the measurement points.

The mTFP loading of the CCMV VLPs is calculated as follows. The absorption at 280 nm (A^{280}) is generated by the absorption of the capsid proteins (CK and wtCP) and by the fluorescent proteins (FP). The absorption at 462 nm (A^{462}) is generated solely by the fluorescent proteins. Lambert–Beer’s law states that absorption, over a path length of 1 cm, is equal to the extinction coefficient (ϵ) times the concentration (c). By combining Equation 3-3 and Equation 3-4 and rearranging the terms, we can calculate the number of fluorescent proteins per VLP (Equation 3-9):

$$\frac{A^{280}}{A^{462}} = \frac{A_{CK+CP}^{280} + A_{FP}^{280}}{A_{FP}^{462}}$$

Equation 3-3

$$A = \epsilon \times C$$

Equation 3-4

$$\frac{A^{280}}{A^{462}} = \frac{\epsilon_{CK+CP}^{280} \times C_{CK+CP} + \epsilon_{FP}^{280} \times C_{FP}}{\epsilon_{FP}^{462} \times C_{FP}}$$

Equation 3-5

$$\frac{A^{280}}{A^{462}} \times \epsilon_{FP}^{462} \times C_{FP} = \epsilon_{CK+CP}^{280} \times C_{CK+CP} + \epsilon_{FP}^{280} \times C_{FP}$$

Equation 3-6

$$\frac{A^{280}}{A^{462}} \times \epsilon_{FP}^{462} - \epsilon_{FP}^{280} = \frac{\epsilon_{CK+CP}^{280} \times C_{CK+CP}}{C_{FP}}$$

Equation 3-7

$$\frac{\frac{A^{280}}{A^{462}} \times \epsilon_{FP}^{462} - \epsilon_{FP}^{280}}{\epsilon_{CK+CP}^{280}} = \frac{C_{CK+CP}}{C_{FP}}$$

Equation 3-8

$$\text{number of fluorescent proteins per VLP} = \left(\frac{C_{CK+CP}}{C_{FP}} \right) // 180$$

Equation 3-9

Practically, this yields to:

$$C_{CK+CP} = \frac{A^{280} - (\epsilon_{FP}^{280} \times C_{FP})}{\epsilon_{CK+CP}^{280}}$$

Equation 3-10

$$C_{FP} = \frac{A^{462}}{\epsilon_{FP}^{462}}$$

Equation 3-11

3.5.8 Steady-state fluorescence spectroscopy

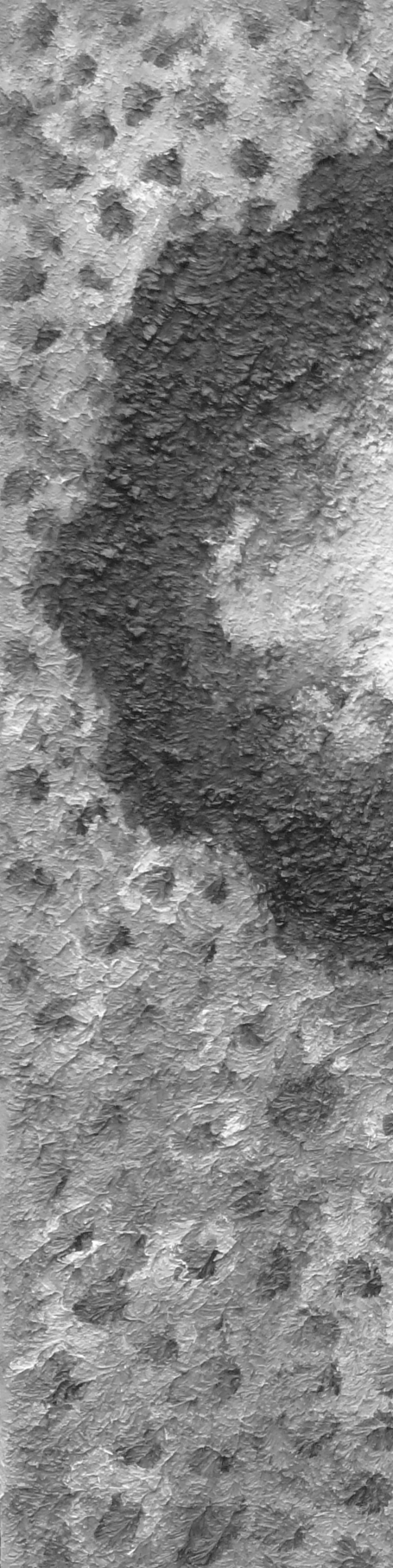
All Steady-state fluorescence spectroscopy measurements were performed on a Perkin Elmer LS55 fluorescence spectrometer. Standard quartz cuvettes with a 1 cm path length were used.

3.6 References

1. Soto, C.M. and B.R. Ratna, *Virus hybrids as nanomaterials for biotechnology*. *Curr Opin Biotechnol*, 2010. **21**(4): p. 426-38.
2. Fischlechner, M. and E. Donath, *Viruses as building blocks for materials and devices*. *Angew Chem Int Ed Engl*, 2007. **46**(18): p. 3184-93.
3. Baumler, H. and R. Georgieva, *Coupled enzyme reactions in multicompartiment microparticles*. *Biomacromolecules*, 2010. **11**(6): p. 1480-7.
4. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids*. *Journal of the American Chemical Society*, 2009. **131**(49): p. 17771-17773.
5. van Dongen, S.F., et al., *A three-enzyme cascade reaction through positional assembly of enzymes in a polymersome nanoreactor*. *Chemistry*, 2009. **15**(5): p. 1107-14.
6. Lavelle, L., J.P. Michel, and M. Gingery, *The disassembly, reassembly and stability of CCMV protein capsids*. *J Virol Methods*, 2007. **146**(1-2): p. 311-6.
7. Lavelle, L., et al., *Phase Diagram of Self-assembled Viral Capsid Protein Polymorphs (dagger)*. *J Phys Chem B*, 2009.
8. O'Neil, A., et al., *Coconfinement of Fluorescent Proteins: Spatially Enforced Communication of GFP and mCherry Encapsulated within the P22 Capsid*. *Biomacromolecules*, 2012. **13**(12): p. 3902-3907.
9. Mullaney, J.M., et al., *Green fluorescent protein as a probe of rotational mobility within bacteriophage T4*. *J Virol Methods*, 2000. **88**(1): p. 35-40.
10. *Viralzone*. Available from: www.expasy.ch/viralzone
11. Speir, J.A., et al., *Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy*. *Structure*, 1995. **3**(1): p. 63-78.
12. Caspar, D.L.D. and A. Klug, *Physical Principles in the Construction of Regular Viruses*. *Cold Spring Harbor Symposia on Quantitative Biology*, 1962. **27**: p. 1-24.
13. Speir, J.A., et al., *Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics*. *J Virol*, 2006. **80**(7): p. 3582-91.

14. Ali, A. and M.J. Roossinck, *Rapid and efficient purification of Cowpea chlorotic mottle virus by sucrose cushion ultracentrifugation*. J Virol Methods, 2007. **141**(1): p. 84-6.
15. Liu, H., et al., *Pseudo-atomic models of swollen CCMV from cryo-electron microscopy data*. J Struct Biol, 2003. **142**(3): p. 356-63.
16. Tama, F. and C.L. Brooks, 3rd, *The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus*. J Mol Biol, 2002. **318**(3): p. 733-47.
17. de la Escosura, A., et al., *Viral capsids as templates for the production of monodisperse Prussian blue nanoparticles*. Chem Commun (Camb), 2008(13): p. 1542-4.
18. Sikkema, F.D., et al., *Monodisperse polymer-virus hybrid nanoparticles*. Org Biomol Chem, 2007. **5**(1): p. 54-7.
19. Minten, I.J., et al., *Catalytic capsids: the art of confinement*. Chemical Science, 2011. **2**(2): p. 358-362.
20. Apostolovic, B. and H.A. Klok, *pH-sensitivity of the E3/K3 heterodimeric coiled coil*. Biomacromolecules, 2008. **9**(11): p. 3173-80.
21. Minten, I.J., R.J.M. Nolte, and J.J.L.M. Cornelissen, *Complex assembly behavior during the encapsulation of green fluorescent protein analogs in virus derived protein capsules*. Macromol Biosci, 2010. **10**(5): p. 539-45.
22. Shaner, N.C., G.H. Patterson, and M.W. Davidson, *Advances in fluorescent protein technology*. J Cell Sci, 2007. **120**(Pt 24): p. 4247-60.
23. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins*. Nat Methods, 2005. **2**(12): p. 905-9.
24. Ai, H.W., et al., *Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging*. Biochem J, 2006. **400**(3): p. 531-40.
25. Sun, Y., et al., *FRET microscopy in 2010: the legacy of Theodor Forster on the 100th anniversary of his birth*. Chemphyschem, 2011. **12**(3): p. 462-74.
26. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. 3rd ed. 2006, New York: Springer. xxvi, 954 p.
27. Shaner, N.C., et al., *Improving the photostability of bright monomeric orange and red fluorescent proteins*. Nat Methods, 2008. **5**(6): p. 545-51.

28. Shaner, N.C., et al., *Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein*. *Nat Biotechnol*, 2004. **22**(12): p. 1567-72.
29. Baird, G.S., D.A. Zacharias, and R.Y. Tsien, *Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral*. *Proc Natl Acad Sci U S A*, 2000. **97**(22): p. 11984-9.
30. Gross, L.A., et al., *The structure of the chromophore within DsRed, a red fluorescent protein from coral*. *Proc Natl Acad Sci U S A*, 2000. **97**(22): p. 11990-5.
31. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor*. *Nat Nanotechnol*, 2007. **2**(10): p. 635-9.



Chapter 4

Controlled loading of TFP fusion protein in CCMV virus-like particles

Based on the findings from the previous chapter, this chapter provides a detailed study towards the controlled loading of fluorescent proteins into CCMV virus-like particles (VLPs). The loading control and efficiency is then used to predict the loading of cargo into CCMV VLPs, demonstrating the accuracy of VLP loading. Fluorescence anisotropy is employed to study the effects of molecular confinement for the differently loaded CCMV VLPs. The effects of confinement and molecular crowding on the fluorescence energy transfer are explained by use of a mathematical model.

Part of this chapter is based on W. Frederik Rurup, Fabian C. Verbij, Melissa S. T. Koay, Christian Blum, Vinod Subramaniam, Jeroen J. L. M. Cornelissen “Controlled loading of virus-like particles” manuscript in preparation

4.1 Introduction

Despite the benefits of the E–K coiled–coil design in guiding fluorescent proteins into CCMV, the E–coils have a strong tendency to form homotrimers at pH 5.0, which significantly reduces the loading efficiency during capsid assembly. As described in the previous chapter, even increasing the intrinsic ratio of TFP cargo (TE) to capsid protein (CK) by promoting cargo dimerization was insufficient to overcome the problems associated with E–coil homotrimer formation. To circumvent this problem, a more stable ‘cargo–capsid’ complex variant was designed in Chapter 3 by engineering a flexible linker whereby the TFP cargo is genetically engineered as a fusion protein to the CCMV capsid protein (HTC). By abolishing the influence of dissociating E–K coiled–coils, it was observed that the loading efficiency of HTC in CCMV VLPs could be significantly improved. Unlike the TECK variants which show maximum loading efficiency with a cargo:CCMV ratio of 30%, the loading of HTC in CCMV VLPs remained linear up to 40%. The linearity confirmed that, not only does the fusion protein avoid problems due to dissociation, but also shows that the cargo can be directed inside CCMV during capsid assembly. In this chapter it is investigated whether the HTC fusion protein can be used for *controlled* loading and how the degree of loading in the virus-like particles (VLP) influences the cargo and the VLP.

4.2 Results and discussion

4.2.1 *VLP loading through capsid–TFP fusion protein*

To investigate the loading capacity and behavior, the HTC loading range was varied from 5% to 100% (and the range of native CCMV capsid protein was varied accordingly from 95% to 0%). Upon increasing the HTC ratio from 5% up to 40%, an almost linear increase in the loading of fluorescent proteins was observed (Figure 4.1, black). Further increasing the HTC concentrations from 40% up to 90% showed almost the same linear trend however, at the same time, there is a dramatic decrease in virus-like particle (VLP) formation (Figure 4.1, grey). At concentrations above 40%, the loading behavior and the formation of CCMV assemblies were much more erratic and resulted in very broad loading distributions. This suggests that CCMV assembly becomes less favorable in the presence of larger amounts of the HTC fusion protein, and is most likely due to steric hindrance of the TFP perturbing CCMV VLP assembly. In the presence of 100% HTC, VLP assembly induced aggregation and precipitation (at pH 5.0), such that no assembled VLPs could be detected by FPLC.

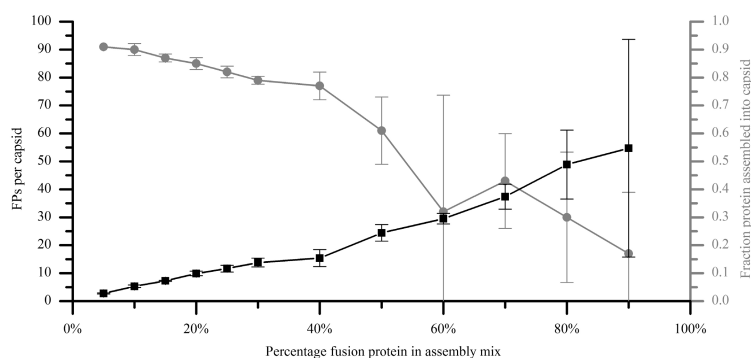


Figure 4.1: Plot of CCMV VLP loading efficiency as a function of the percentage fusion protein (HTC) (black). Plot of the relative fraction of CCMV VLP formation per total amount of protein (wild type CCMV capsid proteins and HTC fusion protein) (grey).

4.2.2 *Directed loading*

To test the controlled loading, the linear HTC loading behavior (from 0% to 40%) was used as a calibration curve to pre-calculate and predict a desired number of TFPs to be encapsulated. Loading of CCMV VLPs with an average of 2, 5, 7, 10, 15 or 20 fluorescent proteins was predicted by calculating the appropriate mixing ratios of HTC with wild type capsid protein (wtCP). In all cases, assembly of the CCMV VLPs at pH 5.0 led to excellent agreement between the predicted number and the obtained number of TFPs encapsulated (Figure 4.2, black). However, it should be noted that the assembly efficiency decreased slightly upon a larger amount of HTC used (Figure 4.2, grey).

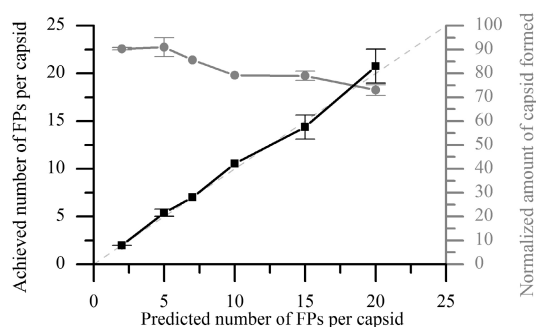


Figure 4.2: The actual CCMV VLP loading as a function of the predicted loading of CCMV with the HTC fusion protein (black). The capsid formation is the amount of protein in the FPLC elution peak corresponding to capsid sized particles normalized to the total amount of protein in each sample (grey).

Overall, although the accuracy of prediction decreased slightly with increasing concentrations, the fusion protein loading system enables accurate directed loading. This confirms, not only that *in vitro* HTC loading is reproducible and reliable, but also that the loading has a high degree of control and can be pre-calculated.

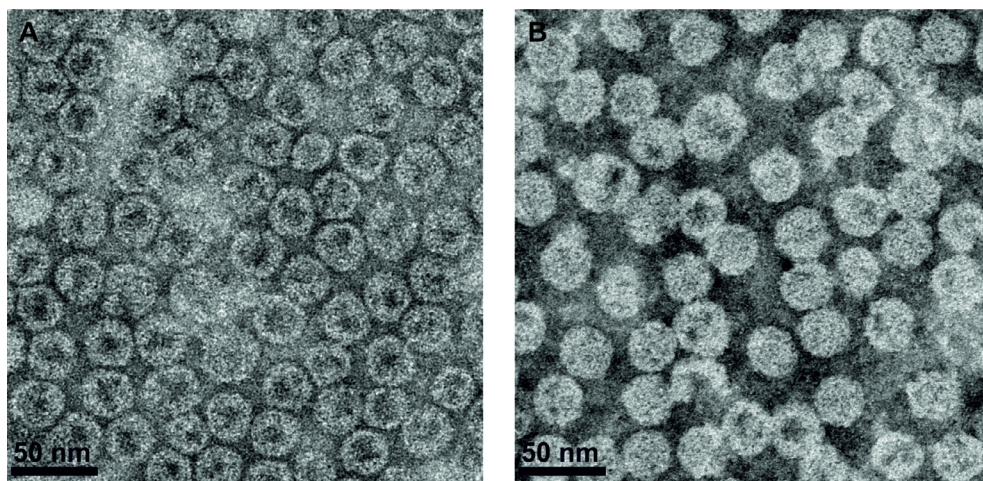


Figure 4.3: TEM images of CCMV loaded with (A) 2 TFPs per capsid and (B) 20 TFPs per capsid. In both cases, the CCMV VLPs showed very well defined and highly monodisperse assemblies, with an average diameter of 27 nm is within the expected size for $T = 3$ icosahedral viruses. These samples were negatively stained with 1% (w/v) uranyl acetate, which can be seen inside the capsids with lower loading (A), whereas the highly loaded capsids show little staining inside the capsids (B).

Transmission Electron Microscopy (TEM) of CCMV VLPs loaded with 2 TFPs confirmed the formation of highly monodisperse particles with an average diameter of ~ 27 nm, which is consistent with native $T = 3$ particles (Figure 4.3A). These particles appear to be mainly hollow, as suggested by the uranyl staining inside the VLPs. However, although the CCMV VLPs filled with 20 TFPs showed mostly intact $T = 3$ assemblies, some irregularities in size and shape are also observed (average diameter 32–38 nm). The presence of partially broken assemblies and the larger diameter of intact assemblies may suggest that some VLPs are partially swollen due to their high cargo loading (Figure 4.3B).

4.2.3 *Fluorescence anisotropy*

To obtain a better understanding of the molecular crowding effects and the rotational freedom of the TFPs localized inside the CCMV capsids, steady-state fluorescence anisotropy measurements were performed. Fluorescence anisotropy

provides information about the fluorophore dynamics by comparing the polarization of the emitted light with the polarization of the excitation light.[1]

The excitation of a fluorophore is highly dependent on the orientation of the fluorophore with respect to the polarization of the excitation light. The probability of excitation is highest when the transition moment of the fluorophore is parallel to the polarization plane of incident light. Depending on the rotational speed of the fluorophore relative to its fluorescence lifetime, the polarization of the emitted light can either be equal to its excitation orientation, when the fluorophore moves very slowly, or completely randomized, if the fluorophore rotates very fast. The molecular motion depends both on molecule size and by the freedom of movement of the molecule. A higher molecular motion influences the fluorescent emission polarization, which is quantified by the fluorescence anisotropy (r). The anisotropy value is influenced by the viscosity of the surrounding solution, the size of the fluorophore and its surroundings (confinement).

The anisotropy value r is measured and calculated according to Equation 4-1, where I_{AB} is the detected intensity with the excitation (A) and emission (B) polarizers either horizontal (H) or vertical (V) (Figure 4.4).

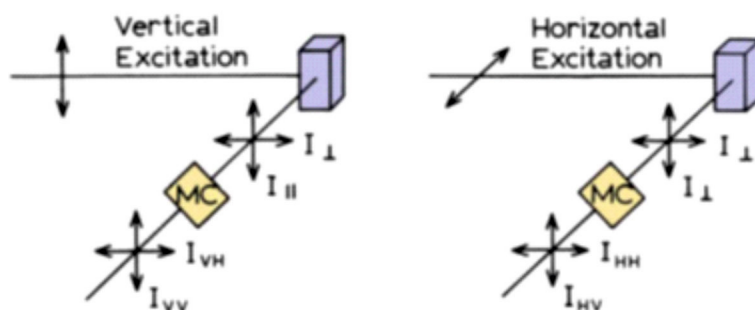


Figure 4.4: Schematic diagram showing the setup of steady-state fluorescence anisotropy measurements (MC, monochromators). (Adopted from [1])

$$r = \frac{I_{VV} - G \cdot I_{VH}}{I_{VV} + 2G \cdot I_{VH}}$$

Equation 4-1

The factor G is the ratio of sensitivity of the detection system for vertically and horizontally polarized light and is given by Equation 4-2.[1]

$$G = \frac{I_{HV}}{I_{HH}}$$

Equation 4-2

The theoretical value of r lies between 0 (fully randomized, highly mobile fluorophore) and 0.4 (highly polarized, almost no molecular movement), due to the fact that absorption and emission of photons is described by a probability function, resulting in an excited-state distribution. This probability of absorption creates an inherent randomization in the detection.

Molecular movement is normally the main cause of changes in anisotropy, however when the fluorophores are crowded close together, another effect (homo-FRET) can influence the anisotropy. Homo-FRET is Förster resonance energy transfer, which occurs only at very small (<10 nm) inter-fluorophore distances between two equal fluorophores (Chapter 3). Because the acceptor fluorescent protein may have a different orientation than the donor protein, the occurrence of homo-FRET can change the polarization of the fluorescent emission.[1] This change is measured as a decrease in anisotropy, since randomization is increased (Figure 4.5). Since fluorescence anisotropy depends on molecular motion and energy transfer processes, it can provide useful information about the TFP (cargo) encapsulation as well as the crowding of TFP inside the CCMV VLPs.

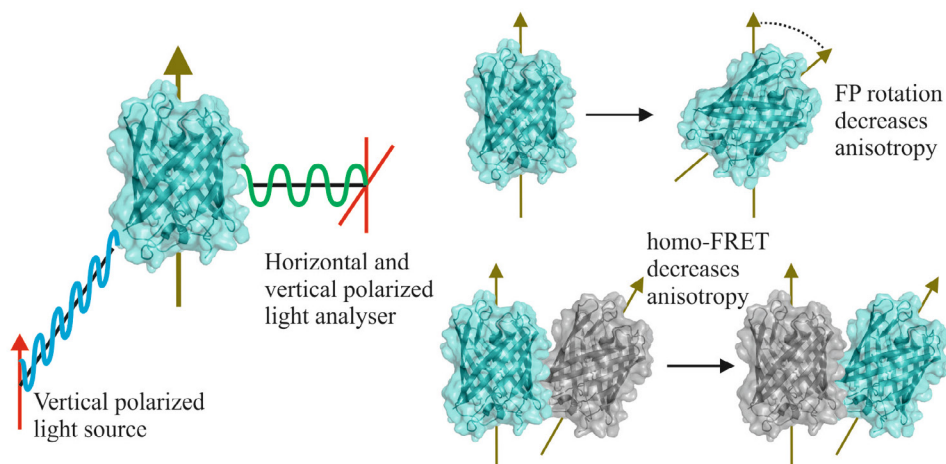


Figure 4.5: Schematic overview of the principles of fluorescence anisotropy and the phenomena that influence the fluorescence emission.

Since different configurations of TFP were available (mTFP, TE, dTE, TECK, dTECK and HTC), all samples were measured in order to obtain a complete interpretation of the anisotropy values. As expected, a higher anisotropy was observed for TE than for mTFP, indicating that the flexible α -helical linker (E-coil) was sufficient to significantly lower the molecular movement of the TFP fluorophore. Similarly, a higher anisotropy was measured for the dimeric TFP E-coil (dTE) than for the TE, suggesting that the restricted rotational freedom due to

dimerization of two TE monomers has a more profound effect than the homo-FRET between the two TFPs in close proximity (Table 4.1). Surprisingly, the fluorescence anisotropy for the TECK and dTECK complexes did not show significant changes compared to the free TE and dTE forms. Instead it was proposed that the increase in fluorescence anisotropy (due to increased restricted rotational freedom) was nullified by the decrease in anisotropy (due to homo-FRET).

Table 4.1: Fluorescence anisotropy values for the different TFP loading components used in this work. The values for free fluorescent proteins, the TFPs complexed the capsid protein and the encapsulated (1–3) TFPs are provided. All values are an average over 490–520 nm with the appropriate standard deviation, for sample concentrations in the order of 1 μ M.

	Sample	Anisotropy (490–520 nm)	St. dev.
Free FPs	mTFP	0.310	0.011
	TE	0.321	0.002
	dTE	0.330	0.002
Complexed FPs	TECK	0.325	0.003
	dTECK	0.325	0.002
	HTC	0.330	0.002
Encapsulated FPs (1–3 FPs per capsid)	TECK encapsulated	0.362	0.006
	dTECK encapsulated	0.343	0.006
	HTC encapsulated	0.351	0.003

Upon encapsulation of low amounts of fluorescent protein (1–3 TFPs) into CCMV VLPs, a distinct increase in anisotropy for all complexes was observed. This increase is consistent with the successful incorporation of TFPs into the CCMV capsid, which gives rise to a decrease in rotational movement. The size of the assembled capsid restricts the movement of the fluorescent proteins, suggesting that the fluorescent proteins may still be attached to the capsid wall. Homo-FRET plays a minor role at these low loading levels. Additional fluorescence anisotropy measurements were performed using the HTC variant, since this material showed the most efficient, predictable and controllable loading. As observed for the TECK complexes, upon increasing the amount of encapsulated HTC, a decrease in the anisotropy value was observed (Figure 4.6). This is likely due to an increase in homo-FRET caused by molecular crowding of the fluorophores within the capsid interior, which has been reported in the past.[2, 3] The anisotropy approaches steady-state for CCMV VLPs with an average loading of 10 fluorescent proteins,

suggesting that the influence of homo-FRET on the anisotropy has already reached its maximum.

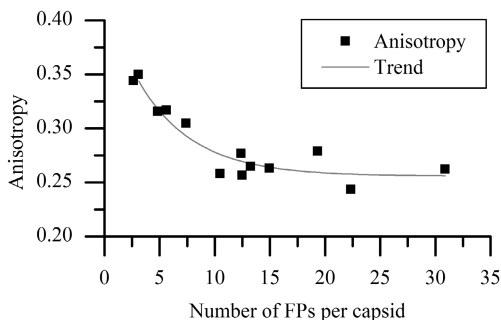


Figure 4.6: Fluorescence anisotropy measurements of encapsulated HTC in CCMV. The grey line indicates the decreasing trend with greater numbers of encapsulated fluorescent proteins.

The efficiency of homo-FRET is strongly dependent on the interfluorophore distance and the fluorophore specific Förster radius (R_0) [1]. Although TFP has no reported Förster radius for homo-FRET, based on interactions with other fluorescent proteins, an $R_0 \sim 5.5$ nm is estimated.[4] From an interfluorophore distance of $1.5R_0$, homo-FRET is known to become significantly more efficient upon further decrease of the interfluorophore distance. For TFP, this efficiency limit ($1.5R_0$) is calculated to be 8.2 nm.

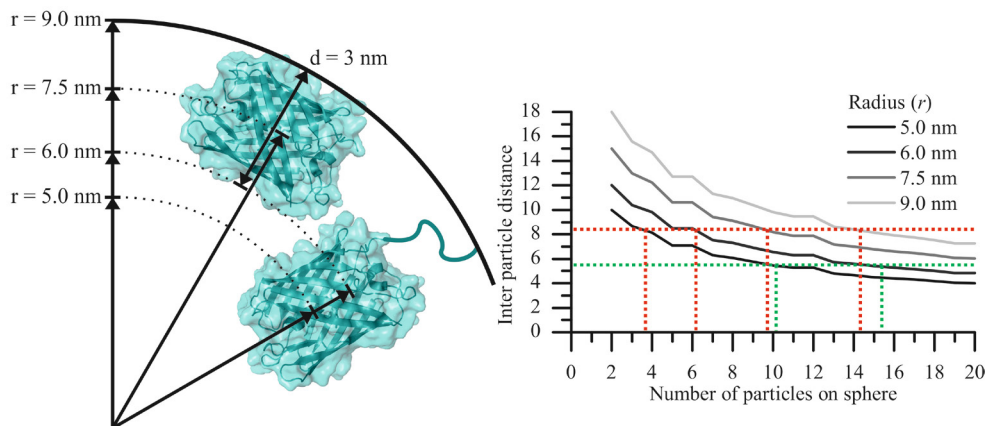


Figure 4.7: Schematic representation of TFP location and orientation inside the 9 nm radius CCMV internal cavity (left). The solutions of the ‘Tammes’ problem for different sphere diameters, the red dotted line illustrates where $1.5R_0$ is reached, the green line where R_0 is reached.

CCMV of $T = 3$ icosahedral symmetry has an average internal diameter of 18 nm. Assuming a perfect sphere for the CCMV interior, the maximum distance in between the fluorophores can be approximated by a ‘Tammes problem’.[5] This mathematical model calculates the maximum possible distance between a number of points on a sphere. Considering the size of TFP (3 nm in diameter, 4 nm in height, [6]) with the fluorophore located in the middle of the protein barrel, the fluorophores can be distributed over an effective capsid shell with a radius (r) of 7.5 nm (Figure 4.7, left). Based on the standard Tammes curves (Figure 4.7, right) that show this maximum possible ‘inter particle distance’ for a certain number of particles, the minimal interfluorophore distance for efficient homo-FRET ($1.5 R_0$) is reached when the 7.5 nm shell is filled with 10 TFPs (shown by the red dotted line). This means that for 10 fluorescent proteins in any position on the shell, the distance between the fluorophores is *maximally* $1.5 R_0$, however, this is most likely to be less.

It should be noted that the 7.5 nm virtual shell is based on the assumption that the TFPs are all aligned flat against CCMV VLP wall. However, the TFPs are more likely to be randomly oriented (*i.e.* non-parallel, tethered by the flexible linker) and a distribution radius of 5–6 nm is more likely to be the case as illustrated in Figure 4.7 (left). Based on a 5–6 nm radius, the Tammes curves now suggests that homo-FRET occurs (maximum ‘inter particle distance’ equals $1.5 R_0$) in the presence of only 4–6 TFPs, respectively (Figure 4.7, right, red dotted line).

Using the 5–6 nm virtual shell radius shows that the maximum ‘inter particle distance’ even reaches R_0 (Figure 4.7, green dotted line) for 10–15 FPs respectively. The interfluorophore distance of R_0 results in 50% FRET efficiency, regarding that it is the maximum interfluorophore distance, the efficiency likely is better. Interestingly, this corresponds nicely with the experimental data presented in Figure 4.6, in which the maximum FRET effect is observed upon loading of up to 10–15 TFPs.

4.2.4 Maximum theoretical fluorescent protein loading

Several approaches are possible for determining the maximum possible loading of a CCMV VLP. The simplest approach is made when fluidity of bodies is assumed. Here, the volume of a sphere was calculated for the CCMV capsid ($\frac{4}{3}\pi \cdot r^3$) assuming a radius of 9 nm. For the β -barrel shaped TFP protein, the volume of a cylinder was assumed ($\pi \cdot r^2 \cdot h$) based on a radius of 1.5 nm and a height of 4 nm. By calculating the relative volumes, 108 fluorescent proteins can occupy the inner

volume of the CCMV VLP. However, this approach assumes that the proteins are rigid bodies.

Therefore, a slightly better approach is to assume that all cylindrical shaped TFPs are oriented perpendicular to the inner CCMV VLP wall (Figure 4.8A). Assuming the full height of TFP occupies the “first packing sphere” of the CCMV shell, the remaining volume in the center with a radius of 5 nm is still available for another ‘shell’ of cylinder shaped TFPs (*i.e.* 2 virtual packing spheres). The most optimal packing of circles on a surface is hexagonal, yielding a 91% surface coverage. The surface of the virtual spheres is $4\pi \cdot r^2$ (with radii 5 nm and 1 nm) and the surface of the circular bottom of the cylinders is $\pi \cdot r^2$ (with a radius of 1.5 nm). This yields 46 fluorescent proteins per capsid. However, here the assumption of full and optimal hexagonal packing on the virtual spheres is questionable.

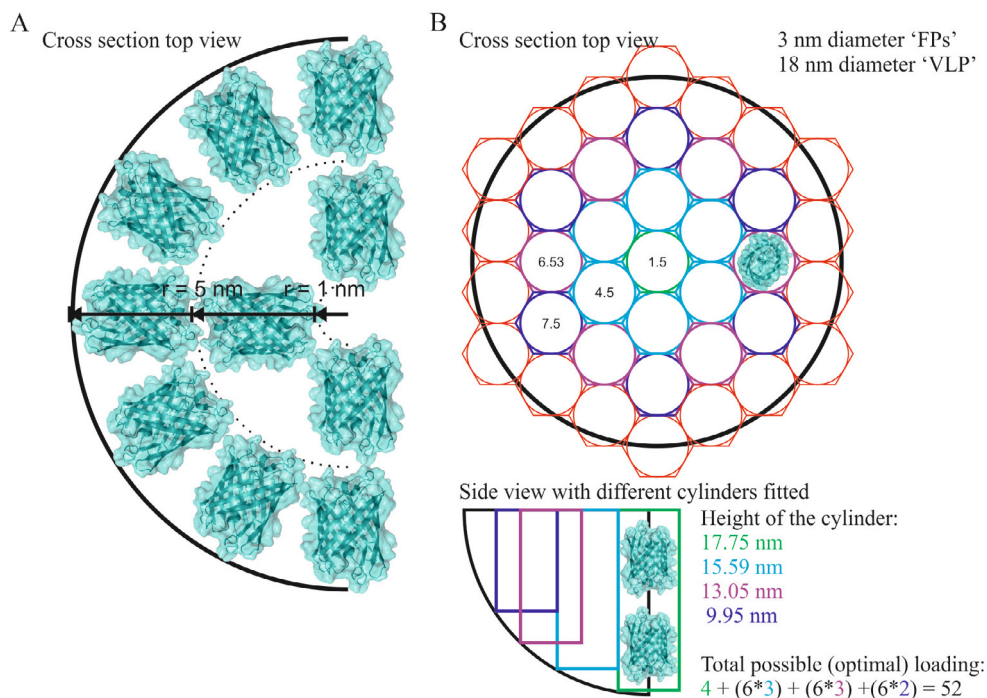


Figure 4.8: Schematic representation of the optimal packing of cylinders (TFPs) within a sphere (CCMV). (A) Orientation of the fluorescent proteins perpendicular to the shell. (B) The numbers in the circles are the distances of the fluorophores from the center. The distance corresponds to a maximum height of stacked cylinders, where only a discrete number can fit. This yields a maximum loading for this approach of 52 cylinders.

A third approach is to assume hexagonal packing in a random cross-section of the sphere. In this case, the optimal surface coverage can be reached and is only

dependent on the height of each cylinder and the number of TFPs (of 4 nm in height) can be arranged in the space volume (Figure 4.8B). Calculations for this approach show that a maximum of 52 TFPs can be packaged inside the CCMV VLP. However, one major drawback is that this model ignores the solvation sphere surrounding the proteins. When a quarter nanometer is added to the cylinder's dimensions in each direction (yielding a cylinder with a diameter of 3.5 nm and a height of 4.5 nm), to account for the solvation shell and repulsion, calculations show that only 27 TFPs can fit within the CCMV VLP sphere.

Based on the two approaches used, the maximum loading of CCMV VLPs is predicted to be significantly less than 52 and 46 TFPs and is even likely to be less than the lower limit of 27 TFPs. These calculations are consistent with the maximum loading observed for HTC (20 TFPs) and may also explain the erratic behavior observed for CCMV VLP formation at high HTC concentrations (Figure 4.1).

4.3 Conclusion

Here, the HTC fusion protein was used to reliably predict and control the loading of up to 20 fluorescent proteins inside CCMV VLPs. Transmission electron microscopy of the formed CCMV assemblies confirmed their morphology and size as $T = 3$ particles. However, higher loading (20 TFPs) appears to lead to slightly swollen CCMV VLP assemblies and distorts the VLPs that are formed. Investigation of the encapsulation effects on crowding and the rotational freedom of the fluorescent proteins through fluorescence anisotropy measurements confirmed that the fluorescent proteins are encapsulated in the CCMV VLPs and that homo-FRET occurs directly upon multiple TFP loading within the confined volume inside CCMV. The homo-FRET efficiency reaches its maximum for 10–15 TFPs per CCMV VLP, suggesting that the interfluorophore distance is significantly less than the Förster radius of the fluorescent protein. Nevertheless, the observed homo-FRET and the increased fluorophore crowding upon higher TFP loading, correlates very well with the mathematically Tammes model. Overall, based on different space-filling models combined with the observed data, it is proposed that a maximum of ~20 TFPs can be loaded in CCMV VLPs. This indicates that the measurements done on P22 [7] and T4 [2] are very difficult to compare to this work. The reported loading is at least 10-fold higher compared to the present studies. Nevertheless it shows that confinement effects, such as (homo-)FRET, occur readily when molecules are encapsulated in VLPs.

4.4 Acknowledgements

F.C. Verbij is gratefully acknowledged for his help performing experiments in this chapter. Dr. C Blum is gratefully acknowledged for his help and fruitful discussions.

4.5 Experimental

4.5.1 *Recombinant protein expression and purification*

All proteins were expressed and purified as described in Chapter 3. The amino acid sequence of HTC is shown in Figure 4.9.

MGSSHHHHH	SSGLVPRGSH	MVSKGEETM	GVIKPMKIK	LKMEGNVNGH	AFVIEGEGEG
KPYDGTNTIN	LEVKEGAPLP	FSYDILTTF	AYGNRAFTKY	PDDIPNYFKQ	SFPEGYSWER
TMTFEDKGIV	KVKSDISMEE	DSFIYEIHLK	GENFPPNGPV	MQKKTGWDA	STERMYVRDG
VLKGDVKHKL	LLEGGGHRV	DFKTIYRAKK	AVKLPDYHFV	DHRIEILNHD	KDYNKVTVYE
SAVARNSTDG	MDELYTGGSG	LVPRGSGSGG	LEMSTVGTGK	LTRAQRRAAA	RKNKRNRVV
QPVIVEPIAS	GQGKAIKAWT	GYSVSKWTAS	CAAAEAKVTS	AITISLPNEL	SSERNKQLKV
GRVLLWLGLL	PSVSGTVKSC	VTETQTAAA	SFQVALAVAD	NSKDVAAMY	PEAFKGITLE
QLAADLTIYL	YSSAALTEGD	VIVHLEVEHV	RPTFDDSFTP	VY*	

Figure 4.9: The amino acid sequence of HTC. The section marked in cyan is the TFP, the section in grey is the CCMV capsid protein. The white sequence in between is the flexible linker (GGSG) containing a thrombin cleavage site (LVPRGS).

4.5.2 *Capsid assembly*

The description of wild type capsid protein production and purification as well as the capsid assembly protocol are described in Chapter 3.

4.5.3 *UV-vis analysis*

All UV-visible measurements were performed on a Perkin Elmer Lambda 850 Spectrometer. Standard quartz cuvettes with a 1 cm path length were used. The slit widths were optimized to achieve the best signal to noise ratio.

4.5.4 *Fluorescence anisotropy*

A Varian Cary Eclipse spectrophotometer was used to measure fluorescence anisotropy of all samples. Samples were measured in quartz cuvettes using 450 nm wavelength light excitation. The anisotropy was averaged over the range of 490 nm to 520 nm.

4.5.5 *Transmission electron microscopy (TEM)*

TEM micrographs were recorded on an analytical FEG-TEM (Philips CM 30) operated at 300 kV acceleration voltages. Samples were prepared by placing 5 μ L on Formvar carbon-coated copper grids (Electron Microscopy Sciences). The sample was left on the grid for 5 minutes, after which time the excess buffer was blotted away with filter paper. Samples were negatively stained by applying 5 μ L

Controlled loading of TFP fusion protein in CCMV virus-like particles

of stain (1% w/v uranyl acetate in MilliQ water) onto the grid and removing the excess stain away after 1 minute with filter paper. The samples were dried overnight before imaging.

4.6 References

1. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. 3rd ed. 2006, New York: Springer. xxvi, 954 p.
2. Mullaney, J.M., et al., *Green fluorescent protein as a probe of rotational mobility within bacteriophage T4*. *J Virol Methods*, 2000. **88**(1): p. 35-40.
3. O'Neil, A., et al., *Genetically programmed in vivo packaging of protein cargo and its controlled release from bacteriophage P22*. *Angew Chem Int Ed Engl*, 2011. **50**(32): p. 7425-8.
4. Sun, Y., et al., *FRET microscopy in 2010: the legacy of Theodor Forster on the 100th anniversary of his birth*. *Chemphyschem*, 2011. **12**(3): p. 462-74.
5. Erber, T. and G.M. Hockney, *Equilibrium-Configurations of N Equal Charges on a Sphere*. *Journal of Physics a-Mathematical and General*, 1991. **24**(23): p. L1369-L1377.
6. Based on the 2HQK PDB entry.
7. O'Neil, A., et al., *Coconfinement of Fluorescent Proteins: Spatially Enforced Communication of GFP and mCherry Encapsulated within the P22 Capsid*. *Biomacromolecules*, 2012. **13**(12): p. 3902-3907.



Chapter 5

Encapsulins: bacterial nanocages

a user-based protocol for expression, purification and characterization

A new class of protein nanocages is introduced: the bacterial encapsulin. The unique properties of two different encapsulins, those of *Brevibacterium linens* and *Thermotoga maritima*, are briefly introduced. This chapter is intended as a user-based protocol for the future design and engineering of these protein cages and therefore provides a detailed description of their recombinant expression, purification and characterization.

Part of this chapter is based on: W. Frederik Rurup, Melissa S. T. Koay, Jeroen J. L. M. Cornelissen, “Engineering Protein Cages as Nanocarriers” in the book series “Methods in Molecular Biology” manuscript in preparation

drug delivery systems and molecular probes. To date, much focus on protein-based cages has been directed towards the loading and redesign of virus-like assemblies [1-6] or, in some cases, *de novo* designed protein cages [7]. Yet there still exists a vast range of stable protein-based cages in nature, that are available for nanotechnology, which remain largely unexplored. For example, a decade passed between the discovery of certain proteins in *Brevibacterium linens* M18 [8] and *Thermotoga maritima* [9], now called ‘encapsulins’ [10] and the realization that their tertiary morphology was a nanocage. Although these nanocages have a substantial sequence homology (Figure 5.1A), the *T. maritima* encapsulin was initially thought to be a protease and the *B. linens* encapsulin was initially referred to as linocin-like protein, a bacteriostatic agent.

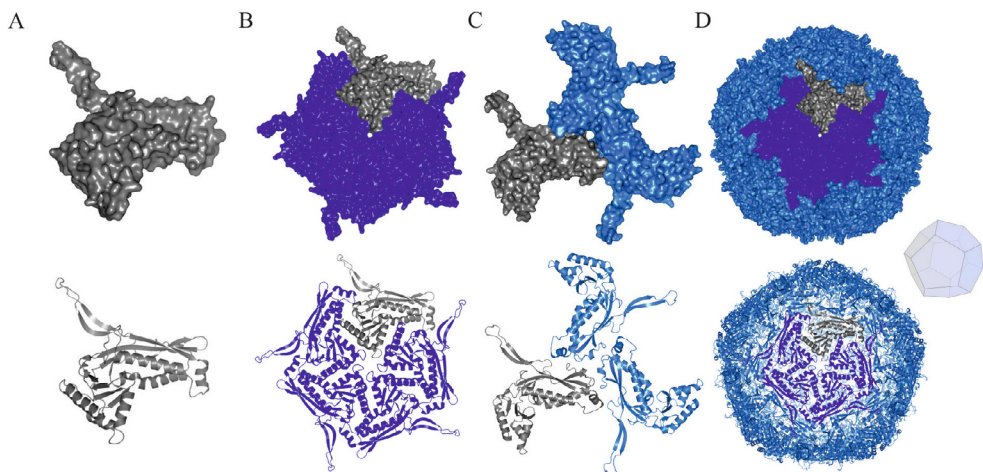


Figure 5.2: The *T. maritima* encapsulin structure showing (A) the monomer, (B) the pentamer, (C) the 3-fold axis where the pentamers connect and (D) the full *T. maritima* encapsulin, with the inset of a dodecahedron. The upper representations are the electron density surface mappings and the lower representations show the secondary protein structures. All images are rendered based on PDB entry 3DKT using PyMol.

The structural details of the *T. maritima* encapsulin were determined by Sutter *et al.* using X-ray crystallographic, biochemical and electron microscopic techniques [10, 11]. The X-ray crystal structure of the *T. maritima* encapsulin revealed that 60 copies of the monomer (Figure 5.1) self-assemble inside the bacteria into a 24 nm diameter protein cage with icosahedral T = 1 symmetry (*i.e.* a dodecahedron of pentamers, Figure 5.2). Each monomer has a molecular mass of 30.8 kDa (268 amino acids) and the formed compartment packages functional proteins within the 2.0–2.5 nm thick shell walls.

Interestingly, the interior of the *T. maritima* encapsulin shell is lined with highly conserved binding sites. These were initially observed in the X-ray structure as it showed an extra area of electron density. The density is localized in a cavity-like part of the encapsulin protein and it was matched to the C-terminus extension of the natural cargo of the *T. maritima* encapsulin (Figure 5.3), suggesting that this C-terminal extension functions as a loading tag for cargo in the *T. maritima* encapsulin.

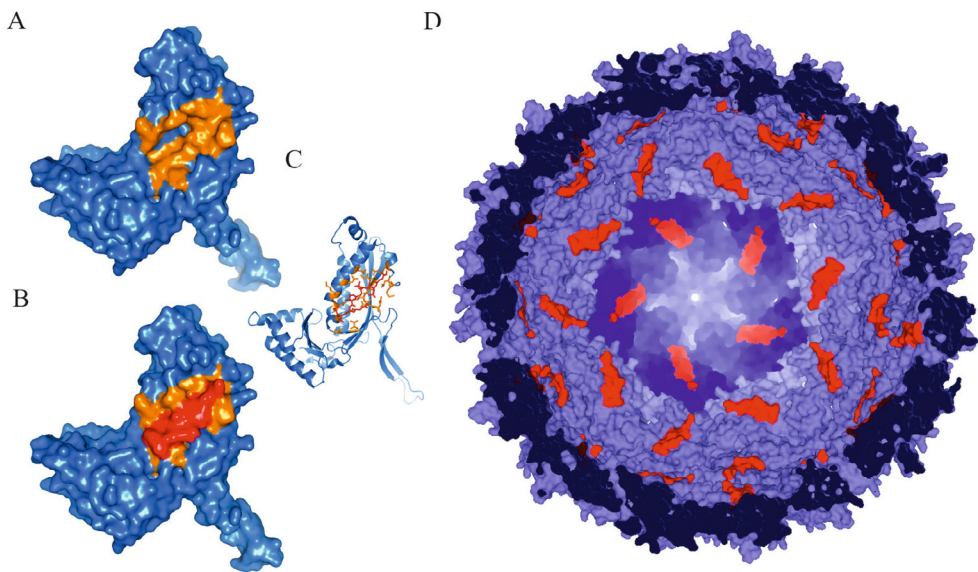


Figure 5.3: (A) The cavity-like binding area of the *T. maritima* encapsulin. (B) The binding area filled with the binding peptide, shown in (C) with all interacting atoms as a stick model. (D) All available binding sites in the *T. maritima* encapsulin. All images are rendered based on PDB entry 3DKT using PyMol.

Molecular modeling based on the sequence homology and electron microscopy data revealed a high similarity between *T. maritima* encapsulin (TM) and *Brevibacterium linens* M18 encapsulin (Blin) [11] and therefore an identical morphology was assumed. The encapsulins of *B. linens* is composed of 28.6 kDa (266 amino acids) monomers, which are arranged identical to the *T. maritima* encapsulin.

Loading of both encapsulins (TM and Blin) with their natural cargo is performed through the specific and highly conserved extension sequence, based on genomic studies (over 30 similar proteins, with practically identical extensions are reported by Sutter *et al.* [11]). Using the crystallographic data of the natural cargo proteins,

simulations show that docking of the C-terminal sequences is guided by the multimerization of the cargo and the similar symmetry of the protein shell subunit arrangement, *i.e.* 3-fold and 5-fold symmetry.

The modeling for the TM content, based on homologous proteins, indicates that ferritin-like proteins (Flp) form pentamers of dimers. The dimensions of the pentamers and thus the location where the C-terminal extension is presented, overlaps with the 5-fold symmetry axis of the TM encapsulin. The same occurs for dye-decolorizing peroxidase (DyP), the natural cargo of the Blin, which forms a trimer of dimers as shown by electron microscopy analysis. Here, the dimensions fit nicely to the 3-fold symmetry axis of the encapsulin (Figure 5.4).[11] This fit for both types of cargo suggests that next to the affinity tag, also multivalency is important for encapsulin loading.

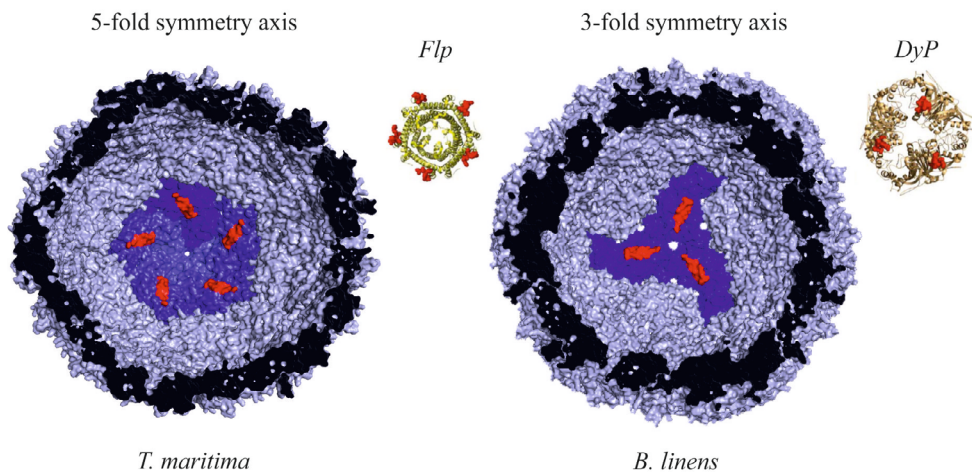


Figure 5.4: Symmetry and dimensions of the cargo targeting sequences (Flp and DyP) compared to the encapsulin shell protein symmetry (TM and Blin). Highlighted in dark blue are the proteins involved in the binding, highlighted in red are the binding sequences, both on the cargo and as they appear on the X-ray structure. Adopted from Sutter *et al.*[11].

Encapsulins, as a new group of protein cages are highly promising candidates for future applications in nanotechnology. Despite a slightly smaller outer size, the inner volumes of encapsulin and CCMV are very similar (Figure 5.5). Unlike viruses, which can undergo reversible pH-driven assembly/disassembly under *in vitro* conditions (CCMV, Chapter 3), encapsulins are highly pH/temperature stable, are assembled entirely *in vivo* and are *non-viral*. Furthermore, the cargo loading in encapsulins is very different from that of virus-based assemblies. Whilst virus capsid proteins contain a highly positively charged N-terminus for the

encapsulation of native genetic cargo, encapsulins contain natural peptide docking sites, which allow the native enzyme cargo to bind.

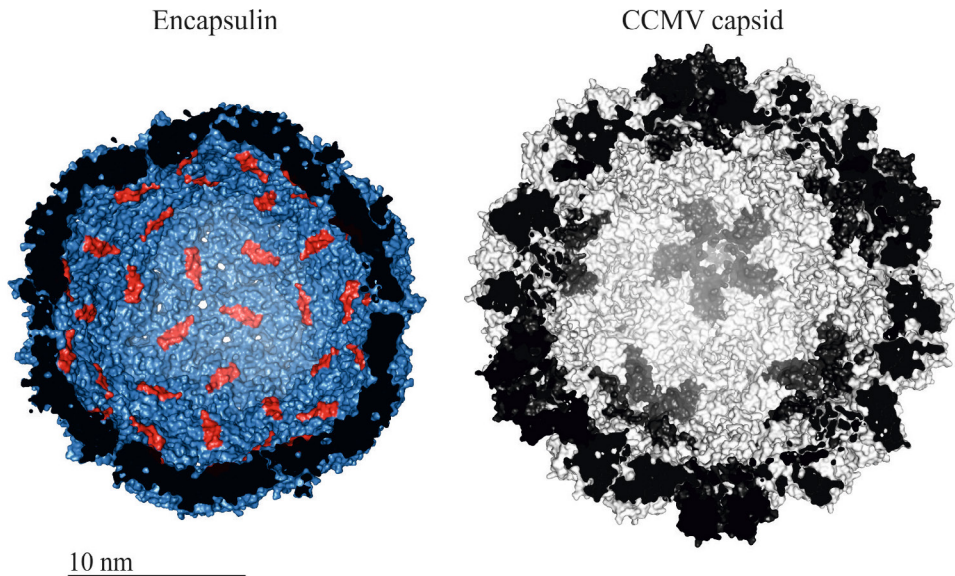


Figure 5.5: Size comparison between TM encapsulin and CCMV. Both images are rendered based on PDB entry 3DKT and 1CWP respectively using PyMol.

The aim of this work is to investigate the use of encapsulins as a new nanocompartment for potential applications in nanotechnology, by guiding and controlling the loading of these highly stable nanocages. In this chapter, the expression and purification of the encapsulins from *B. linens* and *T. maritima* are described.

5.2 Expression and sequences

The vectors and expression hosts were a kind gift from Dr. Markus Sutter and Prof. dr. Nenad Ban (ETH Zurich, Switzerland).

Six expression vectors were numbered accordingly:

#1	pET16b mod TM Enc	<i>T. maritima</i> encapsulin
#2	pET21a Blin	<i>B. linens</i> encapsulin
#3	pET21a Blin_AAA_His	<i>B. linens</i> with C-terminal 'AAALEHHHHHH'
#4	pET21a TM 5p Bac	<i>T. maritima</i> with ferritin-like protein
#5	pET21a Blin His_GFP_li Bac	<i>B. linens</i> encapsulin with GFP-binding sequence
#6	pET21a Blin Perox Enc	<i>B. linens</i> encapsulin with DyP

Vectors #4 and #6 contain the natural operon consisting of cargo *and* encapsulin, allowing for co-expression of both the cargo and the encapsulin under the same inducible promoter. In vector #5 the same operon is used. However, the natural cargo was exchanged for GFP with preservation of the C-terminal extension sequence necessary for binding.

DNA sequencing of all vectors confirmed the correct sequences of all plasmid constructs without mutations. The physical characteristics of the Blin and TM encapsulins were calculated and predicted based on the amino acid sequences using ApE ('a plasmid editor' freeware) and the ExPASy toolkit (<http://web.expasy.org/protparam/>) (Table 5.1).

Table 5.1: Properties of the expression products of the encapsulin vectors #1 – #6.

Vector	Encapsulin monomers		Cargo	
	Mass (kDa)	ϵ_{280} ($M^{-1} cm^{-1}$)	Mass (kDa)	ϵ_{280} ($M^{-1} cm^{-1}$)
#1	30.6	36565	-	-
#2	28.6	29910	-	-
#3	29.9	29910	-	-
#4	30.8	36565	13.3	20970
#5	28.6	29910	34.9	22015
#6	28.6	29910	39.6	20065

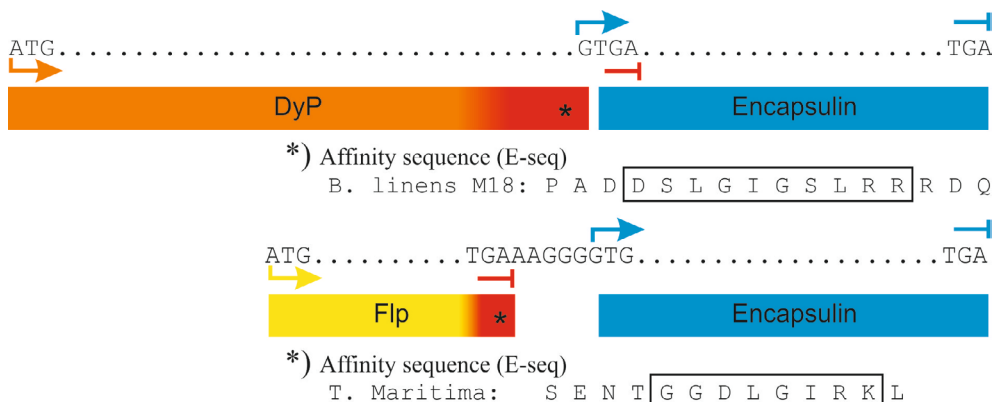


Figure 5.6: Overview of the co-expression operons of *B. linens* (top) and *T. maritima* (bottom) encapsulins. In both cases, the ATG start codon is shown for the cargo proteins DyP and Flp (orange and yellow arrow respectively) and the GTG start codon for the encapsulin is also shown (blue arrow). The amino acid sequence in the box is the docking sequence of respective encapsulin.

It should be noted that the masses for the encapsulin vary slightly when expressed alone or co-expressed with the cargo due to the different start codon used (GTG instead of ATG). The GTG start codon translates to a valine as the N-terminal amino acid residue (117 Da) instead of methionine (149 Da) in the case of the ATG start codon (Figure 5.6).

5.3 Expression and purification protocol

5.3.1 *Encapsulin expression*

The recombinant expression of both TM and Blin encapsulins was performed in *E. coli* Rosetta (DE3). Sterilized LB-medium (Sigma) supplemented with 100 mg/L ampicillin (Amp) and 34 mg/L chloramphenicol (Cam) was used to grow 5–7 mL starter cultures *B. linens* or *T. maritima* encapsulin overnight at 37 °C, shaking at 220 rpm.

After ~18 hours, the starter cultures were used to inoculate 0.5–1.0 L LB containing ampicillin and chloramphenicol. *E. coli* cells were grown at 37 °C with shaking at 150 rpm until the optical density at $\lambda=600$ nm (OD_{600}) reached 0.6–0.8. To improve protein expression, the cultures were cooled to 22 °C while shaking before adding isopropyl- β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM to induce the expression of the encapsulin. The protein expression was left overnight at 22 °C while shaking at 150 rpm.

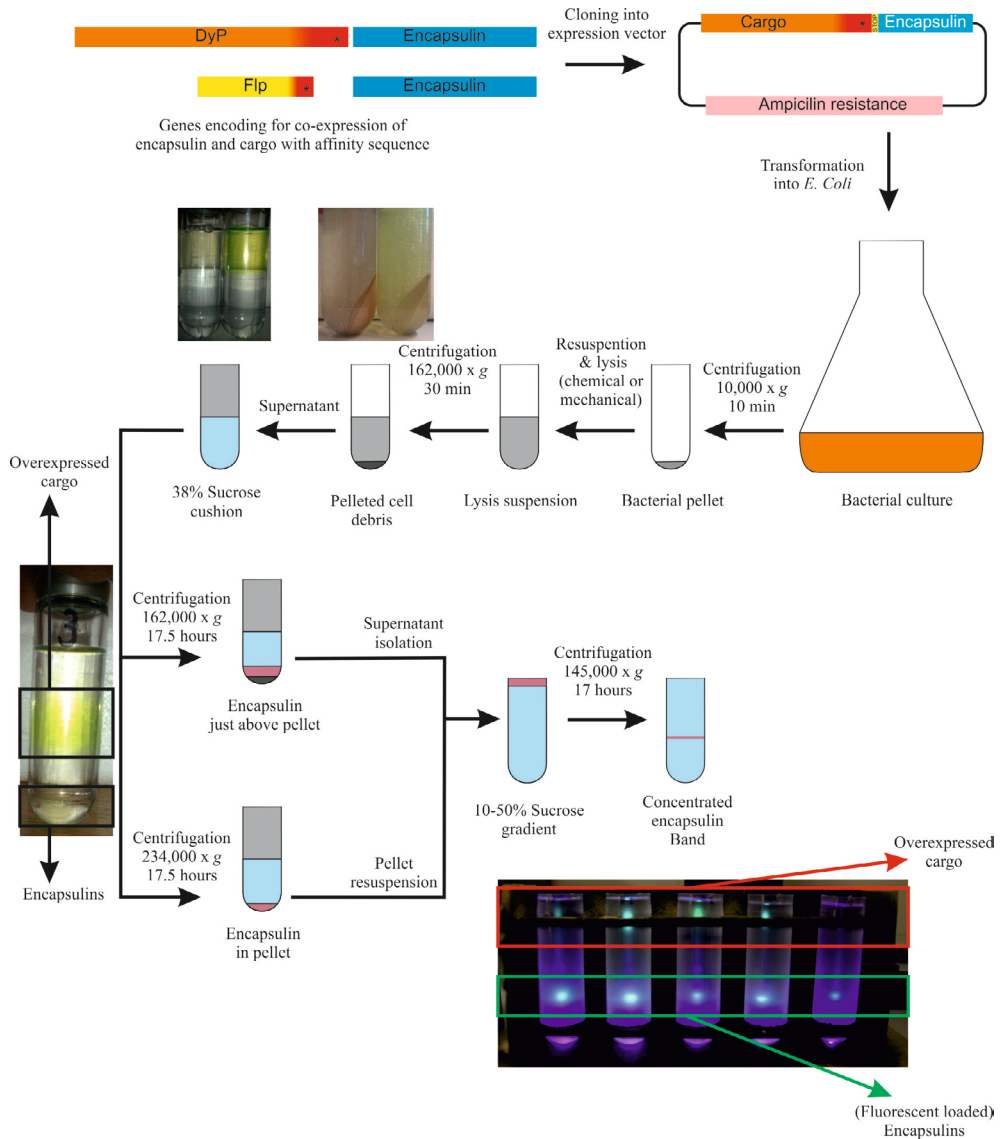


Figure 5.7: Schematic representation of the standard expression and purification of encapsulins. Fluorescent protein loaded encapsulins (Chapter 6) were used to visualize the different fractions.

5.3.2 *Isolation and separation of bacterial encapsulins*

The bacterial cultures were transferred to 250 mL centrifugation tubes and centrifuged for 15 min at $10,000 \times g$ in a TA-10-250 rotor (8,500 rpm) (Figure 5.7). To prevent protein degradation, all centrifugation steps were performed at 4 °C. The pelleted bacterial cells were resuspended in 10 mL ‘Encapsulin buffer’ (20

mM Tris–Cl, 150 mM NH₄Cl, 20 mM MgCl₂, 1 mM β-mercaptoethanol (βME) at pH 7.5). Bacterial cell lysis was induced by an ultrasonic needle, set to full amplitude and power for 2 minutes, to disrupt the cell walls and release the protein content.[†] DNase and RNase were added to induce DNA and RNA digestion (10 μL per 10 mL encapsulin buffer of each) and incubated for 1–2 hours. Alternatively, benzonase (Sigma) was used for both DNA and RNA degradation and lysozyme can be added to improve the bacterial cell lysis.

The insoluble cell debris was pelleted by ultracentrifugation using a Thermo Fisher WX80 ultracentrifuge in a T-865 rotor at 40,000 rpm (162,000 × g) for 15 min at 10 °C[‡].

After the first centrifugation step, the supernatant containing the encapsulins was applied to a ‘sucrose cushion’ to remove any residual contaminants. In this purification step, the sucrose cushion (38% w/v, 10 mL in the encapsulin buffer) was prepared in centrifuge tubes appropriate for the T-865 rotor. The encapsulin supernatant from the first step was carefully layered directly on top of the 38% w/v sucrose cushion before centrifuging for 17 hours at 10 °C at 40,000 rpm (162,000 × g) for Blin encapsulin and at 48,000 rpm (234,000 × g) for TM encapsulin.

5.3.3 *Encapsulin purification by sucrose density gradient*

After the ultracentrifugation step, the Blin encapsulins remain in the lower ~3 mL of the sucrose cushion (directly above the pellet). The 3 mL fraction containing Blin encapsulins was collected and buffer exchanged to encapsulin buffer to remove the sucrose using 100 kDa molecular weight cut-off (MWCO) spin filters, centrifuging at 2500 rpm. After repeated washing and concentrating steps, the Blin encapsulins were concentrated to a final volume of 1 mL in encapsulin buffer. On the other hand, TM encapsulins were pelleted to the bottom of the sucrose cushion during the ultracentrifugation step due to the higher centrifugal speed. Therefore,

[†] It should be noted that chemical lysis using commercially available buffers such as BugBuster are not recommended, as the lysis buffer contains surfactants that interfere with the sucrose cushion in the purification steps. It is strongly recommended to use the Encapsulin buffer at all times during cell lysis and purification however, if BugBuster is used, it is essential that the supernatant is buffer-exchanged to the encapsulin buffer before loading the lysate onto the sucrose cushion.

[‡] For all ultracentrifugation steps, the weight of the tubes are tarred using encapsulin buffer until the mass difference is < 0.01 g.

for TM encapsulins the supernatant can be discarded. Instead, the pellet was carefully rinsed with encapsulin buffer to remove all the sucrose before resuspending the pellet in encapsulin buffer (1 mL).

In both cases, the subsequent purification step involves protein separation by sucrose density fractionation (10–50% w/v). The gradients were prepared by a gradient mixer containing 50% sucrose (17 mL) and 10% sucrose (17 mL), both dissolved in encapsulin buffer. These two solutions were pumped out of the 50% tube into appropriate centrifuge tubes for use in the Surespin 630-36 rotor (Thermo Fisher). To ensure a constant gradient, the gradients were prepared using a peristaltic pump (LBK Bromma, 2232 Microperpex s, set to speed “50”). The partially purified encapsulins that were resuspended in encapsulin buffer in the previous step were then carefully layered on top of the 10–50% sucrose gradient. The samples are then centrifuged at 28,000 rpm ($12,300 \times g$) for 17.5 hours at 10 °C.

The fractions were collected from top to bottom of the tube in the following fractions (in mL): 0–10, 10–15, 15–20, 20–23, 23–26, 26–29, 29–32, and the rest. In general, the encapsulins were found in the 20–23, 23–26 and 26–29 mL fractions.

5.3.4 *Encapsulin analysis*

All of the fractions collected from the sucrose gradient are then analyzed by SDS–PAGE (Figure 5.8A). This shows numerous other bands and, especially for TM, a very faint band corresponding to the encapsulin. Concentration of the fractions containing encapsulins and subsequent FPLC analysis yield two peaks (Figure 5.8B). The 8.4 mL elution peak (and the impurities observed by SDS–PAGE) is attributed to ribosomes which have been reported to co–purify with the encapsulins [11]. SDS–PAGE analysis of the purified peak at 11.7 mL (Figure 5.8C) shows only protein bands corresponding to TM encapsulin and DyP multimers of the proteins. It should be noted that the encapsulins do not appear to fully dissociate upon heating at 99°C for 10 minutes. The appearance of additional bands at approximately 52 kDa and 105 kDa has also been reported by Sutter [11]. In addition, the reason for the high molecular weight of DyP (expected mass of 39.6 kDa, observed 50 kDa) remains unknown but has been reported by Sutter *et al.* [10]. Nevertheless, the fractionation of the encapsulins in the sucrose gradient, the single FPLC elution peak and the appearance of both the encapsulin and

desired cargo protein in the SDS-PAGE fractions is consistent with the successful co-expression and purification of cargo-loaded encapsulins.

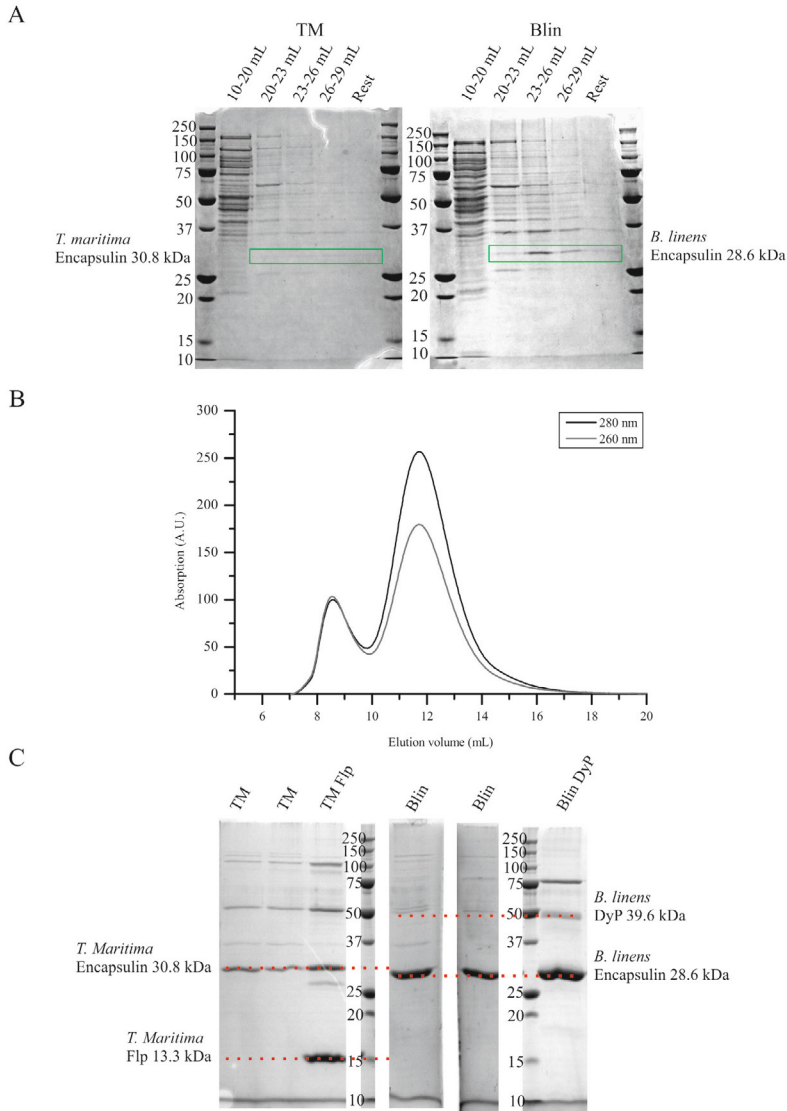


Figure 5.8: Purification of the encapsulins: (A) SDS-PAGE analysis of the sucrose gradient fractions, (B) typical encapsulin FPLC elution profile on a Superose 6 column, (C) SDS-PAGE analysis of the ~ 11.4 mL FPLC elution peak (compilation of different gels, aligned by the marker).

5.3.4.1 Transmission electron microscopy (TEM)

Transmission electron microscopy (TEM) was recorded on an analytical FEG–TEM (Philips CM 30) operated at an acceleration voltage of 300 kV. Samples were prepared by placing 5 μ L of the samples on Formvar carbon–coated copper grids (Electron Microscopy Sciences). The sample was incubated on the grid for 5 minutes, before removing the excess buffer with filter paper. Samples were negatively stained by applying 5 μ L of uranyl acetate (1% w/v in MilliQ water) onto the grid and incubating for 1 minute. The excess stain was removed and the samples were left to dry overnight before imaging.

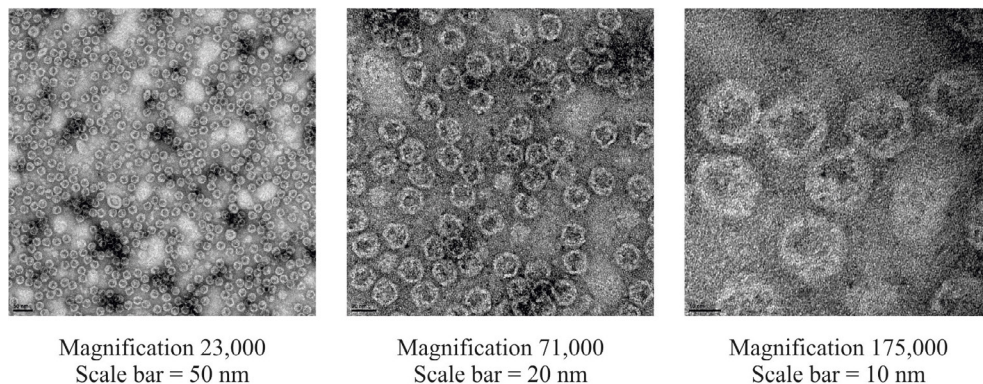


Figure 5.9: TEM micrographs of (empty) Blin at different magnifications reveal highly monodisperse assemblies that are 23 nm in diameter.

Due to the low yield of the TM expression with respect to the Blin expression (Figure 5.8A), only Blin was used for TEM analysis. TEM analysis of the purified Blin encapsulins revealed the presence of well–defined, highly monodisperse particles (Figure 5.9). The diameter of the particles was measured with ImageJ software and was determined to be 23 nm \pm 0.5 nm. The resolution of the assemblies (particularly at 71,000 magnification) shows the 5–sided edges of the encapsulins, which is also consistent with the X–ray crystal structure and EM reconstruction reported by Sutter *et al.* [10].

5.4 Conclusions

In this work, an optimized protocol for the recombinant expression and purification of bacterial encapsulins has been developed. The Expression of Blin yield more encapsulin than TM. Nevertheless, this experimental procedure is highly reproducible and currently describes the isolation of empty encapsulins and/or encapsulins containing their natural cargo (DyP or Flp). However, this protocol is

expected to also be useful for future applications, including encapsulins loaded with non-natural cargo.

5.5 Experienced user protocol

- Grow 5–7 mL overnight cultures of *E. coli* Rosetta (DE3) cells containing *B. linens* or *T. maritima* encapsulin in the presence of ampicillin and chloramphenicol at 37°C, shaking at 210 rpm.
- Use the overnight cultures to inoculate 0.5–1.0 L LB medium containing ampicillin and chloramphenicol and grow these at 37 °C, shaking at 150 rpm, until the optical density (OD₆₀₀) reaches 0.6–0.8.
- Cool the cells to 22 °C while shaking.
- Add IPTG to a final concentration of 0.1 mM and grow overnight at 22 °C, shaking at 150 rpm.

- Harvest bacteria by centrifugation (TA-10-250, 8,500 rpm (10,800 × *g*), 15 min).
- Lyse the bacterial cells in 10 mL encapsulin buffer (20 mM Tris–Cl, 150 mM NH₄Cl, 20 mM MgCl₂, 1 mM βME, pH 7.5) with an ultrasonic probe, full amplitude, full power, 2 min.
- Add DNase and RNase (10 μL of each per 10 mL of encapsulin buffer), incubate for 2 hours.
- Transfer the supernatant to centrifuge tubes suitable for T865 rotor and tare < 10 mg difference.
- Centrifuge using the T865 rotor at 40,000 rpm (162,000 × *g*) for 15 min to precipitate the non-soluble cell debris.
- Prepare T865 centrifuge tubes with 10 mL 38% w/v sucrose in encapsulin buffer.
- Carefully layer 10 mL of the supernatant on top of the 38% w/v sucrose cushion. Centrifuge the sucrose cushion using a T865 rotor. For Blin encapsulins, centrifuge for 17 hours, 40,000 rpm (162,000 × *g*). For TM encapsulins, centrifuge 17 hours, 48,000 rpm (234,000 × *g*).

- For crude Blin encapsulin: collect the last 3 mL of sucrose solution above the pellet. Buffer exchange the fraction to encapsulin buffer by repeat centrifugation/wash steps using 100,000 MWCO spin-filtration (max. 2,500 rpm), end volume ~1 mL.
For crude TM encapsulins: collect the pellet and resuspend in encapsulin buffer, end volume ~1 mL.

Encapsulins: bacterial nanocages

- Prepare a 10–50% sucrose gradient in encapsulin buffer (17 mL each of 10% sucrose and 50% sucrose dissolved in encapsulin buffer).
- Gently layer the ~1 mL of crude encapsulins on top of the sucrose gradient. Centrifuge in a Surespin 630–36 rotor, 28,000 rpm ($12,300 \times g$), 17.5 hours.
- Collect gradient fractions: 0–10, 10–15, 15–20, 20–23, 23–26, 26–29, 29–32, rest

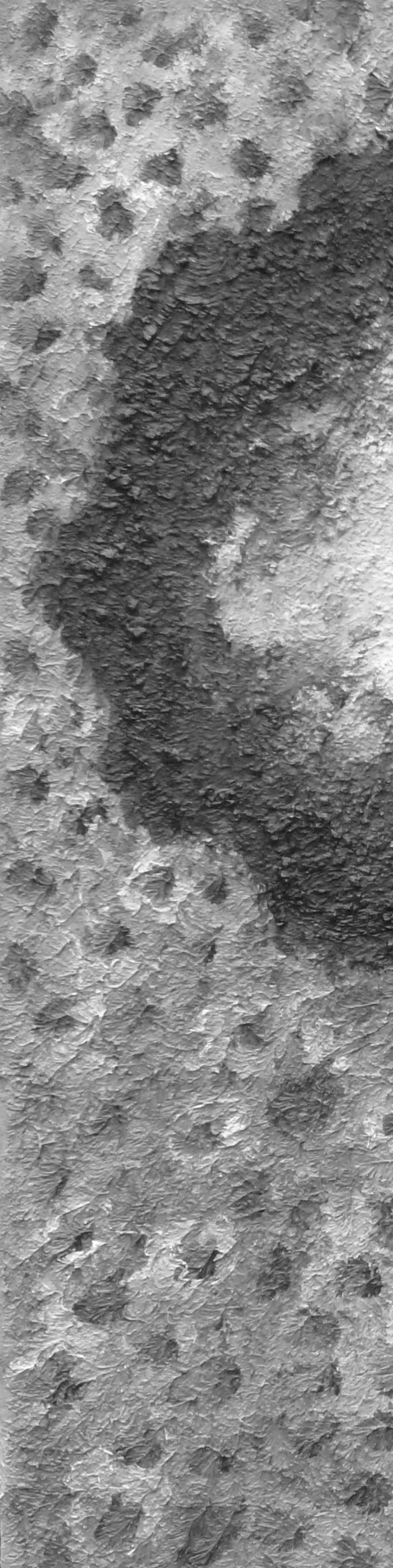
Table 5.2: Encapsulin buffer composition, with addition of the sucrose amounts necessary to build the sucrose gradient.

Encapsulin buffer				
Ingredient	Molar [mol/L]	Volume [L]	Molar mass [g/mol]	Mass [g]
Trizma Base	0.02	1	121.14	2.4228
NH ₄ Cl	0.15	1	53.49	8.0235
MgCl ₂	0.02	1	203.31	4.0662
2-mercaptoethanol	0.001	1	78.13	0.070 ml

For 0.5 L	% mass	Volume [L]		mass [g]
sucrose	50	0.5		250
sucrose	10	0.5		50

5.6 References

1. Fischlechner, M. and E. Donath, *Viruses as building blocks for materials and devices*. *Angew Chem Int Ed Engl*, 2007. **46**(18): p. 3184-93.
2. Kang, S., et al., *Implementation of p22 viral capsids as nanoplatforms*. *Biomacromolecules*, 2010. **11**(10): p. 2804-9.
3. Lee, L., Z. Niu, and Q. Wang, *Viruses and virus-like protein assemblies—Chemically programmable nanoscale building blocks*. *Nano Research*, 2009. **2**(5): p. 349-364.
4. Renggli, K., et al., *Selective and Responsive Nanoreactors*. *Advanced Functional Materials*, 2011. **21**(7): p. 1241-1259.
5. Soto, C.M. and B.R. Ratna, *Virus hybrids as nanomaterials for biotechnology*. *Curr Opin Biotechnol*, 2010. **21**(4): p. 426-38.
6. Uchida, M., et al., *Biological containers: Protein cages as multifunctional nanoplatforms*. *Advanced Materials*, 2007. **19**(8): p. 1025-1042.
7. Worsdorfer, B., K.J. Woycechowsky, and D. Hilvert, *Directed evolution of a protein container*. *Science*, 2011. **331**(6017): p. 589-92.
8. Valdes-Stauber, N. and S. Scherer, *Isolation and characterization of Linocin M18, a bacteriocin produced by Brevibacterium linens*. *Appl Environ Microbiol*, 1994. **60**(10): p. 3809-14.
9. Hicks, P.M., et al., *Homomultimeric protease in the hyperthermophilic bacterium Thermotoga maritima has structural and amino acid sequence homology to bacteriocins in mesophilic bacteria*. *FEBS Lett*, 1998. **440**(3): p. 393-8.
10. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment*. *Nat Struct Mol Biol*, 2008. **15**(9): p. 939-47.
11. Sutter, M., *Structural Basis of Enzyme Encapsulation into a Bacterial Nanocompartment*. 2008, ETH ZÜRICH: ZÜRICH.



Chapter 6

Controlling encapsulin loading

The use of encapsulins for potential applications in nanotechnology is explored. The natural loading pathway of encapsulins is utilized to direct fluorescent protein cargo into the encapsulins as a first attempt to control the loading of these highly novel protein nanocages. Analysis through different techniques hints at a complex loading mechanism, that is highly efficient at encapsulating cargo in the *B. linens* encapsulin.

Part of this chapter is based on W. Frederik Rurup, Joost Snijder, Melissa S. T. Koay, Albert J. R. Heck, Jeroen J. L. M. Cornelissen “Loading of fluorescent proteins in a bacterial nanocompartment” manuscript in preparation.

6.1 Introduction

The origin and recombinant production of the encapsulins of *Brevibacterium linens* and *Thermatoga maritima* have been thoroughly discussed in Chapter 5. As mentioned previously, encapsulins are highly promising candidates as the next generation of protein cages. Similar to viruses and virus-like assemblies, encapsulins are composed entirely of proteins, which make them functionalizable, they can be genetically-engineered and they form highly monodisperse assemblies. However unlike virus-based assemblies, encapsulins offer a distinct advantage over viruses since they are highly pH and temperature stable. Furthermore, they are non-viral which may be of significant relevance for applications in nanomedicine.

6.1.1 *Encapsulin*

The Blin encapsulin is a 60-mer of 28.6 kDa proteins, arranged as a hollow cage of approximately 24 nm in diameter with 2 nm thick walls. The native *B. linens* encapsulin is loaded with a dye-decolorizing peroxidase (DyP), which forms a trimer of dimers. The trimer of dimers bear targeting extension sequences that align precisely over the threefold symmetry binding pockets on the inside of the encapsulin.[1] The targeting sequence docking and alignment suggests that multivalency may play an important role in encapsulin loading. The *in vivo* co-expression of the encapsulin and its native cargo, and its natural loading mechanism via utilization of a specific C-terminal amino acid sequence is described in Chapter 5.

The aim of this work was to investigate whether encapsulins, with their natural loading affinities, can be used as a nanocontainer and how multivalency influences the *in vivo* loading behavior. To do so, the encapsulins were engineered and explored for the controlled loading of artificial cargo. The DyP of *B. linens M18* was replaced by either the monomeric or dimeric Teal Fluorescent Protein (mTFP and dTFP),[2] bearing the appropriate targeting sequence.[1] This work describes the first example of a recombinantly expressed non-viral nanocontainer of bacterial origin, assembled *in vivo*, in which the native cargo was replaced with non-natural cargo.

6.1.2 *Teal Fluorescent Protein cargo*

The mTFP is a genetically engineered variant of the native homotetramer. In Chapter 3, the restoration of one of the TFP dimerization interfaces was described in experimental detail, to yield the dimeric Teal Fluorescent Protein (dTFP), which was confirmed by DNA sequencing. The X-ray crystal structure of mTFP contains

a C-terminus that protrudes outwards from the tertiary structure. This makes them ideal model proteins for loading in encapsulins since genetic fusion of the encapsulin extension peptides to TFP will not perturb sequence docking, and instead the extension peptides will be presented towards the solution with optimal flexibility.

6.2 Results and discussion

As described in Chapter 5, the pET21a based vector encoding the two-gene operon from *B. linens* M18 and DyP was a generous gift from Dr. M. Sutter (ETH Zurich, Switzerland). Here, vector #6 containing encapsulin, the extension sequence and GFP, was used as a template to engineer mTFP and dTFP.[1] Studies by Sutter *et al.* were only used to show co-migration of co-expressed GFP with the *B. linens* encapsulin. No quantitative analysis of loading numbers was performed.

Here, site-directed mutagenesis was used to introduce a BamHI site at the 3'-end of the GFP gene. The unique BamHI and XbaI restriction sites in the plasmid were used to excise the GFP sequence. DNA sequences coding for mTFP and dTFP were multiplied and digested with the same restriction enzymes. The sequences were subsequently cloned into place. The resultant Blin mTFP co-expression construct (BmT) encodes for the monomeric TFP protein bearing the C-terminal affinity tag for the encapsulin and the encapsulin shell protein, whilst BdT encodes the dimeric TFP bearing the natural C-terminal affinity tag and the encapsulin shell protein. DNA sequencing confirmed the creation of these sequences.

Both the BmT and BdT co-expression operons were transformed into *E. coli* Rosetta (DE3) and expressed overnight. As a control, the empty *B. linens* encapsulin was also expressed in *E. coli*. The encapsulins were isolated and purified using the same procedure described in the previous chapter, involving the combined sucrose cushion, sucrose gradient ultracentrifugation and fast protein liquid chromatography (FPLC) steps.

6.2.1 UV-Visible spectroscopy

UV-visible (UV-vis) detection of the FPLC elution fractions showed a single band eluting at 1.4 mL (Figure 6.1, left). The characteristic absorbance peaks for both TFP ($\lambda=462$ nm, 280 nm) and encapsulins ($\lambda=280$ nm) were observed, suggesting that co-expressed TFP was co-assembled and co-elutes with the encapsulins. The absorbance ratio at $\lambda = 280$ nm relative to 260 nm ($A_{280/260} = 1.8$) indicates that the peak is predominantly protein-based material and not nucleic acid based (in which the $A_{280/260}$ would be less than 1.0). The purified encapsulins were analyzed by

Transmission Electron Microscopy (TEM), which confirmed the correct size (24 nm) and morphology (Figure 6.1, right).

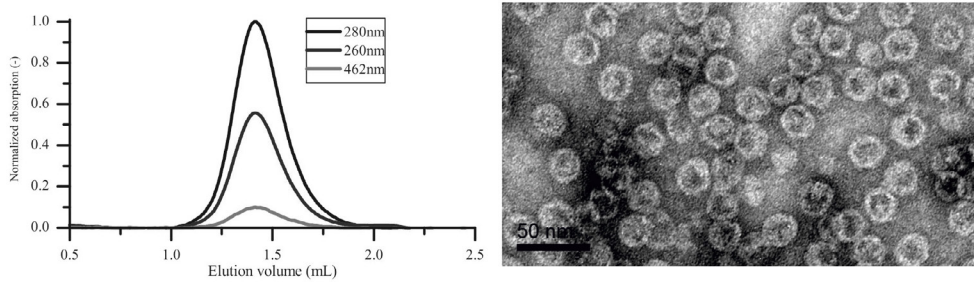


Figure 6.1: Purification and characterization of bacterial encapsulins. BmT is shown as a representable example with (a) The FPLC elution profile shows the co-elution of TFP ($\lambda = 462$ nm, 280 nm) and encapsulin ($\lambda = 280, 260$ nm), (b) TEM image of the purified encapsulins (scale bar 50 nm).

The UV/visible absorbance spectrum of the empty encapsulins showed no absorbance band at 462 nm, thus confirming the absence of fluorescent proteins (Figure 6.2). As described in Chapter 3, the fluorophore specific absorption of TFP at 462 nm can be used to correlate the absorbance at 280 nm due to TFP. By determining the absorbance contribution due to TFP, the remaining absorbance at 280 nm can be attributed to the encapsulin proteins. Based on the relative concentration ratios, the loading of encapsulin and TFP can be determined.

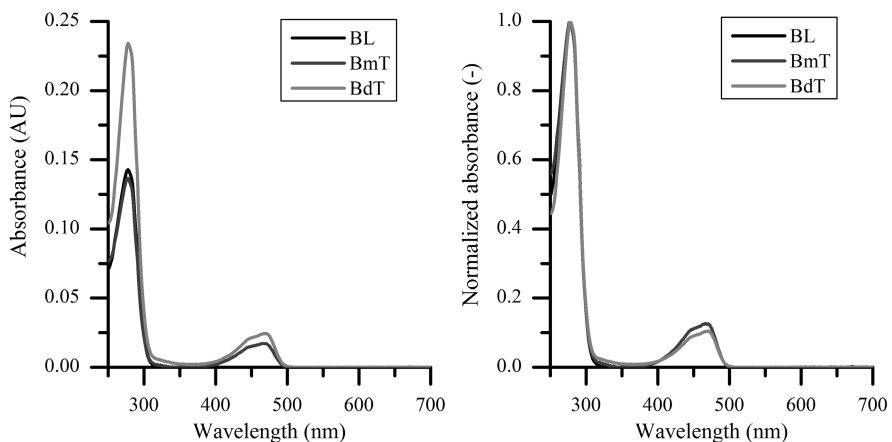


Figure 6.2: UV-visible absorption spectrum of empty and TFP filled encapsulins (left). When normalized to the 280 nm absorption, the loading difference between BmT and BdT is visible.

Figure 6.2 shows a typical absorbance spectrum for encapsulins loaded with TFP. In this particular example, BmT and BdT were determined to have a concentration of 4.1 and 7.2 μM fluorescent protein monomers respectively, suggesting that 3.4 TFPs were loaded per encapsulin for BmT and 2.9 TFPs for BdT. For both variants, the expression, purification and *in vivo* loading were repeated several times (five independent sample batches were eventually characterized). On average, it was found that the loading was almost equal for the monomeric and dimeric TFP, with the average for the dimer even slightly lower than the monomeric cargo. The average loading was calculated to be 4.3 ± 0.8 TFPs for mTFP and 3.1 ± 0.3 TFPs for dTFP.

6.2.2 SDS-PAGE densitometry

Based on the staining intensities of the SDS-PAGE gels of all BmT and BdT batches, a higher loading was determined. The gel analyzer tool of ImageJ software was used to quantify the TFP loading, by comparing the intensity ratios of the bands at 34.7 kDa (corresponding to TFP bearing the targeting sequence) and that at 28.6 kDa (corresponding to *B. linens* encapsulin) (Figure 6.3).

For BmT an average ratio of 1:2.9 for TFP:encapsulin was found. For BdT the average was 1:3.9. Taking into account that 60 encapsulin proteins form one encapsulin, this corresponds to on average 20.2 ± 8.3 mTFP and 15.1 ± 3.8 dTFP monomers in each encapsulin.

It should be noted that the loading numbers of BmT and BdT according to the SDS-PAGE densitometry are not significantly different due to the large margin of error, however, in comparison with the UV-visible absorbance data, the loading determined with the former technique is much higher.

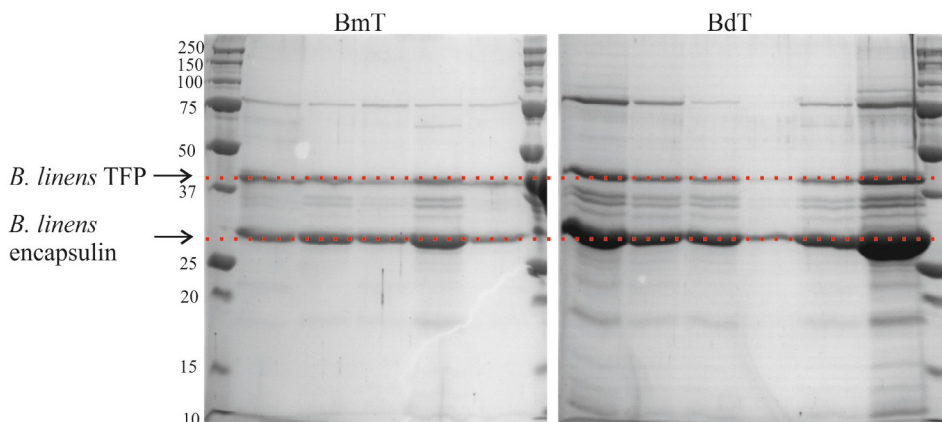


Figure 6.3: SDS-PAGE gels of multiple batches of BmT (left) and BdT (right). Each lane is a different batch of either BmT or BdT.

6.2.3 *Mass spectrometry*

To further confirm the encapsulin loading efficiency and composition (and hence determine the accuracy between UV–visible absorbance quantification vs. quantitative SDS–PAGE), native mass spectrometry (MS) analysis was performed in collaboration with Prof. Albert Heck, Utrecht University, The Netherlands. Based on the expected molecular masses for TFP (34.7 kDa) and the encapsulin monomer (28.6 kDa), the calculated masses correlated to an average of 12.2 ± 2.0 mTFP and 12.4 ± 1.9 dTFP (monomeric units) per encapsulin. Both mTFP and dTFP filled encapsulins show a single charge state distribution with extensive overlap between peaks (Figure 6.4). The peak width suggests heterogeneity amongst the encapsulin assemblies (*i.e.* the presence of more than one mass species).

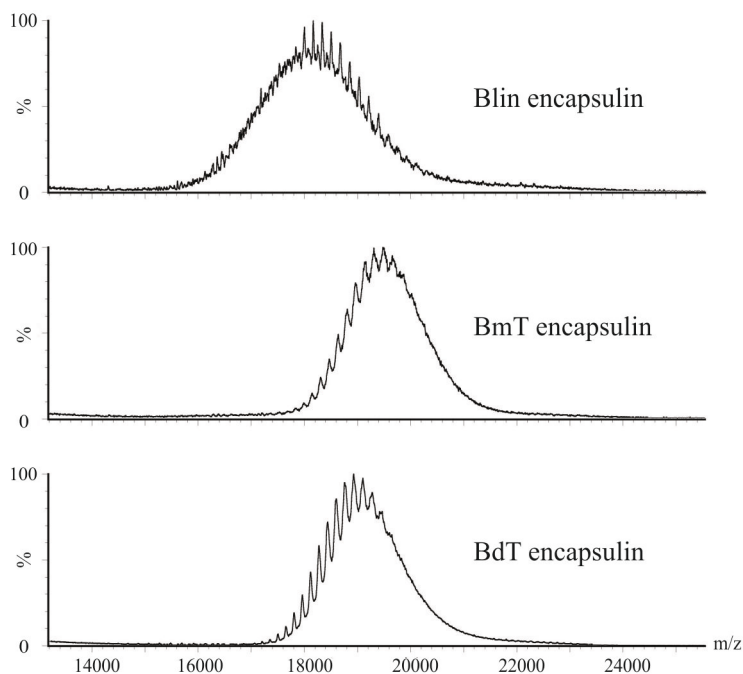


Figure 6.4: Charge state distributions of Blin encapsulin, BmT and BdT.

More detailed and quantitative analysis was performed using collision–induced dissociation tandem MS (CID–MS²). In CID–MS², a specific region of interest of the mass spectrum is selected for collision–induced dissociation. The fragments of this dissociation are subsequently analyzed in the time–of–flight mass analyzer, allowing a more accurate mass assignment.

CID-MS2 revealed additional peaks that could be assigned to the step-wise dissociation of encapsulin (*i.e.* loss of single encapsulin monomers), which is typically observed for large protein complexes (Figure 6.5). Where the intact encapsulin shows only a single series of resolved charge states, two series of peaks can be resolved in the first dissociation product (loss of one subunit), four series of peaks in the second product and one series in the third product. The mass distributions corresponded to BmT encapsulins containing 9.0, 10.0 and 11.0 mTFPs, and BdT encapsulins containing 10.0, 11.0, 12.0 and 13.0 dTFP monomers per encapsulin. Interestingly, an odd number of TFPs were identified for the dimeric TFP.

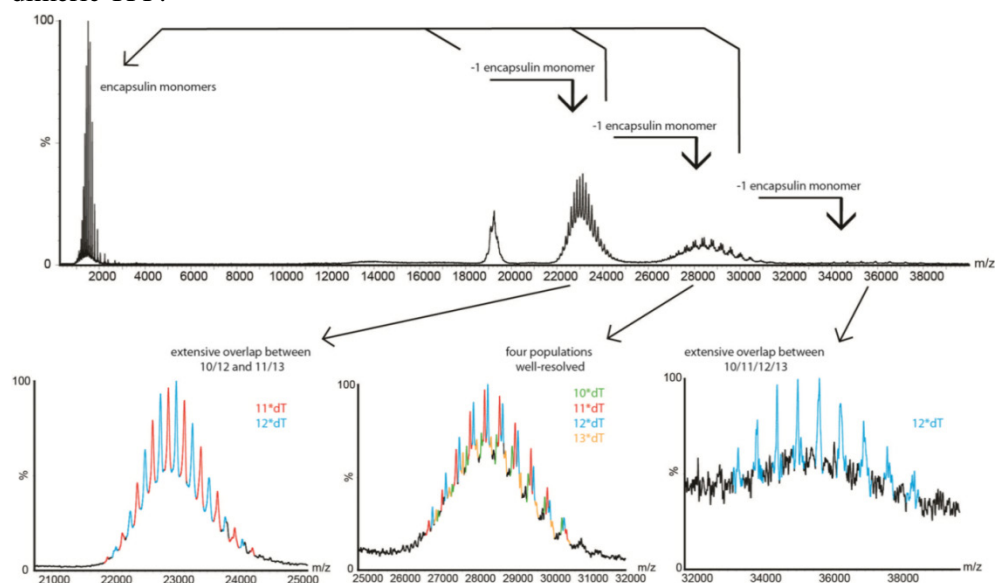


Figure 6.5: CID-MS2 analysis of BdT, showing the different peak clusters for the dissociation products (top) and the individual peak allocation per dissociation product (bottom).

6.2.4 *Comparison of TFP loading analysis*

Analysis of SDS-PAGE densitometry and mass spectrometric data showed consistent loading of mTFP and dTFP. Surprisingly, comparison of monomeric and dimeric TFP did not show any increased loading, suggesting that the native C-terminal affinity tag itself is highly efficient for *in vivo* packaging (Figure 6.6).

Despite the CID-MS2 being resolved for a monomeric loading distribution of 9–11 TFPs per Blin and a dimeric loading of 10–13 TFPs per Blin, the average MS1 data indicate that no significant difference in loading is caused by the dimerization of the TFP.

The very similar loading for dTFP compared to mTFP was rather striking, since cargo dimerization was shown to significantly enhance the loading efficiency in CCMV (Chapter 3). Comparing the two loading strategies for CCMV compared to encapsulins, loading of the encapsulins is based on *in vivo* self-assembly, rather than *in vitro* assembly. Since no loading improvement is observed for dTFP over mTFP and even odd numbers of dimeric TFP are loaded in the encapsulins, this suggests that the driving force for cargo docking and encapsulation is greater than cargo dimerization interactions.

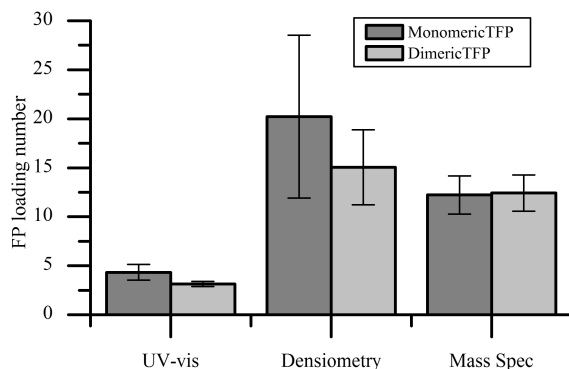


Figure 6.6: Average TFP loading ratios per encapsulin, as determined for each quantitative method.

Interestingly, native MS analysis of the empty Blin encapsulins showed that these particles are not truly empty. There is a mass increase of ~268 kDa compared to empty 60-mer shells, based on the resolved peaks around 18000 m/z (Figure 6.4). This is possibly an encapsulated cytosolic component of the host *E. coli*, which does not interfere with the fluorophore UV-Vis signal and could not be seen on the SDS-PAGE as a clear, single band. To date the origin of this mass is unknown, but it is consistently observed in mass spectrometry and absent in SDS-PAGE analysis. Further studies are necessary to assign the additional mass.

Overall, the CID-MS2 data showed integer loading numbers with high accuracy, which are supported by SDS-PAGE densitometry data of the same order of magnitude. On the other hand, quantification by UV-visible absorbance showed very large discrepancies. Since the mass spectrometry shows that the TFP molecules are present, the most plausible cause is that the chromophore of the TFP does not properly mature upon *in vivo* encapsulation in the encapsulin. Since the absorbance quantification is highly dependent on the chromophore absorbance at

$\lambda=462$ nm, non-matured TFP would not be detected and hence would be severely underestimated. However, this explanation is only speculative and further investigations are currently in process. Strikingly, the mass spectrometry shows (on average) 11–12 fluorescent proteins per encapsulin for both the mTFP and dTFP. These numbers correlate well with the number of homopentamers reported in the encapsulin shell (60 monomers = 12 homopentamers). It remains unknown whether there is a correlation between the cargo loading and the homopentamer arrangement. However, since encapsulin loading and assembly is an *in vivo* driven process, the consistent loading of 11–12 fluorescent proteins (independent of whether mTFP or dTFP is used) is highly intriguing and could suggest that encapsulin assembly follows a very strict loading pathway.

6.2.5 *Different cargo loading in encapsulin*

Attempts were made to genetically engineer and clone mStrawberry or dsRed to replace the mTFP in the BmT encapsulin. The unique NdeI and BsrGI restriction sites were used to replace the mTFP sequence in the BmT vector with that of mStrawberry. Appropriate primers were selected to introduce the necessary restriction sites to the mStrawberry sequence, which was amplified by PCR. Both the plasmid and the PCR product were double digested with NdeI and BsrGI restriction enzymes and purified by agarose gel electrophoresis.

The dsRed sequence was amplified from the dsRed E-coil vector (Chapter 3). Both the PCR product and the BmT plasmid were doubly digested with XbaI and BamHI restriction enzymes and purified by agarose gel electrophoresis.

Despite several attempts, no colonies for co-expression of encapsulin and either mStrawberry or dsRed were obtained. Most sequencing results showed either self-ligation of the plasmid or mTFP still in its original place, suggesting the double digestion of the plasmid vector was incomplete. Increasing the restriction digestion time and careful separation of the plasmid fragments by agarose gel electrophoresis also failed.

6.3 Conclusion

In summary, the efficiency of the natural loading pathway in *B. linens* is shown by the *in vivo* encapsulation of monomeric and dimeric TFP. Interestingly, despite the precedent of the multimeric natural cargo, no significant influence of the cargo multivalency on the loading efficiency was observed. UV-visible spectroscopic analyses gave on average 4.3 mTFP and 3.1 dTFP per Blin, while SDS-PAGE densitometry resulted in 20.2 mTFP and 15.1 dTFP per Blin and mass spectrometry

in 12.2 mTFP and 12.4 dTFP per Blin. These data show significant discrepancies in TFP loading numbers when quantified using each of these different detection methods. This was attributed to the effectiveness of loading: we assume that the *in vivo* loading of the Blin encapsulin is more favorable than the fluorescent protein dimerization and it is probably so effective that the fluorescent proteins are encapsulated before they are properly matured. This may explain the low number of fluorescent proteins observed by UV-vis spectroscopy, while the presence of a higher (integer) number of fluorescent proteins is measured by mass spectrometry in an absolute way. Nevertheless, this work demonstrates the future potential of a very effective loading mechanism for the encapsulation of proteins in non-viral, highly stable nanocages. This opens new routes into the emerging field of protein cages in nanotechnology.

6.4 Acknowledgements

J. Snijder is gratefully acknowledged for fruitful discussions and his help performing all mass spectrometry experiments in this chapter.

6.5 Experimental

6.5.1 *Encapsulin co-expression and purification*

A detailed protocol of the expression and purification of the Blin encapsulin is described in Chapter 5.

6.5.2 *Insertion of cargo in BmT plasmid*

The primers used in this work are listed as follows:

- (1) 5'-GGC ATG GAT GAA CTA TAG GAT CCG GGA GAG CTC GCC GCC G-3'
- (2) 5'- CGG CGG CGA GCT CTC CCG GAT CCT ATA GTT CAT CCA TGC C -3'
- (3) 5'- GGA ATT GTG AGC GGA TAA-3'
- (4) 5'- CTC CCG GAT CCT TGT ACA GCT CGT CCA TGC CG-3'
- (5) 5'-GGA ATT GTG AGC GGA TAA-3'
- (6) 5'-CCG CGG ATC CTA CAG GAA CAG GTG GTG GCG GCC C-3'
- (7) 5'-CCG GTC GCC CAT ATG GTG AGC AAG GGC GAG GAG-3'
- (8) 5'-CCA GGG TCT TGT ACA GCT CGT CCA TGC-3'

To create pET21 BmT, the DNA sequence encoding for GFP was replaced with the mTFP DNA sequence. Using the #5 vector (Chapter 5) as DNA template, a BamHI site was introduced between the GFP gene and the C-terminal affinity tag using primers (1) and (2), according to the standard QuikChange PCR protocol. Double digestion with XbaI and BamHI allowed excision of the GFP sequence. Using primers (3) and (4), the mTFP and dTFP sequences were amplified from the pET15 mTE and pET15 dTE plasmids (described in Chapter 3). Double digestion XbaI and BamHI and subsequent DNA ligation of the appropriate fragments yielded pET21 BmT and pET21 BdT.

The same procedure was used for cloning of pET21 dsRed: pET21 BmT was used as a template and the insert was amplified by PCR using primers (5) and (6). Both pET21 BmT and the PCR products were double digested with XbaI and BamHI and ligated under standard conditions. The mStrawberry sequence was amplified by PCR using primers (7) and (8). The mStrawberry gene was cloned into pET21 Blin mStrawberry by double digestion with NdeI and BsrGI and subsequent DNA ligation.

6.5.3 *UV-visible spectrometry*

All UV-visible measurements were performed on a Perkin Elmer Lambda 850 Spectrometer. Standard quartz cuvettes with a 1 cm path length were used. The slit widths are optimized to achieve the best signal to noise ratio.

6.5.4 SDS-PAGE densitometry

All chemicals for SDS-PAGE analysis were purchased from Sigma. Standard SDS-PAGE was performed with 12% polyacrylamide gels using PeqLab systems. Bio-Safe Coomassie (Bio-Rad laboratories, Inc.) was used to visualize the SDS-PAGE bands. SDS-PAGE densitometry analyses were performed with the ImageJ software version 1.46 (windows version of NIH ImageJ, <http://imagej.nih.gov/ij>) after a background subtraction (rolling ball, $r = 50$ pixels).

6.5.5 Mass spectrometry

All native mass spectrometry was performed in collaboration with the group of Prof. A.J. Heck, Utrecht University, The Netherlands. All samples were prepared for mass spectrometry by repeated dilution and concentration to exchange the buffer to 100 mM ammonium acetate (pH 6.8–7.0), using Vivaspin 500 10K MWCO centrifugal filter units. Aliquots of 1–2 μL were loaded into gold coated borosilicate capillaries, prepared in-house, for nano-electrospray ionization. Samples were analyzed on a modified QToF II instrument optimized for high-mass protein analysis [3]. The instrument operates at elevated pressure in the source region and hexapole ion guide, which results in collisional cooling of the ions and improves transmission. The measurements were performed with Xenon as collision gas. The source conditions for the encapsulin were as follows: capillary 1300–1500V, cone 160V, extraction cone 0V and 10 mbar backing pressure. The pressure in the collision cell was 2×10^{-2} mbar.

6.5.6 Transmission electron microscopy (TEM)

TEM micrographs were recorded on an analytical FEG-TEM (Phillips CM 30) operated at 300 kV acceleration voltages. Samples were prepared by placing a 5 μL drop of the samples on Formvar carbon-coated copper grids (Electron Microscopy Sciences). The sample drop was left on the grid for 5 minutes, after which time the excess buffer was blotted away with filter paper. Samples were negatively stained by applying 5 μL of stain (1% w/v uranyl acetate in MilliQ water) onto the grid and removing the excess stain away after 1 minute with filter paper. The samples were dried overnight before imaging.

6.6 References

1. Sutter, M., *Structural Basis of Enzyme Encapsulation into a Bacterial Nanocompartment*. 2008, ETH ZÜRICH: ZÜRICH.
2. Ai, H.W., et al., *Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging*. *Biochem J*, 2006. **400**(3): p. 531-40.
3. Lorenzen, K., et al., *Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collision gases*. *International Journal of Mass Spectrometry*, 2007. **268**(2-3): p. 198-206.



Chapter 7

Summary

7.1 Summary

Traditional organelles are composed of lipid-based membranes, however, the recent discovery of protein-based microcompartments has generated much interest by suggesting that compartmentalization is important across all domains of life, from humans to plants to bacteria. Yet, very little is understood about how molecular reactions and processes in compartmentalization occur. Although there are many different (synthetic) approaches to mimic the design of natural compartments, the work described in this thesis focuses on the use of natural protein building blocks to assemble nanocages. The most common strategies for the loading of different protein cages have been discussed in Chapter 2, however, such strategies often suffer from a lack of numerical control over cargo loading. The aim of the work in this thesis was to create a controllable methodology to selectively load protein nanocages with the cargo of interest. This thesis focuses on two fundamental approaches to optimize the loading: (1) modifying the cargo by genetically engineering the cargo interface, or (2) exploring different protein cages.

The capsid of the Cowpea Chlorotic Mottle Virus (CCMV) was used as a well-established platform for the design of new cargo loading techniques. The known leucine zipper based loading system, involving non-covalent tethering of the cargo to CCMV capsid proteins, was improved in two ways. First, the EGFP fluorescent cargo was exchanged for a brighter, more pH-stable teal fluorescent protein (TFP), in order to overcome problems due to low pH-stability of the chromophore (which was known for EGFP). Second, TFP was genetically engineered to form homo-dimers or homo-tetramers, in order to increase the intrinsic loading ratios. Extensive studies confirmed the multimeric states improved the overall loading efficiency of cargo into CCMV assemblies. Despite the improvement, it was found that a concentration of 10 μM or more of TFP (and thus zipper coils) induces zipper dissociation and hinders the loading of the capsids, therefore rendering the zipper approach ineffective for controlled loading studies.

To circumvent this problem, a covalent fusion protein of the cargo and the CCMV capsid protein was engineered to induce directed loading. This loading system exhibited a linear loading dependency of up to ~ 20 fluorescent proteins. In a follow-up study, the linear loading dependency of the covalent loading system was used to predict and pre-calculate a desired number of fluorescent proteins to be encapsulated per capsid. These calculations were found to be very accurate up to ~ 20 TFP per capsid and thereby allowing numerical controlled loading of CCMV capsids.

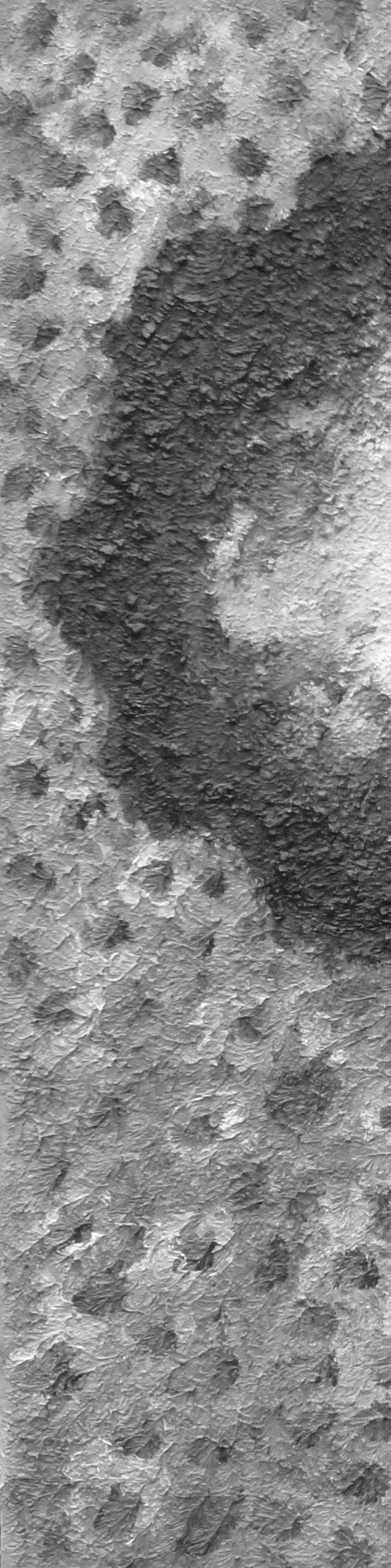
In order to determine the crowding and freedom of movement inside the CCMV capsids, the fluorescent properties of the cargo in these selectively loaded capsids was investigated. Based on steady-state fluorescence anisotropy, it was observed that loading of the fluorescent proteins in the confined environment of the CCMV capsid significantly restricts their freedom of movement. Upon increasing the cargo loading within the single capsid, molecular crowding effects were observed. Loading could be controlled up to ~20 fluorescent proteins per capsid, however, such cargo loading ratios also led to erroneously formed capsids, suggesting steric hindrance plays an important role during capsid assembly.

Finally, a new type of protein nanocage was explored: the bacterial encapsulins. Here, the expression and purification of the *Brevibacterium linens* and *Thermotoga maritima* encapsulins in *E. coli* were optimized. Furthermore, a strategy was developed using the naturally occurring peptide sequence of *B. linens* encapsulins to promote docking and loading of TFP. In contrast to CCMV loading efficiency, the formation of multimeric cargo states (*i.e.* TFP homodimerization) did not significantly improve cargo loading in encapsulins. Since odd integer loading numbers were found, it is assumed that encapsulating interactions are stronger than cargo dimerization interactions. This also explains, most probably, the discrepancies we found in different analysis techniques for determining the loading numbers, where we found that the fluorescent proteins did not mature correctly before encapsulation and apparently cannot mature further after encapsulation.

Overall, the encapsulation of non-natural cargo in encapsulins is still a very novel approach and there remain many things to be investigated. In particular, a clear difference between *in vivo* and *in vitro* encapsulation has to be made. From this thesis, it is apparent that good control over VLP loading can be achieved *in vitro* by purifying and carefully mixing the cargo. By comparison, *in vivo* loading cannot be controlled by adjusting the mixing ratios, but instead, is driven by natural loading mechanisms. This could be a significant drawback, since there is a lack of numerical control for *in vivo* encapsulation (in encapsulins). To fully understand the loading pathway, a different cargo other than mTFP should be investigated. Furthermore, although the inner volume of both a CCMV VLP and encapsulin is about equal, the maximum loading deviates quite significantly (~20 and 12 respectively). It remains unknown whether this difference in loading efficiency is dictated by the cargo or by the protein shell. In addition, the size of the cargo should also be taken into account, since it is unknown whether cargo crowding in the shell is a limiting factor for loading. By understanding which factors drive and

Summary

influence the self-assembly (and loading) of encapsulins, the use of non-viral protein cages, equipped with *in vivo* loading mechanisms, could open up new and exciting opportunities for future applications in nanotechnology.



Chapter 8

Samenvatting

8.1 Samenvatting

Organellen worden altijd beschreven als celcompartimenten met een membraan opgebouwd uit vetten. De recente ontdekking van microcompartimenten die louter bestaan uit eiwitten in verschillende organismen heeft echter veel stof doen opwaaien, omdat dit impliceert dat compartimentalisatie niet alleen voorkomt bij hogere soorten, maar in alle taxonomische domeinen, van mensen tot planten tot bacteriën. Over het belang van compartimenten in organismen, en hoe reacties en processen op moleculair niveau in zulke compartimenten in al deze levensvormen verlopen is nog weinig bekend. Er zijn veel verschillende (synthetische) benaderingen ontworpen om de functie van natuurlijke compartimenten na te bootsen en te onderzoeken, maar het werk in deze thesis beperkt zich tot het gebruik van natuurlijke bouwstenen voor de opbouw van kooi-achtige nanostructuren.

De meest gebruikte methodes voor het laden van verschillende eiwitnanokooien worden besproken in hoofdstuk 2. Al deze ladingmethodes hebben echter als nadeel dat ze geen controle geven over het aantal geladen moleculen. Het doel van deze thesis is juist het creëren van een ladingsmethode waarmee selectief een specifiek aantal moleculen in de nanokooien geladen kan worden. De focus ligt hierbij op twee fundamentele benaderingen: (1) het aanpassen van de lading door genetische manipulatie en (2) het onderzoeken van verschillende eiwitkooien.

De capside van het “cowpea chlorotic mottle virus” CCMV is een bekend en goed gekarakteriseerd platform voor het ontwerpen van nieuwe beladingstechnieken in eiwitnanokooien. Het bekende ladingssysteem, gebaseerd op de zogenaamde ‘leucine-rits’, die een niet-covalente verankering van de lading aan het capside-eiwit hanteert, is op twee manieren verbeterd. Ten eerste is de lading (EGFP) vervangen door een helderder, meer pH-stabiel fluorescent eiwit (TFP). Hiermee werden detectieproblemen opgelost, die veroorzaakt werden door lage pH-stabiliteit van het chromophore (EGFP). Ten tweede is het TFP zodanig gemuteerd dat het homodimeren en homotetrameren vormt, zodat de inherente lading-ratio’s sterk verbeteren. Uitgebreide studies hebben de multimere staat van de fluorescente eiwitten bevestigd en aangetoond dat deze staat de ladingsefficiëntie ten goede komt. Ondanks de verbeterde ladingsefficiëntie werd bevonden dat een concentratie van 10 μM TFP of meer zorgt voor het uiteenvallen van de ‘leucine-rits’. Dit hindert de belading van de capsides en daarmee is gebleken dat het ‘leucine-rits’-systeem niet geschikt is voor gecontroleerde lading van het CCMV-capside.

Om dit probleem te omzeilen is een fusie-eiwit gemaakt dat de domeinen van de lading en het capsid-eiwit covalent aan elkaar koppelt. Dit systeem vertoont een praktisch lineair beladingsverband tussen de aangeboden en geladen fluorescente eiwitten tot een belading van ongeveer 20 fluorescente eiwitten. In een vervolgstudie is dit lineaire gedrag gebruikt om te berekenen en te voorspellen wat de gewenste lading van fluorescente eiwitten per CCMV-capsid zou worden. De berekeningen bleken zeer accuraat tot 20 geladen eiwitten en zorgen daarmee voor een numeriek gecontroleerde belading van CCMV-capsides.

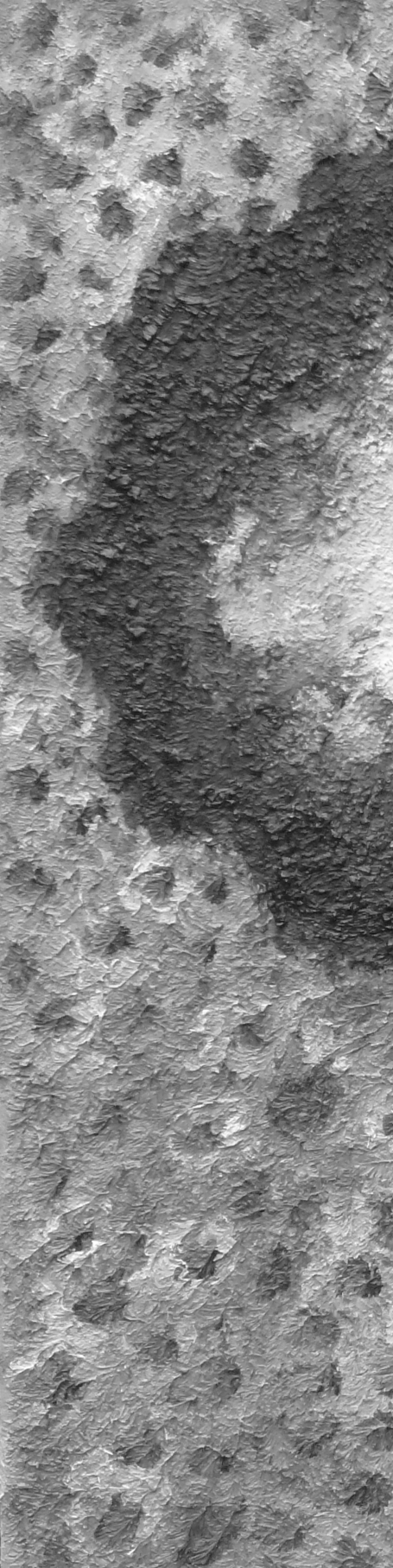
Om de bewegingsvrijheid en de pakkingsdichtheid binnen in de CCMV-capsid te bepalen, zijn de fluorescente eigenschappen van de lading bestudeerd in de selectief geladen capsides. De ‘steady-state’ fluorescentie anisotropiemetingen lieten zien dat de bewegingsvrijheid van de fluorescente eiwitten sterk beperkt wordt door opsluiting in het nanoscopisch kleine CCMV-capsid. Bij een verhoging van het aantal geladen eiwitten in de CCMV-capsid zijn de effecten van de ruimtelijke beperking direct te zien, en ze worden sterker met een hogere lading. Tot 20 fluorescente eiwitten per capsid kan de lading gecontroleerd worden; hogere lading zorgt er echter voor dat de capsides niet correct vormen, wat suggereert dat sterische hinder een belangrijke rol speelt in een correcte capsid-opbouw.

Ook is een nieuw soort eiwitnanokooi onderzocht: bacteriële encapsulines. In deze thesis worden de expressie en purificatie van de *Brevibacterium linens* en *Thermotoga maritima* encapsulines in *E. coli* geoptimaliseerd. Daarnaast is er een strategie ontwikkeld waarbij de natuurlijke peptide van *B. linens* wordt gebruikt om niet-natuurlijke lading (TFP) te laden in deze encapsuline. In vergelijking met de CCMV-ladingsefficiëntie heeft het laden van multimere ladingen (dimeer TFP) geen effect op de encapsuline-beladingsefficiëntie. Voor de dimere lading zijn zelfs oneven beladingsaantallen gemeten, waardoor het lijkt dat de beladingsinteracties sterker zijn dan de dimerisatie-interacties. Dit zou ook verklaren waarom er zulke sterke verschillen zitten tussen de gemeten beladingsaantallen: qua massa werden er veel meer fluorescente eiwitten gemeten dan qua fluorescentie. Als de belading te efficiënt gebeurt, kunnen de fluorescente eiwitten niet volledig vouwen en ontwikkelen, waardoor er geen fluorescentie ontstaat. Voor een goede vouwing is genoeg ruimte nodig, die in een volle encapsuline waarschijnlijk ontbreekt.

Al met al is de encapsulatie van niet-natuurlijke lading in encapsulines een vrij nieuwe discipline en zijn er nog veel aspecten die onderzocht moeten worden. Zo zal er een duidelijk onderscheid gemaakt moeten worden tussen de *in vivo* en *in*

in vitro encapsulatie van lading. In deze thesis wordt een goede controle over de belading van een eiwit-nanokooien via een *in vitro* methode bereikt, door het zuiveren en gedoseerd mengen van de componenten. *In vivo* zal dit niet mogelijk zijn, omdat de belading hier door de natuurlijke beladingsmechanismen wordt gestuurd, zonder controle van buitenaf. Dit zou een significant nadeel voor de *in vivo* belading (van encapsulines) kunnen zijn. Dit is echter alleen voor TFP bepaald. Om hier meer over te kunnen zeggen en om een beter begrip van de beladingsmethode te krijgen, zullen er ook andere ladingen bestudeerd moeten worden. Daarnaast is er een groot verschil tussen de maximaal behaalde lading (ongeveer 20 en 12, respectievelijk voor CCMV en encapsuline), ondanks dat het interne volume van de CCMV-capside en de encapsuline praktisch gelijk is. Tot op heden is het onbekend of dit verschil veroorzaakt wordt door het kapsel of door de lading. In vervolgonderzoek zal ook het formaat van de lading in ogenschouw genomen moeten worden, want het is niet vanzelfsprekend dat er van een kleinere lading meer geladen kan worden.

Als we de factoren die de vorming en belading van eiwit-nanokooien beïnvloeden beter begrijpen, kan dit voor de (bio-)nanotechnologie veel nieuwe deuren openen.



List of abbreviations

9.1 Abbreviations in the text

°C	Degree Celcius
AaLS	Lumazine synthase from <i>Aquifex aeolicus</i>
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
AdhD	Alcohol dehydrogenase D from <i>Pyrococcus furiosus</i>
Amp	Ampicillin
ApE	'A plasmid editor' software
Asp-AMC	Aspartic acid - 7-amino-4-methylcoumarin
BdT	<i>B. linens</i> encapsulin with dTFP cargo
BFP	Blue fluorescent protein
Blin	<i>B. linens M18</i> encapsulin
BmT	<i>B. linens</i> encapsulin with mTFP cargo
Cam	Chloramphenicol
CCMV	Cowpea chlorotic mottle virus
CelB	homotetrameric β -glycosidase
CFP	Cyan fluorescent protein
CK	Recombinant capsid protein with N-terminal K-coil
CPMV	Cowpea mosaic virus
DiFMU	6,8-difluoro-7-hydroxy-4-methylcoumarin octanoate
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
dTE	Dimeric TFP with a C-terminal E-coil
dTECK	Non-covalent complex of dTE with CK
DyP	Dye-decolorizing peroxidase
<i>e.g.</i>	<i>exempli gratia</i> (for example)
E-coil	α -helical repeating unit (negative charged)
EGE	Enhanced green fluorescent protein C-terminal with E-coil
EGECK	Non-covalent complex of EGE and CK
EGFP	Enhanced green fluorescent protein
<i>et al.</i>	<i>et alia</i> (and others)
Eut	Ethanolamine utilization microcompartment
EX	Bacteriophage P22 expanded form
Flp	Ferritin-like protein
FP	Fluorescent protein
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein

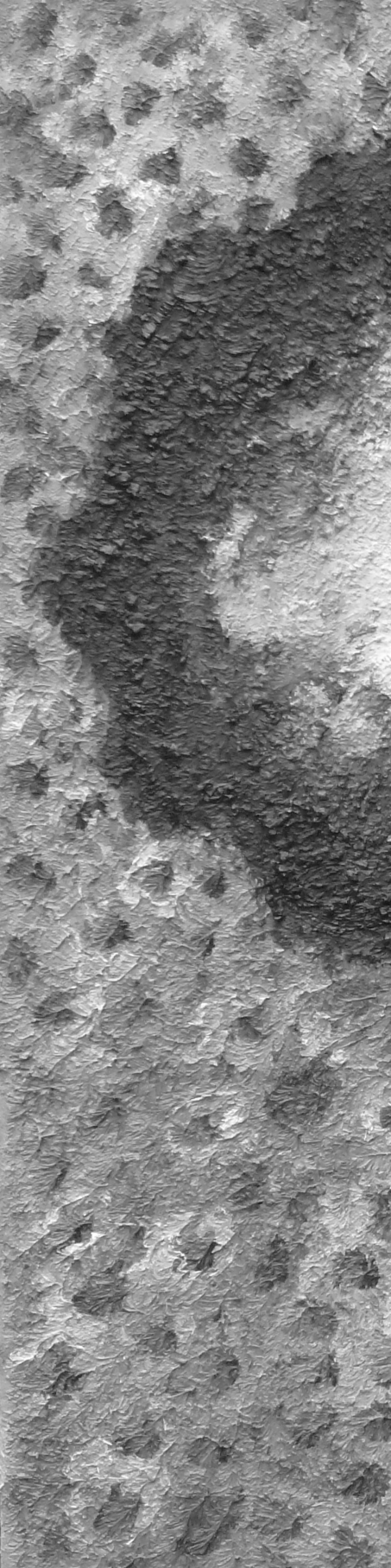
HRP	Horseradish peroxidase
HTC	Fusion protein composed of a his ₆ -tag, mTFP and the CCMV capsid protein
<i>i.e.</i>	<i>id est</i> (that is)
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
K-coil	Positive charged α -helical peptide repeating unit
LB	Lysogeny broth bacterial medium
Luc	Luciferase
<i>m/z</i>	Mass to charge ratio
MALS	Multi-angle light scattering
min	Minutes
mTFP	Monomeric teal fluorescent protein
MWCO	Molecular weight cut-off
Ni-NTA	Nickel nitrilotriacetic acid
PAGE	Poly-acrylamide gel electrophoresis
PalB	Lipase B from <i>Pseudozyma antarctica</i>
PC	Bacteriophage P22 procapsid
PCR	Polymerase chain reaction
PDB	Protein database
Pdu	1,2-propanediol utilization organelles
PepE	Peptidase E, an aspartate dipeptidase
PhoA	Alkaline phosphatase
PNPG	4-Nitrophenyl β -D-glucuronide
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SE	mStrawberry with a C-terminal E-coil
SEC	Size exclusion chromatography
SECK	Non-covalent complex of SE with CK
sfGFP	Superfolder green fluorescent protein
SV40	Simian virus 40
TE	mTFP with a C-terminal E-coil
TECK	Non-covalent complex of TE with CK
TEM	Transmission electron microscopy
TM	<i>T. maritima</i> encapsulin
TMAO	Trimethyl amine N-oxide
TMV	Tobacco mosaic virus
tsLuc	Thermostable luciferase

List of abbreviations

tTE	Tetrameric TFP with a C-terminal E-coil
tTECK	Non-covalent complex of tTE with CK
UV-vis	Ultraviolet - visible spectrometry
VLP	Virus-like particle
WB	Bacteriophage P22 wiffle ball
wtCP	Native wild type CCMV capsid protein, derived from the virus
yCD	Yeast cytosine deaminase
β ME	β -mercaptoethanol

9.2 Amino acids

Ala	alanine (A)
Arg	arginine (R)
Asn	asparagine (N)
Asp	aspartate (D)
Cys	cysteine (C)
Gly	glycine (G)
Glu	glutamate (E)
Gln	glutamine (Q)
His	histidine (H)
Ile	isoleucine (I)
Leu	leucine (L)
Lys	lysine (K)
Met	methionine (M)
Phe	phenylalanine (F)
Pro	proline (P)
Ser	serine (S)
Thr	threonine (T)
Trp	tryptophane (W)
Tyr	tyrosine (Y)
Val	valine (V)



Dankwoord / Acknowledgements

Na vier jaar hard werken aan deze thesis is de zinsnede “het is volbracht” nog nooit zo welgemeend en voldaan uitgesproken, als op het moment dat mijn concept thesis werd goedgekeurd. Er zijn perioden geweest waarin ik de profetische woorden van (reeds gepromoveerde) vrienden, “een promotie is in feite een vier jaar durende cursus ‘omgaan met frustraties’ ”, als enige waarheid kon zien. Maar nu deze thesis af is, zie ik dat elk onderzoek ups en downs heeft, met betrekking tot resultaten, publicaties of toekomst perspectieven en dat je tijdens een promotie daarmee leert omgaan. Dit alles leert je relativiseren en leert je vooral door te gaan en ‘de volgende stap’ te zetten. Mijn volgende stap is nu Enschede verlaten, maar dat doe ik niet zonder een heel aantal mensen te bedanken, zelfs als ik hieronder je naam vergeet te noemen!

Prof. dr. Cornelissen, Jeroen, bedankt dat je me gedurende mijn promotie zoveel vertrouwen hebt gegeven en zoveel waarde hechtte aan mijn kennis en ervaring op moleculair biologisch vlak. Daarnaast ben ik ook erg blij met de wijze waarop je mij en mijn onderzoek gecoacht en gestuurd hebt, mij zelf mijn richting laten bepalen en soms subtiel (of minder subtiel, als ik per se iets wilde uitzoeken dat toch echt niet relevant was) een beetje hebt bijgestuurd. Na ongeveer twee jaar heb je de taak van directe begeleiding overgedragen aan Melissa, maar bleef je wel een vinger aan de pols houden.

Dr. Koay, Melissa, halfway through my PhD you became my co-promoter and the person I could discuss my micro-biology ideas and troubles with. Many discussions followed, sometimes a bit heated if we misunderstood each other and tried to explain the same idea to each other in a different way, but they always ended in good spirit. Thank you for not only being just a supervisor and co-promoter, but also taking interest in me on a more personal note.

De wetenschappelijke staf van het SAnS cluster, ook wel de BNT en MnF groepen: Nathalie, Jurriaan, Pascal en Wim wil ik hartelijk bedanken voor de input en kritische vragen gedurende de colloquia. Alle adviezen aan de koffietafel over organische chemie of mijn onbekende meetmethoden waren erg waardevol.

Naast ons SAnS cluster wil ik ook graag professor Albert Heck bedanken, natuurlijk voor de introductie bij professor Nenad Ban en dr. Markus Sutter, maar ook voor de mogelijkheid om samen te werken aan de encapsulines en de gezellige en nuttige gesprekken in Zurich en Ventura. Naast Albert wil ik ook graag zijn promovendi Dr. Charlotte Uetrecht en Joost Snijder bedanken voor hun expertise in massa spectrometrie en de samenwerking in het encapsuline project.

Within the university of Twente I collaborated closely with dr. Christian Blum and professor Vinod Subramaniam on the fluorescent topics. I like to thank you both for donating materials and a lot of expertise.

Professor Johan Engbersen en dr. Maarten Merckx wil ik bedanken voor het beoordelen van mijn concept proefschrift en voor het zitting nemen in mijn promotiecommissie.

Jasper en Sander wil ik bedanken dat ze mij als paranimfen helpen.

Jasper, als taalgenoot en een van de weinige biologen in onze chemische groep, hebben we vaak van gedachten gewisseld over pittige vraagstukken, een koffiezetter met 'stealth' modus gebouwd en de nodige andere onzinnige dingen gedaan!

Sander, jou ken ik al zo lang, dat mijn keuze voor jou als paranimf gewoon logisch was. Al is het alleen al om alle mooie afleiding tussen het promoveren door, zoals het Scooter concert in Berlijn en de vakantie in Tromsø.

Natuurlijk is een promotietraject in je eentje onmogelijk, daarom wil ik de technici van SAnS, Richard, Marcel, Regine, Bianca en Tieme bedanken voor alle hulp, bestellingen en gezellige koffiepraatjes.

Naast mijn 'eigen' technici, wil ik ook de NBP technici, Yvonne, Kirsten, Irene, Nathalie en Ine bedanken voor de hulp bij de bacteriekweken toen SAnS nog geen eigen ML1 lab had en voor alle microbiologische adviezen gedurende de vier jaar promotie.

Naast de technici hebben ook de secretaresses Isabel en Nicole mij veel hulp gegeven, met name in de bureaucratische jungle op de UT.

I started my PhD project almost at the same time as Melanie. With her I shared all the joys and troubles of research and I want to thank her for all the great moments, good conversations and nice trips we had together.

Anne, Martijn, Yujie, Sarah-Jane, Supitchaya, Chengfen, Piotr, Maarten and Benjamin, I enjoyed working with you guys in the BNT group.

De Nederlandse koffiegroep, Pieter, Rick, Sven, Wies, Janneke, Maarten en Tom wil ik bedanken voor alle gezellige koffiepauzes, lunches en alles wat we daarbuiten hebben gedaan.

Off course I will not forget all the (former) SAnS group members: I liked working with you.

Tijdens de promotie heb ik ook een aantal studenten begeleid.

Fabian, je vond je bachelorproject zo leuk dat je terug kwam voor je masterproject en hebt daar zo goed aan gewerkt dat we het nu gaan publiceren.

Barbara, jouw bachelor project leek zo mooi, maar zoals wel eens gebeurt met onderzoek werd jouw conclusie: dit werkt niet. Desondanks heb je er wel veel tijd in gestoken en daarvoor bedank ik je.

Stan en Iris, jullie hadden dan wel een kort project maar ook daarvan waren de resultaten nuttig voor mijn proefschrift, bedankt.

Buiten het werk in Enschede heb ik veel plezier beleefd aan het hockeyen bij Heren 2 van DHC Drienerlo. Alle teamgenoten wil ik bij dezen bedanken voor 4 jaar sportief plezier. Speciaal wil ik nog Redmar, Thomas en Bram bedanken voor een heel gezellig CHC-jaar en wil ik Thomas, Jorn en Dirk bedanken voor het doorlezen van het manuscript op spelfouten.

Natuurlijk wil ik ook iedereen in Eindhoven bedanken voor het aanhoren van de voor jullie vaak onbegrijpelijke onderzoeksproblemen en het daarop bemoedigend antwoorden.

Noortje, Kim, Rob, Tristan, Sjoerd, Maarten, Marcel, Ruud, Roos, Marleen en alle anderen van UB, KiKa, D&D of daarbuiten, bedankt voor alle goede adviezen, bezoeken aan het ‘verre’ Enschede en leuke weekenden.

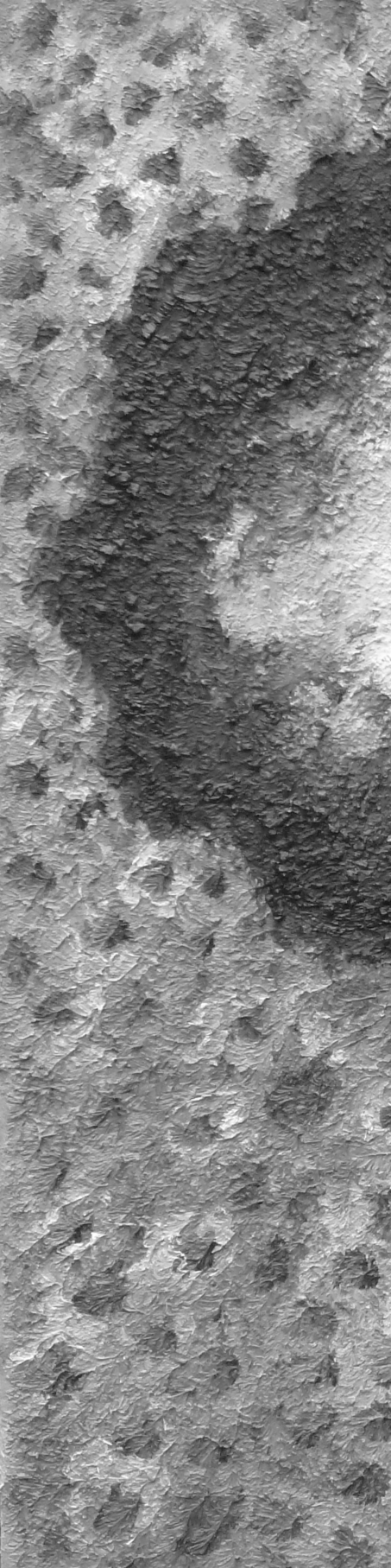
Ruud wil ik nog even speciaal bedanken voor de extra controle op de Nederlandse samenvatting.

Rianne, jij hebt het afgelopen jaar in de afronding van mijn promotie zoveel beren, die ik op de weg zag, verjaagd dat ‘dank je wel’ misschien niet genoeg is. Toch wil ik je bedanken voor de steun en aanmoediging gedurende het laatste half jaar van mijn promotie. Op naar de volgende uitdagingen!

Saske, lief zusje, bedankt dat je me tijdens de 4 jaar hier in Enschede zo gemist hebt. Ik ben erg blij dat jij me op de hoogte houdt van alles wat nieuw is en me daarmee (relatief) jong houdt. Nu ik klaar ben met leren en studeren ga jij beginnen: heel veel succes met jouw studie-pad.

Marjolein, lieve zus, bedankt voor alle steun, al dan niet door de telefoon, aan mij in het verre oosten. Maar ook dat je me een half jaar voor mijn promotie hebt meegesleept naar Kameroen en mij daar aanmoedigde als ik midden in de jungle (bijvoorbeeld op N 02°17.719', E 015°42.768') stukken van deze thesis zat te schrijven.

Thera, Bert, jullie zijn de beste ouders die iemand zich kan wensen: altijd stonden jullie klaar om te helpen, aan te moedigen of goede raad te geven. Zonder jullie aanmoediging en vertrouwen zou ik nooit geworden zijn wie ik nu ben, bedankt!



List of publications

11.1 Publications

W. F. Rurup, F. C. Verbij, M. S. T. Koay, C. Blum, V. Subramaniam, J. J. L. M. Cornelissen, *Controlled loading of Virus-like Particles*, manuscript in preparation

W. F. Rurup, J. Snijder, M. S. Koay, A. J. R. Heck, J. J. L. M. Cornelissen, *Loading of fluorescent proteins in a bacterial nanocompartment*, manuscript in preparation

W. F. Rurup, M. S. T. Koay, J. J. L. M. Cornelissen, “Engineering Protein Cages as Nanocarriers” in the book series “Methods in Molecular Biology” manuscript in preparation

W. F. Rurup, M. S. Koay, J. J. L. M. Cornelissen, *Viruses as enzyme nanocarriers*, chapter in the book “Enzyme Nanocarriers”, submitted

M. Golynskiy, W. F. Rurup, M. Merkx, *Antibody detection using a FRET-based protein conformational switch*, ChemBioChem (2010) 11, 2264-2267

11.2 Conference abstracts and presentations

W. F. Rurup, M. S. Koay, J. J. L. M. Cornelissen, *Loading capsids: less is more*, GRC Physical virology 2013, Ventura, California, USA (Oral presentation)

W. F. Rurup, F. C. Verbij, J. Snijder, M. S. Koay, A. J. R. Heck, J. J. L. M. Cornelissen, *Loading protein shells*, NWO Organic Chemistry & Synthesis Study Group Meeting 2012, Lunteren, The Netherlands

W. F. Rurup, F. C. Verbij, C. Xing, M. S. Koay, J. J. L. M. Cornelissen, *From virus to medicine*, MESA+ meeting on Nanotechnology 2012, Enschede, The Netherlands (Oral presentation)

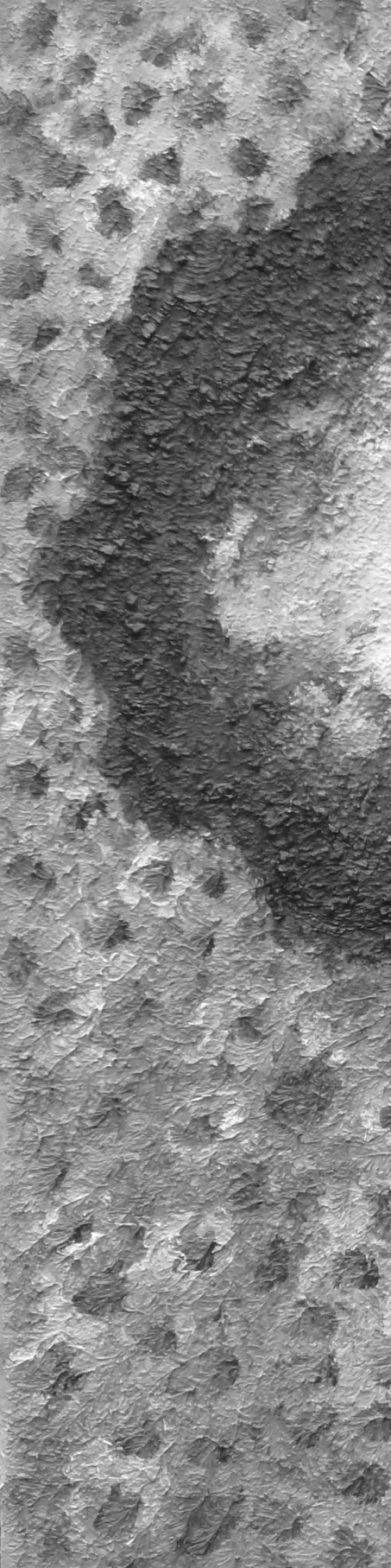
W. F. Rurup, F. C. Verbij, M. S. Koay, J. J. L. M. Cornelissen, *Controlling capsid loading: clever compartment design*, Dutch Polymer Days 2012, Veldhoven, The Netherlands

W. F. Rurup, M. S. Koay, J. J. L. M. Cornelissen, *Directed Loading of Virus-Like Particles*, CHAINS 2011, Utrecht, The Netherlands (Oral presentation)

M. Brasch, W. F. Rurup, A. van der Ham, M. Verwegen, J. J. L. M. Cornelissen, *Biomolecular Nanotechnology*, MESA+ meeting on Nanotechnology 2011, Enschede, The Netherlands

W. F. Rurup, J. J. L. M. Cornelissen, *Directed Loading of Virus-Like Particles*, GRC Physical virology 2011, Ventura, California, USA

W. F. Rurup, J. J. L. M. Cornelissen, *Loading protein cages*, NWO Design and Synthesis, Structure and Reactivity & Biomolecular Chemistry Study Group Meeting 2010, Lunteren, The Netherlands



Curriculum vitae

Rik Rurup was born in Voorburg, The Netherlands, on June 4, 1982. He studied Biomedical Engineering at the Eindhoven University of Technology, The Netherlands, where he received his master degree (ingenieur) in January 2009. The master thesis “Development of a FRET-Based Sensor Protein for Antibody Detection” was performed in the Laboratory of Chemical Biology, Department of Biomedical Engineering of Prof. dr. E.W. Meijer, under the supervision of Dr. M. Merkx and Dr. M.V. Golynskiy.

In the beginning of september 2009, he started as a Ph.D. candidate at the University of Twente under supervision of Prof. dr. J.J.L.M. Cornelissen and Dr. M. S. T. Koay. The aim of his project was to develop a rational loading mechanism for protein cages. The results of this project are described in this thesis.

